

<https://doi.org/10.1038/s42003-025-09410-z>

Land and climate suitability for West Nile virus in Atlantic archipelagos guided by historical data from Europe

Check for updates

Martim A. Geraldes^{1,2}, Marta Giovanetti^{3,4,5}, Mónica V. Cunha^{1,6}✉ & José Lourenço^{1,6,7}✉

West Nile (WNV) is a zoonotic mosquito-borne virus with an expanding geographical and epidemic activity worldwide. Computational studies have contributed to the understanding of factors driving WNV occurrence, particularly in North America and Europe. Archipelagos have largely been overlooked, despite the risk to unique local avian species and human populations. In this study, we apply an ecological niche approach, trained on WNV occurrence and (a)biotic factors from European countries to project ecological suitability for WNV occurrence across several Atlantic archipelagos. The approach gives weight to the temporal dimension, generating novel insights on seasonality both for Europe and the archipelagos. For European countries, modelling results align with previous findings on spatial hotspots and (a)biotic drivers of WNV occurrence, while further unravelling properties of at-risk human populations within dynamically suitable land areas. For Atlantic archipelagos, results constitute a novel and detailed perspective on local ecological suitability for WNV occurrence, providing a data-driven framework that identifies spatial hotspots, defines seasonal patterns and quantifies the local population at risk. The synthetic data generated in this study supports the development of targeted preparedness, surveillance and mitigation plans tailored to the unique ecological and seasonal dynamics of each region under study.

West Nile (WNV) is a mosquito-borne virus of the *Orthoflavivirus* genus, *Flaviviridae* family, and part of the Japanese Encephalitis serocomplex. It is maintained in an enzootic cycle between avian and mosquito species¹. Mosquitoes of the *Culex* genus, particularly *Culex pipiens*, are recognised as the main vectors in Europe and North America, although five other mosquito genera have also been implicated in transmission². Occasionally, epizootic spillover events affect mammals, including humans and equines. First identified in Uganda in 1937, WNV has expanded its geographical transmission range over the past 25 years, after its introduction into North America in 1999, with human infections now recorded on every continent except Antarctica³. Although the majority of human infections are asymptomatic (70–80%), some progress to neuroinvasive disease, manifesting as encephalitis, meningitis or acute flaccid paralysis, with severe cases potentially leading to death⁴.

WNV reporting in the European Union is mandatory, with surveillance efforts focusing on humans, equines, birds and mosquitoes across member states. Epidemic activity in Europe has increased in the past 15 years, spanning an ever wider geographic area, with notable expansion and activity in France, Italy, Greece, Spain, Hungary, Romania, Bulgaria, Serbia and Ukraine^{5,6}. The continent's largest recorded epidemic occurred in 2018, with 2,083 cases surpassing cumulative infections since 2010⁷. In 2020, Spain experienced an unprecedented WNV epidemic marked by high numbers of human infections and deaths, including previously unaffected areas^{8,9}. The second-highest annual count of locally acquired cases was recorded in 2022, with Italy reporting its highest number to date¹⁰. In 2023 alone, seven European Union countries reported over 150 outbreaks among equids and eight reported over 250 outbreaks among birds¹⁰. In 2024, WNV was included in the WHO pathogen prioritisation list, recognising its potential to cause public health emergencies of concern across five continents¹¹.

¹Centre for Ecology, Evolution and Environmental Changes (cE3c) & Global Change and Sustainability Institute (CHANGE), Faculdade de Ciências, Universidade de Lisboa, Lisboa, Portugal. ²Universidade Católica Portuguesa, Católica Medical School, Católica Biomedical Research Center, Oeiras, Portugal. ³Department of Sciences and Technologies for Sustainable Development and One Health, Università Campus Bio-Medico di Roma, Rome, Italy. ⁴Instituto René Rachou, Fundação Oswaldo Cruz, Belo Horizonte, Minas Gerais, Brasil. ⁵Climate amplified diseases and epidemics (CLIMADE) Americas, Rio de Janeiro, Brazil. ⁶Biosystems & Integrative Sciences Institute (BioISI), Faculdade de Ciências, Universidade de Lisboa, Lisboa, Portugal. ⁷Climate amplified diseases and epidemics (CLIMADE) Europe, Lisbon, Portugal. ✉e-mail: mvcunha@ciencias.ulisboa.pt; jlourenco@ucp.pt

Historically during the 20th century, WNV-focused epidemiological studies have mostly explored the African and Middle Eastern settings^{1,12}. The key epidemiological event of WNV introduction into North America in 1999, after which it became endemic across the Americas, marked a turning point in global awareness of the virus. Since then, the range of countries contributing to WNV body of knowledge has widened, with most outputs remaining focused on the North American and European (a)biotic settings¹³, albeit with growing interest in Africa^{14–16}. Within this output, computational approaches have largely focused on large spatial scales, typically covering inland subcontinental regions (e.g. Iberian Peninsula¹⁷), entire countries (e.g., USA¹⁸, Greece¹⁹) or continents (e.g., Europe²⁰, Africa¹⁶). At such scales, a combination of mechanistic, machine learning and statistical approaches (e.g., refs. 14–24) now support the identification of spatio-temporal hotspots for WNV epidemic activity and drivers of such activity, which ultimately have the potential to inform the development of local preparedness, surveillance and mitigation plans. For example, by identifying climate change as a key driver of WNV circulation and expansion in Europe²⁰, and determining ecologically favourable regions for WNV circulation in Africa¹⁴. Climate change is a timely and relevant topic in the context of WNV. Studies are increasingly finding evidence of its role in the recent WNV geographical expansion and growing epidemic activity globally^{25–29}.

In contrast to the aforementioned large-scale regions, archipelagos and islands have generally been overlooked in computational studies, despite accumulating evidence of WNV circulation in these bioclimatic settings (e.g., Caribbean islands^{30–34}, African islands³⁵, Mediterranean islands^{36–39}). Beyond the public health risk WNV may pose to local human populations, two other critical reasons support a deeper understanding of WNV eco-epidemiology on islands. First, islands are avian biodiversity hotspots, with unique indigenous assemblages, and they serve as key stepping stones along avian inter-continental migration routes⁴⁰, potentially acting as hotspots for viral acquisition, global dispersal and genetic mixing¹⁵. Second, many insular avian species already face significant human-driven threats, such that particular concerns on the impact of WNV circulation to their conservation have been raised, following the well reported population decreases in several North American avian species after WNV introduction in 1999⁴¹.

Some archipelagos that have been neglected in field monitoring and computational studies include those in the Atlantic Ocean, such as the Azores (Portugal), Madeira (Portugal), the Canaries (Spain), Cape Verde, and São Tomé and Príncipe. These archipelagos are crucial to the long-distance, inter-continental migration routes of avian species between the Arctic and Antarctic regions⁴². The history and current state of mosquito-borne viruses on these islands, specifically WNV and its mosquito-vectors, are generally incomplete. In the Portuguese archipelagos, no historical reports of WNV occurrence exist, although relevant mosquito-species are established, including *Culex theileri* in Madeira⁴³, and *Culex pipiens* in both Madeira⁴³ and the Azores⁴⁴. The bioclimatic setting of Madeira has previously demonstrated potential for mosquito-borne transmission, when in 2012 the island witnessed the first European sustained outbreak of the dengue virus⁴⁵, supported by a resident population of *Aedes aegypti*. In the Canaries, *Culex pipiens* is also established⁴⁴, and antibodies against WNV have been detected in Eleonora's falcons (*Falco eleonorae*)⁴⁶ and dromedary camels (*Camelus dromedarius*)⁴⁷. In the African archipelagos, São Tomé and Príncipe host persistent populations of vectors for dengue, chikungunya, Zika, and Yellow Fever viruses (*Aedes aegypti*, *Aedes albopictus*)⁴⁸, as well as *Culex quinquefasciatus*, a vector of WNV. Between 2022 and 2023, the first dengue outbreak was reported in São Tomé and Príncipe⁴⁹. The sole local evidence of WNV circulation comes from a serological study that reported cross-reactivity in humans by indirect ELISA⁵⁰. In Cape Verde, mosquito-borne outbreaks of dengue and Zika viruses have been reported in the past⁵¹, but no evidence exists for WNV circulation albeit the presence of *Culex pipiens* and *Culex quinquefasciatus*⁵². While such arboviral histories in the archipelagos confirm their epidemiological relevance, dengue, Zika and Chikungunya viruses are not zoonotic and depend mostly on *Aedes spp.* for

transmission, such that their local transmission success cannot be directly linked to potential for WNV transmission.

In this study, we developed an ecological niche modelling approach with the primary goal of projecting and understanding the ecological suitability for WNV occurrence in Atlantic archipelagos, where mosquito species capable of transmitting the virus are established, but local transmission risk is generally unknown and the relevant (a)biotic factors are largely unidentified. The approach is based on training a series of machine learning models using WNV occurrence and (a)biotic data from European countries, and subsequently projecting ecological suitability for the archipelagos. It estimates seasonal spillover potential as a proxy for local transmission, and critically gives weight to the temporal dimension, allowing to go beyond typical estimations of the spatial landscape of ecological suitability, but also being able to estimate seasonal dynamics in both the European and archipelago areas of study. Ultimately the presented results support the development of tailored preparedness, surveillance and mitigation strategies, addressing the unique ecological and seasonal dynamics of each region under study.

Methods

The study includes data and modelling output for three European countries (Portugal, Spain, Italy), as well as five Atlantic archipelagos - Madeira (one island), Azores (nine islands), Canaries (seven islands), Cape Verde (nine islands) and São Tomé e Príncipe (two islands). The European countries were selected based on the availability of reported WNV occurrence for several consecutive years, and the archipelagos based on the availability of (a)biotic variables at sufficient spatial resolution.

West Nile virus occurrence data

West Nile virus (WNV) occurrence data with associated geographical coordinates (longitude, latitude) was obtained from public sources (Supplementary Table S1).

