SYSTEMATIC REVIEW

# Utilising Mosquito Populations and Machine Learning Algorithms to Predict the Distribution of Mosquito-related Diseases – A Systematic Review

Nur Athen Mohd Hardy Abdullah[1], Nazri Che Dom[1,2], Siti Aekbal Salleh[3], Rahmat Dapari[4,5]

[1] Faculty of Health Sciences, Universiti Teknologi MARA (UiTM), UITM Cawangan Selangor, 42300 Puncak Alam, Selangor, Malaysia
[2] Integrated Mosquito Research Group (I-MeRGe), Universiti Teknologi MARA (UiTM), UITM Cawangan Selangor, 42300 Puncak Alam, Selangor, Malaysia
[3] Institute for Biodiversity and Sustainable Development (IBSD), Universiti Teknologi MARA, 40450, Shah Alam, Selangor, Malaysia
[4] Department of Community Health, Faculty of Medicine and Health Sciences, Universiti Putra Malaysia, Serdang 43400, Malaysia
[5] Integrated Dengue Research and Development, Faculty of Medicine and Health Sciences, Universiti Putra Malaysia, Serdang 43400, Malaysia

## ABSTRACT

**Introduction:** Mosquitoes transmit various diseases and putting over half the global population at risk. The spread of mosquito-borne diseases is increasing, with outbreaks emerging in new and previously controlled areas, driven by factors like climate change and increased human mobility. Therefore, predicting mosquito populations is crucial for public health strategies, given the role of vector abundance in disease outbreaks. This systematic review aims to identify mosquito-related predictors for forecasting diseases or mosquito distribution, as well as to analyses the data sources, outcomes, and machine learning models used in the studies. **Materials and methods:** This review followed the PRISMA guidelines using PubMed and Google Scholar for literature published between January 1, 2015, and July 31, 2024. Literature and systematic reviews, studies without mosquito population data, and association or descriptive studies were excluded from this review. **Results:** Out of 309 studies retrieved from the databases, only 10 were selected for review. The model with the highest performance was Random Forest (RF) (n=3), followed by Artificial Neural Networks (ANN) (n=3). **Conclusion:** The optimal machine learning model was RF (n=3) and has demonstrated the ability to incorporate a wide range of predictors. The main factors affecting the accuracy of forecasting mosquito abundance and disease outbreaks are the selection of machine learning algorithms and their predictors. There were limitations in this study due to the exclusion of preprints and non-English language papers, as well as the use of different predictors and performance indicators in the models, which made it difficult to compare between studies.
*Malaysian Journal of Medicine and Health Sciences* (2025) 21(SUPP7): 183-189. doi:10.47836/mjmhs.21.s7.22

**Keywords:** Systematic review, Machine learning, Mosquito population, Prediction model

**Corresponding Author:**
Nazri Che Dom, PhD
Email: nazricd@uitm.edu.my
Tel : +6019-6143465

## INTRODUCTION

Mosquitoes are vectors of various diseases, including dengue, Zika, malaria, yellow fever, and Japanese encephalitis. Over half of the global population is at risk from these diseases, and the economic impact is significant, with dengue alone costing an estimated US$ 9 billion annually (1,2). The incidence and geographic spread of many mosquito-borne illnesses are thought to be on the rise, with new outbreaks occurring in previously unaffected areas (3,4) and resurging in regions where they had been previously eradicated (5,6). Factors such as climate change, increased transportation, and human mobility are contributing to these trends (7,8).

On the other hand, machine learning is a subset of artificial intelligence that focuses on the development of algorithms, and statistical models that enable computers to perform tasks without being explicitly programmed to do so (9). These predictor model incorporated various inputs using real world and big data to increase the accuracy (10). While most predictive models concentrate on forecasting diseases using environmental, socioeconomic, epidemiological, and demographic factors, few have integrated entomological data. For example, a systematic review by Sylvestre et al. (10) revealed that only 6 out of 119 studies included vector-based data. This scarcity is partly due to the high costs, time demands, and logistical challenges associated with mosquito sampling, particularly in remote areas (11,12).

