
















# The overlapping global distribution of dengue, chikungunya, Zika and yellow fever

Received: 20 December 2024

Accepted: 25 March 2025

Published online: 10 April 2025


 Check for updates

Ahyoung Lim <sup>1,2</sup>, Freya M. Shearer<sup>3,4</sup>, Kara Sewalk<sup>5</sup>, David M. Pigott <sup>6,7</sup>, Joseph Clarke <sup>8</sup>, Azhar Ghouse<sup>1,9</sup>, Ciara Judge <sup>1,2</sup>, Hyolim Kang<sup>1,2,10</sup>, Jane P. Messina <sup>11,12</sup>, Moritz U. G. Kraemer <sup>13,14</sup>, Katy A. M. Gaythorpe <sup>15</sup>, William M. de Souza<sup>16</sup>, Elaine O. Nsoesie <sup>17</sup>, Michael Celone<sup>6</sup>, Nuno Faria <sup>18</sup>, Sadie J. Ryan <sup>19</sup>, Ingrid B. Rabe<sup>20</sup>, Diana P. Rojas<sup>20</sup>, Simon I. Hay <sup>6,7</sup>, John S. Brownstein <sup>21</sup>, Nick Golding <sup>3,4,22</sup> & Oliver J. Brady <sup>1,2</sup> 

Arboviruses transmitted mainly by *Aedes (Stegomyia) aegypti* and *Ae. albopictus*, including dengue, chikungunya, and Zika viruses, and yellow fever virus in urban settings, pose an escalating global threat. Existing risk maps, often hampered by surveillance biases, may underestimate or misrepresent the true distribution of these diseases and do not incorporate epidemiological similarities despite shared vector species. We address this by generating new global environmental suitability maps for *Aedes*-borne arboviruses using a multi-disease ecological niche model with a nested surveillance model fit to a dataset of over 21,000 occurrence points. This reveals a convergence in suitability around a common global distribution with recent spread of chikungunya and Zika closely aligning with areas suitable for dengue. We estimate that 5.66 (95% confidence interval 5.64–5.68) billion people live in areas suitable for dengue, chikungunya and Zika and 1.54 (1.53–1.54) billion people for yellow fever. We find large national and subnational differences in surveillance capabilities with higher income more accessible areas more likely to detect, diagnose and report viral diseases, which may have led to overestimation of risk in the United States and Europe. When combined with estimates of uncertainty, these suitability maps can be used by ministries of health to target limited surveillance and intervention resources in new strategies against these emerging threats.

Arboviruses transmitted by *Aedes aegypti* and *Ae. albopictus* mosquitoes in human-amplified cycles (*Aedes*-borne viruses), including dengue, chikungunya, and Zika viruses, and yellow fever in urban settings, represent a rapidly growing threat on a global scale. In recent decades, their contribution to global mortality and morbidity has grown and their geographical range has significantly expanded<sup>1</sup>. Recent epidemiological data from 2023 and early 2024 reveals an alarming surge in dengue cases across endemic countries, such as Brazil, Peru, and Bangladesh, marking the highest incidence years on record<sup>2–5</sup>. Over the

past decade, chikungunya and Zika emerged and spread throughout the Americas, caused increasingly large outbreaks<sup>6,7</sup> and have spread beyond their historic tropical and subtropical limits to areas such as Mediterranean Europe and the southern United States<sup>8,9</sup>. The current burden of yellow fever is likely underestimated<sup>10,11</sup>, impeding efficient planning, allocation, and evaluation of vaccination programmes. The recent challenges posed by upsurges in *Aedes*-borne arboviruses, coupled with climate, demographic changes, and increased human global mobility all of which are expected to worsen current trends,

A full list of affiliations appears at the end of the paper.  e-mail: [Oliver.Brady@lshtm.ac.uk](mailto:Oliver.Brady@lshtm.ac.uk)

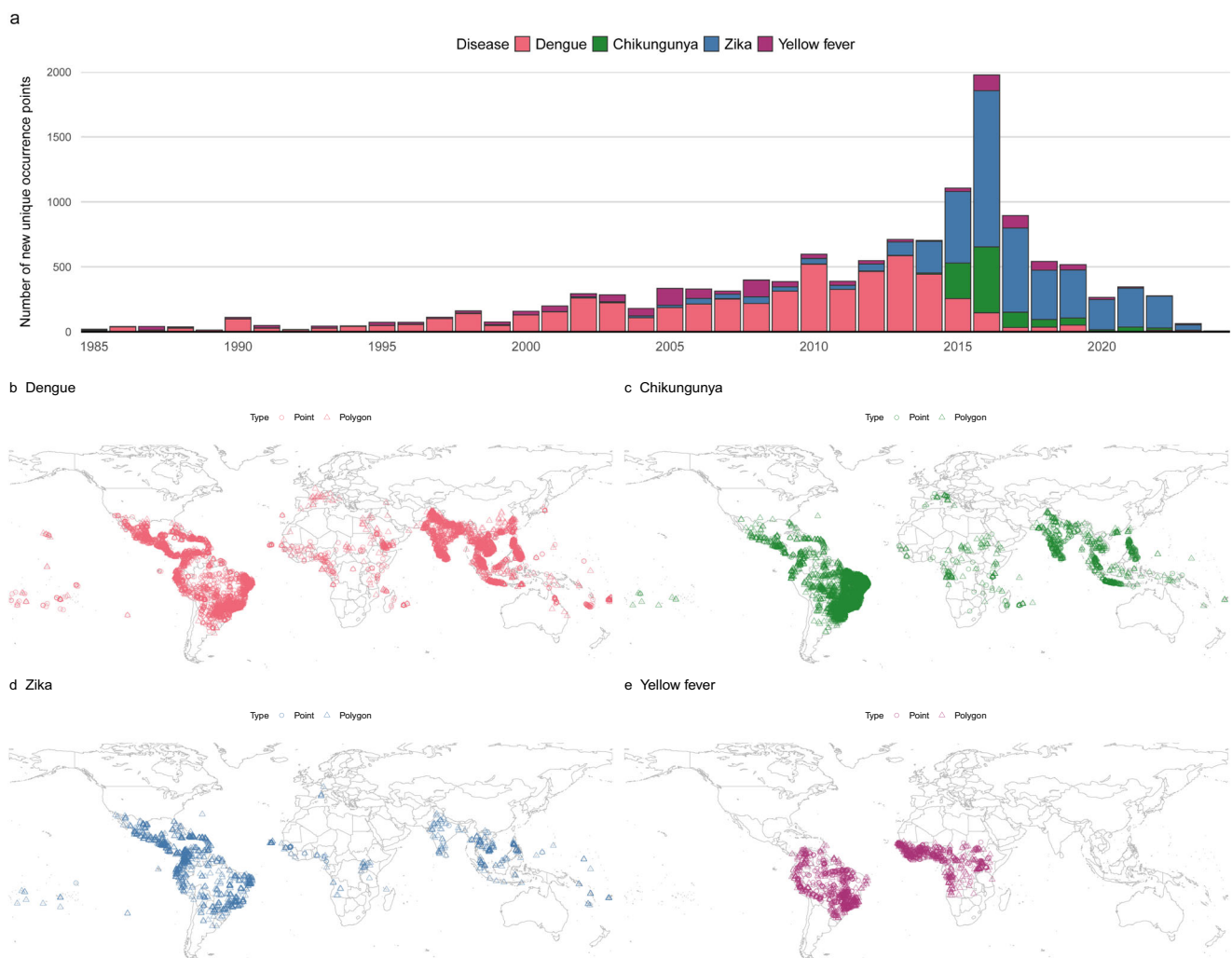
underscore the urgent need to enhance our understanding of the current geographical distribution of these diseases and their potential to expand into new regions<sup>12–14</sup>. In 2022, the WHO Global Arbovirus Initiative was launched to develop an integrated approach to risk monitoring and early detection of *Aedes*-borne arboviruses that would leverage their shared mosquito vectors and environmental drivers<sup>15</sup>.

Global maps of disease risk are valuable tools for identifying vulnerable populations, guiding surveillance, and maximising the impact and efficiency of surveillance and control efforts. The spatial distribution of a vector-borne disease in humans is influenced by the spatial distribution of humans, vectors, the environmental features that shape transmission efficiency and mobility to spread the pathogen to suitable areas. However direct observation of this spatial distribution is rarely possible due to spatial variation in access to care, accurate diagnosis and transparent reporting. A growing availability of historical disease occurrence data and high-resolution remote sensing data, along with advancements in geostatistical modelling techniques<sup>16</sup>, spurred the development of the first generation of global suitability maps for dengue<sup>17</sup>, chikungunya<sup>18</sup>, Zika<sup>19</sup>, and yellow fever<sup>20</sup> as well as their *Aedes* mosquito vectors<sup>21</sup>. However, previous environmental suitability maps have often been challenged by biases in data arising from the spatially varying levels of propensity to detect, diagnose and report infectious diseases across the world, resulting in maps

that represent disparities in surveillance effort rather than, as intended, disease risk.

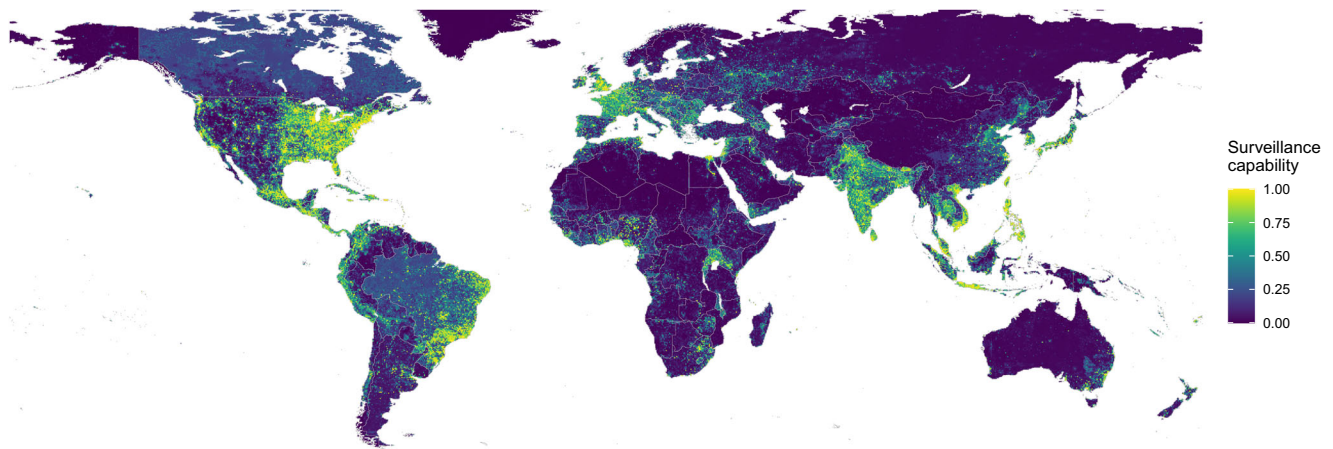
Furthermore, previous mapping studies have been disease-specific, neglecting the epidemiological similarities between *Aedes*-borne diseases through their shared vector species and environmental drivers. The recent spread of chikungunya and Zika has, so far, mostly been limited to a subset of areas where dengue has already been reported, but it remains unknown if these diseases will follow the same path of global expansion as dengue or whether they will be constrained by disease-specific relationships with climatic and environmental factors<sup>22</sup>.

