1

2

3   **A dynamic neural network model for predicting risk of Zika in real-time**

4

5

6   Mahmood Akhtar[1*], Moritz U.G. Kraemer[2,3,4], Lauren M. Gardner[5,1ψ]

7

8   [1]School of Civil and Environment Engineering, University of New South Wales,

9   Sydney, NSW 2052, Australia

10   [2] Department of Zoology, University of Oxford, Oxford OX1 3PS, UK

11   [3] Computational Epidemiology Group, Boston Children's Hospital, Boston, MA, USA

12   [4] Harvard Medical School, Boston, MA, USA

13   [5] Johns Hopkins University, Baltimore, MD, USA

14

15   Corresponding email: *m.akhtar@unsw.edu.au (MA)

16   Corresponding email: ψl.gardner@jhu.edu (LG)

17

18 **Abstract**

19 **Background**

20 In 2015 the Zika virus spread from Brazil throughout the Americas, posing an

21 unprecedented challenge to the public health community. During the epidemic,

22 international public health officials lacked reliable predictions of the outbreak's

23 expected geographic scale and prevalence of cases, and were therefore unable to plan

24 and allocate surveillance resources in a timely and effective manner.

25

26 **Methods**

27 In this work we present a dynamic neural network model to predict the geographic

28 spread of outbreaks in real-time. The modeling framework is flexible in three main

29 dimensions i) selection of the chosen risk indicator, *i.e.,* case counts or incidence rate,

30 ii) risk classification scheme, which defines the high risk group based on a relative or

31 absolute threshold, and iii) prediction forecast window (one up to 12 weeks). The

32 proposed model can be applied dynamically throughout the course of an outbreak to

33 identify the regions expected to be at greatest risk in the future.

34

35 **Results**

36 The model is applied to the recent Zika epidemic in the Americas at a weekly

37 temporal resolution and country spatial resolution, using epidemiological data,

38 passenger air travel volumes, vector habitat suitability, socioeconomic and population

39 data for all affected countries and territories in the Americas. The model performance

40 is quantitatively evaluated based on the predictive accuracy of the model. We show

41 that the model can accurately predict the geographic expansion of Zika in the

42   Americas with the overall average accuracy remaining above 85% even for prediction

43   windows of up to 12 weeks.

44

45   **Conclusions**

46   Sensitivity analysis illustrated the model performance to be robust across a range of

47   features. Critically, the model performed consistently well at various stages

48   throughout the course of the outbreak, indicating its potential value at any time during

49   an epidemic. The predictive capability was superior for shorter forecast windows, and

50   geographically isolated locations that are predominantly connected via air travel.

51   The highly flexible nature of the proposed modeling framework enables policy

52   makers to develop and plan vector control programs and case surveillance strategies

53   which can be tailored to a range of objectives and resource constraints.

54

55   **Keywords**

56   Zika, epidemic risk prediction, dynamic neural network

57

58

59

60

Background

The Zika virus, which is primarily transmitted through the bite of infected *Aedes aegypti* mosquitoes (1), was first discovered in Uganda in 1947 (2) from where it spread to Asia in 1960s, where it since has caused small outbreaks. In 2007 ZIKV caused an island wide outbreak in Yap Island, Micronesia (3), followed by outbreaks in French Polynesia (4) and other Pacific islands between 2013−2014 where attack rates where up to 70% (5-7). It reached Latin America between late 2013 and early 2014, but was not detected by public health authorities until May 2015 (8) and since affected 48 countries and territories in the Americas (9-11). Since there is no vaccination or treatment available for Zika infections (12, 13), the control of *Ae. aegypti* mosquito populations remains the most important intervention to contain the spread of the virus (14). In order to optimally allocate resources to suppress vector populations, it is critical to accurately anticipate the occurrence and arrival time of arboviral infections to detect local transmission (15).

Whereas for dengue, the most common arbovirus infection, prediction has attracted wide attention from researchers employing statistical modelling and machine learning methods to guide vector control (16-29), such real-time machine learning based models do not yet exist for Zika virus. Early warning systems for Thailand, Indonesia, Ecuador and Pakistan have been introduced and are currently in use (30-34). In addition to conventional predictions based on epidemiological and meteorological data (20, 35, 36), more recent models have successfully incorporated search engines (37, 38), land use (39), human mobility information (40, 41) and spatial dynamics (42-44), and various combinations of the above (45) to improve predictions. Whereas

4

85   local spread may be mediated by overland travel, continent wide spread is mostly

86   driven by air passenger travel between climatically synchronous regions (46-52).

87

88   The aims of our work are to 1) present recurrent neural networks for time ahead

89   predictive modelling as a highly flexible tool for outbreak prediction, and 2)

90   implement and evaluate the model performance for the Zika epidemic in the

91   Americas. The application of neural networks for epidemic risk forecasting has

92   previously been applied to dengue forecasting and risk classification (53-58),

93   detection of mosquito presence (59), temporal modeling of the oviposition of *Aedes*

94   *aegypti* mosquito (60), *Aedes* larva identification (61), and epidemiologic time-series

95   modeling through fusion of neural networks, fuzzy systems and genetic algorithms

96   (62). Recently, Jian *et al* (63) performed a comparison of different machine learning

97   models to map the probability of Zika epidemic outbreak using publically available

98   global Zika case data and other known covariates of transmission risk. Their study

99   provides valuable insight into the potential role of machine learning models for

100   understanding Zika transmission; however, it is static in nature, *i.e.,* it does not

101   account for time-series data, and did not account for human mobility, both of which

102   are incorporated in our modelling framework.

103

104   Here, we apply a dynamic neural network model for N-week ahead prediction for the

105   2015-2016 Zika epidemic in the Americas. The model implemented in this work

106   relies on multi-dimensional time-series data at the country (or territory) level,

107   specifically epidemiological data, passenger air travel volumes, vector habitat

108   suitability for the primary spreading vector *Ae. aegypti*, socioeconomic and

109   population data. The modeling framework is flexible in three main dimensions: 1) the

110   preferred risk indictor can be chosen by the policy maker, *e.g.*, we consider outbreak

111   size and incidence rate as two primary indicators of risk for a region, 2) five risk

112   classification schemes are defined, where each classification scheme varies in the

113   (relative or absolute) threshold used to determine the set of countries deemed "high

114   risk", and 3) it can be applied for a range of forecast windows (1 – 12 weeks). Model

115   performance and robustness is evaluated for various combinations of risk indicator,

116   risk classification level, and forecasting windows. Thus, our work represents the first

117   flexible framework of neural networks for epidemic risk forecasting, that allows

118   policy makers to evaluate and weigh the trade-off in prediction accuracy between

119   forecast window and risk classification schemes. Given the availability of the

120   necessary data, the modelling framework proposed here can be applied in real time to

121   future outbreaks of Zika, and other similar vector-borne outbreaks.

122

123   **Materials and Methods**

124

125   *Data*

126   The model relies on socioeconomic, population, epidemiological, travel and mosquito

127   vector suitability data. All data is aggregated to the country level and provided for all

128   countries and territories in the Americas. Each data set and corresponding processing

129   is described in detail below, and summarized in Table 1. All input data is available as

130   Additional files 1-11.

131

132

133

134 **Table 1. Summary of input data**

Description	Original Temporal Resolution	Spatial Resolution	Temporal Disaggregation	Reference
Zika cases (2015)	monthly	country or territory level	estimation, smoothing spline curve fitting	(50, 51, 64)
Zika cases (2016)	weekly	country or territory level	-	Pan American Health Organization (PAHO) (65)
incidence rates	weekly	country or territory level	(case counts/(population / 100,000)	(51)
incoming and outgoing travel volumes (2015)	monthly	country or territory level	smoothing spline curve fitting	International Air Transport Associate (IATA), (64)
incoming and outgoing travel volumes (2016)	monthly	country or territory level	estimation, smoothing spline curve fitting	as previously done (51, 66, 67), (64)
Aedes vector suitability	monthly	country or territory level	smoothing spline curve fitting	(51, 64, 68, 69)
gross domestic product (GDP) per capita	annual	country or territory level	-	World Bank (70), and U.S. Bureau of Economic Analysis (71)
physicians per 1000 people	annual	country or territory level	-	Centre of Disease Control and Prevention (CDC) (72), WHO World Health Statistics report (73), and the PAHO (74)
beds per 1000 people	annual	country or territory level	-	
population densities (people per sq. km of land area)	annual	country or territory level	-	World Bank (75), and the U.S. Bureau of Economic Analysis (71)

135

136 *Epidemiological Data*

137 Weekly Zika infected cases for each country and territory in the Americas were

138 extracted from the Pan American Health Organization (PAHO) (65), as described in

139 previous studies (48, 51) (data available: github.com/andersen-lab/Zika-cases-

140 PAHO). The epidemiological weeks 1 - 78 are labeled herein as EPI weeks,

7

141   corresponding to the dates 29-Jun-2015 - 19-Dec-2016, respectively. Although Zika

142   cases in Brazil were reported as early as May 2015, no case data is available for all of

143   2015 from PAHO because the Brazil Ministry of Health did not declare the Zika cases

144   and associated neurological and congenital syndrome as notifiable conditions until 17

145   February of 2016 (65). The missing numbers of cases from July to December 2015 for

146   Brazil were estimated based on the positive correlation between *Ae. aegypti*

147   abundance (described below) and reported case counts as has been done previously

148   (50, 51). We used smoothing spline (64) to estimate weekly case counts from the

149   monthly reported counts. The weekly country level case counts (Figure 1A) were

150   divided by the total population / 100,000, as previously described (51), to compute

151   weekly *incidence rates* (Figure 1B).

152

153   *Travel Data*

154   Calibrated monthly passenger travel volumes for each airport-to-airport route in the

155   world were provided by the International Air Transport Associate (IATA) (76), as

156   previously used in (51, 66). The data includes origin, destination and stopover airport

157   paths for 84% of global air traffic, and includes over 240 airlines and 3,400 airports.

158   The airport level travel was aggregated to a regional level, to compute monthly

159   movements between all countries and territories in the Americas. The *incoming* and

160   *outgoing travel volumes* for each country and territory, originally available from

161   IATA at a monthly temporal resolution, were curve fitted, again using smoothing

162   spline method (64) to obtain corresponding weekly volumes to match with the

163   temporal resolution of our model. In this study, data and estimates from 2015 were

164   also used for 2016, as was done previously (51, 66, 67).

165

166   *Mosquito Suitability Data*

167   The monthly vector suitability data sets were based on habitat suitability for the

168   principal Zika virus species *Ae. aegypti,* previously used in (51), and initially

169   estimated using original high resolution maps (68) and then enriched to account for

170   seasonal variation in the geographical distribution of *Ae. aegypti* by using time-

171   varying covariate such as temperature persistence, relative humidity, and precipitation

172   as well as static covariates such as urban versus rural areas. The monthly data was

173   translated into weekly data using a smoothing spline (64).

