

Forecasting the abundance of disease vectors with deep learning

Ana Ceia-Hasse^{a,b,*}, Carla A. Sousa^a, Bruna R. Gouveia^{c,d}, César Capinha^b

^a Global Health and Tropical Medicine, Institute of Hygiene and Tropical Medicine, NOVA University of Lisbon, Rua da Junqueira 100, 1349-008 Lisbon, Portugal

^b Centre of Geographical Studies, Institute of Geography and Spatial Planning, University of Lisbon, and Associate Laboratory TERRA, Rua Branca Edmée Marques, 1600-276 Lisbon, Portugal

^c Regional Health Direction, Government of Madeira, Rua 31 de Janeiro 54-55, 9054-511 Funchal, Portugal

^d Interactive Technologies Institute - LARSyS, Caminho da Penteada, 9020-105 Funchal, Portugal

ARTICLE INFO

Keywords:

Machine learning
Mosquito
Dengue
Forecast
Time series classification

ABSTRACT

Arboviral diseases such as dengue, Zika, chikungunya or yellow fever are a worldwide concern. The abundance of vector species plays a key role in the emergence of outbreaks of these diseases, so forecasting these numbers is fundamental in preventive risk assessment. Here we describe and demonstrate a novel approach that uses state-of-the-art deep learning algorithms to forecast disease vector abundances. Unlike classical statistical and machine learning methods, deep learning models use time series data directly as predictors and identify the features that are most relevant from a predictive perspective. We demonstrate for the first time the application of this approach to predict short-term temporal trends in the number of *Aedes aegypti* mosquito eggs across Madeira Island for the period 2013 to 2019. Specifically, we apply the deep learning models to predict whether, in the following week, the number of *Ae. aegypti* eggs will remain unchanged, or whether it will increase or decrease, considering different percentages of change. We obtained high predictive performance for all years considered (mean AUC = 0.92 ± 0.05 SD). Our approach performed better than classical machine learning methods. We also found that the preceding numbers of eggs is a highly informative predictor of future trends. Linking our approach to disease transmission or importation models will contribute to operational, early warning systems of arboviral disease risk.

1. Introduction

Arboviral diseases such as dengue, Zika, chikungunya or yellow fever are a worldwide concern, with large health, social and economic costs. More than half of the world's population is at risk of disease and economic costs are large, with the costs of dengue alone estimated at US\$ 9 billion/year (Mayer et al., 2017; Messina et al., 2019; Shepard et al., 2016). Such diseases are transmitted by mosquitoes of the genus *Aedes* whose distribution is expanding - and projected to expand further in the future - due to climate change and increased transportation and human mobility (Kraemer et al., 2019; Salami et al., 2020b; Santos et al., 2022). These species are now established in many regions of the world, including in several areas of Europe (Capinha et al., 2014; Oliveira et al., 2021), where outbreaks of *Aedes*-borne diseases occurred recently (Brady and Hay, 2019; Sousa et al., 2012). There are no widespread or effective vaccines or treatment for most of these diseases yet, and the

spread of *Aedes* spp. calls for integrated tools to predict their presence and number in space and time, to inform timely control initiatives.

Accurate forecasts of *Aedes* spp. abundances are critical to inform public health decision making and implement preventive measures. Since the abundance of vector species plays a key role in disease outbreak emergence (e.g., Jupille et al., 2016; Li et al., 2019; Messina et al., 2019), forecasting such numbers is essential for preventive risk assessment. Most forecasts of mosquito abundances are performed with “classical” statistical and machine learning models, such as generalized linear models *sensu lato* (e.g., generalized additive models, generalized linear mixed models). These models use tabular-type data (i.e., a one-to-one relationship between values of dependent and predictor variables), which implies that the temporal variation in the predictors must be summarized prior to the modelling. This is often performed by applying generic statistical functions (e.g., means or sums) across fixed temporal windows (e.g., weeks, months, years) (e.g., Chen et al., 2019; Cheng

* Corresponding author at: Global Health and Tropical Medicine, Institute of Hygiene and Tropical Medicine, NOVA University of Lisbon, Rua da Junqueira 100, 1349-008 Lisbon, Portugal.

E-mail addresses: anaceiahasse@edu.ulisboa.pt (A. Ceia-Hasse), casousa@ihmt.unl.pt (C.A. Sousa), bruna.gouveia@madeira.gov.pt (B.R. Gouveia), cesarcapinha@campus.ul.pt (C. Capinha).

<https://doi.org/10.1016/j.ecolinf.2023.102272>

Received 18 October 2022; Received in revised form 14 July 2023; Accepted 18 August 2023

Available online 20 August 2023

1574-9541/© 2023 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

et al., 2020; da Cruz Ferreira et al., 2017; Li et al., 2019; Monaghan et al., 2019; Poh et al., 2019; Ripoche et al., 2019). Despite its widespread use, this approach is likely sub-optimal in a high number of cases as there is no guarantee that the resulting set of predictors will be the most adequate from a predictive perspective. This sub-optimality concerns both the statistical functions used to summarize the data and the time scales represented. Recent studies have identified how even simple changes in these parameters can have a strong impact in the quality of the predictions (e.g., Poh et al., 2019). Autoregressive Integrated Moving Average / Seasonal Autoregressive Integrated Moving Average (ARIMA/SARIMA) models are also used for time-series forecasts (e.g., Balti et al., 2020; Banerjee et al., 2020), including for predicting mosquito abundance and analysing its relationship with environmental drivers (e.g., Roiz et al., 2014; Torina et al., 2023), although they are not necessarily superior in predictive performance (e.g., Appice et al., 2020).

Here we describe and demonstrate the application of a deep learning-based approach, as an alternative to the classical approaches (e.g., Chen et al., 2019; Cheng et al., 2020; da Cruz Ferreira et al., 2017; Li et al., 2019; Monaghan et al., 2019; Poh et al., 2019; Ripoche et al., 2019) for predicting disease vector numbers. While the use of deep learning in ecology is increasing, most studies are for image and sound data classification (e.g., Brodrick et al., 2019; Campos-Taberner et al., 2020; Christin et al., 2019; Mac Aodha et al., 2018). To our knowledge, this is the first time a deep learning-based approach is used for modelling disease vector abundances.

Deep learning models are complex artificial neural networks (ANN; Lek and Guégan, 1999; Olden et al., 2008) composed of a large number of trainable parameters (LeCun et al., 2015), which are able to perform complex tasks with high performance, such as computer vision and natural language processing but also image and sound data classification (Christin et al., 2019; LeCun et al., 2015). In the context of disease vector abundance prediction, a key difference of deep learning models over conventional approaches is that the former allow using time series directly as predictors and the relevance of the features found in these data is evaluated by the models themselves (Capinha et al., 2021; Fawaz et al., 2019). This is in sharp contrast to prior research, where predictors were typically temporally invariant, usually corresponding to a statistical summarization of the variability observed in the time series data (e.g., Capinha, 2019; Dyderski et al., 2018; Priyadarshani et al., 2020). Importantly, the relevance of the features identified by the deep learning approach (both in terms of the type of data transformation to perform and the temporal extent considered) is guided by the capacity of the models to accurately predict the dependent variable. In other words, deep learning models specifically aim at identifying the set of time series features that most accurately predict disease vector abundances, as opposed to conventional models, which rely on human expertise for this purpose.

As case study, we use the yellow fever mosquito (*Aedes aegypti*) on Madeira Island, Portugal. This species was first found on this island in 2005 (Margarita et al., 2006) and from late 2012 to early 2013 it was responsible for an outbreak of dengue fever, the first ever on Madeira and the first recorded in Europe since 1927, which infected >2000 people (Lourenço and Recker, 2014; Sousa et al., 2012). Due to strong socio-economic relations with South America (namely Brazil and Venezuela), the importation of arbovirus from tropical regions has been considered highly probable (Lourenço and Recker, 2014; Seixas et al., 2019), making the occurrence of future outbreaks of this and other arboviruses possible (Lourenço and Recker, 2014). Because of this, the numbers of *Ae. aegypti* across the island have been closely monitored by local health authorities in recent years and are a key indicator of disease outbreak risk.

We describe and analyse a deep learning-based approach that provides accurate week-ahead forecasts of trends in the abundance of *Ae. aegypti* eggs on Madeira Island. We use time series of the species abundance and of weather variables as predictors and integrate these into a

standardized modelling workflow, which could be adapted to provide operational, real-time forecasts, and thus applied in early warning systems of disease risk. We also compare the predictive performance of our deep learning time series classification approach with that of classical machine learning models in predicting the variation in the number of eggs.

