https://doi.org/10.18502/jad.v20i3.23180

Original Article

Satellite-Derived Urbanization and Vegetation Indices Are Associated with Divergent Seasonal Distributions of *Aedes aegypti* and *Ae. albopictus* Larvae in an Urban Landscape of Central India

Deepanker Das[1,2], Siddhartha Maiti[2], *Devojit Kumar Sarma[1]

[1]ICMR-National Institute for Research in Environmental Health, Bhopal, Madhya Pradesh, India
[2]School of Biosciences, Engineering and Technology, VIT Bhopal University, Madhya Pradesh, India

***Corresponding author:** Dr Devojit Kumar Sarma, E-mail: dkbiotek@gmail.com

(Received 05 Jul 2026; accepted 19 Sep 2026)

Abstract

Background: Urbanization and vegetation strongly shape the ecology of the dengue vectors *Aedes aegypti* and *Ae. albopictus*, yet their fine-scale seasonal distributions in rapidly growing Indian cities remain poorly characterised.

Methods: We integrated six satellite and terrain-derived environmental predictors with longitudinal larval data from 29 sites in Bhopal, central India. Seasonal larval distributions and environmental gradients were represented using inverse distance weighting (IDW) and analysed with exploratory Random Forest (RF) regression and SHAP. Environmental associations were further evaluated using partial least squares regression (PLSR) and LMG hierarchical partitioning, while leave-one-site-out (LOSO) validation assessed the complete IDW-RF spatial reconstruction workflow and Moran's I assessed residual spatial autocorrelation.

Results: Exploratory RF/SHAP identified species- and season-specific patterns. *Aedes aegypti* showed stronger associations with urban-intensity variables, particularly population density and built volume, whereas NDVI variables had the highest relative importance for *Ae. albopictus* across both seasons; dry-season NDVI was also important for *Ae. aegypti*. LOSO-RF showed limited out-of-sample performance (R^2= 0.089 to −0.496). PLSR leave-one-out cross-validation showed modest monsoon performance (Q^2= 0.23–0.29) and weak dry-season performance. LMG supported broad species-level differences in relative predictor contribution, while residual spatial autocorrelation was not significant in three of four PLSR models. Seasonal mapping showed broader monsoon hotspots for *Ae. aegypti* and more localized distributions for *Ae. albopictus*.

Conclusion: Integrating remote sensing, GIS, exploratory RF/SHAP and complementary multivariate methods provides a practical framework for spatially explicit characterisation of species-specific environmental associations that may inform vector surveillance.

Keywords: Dengue vectors; Public health entomology; Breeding habitats; Spatial epidemiology; Geospatial analysis

Introduction

Dengue, chikungunya and Zika are rapidly expanding *Aedes*-borne arboviral diseases and a major public health challenge; dengue contributes the greatest burden, with an estimated 390 million infections and 96 million clinically apparent cases annually worldwide (1). *Aedes aegypti* and *Ae. albopictus*, the principal vectors of these diseases, have shown considerable adaptive capacity to anthropogenic environmental change, and their geographic range has expanded substantially with urbanization, climate variability and increased global connectivity (2, 3).

In India, dengue has increased markedly in recent years, with 289,235 confirmed cases and 485 deaths recorded in 2023, the highest in five years, and Madhya Pradesh accounting for 6,979 confirmed cases over the same period (4). Despite the co-occurrence of these two vector species in the semi-arid urban centres of central India, the fine-scale ecological drivers of their distribution remain poorly under-

http://jad.tums.ac.ir
Published Online: September 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.v20i3.23180

stood. Although often found together, the two species occupy distinct ecological niches: *Ae. aegypti* is closely associated with dense human settlement and artificial containers, whereas *Ae. albopictus* is more ecologically plastic, occupying vegetated and peri-urban habitats with natural breeding sites (5, 6). Most Indian studies, however, have modelled the two species together rather than separately by season (3, 7, 8), even though seasonal dynamics complicate the picture considerably monsoon rainfall enables widespread breeding, while the dry season concentrates larvae in persistent artificial water sources (5). Understanding how these seasonal shifts interact with landscape characteristics such as urban form, vegetation, water retention, and population density is therefore important for targeted vector control (9, 10).

Remote sensing (RS) and geographic information systems (GIS) have become valuable tools for describing the fine-scale spatial distribution of environmental factors affecting mosquito vectors (10). The Normalised Difference Vegetation Index (NDVI), derived from Landsat 8 imagery, has been linked to *Ae. albopictus* abundance through vegetative habitat availability, while Global Human Settlement Layer (GHSL) products quantify the built three-dimensional urban fabric associated with *Ae. aegypti* populations (10). Digital-elevation-based models of water accumulation further help identify topographically favourable breeding areas within cities (11). Integrating these multi-source geospatial predictors with ground-based entomological surveillance allows continuous estimation of larval density across entire urban landscapes, including unsampled areas a methodological advantage over conventional point-based regression (12).

Random Forest (RF) machine learning accommodates complex ecological interactions, non-normal environmental variables, and correlated predictors more flexibly than traditional statistical models of mosquito habitat suitability (13–15), and its feature-importance metrics can help identify landscape factors that merit prioritised surveillance (16). RF models built on remote-sensing predictors have shown strong performance for mosquito larval abundance and habitat suitability at pixel-scale resolution in West Africa (17), but comparable work in Indian urban landscapes with their distinct urban morphology, monsoon-driven hydrolog, and high vector diversity remains rare.

Bhopal, the capital of Madhya Pradesh, is well suited to such analysis as its elevation gradient runs from peri-urban forest to high-density urban core, supporting both *Ae. aegypti* and *Ae. albopictus* (18), and both species are established dengue vectors in the city (8, 19). As Bhopal expands in a patchwork around its periphery while older, denser wards intensify, the resulting landscape mosaic is expected to drive divergent species distributions (1). No study, however, has yet combined pixel-scale geospatial predictors with spatial and multivariate modelling to map the comparative larval ecology of *Ae. aegypti* and *Ae. albopictus* across contrasting seasons in Bhopal or a comparable central Indian city. The present study examines the seasonal spatial ecology of both species using larval survey data from 29 sites and six satellite-derived environmental predictors in Bhopal, with spatial patterns interpolated by IDW and environmental associations assessed through RF/SHAP, PLSR and hierarchical variance partitioning. It also assesses co-occurrence across 11 breeding-habitat types to determine which microhabitats are best suited to *Aedes* colonisation.

Materials and Methods

Study area and sampling design

This study was conducted in Bhopal, the capital of Madhya Pradesh, central India a humid subtropical setting of approximately 285.9 km² at 427–610 m elevation, with a diverse landscape of forest, agriculture, high-density built-up area, hillocks and large water bodies surrounded by residential settlement (18). Summers are hot, the monsoon runs June–September and annual rainfall averages approximate-

https://doi.org/10.18502/jad.v20i3.23180

ly 1,260 mm, with forest-associated and humid conditions prevailing in the surrounding vegetation (20). Bhopal's high population density makes it a suitable setting for examining mosquito responses to both climatic and land-use change (18, 21).

A multistage sampling design captured mosquito occurrence across the major land-cover gradients of the study area. A land-use/ land-cover (LULC) map was first prepared by supervised classification (maximum-likelihood algorithm) of Copernicus Global Land Service layers in ArcGIS 10.7, comprising seven classes: water bodies, built-up area, roads, dense forest, open shrub, cropland and rocks/ barren land. The study area was partitioned into ten equal-area sampling blocks, each containing one to five LULC classes; within each block, three locations were selected to represent local land-cover heterogeneity, yielding 29 georeferenced sites for longitudinal entomological collection. Sampling was conducted monthly from October 2020 to September 2022, with all water-holding containers within a 500 m radius of each site inspected at every visit. Data were aggregated into dry and monsoon seasons, reflecting the pronounced seasonal differences in breeding conditions, container longevity and water availability characteristic of central Indian cities (18, 21).