174

175   *Socioeconomic and Human Population Data*

176   For a country, to prevent or manage an outbreak depends on their ability to implement

177   a successful surveillance and vector control programs (77). Due to a lack of global

178   data to quantify vector control at country level, we utilized alternative economic and

179   health related country indicators which have previously been revealed to be critical

180   risk factors for Zika spread (51). A country's economic development can be measured

181   by the *gross domestic product (GDP)* per capita at purchasing power parity (PPP), in

182   international dollars. The figures from World Bank (70) and the U.S. Bureau of

183   Economic Analysis (71) were used to collect GDP data for each country. The number

184   of *physicians* and the number of hospital *beds* per 10,000 people were used to indicate

185   the availability of health infrastructure in each country. These figures for U.S. and

186   other regions in the Americas were obtained from the Centre of Disease Control and

187   Prevention (CDC) (72), WHO World Health Statistics report (73), and the PAHO

188   (74). Finally, the human *population densities* (people per sq. km of land area) for each

189   region were collected from World Bank (75) and the U.S. Bureau of Economic

190   Analysis (71).

9

191   *Connectivity-risk Variables*

192   In addition to the raw input variables, novel connectivity-risk variables are defined and

193   computed for inclusion in the model. These variables are intended to capture the risk

194   posed by potentially infected travelers arriving at a given destination at a given point in

195   time, and in doing so, explicitly capture the dynamic and heterogeneity of the air-traffic

196   network in combination with real-time outbreak status. Two variables are chosen,

197   hereafter referred to as *case-weighted travel risk* and *incidence-weighted travel risk*, as

198   defined in equations (1.a) and (1.b), respectively.

199

200    $CR_j^t = \sum_i C_i^t . V_{i,j}^t \qquad \forall t, \forall j, i \neq j$ \qquad\qquad (1.a)

201

202    $IR_j^t = \sum_i I_i^t . V_{i,j}^t \qquad \forall t, \forall j, i \neq j$ \qquad\qquad (1.b)

203

204   For each region j at time t, CR_j^t and IR_j^t are computed as the sum of product between

205   passenger volume traveling from origin i into destination j at time t ($V_{i,j}^t$) and the state

206   of the outbreak at origin i at time t, namely reported cases, C_i^t, or reported incidence

207   rate, I_i^t. Each of these two variables is computed for all 53 countries or territories for

208   each of the 78 epidemiological weeks. The two dynamic variables, CR_j^t and IR_j^t, are

209   illustrated in Figure 1C and 1D, below the raw case counts and incidence rates,

210   respectively.

211
212   **Neural Network Model**

213   A class of neural architectures based upon Nonlinear Auto Regressive models with

214   eXogenous inputs (NARX) known as NARX neural networks (78-80) is employed

215   herein due to its suitability for modeling of a range of nonlinear systems and

10

216   computational capabilities equivalent to Turing machines (81). The NARX networks,

217   as compared to other recurrent neural network architectures, require limited feedback

218   (i.e., feedback from the output neuron rather than from hidden states) and converge

219   much faster with a better generalization (81, 82). The NARX model can be formalized

220   as follows (81):

221    $$y(t) = f\left(x(t), x(t-1), \dots, x(t-d_x); y(t-1), \dots, y(t-d_y)\right) \qquad (2)$$

222   where $x(t)$ and $y(t)$ denote, respectively, the input and output (or target that should

223   be predicted) of the model at discrete time t, while d_x and d_y (with $d_x \geq 1$, $d_y \geq 1$,

224   and $d_x \leq d_y$) are input and output delays called memory orders (Figure 2). In this

225   work, a NARX model is implemented to provide N-step ahead prediction of a time

226   series, as defined below:

227

228    $y_k(t+N) =$

229    $$f\left(\begin{matrix} \mathbf{x_1}(t), \mathbf{x_1}(t-1), \dots, \mathbf{x_1}(t-d_x), \dots, \mathbf{x_M}(t), \mathbf{x_M}(t-1), \dots, \mathbf{x_M}(t-d_x), \\ y_k(t), y_k(t-1), \dots, y_k(t-d_y) \end{matrix}\right) \qquad (3)$$

230

231   Here, $y_k(t+N)$ is the risk classification predicted for the kth region N weeks ahead

232   (of present time t), which is estimated as a function of $\mathbf{x_m}(t)$ inputs from all $m =$

233    $1, 2, \dots, M$ regions for d_x previous weeks, and the previous risk classification state,

234    $y_k(t)$ for region k for d_y previous weeks. The prediction model is applied at time t,

235   to predict for time $t+N$, and therefore relies on data available up until week t. That is,

236   to predict outbreak risk for epidemiological week X, N-weeks ahead, the model is

237   trained and tested using data available up until week (X − N). For example, 12-week

238   ahead prediction for Epi week 40, is performed using data available up to week 28.

239   The function $f(\cdot)$ is an unknown nonlinear mapping function that is approximated by

240   a Multilayer Perceptron (MLP) to form the NARX recurrent neural network (79, 80).

241   In this work, series-parallel NARX neural network architecture is implemented in

242   Matlab R2018a (The MathWorks, Inc., Natick, Massachusetts, United States) (83).

243

244   In the context of this work, the desired output, $y_k(t + N)$, is a binary risk classifier,

245   *i.e*., classifying a region k as high or low risk at time at time $t+N$, for each region, k, N

246   weeks ahead (of t). The vector of input variables for region m at time t is $\mathbf{x}_m(t)$, and

247   includes both static and dynamic variables. We consider various relative (R) and

248   absolute (A) thresholds to define the set of "high risk" countries at any point in time.

249   We define relative risk thresholds that range uniformly between 10% and 50%, where

250   the 10% scheme classifies the 10% of countries reporting the highest number of cases

251   (or highest incidence rate) during a given week as high risk, and the other 90% as low

252   risk, similar to (45). The relative risk schemes are referred herein as R=0.1, R=0.2,

253   R=0.3, R=0.4, and R=0.5. It is worth noting, for a given percentile, *e.g.,* R=0.1, the

254   relative risk thresholds are dynamic and vary week to week as a function of the scale

255   of the epidemic, while the size of the high risk group remains fixed over time, e.g.,

256   10% of all countries. We also consider absolute thresholds, which rely on case

257   incidence rates to define the "high risk" group. Five absolute thresholds are selected

258   based on the distribution of incidence values over all countries and the entire

259   epidemic. Specifically, the 50th, 60th, 70th, 80th and 90th percentiles were chosen,

260   and are referred herein as A=50, A=60, A=70, A=80, and A=90. These five thresholds

261   correspond to weekly case incidence rates of 0.43, 1.47, 4.05, 9.5 and 32.35 (see

262   Additional file 12: Figure S1), respectively. In contrast to the relative risk scheme,

263   under the absolute risk scheme for a given percentile, *e.g*., A=90, the threshold

264   remains fixed but the size of the high (and low) risk group varies week to week based

12

265   on the scale of the epidemic. The fluctuation in group size for each threshold is

266   illustrated in Additional file 12: Figure S1 for each classification scheme, A=50 to

267   A=90. Critically, our prediction approach differs from (45), in that our model is

268   trained to predict the risk level directly, rather than predict the number of cases, which

269   are post-processed into risk categories. The performance of the model is evaluated by

270   comparing the estimated risk level (high or low) to the actual risk level for all

271   locations at a specified time. The actual risk level is simply defined at each time

272   period t during the outbreak by ranking the regions based on to the number of

273   reported case counts (or incidence rates), and grouping them into high and low risk

274   groups according to the specified threshold and classification scheme.

275

276   The static variables used in the model include GDP PPP, population density, number

277   of physicians, and number of hospital beds for each region. The dynamic variables

278   include mosquito vector suitability, outbreak status (both reported case counts and

279   reported incidence rates), total incoming travel volume, total outgoing travel volume,

280   and the two connectivity-risk variables defined as in Equations (1.a) & (1.b), again for

281   each region. Before applying to the NARX model, all data values are normalized to

282   the range [0, 1].

283

284   A major contribution of this work is the flexible nature of the model, which allows

285   policy makers to be more or less risk averse in their planning and decision making.

286   Firstly, the risk indicator can be chosen by the modeler; in this work we consider two

287   regional risk indicators, i) the number of reported cases and ii) incidence rate. Second,

288   we consider a range of risk classification schemes, which define the set of high-risk

289   countries based on either a relative or absolute threshold that can be chosen at the

13

290   discretion of the modeler, *i.e.,* R=0.1, 0.2, 0.3, 0.4, 0.5, and A=90, 80, 70, 60, 50.

291   Third, the forecast window, *N*, is defined to range from $N = 1, 2, 4, 8$ and 12 weeks.

292   Subsequently, any combination of risk indicator, risk classification scheme and

293   forecasting window can be modelled.

294

295   In initial settings of the series-parallel NARX neural network, a variety numbers of

296   hidden layer neurons and numbers of tapped delay lines (Eq. (2)) were explored for

297   training and testing of the model. Sensitivity analysis revealed minimal difference in

298   performance of the model under different settings. Therefore, for all experiments

299   presented in this work, the numbers of neural network hidden layer neurons and

300   tapped delay lines are kept constant as two and four, respectively.

301

302   To train and test the model, the actual risk classification for each region at each week

303   during the epidemic, $y_k(t)$, was used. For each model run, e.g., a specified risk

304   indicator, risk classification scheme and forecasting window, the input and target

305   vectors are randomly divided into three sets:

306

307   1. 70% for training, to tune model parameters minimizing the mean square error

308   between the outputs and targets,

309   2. 15% for validation, to measure network generalization and to prevent

310   overfitting, by halting training when generalization stops improving (i.e.,

311   mean square error of validation samples starts increasing), and

312   3. 15% for testing, to provide an independent measure of network performance

313   during and after training.

314

315   The performance of the model is measured using two metrics: 1) prediction accuracy

316   (ACC) and 2) receiver operating characteristic (ROC) curves. Prediction accuracy is

317   defined as ACC = (TP + TN) / (TP + FP + TN + FN), where true positive (TP) is the

318   number of high risk locations correctly predicted as high risk, false negative (FN) is

319   the number of high risk locations incorrectly predicted as low risk, true negative (TN)

320   is the number of low risk locations correctly predicted as low risk, and false positive

321   (FP) is the number of low risk locations incorrectly predicted as high risk. The second

322   performance metric, ROC curve (84), explores the effects on TP and FP as the

323   position of an arbitrary decision threshold is varied, which in the context of this

324   prediction problem distinguished low and high risk locations. ROC curve can be

325   characterized as a single number using the area under the ROC curve (AUC), with

326   larger areas having an AUC that approaches one indicating a more accurate detection

327   method. In addition to quantifying model performance using these two metrics, we

328   evaluate the robustness of the predictions by comparing the ACC across multiple runs

329   that vary in their selection of testing and training sets (resulting from the randomized

330   sampling).

331

332   **Results**

333   The model outcome reveals the set of locations expected to be at high risk at a

334   specified date in the future, *i.e., N* weeks ahead of when the prediction is made. We

335   apply the model for all epidemiological weeks throughout the epidemic, and evaluate

336   performance under each combination of i) risk indicator, ii) classification scheme, and

337   iii) forecast window. For each model run, both ACC and ROC AUC are computed.

338

339

340   *Model Performance*

341   Figures 3 and 4 exemplify the output of the proposed model. Figure 3 illustrates the

342   model predictions at a country-level for a *4*-week prediction window, specifically for

343   Epi week 40, *i.e*., using data available up until week 36. Figure 3A illustrates the

344   actual risk percentile each country is assigned to in week 40, based on reported case

345   counts. The results presented in the remaining panels of Figure 3 reveal the risk level

346   (high or low) predicted for each country under the five relative risk classification

347   schemes, namely (B) R=0.1, (C) R=0.2, (D) R=0.3, (E) R=0.4, and (F) R=0.5, and

348   whether or not it was correct. For Panels (B)-(E), green indicates a correctly predicted

349   low risk country (TN), light grey indicates an incorrectly predicted high risk country

350   (FP), dark grey indicates an incorrectly predicted low risk country (FN), and the

351   remaining color indicates a correctly predicted high risk country (TP). The inset

352   highlights the results for the *Caribbean* islands. The figure also presents the average

353   ACC over all regions and ACC for just the Caribbean region (grouped similar to (10))

354   for each classification scheme.

355

356   Figure 4 illustrates the model predictions at a country-level for varying prediction

357   windows, and a fixed classification scheme of R=0.2, again for Epi week 40. Figure

358   4A illustrates the actual risk classification (high or low) each country is assigned to in

359   Epi week 40, based on reported case counts. The results presented in the remaining

360   panels of Figure 4 reveal the risk level (high or low) predicted for each country under

361   the five forecasting windows, specifically (B) N=1, (C) N=2, (D) N=4, (E) N=8, and

362   (F) N=12, and whether or not it was correct. For Panels (B)-(E), red indicates a

363   correctly predicted high risk country (TP), green indicates a correctly predicted low

364   risk country (TN), light grey indicates an incorrectly predicted high risk country (FP),

16

365   dark grey indicates an incorrectly predicted low risk country (FN). The inset

366   highlights the results for the *Caribbean* islands. Similar to Figure 3, for each forecast

367   window, the reported ACC is averaged both over all regions and for just the

368   Caribbean.

369

370   The model's performance and sensitivity to the complete range of input parameters is

371   summarized in Additional file 13: Table S2**.** ACC is presented for each combination

372   of risk indicator (case count and incidence rate), classification scheme (i.e., R = 0.1,

373   0.2, 0.3, 0.4, 0.5 and A = 90, 80, 70, 60, 50) and forecast window (i.e., N = 1, 2, 4, 8

374   and 12), for selected Epi weeks throughout the epidemic. ROC AUC (averaged over

375   all locations and all EPI weeks) is computed for all combinations of risk indicator

376   (case count and incidence rate), classification scheme (i.e., R = 0.1, 0.2, 0.3, 0.4, 0.5

377   and A = 90, 80, 70, 60, 50) and forecast window (i.e., N = 1, 2, 4, 8 and 12).

378
379   Figures 5 and 6 illustrate trends in the model performance as a function of

380   classification scheme and forecast window, aggregated over space and time.

381   Specifically, Figure 5 reveals the model performance (ACC, averaged over all

382   locations and all EPI weeks) for each combination of risk classification scheme (i.e.,

383   R = 0.1, 0.2, 0.3, 0.4 and 0.5) and forecast window (i.e., N = 1, 2, 4, 8 and 12). The

384   aggregated ROC curves (averaged over all locations and all epidemiological weeks)

385   for R=0.4 are presented in Figure 6, and reveal the (expected) increased accuracy of

386   the model as the forecast window is reduced. The ROC AUC results are consistent

387   with ACC results presented in Figure 5, highlighting the superior performance of the

388   1 and 2 week ahead prediction capability of the model. The ROC AUC value remains

389   above 0.91 for N=1, 2 and above 0.83 for N=4, both indicating high predictive

17

390   accuracy of the model. The ROC curves for the other relative risk classification

391   schemes are presented in Additional file 14: Figure S2.

392

393   *Global and Regional Analysis*

394   We further explore the model's performance at a regional level by dividing the

395   countries and territories in the Americas into three groups, namely *Caribbean*, *South*

396   *America* and *Central America,* as in (10), and compare with the *Global* performance,

397   *i.e.,* all countries. For each group the average performance of the model in terms of

398   ACC was evaluated and presented for each combination of risk indicator (case count

399   and incidence rate), classification scheme (i.e., R = 0.1, 0.2, 0.3, 0.4, 0.5 and A = 90,

400   80, 70, 60, 50) and forecast window (i.e., *N* = 1, 2, 4, 8 and 12), aggregated over then

401   entire epidemic period (Table 2).

402   **Table 2. Summary of Global and Regional Model Performance**

Relative Risk Classification Scheme	Prediction Window Size (*N* in weeks)	Overall Prediction Accuracy (ACC)							
		Global		Caribbean		South America		Central America	
		Risk Indicator		Risk Indicator		Risk Indicator		Risk Indicator	
		incidence	cases	incidence	cases	incidence	cases	incidence	cases
R=0.1	1	95.71	96.95	94.63	98.84	93.65	92.28	97.95	94.18
	2	94.29	96.12	92.90	98.86	91.68	90.78	97.01	90.67
	4	91.30	93.13	89.38	97.30	87.11	84.80	95.34	84.14
	8	86.34	90.63	85.72	95.70	74.58	81.97	91.74	76.69
	12	82.57	87.05	81.75	93.59	68.63	75.94	87.99	68.14
R=0.2	1	93.07	93.54	91.73	94.94	92.65	90.16	92.64	87.33
	2	90.01	92.27	88.30	93.93	89.37	88.60	89.26	84.68
	4	84.68	88.09	82.66	89.72	82.77	84.40	82.28	76.49
	8	75.22	81.87	71.58	83.96	69.34	76.27	73.73	65.25
	12	68.96	78.25	65.01	80.92	62.75	71.30	63.73	58.09
R=0.3	1	90.70	93.41	88.30	94.05	91.41	91.41	90.58	87.84
	2	86.74	89.82	85.27	91.06	86.68	86.68	84.15	80.46
	4	80.85	84.31	77.10	85.36	82.63	79.38	78.73	72.76
	8	70.10	76.46	64.73	77.31	69.34	71.34	66.31	58.05
	12	63.37	71.66	56.86	70.66	62.39	64.88	57.35	56.86
R=0.4	1	90.46	91.68	88.25	91.31	93.03	90.54	87.84	86.64
	2	86.79	88.52	84.62	88.24	89.76	86.30	83.10	81.51
	4	79.36	81.67	76.41	81.97	83.04	75.44	72.57	71.83
	8	68.47	72.85	61.67	71.12	73.19	71.19	62.29	54.66
	12	59.82	65.22	51.89	60.33	62.21	60.96	53.19	53.43
R=0.5	1	89.51	91.16	89.67	90.25	87.42	88.42	85.10	90.41
	2	86.21	86.90	84.83	86.13	85.15	82.84	83.10	84.15
	4	77.67	78.46	76.29	77.55	75.44	75.71	70.71	67.72
	8	66.42	68.05	61.99	65.78	69.80	68.26	53.60	47.46
	12	56.16	58.31	48.04	51.81	62.21	54.90	43.38	46.81

403

404   *Model Robustness*

405   Figure 7A and 7B show how the ACC varies over 10 independent runs of the model.

406   This sensitivity analysis was conducted for all combinations risk indicator, relative

407   risk classification schemes, and selected epidemiological weeks (i.e., week number /

408   starting date: 30 / 18-Jan-2016, 40 / 28-Mar-2016, 50 / 6-Jun-2016, 60 / 15-Aug-2016,

409   and 70 / 24-Oct-2016). This time period represents a highly complex period of the

410   outbreak with country level rankings fluctuating substantially, as evidenced in Figure

411   1. Due to computation time, the sensitivity analysis was evaluated for only the 4-

412   week forecast window. The size of the error bars illustrates the robustness of the

413   proposed modeling framework.

414

415   *NARX Feature Selection*

416   While the NARX framework does not provide assigned weights for each input feature

417   as output, sensitivity analysis can be conducted to help identify the key predictive

418   features. We tested the performance of the NARX framework under three different

419   combinations of input features, with the particular objective of quantifying the role of

420   travel data in our outbreak prediction model. We considered i) a simple 'baseline'

421   model using only case count and incidence data, ii) an expanded baseline model that

422   includes case and incidence data, and all non-travel related variables, and iii) the

423   proposed model which includes all features listed in Table 1. The results comparing

424   the performance of these three models with the detailed list of input features for each

425   is provided in Additional file 15: Table S1. The results reveal the case-related data

426   (regional case counts and incidence rates) to be the dominant explanatory variables

427   for predicting outbreak risk in a region, as would be expected. The inclusion of non-

428   travel related variables (regional suitability, regional GDP, regional physicians,

19

429   regional hospital beds, regional population density) is not shown to improve

430   predictive capability over the baseline model, and in fact, sometime performs worse

431   than the baseline model. In contrast, the inclusion of travel data (weekly case-

432   weighted travel risk, weekly incidence-weighted travel risk, weekly incoming travel

433   volume, weekly outgoing travel volume) is revealed to improve the predictive

434   capability, especially for the shorter prediction windows, with a higher AUC ROC for

435   a majority (20 of the 25) of the scenarios tested. These results support the inclusion

436   of the dynamic travel-related variables, which substantially increase the complexity of

437   the model (inputs), and thus, justifies the use of the NARX framework selected.

438

439   **Discussion**

440   Overall, the proposed model is shown to be accurate and robust, especially for shorter

441   prediction windows and higher risk thresholds. As would be expected, the

442   performance of the proposed model decreases as the prediction window increases

443   because of the inherent uncertainty in outbreak evolution over long periods of time.

444   Specifically, the model is almost 80% accurate for 4-week ahead prediction for all

445   classification schemes, and almost 90% accurate for all 2-week ahead prediction

446   scenarios, *i.e.,* the correct risk category of 9 out of 10 locations can always be

447   predicted, indicating strong performance. Although, when the objective is to identify

448   the top 10% of at-risk regions, the average accuracy of the model remains above 87%

449   for prediction up to *12*-weeks in advance. Generally, the model performance is

450   shown to decrease as the risk threshold is reduced, *e.g.*, the size of the high risk group

451   is increased, representing a more risk averse policy. The decrease in performance is

452   likely due to the increased size and fluctuation of the high risk country set over time

453   for lower thresholds. For example, for the absolute risk threshold of A=50, the

20

454　number of countries classified as high risk fluctuates between 1 and 34 throughout the

455　course of the epidemic, compared with A=90, where the set only ranges from 0 to 12

456　(see Additional file 12: Figure S1). These results reveal the trade-off between desired

457　forecast window and precision of the high risk group. The quantifiable trade-off

458　between the two model inputs (classification scheme and forecast window) can be

459　useful for policies which may vary in desired planning objectives.

