Past and future spread of the arbovirus vectors *Aedes aegypti* and *Aedes albopictus*

Kraemer, M.U.G.[1,2,3,$], Reiner Jr., R.C.[4,$], Brady, O.J.[5,$], Messina, J.P.[1,$], Gilbert, M.[6,7,$], Pigott, D.M.[4], Yi, D.[8], Johnson, K.[4], Earl, L.[4], Marczak, L.B.[4], Shirude, S.[4], Davis Weaver, N.[4], Bisanzio, D.[9], Perkins, T.A.[10], Lai, S.[11,12,13], Lu, X.[12,13,14,15], Jones, P.[16], Coelho, G.E.[17], Carvalho, R.G.[17], Van Bortel, W.[18,19], Marsboom, C.[20], Hendrickx, G.[20], Schaffner, F.[21], Moore, C.G.[22], Nax, H.H.[23], Bengtsson, L.[13,15], Wetter, E.[13,24], Tatem, A.J.[12,13,], Brownstein, J.S.[2,3], Smith, D.L.[4,25], Lambrechts, L.[26,27], Cauchemez, S.[28,29], Linard, C.[6, 30], Faria, N.R.[1], Pybus, O.G.[1], Scott, T.W.[31], Liu, Q.[32,33,34,35], Yu, H.[11], Wint, G.R.W.[1,36], Hay, S.I.[4,9§], Golding, N.[37§]

[1]Department of Zoology, University of Oxford, Oxford, UK
[2]Harvard Medical School, Harvard University, Boston, USA
[3]Boston Children's Hospital, Boston, USA
[4]Institute for Health Metrics and Evaluation, University of Washington, WA 98121, USA
[5]Centre for Mathematical Modelling of Infectious Diseases, London School of Hygiene and Tropical Medicine, London, UK
[6]Spatial Epidemiology Lab (SpELL), Universite Libre de Bruxelles, B-1050 Brussels, Belgium
[7]Fonds National de la Recherche Scientifique, B-1000 Brussels, Belgium
[8]Department of Statistics, Harvard University, Cambridge MA, USA
[9]Oxford Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, University of Oxford, Oxford, OX3 7LF, UK
[10]Department of Biological Sciences and Eck Institute for Global Health, University of Notre Dame, USA
[11]School of Health, Fudan University, Key Laboratory of Public Health Safety, Ministry of Education, Shanghai, China
[12]Department of Geography and Environment, University of Southampton, Southampton, UK
[13]Flowminder Foundation, Stockholm, Sweden
[14]College of Information System and Management, National University of Defense Technology, Changsha, China
[15]Department of Public Health Sciences, Karolinska Institutet, Stockholm, Sweden
[16]Waen Associates Ltd, Y Waen, Islaw'r Dref, Dolgellau, Gwynedd LL401TS, UK
[17]National Dengue Control Program, Ministry of Health, Brasilia, DF, Brazil
[18]European Centre for Disease Prevention and Control, Stockholm, Sweden
[19]Institute of Tropical Medicine, Antwerp, Belgium
[20]Avia-GIS, Zoersel, Belgium
[21]Francis Schaffner Consultancy, Riehen, Switzerland
[22]Department of Microbiology, Immunology, and Pathology, Colorado State University, Fort Collins, CO, USA
[23]Computational Social Science, ETH Zurich, Zurich, Switzerland
[24]Stockholm School of Economics, Stockholm, Sweden
[25]Sanaria Institute for Global Health and Tropical Medicine, Rockville, USA
[26]Insect-Virus Interactions Group, Department of Genomes and Genetics, Institut Pasteur, Paris 75015, France
[27]Centre National de la Recherche Scientifique, Unité de Recherche Associée 3012, Paris 75015, France
[28]Mathematical Modelling of Infectious Diseases and Center of Bioinformatics, Biostatistics and Integrative Biology, Institut Pasteur, Paris, France
[29]Centre National de la Recherche Scientifique, URA3012, Paris, France
[30]Department of Geography, Universite de Namur, Belgium
[31]Department of Entomology and Nematology, University of California, Davis, USA
[32]State Key Laboratory for Infectious Disease Prevention and Control, National Institute for Communicable Disease Control and Prevention, Chinese Center for Disease Control and Prevention, Changping, Beijing 102206, China
[33]Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, Hangzhou 310003, China
[34]WHO Collaborating Centre for Vector Surveillance and Management, 155 Changbai Road, Changping, Beijing 102206, China

[35]Centre for Environment and Population Health, Nathan Campus, Griffith University, 170 Kessels Road, Queensland 4111, Nathan, QLD, Australia
[36]Environmental Research Group Oxford (ERGO), Department of Zoology, Oxford University, Oxford, UK
[37]School of BioSciences, University of Melbourne, VIC 3010, Australia

Correspondence to:
Moritz U G Kraemer, DPhil
Department of Zoology, University of Oxford
Oxford, OX13SP, United Kingdom
Email: moritz.kraemer@zoo.ox.ac.uk

Nick Golding, DPhil
School of BioSciences, University of Melbourne
VIC 3010, Australia
nick.golding.research@gmail.com

Prof. Simon I Hay, DSc
Institute for Health Metrics and Evaluation
University of Washington,
WA 98121, United States
sihay@uw.edu

[$] Contributed equally to this work as first authors

[§]These authors jointly supervised this work

One Sentence Summary

Human mobility patterns and climate changes predict the spread of the arbovirus vectors *Aedes aegypti* and *Ae. albopictus*, which transmit viruses such as dengue, yellow fever, chikungunya, and Zika.

Abstract

The global population at risk from mosquito-borne diseases – including dengue, yellow fever, chikungunya, and Zika – is expanding in concert with changes in the distribution of two key vectors, *Aedes aegypti* and *Ae. albopictus*. The distribution of these species is largely driven by both human movement and the presence of suitable climate. Using statistical mapping techniques, we show that human movement patterns explain the spread of both species in Europe and the United States of America (USA) following their introduction. We find that the spread of *Ae. aegypti* is characterised by long distance importations, whilst *Ae. albopictus* has expanded more along the fringes of its current distribution. We describe these processes and predict the future distributions of both species in response to accelerating urbanisation, connectivity, and climate change. Global surveillance and control efforts that aim to mitigate the spread of chikungunya, dengue, yellow fever and Zika viruses must consider the so far unabated spread of these mosquitos. Our maps and predictions offer an opportunity to strategically target surveillance and control programs and thereby augment efforts to reduce arbovirus burden in human populations globally.

Main text

The geographical distributions of the arboviruses dengue, yellow fever, chikungunya, and Zika have expanded, causing severe disease outbreaks in many urban populations.[1–5] Transmission of

these viruses depends, with few exceptions, on the presence of the competent mosquito vectors *Aedes aegypti* and *Ae. albopictus*[6,7]. Previous predictions of the future distributions of *Aedes aegypti [=Stegomyia aegypti]* and *Ae. albopictus [=Stegomyia albopicta]* have focussed solely on climate, despite the known importance of urbanisation and other socioeconomic factors in defining suitable habitat[8]. Moreover, those projections assumed that both species can fully infest all areas of predicted newly suitable habitat[4,9]. Recent trends in the global spread of these species, however, suggest that the process of expansion may be more complex and spatially structured than previously acknowledged[10]. Expansion from the native ranges in *Ae. aegypti* (from African forests) and *Ae. albopictus* (from Asia) was precipitated by a shift from zoophily to anthropophily and by adaptation to container-breeding in domestic or peri-domestic environments[11,12]. Whilst their short flight ranges limit self-powered dispersal[13]. A century of rapid human population growth and international trade has enabled their global spread. Trade in items that provide potential larval development habitats such as tires and potted plants led to inter-continental dissemination of their desiccation-resistant eggs[14–16]. Moreover, the establishment of *Ae. albopictus* in locations with cooler climates has been aided by its ecological plasticity, with eggs able to undergo diapause (dormancy) as one possible explanation for populations persisting through winters too cold for adult survival[17,18].

Whilst the various routes of inter-continental importation are well described[11,19], the processes underlying intra-continental spread of the species remain poorly quantified, preventing informed prediction of future distributions. Modelling of human-mediated range expansion suggests that quantitative models of human movement could, and should, be used to predict intra-continental spread[20–22]. To address this, we developed predictive models of *Ae. aegypti* and *Ae. albopictus* spread and combined these with forecasts of future climatic conditions and urban growth, to

predict the ranges of these medically important vectors from 2015 to 2080 (Extended Data Fig. 1).

We collated spatially- and temporally-explicit data on the distributions of *Ae. aegypti* and *Ae. albopictus* and their spread over time in the USA, and *Ae. albopictus* in Europe (Fig. 1, Extended Data Figs. 2, 3). Extending a previous study[4], we first mapped contemporary habitat suitability for each species together with projected suitability in 2020, 2050, and 2080, under three different Representative Concentration Pathway (RCP) and 17 global climate models (GCMs), as well as under projections of urban growth. We then parameterised quantitative models of human mobility using census data on migration and commuting patterns[23,24], and general movement patterns derived from mobile phone logs (call detail records) (Extended Data Fig. 1)[23–25]. The combined predictions from these different mobility models and datasets capture different aspects of human travel and trade, and their ability to spread *Aedes* eggs and juveniles at different spatial scales.

We tabulated annualised presence records which documented the first detection of each species in 1,567 different locations over 38 years in Europe (225 / 1,588 districts, between 1979 - 2016) and 32 years in the USA (1,342 / 3,134 counties, between 1985 and 2016) (Extended Data Fig. 2a, b, c). These data were used to parameterise statistical models of spatial spread for each species. Detection within a given area was modelled as a function of i) the receptivity of the area (as determined by the habitat suitability models), ii) long-distance importation pressure (from multiple human movement models) and iii) short-distance importation pressure from adjacent areas (to represent natural dispersal). Forward simulation of these fitted models of spatial spread was then used to predict the future spread or recession of each species, considering climate changes, urbanisation, and human-mediated importation. To account for potentially biased

sampling procedures we performed a comprehensive sensitivity analysis assuming different levels of detection for both species (Supplementary Information).

