https://doi.org/10.18502/jad.v20i2.22403

Original Article

Spatio-Temporal Modeling of Malaria Occurrence Probability Using Multi-Sensor Remote Sensing and Machine Learning

***Gholamreza Hassanpour[1], *Fahimeh Youssefi[2], Awat Dehghan[3], Hossein Keshavarz[1,4], Ahmad Ali Hanafi-Bojd[3,5]**

[1]Center for Research of Endemic Parasites of Iran, Tehran University of Medical Sciences, Tehran, Iran
[2]Department of Photogrammetry and Remote Sensing, Faculty of Geodesy and Geomatics Engineering, K.N. Toosi University of Technology, Tehran, Iran
[3]Department of Vector Biology and Control of Diseases, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran
[4]Department of Medical Parasitology and Mycology, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran
[5]Zoonoses Research Center, Tehran University of Medical Sciences, Tehran, Iran

***Corresponding authors:** Dr Gholamreza Hassanpour, E-mail: hassanpour@tums.ac.ir; Dr Fahimeh Youssefi, E-mail: fahimeh.y@gmail.com

(Received 06 Feb 2026; accepted 05 June 2026)

Abstract

Background: Reliable mapping of malaria occurrence requires validation across environmentally distinct years and careful separation of model-derived probability from transmission intensity. This study developed a multi-sensor remote sensing and machine-learning framework for mapping annual malaria case occurrence in southeastern Iran.

Methods: Annual median composites of Soil Water Index (SWI), Normalized Difference Vegetation Index (NDVI) and Land Surface Temperature (LST) were generated from Sentinel-1, Sentinel-2 and Landsat 8/9 at 30m resolution. After quality control, 82, 168, 775 and 483 malaria case locations were retained for the Solar Hijri years 1400–1403 (2021–2025). Five spatially balanced pseudo-absence realizations were generated annually at a 1:1 ratio with a 300m exclusion buffer. Random Forest (RF) and eXtreme Gradient Boosting (XGBoost) were evaluated through complete four-fold leave-one-year-out validation with nested tuning. Performance, bootstrap confidence intervals, uncertainty, Getis–Ord Gi* clusters and SHAP (SHapley Additive exPlanations) were assessed.

Results: Across 20 outer evaluations, RF achieved ROC-AUC 0.685±0.045, PR-AUC 0.693±0.040, balanced accuracy 0.624±0.026 and Brier score 0.230±0.015. XGBoost achieved ROC-AUC 0.687±0.044, PR-AUC 0.691±0.043, balanced accuracy 0.625±0.033 and Brier score 0.231±0.015. Year-specific ROC-AUC ranged from 0.624 to 0.726. LST was the dominant predictor (48.27%), followed by NDVI (37.04%) and SWI (14.69%). Only 3.46–6.10% of pixels exceeded probability 0.70. Mean uncertainty was low (0.030–0.032) but spatially heterogeneous.

Conclusion: Both algorithms showed moderate, comparable performance with year-dependent variation. Maps provide relative probabilities for prioritizing field surveillance, not transmission estimates.

Keywords: Malaria occurrence modeling; Random Forest; XGBoost; Pseudo-absence; Iran

Introduction

Malaria remains a major vector-borne disease and continues to impose substantial health burdens despite sustained progress in diagnosis, treatment and vector control (1). Transmission is produced by interactions among *Plasmodium* parasites, competent *Anopheles* vectors, susceptible human populations, mobility, healthcare access and intervention coverage. Environmental conditions influence only part of this system. Still, they are especially important because they regulate the availability of aquatic habitats, mosquito development and survival,

http://jad.tums.ac.ir
Published Online: June 30, 2026

Copyright © 2026 The Authors. Published by Tehran University of Medical Sciences.
This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International license (https://creativecommons.org/licenses/by-nc/4.0/). Non-commercial uses of the work are permitted, provided the original work is properly cited.

https://doi.org/10.18502/jad.v20i2.22403

adult activity and the duration of parasite development within the vector (2, 3). Temperature commonly produces nonlinear effects: cool conditions may slow immature development and parasite incubation, whereas excessive heat can reduce mosquito survival. Surface water and soil wetness determine whether potential larval habitats form and persist, while vegetation can modify shade, humidity and local microclimate. Environmental mapping can therefore support vector surveillance, although it cannot independently quantify human malaria incidence or transmission intensity (4–6).

These relationships are particularly relevant in arid and semi-arid regions, where apparently dry landscapes may contain localized but epidemiologically important aquatic habitats. Rainfall is irregular, evaporation is high and suitable habitats may be concentrated around rivers, irrigated agricultural areas, wells, tanks, cisterns and other human-made water-storage systems. Small environmental changes can consequently alter the spatial continuity and persistence of vector habitats between years. Southeastern Iran also has cross-border epidemiological importance and remains environmentally receptive to malaria-vector activity even when locally acquired malaria is reduced. Beyond these ecological drivers, genetic diversity within parasite populations poses an additional challenge. Studies have documented considerable genetic diversity in *Plasmodium falciparum* isolates from endemic areas of Iran, particularly in the merozoite surface protein (MSP)-1 and MSP-2 genes, which may facilitate immune evasion and sustain transmission despite control efforts (7). Furthermore, the prevalence of asymptomatic malaria infections has been identified as a critical obstacle to elimination programs. Asymptomatic carriers, who remain undetected by routine passive surveillance, serve as a persistent reservoir for transmission and contribute to the continued occurrence of malaria cases (8). Compounding these biological and epidemiological challenges, molecular analyses have revealed the presence of mutations in the *P. falciparum* dihydrofolate reductase (dhfr) and dihydropteroate synthase (dhps) genes, which are associated with resistance to sulfadoxine-pyrimethamine (SP), a previously widely used anti-malarial drug in Iran (9). The emergence of drug resistance further threatens the effectiveness of therapeutic interventions. Collectively, these findings underscore the need for integrated surveillance strategies that combine environmental risk mapping, climatic monitoring, parasitological assessment and molecular surveillance to achieve malaria elimination in Iran. This setting makes it necessary to distinguish environmental receptivity and recorded malaria case occurrence from the broader concept of malaria transmission, which additionally depends on parasite importation, human movement, surveillance, and control activities.

Anopheles stephensi is a principal malaria vector in southern Iran and is well adapted to both rural and human-modified environments (10–13). The species can exploit diverse larval habitats, including wells, pools, irrigation structures, storage tanks and artificial containers. Previous remote-sensing studies have demonstrated that aquatic *Anopheles* habitats can be detected or predicted using radar, optical, thermal and ancillary environmental information (6, 14). Moisture-sensitive observations may indicate the formation or persistence of water-associated habitats; vegetation can identify irrigated land, riparian corridors, shading and relatively stable microclimates; and surface temperature provides information on thermal conditions at the land–atmosphere interface. However, no single remotely sensed variable captures the complete ecology of *An. stephensi*, and satellite-derived predictors should be interpreted as environmental correlates rather than direct measurements of vector abundance.

Multi-sensor Earth observation is useful because optical, thermal and microwave sensors provide complementary environmental information. Sentinel-1 synthetic aperture radar operates independently of daylight and cloud cover and is sensitive to surface roughness and

http://jad.tums.ac.ir
Published Online: June 30, 2026

https://doi.org/10.18502/jad.v20i2.22403

dielectric properties associated with wetness. Its VV and VH backscatter can therefore contribute information on moisture conditions in locations where short-lived water availability is difficult to observe consistently. Sentinel-2 MultiSpectral Instrument imagery provides 10 m red and near-infrared observations suitable for calculating the Normalized Difference Vegetation Index (NDVI), which summarizes photosynthetically active vegetation and may also indicate irrigation or riparian habitats. Landsat 8/9 thermal observations provide Land Surface Temperature (LST) at a scale appropriate for local environmental mapping. Integrating these sensors offers a reproducible means of characterizing moisture, vegetation and thermal variability, while annual compositing reduces cloud-related data gaps and aligns the predictors with annually grouped occurrence records (4, 14, 15).

Machine-learning algorithms are well suited to these predictors because vector–environment relationships may be nonlinear, interactive and characterized by thresholds. Random Forest (RF) reduces variance by combining many decorrelated decision trees, whereas Extreme Gradient Boosting (XGBoost) constructs regularized trees sequentially to correct residual errors. Both approaches can represent interactions, such as a moisture effect that changes across temperature conditions, without requiring those relationships to be specified in advance. Nevertheless, flexible algorithms can reproduce sampling bias and may appear more accurate than they actually are when training and test observations share the same temporal or spatial structure. In presence–background modeling, performance is also affected by the number and location of pseudo-absences, which are not confirmed biological absences. Selecting a single pseudo-absence sample may therefore conceal an important source of uncertainty (16–18).

Validation design is consequently central to the credibility of spatio-temporal ecological models. Randomly partitioning pooled annual observations allows records from the same year to occur in both training and testing subsets and may reward a model for learning year-specific environmental structure. Such a strategy does not directly test whether a model trained under previous or alternative annual conditions can transfer to an unseen year. Leave-one-year-out (LOYO) validation addresses this issue by withholding an entire annual dataset during model development and repeating the procedure until every year has served as the independent test period (19, 20). A complete LOYO design is particularly informative for semi-arid environments, where annual changes in rainfall, vegetation, surface moisture and extreme temperature can alter both predictor distributions and vector habitats. Repeating every fold across multiple pseudo-absence realizations additionally helps distinguish temporal variation from sensitivity to background selection.