For Portugal, we used a dataset from a recent publication on the historical human and non-human occurrences¹⁰. This dataset was mainly composed of serological evidence. The dataset was thus curated to include solely reported occurrences for which metadata supported a known locality and time of infections. Full data was composed of avian, human and equine occurrences. Reports based on serology and not supported by the presence of symptoms compatible with WNV infection were not considered, since serology on itself is insufficient to support the timing and location of sampling as the same of infection. The filtered dataset used for modelling included a total of 31 unique geographical occurrences (in a total of 34 infections), only in equines, covering the period October 2010 to October 2022.

For Spain, non-human occurrences were obtained from the RASVE dataset⁵³, complemented with human-related occurrences publicly available at the ECDC website¹⁰. The collated data is already declared as being composed of WNV confirmed cases (e.g., supported by VNT testing, viral isolation or with co-occurrence of WNV-compatible symptoms). Curations included occurrences for which the declared geographical location (latitude, longitude) was outside the boundaries of the country or for which dates included errors. The filtered dataset used for modelling included a total of 446 unique geographical occurrences (in a total of 545 infections), including humans, nonhuman vertebrates and mosquitoes, covering the period September 2010 to January 2024.

For Italy, data including human and nonhuman occurrences was obtained from a repository representing the period September 2012 to November 2022⁵⁴. The collated data is already declared as being composed of WNV confirmed cases (e.g., supported by VNT testing, viral isolation or with co-occurrence of WNV-compatible symptoms). Curations included occurrences for which the declared geographical location (latitude, longitude) was outside the boundaries of the country or for which dates included errors. The filtered dataset used for modelling included a total of 2719 unique geographical occurrences (in a total of 5657 infections), including

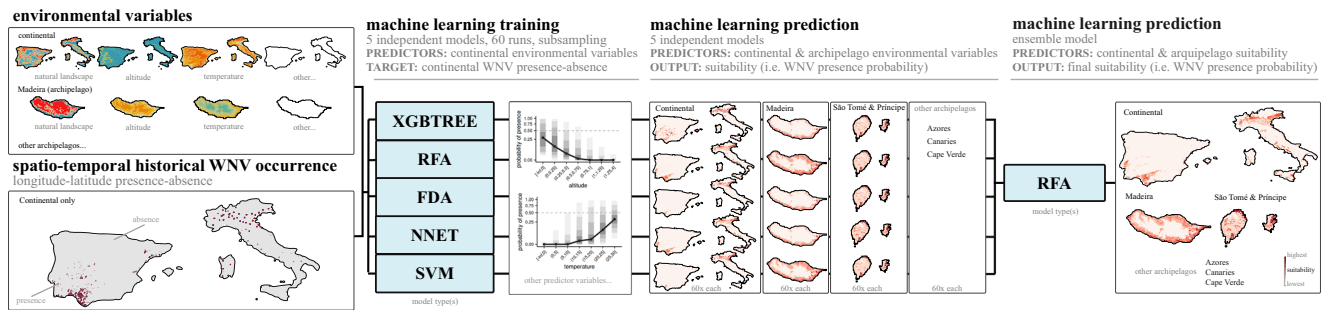


Fig. 1 | Summary diagram of the modelling approach. Summary of the modelling approach, informed by (a) biotic predictor variables and target WNV reported occurrences, towards estimating the probability of WNV “presence” per

geographical location (termed WNV suitability). Model types include: random forests (RFA), extreme gradient boosting trees (XGBTREE), support vector machine (SVM), neural network (NNET), flexible discriminant analysis (FDA).

humans, nonhuman vertebrates and mosquitoes, covering the period September 2012 to November 2022.

(A) biotic data

Several (a) biotic variables were also obtained from public sources. A total of (Supplementary Table S2): altitude, monthly near-surface wind speed, monthly near-surface temperature, monthly near-surface minimum temperature, monthly precipitation, monthly relative humidity, proportion of land associated with natural vegetation, proportion of land associated with crops, proportion of land associated with mixed vegetation (crops and natural), and monthly climate moisture index. Selection of variables was restricted by availability across all studied territories (many datasets of interest do not exist for archipelagos, including e.g., the spatial distributions of host species of interest such as avian species).

Training dataset

We constructed a training dataset based on information from (a) biotic and WNV occurrences of Portugal, Italy, and Spain. Since our climate type (a) biotic variables were defined for each month of a year, we started by collapsing the dates of WNV occurrences into months (i.e. excluding reported day and year). Data entries were created by iterating each month of the year and for each existing WNV occurrence, by: (a) extracting land type (a) biotic values for the reported WNV location (longitude-latitude point), and (b) extracting climate type (a) biotic values for both the reported WNV location (longitude-latitude point) and WNV reported month. For each such “presence” data entry (set of 11 variables: 10 predictors, 1 target = “presence”), steps (a) and (b) were applied to each geographical location without reported WNV occurrence, thus creating matching “pseudo-absence” data entries across the spatial scale and for the same month (set of 11 variables: 10 predictors, 1 target = “pseudo-absence”). The resulting training dataset can thus be imagined as a 12 layer replication (1 per month) of the aggregated spatial dimension of the geographical range of Portugal, Spain and Italy; in which each replicated layer will contain “presence” and “pseudo-absence” data entries. All locations (longitude-latitude point) are thus replicated 12 times, each time with fixed land type (a) biotic variable values but varying climate type (a) biotic variable values, and with “presence” or “pseudo-absence” status depending on reporting per month. Some locations may be represented by a single or multiple “presence” over the 12 replications, while others may be always represented by “pseudo-absence” across the 12 replications. The dataset is imbalanced towards “pseudo-absence” data entries, for which a model multi-run, cross-validation subsampling approach with downscaling to “presence” data entries was implemented (see details in **Methods**).

Ensemble-based modelling

We developed a supervised machine learning (ML) classification approach (summarised in Fig. 1) informed by ten (a) biotic variables (predictor variables, Supplementary Table S1) and WNV reported occurrences (target variable, Supplementary Table S2). The ML approach used five model types,

including random forests (RFA), extreme gradient boosting trees (XGBTREE), support vector machine (SVM), neural networks (NNET), and flexible discriminant analysis (FDA). Further information in **Supplementary Information**.

RFA, XGBTREE, SVM, NNET, and FDA have model-specific trade-offs regarding data and output. RFA and XGBTREE, e.g., excel at generalisation and sensitivity to sampling, handling extremely well high-dimensional and up to intermediately imbalanced data, being able to readily decipher complex data relationships much better than SVM and FDA. NNET also offers high flexibility and generalisation by being able to model non-linear, high-dimensional and hierarchical data relationships with ease, typically above any of the other models. These generalisation and flexibility potentials of RFA, XGBTREE and NNET are however traded-out by providing the lowest interpretability among the 5 models. For example, while predictor variables can be interpreted by their relative importance in learning and quantitative relationship with final predictions - which is one of the main goals of ecological niche modelling - it is not straightforward to spell out exactly why or how a specific prediction is made by RFA, XGBTREE and NNET. The “black-box model” terminology is typically well suited for these models. SVM and FDA present a contrasting scenario, in which they offer top interpretability, but more easily struggle with generalisation and flexibility regarding input data properties such as high-dimensional non-linear relationships, hierarchical structures or the presence of high variation or noise. All of the 5 models are sensitive to sample imbalance, with RFA and XGBTREE being the more robust, followed by NNET, FDA and SVM. Intermediate steps such as resampling (bootstrapping) and cross-validation for each model, as performed in this study, often soften issues related to sampling imbalance, but cannot eliminate the issue. Applying ensemble approaches has become standard in ML-based ecological niche modelling because, when taking into consideration the performance of each independent model, it has the potential to leverage the complementary strengths and pitfalls of each model, balancing interpretability, generalisation, and robustness to sampling imbalance.

Training was performed on all geographical locations (longitude-latitude points) at a spatial resolution of $\sim 5 \text{ Km}^2$ for Portugal, Spain and Italy. Locations with reported WNV occurrences were defined as “presence” and all others as “pseudo-absence”. The training dataset comprised records associated with each location, including: one target variable with a value of “presence” or “pseudo-absence” based on WNV occurrence reports; four predictor variables related to landscape properties specific to the location (constant over time); and six predictor variables related to climate properties specific to both location and month of the year. The five different ML model types were used to independently estimate the probability of WNV “presence” (here termed WNV suitability for occurrence), between 0 and 1, for which the threshold 0.5 was used for both classification and accuracy measurement (e.g.⁵⁵). We considered 60 independent runs of each model under a cross-validation strategy parameterised with 25 folds each and 0.8 subsampling of input data and downscaling for WNV “presence” data points. In effect, each fold of each model type was trained on N locations

(equal to 80% of existing WNV “presence” data points)) representing an equal number of WNV “presence” and “pseudo-absence” data points. In light of the WNV occurrence dataset being unbalanced towards “pseudo-absence”, the large number of model runs and folds guaranteed that the ecological background of each WNV “presence” data point was considered against a wide range of ecological backgrounds of the WNV “pseudo-absence” data points. A similar subsampling strategy was used on a recent ML approach successfully applied to WNV modelling in Portugal²¹.

For each model type, the 60 suitability estimations were generated for each country and archipelago, per month. The weighted mean of the 60 suitability estimations per month was calculated using observed model accuracy, resulting in a single suitability estimation per model type per month. Accuracy was defined as the classic ratio between the number of locations correctly predicted and the total number of locations considered; formally $\text{accuracy} = (\text{TP} + \text{TN}) / (\text{TP} + \text{TN} + \text{FP} + \text{FN})$, where TP = number of true positive locations, TN = true negative, FP = false positive, and FN = false negative. Negative and positive was related to whether a location was deemed suitable under the aforementioned threshold of suitability ≥ 0.5 . The latter were then used as predictor variables to train an RFA model, resulting in a final ensemble estimation of suitability for every territory under study per month. Suitability estimations were generated at a spatial resolution of 5 Km² for countries and 0.5 Km² for archipelagos. Several summary measures across model types (including the ensemble model) are reported, such as (a)biotic variable importance, accuracy, true positive and negative rates. Partial plots were generated by crossing final ensemble suitability estimates with (a)biotic variables per geographical location.

All computational tasks were performed using R V4.1.2 utilising the Caret R-package V6.0-9.0⁵⁶ for all ML computations (which installs specific R-packages for each ML approach). An R script with template code for the ensemble, machine learning pipeline is made available as supplementary material.

