UAB
Universitat Autònoma de Barcelona

Novel approaches to improve mosquito surveillance

María Isabel González Pérez

UAB
Universitat Autònoma
de Barcelona

Novel approaches to improve mosquito surveillance

Tesis doctoral presentada por **María Isabel González Pérez** para acceder al Grado de Doctora en el marco del programa de Doctorado en *Medicina i Sanitat Animal* de la *Facultat de Veterinària* de la *Universitat Autònoma de Barcelona*, bajo la dirección de la **Dra Núria Busquets Martí**, la **Dra Sandra Talavera Forcades**, el **Dr Carles Aranda Pallero** y bajo la tutoría del **Dr Francesc Accensi i Alemany**.

Bellaterra, 2023

IRTA CReSA
Centre de Recerca
en Sanitat Animal

European
Commission

La Dra **Núria Busquets Martí**, la Dra **Sandra Talavera Forcades** y el Dr **Carles Aranda Pallero**, investigadores del *Institut de Recerca i Tecnologia Agroalimentàries* (IRTA)- *Centre de Recerca en Sanitat Animal* (CReSA), y el Dr **Francesc Accensi i Alemany**, profesor del *Departament de Sanitat i Anatomia Animals* de la *Facultat de Veterinària* de la *Universitat Autònoma de Barcelona*,

CERTIFICAN:

Que los trabajos desarrollados en la memoria de la tesis doctoral "Novel approaches to improve mosquito surveillance" presentados por la doctoranda **María Isabel González Pérez** para la obtención del Grado de Doctora en Medicina y Sanidad Animal, se han realizado bajo su dirección y tutoría. Por consiguiente, autorizan su presentación a fin de ser evaluados por la comisión correspondiente.

Para que así conste y a efectos oportunos, firman el presente certificado.

Directora	**Directora**	**Director**
Núria Busquets Martí	**Sandra Talavera Forcades**	**Carles Aranda Pallero**
Tutor	**Doctoranda**	
Francesc Accensi i Alemany	**María Isabel González Pérez**	

Bellaterra (Barcelona), 26 de octubre de 2023.

La presente tesis fue desarrollada en el marco del proyecto europeo VECTRACK: *Earth observation service for preventive control of insect disease vectors* (Grant agreement ID: 853758) financiado por el programa Horizon 2020 de la Comisión Europea.

La impresión de la tesis fue financiada por el Programa de Doctorado en *Medicina i Sanitat Animal* de la *Facultat de Veterinària* de la *Universitat Autònoma de Barcelona*.

Seny i Rauxa

Table of contents

1. ABSTRACT

1.1 English

Mosquitoes (Diptera: Culicidae) act as vectors of several pathogens that cause serious diseases such as malaria, dengue, yellow fever, Zika, chikungunya, or West Nile fever, accounting for hundreds of thousands of human deaths every year. Mosquito surveillance allows to gather relevant information about the vector (presence, distribution, status, abundance, seasonal activity, longevity, pathogen circulation, etc) which is crucial to stablish intervention thresholds for vector control actions and to evaluate their efficacy. Traditional surveillance methods are very labour-intensive and require expert entomologists as well as a minimum of laboratory infrastructure. Besides, they do not allow a fine time-scale monitoring of mosquito populations.

Along the present thesis, an optical sensor prototype coupled to the entrance of a mosquito trap, was trained with machine learning techniques on a large database of mosquito flights from the vector species *Aedes albopictus* and *Culex pipiens,* at different raising and environmental conditions, for different classification tasks: i) the identification of mosquito genus and sex in laboratory conditions; ii) the identification of mosquitoes over other insects and the classification of the genus and sex of those mosquitoes in the field; iii) the estimation of mosquito's age; and iv) the identification of arbovirus-infected mosquitoes. The objective was to develop a ready-to-use reliable tool for the automated remote monitoring of *Aedes* and *Culex* mosquito populations, thus overcoming some of the current limitations of traditional surveillance methods and providing accurate relevant entomological information with high time resolution.

The developed classification system was highly accurate to classify mosquito genus and sex in laboratory conditions, reporting 94.2% of accuracy for genus and 99.4% for sex of *Aedes;* and 100% for sex of *Culex.* In the field, the system was able to distinguish mosquito species from *Aedes* and *Culex* genus over other insects that entered the trap with 95.5% of balanced accuracy and a low rate of false positives; and to classify the genus and sex of those mosquitoes with 88.8% of balanced accuracy. Age classification of female *Cx. pipiens* mosquitoes was assessed in the laboratory showing 74.7% of accuracy to distinguish between young mosquitoes (2-4 days of age) and older mosquitoes (7-16 days of age). In this case, also a molecular technique based on gene profiling was assessed for age-grading in *Cx. pipiens* showing a mean absolute error of 3.8 days in age prediction. For the arbovirus infection assay, the classification system was trained on differentiating Zika-infected from non-infected *Ae.albopictus*, showing 63.0% of classification accuracy.

The accuracy of the developed system was dependent on the classification task pursued, so as in the number of samples used for training the model, and the combination of extracted flight features and machine learning algorithms used. Some classification tasks, such as genus and sex classification, obtained high accuracy results in both laboratory and field conditions, which was a proof of concept of the system’s high-performance to monitor *Aedes* and *Culex* mosquito populations in the field. Other tasks such as the classification of mosquitoes by age or infection status, which were never assessed before by the means of optical sensors, would still require more training samples to improve their classification accuracy, but promising results were obtained. Overall, the present thesis has contributed to improve the current state of art of mosquito surveillance and has paved the way for future vector biology research.

1.2 Spanish

Los mosquitos (Diptera: Culicidae) actúan como vectores de varios patógenos que causan graves enfermedades como la malaria, el dengue, la fiebre amarilla, el Zika, el chikungunya o la fiebre del Nilo Occidental, y son responsables de cientos de miles de muertes humanas cada año. La vigilancia de mosquitos permite recopilar información relevante sobre el vector (presencia, distribución, estado, abundancia, actividad estacional, longevidad, circulación de patógenos, etc.) que es crucial para establecer umbrales de intervención para su control y para evaluar la eficacia de dichas intervenciones. Los métodos tradicionales de vigilancia son muy laboriosos y requieren entomólogos expertos, así como un mínimo de infraestructura de laboratorio. Además, no permiten un seguimiento temporal preciso de las poblaciones de mosquitos.

A lo largo de la presente tesis, un prototipo de sensor óptico acoplado a la entrada de una trampa de mosquitos, fue entrenado con técnicas de aprendizaje automático sobre una gran base de datos de vuelos de mosquitos de las especies vectoras *Aedes albopictus* y *Culex pipiens*, en diferentes condiciones de cría y ambientales, para diferentes tareas de clasificación: i) la identificación del género y sexo del mosquito en condiciones de laboratorio; ii) la identificación de mosquitos sobre otros insectos y la clasificación del género y sexo de esos mosquitos en el campo; iii) la estimación de la edad del mosquito; y iv) la identificación de mosquitos infectados por arbovirus. El objetivo era desarrollar una herramienta fiable y lista para usar para el seguimiento automatizado y remoto de poblaciones de mosquitos *Aedes* y *Culex*, superando así algunas de las limitaciones actuales de los métodos tradicionales de vigilancia y proporcionando información entomológica relevante y precisa con una alta resolución temporal.

El sistema de clasificación desarrollado fue altamente preciso para clasificar el género y el sexo de los mosquitos en condiciones de laboratorio, con un 94,2% de exactitud para el género y un 99,4% para el sexo de *Aedes*; y un 100% para el sexo de *Culex*. En el campo, el sistema fue capaz de distinguir los mosquitos *Aedes* y *Culex* de otros insectos que entraron en la trampa con un 95,5% de precisión y una baja tasa de falsos positivos; y de clasificar el género y el sexo de esos mosquitos con un 88,8% de precisión. La clasificación por edad de las hembras de *Cx. pipiens* fue evaluada en el laboratorio mostrando un 74,7% de precisión para distinguir entre mosquitos jóvenes (de 2 a 4 días de edad) y mosquitos mayores (de 7 a 16 días de edad). En este caso, también se evaluó una técnica molecular basada en perfiles genéticos para la clasificación por edades en *Cx. pipiens*, mostrando un error absoluto medio de 3,8 días en la predicción de la edad. Para el ensayo de infección por arbovirus, el sistema de clasificación se entrenó para diferenciar los *Ae.albopictus* infectados por Zika de los no infectados, mostrando un 63,0% de precisión en la clasificación.

El desempeño del sistema dependió de la tarea de clasificación asignada, así como del número de muestras utilizadas para entrenar el modelo y de la combinación de variables extraídas del vuelo y algoritmos de clasificación utilizados. Algunas tareas, como la clasificación de género y sexo, obtuvieron muy buenos resultados de clasificación tanto en condiciones de laboratorio como de campo, lo que constituyó una prueba de concepto del alto rendimiento del sistema para la vigilancia de poblaciones de *Aedes* y *Culex* en campo. Otras tareas, como la clasificación de mosquitos por edad o estado de infección, que nunca se habían evaluado mediante sensores ópticos, aún requerirían más muestras de entrenamiento para mejorar su precisión, pero los resultados obtenidos fueron prometedores. En conjunto, la presente tesis ha contribuido a mejorar el conocimiento sobre el estado actual de la vigilancia de mosquitos y ha allanado el camino para futuras investigaciones en biología de vectores.

1.3 Catalan

Els mosquits (Diptera: Culicidae) actuen com a vectors de diversos patògens que causen greus malalties com la malària, el dengue, la febre groga, el Zika, el chikungunya o la febre del Nil Occidental, i són responsables de centenars de milers de morts humanes cada any. La vigilància de mosquits permet recopilar informació rellevant sobre el vector (presència, distribució, estat, abundància, activitat estacional, longevitat, circulació de patògens, etc.) que és crucial per a establir llindars d'intervenció per al seu control i per a avaluar l'eficàcia d'aquestes intervencions. Els mètodes tradicionals de vigilància són molt laboriosos i requereixen entomòlegs experts, així com un mínim d'infraestructura de laboratori. A més, no permeten un seguiment temporal precís de les poblacions de mosquits.

Al llarg de la present tesi, un prototip de sensor òptic acoblat a l'entrada d'un parany de mosquits estàndard, va ser entrenat amb tècniques d'aprenentatge automàtic sobre una gran base de dades de vols de mosquits de les espècies vectores *Aedes albopictus* i *Culex pipiens*, en diferents condicions de cria i ambientals, per a diferents tasques de classificació: i) la identificació del gènere i sexe del mosquit en condicions de laboratori; ii) la identificació de mosquits sobre altres insectes, i la classificació del gènere i sexe d'aquests mosquits en el camp; iii) l'estimació de l'edat del mosquit; i iv) la identificació de mosquits infectats per arbovirus. L'objectiu era desenvolupar una eina fiable per al seguiment automatitzat i remot de poblacions de mosquits *Aedes* i *Culex*, superant així algunes de les limitacions actuals dels mètodes tradicionals de vigilància i proporcionant informació entomològica rellevant i precisa amb una elevada resolució temporal.

El sistema de classificació desenvolupat va mostrar una elevada precisió per a classificar el gènere i el sexe dels mosquits en condicions de laboratori, amb un 94,2% d'exactitud per al gènere i un 99,4% per al sexe de *Aedes*; i un 100% per al sexe de *Culex*. En el camp, el sistema va ser capaç de distingir els mosquits *Aedes* i *Culex* d'altres insectes que van entrar en el parany amb un 95,5% de precisió i una baixa taxa de falsos positius; i de classificar el gènere i el sexe d'aquests mosquits amb un 88,8% de precisió. La classificació per edat de les femelles de *Cx. pipiens* va ser avaluada en el laboratori, mostrant un 74,7% de precisió per a distingir entre mosquits joves (de 2 a 4 dies d'edat) i mosquits més vells (de 7 a 16 dies d'edat). En aquest cas, també es va avaluar una tècnica molecular basada en perfils genètics per a la classificació per edats en *Cx. pipiens*, mostrant un error absolut mitjà de 3,8 dies en la predicció de l'edat. Per a l'assaig d'infecció per arbovirus, el sistema de classificació es va entrenar per a diferenciar els *Ae.albopictus* infectats per Zika dels no infectats, mostrant un 63,0% de precisió en la classificació.

La precisió de classificació del sistema va dependre de la tasca de classificació assignada, així com del nombre de mostres utilitzades per a entrenar el model i de la combinació de variables extretes del vol i algoritmes de classificació utilitzats. Algunes tasques, com la classificació de gènere i sexe, van obtenir molt bons resultats de classificació tant en condicions de laboratori com de camp, la qual cosa va constituir una prova de concepte de l'alt rendiment del sistema per a la vigilancia de poblacions de *Aedes* i *Culex* en camp. Altres tasques, com la classificació de mosquits per edat o estat d'infecció, que mai s'havien avaluat mitjançant sensors òptics, encara requeririen més mostres d'entrenament per a millorar la seva precisió, però els resultats obtinguts van ser prometedors. En conjunt, la present tesi ha contribuït a millorar el coneixement sobre l'estat actual de la vigilància de mosquits i ha aplanat el camí per a futures recerques en biologia de vectors.

2. GENERAL INTODUCTION

2.1 The biology of mosquitoes

Mosquitoes (Diptera: Culicidae) are a family of hematophagous insects formed by 41 different genera and more than 3,500 recognised species. This family falls within the suborder Nematocera and is composed by two big subfamilies, Anophelinae and Culicinae, which differ in their morphological characteristics (Wilkerson et al., 2021). Mosquitoes can be found anywhere there is stagnant water, which is required for the development of their aquatic immature stages. They can be found on every continent except Antarctica, from below sea level to above tree line at heights of 3,000 meters or more. They feed primarily on plant nectar and honeydew, which provide them sugar and other non-carbohydrate nutrients for their survival and sexual maturation. However, blood-feeding is, with very few exceptions, an obligated condition for females to obtain the necessary proteins to complete their gonotrophic cycle (Foster & Walker, 2019). Hence, mosquitoes have an unequivocal ecological role as plant pollinators (Peach & Gries, 2020) but also as vectors of pathogens to vertebrate animal and human hosts.

The **holometabolous life cycle** of mosquitoes is composed by four life stages (egg, larva, pupa and adult), which inhabit different environments (aquatic for larvae and pupae and terrestrial for adults) (**Fig 1**). Immature stages develop in a wide range of aquatic habitats, from temporary to permanent surface water bodies (e.g., tidal pools in salt marshes, rain pools, rock pools, flood water, streams, swamps and lakes) and diverse natural and artificial water-holding containers (e.g., tree holes, leaf axils, plant pots, fountain, abandoned pools, tin cans, discarded tires or sewers). Adults are much more mobile than immatures, but they tend to occupy certain resting, foraging and, in some cases, overwintering habitats (Foster & Walker, 2019).