Predictive performance alone is insufficient for operational interpretation. Ensemble tree models can generate accurate rankings without revealing which variables drive a prediction or whether their effects are monotonic. SHapley Additive exPlanations (SHAP) decompose model output into feature-specific contributions and can summarize global importance, direction, nonlinearity and interactions (21, 22). At the same time, pixel-wise variation among repeated pseudo-absence models can identify locations where mapped probabilities are sensitive to background selection. Continuous probability maps, binary threshold maps, uncertainty surfaces, interannual correlations and Getis–Ord Gi* clusters provide complementary rather than interchangeable information (23). Continuous maps preserve relative probability gradients; threshold maps support screening under an explicitly selected cut-off; uncertainty maps identify unstable predictions; and Gi* analysis identifies local clusters of high or low model output rather than confirmed clusters of malaria transmission.

Despite the value of temporally explicit validation, few local studies in southeastern Iran

http://jad.tums.ac.ir
Published Online: June 30, 2026

https://doi.org/10.18502/jad.v20i2.22403

have combined multi-sensor environmental predictors, complete LOYO evaluation, repeated pseudo-absence sampling, prediction uncertainty and model interpretation within a single reproducible framework. This study addressed that gap by aiming to: (i) model annual malaria case occurrence probability using the Soil Water Index (SWI), NDVI and LST; (ii) compare RF and XGBoost through complete four-fold LOYO validation repeated across five pseudo-absence realizations; (iii) quantify year-specific performance, bootstrap confidence intervals, prediction uncertainty and interannual spatial change and (iv) interpret RF predictions using SHAP.

Materials and Methods

Study area

Sarbaz County is located in Sistan and Baluchistan Province in southeastern Iran (Fig. 1), close to the Pakistan border (approximately 25°30′–26°40′ N and 60°30′–61°40′ E). The county has an arid to semi-arid climate, high summer temperatures, irregular precipitation, mountainous terrain, river valleys, irrigated agriculture and scattered settlements. This combination produces strong environmental contrasts over short distances. Riparian areas, irrigation systems, wells, tanks and other water-holding structures may provide habitats for *Anopheles* mosquitoes, including *An. stephensi*, even when the surrounding landscape is dry.

Environmental predictors and annual compositing

Three predictors were generated in Google Earth Engine (GEE). SWI was derived from descending-orbit Sentinel-1 Ground Range Detected observations acquired in interferometric wide-swath mode using VV and VH backscatter. NDVI was calculated from Sentinel-2 MultiSpectral Instrument surface-reflectance imagery as (B8−B4)/(B8+B4). LST was derived from Landsat 8/9 Collection 2 Level-2 surface-temperature observations after applying the product scale and offset factors.

For Sentinel-2, cloud and cloud-shadow contamination was removed with the Scene Classification Layer and quality-assurance information, excluding cloud shadow, medium- and high-probability cloud, cirrus and snow/ice classes before compositing. For Landsat 8/9, QA_PIXEL bit flags were used to remove dilated cloud, cirrus, cloud, cloud shadow and snow; radiometric-saturation pixels identified by QA_RADSAT were also excluded. Sentinel-1 does not require optical cloud masking. All retained observations were summarized using the annual median, which is less sensitive than the mean to residual cloud contamination and extreme observations.

The four annual intervals followed the Solar Hijri reporting years: 1400 (21 March 2021–20 March 2022), 1401 (21 March 2022–20 March 2023), 1402 (21 March 2023–19 March 2024) and 1403 (20 March 2024–20 March 2025). Predictor layers were co-registered to a common 30m grid and clipped to the shared valid-data footprint. Annual compositing matched the temporal resolution of the supplied occurrence records. Because month- or season-specific occurrence dates were unavailable, seasonal predictor matching could not be performed. This limitation is considered explicitly in the discussion. The satellite sources, native spatial information, annual aggregation method, final grid resolution and ecological interpretation of the three predictors are summarized in Table 1.

Malaria case occurrence records and quality control

The response data were annual point locations representing recorded malaria cases. They were not locations of human malaria cases. The supplied files did not contain fields describing indigenous/imported status because that classification applies to human infections, nor did they contain locality names, month/season, or coordinate-accuracy estimates. These unavailable attributes are reported transparently and were not inferred.

For each year, geometries were checked, assigned or verified against the documented co-

http://jad.tums.ac.ir
Published Online: June 30, 2026

https://doi.org/10.18502/jad.v20i2.22403

ordinate reference system, transformed to the annual raster CRS and screened against the raster footprint. The year-1401 file lacked embedded CRS metadata; EPSG: 9707 was assigned based on the documented coordinate system, and its transformed locations were visually verified. Exact duplicate coordinates were collapsed so that repeated records at an identical location did not receive additional weight. Points outside the raster footprint or intersecting No-Data predictor cells were excluded. No aggregation by named locality was possible because no locality field was supplied. The number of raw records, exclusions at each quality-control stage and final retained occurrence locations for each year are reported in Table 2.

Pseudo-absence generation

True absence data were unavailable. Five independent pseudo-absence realizations were therefore generated for each year. Candidate cells were restricted to the valid annual raster footprint and were required to be at least 300 m from every retained occurrence location. The distance rule reduces immediate spatial overlap but does not prove biological absence; consequently, these locations are referred to as pseudo-absences or background points throughout. For every year and realization, the number of pseudo-absences equalled the number of retained occurrences, producing a 1:1 class ratio. Sampling was spatially balanced across the valid footprint and used reproducible random seeds derived from a fixed global seed and the realization number. Repeating the analysis five times allowed sensitivity to pseudo-absence selection to be summarized rather than relying on a single random background sample (18).

Each realization therefore contained 164, 336, 1,550 and 966 observations in years 1400, 1401, 1402 and 1403, respectively. The outer test size was determined by the withheld year; the corresponding training sizes were 2,852, 2,680, 1,466 and 2,050 observations. Because LOYO is year-based, conventional fixed training/validation/test percentages were not used.

Machine-learning algorithms

RF and XGBoost binary classifiers were fitted using SWI, NDVI and LST as predictors and occurrence [1] versus pseudo-absence [0] as the response. RF constructs multiple decision trees from bootstrap samples and averages their predictions to reduce variance and represent nonlinear relationships and interactions. The RF occurrence probability was calculated according to Eq. 1:

$$\hat{p}RF(y=1 \mid x)= B^{-1} \Sigma_{\beta} \hat{p}_{\beta}(y=1 \mid x) \quad \text{Eq. 1}$$

where B is the total number of trees and $\hat{p}_{\beta}(y=1 \mid x)$ is the occurrence probability predicted by tree b for predictor malaria cases x.

XGBoost constructs regularized decision trees sequentially, with each additional tree seeking to reduce the residual classification error of the existing ensemble. The prediction at boosting iteration t was represented by Eq. 2:

$$\hat{y}_i^{(t)}= \Sigma_{k=1}^{t} f_k(x_i), \; f_k \in \mathscr{F} \quad \text{Eq. 2}$$

where $\hat{y}_i^{(t)}$ is the model output for observation i after iteration t, f_k denotes the kth regression tree and $\mathscr{F}$ represents the space of possible trees. The models were implemented in Python 3.10.19 using scikit-learn 1.7.2 and XGBoost 1.7.6. Data processing used NumPy 1.26.4, pandas 2.3.3, GeoPandas 1.1.1 and rasterio 1.4.3; SHAP 0.48.0 was used for interpretation.

Nested hyperparameter optimization

Hyperparameters were optimized inside each outer LOYO training set. The three available training years formed the groups for inner cross-validation, ensuring that the outer withheld year was never used for tuning. Candidate configurations were selected by mean inner ROC-AUC. The RF search covered 300, 600 or 900 trees; maximum depth 8, 16 or unrestricted; sqrt, log2, or all features per split; minimum leaf size 1, 2 or 4; minimum split size 2, 5, or 10; and no or balanced class weighting. The XGBoost search covered 300, 600 or 900 trees; depth 3, 4 or 6; learning rate 0.01, 0.03 or 0.05; minimum child weight 1, 3 or 5; subsample 0.70, 0.85 or 1.00; column subsample 0.70, 0.85 or 1.00; gamma 0, 0.1 or 0.3; and

http://jad.tums.ac.ir
Published Online: June 30, 2026

https://doi.org/10.18502/jad.v20i2.22403

L2 regularization 1, 5 or 10. Tuning was repeated independently for every outer fold and pseudo-absence realization.

For the final full-data models used to produce annual maps, the selected RF settings were 600 trees, maximum depth 8, sqrt feature sampling, minimum leaf size 1 and minimum split size 5. The selected XGBoost settings were 300 trees, maximum depth 3, learning rate 0.03, minimum child weight 5, subsample 1.00, column subsample 0.85, gamma 0.3 and L2 regularization 10.