Further details on environmental data preparation, pseudo-absence generation, and other methodology are provided in **Supplementary Information**.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Results

In preparation for the modelling (Fig. 1), we performed a broad literature review focused on mosquito-species of interest, WNV and other mosquito-borne viral occurrences in the archipelagos under study. This review was translated into a report providing an aggregated historical and contextual perspective, made available in **Supplementary Information**.

Ecological niche modelling captures the spatio-temporal signatures of reported WNV occurrence

Across ML model types, measured predictor variable importance was organised around three groups (Fig. 2A): the first group with the highest importance included *alt* (altitude) and *tempmin* (minimum temperature); the second group, with intermediate importance, included *temp* (mean temperature), *wind*, *prec* (precipitation), *cmi* (moisture index) and *nat* (natural land cover); the third group with the lowest importance included *hum* (humidity), *mixed* (natural and crop land cover) and *crop* (crop land cover). Partial plots of predictor variables versus estimated WNV suitability revealed that some variables were positively associated (e.g., temperatures) while others were negatively associated (e.g., altitude) (Fig. 2B, Figs. S1–10).

Accuracy across model types was generally above 0.8, with the tree-based models (RFA, XGBTREE) being marginally better (Fig. 2C). The ensemble model outperformed the independent results of all other models with accuracy at ~0.92. The tree-based models had higher true negative rates (TNR) and lower true positive rates (TPR) than other models (Fig. 2C). The ensemble model once again outperformed all other models in both rates,

and its TNR of ~0.91 meant that in approximately 9% of all locations (longitude-latitude data points) for which there was no evidence of WNV circulation, the model concluded otherwise, potentially identifying locations where surveillance is missing or insufficient. The generally high accuracy meant that the spatio-temporal distribution of estimated WNV suitability mirrored existing WNV evidence, both in space (Fig. 2D) and time (Fig. 2E). For all continental regions used to train the models (Portugal, Italy, Spain), the temporal signal in estimated suitability mirrored the timing of peaks in local reporting, as well as the timing during which there is no evidence of WNV occurrence (Fig. 2E). In Portugal, for which a recent national blood donor seroprevalence study was performed with samples from 2022⁵⁷, the spatial distribution of yearly median suitability (Fig. S11D) and yearly proportion of land with non-zero suitability (Fig. S11E) were both positively correlated with WNV seroprevalence at the NUTS2 geographical level (0.7 and 0.5, respectively; p-values < 0.05).

Ecological suitability is rich in spatio-temporal dynamics

The seasonality of estimated WNV suitability exhibited similar patterns across the continental regions used to train the models (Portugal, Italy, Spain; Fig. 3A–C). In particular, suitability troughs were generally estimated between January and May, with peaks occurring from August to November, largely driven by shifts in the proportion of land deemed suitable (suitability ≥ 0.5). Seasonally, the proportion of land deemed suitable could go from ~0% (during troughs) to ~50% (during peaks) in Portugal (Fig. 3A), ~0% to ~20% in Spain (Fig. 3B), and ~0% to ~40% in Italy (Fig. 3C). There were also marked time-varying changes in absolute suitability both within land deemed suitable (Fig. 3a1,b1,c1) and land deemed unsuitable (Fig. 3a2,b2,c2), highlighting that seasonality for WNV occurrence is driven by an interplay of changes within both suitable and unsuitable land.

Among the archipelagos, estimated WNV suitability exhibited varying seasonal characteristics (Fig. 3D–H). Only archipelagos with subtropical climates (Azores, Madeira, Canaries) displayed distinct seasons, marked by consecutive months with either troughs or peaks in suitability, along with concurrent oscillations in the proportion of land deemed suitable (Fig. 3D, F, H). These archipelagos were also characterised by longer peak seasons in suitability compared to continental regions, with short troughs occurring between April and June in the Azores (Fig. 3D), March and May in the Canaries (Fig. 3F), and April and August in Madeira (Fig. 3H). Seasonally, the Azores showed the largest variation in the proportion of suitable land, ranging from ~0 to 26%, followed by the Canaries and Madeira, where the proportion varied from 0 to ~21% and ~12%, respectively. The Azores was unique in maintaining a semi-constant suitability level within land classified as suitable, concurrent with a seasonal variation in suitability within unsuitable land (Fig. 3d1–2). In contrast, Madeira and the Canaries showed seasonal variation within both suitable and unsuitable land, similar to continental regions (Fig. 3f1, 2, 3h1, 2). Seasonality in estimated WNV suitability was generally weak for the two tropical archipelagos (Cape Verde, São Tomé & Príncipe; Fig. 3E, 3e1, 2, 3G, 3g1–2). The seasonal variation in the proportion of land deemed suitable in these archipelagos was also significantly different, varying little from ~0% to 4% in São Tomé & Príncipe (Fig. 3E) and widely from ~0% to 32% in Cape Verde (Fig. 3G). In both archipelagos, there was no month of the year presenting zero proportion of land deemed suitable. Notably, Cape Verde’s large variation in the proportion of suitable land was estimated to have a trough between July and November, contrasting to all continental regions.

The association between ecological suitability and human populations reveals subregions of public health importance

Overall, the seasonal characteristics of estimated WNV suitability across the study regions highlighted that seasonality is as much about time (periods with troughs and peaks) as it is about the proportion of land deemed suitable. Based on the spatial distributions of estimated ecological suitability (Figs. S12–19), we thus aimed at identifying and mapping geographical hotspots deemed suitable (Fig. 4A), quantifying their respective proportion of land within the territory over a year (Table S3). The vast majority of land

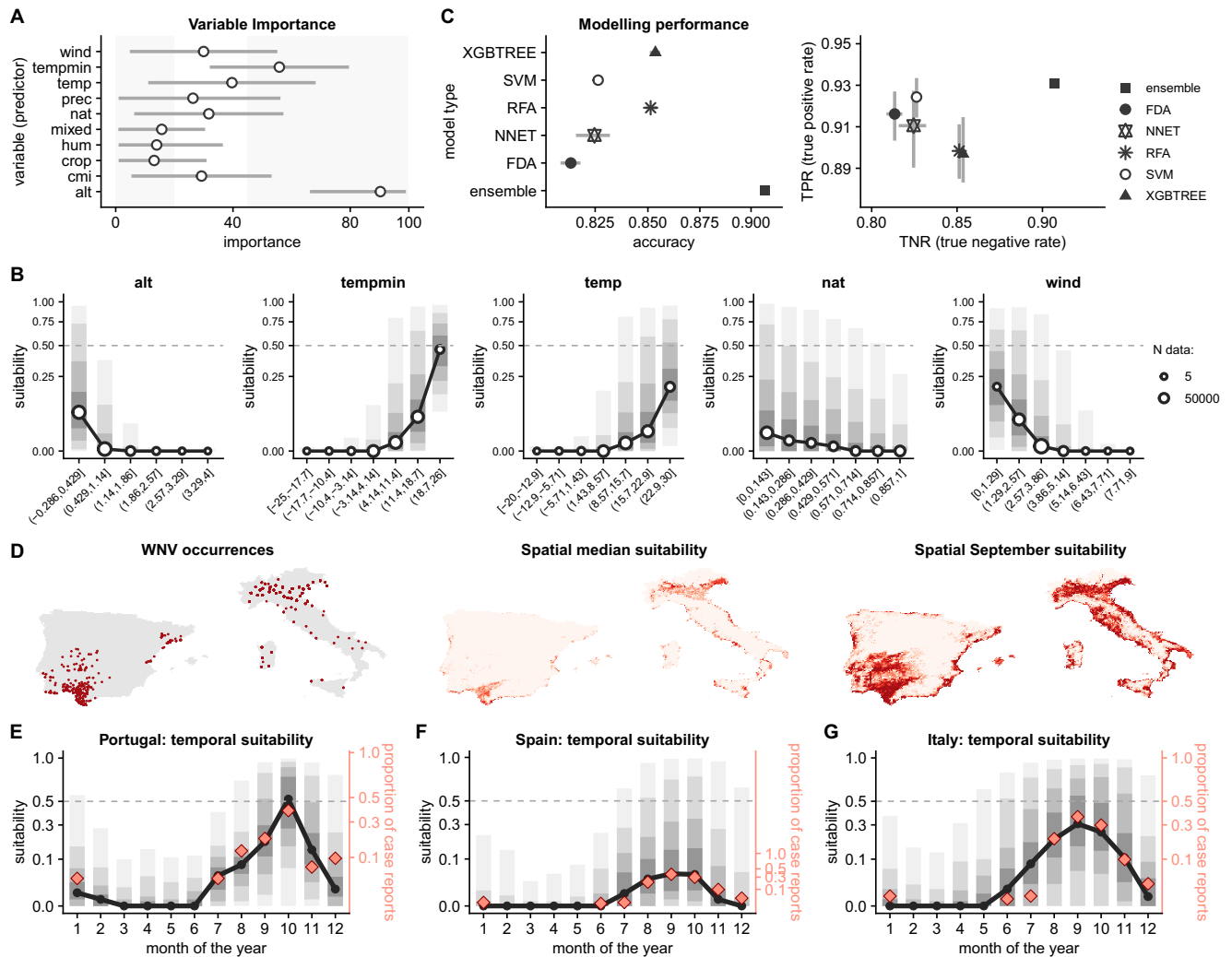


Fig. 2 | Summary of modelling variable importance, performance and estimated suitability. **A** Variable importance across model runs. Points are the median and bars the 95 percentile. **B** Partial plots of estimated suitability versus the top five most important variables, presented as aggregates across continental Portugal, Spain and Italy. Horizontal dashed lines at suitability 0.5 show the threshold for model decision on suitable versus unsuitable. **C** On the left, model performance as measured by accuracy for each model across model runs. On the right, model performance as measured by true positive rate (TPR) and true negative rate (TNR) for each model across model runs. Models are represented by symbols as indicated in the legend to

the right. Points are the median and bars the 95 percentile measured across data points (including all months). **D** Maps presenting the distribution of WNV occurrences used to train the models, and estimated suitability in September and median across time. **E, F, G** Temporal dynamics of estimated suitability for each of the three countries used for training, together with the proportion of case reports per month within each country (orange). Only months with reported occurrences are visually presented (orange), the other months have zero occurrences. For (**C, E–G**), the shaded bars per month present the 25, 50, 75 and 95 percentiles (from darker to lighter), and points and lines present medians.

had at most one month of the year deemed as suitable (grey areas in Fig. 4A). In total, 27% of land across all continental regions, and 23% across all archipelagos was deemed suitable for more than one month (non-grey areas in Fig. 4A), with proportions of 40% for Italy, 32% for Cape Verde, 27% for the Azores, 22% for Portugal, 19% for the Canaries, 18% for Spain, 12% for Madeira and 6% for São Tomé & Príncipe. In continental regions, this proportion of suitable land presented a clear north-south pattern, mirroring the known distribution of WNV occurrence (Fig. 2D), with suitable land predominantly in southern Portugal, the southwestern interior and eastern coast of Spain, and northern and coastal areas of Italy (Fig. 4A). Within archipelagos, the proportion of suitable land varied significantly from 8% (Fogo) to 88% (Sal) in Cape Verde, 13% (La Palma) to 24% (Lanzarote, Fuerteventura) in the Canaries, 13% (São Jorge) to 71% (Graciosa) in the Azores, 12% for the island of Madeira, and 7% for both islands in São Tomé & Príncipe. Of the 29 islands modelled, the five with the largest proportion of suitable land belonged to Cape Verde and the Azores.

