scientific data

OPEN

DATA DESCRIPTOR

Spatial variation of population, density, and composition of mosquitoes in mainland China

Zhenbo Wang[1,3], Xuechun Jia[2,3], Yinan Wang[1] ✉ & Chengdong Xu[1] ✉

Mosquitoes pose a significant threat to global health, impacting over 40% of the world's population. Currently, a dearth of large-scale and long-term data on mosquito populations in China hinders related research and public health endeavors. In response, we meticulously compiled and analyzed existing studies to construct a comprehensive dataset illustrating the spatial variation of mosquito populations, density, and composition across mainland China. This dataset furnishes invaluable information to support further research on mosquitoes and mosquito-borne diseases on broader spatial and temporal scales. The primary aim of this research is to contribute to efforts aimed at controlling mosquito transmission and mosquito-borne diseases, ultimately enhancing human well-being.

Background & Summary

Mosquitoes, classified under the family *Culicidae*, are widely acknowledged as substantial transmitters of diseases at a global level. The main transmitters of these pathogens are primarily from *Aedes*, *Culex*, and *Anopheles*. For instance, *Aedes* transmits dengue fever and Zika, *Culex* transmits Japanese encephalitis and West Nile fever, and *Anopheles* transmits malaria[1,2]. Furthermore, mosquitoes transmit many new viruses, including the Banna virus and Yunnan circovirus[3]. There are 7 genera and 55 species of mosquitoes in China that can transmit or carry viruses. For mosquito-borne viruses, 7 families, 8 genera, and 26 species have been discovered, as well as 13 genera, and 30 species of new viruses[4]. It is suggested that infected patients release acetophenone, a chemical compound that attracts mosquito bites[5]. Mosquitoes transmit pathogens into healthy bodies by biting infected humans or animals[6,7]. Over time, the virus transmitted by mosquitoes progresses from a localized infection to systemic transmission leading to symptoms such as fever, chills, headache, and fatigue, which in severe cases can result in death[3]. Mosquito-borne diseases (MBDs) pose a significant health challenge to over 40% of the world's population[2].

Many researchers have concluded that both environmental factors and human activities contribute to the transmission of MBDs in China[8–10]. These pivotal factors directly orchestrate alterations in mosquito density and distribution[11]. However, studies integrating mosquito density at large spatiotemporal scales remain scarce, potentially for several reasons. On one hand, Chinese studies or reports lack sufficient long-term and large-scale data on mosquito density. On the other hand, the abundant methods for obtaining mosquito density and indices for characterizing mosquito density add to the complexity of the study. According to "Surveillance methods for vector density-Mosquito" (GB/T 23797-2020), adult mosquito surveillance currently includes light trapping, tent trapping, labor hour method, and so on. Larval mosquito surveillance includes the Breteau index (BI), House index (HI), Container index (CI), and so on. Differences in mosquito trapping methods result in disparities in density indices, units, and so on[12,13]. Hence, complete records of mosquitos on a large scale in China were vacant. Recently, some researchers have concentrated on the spatial distribution and occurrence records of mosquitoes or MBDs in China. Wang *et al.*[14] and Atoni *et al.*[4] collected information on the occurrence locations and times of mosquito species and MBDs in mainland China. However, there were no detailed records of mosquito density on a monthly or annual basis, which limited a comprehensive understanding of the ecological habits of mosquitoes in mainland China.

