

Could the recent zika epidemic have been predicted?

Ángel G. Muñoz^{1,2,3*}, Madeleine C. Thomson^{3,4,5}, Anna M. Stewart-Ibarra⁶, Gabriel A. Vecchi^{7,1}, Xandre Chourio⁸, Patricia Nájera⁹, Zelda Moran⁴, Xiaosong Yang²

¹Atmospheric and Oceanic Sciences (AOS). Princeton University. Princeton, NJ. USA

²Geophysical Fluid Dynamics Laboratory (GFDL). Princeton University. Princeton, NJ. USA

³International Research Institute for Climate and Society (IRI). The Earth Institute. Columbia University. New York, NY. USA.

⁴Mailman School of Public Health, Department of Environmental Health Sciences. Columbia University. New York, NY. USA

⁵World Health Organization Collaborating Centre on Early Warning Systems for Malaria and other Climate Sensitive Diseases. Columbia University. New York, NY. USA

⁶Center for Global Health and Translational Science and Department of Medicine. State University of New York Upstate Medical University. Syracuse, NY. USA.

⁷Geosciences Department, and Princeton Environmental Institute (PEI), Princeton University. Princeton, NJ. USA

⁸Latin American Observatory for Climate Events. Centro de Modelado Científico (CMC). Universidad del Zulia. Maracaibo, 4004. Venezuela.

⁹International Health Regulations/Epidemic Alert and Response, and Water Borne Diseases, Communicable Diseases and Health Analysis Department. Pan American Health Organization (PAHO). Washington, DC. USA.

* Correspondence:

Ángel G. Muñoz
agms@princeton.edu

Keywords: zika, *Aedes*-borne diseases, predictability, vector capacity model, climate

Abstract

Given knowledge at the time, the recent 2015-2016 zika virus (ZIKV) epidemic probably could not have been predicted. Without the prior knowledge of ZIKV being already present in South America, and given the lack of understanding of key epidemiologic processes and long-term records of ZIKV cases in the continent, the best related prediction was for potential risk of an *Aedes*-borne disease epidemic. Here we use a recently published two-vector capacity model to assess the predictability of the conditions conducive to epidemics of diseases like zika, chikungunya or dengue, transmitted by the independent or concurrent presence of *Aedes aegypti* and *Aedes albopictus*. We compare the potential risk of transmission forcing the model with the observed climate and with state-of-the-art operational forecasts from the North American Multi Model Ensemble (NMME), finding that the predictive skill of this new seasonal forecast system is highest for multiple countries in Latin America and the Caribbean during the December-February and March-May seasons, and slightly lower –but still of potential use to decision-makers– for the rest of the year. In particular, we find that above-normal *suitable conditions* for the occurrence of the zika epidemic at the beginning of 2015

could have been successfully predicted for several Zika hotspots, and in particular for Northeast Brazil: the heart of the epidemic. Nonetheless, the initiation and spread of an epidemic depends on the effect of multiple factors beyond climate conditions, and thus this type of approach must be considered as a guide and not as a formal predictive tool of vector-borne epidemics.

1 Introduction

Zika virus (ZIKV, family *Flaviviridae*, genus *flavivirus*) disease, a viral illness transmitted primarily by the *Aedes aegypti* and *Aedes albopictus* mosquitoes (1). ZIKV has recently emerged as a major epidemic in Latin America and the Caribbean, with 738,783 suspected and confirmed cases reported to date (2). Prior studies from Yapp Island suggest that the majority of ZIKV infections are asymptomatic or result in mild disease (3), and initial studies from Latin America suggest that the ZIKV infections are less severe and less febrile than chikungunya (CHIKV) or dengue (DENV) infections (4). The spread of ZIKV has been accompanied by severe neurological complications, including children born with microcephaly (5,6) and people with Guillain-Barré syndrome (7,8).

In a previous study (9), our team analyzed the potential contribution of climate signals acting at different timescales in creating the environmental scenario for the current ZIKV epidemic. We found that suitable climate conditions were present, due to the co-occurrence of anomalously high temperatures and persistent below-normal rainfall in several regions of South America, especially in Brazil, the heart of the epidemic.

These suitable conditions are not only favorable for ZIKV, but in general enhance the probability of both *Aedes sp.* reproduction and viral replication. Due to the fact that ZIKV, DENV and CHIKV share the same mosquito vectors and seem to have similar temperature dependence for their extrinsic incubation periods (10), there are advantages in considering the overall eco-epidemiological conditions for the potential risk of transmission of *Aedes*-borne arboviruses rather than focusing on the risk of transmission of only one disease. The effect of rainfall on *Aedes sp.* is more complex than temperature (e.g., (11,12)), because *Aedes* vectors breed in domestic water containers which are more abundant during droughts and water shortages (13). Their presence is also known to increase following unusually high rainfall when peri-domestic breeding sites (discarded containers, flower pots, tires, etc.) are filled with water.

The study of the different environment-virus-vector-human interactions in this field is normally performed using a diversity of mathematical models, most of them based on the Ross-McDonald model (14) or its generalizations. These models are commonly referred to as compartmental models, normally stratifying the human population in susceptible (S), infected (I) and recovered (R) individuals, which explains why they are also known with the alternative name of SIR models. A set of coupled differential equations is used to describe the evolution of each compartment (15,16). These models vary in complexity, and tend to be classified as homogeneous or heterogeneous models; for details, see for example (17).

Although these models are most frequently used to diagnose past or present epidemics, they can also be used in predictive mode, even at seasonal scale (see Thomson et al. (18) and references therein). Predicting environmental suitability conditions presents a complex problem, but it is indeed less complex than predicting the occurrence and transmission of the diseases in human populations. The complexity resides in the non-linear interactions between the different components of the coupled system in consideration, in which the effects of population immunity and susceptibility, or different

possible immunological interactions between the diseases (e.g., co-infections of DENV and ZIKV) are still not well understood.

In this paper, we develop a new seasonal forecast system for suitable climate conditions conducive to enhanced transmission risk of *Ae. aegypti*- and *Ae. albopictus*-borne diseases. We use a two-vector model and state-of-the-art climate forecasts to assess its predictive skill, and we discuss the implications for Latin America and the Caribbean. For brevity, in the following pages we will use “potential risk of transmission” to refer to potential transmission associated with climate conditions suitable for transmission of the diseases mentioned earlier in this paper. Data and general methods are presented in Section 2, the vectorial capacity model is discussed in Section 3, the skill assessment for different seasons of the year is analyzed in Section 4, and the concluding remarks are presented in Section 5.

2 Data and Methods

The domain of study includes Latin America and the Caribbean, and is defined by the boundaries 120°W-25°W and 60°S-32°N.

The observed monthly temperature and rainfall fields for the period 1950-2015 were obtained from the University of East Anglia Climate Research Unit product version 3.24 (CRUv3.24; (19)), available at a horizontal resolution of 0.5 degrees. These datasets were selected to be consistent with our previous study on a similar topic (9). Tests indicated that the results are consistent with other large scale gridded climate datasets, such as the Climate Anomaly Monitoring System (CAMS, Global Historical Climatology Network version 2; see (20) for details) used in (21).

State-of-the-art temperature and rainfall forecasts at monthly timescales were obtained from the North American Multi-Model Ensemble project (NMME; (22)), at a common horizontal resolution of 1 degree. The total of 116 members available was used for the hindcast¹ period of 1982-2010, but only 104 members were used for the December-February 2014-15 forecast, since those were the ones available for that year (no members from the NCAR-CESM1 and NASA-GMAO models). Hindcasts and forecast correspond to the month prior to the target season; for example, for the December-February season, the hindcast and forecast of November was used.