However, accurately predicting mosquito populations is vital for assisting public health decisions and implementing preventive strategies. Given the crucial role that mosquito abundance in an area plays in triggering disease outbreaks (2,13), forecasting these numbers is essential for thorough risk assessment. Therefore, the first aim of this systematic review was to identify all mosquito population related variables used to forecast disease or mosquito distribution, regardless of the region and population. The second aim was to analyse several features of these studies, such as the data sources, the outcome of study, the machine learning algorithms used and their performance indicators.

## MATERIALS AND METHODS

### Data sources

The systematic review on the usage of mosquito population as predictors or outcomes in machine learning models was performed according to PRISMA guidelines (Moher et al, 2015). The indexed articles for this review were searched in the PubMed and Google Scholar databases between January 1, 2015 and July, 31 2024. The bibliographies of the included articles were also reviewed to identify any new articles that fit the criteria.

### Search Strategy

The search terms were structured according to a PICO (population, intervention, comparator, and outcome) question format. The study focused on mosquito population variables, including adult index, mean adult mosquito captures, total larvae at a site, and oviposition activity. The interventions involved machine learning models like linear regression, Naпve Bayes, K-nearest neighbor (kNN), and decision trees. The outcomes were predictions of disease or mosquito distribution. Therefore, search terms such as "mosquito abundance," "mosquito index," "mosquito population," "machine learning," and "predictor model" were used. These terms were selected based on search strings formulated from the exposure, subject, and outcome of the intended study. Boolean operators like "OR," "AND", and "NOT" were applied to refine or expand the search results from the databases.

### Selection of studies

The articles from online databases were further narrowed down based on specific inclusion and exclusion criteria. The inclusion criteria were: (i) studies that used mosquito abundance variables as predictors or outcomes in machine learning models; (ii) studies that considered any stage of the mosquito life cycle; and (iii) studies that provided clear and precise methods for developing machine learning models. The exclusion criteria were: (i) literature and systematic reviews incorporating previously published data; (ii) studies that did not include mosquito abundance variables in their machine learning models; (iii) notifications of outbreaks, clinical disease descriptions, or identification and classification of mosquito species; (iv) association or descriptive-only studies; and (v) duplicate or inaccessible articles, either due to language barriers or lack of full-text access.

### Data extraction

The initial selection by title and abstract was conducted independently by authors based on the inclusion and exclusion criteria. Articles that presented one or more terms with mosquito abundance and machine learning were included in the title selection. Subsequently, the articles were assessed comprehensively to validate the presence of relevant data usable for the systematic review. The disagreements about article inclusion were resolved through discussion between the first and second authors. The information from selected articles were extracted and compiled in Microsoft® Excel™ Open XML Spreadsheet (XLSX) format to ensure systematic and consistent compilation of data. The descriptive analysis contained information such as the article ID, author and year of publication, study location and category of variables used in study. Meanwhile the variables analysis consisted of article ID, data duration or study period, aims of study, predictor variables, outcome variable and mosquito species. Lastly, the machine learning model analysis consisted of machine learning algorithms, type of model, software used for development of machine learning, pre-processing, and machine learning performance indicators.

## RESULTS

### Studies included in systematic reviews

The initial search across the two databases resulted in 309 records. After screening the titles and abstracts, 261 articles were excluded, leaving 48 articles. An additional 37 articles were discarded because they were descriptive studies, systematic or literature reviews, duplicates, focused on machine learning for mosquito species identification, or had unclear methodologies. Finally, 10 publications met the strict inclusion criteria. Fig. 1 illustrates the article selection process for the systematic review, following PRISMA guidelines.