Temporal generalization using leave-one-year-out validation

Temporal transferability was evaluated using complete four-fold LOYO validation. During each outer fold, all observations from one year were reserved as an independent test set, while the remaining three years were used for model development. The four-fold procedure was repeated for each of the five pseudo-absence realizations, producing 20 outer evaluations per algorithm. The withheld year was excluded from fitting, hyperparameter selection and fold-specific threshold estimation. The withheld test year, corresponding training years and sample sizes used in each outer LOYO fold are presented in Table 3.

Threshold selection and performance assessment

Continuous probabilities were the primary model output. Where binary predictions were required, the decision threshold was selected by maximizing Youden's J statistic using training-derived out-of-fold predictions. Youden's J was calculated according to Eq. 3:

$$J = sensitivity + specificity - 1 \quad \text{Eq. 3}$$

The selected threshold was subsequently applied without modification to observations from the corresponding withheld year. The final thresholds estimated from all LOYO out-of-fold predictions were 0.549 for RF and 0.566 for XGBoost. The mean ensemble threshold used to generate the annual binary maps was 0.532.

Model performance was assessed using threshold-independent and threshold-dependent Measures. Threshold-independent discrimination was evaluated using the area under the receiver operating characteristic curve (ROC-AUC) and the area under the precision–recall curve (PR-AUC). Probabilistic accuracy was evaluated using the Brier score. Threshold-dependent performance was assessed using accuracy, balanced accuracy, precision, sensitivity, specificity, F1-score, Matthews correlation coefficient (MCC) and Cohen's kappa.

Metrics were first calculated independently for each of the 20 outer evaluations produced by the four withheld years and five pseudo-absence realizations. Year-specific results were summarized as the mean ± standard deviation across the five pseudo-absence realizations. Overall LOYO performance was summarized as the mean ± standard deviation across all four withheld years and five realizations. Nonparametric 95% confidence intervals were obtained by bootstrap resampling of predictions within each withheld-year test set.

SHapley Additive exPlanations interpretation

The final RF model was interpreted using TreeSHAP. For each prediction, the model output was decomposed into a baseline value and the additive contribution of every predictor. This decomposition is expressed in Eq. 4.

$$f(x) = \varphi_0 + \Sigma_{j=1}^{M} \varphi_j \quad \text{Eq. 4}$$

where f(x) is the model output for observation x, φ_0 is the expected baseline model output, φ_j is the SHAP contribution of predictor j, and M is the total number of predictors. A positive φ_j value indicates that the corresponding predictor increases the modeled occurrence probability relative to the baseline, whereas a negative value indicates a decreasing contribution.

Global feature importance was calculated as the mean absolute SHAP value for each predictor across the evaluated samples. The resulting values were expressed both in model-output units and as a percentage of the sum

http://jad.tums.ac.ir
Published Online: June 30, 2026

https://doi.org/10.18502/jad.v20i2.22403

of mean absolute SHAP values. This procedure allowed the overall influence of LST, NDVI and SWI to be ranked while retaining information about the magnitude of their contributions.

A SHAP summary bar plot was used to present global feature-importance rankings, and a beeswarm plot was used to show the distribution, direction and feature-value gradient of individual contributions. SHAP dependence plots were additionally generated for LST, NDVI and SWI to examine nonlinear response patterns and potential interactions. In each dependence plot, the predictor value was displayed on the horizontal axis and its SHAP contribution on the vertical axis, while point colour represented the strongest displayed interacting predictor.

SHAP values describe the behavior of the fitted RF model within the range and distribution of the analyzed observations. They do not demonstrate causality and apparent response ranges or thresholds should not be interpreted as universal biological limits without independent ecological or experimental validation.

Spatial prediction, uncertainty and Getis–Ord Gi* analysis

Final RF models were fitted separately to each pseudo-absence realization and applied to every valid 30m raster cell. The annual probability map was the mean of the five realization-specific maps, while their pixel-wise standard deviation represented uncertainty attributable to pseudo-absence selection. Binary maps were generated using the ensemble threshold of 0.532. Pearson correlations were calculated among annual continuous probability maps over their common valid footprint.

Local clustering was evaluated with Getis–Ord Gi* after aggregating the 30m probability raster by a factor of 20 to reduce computational burden and avoid treating immediately adjacent high-resolution pixels as independent neighborhoods (23). Cells with $z \geq 1.96$ were classified as hot spots, cells with $z \leq -1.96$ as cold spots and remaining cells as not significant. Gi* identifies spatial clusters in model-derived probabilities, not clusters of confirmed transmission.

Results

Occurrence-data screening and sample balance

Of 2,711 raw records, 1,508 unique malaria case occurrence locations had valid geometries, fell within the predictor footprint and contained complete predictor values. Retention differed among years, especially in 1401, for which 521 records were outside the raster footprint and 268 exact coordinate duplicates were collapsed. The complete year-specific quality-control results are presented in Table 2. Within every pseudo-absence realization, annual occurrence and pseudo-absence counts were equal. The resulting class balance prevents prevalence-driven accuracy inflation but does not convert pseudo-absences into confirmed absences.

Four-fold leave-one-year-out performance

Random Forest and XGBoost showed moderate and broadly comparable temporal discrimination. Across all 20 outer evaluations, ROC-AUC was 0.685±0.045 for RF and 0.687± 0.044 for XGBoost. RF produced slightly higher PR-AUC, precision, specificity and MCC and a marginally lower Brier score, whereas XG Boost produced slightly higher sensitivity and F1-score. The small and inconsistent differences do not support strong overall superiority of either algorithm. The overall performance metrics calculated across all four withheld years and five pseudo-absence realizations are reported in Table 4.

Performance varied by withheld year. Both models performed least strongly for year 1400 and most strongly for year 1401. High sensitivity but low specificity in year 1402 indicates that the selected training-derived thresholds labelled a relatively large part of the 1402 background sample as occurrence. This between-

http://jad.tums.ac.ir
Published Online: June 30, 2026

https://doi.org/10.18502/jad.v20i2.22403

year variation is central to the interpretation of temporal transferability. The complete year-specific LOYO results are provided in Table 5.

The ROC, precision–recall and calibration curves for RF and XGBoost are shown in Fig. 2. The curves indicate differences in discrimination and probability calibration among withheld years. The AUC and average-precision values displayed in Fig. 2 were obtained from predictions pooled across the five pseudo-absence realizations and may therefore differ slightly from the means of the five realization-specific metrics reported in Table 5. The variability of ROC-AUC, PR-AUC, balanced accuracy, F1-score, MCC, Cohen's kappa and Brier score across years and pseudo-absence realizations is further illustrated in Fig. 3.

Annual occurrence-probability patterns and uncertainty

Annual mean probabilities ranged from 0.352 in 1403 to 0.466 in 1402. Contrary to the impression created by the earlier colour scale, only 3.46–6.10% of valid pixels exceeded 0.70. Threshold-positive area was greatest in 1402 (35.61%) and lowest in 1401 (18.17%). Pixel-wise uncertainty attributable to pseudo-absence selection was similar among years, with annual means of approximately 0.030–0.032, although locally elevated uncertainty occurred along transition zones. The annual probability, threshold-positive area and uncertainty statistics are summarized in Table 6.

The spatial distributions of continuous model-derived malaria case occurrence probability for years 1400–1403 are shown in Fig. 4. The maps demonstrate both persistent environmental structures and interannual shifts in predicted occurrence. The corresponding binary predictions produced using the ensemble threshold of 0.532 are presented in Fig. 5. These binary outputs distinguish threshold-positive cells from cells below the selected cut-off but should not be interpreted as verified presence and absence. Spatial variation in prediction uncertainty across the five pseudo-absence realizations is shown in Fig. 6, with locally higher uncertainty occurring primarily near probability-transition zones.

Interannual spatial correspondence and local clustering

Pairwise correlations among annual probability maps were positive but moderate, ranging from 0.33 between 1402 and 1403 to 0.66 between 1401 and 1403. Thus, the maps contained both persistent spatial structure and substantial interannual redistribution; none of the annual maps should be treated as a direct reconstruction or gap-filled version of another year. The complete correlation matrix for the four annual probability maps is presented in Fig. 7.

Getis–Ord Gi* identified extensive hot spots, cold spots, and nonsignificant zones in every year. At the aggregated analysis scale, the numbers of hot-spot cells were 7,955, 8,791, 6,947 and 8,308 for years 1400–1403, respectively; the corresponding numbers of cold-spot cells were 6,808, 8,373, 6,984 and 8,633. The annual spatial distributions of significant hot spots, cold spots and nonsignificant cells are shown in Fig. 8. These clusters represent local concentrations of model-derived occurrence probabilities rather than observed clusters of malaria transmission.

SHapley Additive exPlanations interpretation

Land surface temperature had the greatest global contribution, accounting for 48.27% of total mean absolute SHAP importance, followed by NDVI (37.04%) and SWI (14.69%). This ranking corrects the earlier statement that NDVI was the dominant predictor. The beeswarm distribution showed that all three predictors had nonlinear and interacting relationships with model output. The global mean absolute SHAP ranking and the distribution and direction of feature contributions are presented in Fig. 9.