The vector model used in this work was recently developed by Caminade et al. (21). For the sake of organization, the vectorial capacity model equations are presented in the next section. The model requires climate information, and thus the observations and NMME forecasts mentioned above were used, the first one for diagnostics and the second one for the prognostic set up. The model was coded and executed in Matlab at a monthly timescale for a total of 792 months when forcing it with observed data, and 348 months per member when using the NMME hindcasts; each member was run independently before computing the ensemble and seasonal average. The vectorial capacity model output, forced with both climate observations and hindcasts, is available online in the Latin American Observatory’s Datoteca (23–26): http://datoteca.ole2.org/maproom/Sala_de_Salud-Clima/ContexHist-Map-1/index.html.es.

When analyzing the model forced with observations, standardization was performed with respect to the 1950-2015 period. Anomalies are defined as the value of the variable being analyzed minus its

¹ A hindcast is a retrospective forecast, made using the same methodology of actual forecasts, but for a past period of time. They are usually produced to evaluate forecast skill.

1950-2015 average. To analyze inter-annual variability, a 12-month running average was computed. A linear detrending was used.

Skill was assessed using both Kendall's τ and the 2AFC score (27), computed using the International Research Institute for Climate and Society (IRI) Climate Predictability Tool, CPT (28), version 15.4.7. Kendall's τ is a non-parametric rank correlation coefficient used here to measure the overall association between observations and model output. The 2AFC score indicates the probability of correctly discriminating an observation in a higher category from one in a lower (e.g., an "above-normal" observation from a "normal" observation) given the forecasts expressed in deterministic form (i.e., the actual model values, and not the probabilities associated with them). The following four seasons were considered: December-February, DJF, March-May, MAM, June-August, JJA, and September-November, SON. A cross-validation window of five years was used, for the 1982-2010 period. For each iteration, five years were left out and the remainder years were used to build the statistical model, forecasting the middle year of the five-year window. This window is shifted one year into the future for the next iteration, and so on. The skill reported is the average of the metric computed for each iteration, and it was assessed after magnitude and spatial biases were corrected using a simple Model Output Statistics approach involving a Principal Component Regression (PCR; (29,30)), an option available in the CPT software. For further details see (29).

Maps showing the 2AFC score computed using this methodology were produced for each of the seasons considered. Categories for above normal, normal and below normal were identified in the vector model output using the typical 33.33% and 66.66% thresholds in the corresponding probability density function. Forecast probabilities for each category were computed using the PCR model built with the CPT package.

3 Two-vector ento-epidemiological model

Both *Ae. aegypti* and *Ae. albopictus* are considered the most important vectors in Latin America and the Caribbean for the transmission of ZIKV, CHIKV and DENV (e.g., (31–33)). These vectors are known to have different susceptibilities to these diseases, as well as different feeding characteristics (21). While *Ae. aegypti* and *Ae. albopictus* are considered a domestic mosquito and peri-domestic mosquito respectively, it is in principle possible to find them co-existing in the same place (34,35), something that is expected to be even more common in the near future due to global warming (31,36). Hence, we consider that an actionable seasonal forecast system like the one in consideration should involve at least these two species for Latin America and the Caribbean. This section presents the model equations used by the prediction system.

As it has been shown by other authors (21,37) the equations for the dynamics of a two-vector-one-host capacity model, a generalization of the well-known Ross-McDonald model (14), are

$$\frac{dS_H}{dt} = -\lambda_H S_H \quad (1)$$

$$\frac{dI_H}{dt} = \lambda_H S_H - r I_H \quad (2)$$

$$\frac{dR_H}{dt} = r I_H \quad (3)$$

$$\frac{dS_i}{dt} = \rho_i N_i - \lambda_{Vi} S_i - \mu_i S_i \quad (4)$$

$$\frac{dL_i}{dt} = \lambda_{vi}S_i - (v_i + \mu_i)L_i \quad (5)$$

$$\frac{dI_i}{dt} = v_iL_i - \mu_iI_i \quad (6)$$

where S_H , I_H and R_H are the number of susceptible, infectious and recovered hosts, respectively, associated with the *Aedes*-borne disease of interest. S_i , L_i and I_i are the number of susceptible, latent and infectious vectors of kind $i=1,2$ (*Ae. aegypti* and *Ae. albopictus*, respectively). In addition,

$$\lambda_H = \sum_{i=1,2} \frac{1}{N_i} a_i b_i \phi_i m_i I_i \quad (7)$$

$$\lambda_{vi} = \frac{I_H}{H} a_i \beta_i \phi_i \quad (8)$$

and a_i is the daily biting rate (a function of temperature), b_i is the vector-to-host transmission probability, ϕ_i quantifies the vector's preference for humans, m_i is the vector-to-host ratio (a function of both temperature and rainfall; see (21) for details), β_i is the host-to-vector transmission probability, r is the daily recovery rate, and v_i and μ_i are the inverse of the extrinsic incubation period of the virus in days and the mortality rate², respectively, both a function of temperature. As in (21), the vector-to-host-ratio m_i is defined in terms of the probability of occurrence of the vectors (multiplied by 1000), which was obtained in (34) using maximum and minimum annual rainfall to account for the presence of water-filled containers, and other environmental variables involving temperature and urbanization; for details see the Materials and Methods section in (34). H and N_i are the total number hosts and the total number of the i -th kind of vector, respectively.

This is a 5-compartment model: infectious human host, latent *Ae. aegypti* vectors, latent *Ae. albopictus* vectors, infectious *Ae. aegypti* vectors and infectious *Ae. albopictus* vectors. If Δ and Λ are the new infectious rate appearing in a compartment and the rate at which individuals leave said compartment, respectively, then

$$\Delta = (\lambda_H S_H \quad \lambda_{v1} S_1 \quad \lambda_{v2} S_2 \quad 0 \quad 0)^T \quad (9)$$

$$\Lambda = (r I_H \quad (v_1 + \mu_1) L_1 \quad (v_2 + \mu_2) L_2 \quad -v_1 L_1 + \mu_1 I_1 \quad -v_2 L_2 + \mu_2 I_2)^T \quad (10)$$

The basic reproduction number R_0 is the dominant eigen-value of the next-generation matrix (21)

$$K = \left(\frac{\partial \Delta_m}{\partial x_l} \right)_{x_0} \left(\frac{\partial \Lambda_m}{\partial x_l} \right)_{x_0}^{-1} \quad (11)$$

for $m, l=1..5$ identifying the different compartments, x being a vector with the number of individuals in each compartment, and x_0 denoting the disease-free equilibrium state. The only non-zero elements K_{ml} (new infections in compartment m produced by infectious individuals in compartment l) of K are

² We found a typo in Table 1 for the expression of μ_1 in (21) : the correct value in the argument of the exponential function is $51.4 - 1.3 T$.