Short-range importation between adjacent districts played a greater role in the inferred spread process for *Ae. albopictus* (Fig. 1a, c, d, f) than for *Ae. aegypti* (Fig. 1b, e), which was more frequently imported over longer distances. Historically, most of the observed range expansion of *Ae. aegypti* in the USA originated from southern States (Fig. 1b, Extended Data Fig. 2b). Using thin plate spline regression, we estimated the localised invasion velocity of *Ae. aegypti* spread in the USA to be relatively homogeneous at ~250km per year (Fig. 1b, e). *Aedes albopictus* spread in the USA was fastest between 1990 and 1995 (Fig. 1a, d) and has since slowed to about ~60km per year. In contrast, the estimated rate of spread of *Ae. albopictus* in Europe is faster (~100km per year) rising to ~150km per year over the last five years (Fig. 1c, f, Extended Data Fig. 2c, f, i). The geographic origin of recent *Ae. albopictus* spread in Europe seems to be Italy, with the Alps serving as a dispersal barrier that lowers rates of spread (Extended Data Fig. 2c, f). Once that barrier has been overcome, however, spread rates beyond the Alps are as high as in Italy. This may explain the increased rate of spread in recent years, which also corresponds to the detection of *Ae. albopictus* in areas north of the Alps (Extended Data Fig. 2c, f).

Using human-mobility-driven statistical models we can predict the past spread of both mosquito species with high reliability (Extended Data Fig. 6) and accuracy (out of sample area under the receiver operating characteristic curve [AUC]: 0.7-0.9, Extended Data Fig. 7). Only slight improvements are observed when including human mobility models over models that only included distance and adjacency metrics (Supplementary Information, Extended Data Fig. 12). Further, we evaluated our models' ability to predict the range expansion in Europe using a model fitted to US data (1,149 records) only. This test similarly documented a high degree of predictive

ability (out of sample AUC: 0.8-0.9, Extended Data Fig. 8). In addition, country borders seem not to limit the spread of the mosquitoes (Extended Data Fig. 11) and our spread model is robust even under different assumptions in mosquito sampling strategies but the underlying observational data may impact our estimates of velocity of spread (Supplementary Information). In contrast, the model fitted to only European data was unable to predict the spread in the USA, presumably because of the relatively few *Ae. albopictus* records in Europe compared to the USA (192 records). Therefore we used the model fitted to USA data to project the range of both species into the future (Supplementary Information). Both *Ae. aegypti* and *Ae. albopictus* are anticipated to continue expanding beyond their current distributions (Extended Data Figs. 4, 5). For *Ae. aegypti*, predicted future spread is mostly concentrated within its tropical range and in new temperate areas in the USA and China; reaching as far north as Chicago and Shanghai by 2050 (Figs. 2, 4, Extended Data Fig. 4). At the expansion front in the United States, our model predicts the spread to occur mostly through long-distance introductions in large urban areas (Figs. 2a, b, Extended Data Fig. 10). Even under the most extreme scenarios (RCP8.5 in 2080), *Ae. aegypti* is predicted to establish in Europe in only a few isolated regions of southern Italy and Turkey (Extended Data Fig. 4). By 2080 we predict there will be 159 countries worldwide (range 156 – 162) reporting this species, of which three (range 0-6) will be reporting it for the first time (Extended Data Tab. 8).

By contrast, *Ae. albopictus* is expected to spread broadly through Europe, ultimately reaching wide areas of France and Germany (Fig. 3b). Areas in northern USA and highland regions of South America and East Africa are also projected to see establishment of *Ae. albopictus* over the next 30 years (Figs. 2, 4). At the same time, some areas are predicted to become less suitable for the species, particularly locations in central southern USA (Fig. 2, Extended Data Fig. 5) and

Eastern Europe (Fig. 3) where climate models indicate aridity will increase. Due to *Ae. albopictus* broader distribution in northern latitudes, as in the USA, the spread pressure follows a clear front-like expansion (Figs. 2c, d). In total, 197 countries (range 181-209) are expected to report *Ae. albopictus* by 2080, 20 (range 4-32) of those countries will be reporting its presence for the first time (Extended Data Tab. 8).

Spread of both species over the next 5-15 years is predicted to occur independently of extensive environmental changes as both species continue to expand into their anthropogenic ecological niches through spatial dispersal. *Aedes albopictus* is anticipated to saturate its ecological niche between 2030 and 2050 (Figs. 4d, f), and *Ae. aegypti* by 2020 (Figs. 4a,c). Beyond these dates the predicted expansion of these species will be driven primarily by environmental changes that create new habitat, including changes in climate, especially temperature (Extended Data Tab. 1, 2), as well as exploitation of the increased availability of large human urban environments. Thus efforts to curb or reverse climate change are predicted to be insufficient to prevent fully the expansion of these vector species; significantly greater expansion, however, is predicted, especially between 2050 and 2080, if emissions are not reduced (Fig. 4). At the same time, future human population growth is expected to be concentrated disproportionately within areas where *Ae. aegypti* and *Ae. albopictus* already will be established, leading to large increases in the global population at risk of diseases transmitted by these species.

Overall our predicted expansions will see *Ae. aegypti* invading an estimated 19.96 million km^2 by 2050 (19.91 – 23.45 million km^2, depending on the climate and urbanisation scenarios), placing an estimated 49.13% (48.23 – 58.10%) of the world's population at risk of arbovirus transmission (Figs. 4c, f).

Few countries conduct routine, systematic surveillance for *Ae. aegypti* and *Ae. albopictus*. Consequently our analysis relies on datasets from the USA and Europe that contain spatio-temporal biases in reporting (Extended Data Fig. 2), with an implicit assumption that the processes driving spread in these regions apply elsewhere. These regions have (i) a comparatively high capacity to track establishment and mitigate the spread of these species and (ii) openly available datasets on human movement[26]. Our modeled rate of spread is thus most likely to be biased towards an underestimate of the global rate of spread (Supplementary Information). We did not model potential changes in human mobility which could increase the rate of spread of both species as population mobility increases. Competitive displacement may occur between these two species but this possibility could not be included in this analysis due to a lack of available data[27,28]. However, current ecological literature and ecological theory suggests interspecific competition occurs primarily at localized spatial scales and has not been found to influence species' distributions at a coarser spatial resolution, such as the scale we consider here[29–31]. As both species are already established on every human-inhabited continent on the planet, we did not model spread between continents.

In the context of predicting mosquito-borne viral transmission, *Aedes* distribution maps have already been shown to help predict the local[32], regional[33,34], and international[1,2,6,7,35,36] spread of chikungunya, dengue, yellow fever and Zika viruses. Moreover, local outbreaks of these arboviruses have typically followed within 5-15 years of infestation by *Ae. aegypti* and *Ae. albopictus*, emphasising the importance of vector spread importation as a key risk factor for arbovirus transmission.

There is significant uncertainty surrounding future predictions of changes in climatic conditions. We used an ensemble approach to propagate the uncertainty from climate scenarios through our predictions of both *Aedes* species (Figs. 2, 3, 4, Extended Data Figs. 4, 5).

Even under current climate conditions and population densities, both vector species will continue to spread globally over the coming decades, filling unoccupied suitable habitats and posing a risk to human health in the majority of locations where they survive and reproduce. Thus efforts to prevent their global dissemination in the near future will be most effective if focussed on preventing human-mediated spread and establishment. To prevent introductions, countries should strengthen entomological surveillance, particularly around high-risk introduction routes such as ports and highways and develop rapid response protocols for vector control to prevent introduced mosquitoes from establishing permanent populations[37–41]. We expect such efforts will need to intensify over time as human populations become ever more connected and urban agglomerations grow further[9].

Beyond 2030 and especially 2050, the distributions of both species will continue to expand, co-inciding with niche expansion into climatically suitable urban areas as opposed to the exploration of the current niche. Increased urbanisation worldwide has already put great strains on our ability to prevent the spread of certain disease vectors and has intensified endemic transmission of arboviruses[42]. Some areas may become less suitable for human habitation due to climate change impacts, reducing the number of people living in areas at risk. In the longer term, reducing emission of greenhouse gases would be desirable to limit the increase in *Ae. aegypti* and *Ae. albopictus* suitable habitat. Every effort must be made to limit factors that contribute to the global spread of *Ae. aegypti* and *Ae. albopictus* if we are to limit the future burden of the diseases vectored by these mosquitoes.

257

258

Methods

We used a combination of two approaches to estimate the predicted future distribution of *Ae. aegypti* and *Ae. albopictus:* (1) projecting the environmental suitability of both species using a set of seven environmental covariates and (2) simulating the spread within each continent using the species' past dispersal patterns, human movement data, and between region adjacency matrices (Extended Data Fig. 1). Here we describe the models and data sources for both processes.

1. Data

1.1. Global mosquito occurrence data

We used a previously collated database of 19,930 and 22,137 geopositioned occurrence records for *Ae. aegypti* and *Ae. albopictus* respectively (Extended Data Fig. 3)[43]. Each of these records corresponds to a unique detection of a mosquito population in a given location at a given point in time, as described in detail elsewhere[43]. We excluded records that were classified as temporary presence when such information was available.

1.2. Environmental and socio-economic covariates

Aedes survival is influenced by a variety of climatic and environmental factors such as long term and inter-annual temperature[44,45], water availability (described as relative humidity and precipitation), and degree of urbanisation. We used projections from the "Representative Concentration Pathways" (RCP) developed by the Intergovernmental Panel on Climate Change (IPCC)[46] which represent different assumptions about emission scenarios that might result in a variety of climatic changes over the next 65 years. Here we use RCPs 4.5, 6.0 and 8.5, which

assume emission peaks around 2040, 2080 and increases throughout the 21st century respectively[46]. These time points were chosen because (i) 2020 represents the date when the climate mitigating policies of the Paris Agreement within the United Nations Framework Convention on Climate Change (UNFCCC) will come into action[47], (ii) 2080 corresponds to the date of the emission peaks modelled according to the RCP 6.0 scenario and (iii) 2050 represents the midpoint between these dates. We use an ensemble of 17 GCMs and pattern scaling to produce monthly mean values of maximum and minimum temperature and monthly totals of rainfall as used in MarkSim. Humidity data were calculated from temperature estimates (see details in section 3). To complement the changes in temperature, relative humidity, and precipitation, we modelled a continued process of global urbanisation until 2080 using a probabilistic machine learning algorithm based on Linard et al[48]. Here we use urban growth rates projected by the United Nations as a predictor variable[49] as well as a range of other critical covariates, as described in van Vuuren et al[48].