[1]Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, 100101, P. R. China. [2]School of Ecological Technology and Engineering, Shanghai Institute of Technology, Shanghai, 201418, P. R. China. [3]These authors contributed equally: Zhenbo Wang, Xuechun Jia. ✉e-mail: wynneau@163.com; xucd@lreis.ac.cn

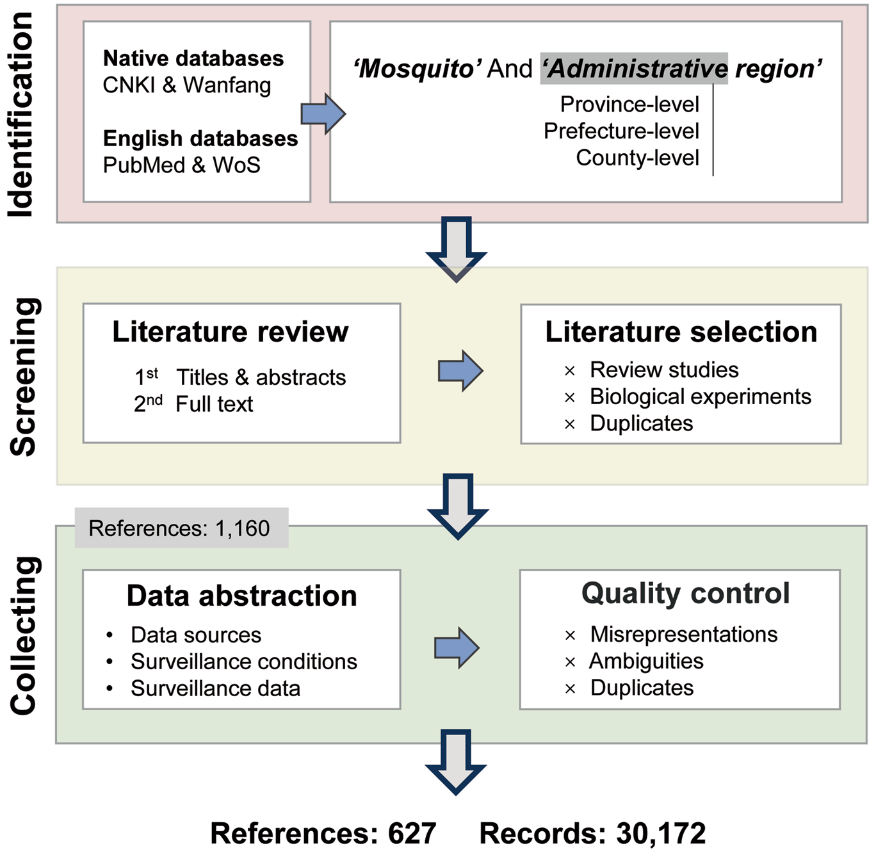

Fig. 1 Literature review flow chart of the search strategy and results.

The present study intended to establish a collection on mosquito density in China. These raw data allow researchers to obtain the density and distribution of various mosquito species, and analyze the causes of mosquito outbreaks. It can also be used to establish conversion formulas for different mosquito density monitoring methods, further addressing the lack of data in China. The collection provides support for modeling the spatial and temporal distribution of mosquitos and lays a scientific basis for developing mosquito control measures and achieving sustainable mosquito management. Furthermore, these data can be used for modeling and predicting identified or emerging MBDs emissions with improved accuracy, which is vital for the reduction of disease, protection of the environment, contribution to human health and well-being, and the achievement of sustainable development.

Methods

Data sources. Due to the lack of authoritative datasets for mosquito surveillance in China, data from previous studies can be reliable sources. Literature searches were conducted on Chinese databases including China Knowledge Network (CNKI, https://www.cnki.net/) and Wanfang Database (https://www.wanfangdata.com.cn/). English databases were used including PubMed (https://www.ncbi.nlm.nih.gov/) and Web of Science (WOS, https://www.webofscience.com/wos/). Literature searches used the term "Mosquito" in conjunction with the names of local administrative regions. To collect more data, each search was systematically conducted at the provincial, prefectural, and county levels. ensuring a comprehensive sweep across all administrative regions. This approach facilitated the accumulation of a robust dataset at various geographical granularities. In each literature search, the first screening process was based on the titles and abstracts of the studies. The selected studies did not include review studies, biological experiments, and repeated publications, as shown in Fig. 1. Thereafter, re-screening was performed based on the content to determine whether mosquito population or density was included. Relevant journal articles, theses, and conference proceedings were retrieved and included. As of August 2024, it has been recorded in all provinces in mainland China, excluding Macau.