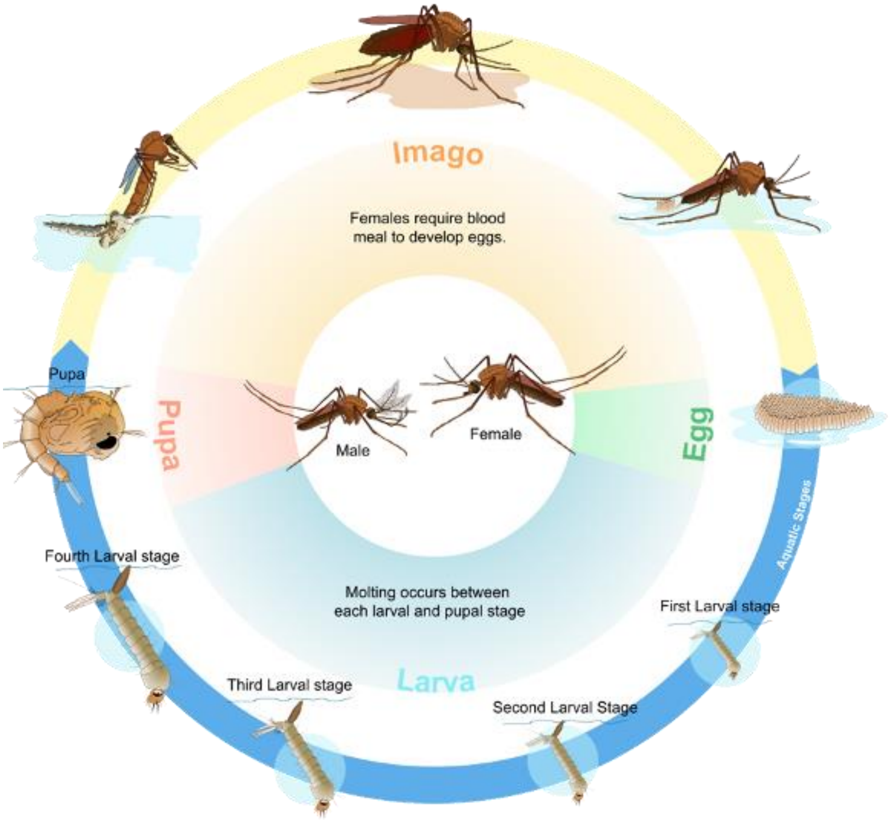


Fig 1. Mosquito life cycle. Extracted from (Iowa State University, n.d.)

Oviposition is made either directly on water or on solid substrates that are likely to be inundated. **Eggs** can be laid individually in the water surface such as the subfamily Anophelinae or clumped in a boat-like rafts such as several genera from Culicinae subfamily (e.g., *Culex*, *Culiseta* or *Coquillettidia*, among others). The tribe Aedini (Culicinae), on the contrary, attach their eggs individually on the substrate. Once the embryonic development of eggs has been completed (2-7 days after oviposition depending on temperature), larvae hatch soon after in species that lay their eggs in water. In case of mosquitoes from *Aedes* genus that oviposit on solid substrate, the eggs remain quiescent until the substrate become inundated. Abiotic signals such as temperature, air moisture or day length also affect hatching inhibition in some *Aedes* such as in the floodwater mosquitoes. When the oxygen levels decrease in the breeding site (signal of increase in the content of organic matter that will nourish larvae), larvae initiate the shell rupture by pressing the so-called egg tooth and emerge.

Larvae pass through four instars which resemble to each other morphologically but differ in size. All four larval stages of culicine and aedine mosquitoes are easy to recognize because of the presence of an elongate air tube used for breathing. However, anopheline larvae lack this are tube and lie parallel to, and just beneath, the surface of the water for breathing. Other genera (e.g., *Coquilletidia* and *Mansonia*) remain submerged in the water column since they possess a siphon which is modified for piercing submerged parts of aquatic plants to obtain oxygen from the aerenchyma. The larval food consists in small particles such as microorganisms, protozoa, algae, invertebrates, and detritus that larvae obtain from filtering, gathering, scraping, shredding, and preying. A special case are mosquitoes of *Toxorhinchites* genus which predate other mosquito larvae. Larvae moult 4 times before reaching the pupal stage, being the duration of this process temperature dependent (Becker et al., 2020).

Pupae are also aquatic and normally remain in the water surface floating motionless unless they are disturbed. They obtain oxygen from the outside through their air trumpets present in the cephalothorax, which are connected to the forming adult mesothoracic spiracles. Some genera like *Coquilletidia* or *Mansonia*, these air trumpets are modified to penetrate plant tissues and pupa remain submerged in the waterbody. Pupal stage usually last 2 days (depending on temperature), during which the metamorphosis of the adults takes place. At the end of this process, the internal pressure of the pupa increases because of forced gas penetration which induces the cephalothoracic cuticle of the pupa to split. Then, adult emerges leaving behind the exuvia.

Adults are small flying midge-like insects with slender bodies, thin legs, and narrow elongate wings. Their body surface is covered with scales, setae, and fine pile, creating the characteristic marks and colours typical of each specie. They have a prominent proboscis projecting anteriorly from the head adapted for sucking. Their mouthparts are sexually dimorphic responding to different feeding strategies (hematophagy in females and phytophagy in males). Females have strong needle-like mouthparts with teeth-bearing maxillae for piercing the host skin; while the maxillae and mandibles of males are much shorter than the proboscis and are considered functionless (Wahid et al., 2002). Sexual dimorphism also applies to the antennae, i.e., long filamentous sensorial organs between the eyes involved in sound recognition during flight, which have longer fibrillae in males giving a more plumose appearance. Also, females have generally a bigger size than males (Lorenz et al., 2017).

Adult mosquitoes remain most part of their time in resting sites such as understory vegetation, tree cavities, rock crevices, and other natural and artificial shelters (Sauer et al., 2021). Mosquitoes stay active when performing tasks of dispersal, foraging, mating, host-seeking, feeding, and ovipositing. Each mosquito specie has a characteristic **daily activity pattern,** which is dictated by the endogenous circadian rhythm entrained by the light-dark photoperiod, being characterised as typically diurnal (e.g., *Aedes*), nocturnal (e.g., *Anopheles*) or crepuscular (e.g., *Culex*). Also, **seasonal activity** patterns are present in temperate areas where overwintering mosquitoes may enter a diapause during unfavourable environmental conditions 0(Mitchell, 2020).

Dispersal of adult mosquitoes varies depending on the specie. Some of them, stay close (a few tens of metres) from their larval habitats to mate and feed, which is the case of domestic species; while others, with specific wind-assisted or light-directed dispersal modes, can move several kilometres from their origin sites (e.g., floodwater mosquitoes) (Verdonschot & Besse-Lototskaya, 2014). This flight behaviour is influenced by temperature, humidity, illumination levels, wind velocity and the physiological stage (Becker et al., 2020).

Mating takes place few days after adult emergence, typically in swarms (eurygamy) but also without (stenogamy). Sexual recognition and successful copula are mediated by the existence of a harmonic convergence between fundamental flight tone of both males and females, which is mutually recognised by the acoustic properties of the sensorial organs in their antennae (Gibson et al., 2010). Females are inseminated once in a lifetime whereas males can inseminate several females. From this single mating event, females store sufficient sperm to fertilize multiple clutches of eggs during their entire life (Cramer et al., 2023). After the copula and usually after sugar-feeding, females seek for a host to blood-feed, which is normally an obligated condition for egg maturation. Several factors such as host specificity, host preference, host availability and host cues (visual, olfactory, and thermal) operate in the success of finding a suitable host (Yan et al., 2021).

Blood-feeding behaviour is enabled by the female elongated mosquito mouthparts adapted for piercing and sucking, which evolved from ancient phytophagy or entomophagy of mosquito ancestors (Peach & Gries, 2020). When females land on a host and find a capillary, the pointed mandibles are used to rupture the skin and allow the rest of the stylets in the proboscis (the food and salivary channels) to penetrate until the blood vessel. To prevent blood coagulation while ingesting, the mosquito injects saliva into the wound which contains anticoagulants among many other compounds. This saliva usually stimulates an immune response in the host which can cause an inflammatory reaction together with irritation, thus leading to the scratching and possible bacterial infection of the wound. Parallelly, if the mosquito was infected with a pathogen from a previous blood meal, it can transmit the infectious agent via saliva while biting to the next host. Mosquitoes that have multiple blood meals, such as some anthropophilic vector from *Aedes* genus, are responsible of an increased risk of pathogen transmission (Scott & Takken, 2012).

2.2 Mosquitoes from a One Health perspective.

Mosquitoes can act as vectors of several pathogens (virus and parasites) of public and animal health importance, causing diseases such as dengue, yellow fever, Zika, chikungunya, West Nile, Rift Valley fever, Japanese encephalitis, malaria or lymphatic filariasis. **Vector-borne diseases** (VBD) account for ~17% of all infectious diseases and cause ~700,000 human deaths every year (WHO, 2020), as well as significant losses in farm livestock and wild animals. From a One Health perspective, VBD, and specifically **mosquito-borne diseases** (MBD), represent a major concern for medical and veterinary health authorities. There is a global burden of MBD emergence and re-emergence that mostly impacts low-income tropical countries (WHO, 2014). In the last decades, there has been an extension of the MBD outbreak range to temperate regions due to socioeconomic and environmental changes that affect the emergence and spread of both diseases and vectors (Semenza, 2013).

From over 3,500 existing mosquito species belonging to 41 different genera, only three of them are involved in MBD transmission: *Anophele*s, *Culex*, and *Aedes* (Foster & Walker, 2019). Among these genera, the most important vectors belong to the following groups of species:

i) ***Anopheles gambiae* complex**. *An. gambiae* s.l. is an African complex formed by eight species, two of the most important ones are *An. gambiae* s.s. and *Anopheles arabiensis*, the main widespread vectors of malaria and lymphatic filariasis in Africa. Both show an anthropophilic, endophilic and endophagic behaviour, which is more pronounced in *An. gambiae* s.s, making it the primary vector of malaria in the continent, a disease that cause over ~400,000 deaths/year, most of them children under 5 years old (WHO, 2023).

ii) ***Culex pipiens* complex**. *Cx. pipiens* s.l is a ubiquitous group with a worldwide distribution, which contains two domestic and peridomestic species: the northern house mosquito *Cx. pipiens* s.s., distributed in temperate areas; and the southern house mosquito *Culex quinquefasciatus,* with a tropical and subtropical distribution. They commonly hybridize at latitudes where their ranges overlap, as in North America, Argentina, Madagascar, Japan and South Korea (Farajollahi et al., 2011). They are vectors of human and animal pathogens such as West Nile virus (WNV), Usutu virus (USUV), Rift Valley fever phlebovirus (RVFV), St. Louis encephalitis virus (SLEV), lymphatic filariasis or avian malaria parasite. Their opportunistic feeding behaviour on a variety of vertebrates hosts makes them play a key role as bridge vectors in the transmission cycle dynamics of some arboviruses such as WNV (G. L. Hamer et al., 2008).

iii) ***Aedes* subgenus *Stegomyia*.** *Aedes* spp. subgenus *Stegomyia* contains several medically important vectors, two of the most concerning ones are: the yellow fever mosquito (*Aedes aegypti*), primary vector of dengue (DENV), yellow fever (YFV), Zika (ZIKV) and chikungunya (CHIKV) viruses, in tropical and subtropical areas; and the Asian tiger mosquito (*Aedes albopictus*), secondary vector of the same arboviruses, that was original from Asia but has notoriously expanded its geographical distribution throughout the world facilitated by human commercial activities. They both exhibit a high anthropophilic behaviour with host-seeking activity during the daylight hours, which makes them very efficient vectors (Kweka et al., 2018).

A mosquito become infected with pathogens mainly during feeding on an infected host. The susceptibility of a vector to oral infection and its ability to transmit a pathogen efficiently (i.e., the infectious agent enters the midgut, replicates within the vector, and disseminates thought the saliva) is called **vector competence** (VC) (**Fig 2**). Experimentally, VC is assessed by estimating the transmission efficiency, i.e., the number of transmitting females (with infectious saliva) over the total of females that were exposed to the pathogen. VC is a multifactorial property and depends on several intrinsic and extrinsic factors such as the midgut and salivary gland barriers, the immune vector responses, the vector and pathogen genotype, the midgut microbiota, the ambient temperature and humidity, the mosquito diet, or the pathogen dose. All these factors contribute to the success of pathogen infection, dissemination, and transmission in vector populations (Agarwal et al., 2017).

The time interval between an infectious blood meal and the transmission of a pathogen via bite (saliva) is called the **extrinsic incubation period** (EIP). The mean duration of the EIP (which can range from several days to a couple of weeks depending on the virus, the pathogen, and the temperature) can be relatively long compared with the life expectancy of wild mosquito populations, thus compromising pathogen transmission (Johnson et al., 2020). A reduction in the EIP of several pathogens with increasing temperature has been reported (Chan & Johansson, 2012; Dohm et al., 2002; Ohm et al., 2018). However, the relationship between the temperature and pathogen transmission is not linear, exhibiting upper and lower thermal limits from which the transmission does not occur (Mordecai et al., 2019).