$$K_{12} = \frac{a_1 b_1 \phi_1 v_1}{(v_1 + \mu_1) \mu_1} \quad (12)$$

$$K_{13} = \frac{a_2 b_2 \phi_2 v_2}{(v_2 + \mu_2) \mu_2} \quad (13)$$

$$K_{14} = \frac{a_1 b_1 \phi_1}{\mu_1} \quad (14)$$

$$K_{15} = \frac{a_2 b_2 \phi_2}{\mu_2} \quad (15)$$

$$K_{21} = \frac{a_1 \beta_1 \phi_1 m_1}{r} \quad (16)$$

$$K_{31} = \frac{a_2 \beta_2 \phi_2 m_2}{r} \quad (17)$$

R_0 is the largest eigen-value solution of the eigen-value problem $|K - R_0 I| = 0$:

$$R_0^4 - R_0^2 (K_{21} K_{12} + K_{31} K_{13}) = 0 \quad (18)$$

or

$$R_0 = \sqrt{\frac{a_1^2 \phi_1^2 b_1 \beta_1 m_1 v_1}{(v_1 + \mu_1) \mu_1 r} + \frac{a_2^2 \phi_2^2 b_2 \beta_2 m_2 v_2}{(v_2 + \mu_2) \mu_2 r}} \quad (19)$$

where, as the indices suggest, the first term in the square root corresponds to *Ae. aegypti* and the second one to *Ae. albopictus*. As in (21), we set $R_0 = 0$ in all locations and times for which the total monthly rainfall has not been at least 80 mm during a minimum of five months, a condition for stable transmission.

This model has been reported (21) to reproduce well the observed basic reproduction number obtained when using the relatively short record of ZIKV cases available in Latin America. Because of this, we have chosen the same values of the parameters and functional dependence on temperature and rainfall that was used in a prior study (21).

The basic reproduction number can be understood as the expected number of new cases generated by a single (typical) infection in a completely susceptible population. It is a dimensionless number that can be associated with the potential risk of transmission of the disease (in the sense of the suitable conditions for transmission, as discussed at the Introduction), considering only basic environmental, entomological and epidemiological information. Only values of $R_0 > 1$ are related to spreading the infection in the population, and thus we focus on that range of values of R_0 in the present study.

The temperature dependence of certain parameters in the model (for example, the mortality rate μ_i ; see Figure 1) strongly controls the spatial and temporal distribution of R_0 . Most of Latin America and the Caribbean typically exhibit high values of the two-vector basic reproduction number (Figure 2). The potential risk of transmission of *Aedes*-borne diseases is higher for the northern half of South America, especially in Brazil, most of Colombia, Venezuela, Guyana, Suriname and the French Guyana, coastal Ecuador and the Ecuadorian and Peruvian Amazon. Central America and the Caribbean, although to a lesser degree, also exhibit high values of R_0 . Furthermore, with the increasing occurrence of high-temperature records, the frontier is extending farther into southern South America, in countries like Uruguay, which reported the first cases of autochthonous dengue

fever in 2016 (38). Nonetheless, places that are too hot decrease the life expectancy of the vectors (roughly speaking, with temperatures above 40°C, see Figure 1), and thus some regions in the future could start seeing a relative decrease in vector number if temperatures keep increasing.

An analysis of the evolution of the suitable conditions for transmission during 2013-2015 (Figure 3) complements the study on the associated temperature and rainfall anomalies performed previously (9). Standardized positive R_0 anomalies covering regions of northern South America and northern Brazil in 2013 became dominant in almost everywhere in northern half of South America, Central America and the Caribbean in 2015, with values exceeding one standard deviation in zones of Brazilian Amazon, the northern Peruvian coast, all of coastal Ecuador, most of northern Colombia and western Venezuela. Standardized anomalies of around two standard deviations occurred in the heart of the Brazilian Amazon.

The neutral standardized anomalies in the Brazilian Nordeste (Northeast), one of the most impacted places in terms of the present ZIKV epidemic, are attributed to the buffering role of the Atlantic Ocean in controlling the local temperatures. Still, neutral standardized anomalies in Nordeste are associated with R_0 ranging between 3.5-5.5, which indicate a very high potential risk of transmission.

The high values of the 2015 standardized anomalies (Figure 3c) are also consistent with the observed behavior of other diseases like dengue; for example, the reported number of dengue cases for Ecuador in 2015 (42,667) was about 3 times larger than the average number of cases for 2011-2014 (14,467.5); for details see (39). Nonetheless, work of our team in Machala (coastal Ecuador) suggests that a high percentage of the 2015 dengue cases reported there are likely to be chikungunya cases. Even if that is the case, the model was able to capture enhanced conditions leading to a higher number of *Aedes*-borne diseases.

The evolution of the spatially-averaged R_0 standardized anomalies for Latin America and the Caribbean exhibits a clear trend between 1950 and 2015 (black curve in Figure 3d), as reported by (21), that we attribute to the persistent increase in temperatures observed in the region. Once the longer-term signals are filtered-out, the inter-annual component of the R_0 standardized anomalies (filled curved in Figure 3d), show a peak in 2015 that is the second-highest on record, after the one in 1998. This contrasts with the analysis performed by (21); overall Figure 3d is telling the same story as Figure 3 in (21), the main differences due to the use of a different dataset and mostly to the use of a 12-month running average in our case (see Section 2 above). Our interpretation is consistent with our previous study on the 2015 climate conditions (9): a superposition of long-term, decadal and inter-annual signals was responsible for the 2015 absolute maximum in the unfiltered time series (black curve in Figure 3d). Although most likely the 2015 El Niño had an important contribution, the maximum cannot be explained only by this inter-annual phenomenon.

4 Skill Assessment and DJF 2014-2015 Forecast

A new seasonal forecast system for potential risk of transmission of *Aedes*-borne diseases can be developed using the ento-epidemiological model discussed in the previous section and multi-model ensemble climate predictions at seasonal scale. For this purpose, we have selected the set of coupled global models participating in the North American Multi-Model Ensemble project (22). Although our focus is Latin America and the Caribbean, the same system can be used for other regions of the

world, and a subset of the NMME models or a completely different seasonal climate forecast system can be used straightforwardly if that provides higher skill for the particular region of interest.

In brief, the system uses the monthly climate information from each one of the 116 (or 104, if the target period is between 2010 and 2015) realizations of the NMME models to compute the associated value of the basic reproduction number for each grid box in our geographical domain. Although the forecast horizon is typically 9 months after the initialization month, skill is normally higher for the first few seasons; to illustrate the approach here we focus on the first season starting immediately after the initialization month (e.g., JJA for forecasts initialized in May). After the multi-model ensemble and the seasonal average is computed, the output is corrected using a simple Principal Component Regression, which provided better results than other methods like Canonical Correlation Analysis or the use of the raw model output. For additional details, see Section 2.

The cross-validated analysis shows that there is relatively high skill ($> 60\%$, as measured by the 2AFC metric) for R_0 for all the seasons considered in most of northern half of South America and several regions of Central and North America, and some Caribbean nations (Figure 4). Overall, the skill is higher in DJF and MAM (with Kendall's τ of 0.199 and 0.191, respectively), and minimum in JJA (0.123), SON being in the middle (0.146). These values of Kendall's τ are typical for rainfall predictions in the region, as can be seen in the Validation Maproom of the Latin American Observatory's Datoteca (23,24,26): http://datoteca.ole2.org/maproom/Sala_de_Validacion/

Regionally speaking, skill is higher in Mexico in JJA, especially in the south (Figure 4). Central American countries exhibit high skill (above 70% for most of them) for DJF and MAM, with the lowest values ($< 50\%$) occurring in JJA and SON for Panama and Costa Rica. The western Caribbean tends to show higher skill during JJA, while the Central Caribbean and Lesser Antilles during MAM.