Here we present new global maps for dengue, chikungunya, Zika and yellow fever that incorporate key data updates and methodological innovations. Our approach builds on an existing ecological niche modelling (ENM) framework, and attempts to address the key challenges of observation bias and shared drivers by: (i) explicitly accounting for spatial variation in surveillance capability for detecting and reporting viral infections; and (ii) joint modelling of multiple *Aedes*-borne arboviruses, leveraging similarities of their shared drivers to overcome disease-specific data gaps. We apply this approach to the latest occurrence records (up to March 2024) for dengue, chikungunya, Zika, and yellow fever. As a result, our maps of different *Aedes*-borne arboviral diseases will be directly comparable, providing the



**Fig. 1 | The temporal and spatial distribution of *Aedes*-borne arbovirus occurrence points.** **a** The global number of new unique occurrence points added each year (i.e. after thinning). Years with sparse ( $n < 100$ ) occurrence records

(1927–1984) are not shown. **b–e** global maps of occurrence data for dengue (**b**), chikungunya (**c**), Zika (**d**) and yellow fever (**e**). The maps were created using public-domain Natural Earth data, accessed through the `rnaturalearth` package in R<sup>86</sup>.



**Fig. 2 | Model-predicted relative surveillance capability for emerging acute viral infectious diseases.** Values close to 1 (yellow) indicate that a viral infection is more likely to be publicly reported. The map was created using public-domain Natural Earth data, accessed through the `rnaturalearth` package in R<sup>86</sup>.

evidence base for a coordinated cross-disease response to their spread and facilitating more effective allocation of resources and interventions across affected regions.

## Results

### The global distribution of occurrence data points

We assembled a total of 58,361 occurrence records between 1927 and March 2024 for dengue, chikungunya, Zika, and yellow fever. These data build on existing occurrence databases<sup>23,24</sup> for each disease adding 23,623 new points for dengue, 12,835 for chikungunya, 4171 for Zika, and 1505 for yellow fever, filling in key gaps like including the spread of Zika after the year 2016 and more recent expansions of chikungunya in South America (Supplementary Fig. 1). The removal of spatial duplicates (thinning) was done by collapsing the presence data over time to prevent repeated occurrences from skewing the geographical representation. This created 13,127 unique locations where dengue ( $n = 5867$ ), chikungunya ( $n = 4727$ ), Zika ( $n = 1138$ ) and yellow fever ( $n = 1395$ ) have occurred across 118, 106, 78 and 34 countries respectively.

The largest increases in the annual number of new unique occurrence points occurred during large outbreaks for chikungunya, Zika and yellow fever (e.g. the 2015–2016 epidemic, Fig. 1a). For dengue, however, the number of novel data points saw a more gradual expansion for decreasing in recent years, suggesting either slowing expansion or stabilisation in the coverage and resolution of dengue reporting. Spatial coverage of the arboviral occurrence data includes all countries that regularly report dengue cases<sup>25</sup>, as well as regions where routine *Aedes*-borne arbovirus surveillance is not performed, but *Aedes*-borne arboviruses are known to occur, such as some regions of Africa<sup>26</sup> (Fig. 1b–e).

We found substantial geographic overlap between dengue, chikungunya and Zika occurrence points with 78.3% (3701/4727) of chikungunya and 83.7% (952/1138) of Zika occurrence points falling within 50 km of a dengue occurrence point, suggesting a high degree of overlap in environmental suitability between these different diseases.

The occurrence data for all viral diseases ( $n = 21,700$ ) covers the vast majority of the populated regions of the world. The density of points was highest in high-income countries, including the United States, European nations, and Japan, but also in many countries in the Global South, including Brazil, India, and Thailand (Supplementary Fig. 2).

### Global map of viral surveillance capability

Our predicted map of surveillance capability for acute viral diseases closely mirrors the all viral diseases dataset with higher values in high-

income countries but also many Global South countries with known extensive surveillance systems like Brazil, Uganda, Philippines, Vietnam, and Indonesia (Fig. 2). We also find important sub-national variations in surveillance capability, with urban areas generally showing higher levels compared to rural areas, particularly evident in countries like Brazil and India. Examining the surveillance model variable relative importance plots (Supplementary Figs. 3 and 4) shows the importance of shorter travel times to healthcare facilities and a higher GDP for predicting higher reporting probabilities.

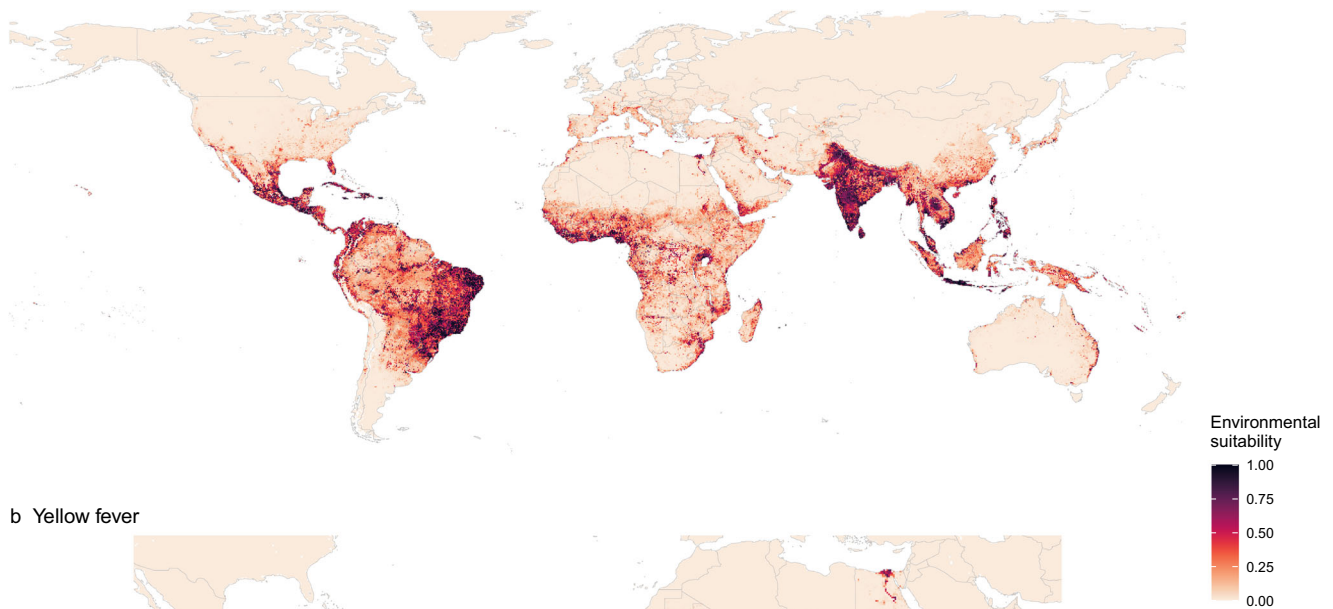
The spatial distribution of uncertainty of the predictions from the surveillance capability model shows very low uncertainty ( $< 0.05$ ) in most areas globally (Supplementary Fig. 5). Validation statistics indicated a high predictive performance of the random forest model evaluated in a 100-fold spatial cross-validation procedure, with an overall AUC of 0.96. Regionally stratified AUC values were good ( $> 0.82$ ) across all regions (Supplementary Fig. 6), with highly populated areas showing higher performance than sparsely populated regions (Supplementary Fig. 7).

### Global maps of arboviral diseases

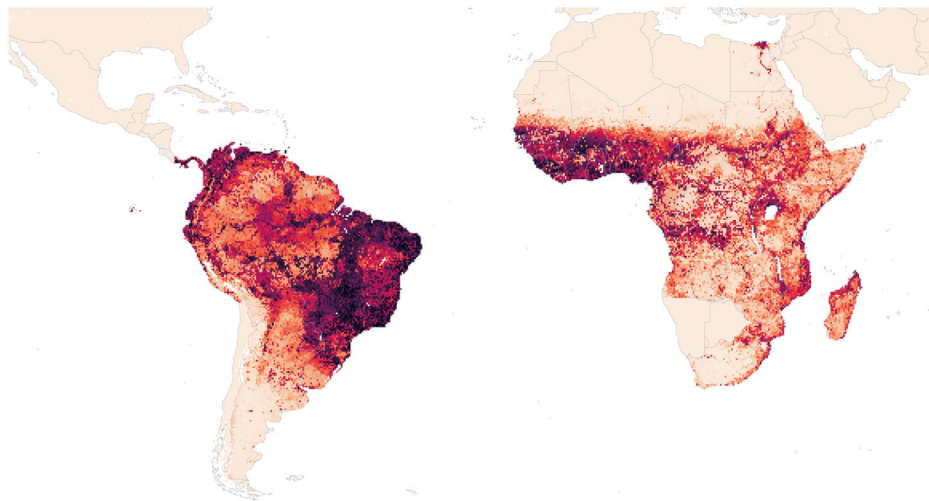
In agreement with the high spatial overlap in datapoints, our models also suggest that dengue, chikungunya, and Zika share a common global distribution. The arboviral disease variable in the arbovirus model contributes the least to decreased node impurity among all variables tested (Supplementary Fig. 8a). It also has a minimal univariate effect in partial dependency plots (Supplementary Fig. 8b) and does not support disease-specific interactions with other variables as evidenced by near-identical predicted suitability maps (Supplementary Fig. 9). The inclusion of disease-specific thermal traits for Zika also reduced the arbovirus model predictive performance for Zika datapoints (Supplementary Fig. 10), further suggesting commonalities. Due to this, we collapsed the maps for dengue, chikungunya, and Zika into a unified map that draws strengths across datasets from all three diseases (Fig. 3a).