1.3. Mosquito spatial spread data

A unique set of time-series occurrence records for both species were abstracted from Kraemer et al.[4,43], and updated with records obtained from Hahn et al[50]. Records were available for *Ae. aegypti* in the United States from 1995 – 2016 with United States county-specific information regarding whether the species was present or absent; for *Ae. albopictus* information was available from the United States (1987 – 2013) and from Europe (1979 -2017) (Fig. 1, Extended Data Fig. 2). We considered these time periods because they show consistent expansion of the species distribution as described in Hahn et al[50].

For the United States, counties were identified as reporting presence of either species in a given year if at least one specimen of any life stage of the mosquito was collected, using any collection

method[50]. Sampling efforts, techniques and temporal resolution were heterogeneous across counties and states in the United States. Therefore, the baseline presence datasets may classify some areas as absent where either of the two *Aedes* species considered may be present.

For Europe, Administrative/Statistical units (NUTS3) were identified as reporting establishment of either species in a given year if immature stages and overwintering were observed, using any collection method. Sampling efforts, techniques, and temporal resolution were heterogeneous across countries and either species may have been absent before investigations were triggered by citizen complaint. Therefore, dates correspond to published reports or expert-shared data (VBORNET, VectorNet), and a species could have established earlier in some locations where regular surveillance had not been implemented. Because we were not able to quantify the sampling biases, we instead employed a sensitivity analysis to account for potential under- or over-reporting (see section 2.4).

1.4. Human mobility datasets

Overland human movements are known to drive the importation of both species[38,39,41]. Therefore we used human movement data to infer the connectivity between regions as a proxy for importation risk of *Ae. aegypti* and *Ae. albopictus*.

US commuting data: For the United States, where both species have been spreading successfully, we obtained data on workforce commuting flows from county to county between 2009 – 2013, conducted by the American Community Survey (ACS). Data are freely available at http://www.census.gov/hhes/commuting/. Here, commuting was defined as a worker's travel between home and workplace, where the latter refers to the geographical location of the worker's job. Daytime population refers to the estimated number of people who are residing and working in an area during "daytime working hours". The data represent 3,134 counties including 50 states

and the District of Columbia (DC) but excluding Puerto Rico. The generalisability of this data has been demonstrated in studies that have successfully approximated human movements derived from mobile phone data and predicted the spread of infectious diseases[24]. As described below in section 2.3 in detail, we considered gravity and radiation movement models as well as nearest neighbour-type movements for human movement. We used the fitted models from the USA to extrapolate to all other regions in the Americas using the movement package in R[51].

European mobile phone data: For Europe, we obtained mobile phone data (or call detail records, or CDRs) from three different countries where *Ae. albopictus* is present or has recently been detected: France[43], Portugal[52], and Spain[43]. CDR data contain the time at which a call was made or a text message was sent, the duration of the call, and the code of the cell in which communication started. The cell corresponds to an area covered by a specific mobile phone tower that serves a particular area. This means that the spatial resolution is restricted to the tower area, the specific location of each individual in the dataset cannot be ascertained. As our analysis was performed at the district level, all users' activity profiles were aggregated up to the district level, which is generally larger than cell tower areas. We thereby obtained a connectivity matrix that shows the connections made between each district *i* to each district *j* within each respective country.

For Portugal, data were available from over one million mobile phone users between April 2006 and March 2007 (12 months). In Spain, CDRs were extracted from 1,034,430 users over three months between November 2007 and January 2008. In France we had the largest sample of 5,695,974 users, collected between September 2007 and mid-October 2007 covering the entire country. Other aspects of the collection and processing methods have been described in detail

348 elsewhere[23]. We used the fitted models from Europe to extrapolate to all other regions in Europe,
349 using the movement package in R[51].

350 Human movement data for Asia: Mobility matrices for Asia are inferred from data from Chinese
351 users of Baidu, the largest location-based service (LBS) in China. Baidu offers a large variety of
352 apps and software for mobile devices and personal computers, mostly for online searching. We
353 extracted GPS data from 23 April 2013 to 30 April 2014 (about 400 million users in China). The
354 raw data was collected at the county level (n = 2,959) and aggregated to the prefecture level (345
355 prefectures). We then estimated daily flows of people between each pair of counties and
356 aggregated this information per year. Movement is recorded in the Baidu data such that on each
357 day if a user was observed at locations A->B->C, then A->B and A->C are counted which may
358 produce biased population flow estimates. To explore potential bias in the data we compared the
359 data derived from Baidu to a complete dataset of taxi-based GPS locations in the capital city of
360 Hunan province, covering a one week period (full details below). The correlation of origin-to-
361 destination flows in the city between the Baidu data and the complete taxi GPS data was very
362 high (R2 = 0.99).

363 Baidu data validation: To verify the validity of the Baidu LBS data, we obtained a complete
364 dataset of GPS locations for all taxis in Changsha city (capital of Hunan Province, population: 7
365 million) in 2014. The location of each taxi is recorded for regulatory reasons using a GPS device
366 in each taxi. The location is updated every 30 seconds. There were approximately 7,000 taxis in
367 Changsha resulting in 20.16 million records (7000*24*60*2) on a daily basis. The status of the
368 cab was also recorded, such as the locations where passengers get on and off. These data are then
369 used to extract the movements between the five districts in the main area of Changsha: Kaifu
370 district, Furong district, Yuhua district, Tianxin district, and Yuelu district. For the purpose of

comparison, one week's data (April 4 to April 17, 2016) were extracted and analysed. The movements were normalized and then compared with the same week in 2014 from the Baidu LBS data. The correlation between the mobility estimates extracted from the Baidu LBS data and from the taxi's GPS data for Changsha city is presented in Extended Data Fig. 9. There is a high level of similarity between the two datasets, with a correlation coefficient of 0.99 ($p=0.001$). We subsequently used the fitted models from China to extrapolate to other regions in Asia and Oceania again using the movement package in R51.

Human movement data for Africa: To calibrate the gravity and radiation models for Africa, we used aggregated and de-identified mobile phone-derived mobility estimates at the constituency level from Namibia between 1 October 2010 and 30 September 2011. These data represent the proportion of time that unique subscriber identity module (SIM) cards in each constituency spend in all other constituencies, as described in detail in Jones & Thornton (2000)53. We used this data set from Namibia because it was openly available and because it offered the best spatial and temporal resolution compared to census-derived data. We then used the fitted models to extrapolate to all other regions in Africa using the movement package in R51. Systematic surveys of cross-border human movements were not available at the time of the study and for the study regions.

It is possible that there are significant differences between regions in terms of mobility, but unfortunately no sufficiently widespread and well-resolved data source was available to test this. Our model captured the spread process of *Aedes* mosquitoes using a variety of human movement data, including both CDR data and commuting data. To assess the generalizability of our results we applied the model fitted to commuting data in the USA to the range expansion process observed in Europe. The predictive ability of this cross-continental validation indicates that the

mobility data used are sufficiently robust in the context of this study (Extended Data Fig. 8). However, we note there may be several limitations to using commuting data to infer vector introductions as they overly emphasize work-related movements. To test whether our model would perform well even in the absence of human movement data, we performed a cross validation that uses only distance and adjacency matrices which only marginally reduces predictability (Extended Data Fig. 12). Despite this, such data has indeed been used in the United States to successfully predict the long distance spread of infectious diseases. We are therefore confident that such data can be applied to predict both short and long distance spread in the USA[54]. Similarly, CDR data has been used to describe the spread of pathogens such as influenza in Europe[23]. As new data become available, our model is flexible enough to incorporate them and estimates of the predicted range expansion of *Ae. aegypti* and *Ae. albopictus* can be updated. There was also no suitable data available on cross border movements that could improve estimates of between-country spread (see section 2.4. for a sensitivity analysis).

2. Model fitting to data

2.1 Description of speed of dispersal:

To understand the past range expansion of both species and to provide basic summary statistics of the speed of dispersal over time in areas where sufficient observations were available, we use the methods of spread rate measurements employed by Tisseuil et al[55]. For each species and study area, the centroids of the spatial units where the species were observed were re-projected in a metric system (epsg 102003 in the US, and epsg 3035 in Europe) and the first date of detection in each centroid was interpolated on a 10 km resolution grid using thin plate spline regression (TPSR). The local slope of the surface was measured by a 3 x 3 moving windows filter, and the resulting friction surface (time / distance) was smoothed by an average 11 x 11 cell filter to

prevent local null frictions values. The local spread rate was then obtained by taking the inverse of the friction. This measure was computed within a mask, which was obtained by kernel density smoothing of the centroids of spatial units where the species were observed. We used the method of Berman and Diggle[56] to determine the optimal bandwidth for the US and EU invasions. In order to have a similar bandwidth for all masks, we used the maximum of the three estimated optimal bandwidths, which was found to be 73.2 km. A density threshold of 2.9 points per 10,000 km^2 was chosen to delineate the mask, which was the maximum threshold value allowing the inclusions of all observation points in the mask in both the US and EU.

2.2. Mosquito environmental niche modelling

To predict the likely future distributions of both species independently (in years 2020, 2050 and 2080), we first fitted species distribution models to data from the present day. This approach built on previous work[4] using the boosted regression tree (BRT) models fit to mosquito occurrence data (section 1.1.). BRTs combine strengths from regression trees and machine learning (gradient boosting) and are able to accommodate non-linear relationships to identify the environmental niche in which the environment is suitable for the species in question. After an initial regression tree is fitted and iteratively improved upon in a forward stepwise manner (boosting) by minimising the variation in the response variable not explained by the model at each iteration. This approach has been shown to simultaneously fit complex non-linear response functions efficiently while guarding against over-fitting.

