



A climate and population dependent diffusion model forecasts the spread of *Aedes Albopictus* mosquitoes in Europe



Sandra Barman ¹✉, Jan C. Semenza ^{2,3,4}, Pratik Singh ^{2,3,4}, Henrik Sjödin ^{2,3}, Joacim Rocklöv ^{2,3,4,6} & Jonas Wallin ^{5,6}

Aedes albopictus, a key vector for Dengue, Chikungunya, Zika, and Yellow Fever, is expanding its range beyond its tropical and subtropical origins, driven by suitable climate, population mobility, trade, and urbanization. Since its introduction to Europe, *Ae. albopictus* has rapidly spread and triggered recurrent outbreaks. Past model attempts have handled vector suitability and vector introduction as independent drivers. Here we develop a highly predictive spatio-temporal vector diffusion model based on climate suitability and human population predictors, integrated in one simultaneous framework. The model explains how short- and long-range spread of *Ae. albopictus* interacts with vector suitability, predicting areas of presence or absence with high accuracy (99% and 79%). These results show that the expansion of *Ae. albopictus* in Europe is predictable and provide a basis for anticipating future outbreaks in situations of dependent interacting co-drivers.

Dengue, chikungunya, zika or yellow fever are caused by arthropod-borne viruses, transmitted by *Aedes* mosquitoes. Currently, almost 4 billion people, in over 129 countries, are at risk for *Aedes*-borne infections¹. Europe has also experienced recurrent dengue and chikungunya outbreaks, due to autochthonous transmission by invasive *Aedes* mosquitoes. To date, there have been 66 autochthonous arbovirus outbreaks on mainland Europe, with over 1200 notifications². While the disease burden cannot be compared to tropical and subtropical countries, the epidemic potential of these diseases is of concern especially in the light of climate change^{3,4}. The introduction, establishment and spread of invasive vector species has been assisted by global environmental change, specifically travel, trade, urbanisation, and climate change. Moreover, viraemic travellers returning from disease endemic countries increasingly introduce dengue and chikungunya viruses into areas with climatically suitable conditions, where competent mosquito vectors are established or re-established. These circumstances have resulted in an upsurge of local outbreaks in Europe².

The mosquito vector responsible for these European outbreaks is *Ae. albopictus*^{2,5}. It originated in tropical forests of South-East Asia and has since invaded all continents⁶. This rapid and widespread expansion has been propelled by its remarkable ecological and physiological plasticity from zoophily to anthropophily, domestic container-breeding, diurnal biting habits, cold-acclimatisation, and desiccation-resistant eggs⁷. The passive

transport of its eggs, aided by globalised travel and trade in e.g., used tires or lucky bamboo, enabled *Ae. albopictus* to invade novel habitats⁸⁻¹⁰. In Europe, *Ae. albopictus* spread along transportation corridors from heavily infested to naïve areas, through passive transport on ground vehicles¹¹⁻¹⁴.

Ae. albopictus was first detected in Albania in 1979 but might have been present there already since 1976¹⁵. It was not until 1990, when it was discovered in another European country, close to a playground in Genova, Italy, where children played with discarded tires¹⁶. Despite an attempted eradication campaign, *Ae. albopictus* became established throughout Italy, in areas below 600 m above sea level, with particularly high densities in urban centres. Subsequently, *Ae. albopictus* was detected in France in 1999, and has since expanded to a number of other European countries¹⁷. Suitable climatic conditions favoured by climate change, urbanisation, and human population mobility, seems to have facilitated the expansion of this invasive mosquito species into novel habitats^{18,19}. However, based on the current introductions and subsequent establishment of *Ae. albopictus* it is not clear what processes and drivers contribute to a successful colonisation of naïve areas and to what extent it is predictable.

The WHO global arbovirus initiative, launched in March 2022, calls for risk monitoring and the mapping of areas vulnerable for arbovirus transmission²⁰. However, defining and predicting which *Ae. albopictus*-free areas will become colonised has proven challenging²¹. Establishment of

¹Product design, Bioeconomy, RISE Research Institutes of Sweden, Göteborg, Sweden. ²Department of Epidemiology and Global Health, Umeå University, Umeå, Sweden. ³Interdisciplinary center for scientific computing, Heidelberg University, Im Neuenheimer Feld 205, Heidelberg, Germany. ⁴Heidelberg Institute of global health, Heidelberg University, Im Neuenheimer Feld 205, Heidelberg, Germany. ⁵Department of statistics, Lund University, Lund, Sweden. ⁶These authors contributed equally: Joacim Rocklöv, Jonas Wallin. ✉e-mail: sandra.barman@ri.se

mosquitoes in an uncolonized area, depends in part on the origin of the imported mosquito species, but also on climatic and environmental conditions, population mobility, and demographic factors^{6,21,22}. Here we present a dynamic, spatio-temporal diffusion model with human population network structures that in combination with the influences of climatic predictors and process-based model predictions of mosquito dynamics, forecasts the newly invaded areas of *Ae. albopictus* in Europe. The model reliably separates and integrates the influence of local vector suitability with vector introduction and predicts the continuing European invasion by first time (non-recurrent) establishment of *Ae. albopictus*. The model framework developed lends itself to be used for public health purposes in Europe and beyond to prepare for the introduction of *Ae. albopictus*, or other similar challenges where the processes of suitability and introduction need to be considered simultaneously. Here we describe the drivers and discuss how insights from such predictions can guide public health efforts and reduce the epidemic risk in unaffected areas where the local population has no immunity to *Aedes*-borne diseases. Predictions can help target awareness and prevention messages to susceptible populations and guide vector control efforts to counteract establishment of vector populations and local outbreaks. Moreover, it can help prepare the healthcare system for epidemics and target novel interventions to areas at risk.

Results

We have developed a spatio-temporal model with a diffusion process that models the spread of *Ae. albopictus* from regions with presence to nearby absence regions. The model predicts the introduction of *Ae. albopictus* into naïve regions of Europe, during the period 2010–2023 (Fig. 1). We integrate the diffusion model with covariates determining the suitability of *Ae. albopictus* establishments using climate and human population data.

The model is a generalised additive mixed (GAM) model with non-linear random effects, which is fitted within the approximated Bayesian spatio-temporal framework INLA based on annual data by NUTS3 areas. We predict a year ahead using a model trained on previous years data. The overall model performance is good, with high area under the receiving operating characteristic curve (AUC) values of 0.99 when evaluated on all observations independent of time of establishment of the vector. It is, however, easier to model the probability of *Ae. albopictus* presence re-occurring in a region as the recurrence rate in the same region is 99.8%. Thus, it is relevant to show model performance statistics only for new observations in a specific year, i.e., region-year locations where *Ae. albopictus* was not present the previous year. We show that the accuracy of our model is high with an AUC of around 0.80 for rolling window cross-validation and AUC 0.73–0.86 for fixed window cross-validation (Table 1). The receiver operator characteristic curve used to estimate rolling window cross-validation AUC values are presented in Fig. 2. More details about model fit are given in Supplementary Table S8.

The GAM model is fitted with a spatio-temporal stochastic partial differential equation (SPDE) term which accounts for extraneous variables not included in the model. A short-range SPDE is used, which (1) is conservative in that a large portion of the fitted values are essentially zero, (2) has a low correlation with the data and (3) has the effect that more weight is put on the covariates which improves the model fit on out-of-sample data (see Supplement, Controlling the SPDE range to improve model predictions).

We calibrate two different GAM models with climate and human population covariates considered in different ways. In one of the models the raw climate (median temperature, minimum temperature, and median relative humidity) and human population covariates are taken into account as non-linear random effects. In the other model these covariates are transformed by a mechanistic mosquito life cycle-model predicting adult *Ae. albopictus* abundance. The predicted adult abundance is then entered into the GAM model as a covariate to which a non-linear random effect. The diffusion process modelling the spread of *Ae. albopictus* from regions with presence to nearby regions, termed the proximity covariate, is included in both versions of the model. Summary statistics for the fit from each of these two models (termed the climate and population- and the mechanistic life

cycle-GAM model, respectively) are presented in Table 1, and fitted non-linear random effects of covariates are shown in Figs. 3–4.

The predictive performance of the model is further broken down by year, by distance to regions with recorded *Ae. albopictus* presence, and by geographical region with slightly reduced predictive performance for some of the subsets, for example, with shorter distance from established areas (Table 2). When broken down by geographical region, the performance is more distinctly reduced for the south-west geographical region -The Iberian peninsula - which has a relatively high number of regions with status “no data” (Fig. 1a).

