ZIKA Virus: Prediction of Aedes Mosquito Larvae Occurrence in Recife (Brazil) using Online Extreme Learning Machine and Neural Networks

Adrian Rubio-Solis*
adrian.rubio@cidesi.edu.mx
Center for Engineering and Industrial Development- CIDESI Mexico

Anwar Musah
a.musah@ucl.ac.uk
IRDR Centre for Digital Public Health in Emergencies (dPHE), University College London (UCL)
London, UK

Wellington P. Dos Santos
wellington.santos@ufpe.br
Department of Biomedical Engineering
Federal University Pernambuco
Brazil

Tiago Massoni
massoni@computacao.ufcg.edu.br
Department of Systems and Computing
Federal University of Campina Grande, Brazil

Georgiana Birjovanu
georgiana.birjovanu.15@ucl.ac.uk
IRDR Centre for Digital Public Health in Emergencies (dPHE), University College London (UCL)
London, UK

Patty Kostkova
p.kostkova@ucl.ac.uk
IRDR Centre for Digital Public Health in Emergencies (dPHE), University College London (UCL)
London, UK

ABSTRACT

Geographical maps showing the abundance of the Aedes species (*A. Aegypti* and *A. Albopictus*) in Latin America plays a crucial role in the fight against the Zika Virus (ZIKV). They aid in the identification of sites that promotes mosquito breeding and transmission of ZIKV. In the case of Brazil, one of the greatest factors that favours rapid mosquito reproduction is the presence of stagnated water in the environment. This could be in the form of non-flowing water filled in tanks, barrels, discarded tires, and many other containers situated in human dwellings. After the ZIKV outbreak from 2015, the environmental agency in Brazil have intensively been engaged with routine surveillance of water bodies present in households and the environment to destroy mosquito breeding hotspots as public health measure to prevent vector-to-human transmission of ZIKV. The objective of this study is to use data from their routine surveillance to showcase how our predictive framework based on Neural Networks and Online Extreme Learning Machine (OELM) can predict for Recife (Brazil) at a health district-level the following: firstly, the spatial distribution of the number of properties with water containers contaminated with the Aedes mosquito larvae responsible for ZIKV; and secondly, the spatial distribution of properties with the Aedes mosquito larvae stratified by type of water container. The ultimate goal for this research is to subsequently implement these models to their real-time surveillance data so as an early warning system is present to flag-out spatially the mosquito hotspots on the fly. This system will be built to guide policy makers for directing resources for controlling the mosquito populations thereby limiting transmission to humans.

DPH' 19, November 20–23, 2019, Marseille, France

ACM ISBN 978-1-4503-7208-4/19/11…$15.00
https://doi.org/10.1145/3357729.3357738

CCS CONCEPTS

• **Computer Information Systems**; • **Pubic Health**; • **Epidemiology** → Digital Health Apps; • **Applied Computing**;

KEYWORDS

ZIKA virus (ZIKV), Neural Networks, Online Learning, Extreme Learning Machines, Fuzzy Logic, Classification, Regression

ACM Reference Format:
Adrian Rubio-Solis, Anwar Musah, Wellington P. Dos Santos, Tiago Massoni, Georgiana Birjovanu, and Patty Kostkova. 2019. ZIKA Virus: Prediction of Aedes Mosquito Larvae Occurrence in Recife (Brazil) using Online Extreme Learning Machine and Neural Networks. In *9th International Digital Public Health Conference (2019) (DPH' 19), November 20–23, 2019, Marseille, France.* , 10 pages. https://doi.org/10.1145/3357729.3357738

1 INTRODUCTION

The Zika virus (ZIKV) infection in humans is a global health priority. It is associated with severe morbidity in adults and children, as well as devastating congenital birth complications in neonates [22]. The ZIKV which is classed as an arbovirus is carried by mosquitoes of the genus Aedes species (*A. Aegypti* and *A. Albopictus spp.*) [7, 15]. It is usually transmitted to humans when bitten by infected Aedes spp mosquitoes; however, the virus can also be spread among humans through blood transfusion, unprotected sexual intercourse with an infected partner and from mother–to–child vertically through the intrauterine tract [22, 26, 33].

The current status of the ZIKV notably remains endemic in some sub–Saharan African and South Asian countries with essentially millions of people who are at risk of coming into contact with the Aedes spp - they typically live in areas that are deemed environmentally suitable for ZIKV transmission [2, 9, 17, 25]. While, this is the case and huge resources are poured into these vulnerable areas to maintain surveillance so as to reduce the endemicity status of ZIKV in sub–Saharan Africa, South Asia and beyond; elsewhere, there is an ongoing epidemic of the infection in the Latin America

with countries such as Venezuela, Colombia and with Brazil experiencing one of the greatest burden of the ZIKV ever since its initial outbreak in 2015. In the case of Brazil, it has been documented that over more than 1.5 million cases have been infected with the ZIKV. The spatial distribution of the ZIKV is very noticeable throughout the whole of Brazil with the reported outcomes being more pronounced in 70% of the districts that make-up the Northeast region of the country [5, 11, 22]. It should be noted that ZIKV is not only the vector borne virus that exist in Brazil, other arbovirus−related infections such as Dengue virus (DENV) and Chikungunya virus (CHIKV) remain endemic with the potential for co−occurrence and thus contributing to this existing public health burden [6, 10, 20].

There are several environmental and climatic risk factors that can contribute to the spread of the arboviruses in Brazil. These include factors such as seasonality, temperature, aridity, rainfall, land surface elevation and many more which cannot be managed nor mitigated [8, 27]. One of the greatest factors that favours rapid mosquito breeding is the presence of stagnated water in the environment. This could be in the form of non−flowing water filled in tanks, barrels, discarded tires, and many other containers situated in human dwellings [1, 3]. This physical factor (i.e. stagnant water) can be prevented or controlled. Since 2009, the environmental agency of Brazil has actively been engaged with routine surveillance of standing water present around residential properties, as well as those in various types of water filled containers found inside households. This process of surveillance has been scaled-up significantly ever since the ZIKV outbreak in 2015 where the health officers have intensively been destroying any standing water sources detected with mosquito larvae as a vector control measure to prevent vector-to-human transmission of ZIKV. In collaboration with entomological researchers at the Laboratory of Immunophatology Keizo Asami (LIKA), health officers take water samples from different type of water filled containers and carry out laboratory analysis to determine the density and type mosquito larvae present in them.

With the ongoing surveillance of monitoring the distribution of mosquitoes and ZIKV in real-time, especially in cities most impacted where ZIKV has the highest rates (i.e. Campina Grande, Recife, Jaboatao dos Guararapes and Olinda), it is imperative that such spatially referenced routine data are integrated to predictive models so as to forecast in the future when and where mosquitoes breeding hotspots become a burden. Thus, we have developed a predictive framework based on Neural Networks and Online Extreme Learning Machine (OELM) to make spatial spatiotemporal predictions of mosquito occurrence. This is important because such outputs, in turn, can ultimately be used to serve as an early detection warning systems which can guide Brazilian health officials and policy makers in the fight against ZIKV [18, 19].

In this research, we therefore showcase how our predictive neural framework based on Online Extreme Learning Machine (OELM) can predict for Recife (Brazil) (1) the spatial distribution of the number of properties with water containers contaminated with the Aedes mosquito larvae responsible for ZIKV; and (2) the spatial distribution of properties with the Aedes mosquito larvae stratified by type of water container.

2 PRELIMINARIES

2.1 An Overview of ZIKV Data

The environmental and entomological data (see table 1) used for this study was collected in Recife city from January 2009 to December 2017, inclusive. Recife, is a micro-region situated along the coasts in the State of Pernambuco (figure 1). Recite city, itself, is one of eight major municipalities that forms the region of Recife. The city has the 4th largest urban agglomeration in Brazil with more than 4 million inhabitants. Recife city was selected (along with neighbouring cities: i.e. Campina Grande, Jaboatao dos Guararapes and Olinda) for this study because they have been noted as the population with the greatest burden of ZIKV. The area of Recife city has been delineated into six health zones for which, each zone has been further sub-divided making a total of 63 health districts [i.e. zone 1(4 districts), zone 2 (10 districts), zone 3 (17 districts), zone 4 (9 districts), zone 5 (15 districts) and zone 6 (8 districts)] for the facilitation of further data collection, mapping and surveillance of ZIKV (see Fig. 1). The data set used in this paper (ZIKV data set) is a collection of 3706 experiments (records) carried out by the Laboratory of immunophatology Keizo Asami (LIKA) in Recife, Brazil. According to Table 1, each input vector consists of 13 input variables where rows A1, A2, B, C, D1, D2 and E are the type of water container as the main hot spot for mosquito breeding. In Table 1, all prevalence rates per capita were reported as $1,000$ per households in Recife.

2.2 Extreme Learning Machine for Single Layer Feedforward Networks (SLFNs)

Extreme Learning Machine (ELM) was initially developed to study a more generalised type of SLFNs whose nodes in the hidden layer need not to be tuned [12, 13, 29]. In other words, ELM overcomes the bottleneck of Gradient Descent approaches (GD) that are usually utilised to train SLFNs whose learning speed is in general far slower. ELM is a fast and efficient learning algorithm where hidden nodes in SLFNs are randomly chosen and the value for the corresponding output weights are analytically determined [13]. According to the basics of ELM theory, a SLFN for $'N'$ distinct samples $(\mathbf{x}_i, \mathbf{t}_i)$, in which $\mathbf{x}_i = [x_{i1}, \ldots, x_{in}]^T \in \mathbf{R}^n$ and $\mathbf{t}_i = [t_{i1}, \ldots, t_{im}]^T \in \mathbf{R}^m$, with M hidden nodes and activation $g_{(x)}$ function can be mathematically expressed as [13, 14]:

$$\sum_{i=1}^{M} \beta_i g_i(x_i) = \sum_{i=1}^{M} \beta_i g(w_i \cdot x_i + b_i) = \mathbf{y}_i \tag{1}$$

in which $\mathbf{w}_i = [w_{i1}, \ldots, w_{in}]^T$ and $\beta_i = [\beta_{i1}, \ldots, \beta_{im}]^T$, where $M \leq N$. A SLFN with $i = 1, \ldots, M$ hidden nodes and activation function $g(\mathbf{x})$ can approximate N samples with zero error means $\sum_{i=1}^{M} \parallel \mathbf{y}_j - \mathbf{t}_j \parallel$, i.e. there exist β_i, $\mathbf{w}_i$ and b_i such that:

$$\sum_{i=1}^{M} \beta_i g_i(x_i) = \mathbf{t}_j, j = 1, \ldots, N \tag{2}$$

where Eq. (2) can be written compactly as $\mathbf{H}\beta = \mathbf{T}$, in which $\mathbf{H}$ is:

$$\mathbf{H} = \begin{pmatrix} g(w_1 \cdot x_1 + b_1) & \cdots & g(w_M \cdot x_1 + b_M) \\ \vdots & \vdots & \vdots \\ g(w_1 \cdot x_N + b_1) & \cdots & g(w_M \cdot x_N + b_M) \end{pmatrix}_{N \times M} \tag{3}$$

Table 1: Definition of Input/output Variables of the ZIKV data set - Recife Brazil.

Variable	Min	Max	Definition
A1	0.0	25.0	Water Tank
A2	0.0	132.0	Deposits at ground level, domestic consumption (barrel, tub, clay deposit, tank, well and cistern)
B	0.0	45.0	Water vases, dishes, drippings, ice containers, drinking fountains in general, small ornamental fountains, deposited building material, religious objects/rituals
C	0.0	19.0	Work tanks, drills and backyards, gutters, slabs and awnings in unevenness, disused sanitary drains, construction debris, untreated pools, ornamental fountains, vases / flower pots, shards of glass in walls, inspection and passage boxes
D1	0.0	15.0	Tires and other rolling stock (stains / chambers
D2	0.0	46.34	Garbage (plastic containers, bottles, cans), scraps on old iron yards and recyclers, debris
E	0.0	9.0	Armpits of leaves (bromeliads, etc.), holes in trees and rocks, bark of animals (hooves and shells)
TB Ae	0.0	3.0	Aedes aegypti mosquitoes
Outros Ae	0.0	125.0	Other mosquito species
TB Al	0.0	15.0	Aedes albopictus Mosquitoes
Outros Al	0.0	15.0	Other mosquito species
Mosquitoes	0.0	163.0	Number of larvae/eggs counted per each distric at different time of the year. i.e. month and year
Traps	0.0	278.8	traps installed in properties by the health agents to let female mosquitoes lay eggs and then get collected, eggs counted in the lab and destroyed.

Where the terms $\beta = (\beta_1, \ldots, \beta_M) \in R^{M \times m}$ and $\mathbf{T} = (\mathbf{t}_1, \ldots, \mathbf{t}_P) \in R^{N \times m}$. $\mathbf{H}$ is the hidden layer output matrix of an SLFN with respect to the inputs $\mathbf{x}_i$. The minimum norm least-squares solution of the linear system $\mathbf{H}\beta = \mathbf{T}$ is unique and can be achieved by calculating the pseudo-inverse $H^{\dagger}$ as follows:

$$\hat{\beta} = \mathbf{H}^{\dagger}\mathbf{T} \tag{4}$$

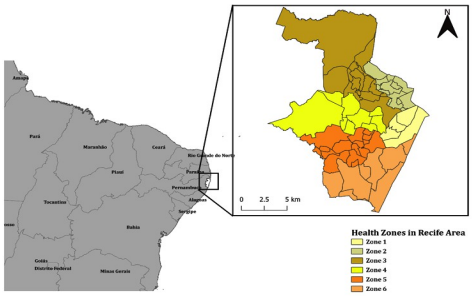

Figure 1: Study area shows the map of Recife, Brazil, situated in the State of Pernambuco delineated into six health zones, and the zones further sub-divided into 63 health districts

2.3 Fuzzy Extreme Learning Machine

As detailed in [28], a class of Fuzzy Inference Systems (FISs) can be viewed as a class of SLFNs used in the area of function approximation and classification problems [16]. According to ELM theory [28, 30], a Fuzzy Inference System (FIS) can be interpreted as an SLFN if for a given number of distinct training samples $(\mathbf{x}_p, \mathbf{t}_p)$ ELM can be directly applied. An FIS with M fuzzy rules is given by:

$$\mathbf{y}_p(\mathbf{x}_p) = \sum_{k=1}^{M} \beta_k G(\mathbf{x}_p, \mathbf{c}_k, a_k) = \mathbf{t}_p, \; p = 1, \ldots, P \tag{5}$$

where, $\mathbf{x}_p = [x_{p1}, \ldots, x_{pN}] \in \mathbf{R}^N$ and $\mathbf{t}_p = [t_{p1}, \ldots, t_{p\tilde{N}}] \in \mathbf{R}^{\tilde{N}}$. An FIS either of Takagi-Sugeno-Kang (TSK) or Mamdani type can be defined by a number of fuzzy rules R^k of the form [29]

$$\begin{aligned} R^k : \text{IF } x_{p1} \text{ is } A_{1k} \text{ AND } x_{p2} \text{ is } A_{2k} \text{ AND } \ldots \\ \text{IF } x_N \text{ is } A_{Nk} \text{ THEN } (y_1 \text{ is } \beta_{k1}) \ldots (y_{\tilde{N}} \text{ is } \beta_{k\tilde{N}}) \end{aligned} \tag{6}$$

where, $A_{sk}(s = 1, \ldots, N, k = 1, \ldots, M)$ are the fuzzy sets that correspond to the sth input variable x_{ps} in the kth rule, where $\tilde{N}$ is the dimension of the pth output vector $\mathbf{y}_p = [y_1, \ldots, y_{\tilde{N}}]$. When an FIS employs a TSK inference engine, β_{kl} $(k = 1, \ldots, M, l = 1, \ldots, \tilde{N})$ is defined by a linear combination of input variables, i.e. $\beta_{kl} = q_{kj,0} + q_{kj,0}x_1 + \ldots q_{kj,N}x_N$, otherwise if the FIS is of Mamdani type, β_{kl} is a crisp value. In Fuzzy Logic System theory (FLS), the degree to which any given input x_{ps} satisfies the quantifier A_{sk} is specified by its Membership Function (MF) $\mu_{A_{ks}}(c_{ks}, a_k)$. By using the symbol $\otimes$ for the representation of fuzzy logic AND operations, firing strength of the kth fuzzy rule can be computed as

$$\begin{aligned} R^k(\mathbf{x}_p; \mathbf{c}_k, a_k) = \mu_{A_{k1}}(x_{p1}, c_{k1}, a_k) \\ \otimes \mu_{A_{k2}}(x_{p2}; c_{k2}, a_k) \otimes \ldots \otimes \mu_{A_{kN}}(x_{pN}; c_{kN}, a_k) \end{aligned} \tag{7}$$

Each fuzzy rule R^k can be normalised as

$$G(\mathbf{x}_p; c_{ks}, c_k) = R^k(\mathbf{x}_p; c_k, a_k) \Big/ \sum_{k=1}^{M} R^k(\mathbf{x}_p; c_k, a_k) \tag{8}$$

Consequent parameters are defined as: $\beta_k = \mathbf{x}_{p,e}^T \mathbf{q}_k$. For a Mamdani fuzzy model, $\mathbf{x}_{p,e} = 1$, and $\mathbf{q}_k = \beta_k = [\beta_{k1}, \ldots, \beta_{k\tilde{N}}]^T$, where $\mathbf{q}_k$ is a weight vector of crips values. For a TSK fuzzy model $\mathbf{x}_{p,e} = [1\ \mathbf{x}_p^T]^T$ is the extended version of $\mathbf{x}_p$.

$$\mathbf{q}_k = \begin{pmatrix} q_{k1,0} & \cdots & q_{k\tilde{N},0} \\ \vdots & & \vdots \\ q_{k1,N} & \cdots & q_{k\tilde{N},N} \end{pmatrix}_{(N+1)\times\tilde{N}} \tag{9}$$

Therefore, Eq. (5) becomes $\mathbf{y}_p(\mathbf{x}_p) = \sum_{k=1}^{M} \mathbf{x}_{p,e}^T \mathbf{q}_k G(\mathbf{x}_p, \mathbf{c}_k, a_k)$. A compact representation for Eq. (5) is then given by $\mathbf{HQ} = \mathbf{T}$, in which, $\mathbf{Q}$ is the matrix of coefficients $q_{kj,s}$.

2.4 Bayesian Extreme Learning Machine

Bayesian Extreme Learning Machine (Bayesian-ELM) is a linear regression technique used to estimate the output weights of SLFNs. Similarly to Bayesian modelling, Bayesian-ELM is carried out into two steps [23, 34]:

1 Inference of the posterior distribution of the model parameters and the likelihood function $P(\mathbf{w}|D) \propto P(\mathbf{w})P(D|\mathbf{w})$, where $'D'$ is the data set.
2 Calculation of the output distribution of the model y_{new} for a new input x_{new}, where the posterior distribution for $\mathbf{w}$ is:

$$P(y_{new}|x_{new}, D) = \int P(y_{new}|x_{new}, \mathbf{w})P(\mathbf{w}|D)d\mathbf{w} \quad (10)$$

From the definition of a linear model $y = \mathbf{h}^t\mathbf{x} + \epsilon$, ϵ follows a normal distribution with zero mean and variance $N(0, \sigma^2)$. Thus, the conditional probability is computed as:

$$p(y|\mathbf{x}, \mathbf{h}, \sigma^2) = N(\mathbf{h}^t\mathbf{x}; \sigma^2) \quad (11)$$

Similarly to [34], parameter distribution is $p(\mathbf{h}|\alpha) = N(\mathbf{0}; \alpha^{-1}\mathbf{I})$, where $\mathbf{I}$ is the identity matrix, and α a hyperparameter. In this study, the prior probability and likelihood as well as the posterior follow a Gaussian distribution, where the mean $\mathbf{m}$ and $\mathbf{S}$ are [34]:

$$\mathbf{m} = \sigma^{-2} \cdot \mathbf{S} \cdot \mathbf{X^T} \cdot \mathbf{y};\ \mathbf{S} = \left(\alpha\mathbf{I} + \sigma^{-2} \cdot \mathbf{X^T}\mathbf{X}\right)^{-1} \quad (12)$$

In this work, Bayesian-ELM is used to train a SLFN based on Radial Basis Functions (RBFs). α is a regularisation term, where each parameter in (13)-(14) are updated using the iterative evidence procedure [34].

$$\gamma = N - \alpha \cdot trace[\mathbf{S}];\ \alpha = \gamma/(\mathbf{m^{-T}m}) \quad (13)$$

$$\sigma^2 = \sum_{i=0}^{P}(y_i - \mathbf{m^T}\mathbf{x}_i)/(P - \gamma) \quad (14)$$

The iterative process of calculating $\mathbf{m}$ is stopped when the difference of the norm of $\mathbf{m}$ between successive iterations falls below a predefined value [34]. Therefore, the output follows the distribution $p(y_{new}|\mathbf{y}, \alpha, \sigma^2) = N(\mathbf{h^T} \cdot \mathbf{x}_{new}; \sigma^2(\mathbf{x}))$ where the term $\sigma^2(\mathbf{x}) = \sigma^2 + \mathbf{x}^{\mathbf{T}}_{new}\mathbf{S} \cdot \mathbf{x}_{new}$, where $\mathbf{x}_{new}$ is a n.

2.5 Interval Type-2 Radial Basis Function Neural Network (IT2-RBFNN)

The IT2-RBFNN is a neural network model based on the fusion between fuzzy logic of Interval Type-2 (IT2) and Radial Basis Function neural networks (RBFNNs) theory. The IT2-RBFNN has proven its efficiency to system modelling in the presence of noisy signals usually outperforming its type-1 counterparts such as the RBFNN and fuzzy T1 SLFNs. In this work, an IT2-RBFNN of Mamdani type having a center-of-sets type reduction, product inference rule and a singleton output space is used. The type-reduced set $[y_l, y_r]$ is obtained by using a Karnik-Mendel algorithm (KM) [35]. Therefore, according to Fig. 2, a matrix representation of an IT2-RBFNN output must be obtained in order to apply ELM as follows:

$$y_f = \frac{1}{2}\left(\mathbf{Y}_l + \mathbf{Y}_r\right)\mathbf{w}^T \quad (15)$$

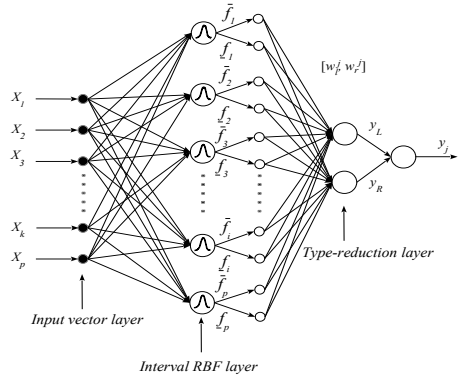

Figure 2: Neural structure of an IT2-RBFNN

in which $y_l = \mathbf{Y}_l\mathbf{w}^T$ and $y_r = \mathbf{Y}_r\mathbf{w}^T$ and

$$\mathbf{Y}_l = \frac{\bar{\mathbf{f}}^T Q^T E_1^T E_1 Q + \underline{\mathbf{f}}^T Q^T E_2^T E_2 Q}{r_l^T Q\bar{\mathbf{f}} + s_l^T Q\underline{\mathbf{f}}} \quad (16)$$

where each value in the reduced set $[\mathbf{Y}_r, \mathbf{Y}_l]$ is calculated using $\mathbf{Y}_l = (\psi_{l,1}, \ldots, \psi_{l,M})$, $E_1 = (e_1|e_2|\ldots|e_L|\mathbf{0}|\ldots|\mathbf{0})^T\ L \times M$, $E_2 = (\mathbf{0}|\ldots|\mathbf{0}|\xi_1|\xi_2|\ldots|\xi_{M-L})^T\ (M-L) \times 1$ [24, 30], $e_j \in R^L$ $(j = 1, \ldots, L)$, $\xi_j \in R^{M-L}$, $j = 1, \ldots, M-L$, and the term $\mathbf{Y}_r$

$$\mathbf{Y}_r = \frac{\underline{\mathbf{f}}^T Q^T E_3^T E_3 Q + \bar{\mathbf{f}}^T Q^T E_4^T E_4 Q}{r_r^T Q\underline{\mathbf{f}} + s_l^T Q\bar{\mathbf{f}}} \quad (17)$$

where $\mathbf{Y}_r = (\psi_{r,1}, \ldots, \psi_{r,M})$, $E_3 = (e_1|e_2|\ldots|e_R|\mathbf{0}|\ldots|\mathbf{0})^T\ R \times M$, $E_4 = (\mathbf{0}|\ldots|\mathbf{0}|\xi_1|\xi_2|\ldots|\xi_{M-R})^T\ (M-R) \times 1$. The terms $e_j \in R^R$ $(j = 1, \ldots, R)$ and $\xi_j \in R^{M-R}$, $j = 1, \ldots, M-R$ as the elementary vectors. $\underline{\mathbf{f}} = (\underline{f}_1, \ldots, \underline{f}_M)^T$, $\bar{\mathbf{f}} = \left(\bar{f}_1, \ldots, \bar{f}_M,\right)^T$.

By using Karnik-Mendel algorithms [35], the reordered consequent weights $\tilde{\mathbf{w}}$ result from a permutation process to find the switching points L and R are calculated $\tilde{\mathbf{w}} = Q\mathbf{w}^T,\ Q \in R^{M \times M}$. In which $\mathbf{w} = (w_1, \ldots, w_M)$ is the set of original rule-ordered consequent weights and Q is the corresponding permutation matrix. A firing strength F^i in the hidden layer of the IT2-RBFNN is a Gaussian function having a fixed mean m_s^i and an **uncertain standard deviation** $[\sigma_i^1, \sigma_i^2]$ when $x_s = x_l'$, such as $F^i := [\underline{f}_i(\vec{x}_p), \bar{f}_i^{\alpha_s}(\vec{x}_p)]$.

2.6 Online Extreme Learning Machine (OLEM)

OELM was assumes that all training data is available [21]. However, for real time applications data usually arrive chunk-by-chunk or one-by-one (a special case of chunk). In genereal, OELM can be applied to a large number of SLFNs including functionally equivalent neural networks [21, 29]. O-ELM usually involves two main steps:

1) Initialisation Step: at this step, training is initialised by using a small chunk of data $\aleph_0 = \{(\mathbf{x}_i, \mathbf{t}_i)\}_{i=1}^{N_0}$ from the training data set $\aleph = \{(\mathbf{x}_i, \mathbf{t}_i)|\mathbf{x}_i \in \mathbf{R}^n, \mathbf{t}_i \in \mathbf{R}^m\}$, $N_0 \geq \tilde{N}$.
   a) Assign random weights a_i and bias b_i (for additive nodes in neural networks) or membership functions parameters $(\mathbf{c}_i, a_i), i = 1, \ldots, L$ for neural fuzzy networks [29, 31].
   b) Calculate the hidden layer output matrix for either neural networks or neural fuzzy networks $\mathbf{H}_0$.

2) **Sequential Learning Phase**. Present the $(k+1)th$ chunk of new observations

$$\aleph_{k+1} = \{\mathbf{x}_i, \mathbf{t}_i\}_{i=(\sum_{j=0}^{k} N_j)+1}^{\sum_{j=0}^{k+1} N_j}$$

in which, the term N_{k+1} denotes the number of samples in the $(k+1)$ chunk. Do the following:

a) Calculate the partial matrix $\mathbf{H}_{k+1}$ that corresponds to the $k+1$ chunk of data $\aleph_{k+1}$ such as for neural networks with additive nodes:

$$\mathbf{H}_{k+1} = \mathbf{H}\left(\mathbf{c}_1, \ldots, \mathbf{c}_L, a_1, \ldots, a_L; \mathbf{x}_{(\sum_{j=0}^{k} N_j)+1}, \ldots, \mathbf{x}_{(\sum_{j=0}^{k+1} N_j)}\right) \tag{18}$$

b) Calculate the output weight $\beta^{(k+1)}$

$$\mathbf{P}_{k+1} = \mathbf{P}_k - \mathbf{P}_k\mathbf{H}_{k+1}^T(\mathbf{I} + \mathbf{H}_{k+1}\mathbf{P}_k\mathbf{H}_{k+1}^T)^{-1}\mathbf{H}_{k+1}\mathbf{P} \tag{19}$$

$$\beta^{(k+1)} = \beta^{(k)} - \mathbf{P}_{k+1}\mathbf{H}_{k+1}^T(\mathbf{T}_{k+1} - \mathbf{H}_{k+1}\beta^{(k)}) \tag{20}$$

c) Set $k = k+1$, then go to step 2.

O-ELM first identifies the parameters of SLFNs by implementing a typical ELM with some initial batch of training data (initialisation step). In the initialisation phase, $\mathbf{H}_0$ is filled up where the size of the data should be at least the number of hidden units. Once this process is completed, O-ELM will learn new training data one-by-one or chunk-by-chunk and then all the training data will be discarded.

3 EXPERIMENTS

In this paper, two types of experiments are performed, namely: a) prediction of the number of properties with water containers contaminated with mosquito larvae (see Fig. 4), and for the b) prediction of the type of water container (see Fig. 5) employed during each experiment at Recife, Brazil. In this paper a number of three different ELM methodologies based on fuzzy logic of interval type-2 [29], logistic regression [13] and Bayesian learning [34] are employed to evaluate the performance of three different OELM approaches, i.e. 1) OELM, 4) Online Fuzzy ELM (OFELM) [28] and 6) an OELM using a Radial Basis Function Neural Network (OELM-RBFNN) as the main neural model [29]. Both type of experiments involve five random simulations where for cross validation, the ZIKV data set is divided into two subsets, i.e. 80% for training, and 20% for testing. To predict the number of containers, each ELM structure is used as a regression model whose inputs are the number of mosquitoes and number of traps as shown in Fig. 3, where the desired output t_j of each ELM is defined as:

$$t_j = TB\ Ae + Outros\ TB\ Ae + TB\ Al + Outros\ TB\ Al \tag{21}$$

For the prediction of the type of water container, two different neural arrangements are implemented. The former involves a Multi-Input-Multi-Output (MIMO) neural structure where the type of container is binary codified as follows:

A1	A2	B	C	D1	D1	E	
↓	↓	↓	↓	↓	↓	↓	
1	0	1	0	0	1	1	= Type of water container

The latter involves a number of seven Multi-Input-Single-Output (MISO) neural structures, each one used to predict the probability of what type of water container was used during the experiments as illustrated in Fig. 4.

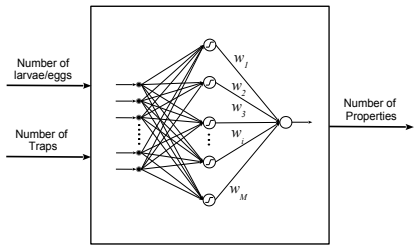

Figure 3: OELM architecture used as a regression model for the prediction of the number of Properties where mosquito larvae were found in water containers.

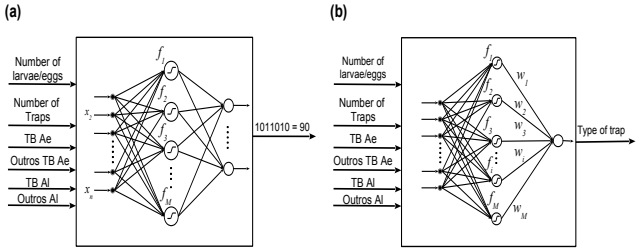

Figure 4: Neural arrangements used to predict the type of water containers. (a) MIMO and (b) MISO neural structures.

This section presents results according to the prediction of the number of properties contaminated with mosquito larvae, and the prediction about the type of water container. For cross-validation purposes, ELM models with the highest trade-off between model simplicity, average testing performance and training time are used as reference models to illustrate prediction results.

3.1 Prediction of the number of Properties with water containers contaminated with mosquito larvae

This section uses the testing performance results obtained by an OELM in order to provide a descriptive interpretation for the prediction of the number of properties with water containers that were contaminated with mosquito larvae during each experiments. To illustrate such performance, the average testing behaviour that corresponds to the year 2009, 2013 and 2017 are presented in Fig. 5. For example, from Fig. 5(a) it can be noticed during the year 2009, the highest suitability for ZIKV is found in northern regions of Recife, where a high density of properties with water containers contaminated with mosquito larvae are detected. Such distribution reduces significantly in 2013 (See Fig. 5(b)) with only three districts having on average of at least 50 properties. By 2017, this distribution increases significantly, whereby the average number of properties with contaminated water containers are more pronounced in the north western parts of Recife (See Fig. 5(c)).

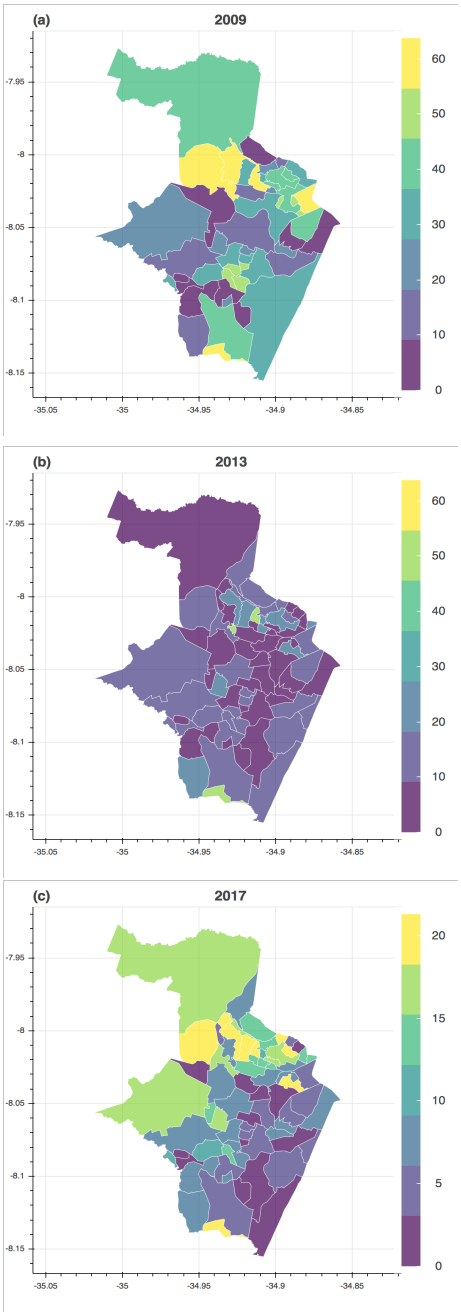

Figure 5: Average prediction of properties where mosquito larvae was found in water containers for the year (a) 2009, (b) 2013 and (c) 2017

Between 2009 and 2017, it can be concluded that the highest average number of properties can be found in some northern and southern districts, making those regions ideal to the spread of ZIKV. In this sense, the highest concentration about the number of properties can be found in a small southern district closely located to the geographic coordinate (−8.15, −34.95). Opposite to this, districts with the smallest number of properties with contaminated water containers can be found in the south east area of Recife. In particular, the district located at the geographical coordinate (-8.05, -34.87). To confirm this performance, a temporal description using time series that correspond to the years 2009, 2013 and 2017 is presented in Figures 6(a), 6(b) and 6(c) respectively.

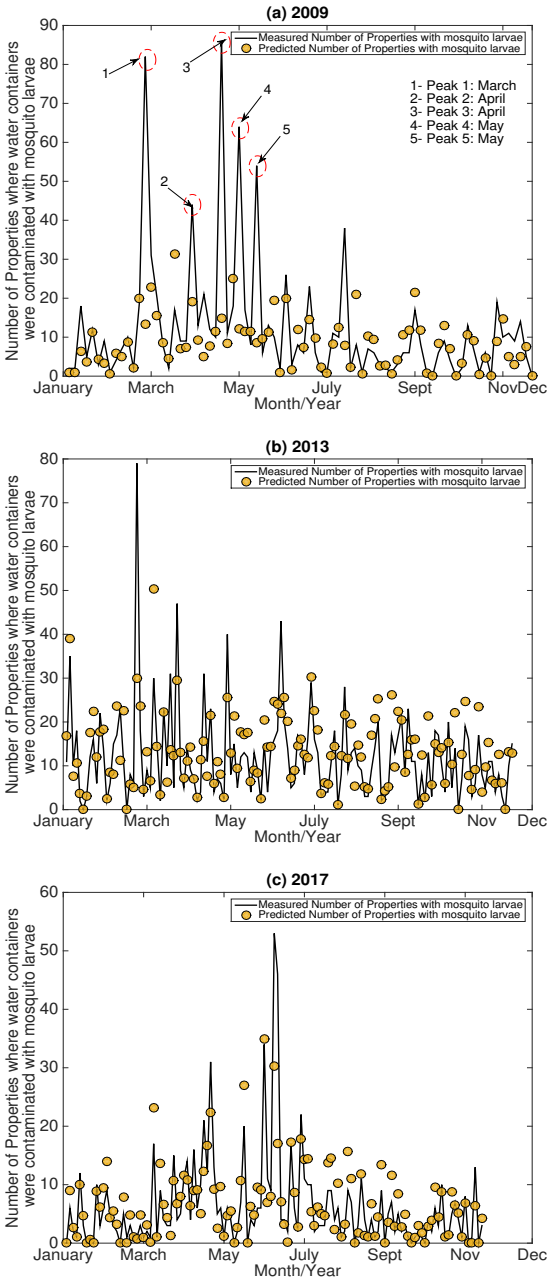

Figure 6: Testing Time series prediction obtained by an OELM for the prediction of the number of properties with water containers contaminated with mosquito larvae that correspond to the years (a) 2009, (b) 2013 and (c) 2017.

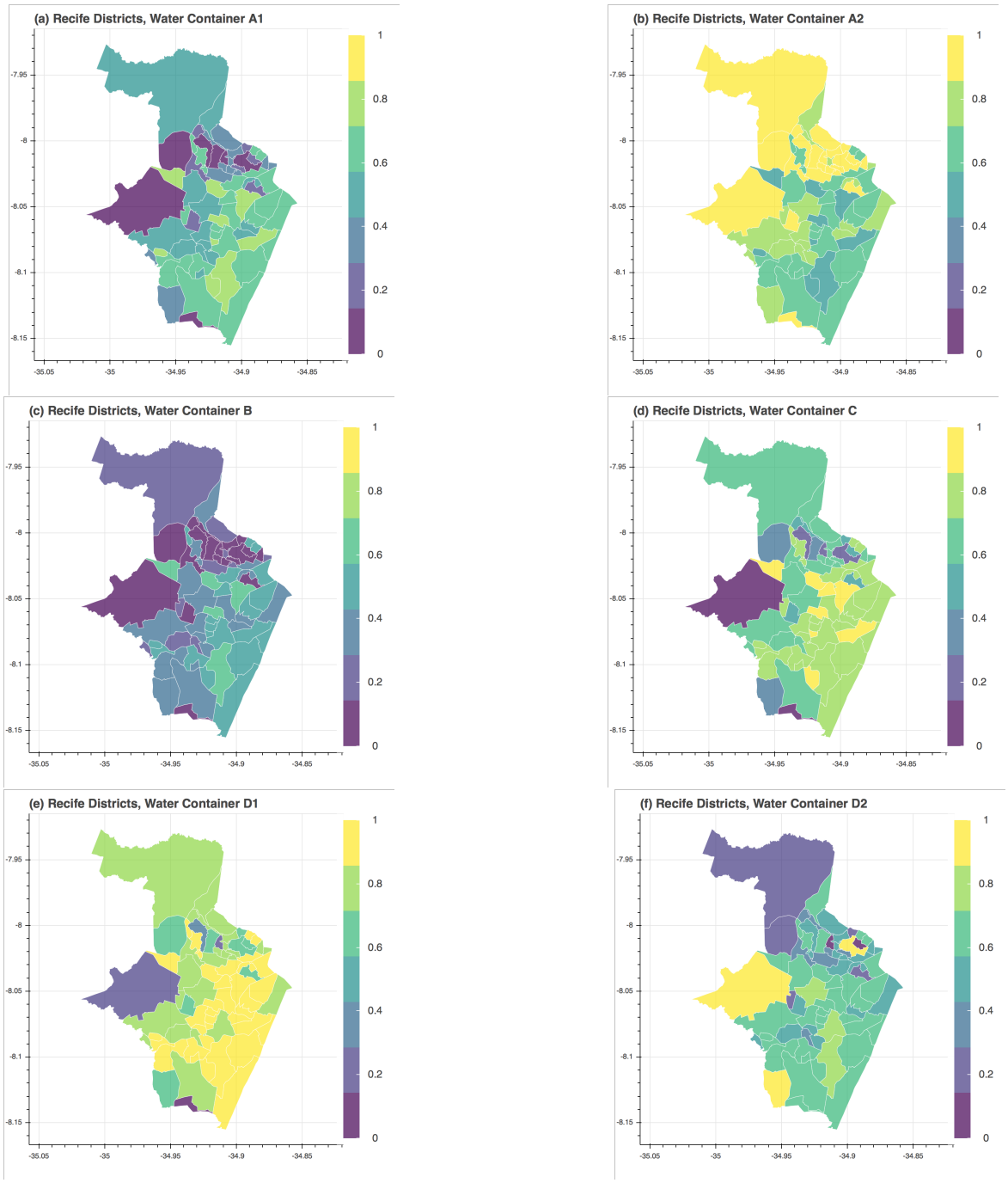

Figure 7: Probability found by an OELM-RBF for the prediction of the water tank of type: (a) A1, (b) A2, (c) B, (d) C, (e) D1 and (f) D2 for the year 2017.

As described in [4], in northern Brazil, most of the year is hot, humid and rainy. According to [4], rains in Brazil are more abundant between December and May. This is reflected in the suitability of environmental conditions to an increase in the number of mosquitoes in a large number of districts in Recife. In general, it can be noticed from Fig. 6, an important increase about the amount of mosquitoes mostly occurs between March and June. For example, in Fig. 6(a), a number of five peaks can be observed particularly between March and May. This trend depicts a similar behaviour for years between 2013 and 2017 with outliers between March and May.

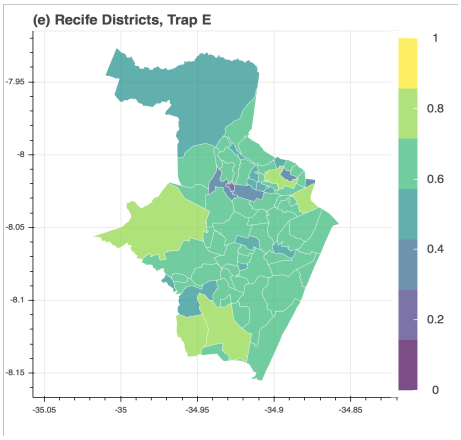

Figure 8: Probability found by an OEL-RBFNN for the prediction of the water container E

.

.

3.2 Prediction of the type of water container

This section provides a probabilistic description about the results obtained for the prediction of the type of water container used at each district in Recife in the year 2017. Such experiment results are obtained by using an Online Extreme Learning Machine (OELM) whose main inference engine is based on a Radial Basis Function Neural Network (OELM-RBFNN) [30, 32]. From Fig. 7, the map for Deposits at ground level, domestic consumption (barrel, tub and clay deposit (water container A2, See Fig. 7(b)) is the most popular type of container, in particular in Northern regions of Recife with a probability that ranges between 0.4 and 1.

In relation to the probability distribution for the water containers of type D1 - Tires, other rolling stock (water container D1) and Garbage (plastic containers, bottles, cans), scraps on old iron yards and debris (container D2), the level of highest occurrence of mosquitoes are particularly concentrated in southern regions. Equally, from the map used to illustrate the level of occurrence for armpits of leaves, holes in trees and rocks and bark of animals (water container E, See Fig. 8), a probability that typically ranges between 0.4 and 0.8 is found in the majority of the districts in Brazil. In general, from Fig. 7 and 8, it can be observed that those regions that are susceptible to contribute to the spread of ZIKV are within those districts where water containers of type A1, C and D1 have been found, while the least popular are of type A2, B and E.

3.3 Model Validation Analysis

This section provides an analysis obtained by the six neural network models implemented for the prediction of the number of properties with water containers contaminated with mosquito larvae and to determine the type of water container employed throughout the experiments (A1, A2, B, C, D1 and D2) between the years 2009 and 2017. First, to evaluate the ability of OELM based neural networks to predict the number of properties with water containers that are contaminated with mosquito larvae at each district in Recife Brazil, in Table 2, the average Root-Mean-Square-Error (RMSE) of six random experiments is presented. Based on Table 2, OELM, OELM-RBFNN and Bayesian-ELM produced the highest trade-off between generalisation performance, training time and model simplicity. ELM is a batch learning method that requires a predefined number of hidden neurons and parameters as well as a predefined number of training samples. Compared to ELM, OELM can learn training data in an online way either one-by-one or chunk by chunk (with fixed or varying size) while updating its model parameters.