We first developed a baseline scenario for the year 2015, using the global dataset of *Ae. aegypti* and *Ae. albopictus* occurrence (section 1.1)[43,57] and a set of environmental and socioeconomic predictors (section 1.2). In a BRT modelling framework pseudo-absences need to be generated to allow for discrimination between areas where the mosquitoes can persist, and to identify biases

in reporting[58]. We used the approach previously described in and applied by Kraemer et al[4] using background points from the Global Biodiversity Information Facility (GBIF) and the inverse of an *Aedes* temperature suitability mask[45] with equal ratio between presence and absence points and no threshold being applied. From that we constructed 100 sub-models to derive the mean prediction map and model-fitting uncertainty using the SEEG-SDM package in R[59,60].

2.3. Human mobility modelling

Given the heterogeneous abundance of both species[61] as well as the low probability of their surviving slower and longer transits, the chance of a species being introduced following any single translocation event is low. Hence we used relatively long time steps (yearly) and generalized human movement models fitted to a variety of data sources to understand the spatial spread patterns of *Ae. aegypti* and *Ae. albopictus*.

We incorporated three distinct human movement models that act at different scales, since we are uncertain *a priori* which type of human movement will be most associated with mosquito spread. We considered (i) a gravity model, (ii) a radiation model, (iii) an adjacency network model and (iv) un-transformed great-circle distance. Each of these models have been shown to be useful depending on the local context to infer regular daily commuting patterns, longer-term movements, and as general descriptions of human mobility[24,62,63]. First, the gravity model, assumes that fluxes between two areas i and j are $T_{i,j} = k\frac{N_i^{\alpha}N_j^{\beta}}{d_{i,j}^{\gamma}}$, where N represents human population size and d is great circle distance between two locations, and k, α, β, and γ are parameters to be fit[64,65]. The gravity model emphasises the attractive power of large population centres. Second, the radiation model assumes fluxes to be $T_{i,j} = T_i\frac{N_iN_j}{(N_i+s_{i,j})(N_i+N_j+s_{i,j})}$, where T_i is the number of individuals leaving area i and s_{ij} is the total population in the circle centered at

462 i with radius $d_{i,j}$ excluding the population of the two areas i and j. The radiation model
463 considers not only distance and population sizes at origin and destination but also the cumulative
464 population at a lesser distance from the origin than the destination[24]. Consequently, this model
465 considers not only the origin and destination but also the landscape of 'intervening opportunities'
466 between them. Third, adjacency networks encode the number of district borders an individual
467 would need to cross to move from one district to another. Thus, this metric reflects the
468 neighbourhood effect. Finally, we computed the great-circle distance between each pair of
469 locations and used that as a metric of mobility in and of itself [32,66].

470 For each second Administrative unit (county/municipality) in the world, we determined the total
471 human population size using gridded population estimates and calculated the great-circle
472 distance between the centroids of each pair of districts within each continent[67]. Gravity and
473 radiation model parameters were fitted by maximum likelihood methods to the empirical data
474 described above using the movement R package[51]. National adjacency networks were computed
475 using administrative boundary data from the GADM dataset (http://www.gadm.org). To account
476 for neighbourhood effects of spread and for the potential importance of within-country and
477 between-country movements, we constructed adjacency matrices that were disaggregated into
478 three binary connectivity matrices with connectivity degrees of one (*i.e.,* districts share a border),
479 two (*i.e.,* districts share a common neighbour), and three (*i.e.,* more than two degrees away).

480 ## 2.4. Mosquito spread modelling

481 Let $x_i(t)$ be the *Aedes* population status of district i at time t (*i.e.*, a binary variable takes the
482 value 1 if there were *Aedes* mosquitoes that time, and 0 otherwise). Given the nature of the
483 dataset collected, we assumed that all data points represented detection of established
484 populations and thus assumed continuous presence of the species for the first and last reported

485 occurrences. We used a standard logistic model to characterize the probability that some district *j*
486 will become occupied at time t:

$$logit\left(P\left(x_j(t) = 1 | x_j(t-1) = 0\right)\right) = \beta_0 + \sum_{k=1}^{n} \beta_k Y_{j,t}^{(k)}$$

487 where $Y_{j,t}^{(k)}$ corresponds to the value of explanatory variable *k* in district *j* at time t. Explanatory
488 variables included in this analysis were the predicted vector habitat suitability (*i.e.* suitability for
489 establishment of an introduced vector, 2.1.) and connectivity between infested and non-infested
490 districts (*i.e.* probability of introduction of a vector). Separate metrics of connectivity were
491 defined for each human movement model (2.2.). From each human movement model, a
492 connectivity matrix $A_{ij}^{(k)}$ was calculated for each location i and j. A corresponding covariate for
493 the occupation model was then computed to represent the global force of importation, exerted
494 from all other infested districts to j: $Y_{j,t}^{(k)} = \sum_i A_{ij}^{(k)} x_i(t-1)$.

495 These models were re-fit in each successive year separately for the North American and
496 European datasets, and for each vector species, using all available data up to that year. Model
497 selection was done through backward selection using Akaike Information Criterion (AIC).[68] The
498 fitted model was then evaluated prospectively over the next year by comparing predicted
499 presence or absence with observations, thereby allowing us to evaluate and validate model
500 performance over time. For model evaluation we considered all locations (*i.e.* 3,134 counties in
501 the USA, 1,587 NUTS in Europe). This model evaluation was used to identify the best
502 explanatory variables to include in the *Aedes* spread model. Model evaluation was performed
503 using receiver operating characteristic curves (ROC curves) (Extended Data Fig. 7) and model
504 accuracy was characterized comparing the predicted probabilities of first detection *vs* the

response (Extended Data Fig. 6). We calculated the probability of first detection p_w predicted by the model for each district-year that had not yet reported mosquitoes. We then partitioned district-years into eight groups with predicted probability in the range of 0-1%, 1-5%, 5-10%, 10-15%, 15-20%, 20-25%, 25-35%, and 35-100%. For each group, we calculated the mean predicted probability and compared it with the proportion of district-years in the group in which range expansion was observed. Our model assumes that each mosquito species will persist in an area once detected, whilst there are some examples of incursions apparently having been successfully eradicated or died out. It is possible that this assumption could result in inflated predictions of the rate of spread, due to an overestimated number of source populations for each potential invasion event. However, it should be noted that this overestimate of the number of source populations would also be present in the training data, and would be at least partially absorbed into estimates of the probabilities of importation. Insufficient data were available to test or account for this potential bias, but based on additional experiments, we do not anticipate our estimates to greatly overpredict *Aedes* presence (see section: sensitivity analyses and sampling bias).

Cross-validation: To test whether the spread between countries is different to the spread within countries, we used the multi-country dataset from *Ae. albopictus* in Europe and varied the relative frequency of within- and between-country mobility by decreasing movement between countries by 20%, 50%, and 70%. The results were then compared with a baseline, in which predicted within-country movement is the same as between-country movement (Extended Data Fig. 11). We also performed sensitivity analyses to evaluate how a model including human movements compares to single variable models that have objective measurements such as great

circle distance and adjacency. A model that includes human movements only slightly increased predictive performance (Extended Data Fig. 12).

Sensitivity analyses and sampling bias: Surveillance efforts to detect *Ae. aegypti* and *Ae. albopictus* may vary in time and space due to gradual progressive improvements as a result of technology trapping technology, general expertise, or in response to specific events. Three types of possible changes in surveillance could bias the estimates of our spread model: (1) spatial expansion of surveillance system coverage to new areas; (2) intensification of sampling effort within areas where the surveillance system already operates; and (3) changes in sampling methods within areas where the surveillance system already operates that make it more or less likely to detect either *Ae. aegypti* or *Ae. albopictus.* To address each of these, we completed sensitivity analyses to understand how possible changes in surveillance may affect the inference about spread in the future.

Expansions of the surveillance system can be definitively distinguished from true known expansions of the vectors by comparing the state transitions of areas in longitudinal datasets, such as our *Ae. albopictus* dataset in Europe between the years of 2013 and 2017. Areas that first report absence of the species (often for multiple years) and later report presence are as close to a clear example of introduction as possible and give a reasonable estimate of the arrival date. Conversely, if an area's first report is presence of the species, the species' arrival date may have been estimated later than it truly occurred.

Firstly, the existence of such longitudinal records in the *Ae. albopictus* database in Europe is strong evidence that the distribution of the species is expanding, however to test if expanding surveillance efforts is a contributing factor to the observed rate of spread we compared our original model fit to the full *Ae. albopictus* in Europe dataset, as used in our main analysis

(model 1), with a model fit only to the data points that have strong evidence for a specific introduction date (i.e., report absence before presence; model 2). We tabulated data from *Ae. albopictus* in Europe where information was available whether there was ongoing surveillance prior to the reporting of the species (transition from absence to presence). Such data was available for 179 out of 600 observations between 2013 – 2018, a time period where 400 new regions reported the presence of the species making our sub-sample about 50% of all new invasions. This data was available at higher spatial resolution that the full *Ae. albopictus* dataset for Europe. 75% of these records are from locations of most recent spread in France and Germany. Finally, as model 2 was fit to data from a narrower date range we also consider a third model (model 3) which was fit to both occurrence and longitudinal data but only from the more recent date range (Extended Data Tab. 3). If expansion of surveillance efforts is a contributing factor to the observed rate of spread in the data, then we would expect Model 2 to predict a significantly lower rate of spread than Models 1 or 3 (our null hypothesis).

Each of these models were fit to the above datasets, then used to simulate *Ae. albopictus* spread from a common baseline (based on occurrence and longitudinal data at the end of 2012) for five years between 2013 and 2017 as described previously. The predicted total number of new districts infested of this period was calculated and is shown in Extended Data Tab. 4. Note that comparison of goodness of fit metrics for these models was not possible since the models were fit to different datasets.

Contrary to the expectation that more precise dates of invasion would lead to conclusions of slower rates of spread, this sensitivity exercise found that restricting the model to just areas where the date of introduction is known significantly increases the predicted rate of spread. Thus, this exercise rejects our above null hypothesis. This effect was also independent of the time

period of the fitting data (similar results for Model 1 and Model 3). These results suggest that it is more likely that true spread of *Ae. albopictus* is outpacing expansion of mosquito surveillance, and if longitudinal surveillance was in place everywhere, the observed rates of spread would be greater.