2. Material and methods

2.1. Mosquito abundance data

Mosquito abundance data were made available by the Regional Health Direction of the Autonomous Region of Madeira (*Direção Regional da Saúde*; <https://www.madeira.gov.pt/drs/>), the regional health authority responsible for the entomological and epidemiological surveillance in the archipelago. These data originated from 140 ovitraps placed in different locations across Madeira Island, including air- and seaports, schools, health units and other public and private places, from 2013 to 2019 (Fig. 1). This network of traps is monitored weekly, recording the number of eggs of *Ae. aegypti* in each trap.

2.2. Environmental predictors

As environmental predictors of change in the number of eggs of *Ae. aegypti*, we considered the following weather variables at a daily resolution: mean (*Tmean*), maximum (*Tmax*) and minimum (*Tmin*) temperature, mean (*RHmean*), maximum (*RHmax*) and minimum (*RHmin*) relative humidity, and accumulated precipitation (*Pre*). These variables were selected because of their known role in driving population dynamics of mosquito species (e.g., Cheng et al., 2020; da Cruz Ferreira et al., 2017; Li et al., 2019; Monaghan et al., 2019; Ripoche et al., 2019). Data for these variables was obtained from local weather stations and made available by the Portuguese Institute for Sea and Atmosphere (*Instituto Português do Mar e da Atmosfera*; www.ipma.pt).

2.3. Forecasting changes in the number of eggs of *Aedes aegypti*

We aimed at forecasting relative changes in the numbers of *Ae. aegypti* eggs instead of their absolute numbers. This a common strategy in risk assessment frameworks (e.g., Carbajo and Vezzani, 2015; Tsuda et al., 2016; Vanwambeke et al., 2011). Accordingly, for each trap and observation week, we classified the variation in the number of eggs into one of 7 classes: no change in the number of eggs; decrease between 1 and 25%; decrease between >25 and 50%; decrease >50%; increase between 1 and 25%; increase between >25 and 50%, and increase >50%.

In addition to the weather predictors (see above), we also used the preceding values of egg numbers as predictor of the response variable. To match with the daily resolution of the weather predictors, we replicated the weekly number of eggs 7 times (i.e., one per each day of the corresponding week).

All predictors represented the six-month period that preceded the week of the observation. In other words, the classes of change in number of eggs were modeled as a function of variation in the predictor variables up to 6-months into the past.

This kind of modelling exercise, where time series data is used to predict one of two or more classes, falls within the scope of time series classification (Capinha et al., 2021; Keogh and Kasetty, 2003). Several deep learning architectures can be used for time series classification, differing in the type of layers they have and on how information flows between them (Fawaz et al., 2019). Because we have no a priori reason to expect that a specific deep learning architecture will perform better than others, we evaluated a large number of models having distinct architectures and parametrizations (the so called ‘automated machine learning’ or ‘autoML’ approach; Guyon et al., 2019; Van Kuppevelt et al., 2020).

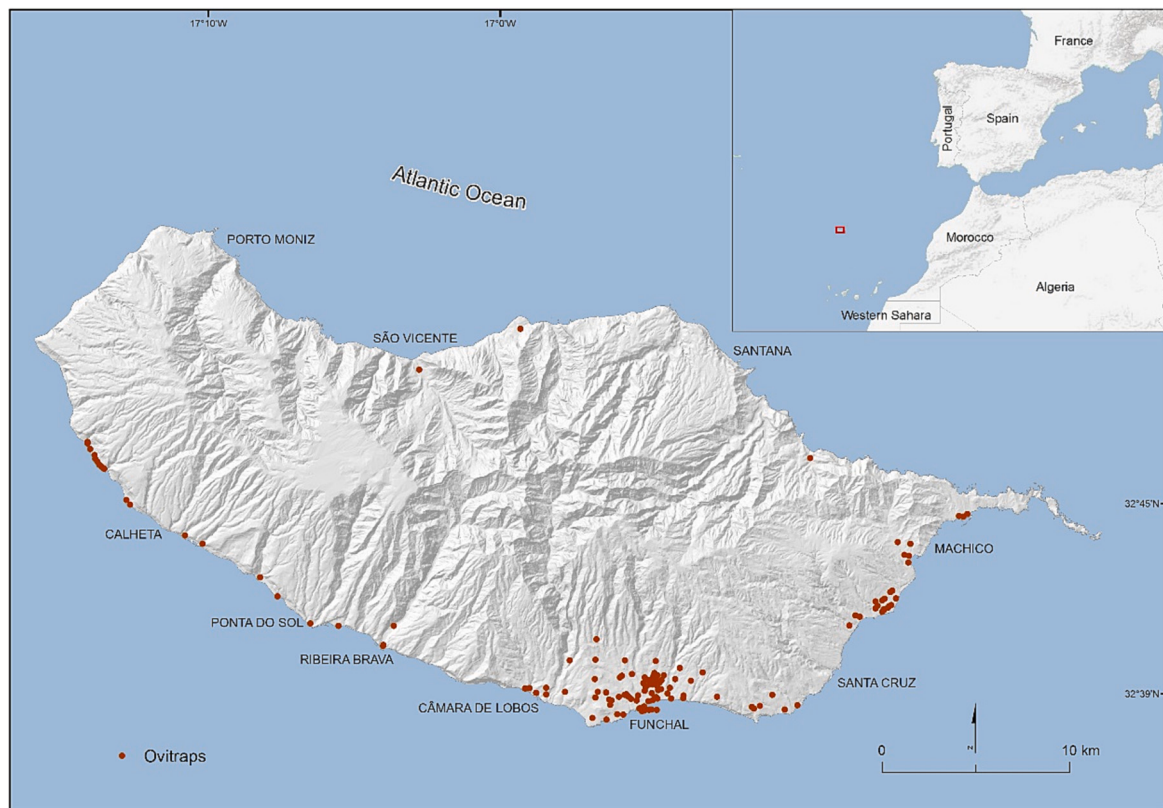


Fig. 1. Location of the ovttraps in the Madeira island.

For this purpose, we use *mcfly* (Van Kuppevelt et al., 2020), a *python* library that allows testing four deep learning architectures frequently used for time series classification: Convolutional Neural Networks (CNN), which have been mostly used for image pattern recognition (Brodrick et al., 2019; Wäldchen and Mäder, 2018) but that are also suited for time series classification (Zhao et al., 2017); Deep Convolutional Long Short-Term Memory networks (DeepConvLSTM), combining convolutional with LSTM recurrent neural networks (Chung et al., 2014), which were developed for sequence-type input data such as time series (Fawaz et al., 2019); Residual Networks (ResNet), which were proposed for image recognition (He et al., 2016) but used recently for time series classification with good performance (Fawaz et al., 2019); and Inception Time networks (InceptionTime), which are a very recent type of architecture proposed explicitly for time series classification (Fawaz et al., 2019). For more details on these modelling architectures see Capinha et al. (2021) and Van Kuppevelt et al. (2020).

2.4. Data partition

The 'autoML' workflow in *mcfly* requires using multiple partitions from the full data set of response and predictor variables. Here we followed the partition scheme of Capinha et al. (2021), consisting of four partitions that are used at the various stages of model selection and training, and a fifth partition that is used to assess the predictive performance of the fully trained ('final') model. To ensure independence between the data used for model training and the data used for model testing, we used the data of each year (i.e., 2013 to 2019) separately for model testing (partition 'T'), and the data for the remaining years for model training. The aim of this procedure is to mimic an operational setting where data available for previous years is used to train a model that will be employed in the real-time forecasting of changes in the number of eggs for the coming week.

The data used for model training and selection were randomly partitioned into: data for training candidate models (25% of the data; A_t);

data for validating candidate models (50%; A_v); data for training the selected candidate model (75%; $B_t = A_t + A_v$); validation data to determine the optimal number of epochs to train the selected candidate model (25%; B_v). Data partition was performed in R with package *dismo* (Hijmans et al., 2017; R Core Team, 2022).