Northern South America shows relatively high skill ($> 70\%$) all year around, with some regions — like Ecuador, northern Peru, southwestern Colombia, northeastern Venezuela and northern Guyana — showing no skill during JJA and SON (Figure 4). The forecast system has in general low skill or no skill at all for southern South America, with some exceptions, e.g., the Bolivian Amazon in DJF, Paraguay and northern Argentina in SON, and northwestern Uruguay in DJF. Most of Brazil exhibits values of the 2AFC metric that are above 50% in all seasons, although southern Brazil has very low skill in MAM. In general, Chile and central and southern Argentina, do not show potential risk of transmission with this model, and thus those regions appear in white in our skill maps (Figure 4).

To illustrate an example of the bias-corrected probabilistic forecasts produced by our system, we now consider the season prior to the first ZIKV case reported in Brazil (May 2015 (40)): DJF 2014-2015. The probabilistic prediction indicates that there were mostly conditions for above-normal risk of transmission in eastern Brazil, which is similar to the observed conditions (Figure 5). Of course, no model is perfect and although, overall, there are similarities between the forecast and observed basic reproduction number for that season — especially in South America — there are also differences. Nonetheless, below-normal conditions were in general no forecast in the ZIKV hotspot places (in Brazil, for example), and as discussed above, the normal category the northern half of South America is already conducive to epidemic conditions. Hence, we claim that this particular forecast, even if not perfect, could have been useful for decision-makers at the time (November 2014), assuming that they already knew there was ZIKV in the region, which was of course not the case.

The previous example also illustrates why this tool can only be used as a guide for the local and international experts, as these diseases involve complex interactions beyond the presence or not of enhanced environmental (climatic) conditions suitable for the occurrence and transmission of *Aedes*-borne epidemics.

5 Concluding remarks

We have discussed the development and predictive skill of new probabilistic forecast system for above-normal, normal and below-normal environmental conditions associated with potential risk of transmission of diseases like ZIKV, DENV and CHIKV. To the best of our knowledge this is the first seasonal forecast system of this type for Latin America and the Caribbean, although it is conceptually similar to a malaria forecast system developed for Africa years ago (18).

From the regional perspective, this forecast system has the potential to help the Pan-American Health Organization (PAHO), the World Health Organization (WHO) and other decision-makers to prepare more detailed epidemiological alerts and guides for Zika's surveillance and other arboviruses; to calculate different levels of population at risk and incidence rates for regional assessment, to prepare vector control guidelines for a more integrated management; to plan and support vector control resources an equipment; to organize and program activities and resource mobilization, as well as improve risk communication materials. One of the co-authors (PN) has already started to explore ways to take advantage of this forecast system at PAHO/WHO.

Our system is a first attempt to provide predictive tools for health practitioners and decision-makers interested in *Aedes*-borne diseases in Latin America and the Caribbean, and can be considered an additional step in the direction followed by previous research groups (21,34,36,41–44).

Indeed, forecasts of health events are designed to change human behavior. Nonetheless, as with the practice of medicine, there are ethical issues to consider. It is possible that there might be negative consequences from an epidemic risk forecast (i.e., incidence, or cases), even if the prediction is skillful. To illustrate this idea, consider that a forecast for ZIKV is provided to the community, indicating that there is above 80% probability of acquiring the disease in Rio de Janeiro during a certain season, but less than 10% probability of infection in Montevideo. People —some of whom could already be infected with ZIKV, or even with a different disease— might decide to travel to Montevideo instead of Rio de Janeiro because of that forecast, thus igniting or being part of a new focus of an epidemic there, that was not predicted and that is partially caused by the original prediction itself. This is an important caveat to be considered by the decision-makers. Another consideration is that a ZIKV forecast may have negative consequences for tourism, leading to livelihood impacts that may have negative health consequences.

There are a number of important limitations related to our forecast system. As indicated earlier in this paper, by itself this kind of system cannot forecast the occurrence and spread of new epidemics, but only partial conditions for that to happen. The model employed here only considers the effect of climatic conditions, through temperature and rainfall, on disease transmission via the vectors and viruses of interest. Direct human to human transmission via sexual intercourse and blood transfusion are outside the scope of this modelling approach. Also, the present version of the model cannot simulate co-infections or mixed states (e.g., a fraction of the population that is recovered from dengue but it is susceptible to Zika).

One particular way in which the model needs to be improved involves how rainfall is considered. The present version of the model only uses rainfall in a rather simplistic way, without really considering its seasonal characteristics. There are examples in the scientific literature that could be used to improve the representation of rainfall in this type of model (see for example, (45–47)). In addition, there is room to consider a better set of realizations in the ensemble of simulations, varying the ento-epidemiological parameters of the model. This will be explored in the future.

6 Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

7 Author Contributions

ÁM and MT established the concept of the study. ÁM obtained the data. All authors undertook the analysis and interpretation of results. ÁM, MT and AS drafted the manuscript. All authors critically reviewed and revised the manuscript and agreed the final submission.

8 Funding

ÁM was supported by National Oceanic and Atmospheric Administration Oceanic and Atmospheric Research, under the auspices of the National Earth System Prediction Capability. XC was supported by the Centro de Modelado Científico's project CMC-CC-Dat-2016.

9 Abbreviations

2AFC: Two-Alternative Forced Choice
 CAMS Climate Anomaly Monitoring System
 CHIKV: chikungunya virus
 CPT: Climate Predictability Tool
 CRU: Climate Research Unit, at East Anglia University
 DENV: dengue virus
 DJF: December-January-February
 JJA: June-July-August
 IRI: International Research Institute for Climate and Society, at Columbia University
 MAM: March-April-May
 NMME: North American Multi-Model Ensemble
 SON: September-October-November
 PAHO: Pan-American Health Organization
 PCR: Principal Component Regression
 WHO: World Health Organization
 ZIKV: Zika virus

10 Acknowledgments

The authors are thankful to the editors of this special number of Frontiers for their kind invitation to contribute a paper, and to Ángel Adames, Hongai Zhang and two anonymous reviewers who helped

improve the clarity of the original manuscript. This research used computational resources from the Latin American Observatory for Climate Events, in particular its Datoteca.