Focusing on land suitable for more than 6 months in a year as a proxy for high transmission risk (magenta areas in Fig. 4A), we estimated a total of

6% in Madeira, 4% in the Azores, Cape Verde and Italy, 3% in São Tomé & Príncipe, and 1% in Portugal, Spain and the Canaries. In total, this equated to ~2% of land for both the continental and archipelago territories. Again, the five islands with the largest proportion of such land belonged only to Cape Verde and the Azores. This proportion of land varied between 0% (Brava) to 11% (Maio) in Cape Verde, 0% (four islands) to 2% (Tenerife) in the Canaries, 1% (São Jorge, Faial) to 19% (Graciosa) in the Azores, 6% on the island of Madeira, and 1% (Príncipe) to 3% (São Tomé) in São Tomé & Príncipe.

As such, across the regions of study, the proportion of land (Table S3) and total land area (Table S4) deemed suitable within each region decreased with the length of time considered (Fig. 4A). In contrast, analysing the distribution of human population sizes within land deemed suitable over increasing periods of time revealed a positive association (Fig. 4B). This implies that while locations deemed suitable for longer periods of time represent the smallest proportion of land, they disproportionately contribute to higher exposure risk to humans by harbouring larger communities (with the exception of Cape Verde) (Table S5).

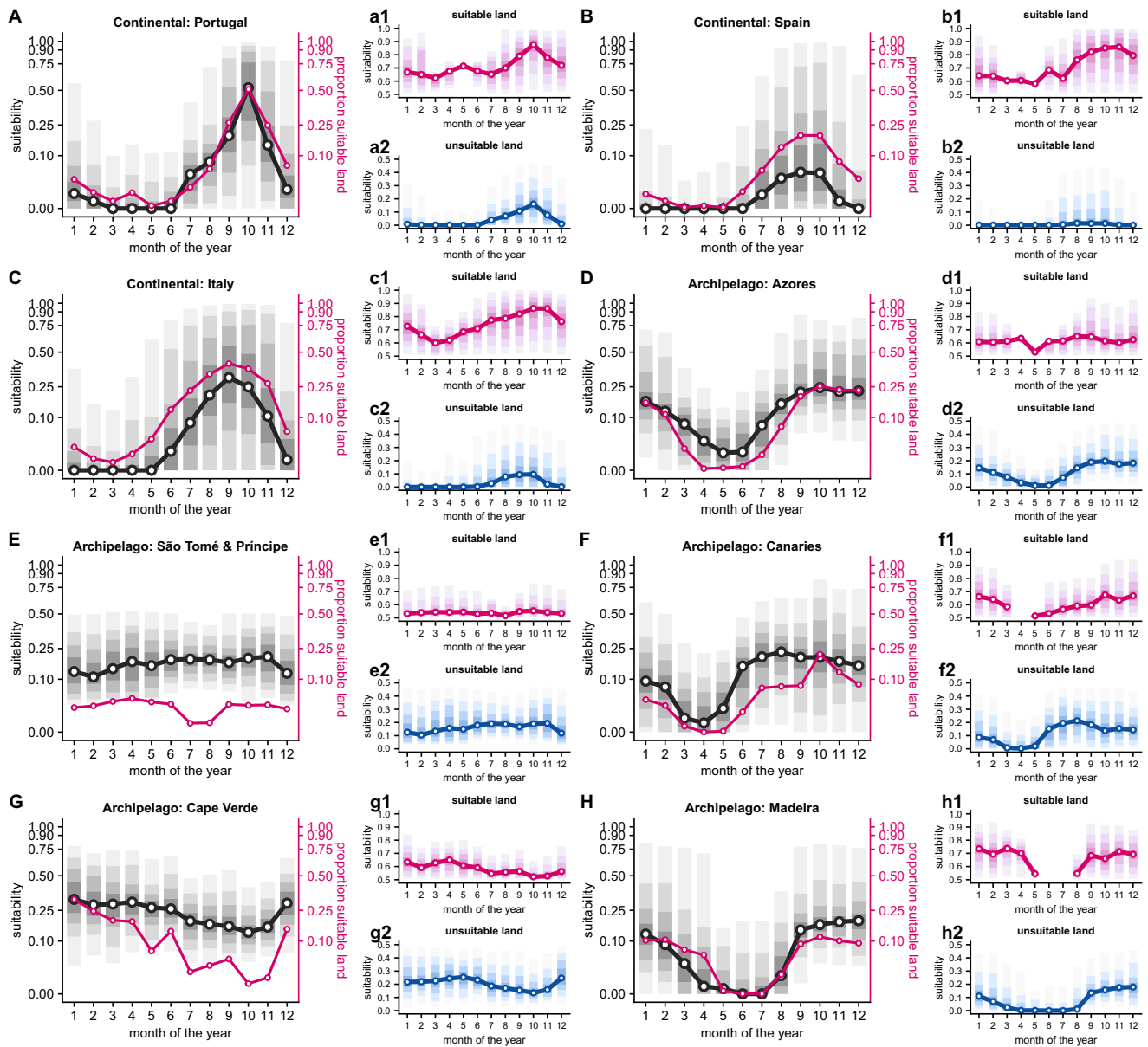


Fig. 3 | Temporal dynamics of estimated suitability across regions of study. For **A** Portugal, **B** Spain, **C** Italy, **D** Azores, **E** São Tomé & Príncipe, **F** Canaries, **G** Cape Verde and **H** Madeira. Each (**A–H**) is composed of: (left) a large subpanel presenting the distribution of monthly suitability and the monthly proportion of suitable land; (right-top) a smaller subpanel presenting the distribution of monthly suitability in

suitable land (spatial pixels with suitability ≥ 0.5); (right-bottom) a smaller sub-panel presenting the distribution of monthly suitability in unsuitable land (spatial pixels with suitability < 0.5). For all sub panels the shaded bars per month present the 25, 50, 75 and 95 percentiles (from darker to lighter), and points and lines present medians.

To formalise the aforementioned relationship between ecological suitability and human population sizes, we defined the *risk to the human population* per subregion as the weighted median of suitability by population size (Fig. S20; full distributions in Table S6). In practice, we considered the estimated spatio-temporal suitability and known population size of all geopixels within well defined subregions, using NUTS III boundaries (Nomenclature of Territorial Units for Statistics, level 3) for European countries and island boundaries for archipelagos (Fig. 5). The spatial distribution of risk to the human population presented patterns of public health relevance. In Portugal, the Algarve presented the highest risk, followed by the subregions within and neighbouring the largest urban centres of Oporto and Lisbon. In Spain, a clear hotspot with highest risk was centred in Seville and neighbouring subregions, followed by Eastern coast subregions including Alicante, València, Castellon, Tarragona and Barcelona presented non-negligible risk. A few hotspots with highest risk were

identified for Italy, including the subregions of Milan and Naples, and a large cluster of Northeastern subregions including Pordenone, Udine, Gorizia, Treviso and Venezia. At the same time, a large number of subregions presented non-negligible to intermediate risk across the entirety of the north, and the islands of Sicily and Sardinia.

Across all islands of the archipelagos under study, none had negligible risk. Azores and Cape Verde had the islands with the highest risk across all archipelagos. In the Azores, the islands of Graciosa and Flores were similar, presenting the highest risk, at approximately double of the lowest risk found for the island São Jorge. In Cape Verde, the islands Sal, Maio and Boavista had the highest risk, also at approximately double of the lowest risk found for Fogo. In the Canaries, La Gomera was found to have the lowest risk across all archipelagos, Lanzarote to have the highest within the archipelago. When compared across the islands of all archipelagos, the island of Madeira had intermediate risk. Finally, in the archipelago of São Tomé and Príncipe,

the island of São Tomé had the highest risk, but overall intermediate in the context of all archipelagos.

Discussion

Driven by increased human mobility, urbanisation and environmental changes, mosquito-borne viruses are emerging as global public health threats. WNV remains a well documented case³⁸, particularly in Europe, where its geographical range and epidemic activity have expanded significantly over the past decade. While recent awareness of the virus has led to a growing body of research, studies have remained largely focused in North America and Europe, with more recent attention to Africa. Archipelagos have insofar been neglected. An overview of historical evidence on mosquitoes and mosquito-borne viruses in Madeira, Azores, Canaries, Cape Verde and São Tomé and Príncipe reveals that most information is available through the research literature but that significant knowledge gaps remain, specifically regarding WNV (**Supplementary Information**). We found that digitalisation and publishing of mosquito-borne virus occurrences was only applied in Cape Verde and São Tomé and Príncipe (which have witnessed sporadic but large mosquito-borne virus outbreaks in the past). Even in these cases, there was no differentiation between autochthonous and imported occurrences and recording of their geo-localisation virtually absent. Future efforts towards digitalisation, geo-location and differentiation between autochthonous and imported infection events in archipelagos are needed and remain critical points for better computational epidemiology research.