2.5. Model selection procedure

The model selection procedure was performed as follows (Capinha et al., 2021; Van Kuppevelt et al., 2020): we randomly generated 5 models for each of the four available deep-ANN architecture types (20 models in total) and trained each one with a small subset of the training data (data partition A_t) for 4 epochs (an "epoch" corresponds to the complete training dataset being passed forward and backward across the network one time; Capinha et al., 2021). The accuracy of candidate models, as provided by *mcfly* (i.e., the "proportion of cases correctly classified"), was then compared using a left-out validation data set (data partition A_v) and the model with the highest performance was selected for training on the full training data (data partition B_t ; $B_t = A_t + A_v$) for up to 30 epochs.

We identified the optimal number of training epochs using data partition B_v . A too low number of epochs may result in underfitting of the model and thus in reduced predictive performance (since important patterns in the data may be missed), while an excessive number of epochs may result in overfitting of the model, which will reduce its generality and ability to classify new data. The predictive performance of the model at each training epoch each was assessed by measuring the area under the receiving operating characteristic curve (AUC), which is usually used in ecology and is not influenced by differences in the prevalence of classes (Dyderski et al., 2018). AUC was computed within R, using package *cvAUC* (LeDell et al., 2014; R Core Team, 2022).

Finally, the performance of the model trained with the optimal number of epochs was evaluated using a test data set (data set T), using AUC as the performance metric as well. This procedure was repeated so

that each year was used as test data (as partition 'T').

2.6. Importance of predictors

We assessed the importance of each predictor variable by performing simulations with randomized test data, where in each simulation one predictor variable was randomized at a time, while keeping the others unchanged (Molnar, 2020), and comparing the AUC from such simulations with the AUC from the original models. We did this for each different test year. We prepared the randomized data in R (using function *sample* from package *base*) (R Core Team, 2022).

See Appendix A for example data and code for implementing data importation and modelling procedures. For more details on the deep learning techniques we used please see Capinha et al. (2021) and Van Kuppevelt et al. (2020).

2.7. Comparison of deep learning with classical machine learning models

To assess how our deep learning time series classification approach performed in relation to classical machine learning methods, we used machine learning models to model the variation in egg numbers using the same data, i.e., the same dependent and predictor variables, the same years, and the same "leave one year out" validation method we had used in the case of the deep learning time series classification approach. For each predictor variable we averaged the six-month time series, and we used the classes of the number of eggs as the dependent variable.

Since we were dealing with a categorical dependent variable, we chose model architectures that are appropriate for classification tasks. We used random forest models, neural networks, and gradient boosting, using packages *caret*, *randomForest*, *nnet*, *deepnet* and *xgboost* (Chen et al., 2023; Kuhn, 2008; Liaw and Wiener, 2002; Rong, 2022; Venables and Ripley, 2002) within R (R Core Team, 2022). For each model we used repeated cross-validation, where one subset of the data is used as validation set, while the other is used to train the model. We perform this procedure (i.e., the data partitioning and validation procedure) 6 times. In all cases we used 75% of the data for model training and 25% for validation. We measured the predictive performance for each model and combination of years and then used the best performing model to measure predictive performance per class of variation in egg number, for each test year, by computing AUC with package *cvAUC* in R (LeDell et al., 2014; R Core Team, 2022). Our aim was to evaluate which approach (deep learning versus classical machine learning) would deliver better predictions.

3. Results

3.1. Observations per class of variation in the number of eggs

Regarding the percentage of observations per class of variation in the number of eggs between subsequent weeks, the class with more observations was "no change", in all the years (Table 1); these correspond mainly to cases where the number of eggs was continuously zero. The second and third most represented classes correspond to the largest (>50%) decrease and increase (respectively), and the classes the least

Table 1

Percentage of observations per class of change in number of eggs between subsequent weeks in each year.

	2013	2014	2015	2016	2017	2018	2019
Increase >50%	7.4	8.9	9.1	10.4	10.3	8.4	10.4
Increase 25–50%	0.8	1.0	1.4	1.2	1.6	1.0	0.9
Increase 1–25%	0.7	0.6	0.7	1.0	0.6	0.7	0.8
No change	81.7	78.0	77.6	74.7	74.9	79.4	75.3
Decrease 1–25%	0.6	0.6	0.8	0.9	0.8	0.5	0.7
Decrease 25–50%	0.3	0.7	0.7	0.4	0.5	0.3	0.5
Decrease >50%	8.5	10.1	9.8	11.5	11.4	9.6	11.4

represented correspond to the lower decreases or increases in the number of eggs (i.e., 1 to 25% and > 25 to 50%; Table 1).

3.2. Deep learning model accuracy, predictive performance, and predictor importance

Concerning the performance of the candidate models, the best validation accuracies were achieved by model 12 in all the years (accuracy = 0.833 ± 0.005 ; mean of years $\pm$ sd) (Fig. 2), a model having a Deep Convolutional Long Short-Term Memory (DeepConvLSTM) architecture.

Model 19, also with a DeepConvLSTM architecture had high validation accuracies too (accuracy = 0.829 ± 0.003 ; mean of years $\pm$ sd), followed by model 14, a Convolutional Neural Network (CNN) architecture reaching a validation accuracy of 0.821 ± 0.005 (mean of years $\pm$ sd).

After full training and identification of the optimal number of training epochs, this model delivered an excellent average predictive performance for all years (mean AUC = 0.92 ± 0.05 sd) used as testing sets. Only in a few cases the performance went below this threshold, namely for class "increase >50%" where performance was fair (i.e., AUC from 0.7 to 0.8) for years 2013 and 2016, and good (AUC 0.8 to 0.9) for the remaining years, and for class "increase 25–50%" for year 2017, where the model performance was good (AUC = 0.88) (Table 2).

Concerning the importance of variables, AUC decreased most notoriously when the predictor 'number of eggs' was randomized, a result observed for all test years (Fig. 3). Slight decreases in performance are also apparent when temperature was randomized (either mean, maximum or minimum) for most test years.

3.3. Classical machine learning model accuracy and predictive performance

Regarding the classical machine learning models, the model with the highest accuracy in all the years was the Extreme Gradient Boosting Tree (*xgbTree*), with a mean accuracy of 0.77 ± 0.0027 (mean of years $\pm$ sd). Among the other model architectures mean accuracy varied between 0.74 ± 0.0027 and 0.75 ± 0.0028 (mean of years $\pm$ sd) (Fig. 4). The parameters tested and specific choices for each algorithm are presented in Table 3.

Predictive performance was poor for all years. AUC per class of variation in the number of eggs was on average low (mean AUC = 0.57 ± 0.11 sd), with a maximum value of 0.83 for class "no change" in the number of eggs in 2013 and 2014 and for class "increase of 25–50%" in 2019, but otherwise having values of 0.5 and below 0.5 for most classes and years (Table 4).

This contrasts with the results obtained with the deep learning time series classification approach, where the validation accuracy of the best model was on average 0.833 (± 0.005) and mean AUC per class was 0.92 (± 0.05), with the lowest value being 0.78 in only 4% of the cases (class/year) (see above).

4. Discussion

Obtaining accurate predictions of disease vector distribution and trends of abundance change is critical to inform decision making and implement timely and effective control actions, in line with the United Nations Sustainable Development Goals 3 and 13 of ending vector-borne disease epidemics and strengthening early warning capacity of global health risks (UN General Assembly, 2015).

In this work we used a deep learning approach to forecast the trends in egg numbers of *Ae. aegypti* on Madeira island. To our knowledge, this is the first time that time series classification with deep learning is used for predicting variation of disease vector abundances. We also showed that the predictive accuracy of the approach was high, with the cross-class performances being good to excellent for all years considered.