The fitted random effects can be compared with empirical relationships between covariates and *Ae. albopictus* presence in Figs. 4 and 5. See also Supplementary Tables S1–S7 and Supplementary Fig. S3 for more details related to Figs. 4 and 5.

We find that the mechanistic life cycle GAM model performs similarly on out-of-sample data compared to with the climate and population GAM model (Table 1). The combined random effects of the climate and human population covariates (Fig. 3a–c, Fig. 4a–d) can be compared to the random effect of the mechanistic life cycle covariate (Fig. 3e, Fig. 4f), as the underlying mechanistic mosquito life cycle-model takes both human population and climate data as input²³.

Climate covariates

The raw climate covariates (median temperature, minimum temperature, and median relative humidity) were binned into relatively few bins to reduce overfitting and to reduce the dependence between the covariates. All but the minimum temperature covariate was computed over the warmer months April–October, and minimum/median values were taken over 30-day averages. The non-linear random effects show the strongest effects overall for median temperature (Fig. 3a, Fig. 4a). The fitted random effect increases at median temperatures up to 24 °C, to then level out and start to decrease at temperatures above 25 °C (Fig. 4a).

We included a random effect of minimum temperature conditioned on high median temperature (Fig. 3b, Fig. 4b). This since the random effect for minimum temperature shows a positive correlation with *Ae. albopictus* presence when conditioned on high median temperatures (median temperature $\geq 19^\circ$), and a negative correlation when conditioned on low median temperatures (Fig. 4b). The empirical *Ae. albopictus* presence supports this conditioned relationship (Fig. 5b). The distribution of *Ae. albopictus* can be seen to follow a pattern that correlates with both median and minimum temperature, having been present for the longest along coastal areas around the Mediterranean (Fig. 1a) which have both high median and minimum temperature.

Ae. albopictus presence is negatively correlated with low levels of relative humidity (Fig. 3c, Fig. 4c). Regions in Spain have relatively low rates of presence, given the risk posed by high temperature and proximity to regions with presence.

Proximity and human population covariates

We hypothesised that *Ae. albopictus* spreads to regions that are geographically close and through human mobility to highly populated regions farther away. We used a spatio-temporal diffusion process to estimate the rate of spread to regions that are geographically close (proximity covariate), and the logarithm of human population as a covariate to capture mobility to highly populated regions further away. The fitted random effects show a substantial impact of geographical proximity through the spatio-temporal diffusion process (Fig. 3e, Fig. 4e) and there is a clear effect of *Ae. albopictus* presence being higher in regions with higher population (Fig. 4d), supporting the hypothesis. Although there is a clear effect of the population covariate, the size of the effect is relatively small compared to the fitted random effects of other covariates. The result that the random effect is relatively small for the population covariate, is also seen when comparing the raw population data with empirical *Ae. albopictus* presence (Fig. 5). There is a clear trend related to population here (Fig. 5d), but the trend is relatively small in size compared to that found for the other covariates (Fig. 5a–c, e, f).

We additionally tried a GAM model including human mobility using a radiation model in the diffusion process. This did not result in any notable improvement in model fit though, and so this GAM model is

not included in the results presented here. Instead, the local diffusion and human population covariate was used to efficiently predict the new establishments.

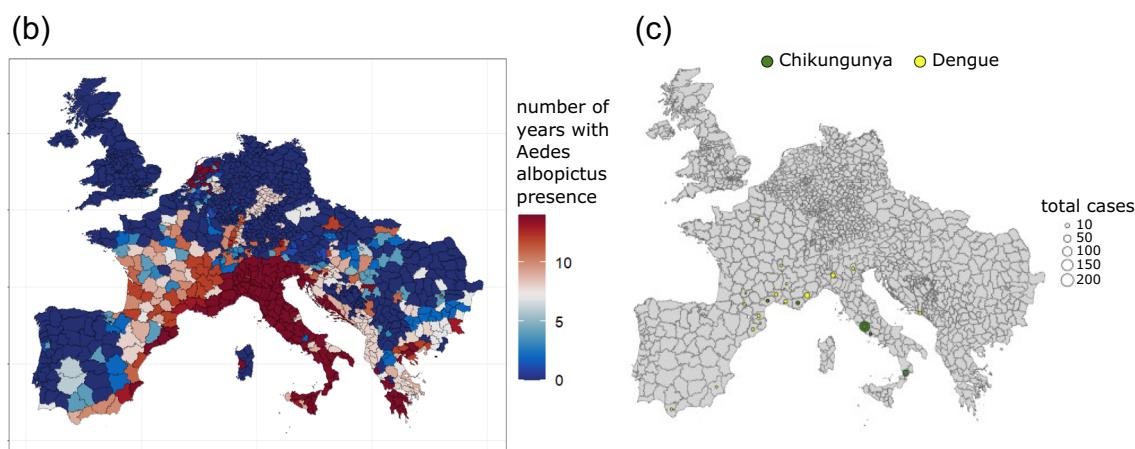
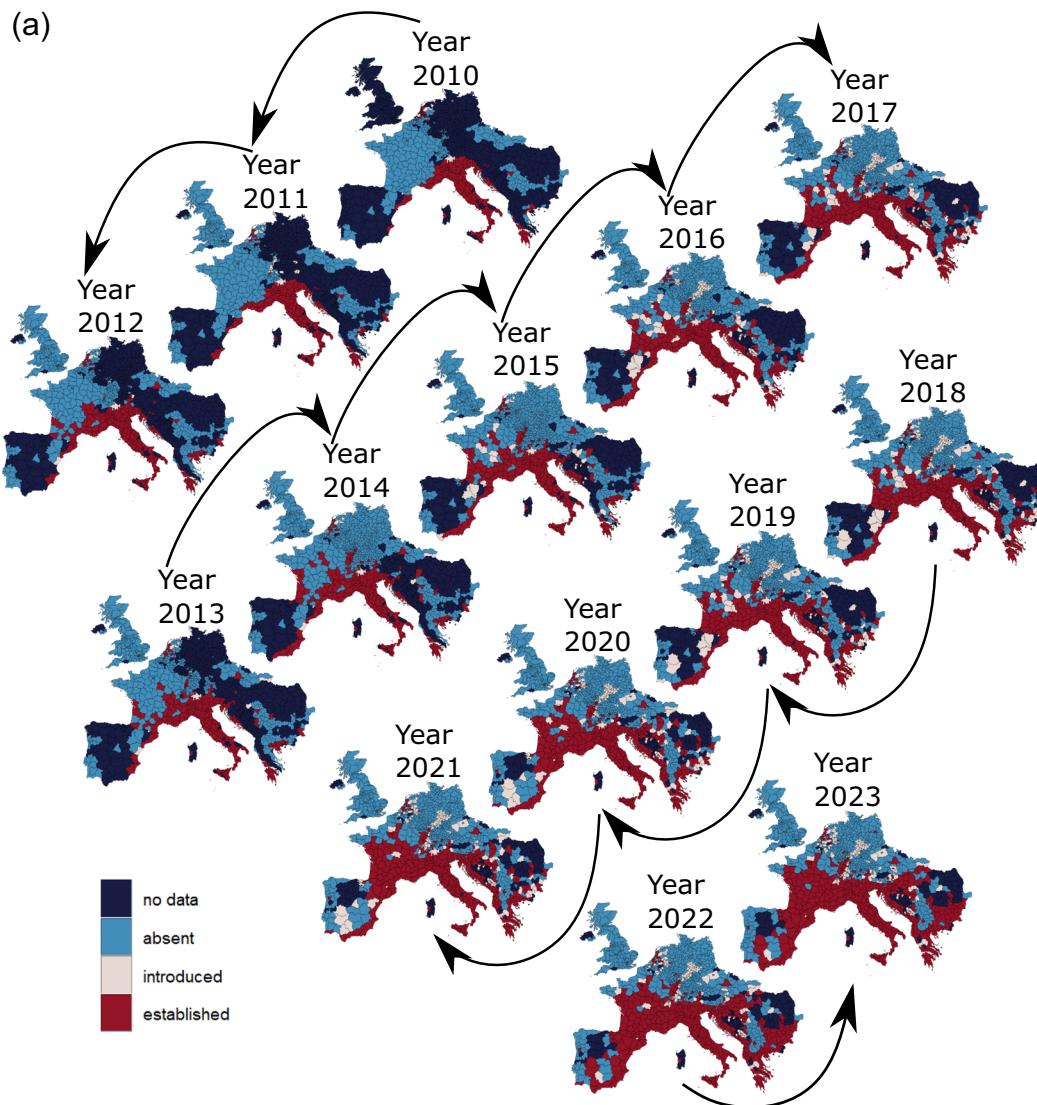


Fig. 1 | *Ae. albopictus* observations and local outbreaks of Dengue and Chikungunya. **a** VectorNet dataset of observations of *Ae. albopictus*, from 2010 to 2023, coloured by *albopictus* status: no data; absent; introduced; and established,

b accumulated *Ae. albopictus* establishment over the years 2010–2023 and **c** local outbreaks of Dengue and Chikungunya.