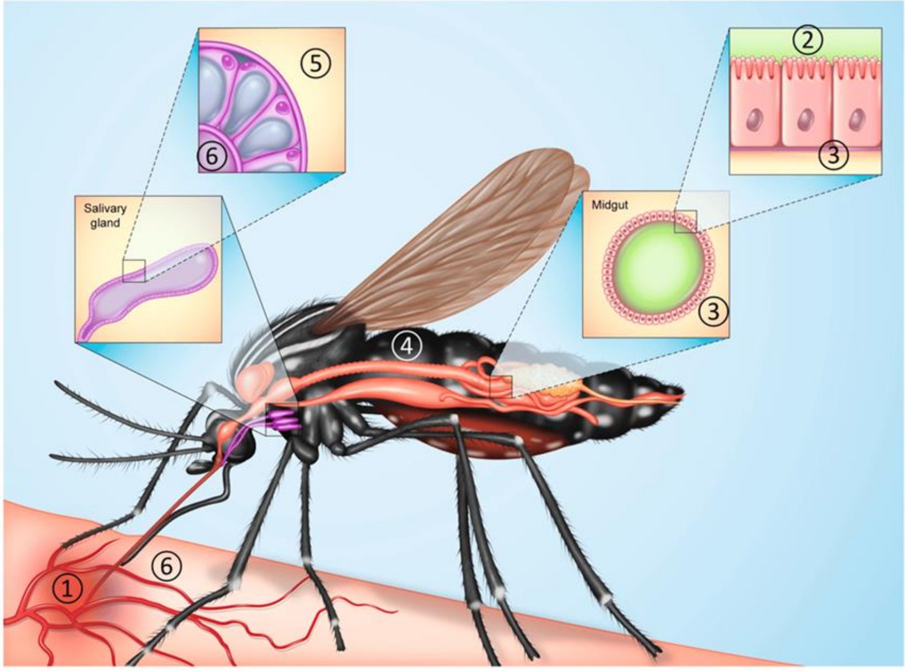


Fig 2. Sequential steps required for a female mosquito to transmit a pathogen. Extracted from (Azar & Weaver, 2019) . **(1)** The pathogen is imbibed by a female mosquito from an infected host; **(2)** the pathogen infects and replicates in midgut epithelial cells overcoming the midgut infection barrier; **(3)** the pathogen escapes from midgut epithelial cells (overcoming the midgut escape barrier) into the mosquito's haemolymph within the hemocoel; **(4)** the infection disseminates in peripheral tissues/organs such as nerves, muscle fibers, or fat body **(5)** the salivary glands become infected (the pathogen overcame the salivary gland infection barrier), and **(6)** the pathogen shedds into the apical cavities of the acinar cells and become present in the saliva for inoculation into subsequent hosts upon feeding.

To fully comprise the transmission potential of vectors, besides VC, the **vectorial capacity** (V) of a vector specie must be considered. V is defined as the total number of potentially infectious bites that would eventually arise from mosquitoes biting a single infectious host on a single day (considering that all mosquitoes become infected when biting the host) (Garrett-Jones & Grab, 1964). It is a multifactorial property that depends on vector density relative to host density; vector daily blood feeding rate; probability of vector daily survival; duration of the EIP; and also, VC (**Fig 3**). High competent mosquitoes may be poor vectors if their frequency of feeding on competent hosts is low or if their life expectancy is short compared to the pathogen EIP. Contrarily, a poor competent vector can sustain or expand a disease outbreak if it is very abundant, its feeding rate is high or the EIP of the pathogen it carries is short (Kramer & Ciota, 2015).

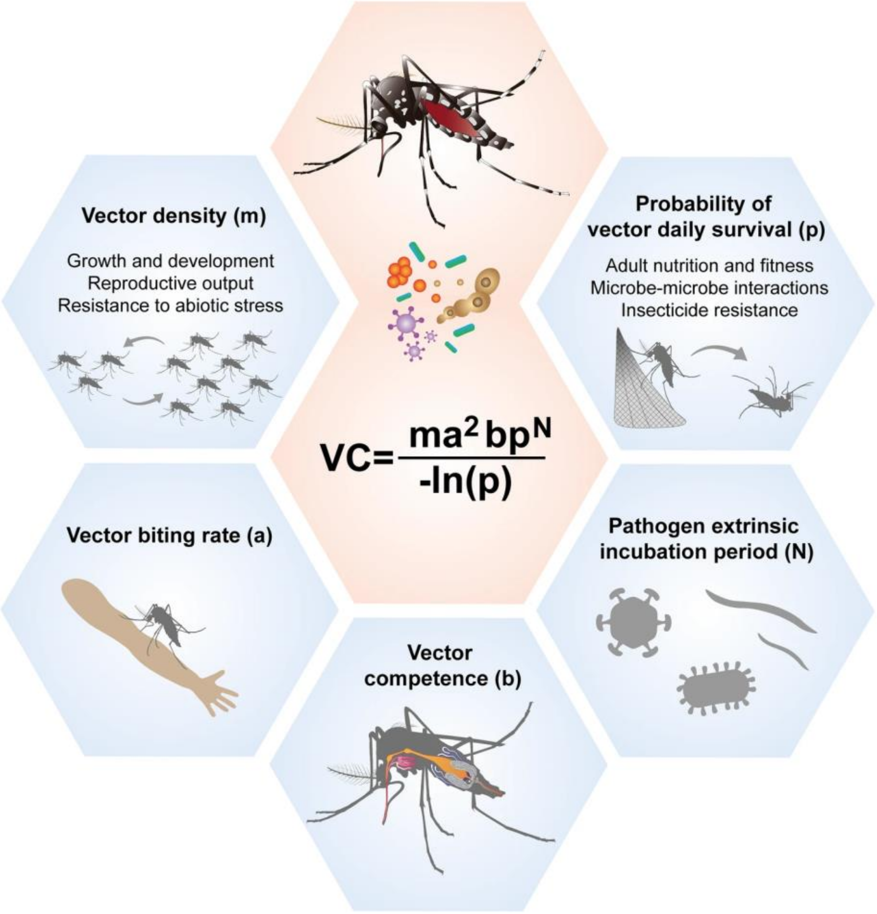


Fig 3. Vector capacity. Extracted from (Cansado-Utrilla et al., 2021)

Pathogen maintenance and amplification cycles in nature involve different strategies, specially: the mentioned **horizontal transmission** (between infected mosquitoes and vertebrate hosts that act as reservoirs or dead-end hosts); and **vertical transmission** (from adult females to their offspring) which has been reported for the main groups of arboviruses (flaviviruses, alphaviruses and bunyaviruses) pointing out the potential relative contribution of this mechanism to virus maintenance during unfavourable environmental conditions, when amplifying hosts responsible for horizontal transmission are not available (Agarwal et al., 2017; Barker & Reisen, 2018). Other types of mechanisms include **venereal transmission** (between males and females during mating), which is considered relatively rare and difficult to identify naturally (Barker & Reisen, 2018).

Understanding vector ecology and the role of vector-host interactions in MBD transmission cycles is necessary to assess the risk of potential outbreaks (Rizzoli et al., 2019; Thongsripong et al., 2021). Some diseases like malaria or dengue have an urban transmission cycle that is entirely sustained by an anthropophagic vector that selectively feeds on humans. The feeding strategies of these anthropophilic vectors like *Anopheles gambiae s.l.* and *Aedes aegypti*, which feed preferentially and frequently on human blood in domestic and peridomestic areas, may result in an increased risk of pathogen transmission (Scott & Takken, 2012). In other cases, as in WNV, the transmission cycle is enzootic (**Fig 4**) and it is sustained by the horizontal transmission between a vector and a reservoir animal, mainly wild birds, that act as primary host. Humans are not an essential component of the transmission cycle (i.e., they do not support a sufficient level of viremia to infect vectors) but rather become incidental hosts when levels of enzootic transmission escalate and become epizootic (Troupin & Colpitts, 2016). In these cases, it may be a spillover from the enzootic cycle to humans as dead-end hosts because of tangential transmission by bridge vectors (Figuerola et al., 2022; Martínez-De La Puente et al., 2018). The degree to which humans are affected may increase if the amplification cycle becomes established among vertebrate animal reservoirs inhabiting urban environments (Figuerola et al., 2022).

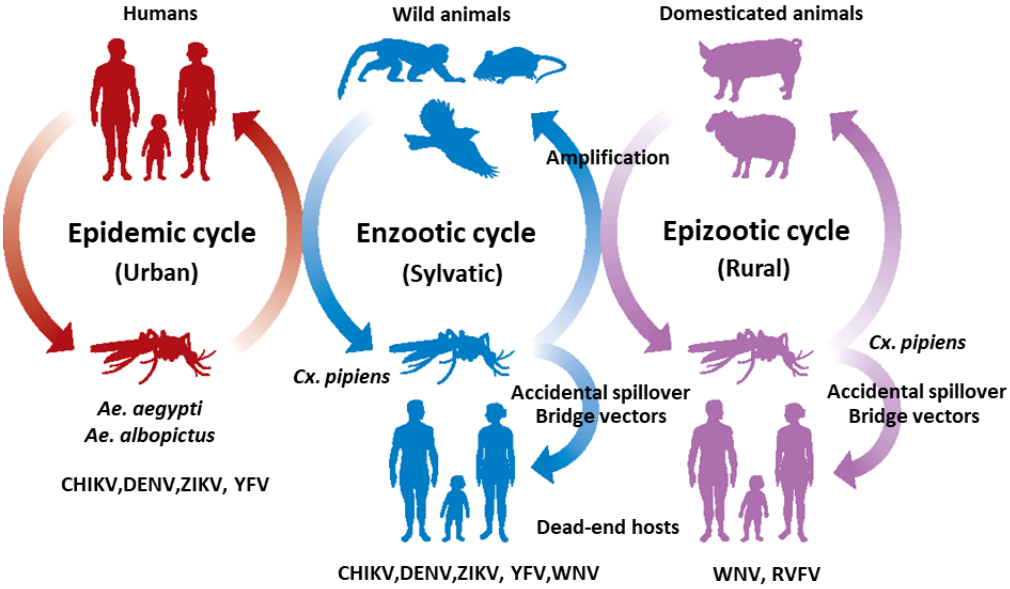


Fig 4. Mosquito borne disease transmission cycles. Adapted from (Weaver et al., 2018).

2.3 Main mosquito vectors in Europe

2.3.1 Aedes albopictus

Aedes (Stegomyia) *albopictus* (Skuse) (Wilkerson et al., 2021), also known as Asian tiger mosquito, is one of the most concerning vector species worldwide. It is a relatively small black mosquito with white scale patches typically distributed in their legs, scutum, and other parts of the body (**Fig 5a**). It is an anthropophilic feeder which bites aggressively during daylight hours, causing serious nuisance to humans. It was original from the tropical forests of South-East Asia but, in the last decades, it has undergone a dramatic global expansion due to human activities, in particular, the trade of used tyres and the passive transportation by ground vehicles (Swan et al., 2022). It is listed as one of the top 100 world's worst invasive alien species by the Invasive Species Specialist Group of the International Union for the Conservation of Nature (Lowe et al., 2000).

Originally described from Calcutta (India), it is now present in all continents except Antarctica. The first record of importation into Europe was in 1979 in Albania and although it became stablished in the country, there were no reports in Europe until 1990, when it was found in Italy. The mosquito expanded rapidly within Italy and to other countries (France, Belgium, Greece, Spain, Balkan countries, and so on) in the following years (ECDC, 2016). The species was found for the first time in the Iberian Peninsula in 2004 (Aranda et al., 2006). The colonization of Europe involved at least three independent introductions in Albania, North Italy and Central Italy that subsequently acted as dispersal centres throughout Europe (Sherpa et al., 2019). Nowadays, *Ae. albopictus* is stablished in all Southern Europe (**Fig 5b**), constituting an important public health issue (Schaffner et al., 2013).

In its original habitat at the edge of the forests, *Ae. albopictus* used to breed in natural spaces such as tree holes, bromeliads, or bamboo stumps. It was considered a rural vector until, as a form of ecological plasticity, it become adapted to urban and periurban environments. There, it started to breed in artificial water-holding containers such as tyres, barrels, catch basins, plant pots, and any other element that support standing-water. Its ecological plasticity also applies for host-choice behaviour. *Aedes albopictus* feeds opportunistically on a wide range of hosts (preferentially mammals) depending on their availability and environmental conditions. If there is a choice though, they prefer to feed on human blood. As a multivoltine specie, it produces several generations per season (Bonizzoni et al., 2013).

Unlike tropical and subtropical populations, which are active throughout the year, temperate strains are adapted to cool temperatures by developing diapausing eggs which allow them to overwinter, supporting their invasion to northern latitudes (Lacour et al., 2015). Generally*, Ae. albopictus* enters diapause when weekly temperatures are below 12.5 ºC to a minimum of 9ºC and photoperiods range from 14 h to 11.2 h. Spring hatching occurs when weekly temperatures are between 10-15 ºC and photoperiods are between 11 h to 12 h (Petrić et al., 2014). The peak of abundance is reached over summer, but the actual timing varies according to eco-climatic variables.

Aedes albopictus

A

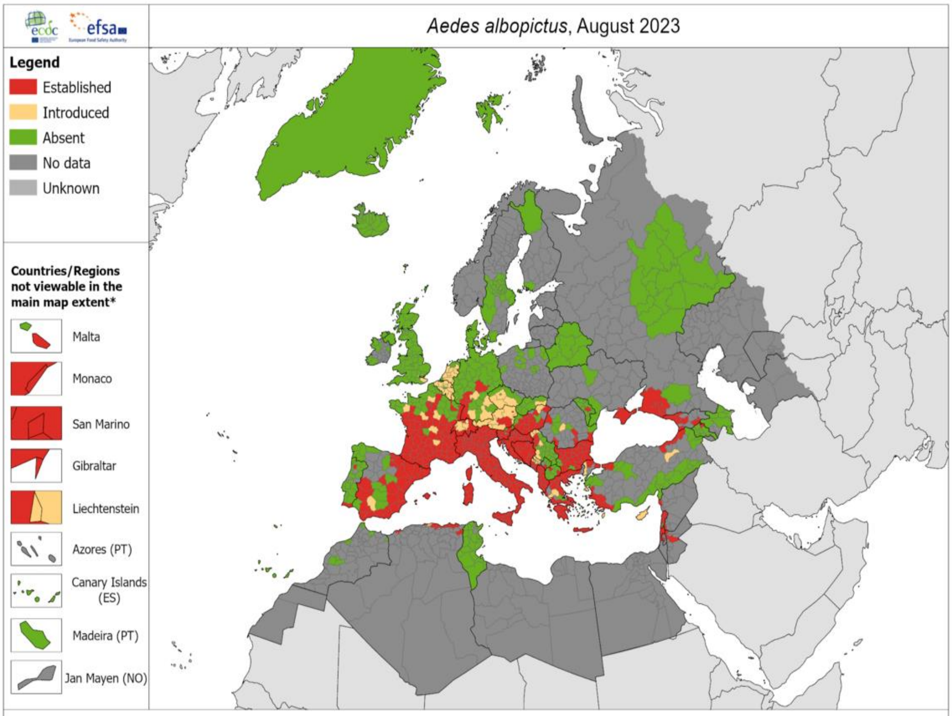
SPECIES NAME: *Aedes (Stegomya) albopictus* Skuse, 1894.

COMMON NAME: Asian tiger mosquito.

ORIGIN: South-East Asia.