We selected the candidate model with the highest validation

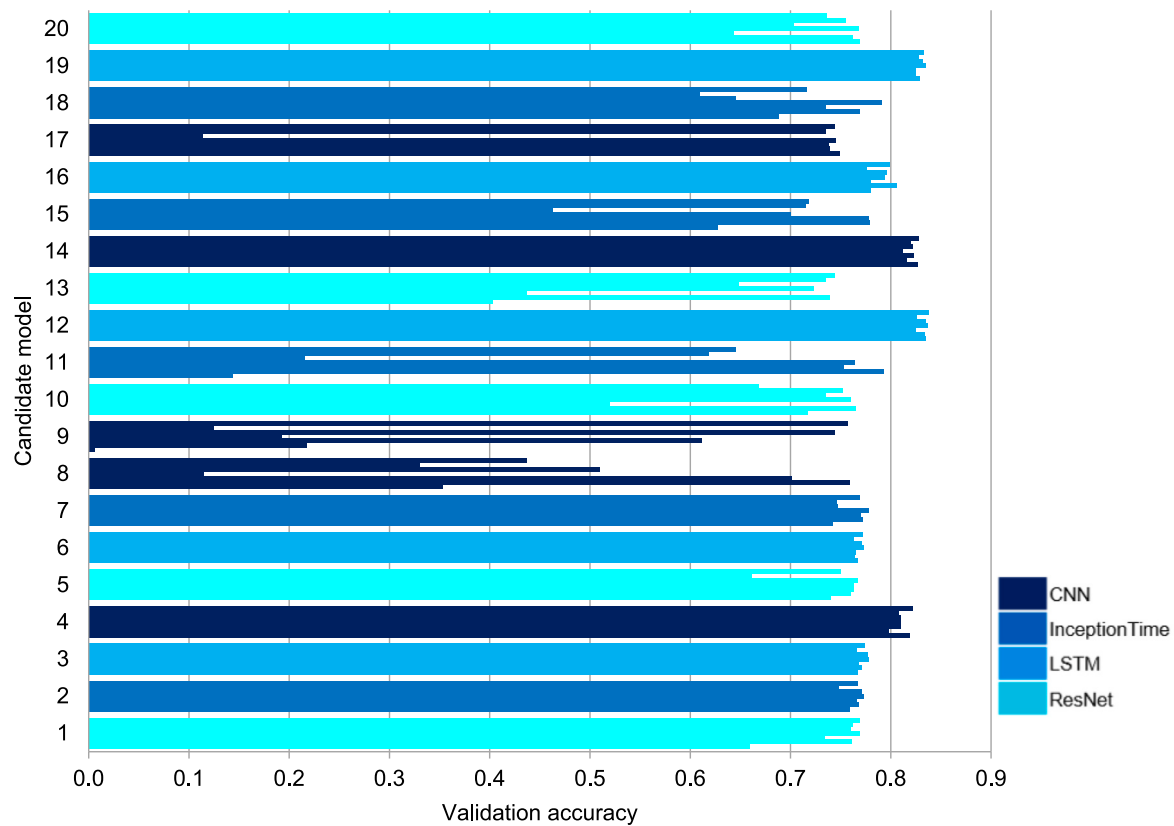


Fig. 2. Validation accuracy for candidate models. For each candidate model (1–20 on the vertical axis), each bar corresponds to one test year (from top to bottom: 2019 to 2013). The best model in all years is model 12, having a DeepConvLSTM architecture.

Table 2

Predictive performance obtained with testing sets. AUC obtained in simulations with the original models: mean (and standard deviation) of classes and AUC per class. The year used for testing model performance is given in the column headers.

		2013	2014	2015	2016	2017	2018	2019
Class	Mean	0.92	0.93	0.92	0.90	0.91	0.91	0.92
	(sd)	(0.06)	(0.04)	(0.04)	(0.06)	(0.05)	(0.05)	(0.05)
	Increase	0.78	0.83	0.83	0.78	0.81	0.80	0.80
	> 50%							
	Increase	0.93	0.93	0.93	0.90	0.88	0.92	0.94
	25–50%							
	Increase	0.95	0.94	0.93	0.91	0.92	0.92	0.94
	1–25%							
	No	0.92	0.94	0.94	0.91	0.93	0.92	0.92
	Change							
	Decrease	0.95	0.96	0.94	0.94	0.94	0.93	0.95
	1–25%							
Decrease	25–50%	0.96	0.94	0.94	0.93	0.95	0.96	0.93
	>50%	0.94	0.95	0.96	0.94	0.93	0.94	0.93

accuracy in all years as the final model to train with the full training data set and obtain the predictions. However, other models also achieved high validation accuracies. Future work could include testing whether making predictions with an ensemble of such models would increase the robustness of predictions. However, for the purpose of demonstrating our approach we considered appropriate to use only the model with the highest validation accuracy.

Validation was conducted by assessing the performance of the models using seven different data sets and time periods independently, providing independent validation and demonstrating the advantages of the approach, given the high predictive accuracy obtained. We provide the code and data files we used as appendix for those who wish to replicate our study.

We also identified that the number of eggs in the previous weeks was the most important predictor of variation in number of eggs, in agreement with previous studies (da Cruz Ferreira et al., 2017), and which suggests this variable has a high intrinsic predictability (Pennekamp et al., 2019). The relative importance of individual predictors representing other factors (temperature, relative humidity, and precipitation) was lower, despite an apparent relevance being also detected for temperature-related time series. *Tmax* was the most important temperature variable, while *RHmax* was more important than *RHmean* or *RHmin* in 42% of the years. We note that in the case of temperature and relative humidity, this apparently lower relevance may also result from correlations between the three measurements used to represent these factors (i.e., minimum, mean, and maximum values). In other words, the

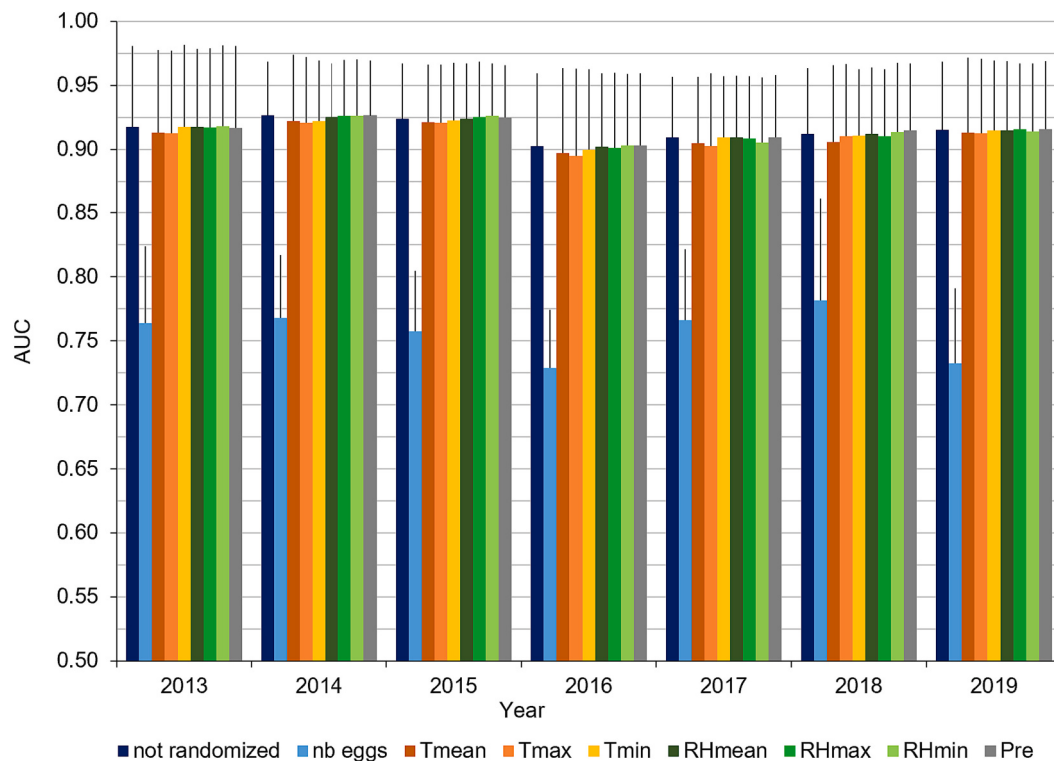


Fig. 3. AUC (mean of classes and standard deviation) for simulations performed with original (not randomized) and randomized predictor variables (e.g., the columns “nb eggs” correspond to the simulations performed with randomized number of eggs, the columns “Tmean” correspond to the simulations performed with randomized mean temperature, and so on), for each year used as test data (indicated in the horizontal axis). *Tmean*, *Tmax*, *Tmin*: mean, maximum and minimum temperature (respectively), *RHmean*, *RHmax*, *RHmin*: mean, maximum and minimum relative humidity (respectively), *Pre*: precipitation.

models may be obtaining most of relevant information from time series representing a different measurement of the same factor.

The approach presented here did not consider spatial structure explicitly. Given that the previous number of eggs had high explanatory power, the past number of eggs in nearby traps as predictors could also be included. We could also consider a leave one trap out method of validation, to see if the developed models are general spatially.