Our model for dengue, chikungunya and Zika predicts a high environmental suitability (hereafter suitability) in many tropical and subtropical regions, with a general higher intensity in South America and Asia but a patchier distribution in Africa (Fig. 3). Our measure of suitability is probability of one or more cases of any of these diseases ever having occurred up to 2024, based on the average environmental conditions of a location over a multi-year period (2010–2020). Areas of predicted high suitability include those where cases and outbreaks have been well documented (e.g. across South America, South East Asia and the Indian sub-continent), but also areas where reporting of arboviral diseases is rare (e.g. West and Central Africa). Such

## a Dengue, chikungunya, and Zika



## b Yellow fever



**Fig. 3 | Model-predicted environmental suitability for dengue, chikungunya, and Zika, and yellow fever after accounting for spatial variation in surveillance capacity.** Suitability values represent the probability of one or more cases of diseases having occurred up to 2024, based on the average environmental conditions of a location over the period 2010–2020. Values close to 1 indicate highly suitable conditions for transmission. **a** Areas without a suitable temperature range for

transmission have been set to 0 for dengue, chikungunya, and Zika. **b** Areas outside the countries at risk, endemic, or potentially at risk for yellow fever as defined by the WHO yellow fever risk assessment working group<sup>31</sup> have been set to 0. The maps were created using public-domain Natural Earth data, accessed through the `rnaturalearth` package in R<sup>86</sup>.

discrepancies could represent underreporting of arboviral diseases (Fig. 2) or identify areas that could be at risk if the virus were introduced. For Zika and chikungunya, our maps of suitability align closely with areas with reported cases, such as the 2015–2016 Zika emergence in the Americas and longer-term distribution of chikungunya in India, but also include areas where these diseases have not yet been reported including several countries in Asia for Zika and Africa for chikungunya. As expected, we predict risk to be patchier in more temperate areas including the United States, Northeast Asia, Europe, and the Middle East where suitability is concentrated in more foci, usually in urban areas.

The yellow fever map shows high suitability across broad regions of central and eastern Brazil, coastal areas of northern South America and West Africa, with smaller foci of risk across South America and sub-Saharan Africa (Fig. 3b). Here, suitability represents the probability of occurrence of a case of yellow fever in the susceptible population, which in many areas has already been reduced, but not eliminated, due to widespread yellow fever vaccination. Our maps are consistent with

the observed distribution of yellow fever outbreaks since the beginning of the 20<sup>th</sup> century (Fig. 1e), including more recent expansions in Paraguay and towards the coast in southeast in Brazil<sup>27</sup>. We do, however, also predict suitability in areas where yellow fever cases are rarely reported (Northeast Brazil, large parts of Southeast Africa) or have never been reported (Argentina and Egypt). These predictions suggest such areas remain at risk despite their distance from areas of active virus circulation.

Uncertainty levels in predictions, as measured by the interquartile range of 100 predictions, are generally very low (<0.05) in most areas for all diseases (Supplementary Fig. 11). We observed generally consistently good predictive performance, with an overall AUC value for each disease being higher than 0.97 (Supplementary Fig. 12). Spatially stratified AUC values were good (>0.79) across all regions, with Africa showing the lowest values for Zika and yellow fever, South America for dengue, and Asia for chikungunya (Supplementary Fig. 13). Urbanisation, population density, and the mean temperature of the coldest month variables consistently improved model accuracy and node



impurity in both arbovirus (dengue, chikungunya, and Zika) and yellow fever models (Supplementary Figs. 8a & 14A). The predicted suitability for *Ae. aegypti* and *Ae. albopictus* (main vectors for dengue, chikungunya, and Zika and yellow fever in urban cycles worldwide) and *Hg. janthinomys* (one of main vectors for enzootic cycle of yellow fever in the Americas) were also important factors in distinguishing areas at risk from those not at risk. The joint model showed comparable performance to individual disease models with slight improvements in some metrics, particularly benefiting Zika, which has sparse data (Supplementary Table 6). Excluding the *Ae. aegypti* covariate from the yellow fever model, on the basis that the vector is not widely involved in transmission in South America<sup>28,29</sup>, led to modest decreases in AUC

**Table 1 | Estimated global population and number of countries at risk for dengue, chikungunya, Zika, and yellow fever**

Region	Dengue, chikungunya and Zika	Yellow fever
<b>Population at risk (in billions)*</b>		
Global	5.66 (5.64–5.68)	1.54 (1.53–1.54)
Africa	1.24 (1.24–1.24)	1.14 (1.14–1.15)
Asia	3.46 (3.45–3.47)	-
Americas	0.73 (0.73–0.73)	0.38 (0.38–0.38)
Europe	0.17 (0.17–0.17)	-
Oceania	0.03 (0.03–0.03)	-
<b>Number of countries (sovereign states) at risk**</b>		
Global	169	54

\* Values show mean population estimates with 95% confidence intervals in parentheses. Surveillance capability scores and at-risk population estimates with 95% confidence intervals for each country are available in the Supplementary Data.

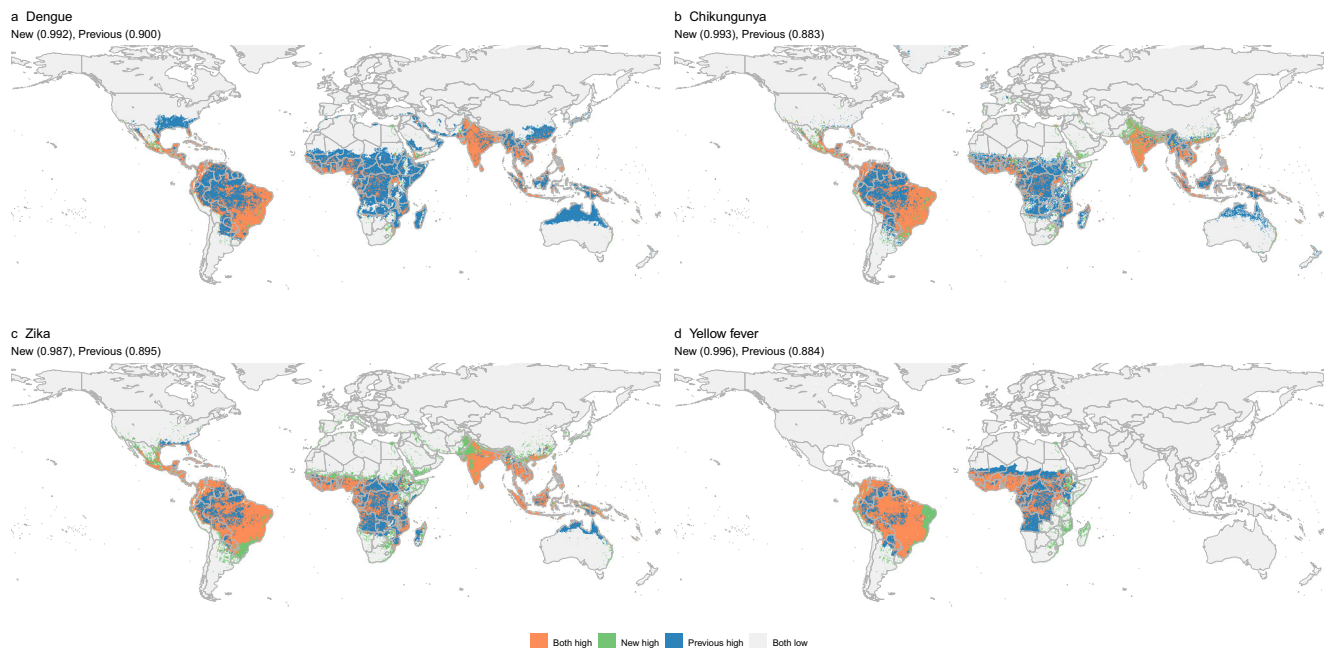
\*\* Countries at risk are defined as UN sovereign states where more than 10% of the population is estimated to be living in at-risk areas.

and minimal changes to the predicted suitability map (Supplementary Fig. 15).

We estimate that globally, 5.66 (95% confidence interval 5.64–5.68) billion people (roughly 73% of the global population in 2022) live in areas that are at-risk of dengue, chikungunya, and Zika, with the vast majority in Asia, followed by Africa and the Americas, encompassing 169 countries (Table 1). We predict that there are 1.54 (1.53–1.54) billion people living in areas at risk of yellow fever, distributed across 54 countries in South America and Africa. The estimated number of people at risk for each country is also provided in the Supplementary Data 1.

### Comparison with previous maps

Our joint models generally predict a more focal distribution of dengue, chikungunya and Zika than previous maps<sup>17–20</sup>, with the most significant reduction in risk observed at the fringes of the global distributions of these diseases (Fig. 4). This refinement better aligns with observed distributions, notably in the United States, western Europe and Northeast Asia, where viral surveillance is robust, yet there are few or no autochthonous case reports. While our maps maintain broad consistency with previous maps, particularly in identifying widespread risks in tropical and subtropical regions (orange in Fig. 4), our analysis also uncovers notable expansions beyond the historical geographical range of these diseases (green in Fig. 4). These areas were not previously identified as at risk but became apparent with the inclusion of additional data from the second half of the 2015–2016 Zika epidemic and subsequent detections in Europe, India, and South East Asia. The expansions into higher latitudes are also notable, particularly in Mexico, Europe, and the Middle East, suggesting including data from more recent expansions is important for observing where expanding distributions of *Aedes* mosquitoes and changing climate may already be driving expansion of arboviral diseases.



**Fig. 4 | Comparison of previously published and current suitability maps for arboviral diseases.** Panels show comparisons between earlier published maps and our newly generated maps for (a) dengue, (b) chikungunya, (c) Zika, and (d) yellow fever. Previous maps were retrieved from Messina et al.<sup>17</sup> for dengue, Nsoesie et al.<sup>18</sup> for chikungunya, Messina et al.<sup>19</sup> for Zika, and Shearer et al.<sup>20</sup> for yellow fever. Area Under the Curve (AUC) values for our map and previous maps are indicated in parentheses. AUC values measure the model's ability to distinguish between occurrence and background points, with values closer to 1 indicating better

predictive performance. The AUC was calculated based on predicted values from both our map and previous maps using presence and background points. Maps were converted into binary format using a threshold that maximised the global sum of sensitivity and specificity. Once binary maps of at-risk areas were generated, we categorised each 5 × 5 km pixel into one of four groups: (1) areas at risk in both our map and the previous maps, (2) areas at risk only in previous maps, (3) areas at risk only in our map, and (4) areas not at risk in either map. The maps were created using public-domain Natural Earth data, accessed through the `rnaturalearth` package in R<sup>86</sup>.

Our map of yellow fever shows a similar distribution in South America, but a patchier distribution in Africa with predicted risk extending to the Southeast of the continent. In previous maps, predictions for coastal Brazil were masked out based on the 2010 expert opinion map by the WHO working group on geographic risk for yellow fever<sup>30</sup>. Following the detection of yellow fever across multiple coastal areas in 2017 and 2018<sup>27</sup> and in consultation with their contemporary equivalent (WHO yellow fever risk assessment working group<sup>31</sup>) their recommendation was now not to mask these areas, but acknowledge that our predictions of risk differ in these areas where they represent the risk of transmission if yellow fever virus were to be introduced.