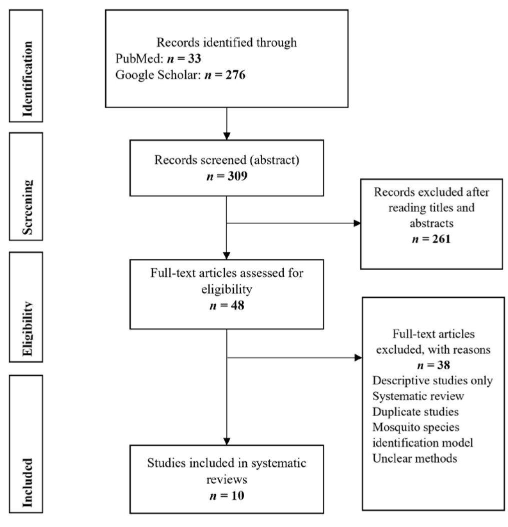


**Fig. 1: Flow chart of articles selection for systematic reviews using PRISMA guidelines**

**Descriptive analysis of selected articles**

The 10 final selected articles were published between 2015 to 2023. Furthermore, the research locations in the articles were conducted in eight different countries. More specifically, the studies were conducted at Korea (n=3), Thailand (n=2), Brazil (n=1), Greece (n=1), Portugal (n=1), USA (n=1) and Vietnam (n=1).

**Variables analysis of selected articles**

The predictor variables in the selected articles were grouped into seven categories. The most common variables were entomological (n=10), followed by meteorological (n=7), environmental (n=6), demographic (n=2), epidemiological (n=2), socioeconomic (n=2), and behavioural (n=1). Furthermore, the majority of the selected studies used mosquito distribution as the outcome in machine learning models (n=9). Other study has dengue morbidity rate as the outcome (n=1). Meanwhile, the mosquito species focused by selected studies were adult Aedes species (n=5), adult Culex species (n=3), generic adult mosquito species (n=2), and Aedes species oviposition activity (n=1).

**Machine learning analysis of selected articles**

The machine learning analysis examined the algorithms applied, the types of models developed, the software used to run these algorithms, and the performance indicators of the models. The most frequently utilised machine learning algorithms were ANN (n=5) and RF (n=5), followed by Support Vector Machine (SVM) (n=4) and kNN (n=1). Other algorithms included Classification and Regression Trees (CART), Generalized Additive Models (GAM), negative binomial regression, gradient boosting, logistic regression, decision tree, and deep learning, each appearing in one study.

The majority of the selected studies used R to develop machine learning models (n=7). Other software included MATLAB, Orange, and Python (each used in one study). The most commonly reported performance indicators were accuracy (n=6), the area under the ROC curve (n=3), and RMSE (n=2). Additional prediction indicators included correlation coefficient (R), Akaike Information Criterion (AIC), mean absolute error (MAE), mean squared error (MSE), and Nash–Sutcliffe model efficiency (NSE), each mentioned in one study. Table I summarised the descriptive, variables and machine learning analysis of selected studies.

**Table I: The descriptive, variables and machine learning analysis of selected studies in systematic reviews**

| Article ID | Descriptive | | Variables | | Machine learning | | |
|---|---|---|---|---|---|---|---|
| | Author (year) | Location of study | Predictors | Outcome | Algorithms | Software | Performance indicator |
| B1 | Kesorn et al. (15) | Nakhon Pathom, Ratchaburi, and Samut Sakhon, Thailand | Demographic<br>Entomological<br>Epidemiological<br>Meteorological | Morbidity rate of dengue | ANN<br>Decision tree<br>kNN<br>SVM:<br>Linear, Polynomial, RBF kernel | - | Accuracy |
| B2 | Kwon et al. (20) | Yeongdeungpo-gu, Seoul, Korea | Entomological<br>Environmental<br>Meteorological | Adult *Culex* species distribution | SVM<br>RF<br>CART | R | Accuracy |
| B3 | Lee et al. (19) | Yeongdeungpo-gu, Seoul, Korea | Entomological<br>Meteorological | Adult *Aedes* species distribution | ANN<br>RF | MATLAB | RMSE<br>Correlation coefficient (R) |
| B4 | Ferreira et al. (14) | Porto Alegre, Brazil | Entomological<br>Meteorological<br>Epidemiological | Generic adult mosquito species distribution | GAM | R | Akaike Information Criterion (AIC) |
| B5 | Chen et al. (16) | North Carolina, USA | Entomological<br>Environmental<br>Socioeconomic | Adult *Aedes* species distribution | ANN<br>SVM<br>kNN | R | Accuracy |
| B6 | Ha et al. (21) | Hanoi, Vietnam | Demographic<br>Entomological<br>Environmental<br>Meteorological | Adult *Culex* species distribution | Negative binomial regression | R | MAE<br>RMSE |
| B7 | Kofidou et al. (22) | Xanthi and Drama, Greece | Entomological<br>Environmental | Generic adult mosquito species distribution | ANN:<br>MLP ANN | R | MSE<br>NSE<br>RMSE |
| B8 | Rahman et al. (23) | Nakhon Pathom, Ratchaburi, and Samut Sakhon, Thailand | Entomological<br>Socioeconomic<br>Behavioural<br>Environmental | Adult *Aedes* species distribution | ANN<br>SVM:<br>Linear, Polynomial, RBF kernel<br>RF<br>kNN<br>Logistic regression | Orange | Accuracy<br>Area under the ROC |
| B9 | Lee et al. (17) | Yeongdeungpo-gu, Seoul, South Korea | Entomological<br>Meteorological<br>Environmental | adult *Culex* species distribution | RF | R | Accuracy<br>Area under the ROC |

