

A new malaria vector in Africa: Predicting the expansion range of *Anopheles stephensi* and identifying the urban populations at risk

Sinka, M. E.^{1*}, Pironon, S.², Massey, N. C.³, Longbottom, J.⁴, Hemingway, J.⁴, Moyes, C. L.^{3†}, Willis, K. J.^{1,2†}

1. Department of Zoology, University of Oxford, Oxford, UK.
2. Royal Botanic Gardens Kew, Richmond, Surrey, UK.
3. Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, University of Oxford, Old Road Campus, Oxford, UK.
4. Department of Vector Biology, Liverpool School of Tropical Medicine, UK.

*Corresponding author

† Senior authors

marianne.sinka@zoo.ox.ac.uk

marianne.sinka@gmail.com

Key words: Vector, urban malaria, ensemble modelling, species distribution model, invasive species, mosquito

Lay summary

Africa suffers the highest incidence of malaria globally, partly due to the highly efficient resident mosquito species that transmit the malaria parasite. These mosquito species tend to avoid polluted locations and consequently malaria transmission is generally higher in rural sites compared to more densely populated urban centres.

In 2012, an unusual outbreak of urban malaria was reported from Djibouti City in the Horn of Africa and increasingly severe outbreaks have been reported annually ever since. Subsequent investigations discovered the presence of an Asian mosquito species; *Anopheles stephensi*, a species known to thrive in urban environments. Since that first report, *An. stephensi* has been identified in Ethiopia and Sudan and this worrying development has prompted the World Health Organization to publish a vector alert calling for active mosquito surveillance in the region.

Using an up-to-date database of published locational records for *An. stephensi* across its full range (Asia, Arabian Peninsula, Horn of Africa) and a set of spatial models that identify the environmental conditions that characterise a species' preferred habitat, we provide evidence-based maps predicting the possible locations across Africa where *An. stephensi* could establish if allowed to spread unchecked. Unsurprisingly, due to this species' close association with man-made habitats, our maps predict a high probability of presence within many urban cities across Africa where our estimates suggest that over 126 million people reside. Our results strongly support the WHO's call for surveillance and targeted vector control, and provide a basis for the prioritisation of surveillance.

Summary

In 2012, an unusual outbreak of malaria occurred in Djibouti City followed by increasingly severe annual outbreaks. Investigations revealed the presence of an Asian mosquito species; Anopheles stephensi, which thrives in urban environments. Anopheles stephensi has since been identified in Ethiopia and Sudan.

By combining data for *An. stephensi* across its full range (Asia, Arabian Peninsula, Horn of Africa) with spatial models that identify the species' preferred habitat, we provide evidence-based maps predicting the possible African locations where *An. stephensi* could establish if allowed to spread. Our results suggest over 126 million people in cities across Africa could be at risk. This supports the WHO's call for targeted *An. stephensi* control and prioritised surveillance.

\body

Introduction

Sub-Saharan Africa suffers the highest global levels of malaria mortality and morbidity [1]. This is a direct consequence of the highly efficient mosquito vector species found on the continent, which include *Anopheles gambiae* ("the most dangerous animal in the world"), *Anopheles coluzzii*, *Anopheles funestus* and *Anopheles arabiensis* [2-5]. These dominant vector species have evolved alongside humans, becoming increasingly specialised in seeking out and feeding on human blood [2, 6].

Malaria has historically been considered to be a disease confined to rural locations [7]. Ongoing urbanisation and immigration into urban centres have resulted in cities with extensive areas of urban agriculture, untended green space and unplanned urban sprawl with poor water management. Such areas share characteristics with rural sites and, consequently, mosquito vectors including both *An. arabiensis* and *An. gambiae* are able to maintain malaria transmission [7], in some cases, at prevalence rates up to 30-40% [8].

In contrast to the endemic African mosquitoes, the Asian malaria vector *Anopheles stephensi* is one of the few anopheline species found in central urban locations. This is presumed to be as a direct consequence of its ability to locate clean water in water storage tanks to lay its eggs [9]. The more transient larval sites, such as puddles or ditches more commonly used by dominant African species, are more likely to be turbid or polluted (e.g. with oil or sewage) in urban settings and repel pollution sensitive anophelines. However, there is increasing evidence that some African species are increasing their tolerance for 'dirty' water. Sinka *et al* [5] summarised the bionomics of the dominant vector species in the published literature and found an equal number of studies (5:5) reported *An. arabiensis* in polluted, turbid water as were found in clear, clean habitats, with a similar result for *An. gambiae* (4:4). Nonetheless urban *Plasmodium falciparum* transmission rates are repeatedly reported as significantly lower than those in peri-urban or rural areas [7, 10]. Hay *et al* [10] conducted a meta-analysis in cities from 22 African countries and reported a mean urban annual *P. falciparum* entomological inoculation rate (APfEIR) of = 18.8 notably lower than the peri-urban (APfEIR = 63.9) or rural (APfEIR = 111.4-141.1) areas and concluded: '...there is clear evidence that urbanization affects anopheline species in the environment — diversity, numbers, survival rates, infection rates with *P. falciparum* and the frequency with which they bite people are all affected'.

Anopheles (Cellia) stephensi (Liston) is capable of transmitting both *P. falciparum* and *Plasmodium vivax* parasites { #17;Sinka, 2011 #9}. It is a pervasive mosquito, with an extensive geographical range reaching from Thailand through the Indian Subcontinent, westwards across the Arabian Peninsula and northwards across southern China [9]. Of the three known forms (*type*, *intermediate* and *mysorensis*) the *type* and *intermediate* forms of *An. stephensi* are efficient vectors in both rural and urban environments [11-13] but it is its capacity to survive and proliferate in urban locations that separates this species from most of the other dominant vectors of malaria [9].

There is increasing evidence that this species is expanding its geographical range, with the *type* form being found for the first time in Sri Lanka in 2016 [14-16] and crossing from the Arabian Peninsula into Africa where it has been reported in Djibouti City in 2012, in Ethiopia in 2016 and 2018 (where it

has now been described as ‘...widely distributed and established in Eastern Ethiopia.’ [17]) and Sudan [18-20]

Phylogenetic analyses of Ethiopian (Kebri Dehar) and Djibouti specimens reveal that these two populations are from different clades that originate from different sources in Pakistan [19]. Despite only recently being detected, it is unknown how long *An. stephensi* may have actually been present in the Horn of Africa. Either it has spread very quickly, or it has remained unnoticed, possibly as a consequence of its preference to feed during twilight hours outdoors (assuming these characteristic behaviours reported from Asian *An. stephensi* are also found in African populations) allowing it to avoid capture using typical African mosquito survey methods. Indeed, Balkew *et al* [17] comment that they captured few *An. stephensi* in their CDC light traps, and this is a method of survey often applied in Ethiopia. However, expert entomologists have surveyed widely in the Horn of Africa over many decades and, despite its known morphological resemblance to *An. arabiensis*, had *An. stephensi* been present it would have been identified (Coetzee, 2019 pers com).

The incursion of *An. stephensi* into Africa is particularly worrying; over 40% of sub-Saharan Africans live in urban environments [21], prompting the WHO to issue a vector alert. They consider the presence of *An. stephensi* to be “...a major threat to malaria control and elimination in Africa...” [22] putting urban populations at significantly increased and potentially new risk of malaria transmission.

The increasing range

Nearly a decade ago, a four year project collated georeferenced locational data for 41 dominant vector species of malaria, including *An. stephensi*. It resulted in the publication of a set of evidence-based species distribution maps derived from models identifying the environmental drivers that described the niche of each species [5, 9, 23]. The maps predicted the probability of where, within its given range, each species was likely to be found.