## Discussion

Global maps of diseases should provide a comprehensive and standardised assessment of risk that can be used to guide health policy and optimise responses to mitigate their impact. One of the priority actions of the Global Arbovirus Initiative at its launch was to develop a global framework to monitor the combined risk of *Aedes*-borne diseases as they share most of the drivers of transmission and an integrated approach is the most effective and logical approach to target integrated arbovirus preparedness, prevention and control activities. Historically, *Aedes*-borne arboviruses have often been mapped separately using different methodologies and approaches<sup>17–20</sup>. This siloed approach has resulted in maps that represent differences in surveillance, testing, and diagnosis as much as the actual distributions of the diseases themselves and under-utilise the similarities between different arboviral diseases. Our approach addresses these by accounting for spatial biases in surveillance and joint modelling of arboviral diseases, respectively to provide more accurate estimates of the global suitability.

Our maps show that recent spread of chikungunya and Zika has been wholly within areas already suitable for dengue. This indicates that any area with dengue is at risk of, or may already have experienced, transmission of Zika or chikungunya viruses. The similarity between the distributions of dengue, chikungunya, and Zika has been obscured in previous maps due to the use of different modelling approaches, which hindered direct comparisons between the diseases. Here, we compared and combined suitability maps of the different *Aedes*-borne diseases by fitting our model to the updated occurrence data for each disease and by using a common joint modelling approach. Given the high potential for misdiagnosis and under-reporting of Zika and chikungunya<sup>32</sup>, this combined map provides the best estimate of their distribution and basis for identifying vulnerable populations where prevention, disease detection, and clinical management should be prioritised.

Our analysis addresses the long-standing issue of spatial surveillance biases in the field of global disease mapping by nesting separate models for surveillance and disease risk. Occurrence data inherently has biases towards areas with more robust surveillance systems, such as the USA and Europe, potentially overestimating disease risk in these areas. Separation of surveillance and arbovirus risk models allows more accurate estimates of the relationships between their respective risk factors and leads to more specific and focal predictions of disease risk, particularly in areas that have been historically oversampled like Europe and the USA. These more focal and refined maps better align with the sporadic nature of dengue, chikungunya and Zika outbreaks in Europe and USA and enable a more targeted approach to detecting and responding to arboviral introduction.

Quantifying spatial uncertainty is increasingly important in risk mapping to show where additional data collection will be most valuable. Our maps of surveillance capability (Fig. 2), model prediction uncertainty (Supplementary Fig. 5), and spatial predictive performance (Supplementary Fig. 7) can be used to fill gaps in surveillance, detect regional outliers where risk factors act differently, and identify further risk factors that may better explain the distribution of these

diseases respectively. This comprehensive assessment of uncertainty serves as a useful starting point for collaboration with researchers and country ministries of health to identify new data sources and targeted data collection activities to prospectively validate the suitability maps. Validation by stakeholders in affected regions will both improve the maps but also build engagement with the maps, increasing the likelihood that they can usefully inform surveillance and control policy changes.

A key limitation of these maps is that they show all areas suitable for transmission not where these viruses are currently circulating which is likely to be more constrained by viral dispersal. Pairing these maps with estimates of human mobility from endemic areas could allow more specific estimates of areas currently experiencing transmission. Targeted testing for disease in areas identified as suitable but with no recent datapoints is also important to differentiate between under detection and true absence. Past and future changes in climate, urbanisation, surveillance capacity and spread of arboviral vectors could all change the resulting risk maps. Due to data sparsity at the global level it is necessary to aggregate data over multiple years, limiting our ability to capture these changes in global risk over time. While previous work has found that such changes will only have minimal and gradual impacts on the environmental niche for dengue, tipping point events, such as invasion of *Ae. aegypti* into mainland Europe could lead to more rapid changes in risk necessitating updating of these maps. Our surveillance model may also be influenced by the spatial distribution of each of the viral diseases used in the training data, limiting our ability to isolate pure surveillance effort. This model also assumed the equivalent detection probabilities across different viral infections despite their different clinical presentations and methods of diagnosis. To minimise these effects we chose a broad range of viral diseases with different global distributions, defining surveillance capability as the wider investment in detecting, diagnosing, and reporting any circulating pathogens, but future studies could incorporate more direct measures of surveillance, like surveillance system assessments<sup>33</sup> or comparisons between reported case data and serologically-detected infections<sup>34</sup>. Our surveillance model did not include data on COVID-19 and the sustaining gains during the pandemic, which has likely since changed the distribution of surveillance and improve the testing and sequencing capacities worldwide, as well as disrupted transmission dynamics<sup>34</sup> and yellow fever vaccination efforts<sup>35</sup>. Our yellow fever model also only included vaccination coverage as of 2020 and did not capture recent vaccination activities<sup>36</sup>. Finally, our models did not include other vector mosquito species that are known to transmit arboviral diseases in enzootic and epizootic cycles in different regions, including *Ae. africanus*, mosquitoes from the genera *Culex*, *Anopheles*, *Haemagogus* and *Sabethes*<sup>37</sup>, which may improve map estimates, particularly for yellow fever.

Here, we produce new global high spatial-resolution maps of environmental suitability for dengue, chikungunya, Zika, and yellow fever and provide up-to-date insights into the evolving landscape of these diseases. Our approach fills key data gaps, addresses issues of spatial biases in surveillance and shares information between *Aedes*-borne viral diseases. The overlap in global environmental suitability among these diseases highlights the interconnected nature of arboviral transmission and underscores the need for integrated and coordinated responses. Our suitability maps offer a starting point for national to local-scale risk assessments over the short to medium term, helping to target interventions and surveillance efforts. These maps should be used in combination with local risk mapping studies that utilise fine-scale spatial information on arboviral disease cases and mosquito counts to further target locally. New data generated from such efforts could be used to iteratively update these maps to improve their contemporariness and operational relevance, contributing to more effective disease control efforts worldwide.

## Methods

To map environmental suitability for *Aedes*-borne arboviral disease, we applied an ENM framework which required: (i) occurrence data; (ii) a set of climatic, environmental and socioeconomic covariates that are known to influence the transmission of these viruses or population dynamics of their vectors; and (iii) a statistical model that parameterises a function which describes the predicted probability of disease being present in each spatial location. The resulting model produced 5 × 5 km spatial-resolution global maps of probability of occurrence of dengue, chikungunya, Zika and yellow fever given long-term average environmental conditions.

### Arbovirus occurrence data

The primary epidemiological data included in our analysis was occurrence data, defined as a unique geographic location where one or more cases of a particular disease been reported at any point in time<sup>23</sup>. We extracted existing occurrence data from previous publications on dengue<sup>17</sup>, chikungunya<sup>18,38</sup>, Zika<sup>19</sup>, and yellow fever<sup>20</sup>. Additional searches were conducted using ProMED mail (<http://www.promedmail.org>) reports between the period 2015–2022 for chikungunya and Zika to fill important temporal gaps following previously established protocol<sup>23,39</sup>. Despite potential variations in coverage and lack of precise geopositioning, ProMED remains valuable for its comprehensive and timely<sup>40,41</sup> tracking of global outbreaks of emerging infectious diseases, supporting traditional surveillance efforts effectively. We also searched through epidemiological bulletins of the ECDC<sup>8,42–44</sup>, WHO regional outbreak updates<sup>45,46</sup> and included occurrence data for countries where diseases had not been reported in the past, but had occurred between 2016 and March 2024 (e.g., dengue in Ibiza, Spain and chikungunya in Uruguay), or where occurrence data were available at the sub-national level (e.g., yellow fever in Bolivia, Brazil, and Peru). Finally, occurrence data for dengue, chikungunya, Zika, and yellow fever from HealthMap platform ([www.healthmap.org](http://www.healthmap.org)) was extracted and inspected by cross-referencing with the peer-reviewed literature<sup>47–51</sup> and the epidemiological bulletins. Additional data uncovered through cross-referencing with these sources was incorporated only if they related to subnational levels. Detailed information about the data sources and the number of occurrence data extracted from each source are presented in Supplementary Table 1.

We included occurrence data in two forms: point and polygon. Point data represents infection in a specific location with a minimum resolution of 0.05 degrees which represents the highest available resolution of all covariate datasets. Polygon data represents occurrence somewhere within an administrative unit area, but with the exact location unknown. With the exception of small island nations, all polygon data is at the first or second administrative unit. Both point and polygon data underwent a spatial standardisation (“thinning<sup>52,53</sup>”) where repeated reports from the same location, irrespective of their reporting time, were removed. This reduced the influence of arbitrary variations in reporting frequency between locations (e.g. monthly vs annually) that would have increased bias in areas that choose to report more frequently. Like other species and disease mapping studies, our approach assumes “niche conservatism”, i.e. the environmental limits that define where transmission is possible, and probable do not change substantially over time. This means that the best estimate of this conserved niche comes from aggregating data over as many years as possible, as opposed to just using contemporary data. The final standardised dataset included 13,127 records, comprising 5337 point locations and 7790 polygon locations. The total number of occurrence data before and after thinning and for each disease are shown in Supplementary Table 2.

### Occurrence data for other viral diseases

To attempt to account for spatial differences in surveillance<sup>17</sup> (see section “Ecological niche model”) we assembled a dataset that

included occurrence data for other viral diseases from the HealthMap platform ([www.healthmap.org](http://www.healthmap.org)). HealthMap is a tool for tracking infectious disease events in 15 functional languages by aggregating data from diverse informal online sources, including news media and social media posts, and employs machine learning algorithms to classify and filter these reports, ensuring relevance and reducing noise<sup>54</sup>. For this analysis, HealthMap data was extracted from a MySQL database using structured query to obtain viral disease occurrences between 2006 and 2019. We carefully reviewed the list of viral infections to identify those under dedicated surveillance programmes and excluded occurrence points for Ebola, HIV/AIDS, measles, and polio, as these may lead to an overestimation of surveillance capability in certain regions, particularly in West African countries. This resulted in a total of 338,005 records, representing occurrence reports for viral diseases that cause acute febrile illness, primarily diagnosed by serology and polymerase chain reaction (PCR). The final standardised dataset after thinning included 21,700 records. The full list of these diseases and their counts before and after thinning is shown in Supplementary Table 3.