CONTINUE

**Table I: The descriptive, variables and machine learning analysis of selected studies in systematic reviews (CONT.)**

| Article ID | Descriptive | | Variables | | Machine learning | | |
|---|---|---|---|---|---|---|---|
| | Author (year) | Location of study | Predictors | Outcome | Algorithms | Software | Performance indicator |
| B10 | Ceia-Hasse et al. (18) | Madeira, Portugal | Entomological<br>Meteorological | *Aedes* species oviposition activity | Classical machine learning:<br>ANN<br>RF<br>Extreme Gradient Boosting Tree<br>Deep learning:<br>Convolutional Neural Networks (CNN)<br>Deep Convolutional Long Short-Term Memory networks<br>Residual Networks<br>Inception Time networks | Python<br>R | Accuracy<br>Area under the ROC |

Best model algorithm and predictor in selected studies

The best-performing machine learning model across the 10 studies was the RF algorithm (n=3). The second-best machine learning model was ANN (n=2). Other machine learning algorithms, such as the GAM, kNN, binomial regression, and Extreme Gradient Boosting Tree, were each found to have the highest predictive power in one study.

The most significant variables for predicting mosquito-borne diseases and mosquito abundance in the selected studies were meteorological variables (n=6). Among these, minimum temperature was the most frequently mentioned (n=3), followed by rainfall (n=2), relative humidity (n=2), maximum temperature (n=1), and LST (n=3). In addition to meteorological factors, entomological (n=3), environmental (n=3), demographic (n=1), and behavioural (n=1) variables were also important. The entomological variables included infection rates in larvae and female mosquitoes (n=1), previous Ae. aegypti vector density (n=1), and the number of mosquito eggs (n=1). Environmental factors involved landscape heterogeneity (n=1), rice cover ratio, forest cover ratio (n=1), and water area (n=1). The demographic variable was human population density (n=1), while the behavioural variable was dengue prevention practices (n=1). Table II summarised the optimal machine learning algorithm, predictors, and model's performance in selected studies.

**Table II: The optimal machine learning algorithm, predictors, and performance of selected articles**

| Article ID | Machine learning | Predictor | Performance |
|---|---|---|---|
| B1 | SVM with RBF kernel | **Entomological:**<br>Infection rate in larvae<br>Infection rate in adult female mosquito | Accuracy = 88.73% |
| B2 | RF | **Meteorological:**<br>Minimum daily temperature (Cluster 1)<br>Maximum daily temperature (Cluster 2&3) | Accuracy:<br>Cluster 1 = 0.80<br>Cluster 2 = 0.71<br>Cluster 3= 0.71 |
| B3 | ANN | Mean relative humidity prior 19 days | R = 0.61<br>RMSE = 14.38 |
| B4 | GAM (M2) | **Entomological:**<br>Previous *Ae. aegypti* vector density<br>**Meteorological:**<br>Minimum temperatures<br>Relative humidity | AIC = 1 628 |
| B5 | kNN | **Environmental:** landscape heterogeneity | Accuracy:<br>Binary = 95%<br>Continuous = 100% |
| B6 | Negative binomial regression | **Demographical:**<br>Human population density variables<br>**Environmental:**<br>Rice cover ratio<br>Forest cover ratio<br>**Meteorological:**<br>Centered rainfall<br>Quadratic term rainfall | RMSE = 24.15<br>MAE = 16.83<br>RMSE = 24.15 |
| B7 | MLP ANN | **Environmental:**<br>LST | MSE = 739<br>NSE = 0.83<br>RMSE = 0.162 |