We therefore believe that the currently implemented model is a conservative estimate of spread of these species that is not highly affected by changes in spatial coverage of surveillance systems and provides the most robust estimates of spread over these time periods given the available data. Given the limited number of years of data available to fit Model 2, we believe that Model 1 provides the most reliable estimates of future spread.

Intensification in sampling effort and technological advancements in collection methods may affect the probability of detection of a species in earlier in their invasion process vs today. Here we test both hypotheses through inclusion of different terms in our spread model regression and compare such models to the null of no changes in surveillance intensity over time (as currently implemented in our main analysis). To represent increases or decreases in surveillance over time, we include the spline-smoothed year of detection as a variable in the regression analysis. To represent step changes in surveillance efforts in response to specific events we include a factor variable; either before the 2003 peak in West Nile Virus cases in the USA, or after 2003 (only for models in USA). Internal cross validation was then used to compare the predictive performance of these three models with evaluation on three-year-lookahead holdout sets, subject to a minimum of 10 consecutive years of data to fit the models. Model predictive performance was then compared using deviance from observed values in the holdout set.

This showed that for all species in all continents, the inclusion of a temporal (Year) term reduced predictive accuracy (increased deviance). This was the case for both gradual change over time

596 (s(Year)) and for breakpoint changes in response to specific events (Year > 2003). As a result,
597 we conclude that there is no evidence for temporal changes in sampling effort in any of the
598 datasets concerned and therefore do not include such terms in our final predictions (Extended
599 Data Tab. 5).

600 Finally, there is a possibility that changes in general vector surveillance strategies could have led
601 to changes that affected the probability of detection of one species more than the other. Such
602 differential biases could undermine our inter-species spread rate comparison. One key period of
603 concern is around the 2003 West Nile Virus (WNV) outbreak in the US where vector
604 surveillance may have prioritized trapping in more rural environments to optimize detection of
605 various *Culex* species. Such a focus on rural environments may have led to relative increases in
606 sampling intensity of *Ae. albopictus* and relative reductions in sampling intensities for *Ae.*
607 *aegypti*.

608 To test this hypothesis, we follow a similar approach to the above analysis, where covariates for
609 "before" and "after" the 2003 WNV outbreak are included in the USA spread model for each
610 species. If the above hypothesis is true, such terms should i) have larger "after" values than
611 "before" values in the *Ae. albopictus* model and vice versa in the *Ae. aegypti* model, and ii)
612 improve model prediction accuracy.

613 The best fits from the *Ae. aegypti* and *Ae. albopictus* spread models in the USA show that
614 detection of *Ae. aegypti* marginally increased relative to *Ae. albopictus* (positive model
615 coefficients for post-2003 term in *Ae. aegypti,* negative in *Ae. albopictus*) (Extended Data Tab.
616 6). However, as previously stated, inclusion of such changes in surveillance quality over time
617 reduces the model predictive performance (increase in deviance for both species) and therefore
618 may not provide a better time period to mirror the spread of the species in the United States.

2.5. Classifying the ranges of each mosquito species and incorporating uncertainty

Current reported distributions of *Ae. aegypti* and *Ae. albopictus* are unlikely to be fully representative of their actual distributions because of logistical and financial constraints on vector surveillance.[37] Therefore we used the following method to estimate the current-day global distribution (realised niche) of each mosquito species by comparing environmental suitability maps with occurrence data. We extracted the predicted environmental suitability value at each of the locations where the mosquito species has been reported, and the value of environmental suitability that encompassed 90% of these reported locations was chosen as the range threshold. Every value above or equal to this threshold was defined as within the range of the mosquito species (Extended Data Fig. 13). This approach assumes that the 10% of occurrences outside of the predicted range represent temporary introductions that do not persist longer than one year and are not representative of the long-term distribution of the species. As there is uncertainty in what proportion of the data are representative of these transient identifications (given that the majority of the data are cross-sectional not longitudinal), we undertook a sensitivity analysis that varied this threshold from 85% to 95%, thereby creating 96 different possible range maps that represent different realisations of the current distribution of each species. In doing so, we capture locations that have the conditions for mosquito presence and where there is potential for onward spread. We did not include international shipping as a contributor to infrequent long-distance importation events between continents since both species are already well established on each continent and therefore new occurrences are more likely to be driven by intra-continental importation pressure.

3. Future projections

3.1. Projecting environmental and socioeconomic covariates

We used 17 GCMs to estimate 30 arc-sec images for monthly mean climate data. Extended Data Table 7 provides the designation, origin, references and number of replicate runs for each model. The procedures are described in detail in MarkSim documentation[65]. For each GCM the baseline monthly climate was derived from the historic runs for temperatures and rainfall, the monthly means were calculated for each GCM for the years 2000 to 2095, and the difference 'delta' for each month was calculated by subtracting the specific GCM baseline. The deltas were interpolated from the native GCM pixel (Extended Data Tab. 7) to a one degree by one degree pixel for the globe. The data were pattern scaled to WorldClim 1.03[64] for each one degree pixel, RCP, and month. For each variant a fourth order polynomial regression was fitted over the 96 years of data and through the origin at 1985 (1985 being the mean midpoint of the data used in the WorldClim construction) to calculate one output per model per year per scenario.

Humidity data were estimated directly at the 30 arc-sec level from dewpoint calculated by the tabular method of Linacre[69] and the mean temperature. To fully propagate the variation between the climate models through our predictions we used the outputs of 17 GCM, for all 3 years, and 3 scenarios.

Global temperature estimates were converted into temperature suitability for mosquito population persistence (separate metrics for each vector species), hereafter referred to as temperature suitability, using temperature-based mathematical models from Riahi et al[44] and Fujino et al[45]. These show the effects of diurnal and seasonal changes in temperatures on the generation time of the mosquito and its resultant effects on the persistence of a population.

As a highly anthropophilic mosquito species, the future distribution of the *Aedes* is likely to depend critically on both environmental and human socioeconomic factors that modify the availability of its habitat[8]. To incorporate these features, we also modelled the continued process of global urbanisation until 2080 using a probabilistic machine learning algorithm based on the work of Linard et al[48]. Here we use urban growth rates predicted by the United Nations as a predictor variable[49] as well as a range of other covariates as previously described in van Vuuren et al[48].

3.2. Projecting future niche of *Ae. aegypti* and *Ae. albopictus*

Although niche shifts might occur over long time-periods, the future effects remains unclear for *Ae. aegypti* and *Ae. albopictus* since their expansion from their native range[70]. Therefore, we assume niche conservatism, implying that the mosquitoes tend to establish and survive under similar environmental conditions in native and invaded ranges in the future[4,71,72].

Our final aim was to produce 18 maps predicting *Ae.* aegypti and *Ae. albopictus* suitability in the years 2020, 2050 and 2080 under three different emissions scenarios (RCPs). Each of these 18 maps were composed of 100 ensemble predictions that randomly sampled (with replacement) the following aspects of the analysis:

1. The fitted *Aedes* BRT model (from a choice of 100 BRT models fitted to 2015 data)
2. The predicted temperature suitability for *Aedes* survival (from a choice of 17 GCMs)
3. The predicted minimum precipitation (from a choice of 17 GCMs)
4. The predicted relative humidity (from a choice of 17 GCMs)
5. The predicted minimum precipitation (from a choice of 17 GCMs)
6. The predicted geographic expansion via land from the spread models (section 3.3).

This approach sought to fully propagate the uncertainty in the climate, *Aedes* temperature suitability and *Aedes* models through to the final prediction. These 100 predictions were then summarised by mean and 95% credible intervals to give the final prediction for each year RCP combination. Uncertainties are shown in all maps along the X-axes.

Our baseline map modelling is different from previously published maps in so far that it uses only projectable environmental and socio-demographic variables and does not use the Enhanced Vegetation Index (EVI), as the EVI is a direct empirical measure of the Earth's current greenness[4]. To minimise potential reduction in the predictive ability of the model by omitting this covariate, we include precipitation and relative humidity as predictors for suitability for green vegetation growth in both the present day and future models.

3.3. Projecting mosquito spread

To derive yearly model-based estimates of the possible expansion of both species by 2080 we forward-simulated the geographic spread model based on the equation in 2.4. To account for the spatio-temporal dependence in first detection probabilities (each district's probability is a function of every other district that was infested the year before), we run 1,000 simulations forward in time. Within each simulation we estimate the probability of infestation to each district that had yet to detect the species. We then drew a Bernoulli random variable with that probability of '1' (i.e., invasion) and imputed those results for each potential detection. Using these imputed invasions as well as all districts that had previously been infested, we repeat the estimation of range expansion for the next year. This process is repeated up to the desired forecast horizon. This represents a single simulation. It is important to note that we did not allow for the situation where an already infested district will 'lose' its infection status (*i.e.*, if $x_i(t-1) = 1$ for district

706 i, we force $x_i(t) = 1$). We then combine the results of the 1,000 simulations to identify which
707 districts were most likely to have a positive species presence at any point.

708 *3.4. Calculating population at risk and area expansion*

709 To classify areas as at risk or not at risk of *Ae. aegypti* and *Ae. albopictus* a threshold was
710 defined for the continuous *Aedes* suitability maps by the value that maximised sensitivity and
711 specificity when classifying the occurrence and background data using the 2015 map. This value
712 was found to be 0.47 and 0.51 for *Ae. aegypti* and *Ae. albopictus* respectively. Any pixel with a
713 predicted suitability value above that was considered at risk and the same threshold was applied
714 to each time point and scenario to calculate the population and area at risk in each global region.
715 The final maps for 2020, 2050, 2080 are then overlaid with contemporary estimates of human
716 populations at 5 km resolution and extracted the relevant population at risk was estimated using
717 the raster package in R. We paired the climatic scenarios based on Shared Socioeconomic
718 Pathways (SSPs) that were defined by O'Neill *et al.* in 2014[73]. They represent reference
719 pathways that describe plausible alternate trends in the evolution of society and ecosystems over
720 a century, in the absence of climate change or climate policies. SSPs are predicated on possible
721 outcomes that would make it more or less difficult to respond to climate change challenges. Each
722 SSP consists of quantified population and Gross Domestic Product (GDP) trajectories, serving as
723 the starting points for various organisations to model these factors and to provide projections for
724 demographic and economic development variables. The Integrated Assessment Modelling
725 Consortium (IAMC) made available certain peer-reviewed projections via the International
726 Institute for Applied Systems Analysis (IIASA, http://www.iiasa.ac.at), whereby the SSP
727 storylines were converted into population and GDP projections for 195 countries[74] for every
728 decade between the years 2010 and 2100.