Larval collection and processing

Larval surveys targeted container habitats and other small breeding sites in and around human habitation, following standard field methods focused on inspecting domestic and peridomestic containers, recording habitat type and collecting larvae for identification (22). Containers were checked thoroughly; larvae were collected by conventional field techniques and reared in the laboratory under a constant 12: 12 h light/dark cycle with a 1 h dusk/dawn transition, 26–28 °C and 70–80% relative humidity (23) until reaching a stage suitable for morphological identification. Species identification used established taxonomic keys for Indian Culicidae (24, 25) and was verified by a second observer for consistency. The primary entomological outcome was container positivity, each inspected container classified as positive or negative for *Aeedes* larvae, irrespective of larval abundance. Identified taxa included *Ae. aegypti*, *Ae. albopictus*, *Ae. vittatus*, *Anopheles*, *Armigeres* and *Culex*. The two *Aedes* vectors were the focus of primary analysis given their established role in local dengue transmission (8).

Seasonal averaging of site data

Monthly positive-container records at each of the 29 sites were aggregated into two seasons: monsoon (June–October) and dry (November–May). Seasonal boundaries were defined using local precipitation data recorded at Ambient Weather stations (Model: WS-2000) installed at five larval collection points (Supplementary Fig. S1), rather than fixed calendar thresholds: the monsoon period ran from the month in which rainfall rose markedly from a near-zero baseline (~0.6 mm in May) through October, when it returned to approximately 2 mm and November onward consistently recorded negligible precipitation and was assigned to the dry season. This hydrometeorological delineation was consistent across both sampling years (2020–2022). For each site and season, the mean of monthly positive-container counts was computed, yielding one seasonal value per site per species reducing short-term sampling noise while preserving ecologically meaningful spatial variation, and capturing the strong seasonality of water availability, container longevity and habitat suitability that characterises urban mosquito breeding in central India.

Geospatial data acquisition

All environmental predictor source layers (Table 1) were derived and processed in Google Earth Engine (GEE) or from DEM-based spatial processing and prepared at a common 100 m resolution in UTM Zone 43N (EPSG:32643), clipped to the Bhopal urban boundary. Predictor values for the 29 sampling locations were

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

https://doi.org/10.18502/jad.v20i3.23180

extracted from these source layers and formed the environmental modelling dataset. Seasonal NDVI composites were generated from Landsat 8 Collection 2 Tier 1 Level 2 imagery (LANDSAT/LC08/C02/T1_L2). NDVI was calculated from the surface-reflectance bands as NDVI= (SR_B5−SR_B4)/(SR_B5+SR_B4), where SR_B5 represents near-infrared reflectance and SR_B4 represents red reflectance. Collection 2 Level-2 surface-reflectance digital numbers were converted to reflectance using the scale factor 0.0000275 and additive offset −0.2 before NDVI calculation. Landsat 8 imagery spanning the 2020–2022 survey period was filtered at the scene level using a cloud-cover threshold of <50% for both the monsoon and extended dry-season composites. Pixels flagged in QA_PIXEL bit 1 (dilated cloud), bit 3 (cloud), or bit 4 (cloud shadow), or flagged by QA_RADSAT, were masked before compositing. The monsoon composite (June–October) and the dry composite (November–May, matching the entomological dry-season definition above) were each built from a per-year extended-median composite. Residual gaps in the monsoon composite were filled using a quality mosaic ranked by the QA-derived "clear_score" variable (Supplementary Methods, Section Quality-mosaic ranking variable), which assigned higher values to pixels without cloud or cloud-shadow flags. The widened acquisition window used for this gap-filling step extended from 1 May to 30 November of each year. Annual composites were averaged across 2020–2022 and downscaled from 30 m to 100 m using a mean reducer with explicit projection declaration prior to resolution reduction. Built height and built volume were sourced from the Global Human Settlement Layer (JRC GHSL P2023A, 100 m; 2018 and 2020 respectively) and population from WorldPop Global High Resolution Population Denominators (100 m, India, 2020). Surface water extent was derived from ESA WorldCover v100 (2020), with permanent water bodies used as a binary mask (nearest-neighbour resampling, to preserve the binary classification) to exclude open-water pixels from the spatial representations and modelling domain. A distinct, continuous water-accumulation predictor (Table 1) was separately derived from a digital elevation model using hydrological flow-direction and flow-accumulation modelling, log-normalised to a 0–1 range; this DEM-derived variable is conceptually and computationally independent of the binary WorldCover exclusion mask, representing terrain-driven runoff-pooling potential rather than remotely sensed permanent-water presence (34, 35). Continuous source layers were resampled to the common 100 m grid using bilinear interpolation, whereas the binary water-exclusion mask used nearest-neighbour resampling throughout (27, 28).

Spatial interpolation of larval surfaces

Seasonal site-level positive-container values for each species, together with the corresponding site-level environmental predictor values used in the modelling dataset, were represented on a common spatial grid using min-max scaling to a 0–1 range prior to interpolation, allowing comparable visualisation and weighting across variables of differing units and ranges. Spatially continuous representations were generated using inverse distance weighting (IDW) interpolation with a power parameter of 2, implemented in Python using NumPy, on a regular 300×300 geographic-coordinate grid in WGS84 (EPSG:4326), spanning the extent of the 29 sampling sites with a 0.01° buffer on all four sides. The grid extended from 23.145584 to 23.311451° N and from 77.327956 to 77.499438° E, with the upper-left origin at 77.327956° E, 23.311451° N and regularly spaced longitude and latitude coordinates at approximately 0.000574° and 0.000555°, respectively. The environmental predictor values were first extracted at the 29 sampling sites from the 100 m UTM Zone 43N (EPSG:32643) source layers and were subsequently spatially represented on the WGS84 interpolation grid using IDW; the 100 m predictor rasters were not directly resampled to the 300×300 grid. The binary water-exclusion

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

https://doi.org/10.18502/jad.v20i3.23180

mask was separately reprojected to the same WGS84 grid using nearest-neighbour resampling and was used to exclude water cells from the spatial representations and modelling domain. A power of 2 was retained as a conventional distance-decay specification that gives greater influence to nearby observations while retaining contributions from more distant sites. The 300×300 grid was selected to provide a continuous spatial representation of the 29 sampling locations for exploratory spatial visualisation. IDW estimated values at unsampled locations as a distance-weighted average of observed site values, with weights decreasing as a function of distance raised to the power of 2 (36, 37). The resulting spatial representations included both the positive-container response surfaces and the site-derived environmental predictor gradients used in the exploratory RF analysis. Because IDW served as an intermediate spatial representation rather than the inferential statistical model, its parameterisation was not treated as an optimisation step for ecological inference. The principal site-level inference was based on the original 29-site dataset using PLSR, LMG hierarchical partitioning, leave-one-site-out validation and "Moran's I" (38).

Random Forest modelling (exploratory spatial analysis)

Random Forest regression models were fitted to the normalised, IDW-spatially represented positive-container response at each valid grid cell using the corresponding spatially represented environmental predictor values at that location, for each of four species–season combinations. Each model used 300 trees, square-root feature sub-sampling at each split (max_features= 'sqrt'), a minimum leaf size of 5, and a fixed random seed of 42 (39, 40). Only land pixels with finite predictor and response values and a positive value for at least one of the four mosquito response surfaces were retained. To provide a deterministic computational upper bound, the valid pixel set was limited to 50,000 pixels when more than 50,000 were available, using a fixed random seed of 42; this cap did not alter the RF hyperparameters. The final RF models were fitted to the selected pixel dataset and evaluated using in-sample R^2, RMSE and MAE. These metrics describe reconstruction of the spatially smoothed IDW surfaces and were not treated as independent predictive performance (Table 2).

Because both the response surface and the spatially represented environmental predictor surfaces were generated from the same 29 sampling locations, predictor and response pixels share the same underlying spatial support, and the resulting near-perfect in-sample R^2 values (Table 2) reflect the model's capacity to reproduce mathematically smooth interpolated surfaces rather than constituting independent, site-level ecological prediction. We therefore treat RF/SHAP results throughout as an exploratory spatial visualisation of predictor–response gradients across the interpolated surface, and ground genuine site-level explanatory inference in the leave-one-site-out validation, PLSR and hierarchical variance-partitioning analyses described below (39, 43).