DISTRIBUTION: Europe, Middle East, Asia & Australasia, North America, Central America, Caribbean, South America and Africa.

B



ECDC and EFSA, map produced on 25 Aug 2023. Data presented in this map are collected by the VectorNet project. Maps are validated by external experts prior to publication. Please note that the depicted data do not reflect the official views of the countries.
* Countries/Regions are displayed at different scales to facilitate their visualisation.The boundaries and names shown on this map do not imply official endorsement or acceptance by the European Union.Administrative boundaries © EuroGeographics, UNFAO.

Fig 6. *Aedes albopictus* technical sheet. (**A**) Female of *Ae. albopictus* feeding on a human host. Extracted from (ECDC, 2016); (**B**) Current know distribution of *Ae. albopictus* in Europe (August 2023). Extracted from (ECDC, 2023a).

Vector competence studies have shown that *Ae. albopictus* can transmit at least 26 viruses belonging to *Flaviviridae*, *Togaviridae*, *Bunyaviridae*, *Reoviridae* and *Nodaviridae* families. Some of them (DENV, WNV, ZIKV, Japanese encephalitis virus, CHIKV, Eastern equine encephalitis virus, Potosi virus, Tensaw virus, Keystone virus, La Crosse virus and Jamestown Canyon virus), have been isolated from wild-caught *Ae. albopictus*, however their role in transmission is uncertain (Paupy et al., 2009). Currently, *Ae. albopictus* is the primary vector of CHIKV, DENV, ZIKV and dirofilariasis in continental Europe (ECDC, 2016).

The first autochthonous MBD in Europe attributed to *Ae. albopictus* was in Italy in 2007, with 205 confirmed cases of CHIKV from an initial imported case from India (Rezza et al., 2007). The next cases of autochthonous transmission were reported in 2010, first in France (with two cases of CHIKV and two of DENV) (Gould et al., 2010) and later in Croatia (with more cases of DENV) (Gjenero-Margan et al., 2011). These were the first reports of local DENV transmission in Europe since the major dengue epidemic of 1927–1928 in Greece where *Ae. aegypti* was the vector (Louis, 2012). Since then, more cases of DENV have arose year after year in mainland Europe in France (2013, 2014, 2015, 2018, 2019, 2020, 2021, 2022, 2023), Spain (2018, 2019, 2022, 2023) and Italy (2020, 2023). The number of reports per European country has remained at low levels to date, being the outbreak of 2023 the year with the major number of autochthonous reported cases (31 in France and 42 in Italy) (ECDC, 2023c). Also, cases of CHIKV have periodically been reported in France (2014, 2017) and a big last outbreak in Italy (2017), with 270 confirmed and 219 probable cases (ECDC, 2023b). From the beginning of ZIKV surveillance in Europe in 2016, imported cases from returning travellers have been reported year after year. In 2019, two cases of autochthonous transmission due to *Ae. albopictus* were registered in France (ECDC, 2021).

Generally, *Ae. albopictus* is considered to have a lower vector capacity than *Ae. aegypti* for DENV, CHIKV and ZIKV. However, it is the primary vector of these arboviruses in places where *Ae. aegypti* is absent, leading to the occurrence of explosive disease outbreaks in case of CHIKV (Charrel et al., 2007). Phylogenetic analysis of viral sequences of isolates from these outbreaks revealed an example of rapid (1–2 years) convergent evolution between CHIKV and the vector, which resulted from a virus mutation associated with improved replication and transmission efficiency in *Ae. albopictus* (De Lamballerie et al., 2008). In a global context of expansion of *Ae. albopictus*, this evolutionary convergence between arbovirus and vectors could have serious implications for public health.

Aedes albopictus shows ecological plasticity in different traits such as larval breeding sites, feeding behaviour, and climatic adaptation which increase their potential for spread and adaptation to new environments, and influence their coexistence with other vector species. Its anthropophilic behaviour and adaptation to urban environments may contribute to amplify the virus transmission cycle. Moreover, as opportunistic feeders they have a potential role as bridge vectors (Pereira-Dos-Santos et al., 2020). Global change associated with movements of goods and humans together with changes in land uses, socioeconomics and climate contribute to expand the range of this vector thus amplifying its impact on global health.

2.3.2 *Culex pipiens*

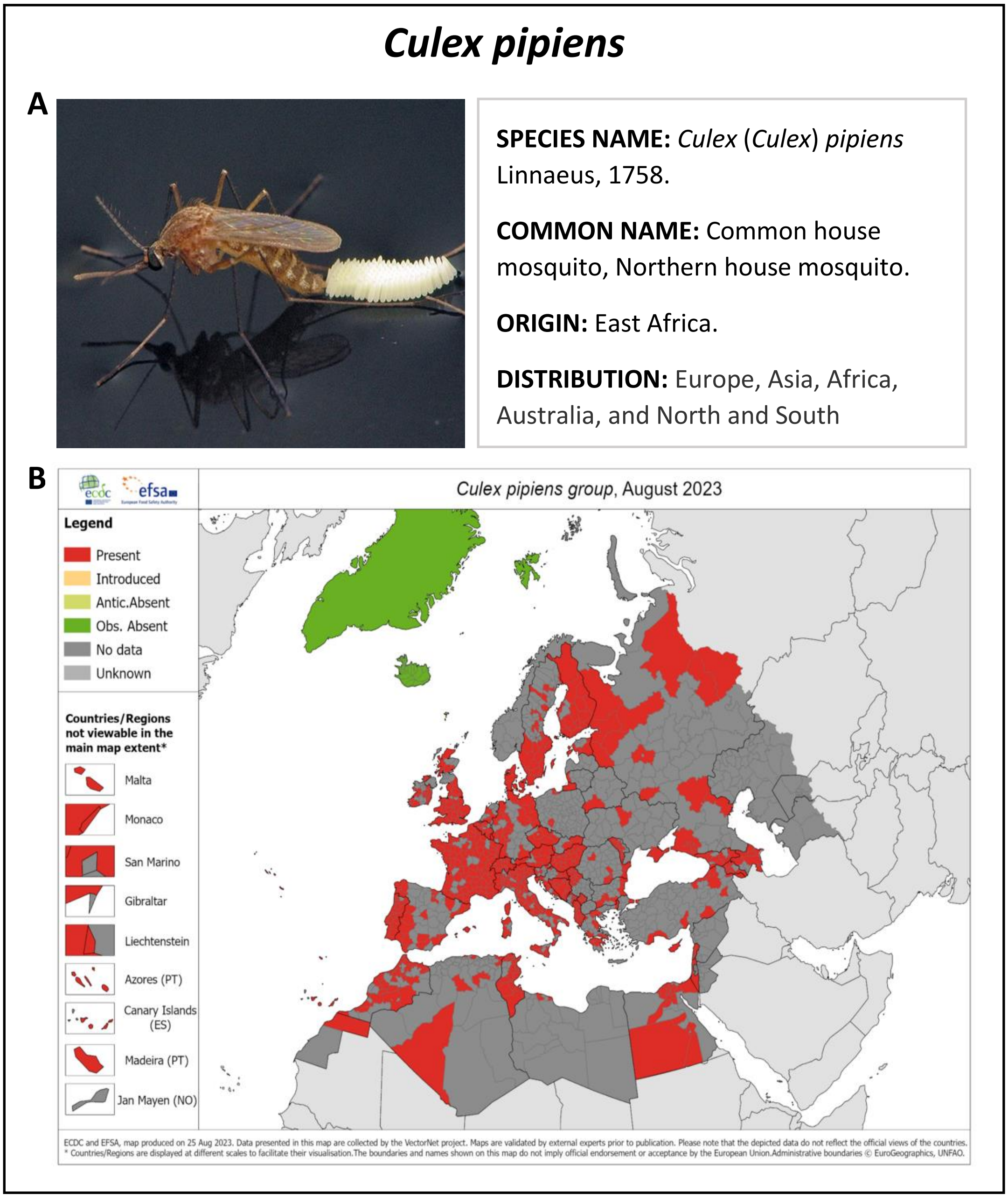
Culex (*Culex*) *pipiens* Linnaeus, 1758, also known as “Common House mosquito” or “Northern house mosquito” is a worldwide distributed mosquito, one of the most important disease vectors in temperate zones and considered a pest in urban environments. It is a medium-size mosquito (4-10 mm), brownish-coloured overall, without any obvious specific pattern (**Fig 6a**). *Culex pipiens* s.s. is a polytypic species that is member of the *Cx. pipiens* complex and has 3 sibling species: *Culex quinquefasciatus* (tropical distribution), *Culex australicus* (Australia) and *Culex globocoxitus* (Australia). The taxon has also one subspecies (*Culex pipiens pallens*) and two forms or biotypes (*Cx. pipiens* f. *pipiens* and *Cx. pipiens* f. *molestus*) (Aardema et al., 2022). *Culex pipiens* s.s is native to Africa, Asia, and Europe although nowadays it is widely distributed in temperate regions of Europe, Asia, Africa, Australia, and North and South America (**Fig 6b**).

The ecology and behaviour of *Cx. pipiens* s.s differ considerably depending on the biotype. *Cx. pipiens* f. *pipiens* typically inhabits rural and urban habitats aboveground, primarily feeds on birds (ornithophilic), mate in swarms in open spaces (eurygamous), require a blood meal to synthesize eggs (anautogenous), and enter a reproductive diapause in winter (heterodynamic). On the contrary, *Cx. pipiens* f. *molestus* typically inhabits belowground urban settles, feeds primarily on mammals (mammophilic) and humans (anthropophilic), mate in enclosed habitats (stenogamous), do not require a blood meal to lay the first clutch of eggs (autogenous) and are active throughout the year (homodynamic) (Haba & McBride, 2022).

The origins of the taxa may have been in East Africa, from a last common ancestor of *Cx. pipiens* s.l. which was probably ornithophilic, eurygamous and homodynamic, such as some contemporary populations of *Cx. pipiens* f. *pipiens* in sub-Saharan Africa. Higher altitudes, such as those found in Ethiopia, with colder weather conditions may have contributed to the evolution of heterodynamy and reproductive diapause. Then, their natural dispersal to cooler climates in Europe and Asia may have given rise to the modern form of *Cx. pipiens* f. *pipiens*. The expansion into the Mediterranean or Near East with more stable human settlements and urbanization may have promoted the evolution of the form *Cx. pipiens* f. *molestus*, adapted to stenogamy and mammophily (Aardema et al., 2022).

Nowadays, both ecotypes coexist in northern Europe and Asia, with two reproductively isolated populations of *Cx. pipiens* f. *pipiens* aboveground and *Cx. pipiens* f. *molestus* in human-made belowground environments. Gene flow between them increases within a latitudinal gradient, with higher hybridisation rates from northern to southern latitudes. Hybridization may have important implications for pathogen transmission by promoting a more opportunistic feeding behaviour in *Cx. pipiens* s.s. that would enhance its role as bridge vector of some arboviruses such as WNV. In fact, feeding preferences has been described as one of the most influential parameters driving intensity and timing of WNV infection (Rizzoli et al., 2015). Therefore, the occurrence of natural hybrid populations leads to a complex ecological mosaic that complicates predictions of vectorial capacity for competent arboviruses (Haba & McBride, 2022).

Culex pipiens s.s. is considered the major vector in Europe of WNV, an endemo-epidemic arbovirus with the longest history of circulation in the continent. The virus is transmitted among birds within an amplification zoonotic cycle via the bite of infected mosquitoes and incidentally infects humans and horses when there is a spillover to an epizootic cycle. The virus was first isolated in 1937 from a patient with fever in the West Nile district of Uganda (Smithburn et al., 1940). Later in 1955, an enzootic transmission cycle was identified involving indigenous wild birds of the Nile Delta as natural reservoirs for the virus and *Culex* mosquitoes as the primary vector group (Work et al., 1955). Since then, many countries from Sub-Saharan and North Africa have reported evidence for WNV. The translocation of the virus from Africa to northern latitudes is attributed to the migratory behaviour of these wild birds travelling north (Brugman et al., 2018).



SPECIES NAME: *Culex* (*Culex*) *pipiens* Linnaeus, 1758.

COMMON NAME: Common house mosquito, Northern house mosquito.

ORIGIN: East Africa.

DISTRIBUTION: Europe, Asia, Africa, Australia, and North and South

ECDC and EFSA, map produced on 25 Aug 2023. Data presented in this map are collected by the VectorNet project. Maps are validated by external experts prior to publication. Please note that the depicted data do not reflect the official views of the countries.
* Countries/Regions are displayed at different scales to facilitate their visualisation.The boundaries and names shown on this map do not imply official endorsement or acceptance by the European Union.Administrative boundaries © EuroGeographics, UNFAO.

Fig 7. *Culex pipiens* technical sheet. (**A**) Female *Cx. pipiens* laying an egg raft on the water surface. Photograph by Sean McCann (**B**) Current know distribution of *Cx.pipiens* in Europe (August 2023). Extracted from (ECDC, 2023b).

The first European outbreak of WNV in humans occurred in 1962 in France, but the first major epidemic was in Romania in 1996, with 393 cases and 17 deaths (Sejvar, 2003). During the first decade of the 21st century, the number of outbreak detections increased in Europe. It remains unclear if it was due to a more frequent annual re-introduction of WNV, a greater focus on surveillance in the Mediterranean Basin, or both. Anyhow, WNV cases have been continuously reported in eastern, western and southern Europe, being 2018 the transmission season with the highest number of cases registered to date (1548 locally acquired cases and 166 deaths, mostly concentrated in Greece, Italy and Romania) (ECDC, 2019). In Spain, the first diagnosed human case of WNV was in 2004 (Kaptoul et al., 2007). Since then, several cases have been annually reported until 2020 where the largest WNV outbreak was declared in southwest Spain with 77 human cases (Macias et al., 2023). The geographic expansion of WNV has continued in Europe towards central and northern countries (Bakonyi & Haussig, 2020), being detected in Germany (in birds and horses in 2018 and in humans in 2019) (Ziegler et al., 2020) and Netherlands (with six human cases in 2020) (Vlaskamp et al., 2020). The identification of WNV in overwintering mosquitoes points out virus persistence in mosquitoes thorough the winter season in Europe (Rudolf et al., 2017).