Information extracted. Data sources of studies are recorded during the construction of the dataset, including titles, years, and DOIs. As some Chinese studies do not provide DOIs, they are supplemented with URLs. They can be used for tracking the monitoring process, verifying the data, and providing solid foundations for subsequent analysis and application. Due to the different surveillance conditions, detailed records were kept of the time, location, and method records of surveillance. Time records included the specific year and month of surveillance, ensuring the timeliness of the data and the accuracy of the seasonal analysis. Location data provides the exact geographic coordinates of the monitoring site, which is crucial for analyzing changes in mosquito

Information	Details
Data sources	Names of studies, DOI, Publication Date
Surveillance conditions	Administrative units, habitats, latitude, longitude, time of sampling, sampling methods
Surveillance data	Mosquito species, mosquito density, density units, population, number of mosquitoes, frequency of repetitions

Table 1. Basic information of data collection.

distribution and density. Records of surveillance methods address the specific techniques and tools used, which contribute to the comparability and reproducibility of the data. In addition, the sites for mosquito surveillance are differentiated into administrative districts and habitats. The average mosquito density from multiple habitats within a given level of administrative region is considered to be the density for that area. This approach strikes a balance between data representation and comprehensive surveillance. It more accurately reflects the overall distribution of mosquitoes across the entire region, avoiding biases that could arise from focusing on a single habitat. Additionally, by integrating data from multiple habitats, the method captures a broader range of ecological conditions and mosquito behavioral patterns, enhancing the depth and reliability of surveillance efforts. These elements in Table 1 make up the new dataset.

A team of 4 people with expertise in ecology, environment, and geography was recruited to collect these data. Upon agreeing to contribute to this project, they were (1) invited to join the project's online social network, (2) invited to attend an orientation meeting including the workflow, entry requirements, and use of the software, (3) provided with data entry templates, several randomly assigned studies and step-by-step instructions, (4) asked to report on the progress of their work and be reviewed, and (5) requested to provide prompt feedback on problems encountered and follow instructions from the person in charge. In this process, members need to accurately document data sources, surveillance conditions, and surveillance data based on the description of the original studies. To refine the information, coordinates of surveillance sites were added to the dataset. Some of these were sourced from the original study and the rest were sourced from Amap (https://www.amap.com/) according to prefecture and county. All members need to report on the reasons for the exclusion of studies, including misrepresentations, ambiguities, or duplicates. Records of the population, density, and composition of mosquitoes were presented in the study in the form of text, tables, and figures, as shown in Fig. 1. Among them, raw data collected from figures were used by GetData Graph Digitizer (http://www.getdata-graph-digitizer.com/). Mosquito trapping methods and measurement units are required to be consistent with data sources, and judgment are needed as to whether the monitoring site is a habitat or an administrative region. All data are summarized together in a monthly and annual format. If there were several monitoring attempts in a month, the average values were computed and the frequency of repetitions was documented. In addition, multiple monitoring methods in one study were conserved to ensure the integrity of the data. These efforts finished in August 2024. Ultimately, all assignments were aggregated into the dataset.

Data Records

The dataset is available at Figshare[15]. The "Literature Information.csv" file contains information regarding the 627 included studies to prove authenticity, including the title of the study, year of publication, and web links. The description of column headings of "Literature Information.csv" is shown in Table 2. The "Monitoring Information.csv" file contains 30,172 records of mosquito surveillance. Each record represents the density or number of a mosquito species in a specific location as described at a set time-point, including information on monitoring time, location, species, density, and others. The description of the column headings of "Monitoring Information.csv" is shown in Table 2. And the records in the "Monitoring Information.csv" file are linked with the "Literature Information.csv" file by the column "STUDIES_ID".