References

1. Abushouk AI, Negida A, Ahmed H. An updated review of Zika virus. *Journal of Clinical Virology*. 2016. p. 53–8.
2. PAHO WHO. PAHO WHO | Zika Cumulative Cases [Internet]. [cited 2017 Jan 20]. Available from: http://www.paho.org/hq/index.php?option=com_content&view=article&id=12390&Itemid=42090
3. Duffy MR, Chen T-H, Hancock WT, Powers AM, Kool JL, Lanciotti RS, et al. Zika virus outbreak on Yap Island, federated states of Micronesia. *N Engl J Med*. 2009;2009(360):2536–2543.
4. Waggoner JJ, Gresh L, Vargas MJ, Ballesteros G, Tellez Y, Soda KJ, et al. Viremia and Clinical Presentation in Nicaraguan Patients Infected with Zika Virus, Chikungunya Virus, and Dengue Virus. *Clin Infect Dis*. Oxford University Press; 2016 Dec 15;63(12):1–7.
5. Schuler-Faccini L. Possible association between Zika virus infection and microcephaly—Brazil, 2015. *MMWR Morb Mortal Wkly Rep*. 2016;65.
6. Calvet G, Aguiar RS, Melo ASO, Sampaio SA, de Filippis I, Fabri A, et al. Detection and sequencing of Zika virus from amniotic fluid of fetuses with microcephaly in Brazil: a case study. *Lancet Infect Dis*. 2016 Jun;16(6):653–60.
7. Cao-Lormeau V-M, Blake A, Mons S, Lastère S, Roche C, Vanhomwegen J, et al. Guillain-Barré Syndrome outbreak associated with Zika virus infection in French Polynesia: a case-control study. *Lancet*. Elsevier; 2016 Mar;
8. PAHO. WHO | Zika situation report. World Health Organization; 2016.
9. Muñoz ÁG, Thomson MC, Goddard L, Aldighieri S. Analyzing climate variations at multiple timescales can guide Zika virus response measures. *Gigascience*. BioMed Central; 2016 Dec 6;5(1):41.
10. Mordecai EA, Cohen JM, Evans M V., Gudapati P, Johnson LR, Lippi CA, et al. Detecting the impact of temperature on transmission of Zika, dengue, and chikungunya using mechanistic models. Althouse B, editor. *PLoS Negl Trop Dis*. R Foundation for Statistical Computing; 2017 Apr 27;11(4):e0005568.
11. Stewart Ibarra AM, Ryan SJ, Beltrán E, Mejía R, Silva M, Muñoz A. Dengue vector dynamics (*Aedes aegypti*) influenced by climate and social factors in Ecuador: implications for targeted control. *PLoS One*. 2013 Jan;8(11):e78263.
12. Stewart-Ibarra AM, Lowe R. Climate and non-climate drivers of dengue epidemics in southern coastal Ecuador. *Am J Trop Med Hyg*. 2013 May 11;88(5):971–81.

- 469 13. Chretien J-P, Anyamba A, Bedno SA, Breiman RF, Sang R, Seron K, et al. Drought-
470 associated chikungunya emergence along coastal East Africa. *Am J Trop Med Hyg. American*
471 *Society of Tropical Medicine and Hygiene*; 2007 Mar;76(3):405–7.
- 472 14. Smith DL, Battle KE, Hay SI, Barker CM, Scott TW, McKenzie FE, et al. Ross, Macdonald,
473 and a Theory for the Dynamics and Control of Mosquito-Transmitted Pathogens. Chitnis CE,
474 editor. *PLoS Pathog. Public Library of Science*; 2012 Apr 5;8(4):e1002588.
- 475 15. Murray JD (James D. Mathematical biology. Springer; 2002.
- 476 16. Anderson RM, May RM. *Infectious Diseases of Humans: Dynamics and Control*. Oxford
477 University Press; 1991. 757 p.
- 478 17. Moreno Y, Pastor-Satorras R, Vespignani A. Epidemic outbreaks in complex heterogeneous
479 networks. *Eur Phys J B. EDP Sciences, Springer-Verlag*; 2002 Apr;26(4):521–9.
- 480 18. Thomson MC, Doblas-Reyes FJ, Mason SJ, Hagedorn R, Connor SJ, Phindela T, et al. Malaria
481 early warnings based on seasonal climate forecasts from multi-model ensembles. *Nature*.
482 Nature Publishing Group; 2006 Feb 2;439(7076):576–9.
- 483 19. Harris I, Jones PD, Osborn TJ, Lister DH. Updated high-resolution grids of monthly climatic
484 observations - the CRU TS3.10 Dataset. *Int J Climatol*. 2014 Mar 15;34(3):623–42.
- 485 20. Fan Y, van den Dool H. A global monthly land surface air temperature analysis for 1948–
486 present. *J Geophys Res Atmos*. 2008;113(D1).
- 487 21. Caminade C, Turner J, Metelmann S, Hesson JC, Blagrove MSC, Solomon T, et al. Global
488 risk model for vector-borne transmission of Zika virus reveals the role of El Niño 2015. *Proc*
489 *Natl Acad Sci. National Academy of Sciences*; 2017 Jan 3;114(1):119–24.
- 490 22. Kirtman BP, Min D, Infanti JM, James L. Kinter I, Paolino DA, Zhang Q, et al. The North
491 American Multimodel Ensemble: Phase-1 Seasonal-to-Interannual Prediction; Phase-2 toward
492 Developing Intraseasonal Prediction. *American Meteorological Society*; 2014 Jun 4;
- 493 23. Muñoz ÁG, López P, Velásquez R, Monterrey L, León G, Ruiz F, et al. An Environmental
494 Watch System for the Andean Countries: El Observatorio Andino. *Bull Am Meteorol Soc*.
495 2010 Dec;91(12):1645–52.
- 496 24. Muñoz ÁG, Ruiz D, Ramírez P, León G, Quintana J, Bonilla A, et al. Risk Management -
497 Current Issues and Challenges. Banaitiene N, editor. *Risk Management - Current Issues and*
498 *Challenges*. InTech; 2012. 532-556 p.
- 499 25. Muñoz ÁG, Thomson MC, Goddard L, Aldighieri S. Supporting data for “Analyzing climate
500 variations on multiple timescales can guide Zika virus response measures” [Internet].
501 GigaScience Database; 2016. Available from: <http://dx.doi.org/10.5524/100243>
- 502 26. Chourio X. The Latin American Observatory’s Datoteca. *Climate Service Partnership*
503 *Newsletter*. 2016 Apr;6.
- 504 27. Mason SJ, Weigel AP. A Generic Forecast Verification Framework for Administrative

- 505 Purposes. Mon Weather Rev. 2009 Jan 25;137(1):331–49.
- 506 28. Mason SJ, Tippet MK. Climate Predictability Tool version 15.3. 2016;
- 507 29. Mason SJ, Baddour O. Statistical Modelling. In: Troccoli A, Harrison M, Anderson DLT,
508 Mason SJ, editors. Seasonal Climate: Forecasting and Managing Risk. Dordrecht: Springer
509 Netherlands; 2008.
- 510 30. Jolliffe IT, Stephenson DB, (Eds). Forecast Verification: a practitioner's guide in atmospheric
511 science. 2nd ed. Chichester: Wiley and Sons; 2012. 292 p.
- 512 31. Gardner LM, Chen N, Sarkar S. Global risk of Zika virus depends critically on vector status of
513 Aedes albopictus. Lancet Infect Dis. 2016;16.
- 514 32. Lambrechts L, Scott TW, Gubler DJ. Consequences of the expanding global distribution of
515 aedes albopictus for dengue virus transmission. Halstead SB, editor. PLoS Neglected Tropical
516 Diseases. W. H. Freeman and Co; 2010. p. e646.
- 517 33. Muñoz ÁG, Thomson MC, Goddard L, Aldighieri S. Analyzing climate variations at multiple
518 timescales can guide Zika virus response measures. Gigascience. 2016 Dec 6;5(1):41.
- 519 34. Kraemer MUG, Sinka ME, Duda KA, Mylne AQN, Shearer FM, Barker CM, et al. The global
520 distribution of the arbovirus vectors Aedes aegypti and Ae. albopictus. Jit M, editor. Elife.
521 eLife Sciences Publications, Ltd; 2015 Jun;4:e08347.
- 522 35. Li Y, Kamara F, Zhou G, Puthiyakunnon S, Li C, Liu Y, et al. Urbanization Increases Aedes
523 albopictus Larval Habitats and Accelerates Mosquito Development and Survivorship.
524 Kittayapong P, editor. PLoS Negl Trop Dis. 2014 Nov 13;8(11):e3301.
- 525 36. Lessler J, Chaisson LH, Kucirka LM, Bi Q, Grantz K, Salje H, et al. Assessing the global
526 threat from Zika virus. Science. American Association for the Advancement of Science; 2016
527 Jul 14;46(24):601–4.
- 528 37. Turner J, Bowers RG, Baylis M, Diekmann O, Heesterbeek J, Roberts M, et al. Two-Host,
529 Two-Vector Basic Reproduction Ratio (R0) for Bluetongue. Gomez-Gardenes J, editor. PLoS
530 One. Public Library of Science; 2013 Jan 8;8(1):e53128.
- 531 38. WHO | Dengue Fever – Uruguay. WHO. World Health Organization; 2016;
- 532 39. PAHO WHO. PAHO WHO | Dengue | Annual Cases Reported of Dengue | PAHO/WHO
533 Data, Maps and Statistics [Internet]. Dengue: PAHO/WHO Data, Maps and Statistics. 2016
534 [cited 2017 Jan 20]. Available from:
535 http://www.paho.org/hq/index.php?option=com_topics&view=rdmore&cid=6290&Itemid=40
536 734
- 537 40. Faria NR, Azevedo R do S da S, Kraemer MUG, Souza R, Cunha MS, Hill SC, et al. Zika
538 virus in the Americas: Early epidemiological and genetic findings. Science. American
539 Association for the Advancement of Science; 2016 Mar 24;aaf5036.
- 540 41. Messina JP, Kraemer MU, Brady OJ, Pigott DM, Shearer FM, Weiss DJ, et al. Mapping global