Leveraging abundant WNV occurrence data from Portugal, Spain and Italy, we aimed at estimating and better understanding the ecological suitability for WNV occurrence in Atlantic archipelagos, where competent mosquito-species are present but little is known about local transmission risk. Adding to the typical modelling output of estimating the spatial landscape of WNV suitability for the regions under study, we offer novel insights into temporal properties of seasonality and risk to the human population for both the countries and archipelagos. In general, we find that the machine learning approach was able to replicate both the seasonal timings and spatial distribution of historical WNV occurrence in the continental territories of European countries. Consistent with other studies, the ensemble model, integrating the predictions from five different models, outperformed all individual models. Although the modelling approach does not provide evidence of causation per se, results related to predictor variable importance and its relationship with estimated suitability for WNV occurrence were generally compatible with existing knowledge regarding the enzootic cycle and spillover dynamics of WNV.

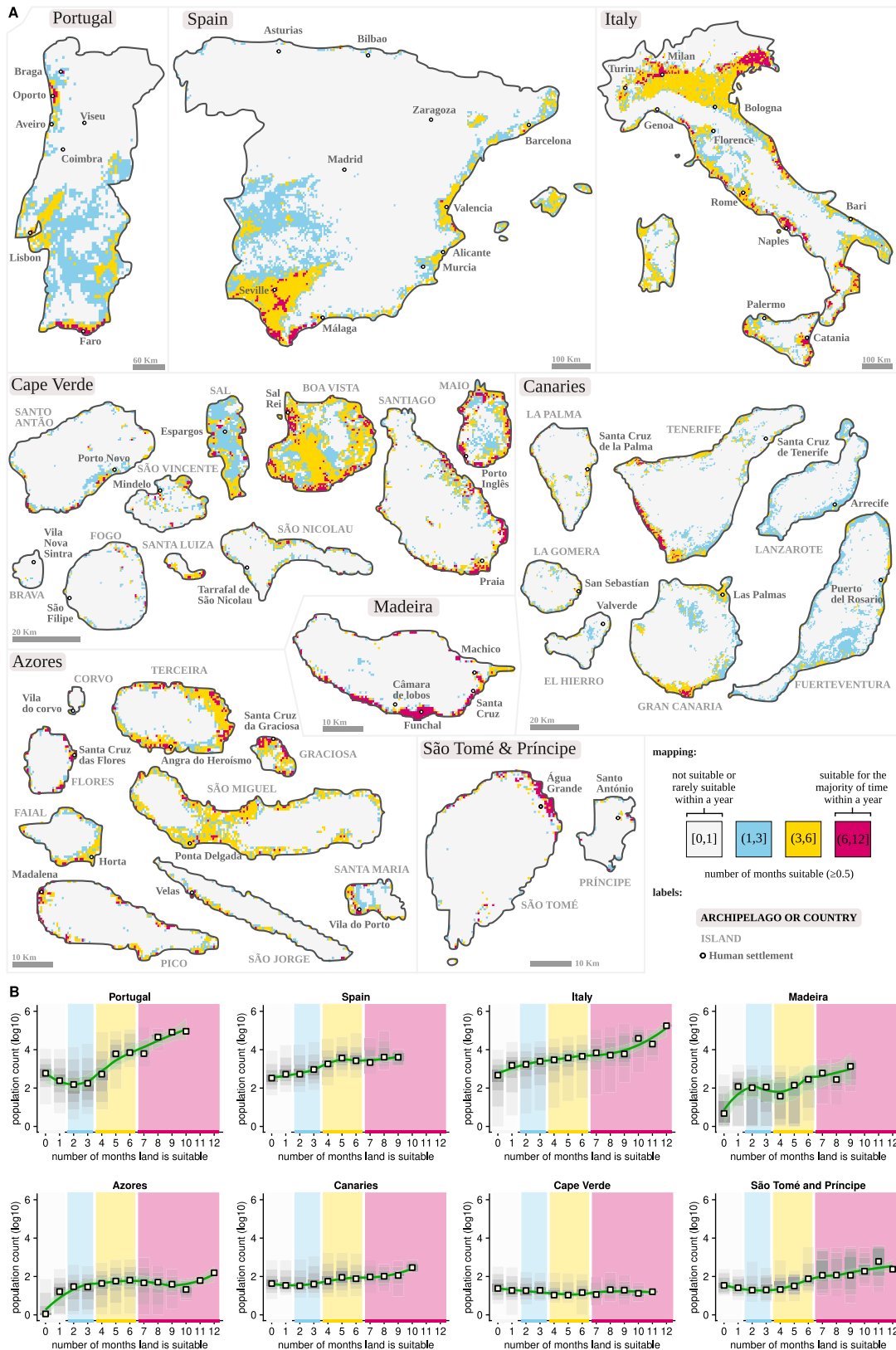
Key predictor variables included altitude, mean and minimum temperature. Altitude is a known abiotic factor that affects mosquito traits (e.g., wing size^{19,59}) influencing mosquito abundance⁶⁰ and consequently virus transmission potential. Several studies have shown that WNV occurrence is negatively associated with altitude (e.g., refs. 61–63), and accordingly, altitude showed a consistent negative association with estimated suitability for WNV occurrence in this study. Temperature is one of the most extensively studied climatic variables associated with WNV seasonality and spatial distribution. It directly affects mosquito thermal performance curves and consequently key traits⁶⁴, such as mortality rate and viral incubation period, influencing population size and vectorial capacity (i.e., viral transmission potential). For example, temperatures below 10 degree Celsius are considered unsustainable for WNV transmission by *Culex spp.*⁶⁵. Our modelling outputs confirmed the expected positive relationship between temperature and suitability for WNV occurrence. To the best of our knowledge, wind has not commonly featured in quantitative approaches to WNV occurrence. However, it has been suggested that wind may positively influence WNV occurrence and dispersal dynamics⁶⁶, e.g. through its effect on mosquito and avian dispersal⁶⁷, although negative associations have been demonstrated e.g. in Greece⁶⁸. In this study, wind was consistently negatively associated with suitability across Portugal, Spain and Italy. The role of precipitation on WNV occurrence remains complex, with contrasting conclusions frequently reported. While higher precipitation can theoretically increase

mosquito abundance, lower precipitation can also force the accumulation of organic material that favours larval survival and development, resulting in higher larval survival and thus mosquito abundance⁶⁷. In general, a weak negative relationship between precipitation and suitability was recovered in this study, in line with some studies (e.g., refs. 22,23) and in contrast with others (e.g., refs. 66,69). Similarly to precipitation, the current understanding of the relationship between humidity and WNV occurrence remains inconclusive⁶⁷, with studies e.g. reporting negative associations in Greece⁶⁸ and Romania⁶³, but positive association in Israel⁷⁰. In this study, a negative association between humidity and estimated suitability was found, which was clearer than that of precipitation but weaker than all other predictor variables. Rather unique was the inclusion of the climate moisture index, which measures local balance between moisture supply (precipitation) and demand (potential evapotranspiration). Overall, the relationship of the index with estimated suitability for WNV occurrence was negative, with a clear trend for Italy compared to Portugal and Spain. It had a higher predictive value than both humidity and precipitation, suggesting its usefulness in future predictive modelling. Regarding land use, estimated suitability typically had a consistent negative association with the proportion of natural cover, and a positive association with the proportion of crop and mixed covers. These observations were in accordance with various sources of evidence, suggesting a link between WNV occurrence and land associated with agriculture^{60,71}.

From the model's potential to estimate time-varying outputs, both the absolute suitability levels and the proportion of suitable land within a region presented relevant seasonal patterns. Three general behaviours emerged from WNV suitability. Temperate regions (Portugal, Spain and Italy) had short seasonal windows peaking in late summer and early autumn, aligning with reported WNV seasonality. Subtropical archipelagos (Azores, Canaries and Madeira), had longer seasonal windows with peaks distributed across most of the year except spring and early summer. Tropical archipelagos (Cape Verde, São Tomé and Príncipe), showed minimal seasonal variation with weak or no clear troughs and peaks. Regarding the seasonality of the proportion of land deemed suitable, similar conclusions were reached across study regions, albeit with Cape Verde behaving more like subtropical regions despite its tropical climate. In general, these results highlighted that WNV seasonality, both spatial and temporal, generally weakens as climates shift from temperate, to subtropical to tropical.

WNV is maintained in a zoonotic cycle dependent on avian and mosquito species. Within this cycle, some *Culex spp.* involved in mammalian transmission are less anthropophilic than *Aedes spp.* which are key vectors for other arboviruses. It is thus tempting to associate the human risk for WNV transmission to areas with smaller or less dense human populations where avian and *Culex spp.* more likely have better established population distributions. This association, however, has had contrasting evidence, with WNV occurrence being associated with agricultural land cover in certain regions (e.g., refs. 9,71) and associated with urban land cover in others (e.g., refs. 71,72). Differences between studies in the type of WNV occurrence data used (i.e. unique or combinations of human, equine, mosquito and avian sources) likely contribute to such contrasting evidence. As highlighted in the context of the 2020 large WNV outbreak in Spain, other factors may contribute to discrepancies⁹. For example, that transmission is likely much wider than is reflected by limited data sources whose information is itself influenced by constant avian movement between urban and non-urban areas, or that there is varying avian biodiversity of importance for WNV transmission between closely located regions, and that there is uneven contribution of mosquito species with slightly different environmental preferences, susceptibility, infectivity and biting habits. Similar to some studies and in contrast to others, the current study found that agriculture-related land cover was most informative under a positive association with estimated suitability for WNV occurrence.