Land surface temperature values of approximately 33–36 °C generally made positive contributions, whereas cooler conditions and temperatures above approximately 37–38 °C tend-

http://jad.tums.ac.ir
Published Online: June 30, 2026

https://doi.org/10.18502/jad.v20i2.22403

ed to reduce predictions; very high LST values above approximately 41 °C were strongly negative. Normalized difference vegetation index values above roughly 0.10–0.15 generally increased model output, whereas low-to-intermediate values around 0.04–0.09 frequently had negative contributions. Moderate SWI values around 0.60–0.70 produced small positive contributions, while values above approximately 0.75 were generally negative. The SHAP dependence relationships for LST, NDVI and SWI are shown in Fig. 10. These ranges describe the fitted model's behavior within the sampled data and should not be interpreted as universal biological thresholds or causal effects.

Table 1. Remote-sensing predictors used in the analysis. All annual summaries were medians and all final layers were aligned to a 30m grid, Sarbaz County, Sistan and Baluchistan Province, Southeastern Iran

Variable	Source	Native information	Annual summary	Final grid	Ecological interpretation
SWI	Sentinel-1 SAR (VV/VH)	~10 m	Median	30 m	Surface-wetness proxy and potential persistence of aquatic habitats
NDVI	Sentinel-2 MSI (B8, B4)	10 m	Median	30 m	Vegetation cover, shading, irrigation and habitat stability
LST	Landsat 8/9 OLI–TIRS Level-2	30 m	Median	30 m	Surface thermal conditions affecting vector development and survival

SWI: Soil Water Index, NDVI: Normalized Difference Vegetation Index, LST: Land Surface Temperature

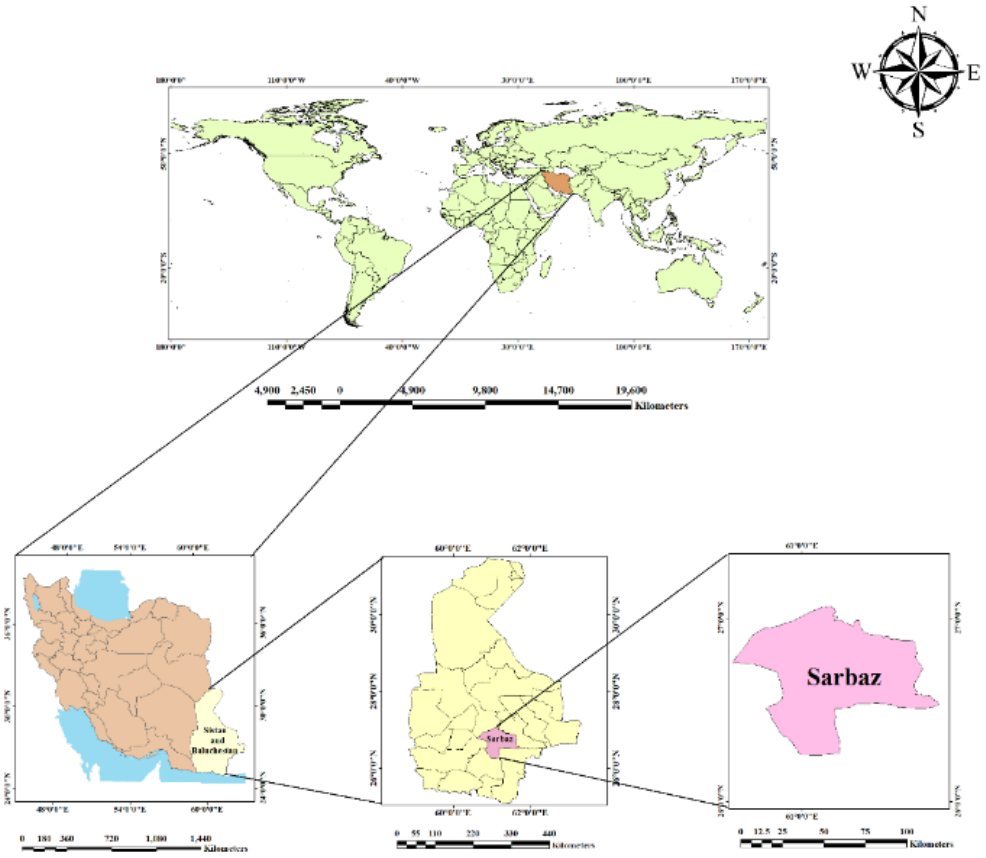

Fig. 1. Location of Sarbaz County, Sistan and Baluchistan Province, in southeastern Iran

http://jad.tums.ac.ir
Published Online: June 30, 2026

https://doi.org/10.18502/jad.v20i2.22403

Table 2. Quality control of annual malaria occurrence records, Sarbaz County, Sistan and Baluchistan Province, southeast of Iran, 1400–1404 Solar Hijri years

Year	Raw	Outside footprint	Duplicate coordinates	NoData	Retained occurrences
1400	94	4	4	4	82
1401	969	521	268	12	168
1402	969	6	98	90	775
1403	679	13	136	47	483
Total	2,711	544	506	153	1,508

Table 3. Definition of the four outer leave-one-year-out (LOYO) folds, Sarbaz County, Sistan and Baluchistan Province, southeast of Iran, 1400−1403 Solar Hijri years

Fold	Withheld test year	Training years	Test n	Training n
1	1400	1401, 1402, 1403	164	2,852
2	1401	1400, 1402, 1403	336	2,680
3	1402	1400, 1401, 1403	1,550	1,466
4	1403	1400, 1401, 1402	966	2,050

Table 4. Overall performance across four withheld years and five pseudo-absence realizations (mean ± standard deviation; n= 20 outer evaluations per model), Sarbaz County, Sistan and Baluchistan Province, southeast of Iran, 1400−1403 Solar Hijri years

Metric	Random Forest	XGBoost
ROC-AUC	0.685 ± 0.045	0.687 ± 0.044
PR-AUC	0.693 ± 0.040	0.691 ± 0.043
Balanced accuracy	0.624 ± 0.026	0.625 ± 0.033
Precision	0.660 ± 0.066	0.647 ± 0.067
Sensitivity	0.581 ± 0.192	0.617 ± 0.168
Specificity	0.666 ± 0.212	0.633 ± 0.195
F1-score	0.594 ± 0.074	0.613 ± 0.062
MCC	0.270 ± 0.046	0.267 ± 0.063
Cohen's kappa	0.248 ± 0.052	0.250 ± 0.065
Brier score	0.230 ± 0.015	0.231 ± 0.015

Table 5. Year-specific leave-one-year-out performance (mean ± standard deviation across five pseudo-absence realizations), Sarbaz County, Sistan and Baluchistan Province, southeast of Iran, 1400−1403 Solar Hijri years

Model	Year	ROC-AUC	PR-AUC	Bal. acc.	Sensitivity	Specificity	F1	Brier
RF	1400	0.624 ± 0.021	0.663 ± 0.015	0.601 ± 0.009	0.432 ± 0.103	0.771 ± 0.086	0.515 ± 0.058	0.244 ± 0.006
RF	1401	0.726 ± 0.010	0.745 ± 0.012	0.651 ± 0.018	0.493 ± 0.080	0.810 ± 0.056	0.582 ± 0.052	0.216 ± 0.005
RF	1402	0.668 ± 0.011	0.655 ± 0.021	0.602 ± 0.010	0.883 ± 0.025	0.321 ± 0.030	0.689 ± 0.008	0.243 ± 0.006
RF	1403	0.721 ± 0.009	0.710 ± 0.011	0.641 ± 0.011	0.518 ± 0.035	0.764 ± 0.036	0.590 ± 0.020	0.218 ± 0.004
XGB	1400	0.635 ± 0.024	0.666 ± 0.016	0.595 ± 0.020	0.554 ± 0.080	0.637 ± 0.074	0.575 ± 0.041	0.243 ± 0.009
XGB	1401	0.732 ± 0.005	0.748 ± 0.006	0.670 ± 0.006	0.588 ± 0.073	0.752 ± 0.067	0.639 ± 0.032	0.214 ± 0.002
XGB	1402	0.660 ± 0.012	0.646 ± 0.025	0.602 ± 0.009	0.869 ± 0.024	0.336 ± 0.038	0.686 ± 0.005	0.246 ± 0.005
XGB	1403	0.719 ± 0.013	0.704 ± 0.013	0.632 ± 0.014	0.457 ± 0.065	0.807 ± 0.046	0.552 ± 0.042	0.219 ± 0.004