- environmental suitability for Zika virus. *Elife*. eLife Sciences Publications Limited; 2016 Apr 19;5:e15272.
42. Carlson CJ, Dougherty ER, Getz W, Attar N, Dick G, Kitchen S, et al. An Ecological Assessment of the Pandemic Threat of Zika Virus. Johansson MA, editor. *PLoS Negl Trop Dis*. Public Library of Science; 2016 Aug 26;10(8):e0004968.
 43. Samy AM, Thomas SM, Wahed AA El, Cohoon KP, Peterson AT. Mapping the global geographic potential of Zika virus spread. *Mem Inst Oswaldo Cruz*. Instituto Oswaldo Cruz; 2016 Sep;111(9):559–60.
 44. Monaghan AJ, Morin CW, Steinhoff DF, Wilhelmi O, Hayden M, Quattrochi DA, et al. On the Seasonal Occurrence and Abundance of the Zika Virus Vector Mosquito *Aedes Aegypti* in the Contiguous United States. *PLoS Curr*. Public Library of Science; 2016;8(August):1–31.
 45. Santos LBL, Costa MC, Pinho STR, Andrade RFS, Barreto FR, Teixeira MG, et al. Periodic forcing in a three-level cellular automata model for a vector-transmitted disease. *Phys Rev E - Stat Nonlinear, Soft Matter Phys*. American Physical Society; 2009 Jul 6;80(1):16102.
 46. Magori K, Legros M, Puente ME, Focks DA, Scott TW, Lloyd AL, et al. Skeeter Buster: A Stochastic, Spatially Explicit Modeling Tool for Studying *Aedes aegypti* Population Replacement and Population Suppression Strategies. Kittayapong P, editor. *PLoS Negl Trop Dis*. Florida Medical Entomology Laboratory; 2009 Sep 1;3(9):e508.
 47. Morin CW, Comrie AC. Modeled response of the West Nile virus vector *Culex quinquefasciatus* to changing climate using the dynamic mosquito simulation model. *Int J Biometeorol*. 2010 Sep;54(5):517–29.

FIGURE CAPTIONS

Figure 1. Daily vector mortality rate as a function of mean temperature (in Celsius).

Figure 2. Observed climatology of R_0 considering all months in the period 1982-2010. Only $R_0 > 1$ values are plotted. There is no data over the oceans.

Figure 3. Spatial evolution of standardized R_0 yearly anomalies for (a) 2013, (b) 2014 and (c) 2015. (d) Average evolution of standardized R_0 anomalies (in units of standard deviations, σ) for Latin America and the Caribbean (domain in panel (a)) for the 1950-2015 period. Black empty curve and filled curve show the raw and linearly detrended standardized anomalies, respectively. A 12-month running average filter was applied to both curves to better capture the inter-annual variability. There is no data over the oceans.

Figure 4. 2AFC skill score for the seasonal forecast system for each one of the four seasons selected: (a) DJF, (b) MAM, (c) JJA and (d) SON. Units in %. The 2AFC score is an indication of how often the forecasts are correct; it also measures how well the system can distinguish between the above-normal, normal and below-normal categories.

Figure 5. (a) Observed terciles (above normal, normal, below normal) for the basic reproduction number (R_0), computed using observed climate data for DJF 2014-2015 and the model presented in

section 3. (b) Forecast probabilities (in %) for R_0 for the same DJF season, computed using predicted climate data, the vector model presented in section 3 and the probabilistic Principal Component Regression model described in section 2.

