Reviewer's report

Title: A dynamic neural network model for predicting risk of Zika in real-time

Version: 0 Date: 17 Dec 2018

Reviewer: Stefano Merler

Reviewer's report:

In this paper the authors use a multilayer perceptron classification model to predict the risk of ZIKV transmission in the Americas. For each country, the risk is defined by a binary variable taking two values, namely "low risk" or "high risk", according to whether the country is ranked in the top R% of countries in terms of number of reported cases (or incidence). The model is informed by mosquito suitability maps, socioeconomic and demographic indicators, travel data. The prediction window ranges from 1 to 12 weeks and different threshold values R (from R=10% to R=50%) are considered. Model performances (prediction accuracy and ROC curves) are evaluated on the 15% of the available training material (randomly chosen).

The topic of the paper is clearly interesting - new methods to improve our ability to predict the spread of infectious diseases are more than welcome. I have a few methodological issues:

- I think that the adopted definition of risk is not very adequate in this domain. As far as I understand, the model provides predictions of the ranking, but the risk should be somehow defined as an "absolute" measure. For instance, in case of widespread sustained transmission, all countries should be classified at "high risk", no matter the ranking. For this reason, I think that the definition of risk should be based on case counts or incidence rates. And a binary classification scheme could work in this case as well - it is sufficient to consider a reasonable set of threshold values.

- Since the aim of this study is to show that the proposed model works well, prediction accuracy should be compared to that of other models. The authors report an overall prediction accuracy above 85% for prediction windows of up to 12 weeks. How well (or bad) do simpler models (e.g. linear models) perform? Starting from that, models of increasing complexity should be considered, and different machine learning approaches should be considered as well (e.g. support vector machines, classification trees, etc.).
Performances are evaluated on the 15% of the data. I think that more consistent estimates of the accuracy could be obtained by using a cross-validation scheme (e.g. 10-fold) or bootstrap. Just to make it sure that reported model accuracy does not depend on a "lucky" random choice of the test data.

The spread of Zika can be highly heterogeneous in space, with sustained transmission in certain areas and no transmission at all in other areas. Consequently, a country can be at "low risk" on average (in terms of both reported cases or incidence) but with areas among those at highest risk. So, I think that the spatial resolution of this analysis (just one risk value associated to the entire country) might be not appropriate. This might be relevant for large countries (e.g. Brazil).

Are the methods appropriate and well described?
If not, please specify what is required in your comments to the authors.

No

Does the work include the necessary controls?
If not, please specify which controls are required in your comments to the authors.

Not applicable

Are the conclusions drawn adequately supported by the data shown?
If not, please explain in your comments to the authors.

No

Are you able to assess any statistics in the manuscript or would you recommend an additional statistical review?
If an additional statistical review is recommended, please specify what aspects require further assessment in your comments to the editors.

I am able to assess the statistics

Quality of written English
Please indicate the quality of language in the manuscript:

Acceptable
Declaration of competing interests
Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?
If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

I declare that I have no competing interests

**Statement on potential review bias**
Please complete a statement on potential review bias, considering the following questions:

1. Did you co-author any publication with an author of this manuscript in the last 5 years?

2. Are you currently or recently affiliated at the same institution as an author of this manuscript?

If you can answer no to all of the above, write 'I declare that I did not publish with these authors in the last 5 years and also meet the affiliation criteria". If your reply is yes to any, please give details below.

I declare that I did not publish with these authors in the last 5 years and also meet the affiliation criteria

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal