Accepted Manuscript

Model-Informed Risk Assessment for Zika Virus Outbreaks in the Asia-Pacific Regions

Yue Teng, Dehua Bi, Guigang Xie, Yuan Jin, Yong Huang, Baihan Lin, Xiaoping An, Yigang Tong, Dan Feng

PII: S0163-4453(17)30049-X
DOI: 10.1016/j.jinf.2017.01.015
Reference: YJINF 3883

To appear in: Journal of Infection

Received Date: 2 November 2016
Revised Date: 27 December 2016
Accepted Date: 9 January 2017


This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.
Model-Informed Risk Assessment for Zika Virus Outbreaks in the Asia-Pacific Regions

Yue Teng1, 2, *, †, Dehua Bi2, 3, *, Guigang Xie2, Yuan Jin2, 4, Yong Huang1, 2, Baihan Lin5, Xiaoping An1, 2, Yigang Tong1, 2, †, Dan Feng6, †

1 Beijing Institute of Microbiology and Epidemiology, Beijing 100071, China;
2 State Key Laboratory of Pathogen and Biosecurity, Beijing 100071, China;
3 Department of Mechanical and Mechatronics Engineering, University of Waterloo, 200 University Avenue West, Waterloo, Ontario N2L 3G1, Canada;
4 Beijing Institute of Biotechnology, Beijing 100071, China;
5 Computational Neuroscience Program, Department of Psychology, Physics, and Computer Science and Engineering; Institute for Protein Design, University of Washington, Seattle, WA 98195, USA;
6 Division of Standard Operational Management, Institute of Hospital Management, Chinese PLA General Hospital, Beijing 100853, China.

* Y.T. and D.H.B. contributed equally to this work.
† Correspondence:
Yue Teng Ph.D., M.D., State Key Laboratory of Pathogen and Biosecurity, Beijing Institute of Microbiology and Epidemiology, 20 Dong-Da Street, Fengtai District, Beijing 100071, China.
Telephone/ Fax: +86-10-68164807; E-mail: yueteng@sklpb.org;
Yigang Tong Ph.D., State Key Laboratory of Pathogen and Biosecurity, Beijing Institute of Microbiology and Epidemiology, 20 Dong-Da Street, Fengtai District, Beijing 100071, China. Telephone/ Fax:  
Abstract

Recently, Zika virus (ZIKV) has been recognized as a significant threat to global public health. The disease was present in large parts of the Americas, the Caribbean, and also the western Pacific area with southern Asia during 2015 and 2016. However, little is known about the factors affecting the transmission of ZIKV. We used Gradient Boosted Regression Tree models to investigate the effects of various potential explanatory variables on the spread of ZIKV, and used current with historical information from a range of sources to assess the risks of future ZIKV outbreaks. Our results indicated that the probability of ZIKV outbreaks increases with vapor pressure, the occurrence of Dengue virus, and population density but decreases as health expenditure, GDP, and numbers of travelers. The predictive results revealed the potential risk countries of ZIKV infection in the Asia-Pacific regions between October 2016 and January 2017. We believe that the high-risk conditions would continue in South Asia and Australia over this period. By integrating information on eco-environmental, social-economical, and ZIKV-related niche factors, this study estimated the probability for locally acquired
mosquito-borne ZIKV infections in the Asia-Pacific region and improves the ability to forecast, and possibly even prevent, future outbreaks of ZIKV.

**Keywords:** Zika virus; endemic; risk factors; modeling; forecasting; surveillance

**Introduction**

Zika virus (ZIKV) was first identified in a monkey in Africa in 1947, and was first reported in Asia in 1969[1, 2]. The first documented outbreak of ZIKV in humans occurred on Yap State (Federated States of Micronesia) in 2007 [3]. The most recent outbreak of ZIKV began in early 2015 in South America, and then spread to other parts of South and North America and the Caribbean [4]. It also reached several Pacific islands and Singapore by mid-2016 [5]. The World Health Organization declared it a Public Health Emergency of International Concern in February 2016[6]. Confirmed cases were reported in 40 countries and territories in America and the Caribbean, as well as 16 in the western Pacific and one in Africa, in 2015 and 2016[7].

ZIKV belongs to the virus family *Flaviviridae* and the genus *Flavivirus*, which also contains West Nile virus, Dengue virus, Yellow fever virus, and Japanese encephalitis virus [8]. Zika virus disease (ZVD) also causes an illness that resembles a mild form of dengue fever [9]. Currently, there is no medication or vaccine that can
prevent the ZIKV infection. ZIKV is mainly spread by the *Aedes aegypti* and *Aedes albopictus* mosquitoes, which are common throughout tropical and subtropical regions of the world [10]. Previous publications have highlighted a complex set of human and environmental factors that determine the spatial distribution of this disease [11]. Additionally, researchers found that travel from ZIKV-affected areas and socioeconomic conditions in countries facing potential abundant mosquito populations may play important roles in the ZIKV transmission [12]. Finally, researchers believe that high risk factors for transmission exist along the borders of affected countries [13].

However, the various risk factors that affected the progress of the ZIKV outbreaks remain unidentified [14, 15]. We, therefore, collected data from 50 countries with evidence of local mosquito-borne ZIKV infections, and 80 countries without ZIKV infection reports, to investigate potential explanatory variables. We considered ecological, environmental, meteorological and social-economical niche factors. These included precipitation, vapor pressure, population density, the number of travelers, temperature, health expenditure *per capita*, Gross Domestic Product *per capita*, water coverage and ZIKV transmission in nearby countries. Because the ZIKV-related mosquitoes often carried other viruses, we also extracted occurrence data for these mosquito species and Dengue virus (DENV), West Nile virus (WNV), Yellow Fever virus (YFV) and Chikungunya virus (CHIKV). We used a
Gradient Boosted Regression Tree (GBRT) model to evaluate the effect of these factors on the probability of ZIKV transmission and to estimate risk levels for ZVD outbreaks, at the country level, in Asia and two countries in Oceania between October 2016 and January 2017. Our approach estimates the relative influence of these potential risk factors and their importance for surveillance systems that attempt to provide early warning of outbreaks and inform decision-making in countries within the Asia-Pacific region.

Materials and Methods

Data collection

The countries with locally acquired mosquito-borne ZIKV infections, and the number of confirmed cases during the 2015–2016 outbreaks, were obtained from the World Health Organization (WHO)’s latest Zika situation report (http://www.who.int/emergencies/zika-virus/situation-report/8-september-2016/en/) and the Pan America Health Organization (http://www.paho.org/hq/index.php?option=com_content&view=article&id=11599&Itemid=41691&lang=en). The distribution of confirmed cases was linked to a global geographic map to plot the thematic figure using GIS technologies. The map was created in ArcGIS 9.3 software (ESRI Inc., Redlands, CA, USA) (http://www.esri.com/). We collected data from 50 countries with evidence of local
mosquito-borne ZIKV infections and 80 countries without ZIKV infection reports. The data concerning ecological, environmental, meteorological and social-economical niche factors, in addition to various factors concerning phylogenetic relatedness to ZIKV, were included in the mathematical model used in this study.

Among these data, the ecological niche data of monthly temperature, vapor pressure and precipitation were from the dataset of CRU-3.23 (http://iridl.ldeo.columbia.edu/SOURCES/.UEA/.CRU/.TS3p23/) and the water coverage and population density were extracted from the World Factbook (https://www.cia.gov/library/publications/the-world-factbook/). We further examined the borders of the countries in this database to identify neighboring countries to those affected by ZIKV (a value of 1 indicated a country that neighbored a ZIKV affected country, and 0 meant it had no such neighbor). GDP, health expenditure per capita and the annual number of travelers for each country were downloaded from the World Bank dataset (http://data.worldbank.org/). The occurrence data for mosquitos, Dengue virus, Chikungunya virus, Yellow Fever virus and West Nile virus was obtained from a previous publication [16, 17], WHO’s report (http://www.who.int/csr/don/archive/disease/yellow_fever/en/) and the CDC (http://www.cdc.gov/westnile/statsmaps/). The detailed list of the factors with their descriptions was given in Table 1 and Table S2. The data was split into a
learning set (40 ZIKV affected countries and 70 none ZIKV countries) and a prediction set (10 ZIKV affected countries and 10 none ZIKV countries) for the prediction model.

**Model summary**

A Gradient Boosting Regression Tree (GBRT) model was built on the learning dataset at the country level in this study to examine risk factors for the endemic spreading of ZIKV, and to predict the potential risk regions. The 40 countries with endemic transmission and the other 70 countries, without ZIKV in the learning set, were considered as the positive and negative samples, respectively. The GBRT model is efficient for predicting distributions of organisms while accounting for non-linear relationships and interactions between covariates [18, 19]. For the GBRT model, a bootstrapping procedure was utilized to provide a robust estimation of model parameters. A tree complexity of 2, learning rate of 0.06, estimation iteration of 340 and a bag fraction of 75% were used to identify the optimal tree for each bootstrap data. The weight of each variable was estimated from the identified trees and used as an indicator to show the importance of each variable for predicting the probability of ZIKV spreading in the country. Note that these weights are not absolute metrics and the weights of all variables sum to 1.

The following sequential steps were repeated 600,000 times in the bootstrapping procedure. Data from 100 countries were randomly selected with
replacement from the learning set, which consist of 110 countries. These data were divided randomly into 75% training and 25% testing sets. Then, a GBRT model was built on the training data, and validated with both the testing and the prediction data using the receiver-operating characteristic (ROC) curves and the areas under the curve (AUC). The mean and standard deviation, over the 50 iterations, of each parameter was also calculated. Risk function values greater than 0 mean that the factor has a positive influence on the spreading of the ZIKV, while values less than 0 correspond to negative influences.

**Results**

The distribution of confirmed cases of the Zika virus, in countries with possible endemic transmission or evidence of local mosquito-borne Zika infections during the 2015-2016 outbreaks, was shown in Figure 1A [7]. Regions of local transmission of Zika virus have been identified in 40 countries or territories in South America, North America, and the Caribbean (Figure 1B), as well as 16 in the western Pacific and one in Africa (Figure 1C), since the beginning of 2015. The ZIKV epidemic in Brazil, Puerto Rico, and Colombia was most severe in these regions, with 78421, 17935 and 8826 confirmed cases, respectively (Table S1). It was worth noting that these areas were in tropical and sub-tropical zones with climates suitable for the relevant species of mosquitoes.
However, there was still uncertainty about which other areas could support ZIKV. We collected data on variables that might be associated with the rapid spread of ZIKV at the country level, such as precipitation and temperature fluctuation, by calculating monthly averaged temperatures, vapor pressures and precipitation (Table 1). Other socio-economic factors were also considered (water coverage percentage, population density, number of travelers, average GDP, government health expenditure, and the occurrence of mosquitoes). Variables describing the prevalence of four closely related viruses (DENV, YFV, WNV and CHIKV) were also considered as indicators of the potential for ZIKV outbreaks. A Gradient Boosted Regression Tree (GBRT) was constructed with these factors and a learning data set containing 40 ZIKV affected countries as positive samples and 70 non-ZIKV countries as negative samples (Table S2) to examine the contribution of each variable to the occurrence of ZIKV infection. The importance of each factor was evaluated by estimating their weights in the constructed tree-based model, with the sum of all variable weights set equal to 1.

The results of this model indicated that temperature, vapor pressure, precipitation, water coverage, number of travelers, population density, GDP, health expenditures and the occurrence of Dengue virus were significantly associated with the occurrence of ZIKV infection (all GBRT mean weights > 5.0%, Table 2). Remarkably, vapor pressure had the highest mean weight (17.2%) of all these
variables, in the forecasting of riskiness. Population density, temperature, health expenditure, and the occurrence of Dengue virus had mean weights >10% for ZIKV infection (13.7%, 13.9%, 10.7% and 10.1%, respectively). GDP, precipitation, water coverage and the number of travelers also had some impact on ZIKV outbreaks yet (mean weights are 9.4%, 8.2%, 10.0% and 5.6%, respectively). The estimated effects of these variables are shown in Figure 2. This shows that the probability of ZIKV transmission increased with vapor pressure, the occurrence of Dengue virus and population density (Figure 2A, 2B and 2E), but was negatively correlated with health expenditure, GDP and the number of travelers (Figure 2D, 2F and 2I). The probability of ZIKV outbreaks appears to increase dramatically when the vapor pressure exceeds 18hPa (Figure 2A). The results also suggest that there is a positive correlation between the presence of Dengue virus and the appearance of ZIKV, which implied the high risk for ZIKV occurring in Dengue virus epidemic regions (DENV = 1.0, Figure 2E). The highest risk of ZIKV infections appears to occur between 22 to 28 Degrees Celsius (Figure 2C). These results agree with the earlier publication [10], that water-filled environments with stable warm temperatures, which offer perfect breeding grounds for mosquitoes and suitable conditions for the mosquito's entire lifestyle, also promote the spread of ZIKV.