CONTINUE

**Table II: The optimal machine learning algorithm, predictors, and performance of selected articles (CONT.)**

| Article ID | Machine learning | Predictor | Performance |
|---|---|---|---|
| B8 | RF | **Behaviour:** Dengue prevention practice<br>**Environmental:** Rice crops<br>**Socioeconomic:** House crowding index | Accuracy = 0.86<br>Area under the ROC curve = 0.93 |
| B9 | RF | **Meteorological:**<br>Cumulative minimum temperature over 56 days (Cluster 1)<br>Cumulative rainfall over 60 days (Cluster 1)<br>Cumulative number of rainy days over 43 days (Cluster 1)<br>Cumulative minimum temperature over 13 days (Cluster 2)<br>Cumulative number of rainy days over 60 days (Cluster 2)<br>**Environmental:** Extent of the residential area (Cluster 2) | Accuracy:<br>Cluster 1 = 0.895<br>Cluster 2 = 0.870<br>Area under the ROC curve:<br>Cluster 1 = 0.827<br>Cluster 2 = 0.808 |
| B10 | Deep Convolutional Long Short-Term | **Entomological:** Number of eggs predictor | Accuracy = 0.833 |

## DISCUSSION

Based on the studies reviewed, one of the factors that affect the efficiency of different predictive models in mosquito distribution and disease prediction is the model selection. In the modelling comparison study, several studies have demonstrated that the SVM has higher prediction accuracy compared to ANN (15,16). This is because the NN, when learning too much complexity, can result in overfitting the training set (24). Furthermore, the time lag variables had significantly affected to the performance of the ANN model (14). Nonetheless, ANN show promise in forecasting changes in mosquito populations and may outperform traditional statistical methods (17). Meanwhile, RF has shown to be the optimal model in several studies (20,23). This due to the RF model offers several advantages over other statistical methods, such as high classification accuracy, has method for determining variable importance, and the ability to model complex interactions among predictor variables (25). Therefore, RF provides a powerful alternative to traditional statistical methods for the analysis of ecological data (20). When comparing the performance between deep learning time series classification approach performed to classical machine learning methods by developing model to predict the variation in egg numbers using the same data. The deep learning time series classification demonstrated better prediction accuracy than classical machine learning (18).

The second factor affecting the efficiency of different predictive models in mosquito abundance and disease prediction is the selection of predictors. Many review studies have concluded that environmental and meteorological factors are key predictors in machine learning models. This is because both factors are crucial for predicting mosquito population. For instance, rice crops can cause major land-use changes, and a previous study reported that rice cultivation significantly impacts human health and can sharply increase mosquito abundance and vector-borne diseases (26). Conversely, a 1% increase in forest cover ratio was associated with a 97.5% reduction in mosquito abundance (21). Mosquitoes are unlikely to survive in dense forests due to unfavourable environmental conditions (27). Forests with bushy vegetation can block sunlight, reducing the temperature of small water bodies and thereby preventing larval growth (28). Additionally, meteorological variables such as rainfall between 80 mm and 120 mm are conducive to mosquito growth and breeding (29). While some studies implied that areas with higher precipitation are more conducive to Culex mosquitoes carrying the Japanese encephalitis virus, excessive rainfall can wash away mosquito habitats and larvae (29,30). The prior entomological data as predictors also had a significant impact on the model prediction accuracy (18). This is due to the entomological data is directly associated with mosquito distribution by provides information about the presence, abundance, and behaviour of mosquito populations in specific areas over time.