From Table 2, it can also be observed that the number of samples that are required to train an ELM and IT2-ELM corresponds to one chunk of 2965 random samples. From Table 2, it is evident that for the prediction of the number of water containers, OELM-based strategies and Bayesian ELM represent the best option, not only because the number of hidden units is smaller compared to traditional ELM and the IT2-RBFNN, but also their are more accurate and faster. In this sense, the correlation between the ZIKV data set and the predicted number of properties with infected containers using an OELM is shown in Fig. 9. According to From Fig. 9, OELM is able to predict more efficiently those points that are defined in the interval [0-60], where predicted number of properties between the year 2009 and 2017 are indicated by purple points. From Fig. 9(b), it can be noted that a certain level of outlierness is present in the ZIKV data set. Particularly to predict unseen data. In relation to the prediction of the type of water containers, in Table 3 and 4, the average performance using MIMO and MISO neural arrangements are presented respectively. Opposite to Table 2, in Table 3, MISO models OELFM and OELM-RBFNN are the most suitable learning methodologies to provide the highest trade off between model accuracy and average training time. In Table 4, the average performance of 7 MISO neural arrangements is presented, where the highest performance is achieved by OELM and OELM-RBFNN.

4 DISCUSSION AND FUTURE WORK

According to our experiments, a high percentage of properties with water containers is more clear between the months March and May when the rainy season has reached its peak. Based on the accuracy performance obtained by each neural structure, it is clear that OELM offers the highest trade-off between computational load, model simplicity and generalisation performance. However, a level of outlierness is observed by using any type of ELM-based technique. In this sense, even the ZIKV data set is highly sparse and shows a certain level of outlierness, the prediction accuracy reached by OELMs as MISO engines is about 80%. This favours an appropriate prediction of an average number of properties that ranges between 0-60. Moreover, the number of properties with water containers infected with mosquito larvae are high, specially in the years 2009 and 2017. In terms of water container prediction, it can be concluded that deposits at ground level (A1) and plastic containers (D2) are the most common type of container where mosquito breeds, while the less popular place for breeding among mosquitoes are water tanks (A1) and water vases and fountains (B). Thus, advantages/disadvantages are listed below:

- State-of-the-art online methods allow collating mosquito data in real-time spanning over a period of time to make spatiotemporal predictions of mosquito occurrence in Recife.

Table 2: Average RMSE achieved by OELM and ELM for the prediction of the number of properties.

Model	# hidden units	# samples per training chunk	Avg. Training time (sec)	Avg. Training RMSE	Avg. Testing RMSE
ELM	200	2965	4.54s	9.07	10.01
IT2-FELM	210	2965	7.82s	8.76	8.91
OELM	10	300	1.39s	8.02	8.31
OFELM	10	300	1.41s	8.36	10.49
OELM-RBFNN	10	2965	1.37s	8.27	8.53
Bayesian-ELM	12	2965	1.40s	8.55	8.72

Table 3: Average accuracy achieved by MIMO neural networks for the classification of the type of water containers.

Model	# hidden units	# samples per training chunk	Avg. Training Time (sec)	Avg. Training Accuracy	Avg. Testing Accuracy
ELM	300	2965	7.23s	72.19%	69.09%
IT2-FELM	150	2965	8.79s	74.10%	68.29%
OELM	10	100	1.41s	69.87%	68.93%
OFELM	10	110	1.03s	70.65%	70.02%
OELM-RBFNN	10	100	1.12s	71.80%	71.32%
Bayesian-ELM	30	2965	2.59s	73.22%	68.08%

Table 4: Average accuracy achieved by MISO neural networks for the classification of the type of water containers.

Model	# hidden units	# samples per training chunk	Avg. Training Time (sec)	Avg. Training Accuracy	Avg. Testing Accuracy
ELM	300	2965	8.22s	80.21%	79.04%
IT2-FELM	180	2965	9.10s	82.01%	81.29%
OELM	10	120	1.55s	83.02%	82.87%
OFELM	10	120	1.13s	80.04%	79.22%
OELM-RBFNN	10	110	1.22s	83.24%	81.17%
Bayesian-ELM	30	2965	3.11s	82.03%	80.11%

- This study uses a robust and interdisciplinary method which connects different theories from machine learning.
- One limitation of this study is that the predictions are made at district-level (and not disaggregated to a household) - because interpretations are made at a district-level (instead of household-level) one form of bias that might be present in our result is called ecological fallacy.
- There are several other risk factors that must be considered in future studies - these include vegetation index, temperature etc. so there is residual confounding.
- In its current form, the neural structures used for ZIKV prediction are data-driven models based on the knowledge acquired from a number of scientific experiments.

Future research work, involves the development of m-gamification applications in mobile technology based on OELM and neural networks, as well as other type of geostatistical approaches that include other variables such as climate, vegetation, population density, etc. Thus, the results presented in this paper are the basis to create a Geographic Information System (GIS) where mosquito data can be collated in real-time spanning over a period to make spatial spatio temporal predictions of mosquito occurrence in Recife.

5 CONCLUSIONS

In this study, a machine learning framework that is based on neural networks and Online Extreme Learning Machine (OELM) is applied to the prediction of susceptible areas in Recife, Brazil, where new cases of ZIKA Virus may appear. Two key factors are predicted - first, the number of properties across Recife with water containers that are contaminated with different type of mosquitoes such as aedes aegypti and aedes albopictus. Secondly, this study aims at predicting the type of water container at each district where usually mosquitoes breed. Data that results from a health surveillance between 2009 and 2017 in Recife, and carried by the laboratory of immunophatology Keizo Asami (LIKA) is employed to train the proposed predictive framework. In order to identify a model that offers the highest trade-off between computational load, model simplicity and generalisation accuracy, five different neural structures whose training is based on ELM and OELM are implemented.

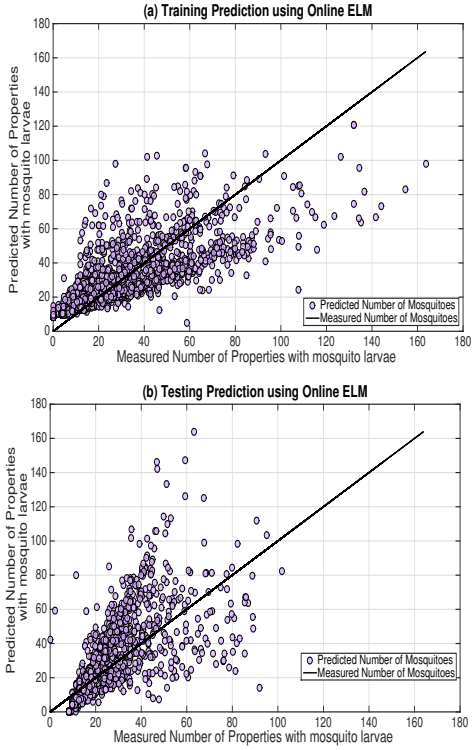

Figure 9: Random (a) training and (b) testing data fit obtained by an OELM for the Prediction of the number of properties with water containers contaminated with mosquito larvae.

REFERENCES

[1] Cesar Basso, Ruben M Caffera, Garcia da Rosa, Rosario Lairihoy, Cristina Gonzalez, Walter Norbis, and Ingrid Roche. 2012. Mosquito-Producing Containers, Spatial Distribution, and Relationship between Aedes aegypti Population Indices on the Southern Boundary of its Distribution in South America (Salto, Uruguay). *The American Journal of Tropical Medicine and Hygiene* 87, 6 (2012), 1083–1088.

[2] Juan D Beltrán, Andrei Boscor, Wellington P dos Santos, Tiago Massoni, and Patty Kostkova. 2018. 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. In *Proceedings of the 2018 International Conference on Digital Health*. ACM, 90–94.

[3] Giovanni Benelli and Heinz Mehlhorn. 2016. Declining malaria, rising of Dengue and Zika virus: insights for mosquito vector control. *Parasitology Research* 115, 5 (2016), 1747–1754.

