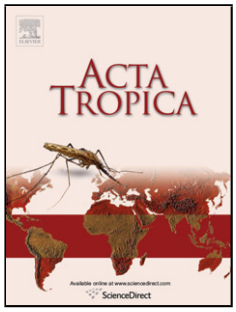

Title: Modeling Dengue Vector Population Using Remotely Sensed Data and Machine Learning

Author: Juan M. Scavuzzo Francisco Trucco Manuel Espinosa Carolina B. Tauro Marcelo Abril Carlos M. Scavuzzo Alejandro C. Frery

PII: S0001-706X(17)31211-1
DOI: https://doi.org/doi:10.1016/j.actatropica.2018.05.003
Reference: ACTROP 4660

To appear in: *Acta Tropica*

Received date: 10-10-2017
Revised date: 3-5-2018
Accepted date: 3-5-2018

Please cite this article as: Juan M. Scavuzzo, Francisco Trucco, Manuel Espinosa, Carolina B. Tauro, Marcelo Abril, Carlos M. Scavuzzo, Alejandro C. Frery, Modeling Dengue Vector Population Using Remotely Sensed Data and Machine Learning, *<![CDATA[Acta Tropica]]>* (2018), https://doi.org/10.1016/j.actatropica.2018.05.003

Modeling Dengue Vector Population Using Remotely Sensed Data and Machine Learning

Juan M. Scavuzzo[a], Francisco Trucco[a], Manuel Espinosa[c], Carolina B. Tauro[b], Marcelo Abril[c], Carlos M. Scavuzzo[b], Alejandro C. Frery[d]

[a]*Facultad de Maremática, Atronomía, Física y Computación, Universidad Nacional de Córdoba*
[b]*Instituto de Altos Estudios Espaciales Mario Gulich, Universidad Nacional de Córdoba-Comisión Nacional de Actividades Espaciales*
[c]*Fundación Mundo Sano, Buenos Aires, Argentina*
[d]*Universidade Federal de Alagoas, Brazil*

Abstract

Mosquitoes are vectors of many human diseases. In particular, *Aedes ægypti* (Linnaeus) is the main vector for Chikungunya, Dengue, and Zika viruses in Latin America and it represents a global threat. Public health policies that aim at combating this vector require dependable and timely information, which is usually expensive to obtain with field campaigns. For this reason, several efforts have been done to use remote sensing due to its reduced cost. The present work includes the temporal modeling of the oviposition activity (measured weekly on 50 ovitraps in a north Argentinean city) of *Aedes ægypti* (Linnaeus), based on time series of data extracted from operational earth observation satellite images. We use are NDVI, NDWI, LST night, LST day and TRMM-GPM rain from 2012 to 2016 as predictive variables. In contrast to previous works which use linear models, we employ Machine Learning techniques using completely accessible open source toolkits. These models have the advantages of being non-parametric and capable of describing nonlinear relationships between variables. Specifically, in addition to two linear approaches, we assess a Support Vector Machine, an Artificial Neural Networks, a K-nearest neighbors and a Decision Tree Regressor. Considerations are made on parameter tuning and the validation and training approach. The results are compared to linear models used in previous works with similar data sets for generating temporal predictive models. These new tools perform better than linear approaches, in particular Nearest Neighbor Regression (KNNR) performs the best. These results provide better alterna-