http://jad.tums.ac.ir
Published Online: June 30, 2026

https://doi.org/10.18502/jad.v20i2.22403

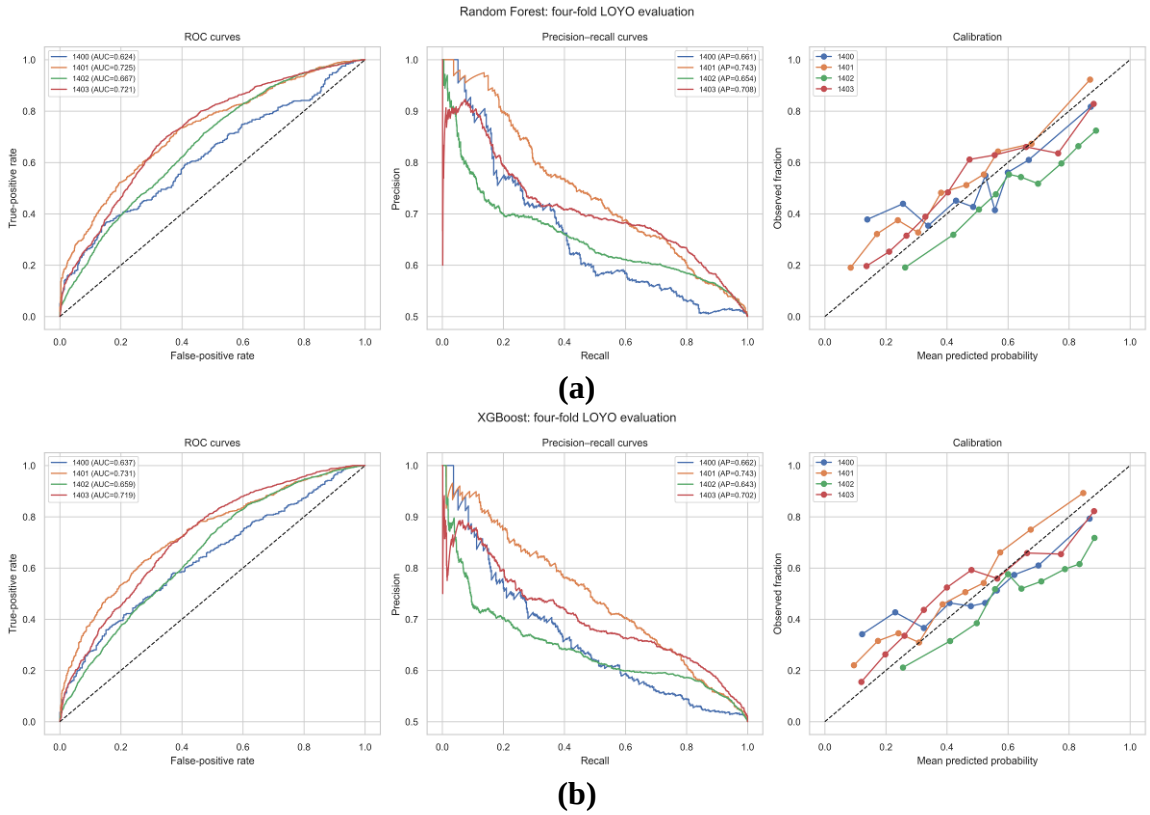

Fig. 2. Four-fold LOYO ROC, precision–recall, and calibration curves for (a) Random Forest and (b) Extreme Gradient Boosting, Sarbaz County, Sistan and Baluchistan Province, southeast of Iran, 1400−1403 Solar Hijri years

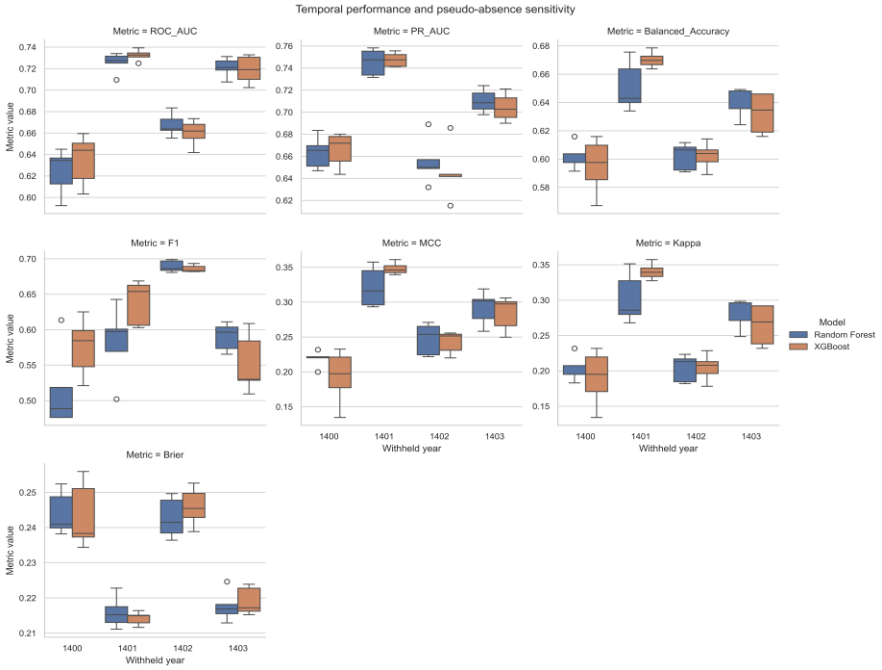

Fig. 3. Temporal performance and sensitivity to pseudo-absence realization, Sarbaz County, Sistan and Baluchistan Province, southeast of Iran, 1400−1403 Solar Hijri years

http://jad.tums.ac.ir
Published Online: June 30, 2026

https://doi.org/10.18502/jad.v20i2.22403

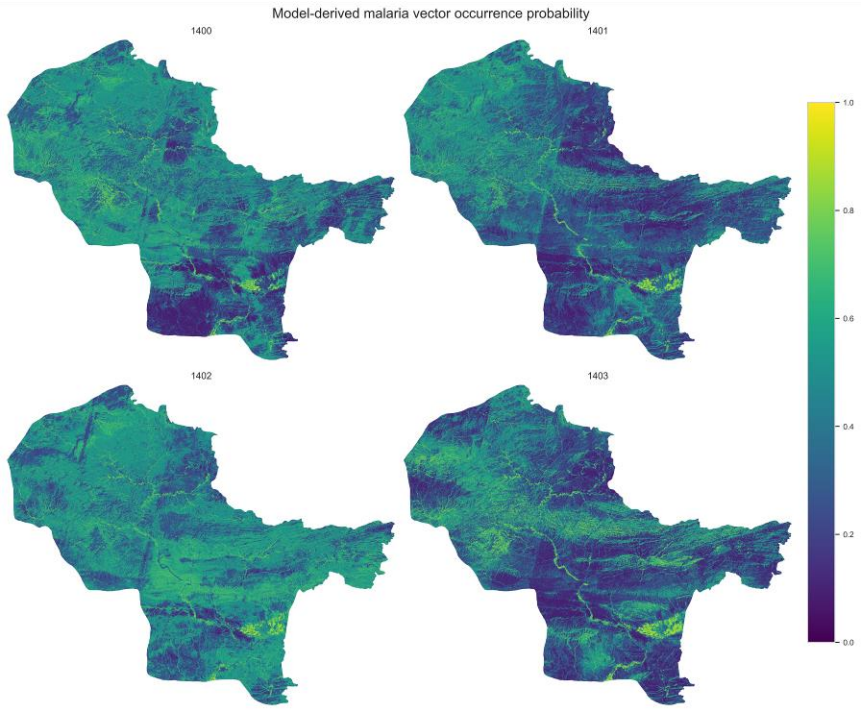

Fig. 4. Annual model-derived malaria case occurrence probabilities for 1400–1403 Solar Hijri years, Sarbaz County, Sistan and Baluchistan Province, southeast of Iran. Values range from 0 to 1 and represent RF ensemble probabilities conditional on recorded cases, soil water index, normalized difference vegetation index, land surface temperature and the pseudo-absence design; they are not direct measurements of transmission intensity

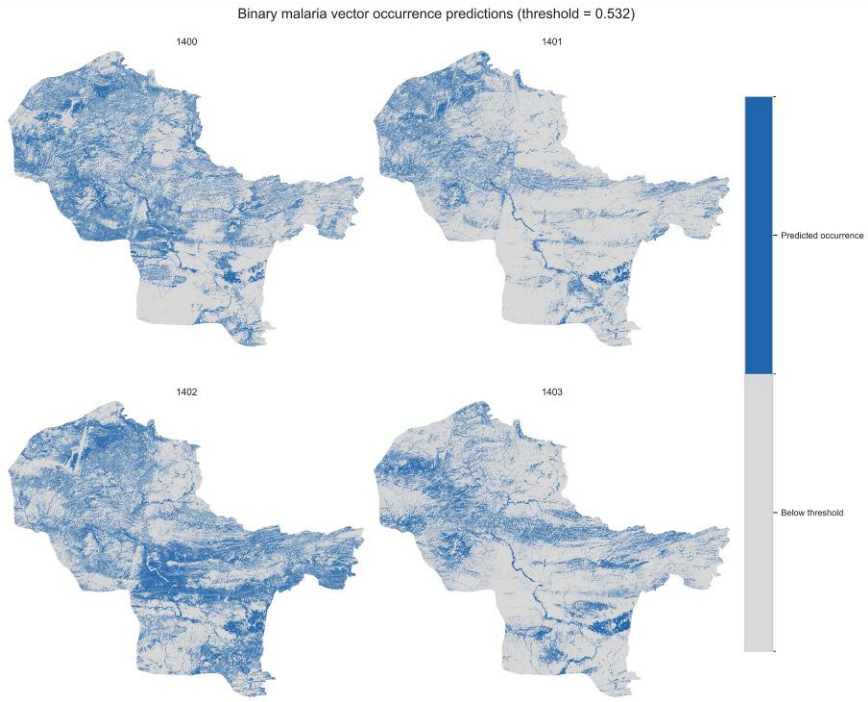

Fig. 5. Binary annual malaria case occurrence predictions obtained with the ensemble threshold of 0.532, Sarbaz County, Sistan and Baluchistan Province, southeast of Iran for 1400–1403 Solar Hijri years. Blue cells indicate predicted occurrence and light-grey cells indicate probabilities below the threshold

http://jad.tums.ac.ir
Published Online: June 30, 2026

https://doi.org/10.18502/jad.v20i2.22403

Table 6. Summary of annual Random Forest occurrence-probability maps, Sarbaz County, Sistan and Baluchistan Province, southeast of Iran, 1400−1403 Solar Hijri years