[4] RV Calheiros and I Zawadzki. 1987. Reflectivity-rain rate relationships for radar hydrology in Brazil. *Journal of climate and applied meteorology* 26, 1 (1987), 118–132.

[5] T Nunes de Lima-Camara and N Alves Honorio. 2016. Climate change and its effect on urban mosquitoes in South America. *Climate change impacts on urban pests* 10 (2016), 127.

[6] Ed De Quincey and Patty Kostkova. 2009. Early warning and outbreak detection using social networking websites: The potential of twitter. In *International Conference on Electronic Healthcare*. Springer, 21–24.

[7] Anthony S Fauci and David M Morens. 2016. Zika virus in the Americas—yet another arbovirus threat. *New England Journal of Medicine* 374, 7 (2016), 601–604.

[8] Trevon L Fuller, Guilherme Calvet, Camilla Genaro Estevam, et al. [n.d.]. Behavioral, climatic, and environmental risk factors for Zika and Chikungunya virus infections in Rio de Janeiro, Brazil, 2015-16. *PloS ONE* 12, 11 ([n. d.]), e0188002.

[9] Trevon L Fuller, Guilherme Calvet, Camila Genaro Estevam, Jussara Rafael Angelo, Gbenga J Abiodun, Umme-Aiman Halai, Bianca De Santis, Patricia Carvalho Sequeira, Eliane Machado Araujo, Simone Alves Sampaio, et al. 2017. Behavioral, climatic, and environmental risk factors for Zika and Chikungunya virus infections in Rio de Janeiro, Brazil, 2015-16. *PloS one* 12, 11 (2017), e0188002.

[10] Anne Gulland. 2016. WHO urges countries in dengue belt to look out for Zika.

[11] Jorg Heukelbach, Carlos Henrique Alencar, Alyson Ann Kelvin, Wanderson Kleber de Oliveira, and Luciano Pamplona de Góes Cavalcanti. 2016. Zika virus outbreak in Brazil. *The Journal of Infection in Developing Countries* 10, 02 (2016), 116–120.

[12] Guang-Bin Huang, Hongming Zhou, Xiaojian Ding, and Rui Zhang. 2012. Extreme learning machine for regression and multiclass classification. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 42, 2 (2012), 513–529.

[13] Guang-Bin Huang, Qin-Yu Zhu, and Chee-Kheong Siew. 2006. Extreme learning machine: theory and applications. *Neurocomputing* 70, 1-3 (2006), 489–501.

[14] Guang-Bin Huang, Qin-Yu Zhu, Chee-Kheong Siew, et al. 2004. Extreme learning machine: a new learning scheme of feedforward neural networks. *Neural networks* 2 (2004), 985–990.

[15] Farakh Javed, Khanzadi N Manzoor, Mubashar Ali, Irshad U Haq, Abid A Khan, Assad Zaib, and Sobia Manzoor. 2018. Zika virus: what we need to know? *Journal of basic microbiology* 58, 1 (2018), 3–16.

[16] Wu Jun, Wang Shitong, and Fu-lai Chung. 2011. Positive and negative fuzzy rule system, extreme learning machine and image classification. *International Journal of Machine Learning and Cybernetics* 2, 4 (2011), 261–271.

[17] Mahmoud Kamal, Mohamed A Kenawy, Magda Hassan Rady, Amany Soliman Khaled, and Abdallah M Samy. 2018. Mapping the global potential distributions of two arboviral vectors Aedes aegypti and Ae. albopictus under changing climate. *PloS one* 13, 12 (2018), e0210122.

[18] Patty Kostkova. 2013. A roadmap to integrated digital public health surveillance: the vision and the challenges. In *Proceedings of the 22nd International Conference on World Wide Web*. ACM, 687–694.

[19] Patty Kostkova. 2018. Disease surveillance data sharing for public health: the next ethical frontiers. *Life sciences, society and policy* 14, 1 (2018), 16.

[20] Patty Kostkova, Martin Szomszor, and Connie St Louis. 2014. # swineflu: The use of twitter as an early warning and risk communication tool in the 2009 swine flu pandemic. *ACM Transactions on Management Information Systems (TMIS)* 5, 2 (2014), 8.

[21] Nan-Ying Liang, Guang-Bin Huang, Paramasivan Saratchandran, and Narasimhan Sundararajan. 2006. A fast and accurate online sequential learning algorithm for feedforward networks. *IEEE Transactions on neural networks* 17, 6 (2006), 1411–1423.

[22] Roosecelis Brasil Martines. 2016. Notes from the field: evidence of Zika virus infection in brain and placental tissues from two congenitally infected newborns and two fetal losses—Brazil, 2015. *MMWR. Morbidity and mortality weekly report* 65 (2016).

[23] Uriel Martinez-Hernandez, Adrian Rubio-Solis, and Tony J Prescott. 2016. Bayesian perception of touch for control of robot emotion. In *2016 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 4927–4933.

[24] Jerry M Mendel. 2013. General type-2 fuzzy logic systems made simple: a tutorial. *IEEE Transactions on Fuzzy Systems* 22, 5 (2013), 1162–1182.

[25] Jane P Messina, Moritz UG Kraemer, Oliver J Brady, David M Pigott, Freya M Shearer, Daniel J Weiss, Nick Golding, Corrine W Ruktanonchai, Peter W Gething, Emily Cohn, et al. 2016. Mapping global environmental suitability for Zika virus. *Elife* 5 (2016), e15272.

[26] Didier Musso, Susan L Stramer, and Michael P Busch. 2016. Zika virus: a new challenge for blood transfusion. *The Lancet* 387, 10032 (2016), 1993–1994.

[27] Paul Reiter. 2001. Climate change and mosquito-borne disease. *Environmental Health Perspectives* 109, Suppl 1 (2001), 141–161.

[28] Hai-Jun Rong, Guang-Bin Huang, Narasimhan Sundararajan, and Paramasivan Saratchandran. 2009. Online sequential fuzzy extreme learning machine for function approximation and classification problems. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 39, 4 (2009), 1067–1072.

[29] Adrian Rubio-Solis, Uriel Martinez-Hernandez, and George Panoutsos. 2018. Evolutionary extreme learning machine for the interval type-2 radial basis function neural network: A fuzzy modelling approach. In *2018 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*. IEEE, 1–8.

[30] Adrian Rubio-Solis, Patricia Melin, Uriel Martinez-Hernandez, and George Panoutsos. 2019. General Type-2 Radial Basis Function Neural Network: A Data-Driven Fuzzy Model. *IEEE Transactions on Fuzzy Systems* 27, 2 (2019), 333–347.

[31] Adrian Rubio-Solis and George Panoutsos. 2015. Interval type-2 radial basis function neural network: a modeling framework. *IEEE Transactions on Fuzzy Systems* 23, 2 (2015), 457–473.

[32] Adrian Rubio-Solis and George Panoutsos. 2017. An ensemble data-driven fuzzy network for laser welding quality prediction. In *2017 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*. IEEE, 1–6.

[33] Hercules Sakkas, Petros Bozidis, Xenofon Giannakopoulos, Nikolaos Sofikitis, and Chrissanthy Papadopoulou. 2018. An update on sexual transmission of Zika Virus. *Pathogens* 7, 3 (2018), 66.

[34] Emilio Soria-Olivas, Juan Gomez-Sanchis, José D Martin, Joan Vila-Frances, Marcelino Martinez, José R Magdalena, and Antonio J Serrano. 2011. BELM: Bayesian extreme learning machine. *IEEE Transactions on Neural Networks* 22, 3 (2011), 505–509.

[35] Dongrui Wu and Jerry M Mendel. 2008. Enhanced karnik–mendel algorithms. *IEEE transactions on fuzzy systems* 17, 4 (2008), 923–934.