To improve the stability of the GBRT model, and the reproducibility of its results, we performed a bootstrapping procedure. This provided stable and robust
predictions for forecasting future outcomes. The receiver-operating characteristic (ROC) curve was produced for the training and testing data from the learning set, as well as the estimated result on the predictive set. The area under the curve (AUC) was also calculated to test the discriminatory ability of the GBRT model (Figure 3). An AUC of 0.92 (95% CI 0.90 – 0.94) was achieved on the predictive set. The AUC value for the training set and the testing set were 1.0 (95% CI 1.0 – 1.0) and 0.99 (95% CI 0.98 – 1.0), respectively. The results suggested that the GBRT model containing these variables not only has a good predictive power for the learning dataset the model was built on, but also provides robust predictions which can be generalized to predict future risks of ZIKV infection.

Furthermore, we ran the GBRT model with the bootstrapping procedure to estimate the probability of transmission for ZIKV infections, between October 2016 and January 2017, in Asian countries and also two countries in Oceania (Australia and New Zealand; Table S3). Figure 4 contains maps illustrating the dynamic predictive situation and the risks of ZIKV endemic in each month. The results suggest that the high-risk countries were clustered at South Asia in October 2016 (India, Sri Lanka, Vietnam, Laos, Thailand, Philippine, Singapore, Indonesia, Malaysia; all risk probabilities > 80%, Figure 4A), and this risk status would continue until November 2016. In contrast, Australia would reach a relatively high riskiness for ZIKV infection in November 2016 (Risk probability: 84%) and this would enhance in
December 2016 and January 2017 (Risk probabilities: 89% and 89%, respectively. Figure 4B-D). As expected, New Zealand would enter warning status (Risk probability: 39%) for ZIKV infections in January 2017 with the high-risk situation in Australia (Figure 4D). We also found that countries in West and Central Asia would have middle-risk probabilities for ZIKV outbreaks in these 4 months (United Arab Emirates, Afghanistan, Turkmenistan and Uzbekistan; all Risk probabilities < 60%, Figure 4A-D). The East Asian countries, including China, Japan and South Korea, were predicted to have low probabilities of experiencing ZIKV epidemics (Risk probabilities: < 20%) from October 2016 to January 2017 (Figure 4A-D). These geographical patterns of predicted riskiness were consistent with the above results of the analysis of the various risk contributors for endemic ZIKV.

Discussion

The current ZIKV prevalence in the Americas poses a severe worldwide threat to public health, and more countries than previously assumed could soon be grappling with ZIKV. The two species of mosquitos appear to be equally effective carriers of the disease [8, 9]. Figure 1 displayed the spatial distribution of local ZIKV transmission, with the number of confirmed cases in worldwide during 2015 and 2016. To measure the potential risk of circulating ZIKV epidemics, we combined data on various factors associated with the rapid spread of ZIKV at the country level
and used a Gradient Boosted Regression Tree (GBRT) model to estimate the importance of each of these contributors to the probability of ZIKV transmission in countries within the Asia-Pacific region between October 2016 and January 2017. Our results showed that temperature, vapor pressure, precipitation, water coverage, the number of travelers, population density, GDP, health expenditures, and the occurrence of Dengue virus were all significantly associated with the occurrence of ZIKV infection (all GBRT mean weights > 5.0%, Table 2). Among these factors, the increasing of vapor pressure, occurrence of Dengue virus and population density appear to improve the probability of ZIKV transmission (Figure 2A, 2B and 2E); however, the high level of health expenditure, GDP and numbers of travelers (Figure 2D, 2F and 2I) may help to reduce the risks. According with above finding, the results revealed that water-filled environments (vapor pressure > 18hPa) with stable warm temperatures (22 to 28 Degrees Celsius) would provide ideal conditions for mosquitoes to transmit ZIKV. Furthermore, we bootstrapped the GBRT model to obtain robust predictions of the probability of ZIKV infections occurring between October 2016 and January 2017 in Asian countries and two countries in Oceania (Australia and New Zealand). We noticed that the high-risk status in South Asia and Australia would continue (Figure 4A-D), because the conditions in these regions would be steadily suitable for the mosquitoes. Generally, once a mosquito has acquired the ZIKV from an infected person, it needs to live long enough (always
more than one week, though this depends on environmental temperatures), for the ZIKV to move from the mosquito’s mid-gut to its salivary glands. The mosquito can then transmit the ZIKV in the saliva to another person [8]. Conditions in the southern countries of Asia would be most suitable for the mosquitoes in the next 4 months in these predictive areas, though the peak times will vary by country. In the Oceania countries, conditions seem likely to remain suitable as late as spring 2017. As springtime weather warms in southern hemisphere, the potential abundance of the mosquito thus begins to increase in Australia and New Zealand. However, fall conditions during October and November 2016 can only support low and moderate populations of the mosquito in the countries of Western and Central Asia. This produced a middle riskiness of ZIKV infection in our results (Figure 4A-D). And, the winter is too cold for the species of the mosquito to survive in northern countries of Asia, such as China, Japan and South Korea. The Asia-Pacific region is a tropical region where many tropical diseases are endemic, and has a very high prevalence of dengue and chikungunya fever [20]. The main reason for this is the abundance of mosquito vectors in this area [21]. Countries in Asia-Pacific region may therefore be the next hot spot of ZIKV outbreaks. Our estimates of the relative importance of potential risk factors can be used to overcome known ZIKV preparedness challenges [22, 23] by strengthening the design of early warning and surveillance systems in countries in the Asia-Pacific region.
ACKNOWLEDGEMENTS