The modelling methodology combined environmental covariates and occurrence data with an expert opinion (EO) species range, used to provide a boundary to the species’ distributions. These EO polygons encompassed all the known presence data and were agreed to be representative of the species’ accepted range by an international group of vector experts. The EO ranges for the Asian-Pacific species were created in 2011 and were developed from the expert opinion range presented by White [24] in 1989.

As part of an ongoing study [25], we have updated our vector database of occurrences for the 19 species/species complexes of the Asian-Pacific region. It now includes presence data reported between 1985-2016 and has significantly increased the number of geolocated sites for *An. stephensi* from 251 to 358 confirmed locations.

To assess how far *An. stephensi* has expanded its range over the past decade, we overlaid the updated *An. stephensi* occurrence data, including points from Africa and Sri Lanka, on the 2011 EO range. The resulting map (Figure 1) illustrates the new reports of *An. stephensi* across the Arabian peninsula and into the Horn of Africa. It starkly highlights the need for a contemporary assessment of this species’ range.

A new map of environmental suitability

Using the updated *An. stephensi* occurrence data, a set of background data (see methods), and *biomod2* (a species distribution modelling platform run in R [26-29]), we created ensemble map projections of *An. stephensi* across its range and into Africa. The *biomod2* package allows comparison between multiple species distribution models; those performing well are then combined

to create the final ensemble map, thus addressing uncertainties inherent in individual model methodologies. Our maps are composed of outputs from 500 model runs, using five species distribution modelling algorithms (Maximum Entropy (MAXENT) [30, 31], random forest [32], generalised boosted models [33], generalised additive models [34], and multiple adaptive regression splines (MARS) [35]).

Most species distribution models assume some level of equilibrium between the species and its environment [36] [37] whereas species that are invading and expanding into new locations are not stable. An ensemble model methodology will mitigate this instability to some extent [36, 38, 39] but it is important to note that our outputs include some extrapolation into novel environments and we are predicting a species not yet at equilibrium.

To establish whether i) the invading African *An. stephensi* is occupying a similar ecological niche to its Asian counterparts (i.e. is the invasion a consequence of a behavioural or ecological divergence), and ii) our models are accurately predicting suitable habitats in this new continent, we created two separate maps of habitat suitability in Africa: the first, our *exclusive* map, used all occurrence data except those from Africa and used the environmental associations found in Asia to predict habitat suitability in Africa. The second map, our *inclusive* projection, used the full occurrence database including those data from Africa.

We predict across the continent to highlight the potential environments this species could occupy if no action is taken. We discuss the maps in the context of the potential increased populations at risk (PAR) this expansion could cause as well as the likelihood of this 'worst case scenario'.

Results

Our map without African data (the *exclusive* map) used 343 occurrences. In comparison, the *inclusive* map used all 358 occurrences. Both models were provided with a set of seven environmental covariates, chosen as relevant to the ecology and bionomics of the dominant vectors of the Asian-Pacific region (SI Appendix (Table S1 & Table S2)). Both the *inclusive* and *exclusive* ensemble models selected the annual mean temperature to be the most influential (SI Appendix (Table S1)), with 86% of the models of the *exclusive* ensemble and 77.3% of the *inclusive* models selecting this as the primary influencing covariate. Human population density appears to be more important when the models include the African data, with 23.2% of the *inclusive* models selecting this variable first compared to 13.2% of the *exclusive* models. Temperature and human population density were most consistently selected either first or second in both the *inclusive* and *exclusive* models, and therefore of the covariates provided, are predicted to have the greatest influence on the niche of *An. stephensi*. Our evaluation statistics indicate the ensemble model is marginally better at predicting species distribution when the African data are included (*Inclusive* ensemble TSS: 0.907, ROC: 0.987; *Exclusive* ensemble: TSS: 0.897, ROC: 0.985).

To assess how the explanatory variables may influence the model outputs, response curves were generated for both the *inclusive* and *exclusive* models for all seven covariates (SI Appendix, Fig S2)). Mean annual temperature produced an expected uni-modal response for both models, although the few additional African data appear to cause the model to predict a slightly increased optimal temperature range. The mean temperature calculated across all *An. stephensi* sites showed a slight increase when the 15 African sites were included (without African sites, $x = 25.77^{\circ}\text{C}$, $n = 343$). with African sites, $x = 25.94^{\circ}\text{C}$, $n = 358$). The population density curve suggests an almost a binary response; the presence of people equates to the presence of *An. stephensi* however; mosquito vector surveys are unlikely to be conducted at uninhabited locations.

Overlaying the confirmed occurrence points in Africa over the *exclusive* ensemble map (Figure 2) strongly suggests that the recent western expansion of *An. stephensi*, specifically inland from the established coastal populations, is due to some behavioural or niche adaptation. The model predicts the confirmed *An. stephensi* sites in Djibouti and in Semera and Jigjiga in Ethiopia well with high predicted suitability but fails to predict high suitability in Degehabur, Kebridehar or Godey in Ethiopia. Coastal locations in Sudan are predicted as highly suitable, but inland sites are not.

Populations at risk (PAR)

Our final *inclusive* ensemble model was created using the full dataset, including those occurrences in Djibouti, Ethiopia and Sudan. The *inclusive* map (Figure 3), when compared to the *exclusive* output (Figure 2), predicts all the confirmed locations in Africa, including the inland sites in Ethiopia and Sudan. It demonstrates an increased area of predicted suitable habitat as a consequence of the wider range of conditions at locations where African *An. stephensi* have been found. Most of the locations predicted to be highly suitable are urban, densely populated cities that traditionally incur low levels of malaria transmission. Many of these areas are, however, within malaria endemic zones, i.e. the surrounding rural and peri-urban locations have a parasite reservoir present within the local human populations and provide suitable habitats for the efficient dominant vector species primarily responsible for sub-Saharan Africa's high malaria burden. The city environment currently provides a barrier to, or reduction in, malaria transmission but this would be removed when or if *An. stephensi* were able to establish. These densely populated areas may hold a significant reservoir of epidemiologically naïve people at high risk of malaria (depending on the level of population movement).

Using the Brinkoff City Population database [40], filtered to include only those cities with a population equal to or greater than one million, we overlaid the city locations on the inclusive ensemble map (Figure 3, SI Appendix (Fig. S2)), with the malaria endemic areas of Africa shown in more detail in Figure 4, which illustrate the predicted urban preference by *An. stephensi* in Africa. All African cities with 1M inhabitants or more were assessed using the habitat suitability probability values generated by this study, their distance from the current records of *An. stephensi* collated by this study, and their distance from the malaria endemic zone. Table 1 is an abridged table, showing the top fifteen cities, ordered by the increasing distance from the nearest African site where *An. stephensi* has been confirmed as present (Djibouti, Ethiopia or Sudan) that vary according to the predicted probability of *An. stephensi* establishing populations. The full table, listing 68 cities, is shown in the SI Appendix (Table S3). The median and quartile values for the suitability predictions shown in Figure 3 and 4, within the current extent of this species, were used to classify the habitat suitability within each city. Most (44/68) are predicted as highly suitable locations for *An. stephensi*. Of these 44 highly suitable cities, 36 are found within the current malaria endemic zone (Figure 4 and SI Appendix (Table S3)). As such, there could be an estimated additional 126 million people at risk of malaria if *An. stephensi* is able to continue its invasion across the continent.