729

References

1. Nsoesie, E. O. *et al.* Global distribution and environmental suitability for chikungunya virus, 1952 to 2015. *Eurosurveillance* **21,** pii=30234 (2015).
2. Messina, J. P. *et al.* Mapping global environmental suitability for Zika virus. *Elife* **5,** e15272 (2016).
3. Bhatt, S. *et al.* The global distribution and burden of dengue. *Nature* **496,** 504–7 (2013).
4. Kraemer, M. U. G. *et al.* The global distribution of the arbovirus vectors Aedes aegypti and Ae. albopictus. *Elife* **4,** e08347 (2015).
5. Lessler, J. *et al.* Assessing the global threat from Zika virus. *Science* **353,** aaf8160 (2016).
6. Faria, N. R. *et al.* Epidemic establishment and cryptic transmission of Zika virus in Brazil and the Americas. *Nature* **546,** 406–10 (2017).
7. Grubaugh, N. D., Ladner, J. T., Kraemer, M. U. G., Dudas, G. & Tan, A. L. Genomic epidemiology reveals multiple introductions of Zika virus into the United States. *Nature* **546,** 401–5 (2017).
8. Messina, J. P. *et al.* The many projected futures of dengue. *Nat. Rev. Microbiol.* **13,** 230–9 (2015).
9. Kraemer, M. U. G. *et al.* Progress and challenges in infectious disease cartography. *Trends Parasitol.* **32,** 19–29 (2016).
10. Powell, J. R. Mosquitoes on the move. *Science* **354,** 971–972 (2016).
11. Powell, J. R. & Tabachnick, W. J. History of domestication and spread of Aedes aegypti - a review. *Mem. Inst. Oswaldo Cruz* **108 Suppl,** 11–7 (2013).
12. Scott, T. W. *et al.* Longitudinal Studies of Aedes aegypti (Diptera: Culicidae) in Thailand and Puerto Rico: Blood Feeding Frequency. *J. Med. Entomol.* **37,** 89–101 (2000).
13. Guerra, C. A. *et al.* A global assembly of adult female mosquito mark-release-recapture data to inform the control of mosquito-borne pathogens. *Parasites and Vectors* **7,** 276 (2014).
14. Hawley, W. A., Pumpuni, C. B., Brady, R. H. & Craig, G. B. Overwintering survival of Aedes albopictus (Diptera, Culicidae) eggs in Indiana. *J. Med. Entomol.* **26,** 122–129 (1989).
15. Maynard, A. J. *et al.* Tiger on the prowl: Invasion history and spatio-temporal genetic structure of the Asian tiger mosquito Aedes albopictus (Skuse 1894) in the Indo-Pacific. *PLoS Negl Trop Dis* **11,** e (2017).
16. Benedict, M. Q., Levine, R. S., Hawley, W. a & Lounibos, L. P. Spread of the tiger: global risk of invasion by the mosquito Aedes albopictus. *Vector borne zoonotic Dis.* **7,** 76–85 (2007).
17. Armbruster, P. A. Review Photoperiodic Diapause and the Establishment of Aedes albopictus (Diptera: Culicidae) in North America. *J. Med. Entomol.* **53,** 1013–1023 (2016).
18. Vega-rua, A. *et al.* High Efficiency of Temperate Aedes albopictus to Transmit

Chikungunya and Dengue Viruses in the Southeast of France. *PLoS One* **8,** e59716 (2013).

19. Gloria-soria, A. *et al.* Global genetic diversity of Aedes aegypti. *Mol. Ecol.* **25,** 5377–5395 (2016).

20. Padilla, D. K., Chotkowski, M. A. & Buchan, L. A. J. Predicting the Spread of Zebra Mussels (Dreissena polymorpha) to Inland Waters Using Boater Movement Patterns. *Glob. Ecol. Biogeogr. Lett.* **5,** 353–359 (1996).

21. Roche, B. *et al.* The Spread of Aedes albopictus in Metropolitan France: Contribution of Environmental Drivers and Human Activities and Predictions for a Near Future. *PLoS One* **10,** e0125600 (2015).

22. Moore, C. G., Medicine, V., Sciences, B. & State, C. Using Geographic Information Systems to Analyze the Distribution and Abundance of Aedes aegypti in Africa : The Potential Role of Human Travel in Determining the Intensity of Mosquito Infestation. *Int. J. Appl. Geospatial Res.* **4,** 9–38 (2013).

23. Tizzoni, M. *et al.* On the use of human mobility proxies for modeling epidemics. *PLoS Comput. Biol.* **10,** e1003716 (2014).

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

25. Simini, F., Maritan, A. & Néda, Z. Human mobility in a continuum approach. *PLoS One* **8,** e60069 (2013).

26. Canali, M., Rivas-Morales, S., Beutels, P. & Venturelli, C. The Cost of Arbovirus Disease Prevention in Europe: Area-Wide Integrated Control of Tiger Mosquito, Aedes albopictus, in Emilia-Romagna, Northern Italy. *Int. J. Environ. Res. Public Health* **14,** 444 (2017).

27. Leishnam, P. T., Lounibos, L. P., O'Meara, G. F. & Juliano, S. A. Interpopulation divergence in competitive interactions of the mosquito Aedes albopictus. *Ecology* **90,** 2405–13 (2009).

28. Camara, D. C. P. *et al.* Seasonal differences in density but similar competitive impact of Ae. albopictus (Skuse) and Aedes aegypti (L.) in Rio de Janeiro, Brazil. *PLoS One* **11,** e0157120 (2016).

29. Lounibos, L. P. *et al.* Does temperature affect the outcome of larval competition between Aedes aegypti and Aedes albopictus? *J. Vector Ecol.* **27,** 86–95 (2002).

30. Prinzing, A., Durka, W., Klotz, S. & Brandl, R. Geographic variability of ecological niches of plant species: Are competition and stress relevant? *Ecography (Cop.).* **25,** 721–729 (2002).

31. Soberón, J. Grinnellian and Eltonian niches and geographic distributions of species. *Ecol. Lett.* **10,** 1115–1123 (2007).

32. Kraemer, M. U. G. *et al.* Big city, small world: density, contact rates, and transmission of dengue across Pakistan. *J. R. Soc. Interface* **12,** 20150468 (2015).

33. Kraemer, M. U. G. *et al.* Spread of yellow fever virus outbreak in Angola and the Democratic Republic of the Congo 2015–16: a modelling study. *Lancet Infect. Dis.* **17,**

807 330–338 (2017).

808 34. Perkins, T. A., Siraj, A. S., Ruktanonchai, C. W., Kraemer, M. U. G. & Tatem, A. J.
809 Model-based projections of Zika virus infections in childbearing women in the Americas.
810 *Nat. Microbiol.* **16126,** (2016).

811 35. Bogoch, I. I. *et al.* Potential for Zika virus introduction and transmission in resource
812 limited countries in Africa and Asia-Pacific. *Lancet Infect. Dis.* **16,** 1237–45 (2016).

813 36. Bogoch, I. I. *et al.* Anticipating the international spread of Zika virus from Brazil. *Lancet*
814 **387,** 335–336 (2016).

815 37. Schaffner, F. *et al.* Development of guidelines for the surveillance of invasive mosquitoes
816 in Europe. *Parasit. Vectors* **6,** 209 (2013).

817 38. Moore, C. G. & Mitchell, C. J. Aedes albopictus in the United States: ten-year presence
818 and public health implications. *Emerg. Infect. Dis.* **3,** 329–34 (1997).

819 39. Flacio, E., Engeler, L., Tonolla, M., Lüthy, P. & Patocchi, N. Strategies of a thirteen year
820 surveillance programme on Aedes albopictus (Stegomyia albopicta) in southern
821 Switzerland. *Parasites and Vectors* **8,** (2015).

822 40. Collantes, F. *et al.* Review of ten-years presence of Aedes albopictu s in Spain 2004 –
823 2014 : known distribution and public health concerns. *Parasit. Vectors* **8,** (2015).

824 41. Eritja, R., Palmer, J. R. B., Roiz, D., Sanpera-calbet, I. & Bartumeus, F. Direct Evidence
825 of Adult Aedes albopictus Dispersal by Car. *Sci. Rep.* **7,** 14399 (2017).

826 42. Salje, H. *et al.* Dengue diversity across spatial and temporal scales: Local structure and the
827 effect of host population size. *Science* **355,** 1302–1306 (2017).

828 43. Kraemer, M. U. G. *et al.* The global compendium of Aedes aegypti and Ae . albopictus
829 occurrence. *Sci. Data* **2,** 150035 (2015).

830 44. Brady, O. J. *et al.* Global temperature constraints on Aedes aegypti and Ae. albopictus
831 persistence and competence for dengue virus transmission. *Parasit. Vectors* **7,** 338 (2014).

832 45. Brady, O. J. *et al.* Modelling adult Aedes aegypti and Aedes albopictus survival at
833 different temperatures in laboratory and field settings. *Parasit. Vectors* **6,** 351 (2013).

834 46. IPCC. *Climate Change 2013: the physical science basis: contribution of working group I*
835 *to the Fifth Assessment Report on The Intergovernmental Panel on Climate Change, IPCC*
836 *Fifth Assessment Report: Climate Change 2013 (AR5).* (Cambridge University Press,
837 2013).

838 47. UNFCCC (United Nations Framework Convention on Climate Change). Adoption of the
839 Paris Agreement. *21st Conf. Parties* (2015).

840 48. Linard, C., Tatem, A. J. & Gilbert, M. Modelling spatial patterns of urban growth in
841 Africa. *Appl. Geogr.* **44,** 23–32 (2013).

842 49. United Nations Population Division. *World Urbanization Prospects: The 2014 Revision*.
843 (United Nations, 2014).