A related flavivirus transmitted by *Cx. pipiens s.s.*, USUV, has also emerged and spread across Europe around the same time as WNV. It was first isolated in 1959 in the Usutu River in Swaziland and was previously restricted to sub-Saharan Africa until it was introduced in

Europe in 1996 (Vilibic-Cavlek et al., 2020). The virus is sustained within an enzootic amplification cycle in birds, mainly Passeriformes, and have an occasional spillover to humans. USUV has shown to be highly pathogenic for several bird species in Europe and it has become a potential public health concern given the increasing number of reported human infections (Cadar et al., 2017). Although the USUV diversity in Europe appeared in the last decade, phylogenetic analysis suggests a long-term history of virus circulation in the region, with multiple introductions and endemicity in some countries. Migrating birds are the most likely mechanism of translocation over long and short distances (Brugman et al., 2018).

Other viruses such as Sindbis virus, Tahyna virus or Batai virus have been occasionally identified from naturally infected field caught female *Cx. pipiens* s.s in Europe although they are generally poorly reported and their public health importance in Europe is still unknown (Napp et al., 2018). Pathogens of veterinary importance such as dirofilarial worms or avian malaria parasites that are currently circulating in Europe are considered to be transmitted by *Culex pipiens* s.s. as well (ECDC, 2020).

2.4 Epidemiology of mosquito borne diseases.

MBD constitute a major threat for public health authorities worldwide. According to the International Health Regulations (2005) stablished by the World Health Assembly, it is required that all the member states have the ability to detect, assess, report, and respond to public health events (International Health Regulations (2005), 2008); specially if such health events constitute a potential public health emergency of international concern, as it was declared for ZIKV in 2016 (Wilder-Smith & Osman, 2020). In this regard, the **epidemiology** plays a key role by investigating the factors contributing to disease occurrence in a population.

From a public health point of view, different **levels of a disease** can be identified regarding their prevalence in human populations and their geographical and seasonal distribution. Diseases that are always present or reappear consistently at a similar level during a specific transmission season in a certain location are considered to be **endemic** (Dicker et al., 2006). When the disease levels rise above the expected generating an increase, often sudden, in the number of cases within a community, an **epidemic** is declared. An **outbreak** carries the same definition of epidemic but is often used for a more limited geographic area (Dicker et al., 2006).

The disease status is going to determine the levels of intervention and actions carried out by the governments to mitigate the scope of the disease. The frontier between endemic and epidemic, however, is not always clear and usually respond to subjective criteria like disease risk perception (Medley & Vassall, 2017). For some arbovirus such as WNV, even there is virus circulation, institutional measures are not triggered until there are confirmed human cases, which evidence a lack of a One Health perspective in the management of MBD (Dente et al., 2019).

The epidemiology of MBD relies on the study of vector-host-pathogen dynamics and the environmental factors that affect them. It is difficult to elucidate the complex interactions underlying these dynamics since they act at different temporal and spatial scales and usually in a nonlinear manner (Gallana et al., 2013). Therefore, identifying a single key factor or epidemiological determinant responsible of the dynamics of a disease is not realistic denoting the need of an integrative approach for the study of MBD.

Herein, lies the utility of mathematical epidemiology, which objective is providing models of infectious diseases which reflect the interaction between the different components that integrate the whole system (Focks & Barrera, 2006). Two relevant insights of it are: **the mass action principal** (W. H. Hamer et al., 1906) which states that the course of an epidemic depends on the rate of contact between infectious and susceptible individuals; and the **threshold theory** (Kermack & Mckendrick, 1927), which states that that the introduction of a few infectious individuals into a community of susceptible hosts will not give rise to an epidemic unless the density of infectious exceeds a certain critical level.

For infectious diseases, such as VBD, the **basic reproduction number** (R_0) is a common epidemiological metric used to assess the contagiousness or transmissibility of an infectious agent (Macdonald et al., 1956). It is described as the number of secondary cases that will arise from an infected individual in a completely susceptible population (Dietz, 1993); and it depends on the duration of the infectious period, the probability of infection during contact between susceptible and infectious individuals, and the contact rate. When R_0 is greater than one, an outbreak it is expected to continue. It must be applied with caution since it is affected by numerous biological, socio-behavioural, and environmental factors (Delamater et al., 2019).

Stratifying the risk of VBD within a population is a crucial starting point for disease risk assessment. Depending on the presence or absence of either pathogens, vectors, or endemic disease cases in humans, five **VBD contexts** can be identified (M. Braks et al., 2011): Endemic disease cases in humans arise because of pathogen transmission by indigenous vectors; 2) Both the pathogen and the competent vector are present but there are not human cases, e.g. there is pathogen circulation sustained within a sylvatic cycle but it has not been a spillover to a urban cycle; 3) There is a competent vector but the pathogen has not been introduced and consequently there are not locally transmitted human cases; 4) A pathogen is frequently imported into a location where there are not competent vectors, so there is not local transmission between humans; and 5) Neither the vector nor the pathogen are present and there is not a disease burden.

In order to estimate the magnitude of a health problem and consequently implement the corresponding measures for prevention and control, it is necessary to develop an appropriate surveillance program, which must provide quality and timely information to key stakeholders for decision-making (WHO, 2017). Given the complexity of arboviral transmission, different epidemiological scenarios or VBD contexts must require different surveillance strategies, specifically designed with all the available information about the ecology of the disease (Gu et al., 2008).

VBD surveillance gathers information from different data sources (Barker & Reisen, 2018) that are: **pathogen circulation data**, which focuses on the detection and identification of pathogens in humans, animal hosts, and vectors; **serological data**, that aims to monitor the evolution in the immunological responses in human and vertebrate animal hosts that have been exposed to a pathogen; **clinical data**, that usually refers to the passive surveillance in medical and veterinary institutions from humans and animals, which show a clinical syndrome of a disease; and **risk data**, which is based on the detection of risk factors for VBD transmission such as the distribution and abundance of vector populations, the ratios of vector-host contact, and other environmental factors susceptible to affect them.

Timely analysis of local VBD surveillance data will provide the basis for the construction of **early-warning systems** based on thresholds and other alarm signals to give an early response in case of outbreak. Relevant early-warning indicators are the identification of a disease outbreak in neighbouring areas; the increase in the percentage of positive serology; or the increase in the number of syndromic cases. Other alarm signals to be considered are the introduction of new predominant serotypes; the identification of clusters of MBD cases through geographical information systems; climatic changes in rainfall, temperature and humidity; and the increase in the vector presence (WHO, 2016).

Since VBD require, by definition, a vector for pathogen transmission, it seems unquestionable that the eradication or control of vector populations is essential for VBD management, particularly for diseases for which vaccines or specific treatment are not available. **Mosquito control** activities can be based on: **source reduction**, a traditional reliable tool that consists in the elimination of potential breeding sites by removing standing water and modifying habitats to make them unsuitable for mosquitoes to breed; **chemical control**, which has been broadly applied to fight neglected MBD and consists in the application of larvicides (mineral oils, insect growth regulators, organophosphates) and adulticides (specially pyrethroids); **biological control**, as an alternative to chemical control, based on the introduction of predator species (e.g., fish and aquatic invertebrates that feed on mosquito larvae), entomopathogenic fungi, pathogenic microorganisms (e.g., the bacterium *Bacillus thuringiensis var. israeliensis*, used extensively in larval control) or *Wolbachia* endosymbiont bacteria (Huang et al., 2017); and new approaches based on **genetic control**, such as the release in the nature of sterilised males and transgenic mosquitoes (with dominant lethality or gene drives) that will produce unviable offspring (G. H. Wang et al., 2021).

The reduction of vector populations through mosquito control interventions has been the principal method of preventing MBD for over 100 years and remains highly effective in present, contributing to decrease the global burden of many diseases (Wilson Id et al., 2020). However, the widespread development of insecticide resistance in mosquitoes due to the massive spray of chemicals for their control have resulted in the reemergence of MBD in many parts of the world (N. Liu, 2015). Also, the effect of global change regarding land uses, socioeconomics and climate has an impact on MBD by altering the distribution of competent vectors (Franklinos et al., 2019).

Therefore, it becomes clear the necessity of an effective and environmentally sensitive approach to mosquito management that relies on a combination of common-sense and innovative practices through a comprehensive understanding of mosquito biology, mosquito life cycle and pathogen transmission cycles (Fouet & Kamdem, 2018). The basic components of this **integrated mosquito management** include surveillance, measures for personal protection to prevent mosquito bites, actions of source reduction, an integrated control of all mosquito life-stages, insecticide resistance testing, public education, community involvement and evaluation of the actions taken (CDC, 2020).

From an operational perspective, an integrative approach to MBD management requires all the actors involved (medical and veterinarian entomologists and public health professionals) to work coordinately to strengthen the preparedness and response towards health threats (M. Braks et al., 2011). VBD surveillance plans, properly design according to the context, will provide the basis to stratify the risk of a disease and give a proactive response. Early-warning systems with evidence-based action thresholds (based on epidemiological or entomological data) are essential for public health agencies to initiate or intensify control activities aiming at reducing mosquito populations (Aryaprema et al., 2023). Moreover, to measure the effect of vector control, indicators on the presence/absence and abundance of the vector in a given location are necessary, which are also obtained through the implementation of mosquito surveillance plans.

2.5. The role of mosquito surveillance.

Mosquito surveillance consists in monitoring mosquito populations over a temporal or spatial gradient with the objective of determining the presence/absence or abundance of certain species, identifying the presence of pathogen circulation among vector species, and characterizing certain biological, ecological, or behavioural traits of mosquitoes. Mosquito surveillance programs are essential to guide vector-control operations and inform communities and policy-makers against the spread of MBD, highlighting the need of high-quality data (Caputo & Manica, 2020).

The design of an appropriate **mosquito sampling program** can be a complex task and must be ideally guided by the vector species targeted, the specific environmental characteristics of the surveyed site and the objective or ecological problem that is intended to be assessed (Silver, 2008). All this will define the attributes of the sampling such as: timing and duration; geographical coverage; spatial or habitat variability; sampling methods; sampling effort; or sample size. However, sampling design is often more dictated by practical constraints like time availability, resources, and budget limitations, which may limit the quality of the data obtained.

With the aim to support the implementation of a harmonised surveillance of mosquito species of medical and veterinary importance in Europe, the European Centre for Disease Prevention and Control has developed two **reference guidelines** for the surveillance of both native and invasive mosquito species (ECDC, 2012; Schaffner et al., 2014). These guidelines respond to three specific objectives: assessment of nuisance and disease risks to human and animal health; implementation of mosquito control measures and evaluation of their efficacy; and in case of invasive species, early detection of their introduction to new territories, and survey of its possible establishment and spread.

There are two basic approaches to mosquito surveillance: **active surveillance**, which is performed by professionals and refers to sampling methods specifically selected to collect the targeted species at the sites where they may occur, according to the available data or scientific knowledge about the ecology of the mosquitoes; and **passive surveillance**, which refers to data gathered by existing databases and in latest years, submission of reports from non-professionals usually through citizen science platforms. Lately, community-based approaches for passive mosquito surveillance through citizen science initiatives have gained popularity across Europe (Kampen et al., 2015) and can be a good complement of targeted surveillance (Palmer et al., 2017), which remain the main official approach to mosquito sampling, allowing a carefully selected sampling design applied to a specific aim and context.

Active surveillance is deployed by a broad range of sampling methods which include mainly the use of physical traps to attract host-seeking or gravid female mosquitoes, but also the use of aspirators for human landing collections and resting mosquitoes, or dippers to sample larvae and pupae at mosquito breeding sites (Farlow et al., 2020; Q.-M. Liu et al., 2023). Depending on the required entomological information (presence/absence, distribution, abundance and seasonality, biting behaviour, adult resting behaviour, larval habitats, flight range and dispersal, longevity, pathogen screening, etc) one sampling method would be more suitable than others (**Table 1**). The capture efficacy of the different methods in targeting certain mosquito species would depend on its ecology and behaviour (habitat, host-preferences, daily activity patterns, etc).

Table 1 Recommended mosquito collection methods with trapping frequency and trapping period by required entomological information. Adapted from (ECDC, 2014).

Required information	Recommended methods	Frequency of trapping	Period of trapping
Presence, distribution and status	Larval dipping Any adult trapping method	Once or twice a month	Apr–Nov
Adult abundance	CO2-baited traps	Twice a month	Apr–Nov
Seasonal activity	CO2-baited traps Larval dipping Ovitraps	Twice a month	Apr–Nov
Adult resting sites and behaviour	Site visual inspection Aspiration Resting boxes	Twice a month Daily	Jan–Dec Apr-Nov
Host bitting rate	Host-baited traps Host landing collection Human landing collection	Weekly	Jun–Oct
Larval habitats	Larval dipping	Monthly	Mar-Nov
Quality and efficacy of control methods	Larval dipping CO2-baited traps Ovitraps	Before and after the treatments	Jun-Sep
Pathogen screening	Gravid traps CO2-baited traps	Weekly	During and after outbreaks
Longevity	Adult trapping + Mark-release-recapture Adult trapping + laboratory age grading methods	Daily Weekly	Apr–Nov
Dispersal	Adult trapping + Mark-release-recapture	Daily	Apr–Nov

Traditional **taxonomical identification** of the field collected specimens is based on the existence of recognisable morphological differences between species. It requires the expertise of trained entomologists with access to laboratory facilities fully equipped, especially with optical tools such as stereoscopic microscopes. Several morphological **identification keys** for adults and larvae are available but usually restricted to specific geographic areas and genera. In Europe, efforts have been made to develop some reference keys to guide the taxonomical identification of European mosquitoes based on descriptive dichotomous choices as in (Becker et al., 2020), computer-aided pictorical keys based on images as in (Schaffner et al., 2001) or (Gunay et al., 2018) and reverse identification keys as in (ECDC, 2022).

Since female mosquitoes are the ones that are involved in pathogen transmission, many mosquito sampling strategies are designed to target them and consequently, a big part of the taxonomical identification efforts are focused on adult mosquitoes. The main diagnostic features for species identification in adult mosquitoes are the shape of the scutum in the thorax, the pattern of bands in the dorsal part of the abdomen, and the presence and distribution of pale bands in the hind legs, among others (**Fig 8**). To identify the sex, the main features are the relative length of the palps with respect to proboscis, the amount of antennal hair, and the genitalia.