Bionomics

Anopheles stephensi is a common laboratory species (as a consequence of its ability to breed in small containers), yet there is little quantified data describing the bionomics of *An. stephensi* in the wild. No reported entomological inoculation rate (EIR – the number of infected mosquito bites received per unit time) and only a single human biting rate value (HBR – the number of mosquito bites per unit time) could be found in the literature [41]. In a brief summary of Asian species bionomics, Sinka

et al. [9] indicate *An. stephensi* primarily prefers to feed on animals (zoophilic) with examples of human biting (anthropophagy) indoors (endophagic) with outdoor (endophilic) resting behaviour, but with examples of outdoor biting (exophagic) behaviour, with a preference for biting during dusk (crepuscular) and the night. Similar to many of the Asian dominant vector species, *An. stephensi* appears to have some level of plasticity in its behaviour, specifically in host choice, with rural *An. stephensi* tending to be more zoophilic than urban populations [42] [41]. The mosquito's biting behaviour is driven by host availability including human activity patterns and livestock sheltering practices. Fed female mosquitoes are reported as preferring to rest in poorly constructed buildings or animal sheds [41-44]. Such structures are characteristic of unplanned urban sites and can be found around and within many rapidly expanding cities in Africa [45, 46].

Many malaria vector surveys include hourly mosquito catches that run between 6pm and 6am, when anopheline vectors are typically most active. Using data from the Malaria Atlas Project's bionomics database [47], nightly activity for the three primary African vector species (*An. gambiae*, *An. arabiensis* and *An. funestus*) are compared with *An. stephensi* (Figure 5). This simple plot illustrates that *An. stephensi* has peak activity (i.e. is actively seeking a blood meal) earlier in the evening than any of the African DVS, with over half of the total bites from this species occurring within the first quarter of the night (6pm – 9pm). Activity can be observed later, but at a much lower rate [41]. African vector species are known to prefer to bite when humans are asleep and vulnerable, and as expected, *An. gambiae* and *An. funestus* both have peak activity well into the night (3rd quarter: midnight to 3am). Whether newly establishing African *An. stephensi* populations maintain a preference for biting in the evening remains to be seen, and there is some indication that it may change its outdoor biting activity depending on the season [41]. However, its crepuscular biting preference would allow *An. stephensi* to bypass insecticide treated bednets; the most widely used (and currently most effective) mosquito interventions used across Africa [48, 49]. *Anopheles stephensi* larvae have been found coexisting in water storage containers with *Aedes aegypti* (a vector of yellow fever and dengue amongst others) and *Culex quinquefasciatus* (the southern house mosquito – vector of lymphatic filariasis and West Nile fever) so interventions already in place to target these mosquitoes, as well as better management of urban infrastructure including water and green spaces, could help control *An. stephensi* [50-52]

Anopheles stephensi is known to have a seasonal peak in abundance, normally in the late spring/early summer months [41, 53] [42] and shows some regulation by climatic conditions and here we have seen annual temperature is the most influential covariate in describing this species' niche. In urban areas *An. stephensi* has been found in roof gutters, domestic wells, air conditioning units and cisterns [54] [9] [51]. It is not reliant on rainfall to fill its larval sites like many other species (e.g. *An. gambiae* that makes use of temporary rain filled puddles), but takes advantage of human built and maintained water storage facilities, particularly cement or brick structures [19, 20, 41, 42, 51, 54, 55]; a necessity to the human urban populations during the dry summer months. Indeed it is *An. stephensi*'s ability to utilise such human built water storage tanks that may have facilitated its ability to expand its range [15]. Within our model, human population density was also a highly influential variable and our final ensemble map suggests that many large urban expanses are highly suitable if this species continues to invade Africa, whereas most rural areas are not.

Discussion

Anopheles stephensi is evidently expanding its geographic range (Figure 1). Whether this is a recent event or has occurred over the longer term is still unclear. Nonetheless, this invasion has already been implicated in recent outbreaks of urban malaria in Djibouti [20]. Djibouti was entering the pre-

elimination stage, with the last outbreak of malaria reported in 1999. However, in January 2013, outbreaks of *P. falciparum* were reported, predominantly in Djibouti City, including historically malaria-free locations that were considered to lack suitable anopheline larval habitats [20]. Since 2013, outbreaks have recurred annually in the city with increasing intensity. Reported and confirmed cases in the country have increased from only 25 in 2012 to 9473 in 2015 { #49} and 14810 cases reported in 2017 [18]. *Anopheles stephensi* is highly susceptible to infection by African *P. falciparum* isolates [56] and also *P. vivax*. Vivax transmission was reported in Djibouti for the first time in 2016, and again in 2017 [18]. Further, there is some indication that *An. stephensi* may already be adapting to the African environment, changing from predominantly seasonal activity reported between 2013-2016 to year round activity from 2017 [18].

Carter et al [19] provide genetic evidence that the Djibouti and Ethiopian *An. stephensi* populations are a result of separate introduction events, both originating from Pakistan. Our *exclusive* model was able to predict the Djibouti populations but not the full extent of the Ethiopian sites indicating some variation in the niche characteristics between the Djibouti and Ethiopian populations. Whether these adaptations are intrinsic in the original Pakistan populations (suggesting the full extent of the Asian niche has yet to be described) or have occurred since its expansion into Africa, can only be teased apart with further phylogenetic analyses. However, when comparing the *exclusive* and *inclusive* ensemble maps over Asia, there appears to be very little difference in the predicted extent, suggesting these additional data are not adding new niche information to the predictions in Asia.

It is likely that *An. stephensi* individuals have been inadvertently introduced into Africa on multiple occasions, but it appears that in Djibouti at least, this species has only recently established persistent populations. The situation in Ethiopia remains less clear; Balkew *et al.* [17] report *An. stephensi* as widespread across eastern Ethiopia and hypothesise that either *An. stephensi* is a recent introduction or that it has been present but remained undetected due to morphological similarity to *An. arabiensis*. Whether the specimens from Sudan clarify the situation or raise further questions is yet to be seen.

Species distribution models provide a snapshot of the species' suitable niche based on a set of criteria defined by the covariates provided to the model at the locations where the species has been shown to be present. This correlative approach does not, however, provide a mechanistic interpretation of how the species came to exist within that niche and the predicted output does not always consider biological and physical barriers to species movement.

Our database is the most comprehensive set of dated and geolocated, species-specific occurrence records currently available. However, as with all presence data, it provides information about a single point in space and time; although we know the date of sampling, we do not know how long prior to that positive occurrence *An. stephensi* has been inhabiting that site, nor if it continues to exist there. These records can be considered opportunistic sampling and do not come from continuous monitoring programmes. They do not fully represent the geographical and environmental space considered by our study. Thus we cannot use this modelling technique to help answer questions about how and why *An. stephensi* appears to be expanding its range.

Nonetheless, generating response curves provides some indication of the potential niche width and how *An. stephensi* interacts with specific environmental covariates supplied to the model. Here, as expected, we see a unimodal response to temperature, suggesting that there are sites represented that are beyond the optimal temperature for the species to exist. Moreover, we see that the few additional sites in Africa appear to expand the optimal temperature range, which could suggest a

recent potential expansion of its thermal niche (although as this is based on few data, it should be treated with caution).

A recent study that modelled the environmental suitability of malaria transmission and temperature using *An. stephensi* data compared to multi-species models (primarily using *An. gambiae* data), reported a wider thermal niche in the *An. stephensi* model [57]. It implicated *An. stephensi* as having a thermal minimum 3.4 °C lower and thermal maximum 3.4 °C higher than the multispecies estimates. *Anopheles stephensi* is active at dusk and as such may have more tolerance for higher temperature than typical anophelines that tend to become active later in the evening. Moreover, as an urban species, *An. stephensi* are routinely exposed to higher temperatures generated within cities (the urban heat island effect) compared to rural counterparts and could be more buffered from the influence of climate than rural species.