Predictor importance and interpretation (RF/SHAP)

Feature importance values were extracted from each RF model using Mean Decrease in Impurity (MDI); MDI should be interpreted as a relative contribution to model fit rather than a causal effect size, and may under-represent the collective influence of correlated predictors such as built volume, built height, and population density, which share overlapping urban gradients (43). SHAP (SHapley Additive exPlanations) values were additionally computed using TreeExplainer on a deterministic subset of up to 5,000 observations from the RF modelling dataset to characterise the magnitude and direction of individual predictor contributions (43). Together, MDI and SHAP provide complementary perspectives on predictor ranking and response directionality within the exploratory RF spatial reconstruc-

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

https://doi.org/10.18502/jad.v20i3.23180

tion; the corresponding site-level, cross-validated inference is presented in the following sections.

Leave-one-site-out cross-validation

To assess out-of-sample performance of the complete IDW-RF spatial reconstruction workflow, we conducted leave-one-site-out (LOSO) cross-validation. For each of the 29 sites in turn, predictor and response values were min-max normalised using only the remaining 28 training sites. IDW surfaces were then reconstructed from the 28 training sites alone, with the same water-masking procedure described above, and the RF model was refitted to the resulting grid. The held-out site's response was predicted at its sampling coordinates using predictor values interpolated from the training sites only. Performance was assessed from the pooled held-out predictions using R^2, RMSE, and MAE for each species–season combination (Supplementary Table S1). Because the environmental predictor values at the held-out site were reconstructed from the training sites rather than independently extracted from the original environmental rasters, this LOSO procedure assesses the combined spatial-reconstruction and RF workflow rather than a pure ecological prediction test using independently observed site-level predictors. Partial least squares regression and variable importance.

Because the six environmental predictors are moderately to strongly correlated (variance inflation factors up to 15.7; Supplementary Table S2), ordinary multiple regression coefficients are unstable at n= 29. We therefore used partial least squares regression (PLSR) as the primary site-level explanatory method, since PLSR is suited to small samples with correlated predictors: it extracts latent components that maximise covariance between predictors and response rather than requiring predictor independence. For each of the four species–season models, standardised predictors (z-scored) were evaluated with 1–6 PLSR components using leave-one-out cross-validation (LOO-CV). The number of components yielding the lowest LOO prediction RMSE was selected, and the corresponding pooled cross-validated Q^2 (LOO Q^2) was reported for the selected model. The selected PLSR model was then refit on all 29 sites to obtain variable importance in projection (VIP) scores for each predictor, with VIP > 1.0 indicating above-average importance (Supplementary Table S3). Hierarchical partitioning (LMG) with bootstrap confidence intervals.

As a complementary, non-black-box assessment of relative predictor contribution, we computed exact LMG (Lindeman, Merenda and Gold) hierarchical R^2 decomposition for linear regression across all 2^6 predictor subsets. LMG apportions the model’s explained variance among predictors by averaging their incremental contributions over predictor-entry orders and therefore provides an order-averaged measure of relative contribution when predictors are correlated. Because the allocation depends on the correlation structure among predictors, LMG was interpreted as a descriptive measure of relative importance rather than as evidence of causal precedence. Ninety per cent confidence intervals for each predictor's percentage contribution were obtained by bootstrap resampling of sites (300 replicates) (Supplementary Table S4). Spatial autocorrelation of model residuals (Moran's I)

To assess whether unmeasured spatial structure remained in the site-level PLSR residuals, we computed global Moran's I with inverse-distance spatial weights and a 999-permutation significance test for each of the four species–season models (Supplementary Table S5).

Statistical workflow and software

The overall workflow spanned field entomology, geospatial processing, spatial interpolation, exploratory machine learning and site-level multivariate inference. Monthly site-level positive-container records were first aggregated seasonally. The site-level response and predictor values were then spatially represented using IDW for exploratory RF/SHAP

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

https://doi.org/10.18502/jad.v20i3.23180

analysis, while the original 29-site dataset was analysed independently using PLSR, LMG hierarchical partitioning, leave-one-site-out cross-validation and Moran's I. The main RF/SHAP pipeline was implemented in Python and executed in Google Colab, with deterministic pixel selection, IDW interpolation, RF fitting, feature-importance extraction and SHAP computation documented in the analysis notebooks. The site-level validation and residual spatial-autocorrelation analyses were conducted as complementary analyses. Output products included model performance tables, feature-importance plots, SHAP visualisations, PLSR/VIP and LMG figures, and Moran's I diagnostics. All Python scripts and GEE code will be made available via a public repository upon acceptance; in the interim they are available from the corresponding author upon reasonable request.

Results

The IDW-interpolated positive-container surfaces (Fig. 1a–d), viewed alongside the satellite-derived environmental predictor layers (Fig. 1e–j), reveal clear differences in the seasonal distribution of the two species across Bhopal. *Ae. aegypti* produced broader and more prominent hotspots across the central and western parts of the city during the monsoon (Fig. 1a), with the dry season showing a notable contraction of these high-density areas (Fig. 1c) a pattern that mirrors the spatial footprint of dense urban settlement visible in the built-environment and population layers (Fig. 1h–j). *Ae. albopictus*, by contrast, showed more localized and patchy hotspots in both seasons (Fig. 1b, d), consistent with a species whose distribution is shaped less by broad urban gradients and more by specific local conditions.

Figure 2 shows *Ae. aegypti* recording higher mean positive-container counts than *Ae. albopictus* across most of the 29 sites, with monsoon values consistently exceeding dry-season values across much of the sampling network. *Ae. albopictus* counts, plotted on the inverted axis, were overall lower in magnitude and more variable site-to-site, with a pronounced reduction around sites S08–S09 in both seasons before recovering across the eastern portion of the network. Both species showed a clear monsoon peak relative to the dry season, but the seasonal gap was more pronounced and spatially consistent for *Ae. aegypti*, whereas *Ae. albopictus* responded more unevenly across sites suggesting the two co-occurring species respond to different combinations of local environmental conditions, a question addressed by the modelling framework below.

The Random Forest models fitted to the IDW-spatially represented positive-container surfaces produced extremely high in-sample R^2 values across all four species–season combinations: 0.9997 for *Ae. aegypti* monsoon, 0.9994 for *Ae. albopictus* monsoon, 0.9996 for *Ae. aegypti* dry season, and 0.9994 for *Ae. albopictus* dry season (Table 2). The corresponding in-sample RMSE values were 0.0145, 0.0143, 0.0092 and 0.0094, respectively. As set out in Methods, these values are expected when a flexible RF model reconstructs a mathematically smooth IDW-derived response surface and are therefore not interpreted as independent ecological predictive accuracy. The RF results are retained as exploratory spatial reconstruction and predictor-ranking evidence rather than as validation of generalisation to unsampled sites.

Leave-one-site-out (LOSO) cross-validation was conducted as a complementary site-level assessment of the IDW-RF spatial reconstruction workflow (Supplementary Table S1). The pooled LOSO-RF R^2 values were 0.089 for *Ae. aegypti* during the monsoon, −0.496 for *Ae. albopictus* during the monsoon, −0.218 for *Ae. aegypti* during the dry season, and −0.451 for *Ae. albopictus* during the dry season. The corresponding LOSO RMSE values were 1.579, 1.830, 1.146 and 1.105, respectively, while MAE values were 1.255, 1.226, 0.938 and 0.680, respectively. These results indicate limited out-

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

https://doi.org/10.18502/jad.v20i3.23180

of-sample performance of the complete IDW–RF workflow at the 29 sampling sites. The pooled out-of-fold predictions are provided as an additional CSV data file. Independent of the LOSO-RF analysis, PLSR models were evaluated by LOO-CV for selection of the number of latent components, with one component selected for *Ae. aegypti* during the monsoon and two components for the other three species–season models. The corresponding LOO Q^2 values were 0.231, 0.287, 0.031 and −0.217, respectively (Supplementary Table S3). For *Ae. albopictus*, the two NDVI predictors were consistently the leading variables in both seasons, with VIP values of 1.22 and 1.57 during the monsoon and 1.27 and 1.48 during the dry season. LMG hierarchical partitioning independently converged on the same broad pattern: population density and built volume together accounted for 57.2% of explained variance for *Ae. aegypti* during the monsoon and 45.6% during the dry season, whereas the two NDVI predictors together accounted for 60.8% and 68.5% of explained variance for *Ae. albopictus* during the monsoon and dry seasons, respectively (Supplementary Table S4). This convergence across RF/SHAP, PLSR-VIP and LMG provides convergent evidence for species-level environmental differentiation rather than relying on any single method alone.