Year	Mean probability	Median probability	Pixels >0.70 (%)	Threshold-positive (%)	Mean uncertainty
1400	0.426	0.459	5.08	29.86	0.0311
1401	0.355	0.343	3.46	18.17	0.0304
1402	0.466	0.492	5.75	35.61	0.0317
1403	0.352	0.317	6.10	20.96	0.0319

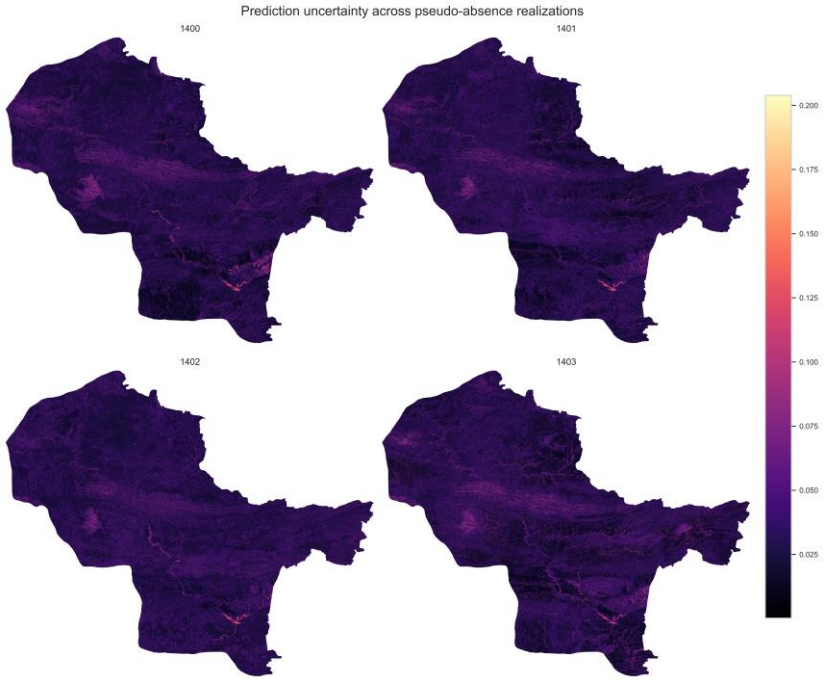

Fig. 6. Pixel-wise standard deviation of Random Forest occurrence probability across five pseudo-absence realizations. Brighter cells indicate greater sensitivity to pseudo-absence selection, Sarbaz County, Sistan and Baluchistan Province, southeast of Iran

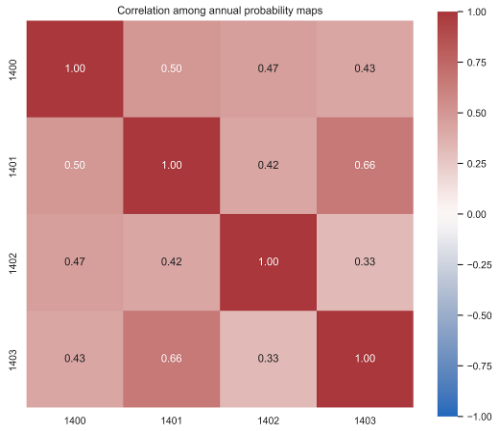

Fig. 7. Pearson correlations among annual Random Forest occurrence-probability maps calculated over the common valid-data footprint, Sarbaz County, Sistan and Baluchistan Province of Iran, 1400–1403 Solar Hijri years. The strongest correspondence occurred between years 1401 and 1403 (r= 0.66) and the weakest between years 1402 and 1403 (r= 0.33)

http://jad.tums.ac.ir
Published Online: June 30, 2026

https://doi.org/10.18502/jad.v20i2.22403

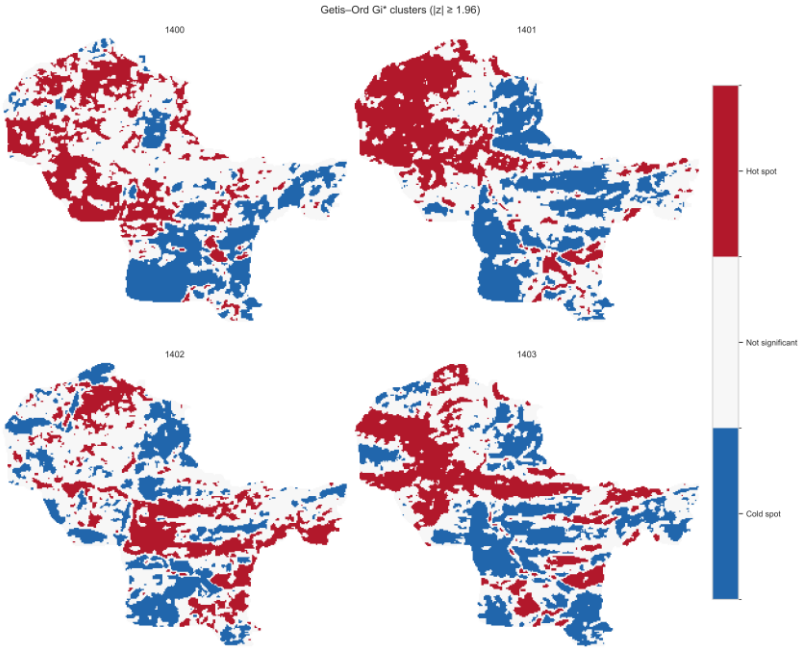

Fig. 8. Getis–Ord Gi* classification of annual Random Forest occurrence-probability maps after 20×20 cell aggregation. Red denotes statistically significant hot spots (z≥ 1.96), blue denotes cold spots (z≤ −1.96), white denotes nonsignificant cells, and areas outside the valid raster footprint are left blank. Sarbaz County, Sistan and Baluchistan Province, southeast of Iran

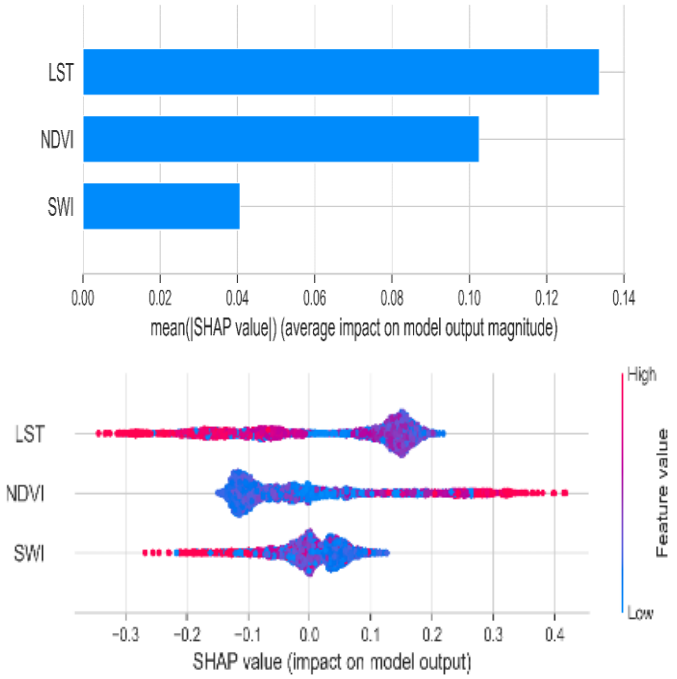

Fig. 9. Global SHapley Additive exPlanations (SHAP) interpretation of the final Random Forest model, Sarbaz County, Sistan and Baluchistan Province, southeast of Iran. The bar plot ranks predictors by mean absolute SHAP value, while the beeswarm plot shows the distribution, direction, and feature-value gradient of their contributions. Positive SHAP values increase the modeled occurrence probability relative to the baseline

http://jad.tums.ac.ir
Published Online: June 30, 2026

https://doi.org/10.18502/jad.v20i2.22403

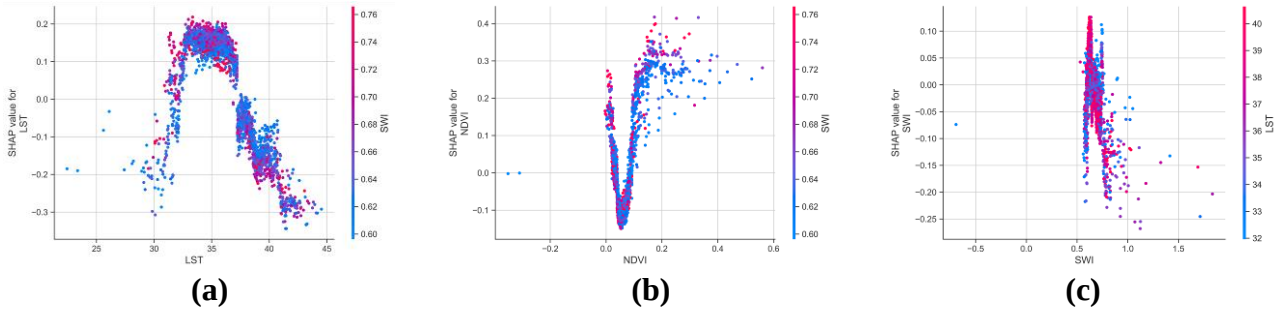

Fig. 10. SHapley Additive exPlanations dependence plots for (a) Land surface temperature, (b) Normalized difference vegetation index and (c) Soil water index, Sarbaz County, Sistan and Baluchistan Province, southeast of Iran. Each point is an evaluated sample; the horizontal axis is the predictor value, and the vertical axis is its SHAP contribution. Point color represents the strongest displayed interacting feature. Apparent response ranges are model-derived associations, not causal thresholds