Socioeconomic variables, such as the house crowding index, are also important predictors, as a higher index provides a suitable environment for human-vector contact, potentially increasing dengue transmission in Thailand and other dengue-endemic countries (23). Conversely, demographic variables like population density have shown that the decrease in mosquito populations in densely populated areas can be attributed to consistent mosquito control measures, enhanced urban planning, and improved sanitation (31,32). Meanwhile, behavioural variables such as dengue prevention practices have the highest contribution to reducing mosquito abundance, likely because the elimination of larval habitats is the simplest and most effective method of managing dengue vector populations (33). Overall, the integration of all predictors enhances model prediction accuracy.

By selecting optimal machine learning algorithms and key predictors of mosquito abundance or disease, public health systems can develop more accurate prediction models for mosquito-borne disease outbreaks. This enables better surveillance and early warning systems, preventing outbreaks from escalating. Additionally, precise predictions of mosquito hotspots or potential disease outbreaks allow for more efficient allocation of resources, such as insecticides, larvicides, and personnel, by concentrating efforts on high-risk areas during peak transmission seasons. Most significantly, accurate forecast on mosquito abundance or disease outbreaks enables healthcare systems to shift from reactive to proactive disease management.

There were several limitations in this study. First, preprints and papers in languages other than English were not included, which may have significantly narrowed the range of studies considered for this review. Second, the selected studies employed a variety of machine learning algorithms, performance metrics, and predictor variables. These differences made it challenging to compare studies directly, limiting the ability to draw robust and consistent conclusions. To improve future research, expanding the range of databases searched and including studies in multiple languages would help broaden the evidence base. Additionally, focusing on one type of machine learning algorithm per study could help identify factors that contribute to variations within that specific algorithm type, leading to more nuanced and precise conclusions.

## CONCLUSION

In summary, the choice of model and predictors can affect the accuracy of mosquito abundance and disease predictions. Random Forest models are generally effective, particularly when incorporating a wide range of predictors. Meteorological and environmental factors are consistently important but integrating socioeconomic and behavioural data can further enhance predictive performance. The study's findings contribute to the development of more accurate prediction models for mosquito abundance and disease outbreaks, which enables more proactive and effective vector control efforts.

## ACKNOWLEDGEMENT

This work is supported by the Fundamental Research Grant Scheme, Ministry of Higher Education Malaysia (600-RMC/FRGS 5/3 (141/2021).