Increasing or changing climatic temperature does not appear to explain the movement of this mosquito across the Arabian peninsula and into Africa. It is more likely to be due to the movement of people and livestock. Indeed presences reported from Ethiopia appear to be located along primary transportation routes (pers com, S. Irish, 2020), although a full randomised survey would be needed to confirm these observations. With African cities growing at an extraordinary rate and 43% of Africans now living in urban areas [58], the movement of people into urban sites and the observation that *An. stephensi* is spreading via key transportation routes, mean it is not inconceivable that this species could be transported large distances and that this may not remain a problem centred on the Horn of Africa. As illustrated in Figure 3, many cities across the whole of Africa contain potentially suitable habitats for *An. stephensi*. For example, due to its high population density, Nigeria, appears to be particularly suitable.

The escalation of urban Falciparum malaria as well as emerging Vivax in Djibouti [18, 20] needs to act as a stark warning, one succinctly summarised by Takken & Lindsay [59]: “If... [*An. stephensi*] invades large cities, such as Khartoum, Sudan; Mombasa, Kenya; and Dar es Salaam, the region could face malaria outbreaks of unprecedented size.” Once established, *An. stephensi* is difficult to control. Traditional mosquito interventions via indoor residual spraying and treated bednets are notably more difficult to implement in cities, and the crepuscular biting habits of *An. stephensi* suggest they may have less impact on this species than the dominant African vectors.

Anopheles stephensi is not the first malaria vector to invade a new continent. In one notable historic example, a member of the *An. gambiae* complex (since shown to be *An. arabiensis* [60]) established in the city of Natal, north-eastern Brazil. Similar to the current situation, the presence of *An. arabiensis* was accompanied by a significant increase in malaria transmission [61]. Although *An. arabiensis* was quickly eliminated from the city, the larval treatment used was not extended to the surrounding areas and thus *An. arabiensis* was able to move out of the city unnoticed (dubbed the ‘silent spread’) and increase its population until eventually a ‘catastrophic malaria epidemic’ [62] began in 1938 within the states of Rio Grande do Norte and Ceará [61, 62]. It took the formation of a new anti-malaria service and a multi-million dollar budget to finally eradicate this mosquito from Brazil [62]. This is a lesson in history that would be advisable to note when considering the future course of action against *An. stephensi* in Africa.

These recent reports of *An. stephensi* in Africa combined with our ensemble maps presented here highlight a substantial future threat to urban African populations [63]. The WHO are urging member states in and around the Horn of Africa, the Republic of the Sudan and surrounding geographical areas to take immediate action including active vector surveillance and reporting, assessing the ecological characteristics if the mosquito is found, as well as assessing insecticide resistance [22]. Here our *inclusive* map provides an evidence based indication of where the environmental

conditions may be suitable for *An. stephensi* to thrive. We distinguish cities that are predicted to be highly suitable within the current malaria endemic zone and those outside it. For example, in the case of cities such as Cairo that are distant from malaria endemic regions in Africa and the Middle East, the potential threat will only become a reality if both the vector and the parasite are introduced [64].

Targeted surveillance is now needed to monitor the existing populations in Africa, optimise the strategies for its control and ensure that this species does not continue its invasion across Africa.

Conclusion

Using a comprehensive and updated dataset incorporating the recently identified occurrence sites in Africa, our predictive map provides an evidence-base for the possible expansion of *An. stephensi* across Africa and, as expected, the probability of occurrence is much higher in the more densely populated areas where huge numbers of people could be at risk. We predict an increase PAR of up to 126 million people if *An. stephensi* were to occupy this predicted niche. If it continues its incursion into the African continent unchecked, there is a very real possibility of mass outbreaks of malaria amongst naïve populations in areas of vector-parasite coexistence. In a continent striving to improve and strengthen its health systems, such a huge burden could be catastrophic. Targeted vector surveillance is therefore urgently needed.

Materials and Methods

Our maps were created using an updated database of *An. stephensi* occurrence ranging from 1985 – 2019, including those sites reported from the Horn of Africa (see SI Appendix). Our background dataset included presences for anopheline species (excluding *An. stephensi*), *Aedes* and *Culex* species across Asia and within the Horn of Africa. We provided a final set of seven relevant environmental covariates, refined from an initial set of 19, and run our models using the BIOMED2 platform [28, 29] in R studio (version 1.2.1335) [27] using R (version 3.6.1) [26]. This platform allows the comparison of multiple distribution model methodologies to provide a consensus map based on those evaluated (True Skill Statistic (TSS) and the area under the receiver operating curve (ROC)) to have the best predictive performance. We provide a Coefficient of Variation map alongside our consensus map. This is a simple evaluation that illustrates areas where the confidence in the model predictions are lower (i.e. where data are sparse). We also provide the ranked, model selected influential variables, generated in *biomod2*, to show which covariates are having the highest influence within the model and the final consensus map.

To evaluate the potential effect of the spread of *An. stephensi*, we estimated the populations at risk for all cities with over 1 million inhabitants based on our predicted maximum suitability values from the pixels within the city. Using QGIS [65] the distance to locations confirmed as having *An. stephensi* were calculated as well as the distance from the combined *P. falciparum* and *P. vivax* endemic zone [16].

Data Availability

The full occurrence dataset used to generate our maps are available via Dryad (Sinka, Marianne *et al.* (2020), Anopheles stephensi occurrence data 1985 - 2019, Dryad, Dataset, <https://doi.org/10.5061/dryad.3xsi3txcx>) or directly from the lead author.

Acknowledgements

This work was funded by the Google Impact Challenge award (Mosquito Detection Project DFR01520 - HumBug) and Wellcome grant 108440/Z/15/Z. Additional thanks to Ian Ondo for mapping advice, Seth Irish, Dave Weetman and colleagues for providing additional data, Marybel Soto Gomez for

assisting in running our final models and to Dan Strickman for a query that inspired this work in the first place.

Author contributions

Marianne Sinka conceived the study, ran the preliminary models and wrote the manuscript. Sam Pironon wrote the modelling code underlying this work and ran and produced the final models. Claire Massey updated the South-East Asian vector occurrence database to include all sibling species and data up to 2016. Josh Longbottom provided the background data points used in the model. Janet Hemingway provided technical advice and Catherine Moyes created the PAR table and was fundamental to the instigation of the project and provided valuable feedback. Kathy Willis provided valuable feedback on the manuscript and support throughout. All authors commented and had input in the final manuscript.

Figure Legends

Figure 1: The new 'out of range' occurrence of *An. stephensi* in the Arabian Peninsula and Horn of Africa showing the 358 *An. stephensi* site locations used in our final SDMs colour coded by the decade in which they were sampled. The yellow shaded area shows the 2011 expert opinion range based on data published up to 31 October 2009 [9]. Data showing the presence of *An. stephensi* more westerly across the Arabian peninsula (sampled in 2005-2006) were published after 2010. Thus *Anopheles stephensi* may have been present but unreported or been expanding its range into western areas of the Arabian peninsula over the last 30 years.

