SHORT TERM SCIENTIFIC MISSION (STSM)
SCIENTIFIC REPORT

This report is submitted for approval by the STSM applicant to the STSM coordinator

Action number: CA17108 - Aedes Invasive Mosquitoes
STSM title: Training on MosquitoAlert data management and analysis
STSM start and end date: 13/09/2021 to 17/09/2021
Grantee name: Chiara Virgillito

PURPOSE OF THE STSM:

During the Short Term Scientific Mission (STSM) the first objective was to learned from MosquitoAlert developers the management of the app data in order to model MosquitoAlert data by novel spatio-temporal statistical models. The second objective was to learned how extrapolate the data from server for Italian records and how calculate the sampling effort. This STSM allowed me the opportunity to understand type of MosquitoAlert data I need to do the statistical analysis.

DESCRIPTION OF WORK CARRIED OUT DURING THE STSMS

During the first day (13/09/2021) focus was mainly to presented the methodology of analysis the MosquitoAlert data in order to understand the procedure to calculate: a) the propensity score: an estimation from model of reporting activity as function of time elapsed since the participant first downloaded the app and registered with the project 2) sampling effort: the sum of participant sighting in the sampling cell during the time weighted by each participant reporting propensity score at the time of sighting; 3) the human-mosquito encounters as proxy of the probability of at least one reliable report sent from sampling cell during the time giving sampling effort.
During the during next days I learned the Bayesian framework model in order to predict: a) propensity score calculated as the marginal probability randomly drawn participant ever sending a report multiplied by the probability of the participant sending at least one report on a given days conditional on the participant ever sending a report; 2) probability of tiger mosquito eggs presence at ovitrap placed for x days in a sampling cell and how to compare the ovitrap and MosquitoAlert results using ROC curves calculated with pROC package for R. Moreover I participate at Mosquito Alert conference where teams of entomologist, modellers and data management discussed about Entolab system and how to implement the entomological validations.

DESCRIPTION OF THE MAIN RESULTS OBTAINED

During my stay at the University of Pompeu Fabra, Barcelona I believe that I have accomplished the main goal of my Short Term Scientific Mission, which was to get familiar with MosquitoAlert data and the track geolocalization of the participants, the Bayesian framework in order to model the propensity score, sampling effort and human mosquito encounters probabilities. I learned how to explore mosquito alert data by using Hamilton Monte Carlo simulations implemented with Stan and R using the rstanarm package. Moreover, my participation in Mosquito Alert conference was a great experience for me in order to learn: the Entolab process, the identification of reports, the expert validations process, the expert report -based scores, the database and data management and how mosquito alert developers collect anonymous information on the

geographic distribution of its citizen scientist participant in order to correct for biases caused by uneven sampling effort.

FUTURE COLLABORATIONS (if applicable)

During the last days of my STMS we have discussed future analysis to be implemented using Italian reports. The main objectives of these analysis will be:

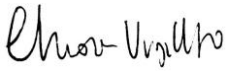
1) Explore the propensity score of Italian users. First, we will model reporting activity as function of time elapsed since the participant first downloaded the app and registered with the project and intrinsic motivation following the approach in Palmer et al., 2017.

2) Calculate sampling effort in Italian grids using data from: a) propensity score in function of time obtain Italy (see model described above) and b) the sum of participant sighting in the sampling cell during time weighted by participants reporting propensity score at the time of sighting.

3) Predict the probability of at least one reliable (confirmed or probable of Aedes albopictus) report sent from sampling cell during the time as function of the counts of eggs per trap/days. We will implement this model using Italian reports from MosquitoAlert app and eggs counted in ovitraps. We will obtain ovitrap data from several monitoring surveillance going in 2021 in Italy: 1) the monitoring database obtained in the frame of AIM cost surveillance in Roma; 2) several other databases obtained during surveillance of alien mosquito in different regions/sites in Italy from collaborators of University of Rome Sapienza.

We will develop a bayesian multilevel logistic regression to estimate what we term “alert probability”, the probability of at least reliable tiger mosquito report being sent through MosquitoAlert from a given geographic cell during a time period conditional on sampling effort. Controlling for sampling will be crucial to making sense of reporting data and sampling effort is itself modelled as a function of time elapsed since the participant downloaded the app as well as intrinsic participant motivation (modelled as random intercept). We will try to implement the model using INLA.

The STSM grantee



Chiara Virgillito
Date: 20/09/2021