## REFERENCES

1. Mayer SV, Tesh RB, Vasilakis N. The emergence of arthropod-borne viral diseases: A global prospective on dengue, chikungunya and zika fevers. Acta Trop. 2017 Feb;166:155–63. doi: 10.1016/j.actatropica.2016.11.020.
2. Messina JP, Brady OJ, Golding N, Kraemer MUG, Wint GRW, Ray SE, et al. The current and future global distribution and population at risk of dengue. Nat Microbiol. 2019 Jun 10;4(9):1508–15. doi: 10.1038/s41564-019-0476-8.
3. Oliveira S, Rocha J, Sousa CA, Capinha C. Wide and increasing suitability for Aedes albopictus in Europe is congruent across distribution models. Sci Rep. 2021 May 10;11(1). doi: 10.1038/s41598-021-89096-5.
4. Paixro ES, Teixeira MG, Rodrigues LC. Zika, chikungunya and dengue: the causes and threats of new and re-emerging arboviral diseases. BMJ Glob Health. 2017 Dec 1;3(Suppl 1):e000530. doi: 10.1136/bmjgh-2017-000530.
5. Watts N, Amann M, Arnell N, Ayeb-Karlsson S, Belesova K, Berry H, et al. The 2018 report of the Lancet Countdown on health and climate change: shaping the health of nations for centuries to come. Lancet. 2018 Dec;392(10163):2479–514. doi: 10.1016/S0140-6736(18)32594-7.
6. World Health Organization. World malaria report 2017. Geneva: World Health Organization; 2017.
7. Santos JM, Capinha C, Rocha J, Sousa CA. The current and future distribution of the yellow fever mosquito (*Aedes aegypti*) on Madeira Island. Wu JT, editor. PLoS Negl Trop Dis. 2022 Sep 12;16(9):e0010715. doi: 10.1371/journal.pntd.0010715.
8. Ryan SJ, Carlson CJ, Mordecai EA, Johnson LR. Global expansion and redistribution of Aedes-borne virus transmission risk with climate change. Han BA, editor. PLoS Negl Trop Dis. 2019 Mar 28;13(3):e0007213. doi: 10.1371/journal.pntd.0007213.
9. Samuel AL. Some Studies in Machine Learning Using the Game of Checkers. IBM J Res Dev. 1959 Jul;3(3):210–29. doi: 10.1147/rd.33.0210.
10. Sylvestre E, Joachim C, Cécilia-Joseph E, Bouzillé G, Campillo-Gimenez B, Cuggia M, et al. Data-driven methods for dengue prediction and surveillance using real-world and Big Data: A systematic review. Santos VS, editor. PLoS Negl Trop Dis. 2022 Jan 7;16(1):e0010056. doi: 10.1371/journal.pntd.0010056.
11. Ong SQ, Dawood MM, Rahman H, Alias MF, Moideen MA, Lee PC, et al. *A protocol and training guidelines for mosquito sampling in remote areas with limited power supply*. MethodsX. 2024 Jan 9;12:102563–3. doi: 10.1016/j.mex.2024.102563.
12. Coulibaly ZI, Gowelo S, Traore I, Mbewe RB, Ngulube W, Olanga EA, et al. Strengthening adult mosquito surveillance in Africa for disease control: learning from the present. Curr Opin Insect Sci. 2023 Dec:60:101110. doi: 10.1016/j.cois.2023.
13. Li R, Xu L, Bjɯrnstad ON, Liu K, Song T, Chen A, et al. Climate-driven variation in mosquito density predicts the spatiotemporal dynamics of dengue. Proc Natl Acad Sci U S A. 2019 Feb 26;116(9):3624-3629. doi: 10.1073/pnas.1806094116.
14. Lee KY, Chung N, Hwang S. Application of an artificial neural network (ANN) model for predicting mosquito abundances in urban areas. Ecol Inform. 2016 Nov;36:172–80.doi: 10.1016/j.ecoinf.2015.08.011.
15. Ferreira DA da C, Degener CM, de Almeida Marques-Toledo C, Bendati MM, Fetzer LO,

Teixeira CP, et al. Meteorological variables and mosquito monitoring are good predictors for infestation trends of *Aedes aegypti*, the vector of dengue, chikungunya and Zika. Parasit Vector. 2017 Feb 13;10(1):78. doi: 10.1186/s13071-017-2025-8.