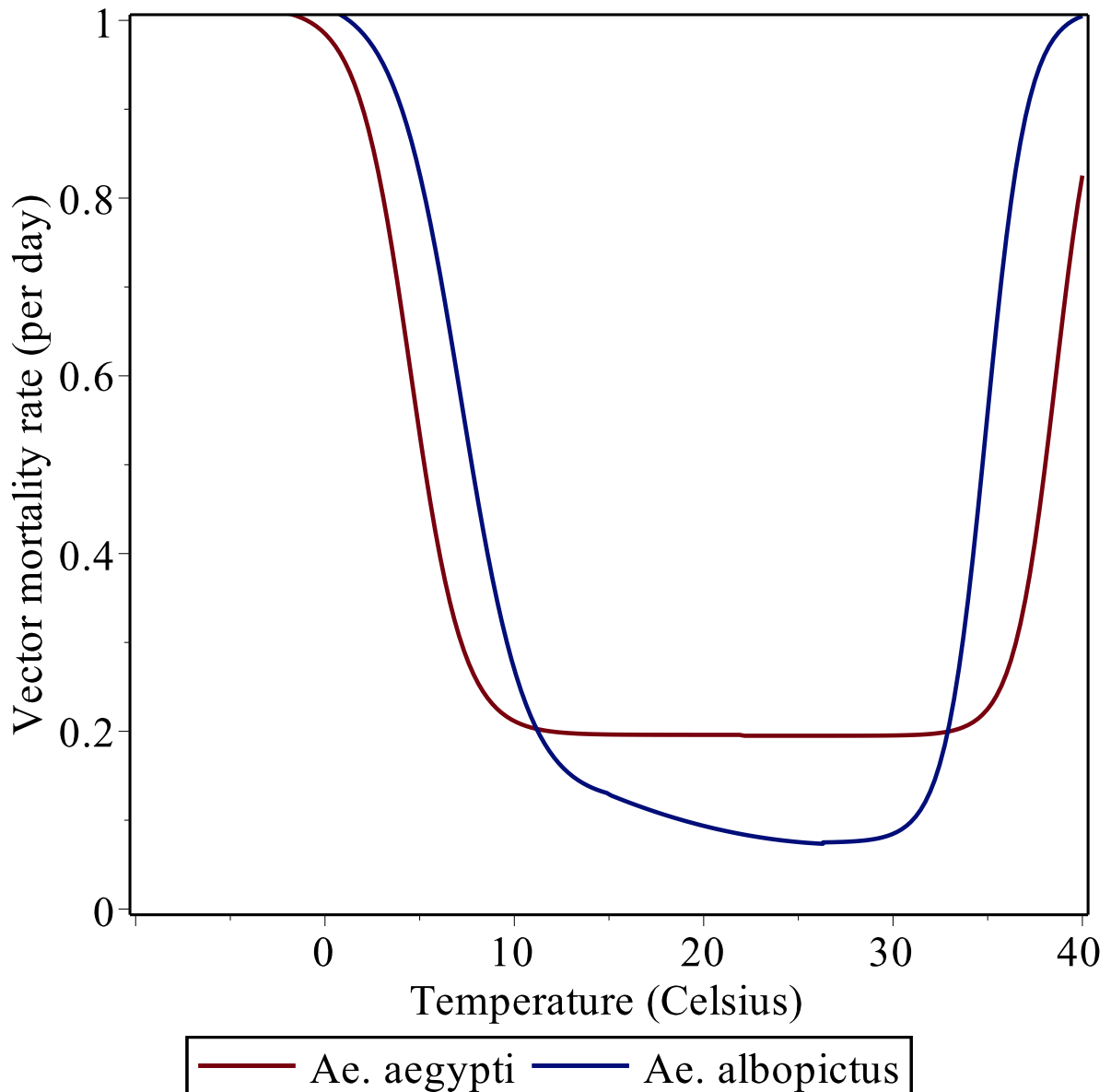


Figure 1

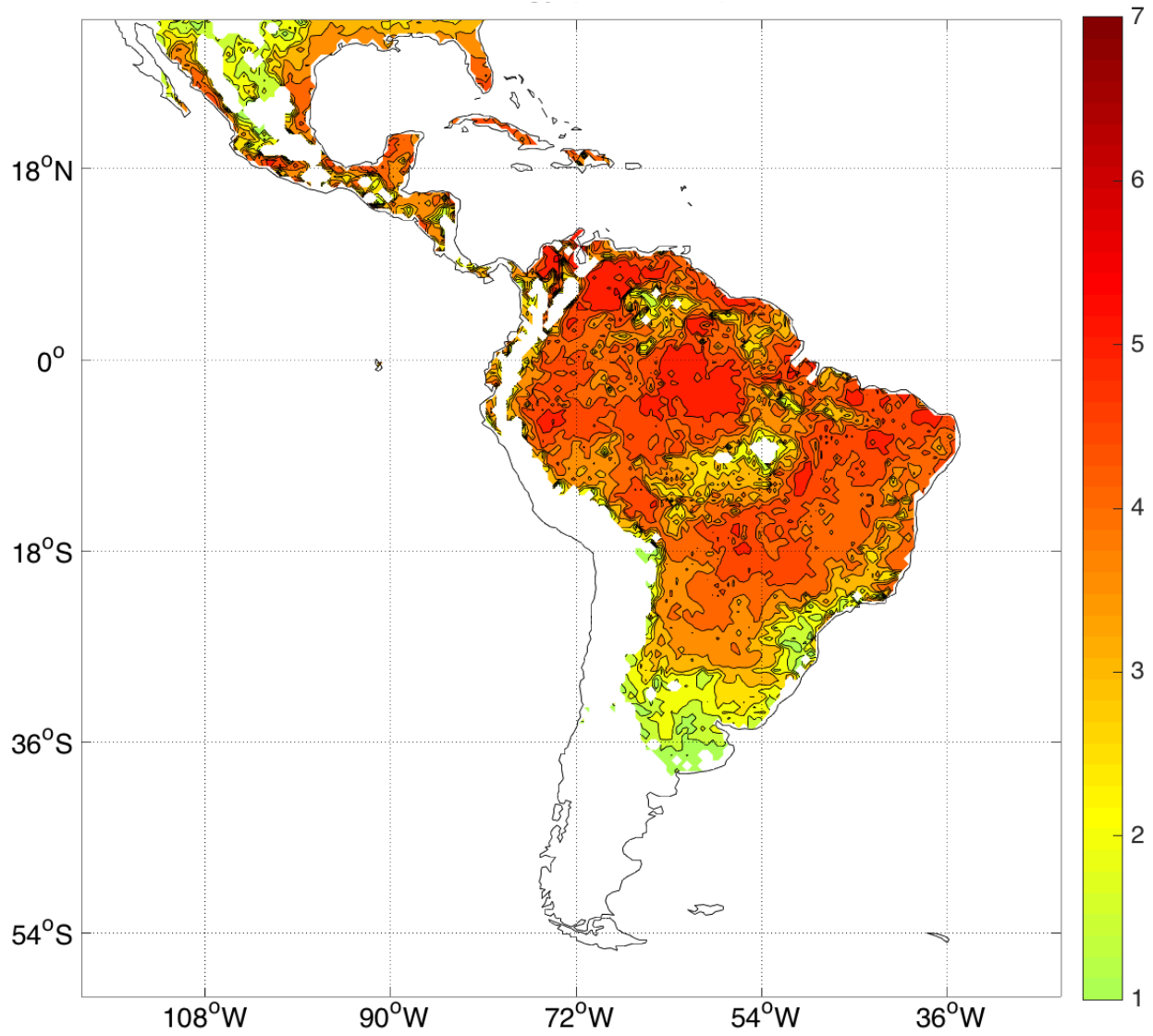


Figure 2

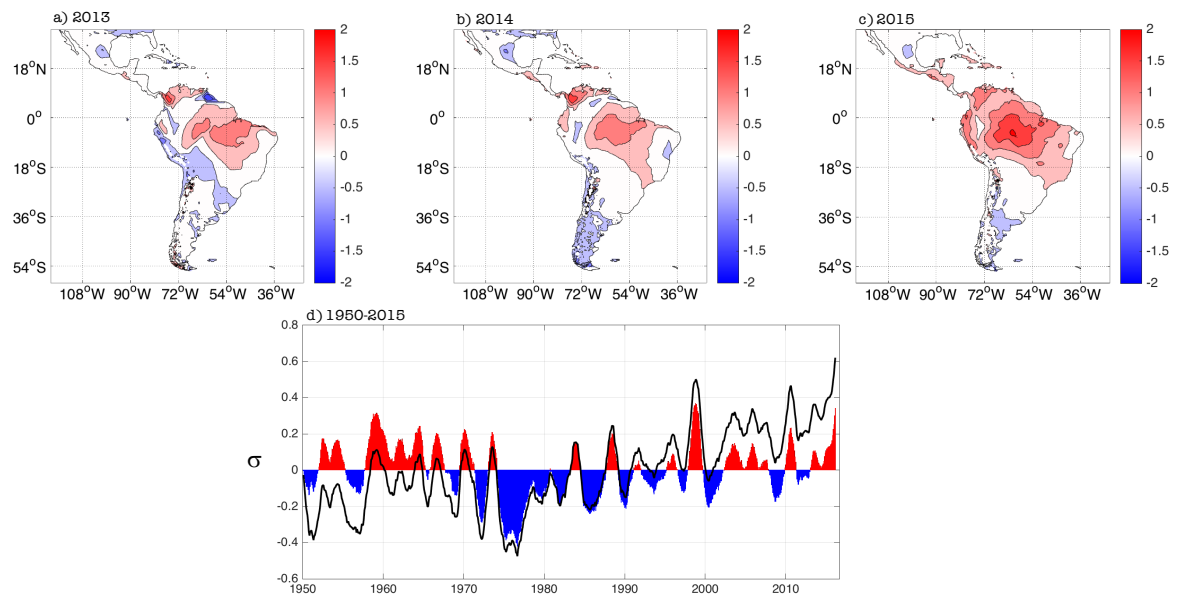