### Background points

No globally representative survey data are available for these diseases, making classical geostatistical models and their assumptions of unbiased sampling from a clear denominator population, unsuitable<sup>16</sup>. We instead used a “presence-background” or “presence-only” modelling approach<sup>55</sup> where observed ‘presence’ points are supplemented with randomly generated background (absence) across all land surfaces. These background points represent areas where we assume the disease is absent, providing a contrast to the observed presence points in the model. To reduce label-imbalance which can lead to bias, the number of background points selected was proportionate to the number of dengue, chikungunya, Zika and yellow fever occurrence points separately in order to maintain a 1:1 ratio between presence and background points with each disease (sometimes referred to as “down-sampling<sup>56</sup>”).

### Covariates

We included global raster layers associated with a range of factors hypothesised to be associated either with transmission of *Aedes*-borne arboviruses or with capacity to detect, diagnose and report cases. Our choice of covariates was based on a previous systematic review of arbovirus risk mapping studies<sup>57</sup>. Covariate data sources were prioritised by those that gave the highest spatial resolution and covered the 2010–2020 time period where most of our occurrence point data are concentrated. Further details of coverage and resolution of each covariate are provided in Supplementary Table 4. Where multiple observations were available across multiple years, mean values for each pixels were calculated to produce a synoptic raster layer representing the average over time period covered. Most of the covariates were available at 0.05 degree (5 × 5 km at the equator) resolution or higher. For covariates available only at a national scale (e.g., treatment seeking, child mortality, government effectiveness and physicians density), we computed the average value for each World Bank income group and assigned these values to countries with missing data, thus ensuring a globally complete raster layer.

Covariates were resampled to a consistent 0.05 degree grid with a common extent and land/sea mask with lakes and major water bodies removed. The log transformation was optionally applied based on the distributions of each covariate, and all covariates were scaled and centred to have a zero mean and variance of 1. Covariate values for each occurrence point were extracted with mean values across administrative unit areas used for polygon data. Maps of each covariate layer are provided in Supplementary Figs. 16 and 17.



### Covariates for surveillance model

To explicitly model spatial variation in the probability of reporting of an arboviral disease (if present in a given location), we considered nine covariates that are known to be related to the likelihood of detection, diagnosis, and reporting of acute viral infections: (i) gross domestic product (GDP) (5 × 5 km resolution and aggregated national level)<sup>58</sup>; (ii) the fraction of urban land<sup>59</sup>; (iii) travel time to healthcare facilities by walk<sup>60</sup>; (iv) travel time to cities (>50,000 people, any travel mode)<sup>60</sup>; (v) treatment-seeking for fever in children under five years old<sup>61</sup>; (vi) child mortality under five years old<sup>62,63</sup>; (vii) physicians density<sup>62,64</sup>; and (viii) government effectiveness<sup>62,65</sup>.

### Covariates for arbovirus model

We included nine covariates in our arbovirus model that reflect known drivers of transmission and vector dynamics: (i) temperature suitability for dengue virus transmission<sup>66</sup>; (ii) mean temperature of the coldest month<sup>67</sup> (iii) annual cumulative precipitation<sup>67</sup>; (iv) Normalised Difference Vegetation Index (NDVI)<sup>68</sup>; (v) Dynamic Habitat Indices (DHI)<sup>69</sup>; (vi) predicted suitability for *Ae. albopictus*<sup>69</sup>; (vii) predicted suitability for *Ae. aegypti*<sup>69</sup>; (viii) GDP (aggregated at national level)<sup>58</sup> and (ix) human population density<sup>70</sup>. Three additional covariates were included in yellow fever model: (i) predicted suitability for *Haemagogus janthinomys* in South America<sup>71</sup>; (ii) distribution of non-human primates (NHPs)<sup>20</sup> and (iii) yellow fever vaccination coverage<sup>72</sup>. Predicted suitability for *Ae. albopictus* was not included in the yellow fever model.

### Ecological niche model

**Approach.** We used a machine learning model which has previously proven useful for global environmental suitability mapping applications<sup>57</sup>, including dengue<sup>17,39</sup>, chikungunya<sup>18</sup>, Zika<sup>19</sup>, and yellow fever<sup>20</sup>, as well as the global distribution map of *Aedes* vectors<sup>21</sup>. Specifically we used a down sampled random forest (RF) approach that balances the presence and background points at a 1:1 ratio, as RF model has shown to outperform many other machine-learning approaches for the modelling of presence-only data across a range of examples<sup>56,73</sup>.

Our approach involves the development of two separate models: a surveillance model and an arbovirus model. The surveillance model aims to capture the between and within country differences in the long-term average probability that a person infected with an acute viral infection seeks treatment, is correctly diagnosed and is reported in the public domain. It should be noted that such estimates may not correlate with ability to detect unfamiliar newly emerging diseases but will represent longer-term differences in surveillance capacity once locally circulating diseases have been appropriately characterised. The predictions from this surveillance model are then used as an offset in the arbovirus model, allowing the arbovirus model to attribute spatial variations in arbovirus occurrence data solely to drivers of transmission risk.

The surveillance model uses occurrence data on all emerging acute febrile diseases including arboviral diseases (Supplementary Table 3) with an equal ratio of randomly sampled background points, following a uniform distribution across the global landscape. The model formula is as follows, with covariates listed in the same order as in Supplementary Table 4.

$$\begin{aligned} occ_{viral_i} &\sim \text{Bernoulli}(P(occ_{viral_i})) \\ \text{logit}(P(occ_{viral_i})) &= f(GDP_i, GDP_{National_i}, \\ &\quad Urban_i, travel_{health_i}, travel_{cities_i}, \\ &\quad treatmentseeking_i, childmortality_i, goveffectiveness_i, physician_i) \end{aligned} \quad (1)$$

$P(occ_{viral_i})$  is the location-specific probability that a record  $i$  is an occurrence record of an emerging acute febrile disease, rather than a

background point, and  $f$  denotes the non-linear and interacting function of all covariates at the same spatial location as record  $i$ .

The arbovirus model includes occurrence data for all arboviral diseases (dengue, chikungunya, Zika and yellow fever) combined and includes arboviral disease as a categorical variable (*arbovirus<sub>i</sub>*). This allows the model to share information between arboviral diseases, but also generate distinct relationships between risk of each disease and the environmental covariates if they are justified<sup>74</sup>, e.g. different vector competencies for different viruses. The overall arbovirus model equation is as follows:

$$\begin{aligned} \text{logit}(P(occ_{arbovirus_i})) &= f(arbovirus_i, \\ &\quad Temp_i, Tcold_i, Precip_i, NDVI_i, DHI_i, \\ &\quad aegypti_i, albo_i, Haemagogus_i, NHP_i, \quad (2) \\ &\quad GDP_{National_i} Urban_i, Pop_i) + \\ &\quad \log(P(occ_{viral_i}) * (1 - Vaccine_i)) \end{aligned}$$

where  $\log(P(occ_{viral_i}) * (1 - Vaccine_i))$  is an offset term. Two different versions of the arbovirus model are fit using the same data but with different covariates: i) an arbovirus model for dengue, chikungunya and Zika which omits covariates from  $NHP_i$  and  $Haemagogus_i$ ; and sets  $Vaccine_i$  to 0 and ii) a yellow fever arbovirus model that omits the *Ae. albopictus* covariate ( $albo_i$ ).

The use of a modelled offset for spatial sampling bias based on detections of a wide range of other related diseases is analogous to the commonly used Target Group Background approach<sup>75</sup>, but further enables us to explicitly map the estimated spatial variation in reporting rates. Unlike other approaches that have been proposed in ecology to infer spatial bias in a single model<sup>76</sup>, the two-stage approach we employ separates the modelling of surveillance and disease processes. This distinction minimises the potential for overlap in the effects of covariates (e.g. urbanness and GDP) that influence both surveillance and disease dynamics, which is particularly important for the arboviruses we consider.

**Spatial block cross-validation.** Randomly repeated cross validation may lead to an over-optimistic model performance metrics if the independence assumption between training and test data is violated. This is especially common when dealing with spatially structured (clustered) data where training observations are often in close proximity to test observations, introducing spatial autocorrelation<sup>77</sup>. To maintain independence between training and test data, a spatial block 100-fold cross-validation was employed to assess the overall and spatially stratified predictive performance of the model. This approach considers the spatial dependence structure of the data when estimating predictive performance, resulting in a final metric that is typically lower than would be returned by conventional cross-validation procedures. Initially, a grid of ~1000 equal-sized square blocks, each with a size of 500 × 500 km, was generated using the `cv_spatial` function from the “blockCV” package<sup>78</sup>. Each block was then randomly assigned to 100 roughly equal-sized folds, maintaining the balanced number of presences versus background records in each fold. In each iteration, we reserved one-fold (1%) for validation, which was kept out of the model fitting process (99%) and only used to validate predictions (‘out-of-sample’ validation). By cycling through the data, each fold was used as a validation fold exactly once and then we averaged model performance statistics across the 100 repeated runs, using the area under the curve (AUC) of the receiver-operating characteristic plot as a metric of goodness-of-fit. AUC is a confusion matrix-based model performance metrics that measures how well the model can discriminate presence from background points<sup>79</sup>.

**Spatially stratified AUCs.** To understand and visualise the model’s performance in different parts of the world, the spatial map of AUC



values was produced using a geospatial approach. We calculated AUC values for a set of validation polygons that are defined by a 250 km radius around each presence or background (PB) point. ROC curves and AUC statistics were then calculated within each of these polygons which was summarised using the coordinates of the centroid for visualisation purposes. The resulting values were then averaged to derive a comprehensive AUC value for each specific PB point. Additionally, regionally stratified AUCs are calculated by averaging the AUC values of PB points within regional boundaries. This iterative approach was chosen to balance spatial continuity of the map while mitigating the impact of small spatial units on AUC values.

**Prediction and uncertainty levels.** To increase the robustness of model predictions and quantify model uncertainty, we obtained 100 predictions by using RF models that were iteratively calibrated during spatial cross-validation as described above. Each of the 100 fitted sub-models predicted environmental suitability on a scale from 0 to 1. Predictions were made separately for each disease, by updating the “arboviral disease” variable (see section “Ecological niche model”) to represent different diseases. To generate the final prediction raster, a weighted average prediction was calculated for each  $5 \times 5$  km pixel based on the following equation:

$$P_k = \frac{\sum_{i=1}^n (p_{k,i} \times w_i)}{\sum_{i=1}^n w_i} \quad (3)$$

where  $P_k$  represents the weighted average of predicted risk for pixel  $k$ ,  $p_{k,i}$  represents the predicted risk for pixel  $k$  derived from the  $i$ -th sub-model,  $w_i$  represents the AUC of the  $i$ -th sub-model, and  $n$  is total number of sub-models (here  $n = 100$ ). We calculated the interquartile range (IQR) for 100 model predictions at each location to quantify uncertainty. We also generated 1000 bootstrap samples for weighted average prediction for each pixel and determined the confidence intervals by computing 2.5th and 97.5th percentiles from their distributions.