16. Chen S, Whiteman A, Li A, Rapp T, Delmelle E, Chen G, et al. An operational machine learning approach to predict mosquito abundance based on socioeconomic and landscape patterns. Landscape Ecol. 2019 May 27;34(6):1295–311. doi: 10.1007/s10980-019-00839-2
17. Lee DS, Lee DY, Park YS. Interpretable machine learning approach to analyze the effects of landscape and meteorological factors on mosquito occurrences in Seoul, South Korea. Environ Sci Pollut Res Int. 2023 Jan;30(1):532-546. doi: 10.1007/s11356-022-22099-5.
18. Ceia-Hasse A, Sousa C, Gouveia BR, Capinha C. Forecasting the abundance of disease vectors with deep learning. Ecol Inform. 2023 Dec 1;78:102272–2. doi: 10.1016/j.ecoinf.2023.102272
19. Kesorn K, Ongruk P, Chompoosri J, Phumee A, Thavara U, Tawatsin A, et al. Morbidity rate prediction of dengue hemorrhagic fever (DHF) using the support vector machine and the *Aedes aegypti* infection rate in similar climates and geographical areas. PLoS One. 2015 May 11;10(5):e0125049. doi: 10.1371/journal.pone.0125049. eCollection 2015.
20. Kwon YS, Bae MJ, Chung N, Lee YR, Hwang S, Kim SA, et al. Modeling Occurrence of urban mosquitos based on land use types and meteorological factors in Korea. Int J Environ Res Public Health. 2015 Oct 20;12(10):13131-47. doi: 10.3390/ijerph121013131.
21. Ha TV, Kim W, Nguyen-Tien T, Lindahl J, Nguyen-Viet H, Thi NQ, et al. Spatial distribution of Culex mosquito abundance and associated risk factors in Hanoi, Vietnam. PLoS Negl Trop Dis. 2021 Jun 21;15(6):e0009497. doi: 10.1371/journal.pntd.0009497.
22. Kofidou M, de Courcy Williams M, Nearchou A, Veletza S, Gemitzi A, Karakasiliotis I. Applying remotely sensed environmental information to model mosquito populations. Sustainability. 2021 Jan 1;13(14):7655. doi: 10.3390/su13147655.
23. Rahman MS, Pientong C, Zafar S, Ekalaksananan T, Paul RE, Haque U, et al. Mapping the spatial distribution of the dengue vector *Aedes aegypti* and predicting its abundance in northeastern Thailand using machine-learning approach. One Health. 2021 Dec 1;13:100358. doi: 10.1016/j.onehlt.2021.100358.
24. Cao LJ, Tay FEH. Support vector machine with adaptive parameters in financial time series forecasting. IEEE Trans Neural Netw. 2003 Nov;14(6):1506–18. doi: 10.1109/TNN.2003.820556.
25. Cutler DR, Edwards TC, Beard KH, Cutler A, Hess KT, Gibson J, et al. Random forest for classification in ecology. Ecology. 2007 Nov;88(11):2783-92. doi: 10.1890/07-0539.1.
26. Diuk-Wasser MA, Sogoba N, Bagayoko M, Taylor CE, Dolo G, Toure MB, et al. Vector abundance and malaria transmission in rice-growing villages in Mali. Am J Trop Med Hyg. 2005 Jun;72(6):725-31.
27. Thongsripong P, Green A, Kittayapong P, Kapan D, Wilcox B, Bennett S. Mosquito vector diversity across habitats in Central Thailand Endemic for dengue and other arthropod-borne diseases. Barrera R, editor. PLoS Negl Trop Dis. 2013 Oct 31;7(10):e2507. doi: 10.1371/journal.pntd.0002507
28. Jacob BG, Muturi EJ, Mwangangi JM, Funes J, Caamano EX, Muriu S, et al. Remote and field level quantification of vegetation covariates for malaria mapping in three rice agro-village complexes in Central Kenya. Int J Health Geogr. 2007 Jun 5:6:21. doi: 10.1186/1476-072X-6-21.
29. Wang L, Hu W, Magalhaes SRJ, Bi P, Ding F, Sun H, et al. The role of environmental factors in the spatial distribution of Japanese encephalitis in mainland China. Environ Int. 2014 Dec:73:1-9. doi: 10.1016/j.envint.2014.07.004.
30. Gubler DJ, Reiter P, Ebi KL, Yap W, Nasci R, Patz JA. Climate variability and change in the United States: potential impacts on vector- and rodent-borne diseases. Environ Health Perspect. 2001 May;109 Suppl 2(Suppl 2):223-33. doi: 10.1289/ehp.109-1240669.
31. Lindsay SW, Wilson A, Golding N, Scott TW, Takken W. Improving the built environment in urban areas to control *Aedes aegypti*-borne diseases. Bull World Health Organ. 2017 Aug 1;95(8):607-608. doi: 10.2471/BLT.16.189688.
32. Cappello V, Marchetti L, Parlanti P, Landi S, Tonazzini I, Cecchini M, et al. Ultrastructural characterization of the lower motor system in a mouse model of krabbe disease. Sci Rep. 2016 Dec;6(1). doi: 10.1038/s41598-016-0001-8.
33. Dhar-Chowdhury P, Haque CE, Lindsay R, Hossain S. Socioeconomic and ecological factors influencing *Aedes aegypti* prevalence, abundance, and distribution in Dhaka, Bangladesh. Am J Trop Med Hyg. 2016 Jun 1;94(6):1223–1233. doi: 10.4269/ajtmh.15-0639.