Table 1 | Model fit for the mechanistic life cycle and climate /population GAM models

GAM model	Rolling window		Fixed window	
	Deviance, 2017–2023	AUC, 2017–2023	AUC 2013–2017	AUC, 2019–2023
Mechanistic life cycle covariate set	190	0.79	0.73	0.84
Climate and population covariate set	191	0.80	0.73	0.86

Comparison of the model fit for the two sets of covariates. Both covariate sets include the proximity, presence previous year and first year covariate, see Section Methods. For rolling window cross-validation, deviance and AUC are computed on test data over seven cross-validation test sets, withholding one year ahead at a time for each of the years 2017–2023. For fixed window cross-validation, AUC is computed on joint prediction over five consecutive years (2013–2017 and 2019–2023). Only regions with confirmed *Ae. albopictus* presence or absence, but no recorded presence previous years are counted. Bold font indicates the value with the best fit.

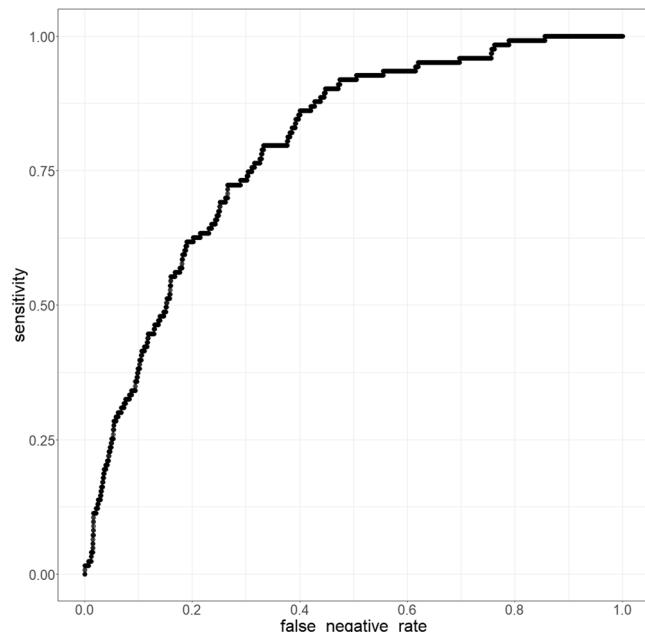


Fig. 2 | Model fit for the mechanistic life cycle GAM model. Receiver operator characteristic curve for the mechanistic life cycle GAM model, with region-year locations being left out one year ahead, with each of the years 2017 to 2023 being left out in turn. Only regions with confirmed *Ae. albopictus* presence or absence, but no recorded presence previous years are counted.

Mechanistic life cycle covariate

The fitted random effect for the mechanistic life cycle covariate (Fig. 3e) shows similar effects for regions close to the Atlantic and around the Mediterranean Sea as discussed in relation to the climate covariates. Further, there is a relatively high risk of *Ae. albopictus* presence in the northern part of Europe, namely in the Low Countries (Belgium, Luxembourg and the Netherlands), and in parts of Germany and northern France. Comparing the fitted random effect from the mechanistic life cycle covariate (Fig. 3e) and the accumulated *Ae. albopictus* presence over the full dataset (Fig. 1b), we see some striking similarities. Note the *Ae. albopictus* presence in Berlin in year 2023, which is the first noted presence in the northern part of Germany.

Discussion

The introduction, establishment, and expansion of *Ae. albopictus* in Europe are the primary causes of recurrent dengue and chikungunya outbreaks². Monitoring this expansion, through labour- and time-intensive vector surveillance has proven to be cumbersome and costly²⁴. However, our study demonstrates that this expansion is predictable, based on a vector population spatial diffusion process, with climate and human population covariates depicting suitability. This innovative model performs well, also on data withheld from model fitting, and will prove to be a valuable tool for preparedness and response to *Aedes*-borne infections. Notably, model

evaluation reveals that new introduction of *Ae. albopictus* into naïve areas, are very well predicted, which has not been achieved before with this type of model. This model can be integrated into early warning systems and help delineate areas at risk for the introduction and establishment of *Ae. albopictus*. These areas can then be targeted for vector control interventions to prevent potential outbreaks²⁵. Thus, our model can accelerate the detection of arboviral outbreaks by helping to focus limited resources to high-risk areas and thereby accelerate the public health response, which includes enforcing donor blood deferral and supplementing screening of the blood banks in areas at risk. Moreover, insights from our model could support strategic planning for the deployment of vaccines against arboviral diseases. Currently, there are three arboviral vaccines available in Europe (dengue vaccines: Dengvaxia and Qdenga; and chikungunya vaccine: Ixchiq) that could be deployed in at risk areas prior to an outbreak. Our model can help public health authorities prioritise vaccination campaigns to mitigate the impact of diseases like chikungunya and dengue in vulnerable populations. While there is no evidence of Zika virus transmission in Europe to date, these strategies would apply to all pathogens transmitted by *Ae. albopictus* in Europe.

We find that temperature, relative humidity, long-distance mobility and short-distance diffusion is enough to predict presence/absence and more interestingly the spatio-temporal patterns of new establishment of *Ae. albopictus* in Europe. In fact, for the first time, the models document an ability not only to predict presence/absence, but the changing risk landscape by predicting areas at risk for new establishment which is a more relevant outcome for timely early warnings and adaptation priorities. We calibrated two different models using the spatio-temporal Bayesian INLA modelling framework describing the patterns of the new establishments in Europe. The first model uses simply climate covariates, but the second model uses a climate-driven prediction from a mechanistic mosquito life cycle model as covariate. Both models include a diffusion process capturing how new establishments spread from regions with *Ae. albopictus* presence. We find both models provide reliable predictions on data withheld from model fitting with a classification error of around 20% when applied to naïve regions not previously colonised by *Ae. albopictus*.

In detail, we incorporate climate and demographic components in the model, such as relative humidity, temperature, rainfall, and human population data. We introduce a continuous spatio-temporal diffusion process for vector introductions from regions with *Ae. albopictus* presence, mainly contributing to the spread at shorter distances. Human mobility, which can be a key factor in the spread at both short and long distances, is included implicitly with a human population covariate. These covariates perform well together and show that all of these factors are important for the establishment of vectors in new areas, as is the mechanistic mosquito life cycle covariate which captures all of these processes in one covariate.

The finding that the mobility diffusion did not add notably to the model performance fits with previous observations⁶, where including mobility was found to only give slight improvements in model fit. In our analysis we therefore chose the simpler model where the diffusion process (proximity covariate) and human population are included as separate covariates. In a way, we can say that we cannot show any interaction effects between geographical proximity and human population (which is captured by the mobility diffusion), but that when including both proximity and