Figure 2: Exclusive Map: Environmental suitability map of *An. stephensi* using the updated occurrence database but without African sites. Red indicates a higher probability of environmental suitability whereas the blue indicates environments with a lower probability, i.e. more likely to be unsuitable for the species to occur. The environmental variables selected by the model as relevant to *An. stephensi* habitat suitability, in descending order (based on correlation score): Ann. Mean Temp. = 0.459, Human Popⁿ Dens. = 0.325, Precip (season) = 0.171, EVI = 0.161, Irrigation = 0.155, TCW = 0.110, Crop mosaic = 0.011. Dark grey circles indicate the location of all 358 occurrence data used in the final inclusive map (Figure 3) including those in Africa (not used in this exclusive model). The thumbnail map shows the coefficient of variation calculated per pixel across the predicted range indicating where the ensemble model provides the most reliable (higher confidence: dark green) and least reliable (lower confidence: red) predictions.

Figure 3: Inclusive Map: Environmental suitability map of *An. stephensi* using the updated occurrence database including all African sites. Red indicates a higher probability of environmental suitability whereas the blue indicates environments with a lower probability, i.e. more likely to be unsuitable for the species to occur. The environmental variables selected by the model as relevant to *An. stephensi* habitat suitability, in descending order (based on correlation score): Ann. Mean Temp. = 0.461, Human Popⁿ Dens. = 0.370, EVI = 0.174, Precip (season) = 0.161, TCW = 0.134, Irrigation = 0.130, Crop mosaic = 0.010. Turquoise circles indicate the location of cities with a population > 1 million. The thumbnail map shows the coefficient of variation calculated per pixel across the predicted range indicating where the ensemble model provides the most reliable (higher confidence: dark green) and least reliable (lower confidence: red) predictions.

Figure 4: The inclusive map focussing on the malaria (*P. falciparum* and *P. vivax*) endemic areas in Africa: Environmental suitability map of *An. stephensi* using the updated occurrence database including all African sites, masked (dark grey) using the MAP *PfPv* endemicity mask [16]. Red indicates a higher probability of environmental suitability whereas the blue indicates environments

with a lower probability, i.e. more likely to be unsuitable for the species to occur. Turquoise circles indicate the location of cities with a population > 1 million.

Figure 5: Mosquito biting activity summarised from published studies (MAP bionomics database) comparing the Asian (open triangles) *An. stephensi* with the African (open circles) DVS (*An. arabiensis*, *An. gambiae* and *An. funestus*) showing the peak activity periods as the proportion of mosquitoes biting in the 1st, 2nd, 3rd and 4th quarters of the night (from 6pm – 6am).

References

1. World Health Organization: **World Malaria Report**. pp. 210: World Health Organization (WHO); 2018:210.
2. Coetsee M: **Distribution of the African malaria vectors of the *Anopheles gambiae* complex**. *Am J Trop Med Hyg* 2004, **70**:103-104.
3. Coluzzi M: **The clay feet of the malaria giant and its African roots: hypotheses and inferences about origin, spread and control of *Plasmodium falciparum***. *Parassitologia* 1999, **41**:277-283.
4. Gillies MT, de Meillon B: *The Anophelinae of Africa South of the Sahara (Ethiopian zoogeographical region)*. 2nd edn. Johannesburg: The South African Institute for Medical Research; 1968.
5. Sinka ME, Bangs MJ, Manguin S, Coetsee M, Mbogo CM, Hemingway J, Patil AP, Temperley WH, Gething PW, Kabaria CW, et al: **The dominant *Anopheles* vectors of human malaria in Africa, Europe and the Middle East: occurrence data, distribution maps and bionomic precis**. *Parasit Vectors* 2010, **3**:117.
6. Charlwood JD, Smith T, Kihonda J, Heiz B, Billingsley PF, Takken W: **Density-independent feeding success of malaria vectors (Diptera: Culicidae) in Tanzania**. *Bull Entomol Res* 1995, **85**:29-35.
7. Robert V, Macintyre K, Keating J, Trape JF, Duchemin JB, Warren M, Beier JC: **Malaria transmission in urban sub-Saharan Africa**. *Am J Trop Med Hyg* 2003, **68**:169-176.
8. Overgaard HJ, Reddy VP, Abaga S, Matias A, Reddy MR, Kulkarni V, Schwabe C, Segura L, Kleinschmidt I, Slotman MA: **Malaria transmission after five years of vector control on Bioko Island, Equatorial Guinea**. *Parasit Vectors* 2012, **5**:253.
9. Sinka ME, Bangs MJ, Manguin S, Chareonviriyaphap T, Patil AP, Temperley WH, Gething PW, Elyazar IR, Kabaria CW, Harbach RE, Hay SI: **The dominant *Anopheles* vectors of human malaria in the Asia-Pacific region: occurrence data, distribution maps and bionomic precis**. *Parasit Vectors* 2011, **4**:89.
10. Hay SI, Guerra CA, Tatem AJ, Atkinson PM, Snow RW: **Urbanization, malaria transmission and disease burden in Africa**. *Nat Rev Microbiol* 2005, **3**:81-90.
11. Rao BA, Sweet WC, Subba Rao AM: **Ova measurements of *A. stephensi* type and *A. stephensi* var. *mysorensis***. *J Malar Inst India* 1938, **1**:261-266.
12. Subbarao SK, Vasantha K, Adak T, Sharma VP, Curtis CF: **Egg-float ridge number in *Anopheles stephensi*: Ecological variation and genetic analysis**. *Med Vet Entomol* 1987, **1**:265-271.
13. Sweet WC, Rao BA: **Races of *A. stephensi* Liston, 1901**. *Ind Med Gaz* 1937, **72**:665-674.
14. Surendran SN, Sivabalakrishnan K, Gajapathy K, Arthiyan S, Jayadas TTP, Karvannan K, Raveendran S, Parakrama Karunaratne SHP, Ramasamy R: **Genotype and biotype of invasive *Anopheles stephensi* in Mannar Island of Sri Lanka**. *Parasit Vectors* 2018, **11**:3.
15. Surendran SN, Sivabalakrishnan K, Sivasingham A, Jayadas TTP, Karvannan K, Santhirasegaram S, Gajapathy K, Senthilnathanan M, Karunaratne SP, Ramasamy R: **Anthropogenic Factors Driving Recent Range Expansion of the Malaria Vector *Anopheles stephensi***. *Front Public Health* 2019, **7**:53.
16. **Malaria Atlas Project (MAP)** [<https://malariaatlas.org/>]