**Variable importance and partial dependence plots.** The metrics employed for variable importance included mean decrease in accuracy and mean decrease in node impurity (measured by the Gini index). The variable importance values were extracted and normalised by dividing each value by the sum of all feature importance values and averaged across 100 folds. Partial dependence values for each model were also extracted using the “pdp” package and aggregated across 100 models by calculating the mean and the 95% confidence interval.

### Masking

Following previous global mapping studies<sup>17,21</sup>, to limit extrapolation of risk predictions far beyond their original fitting datasets we post-hoc mask out (set risk to 0) model-based risk predictions in areas where alternative forms of evidence suggest that transmission is extremely unlikely. This included where temperature profiles were estimated as too cold to allow mosquitoes to survive long enough to complete the extrinsic incubation period of the dengue virus<sup>66</sup>. Temperature suitability values were calculated using the mean temperature of 2010–2020<sup>67</sup> and converted to binary range maps using threshold of  $14/365 = 0.04$  as they were unlikely to support transmission over the two week serial interval of autochthonous arboviral disease transmission. Environmental suitability predictions for dengue, chikungunya and Zika in areas with temperature suitability values of less than 0.04 were set to 0. Risk predictions for yellow fever were set to 0 in areas outside the countries at risk, endemic, or potentially at risk for yellow fever as defined by the WHO yellow fever risk assessment working group which excluded countries outside of South America and Africa<sup>31</sup>. Masking layers and unmasked versions of environmental suitability

maps are provided in Supplementary Figs. 18 and 19 to show the regions of the world affected by the masking process.

### Estimating distribution and population at risk

The continuous suitability maps were converted into binary distribution maps using a threshold value above which an area was classified as at-risk. We defined the threshold as the suitability value that maximised the global sum of sensitivity and specificity when compared the original presence and background points. Separate thresholds were generated for each disease, with the following suitability values: dengue (0.37), chikungunya (0.21), Zika (0.14), and yellow fever (0.32), to produce individual binary maps. An aggregate binary map for dengue, chikungunya and Zika was then created by combining these individual binary maps from the three diseases, such that any area predicted to be at risk for any of these diseases was considered at risk. Using these binary maps and the global population grid<sup>70</sup>, we calculated population at a global, continent and country (UN member state; 193 in total) level. For the purpose of estimating total countries at risk, a country was only included if more than 10% of its total population was identified as being at risk of disease.

### Sensitivity analyses

To test if the maps for Zika were improved with a disease-specific temperature suitability covariate, the dengue temperature suitability covariate was replaced with a Zika temperature suitability covariate based on temperature relationships from Ryan et al.<sup>80</sup> The improvement of the maps was assessed using a 50-fold block cross validation approach comparing overall and regionally-stratified model performance metrics (AUC, sensitivity and specificity) of the base model (as presented in the section “Ecological niche model”) against models with alternative specifications for the relationship between temperature and transmission risk<sup>80–83</sup>.

Given that *Ae. aegypti* play a minimal role in modern transmission of yellow fever virus in South America, where transmission primarily occurs through sylvatic spillover rather than urban cycle<sup>28,29</sup>, an alternative version of the yellow fever arbovirus model was fit without the *Ae. aegypti* covariate. The improvement of this model was assessed by using a 50-fold block cross validation approach and comparing overall and South America-specific model performance metrics with those of the base model.

To assess the benefits of joint modelling over disease-specific models, we re-ran individual models for dengue, chikungunya, Zika, and yellow fever separately and compared the model performance metrics (AUC, sensitivity, and specificity) with those of the joint model. Model performance was evaluated using 10-fold cross-validation with metrics calculated based on a cutoff value where the sum of sensitivity and specificity is maximised.

### Comparison between diseases and with previously published maps

We obtained previously published suitability maps from Messina et al.<sup>17</sup> for dengue, Nsoesie et al.<sup>18</sup> for chikungunya, Messina et al.<sup>19</sup> for Zika, and Shearer et al.<sup>20</sup> for yellow fever. These maps were chosen because they represent recent global predictions of disease risk and also utilise occurrence data and species distribution modelling approaches<sup>57</sup>, making a more useful assessment of the advances made in this work. The same threshold calculation above was performed for the previous suitability maps using the same set of presence and background points to convert them into binary maps. This was a necessary calibration step to standardise the scales of predicted suitability. Once the binary maps of at-risk areas were constructed, our map and previously published maps were compared and every  $5 \times 5$  km pixel was categorized into four distinct groups: those at-risk in both our map and previous maps, those at-risk only in previous maps, those at-risk only in our map, and areas deemed not at-risk in either.

## Independent validation

As part of the validation process, the initial versions of the suitability maps were presented in June 2023<sup>84</sup> to the Technical Advisory Group on Arbovirus (TAG-Arbovirus<sup>85</sup>), a multidisciplinary group of experts supported by WHO Arbovirus Secretariat. Following the presentation, we invited feedback by providing the TAG with an interactive map of occurrence records for each disease ([https://ahyounglim.shinyapps.io/multi\\_arbo\\_mapping/](https://ahyounglim.shinyapps.io/multi_arbo_mapping/)) and a concise questionnaire (Supplementary Table 5). This questionnaire addressed specific questions, particularly regarding any discrepancies between our model estimates and their context-specific knowledge, as well as additional drivers of arbovirus risk that should be taken into consideration. This resulted in two main changes to our models: i) identified additional data sources for chikungunya occurrences in Brazil<sup>86</sup>; ii) included additional covariates, such as *Haemagogus janthinomys*<sup>71</sup> for yellow fever model and under-five mortality, government effectiveness, and physicians density for the surveillance capability model (see section “Covariates”).

## Reporting summary

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

## Data availability

Disease occurrence data is available from previous publications on dengue<sup>17</sup>, chikungunya<sup>18,38</sup>, Zika<sup>19</sup>, and yellow fever<sup>20</sup>. Additional data was extracted from publicly available sources, including WHO regional outbreak updates (<https://www.who.int/emergencies/disease-outbreak-news>); ECDC website (<https://www.ecdc.europa.eu/en/>); and ProMED mail reports (<http://www.promedmail.org>) and the HealthMap platform ([www.healthmap.org](http://www.healthmap.org)). The maps with administrative boundaries were created using public-domain Natural Earth data, accessed via the `rnaturalearth` package in R<sup>86</sup>.

Climate and environmental covariates are freely available from previous publications (GDP<sup>58</sup>, urbanisation<sup>59</sup>, temperature suitability<sup>66</sup>, treatment-seeking for fever<sup>61</sup>, dynamic habitat indices<sup>69</sup>, and yellow fever vaccination coverage<sup>73</sup>). Surface travel time covariates are available from the Malaria Atlas Project (<https://data.malariaatlas.org/maps>). Child mortality, physicians density, and the government effectiveness estimates can be freely downloaded via European Commission Disaster Risk Management Knowledge Centre (<https://drmkc.jrc.ec.europa.eu/inform-index/>). High resolution population data can be freely obtained from LandScan programme (<https://landscan.ornl.gov/about>). Global climate data can be downloaded from TerraClimate (<https://www.climatologylab.org/terraclimate.html>). Normalised Difference Vegetation Index data is freely available from NASA Earth Observation Data (<https://www.earthdata.nasa.gov/>) and can be downloaded using R `MODIS` package (<https://github.com/ropensci/MODISr>). Predicted suitability for *Ae. albopictus*, *Ae. aegypti*, *Haemagogus janthinomys*, and non-human primates are not publicly available but can be obtained by contacting the authors of the cited papers<sup>20,69,71</sup>. A detailed description of data sources can be found in Supplementary Tables 1 and 4.

Processed versions of these datasets used in our analyses are available in two repositories: the study github repository ([https://github.com/ahyoung-lim/Arbo\\_riskmaps\\_public](https://github.com/ahyoung-lim/Arbo_riskmaps_public)) for past and current versions, and the Figshare repository (<https://doi.org/10.6084/m9.figshare.26172934>) for the version that has been peer-reviewed and described in this article.

## Code availability

Data analyses were carried out using R version 4.3.0 using the following packages: `doParallel` (v1.0.17), `foreach` (v1.5.2), `pdp` (v0.8.1), `matrixStats` (v1.3.0), `Metrics` (v0.1.4), `cutpointr` (v1.1.2), `boot` (v1.3-28.1), `blockCV` (v3.1-3), `rsample` (v1.2.0), `randomForest` (v4.7-1.1), `terra` (v1.7-29), `rnaturalearth` (v0.3.4), `raster` (v3.6-23), `sp` (v1.6-0), `exactextractr` (v0.9.1), `sf`

(v1.0-14), `tidyterra` (v0.5.2), `countrycode` (v1.5.0), `dplyr` (v1.1.2), `data-table` (v1.14.8). All code and processed datasets used for the analyses are publicly available online in the study github repository ([https://github.com/ahyoung-lim/Arbo\\_riskmaps\\_public](https://github.com/ahyoung-lim/Arbo_riskmaps_public)) for past and current versions and the Figshare repository (<https://doi.org/10.6084/m9.figshare.26172934>) for the version that has been peer-reviewed and described in this article.