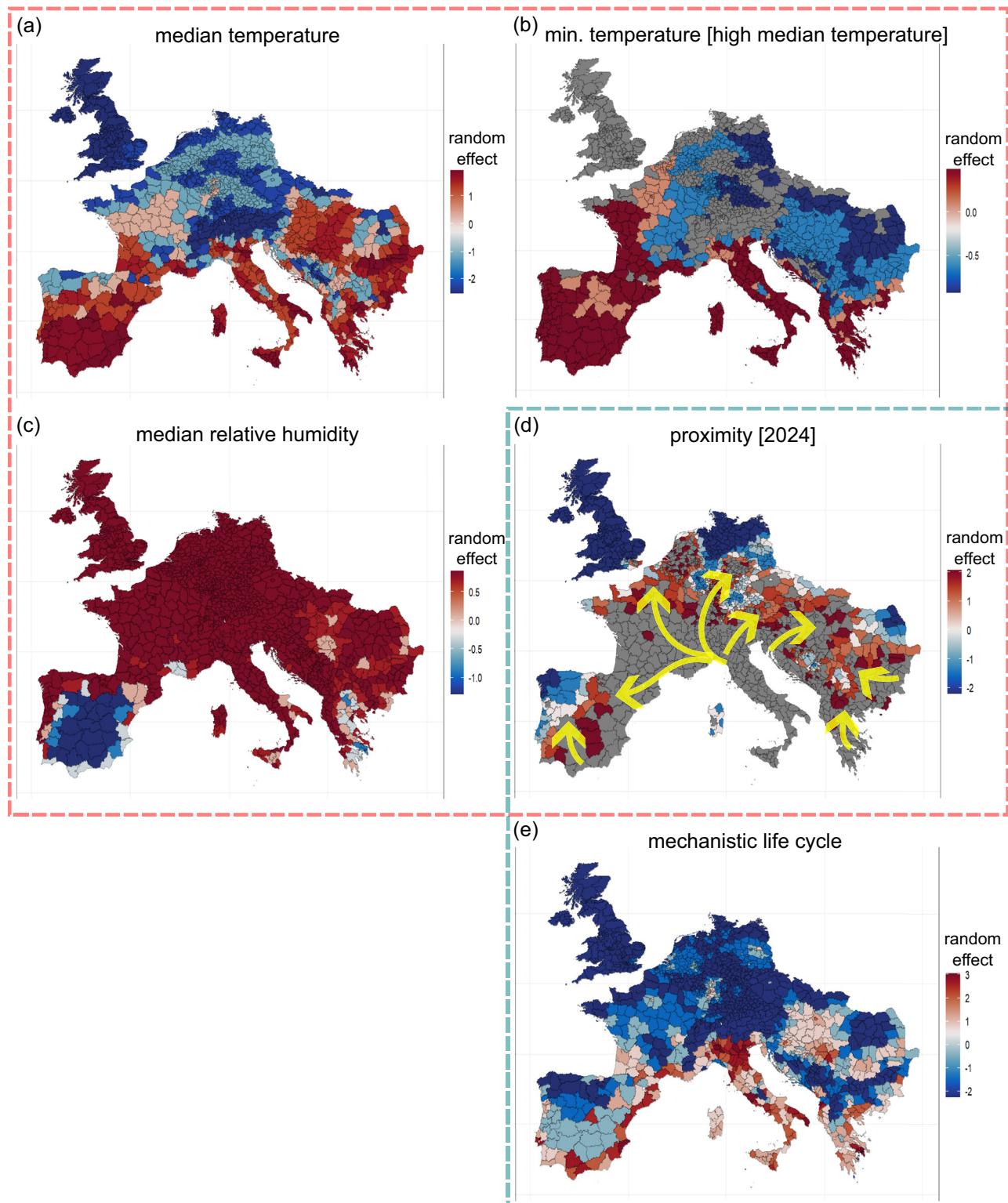


Fig. 3 | Model contribution from each covariate, geographical form. Random effect shown on geographical form, for a the median temperature; b the minimum temperature, conditioned on high median temperature; c the median relative humidity; d the proximity; and e the mechanistic life cycle covariate, showing the contribution from each covariate to the GAM model. The proximity covariate is different from year to year, it is here shown at year 2024. Regions in grey in the

proximity map are regions that have recorded *Ae. albopictus* presence in 2023. Yellow arrows illustrate visually how the proximity covariate models risk of *Ae. albopictus* spreading from regions with presence to nearby regions with no presence. The maps indicated by the red and green dashed lines show covariates included in the climate and population GAM model (a-d) and the mechanistic life cycle GAM model (d, e), respectively.

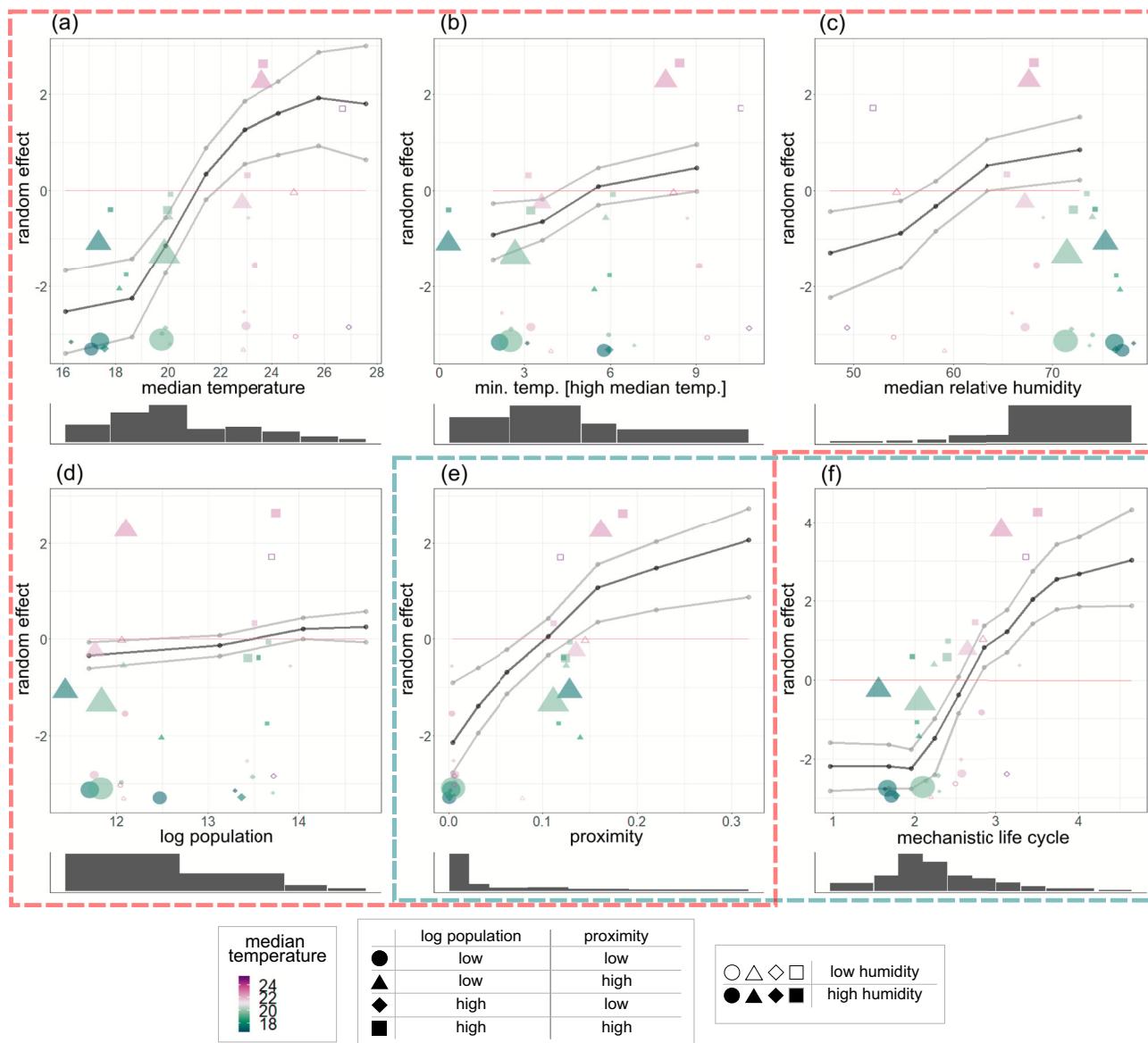


Fig. 4 | Model contribution from each covariate, functional form. Random effect for **a** the median temperature; **b** the minimum temperature, conditioned on high median temperature; **c** the median relative humidity; **d** the log human population; **e** the proximity; and **f** the mechanistic life cycle covariate, showing the contribution from each covariate to the fitted GAM model. Black lines indicate the random effect, and grey lines indicate the lower and upper limits of 95% credible intervals. Bar plots below each panel show the number of region-year locations within each covariate bin, see Supplementary Methods, GAM modelling procedure for details. Scatter

plots of empirical *Ae. albopictus* ratios, by a five-way grouping of covariates are super-imposed on the plots. Group category is indicated by marker type and colouring in the scatter plots, see the legend, and the size of the marker is proportional to the number of region-year locations in the corresponding group. For more detailed information about the scatter plots see Supplementary Table S7 and Supplementary Fig. S3. The panels indicated by the red and green dashed lines show covariates included in the climate and population GAM model (a–e) and the mechanistic life cycle GAM model (e, f), respectively.

human population as covariates they can together explain the effect of human mobility on new establishments.