460

461　The results in Figures 3 and 4, as well as Table 2 reveal a similar trend at the regional

462　level as was seen at the global level, with a decrease in predictive accuracy as the

463　forecast window increases in length, and the and high risk group increases in size.

464　As shown in Figure 3, the ACC remains above 90% for R < 0.3, indicating superior

465　model performance. For example, at Epi week 40, R = 0.3 and N=4 (using outbreak

466　data and other model variables up to Epi week 36), there were 16 total regions

467　classified as high risk, of which the model correctly identified 13. Furthermore, of the

468　16 high risk regions, 8 were in the *Caribbean* (i.e., Aruba, Curacao, Dominican

469　Republic, Guadeloupe, Haiti, Jamaica, Martinique, and Puerto Rico), of which the

470　model correctly identified 7. Aruba in the only Caribbean, and Honduras and Panama

471　were the only regions incorrectly predicted as low risk in this scenario; accurately

472　classifying low risk regions is also important (and assuring the model is not too risk

473　averse). For the same scenario, *i.e.,* Epi week 40, R = 0.3 and N=4, all 18 low risk

474　*Caribbean* locations and 17 of the 19 low risk non-*Caribbean* locations were

475　accurately classified by the model. Paraguay and Suriname were the only regions

476　incorrectly predicted as high risk. These results are consistent with the high reported

477　accuracy of the model, i.e., Overall ACC = 90.15%; *Caribbean* ACC = 96.15%.

478

479 Figure 4 reveals that the performance of model, expectedly, deteriorates as the

480 forecast window increases; however, the average accuracy remains above 80% for

481 prediction up to *8*-weeks ahead, and well about 90% for up to 4-weeks ahead. The

482 prediction accuracy for the Caribbean slightly lags the average performance in the

483 Americas. Specifically, for R=0.2, 5 of the 11 *Caribbean* regions were designated as

484 HIGH risk locations at Epi week 40*, i.e.,* Dominican Republic, Guadeloupe, Jamaica,

485 Martinique, Puerto Rico. For a one-week prediction window, N=1, the model was

486 able to correctly predict 3 of the high risk regions (*i.e.,* Jamaica, Martinique, Puerto

487 Rico), for N=2 it correctly identified two (*i.e.,* Martinique, Puerto Rico), and for N=4,

488 it again correctly identified three (*i.e.,* Guadeloupe, Martinique, Puerto Rico).

489 However, the model did not correctly predict any high risk locations in the Caribbean

490 at N=8 and N=12 window lengths. This error is due to the low and sporadic reporting

491 of Zika cases in the region around week 30, and the high variability of the outbreak

492 over the 8 and 12 week period. Similar prediction capability is illustrated for R=0.5

493 (not shown in the figure), in which case out of the 13 *Caribbean* HIGH risk locations,

494 the model correctly identifies all locations at N=1, 2 and 4, 10 of the 13 locations at

495 N=8, and only 1 of the 13 at N=12.

496

497 When comparing performance across regions (see Table 2) results reveal the

498 predictive accuracy is best for the *Caribbean* region, while predictions for *Central*

499 *America* were consistently the worst; the discrepancy in performance between these

500 groups increases as the forecast window increases. The difference in performance

501 across regions can be attributed to the high spatial heterogeneity of the outbreak

502 patterns, the relative ability of air travel to accurately capture connectivity between

503 locations, and errors in case reporting that may vary by region. For example, the

22

504   *Caribbean,* which consists of more than twice as many locations as any other group,

505   first reported cases around week 25, and remained affected throughout the epidemic.

506   In contrast, *Central America* experienced a slow start to the outbreak (at least

507   according to case reports) with two exceptions, namely Honduras and El Salvador.

508   The large number of affected region in the Caribbean, with more reported cases

509   distributed over a longer time period contributed to the training of the model, thus

510   improving the predictive capability for these regions. Additionally, the geographically

511   isolated nature of Caribbean islands enables air travel to more accurately capture

512   incoming travel risk, unlike countries in Central and South America, where

513   individuals can also move about using alternative modes, which are not accounted for

514   in this study. These factors combined explain the higher predictive accuracy of the

515   model for the Caribbean region, and importantly, helps to identify the critical features

516   and types of settings under which this model is expected to perform best.

517

518   Finally, the robustness of the model predictions is illustrated by the short error bars in

519   Figure 7. The model is also demonstrated to perform consistently throughout the

520   course of the epidemic, with the exception of week 30, at which time there was

521   limited information available to train the model, *e.g.,* the outbreak was not yet

522   reported in a majority of the affected countries. Comparing Figure 7A and 7B reveals

523   relatively similar performance for both risk indicators, and Additional File 13: Table 2

524   demonstrating the model's flexibility and adaptability with respect to both the risk

525   scheme chosen, *i.e.,* relative or absolute, and the metric used to classify outbreak risk,

526   *i.e.,* number of cases or incidence rate in a region.

527

528

529   *Limitations*

530   There are several limitations of this work. The underlying data on case reporting vary

531   by country and may not represent the true transmission patterns (85). However, the

532   framework presented was flexible enough to account for these biases and we

533   anticipate will only be improved as data become more robust. Additionally, 2015

534   travel data was used in place of 2016 data, as has been done previously (51, 66, 67),

535   which may not be fully representative of travel behaviour. Furthermore, air travel is

536   the only mode of travel accounted for, thus, additional person movements between

537   country pairs that share land borders are unaccounted for, and as a result, the model

538   likely underestimates the risk posed to some regions. This limitation may partially

539   explain the increased model performance for the geographically isolated Caribbean

540   Islands, which represent a large proportion of ZIKV affected regions. This study does

541   not account for species of mosquitos other than *Ae. Aegypti,* such as *Ae. Albopictus*,

542   which can also spread ZIKV; however, *Ae. Aegypti* are known to be the primary

543   spreading vector, and responsible for the majority of the ZIKV epidemic in the

544   Americas (86). Additionally, alternative non-vector-borne mechanisms of

545   transmission are ignored. Lastly, due to the lack of spatial resolution of case reports,

546   we were limited to make country to country spread estimates. We do however

547   appreciate that there is considerable spatial variation within countries (i.e., northern

548   vs. southern Brazil) and that this may influence the weekly covariates used in this

549   study. We again hypothesise that models will become better as the spatial resolution

550   of available data increases.

551

552

553

554 **Conclusions**

555 We have introduced a flexible, predictive modelling framework to forecast outbreak

556 risk in real-time that can be scaled and readily applied in future outbreaks. An

557 application of the model was applied to the Zika epidemic in the Americas at a

558 weekly temporal resolution, and country-level spatial resolution, using a combination

559 of population, socioeconomic, epidemiological, travel patterns and vector suitability

560 data. The model performance was evaluated for various risk classification schemes,

561 forecast windows and risk indicators, and illustrated to be accurate and robust across a

562 broad range of these features. First, the model is more accurate for shorter prediction

563 windows and restrictive risk classification schemes. Secondly, regional analysis

564 reveals superior predictive accuracy for the Caribbean, suggesting the model to be

565 best suited to geographically isolated locations that are predominantly connected via

566 air travel. Predicting the spread to areas that are relatively isolated has previously

567 been shown to be difficult due to the stochastic nature of infectious disease spread

568 (87). Thirdly, the model performed consistently well at various stages throughout the

569 course of the outbreak, indicating its potential value at the early stages of an epidemic.

570 The outcomes from the model can be used to better guide outbreak resource allocation

571 decisions, and can be easily adapted to model other vector-borne epidemics.

572

573

574 **Additional files**

575 **Additional file 1: Data (cases).** Country or territory level weekly Zika cases.
576
577 **Additional file 2: Data (incidence).** Country or territory level weekly Zika incidence
578 rates.
579
580 **Additional file 3: Data (incoming_travel).** Country or territory level, weekly
581 incoming travel volume.
582
583 **Additional file 4: Data (outgoing_travel).** Country or territory level weekly
584 outgoing travel volume.
585
586 **Additional file 5: Data (suitability).** Country or territory level weekly *Aedes* vector
587 suitability.
588
589 **Additional file 6: Data (gdp).** Country or territory level GDP per capita.
590
591 **Additional file 7: Data (physicians).** Country or territory level physicians per 1000
592 people.
593
594 **Additional file 8: Data (beds).** Country or territory level beds per 1000 people.
595
596 **Additional file 9: Data (pop_density).** Country or territory level population densities
597 (people per sq. km of land area).
598
599 **Additional file 10: Data (case_weighted_travel_risk).** Country or territory level
600 weekly case-weighted travel risk.
601
602 **Additional file 11: Data (incidence_weighted_travel_risk).** Country or territory
603 level weekly incidence-weighted travel risk.
604
605 **Additional file 12: Figure S1. Number of high risk countries each week under all**
606 **absolute risk classification schemes.** The number of countries classified as high risk
607 each week for each absolute case incidence threshold, ranging from A=50 to A=90. In
608 parentheses is the weekly incidence rate defining the high risk threshold based on the
609 percentile (A) specified.
610
611 **Additional file 13: Table S2. Summary of model performance.** ACC is presented
612 for each combination of risk indicator (case count and incidence rate), classification
613 scheme (i.e., R = 0.1, 0.2, 0.3, 0.4, 0.5 and A = 90, 80, 70, 60, 50) and forecast
614 window (i.e., N = 1, 2, 4, 8 and 12), for selected Epi weeks throughout the epidemic.
615 ROC AUC (averaged over all locations and all EPI weeks) is computed for all
616 combinations of risk indicator (case count and incidence rate), classification scheme
617 (i.e., R = 0.1, 0.2, 0.3, 0.4, 0.5 and A = 90, 80, 70, 60, 50) and forecast window (i.e.,
618 N = 1, 2, 4, 8 and 12).
619
620

621 **Additional files 14: Figure S2. Aggregate model performance measured by ROC**
622 **AUC.** The ROC AUC is averaged over all locations and all weeks, for each relative
623 risk classification scheme, *i.e*., R = 0.1, 0.2, 0.3, 0.4, 0.5 and forecast window *i.e., N =*
624 1, 2, 4, 8 and 12. For the results shown the risk indicator is case counts.
625
626 **Additional file 15: Table S1. Summary of model sensitivity to feature selection.**
627 The ACC and ROC AUC performance of the model is computed and presented under
628 different combinations of input data features. The proposed model is compared
629 against two baseline models; one includes only case (and incidence) data, and the
630 second includes case and all non-travel related data, while the final proposed model
631 includes all features. The results presented are for the absolute risk classification
632 scheme, where the risk indicator is incidence rate.
633
634 **Abbreviations**

635 **ACC:** Prediction accuracy

636 **AUC:** Area under the curve

637 **CDC:** Centre of disease control and prevention

638 **FN:** False negative

639 **FP:** False positive

640 **GDP:** Gross domestic product

641 **IATA:** International air transport associate

642 **MLP:** Multilayer perceptron

643 **NARX:** Nonlinear autoregressive models with exogenous inputs

644 **PAHO:** Pan American health organization

645 **PPP:** Purchasing power parity

646 **ROC:** Receiver operating characteristic

647 **TN:** True negative

648 **TP:** True positive

649 **ZIKV:** Zika virus

650

651 **Ethics approval and consent to participate.** Not applicable.