## References

- GBD 2021 Risk Factors Collaborators. Global burden and strength of evidence for 88 risk factors in 204 countries and 811 subnational locations, 1990–2021: a systematic analysis for the global burden of disease study 2021. *Lancet* **403**, 2162–2203 (2024).
- World Health Organization. Dengue- Global situation. <https://www.who.int/emergencies/disease-outbreak-news/item/2023-DON498> (2023).
- Munayco, C. V. et al. Notes from the field: dengue outbreak - Peru, 2023. *MMWR Morb. Mortal. Wkly Rep.* **73**, 86–88 (2024).
- Mahmud, A. S., Bhattacharjee, J., Baker, R. E. & Martinez, P. P. Alarming trends in dengue incidence and mortality in Bangladesh. *J. Infect. Dis.* **229**, 4–6 (2024).
- Pan American Health Organization / World Health Organization (PAHO/WHO). dengue epidemiological situation in the region of the Americas - epidemiological week 08, 2024. <https://www.paho.org/en/documents/situation-report-no-9-dengue-epidemiological-situation-region-americas-epidemiological> (2024).
- Ferreira de Almeida, I. et al. The expansion of chikungunya in Brazil. *Lancet Reg. Health Am.* **25**, 100571 (2023).
- Torales, M. et al. Notes from the field: chikungunya outbreak - Paraguay, 2022–2023. *MMWR Morb. Mortal. Wkly Rep.* **72**, 636–638 (2023).
- ECDC. Autochthonous vectorial transmission of dengue virus in mainland EU/EEA, 2010-present. <https://www.ecdc.europa.eu/en/all-topics-z/dengue/surveillance-and-disease-data/autochthonous-transmission-dengue-virus-eueea> (2023).
- Castro, L. A. et al. Assessing real-time Zika risk in the United States. *BMC Infect. Dis.* **17**, 284 (2017).
- Garske, T. et al. Yellow fever in Africa: estimating the burden of disease and impact of mass vaccination from outbreak and serological data. *PLoS Med.* **11**, e1001638 (2014).
- Gaythorpe, K. A. et al. The global burden of yellow fever. *eLife* **10**, <https://doi.org/10.7554/eLife.64670> (2021).
- Liu-Helmersson, J., Brännström, Å., Sewe, M. O., Semenza, J. C. & Rocklöv, J. Estimating past, present, and future trends in the global distribution and abundance of the arbovirus vector *Aedes aegypti* under climate change scenarios. *Front. Public Health* **7**, 148 (2019).
- Kolimenakis, A. et al. The role of urbanisation in the spread of *Aedes* mosquitoes and the diseases they transmit—a systematic review. *PLoS Negl. Trop. Dis.* **15**, e0009631 (2021).
- Lee, S. A., Jarvis, C. I., Edmunds, W. J., Economou, T. & Lowe, R. Spatial connectivity in mosquito-borne disease models: a systematic review of methods and assumptions. *J. R. Soc. Interface* **18**, 20210096 (2021).
- Balakrishnan, V. S. WHO launches global initiative for arboviral diseases. *Lancet Microbe* **3**, e407 (2022).
- Diggle, P. J. & Giorgi, E. *Model-based Geostatistics for Global Public Health: Methods and Applications*. (Chapman and Hall/CRC, 2019).
- Messina, J. P. et al. The current and future global distribution and population at risk of dengue. *Nat. Microbiol.* **4**, 1508–1515 (2019).
- Nsoesie, E. O. et al. Global distribution and environmental suitability for chikungunya virus, 1952 to 2015. *Euro Surveill.* **21**, <https://doi.org/10.2807/1560-7917.es.2016.21.20.30234> (2016).
- Messina, J. P. et al. Mapping global environmental suitability for Zika virus. *eLife* **5**, e15272 (2016).

20. Shearer, F. M. et al. Existing and potential infection risk zones of yellow fever worldwide: a modelling analysis. *Lancet Glob. Health* **6**, e270–e278 (2018).
21. Kraemer, M. U. G. et al. Past and future spread of the arbovirus vectors *Aedes aegypti* and *Aedes albopictus*. *Nat. Microbiol.* **4**, 854–863 (2019).
22. Mordecai, E. A. et al. Detecting the impact of temperature on transmission of Zika, dengue, and chikungunya using mechanistic models. *PLoS Negl. Trop. Dis.* **11**, e0005568 (2017).
23. Messina, J. P. et al. A global compendium of human dengue virus occurrence. *Sci. Data* **1**, 140004 (2014).
24. Messina, J. & Shearer, F. Global compendium of human Zika virus occurrence. *Figshare* <https://doi.org/10.6084/m9.figshare.2573629.v1> (2016).
25. Clarke, J. et al. A global dataset of publicly available dengue case count data. *Sci. Data* **11**, 296 (2024).
26. Eltom, K., Enan, K., El Hussein, A. R. M. & Elkhidir, I. M. Dengue virus infection in sub-saharan Africa between 2010 and 2020: a systematic review and meta-analysis. *Front. Cell. Infect. Microbiol.* **11**, 678945 (2021).
27. Hill, S. C. et al. Climate and land-use shape the spread of zoonotic yellow fever virus. *medRxiv* <https://doi.org/10.1101/2022.08.25.22278983> (2022).
28. Abreu, F. V. Sde et al. Ecological and environmental factors affecting transmission of sylvatic yellow fever in the 2017-2019 outbreak in the Atlantic Forest, Brazil. *Parasit. Vectors* **15**, 23 (2022).
29. Chippaux, J.-P. & Chippaux, A. Yellow fever in Africa and the Americas: a historical and epidemiological perspective. *J. Venom. Anim. Toxins Incl. Trop. Dis.* **24**, 20 (2018).
30. Jentes, E. S. et al. The revised global yellow fever risk map and recommendations for vaccination, 2010: consensus of the Informal WHO working group on geographic risk for yellow fever. *Lancet Infect. Dis.* **11**, 622–632 (2011).
31. Eliminate Yellow fever Epidemics (EYE). a global strategy, 2017–2026. *Wkly Epidemiol. Rec.* **92**, 193–204 (2017).
32. Rufalco-Moutinho, P. et al. Evidence of co-circulation of multiple arboviruses transmitted by *Aedes* species based on laboratory syndromic surveillance at a health unit in a slum of the Federal District, Brazil. *Parasit. Vectors* **14**, 610 (2021).
33. World Health Organization. *Surveillance and control of arboviral diseases in the WHO African region: assessment of country capacities*. <https://www.who.int/publications/i/item/9789240052918> (2022).
34. Chen, Y. et al. Measuring the effects of COVID-19-related disruption on dengue transmission in southeast Asia and Latin America: a statistical modelling study. *Lancet Infect. Dis.* **22**, 657–667 (2022).
35. Gaythorpe, K. A. et al. Impact of COVID-19-related disruptions to measles, meningococcal A, and yellow fever vaccination in 10 countries. *eLife* **10**, (2021).
36. World Health Organization. Yellow fever – African Region (AFRO). <https://www.who.int/emergencies/disease-outbreak-news/item/2024-DON510> (2024).
37. Pougou, N. et al. Mosquito-borne arboviruses occurrence and distribution in the last three decades in central Africa: a systematic literature review. *Microorganisms* **12**, 4 (2023).
38. de Souza, W. M. et al. Spatiotemporal dynamics and recurrence of chikungunya virus in Brazil: an epidemiological study. *Lancet Microbe* **4**, e319–e329 (2023).
39. Bhatt, S. et al. The global distribution and burden of dengue. *Nature* **496**, 504–507 (2013).
40. Carrion, M. & Madoff, L. C. ProMED-mail: 22 years of digital surveillance of emerging infectious diseases. *Int. Health* **9**, 177–183 (2017).
41. Bijkerk, P. et al. ECDC Round Table Report and ProMed-mail most useful international information sources for the Netherlands Early Warning Committee. *Euro Surveill.* **22**, (2017).
42. ECDC. Chikungunya worldwide overview. <https://www.ecdc.europa.eu/en/chikungunya-monthly> (2023).
43. ECDC. Autochthonous transmission of chikungunya virus in mainland EU/EEA, 2007–present. <https://www.ecdc.europa.eu/en/infectious-disease-topics/z-disease-list/chikungunya-virus-disease/surveillance-threats-and> (2019).
44. ECDC. Zika virus disease - Annual Epidemiological Report for 2021. <https://www.ecdc.europa.eu/en/publications-data/zika-virus-disease-annual-epidemiological-report-2021> (2023).
45. Pan American Health Organization / World Health Organization (PAHO/WHO). Epidemiological update - yellow fever in the region of the Americas - 25 April 2023. <https://www.paho.org/en/documents/epidemiological-update-yellow-fever-region-americas-25-april-2023> (2023).
46. Disease Outbreak News. <https://www.who.int/emergencies/disease-outbreak-news> (2024).
47. Wu, D. et al. Chikungunya outbreak in Guangdong Province, China, 2010. *Emerg. Infect. Dis.* **18**, 493–495 (2012).
48. Yu, Y., Liu, Y., Ling, F., Sun, J. & Jiang, J. Epidemiological characteristics and economic burden of dengue in Zhejiang province, China. *Viruses* **15**, 1731 (2023).
49. Sang, S., Yue, Y., Wang, Y. & Zhang, X. The epidemiology and evolutionary dynamics of massive dengue outbreak in China, 2019. *Front. Microbiol.* **14**, 1156176 (2023).
50. Wang, J. Z. et al. Frequent Import and multiple sources of dengue fever have changed the epidemic situation of the disease in Fujian Province, China. *Biomed. Environ. Sci.* **33**, 123–132 (2020).
51. Ren, J. et al. Epidemiology of Indigenous dengue cases in Zhejiang province, southeast China. *Front. Public Health* **10**, 857911 (2022).
52. Cuervo, P. F., Artigas, P., Lorenzo-Morales, J., Bargues, M. D. & Mas-Coma, S. Ecological niche modelling approaches: challenges and applications in vector-borne diseases. *Trop. Med. Infect. Dis.* **8**, 187 (2023).
53. Steen, V. A., Tingley, M. W., Paton, P. W. C. & Elphick, C. S. Spatial thinning and class balancing: key choices lead to variation in the performance of species distribution models with citizen science data. *Methods Ecol. Evol.* **12**, 216–226 (2021).
54. Freifeld, C. C., Mandl, K. D., Reis, B. Y. & Brownstein, J. S. HealthMap: global infectious disease monitoring through automated classification and visualization of Internet media reports. *J. Am. Med. Assoc.* **15**, 150–157 (2008).
55. Barbet-Massin, M., Jiguet, F., Albert, C. H. & Thuiller, W. Selecting pseudo-absences for species distribution models: how, where and how many? *Methods Ecol. Evolution* **3**, 327–338 (2012).
56. Valavi, R., Guillera-Aroita, G., Lahoz-Monfort, J. J. & Elith, J. Predictive performance of presence-only species distribution models: a benchmark study with reproducible code. *Ecological Monographs* **92**, e01486 (2022).
57. Lim, A.-Y. et al. A systematic review of the data, methods and environmental covariates used to map *Aedes*-borne arbovirus transmission risk. *BMC Infect. Dis.* **23**, 708 (2023).
58. Chen, J. et al. Global 1 km × 1 km gridded revised real gross domestic product and electricity consumption during 1992–2019 based on calibrated nighttime light data. *Sci. Data* **9**, 202 (2022).
59. Gao, J. & O'Neill, B. C. Mapping global urban land for the 21st century with data-driven simulations and Shared Socioeconomic Pathways. *Nat. Commun.* **11**, 2302 (2020).
60. Weiss, D. J. et al. Global maps of travel time to healthcare facilities. *Nat. Med.* **26**, 1835–1838 (2020).
61. Nguyen, M. et al. Trends in treatment-seeking for fever in children under five years old in 151 countries from 1990 to 2020. *PLOS Glob Public Health* **3**, e0002134 (2023).