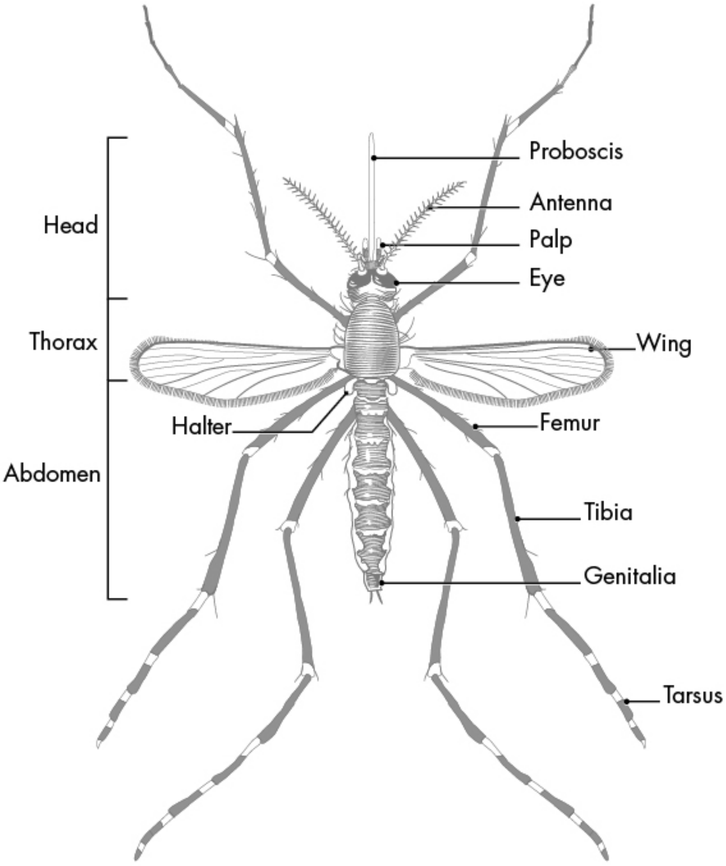


Fig 8. Body parts of an adult mosquito. Extracted from CDC, 2023

When captured in the field, specimens got usually damaged by the sampling procedure and eventually loose part of their scales or entire body parts such as legs, compromising the task of identification. Fresh specimens are always preferred, especially if they are being latter used for pathogen screening and isolation. However, this is not always feasible during routine monitoring, where samples are frequently collected from traps each 7 to 15 days, thus assuming a certain degree of sample damage due to desiccation, predation and/or climatic events.

In this case, so as in others where the morphological classification through visual clues is not possible, e.g., there are sibling species from a species complex that are morphologically indistinguishable, it becomes necessary a molecular analysis approach. These molecular techniques require well-equipped laboratory and skilled personnel, which is costly, particularly in large-scale studies.

Novel approaches for mosquito surveillance point towards the use of machine learning (ML) techniques and a variety of sensing devices for the automated remote classification of mosquitoes through the so-called intelligent sensors (Joshi & Miller, 2021; Santos et al., 2019). In the recent years, there has been an upward trend of works that include the use of these new technologies for mosquito management with emphasis in *Aedes* mosquitoes and to a lesser extent, to *Culex* and *Anopheles* genera (Joshi & Miller, 2021). These new techniques, based on image (J. Park et al., 2020), acoustic ((Vasconcelos et al., 2019) and pseudo-acoustic systems (Genoud et al., 2018), offer high accuracy results for mosquito classification mainly in laboratory conditions. These intelligent sensors have the potential to overcome some of the major disadvantages of traditional surveillance methods; thus, related with the costs in terms of time and resources involved in the tasks of trap setting, sample collection, labelling, transportation, preservation, manual inspection for taxonomical identification and process of the results. Besides, with the expansion of the internet of things (IoT), they offer the possibility of gather data about vector populations in real-time which may contribute to a more time accurate MBD assessment (Nayak et al., 2023).

Currently, there are very few studies that have tested the performance of these devices in field conditions (Day et al., 2020; Lai et al., 2022) where there are uncontrolled environmental variables which may affect their efficacy. The progress in the field of remote automated mosquito surveillance requires to perform field tests to evaluate if the accuracy classification results obtained in the laboratory can be extrapolated to a field-work scenario with the ultimate objective of being further implemented into the routine of entomological surveillance.

3. OBJECTIVES

The main objective of the present thesis is **to develop a ready-to-use reliable tool for the automated remote monitoring of *Aedes* and *Culex* mosquito populations in real-time**. This novel technological approach, based on an entomological optical sensor coupled to a mosquito trap and trained with machine learning, aims to overcome the current limitations of traditional mosquito surveillance by reducing the associated operational costs (time, personnel, and resources) and providing accurate relevant entomological information with an unprecedented time resolution. With this goal, five specific objectives have been defined and are listed below:

1. **To stablish a conceptual framework about the use of novel technologies applied to mosquito surveillance based on machine learning techniques, with a focus on optical sensing systems.** To that purpose, a review of the current scientific literature was performed, comparing traditional versus novel approaches for mosquito surveillance. This objective was assessed in the **Chapter 1** of the present thesis.

2. **To train a new optical sensor prototype with machine learning techniques to differentiate the genus and sex of *Aedes* and *Culex* mosquitoes in laboratory conditions.** To that end, the flight features extracted from mosquito flight recordings made by the sensor were trained with different machine learning algorithms to find the best classification accuracy for genus and

sex. This goal was evaluated in the **Chapter 2** of the present thesis.

3. To evaluate the efficacy of a new sensor system for the automatic counting and identification of the genus and sex of *Aedes* and *Culex* mosquitoes in field conditions. To do so, the performance of the sensor for mosquito monitoring was compared to the manual inspection performed by entomologists in two field sites. This aim was addressed in the **Chapter 3** of the present thesis.

4. To determine the chronological age of *Culex pipiens* mosquitoes. Two approaches were employed for this aim: i) the transcriptional age grading technique based on the differential expression of age-related genes; and ii) the use of an optical sensor for the automated discrimination of mosquito age based on differences in the flight pattern. This objective was assessed in the **Chapter 4** of the present thesis.

5. To automatedly identify *Aedes albopictus* infected with Zika virus using an optical sensor system. To that end, the sensor was trained on a dataset of infected and non-infected mosquitoes to build a machine learning model for the automated classification. This objective was assessed in the **Chapter 5** of the present thesis.

4. CHAPTERS

4.1 CHAPTER 1: Novel technologies for mosquito surveillance

The use of artificial intelligence and automatic remote monitoring for mosquito surveillance (González et al., 2022).

The use of artificial intelligence and automatic remote monitoring for mosquito surveillance

María I. González[1], Joao Encarnação[2], Carles Aranda[1,3], Hugo Osório[4], Tomás Montalvo[5,6] and Sandra Talavera[1*]

[1]IRTA, Centre de Recerca en Sanitat Animal (CReSA, IRTA-UAB), Campus de la Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain.

[2]Irideon SL, Av. Josep Tarradellas 38, 08029 Barcelona, Spain.

[3]Servei de Control de Mosquits del Consell Comarcal del Baix Llobregat, Barcelona, Spain.

[4]Centre for Vectors and Infectious Diseases Research/National Institute of Health Doutor Ricardo Jorge, Avenida da Liberdade 5, 2965-575 Águas de Moura, Portugal.

[5]Servei de Vigilància i Control de Plagues Urbanes, Agència de Salut Pública de Barcelona, Barcelona, Spain.

[6]CIBER Epidemiologia y Salud Publica/CIBERESP, Av. Monforte de Lemos 3-5, Pabellón 11, Planta 028029 Madrid, Spain.

*Correspondence: sandra.talavera@irta.cat

Abstract

Mosquito surveillance consists in the routine monitoring of mosquito populations: to determine the presence/absence of certain mosquito species; to identify changes in the abundance and/or composition of mosquito populations; to detect the presence of invasive species; to screen for mosquito-borne pathogens; and, finally, to evaluate the effectiveness of control measures. This kind of surveillance is typically performed by means of traps, which are regularly collected and manually inspected by expert entomologists for the taxonomical identification of the samples. The main problems with traditional surveillance systems are the cost in terms of time and human resources and the lag that is created between the time the trap is placed and collected. This lag can be crucial for the accurate time monitoring of mosquito population dynamics in the field, which is determinant for the precise design and implementation of risk assessment programs. New perspectives in this field include the use of smart traps and remote monitoring systems, which generate data completely interoperable and thus available for the automatic running of prediction models; the performance of risk assessments; the issuing of warnings; and the undertaking of historical analyses of infested areas. In this way, entomological surveillance could be done automatically with unprecedented accuracy and responsiveness, overcoming the problem of manual inspection labour costs. As a result, disease vector species could be detected earlier and with greater precision, enabling an improved control of outbreaks and a greater protection from diseases, thereby saving lives and millions of Euros in health costs.

Keywords: mosquito monitoring, remote surveillance, acoustic sensor, optical sensor, intelligent sensor, smart trap, machine learning, Internet of Things (IoT).

Mosquito surveillance and traditional monitoring methods

Mosquitoes (Diptera, Culicidae) are responsible for the transmission of diverse medically and veterinary important disease agents (viruses, protozoans and other parasites) which cause serious diseases in humans and animals, such as malaria, dengue, Zika, yellow fever, chikungunya, West Nile virus, Eastern equine encephalitis or avian malaria. Entomological surveillance plays a key role in human and veterinary disease surveillance within the framework of the 'One Health' concept, where interdisciplinary collaboration and communication in healthcare is crucial (ECDC, 2012; Schaffner et al., 2014; WHO, 2017). A paradigmatic example of this 'One Health' approach would be the West Nile Virus (WNV) surveillance. This implies a coordinated strategy of Public Health actors that carry out the diagnoses of possible infected horses and humans that are dead end hosts of the pathogen; the monitoring of *Culex* mosquitoes that may act as vectors in the areas with WNV cases; and the detection of possible infected birds which may act as reservoirs of the virus.

Mosquito surveillance methods should provide clear and meaningful information for program managers and policy-makers for the purpose of: (1) determining and quantifying the composition of mosquito populations which are present in a specific area; (2) monitoring changes in mosquito populations; (3) identifying the presence of new invasive mosquito species which can act as disease vectors; (4) detecting mosquito-borne diseases; (5) determining which control measures need to be conducted; (6) performing the quality assessment of control measures; and (7) designing accurate risk assessment programs in order to prevent and manage potential disease outbreaks (Flores, 2015; Schaffner et al., 2014).

Mosquito surveillance can be understood as a task involving the routine monitoring of immature stages and adult mosquito populations over the course of an entire mosquito season (Flores, 2015; Silver, 2008). Several methodologies have been developed to sample and analyse different stages of the biological cycle of mosquitoes (egg, larvae and adults), although most of them mainly target adults since only adult female mosquitoes are responsible for disease transmission (Focks, 2003; Sivagnaname & Gunasekaran, 2012). Thus, with the exception of pathogen monitoring in immature stages to investigate vertical transmission, adult mosquito surveillance is probably the most precise approach to properly monitor mosquito populations for vector-borne disease (VBD) risk assessment. While some methodologies focus on resting mosquitoes, such as aspiration in vegetation that is performed with entomological aspirators, most have been developed to catch flying mosquito females when seeking hosts for blood feeding or gravid females when seeking oviposition sites (Becker et al., 2020; Service, 1993).

To allow standardised monitoring of adult mosquito populations, many types of traps have been developed to attract different target species. Some rely solely on a conventional incandescent filament light bulb as the main source of attraction or use an ultra-violet light source while others add CO_2 or chemical attractants to the light source. Various models are commercially available. The most popular are adapted models of CDC mosquito light-traps, EVS trap (**Fig 1**) and in the last decade, BG sentinel traps (**Fig 2**) with different combinations of light, CO_2 and chemical lure (EMCA, n.d.). Other traps include the Reiter trap for gravid females, and even types developed not just for mosquito sampling but for mosquito control as well, such as the Mosquito Magnet™, among others. Despite these methodologies, during the first decade of the 21st century the need became evident for a much greater effort to develop, manufacture and market new tools that would be effective for different species and environmental conditions and that could be standardised in different countries in order to obtain more significant and comparable data (Qiu et al., 2007).

Several studies have compared the efficacy of different commercial trapping devices, reporting differences in both performance and efficacy depending on the target mosquito species, the type of attractant and other environmental factors (Brown et al., 2008; Y. Li et al., 2016; Lühken et al., 2014; Roiz et al., 2012). Generally, BG-traps have shown better, or at least a similar performance, compared to CDC, EVS or MM traps (Y. Li et al., 2016; Lühken et al., 2014), but the results have been dependent on multiple factors and varied from one study to other. It is important to consider, as pointed out by Brown et al. (2008), that differences between traps could affect the estimations of species abundance and composition.

Fig 1. Classic EVS mosquito trap baited with a container with CO 2 pellets.

Fig 2. BG sentinel trap.

Traditional surveillance methods have two main limitations. The first is the cost in terms of time and professionals involved in the surveillance (trap placement, sample collection and transport of the sample to the laboratory for the counting and identification of captured mosquitoes). The second limitation is the inevitable time lag between the moment that the trap is placed in the field and the moment the sample is collected. This lag can be crucial, potentially resulting in the dynamics of mosquito populations in the field not being accurately and timely monitored (Focks, 2003).

In this scenario, artificial intelligence (AI) is forging a path in improving traditional entomological surveillance methods by generating new techniques for the automated remote monitoring of mosquito populations. These new approaches include the emergence of automated electronic devices which remotely classify mosquitoes based on the analysis of their flight pattern (Potamitis, 2014; Santos et al., 2019).

In addition, the use of the 'Internet of Things' (IoT) is enabling that the information collected remotely in the field could be sent wirelessly to a cloud server in real time (P. Eliopoulos et al., 2018; Geier, Weber, Rose, Gordon, et al., 2016; Potamitis et al., 2017). Thus, eliminating the gap between trap installation and collection, representing mosquito population dynamics much more accurately.

New technological approaches for remote mosquito surveillance through the perspective of artificial intelligence

Acoustic sensing technology

Mosquito flight tones have been extensively studied since the first half of the 20th century, mainly through the use of acoustic methods such as microphones (Kahn & Offenhauser, 1949). Mosquito flights produce a tone as a side effect of wing movement. This tone is also a communication signal that is frequency-modulated during courtship and can be detected by other mosquitoes thanks to certain properties of their antennae including Johnston's organ at the base of each antenna (Cator et al., 2009; Gibson et al., 2010). Rapid frequency modulation flight in males occurs as a response to female wing beat frequency and is likely to represent a pre-copulatory controlled flight to maintain a close-range position while attempting to seize and engage terminalia with the female (Simões et al., 2016). Females have the ability to reject or accept the male mating attempt. In the event that the interaction between male-female pairs is successful, copulation will take place preceded by an acoustic harmonic convergence (Aldersley et al., 2016; Aldersley & Cator, 2019).