When analysing the number of months in a year that locations remain suitable for WNV occurrence, a striking positive association with local human population size was observed across all study regions (except for Cape Verde). When weighting local estimated suitability by community size



as a proxy for risk to human populations, we were able to describe the spatial heterogeneity of risk across subregions of the countries and islands of the archipelagos. For countries, as expected, many subregions with large human populations were identified as having relevant risk levels, but other subregions with smaller populations were revealed as having higher risk than

expected from solely considering suitability. This reiterated the conclusions from some of the intermediate exercises performed in this study, which revealed that locations deemed suitable for WNV occurrence for longer periods of time were seemingly associated with larger local human populations. In turn, this highlighted the importance of exploring the actual

Fig. 4 | Spatial hotspots and human population under land deemed suitable for varying periods of the year. A Maps of all studied regions are presented, coloured according to the legend on the bottom right: univariate categorised 1×4 colour scale representing the total number of months within a year for which a location is suitable (the machine learning classification threshold of suitability 0.5). For countries, the main human settlements (cities, towns, etc) of the top ten metropolitan regions are identified (except for Portugal, for which only seven exist according to Eurostat). For archipelagos, the name of each island (capital, light grey) and largest human

settlement (dark grey, open circles) are identified. For visualisation, the position of some islands do not reflect their real location, and some areas of sea were removed such that the spatial scales presented in kilometres refer solely to land area. **B** Distribution of human population size under suitable land (suitability >= 0.5) according to the number of months the land is suitable. For all sub panels, the shaded bars per month present the 25, 50, 75 and 95 percentiles (from darker to lighter) and Loess curves (green lines) represent the trend of the median values (squares). The colours in the x-axis follow the colour scale of (A).

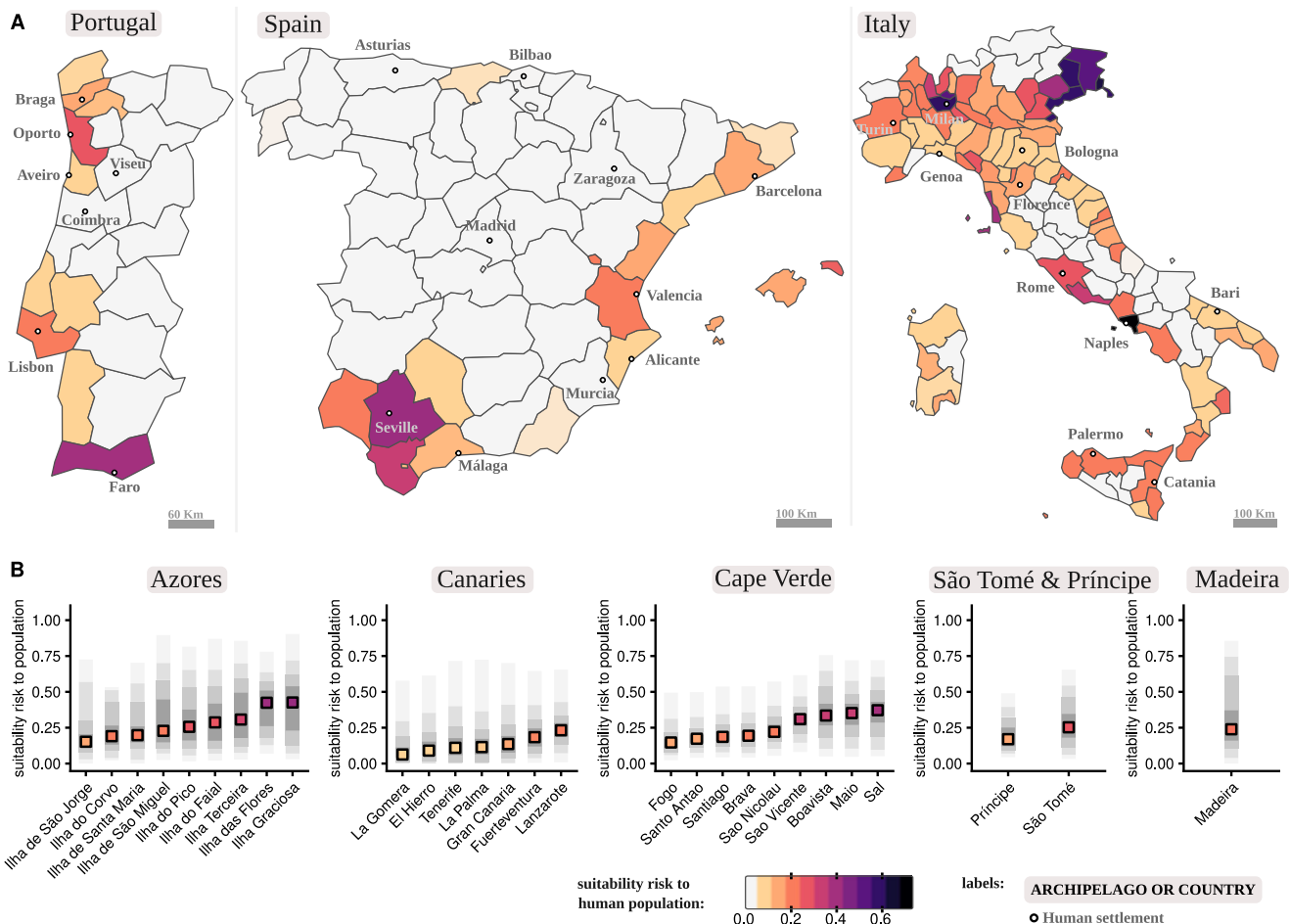


Fig. 5 | Ecological suitability risk to human populations. A European countries (Italy, Portugal, Spain) are mapped according to NUTS III regions (Nomenclature of Territorial Units for Statistics, level 3). Regions are coloured according to the estimated median suitability risk to human populations (colour scale at the bottom). **B** Archipelagos (Azores, Canaries, Cape Verde, São Tomé and Príncipe, Madeira)

are presented in XY plots showing the distribution of suitability risk to human populations per island (25, 50, 75 and 95 percentiles presented as bars from darker to lighter, respectively) including the median values (squares, colour scale at the bottom). The distributions for all NUTS III and islands are included in Supplementary Fig. S20, also available as Supplementary Table S6.

association between estimations of pathogen transmission potential and the distribution of relevant hosts, since e.g. regions with higher estimated transmission potential do not necessarily overlap with a significant presence of hosts relevant for public health (and vice-versa).

Limitations of this work include e.g., those related to data sources. Namely, available satellite data at adequate spatial resolution for archipelago modelling included climatic variables representative of monthly averages for the period between 1981 and 2010 and snapshots for particular years regarding land types. As such the data used for training and projection did not specifically represent the dates of WNV occurrences, and allowed only to perform suitability projections per typical month of the year (Jan–Dec) rather than for every day, month and year included in the WNV occurrence data. While this allowed us to estimate monthly spatial landscapes of suitability and generate novel insights into the expected variation of seasonality within a typical year, future research should implement higher temporal

resolution outputs once satellite data is available at better spatio-temporal resolutions, specially for the archipelagos. It should also be noted that mapping of satellite variables on islands is potentially less accurate than mapping over continental zones due to challenges in prediction near the land-sea interface, which may have introduced spatial artifacts in the suitability outputs. Avian and mosquito species’ distribution or mobility/migration datasets are also not available for archipelagos, although they could be of significant value for WNV ecological niche modelling and should be explored in future modelling studies when available. Furthermore, while the WNV occurrence data used for model training from the three European countries included mosquito, human, equine and avian sources, it was nonetheless enriched in human and equine sources and training was based on purely environmental variables. As such, the resulting suitability estimations could potentially be biased to represent a summary measure of spillover potential as driven by environmental backgrounds; i.e. outputs are

potentially more informative towards a better understanding of the landscape of spillover risk to humans and equines rather than viral maintenance in the natural avian-mosquito reservoir, and may also be limited to the contribution of the environment to risk, not addressing other potential human-related factors (e.g. behaviour) that may modulate individual risk. Still on the topic of occurrence data, it is of relevance to note that when using presence-only occurrence data for ecological niche modelling, reporting bias may be a relevant factor for model output. Typically, combinations of environmental conditions (predictor variables) may be overrepresented in model training (e.g. leading to autocorrelation), since occurrence (infection) reports may be enriched to particular time periods and/or geographical regions due to inherent bias in surveillance, research or public health efforts. As detailed in **Supplementary Information**, we considered five approaches that can hamper sampling bias effects but not all could be applied. Another relevant factor in ecological niche modelling is the risk of overfitting some or all of the models when using presence-only occurrence data. Significant overfitting would have resulted e.g. in outstandingly high model accuracies and the spatial distribution of estimated suitability mirroring the locations of reported WNV used for training. None of these expectations were observed. For example, the independent accuracy of the five models was approximately between 0.8 and 0.85, and of the ensemble model was ~ 0.9 . Considering environmental bias correction of pseudo-absence sampling for model training by ecological buffering reduced the ensemble model's capacity to reproduce temporal patterns and in fact was suggestive of overfitting by increasing accuracy by ~ 0.2 per doubling buffer radius (Figs. S11–2). At the same time, using Portugal as an example for which a national serosurvey was available, the spatial distribution of estimated suitability did not mirror the locations of reported WNV used for training while critically being statistically correlated with WNV seroprevalence at a spatial level (Fig. S11). Regarding the use of predictor variables, it is generally the case that geographical regions rich in historical WNV occurrence data that serve as training grounds for modelling approaches also typically provide varied data sources for predictor variables. However, difficulty in assembling sets of predictor variables that exist for both such training regions and archipelagos remains one of the main reasons why islands do not commonly feature in ecological modelling studies of arboviruses. In this study, an effort was made to not exclude predictor variables from the already small list that was available in both of these geographical contexts. To this effort, feature selection was not performed, and we considered an ensemble approach based on five machine learning algorithms that are either insensitive (RFA, XGBTREE) or only moderately sensitive (SVM, NNET, FDA) to predictor multicollinearity, excluding highly sensitive approaches such as generalised linear models. For those moderately sensitive to multicollinearity, there are possible implications for reduced precision and both variation in predictor importance and prediction among independent runs. To counterbalance this expected variation in output, we considered 60 independent runs from each approach, which were considered for the final ensemble output, weighted by their measured accuracy. The insensitive approaches (RFA, XGBTREE) were identified as the most accurate and thus weighted more on the ensemble solution. The selection of only insensitive and moderately sensitive models, use of a large number of independent model runs, and an ensemble output weighting more on the insensitive models, should have minimised the possible detrimental effects of predictor multicollinearity in a scenario in which availability of predictor variables is already extremely limited. Another limitation is the inability to infer causation, such that all outputs, specifically those regarding predictor variables, should be interpreted purely as informative associations. This, however, is a general limitation of machine learning approaches and is not specific to the current study. As discussed in more detail in the **Supplementary Information**, the need to downscale predictor variables to be able to work with the low spatial areas of some of the inhabited islands of the five archipelagos may have resulted in reduced model sensitivity to localised habitat ecological conditions. The results herein presented should therefore be interpreted as a first attempt to map local suitability, which still require further validation once environmental data sources with higher spatial resolutions become

available. Another limitation is that while we have previously shown that avian biodiversity is informative to predict the WNV occurrence spatial landscape in Portugal, and others have shown that mosquito distributions play a role in observed WNV occurrence spatial landscapes elsewhere, such mosquito and avian data was not included in the current study. The reason for this was a lack of available data sources, especially for the archipelagos (that is, even collating such data for the European training regions would not suffice since model projections for archipelagos would require such data to also be available for all islands). For mosquitoes, methods that combine remote sensing with entomological field data are showing promise in predicting mosquito distributions by correlating field data (e.g. trap counts, infestation levels, etc.) with environmental variables (e.g., vegetation indices, landscape, climate, etc.); but these have so far not been applied to the studied archipelagos. A possible future venue is to pursue reanalyses of archipelago suitabilities where and when new avian and/or mosquito-related data becomes available.