Our models use tabular-type data and thus the temporal variation in the predictors was discretized prior to modelling. Although not the objective of our work, there are statistical models that can estimate lags between time series and disease outcomes, allowing the exploration of temporal lags without specifying lags a priori (e.g., Davis et al., 2018; Smith et al., 2020). Longer time series could also be considered as this has been reported for infectious disease predictions (e.g., Smith et al., 2020).

Additionally, model forecasts were for one week ahead. In other settings, performance could be analysed across longer windows to see if and where this method stops being useful. Here, we did not test for longer windows because our objective was to mimic an operational setting using our models to produce real-time forecasts of changes in the number of eggs for the coming week, to support health authorities in decision making regarding vector control and human population health.

Deep learning is a promising avenue but there is room for improvement and limitations to be mentioned and addressed in future studies. On one hand, the approach presented can be adapted for other specific classification tasks, and model generation and selection in *mcfly* can be fine-tuned. For example, the selection of candidate models can be adjusted to generate only one type of model architecture (Capinha et al., 2021; Van Kuppevelt et al., 2020).

Also, deep learning approaches should be compared with the more classical statistical and machine learning approaches to assess how model performances differ between methods. In this sense, we complemented our study design with a comparison with classical machine

learning methods in predicting variation in the number of eggs. We observed that the deep learning time series classification approach performed better, with the classical machine learning models having poor predictive performance. This contrasts with the high performance of the deep learning time series classification approach and demonstrates the potential of the latter. We based our comparison between the deep learning and the classical machine learning models solely on the predictive performance of the models, which is one of the most important and commonly assessed criteria, especially in the context of early warning systems (e.g., Monaghan et al., 2019; Poh et al., 2019; Ripoché et al., 2019). Nevertheless, several other factors may also be considered in the future to provide a more complete picture of their merits and drawbacks when comparing the approaches, including response times, model complexity, construction complexity, iteration complexity, or complexity of incorporating new data.

Moreover, it would be interesting to gain additional insights into the predictors used in the study beyond their relative importance, for example by assessing the shapes of relationships identified by the best-performing model and comparing them with those from the classical machine learning models and to see what parts of the relationships those techniques missed such that their predictive accuracy was lower. The features uncovered from deep learning models likely differ substantially from those identified by conventional approaches relying on temporally invariant pre-processed predictors. Deep learning models learn discriminative features directly from high dimensional, “raw” predictor data, and thus the identification and interpretation of the features they learn represents a major opportunity to obtain new understanding about the systems being modeled. However, the ecological interpretability of deep learning models and their outputs has yet to be improved (Reichstein et al., 2019). This is being now tackled by the field of “explainable artificial intelligence” to understand how the models work and explain their outcomes (Shickel and Rashidi, 2020; Siddiqui et al., 2019).

When drawing parallels with other studies, we obtained a high

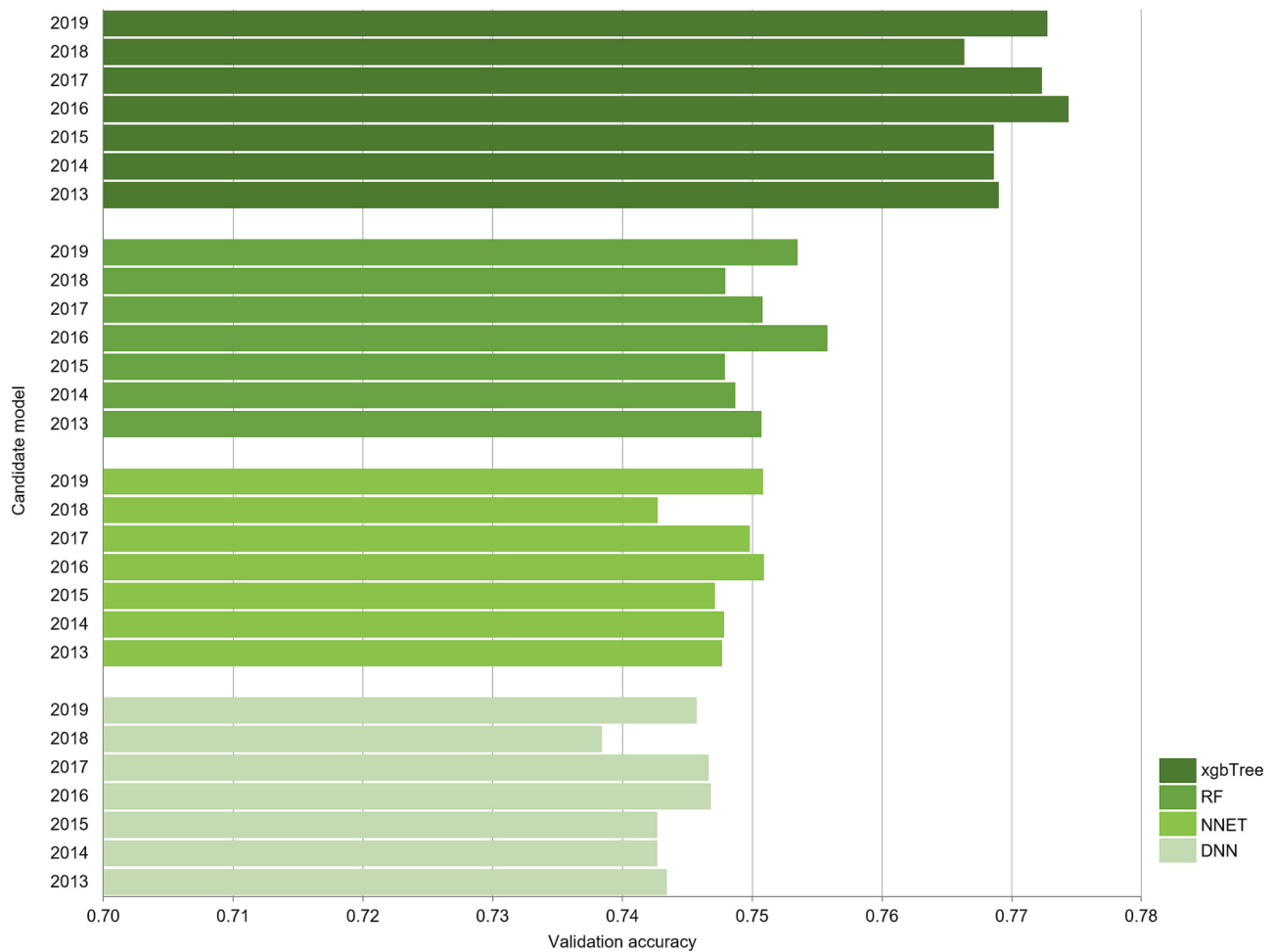


Fig. 4. Validation accuracy for classical machine learning models. For each model architecture (represented in different shades of green; *xgbTree* = extreme gradient boosting tree, *RF* = random forest, *NNET* = neural network, *DNN* = deep neural network), each bar corresponds to one test year as indicated in the vertical axis. The best model in all years is an extreme gradient boosting tree model. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 3

Parameters tested and values for each classical machine learning model used and each test year. *xgbTree* = extreme gradient boosting tree, *RF* = random forest, *NNET* = neural network, *DNN* = deep neural network.

Model	Parameter	2013	2014	2015	2016	2017	2018	2019
xgbTree	Nrounds	150	150	150	150	150	150	150
	max_depth	3	3	3	3	3	3	3
	Eta	0.4	0.4	0.4	0.4	0.4	0.4	0.3
	Gamma	0	0	0	0	0	0	0
	colsample_bytree	0.8	0.6	0.6	0.6	0.8	0.6	0.6
	min_child_weight	1	1	1	1	1	1	1
RF	Subsample	1	1	1	1	1	1	1
	Mtry	2	2	2	2	2	2	2
	Size	5	5	5	5	3	5	5
NNET	Decay	0.1	0.1	0.1	0.1	0.1	0.1	0.1
DNN	layer1	1	1	1	1	1	1	1
	layer2	0	0	0	0	0	0	0
	layer3	0	0	0	0	0	0	0
	hidden_dropout	0	0	0	0	0	0	0
	visible_dropout	0	0	0	0	0	0	0

performance (AUC = 0.92) in predicting variation in the number of *Ae. aegypti* eggs with deep learning time series classification. This result compares well with a recent study where two of us participated (Capi-nha et al., 2021) and where we used a similar modelling approach, obtaining high performances in all the case studies included, from insect species identification (AUC = 0.88–1) to species distribution modelling

(AUC = 0.95) and forecasting mushroom phenology (AUC = 0.91).