844 50. Hahn, M. B. *et al.* Reported Distribution of Aedes (Stegomyia) aegypti and Aedes

(Stegomyia) albopictus in the United States, 1995-2016 (Diptera: Culicidae). *J. Med. Entomol.* 1–7 (2016). doi:10.1093/jme/tjw072

51. Golding, N., Schofield, A. & Kraemer, M. U. G. Movement: Functions for the analysis of movement data in disease modelling and mapping. *R Packag. version 0.2* (2015).

52. Osório, H. *et al.* Detection of the Invasive Mosquito Species Aedes (Stegomyia) albopictus (Diptera: Culicidae) in Portugal. *Int. J. Environ. Res. Public Heal. 2018, Vol. 15, Page 820* **15,** 820 (2018).

53. Ruktanonchai, N. W. *et al.* Identifying Malaria Transmission Foci for Elimination Using Human Mobility Data. *PLOS Comput. Biol.* **12,** e1004846 (2016).

54. Viboud, C. *et al.* Synchrony, waves, and spatial hierarchies in the spread of influenza. *Science* **312,** 447–51 (2006).

55. Tisseuil, C. *et al.* Evaluating methods to quantify spatial variation in the velocity of biological invasions. *Ecography (Cop.).* **39,** 409–418 (2015).

56. Berman, M. & Diggle, P. Estimating weighted integrals of the second-order intensity of a spatial point process. *J. R. Stat. Soc. Ser. B* **51,** 81–92 (1989).

57. Kraemer, M. U. G. *et al.* Data from: the global compendium of Aedes aegypti and Ae. albopictus occurrence. *Dryad Digit. Repositry* (2015). doi:dx.doi.org/10.5061/dryad.47v3c

58. Phillips, S. J. *et al.* Sample selection bias and presence-only distribution model: implications for background and pseudo-absence data. *Ecol. Appl.* **19,** 181–97 (2009).

59. R Core Team. R: A language and environment for computing. Vienna, Austria. *R Found. Stat. Comput.* (2016).

60. Golding, N. Streamlined functions for species distribution modelling in the seeg research group. *R Packag. version 0.1-3* (2014).

61. Perkins, T. A., Scott, T. W., Le Menach, A. & Smith, D. L. Heterogeneity, mixing, and the spatial scales of mosquito-borne Ppathogen transmission. *PLoS Comput. Biol.* **9,** e1003327 (2013).

62. Brockmann, D. & Helbing, D. The hidden geometry of complex, network-driven contagion phenomena. *Science* **342,** 1337–42 (2013).

63. Jongejans, E. *et al.* A unifying gravity framework for dispersal. *Theor. Ecol.* **8,** 207–223 (2015).

64. Wesolowski, A., O'Meara, W. P., Eagle, N., Tatem, A. J. & Buckee, C. O. Evaluating Spatial Interaction Models for Regional Mobility in Sub-Saharan Africa. *PLOS Comput. Biol.* **11,** e1004267 (2015).

65. Wesolowski, A. *et al.* Commentary: Containing the Ebola outbreak – the potential and challenge of mobile network data. *PLOS Curr. Outbreaks* **Sep 29,** (2014).

66. Tatem, A. J., Hemelaar, J., Gray, R. R. & Salemi, M. Spatial accessibility and the spread of HIV-1 subtypes and recombinants. *AIDS* **26,** 2351–60 (2012).

67. WorldPop project. WorldPop. Available at: http://worldpop.org.uk/.

883 68. Hastie, T. J. & Tibshirani, R. J. *Generalized additive models*. (CRC Press, 1990).

884 69. Linacre, E. T. A simple formula for estimating evaporation rates in various climates, using
885 temperature data alone. *Agric. Meteorol.* **18,** 409–424 (1977).

886 70. Medley, K. A. Niche shifts during the global invasion of the Asian tiger mosquito, Aedes
887 albopictus Skuse (Culicidae), revealed by reciprocal distribution models. *Glob. Ecol.*
888 *Biogeogr.* **19,** 122–33 (2010).

889 71. Wiens, J. J. & Graham, C. H. Niche conservatism: integrating evolution, ecology, and
890 conservation biology. *Annu. Rev. Ecol. Evol. Syst.* **36,** 519–39 (2005).

891 72. Petitpierre, B. *et al.* Climatic Niche Shifts Are Rare Among Terrestrial Plant Invaders.
892 *Science* **338,** 1344–8 (2012).

893 73. O'Neill, B. C.*et al.* A new scenario framework for climate change research : the concept of
894 shared socioeconomic pathways. *Clim. Change* **122,** 387–400 (2014).

895 74. Kc, S. & Lutz, W. Demographic scenarios by age, sex and education. *Popul. Environ.* **35,**
896 243–260 (2014).

897 75. Elith, J., Leathwick, J. R. & Hastie, T. A working guide to boosted regression trees. *J.*
898 *Anim. Ecol.* **77,** 802–13 (2008).

899 76. Rochlin, I., Ninivaggi, D. V, Hutchinson, M. L. & Farajollahi, A. Climate Change and
900 Range Expansion of the Asian Tiger Mosquito (Aedes albopictus) in Northeastern USA :
901 Implications for Public Health Practitioners. *PLoS One* **8,** e60874 (2013).

902 77. Ogden, N. H., Milka, R., Caminade, C. & Gachon, P. Recent and projected future climatic
903 suitability of North America for the Asian tiger mosquito Aedes albopictus. *Parasit.*
904 *Vectors* **7,** in press (2014).

905 78. Butterworth, M. K., Morin, C. W. & Comrie, A. C. An Analysis of the Potential Impact of
906 Climate Change on Dengue Transmission in the Southeastern United States. *Environ.*
907 *Health Perspect.* **125,** 579–585 (2017).

908

Figures

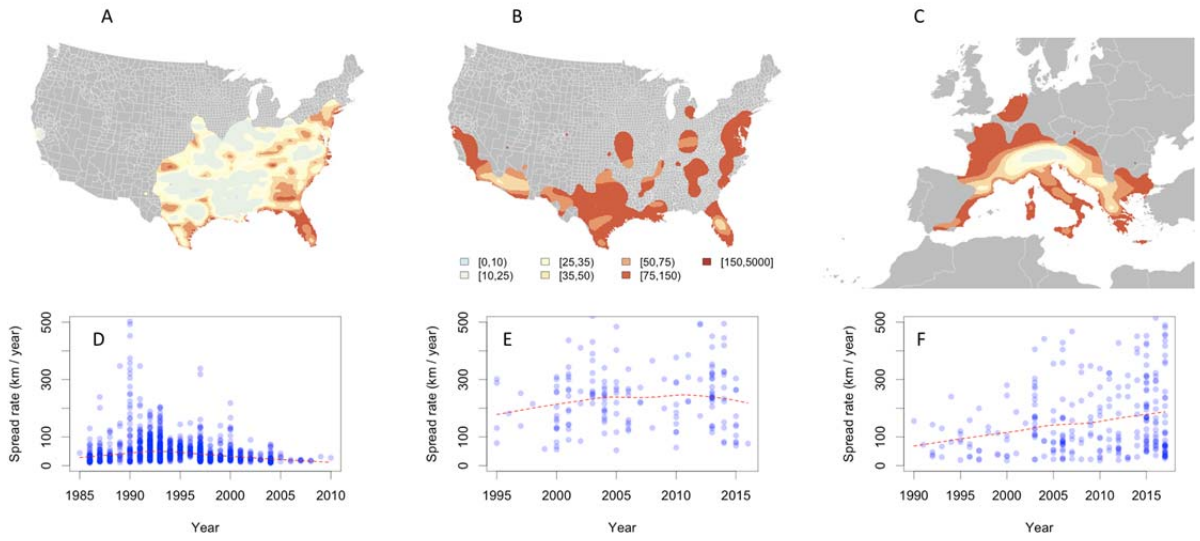


Fig. 1: Reconstruction of *Ae. albopictus* and *Ae. aegypti* spread in the United States (a and b respectively), and *Ae. albopictus* in Europe (c). Estimates of speed of spread in km/year are based on thin spline regression on mosquito observations since their earliest detection in each continent. Red indicates fast dispersal (km/year) whereas yellow and white indicate slower spread (km/year) velocity (see legend below panel b). Areas highlighted in grey have no reported mosquito presence. Panels d – f summarise the speed of dispersal of *Ae. albopictus* and *Ae. aegypti* spread in the United States (d, e) and of *Ae. albopictus* in Europe (f) starting from their date of first detection until 2017. The red line indicates the average velocity per year across all districts using the thin spline regression model.