Visual inspection of the accumulated *Ae. albopictus* presence indicates a clear effect of human mobility when comparing the presence in large cities (Berlin, London, Madrid, Porto, Prague and Thessalonique) with their surrounding regions (Fig. 1b, Supplementary Figs. S1–S2). Thus, there seems to be a clear effect of human population size and mobility in the raw data, though with the metrics we are using the measured effect is relatively small.

Interestingly the mechanistic mosquito life cycle covariate, with the underlying process-based life cycle model driven by climate and human population, does capture the patterns as well as the more flexible climate covariates. The same process-based model used is frequently applied to larger climate assessments in the *Lancet* countdown on climate change and

health^{26–28}. These findings provide insights into the reliability of such more complex process-based models and provide proof of its validity.

The approach presented here is using the most recent data in a statistical framework that quantifies the contribution of known factors to the spread of *Ae. albopictus*. The work further shows that human population and climate along with a spatio-temporal diffusion can be integrated in a simultaneous modelling framework and that it can robustly predict the establishment through accounting for vector suitability and vector introduction processes. Compared to previous prediction models, the present model is reducing the model complexity substantially, yet has similar model performance (AUC: 0.73–0.86). More specifically, the model performance seen in Table 1 for fixed window cross-validation (AUC 0.73, training (test) set 2010–2012 (2013–2017), AUC 0.84–0.86, training (test) set 2010–2018 (2019–2023)) is similar to that found in previously, where AUC values up to

Table 2 | Model fit for models restricted in time/space

(a)		AUC
2017		0.77
2018		0.75
2019		-
2020		0.86
2021		0.75
2022		0.88
2023		0.74

(b)		AUC	# regions
50.000		0.73	438
100.000		0.75	523
150.000		0.75	570
200.000		0.77	606
500.000		0.79	683

Region	Time period	AUC	# regions
north-east	2017–2020	0.84	208
north-east	2020–2023	0.85	208
north-west	2017–2020	0.79	496
north-west	2020–2023	0.81	496
south-east	2017–2020	0.72	429
south-east	2020–2023	0.64	429
south-west	2017–2020	0.60	202
south-west	2020–2023	0.69	202

AUC reported for subdivisions of the data into year/geographical region for the mechanistic life cycle GAM model. AUC was computed on test sets withholding one year ahead for each of the years 2017 to 2023. Only regions with confirmed *Ae. albopictus* presence or absence, but no recorded presence previous years are counted. (a) AUC for each of the years 2017–2023. 2019 had no new observations of *Ae. albopictus* presence. (b) AUC for the years 2017–2023, divided by distance to regions with recorded presence the previous year. (c) AUC by geographical region and by time period, with north: latitude >=50, south: latitude <50, east: longitude >= 10, west: longitude <10.

0.82 (training (test) set 1979–2012 (2013–2017)) were obtained for models fitted to *Ae. albopictus* data from Europe (see Supplementary Fig. 7(B)⁶). Note that better results, with AUC up to 0.91, were obtained in that work for a model fitted to a larger US dataset and used to predict European data⁶, Supplementary Fig. 8. We note additionally, that the new model and predictions provided here (Fig. 3) predicts correctly several areas which was predicted not at risk of expansion, or as missing, in the previous work, for example, in the region around Paris (Fig. 3⁶). Although past modelling attempts of *Ae. albopictus* in Europe used climate and human population as well as a larger dataset of environmental and social data⁶, the nature of the previous model didn't include statistical inference in terms of the estimated explicit relationships between model components and risk of new *Ae. albopictus* establishments. Especially, there is a risk of covariate bias as it lacked a simultaneous estimation of the key drivers of vector suitability and vector introduction in one model framework. We additionally explore and describe both empirical and fitted non-linear relationships between each model component and *Ae. albopictus* presence, providing insight into the dynamics of new establishments. This can help to better understand and predict emergence and by doing so guide future surveillance efforts.

Additionally, to the insights, the model framework and model may lend itself for early warning systems predicting the emergence of *Ae. Albopictus*, and other threats, in naïve areas. It could also be deployed as a climate-service using long-term projections or inter-annual early warnings to enhance public health preparedness within Europe²⁹.

Limitations of inferences drawn from the model come partly from *Ae. albopictus* surveillance, where we have regional differences in surveillance

efforts across countries, and where we expect more effort close to regions with recorded presence as well as in more densely populated regions. From the modelling perspective, we have rather coarse data with presence/absence on a yearly basis and on the NUTS3/GAUL spatial resolution, which limits the inference. On the other hand, the study does not make use of pseudo-absence observations which may introduce bias.

Regarding the limitation of poor data quality in some regions, such as on the Iberian peninsula where many regions have status “no data”, this could explain the relatively poor performance of the model in the south-west geographical region (Table 2c). The model interpolates information from regions with similar conditions when making predictions for regions with “no data”. However, if the patterns in the unseen data are different from those regions with similar conditions where there is data, this can introduce bias in the model predictions. Given more data on the Iberian peninsula, e.g., this could increase the model performance for regions such as these with high temperature and low humidity. However, entomological data gaps can be complemented with Citizen Science activities, such as MosquitoAlert, a mobile app which has proven to be reliable at collecting timely and large-scale data on the presence and distribution of mosquito vectors, such as *Ae. albopictus*. Citizen scientists can participate in research projects and register mosquito nuisance, records of bites, breeding sites and even upload photos of targeted mosquito species, to be validated by expert entomologists. These data can be incorporated into our model and be used for validation purposes^{30–33}.

In conclusion, our approach, using the most recent data and a streamlined statistical framework, offers several advantages: updated vector