Other studies using classical machine learning methods to predict mosquito disease vector abundance obtained as low as 63% predictive ability (Monaghan et al., 2019) but highlighted the importance of weather (namely temperature and precipitation) (e.g., Li et al., 2019; Ripoche et al., 2019) and weather variation (Poh et al., 2019).

Table 4
Predictive performance obtained with the *xgbTree* model: mean AUC (and standard deviation) of classes and AUC per class. The year used for testing model performance is given in the column headers.

		2013	2014	2015	2016	2017	2018	2019
Class	Mean	0.58	0.59	0.49	0.56	0.57	0.58	0.60
	(sd)	(0.13)	(0.13)	(0.04)	(0.10)	(0.11)	(0.12)	(0.13)
	Increase >50%	0.58	0.62	0.58	0.55	0.54	0.59	0.53
	Increase 25–50%	0.49	0.49	0.49	0.49	0.49	0.49	0.83
	Increase 1–25%	0.49	0.50	0.50	0.49	0.50	0.50	0.50
	No Change	0.83	0.83	0.43	0.76	0.79	0.79	0.72
	Decrease 1–25%	0.50	0.50	0.50	0.50	0.50	0.50	0.50
	Decrease 25–50%	0.50	0.50	0.50	0.50	0.50	0.50	0.50
	Decrease >50%	0.66	0.69	0.47	0.65	0.64	0.66	0.64

Importantly, da Cruz Ferreira et al. (2017) also found that including previous mosquito abundance as a predictor greatly improved model performance in predicting infestation trends of *Ae. aegypti* in Brazil, in line with our finding that the number of eggs in the preceding weeks was the most important predictor of the variation in the number of *Ae. aegypti* eggs.

There are several ways in which uncertainty can be addressed, some that we employed here and some that could be included in future studies, which we discussed above and now summarize here: including spatial structure and other predictor variables than may influence oviposition and egg survival and development; exploring different temporal lags; using longer time series; analysing performance across longer forecast windows; using model ensembles to produce the forecasts; comparing performance and computation time with other modelling approaches; increasing the ecological explainability of the models.

5. Conclusions and broader implications of our work

Our results are supportive of a wider testing and application of deep learning for predicting disease vector abundances. Other applications of this approach could include, for example, species distribution modelling, forecasting algal blooms, the phenology and abundance of crop pests, or the phenology of species with commercial interest such as edible fruits and mushrooms (e.g., Capinha et al., 2021).

A defining feature of this approach is that it allows using raw time series data as predictors, being able to automatically identify relevant information such as thresholds and lag effects (Fawaz et al., 2019; Ryo et al., 2019) that could otherwise be missed when using classical approaches, which rely on temporally aggregated (and thus simplified) variables (Capinha et al., 2021). This capacity of deep learning models can be an important asset in face of the growing availability of time series data, namely from mosquito monitoring programs and environmental sensors (e.g., satellite imagery and meteorological stations; Reichstein et al., 2019), that can be directly fed into the models (Capinha et al., 2021).

Despite the challenges mentioned, we presented a well performing modelling framework that can be used by non-experts in deep learning (Van Kuppevelt et al., 2020) to predict disease vector abundances. Our work has a clear public health relevance. Integrating this approach into existing disease transmission or importation models (Lieberthal and Gardner, 2021; Salami et al., 2020a; Salami et al., 2020b) or with disease-related Internet search activity (Aiken et al., 2020; Yang et al., 2017) and using high resolution environmental data will likely contribute to improve operational, early warning systems of disease risk and guide the implementation of mosquito control measures.

Data availability

Data and code are available as supplementary material to the article

Acknowledgements

We thank the personnel of the Regional Health Direction of the Autonomous Region of Madeira (*Direção Regional da Saúde*) involved in egg collection and analysis and the Portuguese Institute for Sea and Atmosphere (*Instituto Português do Mar e da Atmosfera*), namely Dr. Victor Prior, for providing the weather data. ACH, CAS and CC were supported by Portuguese National Funds through Fundação para a Ciência e a Tecnologia (ACH and CAS: PTDC/SAU-PUB/30089/2017 and GHTM-UID/Multi/04413/2013; CC: CEECIND/02037/2017, UIDB/00295/2020 and UIDP/00295/2020). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ecoinf.2023.102272>.

References

Aiken, E.L., McGough, S.F., Majumder, M.S., Wachtel, G., Nguyen, A.T., Viboud, C., Santillana, M., 2020. Real-time estimation of disease activity in emerging outbreaks using internet search information. *PLoS Comput. Biol.* 16, 1–19. <https://doi.org/10.1371/journal.pcbi.1008117>.

Appice, A., Gel, Y.R., Iliev, I., Lyubchich, V., Malerba, D., 2020. A multi-stage machine learning approach to predict dengue incidence: a case study in Mexico. *IEEE Access* 8, 52713–52725. <https://doi.org/10.1109/ACCESS.2020.2980634>.

Balti, H., Ben Abbes, A., Mellouli, N., Farah, I.R., Sang, Y., Lamolle, M., 2020. A review of drought monitoring with big data: issues, methods, challenges and research directions. *Ecolog. Informat.* 60, 101136 <https://doi.org/10.1016/j.ecoinf.2020.101136>.

Banerjee, A., Chakrabarty, M., Bandyopadhyay, G., Roy, P.K., Ray, S., 2020. Forecasting environmental factors and zooplankton of Bakreswar reservoir in India using time series model. *Ecolog. Informat.* 60, 101157 <https://doi.org/10.1016/j.ecoinf.2020.101157>.

Brady, O.J., Hay, S.I., 2019. The first local cases of Zika virus in Europe. *Lancet* 394, 1991–1992. [https://doi.org/10.1016/S0140-6736\(19\)32790-4](https://doi.org/10.1016/S0140-6736(19)32790-4).

Brodrick, P.G., Davies, A.B., Asner, G.P., 2019. Uncovering ecological patterns with convolutional neural networks. *Trends Ecol. Evol.* 34, 734–745. <https://doi.org/10.1016/j.tree.2019.03.006>.

Campos-Taberner, M., García-Haro, F.J., Martínez, B., Izquierdo-Verdiguier, E., Atzberger, C., Camps-Valls, G., Gilabert, M.A., 2020. Understanding deep learning in land use classification based on Sentinel-2 time series. *Sci. Rep.* 10, 17188. <https://doi.org/10.1038/s41598-020-74215-5>.

Capinha, C., 2019. Predicting the timing of ecological phenomena using dates of species occurrence records: a methodological approach and test case with mushrooms. *Int. J. Biometeorol.* 63, 1015–1024. <https://doi.org/10.1007/s00484-019-01714-0>.