62. Disaster Risk Management Knowledge Centre. INFORM Risk - Results and data. <https://drmkc.jrc.ec.europa.eu/inform-index/INFORM-Risk/Results-and-data/moduleId/1782/id/469/controller/Admin/action/Results> (2023).
63. United Nations Inter-agency Group for Child Mortality Estimation. *Levels and trends in child mortality: report 2022*. (2023).
64. World Bank. Physicians (per 1000 people). <https://data.worldbank.org/indicator/SH.MED.PHYS.ZS> (2023).
65. Kaufmann, D. & Kraay, A. Worldwide Governance Indicators, 2023 Update. <http://www.govindicators.org>.
66. Brady, O. J. et al. Global temperature constraints on *Aedes aegypti* and *Ae. albopictus* persistence and competence for dengue virus transmission. *Parasit. Vectors* **7**, 338 (2014).
67. Abatzoglou, J. T., Dobrowski, S. Z., Parks, S. A. & Hegewisch, K. C. TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958-2015. *Sci. Data* **5**, 170191 (2018).
68. Didan, K. MOD13C2 MODIS/Terra Vegetation Indices Monthly L3 Global 0.05Deg CMG V006. NASA EOSDIS Land Processes DAAC <https://doi.org/10.5067/modis/mod13c2.006> (2015).
69. Coops, N. C., Kearney, S. P., Bolton, D. K. & Radeloff, V. C. Remotely-sensed productivity clusters capture global biodiversity patterns. *Sci. Rep.* **8**, 16261 (2018).
70. Sims, K. et al. LandScan Global 2022. *Oak Ridge National Laboratory* <https://doi.org/10.48690/1529167> (2023).
71. Celone, M. et al. An ecological niche model to predict the geographic distribution of *Haemagogus janthinomys*, Dyar, 1921 a yellow fever and Mayaro virus vector, in South America. *PLoS Negl. Trop. Dis.* **16**, e0010564 (2022).
72. Hamlet, A. et al. POLICI: a web application for visualising and extracting yellow fever vaccination coverage in Africa. *Vaccine* **37**, 1384–1388 (2019).
73. Grimmett, L., Whitsed, R. & Horta, A. Presence-only species distribution models are sensitive to sample prevalence: evaluating models using spatial prediction stability and accuracy metrics. *Ecol. Modell.* **431**, 109194 (2020).
74. Shearer, F. M. et al. Estimating geographical variation in the risk of zoonotic plasmodium knowlesi infection in countries eliminating malaria. *PLoS Negl. Trop. Dis.* **10**, e0004915 (2016).
75. Phillips, S. J. et al. Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. *Ecol. Appl.* **19**, 181–197 (2009).
76. Warton, D. I., Renner, I. W. & Ramp, D. Model-based control of observer bias for the analysis of presence-only data in ecology. *PLoS ONE* **8**, e79168 (2013).
77. Cattarino, L., Rodriguez-Barraquer, I., Imai, N., Cummings, D. A. T. & Ferguson, N. M. Mapping global variation in dengue transmission intensity. *Sci. Transl. Med.* **12**, (2020).
78. Valavi, R., Elith, J., Lahoz-Monfort, J. J. & Guillerá-Arroita, G. blockCV: an R package for generating spatially or environmentally separated folds for k-fold cross-validation of species distribution models. *Methods Ecol. Evol.* **10**, 225–232 (2018).
79. Abrego, N. & Ovaskainen, O. Evaluating the predictive performance of presence-absence models: why can the same model appear excellent or poor? *Ecol. Evol.* **13**, e10784 (2023).
80. Ryan, S. J. et al. Warming temperatures could expose more than 1.3 billion new people to Zika virus risk by 2050. *Glob. Chang. Biol.* **27**, 84–93 (2021).
81. Ryan, S. Global current *Aedes aegypti* suitability for dengue transmission at 97.5% CI (5 arc minutes). *Harvard Dataverse* <https://doi.org/10.7910/dvn/nsg5uh> (2019).
82. Ryan, S. Global current suitability for *Aedes aegypti* transmitted Zika at 97.5% CI (5 arc minutes). *Harvard Dataverse*. <https://doi.org/10.7910/dvn/tk041g> (2020).
83. Ryan, S. J., Carlson, C. J., Mordecai, E. A. & Johnson, L. R. Global expansion and redistribution of *Aedes*-borne virus transmission risk with climate change. *PLoS Negl. Trop. Dis.* **13**, e0007213 (2019).
84. World Health Organization. Technical Advisory Group on Arbovirus (TAG) meeting, Accra, Ghana. [https://www.who.int/news/item/02-08-2023-technical-advisory-group-on-arbovirus-\(tag\)-meeting--accra-ghana](https://www.who.int/news/item/02-08-2023-technical-advisory-group-on-arbovirus-(tag)-meeting--accra-ghana) (2023).
85. World Health Organization. Technical Advisory Group on Arbovirus. <https://www.who.int/groups/technical-advisory-group-on-arbovirus> (2021).
86. South, A., Michael, S. & Massicotte, P. *naturalearthdata: World Vector Map Data from Natural Earth Used in "rnatualearth."* (2024).

## Acknowledgements

This work was discussed with the Technical Advisory Group on arboviruses (TAG-Arbovirus) and WHO yellow fever risk assessment working group. This work was supported by the World Health Organisation Global Arbovirus Initiative, partially funded through a cooperative agreement with the U.S. CDC. OJB was supported by a UK Medical Research Council Career Development Award (MR/V031112/1) which also supports AL. AL was additionally supported by the Basic Science Research Programme through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2022R1A6A3A03061207). M.U.G.K. acknowledges funding from The Rockefeller Foundation (PC-2022-POP-005), Google.org, the Oxford Martin School Programmes in Pandemic Genomics & Digital Pandemic Preparedness, European Union's Horizon Europe programme projects MOOD (#874850) and E4Warning (#101086640), Wellcome Trust grants 303666/Z/23/Z, 226052/Z/22/Z & 228186/Z/23/Z, the United Kingdom Research and Innovation (#APP8583), the Medical Research Foundation (MRF-RG-ICCH-2022-100069), UK International Development (301542-403), the Bill & Melinda Gates Foundation (INV-063472) and Novo Nordisk Foundation (NNF24OC0094346). WMdS was supported by a Global Virus Network fellowship, the Burroughs Wellcome Fund (#1022448), and a Wellcome Trust–Digital Technology Development award (Climate Sensitive Infectious Disease Modelling; 226075/Z/22/Z). CJ was supported by a Bloomsbury Colleges PhD Studentship and a National University of Ireland Denis Phelan Scholarship. SJR was supported by NSF CIBR: VectorByte: A Global Informatics Platform for studying the Ecology of Vector-Borne Diseases (NSF DBI 2016265). SJR was additionally supported by funding to Verena (viralemergence.org), including NSF BII 2021909 and NSF BII 2213854. MC was supported by Bill and Melinda Gates Foundation. KAMG received funding from Gavi [Grant ID: 226727\_Z\_22\_Z], BMGF [Grant Numbers INV-034281 and INV-009125 / OPP1157270] and/or the Wellcome Trust via the Vaccine Impact Modelling Consortium during the course of the study. KAMG also acknowledges funding from the MRC Centre for Global Infectious Disease Analysis (reference MR/X020258/1), funded by the UK Medical Research Council (MRC). This UK funded award is carried out in the frame of the Global Health EDCTP3 Joint Undertaking. KAMG reports speaker fees from Sanofi Pasteur outside the submitted work.

## Author contributions

O.J.B. and N.G. conceived and planned the study. A.L., J.C., A.G., C.J., H.K., W.Mde.S., K.S., and J.S.B. identified and obtained data for analysis. A.L., J.C., A.G., C.J., H.K., K.S., and J.S.B. extracted, processed and geopositioned the data. A.L. and O.J.B. carried out the statistical analyses with assistance and input from F.M.S., D.M.P., S.J.R. and N.G. A.L. and O.J.B. prepared the first draft of the manuscript with assistance from F.M.S., K.S., D.M.P., C.J., H.K., J.P.M., M.U.G.K., K.A.M.G., W.Mde.S., E.O.N., M.C., N.F., S.J.R., I.B.R., D.P.R., S.I.H. and N.G.

## 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/s41467-025-58609-5>.

**Correspondence** and requests for materials should be addressed to Oliver J. Brady.

**Peer review information** *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work. 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

---

<sup>1</sup>Department of Infectious Disease Epidemiology and Dynamics, Faculty of Epidemiology and Population Health, London School of Hygiene & Tropical Medicine, London, UK. <sup>2</sup>Centre for Mathematical Modelling of Infectious Diseases, Faculty of Epidemiology and Population Health, London School of Hygiene & Tropical Medicine, London, UK. <sup>3</sup>Infectious Disease Dynamics Unit, Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, Australia. <sup>4</sup>The Kids Research Institute Australia, Perth Children's Hospital, Perth, Australia. <sup>5</sup>Boston Children's Hospital, Boston, MA, USA. <sup>6</sup>Institute for Health Metrics and Evaluation, University of Washington, Seattle, WA, USA. <sup>7</sup>Department of Health Metrics Sciences, School of Medicine, University of Washington, Seattle, WA, USA. <sup>8</sup>University of Cambridge, Cambridge, UK. <sup>9</sup>Epidemiology Unit, Ministry of Health, Colombo, Sri Lanka. <sup>10</sup>School of Tropical Medicine and Global Health, Nagasaki University, Nagasaki, Japan. <sup>11</sup>School of Geography and the Environment, University of Oxford, Oxford, UK. <sup>12</sup>Oxford School of Global and Area Studies, University of Oxford, Oxford, UK. <sup>13</sup>Department of Biology, University of Oxford, Oxford, UK. <sup>14</sup>Pandemic Sciences Institute, University of Oxford, Oxford, UK. <sup>15</sup>MRC Centre for Global Infectious Disease Analysis, School of Public Health, Imperial College London, London, UK. <sup>16</sup>Department of Microbiology, Immunology and Molecular Genetics, University of Kentucky, College of Medicine, Lexington, KY, USA. <sup>17</sup>Department of Global Health, School of Public Health, Boston University, Boston, MA, USA. <sup>18</sup>Virus Genomic Epidemiology, Faculty of Medicine, School of Public Health, Imperial College, London, Imperial College London, London, UK. <sup>19</sup>Department of Geography and the Emerging Pathogens Institute, University of Florida, Gainesville, FL, USA. <sup>20</sup>Department of Epidemic and Pandemic Preparedness and Prevention, World Health Organization, Geneva, Switzerland. <sup>21</sup>Department of Pediatrics, Harvard Medical School, Boston, MA, USA. <sup>22</sup>Curtin University, Perth, Australia.

✉ e-mail: [Oliver.Brady@lshtm.ac.uk](mailto:Oliver.Brady@lshtm.ac.uk)