Residual spatial autocorrelation in the PLSR models, assessed using Moran's I, was not significant for three of the four models (*Ae. aegypti* monsoon: I= 0.025, p= 0.704, *Ae. aegypti* dry: I= −0.058, p= 0.366, *Ae. albopictus* dry: I= −0.082, p= 0.120), indicating no significant residual spatial autocorrelation detected in these cases. *Ae. albopictus* monsoon showed significant negative autocorrelation (I= −0.175, p= 0.008), consistent with a local, single-site outlier (site S08, whose *Ae. albopictus* count was markedly higher than any other site) rather than a broad unmodelled spatial trend (Supplementary Table S5).

The feature-importance profiles from the exploratory RF analysis (Fig. 3a–d) showed broadly consistent species-level separation but also highlighted seasonal differences. For *Ae. aegypti*, population density was the leading predictor during the monsoon (0.292), followed by built volume (0.200) and water accumulation (0.160). During the dry season, dry-season NDVI (0.238) and population density (0.235) were the two leading predictors, followed by built volume (0.164). For *Ae. albopictus*, NDVI variables had the highest relative importance across both seasons: during the monsoon, dry-season NDVI ranked first (0.336) followed by monsoon NDVI (0.230), while during the dry season dry-season NDVI (0.227) and monsoon NDVI (0.207) ranked first and second. These exploratory RF rankings were interpreted as relative contributions to reconstruction rather than causal effects.

The SHAP beeswarm plots (Fig. 3e–h) broadly supported the RF rankings and added directional context. For *Ae. aegypti*, population and built-volume values showed substantial mean absolute SHAP contributions in both seasons, with dry-season NDVI also contributing strongly during the dry-season model. For *Ae. albopictus*, dry-season and monsoon NDVI produced the largest mean absolute SHAP contributions in both seasons, while built height showed a secondary contribution during the dry season. These patterns were consistent with the VIP and LMG results and support interpretation of the species as occupying different landscape-associated environmental niches. Because SHAP values here describe the exploratory RF reconstruction, they are interpreted as predictor contributions rather than causal effects. These field-scale associations between container positivity and environmental gradients are further contextualised by the observed distribution of breeding container types, examined next.

The container habitat records provide field-level context for the survey (Fig. 4). All positive containers were found in artificial habitats across 13 categories: bottles, construction blocks, discarded containers (clay, metallic,

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

https://doi.org/10.18502/jad.v20i3.23180

and plastic), discarded toys, drums, flowerpots, other, overhead tanks, tires and water storage (in-house and outside). *Ae. aegypti* (n= 296) was most frequently recorded in discarded clay containers, water storage outside, drums, tires and in-house water storage container types characteristic of domestic and peri-domestic water retention in densely settled areas. *Ae. albopictus* (n= 119) showed a broader spread across discarded clay, plastic and metallic containers, water storage outside and tires, reflecting a wider range of container microhabitats. *Ae. aegypti* and *Ae. albopictus* co-occurred within the same container in 82 instances, predominantly in discarded clay containers and outside water storage, showing that the two species can and do share breeding sites despite their divergent landscape-scale environmental associations. Additional multi-species combinations involving *Ae. vittatus*, *Anopheles*, *Armigeres* and *Culex* were recorded at considerably lower frequency. While these patterns reflect observed occurrence rather than confirmed habitat preference total containers inspected per type were not recorded they broadly align with the modelling results, with *Ae. aegypti* linked to high-density domestic settings and *Ae. albopictus* to more heterogeneous breeding environments.

Table 1. Environmental predictor datasets used to derive six site-level environmental variables for the seasonal *Aedes* larval analyses across 29 sampling sites in Bhopal, central India, during the 2020–2022 survey period

Variable	GEE dataset / source	Temporal coverage	Citation
NDVI (Monsoon)	Landsat 8 C2 T1 L2	Jun–Oct composite	(29)
NDVI (Dry)	Landsat 8 C2 T1 L2	Nov–May composite	(29)
Water Accumulation	DEM-derived flow-accumulation model (pysheds; D8 flow direction → accumulation, log-normalised)	Static, derived from SRTM-based DEM	(34, 35)
Built Height	JRC/GHSL/P2023A/GHS_BUILT_H/2018	2018 layer	(31)
Built Volume	JRC/GHSL/P2023A/GHS_BUILT_V/2020	2020 layer	(32)
Population	WorldPop/GP/100m/pop (IND)	2020 layer	(33)

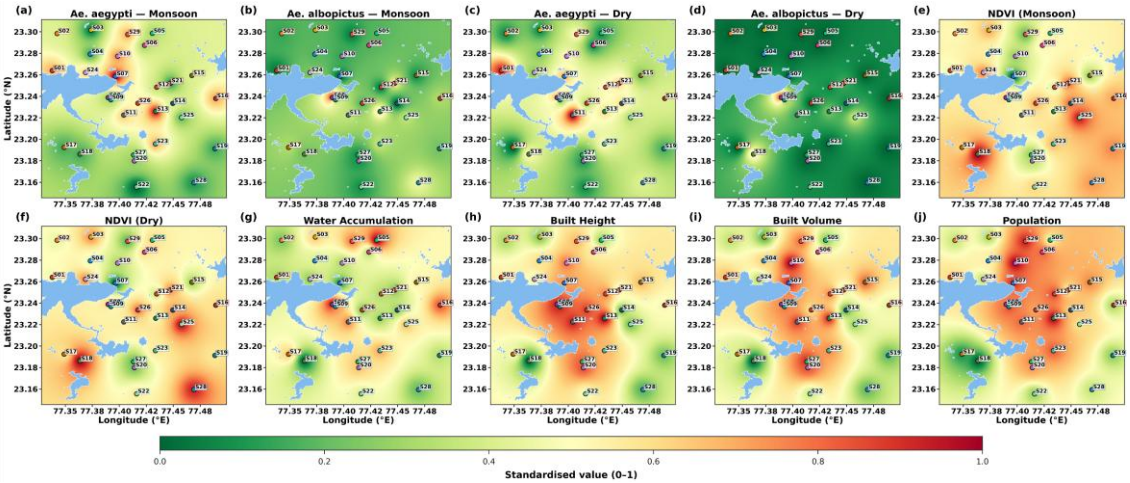


Fig. 1. Inverse Distance Weighted (IDW) spatial representations across 29 sampling sites in Bhopal, with all modelling variables standardised on a 0–1 scale and water masked. IDW-spatially represented seasonal positive-container surfaces for (a) *Aedes aegypti* monsoon, (b) *Ae. albopictus* monsoon, (c) *Ae. aegypti* dry season, and (d) *Ae. albopictus* dry season; and corresponding spatial representations of the environmental predictor gradients derived from the satellite- and DEM-based source layers for (e) NDVI monsoon, (f) NDVI dry season (November–May), (g) water accumulation, (h) built height, (i) built volume, and (j) population density. The source environmental layers were prepared at 100 m resolution in UTM Zone 43N (EPSG:32643) before site-level values were extracted and represented on the common exploratory analysis grid