652 **Consent for publication.** Not applicable.

653

654 **Availability of data and material.** All data used in this study is provided as
655 Additional files.

656

657 **Competing Interests.** We have no competing interests.

658

659 **Authors' Contributions.** LG and MA conceived the study, designed the experiments,
660 analyzed the model results, and drafted the original manuscript. MA developed the
661 model and performed the computational analysis. MUGK contributed vector
662 distribution data. All authors contributed to data curation and editing of the
663 manuscript. LG supervised the study.

668

669

670

References

1. Chouin-Carneiro T, Vega-Rua A, Vazeille M, Yebakima A, Girod R, Goindin D, et al. Differential Susceptibilities of Aedes aegypti and Aedes albopictus from the Americas to Zika Virus. PLoS Negl Trop Dis. 2016;10(3):1-11.

2. Dick GW. Zika virus. II. Pathogenicity and physical properties. Trans R Soc Trop Med Hyg. 1952;46(5):521-34.

3. Duffy MR, Chen TH, 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;360(24):2536-43.

4. Hancock WT, Marfel M, Bel M. Zika virus, French Polynesia, South Pacific, 2013. Emerg Infect Dis. 2014;20(11):1960.

5. Dupont-Rouzeyrol M, O'Connor O, Calvez E, Daures M, John M, Grangeon JP, et al. Co-infection with Zika and dengue viruses in 2 patients, New Caledonia, 2014. Emerg Infect Dis. 2015;21(2):381-2.

6. Musso D, Nilles EJ, Cao-Lormeau VM. Rapid spread of emerging Zika virus in the Pacific area. Clin Microbiol Infect. 2014;20(10):O595-6.

7. Tognarelli J, Ulloa S, Villagra E, Lagos J, Aguayo C, Fasce R, et al. A report on the outbreak of Zika virus on Easter Island, South Pacific, 2014. Arch Virol. 2016;161(3):665-8.

8. Faria NR, Azevedo RdSdS, Kraemer MUG, Souza R, Cunha MS, Hill SC, et al. Zika virus in the Americas: Early epidemiological and genetic findings. Science. 2016.

9. Campos GS, Bandeira AC, Sardi SI. Zika Virus Outbreak, Bahia, Brazil. Emerg Infect Dis. 2015;21(10):1885-6.

10. PAHO. Regional Zika Epidemiological Update (Americas). In: World Health Organization PAHO, editor. Washington DC2017.

11. Zanluca C, Melo VC, Mosimann AL, Santos GI, Santos CN, Luz K. First report of autochthonous transmission of Zika virus in Brazil. Mem Inst Oswaldo Cruz. 2015;110(4):569-72.

12. Scott TW, Morrison AC. Vector dynamics and transmission of dengue virus: implications for dengue surveillance and prevention strategies: vector dynamics and dengue prevention. Current topics in microbiology and immunology. 2010;338:115-28.

13. Achee NL, Gould F, Perkins TA, Reiner RC, Jr., Morrison AC, Ritchie SA, et al. A critical assessment of vector control for dengue prevention. PLoS Negl Trop Dis. 2015;9(5):e0003655.

14. Vector control with a focus on Aedes aegypti and Aedes albopictus mosquitoes: literature review and analysis of information. Stockholm: ECDC: European Centre for Disease Prevention and Control; 2017.

15. McGough SF, Brownstein JS, Hawkins JB, Santillana M. Forecasting Zika Incidence in the 2016 Latin America Outbreak Combining Traditional Disease Surveillance with Search, Social Media, and News Report Data. PLoS Negl Trop Dis. 2017;11(1):e0005295.

16. Martínez-Bello DA, López-Quílez A, Torres-Prieto A. Bayesian dynamic modeling of time series of dengue disease case counts. PLOS Neglected Tropical Diseases. 2017;11(7):e0005696.

17. Guo P, Liu T, Zhang Q, Wang L, Xiao J, Zhang Q, et al. Developing a dengue forecast model using machine learning: A case study in China. PLOS Neglected Tropical Diseases. 2017;11(10):e0005973.

29

18.   Johansson MA, Reich NG, Hota A, Brownstein JS, Santillana M. Evaluating the performance of infectious disease forecasts: A comparison of climate-driven and seasonal dengue forecasts for Mexico. Sci Rep. 2016;6:33707.

19.   Earnest A, Tan SB, Wilder-Smith A, Machin D. Comparing Statistical Models to Predict Dengue Fever Notifications. Computational and Mathematical Methods in Medicine. 2012;2012:6.

20.   Hii YL, Zhu H, Ng N, Ng LC, Rocklöv J. Forecast of Dengue Incidence Using Temperature and Rainfall. PLOS Neglected Tropical Diseases. 2012;6(11):e1908.

21.   Shi Y, Liu X, Kok SY, Rajarethinam J, Liang S, Yap G, et al. Three-Month Real-Time Dengue Forecast Models: An Early Warning System for Outbreak Alerts and Policy Decision Support in Singapore. Environ Health Perspect. 2016;124(9):1369-75.

22.   Teng Y, Bi D, Xie G, Jin Y, Huang Y, Lin B, et al. Dynamic Forecasting of Zika Epidemics Using Google Trends. PLOS ONE. 2017;12(1):e0165085.

23.   Althouse BM, Ng YY, Cummings DAT. Prediction of Dengue Incidence Using Search Query Surveillance. PLOS Neglected Tropical Diseases. 2011;5(8):e1258.

24.   Morsy S, Dang TN, Kamel MG, Zayan AH, Makram OM, Elhady M, et al. Prediction of Zika-confirmed cases in Brazil and Colombia using Google Trends. Epidemiology and Infection. 2018;146(13):1625-7.

25.   Kraemer MUG, Faria NR, Reiner RC, Jr., Golding N, Nikolay B, Stasse S, et al. Spread of yellow fever virus outbreak in Angola and the Democratic Republic of the Congo 2015-16: a modelling study. The Lancet Infectious diseases. 2017;17(3):330-8.

26.   Zhang Q, Sun K, Chinazzi M, Pastore YPA, Dean NE, Rojas DP, et al. Spread of Zika virus in the Americas. Proc Natl Acad Sci U S A. 2017;114(22):E4334-E43.

27.   Ahmadi S, Bempong N-E, De Santis O, Sheath D, Flahault A. The role of digital technologies in tackling the Zika outbreak: a scoping review. Journal of Public Health and Emergency. 2018;2(6).

28.   Majumder MS, Santillana M, Mekaru SR, McGinnis DP, Khan K, Brownstein JS. Utilizing Nontraditional Data Sources for Near Real-Time Estimation of Transmission Dynamics During the 2015-2016 Colombian Zika Virus Disease Outbreak. JMIR Public Health Surveill. 2016;2(1):e30.

29.   Beltr JD, #225, Boscor A, Santos WPd, Massoni T, Kostkova P. ZIKA: A New System to Empower Health Workers and Local Communities to Improve Surveillance Protocols by E-learning and to Forecast Zika Virus in Real Time in Brazil. Proceedings of the 2018 International Conference on Digital Health; Lyon, France. 3194683: ACM; 2018. p. 90-4.

30.   Cortes F, Turchi Martelli CM, Arraes de Alencar Ximenes R, Montarroyos UR, Siqueira Junior JB, Goncalves Cruz O, et al. Time series analysis of dengue surveillance data in two Brazilian cities. Acta Trop. 2018;182:190-7.

31.   Abdur Rehman N, Kalyanaraman S, Ahmad T, Pervaiz F, Saif U, Subramanian L. Fine-grained dengue forecasting using telephone triage services. Science Advances. 2016;2.

32.   Lowe R, Stewart-Ibarra AM, Petrova D, Garcia-Diez M, Borbor-Cordova MJ, Mejia R, et al. Climate services for health: predicting the evolution of the 2016 dengue season in Machala, Ecuador. Lancet Planet Health. 2017;1(4):e142-e51.

33.   Ramadona AL, Lazuardi L, Hii YL, Holmner A, Kusnanto H, Rocklov J. Prediction of Dengue Outbreaks Based on Disease Surveillance and Meteorological Data. PLoS One. 2016;11(3):e0152688.

771   34.   Lauer SA, Sakrejda K, Ray EL, Keegan LT, Bi Q, Suangtho P, et al.
772   Prospective forecasts of annual dengue hemorrhagic fever incidence in Thailand,
773   2010-2014. Proc Natl Acad Sci U S A. 2018;115(10):E2175-E82.
774   35.   Baquero OS, Santana LMR, Chiaravalloti-Neto F. Dengue forecasting in Sao
775   Paulo city with generalized additive models, artificial neural networks and seasonal
776   autoregressive integrated moving average models. PLoS One. 2018;13(4):e0195065.
777   36.   Sirisena P, Noordeen F, Kurukulasuriya H, Romesh TA, Fernando L. Effect of
778   Climatic Factors and Population Density on the Distribution of Dengue in Sri Lanka:
779   A GIS Based Evaluation for Prediction of Outbreaks. PLoS One.
780   2017;12(1):e0166806.
781   37.   Anggraeni W, Aristiani L, editors. Using Google Trend data in forecasting
782   number of dengue fever cases with ARIMAX method case study: Surabaya,
783   Indonesia. 2016 International Conference on Information & Communication
784   Technology and Systems (ICTS); 2016 12-12 Oct. 2016.
785   38.   Marques-Toledo CA, Degener CM, Vinhal L, Coelho G, Meira W, Codeco
786   CT, et al. Dengue prediction by the web: Tweets are a useful tool for estimating and
787   forecasting Dengue at country and city level. PLoS Negl Trop Dis.
788   2017;11(7):e0005729.
789   39.   Cheong YL, Leitão PJ, Lakes T. Assessment of land use factors associated
790   with dengue cases in Malaysia using Boosted Regression Trees. Spatial and Spatio-
791   temporal Epidemiology. 2014;10:75-84.
792   40.   Wesolowski A, Qureshi T, Boni MF, Sundsoy PR, Johansson MA, Rasheed
793   SB, et al. Impact of human mobility on the emergence of dengue epidemics in
794   Pakistan. Proc Natl Acad Sci U S A. 2015;112(38):11887-92.
795   41.   Zhu G, Liu J, Tan Q, Shi B. Inferring the Spatio-temporal Patterns of Dengue
796   Transmission from Surveillance Data in Guangzhou, China. PLoS Negl Trop Dis.
797   2016;10(4):e0004633.
798   42.   Zhu G, Xiao J, Zhang B, Liu T, Lin H, Li X, et al. The spatiotemporal
799   transmission of dengue and its driving mechanism: A case study on the 2014 dengue
800   outbreak in Guangdong, China. Sci Total Environ. 2018;622-623:252-9.
801   43.   Liu K, Zhu Y, Xia Y, Zhang Y, Huang X, Huang J, et al. Dynamic
802   spatiotemporal analysis of indigenous dengue fever at street-level in Guangzhou city,
803   China. PLOS Neglected Tropical Diseases. 2018;12(3):e0006318.
804   44.   Li Q, Cao W, Ren H, Ji Z, Jiang H. Spatiotemporal responses of dengue fever
805   transmission to the road network in an urban area. Acta Trop. 2018;183:8-13.
806   45.   Chen Y, Ong JHY, Rajarethinam J, Yap G, Ng LC, Cook AR.
807   Neighbourhood level real-time forecasting of dengue cases in tropical urban
808   Singapore. BMC Medicine. 2018;16(1):129.
809   46.   Gardner L, Sarkar S. A global airport-based risk model for the spread of
810   dengue infection via the air transport network. PLoS One. 2013;8(8):e72129.
811   47.   Lauren M. Gardner DF, S. Travis Waller, Ophelia Wang and Sahotra Sarkar.
812   A Predictive Spatial Model to Quantify the Risk of Air-Travel-Associated Dengue
813   Importation into the United States and Europe. Journal of Tropical Medicine.
814   2012;2012.
815   48.   Grubaugh ND, Ladner JT, Kraemer MUG, Dudas G, Tan AL, Gangavarapu K,
816   et al. Genomic epidemiology reveals multiple introductions of Zika virus into the
817   United States. Nature. 2017;546:401.
818   49.   Wilder-Smith A, Gubler DJ. Geographic expansion of dengue: the impact of
819   international travel. Med Clin North Am. 2008;92(6):1377-90, x.