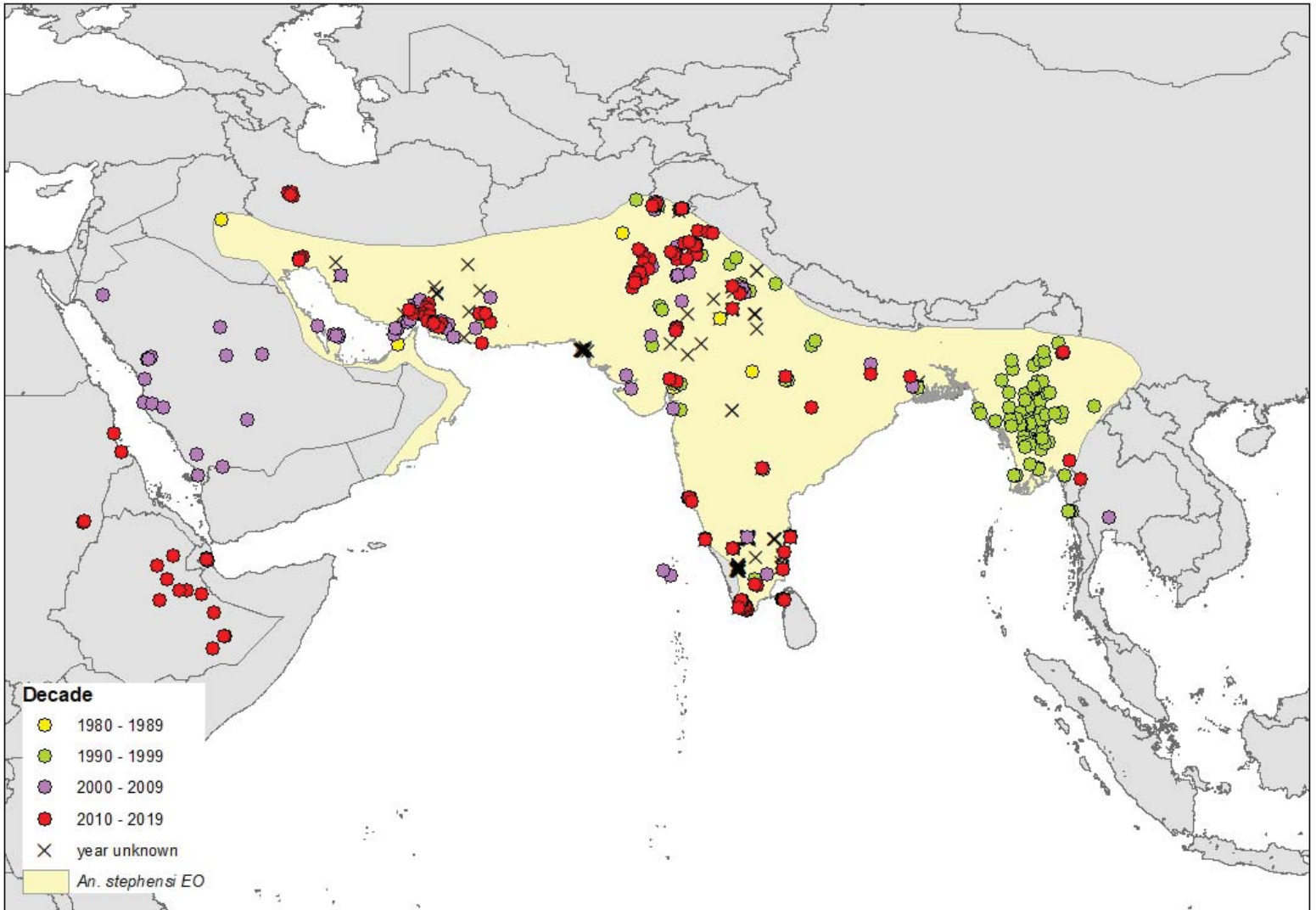
17. Balkew M, Mumba P, Dengela D, Yohannes G, Getachew D, Yared S, Chibsa S, Murphy M, George K, Lopez K, et al: **Geographical distribution of *Anopheles stephensi* in eastern Ethiopia.** *bioRxiv* 2019:802587.
18. Seyfarth M, Khaireh BA, Abdi AA, Bouh SM, Faulde MK: **Five years following first detection of *Anopheles stephensi* (Diptera: Culicidae) in Djibouti, Horn of Africa: populations established-malaria emerging.** *Parasitol Res* 2019, **118**:725-732.
19. Carter TE, Yared S, Gebresilassie A, Bonnell V, Damodaran L, Lopez K, Ibrahim M, Mohammed S, Janies D: **First detection of *Anopheles stephensi* Liston, 1901 (Diptera: culicidae) in Ethiopia using molecular and morphological approaches.** *Acta Trop* 2018, **188**:180-186.
20. Faulde MK, Rueda LM, Khaireh BA: **First record of the Asian malaria vector *Anopheles stephensi* and its possible role in the resurgence of malaria in Djibouti, Horn of Africa.** *Acta Trop* 2014, **139**:39-43.
21. World Bank [<https://data.worldbank.org/indicator/SP.URB.TOTL.IN.ZS?locations=ZG>]
22. World Health Organization: **Vector Alert: *Anopheles stephensi* invasion and spread** [<https://www.who.int/news-room/detail/26-08-2019-vector-alert-anopheles-stephensi-invasion-and-spread>]
23. Sinka ME, Rubio-Palis Y, Manguin S, Patil AP, Temperley WH, Gething PW, Van Boeckel T, Kabaria CW, Harbach RE, Hay SI: **The dominant *Anopheles* vectors of human malaria in the Americas: occurrence data, distribution maps and bionomic precis.** *Parasit Vectors* 2010, **3**:72.
24. White GB: **Malaria. Geographical distribution of arthropod-borne diseases and their principal vectors** pp. 7-22. Geneva: World Health Organization, Division of Vector Biology and Control; 1989:7-22.
25. HumBug [<http://humbug.ac.uk/>]
26. The R Foundation for Statistical Computing: "Action of the Toes" (Copyright (C)) (2019)
27. R Studio Inc, R Studio
28. Thuiller W: **BIOMOD - optimizing predictions of species distributions and projecting potential future shifts under global change.** *Global Change Biology* 2003, **9**:1353-1362.
29. Thuiller W, Lafourcade B, Engler R, Araujo MB: **BIOMOD - a platform for ensemble forecasting of species distributions.** *Ecography* 2009, **32**:369-373.
30. Phillips SJ, Anderson RP, Schapire RE: **Maximum entropy modeling of species geographic distributions.** *Ecol Model* 2006, **190**:231-259.
31. Elith J, Phillips SJ, Hastie T, Dudik M, Chee YE, Yates CJ: **A statistical explanation of MaxEnt for ecologists.** *Divers Distrib* 2011, **17**:43-57.
32. Breiman L: **Random forests.** *Mach Learn* 2001, **45**:5-32.
33. Elith J, Leathwick JR, Hastie T: **A working guide to boosted regression trees.** *J Anim Ecol* 2008, **77**:802-813.
34. Hastie TJ, Tibshirani RJ: *Generalized Additive Models.* Chapman and Hall/CRC; 1990.
35. Friedman JH: **Multivariate Adaptive Regression Splines.** *Annals of Statistics* 1991, **19**:1-67.
36. Elith J, Kearney M, Phillips S: **The art of modelling range-shifting species.** *Methods Ecol Evol* 2010, **1**:330-342.
37. Broennimann O, Guisan A: **Predicting current and future biological invasions: both native and invaded ranges matter.** *Biol Letters* 2008, **4**:585-589.
38. Engler R, Waser LT, Zimmermann NE, Schaub M, Berdos S, Ginzler C, Psomas A: **Combining ensemble modeling and remote sensing for mapping individual tree species at high spatial resolution.** *Forest Ecol Manag* 2013, **310**:64-73.
39. Marmion M, Parviainen M, Luoto M, Heikkinen RK, Thuiller W: **Evaluation of consensus methods in predictive species distribution modelling.** *Divers Distrib* 2009, **15**:59-69.
40. Brinkoff, T., City Population [<http://www.citypopulation.de/>]