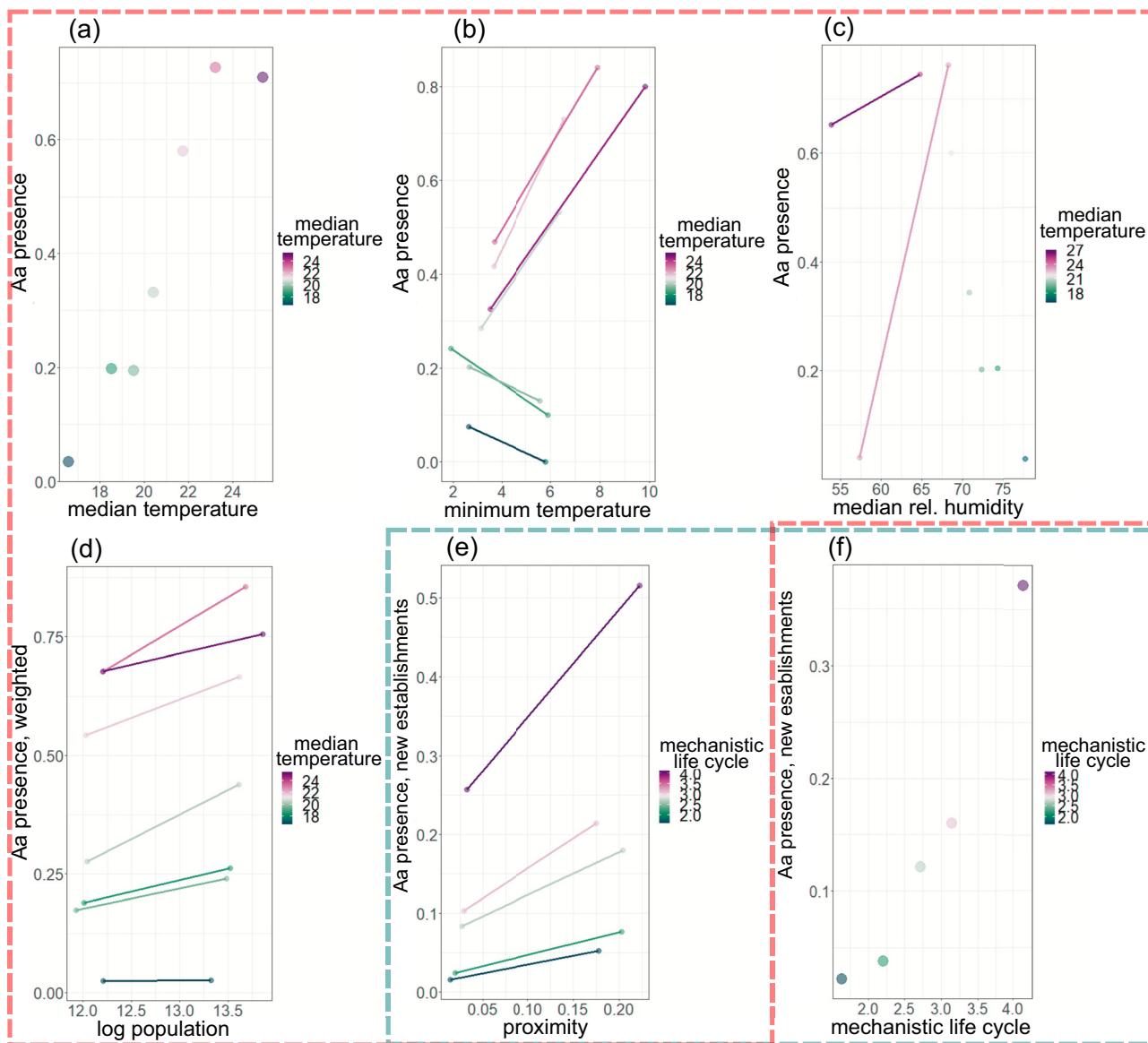


Fig. 5 | Empirical ratio of *Ae. albopictus* presence. The empirical ratio of *Ae. albopictus* presence for sub-groups by (a) the median temperature; (b) the median temperature crossed with minimum temperature; and (c) the median temperature crossed with the median relative humidity; (d) the log human population crossed with the median temperature; (e) the mechanistic life cycle covariate crossed with the proximity; and (f) the mechanistic life cycle covariate. The groups and data used to create these plots are provided in Supplementary Tables S1–S6, with each table corresponding to one of the (a–f). The same median temperature groups have been

used in (a–c), and the same mechanistic life cycle covariate groups have been used in (d, e). Pairs within the same median temperature or mechanistic life cycle covariate group are linked with lines. The sub-groups in (a–d) include the whole dataset, whereas the sub-groups in (e, f) include only new establishments. The panels indicated by the red and green dashed rectangles show covariates included in the climate and population GAM model (a–e) and the mechanistic life cycle GAM model (e, f), respectively.

surveillance data, a simpler model structure, integration into a coherent simultaneously modelling framework capturing both vector suitability and vector introduction processes, statistical inference capabilities, and a focus on predicting new establishments. This model can serve as an early warning system, predicting the emergence of *Ae. albopictus* in new areas and enhancing public health preparedness in Europe without having to rely on vector surveillance data³⁴. Additionally, this framework could be adapted for use in other regions and for other vector-borne diseases³⁵.

Methods

Dataset

***Aedes albopictus*.** The data comes from the project VectorNet, funded by the European Food Safety Authority and the European Centre for Disease Prevention and Control. The project collects data from different

sources, and records confirmed observations of presence (established or introduced), absence and unknown/no data status of *Ae. albopictus* within administrative units in European countries and within some countries outside Europe. For countries belonging to the European Union (EU), the administrative units are NUTS (Nomenclature of Territorial Units for Statistics) regions, and for the countries outside the EU the administrative units are GAUL (Global Administrative Unit Layers) regions. We have chosen to work with those countries in Europe where there has been at least one observation of *Ae. albopictus* presence.

When cleaning the data, some observations had no associated geometry. This was solved by downloading maps of regions from NUTS version 2010. Some regions in the joined map were overlapping. Most often, this was with one set of regions having unknown/no data status and the other set having observations of presence or absence of *Ae. albopictus*. In Greece,

however, some regions were located on top of each other with contradicting data of presence and absence. In this case, we choose the observations with presence over absence. The full dataset is shown in Fig. 1a, b.

Climate. Historical climate data for the year 2010–2023 were obtained from the Copernicus Climate Data Store³⁶). Near-surface air temperature (°C) and total precipitation (mm) are obtained from the fifth generation of the European Centre for Medium-Range Weather Forecasts atmospheric reanalyses (ERA5). The gridded ERA5-land hourly climate dataset³⁷ is converted to daily estimates at a resolution 0.25×0.25 degree which approximately covers an area of 25 km^2 . Daily mean, minimum, maximum temperature and daily total rainfall were extracted from hourly ERA5-land dataset. All temporal and spatial data aggregation processes were conducted by utilising the Climate Data Operators (CDO) software³⁸. The human population density data and population count is obtained from GPWv4 data set for the year 2010, 2015, and 2020³⁹. The population density of the rest of the years in between 2010–2020 and 2021–2023 is retrieved using linear interpolation and extrapolation respectively.

Human population. The human population data that was used as a covariate comes from Eurostat⁴⁰), complemented with population for GAUL regions in the dataset which do not have population data available from Eurostat.

Model

We use a generalised additive mixed model with a spatio-temporal component⁴¹ and covariates presented below. The model captures both non-linear and linear effects related to the covariates, and a spatio-temporal Matérn field formulated as an SPDE⁴² is used to control for extraneous variables. The model is constructed as follows: For region i and year t we denote *Ae. albopictus* presence/absence as $y_{i,t} = 1$ for presence (introduced or established status) and $y_{i,t} = 0$ for absence. Any locations (i, t) with no observation (denoted no data in Supplementary Fig. S1) were excluded from the model fitting. The model can be written as

$$y_{i,t} \sim \text{Bernoulli}\left(p_{i,t}\right), i = 1, \dots, n_{\text{regions}}, t = 2010, \dots, 2023$$

with $p_{i,t}$ the probability of *Ae. albopictus* being recorded present. The log-odds are modelled as a function of the covariates, as

$$g(p_{i,t}) = \mathbf{v}_{i,t}^\top \boldsymbol{\beta} + \sum_h f_h(w_{h,i,t}) + f_s(i, t).$$

Here g is the logistic link function, \mathbf{v} are covariates that enter the model as linear effects, w are covariates that enter the model as non-linear random effects. The non-linear random effects f_h , and the SPDE-based process f_s , are described in the Supplementary Methods, GAM modelling procedure.

For the covariates, continuous covariates were binned by first dividing the covariate into evenly spaced intervals. Bins with few locations (i, t) were merged in a second step to make sure that each bin had a sufficient amount of locations to make model fits relatively stable, meaning that the fitted random effect was consistent between models with different covariates (see Supplementary Fig. S6 for an example). In cases where it is reasonable to expect the effect to be increasing, the number of bins were reduced in a final step until the fitted random effect was reasonably increasing.

Here follows a description of the covariates.

Mechanistic mosquito life cycle-model. We estimated the mosquito abundance of adult *Ae. albopictus* using a stage-structured, climate-data driven dynamic mechanistic model based on the life cycle of dengue vector's life stages (see Supplementary Methods, Mechanistic life cycle model details, Supplementary Table S11)^{43–45}. The mechanistic mosquito

life cycle-model compute the population-density (number of mosquitoes per hectare) of mosquitoes independently in each grid cell with the assumption of well-mixed mosquito populations within 0.25×0.25 latitude longitude grid-cells⁴⁶. Daily time steps simulations are used for solving the model and then we aggregated the output to yearly values. To account for diurnal temperature variation, simulation for a single day is divided into 100-time steps according to numerical solver (deSolve4 package in R) as 0.14, 0.19, ..., 24.00, and temperature at each of these hours is used to simulate the development and mortality rates for one day. For more details on simulation of model and a simple example model simulation code (in octave v4.2.1), see Metelmann et al.⁴⁷. Lastly, spatial aggregation at the NUTS3 (Europe) level was executed in R version 4.1 utilising the rasterR6 package⁴⁸.

Climate. The climate data, in addition to being used to compute the mechanistic mosquito life cycle- covariate, used in the mosquito life cycle GAM model, was also used directly as covariates in the climate and population GAM model. The covariates are constructed from 30-day moving average of precipitation and temperature, and from which we compute the yearly median precipitation, median temperature, and minimum temperature denoted simply *median precipitation*, *median temperature* and *minimum temperature*. These climate summaries are aggregated to the NUTS/GAUL regions by taking the mean value over each region.

Human population. Human population plays an important role in the spread and habitat suitability of *Ae. albopictus*⁴⁶. We use the total population of each NUTS/GAUL region, transformed to a logarithmic scale as covariate.