820 50.   Faria NR, Azevedo R, Kraemer MUG, Souza R, Cunha MS, Hill SC, et al.
821 Zika virus in the Americas: Early epidemiological and genetic findings. Science.
822 2016;352(6283):345-9.
823 51.   Gardner LM, Bota A, Gangavarapu K, Kraemer MUG, Grubaugh ND.
824 Inferring the risk factors behind the geographical spread and transmission of Zika in
825 the Americas. PLoS Negl Trop Dis. 2018;12(1):e0006194.
826 52.   Tatem AJ, Hay SI. Climatic similarity and biological exchange in the
827 worldwide airline transportation network. Proceedings of the Royal Society B:
828 Biological Sciences. 2007;274(1617):1489.
829 53.   Siriyasatien P, Phumee A, Ongruk P, Jampachaisri K, Kesorn K. Analysis of
830 significant factors for dengue fever incidence prediction. BMC Bioinformatics.
831 2016;17(1):166.
832 54.   Nishanthi P h m Herath AaiPaHpW. Prediction of Dengue Outbreaks in Sri
833 Lanka using Artificial Neural Networks. International Journal of Computer
834 Applications. 2014;101(15):1-5.
835 55.   Aburas HM, Cetiner BG, Sari M. Dengue confirmed-cases prediction: A
836 neural network model. Expert Systems with Applications. 2010;37(6):4256-60.
837 56.   Baquero OS, Santana LMR, Chiaravalloti-Neto F. Dengue forecasting in São
838 Paulo city with generalized additive models, artificial neural networks and seasonal
839 autoregressive integrated moving average models. PLOS ONE. 2018;13(4):e0195065.
840 57.   Faisal T, Taib MN, Ibrahim F. Neural network diagnostic system for dengue
841 patients risk classification. Journal of medical systems. 2012;36(2):661-76.
842 58.   Laureano-Rosario EA, Duncan PA, Mendez-Lazaro AP, Garcia-Rejon EJ,
843 Gomez-Carro S, Farfan-Ale J, et al. Application of Artificial Neural Networks for
844 Dengue Fever Outbreak Predictions in the Northwest Coast of Yucatan, Mexico and
845 San Juan, Puerto Rico. Tropical Medicine and Infectious Disease. 2018;3(1).
846 59.   Kiskin I OB, Windebank T, Zilli D, Sinka M, Willis K, Roberts S. Mosquito
847 detection with neural networks: the buzz of deep learning. arXiv. 2017.
848 60.   Scavuzzo JM, Trucco FC, Tauro CB, German A, Espinosa M, Abril M,
849 editors. Modeling the temporal pattern of Dengue, Chicungunya and Zika vector
850 using satellite data and neural networks. 2017 XVII Workshop on Information
851 Processing and Control (RPIC); 2017 20-22 Sept. 2017.
852 61.   Sanchez-Ortiz A, Fierro-Radilla A, Arista-Jalife A, Cedillo-Hernandez M,
853 Nakano-Miyatake M, Robles-Camarillo D, et al., editors. Mosquito larva
854 classification method based on convolutional neural networks. 2017 International
855 Conference on Electronics, Communications and Computers (CONIELECOMP);
856 2017 22-24 Feb. 2017.
857 62.   Nguyen T, Khosravi A, Creighton D, Nahavandi S, editors. Epidemiological
858 dynamics modeling by fusion of soft computing techniques. The 2013 International
859 Joint Conference on Neural Networks (IJCNN); 2013 4-9 Aug. 2013.
860 63.   Jiang D, Hao M, Ding F, Fu J, Li M. Mapping the transmission risk of Zika
861 virus using machine learning models. Acta Tropica. 2018;185:391-9.
862 64.   Wahba G. Spline Models for Observational Data: Society for Industrial and
863 Applied Mathematics; 1990. 177 p.
864 65.   PAHO. Countries and territories with autochthonous transmission in the
865 Americas reported in 2015-2017: Washington DC: World Health Organization, Pan
866 American Health Organization; 2017 [Available from:
867 http://www.paho.org/hq/index.php?option=com_content&view=article&id=11603&It
868 emid=41696&lang=en.

869   66.   Gardner L, Chen N, Sarkar S. Vector status of Aedes species determines
870   geographical risk of autochthonous Zika virus establishment. PLoS Negl Trop Dis.
871   2017;11(3):e0005487.
872   67.   Gardner LM, Chen N, Sarkar S. Global risk of Zika virus depends critically on
873   vector status of Aedes albopictus. The Lancet Infectious diseases. 2016;16(5):522-3.
874   68.   Kraemer MU, Sinka ME, Duda KA, Mylne AQ, Shearer FM, Barker CM, et
875   al. The global distribution of the arbovirus vectors Aedes aegypti and Ae. albopictus.
876   Elife. 2015;4:e08347.
877   69.   Theze J, Li T, du Plessis L, Bouquet J, Kraemer MUG, Somasekar S, et al.
878   Genomic Epidemiology Reconstructs the Introduction and Spread of Zika Virus in
879   Central America and Mexico. Cell host & microbe. 2018;23(6):855-64.e7.
880   70.   WorldBank. International Comparison Program database. GDP per capita,
881   PPP 2016 [Available from:
882   https://data.worldbank.org/indicator/NY.GDP.PCAP.PP.CD.
883   71.   Analysis USBoE. Widespread Economic Growth Across States In 2011 2011 [
884   72.   Services USDoHaH. Health, United States, 2015 2015 [Available from:
885   https://www.cdc.gov/nchs/data/hus/hus15.pdf.
886   73.   (WHO) WHO. WHO World Health Statistics 2015 2015 [Available from:
887   http://www.who.int/gho/publications/world_health_statistics/2015/en/.
888   74.   PAHO. PLISA Health Indication Platform for the Americas: Washington DC:
889   World Health Organization, Pan American Health Organization; 2017 [Available
890   from: http://www.paho.org/data/index.php/en/.
891   75.   Bank W. International Comparison Program database. Population density
892   (people per sq. km of land area) 2016 [Available from:
893   http://data.worldbank.org/indicator/EN.POP.DNST.
894   76.   International Air Travel Association (IATA)- Passenger Intelligence Services
895   (PaxIS) [Available from:
896   http://www.iata.org/services/statistics/intelligence/paxis/Pages/index.aspx.
897   77.   Pigott D, Deshpande A, Letourneau I, Morozoff C, Reiner Jr R, Kraemer M, et
898   al. Local, national, and regional viral haemorrhagic fever pandemic potential in
899   Africa: a multistage analysis. Lancet. 2017;390(10113):2662-72.
900   78.   Leontaritis IJ, Billings SA. Input-output parametric models for non-linear
901   systems Part I: deterministic non-linear systems. International Journal of Control.
902   1985;41(2):303-28.
903   79.   Narendra KS, Parthasarathy K. Identification and control of dynamical
904   systems using neural networks. IEEE Transactions on Neural Networks. 1990;1(1):4-
905   27.
906   80.   Chen S, Billings SA, Grant PM. Non-linear system identification using neural
907   networks. International Journal of Control. 1990;51(6):1191-214.
908   81.   Siegelmann HT, Horne BG, Giles CL. Computational capabilities of recurrent
909   NARX neural networks. IEEE Trans Syst Man Cybern B Cybern. 1997;27(2):208-15.
910   82.   Tsungnan L, Bill GH, Peter T, Giles CL. Learning long-term dependencies is
911   not as difficult with NARX recurrent neural networks. University of Maryland at
912   College Park; 1995. p. 23.
913   83.   MATLAB and Neural Network Toolbox Release 2018a Natick,
914   Massachusetts, United States: MathWorks; [Available from:
915   https://au.mathworks.com/help/pdf_doc/nnet/nnet_ug.pdf.
916   84.   Fawcett T. ROC graphs: Notes and practical considerations for researchers.
917   Machine Learning. 2004;31:1-38.