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

https://doi.org/10.18502/jad.v20i3.23180

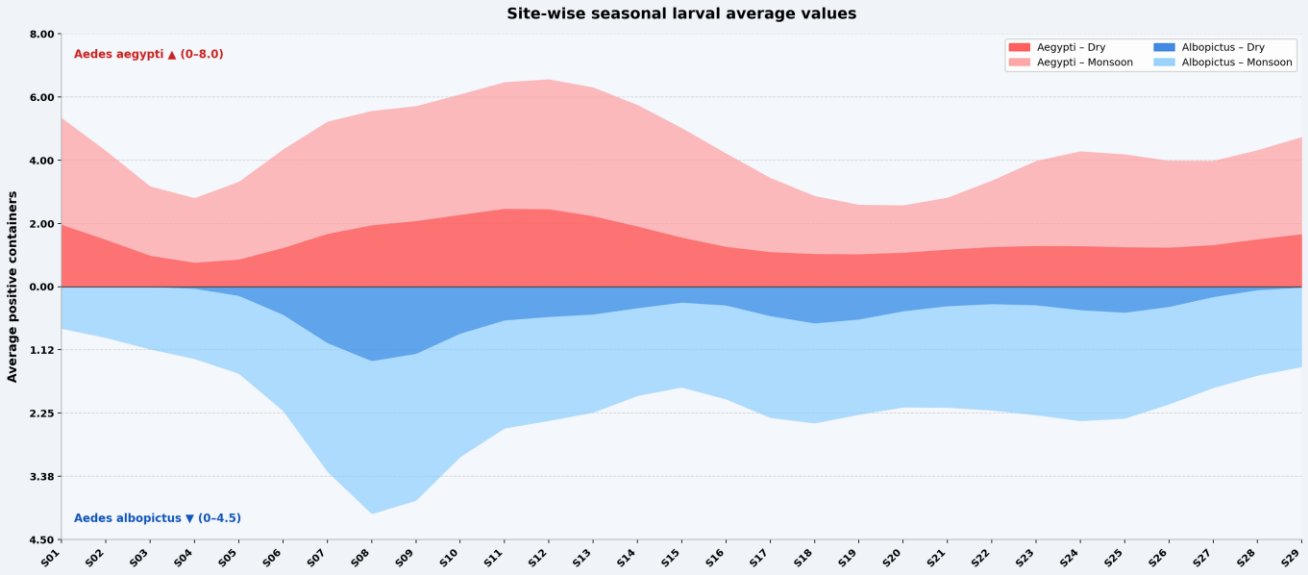


Fig. 2. Site-wise seasonal average of positive containers for *Aedes aegypti* (positive axis) and *Ae. albopictus* (inverted axis) across 29 sampling sites during the dry and monsoon seasons

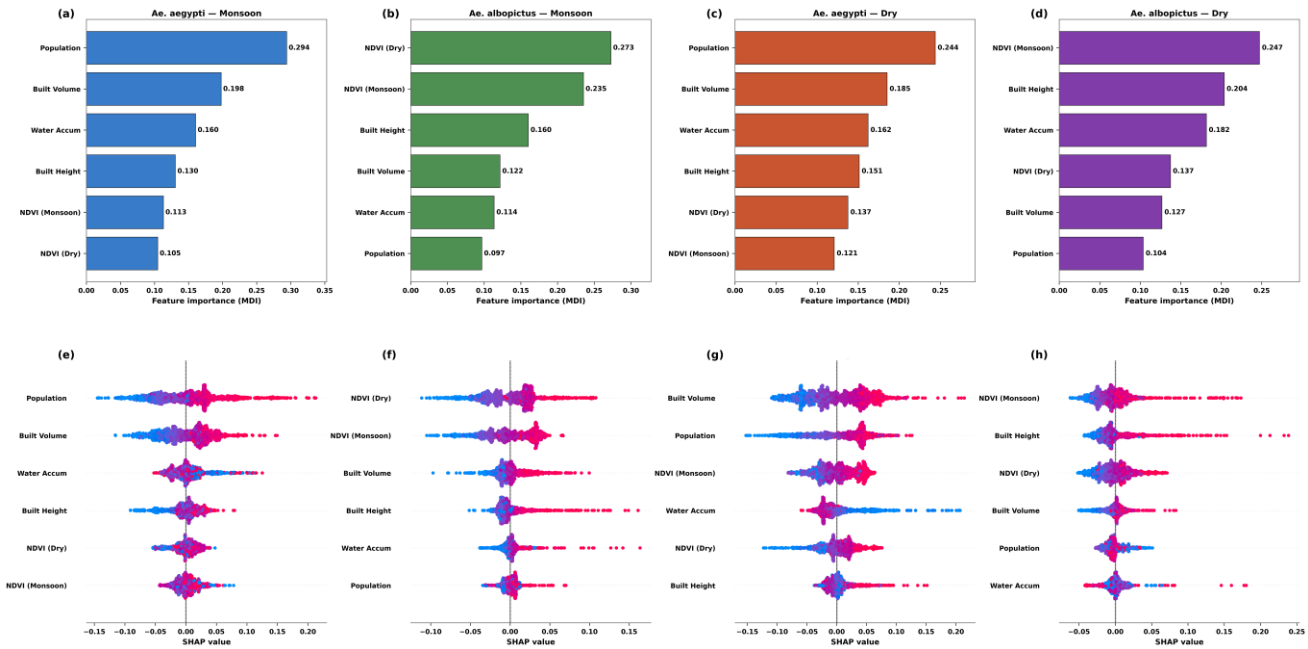


Fig. 3. Random Forest feature importance and SHAP beeswarm plots for environmental predictors of *Aedes* larval density in Bhopal, India (exploratory spatial analysis). The top row shows Mean Decrease in Impurity (MDI) feature importance for (a) *Ae. aegypti* monsoon, (b) *Ae. albopictus* monsoon, (c) *Ae. aegypti* dry season, and (d) *Ae. albopictus* dry season. The bottom row shows corresponding SHAP beeswarm plots for (e) *Ae. aegypti* monsoon, (f) *Ae. albopictus* monsoon, (g) *Ae. aegypti* dry season, and (h) *Ae. albopictus* dry season. See Supplementary Tables S3–S4 for the corresponding PLSR-VIP and LMG hierarchical-partitioning results

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

https://doi.org/10.18502/jad.v20i3.23180

Table 2. Performance metrics of Random Forest regression models for *Aedes aegypti* and *Ae. albopictus* larval positive-container surfaces across monsoon and dry seasons in Bhopal, central India. Metrics are in-sample reconstruction measures from the exploratory RF analysis; see Supplementary Table S1 for leave-one-site-out site-level validation

Model	N pixels	R^2 (in-sample exploratory)	RMSE	MAE
***Ae. aegypti* Monsoon**	50,000	0.9997	0.0145	0.0072
***Ae. albopictus* Monsoon**	50,000	0.9994	0.0143	0.0062
***Ae. aegypti* Dry**	50,000	0.9996	0.0092	0.0045
***Ae. albopictus* Dry**	50,000	0.9994	0.0094	0.0040