Presence previous year. If *Ae. albopictus* was recorded as present in a region at year t , then in 99.8% of the times it was also recorded as present at year $t + 1$. Thus, it is simple to model the probability of *Ae. albopictus* re-occurring in a region for this dataset. Re-occurrence is modelled here by adding a covariate *presence previous year*, which at year $t > 1$ for region i equals one if *Ae. albopictus* was present in region i the previous year, and zero otherwise. It equals zero for all regions at the first year $t = 2010$.

First year. The first year, 2010, is different from the later years in the dataset in that it includes an accumulated presence of *Ae. albopictus* from previous years. To account for this difference, we include the first year as a covariate.

Proximity. To model how *Ae. albopictus* spreads from regions with presence to neighbouring regions, we developed a spatio-temporal covariate that has the property of diffusing out from regions with *Ae. albopictus* presence. Specifically, we define a Gaussian random field over a network and set the proximity covariate to the conditional expectation of the random field given observations of *Ae. albopictus* presence the previous year. The chosen network-structure (Supplementary Fig. S4, left and Fig. S5) then controls how the risk of *Ae. albopictus* presence is diffusing out from regions with presence.

Proximity with mobility. We used a radiation model based on diffusion dynamics⁴⁹, to account for how human mobility impacts the spread of *Ae. albopictus*. In particular, we used a version of the radiation model which corrects for finite systems⁵⁰. The radiation model takes as input the distance between regions and their total human population and gives the commuting flow between all pairs of regions. We extracted a mobility network by only keeping the flow that is higher than a specified threshold. Mobility links corresponding to high commuting flow were added to the proximity network (Supplementary Fig. S4, right), which was then used to create a proximity with mobility-covariate. More details are given in Supplementary Methods, Constructing the proximity covariate.

Model setup. We use two combinations of covariates. The first set consists only of the mechanistic covariate. The second set, termed the climate covariates, consists of the median temperature, the minimum temperature (conditioned on high median temperature), the median relative humidity and the log human population covariate. These are combined to form two GAM models, with both models including the proximity-, presence previous year- and first year-covariate.

Model fitting

The GAM models were fitted in a Bayesian setting using the R package R-INLA⁴¹, available at www.r-inla.org. Cross-validation was used to assess model fit. Regions where *Ae. albopictus* has been observed in previous years were excluded when computing model performance, with the exception of the AUC 0.99 mentioned early in the results. Rolling window cross-validation with one year ahead in time being left out, from 2017 until 2023, was used. Summary statistics for the model fits were calculated as the average over the test sets in the case of DIC and deviance. A receiver operator characteristic curve and area under the curve (AUC) statistics were computed on the joint set of regions left out over the test sets, encompassing seven test sets for the years 2017–2023 (Fig. 2). Summaries are listed in Table 1 and Supplementary Table S8.

To compare with previous work⁶, we also include fixed window cross-validation with the test set consisting of five years following the years included in the training set. We do this with two sets of training sets (test sets): 2010–2012 (2013–2017) and 2010–2018 (2019–2023). The probability $p_{i,t}$ of observing *Ae. albopictus* is computed using the fitted random effects, which include the proximity. We need presence/absence data from the previous year to compute the proximity covariate for the following year. This means that for this cross-validation setup, we need to use Monte Carlo sampling by simulating presence/absence one year at a time to estimate the probabilities $p_{i,t}$ of observing *Ae. albopictus*. For t_1, \dots, t_5 , the five test set years, we use Monte Carlo sampling to compute estimated probabilities \hat{p}_{i,t_j} , $i = 1, \dots, n_{regions}$, $j = 1, \dots, 5$, of observing *Ae. albopictus*. First, simulating *Ae. albopictus* presence/absence \hat{y}_{i,t_1} for the first year t_1 is done using the proximity covariate computed from presence/absence y_{i,t_1-1} , which is included in the training set. For the second year t_2 , the proximity is computed from the simulated \hat{y}_{i,t_1} , which allows us to compute p_{i,t_2} using the fitted GAM model, based on which we simulate presence/absence \hat{y}_{i,t_2} . Repeating for t_3, \dots, t_5 , we get simulated presence/absence $\hat{y}_{i,t_1}, \dots, \hat{y}_{i,t_5}$ for all years in the test set. Repeating the procedure 100 times, probabilities $\hat{p}_{i,t_1}, \dots, \hat{p}_{i,t_5}$ are then estimated as the average of the simulated presence/absence.

The range- and precision-parameters of the SPDE were explored in terms of the deviance on test sets (see Supplementary Methods, Controlling the SPDE range to improve model predictions). When assessing model fit, we found that the models performed better when restricting the range of the spatio-temporal SPDE. Comparing model performance for different choices of a parameter that controls the SPDE range, there was a clear trend where shorter range yielded better model performance, with the same trend seen for both the mechanistic life cycle and the climate and population GAM model (Supplementary Table S9–S10, Fig. S7). The reason for this seems to be that the fitted spatio-temporal SPDE is more conservative when we use a shorter range, which has the effect that high values of the fitted SPDE are more concentrated to regions where the covariates do not explain the *Ae. albopictus* presence so well (Supplementary Figs. S8–S9).

Reporting summary

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

Data availability

The data on *Ae. albopictus* presence is freely available and can be obtained upon request from the European Center for Disease Prevention and Control, but it cannot be made available elsewhere. The climate, mechanistic life cycle and human population covariates are included in the Supplementary at bioRxiv, <https://doi.org/10.1101/2024.09.20.614113>.

Code availability

Code used to fit the spatio-temporal GAM model and create figures is included in the Supplementary at bioRxiv, <https://doi.org/10.1101/2024.09.20.614113>. Since the data on *Ae. albopictus* presence cannot be made available, the code for fitting the GAM model runs on a partially simulated dataset.

Received: 25 November 2024; Accepted: 7 March 2025;

Published online: 09 April 2025

References

- WHO, vector borne diseases. <https://www.who.int/news-room/fact-sheets/detail/vector-borne-diseases>.
- Farooq, Z. et al. Impact of Climate and *Aedes albopictus* Establishment on Arbovirus Outbreaks in Europe: A Time-to-Event Analysis. *Lancet Planetary Health* (in press).
- Local transmission of dengue virus in mainland EU/EEA, 2010–present. <https://www.ecdc.europa.eu/en/all-topicsz/dengue/surveillance-and-disease-data/autochthonous-transmission-dengue-virus-eueea>.
- Local transmission of chikungunya virus in mainland EU/EEA, 2007–present. <https://www.ecdc.europa.eu/en/infectiousdisease-topics/zdisease-list/chikungunya-virus-disease/surveillance-threats-and>.
- Skuse, F. A. The banded mosquito of Bengal. *Indian Mus. Notes* **3**, 20 (1894).
- Kraemer, M. U. et al. Past and future spread of the arbovirus vectors *Aedes aegypti* and *Aedes albopictus*. *Nat. Microbiol.* **4**, 854–863 (2019).
- Benelli, G., Wilke, A. B. & Beier, J. C. *Aedes albopictus* (Asian tiger mosquito). *Trends Parasitol.* **36**, 942–943 (2020).
- Madon, M. B., Mulla, M. S., Shaw, M. W., Kluh, S. & Hazelrigg, J. E. Introduction of *Aedes albopictus* (Skuse) in southern California and potential for its establishment. *J. Vector Ecol. J. Soc. Vector Ecol.* **27**, 149–154 (2002).
- Scholte, E. J. et al. Accidental importation of the mosquito *Aedes albopictus* into the Netherlands: a survey of mosquito distribution and the presence of dengue virus. *Med. Vet. Entomol.* **22**, 352–358 (2008).
- Derneulemeester, J. et al. First interception of *Aedes* (Stegomyia) *albopictus* in Lucky bamboo shipments in Belgium. *J. Eur. Mosq. Control Assoc.* **32**, 14–16 (2014).
- Aranda, C., Eritja, R. & Roiz, D. First record and establishment of the mosquito *Aedes albopictus* in Spain. *Med. Vet. Entomol.* **20**, 150–152 (2006).
- Scholte, E. J. & Schaffner, F. 14. Waiting for the tiger: establishment and spread of the *Aedes albopictus* mosquito in Europe. *Emerg. Pests vector-borne Dis. Eur.* **1**, 241 (2007).
- Sebesta, O., Rudolf, I., Betasova, L., Pesko, J. & Hubalek, Z. An invasive mosquito species *Aedes albopictus* found in the Czech Republic, 2012. *Eur. Surveill.* **17**, 20301 (2012).
- Kampen, H., Kronefeld, M., Zielke, D. & Werner, D. Further specimens of the Asian tiger mosquito *Aedes albopictus* (Diptera, Culicidae) trapped in southwest Germany. *Parasitol. Res.* **112**, 905–907 (2013).
- Adhami, J. & Reiter, P. Introduction and establishment of *Aedes* (Stegomyia) *albopictus* skuse (Diptera: Culicidae) in Albania. *J. Am. Mosq. Control Assoc.* **14**, 340–343 (1998).
- Sabatini, A., Rainieri, V., Trovato, G. & Coluzzi, M. *Aedes albopictus* in Italy and possible diffusion of the species into the Mediterranean area. *Parassitologia* **32**, 301–304 (1990).
- ECDC. *Aedes albopictus* - current known distribution: October 2023. <https://www.ecdc.europa.eu/en/publications-data/aedes-albopictus-current-known-distribution-october-2023> (2023).
- Metelmann, S. et al. The UK's suitability for *Aedes albopictus* in current and future. *J. R. Soc. Interface* **16**, 20180761 (2019).
- Torina, A. et al. Modelling time-series *Aedes albopictus* abundance as a forecasting tool in urban environments. *Ecol. Indic.* **150**, 110232 (2023).
- Launch of the Global Arbovirus Initiative. <https://www.who.int/news-room/events/detail/2022/03/31/default-calendar/global-arbovirus-initiative>.