Discussion

The complete four-fold LOYO analysis provides a stricter picture of temporal generalization than the previously reported pooled or single-year split. Both RF and XGBoost achieved moderate discrimination, with overall ROC-AUC near 0.69 and balanced accuracy near 0.62. Their performance was comparable, but it changed by withheld year. This variation indicates that relationships learned from SWI, NDVI, and LST were partly transferable but not temporally invariant. This degree of performance is consistent with other remote-sensing-based malaria studies in Iran. For example, Rezvan et al. (17) reported ROC-AUC values of 0.70–0.78 for *Anopheles* larval habitat mapping using Sentinel-1 time-series data, while Youssefi et al. (6) achieved an ROC-AUC of approximately 0.75 using optical and radar predictors. The slightly lower performance in the present study may reflect the use of annual rather than seasonal predictors, the coarser temporal resolution, or the greater challenge of predicting human malaria case locations (which depend on additional factors beyond vector habitat) rather than vector presence alone.

The annual composite framework was a deliberate choice necessitated by the temporal aggregation of the supplied occurrence records. Because month- or season-specific dates were unavailable, matching each observation to con temporaneous seasonal predictors was not possible. This limitation is substantial: annual medians suppress seasonal peaks, ephemeral water bodies and lagged ecological responses that are critical for *Anopheles* habitat dynamics in semi-arid environments. The moderate inter-annual correlations observed (r= 0.33–0.66) confirm that each annual map captured distinct environmental conditions rather than serving as a gap-filled reconstruction of other years. This reinforces the importance of year-specific validation and suggests that annual mapping, while imperfect, provides a practical baseline for surveillance planning when higher-temporal-resolution data are unavailable. Future surveillance systems should prioritize recording collection dates to enable seasonal or lagged time-series approaches.

The response variable requires equally careful interpretation. Because the observations were malaria case locations, the continuous raster represents a model-derived relative probability of recorded malaria case occurrence under the available environmental covariates and sampling design. It is not a direct estimate of human malaria incidence, parasite prevalence, vector abundance, or transmission intensity. Human movement, imported infections, parasite reservoirs, intervention coverage, healthcare access, housing, vector-control practices and sur-

http://jad.tums.ac.ir
Published Online: June 30, 2026

https://doi.org/10.18502/jad.v20i2.22403

veillance effort can alter malaria transmission independently of the three satellite predictors. The maps are therefore best used for prioritization and hypothesis generation alongside epidemiological and entomological evidence. This is particularly important in southeastern Iran, where cross-border malaria importation, asymptomatic carriers and drug-resistant parasite strains (7–9) complicate the relationship between environmental suitability and observed cases.

The SHAP findings were ecologically plausible but not causal. The positive LST contribution around 33–36 °C and decreasing contributions under hotter conditions are consistent with nonlinear thermal constraints on mosquito development and survival (2,10,12). Higher NDVI may indicate irrigated vegetation, shade, or more stable microclimates, while SWI may capture aspects of wetness related to aquatic habitat formation. *Anopheles stephensi* is adaptable and can exploit artificial containers, wells, cisterns and irrigation infrastructure (18); these habitats may not be represented fully by 30m SWI, NDVI and LST. The lower importance of SWI should therefore not be interpreted as evidence that water availability is biologically unimportant. Several factors may explain the relatively modest SWI contribution: (i) the 30m spatial resolution is too coarse to detect small artificial containers or narrow drainage channels, (ii) annual median compositing masks short-lived or ephemeral water bodies that appear and disappear within a single season and (iii) *An. stephensi* often exploits human-made water storage structures that are not detectable by radar backscatter. Future studies should consider higher-resolution moisture indicators, such as Sentinel-1 time-series metrics (for example, minimum or variance), or include distance to known water features as a predictor.

The operational utility of the 30m resolution maps warrants consideration for Sarbaz County's specific context. Sarbaz County encompasses mountainous terrain, river valleys, irrigated agriculture and scattered settlements across an arid to semi-arid landscape. The 30 m grid captures landscape-scale environmental variation relevant to village-level and district-level surveillance planning. However, because *An. stephensi* can breed in small artificial containers no larger than a few meters across, these maps should be interpreted as indicating areas of higher environmental receptivity rather than pinpointing individual breeding sites. This scale is appropriate for guiding entomological survey teams to priority areas but insufficient for micro-scale intervention targeting without ground validation. Integration with local knowledge of water-storage practices, irrigation schedules and settlement patterns would enhance operational interpretation.

Annual median compositing reduced cloud contamination, aligned predictors with the annual grouping of the supplied malaria case records and made interannual LOYO feasible. Nevertheless, aggregation suppresses seasonal peaks, short-lived water bodies and lagged responses. Because the source occurrence files contained no month or season, matching each observation to contemporaneous monthly or seasonal predictors was not possible. Future surveillance should record collection date, sampling method, effort, vector species and life stage, coordinate accuracy and locality. Such data would support seasonal or lagged time-series models and separate spring–summer from autumn–winter dynamics.

The predictor set was deliberately parsimonious but incomplete. Elevation, slope, aspect, topographic position, rainfall, distance to rivers and water-storage structures, land use, settlements, population density, road access and intervention data could improve spatial discrimination (16). A DEM was not available in the supplied analysis and could not be added retrospectively without changing the data inputs and rerunning the entire workflow. Topographic variables should be prioritized in future work because they may influence temperature, drainage, accessibility and habitat persistence. The relatively modest MCC and kappa values also suggest substantial unexplained heterogenei-

http://jad.tums.ac.ir
Published Online: June 30, 2026

https://doi.org/10.18502/jad.v20i2.22403

ty, likely arising from missing predictors, incomplete surveillance and the inherent stochasticity of malaria transmission. Given Sarbaz County's border location, including distance to international crossing points or mobility indicators could partially account for imported case dynamics.

Pseudo-absence sampling is another important limitation. A 300m exclusion distance reduces direct overlap with known occurrences but cannot guarantee true non-presence, particularly where surveillance is incomplete. Five independent realizations quantified only one component of sampling uncertainty. The generally low mean uncertainty does not capture uncertainty caused by observation error, missing covariates, temporal aggregation, spatial autocorrelation, or model specification. Spatial autocorrelation in the occurrence records was not explicitly assessed or accounted for in the model evaluation. While the 300m buffer and spatially balanced pseudo-absence sampling reduce spatial overlap, residual spatial clustering may still inflate performance metrics. Future work should consider spatial cross-validation or spatially explicit modeling approaches (19). Moreover, coordinate accuracy was not documented and exact duplicates were collapsed, but unidentified errors may remain.

The annual probability summaries clarify that most of Sarbaz County did not exceed 0.70: only 3.46–6.10% of valid pixels did so. Continuous probabilities, threshold-positive areas, uncertainty surfaces and Gi* clusters have distinct meanings. Gi* hot spots indicate local clusters of high model output, not confirmed transmission hot spots. Field validation is necessary before operational interventions are targeted at specific communities. A pragmatic approach for integrating these maps into Sarbaz County's existing malaria surveillance system would be to: (i) prioritize areas exceeding both the 0.70 probability threshold and containing significant Gi* hot spots for entomological surveys, (ii) treat areas with high uncertainty (pixel-wise standard deviation >0.05) as requiring resurveying or cautious interpretation, and (iii) use the annual map series to identify locations with persistently elevated probabilities across multiple years for long-term sentinel surveillance. This tiered approach acknowledges the maps' role as screening tools rather than definitive diagnostic instruments. The persistence of some spatial structures across years, combined with substantial interannual redistribution, suggests that both stable environmental features (for example, river corridors, irrigated agricultural areas) and year-specific conditions for example, rainfall patterns, temperature anomalies) shape occurrence probabilities. Surveillance planning should therefore consider both the multi-year consensus map and the most recent annual map.