In brief, we have developed a machine learning approach trained on European data to project WNV ecological suitability in several Atlantic archipelagos where competent mosquito-species are present but WNV transmission potential remains largely unknown. Through ecological suitability analyses, unique quantitative and dynamic insights are provided related to the spatio-temporal suitability for WNV occurrence and risk to the human population across Portugal, Italy, Spain, Madeira, Azores, Canaries, Cape Verde and São Tomé and Príncipe. Towards capacity building, this work also produces unique open-access resources. For example, we provide an in depth review of the literature on existing historical evidence on WNV and mosquito-species of interest per archipelago, as well as datasets with estimated spatio-temporal ecological suitability and human-related risk estimates per country subregion and per island across the archipelagos. In the context of the archipelagos, these resources are novel and unique, and could fundamentally transform integrated surveillance systems in such currently data-poor settings. For example, they provide baseline suitability maps that could enable strategic placement of entomological monitoring efforts in high-suitability areas, thereby optimising resource allocation for mosquito surveillance. The quantitative temporal dynamics outputs could also allow the public health sector to raise clinical awareness and implement adequate diagnostic protocols during estimated peak transmission periods, enhancing syndromic case detection efficiency when laboratory capacity is constrained. Furthermore, the baseline suitability maps could guide the integration of early warning triggers, creating automated alert systems that prompt intensified mosquito collection and human syndromic monitoring when ecological conditions favour WNV transmission. In summary, these resources would enable the establishment of evidence-based sentinel surveillance networks that combine strategic mosquito monitoring with targeted clinical surveillance in human populations at highest risk, creating a cost-effective integrated system particularly valuable in island contexts where comprehensive surveillance coverage is logistically challenging and resource-constrained.

Ethical approval

No ethical approval was needed for this study, since all data sources are of public access, none of the generated data is at the individual level nor can it identify individuals, and WNV occurrence data used were already anonymised (as made available in the public repositories or literature).

Supplementary Information

Supplementary Information File, including review of literature, additional methodological details and modelling outputs, Supplementary Tables S1–2 and Figures S1–20. Supplementary Data 1, including Supplementary Tables S3–6 with data related to figures. Supplementary Data 2, including the code template for the ensemble, machine learning pipeline.

Data availability

Ecological suitability for WNV occurrence is made available within a Figshare repository (DOI:10.6084/m9.figshare.c.7697918) in GeoTiff format

(.tif), which is compatible with the Terra R-package⁷³. The most efficient way to access these resources in Figshare is to search by title “GeoTiffs with estimates of ecological suitability for WNV occurrence for regions stated in the filenames”. Curated WNV occurrence data is also provided in the same Figshare repository. The most efficient way to access these resources in Figshare is to search by title “CSVs with curated WNV occurrence records for Spain, Portugal and Italy”. Data otherwise presented in Figures is provided in table format in Supplementary Data. Input data is not made directly available due to resharing restrictions but all data is freely accessible on the cited repositories (see Table S2).

Code availability

Code template (in R) for the ensemble, machine learning pipeline is provided in **Supplementary Code File**.

Received: 10 April 2025; Accepted: 10 December 2025;

Published online: 03 January 2026

References

- West Nile Virus: Epidemiology and Ecology in North America. in *Advances in Virus Research* vol. 61 185–234 (Academic Press, 2003).
- Fall, G. et al. Biological and phylogenetic characteristics of West African lineages of West Nile virus. *PLoS Negl. Trop. Dis.* **11**, e0006078 (2017).
- Chancey, C., Grinev, A., Volkova, E. & Rios, M. The global ecology and epidemiology of West Nile Virus. *Biomed Res. Int.* **2015**, 376230 (2015).
- McDonald, E. Surveillance for West Nile Virus Disease—United States, 2009–2018. *MMWR Surveill. Summ.* **70**, 1–15 (2021).
- Haussig, J. M. et al. Early start of the West Nile fever transmission season 2018 in Europe. *Euro Surveill.* **23**, 1800428 (2018).
- Riccardo, F. et al. West Nile virus in Europe: after action reviews of preparedness and response to the 2018 transmission season in Italy, Slovenia, Serbia and Greece. *Glob. Health* **16**, 1–13 (2020).
- Bakonyi, T. & Haussig, J. M. West Nile virus keeps on moving up in Europe. *Eurosurveillance* **25**, 2001938 (2020).
- García San Miguel Rodríguez-Alarcón, L. et al. Unprecedented increase of West Nile virus neuroinvasive disease, Spain, summer 2020. *Euro Surveill.* **26**, 2002010 (2021).
- Figuerola, J. et al. A One Health view of the West Nile virus outbreak in Andalusia (Spain) in 2020. *Emerg. Microbes Infect.* (2022) <https://doi.org/10.1080/22221751.2022.2134055> (2022).
- Historical data by year - West Nile virus seasonal surveillance. *European Centre for Disease Prevention and Control* <https://www.ecdc.europa.eu/en/west-nile-fever/surveillance-and-disease-data/historical> (ECDC, 2024).
- Prioritizing diseases for research and development in emergency contexts. <https://www.who.int/activities/prioritizing-diseases-for-research-and-development-in-emergency-contexts>.
- Murgue, B., Zeller, H. & Deubel, V. The ecology and epidemiology of West Nile Virus in Africa, Europe and Asia. *Japanese Encephalitis and West Nile Viruses 195–221* (Springer, 2002).
- Klingelhöfer, D. et al. A virus becomes a global concern: research activities on West-Nile virus. *Emerg. Microbes Infect.* (2023) <https://doi.org/10.1080/22221751.2023.2256424>.
- García-Carrasco, J.-M., Muñoz, A.-R., Olivero, J., Segura, M. & Real, R. Mapping the risk for West Nile Virus Transmission, Africa - Volume 28,2022 - *Emerging Infect. Dis. J.* - CDC. <https://doi.org/10.3201/eid2804.211103>.
- Mencattelli, G. et al. Spatial and temporal dynamics of West Nile virus between Africa and Europe. *Nat. Commun.* **14**, 6440 (2023).
- García-Carrasco, J.-M. et al. Present and future situation of West Nile virus in the Afro-Palaeartic pathogeographic system. *Ecography* <https://doi.org/10.1111/ecog.06941> (2024).
- García-Carrasco, J.-M. et al. West Nile virus in the Iberian Peninsula: using equine cases to identify high-risk areas for humans. *Euro Surveill.* **28**, 2200844 (2023).
- Harrigan, R. J., Thomassen, H. A., Buermann, W. & Smith, T. B. A continental risk assessment of West Nile virus under climate change. *Glob. Chang. Biol.* **20**, 2417–2425 (2014).
- Valiakos, G. et al. Use of wild bird surveillance, human case data and GIS spatial analysis for predicting spatial distributions of West Nile virus in Greece. *PLoS One* **9**, e96935 (2014).
- Erazo, D. et al. Contribution of climate change to the spatial expansion of West Nile virus in Europe. *Nat. Commun.* **15**, 1–10 (2024).
- The historical ecological background of West Nile virus in Portugal indicates One Health opportunities. *Sci. Total Environ.* **944**, 173875 (2024).
- Myer, M. H. & Johnston, J. M. Spatiotemporal Bayesian modeling of West Nile virus: identifying risk of infection in mosquitoes with local-scale predictors. *Sci. Total Environ.* **650**, 2818–2829 (2019).
- Farooq, Z. et al. Artificial intelligence to predict West Nile virus outbreaks with eco-climatic drivers. *Lancet Reg. Health Eur.* **17**, 100370 (2022).
- Heidecke, J., Lavarello Schettini, A. & Rocklöv, J. West Nile virus eco-epidemiology and climate change. *PLoS Clim.* **2**, e0000129 (2023).
- Morin, C. W. & Comrie, A. C. Regional and seasonal response of a West Nile virus vector to climate change. *Proc. Natl. Acad. Sci. USA* **110**, 15620–15625 (2013).
- Farooq, Z. et al. European projections of West Nile virus transmission under climate change scenarios. *One Health* **16**, 100509 (2023).
- Heidecke, J. et al. Uncovering temperature sensitivity of West Nile virus transmission: Novel computational approaches to mosquito-pathogen trait responses. *PLoS Comput. Biol.* **21**, e1012866 (2025).
- Brown, H. E. et al. Projection of climate change influences on U.S. West Nile Virus Vectors. *Earth Interact.* **19**, 18 (2015).
- Belova, A. et al. Impacts of increasing temperature on the future incidence of west Nile neuroinvasive disease in the United States. *Am. J. Clim. Change* **06**, 166–216 (2017).
- Barrera, R. et al. First isolation of West Nile virus in the Caribbean. *Am. J. Trop. Med. Hyg.* **78**, 666–668 (2008).
- Dupuis, A. P. 2nd, Marra, P. P. & Kramer, L. D. Serologic evidence of West Nile virus transmission, Jamaica, West Indies. *Emerg. Infect. Dis.* **9**, 860–863 (2003).
- Dupuis, A. P. et al. Serologic evidence for West Nile virus transmission in Puerto Rico and Cuba. *Am. J. Trop. Med. Hyg.* **73**, 474–476 (2005).
- Quirin, R. et al. West Nile Virus, Guadeloupe - Volume 10,2004 - *Emerging Infectious Diseases journal - CDC*. <https://doi.org/10.3201/eid1004.030465>.
- Komar, O. et al. West Nile virus survey of birds and mosquitoes in the Dominican Republic. *Vector Borne Zoonotic Dis.* **5**, 120–126 (2005).
- Cardinale, E. et al. West Nile virus infection in horses, Indian ocean. *Comp. Immunol. Microbiol. Infect. Dis.* **53**, 45–49 (2017).
- MariéJean-Lou & Leparç-GoffartIsabelle. First Serological Evidence of West Nile Virus in Horses and Dogs from Corsica Island, France. *Vector-Borne and Zoonotic Diseases* <https://doi.org/10.1089/vbz.2016.2024> (2017).
- Monaco, F. et al. The 2011 West Nile disease outbreak in Sardinia region, Italy. *Vet. Ital.* **51**, 5–16 (2015).
- Macaluso, G. et al. First neuroinvasive human case of West Nile Disease in Southern Italy: results of the ‘One Health’ approach. *Vet. Med. Sci.* **7**, 2463–2472 (2021).
- Pallari, C. T. et al. Evidence of West Nile virus seropositivity in wild birds on the island of Cyprus. *Comp. Immunol. Microbiol. Infect. Dis.* **74**, 101592 (2021).
- Ferretti, A., Maggini, I. & Fusani, L. How to recover after sea crossing: the importance of small islands for passerines during spring migration.