21. Yang, B. et al. Modelling distributions of *Aedes aegypti* and *Aedes albopictus* using climate, host density and interspecies competition. *PLoS Neglected Trop. Dis.* **15**, e0009063 (2021).

22. Fischer, D., Thomas, S. M., Niemitz, F., Reineking, B. & Beierkuhnlein, C. Projection of climatic suitability for *Aedes albopictus* Skuse (Culicidae) in Europe under climate change conditions. *Glob. Planet. Change* **78**, 54–64 (2011).

23. Yang, H. M. et al. Abiotic effects on population dynamics of mosquitoes and their influence on dengue transmission. in ecological modelling applied to entomology (eds. Ferreira, C. P. & Godoy, W. A. C.) 39–79 (Springer International Publishing, Cham, 2014). https://doi.org/10.1007/978-3-319-06877-0_3.

24. Guidelines for mosquito surveillance. <https://www.ecdc.europa.eu/en/disease-vectors/surveillance-and-disease-data/guidelines-mosquito>.

25. Semenza, J. C. & Zeller, H. Integrated surveillance for prevention and control of emerging vector-borne diseases in Europe. *Eur. Surveill.* **19**, 20757 (2014).

26. Romanello, M. et al. The 2023 report of the Lancet Countdown on health and climate change: the imperative for a health-centred response in a world facing irreversible harms. *Lancet* **402**, 2346–2394 (2023).

27. Romanello, M., et al. The 2024 report of the Lancet Countdown on health and climate change: facing record-breaking threats from delayed action. *Lancet* **404**, 1847–1896 (2024).

28. van Daalen, K. R. et al. The 2024 Europe report of the Lancet Countdown on health and climate change: unprecedented warming demands unprecedented action. *Lancet Public Health* **10**, S2468–S2667 (2024).

29. Liu-Helmersson, J., Rocklöv, J., Sewe, M. & Brännström, Å. Climate change may enable *Aedes aegypti* infestation in major European cities by 2100. *Environ. Res.* **172**, 693–699 (2019).

30. Carney, R. M. et al. Integrating Global Citizen science platforms to enable next-generation surveillance of invasive and vector mosquitoes. *Insects* **13**, 675 (2022).

31. Južnič-Zonta, Ž. et al. Mosquito Alert Digital Entomology Network; Mosquito Alert Community. Mosquito alert: leveraging citizen science to create a GBIF mosquito occurrence dataset. *GigaByte*. gigabyte54 <https://doi.org/10.46471/gigabyte.54> (2022).

32. Pataki, B. A. et al. Deep learning identification for citizen science surveillance of tiger mosquitoes. *Sci. Rep.* **11**, 4718 (2021).

33. Virgillito, C. et al. Involving citizen scientists in monitoring arthropod vectors of human and zoonotic diseases: the case of Mosquito Alert in Italy. *Sci. Total Environ.* **948**, 174847 (2024).

34. Zardini, A. et al. Estimating the potential risk of transmission of arboviruses in the Americas and Europe: a modelling study. *Lancet Planet. Health* **8**, e30–e40 (2024).

35. Longbottom, J. et al. *Aedes albopictus* invasion across Africa: the time is now for cross-country collaboration and control. *The Lancet, Glob. Health* **11**, e623–e628 (2023).

36. Copernicus Climate Data Store. <https://cds.climate.copernicus.eu/>.

37. ERA5-Land hourly data from 1950 to present. <https://cds.climate.copernicus.eu/datasets/reanalysis-era5-land?tab=overview>.

38. Schulzweida, U. CDO User Guide. <https://doi.org/10.5281/zenodo.4246983> (2020).

39. Center For International Earth Science Information Network-CIESIN-Columbia University. Gridded Population of the World, Version 4 (GPWv4): Population Count, Revision 11. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC) <https://doi.org/10.7927/H4JW8BX5> (2018).

40. Eurostat. <https://ec.europa.eu/eurostat>.

41. Rue, H., Martino, S. & Chopin, N. Approximate Bayesian inference for latent Gaussian models using integrated nested Laplace approximations (with discussion). *Journal of the Royal Statistical Society, Ser. B* **71**, 319–392 (2009).

42. Lindgren, F., Rue, H. & Lindström, J. An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. *J. R. Stat. Soc. Ser. B: Stat. Methodol.* **73**, 423–498 (2011).

43. Tran, A. et al. A Rainfall- and Temperature-Driven Abundance Model for *Aedes albopictus* Populations. *Int. J. Environ. Res. Public Health* **10**, 1698–1719 (2013).

44. Jia, P. et al. A climate-driven mechanistic population model of *Aedes albopictus* with diapause. *Parasites Vectors* **9**, 175 (2016).

45. Cailly, P. et al. A climate-driven abundance model to assess mosquito control strategies. *Ecol. Model.* **227**, 7–17 (2012).

46. Erguler, K. et al. Large-scale modelling of the environmentally-driven population dynamics of temperate *Aedes albopictus* (Skuse). *PLoS ONE* **11**, e0149282 (2016).

47. Metelmann, S. et al. The UK's suitability for *Aedes albopictus* in current and future climates. *J. R. Soc. Interface* **16**, 20180761 (2019).

48. Soetaert, K., Petzoldt, T. & Setzer, R. W. Solving Differential Equations in R: Package deSolve. *J. Stat. Softw.* **33**, 1–25 (2010).

49. Simini, F., González, M. C., Maritan, A. & Barabási, A.-L. A universal model for mobility and migration patterns. *Nature* **484**, 96–100 (2012).

50. Masucci, A. P., Serras, J., Johansson, A. & Batty, M. Gravity versus radiation models: On the importance of scale and heterogeneity in commuting flows. *Phys. Rev. E* **88**, 022812 (2013).

Acknowledgements

The study received support from the Swedish research council Vinnova through the CLAIRE project (ref 2020-03367). JR received support from the Alexander von Humboldt foundation through the Alexander von Humboldt Professorship funding scheme.

Author contributions

Sandra Barman: Conceptualisation, Data curation, Formal analysis, Methodology, Resources, Software, Validation, Visualisation, Writing – original draft and Writing – review & editing. Jan C. Semenza: Conceptualization, Funding acquisition, Resources, Supervision, Writing – original draft and Writing – review & editing. Pratik Singh: Data curation, Formal analysis, Resources, Software, Validation, Visualisation and Writing – review & editing. Henrik Sjödin: Software and Writing – review & editing. Joacim Rocklöv: Conceptualization, Funding acquisition, Methodology, Supervision, Writing – original draft and Writing – review & editing. Jonas Wallin: Conceptualization, Funding acquisition, Methodology, Software, Supervision, Writing – original draft and Writing – review & editing.

Funding

Open access funding provided by RISE Research Institutes of Sweden.

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/s43247-025-02199-z>.

Correspondence and requests for materials should be addressed to Sandra Barman.

Peer review information *Communications Earth & Environment* thanks André Ricardo Ribas Freitas and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editors: Heike Langenberg. [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 4.0 International License, which permits use, sharing, adaptation, 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 changes were made. 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/4.0/>.

© The Author(s) 2025