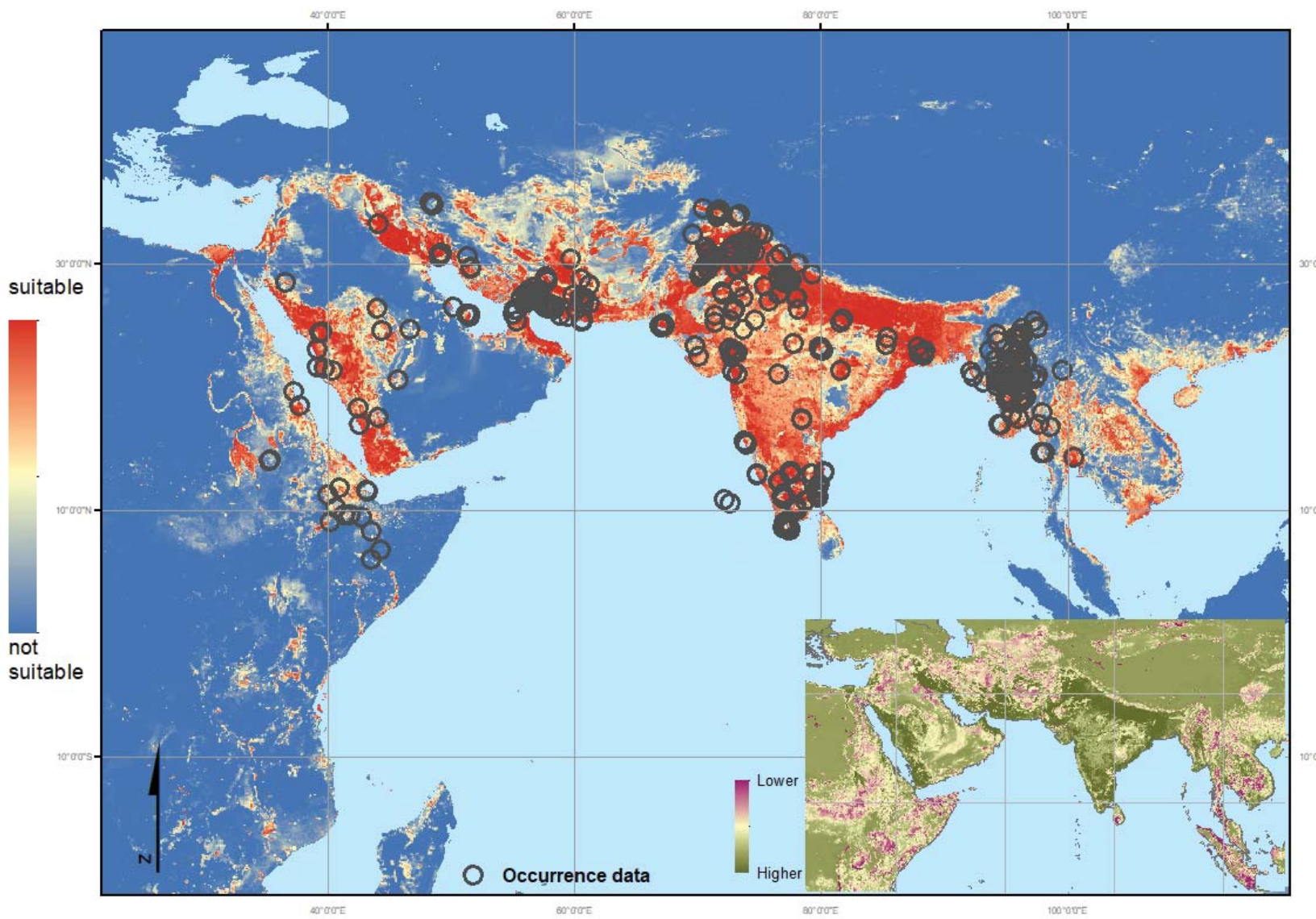
41. Singh OP: **Bionomics of malaria vectors in India**. In *Vector Biology*. pp. 19-31: Malaria Research Centre; 2002:19-31.
42. Batra CP, Adak T, Sharma VP, Mittal PK: **Impact of urbanization on bionomics of *An. culicifacies* and *An. stephensi* in Delhi**. *Indian J Malariol* 2001, **38**:61-75.
43. Chatterjee KK, Biswas D, Choudhuri DK, Mukherjee H, Hati AK: **Resting sites of *Anopheles stephensi* Liston in Calcutta**. *Indian J Malariol* 1993, **30**:109-112.
44. Hati AK, Chatterjee KK, Biswas D: **Daytime resting habits of *Anopheles stephensi* in an area of Calcutta**. *Indian J Malariol* 1987, **24**:85-87.
45. **United Nations - African Renewal** [<https://www.un.org/africarenewal/magazine/april-2012/towards-african-cities-without-slums>]
46. UNHABITAT: *The State of African Cities 2018 - The geography of African Investment*. Erasmus University Rotterdam (EUR),: United Nations Human Settlements Programme (UN-Habitat); 2018.
47. Massey NC, Garrod G, Wiebe A, Henry AJ, Huang Z, Moyes CL, Sinka ME: **A global bionomic database for the dominant vectors of human malaria**. *Sci Data* 2016, **3**:160014.
48. Bhatt S, Weiss DJ, Cameron E, Bisanzio D, Mappin B, Dalrymple U, Battle K, Moyes CL, Henry A, Eckhoff PA, et al: **The effect of malaria control on *Plasmodium falciparum* in Africa between 2000 and 2015**. *Nature* 2015, **526**:207-211.
49. Killeen GF, Kiware SS, Okumu FO, Sinka ME, Moyes CL, Massey NC, Gething PW, Marshall JM, Chaccour CJ, Tusting LS: **Going beyond personal protection against mosquito bites to eliminate malaria transmission: population suppression of malaria vectors that exploit both human and animal blood**. *BMJ Glob Health* 2017, **2**:e000198.
50. Chakraborty S, Ray S, Tandon N: **Seasonal prevalence of *Anopheles stephensi* larvae and existence of two forms of the species in an urban garden in Calcutta City**. *Indian J Malariol* 1998, **35**:8-14.
51. Mariappan T, Thenmozhi, V., Udayakumar, P., Bhavaniumadevi, V., Tyagi, B. K.: **An observation on breeding behaviour of three different vector species (*Aedes aegypti* Linnaeus 1762, *Anopheles stephensi* Liston 1901 and *Culex quinquefasciatus* Say 1823) in wells in the coastal region of Ramanathapuram district, Tamil Nadu, India** *Int J Mosq Res* 2015, **2**:42-44.
52. Weetman D, Kamgang B, Badolo A, Moyes CL, Shearer FM, Coulibaly M, Pinto J, Lambrechts L, McCall PJ: ***Aedes* Mosquitoes and *Aedes*-Borne Arboviruses in Africa: Current and Future Threats**. *Int J Environ Res Public Health* 2018, **15**.
53. Mehravaran A, Vatandoost H, Oshaghi MA, Abai MR, Edalat H, Javadian E, Mashayekhi M, Piazak N, Hanafi-Bojd AA: **Ecology of *Anopheles stephensi* in a malarious area, southeast of Iran**. *Acta Med Iran* 2012, **50**:61-65.
54. Biswas D, Dutta RN, Ghosh SK, Chatterjee KK, Hati AK: **Breeding habits of *Anopheles stephensi* Liston in an area of Calcutta**. *Indian J Malariol* 1992, **29**:195-198.
55. Sharma RS: **Urban malaria and its vectors *Anopheles stephensi* and *Anopheles culicifacies* (Diptera : Culicidae) in Gurgaon, India**. *Southeast Asian J Trop Med Public Health* 1995, **26**:172-176.
56. Hume JC, Tunnicliff M, Ranford-Cartwright LC, Day KP: **Susceptibility of *Anopheles gambiae* and *Anopheles stephensi* to tropical isolates of *Plasmodium falciparum***. *Malar J* 2007, **6**:139.
57. Miazgowicz KL, Mordecai E, Ryan SJ, Hall RJ, Owen J, Adanlawo T, Balaji K, Murdock CC: **[preprint] Mosquito species and age influence thermal performance of traits relevant to malaria transmission** biorxiv. org2019.
58. United Nations DoEaSA, Population Division: **World Urbanization Prospects: The 2018 Revision** New York: United Nations.; 2019.
59. Takken W, Lindsay S: **Increased Threat of Urban Malaria from *Anopheles stephensi* Mosquitoes, Africa**. *Emerg Infect Dis* 2019, **25**:1431-1433.