Figure 3

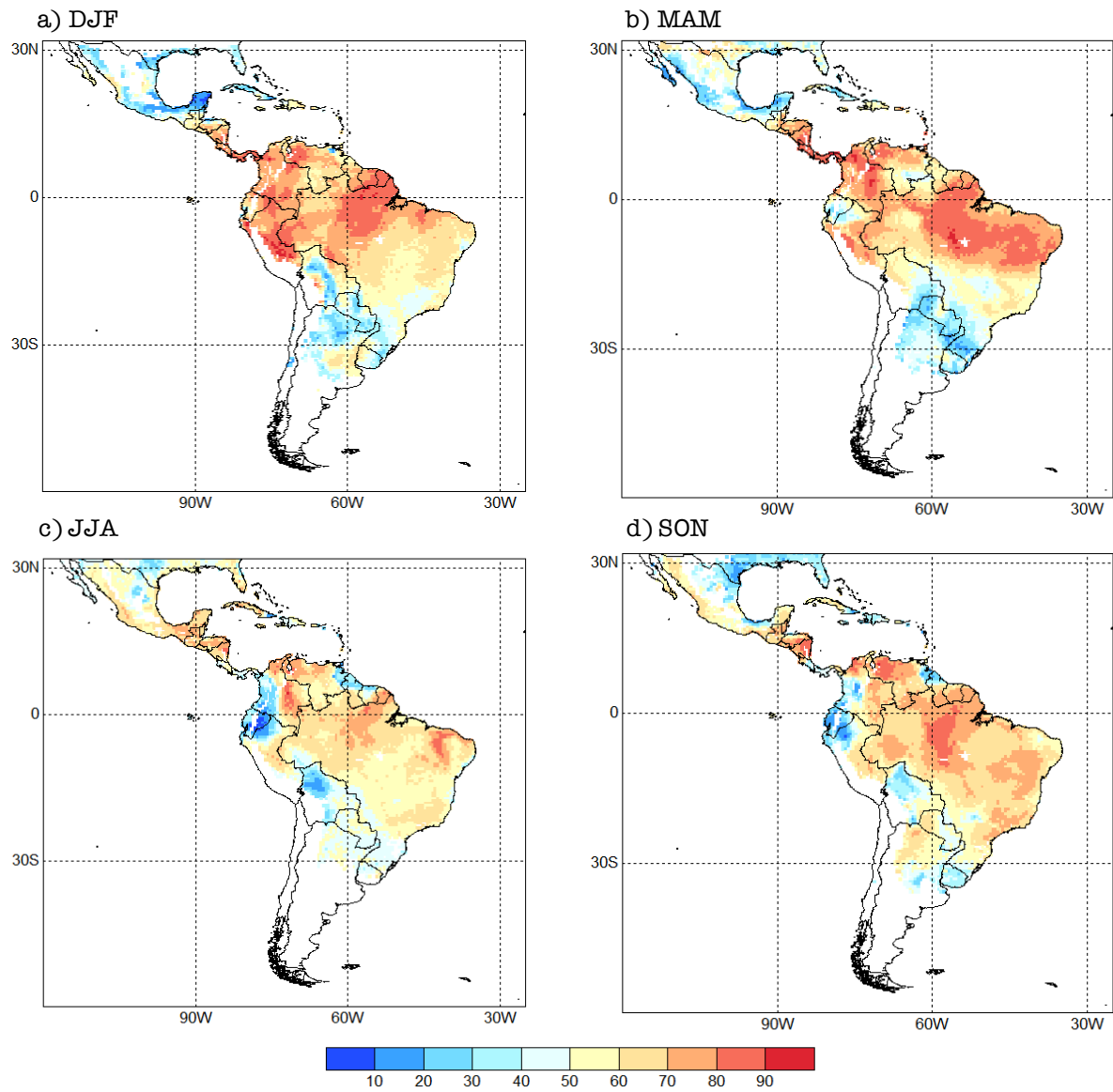


Figure 4

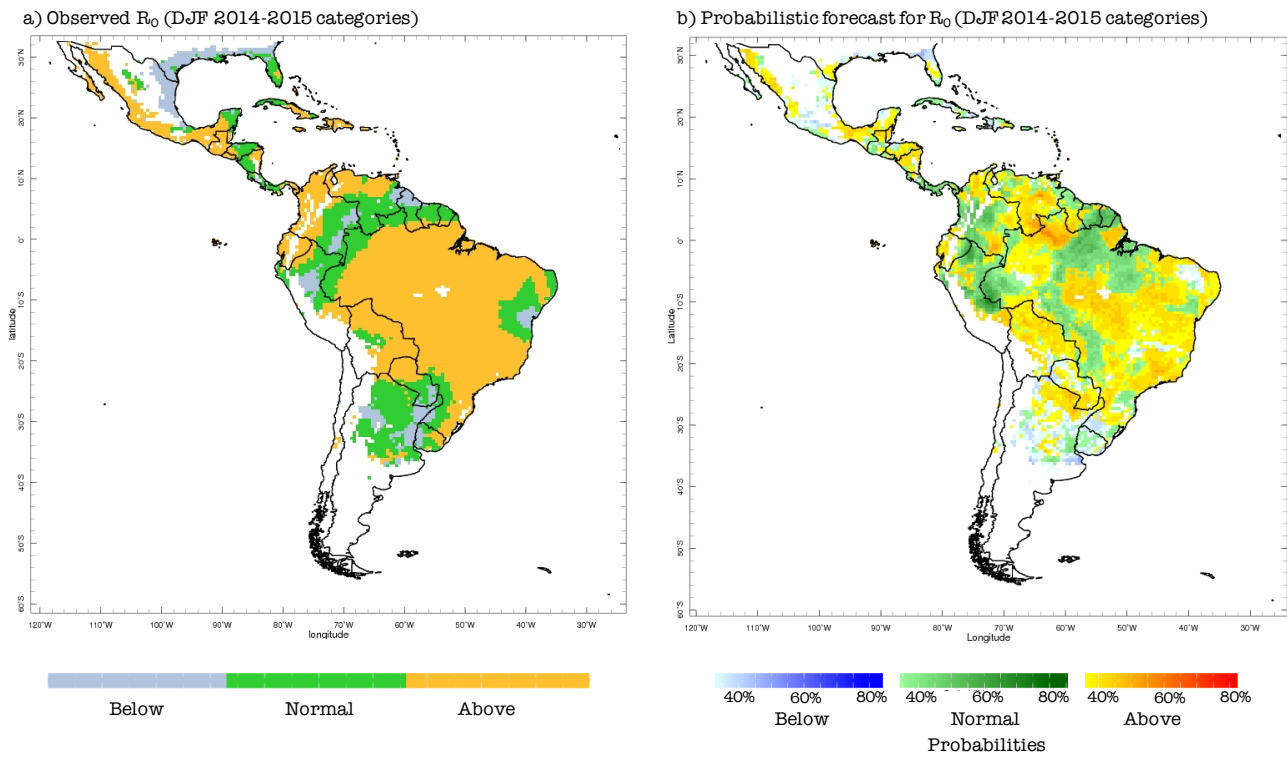


Figure 5