Overview of the data. Figure 2 provides information on the mosquito surveillance efforts conducted in mainland China. Figure 2a shows the times of mosquito surveillance each year. The earliest available surveillance data on mosquitoes in China appeared in 1955, while the next 20 years were almost undocumented. A significant shift occurred in the 1980s, peaking in 2018. Due to inconsistencies in spatial and temporal scales, the frequency of main species, genera, habitats, and trapping methods for mosquito surveillance are considered on an annual basis for each province. Figure 2b and Fig. 2c highlight the most frequently detected mosquito genera and species. The frequency was calculated by dividing the number of years each species or genus was present in each province by the total number of years, with duplications removed. The dominance of *Culex*, *Anopheles*, *Aedes*, and *Armigeres* genera is noteworthy. Particularly, *Anopheles sinensis*, *Culex tritaeniorhynchus*, *Aedes albopictus*, *Culex pipiens pallens*, and *Armigeres subalbatus* are the most common mosquito species. Figure 2d shows the most prevalent habitats where mosquito presence is detected in China. These habitats have been meticulously categorized. The frequency was the number of years in which each habitat occurred in each province divided by the total number of years, with duplications removed. It shows that residential areas, livestock sheds, green spaces, and hospitals are the most common habitats for mosquitoes. Figure 2e presents the most frequently employed surveillance methods for mosquitoes in China. The frequency was the number of years for each trapping method used in each province from previous studies divided by the total number of years, with duplications removed. For adult mosquitoes, the light trapping method emerges as the predominant technique, followed by the labor hour method and the tent trapping method. BI and CI are the main methods for larvae.

Figure 3 shows the provinces and locations where mosquito surveillance records predominantly appear. Coordinate points represent mosquito monitoring sites. The fill color represents the number of years in which

Information	Fields	Details
Studies	Studies_ID	Using an abbreviation of a province followed by a number for each study. The number represents the order in which the study was selected for that province. Each study has only one of these fields. If the study covers more than one province, the field starts with "MIX".
	Title_C	Titles of studies (in Chinese)
	Title_E	Titles of studies (in English)
	Studies_ PUB	Years of publication of studies
	Studies_DOI	DOI of studies
	Studies_URL	Links to some Chinese studies without DOIs
Monitoring information	Site_L1	Provinces where the site is located
	Site_L2	Municipalities where the site is located
	Site_L3	County districts where the site is located
	Site_long	Longitudes of the sites
	Site_lat	Latitudes of the sites
	Site_year	Sampling years
	Site_month	Sampling months
	Site_method	Sampling methods
	N	Repetition number
Mosquito species and population characterization	Species	Species of mosquito
	Den_admin	Density of mosquitos in administrative districts
	Num_admin	Numbers of mosquitos in administrative districts
	Habitat	Habitats of the site
	Den_hab	Density of mosquitos in Habitats
	Num_hab	Numbers of mosquitos in Habitats
	Den_unit	Units of density

Table 2. Field names.

a specific mosquito species was detected in various provinces. The deeper the color, the higher the frequency of surveillance. For mosquito records not specifying the exact species in Fig. 3a, provinces with more monitoring years or monitoring sites are mainly located in the southwest and eastern regions, while the western and northern regions are relatively fewer. Figure 3b-k are arranged in descending order of the frequency at which each mosquito species is monitored in China. It can be observed that *An. sinensis* and *Cx. tritaeniorhynchus* are distributed across China except in the northwest. *Ae. albopictus* and *Ar. subalbatus* are distributed across China but absent in the northwest and northeast parts. *Cx. pipiens pallens* is primarily distributed in the north, whereas *Culex pipiens quinquefasciatus* is mainly found in the south, with the central region being where the two subspecies coexist. *Aedes vexans* is primarily found in the north and southwest. *Anopheles minimus* and *Anopheles anthropophagus* predominantly appear in the south.