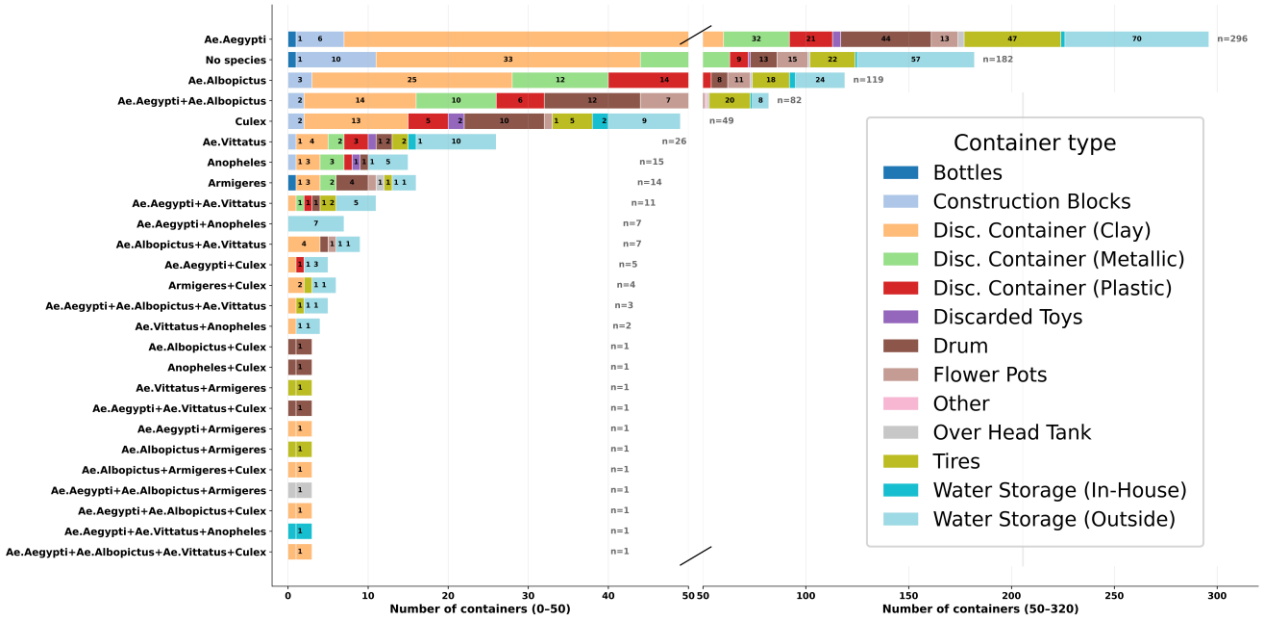


Fig. 4. Distribution and total count of positive breeding containers categorized by specific container types across isolated and co-inhabiting mosquito species, featuring an axis break to accommodate high container volume

Discussion

The results indicate species-specific and seasonal differences in positive-container distribution across Bhopal, with *Ae. aegypti* showing broader spatial coverage across the urban core and *Ae. albopictus* a more localised, patch-dependent pattern (Fig. 1a–d, Fig. 2). The IDW-spatial representations helped translate discrete site-level observations and environmental gradients into continuous visual representations, making it easier to relate container positivity to the surrounding urban landscape; comparable approaches have been used in other site-based entomological studies in Indian urban settings to map larval occurrence gradients (3). IDW produces smoothed spatial trends rather than precise ecological boundaries, and the resulting surfaces should be interpreted accordingly. This smoothing property, together with the shared site-level origin of the response and predictor surfaces, explains the very high in-sample RF R^2 values in Table 2 and reinforces why the RF/SHAP component is treated as exploratory rather than predictive. Site-level analyses therefore provide complementary evidence: LOSO assesses the out-of-sample performance of the IDW–RF spatial reconstruction workflow, while PLSR and LMG provide multivariate assessment of environmental associations and Moran's I evaluates residual spatial structure.

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

https://doi.org/10.18502/jad.v20i3.23180

Both species showed higher positive-container counts during the monsoon than the dry season, consistent with the established relationship between rainfall, water availability and container-breeding *Aedes* activity in urban environments (3, 42). For *Ae. aegypti*, the broader spatial extent of monsoon hotspots across the urban core likely reflects a seasonal expansion in water-filled domestic containers in densely populated areas, rather than rainfall acting as a direct driver in isolation. For *Ae. albopictus*, the persistence of localised clusters across both seasons points toward greater dependence on site-specific microhabitat conditions shaded, vegetated, or structurally sheltered settings that maintain suitable breeding conditions beyond the peak monsoon period (3, 46). These contrasting seasonal responses are consistent with the divergent environmental associations identified across all three modelling approaches and suggest the two co-occurring species respond to seasonal change through different ecological pathways.

The predictor associations converged across the exploratory RF/SHAP analysis, PLSR-VIP and LMG hierarchical partitioning, although the relative ranking varied by season. *Aedes aegypti* was most strongly associated with population density and built volume during the monsoon, while dry-season NDVI became an additional leading predictor during the dry season. *Aedes albopictus* showed the most consistent association with NDVI across both seasons (Fig. 3a–d; Supplementary Tables S3–S4). We deliberately use "important predictors" rather than "dominant drivers" here: because built volume, built height and population density are moderately to strongly correlated (VIF up to 15.7; Supplementary Table S2), no single method including RF/SHAP can cleanly assign causal precedence among them. LMG was therefore used as an order-averaged decomposition of explained variance to describe relative predictor contribution, while recognising that its allocations depend on the correlation structure among predictors. The seasonal convergence across methods suggests a repeatable ecological signal, while the modest cross-validated PLSR Q^2 for monsoon models (0.23–0.29) and weak dry-season Q^2, including a negative value for *Ae. albopictus*, indicate that these associations should not be interpreted as strong predictive validation.

The environmental predictor layers were derived from datasets with different temporal reference periods: NDVI represented seasonal conditions across 2020–2022, whereas built height, built volume, population density, and the WorldCover-derived water mask corresponded to 2018 or 2020 (Table 1). This temporal mismatch may introduce some measurement mismatch when relating environmental characteristics to the longitudinal entomological observations. The results should therefore be interpreted as landscape-scale environmental associations rather than year-specific causal effects. The use of seasonally resolved NDVI together with relatively persistent structural urban variables nevertheless provides a consistent representation of the principal landscape gradients considered in the present analysis.

The role of population density as a leading predictor for *Ae. aegypti* warrants contextual interpretation. In this setting, population density more plausibly operates as a proxy for housing density, water-storage practices, waste accumulation and the overall availability of artificial containers in domestic and peri-domestic environments, rather than as a direct measure of human presence. This interpretation is supported by the container habitat records, where *Ae. aegypti* was most frequently recorded in water storage containers, overhead tanks and drums container types characteristic of densely populated urban households (Fig. 4). Similar associations between urbanisation intensity and *Ae. aegypti* occurrence have been reported across other Indian and tropical urban settings (19, 47). For *Ae. albopictus*, the consistent NDVI signal across both seasons and all three methods is consistent with eco-

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

https://doi.org/10.18502/jad.v20i3.23180

logical niche studies showing positive responses to vegetated, edge and mixed urban-natural habitats (26, 45), echoed at the field level by the species' comparatively broader spread across discarded metallic and plastic containers more commonly encountered in peri-domestic and vegetated microhabitats (Fig. 4). These associations align with the established habitat ecology of both species (48, 49), and comparable frameworks integrating remotely sensed predictors with urban vector surveillance have reported similar species-level environmental separation in other tropical settings (50, 51).

The container habitat data offer additional ecological resolution at the microhabitat scale. Discarded clay containers were the dominant breeding substrate across virtually all species and multi-species combinations, underscoring their systemic importance as larval habitats, likely reflecting their widespread domestic use and surface properties favourable to oviposition. *Aedes aegypti* showed a pronounced association with larger-volume, stable water bodies particularly outside water storage and drums consistent with its documented preference for contained, relatively undisturbed water close to human habitation. *Ae. albopictus* exhibited a comparatively even distribution across container types, with relatively stronger representation in discarded metallic and plastic containers, reflecting broader ecological plasticity. Tires emerged as a consistently productive breeding site across multiple species combinations, reinforcing their recognised role as generalist containers. The co-occurrence of both species in 82 containers, concentrated in discarded clay containers and outside water storage, illustrates that while the two species carry distinct landscape-scale environmental associations, they are not ecologically segregated at the container level. Multi-species combinations involving *Ae. vittatus*, *Anopheles*, *Armigeres* and *Culex* occurred at very low frequency and almost exclusively in discarded clay containers, suggesting opportunistic rather than structured co-breeding.

Overall, the results point toward a landscape-scale ecological partitioning between the two primary vector species in Bhopal, with *Ae. aegypti* more closely associated with densely built and populated urban environments and *Ae. albopictus* with vegetated and structurally heterogeneous settings a distinction that persists across seasons and independent analytical methods, with some seasonal shifts in the relative importance of specific predictors. Given the modest cross-validated Q^2 for monsoon models and the weaker dry-season evidence, the present framework is best understood as an exploratory spatial characterisation of these environmental associations rather than a validated predictive model. Nevertheless, the convergence of the species-level signal across exploratory RF/SHAP, PLSR-VIP, LMG partitioning, two seasons, and field-level container observations provides a useful basis for species-specific surveillance hypotheses in central Indian cities.