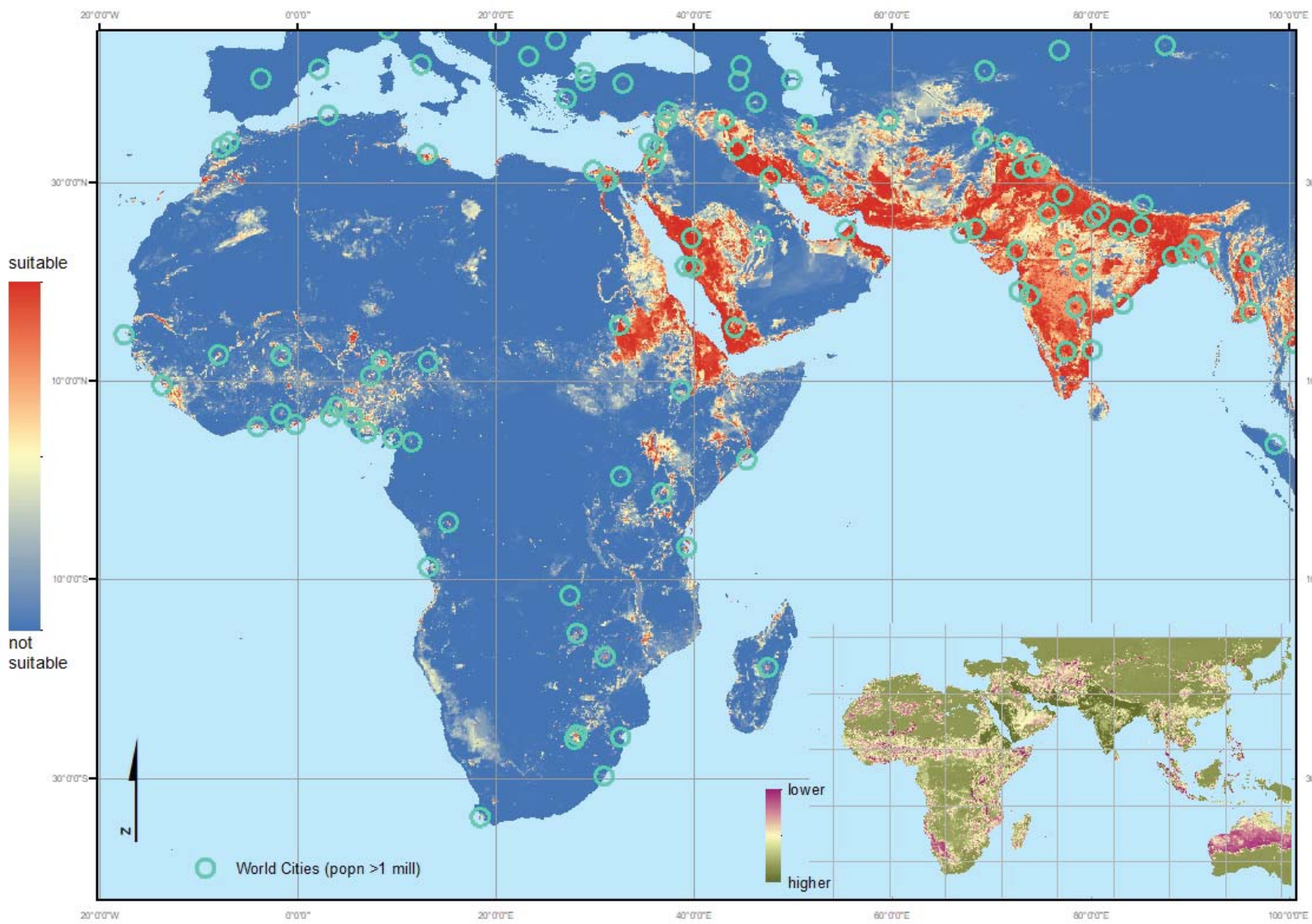
60. Parmakelis A, Russello MA, Caccone A, Marcondes CB, Costa J, Forattini OP, Sallum MA, Wilkerson RC, Powell JR: **Historical analysis of a near disaster: *Anopheles gambiae* in Brazil.** *Am J Trop Med Hyg* 2008, **78**:176-178.
61. Lopes G: ***Anopheles gambiae* in Brazil: the background to a "silent spread," 1930-1932.** *Hist Cienc Saude Manguinhos* 2019, **26**:823-839.
62. Killeen GF, Fillinger U, Kiche I, Gouagna LC, Knols BG: **Eradication of *Anopheles gambiae* from Brazil: lessons for malaria control in Africa?** *Lancet Infect Dis* 2002, **2**:618-627.
63. Qiuyin Q, Guerra CA, Moyes CL, Elyazar IR, Gething PW, Hay SI, Tatem AJ: **The effects of urbanization on global *Plasmodium vivax* malaria transmission.** *Malar J* 2012, **11**:403.
64. Reisen WK: **Landscape epidemiology of vector-borne diseases.** *Annu Rev Entomol* 2010, **55**:461-483.
65. **QGIS** [<https://qgis.org/en/site/>]

Table 1: Abridged table showing the populations at risk if *An. stephensi* were to establish in urban cities of Africa (full table in Supplemental file). The cities are listed in order of increasing distance from confirmed *An. stephensi* presences in Africa and many of the cities that are not shown in the abridged version are identified as highly suitable (Table ED3). The habitat suitability for *An. stephensi* classification was evaluated from the predicted habitat suitability map (Fig. 3). The quartiles for the full distribution of predicted values within the current species range were calculated and used to assign suitability class. Class 4 (below the lower quartile, indicating lowest suitability) is not found in any listed city (full table in Extended Data: table ED3). Class 3 and 2 indicate increasing increments of suitability, with class 1 indicating the highest suitability (i.e. predictions greater than the upper quartile value). Population data are taken from Brinkoff [42] listing urban agglomerations with ≥ 1 million inhabitants. The distance from *An. stephensi* records (data collated by this study) was calculated in QGIS [43] using UN sourced city coordinates which were also used to evaluate the distance from *P. falciparum* and *P. vivax* endemic zones [17]. The eleven cities closest to the current *An. stephensi* distribution are all within the malaria endemic zone. *Addis Ababa is within the area of the malaria endemic zone, but due to its altitude (~ 2300 m) is considered non-endemic.

City	Country	Population	Distance from <i>An. stephensi</i> records (km)	Distance from malaria endemic zone (km)	Habitat suitability class
Djibouti City	Djibouti	<1M	0	0	1
Addis Ababa	Ethiopia	3,725,000	160	0*	2
Asmara	Eritrea	<1M	450	0	1
Muqdisho (Mogadishu)	Somalia	1,890,000	480	0	1
Al-Khartum (Khartoum)	Sudan	6,150,000	940	0	1
Nairobi	Kenya	5,950,000	1100	0	1
Mombasa	Kenya	1,240,000	1190	0	1
Kampala	Uganda	3,400,000	1280	0	2
Dar es Salaam	Tanzania	6,150,000	1490	0	1
Kigali	Rwanda	1,140,000	1650	0	2
Kisangani	DRC	1,120,000	1900	0	3
Al-Qahirah (Cairo)	Egypt	20,500,000	2260	1490	1
Bangui	CAR	1,160,000	2440	0	2
Lilongwe	Malawi	1,020,000	2450	0	1
Mbuji-Mayi	DRC	2,000,000	2490	0	2







suitable



not
suitable



 World Cities (popn >1 mill)