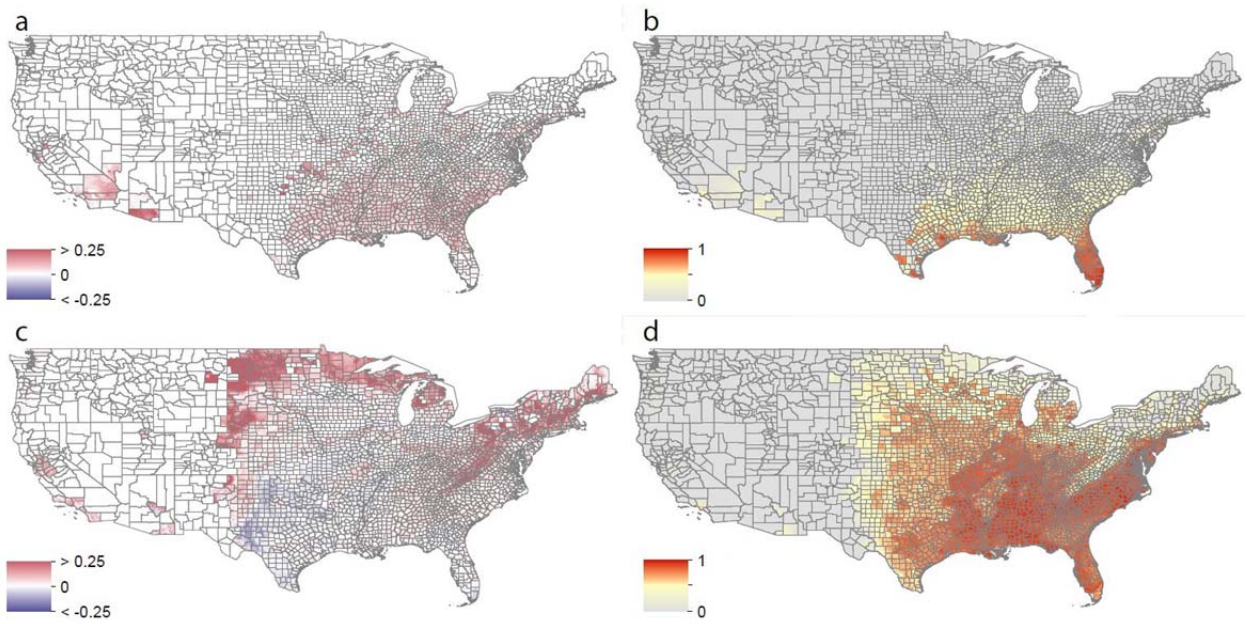


Fig. 2: Predicted future spread of *Aedes aegypti* and *Aedes albopictus* in the United States, estimated using human-mobility metrics and ecological determinants fitted to past occurrence data. Panel A shows the forecasted change in the distribution of *Ae. aegypti* between 2020 and 2050 using the medium climatic scenario Representative Concentration Pathways 6.0 at the United States county level ranging from -0.25 (blue) to 0.25 (red). Red indicates expansion and dark blue contraction of the *Aedes* range distribution between 2020 and 2050. Panel b shows the predicted suitability of presence of *Ae. aegypti* in 2050. Pixels with no predicted suitability are coloured in grey. Panels c and d show the corresponding results for *Ae. albopictus.*

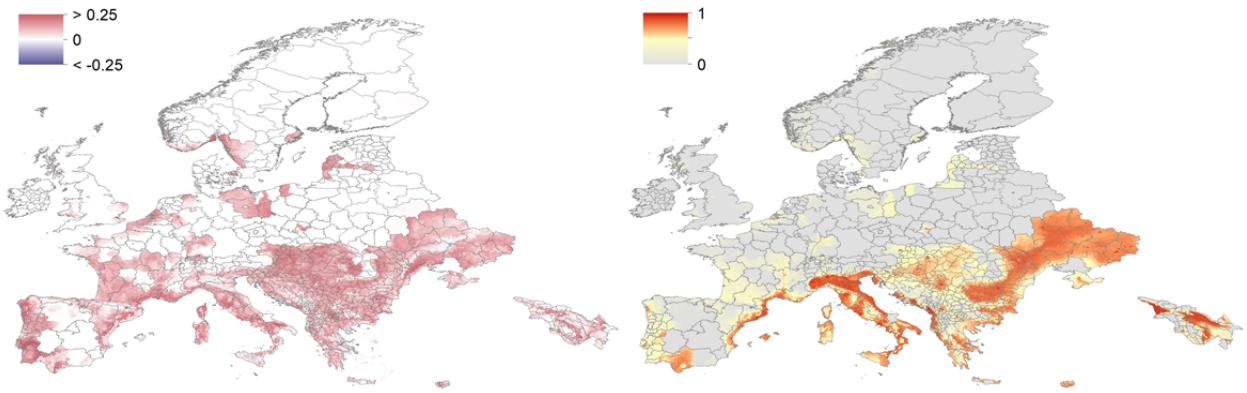


Fig. 3: Predicted future spread of *Aedes albopictus* in Europe. Panel a shows the expansion (red) and contraction (blue) of *Ae. albopictus* between 2020 and 2050 under the medium climate scenario RCP6.0 with emissions peaking in 2080. Panel b shows the predicted distribution of *Ae. albopictus*. Panel b shows the predicted suitability of presence of *Ae. albopictus* in 2050. Pixels with no predicted suitability are coloured in grey.

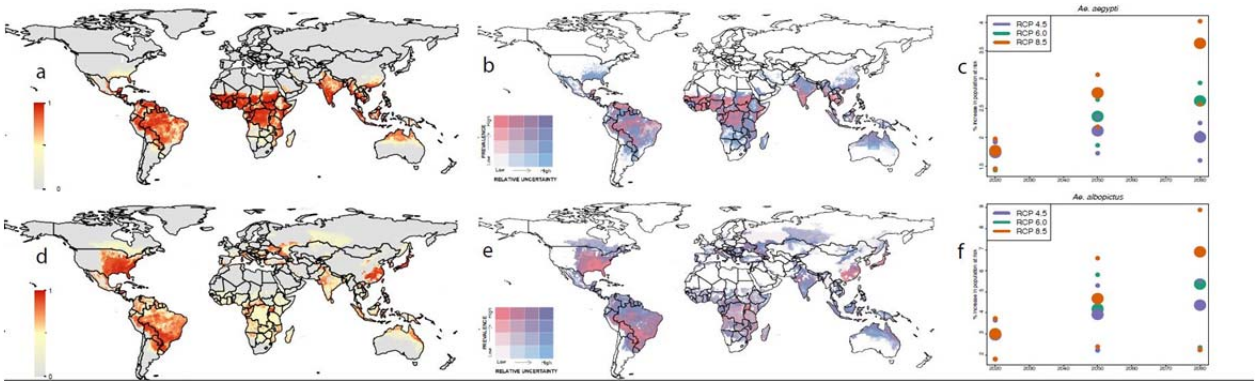


Fig. 4: Predicted global geographic distribution of *Ae*. *aegypti* (a) and *Ae*. *albopictus* (c) in 2050 under the medium climatic scenario RCP6.0 and uncertainty for *Ae. aegypti* (b) and *Ae. albopictus* (e). Predicted suitability of Ae. *aegypti* quantile cutoff points were 0.24, 0.66, 0.88. Relative uncertainty was computed as the ratio of the 95% uncertainty intervals and predicted *Ae. aegypti* suitability for each pixel. Cutoff points for uncertainty were 0.08, 0.18, 0.31. The lowest quantile of predicted suitability is shown in white, and the highest in dark pink. The lowest quantile for uncertainty is white and the highest is blue. The colours overlap such that areas coloured purple have both high predicted suitability of *Ae. aegypti* and high relative uncertainty. Pixels with no predicted suitability are coloured in grey. Panel c show the corresponding results for *Ae. albopictus.* Predicted suitability of Ae. *albopictus* quantile cutoff points were 0.13, 0.41, 0.70. Cutoff points for uncertainty for *Ae. albopictus* were 0.16, 0.36, 0.53. The global population predicted to live in areas suitable for *Ae. aegypti* (b) and *Ae. albopictus* (c) under the conservative (RCP4.5), medium (RCP6.0), and worst-case scenario (RCP8.5) using the binary cutoff values of suitability of 0.46 and 0.51 for both species respectively.

• *Acknowledgements:* We want to thank Sarah Ray for editorial review. MUGK acknowledges funding from the Society in Science, The Branco Weiss Fellowship, administered by the ETH Zurich. MUGK also acknowledges funding from the Training Grant from the National Institute of Child Health and Human Development (T32HD040128). MUGK and SIH acknowledge funding from the International Research Consortium on Dengue Risk Assessment Management and Surveillance (IDAMS; European Commission 7th Framework Programme #21893). SIH received a grant from the Research for Health in Humanitarian Crises (R2HC) Programme, managed by ELRHA (#13468) which also supported MUGK and NG. The R2HC programme aims to improve health outcomes by strengthening the evidence base for public health interventions in humanitarian crises. The £8 million R2HC programme is funded equally by the Wellcome Trust and Department of International Development (DFiD), with Enhancing Learning and Research for Humanitarian Assistance (ELRHA) overseeing the programme's execution and management. SIH was also funded by a Senior Research Fellowship from the Wellcome Trust (#95066) and grants from the Bill & Melinda Gates Foundation (OPP1106023, OPP1093011, OPP1132415 and OPP1159934). This study was made possible by the support of the American people through the United States Agency for International Development Emerging Pandemic Threats Program-2 PREDICT-2 (Cooperative Agreement number AID-OAA-A-14-00102), which also supported MUGK. JSB is supported by the National Library of Medicine of the National Institutes of Health (R01LM010812, R01LM011965), which also supports MUGK. DLS is funded by the National Institutes of Health and National Institute of Allergy and Infectious Diseases (#U10AI089674). HHN was funded by the European Commission through the European Research Council Advanced Investigator Grant `Momentum' 324247. LL received funding from the French Government's Investissement d'Avenir program, Laboratoire d'Excellence Integrative Biology of Emerging Infectious Diseases (grant ANR-10-LABX-62-IBEID), the French Agence Nationale de la Recherche (grant ANR-16-CE35-0004), the City of Paris Emergence(s) programme in Biomedical Research, and the European Union's Horizon 2020 research and innovation programme under ZikaPLAN grant agreement No 734584. NG is supported by a University of Melbourne McKenzie fellowship. WVB, GH, and FS acknowledge funding from VBORNET and VectorNet, an ECDC and EFSA funded project (#ECDC/09/018 and OC/EFSA/AHAW/2013/02), and thank all contributing VBORNET/VectorNet experts for data sharing. TWS, RCR, and LL received funding from the National Institutes of Health Program Project grant (#P01 AI098670). XL is supported by the Natural Science Foundation of China (71522014, 71725001, 71690233 and 71731009).

Author Contributions: All contributions are listed in order of authorship. Designed the experiments: MUGK, RCR, OJB, SIH, NG; Provided data: SL, XL, PJ, LB, EW, AJT, GEC, RGC, WVB, GH, FS, CGM, HY, QL; Analyzed the data: MUGK, RCR, OJB, JPM, MG; Interpreted the results: MUGK, RCR, OJB, JPM, MG, DY, DB, TAP, HHN, DLS, LL, SC, NRF, OGP, TWS, GRWW, SIH, NG; Edited the manuscript: JPM, LBM, NDW, SS, GRWW, SIH; Wrote the manuscript: MUGK, OJB, OGP, SIH, NG; All authors read and approved the content of the manuscript.

Conflicts of Interest: We declare no competing financial interests.

Reprints and permissions information is available at www.nature.com/reprints

Correspondence and requests for materials should be addressed to moritz.kraemer@zoo.ox.ac.uk