- Capinha, C., Rocha, J., Sousa, C.A., 2014. Macroclimate determines the global range limit of *Aedes aegypti*. *EcoHealth* 11, 420–428. <https://doi.org/10.1007/s10393-014-0918-y>.
- Capinha, C., Ceia-Hasse, A., Kramer, A.M., Meijer, C., 2021. Deep learning for supervised classification of temporal data in ecology. *Ecolog. Informat.* 61, 101252 <https://doi.org/10.1016/j.ecoinf.2021.101252>.
- Carbajo, A.E., Vezzani, D., 2015. Waiting for chikungunya fever in Argentina: spatio-temporal risk maps. *Mem. Inst. Oswaldo Cruz* 110, 259–262.
- Chen, S., Whiteman, A., Li, A., Rapp, T., Delmelle, E., Chen, G., Brown, C.L., Robinson, P., Coffman, M.J., Janies, D., Dulin, M., 2019. An operational machine learning approach to predict mosquito abundance based on socioeconomic and landscape patterns. *Landsc. Ecol.* 34, 1295–1311. <https://doi.org/10.1007/s10980-019-00839-2>.
- Cheng, J., Bambrick, H., Yakob, L., Devine, G., Frentiu, F.D., Toan, D.T.T., Thai, P.Q., Xu, Z., Hu, W., 2020. Heatwaves and dengue outbreaks in Hanoi, Vietnam: new evidence on early warning. *PLoS Negl. Trop. Dis.* 14, 1–15. <https://doi.org/10.1371/journal.pntd.0007997>.
- Chen, T., He, T., Benesty, M., Khotilovich, V., Tang, Y., Cho, H., Chen, K., Mitchell, R., Cano, I., Zhou, T., Li, M., Xie, J., Lin, M., Geng, Y., Li, Y., 2023. xgboost: Extreme Gradient Boosting. R package version 1.7.5.1. Yuan J. <https://CRAN.R-project.org/package=xgboost>.
- Christin, S., Hervet, É., Lecomte, N., 2019. Applications for deep learning in ecology. *Methods Ecol. Evol.* 10, 1632–1644. <https://doi.org/10.1111/2041-210X.13256>.
- Chung, J., Gulchere, C., Cho, K., Bengio, Y., 2014. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling. <https://doi.org/10.48550/arXiv.1412.3555>.
- da Cruz Ferreira, D.A., Degener, C.M., de Almeida Marques-Toledo, C., Bendati, M.M., Fetuz, L.O., Teixeira, C.P., Eiras, A.E., 2017. Meteorological variables and mosquito monitoring are good predictors for infestation trends of *Aedes aegypti*, the vector of dengue, chikungunya and Zika. *Parasit. Vectors* 10, 78. <https://doi.org/10.1186/s13071-017-2025-8>.
- Davis, J.K., Vincent, G.P., Hildreth, M.B., Kightlinger, L., Carlson, C., Wimberly, M.C., 2018. Improving the prediction of arbovirus outbreaks: a comparison of climate-driven models for West Nile virus in an endemic region of the United States. *Acta Trop.* 185, 242–250. <https://doi.org/10.1016/j.actatropica.2018.04.028>.
- Dyderski, M.K., Paž, S., Frelich, L.E., Jagodziński, A.M., 2018. How much does climate change threaten European forest tree species distributions? *Glob. Chang. Biol.* 24, 1150–1163. <https://doi.org/10.1111/gcb.13925>.
- Fawaz, H.I., Forestier, G., Weber, J., Idoumghar, L., Muller, P.-A., 2019. Deep learning for time series classification: a review. *Data Min. Knowl. Disc.* 33, 917–963. <https://doi.org/10.1007/s10618-019-00619-1>.
- Guyon, I., Sun-Hosoya, L., Boullé, M., Escalante, H.J., Escalera, S., Liu, Z., Jajetic, D., Ray, B., Saeed, M., Sebag, M., Statnikov, A., Tu, W.-W., Viegas, E., 2019. Analysis of the AutoML challenge series 2015–2018. In: Hutter, F., Kotthoff, L., Vanschoren, J. (Eds.), *Automated Machine Learning: Methods, Systems, Challenges*. Springer International Publishing, Cham, pp. 177–219. https://doi.org/10.1007/978-3-030-05318-5_10.
- He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770–778. <https://doi.org/10.1109/CVPR.2016.90>.
- Hijmans, R.J., Phillips, S., Leathwick, J., Elith, J., 2017. Dismo: species distribution modeling. In: R package version 1, 1–1.
- Jupille, H., Seixas, G., Mousson, L., Sousa, C.A., Failloux, A.-B., 2016. Zika virus, a new threat for Europe? *PLoS Negl. Trop. Dis.* 10, 1–8. <https://doi.org/10.1371/journal.pntd.0004901>.
- Keogh, E., Kasetty, S., 2003. On the need for time series data mining benchmarks: a survey and empirical demonstration. *Data Min. Knowl. Disc.* 7, 349–371. <https://doi.org/10.1023/A:1024988512476>.
- Kraemer, M.U.G., Reiner, R.C., Brady, O.J., Messina, J.P., Gilbert, M., Pigott, D.M., Yi, D., Johnson, K., Earl, L., Marczak, L.B., Shirude, S., Davis Weaver, N., Bisanzio, D., Perkins, T.A., Lai, S., Lu, X., Jones, P., Coelho, G.E., Carvalho, R.G., Van Bortel, W., Marsboom, C., Hendrickx, G., Schaffner, F., Moore, C.G., Nax, H.H., Bengtsson, L., Wetter, E., Tatem, A.J., Brownstein, J.S., Smith, D.L., Lambrechts, L., Cauchemez, S., Linard, C., Faria, N.R., Pybus, O.G., Scott, T.W., Liu, Q., Yu, H., Wint, G.R.W., Hay, S. I., Golding, N., 2019. Past and future spread of the arbovirus vectors *Aedes aegypti* and *Aedes albopictus*. *Nat. Microbiol.* 4, 854–863. <https://doi.org/10.1038/s41564-019-0376-y>.
- Kuhn, M., 2008. Building Predictive Models in R Using the caret Package. *Journal of Statistical Software* 28 (5), 1–26. <https://doi.org/10.18637/jss.v028.i05>.
- LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. *Nature* 521, 436–444. <https://doi.org/10.1038/nature14539>.
- LeDell, E., Petersen, M., van der Laan, M., 2014. cvAUC: Cross-Validated Area under the ROC Curve Confidence Intervals. R Package Version 1.
- Lek, S., Guégan, J.F., 1999. Artificial neural networks as a tool in ecological modelling, an introduction. *Ecol. Model.* 120, 65–73. [https://doi.org/10.1016/S0304-3800\(99\)00092-7](https://doi.org/10.1016/S0304-3800(99)00092-7).
- Li, R., Xu, L., Bjørnstad, O.N., Liu, K., Song, T., Chen, A., Xu, B., Liu, Q., Stenseth, N.C., 2019. Climate-driven variation in mosquito density predicts the spatiotemporal dynamics of dengue. *Proc. Natl. Acad. Sci.* 116, 3624–3629. <https://doi.org/10.1073/pnas.1806094116>.
- Liaw, A., Wiener, M., 2002. Classification and Regression by randomForest. *R News* 2 (3), 18–22.
- Lieberthal, B., Gardner, A.M., 2021. Connectivity, reproduction number, and mobility interact to determine communities' epidemiological superspreader potential in a metapopulation network. *PLoS Comput. Biol.* 17, 1–22. <https://doi.org/10.1371/journal.pcbi.1008674>.
- Lourenço, J., Recker, M., 2014. The 2012 Madeira dengue outbreak: epidemiological determinants and future epidemic potential. *PLoS Negl. Trop. Dis.* 8, 1–11. <https://doi.org/10.1371/journal.pntd.0003083>.
- Mac Aodha, O., Gibb, R., Barlow, K.E., Browning, E., Firman, M., Freeman, R., Harder, B., Kinsey, L., Mead, G.R., Newson, S.E., Pandourski, I., Parsons, S., Russ, J., Szodoray-Paradi, A., Szodoray-Paradi, F., Tilova, E., Girolami, M., Brostow, G., Jones, K.E., 2018. Bat detective - deep learning tools for bat acoustic signal detection. *PLoS Comput. Biol.* 14, e1005995 <https://doi.org/10.1371/journal.pcbi.1005995>.
- Margarita, Y., Grácio, A., Lencastre, I., Silva, A., Novo, M., Sousa, C., 2006. First record of *Aedes (Stegomyia) aegypti* (Linnaeus, 1762) (Diptera, Culicidae) in Madeira Island-Portugal. *Acta Parasitológica Portuguesa* 13, 59–61.
- Mayer, S.V., Tesh, R.B., Vasilakis, N., 2017. The emergence of arthropod-borne viral diseases: a global perspective on dengue, chikungunya and Zika fevers. *Acta Trop.* 166, 155–163. <https://doi.org/10.1016/j.actatropica.2016.11.020>.
- Messina, J.P., Brady, O.J., Golding, N., Kraemer, M.U.G., Wint, G.R.W., Ray, S.E., Pigott, D.M., Shearer, F.M., Johnson, K., Earl, L., Marczak, L.B., Shirude, S., Davis Weaver, N., Gilbert, M., Velayudhan, R., Jones, P., Jaenisch, T., Scott, T.W., Reiner, R.C., Hay, S.I., 2019. The current and future global distribution and population at risk of dengue. *Nat. Microbiol.* 4, 1508–1515. <https://doi.org/10.1038/s41564-019-0476-8>.
- Molnar, C., 2020. Interpretable Machine Learning (Lulu. com).
- Monaghan, A.J., Schmidt, C.A., Hayden, M.H., Smith, K.A., Reiskind, M.H., Cabell, R., Ernst, K.C., 2019. A simple model to predict the potential abundance of *Aedes aegypti* mosquitoes one month in advance. *Am. J. Trop. Med. Hyg.* 100, 434–437. <https://doi.org/10.4269/ajtmh.17-0860>.
- Olden, J.D., Lawler, J.J., Poff, N.L., 2008. Machine learning methods without tears: a primer for ecologists. *Q. Rev. Biol.* 83, 171–193. <https://doi.org/10.1086/587826>.
- Oliveira, S., Rocha, J., Sousa, C.A., Capinha, C., 2021. Wide and increasing suitability for *Aedes albopictus* in Europe is congruent across distribution models. *Sci. Rep.* 11, 9916. <https://doi.org/10.1038/s41598-021-89096-5>.
- Pennekamp, F., Iles, A.C., Garland, J., Brennan, C., Brose, U., Gaedke, U., Jacob, U., Kratina, P., Matthews, B., Munch, S., Novak, M., Palamara, G.M., Rall, B.C., Rosenbaum, B., Tabi, A., Ward, C., Williams, R., Ye, H., Petchey, O.L., 2019. The intrinsic predictability of ecological time series and its potential to guide forecasting. *Ecol. Monogr.* 89, e01359 <https://doi.org/10.1002/ecm.1359>.
- Poh, K.C., Chaves, L.F., Reyna-Nava, M., Roberts, C.M., Fredregill, C., Bueno, R., Debboun, M., Hamer, G.L., 2019. The influence of weather and weather variability on mosquito abundance and infection with West Nile virus in Harris County, Texas, USA. *Sci. Total Environ.* 675, 260–272. <https://doi.org/10.1016/j.scitotenv.2019.04.109>.
- Priyadarshani, N., Marsland, S., Juodakis, J., Castro, I., Listanti, V., 2020. Wavelet filter for automated recognition of birdsong in long-time field recordings. *Methods Ecol. Evol.* 11, 403–417. <https://doi.org/10.1111/2041-210X.13357>.
- R Core Team, 2022. R Core Team: A Language and Environment for Statistical Computing R Foundation for Statistical Computing. Vienna, Austria.
- Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., Prabhat, 2019. Deep learning and process understanding for data-driven earth system science. *Nature* 566, 195–204. <https://doi.org/10.1038/s41586-019-0912-1>.
- Ripoche, M., Campagna, C., Ludwig, A., Ogden, N.H., Leighton, P.A., 2019. Short-term forecasting of daily abundance of West Nile virus vectors *Culex pipiens-restuans* (Diptera: Culicidae) and *Aedes vexans* based on weather conditions in southern Québec (Canada). *J. Med. Entomol.* 56, 859–872. <https://doi.org/10.1093/jme/tjz002>.
- Roiz, D., Ruiz, S., Soriguer, R., Figuerola, J., 2014. Climatic effects on mosquito abundance in Mediterranean wetlands. *Parasit. Vectors* 7, 333. <https://doi.org/10.1186/1756-3305-7-333>.
- Rong, X., 2022. deeptnet: Deep Learning Toolkit in R. R package version 0.2.1. <https://CRAN.R-project.org/package=deeptnet>.
- Ryo, M., Aguilar-Trigueros, C.A., Pinek, L., Muller, L.A.H., Rillig, M.C., 2019. Basic principles of temporal dynamics. *Trends Ecol. Evol.* 34, 723–733. <https://doi.org/10.1016/j.tree.2019.03.007>.
- Salami, D., Capinha, C., Sousa, C.A., Martins, M., Lord, C., 2020a. Simulation models of dengue transmission in Funchal, Madeira Island: influence of seasonality. *PLoS Negl. Trop. Dis.* 14, 1–21. <https://doi.org/10.1371/journal.pntd.0008679>.
- Salami, D., Sousa, C.A., Martins, M., Capinha, C., 2020b. Predicting dengue importation into Europe, using machine learning and model-agnostic methods. *Sci. Rep.* 10, 9689. <https://doi.org/10.1038/s41598-020-66650-1>.
- Santos, J.M., Capinha, C., Rocha, J., Sousa, C.A., 2022. The current and future distribution of the yellow fever mosquito (*Aedes aegypti*) on Madeira Island. *PLoS Negl. Trop. Dis.* 16, 1–19. <https://doi.org/10.1371/journal.pntd.0010715>.
- Seixas, G., Salgueiro, P., Bronzato-Badial, A., Gonçalves, Y., Reyes-Lugo, M., Gordicho, V., Ribolla, P., Viveiros, B., Silva, A.C., Pinto, J., Sousa, C.A., 2019. Origin and expansion of the mosquito *Aedes aegypti* in Madeira Island (Portugal). *Sci. Rep.* 9, 2241. <https://doi.org/10.1038/s41598-018-38373-x>.
- Shepard, D.S., Undurraga, E.A., Halasa, Y.A., Stanaway, J.D., 2016. The global economic burden of dengue: a systematic analysis. *Lancet Infect. Dis.* 16, 935–941. [https://doi.org/10.1016/S1473-3099\(16\)00146-8](https://doi.org/10.1016/S1473-3099(16)00146-8).
- Shickel, B., Rashidi, P., 2020. Sequential Interpretability: Methods, Applications, and Future Direction for Understanding Deep Learning Models in the Context of Sequential Data arXiv preprint arXiv:2004.12524.
- Siddiqui, S.A., Mercier, D., Munir, M., Dengel, A., Ahmed, S., 2019. TSViz: demystification of deep learning models for time-series analysis. *IEEE Access* 7, 67027–67040. <https://doi.org/10.1109/ACCESS.2019.2912823>.
- Smith, K.H., Tyre, A.J., Hamik, J., Hayes, M.J., Zhou, Y., Dai, L., 2020. Using climate to explain and predict West Nile virus risk in Nebraska. *GeoHealth* 4. <https://doi.org/10.1029/2020GH000244>.