918   85.   Faria NR, Quick J, Claro IM, Thézé J, de Jesus JG, Giovanetti M, et al.
919   Establishment and cryptic transmission of Zika virus in Brazil and the Americas.
920   Nature. 2017;546:406.
921   86.   Gardner L, Chen N, Sarkar S. Global risk of Zika virus depends critically on
922   vector status of Aedes albopictus. Lancet Infectious Diseases. 2016;16(5):522-3.
923   87.   Brockmann D, Helbing D. The Hidden Geometry of Complex, Network-
924   Driven Contagion Phenomena. Science. 2013;342:1337-42.
925   88.   Boussaada Z, Curea O, Remaci A, Camblong H, Mrabet Bellaaj N. A
926   Nonlinear Autoregressive Exogenous (NARX) Neural Network Model for the
927   Prediction of the Daily Direct Solar Radiation. Energies. 2018;11(3).
928
929

930
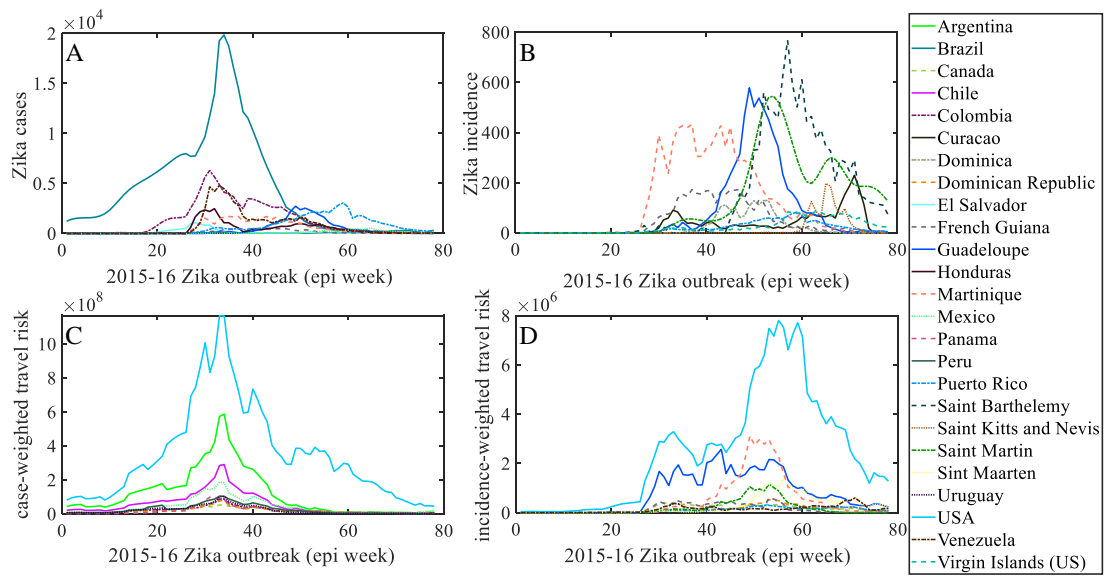
931
932
933 **Fig 1. Weekly distribution of case and connectivity-risk variables**. (A) Zika cases
934 (B) incidence rates in the Americas, (C) case-weighted travel risk CR_j^t, and (D)
935 incidence weighted travel risk IR_j^t, for top 10 ranked countries and territories in the
936 Americas for each respective variable.

937
938
939
940
941
942

943
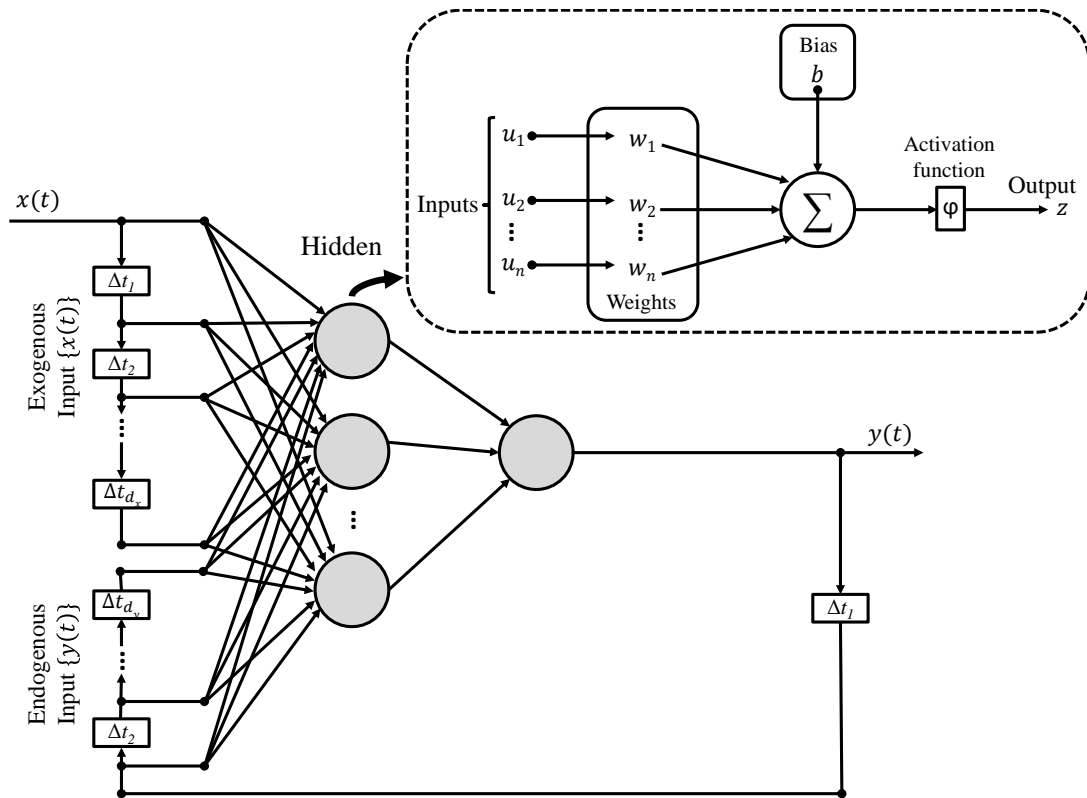
Fig 2. Schematic of NARX network with $\mathbf{d_x}$ input and $\mathbf{d_y}$ output delays: Each neuron produces a single output based on several real-valued inputs to that neuron by forming a linear combination using its input weights and sometimes passing the output through a nonlinear activation function: $\mathbf{z} = \boldsymbol{\varphi}(\sum_{i=1}^{n} \mathbf{w_i u_i} + \mathbf{b}) = \boldsymbol{\varphi}(\mathbf{w^T x} + \mathbf{b})$, where $\mathbf{w}$ denotes the vector of weights, $\mathbf{u}$ is the vector of inputs, $\mathbf{b}$ is the bias and $\boldsymbol{\varphi}$ is a linear or nonlinear activation function (e.g., Linear, Sigmoid, and Hyperbolic tangent (88)).