With these acoustic properties in mind, entomologists have been pursuing the control of mosquitoes by means of sound traps for many decades (Kahn & Offenhauser, 1949) and continue to do so (Diabate & Tripet, 2015; Rohde et al., 2019). Sound traps, such as the Sound Gravid Aedes Trap (SGAT), the Male Aedes Sound Trap (MAST) (Staunton, Crawford, et al., 2020), or other modified commercial traps with an acoustic basis, are nowadays being used as cost-effective alternatives for field use in areas with sterile male mosquito rear-and-release programs (Johnson & Ritchie, 2016; Rohde et al., 2019; Staunton, Rohde, et al., 2020).

The acoustic detection of insects is a highly active research field, especially in its application to food crops and stored grain pests (Eliopoulos et al., 2016; Hagstrum et al., 2012; Potamitis et al., 2009) but also with respect to pests of medical importance, such as mosquitoes (Salim et al., 2017; Vasconcelos et al., 2019). In recent years, so-called deep learning techniques have become widely used in bioacoustic classification tasks based on the analysis of mosquito wing beat frequency. However, since mosquitoes from different species can actually have overlapping frequency distributions, it seems insufficient to use the fundamental wing beat frequency as the sole distinguishing characteristic between species (Chen et al., 2014). To improve the classification method, metadata such as time or place of recording can be used as additional features to differentiate between mosquitoes with varying circadian activity or geographic distribution.

Current approaches for mosquito wing-beat analysis and classification through acoustic sensors include the use of mobile phones as an easily available tool for entomological surveillance (Fernandes et al., 2021; Y.Li et al., 2017; Mukundarajan et al., 2017). Mobile phones offer the advantage of automatically registering time and location stamps for acoustic data and allow the collection of other metadata such as photographs which can support identification. Studies based on mobile phone-based bioacoustics demonstrate that even low-cost smartphones are capable of accurately recording mosquito wing-beat frequencies, enabling continuous and large-scale data mapping which can be particularly useful in resource-constrained areas (Mukundarajan et al., 2017). In this sense, there are some open data platforms that rely on the participation of non-expert volunteers to record the wing-beat sound of the mosquitoes. Two of the most popular ones are 'ABUZZ' (Mukundarajan & …, 2019) and 'Humbug Zooniverse' (Kiskin et al., 2020).

The inconvenience of acoustic methods is the limitation to the quality of the microphone recordings of the insects in field conditions. Many mosquito bioacoustics experiments are undertaken in unnatural conditions with tethered individuals or in acoustically isolated spaces, thus leading to difficulties to apply these models in in field conditions (Arthur et al., 2014). Given this difficulty in microphone-sourced field recordings, classification models based on machine learning algorithms commonly suffer from scarce and poor-quality data.

Chen et al. (2014) reported a 'lack of progress' in acoustic technology applied to the automatic classification of insects. This can be attributed to limitations of the microphones themselves. One such limitation is microphone sensitivity. The sound attenuates with the distance from the microphone according to an inverse squared law, which means that if an insect is flying three times more distant from the microphone, the sound intensity will drop to one ninth. When increasing the microphone sensitivity to mitigate this effect, any surrounding noise will saturate the signal. Filtering insect detection can then become a complex task, as well as requiring more system processing power. Besides, systems based on a microphone and recorder set spend the entire experiment running time making recordings, thus increasing power consumption.

The foremost challenges for acoustic sensing approaches are related to dealing with the problem of the signal-to-noise ratio of recorded audio and power consumption. As a result, optical approaches for remote sensing and automatic classification have gained in popularity as they offer significant performance advantages (Potamitis & Rigakis, 2015; Santos et al., 2018, 2019a).

Optical sensing technology

Optical technology for mosquito wing beat analyses dates back to the second half of the 20thcentury when the first photoelectric cell was discovered to detect the light modulation of a flying insect crossing its field of detection (Richards, 1955). This was the starting point for the implementation of numerous studies on the use of optical sensors to monitor mosquito flight patterns which continue to the present day (Gibson et al., 2010; Kirkeby et al., 2016; Ouyang et al., 2015; Potamitis, 2014; Potamitis & Rigakis, 2016).

The diverse light source options for optical sensing include laser and LED (light emitting diodes). Potamitis and Rigakis (2015) developed a novel noise-robust optical sensor to record insect wing beats and analysed the recording performance of both types of light sources, comparing them to the recordings of an acoustic sensor . The results showed that both performed as well or even better than the acoustical sensing approach in any ambient light condition. Unlike acoustic sensors, optoelectronic sensors only record when triggered by flying insects, allowing large savings in power consumption. In addition, optoelectronic sensors are capable of modulating the optical signal at high frequencies, thus eliminating major optical interference sources and increasing sensor efficiency without further data processing requirements (Santos et al., 2018).

Optical sensors basically comprise an optical emitter (a laser beam or a LED array) and an optical receiver (a phototransistor, mainly photodiodes) creating a FOV (field of view). When an insect crosses the FOV, fluctuations in light intensity (caused by the partial occlusion of the light from the wing's movement) are perceived by the optical receiver. The signal containing information on the detected insect's wing beat frequency is then amplified, filtered, and demodulated in an audio signal (G. E. A. P. A. Batista et al., 2011; Potamitis & Rigakis, 2016b). The conversion of the optical signal into audio data allows comparison of the results obtained with those available in the literature for acoustic systems.

The practical applications of these new findings involve extending the use of optical sensors from laboratory tests to the production of massive datasets and the creation of smart insect traps that can count, recognise, and alert for the presence of insects of economic and public health importance (Potamitis et al., 2018). Novel optoelectronic sensor prototypes are being trained with several machine learning algorithms, mainly Bayesian classifiers, to learn how to distinguish between mosquito species and mosquito gender (male and female) based on their wing beat frequency (G. E. A. P. A. Batista et al., 2011; Genoud et al., 2018, 2019; Ouyang et al., 2015; Potamitis & Rigakis, 2016b). While high accuracy values in gender discrimination are now commonly obtained, classification to species level is still challenging (Genoud et al., 2018), although the use of deep learning techniques has shown promising levels of precision (Fanioudakis et al., 2018).

The biggest difficulty appears when trying to distinguish two different mosquito species from the same genus as they may have overlapping frequency spectrums. This suggests that the fundamental wing beat frequency alone, although it may be sufficient to distinguish the mosquito genus or gender, it may be insufficient on its own to properly classify mosquito species. This inefficacy will be even more apparent in the context of field measurements, where plenty of mosquito species, Diptera and other insects may be present. A common way to improve identification accuracy is to add other predictor variables in addition to fundamental wing beat frequency (G. E. A. P. A. Batista et al., 2011; Chen et al., 2014; Genoud et al., 2019). For instance, Genoud et al. (2019) proposed the use of the depolarisation ratio of the mosquito body together with the wing-beat frequency to distinguish gravid from non-gravid females, which reported high accuracy results.

Another option to increase the accuracy of automated taxonomical classification of mosquitoes in field studies may be the use metadata (Chen et al., 2014): meteorological features (temperature, humidity, and air pressure), spatiotemporal features (distance from freshwater, land cover type, human/livestock population density, local agricultural type, time of year, time of day, etc.) and circadian rhythms. Certain species are more adapted to survive in particular environmental conditions, e.g. many mosquitoes are native to tropical and subtropical regions, where the climate is typically warm and wet. The ambient temperature can be determinant in insect classification since it influences insect metabolism, leading to an increase in the wing beat frequency. Villarreal et al. (2017) reported an increase of 8-13 Hz per degree Celsius (°C) in females of *Aedes aegypti*, revealing a highly dependent relationship between these factors(Villarreal et al., 2017). Circadian rhythm is also an important feature to be considered since mosquitoes have different peaks of activity throughout the day which can be of help to distinguish between species. However, circadian rhythm cannot be used without at least a rough estimate of the population of each considered species (Genoud et al., 2019). If a species with a small population has an activity peak at the same time as another with a much larger population but with lower activity, although their probability of interaction with the sensing instrument may be equal, the classification system will consider the former to be much more likely, thus inducing a bias in the results.

New optoelectronic devices for remote sensing include, in addition to insect counts and classification, the use of IoT technology. This allows that the entire information that is being registered remotely in the field, is also being transmitted wirelessly to a central monitoring agency in real time for risk assessment analysis. In this way, novel optoelectronic sensors can be self-organised in networks that collectively report data at local, regional, country, continental, and global scales. The emergence of so-called e-traps has the potential to profoundly impact entomological surveillance and pest control (Potamitis et al., 2017).

Smart trap technology

Novel smart traps entail the possibility of automating everything that is still presently done manually (collecting insect information in the field, processing that information, and sending it to vector control technicians) thanks to the use of IoT technology. The development of IoT solutions using conventional approaches is complex and time consuming due to the lack of common architectures and languages, and the widespread use of non-standard, proprietary interfaces and sensor data formats. Numerous developers, companies and R&D groups have been using state-of-the art commercial platforms like Arduino (Italy), Raspberry Pi (UK) or BeagleBone (USA), which are capable of prototyping straightforward sensor applications with low technology readiness levels (TRL) of between 1 and 4. However, such platforms may be insufficient if advances are to be made to TRL 5 prototyping and above, especially if dealing with sensors that are not off-the-shelf. This implies that off-the-shelf platforms offer limitations to reach TRL 9 (go-to-market), where manufacturers will be fighting issues of functionality, cost, power consumption, scalability, margin, manufacturability, testability, packaging, mechanical robustness and working conditions (e.g. temperature, humidity), etc.

To address this, IRIDEON (Spain) has developed SENSCAPE®, a disruptive, modular, standards-based framework for the development of fast IoT time-to-market solutions. There are several advantages to developing an IoT application with SENSCAPE: (1) ready-made hardware platforms – static and mobile; (2) standards based, (3) interoperability; (4) scalability; (5) low power consumption; (6) reduced costs; (7) smartphone integration; (8) customisation and (9) cloud-ready. The general idea is to use SENSCAPE® as the platform to combine a sensor capable of capturing physical information from flying insect, with two emerging disruptive technologies: IoT and AI.

The IoT refers to systems of physical devices that receive and transfer data through wireless networks without human intervention, while AI refers to the combination of algorithms developed to have machines reasoning like human beings. The combination of these various elements could lead to a solution where each trap acts as an interconnected device that can remotely analyse each captured flying insect, just as a professional entomologist would do.

For the moment, only one optical sensor product designed for the remote monitoring of mosquito populations is commercially available, the Biogents BG-Counter (Germany) (Geier, Weber, Rose, Gordon, et al., 2016). This device is able to distinguish between mosquitoes and other different insects, and to count mosquitoes, but does not provide further information on mosquito species, sex or other attributes. In parallel, there is another optoelectronic sensor prototype that has been created by IRIDEON (Spain), which is already capable of distinguishing between species, sex and age of mosquitoes in laboratory conditions (Brosa, 2018). The mosquito sensor is an optoelectronic device comprising an emitter, an array of LEDs, and an array of phototransistors acting as photoreceptors connected in parallel. This optical setup generates a light field. The sensor constantly captures the input from the sensor but only processes the samples when a triggering event occurs, i.e. when there is a perturbation of the light field. Optical sensor, microprocessor, and wireless communications are integrated into the electronics module.

Smart trap stations can be deployed as a wireless sensor network (WSN) with bidirectional management of data between sensors and a cloud application framework. When an insect is drawn into the trap equipped with a sensor, its characteristic wing flapping modulates the light field. Captured signals are sampled at a rate sufficient to resolve the fundamental frequency of the wing-flap as well as several overtones. The light field is also perturbed by other physical elements associated to the flying insect: the kinetics of flight (speed, direction, and trajectory) and morphology (body/wing size and shape). Each of these physical elements of flying insects that cause a perturbation of the light field leads to a species-specific signature. The signal of this signature is filtered, amplified, acquired and processed using a combination of AI methods (e.g. rule-based systems, genetic algorithms, artificial neural networks and fuzzy models).

Depending on the tests performed, these methods can be used to count each event that perturbs the light volume, determine if the event is caused by a flying insect, analyse if the flying insect is a mosquito or not, classify the genus of the insect, identify the species, identify the sex and estimate the age in days (**Fig 3** and **Fig 4**).

These assets have been benchmarked by experts and judged to be at TRL 7 for genus (accuracy 90-92%), species (accuracy 75-80%) and sex classification (accuracy 93-99%) of *Aedes albopictus*, *Ae. aegypti* and *Culex pipiens*. The sensor also achieved TRL 5 for age classification of each of these species (accuracy 61-95%), giving an overall TRL 6 (Brosa, 2018). Further work is being done to improve the overall accuracy of the solution to reach TRL 9 by 2022.

Fig 3. Vector control technician installing a trap with IRIDEON's smart mosquito sensor.

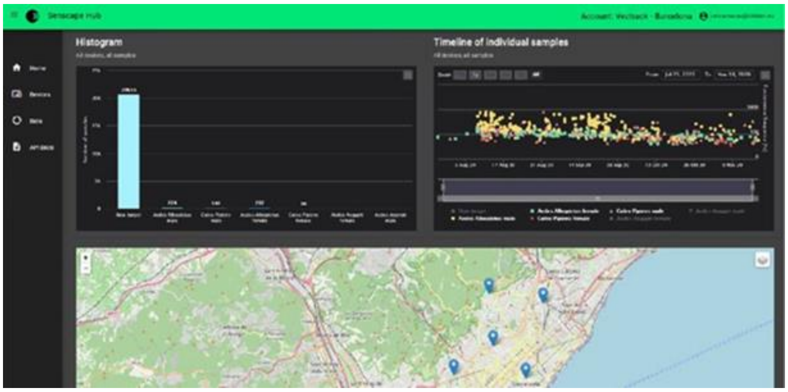


Fig 4. Dashboard of the software cloud application of the smart mosquito sensor.

How intelligent traps can improve mosquito monitoring and arbovirus control programs

Integrated pest management (IPM) relies on the accuracy of pest population monitoring (Freier & Boller, 2009). Without gathering information of population dynamics, and related ecological factors, it is almost impossible to execute an appropriate control at the right place and time. Mosquitoes are usually spread across large areas and boundaries, and the use of traditional surveillance methods which are strongly dependent on human labour is unsuitable for efficient large-scale monitoring (BiPRO & EC, 2009). Fully automated remote monitoring could be the key in this context.