The observed year-dependent performance (ROC-AUC ranging from 0.624 to 0.726) carries practical implications for operational use. Models trained on one set of years may not perform optimally when transferred to a different year, particularly in years with anomalous environmental conditions or surveillance effort. The high sensitivity but low specificity observed for year 1402 suggests that the training-derived threshold labelled a large portion of background samples as occurrence, possibly reflecting environmental conditions in 1402 that differed substantially from the training years. This cautions against deploying a fixed model year after year without recalibration. At minimum, annual updates to the model with the most recent year's occurrence data are recommended, and LOYO-style validation should be repeated as new years of data accumulate to monitor temporal stability.

Conclusion

This study developed a reproducible multi-sensor remote sensing and machine learning framework for annual malaria case occurrence mapping in Sarbaz County using Sentinel-1 SWI, Sentinel-2 NDVI and Landsat LST. Complete four-fold LOYO validation across five

http://jad.tums.ac.ir
Published Online: June 30, 2026

https://doi.org/10.18502/jad.v20i2.22403

pseudo-absence realizations revealed moderate and year-dependent performance for both RF and XGBoost, with overall ROC-AUC values of 0.685±0.045 and 0.687±0.044, respectively and balanced accuracy of approximately 0.62 for both algorithms. Neither model consistently outperformed the other, and performance varied substantially by withheld year, indicating that environmental-vector relationships are not temporally invariant in this semi-arid region.

SHapley Additive exPlanations analysis identified LST as the dominant predictor (48.27% of mean absolute SHAP importance), followed by NDVI (37.04%) and SWI (14.69%). The fitted model showed nonlinear associations, including positive LST contributions around 33–36 °C and increasing contributions at higher NDVI values. Continuous probability maps, binary threshold maps, uncertainty surfaces and Getis–Ord Gi* clusters provided complementary spatial information. Only 3.46–6.10% of valid pixels exceeded a predicted probability of 0.70, while threshold-positive area ranged from 18.17 % to 35.61% across years. Prediction uncertainty attributable to pseudo-absence selection was generally low but spatially heterogeneous.

The maps represent relative probabilities of recorded malaria case occurrence conditional on the available environmental predictors and sampling design. They are not direct estimates of vector abundance, human malaria incidence, or transmission intensity. Their primary utility lies in spatial prioritization for entomological surveys, field verification and vector surveillance when combined with epidemiological and local operational knowledge. Future work should incorporate dated and effort-standardized vector observations, additional predictors (for example, elevation, rainfall, distance to water bodies, land use) and independent field validation. Overall, this framework provides a cautious and transferable basis for integrating Earth observation, temporally explicit validation and uncertainty analysis into local malaria surveillance.

Data and code availability

The analysis configuration, software versions, quality-control tables, LOYO fold definitions, fold-specific predictions and bootstrap confidence intervals, hyperparameter-tuning results, final model settings, SHAP summaries and output manifest are available from the authors with the revised analysis package. Access to precise occurrence coordinates remains subject to the data custodian’s requirements.

Acknowledgements

We want to thank Dr Fatemeh Nikpour for her cooperation in providing information on the distribution of malaria patients in Sarbaz County. This study has been financially supported by the Deputy of Research, Tehran University of Medical Sciences, grant No. 1401-3-211-63286.

Conflicts of interest statement

The authors declare that there is no conflict of interest.

Ethical consideration

This research has been approved by the Ethics Committee of Tehran University of Medical Sciences with the number: IR.TUMS.SPH.REC.1401-219.

References

1. World Health Organization (2023) World malaria report 2023. WHO, Geneva.
2. Mordecai EA, Caldwell JM, Grossman MK, Lippi CA, Johnson LR, Neira M, Rohr JR, Ryan SJ, Savage V, Shocket MS, Sippy R, Stewart Ibarra AM, Thomas MB, Villena O (2019) Thermal biology of mosquito-borne disease. Ecol Lett. 22(10): 1690–1708.

https://doi.org/10.18502/jad.v20i2.22403

3. Bellone R, Failloux AB (2020) The role of temperature in shaping mosquito-borne viruses transmission. Front Microbiol. 11: 584846.
4. Parselia E, Kontoes C, Tsouni A, Hadjichristodoulou C, Kioutsioukis I, Magiorkinis G, Stilianakis NI (2019) Satellite earth observation data in epidemiological modeling of malaria, dengue and West Nile virus: a scoping review. Remote Sens. 11 (16): 1862.
5. Beloconi A, Nyawanda BO, Bigogo G, Khagayi S, Obor D, Danquah I, Kariuki S, Munga S, Vounatsou P (2023) Malaria, climate variability and interventions: modelling transmission dynamics. Sci Rep. 13(1): 7367.
6. Youssefi F, Zoej MJ, Hanafi-Bojd AA, Dariane AB, Khaki M, Safdarinezhad A (2022) Predicting the location of larval habitats of *Anopheles* mosquitoes using remote sensing and soil type data. Int J Appl Earth Obs Geoinf. 108: 102746.
7. Heidari A, Keshavarz H, Rokni MB, Jelinek T (2007) Genetic diversity in merozoite surface protein (MSP)-1 and MSP-2 genes of *Plasmodium falciparum* in a major endemic region of Iran. Korean J Parasitol. 45(1): 59–63.
8. Hassanpour G, Mohebali M, Zeraati H, Raeisi A, Keshavarz H (2017) Asymptomatic malaria and its challenges in the malaria elimination program in Iran: a systematic review. J Arthropod-Borne Dis. 11(2): 172–181.
9. Eskandarian AA, Keshavarz H, Basco LK, Mahboudi F (2002) Do mutations in *Plasmodium falciparum* dihydropteroate synthase and dihydrofolate reductase confer resistance to sulfadoxine-pyrimethamine in Iran?. Trans Roy Soc Trop Med Hyg. 96(1): 96–98.
10. Hanafi-Bojd AA, Vatandoost H, Yaghoobi-Ershadi MR (2020) Climate change and the risk of malaria transmission in Iran. J Med Entomol. 57(1): 50–64.
11. Sinka ME, Pironon S, Massey NC, Longbottom J, Hemingway J, Moyes CL, Willis KJ (2020) A new malaria vector in Africa: Predicting the expansion range of *Anopheles stephensi* and identifying the urban populations at risk. Proc Natl Acad Sci U S A. 117(40): 24900–24908.
12. Abbasi M, Foroushani AR, Jafari-Koshki T, Pakdad K, Vatandoost H, Hanafi-Bojd AA (2020) The impact of climatic variables on the population dynamics of the main malaria vector, *Anopheles stephensi* Liston (Diptera: Culicidae), in southern Iran. Asian Pac J Trop Med. 13(10): 448–455.
13. Zaim M, Parkhideh SZ, Zamani G, Enayati A, Nikpour F, Hanafi-Bojd AA, Sedaghat MM (2026) Historical overview of the bionomics of *Anopheles stephensi* (Diptera: Culicidae) in Iran: insights and implications. J Arthropod Borne Dis. 20(1):1–23.
14. Rezvan H, Valadan Zoej MJ, Hassanpour G, Youssefi F, Hanafi-Bojd AA (2024) Identification of aquatic habitats of *Anopheles* mosquito using time-series analysis of Sentinel-1 data through Google Earth Engine. Int Arch Photogramm Remote Sens Spat Inf. 48: 63–68.
15. Weiss DJ, Dzianach PA, Saddler A, Lubinda J, Browne A, McPhail M, Rumisha SF, Sanna F, Gelaw Y, Kiss JB, Hafsia S (2025) Mapping the global prevalence, incidence and mortality of *Plasmodium falciparum* and *Plasmodium vivax* malaria, 2000–22: a spatial and temporal modelling study. Lancet. 405(10483): 979–990.
16. Guillera-Arroita G, Lahoz-Monfort JJ, Elith J, Gordon A, Kujala H, Lentini PE, McCarthy MA, Tingley R, Wintle BA (2015) Is my species distribution model fit for purpose? Matching data and models to applications. Glob Ecol Biogeogr. 24 (3): 276–292.
17. Rezvan H, Valadan Zoej M, Hassanpour G, Youssefi F, Hanafi-Bojd AA, Ghaderpour E (2026) A two-steps remote sensing and machine learning framework for predicting malaria prone areas by mapping

https://doi.org/10.18502/jad.v20i2.22403

Anopheles larval habitats. Earth Syst Environ. 10: 5977–5998.
18. Barbet-Massin M, Jiguet F, Albert CH, Thuiller W (2012) Selecting pseudo-absences for species distribution models: how, where and how many? Methods Ecol Evol. 3(2): 327–338.
19. Roberts DR, Bahn V, Ciuti S, Boyce MS, Elith J, Guillera-Arroita G, Hauenstein S, Lahoz-Monfort JJ, Schröder B, Thuiller W, Warton DI (2017) Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. Ecography. 40(8): 913–929.
20. Meyer H, Reudenbach C, Hengl T, Katurji M, Nauss T (2019) Improving performance of spatio-temporal machine learning models using forward feature selection and target-oriented validation. Environ Model Softw. 101: 1–9.
21. Lundberg SM, Lee SI (2017) A unified approach to interpreting model predictions. Adv Neural Inf Process Syst. 30: 4765–4774.
22. Molnar C (2020) Interpretable Machine Learning. 2nd Ed. Lulu.com.
23. Getis A, Ord JK (1992) The analysis of spatial association by use of distance statistics. Geogr Anal. 24(3): 189–206.