951

36
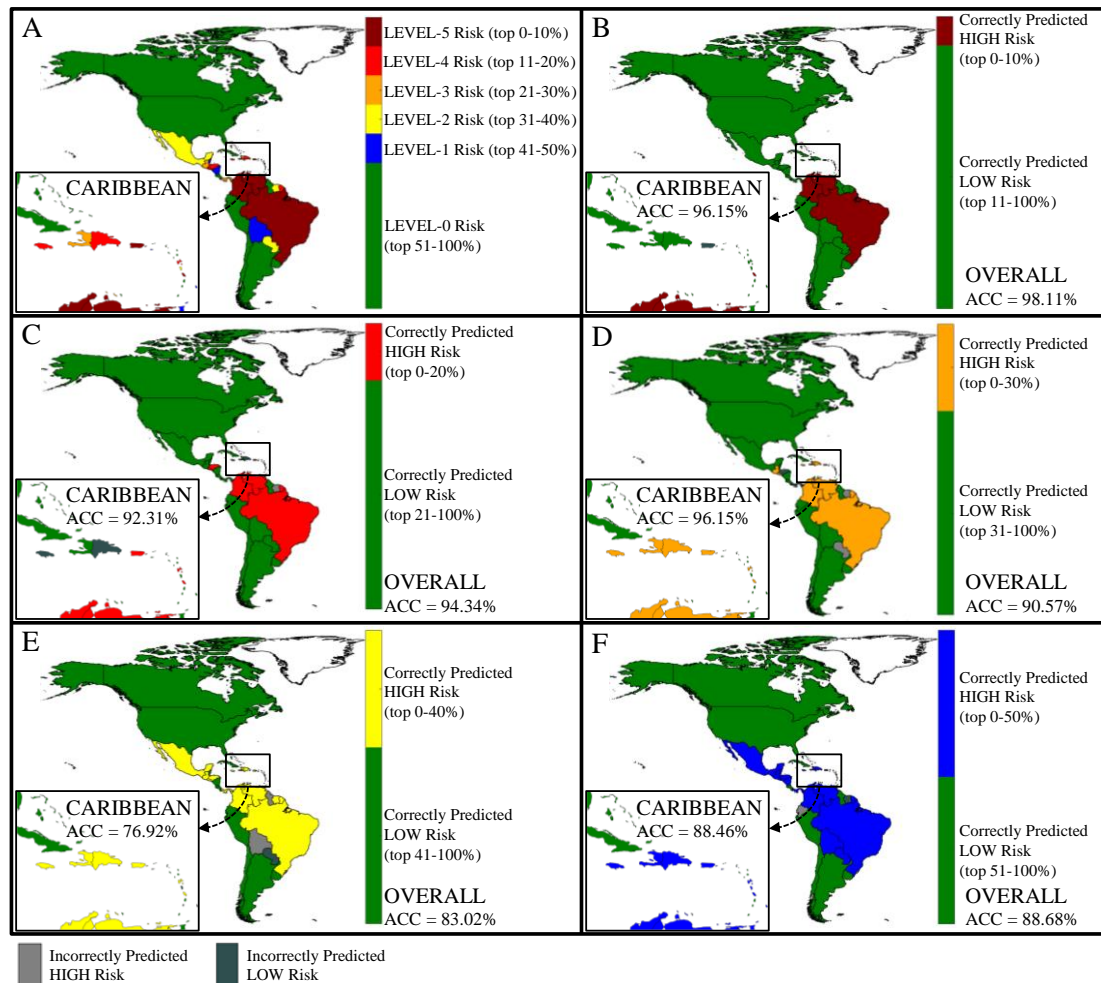
Fig 3. Country prediction accuracy by relative risk level. Panel (A) illustrates the actual relative risk level assigned to each country at Epi week 40 for a fixed forecast window, N=4. Panels (B)-(E) each corresponds to a different classification scheme, specifically (B) R=0.1, (C) R=0.2, (D) R=0.3, (E) R=0.4, and (F) R=0.5. The inset shown by the small rectangle highlights the actual and predicted risk in Caribbean islands. For Panels (B)-(E), green indicates a correctly predicted low risk country, light grey indicates an incorrectly predicted high risk country, and dark grey indicates an incorrectly predicted low risk country. The risk indicator used is case counts.
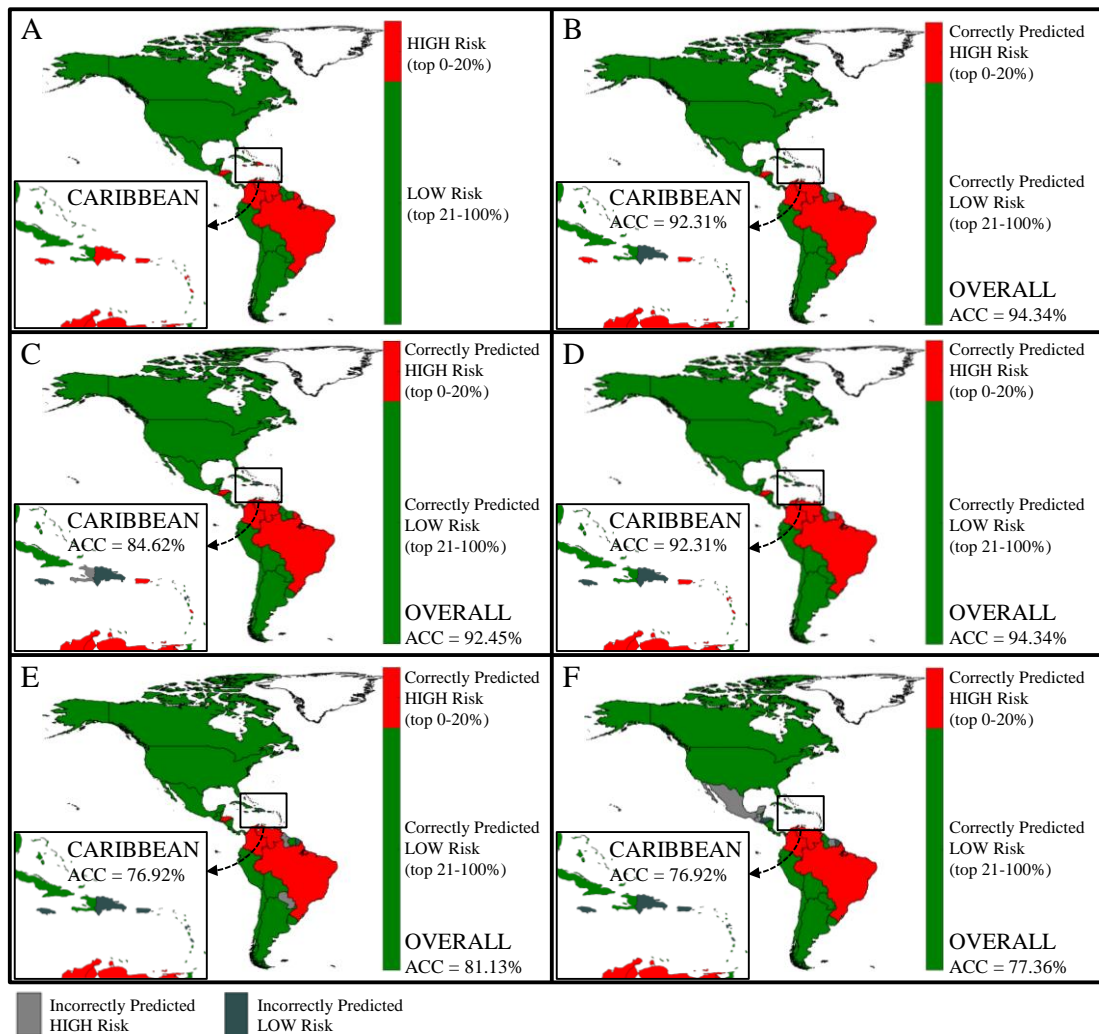
Fig 4. Country prediction accuracy by forecast window. Panel (A) illustrates the actual relative risk level assigned to each country at Epi week 40 for a fixed classification scheme, R=0.2. Panels (B)-(E) each corresponds to different forecast windows, specifically (B) N=1, (C) N=2, (D) N=4, (E) N=8, and (F) N=12. The inset shown by the small rectangle highlights the actual and predicted risk in Caribbean islands. For Panels (B)-(E), the red indicates a correctly predicted high risk country and green indicates a correctly predicted low risk country. Light grey indicates an incorrectly predicted high risk country, and dark grey indicates an incorrectly predicted low risk country. The risk indicator used is case counts.
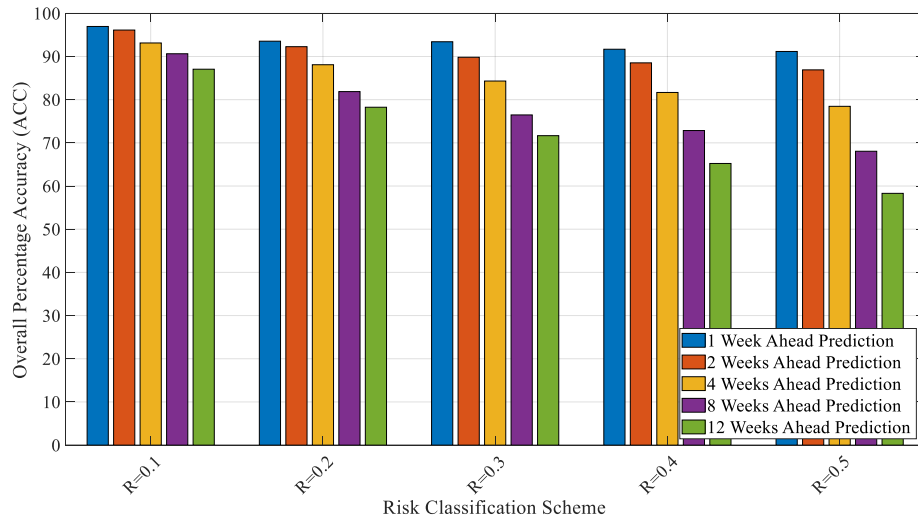
Fig 5. Aggregate model performance measured by ACC (averaged over all locations and all weeks) for all combinations of relative risk classification schemes (i.e., R = 0.1, 0.2, 0.3, 0.4 and 0.5) and forecast windows (i.e., N = 1, 2, 4, 8 and 12), where the risk indicator is case counts.
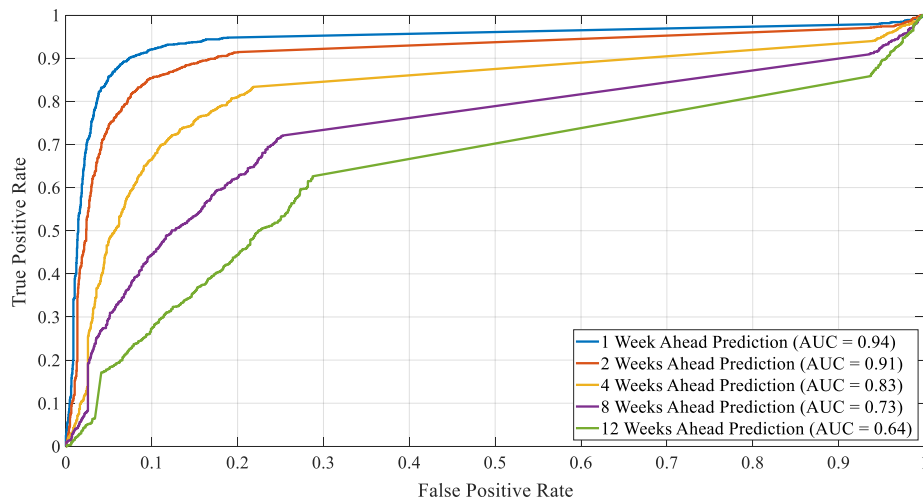
Fig 6. Aggregate model performance measured by ROC AUC (averaged over all locations and all weeks) for a fixed relative risk classification scheme, *i.e.*, R = 0.4, and forecast windows (i.e., N = 1, 2, 4, 8 and 12), where the risk indicator is case counts.

(a)

(b)
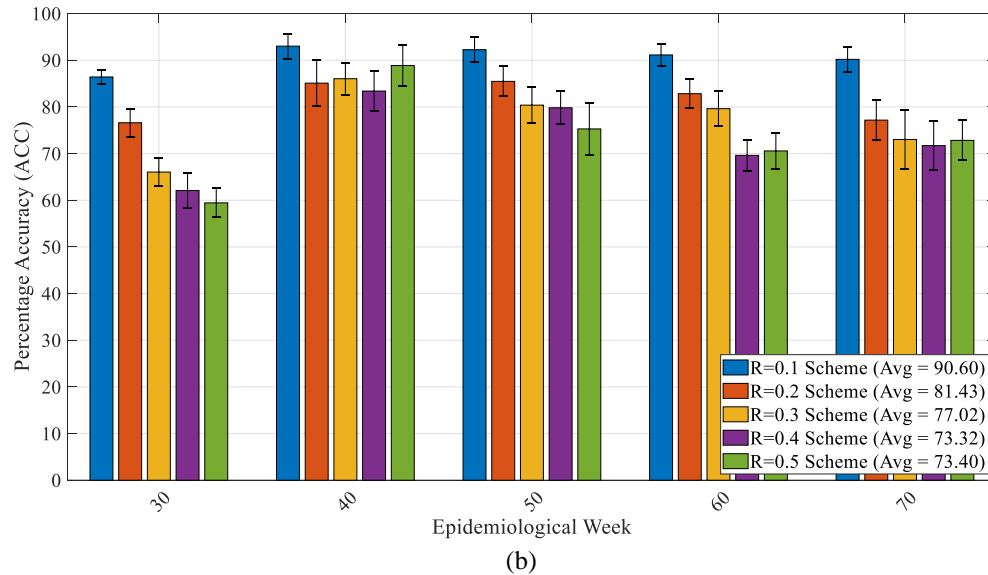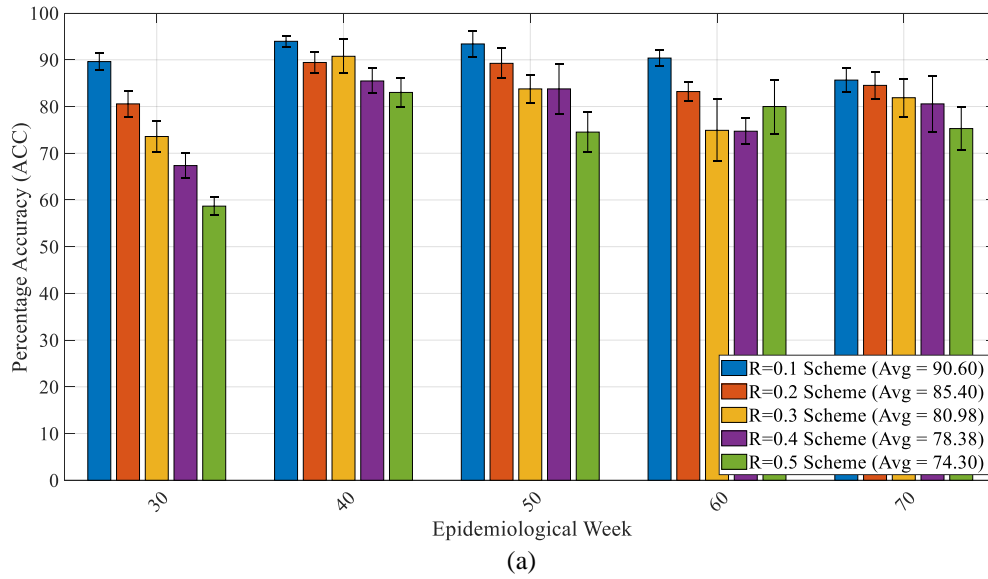
987
988 **Fig 7. Model performance and robustness.** ACC is averaged over all locations for
989 selected epidemiological weeks when risk indicator is (a) case counts and (b)
990 incidence rate, and a fixed forecast windows (i.e., N = 4). The error bars represent the
991 variability in expected ACC across ten runs for each combination.