The dataset contains reliable mosquito surveillance records from 1955 to 2023, encompassing all of mainland China except for Macao. A total of 32 genera and 208 mosquito species have been identified. These records are conducive to analyzing the relationship between mosquitoes and the above factors in mainland China in a larger spatial and temporal pattern, thus realizing the prediction of mosquito density or distribution. Since some mosquito species transmit specific diseases, accurate prediction can serve as an early warning of certain diseases. This means that mosquito control strategies need to be adapted to geographical locations. It helps public health departments to identify priority areas for surveillance and control and to optimize the allocation of resources as a means of improving control efficiency.

Technical Validation

Herein, this dataset contains 30,172 records that were extracted from 627 studies. Literature searches lasted until August 2024. During the literature search, when an administrative area is searched and screened, basic information about the study is recorded, including title, province, and DOI. Another colleague would then be assigned to repeat the task to avoid the omission of available studies. In the process of extracting the data, each member was assigned to all studies in at least one province, so that cases where the same monitoring work was published in different studies could be easily identified and eliminated. Thereafter, one team member compiled and cross-checked the data. The data included are mosquito density on an annual or monthly basis, with administrative regions covering provinces, prefectures, and counties. Average values of mosquito density covering multiple years were not included to facilitate integration with other indicators.

Due to the difference in the description of mosquito trapping methods between studies, the description is amended according to "Surveillance methods for vector density-Mosquito" (GB/T 23797-2020). Where multiple mosquito trapping methods are utilized or not explicitly described, they are annotated as "Combination". Mosquito trapping methods not included in this standard have been retained to the extent possible. They may include obsolete, newly invented, and reasonably improved methods. This approach reflects support for innovation and historical experience, respecting the diversity of methods.