We thank Dr. Wuchun Cao in State Key Laboratory of Pathogen and Biosecurity for discussions and our colleagues in the Beijing Institute of Microbiology and Epidemiology for help in technical assistance.

Study funded by grants from the State Key Laboratory of Pathogen and Biosecurity Program (No. SKLPBS1408 and No. SKLPBS1451).

Conflict of interest

The authors declare that they have no conflict of interest. The funding source was not involved in the current analyses or in the preparation of this report. The corresponding author had full access to all data and had the final responsibility for the decision to submit for publication.

Author Contributions

The manuscript was written by Y.T. and D.H.B.; data analyses were performed by D.H.B., G.G.X., B.H.L. and Y.T.; the study was designed by Y.T., D.H.B., Y.J., and D.F.

References

Epub 2016/03/22. eng.


**Figure Legends**

**Figure 1.** Global spatial distribution of locally acquired ZIKV infections during 2015 and 2016 outbreaks. The darker red color indicated a greater number of infected cases.

**Figure 2.** Association between risk variables and the probability of ZIKV infection. The ZIKV risk based on the GBRT model is plotted as a function of (A) vapor
pressure (VP) (B) population density (PD) (C) temperature (TP) (D) health expenditure per capita (HE) (E) Dengue virus occurrence (DENV) (F) Gross Domestic Product per capita (GDP) (G) precipitation (PP) (H) water coverage percentage (WC) (I) the number of travelers (NT).

Figure 3. Receiver operating characteristic (ROC) curves of the estimated riskiness. The solid black, blue, red lines and the dashed grey line indicated the average ROC curves of 50 repeats based on the bootstrapping procedure for the training set, test set, prediction set, and the random guess line.

Figure 4. Dynamic forecasting maps of the probability for ZIKV infection in the Asia-Pacific region. The risk of ZIKV spreading is displayed by different color grades, according to the GBRT model. Red colors indicate a high risk, yellow colors indicate an intermediate risk and white indicates no risk. (A) in October 2016, (B) in November 2016, (C) in December 2016 and (D) in January 2017.

Table 1. Description of potential risk variable used in the GBRT analyses.

Table 2. Results of the gradient boosted regression trees (GBRT) model.

Table S1. Cumulative number of confirmed cases for ZIKV endemic transmission countries.

Table S2. Data used the GBRT Model for training, testing and validating.