- Sousa, C., Clairouin, M., Seixas, G., Viveiros, B., Novo, M., Silva, A., Escoval, M., Economopoulou, A., 2012. Ongoing outbreak of dengue type 1 in the autonomous region of Madeira, Portugal: preliminary report. *Eurosurveillance* 17, 20333.
- Torina, A., Russa, F.L., Blanda, V., Peralbo-Moreno, A., Casades-Martí, L., Pasquale, L.D., Bongiorno, C., Badaco, V.V., Toma, L., Ruiz-Fons, F., 2023. Modelling time-series *Aedes albopictus* abundance as a forecasting tool in urban environments. *Ecol. Indic.* 150, 110232 <https://doi.org/10.1016/j.ecolind.2023.110232>.
- Tsuda, Y., Maekawa, Y., Ogawa, K., Itokawa, K., Komagata, O., Sasaki, T., Isawa, H., Tomita, T., Sawabe, K., 2016. Biting density and distribution of *Aedes albopictus* during the September 2014 outbreak of dengue fever in Yoyogi Park and the vicinity of Tokyo Metropolis, Japan. *Jpn. J. Infect. Dis.* 69, 1–5. <https://doi.org/10.7883/yoken.JJID.2014.576>.
- UN General Assembly, 2015. Transforming our World: The 2030 Agenda for Sustainable Development (a/RES/70/1). United Nations.
- Van Kuppevelt, D., Meijer, C., Huber, F., van der Ploeg, A., Georgievskia, S., van Hees, V. T., 2020. Mcfly: automated deep learning on time series. *SoftwareX* 12, 100548. <https://doi.org/10.1016/j.softx.2020.100548>.
- Vanwambeke, S.O., Bennett, S.N., Kapan, D.D., 2011. Spatially disaggregated disease transmission risk: land cover, land use and risk of dengue transmission on the island of Oahu. *Tropical Med. Int. Health* 16, 174–185. <https://doi.org/10.1111/j.1365-3156.2010.02671.x>.
- Venables, W.N., Ripley, B.D., 2002. *Modern Applied Statistics with S, Fourth Edition*. Springer, New York. ISBN 0-387-95457-0.
- Wäldchen, J., Mäder, P., 2018. Machine learning for image based species identification. *Methods Ecol. Evol.* 9, 2216–2225. <https://doi.org/10.1111/2041-210X.13075>.
- Yang, S., Kou, S.C., Lu, F., Brownstein, J.S., Brooke, N., Santillana, M., 2017. Advances in using internet searches to track dengue. *PLoS Comput. Biol.* 13, 1–14. <https://doi.org/10.1371/journal.pcbi.1005607>.
- Zhao, B., Lu, H., Chen, S., Liu, J., Wu, D., 2017. Convolutional neural networks for time series classification. *J. Syst. Eng. Electron.* 28, 162–169. <https://doi.org/10.21629/JSEE.2017.01.18>.