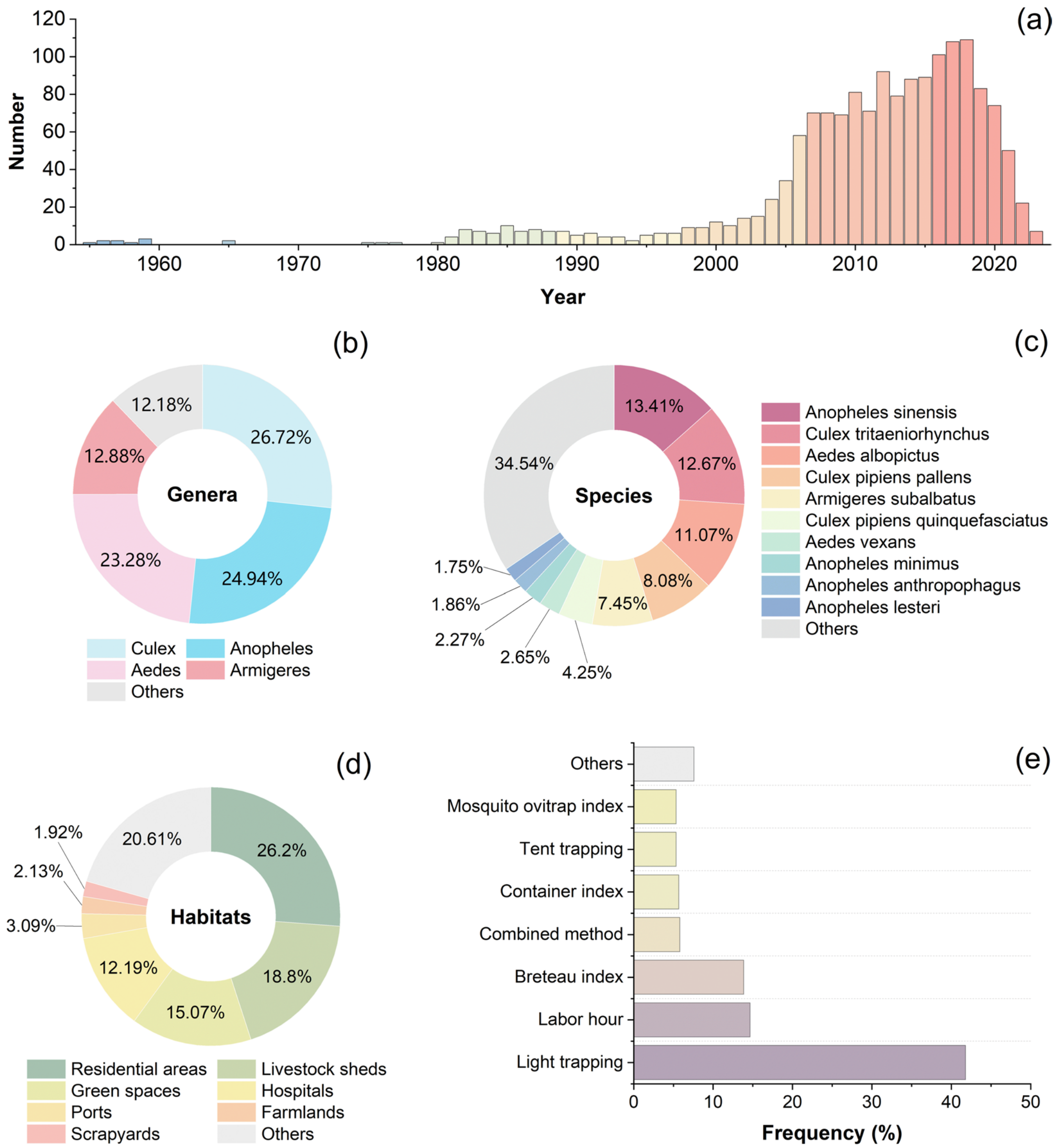

Fig. 2 Information on mosquito surveillance in mainland China. (**a**) Number of mosquito surveillance conducted annually, (**b**) Frequency of mosquito genera monitored, (**c**) Frequency of mosquito species monitored, (**d**) Frequency of habitats where mosquitoes are more prevalent, (**e**) Frequency of use of mosquito trapping methods.

Data visualization can verify all locations of mosquito species were appropriately geo-referenced. It also allows for comparison with other relevant datasets, thereby confirming the authenticity and superiority of the new dataset. Mosquitoes in mainland China are more frequently recorded in the southwest and east of the country, which is consistent with Atoni *et al.*[4] and Wang *et al.*[14]. In contrast to previous studies, this dataset provides the density or number of mosquito species. It is not restricted to mosquitoes with a high association with the virus and offers detailed geographic information and time scales. In addition, the literature search for this work was more detailed and supplemented with more recent research. It fills in the gaps without monitoring data and updates the geographic distribution of mosquitoes in mainland China. It also provides detailed records of mosquito density over time and is an empirical resource for mosquito ecology.

Usage Notes

The database was retrieved up to August 2024. The present dataset contains the monthly and annual density of mosquitos across mainland China, which can be used for mosquito prevention and control efforts. In particular, the years of surveillance in each place may not be consistent, and the standards may be biased over time. However, it still compensates for the absence of mosquito density datasets in mainland China. It is recommended to combine mosquito density with meteorological, land use, and socioeconomic data. It allows for the