Table S3. Estimates of the potential risk in the Asia-Pacific Regions calculated from the GBRT Model for each month between October 2016 and January 2017.
<table>
<thead>
<tr>
<th>Abbr.</th>
<th>Variables</th>
<th>Description (Unit)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PP</td>
<td>Precipitation</td>
<td>Monthly precipitation (mm)</td>
</tr>
<tr>
<td>VP</td>
<td>Vapor Pressure</td>
<td>Monthly vapor pressure (hPa)</td>
</tr>
<tr>
<td>PD</td>
<td>Population Density</td>
<td>Human population density for each county (1,000 persons per km²)</td>
</tr>
<tr>
<td>NT</td>
<td>Number of Travelers</td>
<td>Yearly average travelers by airplane to the country (person)</td>
</tr>
<tr>
<td>TP</td>
<td>Temperature</td>
<td>Monthly temperature (degree Celsius)</td>
</tr>
<tr>
<td>HE</td>
<td>Health Expenditure per capita</td>
<td>Health Expenditure per capita ($US)</td>
</tr>
<tr>
<td>GDP</td>
<td>Gross Domestic Product per capita</td>
<td>GDP per capita ($US)</td>
</tr>
<tr>
<td>WC</td>
<td>Water Coverage</td>
<td>Percentage coverage of water body for each county (1%)</td>
</tr>
<tr>
<td>ZTNC</td>
<td>ZIKV Transmission in Nearby Countries</td>
<td>The country neighbored the countries in ZIKV endemic</td>
</tr>
<tr>
<td>MO</td>
<td>Mosquito Occurrence</td>
<td>Detailed data on the present distribution of mosquito vectors</td>
</tr>
<tr>
<td>CIZC</td>
<td>Confirmed Imported ZIKV Case</td>
<td>Imported confirmed ZIKV cases in the country</td>
</tr>
<tr>
<td>DENV</td>
<td>Dengue Occurrence</td>
<td>Detailed data on the present distribution of Dengue virus</td>
</tr>
<tr>
<td>WNV</td>
<td>West Nile Occurrence</td>
<td>Detailed data on the present distribution of West Nile virus</td>
</tr>
<tr>
<td>YFV</td>
<td>Yellow Fever Occurrence</td>
<td>Detailed data on the present distribution of Yellow Fever virus</td>
</tr>
<tr>
<td>CHIKV</td>
<td>Chikungunya Occurrence</td>
<td>Detailed data of the present distribution of Chikungunya virus</td>
</tr>
</tbody>
</table>
Table 2 Results of relative contribution for risk variables in the GBRT analysis during the current ZIKV endemic

<table>
<thead>
<tr>
<th>Risk factors†</th>
<th>Relative contribution</th>
<th>ZIKV infection</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (%)</td>
<td>SD</td>
</tr>
<tr>
<td>Vapor Pressure</td>
<td>17.19</td>
<td>8.3e-3</td>
</tr>
<tr>
<td>Population Density</td>
<td>13.73</td>
<td>1.4e-2</td>
</tr>
<tr>
<td>Temperature</td>
<td>13.86</td>
<td>1.0e-2</td>
</tr>
<tr>
<td>Dengue Occurrence</td>
<td>10.72</td>
<td>1.1e-2</td>
</tr>
<tr>
<td>Health Expenditure per capita</td>
<td>10.10</td>
<td>1.0e-2</td>
</tr>
<tr>
<td>Gross Domestic Product per capita</td>
<td>9.37</td>
<td>1.1e-2</td>
</tr>
<tr>
<td>Precipitation</td>
<td>8.21</td>
<td>8.8e-3</td>
</tr>
<tr>
<td>Water Coverage</td>
<td>6.98</td>
<td>1.2e-2</td>
</tr>
<tr>
<td>Number of Travelers</td>
<td>5.61</td>
<td>1.2e-2</td>
</tr>
<tr>
<td>ZIKV Transmission in Nearby Countries</td>
<td>2.00</td>
<td>1.6e-3</td>
</tr>
<tr>
<td>Yellow Fever Occurrence</td>
<td>NS</td>
<td>-</td>
</tr>
<tr>
<td>Chikungunya Occurrence</td>
<td>NS</td>
<td>-</td>
</tr>
<tr>
<td>West Nile Occurrence</td>
<td>NS</td>
<td>-</td>
</tr>
<tr>
<td>Mosquito Occurrence</td>
<td>NS</td>
<td>-</td>
</tr>
<tr>
<td>Confirmed Imported ZIKV Case</td>
<td>NS</td>
<td>-</td>
</tr>
</tbody>
</table>

†Variables with mean weights ≥ 5% were considered as significant contributors to the occurrence of human infections.

“NS”: these variables were excluded from the final model due to small BRT weights (< 2.0%)
A plot of the Receiver Operating Characteristic (ROC) curve showing the performance of a binary classifier. The x-axis represents 1-Specificity, and the y-axis represents Specificity. The plot includes three curves:

- **Train (AUC = 1.00)**: The solid black line at the top of the plot, indicating perfect classification.
- **Test (AUC = 0.99)**: The blue line, slightly below the Train line, indicating high classification performance.
- **Predict (AUC = 0.92)**: The red line, showing lower performance compared to the Test curve.

The dashed line represents the no-signal line, indicating no predictive value.