Conclusion

This study provides a species-specific, seasonally stratified characterisation of the environmental associations underlying *Aedes* larval ecology in Bhopal, central India. Integrating field-based container surveys, spatial interpolation, exploratory Random Forest/SHAP analysis and complementary multivariate methods (PLSR and LMG hierarchical partitioning) revealed an ecological differentiation between the two primary vector species: *Ae. aegypti* was most closely aligned with urban-intensity gradients involving population density and built volume, with dry-season NDVI also contributing strongly, whereas *Ae. albopictus* showed a consistent association with vegetation-related gradients. These patterns were supported across exploratory and site-level analytical approaches and were corroborated by field-level container habitat observations. They should nevertheless be understood as ecologically indicative rather than predic-

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

https://doi.org/10.18502/jad.v20i3.23180

tively definitive: LOSO assessment of the IDW–RF workflow showed limited site-level out-of-sample performance, whereas PLSR leave-one-out cross-validation yielded Q^2 values of 0.23–0.29 for the monsoon models and weaker performance during the dry season. Accordingly, the present framework is best interpreted as an exploratory spatial characterisation of species-specific environmental associations rather than as a validated predictive model.

The observed divergence in environmental associations suggests that surveillance prioritisation could consider high-density domestic environments for *Ae. aegypti* and vegetated or structurally heterogeneous landscapes for *Ae. albopictus*, but these associations should be evaluated against additional field surveillance before being used for operational targeting. The analytical framework, combining freely available remote-sensing layers with accessible machine-learning and multivariate tools, provides a transferable exploratory approach for cities where fine-scale entomological data remain limited.

Acknowledgements

We extend our gratitude to Mr Kamlesh Mewada and Mr Mahendra Damle for their help in collecting the *Aedes* larvae.

Ethical approval

This study was approved by the Internal Ethics Committee of the ICMR-National Institute for Research in Environmental Health (NIREH), Bhopal (Approval No. NIREH/BPL/IEC/2018-19/3130, dated March 18, 2019).

Conflict of interest statement

The authors declare there is no conflict of interest.

Generative AI statement

The authors declare the use of Claude AI as an assistive tool for developing the Google Earth Engine (GEE) scripts and Google Colab Python code utilized in this study, including the exploratory Random Forest/SHAP pipeline and the subsequent leave-one-site-out validation, PLSR/VIP, LMG hierarchical partitioning and Moran's I analyses developed during peer review. Following the use of this tool, the authors thoroughly reviewed, executed, and edited the code to verify its functionality and scientific validity.

Supplementary information

Supplementary information, including Supplementary Tables S1–S5 (leave-one-site-out validation, predictor variance inflation factors, PLSR variable importance in projection, LMG bootstrap contributions and Moran's I diagnostics) and Supplementary Figures S1–S3, is available in the ESM1 file.

References

1. Bhatt S, Gething PW, Brady OJ, Messina JP, Farlow AW, Moyes CL, Drake JM, Brownstein JS, Hoen AG, Sankoh O, Myers MF, George DB, Jaenisch T, Wint GR, Simmons CP, Scott TW, Farrar JJ, Hay SI (2013) The global distribution and burden of dengue. Nature. 496(7446): 504–507.
2. Laporta GZ, Potter AM, Oliveira JF, Bourke BP, Pecor DB, Linton YM (2023) Global distribution of *Aedes aegypti* and *Aedes albopictus* in a climate change scenario of regional rivalry. Insects. 14(1): 49.
3. Nayak PP, Pai BJ, Govindan S, Babu NN (2025) Influence of climatic and land use factors on post-monsoon distribution of *Aedes* mosquito vectors in Udupi taluk. Sci Rep. 15(1): 36649.

https://doi.org/10.18502/jad.v20i3.23180

4. Ministry of Health and Family Welfare, Government of India (2026) Dengue Situation in India. National Center for Vector Borne Diseases Control (NCVBDC), New Delhi, India. Available at: https://ncvbdc.mohfw.gov.in/index4.php?lang=1&level=0&linkid=431&lid=3715.
5. Delatte H, Gimonneau G, Triboire A, Fontenille D (2009) Influence of temperature on immature development, survival, longevity, fecundity and gonotrophic cycles of *Aedes albopictus*, vector of chikungunya and dengue in the Indian Ocean. J Med Entomol. 46(1): 33–41.
6. Vanlandingham DL, Higgs S, Huang Y-JS (2016) *Aedes albopictus* (Diptera: Culicidae) and mosquito-borne viruses in the United States. J Med Entomol. 53 (5): 1024–1028.
7. Kumar A, Prasad A. (2021) Surveillance of dengue vector, *Aedes aegypti* (L.) mosquitoes in Udaipur District of South Rajasthan (India). Bull Pure Appl Sci Zool. 40A(2): 238–249.
8. Sarma DK, Rathod L, Mishra S, Das D, Agarwal A, Sharma G, Singh TA, Kumawat M, Singh S, Verma V, Kumar M, Shubham S, Tiwari RR, Prakash A. (2023) Molecular surveillance of dengue virus in field-collected *Aedes* mosquitoes from Bhopal, central India: evidence of circulation of a new lineage of serotype 2. Front Microbiol. 14: 1260812.
9. Kraemer MUG, Reiner RC Jr, Brady OJ, Messina JP, Gilbert M, Pigott DM, Yi D, Johnson K, Earl L, Marczak LB, Shirude S, Weaver ND, Bisanzio D, Perkins TA, Lai S, Lu X, Jones P, Coelho GE, Carvalho RG, Van Bortel W, Marsboom C, Hendrickx G, Schaffner F, Moore CG, Nax HH, Bengtsson L, Wetter E, Tatem AJ, Brownstein JS, Smith DL, Lambrechts L, Cauchemez S, Linard C, Faria NR, Pybus OG, Scott TW, Liu Q, Yu H, Wint GRW, Hay SI, Golding N (2019) Past and future spread of the arbovirus vectors *Aedes aegypti* and *Aedes albopictus*. Nat Microbiol. 4(5): 854–863.
10. Roiz D, Ruiz S, Soriguer R, Figuerola J (2015) Landscape effects on the presence, abundance and diversity of mosquitoes in Mediterranean wetlands. PLoS One. 10(6): e0128112.
11. Avramov M, Thaivalappil A, Ludwig A, Miner L, Cullingham CI, Waddell L, Lapen DR (2024) Relationships between water quality and mosquito presence and abundance: a systematic review and meta-analysis. J Med Entomol. 61(1): 1–33.
12. Harsha G, Anish TS, Rajaneesh A, Prasad MK, Mathew R, Mammen PC, Ajin RS, Kuriakose SL (2023) Dengue risk zone mapping of Thiruvananthapuram district, India: a comparison of the AHP and F-AHP methods. GeoJournal. 88(3): 2449–2470.
13. O'Dell N, Bolling BG, Dacko N, Carr JT, Hambrick B, Chaves LF, McMillan JR (2025) Identifying environmental drivers of *Aedes aegypti* and *Aedes albopictus* abundance in the Dallas-Fort Worth metroplex using Random Forest modeling. J Med Entomol. 62(4): 789–799.
14. Breiman L. (2001) Random forests. Mach Learn. 45(1): 5–32.
15. Dormann CF, Elith J, Bacher S, Buchmann C, Carl G, Carré G, García Marquéz JR, Gruber B, Lafourcade B, Leit ã o PJ, Münkemüller T, McClean C, Osborne PE, Reineking B, Schröder B, Skidmore AK, Zurell D, Lautenbach S (2013) Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography. 36(1): 27–46.
16. Lundberg SM, Lee SI (2017) A unified approach to interpreting model predictions. Adv Neural Inf Process Syst. 30.
17. Huang C, Bondre R, Chan R, Liu A, Sangra N (2023) A Random Forest Analysis of Remote Sensing Driven Mosquito Habitat Prediction in West Africa. AGU