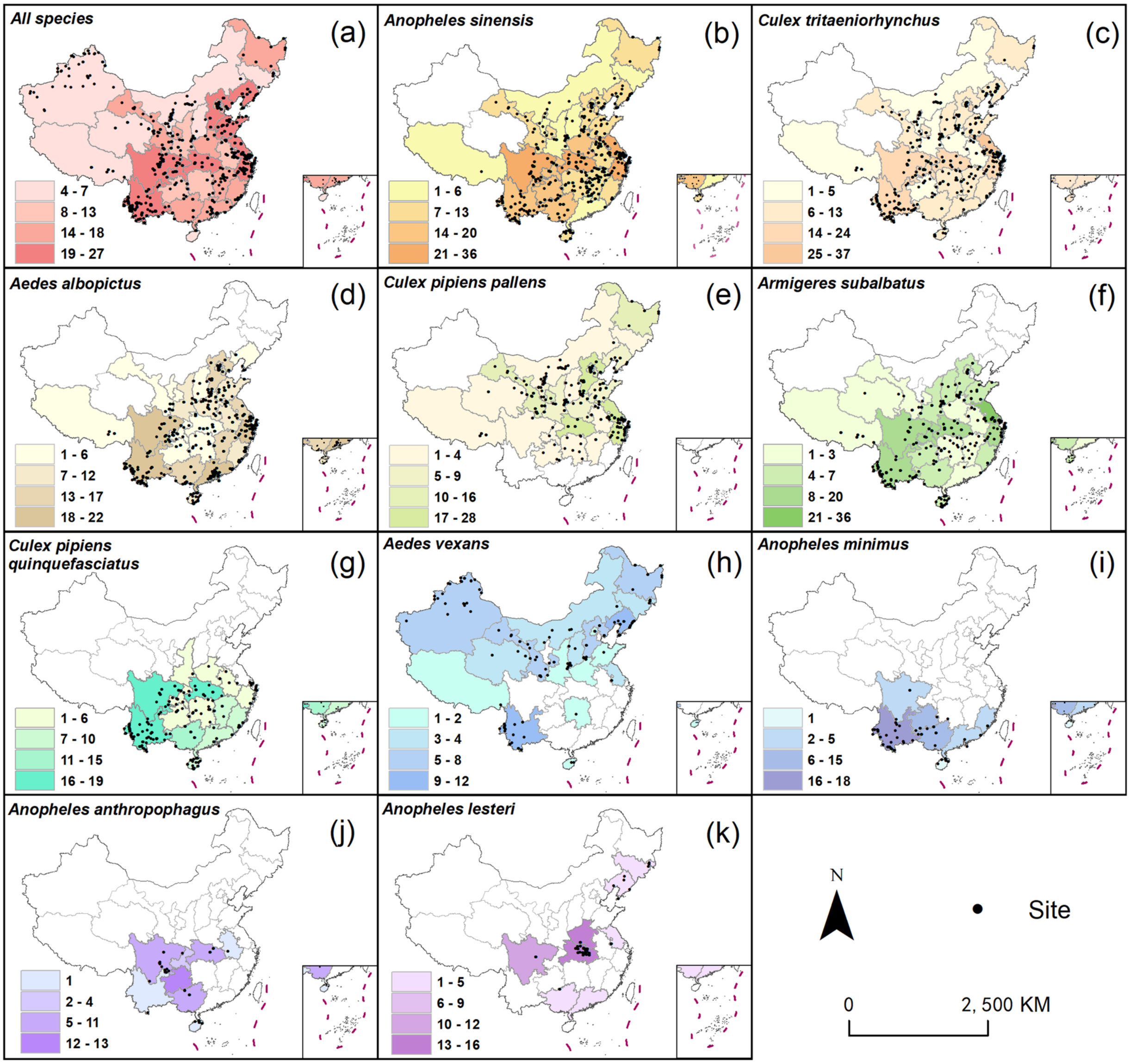

Fig. 3 The number of years of mosquito species monitored in each province and the distribution of sites.

simulation of mosquito density and distribution in larger spatial and temporal patterns within China. In addition, the standardization of units and the conversion between indicators may contribute to better data support.

Code availability

No computer code was used in the data curation process.

