Modeling Mosquito Vector Migration Under Climate Change

Siba Radwan Nayef Al Jarrah
Department of Data Science
Jordan University of Science and Technology
Irbid, Jordan
sraljarrah22@cit.just.edu.jo

Mohammad Ali Mohammad Othman
Department of Data Science
Jordan University of Science and Technology
Irbid, Jordan
maothman22@cit.just.edu.jo

Dana Amjad Mowafaq Al Naimi
Department of Data Science
Jordan University of Science and Technology
Irbid, Jordan
daalnaimi22@cit.just.edu.jo

Shehab Aldeen Ramadan Hussein Albattah
Department of Data Science
Jordan University of Science and Technology
Irbid, Jordan
sralbattah23@cit.just.edu.jo

Abstract—*Climate change is driving mosquito vectors into new geographic regions by shifting temperature and precipitation patterns. This expansion increases the risk of mosquito-borne diseases in previously unaffected populations. In this project, we processed and integrated four datasets—GBIF mosquito occurrence data, historical temperature data, precipitation data, and disease incidence datasets—to build the foundation for a geospatial modeling pipeline. The work completed includes dataset cleaning, harmonization, ISO3 country code alignment, temporal filtering, reshaping climate tables, and constructing a unified master table combining mosquito activity with climate and disease variables. Preliminary analysis shows clear climate trends and regions exhibiting rising mosquito presence. Future work includes habitat suitability modeling, climate-scenario forecasting, and population exposure estimation.*

Keywords—climate change, mosquito vectors, GBIF, WorldClim, habitat suitability, disease risk.

I. Introduction

Climate change is reshaping ecological conditions worldwide. In particular, mosquito vectors such as Aedes aegypti and Aedes albopictus are expanding into new regions as temperatures rise and precipitation patterns shift. These ecological changes enable mosquitoes to survive in areas that were historically too cold or too dry, posing new public health concerns. Understanding how climate change affects mosquito migration is critical for predicting emerging disease risks and preparing health systems for potential outbreaks.

This research aims to build a data-driven modeling pipeline that predicts future mosquito habitat suitability under climate change scenarios and identifies populations that may become newly at risk.

II. Methods

A. Dataset Collection

The following datasets were collected from publicly accessible sources:

1. *GBIF Mosquito Occurrence Data*

Contains spatial and temporal mosquito sightings, species-level observations, and collection metadata.

2. *WorldClim Historical Temperature Data*

Provides monthly temperature grids for the past decades.

3. *Global Precipitation Dataset*

Includes yearly precipitation averages for all countries.

4. *Disease Incidence & Mortality Data (IHME)*

Includes global disease counts for mosquito-borne illnesses.

B. Data Cleaning and Harmonization

Substantial preprocessing was required to align all datasets:

Removal of duplicates and corrupted entries.

Standardization of country names and conversion to ISO3 codes.

Filtering all datasets to a common temporal range (1980–2023).

Reshaping temperature tables using melt() to convert yearly columns into rows

Assigning missing country codes using PyCountry.

Handling missing temperature/precipitation values and removing non-recorded future months.

Extracting only relevant mosquito fields from the GBIF dataset.

Generating synthetic data due to a lack of consistent historical recording in mosquito dataset

C. Dataset Integration Procedure

A master table was constructed by merging the following layers:

Mosquito activity

Monthly temperature

Annual precipitation

Annual disease incidence per disease

Merge keys used: iso_code, year, month.

This unified table will serve as the basis for modeling mosquito habitat suitability.

III. Current Progre

A. Completed Work

Downloaded and cleaned all major datasets

Unified country identifiers across datasets

Standardized temporal range (1980–2023)

Created wide-format disease tables

Built a merged master dataset

Conducted preliminary climate trend visualizations

Identified missing data, outliers, and structural inconsistencies

Prepared the pipeline for future modeling.

B. Preliminary Observations

- Temperature shows an upward trend globally
- Some regions exhibit increasing mosquito observations.
- Disease incidence matches climate-dependent patterns for known vector species..

C. Next Steps:

Download WorldClim future climate scenarios (2050/2070).

Implement habitat suitability modeling (MaxEnt, Random Forest).

Predict mosquito migration and visualize on heatmaps.

Generate synthetic data for mosquito counts

Add population datasets to estimate exposure and vulnerability..

IV. Expected Outcomes

The final system will generate:

1. Predictive maps showing where mosquito vectors are likely to expand.
2. Climate suitability curves for major mosquito species.
3. Population-at-risk estimates under different climate scenarios
4. Data-driven insights for public health surveillance.
5.

V. Conclusion

SIGNIFICANT PROGRESS HAS BEEN MADE IN COLLECTING, CLEANING, AND INTEGRATING GLOBAL DATASETS RELATED TO MOSQUITO OCCURRENCE, CLIMATE VARIABLES, AND DISEASE INCIDENCE. THE UNIFIED MASTER DATASET PROVIDES A STRONG FOUNDATION FOR ADVANCED MODELING. FUTURE PHASES WILL APPLY GEOSPATIAL AND MACHINE LEARNING TECHNIQUES TO FORECAST MOSQUITO MIGRATION AND IDENTIFY EMERGING PUBLIC HEALTH RISKS DRIVEN BY CLIMATE CHANGE.

Acknowledgment

This research is conducted as part of Group 5 in the Data Science program at Jordan University of Science and Technology.

References

[1] Later………

Identify applicable funding agency here. If none, delete this text box.