https://doi.org/10.18502/jad.v20i3.23180

Fall Meeting, 2023, San Francisco, USA.
18. Government of Madhya Pradesh (2022) Demography, District Bhopal. Government of Madhya Pradesh, India. Available at: https://bhopal.nic.in/en/demography/.
19. Das D, Maiti S, Sarma DK. (2025) A high-resolution GIS and machine learning approach for targeted disease management and localized risk assessment in an urban setup: A case study from Bhopal, Central India. Acta Trop. 267: 107662.
20. Central Ground Water Board (CGWB) (2013) Bhopal District Profile. Ministry of Water Resources, Government of India, Faridabad, India.
21. Bhopal Division, Government of Madhya Pradesh (2022) Demography, Bhopal Division. Government of Madhya Pradesh, India. Available at: https://bhopaldivisionmp.nic.in/en/demography/.
22. Manrique-Saide P (2011) Operational Guide for Assessing the Productivity of *Aedes aegypti* Breeding Sites. World Health Organization, Geneva.
23. Suh E, Grossman MK, Waite JL, Dennington NL, Sherrard-Smith E, Churcher TS, Thomas MB (2020) The influence of feeding behaviour and temperature on the capacity of mosquitoes to transmit malaria. Nat Ecol Evol. 4(7): 940–951.
24. Kumar A, Prasad A (2021) Surveillance of Dengue Vector, *Aedes aegypti* (L.) Mosquitoes in Udaipur District of South Rajasthan (India). Bull Pure Appl Sci Zool. 40A(2): 238–249.
25. Tyagi B, Munirathinam A, Venkatesh A (2015) A catalogue of Indian mosquitoes. Int J Mosq Res. 2(2): 50–97.
26. Westby KM, Adalsteinsson SA, Biro EG, Beckermann AJ, Medley KA (2021) *Aedes albopictus* populations and larval habitat characteristics across the landscape: significant differences exist between urban and rural land use types. Insects. 12(3).
27. Esch T, Brzoska E, Dech S, Leutner B, Palacios-Lopez D, Metz-Marconcini A, Marconcini M, Roth A, Zeidler J (2022) World Settlement Footprint 3D-A first three-dimensional survey of the global building stock. Remote Sens Environ. 270: 112877.
28. Gorelick N, Hancher M, Dixon M, Ilyushchenko S, Thau D, Moore R (2017) Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sens Environ. 202: 18–27.
29. Earth Resources Observation and Science (EROS) Center (2020) Landsat 8-9 Operational Land Imager/Thermal Infrared Sensor Level-2, Collection 2. United States Geological Survey, Reston, USA.
30. Zanaga D, Van De Kerchove R, De Keersmaecker W, Souverijns N, Brockmann C, Quast R, Wewers J, Grosu A, Paccini A, Vergnaud S, Cartus O, Santoro M, Fritz S, Georgieva I, Levis M, Carter S, Herold M, Li L, Tsendbazar N, Ramoino F, Arino O (2021) ESA WorldCover 10 m 2020 v100. Zenodo, European Space Agency. Available at: https://doi.org/10.5281/zenodo.5571936.
31. Pesaresi M, Politis P (2023) GHS-BUILT-H R2023A - GHS Building Height, Derived from AW3D30, SRTM30 and Sentinel2 Composite (2018) European Commission, Joint Research Centre, Ispra, Italy.
32. Pesaresi M, Politis P (2023) GHS-BUILT-V R2023A - GHS Built-up Volume Grids Derived from Joint Assessment of Sentinel2, Landsat and Global DEM Data, Multitemporal (1975–2030). European Commission, Joint Research Centre, Ispra, Italy.
33. WorldPop (2017) India 100m Population. Version 2.0 Estimates for Numbers of People per Pixel (ppp) and People per Hectare (pph), for 2010, 2015 and 2020. School of Geography and Environmental Science, University of Southampton, Southampton, UK.

https://doi.org/10.18502/jad.v20i3.23180

34. Farr TG, Rosen PA, Caro E, Crippen R, Duren R, Hensley S, Kobrick M, Paller M, Rodriguez E, Roth L, Seal D, Shaffer S, Shimada J, Umland J, Werner M, Oskin M, Burband D, Alsdorf D (2007) The Shuttle Radar Topography Mission. Rev Geophys. 45(2): RG2004.
35. Bartos M (2020) pysheds: simple and fast watershed delineation in Python. Zenodo.
36. Lu GY, Wong DW (2008) An adaptive inverse-distance weighting spatial interpolation technique. Comput Geosci. 34 (9): 1044–1055.
37. Wong DW (2016) Interpolation: Inverse-distance weighting. In: International Encyclopedia of Geography: People, the Earth, Environment and Technology (eds Richardson D, Castree N, Goodchild MF, Kobayashi A, Liu W, Marston RA): 1–7.
38. Zarco-Perello S, Simões N (2017) Ordinary kriging vs inverse distance weighting: spatial interpolation of the sessile community of Madagascar reef, Gulf of Mexico. PeerJ. 5: e4078.
39. Valavi R, Elith J, Lahoz-Monfort JJ, Guillera-Arroita G (2021) Modelling species presence-only data with random forests. Ecography. 44(12): 1731–1742.
40. Akselrud CIA (2024) Random forest regression models in ecology: accounting for messy biological data and producing predictions with uncertainty. Fish Res. 280: 107161.
41. Cutler DR, Edwards Jr TC, Beard KH, Cutler A, Hess KT, Gibson J, Lawler JJ (2007) Random forests for classification in ecology. Ecology. 88(11): 2783–2792.
42. Ray S, Tandon N (1999) Breeding habitats and seasonal variation in the larval density of *Aedes aegypti* (L) and *Ae. albopictus* (Skuse) in an urban garden in Calcutta City. Indian J Med Res. 109: 221.
43. Simon SM, Glaum P, Valdovinos FS (2023) Interpreting random forest analysis of ecological models to move from prediction to explanation. Sci Rep. 13 (1): 3881.
44. Getis A, Morrison AC, Gray K, Scott TW (2003) Characteristics of the spatial pattern of the dengue vector, *Aedes aegypti*, in Iquitos, Peru. Am J Trop Med Hyg. 69(5): 494–505.
45. Reiskind MH, Lounibos LP (2009) Effects of intraspecific larval competition on adult longevity in the mosquitoes *Aedes aegypti* and *Aedes albopictus*. Med Vet Entomol. 23(1): 62–68.
46. Rajarethinam J, Ong J, Neo ZW, Ng LC, Aik J (2020) Distribution and seasonal fluctuations of *Ae. aegypti* and *Ae. albopictus* larval and pupae in residential areas in an urban landscape. PLoS Negl Trop Dis. 14(4): e0008209.
47. Knoblauch S, Mukaratirwa RT, Pimenta PFP, de A Rocha AA, Yin MS, Randhawa S, Lautenbach S, Wilder-Smith A, Rocklöv J, Brady OJ, Biljecki F, Dambach P, Jänisch T, Resch B, Haddawy P, Bärnighausen T, Zipf A (2025) Urban *Aedes aegypti* suitability indicators: a study in Rio de Janeiro, Brazil. Lancet Planet Health. 9(4): e264–e273.
48. Yao J, Zhou Z, Liu H, Yao S, Wu J (2026) Analyzing the influence of environment, demographic and socio-economic factors on *Aedes albopictus* (Diptera: Culicidae) mosquito density at the micro-level using XGBoost and SHAP. Parasit Vectors. 19(1): 71.
49. Siddiqui T, Malysheva N, Hartner A-M, Butyrin S, Parreira D, Genger J-W, Irrgang C (2026) Suitable seasons: Global monthly habitat suitability for the arbovirus vectors *Aedes aegypti* and *Aedes albopictus* in 1975–2024. bioRxiv. 2026.04.17.719149.
50. Luan J, Zhang C, Xu B, Xue Y, Ren Y (2020) The predictive performances of

https://doi.org/10.18502/jad.v20i3.23180

random forest models with limited sample size and different species traits. Fish Res. 227: 105534.

51. Georgiades P, Proestos Y, Lelieveld J, Erguler K (2023) Machine learning modeling of *Aedes albopictus* habitat suitability in the 21st century. Insects. 14(5).