Received: 18 April 2024; Accepted: 2 January 2025;
Published online: 07 January 2025

References

1. Franklinos, L. H. V., Jones, K. E., Redding, D. W. & Abubakar, I. The effect of global change on mosquito-borne disease. *Lancet Infect Dis* **19**, 302–312, https://doi.org/10.1016/s1473-3099(19)30161-6 (2019).
2. Atoni, E. *et al*. The discovery and global distribution of novel mosquito-associated viruses in the last decade (2007–2017). *Rev. Med. Virol.* **29**, 21, https://doi.org/10.1002/rmv.2079 (2019).
3. Xia, H., Wang, Y., Atoni, E., Zhang, B. & Yuan, Z. Mosquito-associated viruses in China. *Virol Sin* **33**, 5–20, https://doi.org/10.1007/s12250-018-0002-9 (2018).
4. Atoni, E. *et al*. A dataset of distribution and diversity of mosquito-associated viruses and their mosquito vectors in China. *Sci. Data* **7**, 342, https://doi.org/10.1038/s41597-020-00687-9 (2020).
5. Zhang, H. *et al*. A volatile from the skin microbiota of flavivirus-infected hosts promotes mosquito attractiveness. *Cell* **185**, 2510–2522.e2516, https://doi.org/10.1016/j.cell.2022.05.016 (2022).
6. Wu, P., Yu, X., Wang, P. H. & Cheng, G. Arbovirus lifecycle in mosquito: acquisition, propagation and transmission. *Expert Rev. Mol. Med.* **21**, 6, https://doi.org/10.1017/erm.2018.6 (2019).
7. Wang, Z. Y., Nie, K. X., Niu, J. C. & Cheng, G. Research progress toward the influence of mosquito salivary proteins on the transmission of mosquito-borne viruses. *Insect Sci*., 11 https://doi.org/10.1111/1744-7917.13193 (2023).

8. Quan, Y. *et al*. How do temperature and precipitation drive dengue transmission in nine cities, in Guangdong Province, China: a Bayesian spatio-temporal model analysis. *Air Quality, Atmosphere & Health*, 1–11, https://doi.org/10.1007/s11869-023-01331-2 (2023).
9. Li, F. *et al*. The spatial-temporal pattern of Japanese encephalitis and its influencing factors in Guangxi, China. *Infection, Genetics and Evolution* **111**, 105433, https://doi.org/10.1016/j.meegid.2023.105433 (2023).
10. Huang, F. *et al*. A retrospective analysis of malaria epidemiological characteristics in Yingjiang County on the China–Myanmar border. *Sci. Rep.* **11**, 14129, https://doi.org/10.1038/s41598-021-93734-3 (2021).
11. Liu, Q. Sustainable Pest Management for Health and Well-Being. *China CDC Weekly* **2**, 438–442, https://doi.org/10.46234/ccdcw2020.112 (2020).
12. Hou, J. *et al*. Field evaluation of two mosquito traps in Zhejiang Province, China. *Sci. Rep.* **11**, 294, https://doi.org/10.1038/s41598-020-80618-1 (2021).
13. Wu, Y. Y. *et al*. Effect of different carbon dioxide (CO_2) flows on trapping Aedes albopictus with BG traps in the field in Zhejiang Province, China. *PLoS One* **15**, 11, https://doi.org/10.1371/journal.pone.0243061 (2020).
14. Wang, T. *et al*. Mapping the distributions of mosquitoes and mosquito-borne arboviruses in China. *Viruses-Basel* **14**, 24, https://doi.org/10.3390/v14040691 (2022).
15. Wang, Z., Jia, X., Wang, Y. & Xu, C. Spatial variation of population, density, and composition of mosquitoes in mainland China. *Figshare* https://doi.org/10.6084/m9.figshare.27266409 (2024).

Acknowledgements
This work was financially sponsored by the National Natural Science Foundation of China (42130713) and The Second Tibetan Plateau Scientific Expedition (2019QZKK1005).

Author contributions
Zhenbo Wang: Project administration, Funding acquisition, and Writing- review & editing. Xuechun Jia: Data Curation, Formal Analysis, and Writing-Original Draft. Yinan Wang: Methodology, Visualization, and Supervision. Chengdong Xu: Conceptualization and Investigation.

Competing interests
The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Additional information
Correspondence and requests for materials should be addressed to Y.W. or C.X.