- Ethol. Ecol. Evol.* <https://doi.org/10.1080/03949370.2021.1886181> (2021).
41. Kilpatrick, A. M. & Wheeler, S. S. Impact of West Nile Virus on Bird populations: limited lasting effects, evidence for recovery, and gaps in our understanding of impacts on ecosystems. *J. Med. Entomol.* **56**, 1491–1497 (2019).
 42. Lovette, I. J. & Fitzpatrick, J. W. *Handbook of Bird Biology*. (Wiley, 2016).
 43. Santa-Ana, M., Khadem, M. & Capela, R. Natural infection of *Culex theileri* (Diptera: Culicidae) with *Dirofilaria immitis* (Nematoda: Filarioidea) on Madeira Island, Portugal. *J. Med. Entomol.* **43**, 104–106 (2006).
 44. *Culex pipiens* group - current known distribution: October 2023. *European Centre for Disease Prevention and Control* <https://www.ecdc.europa.eu/en/publications-data/culex-pipiens-group-current-known-distribution-october-2023> (2023).
 45. Lourenço, J. & Recker, M. The 2012 Madeira dengue outbreak: epidemiological determinants and future epidemic potential. *PLoS Negl. Trop. Dis.* **8**, e3083 (2014).
 46. Gangoso, L. et al. Prevalence of neutralizing antibodies to West Nile virus in Eleonora's Falcons in the Canary Islands. *J. Wildl. Dis.* **46**, 1321–1324 (2010).
 47. Mentaberre, G. et al. A transversal study on antibodies against selected pathogens in dromedary camels in the Canary Islands, Spain. *Vet. Microbiol.* **167**, 468–473 (2013).
 48. Biodiversity of the Gulf of Guinea Oceanic Islands: Science and Conservation. (Springer, 2022).
 49. Kamgang, B. et al. Entomological surveys and insecticide susceptibility profile of *Aedes aegypti* during the dengue outbreak in Sao Tome and Principe in 2022. *PLoS Negl. Trop. Dis.* **18**, e0011903 (2024).
 50. Yen, T.-Y. et al. Seroprevalence of antibodies against dengue virus among pregnant women in the Democratic Republic of Sao Tome and Principe. *Acta Trop.* **155**, 58–62 (2016).
 51. Franco, L. et al. Recent expansion of dengue virus serotype 3 in West Africa. *Euro Surveill.* **15**, 19490 (2010).
 52. Da Veiga Leal, S. et al. Update on the composition and distribution of the mosquito fauna (Diptera: Culicidae) in Cabo Verde, a country at risk for mosquito-borne diseases. *J. Med. Entomol.* **61**, 919–924 (2024).
 53. [Consulta de notificación de enfermedades de los animales de declaración obligatoria] - Ganadería - mapama.gob.es. <https://servicio.mapa.gob.es/rasve/Publico/Publico/BuscadorFocos.aspx>.
 54. Mingione, M., Branda, F., Maruotti, A., Ciccozzi, M. & Mazzoli, S. Monitoring the West Nile virus outbreaks in Italy using open access data. *Sci. Data* **10**, 1–11 (2023).
 55. Klitting, R. et al. Predicting the evolution of the Lassa virus endemic area and population at risk over the next decades. *Nat. Commun.* **13**, 5596 (2022).
 56. Kuhn, M. Building predictive models in using the care package. *J. Stat. Softw.* **28**, 5 (2008).
 57. Rocha, R., Kurum, E., Charrel, R., Ayhan, N. & Maia, C. Seroprevalence of West Nile Virus in blood donors in Mainland Portugal. *Trop. Med. Infect. Dis.* **10**, 229 (2025).
 58. Kilpatrick, A. M. Globalization, land use, and the invasion of West Nile virus. *Science* **334**, 323–327 (2011).
 59. Leyton Ramos, L. M., Aguirre Obando, O. A., Duque, J. E. & García-Merchán, V. H. Effect of altitude on wing metric variation of *Aedes aegypti* (Diptera: Culicidae) in a region of the Colombian Central Andes. *PLoS One* **15**, e0228975 (2020).
 60. Asigau, S., Hartman, D. A., Higashiguchi, J. M. & Parker, P. G. The distribution of mosquitoes across an altitudinal gradient in the Galapagos Islands. *J. Vector Ecol.* **42**, 243–253 (2017).
 61. Mughini-Gras, L. et al. Ecological niche modelling of potential West Nile virus vector mosquito species and their geographical association with equine epizootics in Italy. *Ecohealth* **11**, 120–132 (2014).
 62. Conte, A. et al. Spatio-temporal identification of areas suitable for west nile disease in the mediterranean Basin and Central Europe. *PLoS One* **10**, e0146024 (2015).
 63. Coroian, M. et al. Human West Nile Meningo-encephalitis in a highly endemic country: a complex epidemiological analysis on biotic and abiotic risk factors. *Int. J. Environ. Res. Public Health* **17**, 8250 (2020).
 64. Couper, L. I., Farmer, J. E., Lyberger, K. P., Lee, A. S. & Mordecai, E. A. Mosquito thermal tolerance is remarkably constrained across a large climatic range. *Proc. Biol. Sci.* **291**, 20232457 (2024).
 65. Di Pol, G., Crotta, M. & Taylor, R. A. Modelling the temperature suitability for the risk of West Nile Virus establishment in European *Culex pipiens* populations. *Transbound. Emerg. Dis.* **69**, e1787–e1799 (2022).
 66. Rehbein, M. M., Viadero, R. Jr, Hunt, J. R. & Miller, C. The role of temperature, wind speed, and precipitation on the abundance of *Culex* species and West Nile virus infection rate in rural West-Central Illinois. *J. Am. Mosq. Control Assoc.* **40**, 1–10 (2024).
 67. Paz, S. Climate change impacts on West Nile virus transmission in a global context. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* <https://doi.org/10.1098/rstb.2013.0561> (2015).
 68. Stilianakis, N. I. et al. Identification of climatic factors affecting the epidemiology of human West Nile Virus infections in Northern Greece. *PLoS One* **11**, e0161510 (2016).
 69. Magallanes, S. et al. Long-term serological surveillance for West Nile and Usutu virus in horses in south-West Spain. *One Health* **17**, 100578 (2023).
 70. Paz, S. The West Nile Virus outbreak in Israel (2000) from a new perspective: the regional impact of climate change. *Int. J. Environ. Health Res.* **16**, 1–13 (2006).
 71. Bowden, S. E., Magori, K. & Drake, J. M. Regional differences in the association between land cover and West Nile virus disease incidence in humans in the United States. *Am. J. Trop. Med. Hyg.* **84**, 234–238 (2011).
 72. Ruiz, M. O., Walker, E. D., Foster, E. S., Haramis, L. D. & Kitron, U. D. Association of West Nile virus illness and urban landscapes in Chicago and Detroit. *Int. J. Health Geogr.* **6**, 10 (2007).
 73. terra. <https://rspatial.github.io/terra/>.

Acknowledgements

This work was supported by Fundação para a Ciência e Tecnologia (FCT) in the scope of institutional support from FCT to CE3C (DOI:10.54499/UIDB/00329/2020, UID/00329/2025), BioISI (DOI:10.54499/UIDB/04046/2020, UID/04046/2025), CHANGE (LAP/0121/2020), and CBR (DOI:10.54499/UIB/06497/2025, UID/06497/2025). We acknowledge Colégio Tropical (CTROP) from the University of Lisbon for supporting a WNV research initiative in São Tomé linked to this work. MVC acknowledges institutional support from FCT (DOI:10.54499/CEECINST/00032/2018/CP1523/CT0005).

Author contributions

M.A.G.: Data Curation, Methodology, Formal analysis, Investigation, Writing - Original Draft, Writing - Review & Editing. J.L.: Data Curation, Resources, Methodology, Conceptualisation, Formal analysis, Investigation, Writing - Original Draft, Writing - Review & Editing, Supervision. M.G.: Conceptualisation, Investigation, Writing - Original Draft, Writing - Review & Editing, Supervision. M.V.C.: Resources, Conceptualisation, Investigation, Writing - Original Draft, Writing - Review & Editing, Supervision.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s42003-025-09410-z>.

Correspondence and requests for materials should be addressed to Mónica V. Cunha or José. Lourenço.

Peer review information *Communications Biology* thanks Diana Erazo and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editors: Hannes Schuler and Tobias Goris. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2026