Earth observation service for preventive control of insect disease vectors – the VECTRACK project

Obtaining high quality field information is notoriously costly and time consuming. The amount of money required can significantly be reduced by combining cost-efficient sampling strategies, remote sensing, and spatial modelling techniques to compute risk maps of vector presence and abundance, as well as maps indicating high-risk zones for the establishment of exotic species at local or regional level. Such maps could then serve as a basis for targeted surveillance and VBD risk assessments. To address this, IRIDEON is heading a Horizon 2020 (H2020) project called VECTRACK, in collaboration with AVIA-GIS (Belgium), the IRTA-CReSA research institute (Spain) and the public health institute CEVDI/INSA (Portugal). VECTRACK constitutes a novel and unique opportunity, integrating the added value of Earth observation (EO), spatial-positioning and information and communications technology (ICT) technologies: Copernicus data + operational vector mapping with spatial modelling + IoT ground sensors + IoT smart mosquito traps + IoT interoperable disease vector data cloud application. The proposed innovation is a service platform for which Copernicus is a critical part of the solution. The main objective is to develop and validate a new Copernicus-based EO service to monitor disease vectors, associated to a novel ground wireless sensor network comprising miniaturised nodes measuring micro-environmental data (T °C, %RH, etc.), together with a smart trap station acting as a gateway.

Earth observation platform can measure land surface temperature and vegetation, which act as the main drivers of vector population (C3S 2020). Given the importance of the evolution of the meteorological parameters, the technical requirements for these satellites are: (1) high temporal resolution (1 day); (2) medium spatial resolution (1 km); and (3) measurement in the visible/near infrared part of the electromagnetic spectrum for derivation of vegetation indices and in thermal infrared for temperature.

In this context, it is important to mention the contribution of AVIA-GIS in their development of VECMAP, a seamless system and service that integrates the entire process of producing risk maps into a single package that supports all the steps required to map and model, at various scales, the distribution of vectors and to plan surveillance and control programs. This system provides all the satellite data required to obtain the risk maps, however, is limited by the fact that it uses data from periodical manual trap inspections. This value proposition is strengthened by IRIDEON's smart IoT ground sensors, deployed in the field integrated with standard commercial mosquito traps. With the combination of all approaches, it is finally possible to remotely and automatically acquire near real-time ground data on mosquito counts, sex, species, age and local micro-environmental parameters. This data is invaluable as an automatic and direct input to feed mosquito-borne epidemic models.

Future approaches

With the use of novel smart traps, new challenges will appear; the automated identification of different mosquito species should be improved to the same level as when it is performed by a skilled entomologist and should be supervised until this degree of accuracy is reached. New maintenance and logistic protocols will need to be developed, as traps will go from being mobile and temporary to fixed and permanent.

With new methodologies, surveillance and control programs can be significantly affected as they require important scientific and logistic efforts for the management of large amounts of mosquito traps and collected samples. With the use of remote monitoring systems, once the system has been developed, these efforts can be redirected to other areas and most of the classification work would be done in an automatic way, but always with an accurate quality control system. The data will be completely interoperable and thus available for the automatic running of prediction models, the performance of risk assessments, the issuing of warnings and the undertaking of historical analyses of infested areas. In this way, vector control professionals could establish automatic surveillance programs with unprecedented accuracy and responsiveness, overcoming the labour costs of manual inspections. As a result, disease vector species will be detected earlier with greater precision, enabling improved control of outbreaks and a lower risk of disease transmission.

4.2 CHAPTER 2: Training the sensor in laboratory conditions

A novel optical sensor system for the automatic classification of mosquitoes by genus and sex with high levels of accuracy (González-Pérez et al., 2022).

A novel optical sensor system for the automatic classification of mosquitoes by genus and sex with high levels of accuracy.

María I. Gonzalez-Perez[1†], Bastian Faulhaber[2†], Mark Williams[2], Josep Brosa[1], Carles Aranda[1,3], Nuria Pujol[1], Marta Verdún[1], Pancraç Villalonga[2], Joao Encarnação[2], Núria Busquets[1†], Sandra Talavera[1†]*

[1] IRTA, Centre de Recerca en Sanitat Animal (CReSA, IRTA-UAB), Campus de la Universitat Autònoma de Barcelona, Bellaterra, Cerdanyola del Vallès, Spain

[2] Irideon S.L., Barcelona, Spain.

[3] Servei de Control de Mosquits del Consell Comarcal del Baix Llobregat, Barcelona, Spain.

† María I. Gonzalez-Perez and Bastian Faulhaber contributed equally to this work.

† Núria Busquets and Sandra Talavera contributed equally to this work.

*Correspondence: sandra.talavera@irta.cat

Abstract

Background

Every year, more than one billion people are infected and almost one million die due to vector borne diseases mainly transmitted by mosquitoes. Vector surveillance plays a major role in the control of these diseases that includes, as a key factor, a suitable and rapid taxonomical identification. New approaches for mosquito surveillance include the use of acoustic and optical sensors in combination with machine learning techniques, to provide an automatic classification of mosquitoes based on their flight characteristics, including wingbeat frequency. The development and application of these methods could enable the remote monitoring of mosquito populations in the field, which could lead to significant improvements in vector surveillance.

Methods

A novel optical sensor prototype coupled to a commercial mosquito trap was tested in laboratory conditions for the automatic classification of mosquitoes by genus and sex. Recordings of more than 4300 laboratory-reared mosquitoes of *Aedes* and *Culex* genera were made using the sensor. Five features were extracted from each recording in balanced datasets and used for the training and evaluation of five different machine learning algorithms to achieve the best model for mosquito classification.

Results

The best accuracy results achieved using machine learning were: 94.2% for genus classification; 99.4% for sex classification of *Aedes;* and 100% for sex classification of *Culex*. The best algorithms and features were; for genus classification: deep neural network with spectrogram; for *Aedes* sex classification: logistic regression with spectrogram or Mel Frequency Cepstrum Coefficients and gradient boosting with Mel Frequency Cepstrum Coefficients; for *Culex* sex classification: various features and algorithms gave the best performance.

Conclusions

To our knowledge, this is the first time a sensor coupled to a standard suction trap provides automatic classification of mosquito genus and sex with high accuracy using a large number of unique samples with class balance. This system represents an improvement of the state of the art in mosquito surveillance and encourages future use of the sensor for remote, real-time characterization of mosquito populations.

Keywords

Mosquito trap, automatic classification, optical sensor, machine learning, deep learning, Aedes, Culex, genus and sex classification, mosquito surveillance.

Background

Approximately 80% of the world's human population lives with the risk of one or more vector-borne diseases (VBDs), and every year > 700,000 people die as a result (WHO, 2020).In an increasingly connected world, travel and trade contribute to the spread of VBDs. Furthermore, a global warming scenario may lead to more favourable conditions for the survival and life cycle completion of the vectors (Rossati, 2017) and may affect their abundance and distribution (Khasnis & Nettleman, 2005). Mosquitoes (Diptera: Culicidae), particularly those belonging to *Aedes, Anopheles* and *Culex* genera, are one of the deadliest vectors worldwide. Mosquito species can transmit diseases such as malaria, dengue, yellow fever, West Nile fever, Zika, Chikungunya and others (WHO, 2014). According to World Health Organization directives (ECDC, 2014) and European Centre for Disease Prevention and Control guidelines (ECDC, 2012, 2018) appropriate surveillance methods and indicators are needed to: determine the composition and monitor changes in mosquito populations, identify the presence of new invasive species, monitor mosquito-borne diseases, quantify the transmission potential of vectors and enable the design of accurate control programs.

A range of insect trap types and methods are used in regular monitoring and surveillance of immature and/or adult mosquito populations (Focks, 2003; Romero-Vivas & Falconar, 2005). Although immature stage monitoring can be easier to set up, it is not useful for estimating adult abundance due to the lack of correlation between egg, larval and pupal density indices, and adult indices (Romero-Vivas & Falconar, 2005).

Studies show that the seasonal variation in mosquito abundance is better represented by adult trap monitoring than by other indices (e.g. House Index) based on immature stages (Codeço et al., 2015). Therefore, adult mosquito surveillance is generally the most widely applicable and accurate solution, especially for VBD risk assessment (Krökel et al., 2006). Many adult mosquito monitoring systems rely on traps using light, chemical attractants or CO_2 as a bait. Most traps include a suction fan to draw approaching insects into a catch bag within the trap, and such types have been successfully used in many studies (Farajollahi et al., 2009; Lühken et al., 2014). However, they require the catch bag to be periodically collected in the field, followed by a time-consuming process of identification of the collected specimens by entomologists. The time delay between insect trapping and analysis may limit the correct characterization of the temporal dynamics of mosquito populations. Such delays may also result in degradation of the insects in the catch bag because of desiccation or predation. New approaches for entomological surveillance include novel optical sensors to sense the characteristics of flying mosquitoes and analysis methods including machine learning methods to enable classification of mosquitoes in near real-time (Chen et al., 2014; Genoud et al., 2018, 2020; Potamitis, 2014; Potamitis et al., 2015; Santos et al., 2019), which is crucial for surveillance programs.

Since the 1940s, microphones have been used to sense the audible flight tones emitted by flying mosquitoes which may be associated with a particular mosquito genus, species, or sex (Offenhauser & Kahn, 1949). Acoustic methods are still employed today in applications such as sound traps which emit species and sex specific sound frequencies to attract mosquitoes (Johnson et al., 2018)and in classification systems such as those in which citizen scientists use their mobile phones to record mosquitoes (Li et al., 2017; Mukundarajan et al., 2017). However, it is hard to obtain acceptable quality audio recordings of free flying insects in the field due to the presence of background noise (Chen et al., 2014). To address this, optical methods have been employed, in which a light source is used to illuminate the flying insect and a light sensor is used to detect the light reflected and scattered, or attenuated, by the insect in flight (Brydegaard & Lazzari, 2015; Kirkeby et al., 2016; Mullen et al., 2016; Ouyang et al., 2015; Potamitis & Rigakis, 2016a, 2016b; Song et al., 2019). The use of optical methods in this field began in 1955 when a photoelectric cell was used to detect the light modulation produced by a flying insect crossing its field of view (Richards, 1955). In recent years, several optoelectronic sensors have been developed and used in conjunction with machine learning techniques to classify flying mosquitoes, with promising levels of accuracy (G. E. A. P. A. Batista et al., 2011; Fanioudakis et al., 2018; Genoud et al., 2018, 2020; Silva et al., 2015).

Variables known to condition mosquito wingbeat fundamental frequency or its detection include taxonomy, sex, parity status, size, and environmental temperature (Genoud et al., 2019; Gibson et al., 2010; Staunton et al., 2019; Villarreal et al., 2017). Historically, wingbeat frequency has been used as the only predictor variable for mosquito classification, but it appears insufficient on its own to differentiate between mosquito species, especially those of the same genus (Chen et al., 2014). This could limit field applications, where different mosquito species can coexist, with the possibility of overlap in wingbeat frequency distributions (Kim et al., 2021). Efforts have been made in recent years to improve classification methods to distinguish among mosquito species, sex and even parity status (Genoud et al., 2018, 2019, 2020). In some cases, more advanced optical approaches have been used, for example to determine insect body and wing depolarization ratio, to improve the accuracy of classification (Genoud et al., 2020).

In addition to the selection of the proper predictor variables and machine learning algorithms, the use of metadata such as the climatic conditions, the spatiotemporal localization and other ecological features accompanying mosquito captures may also be relevant for remote mosquito classification in the field (Chen et al., 2014; Fanioudakis et al., 2018), since different mosquito species have different behaviour and ecological needs (geographical distribution, climatic range, circadian rhythm, peaks of activity, etc.). According to new paradigms of remote mosquito surveillance, wingbeat sensor information and metadata could be sent wirelessly in real-time to a server using *Internet of Things* (IoT) technology (Eliopoulos et al., 2018; Geier, Weber, Rose, Obermayr, et al., 2016; Potamitis et al., 2017) with the potential to improve entomological surveillance.

Currently there is only one commercial optical sensor product available for the remote monitoring of mosquito populations (Geier, Weber, Rose, Obermayr, et al., 2016). It is called the BG-Counter (Biogents, Germany), which according to the company, can distinguish mosquitoes from other insects and count mosquitoes. However, the sensor does not provide information about mosquito genus, species, sex or other attributes.

In this study, we present the results of a prototype optical sensor, which is coupled to the entrance of a commercial mosquito trap. The trap is of a type widely used for mosquito surveillance in the field and contains a suction fan. The fan causes the mosquitoes to pass through the sensor more quickly and with a more perturbed wingbeat compared to free flight conditions as described in another work (J. Wang et al., 2020). For the present work, 4335 flights from mosquitoes of *Aedes* and *Culex* genus were recorded using the sensor. The three species for the study*, Aedes albopictus*, *Aedes aegypti* and *Culex pipiens*, were chosen because they are major vectors of arboviruses, have a significant impact on public health and are a focus of vector surveillance and control programs in many parts of the world. A set of features were extracted from each recording and used to train a series of machine learning algorithms to determine which combination of feature and algorithm gave the best performance in classifying mosquitoes by genus and sex. Whilst the scope of this work is limited to the classification of genus (Aedes/Culex) and sex (female/male), the inclusion of the two Aedes species in this study improves the genetic variability and permits future work on species classification using the data set from the present work.

Methods

Mosquito rearing conditions

Three species of mosquitoes, from two genera, were used to generate the dataset:

i. *Culex pipiens*, population of Gavà (2012), Barcelona, Spain (41.3000º, 2.0167º)
ii. *Aedes albopictus*, population of Sant Cugat del Vallès (2005), Barcelona, Spain (41.4667º, 2.0833º)
iii. *Aedes aegypti*, population of Paea (1994), Tahiti, French Polynesia (-17.6889°, -149.5869°)

The mosquito populations were all reared under controlled environmental conditions in a climatic chamber at a temperature of 28 ºC and a relative humidity of 80%, with a light:dark photoperiod of 12:12 hours, except for *Cx. pipiens* (with a light:dark photoperiod of 11:11 hours plus 1h of dusk and 1h of dawn). *Culex pipiens* and *Ae. albopictus* were reared in biosafety level 2 (BSL2) laboratory and *Ae. aegypt*i in a biosafety level 3 (BSL3) laboratory at IRTA-CReSA facilities. Larvae were maintained in plastic trays with 750 ml of dechlorinated tap water (renewed three times per week) and were fed with fish pellets (Goldfish Sticks-TETRA, Melle, Germany) *ad libitum*. Pupae, upon appearance, were immediately placed in insect cages (BugDorm-1 Insect Rearing Cage W30 x D30 x H30 cm, MegaView Science, Talchung, Taiwan). After metamorphosis, adults were fed with sucrose solution (10%) *ad libitum*. Females were not fed with blood to avoid any body-size or flight variation. For *Aedes* females, the sucrose solution was removed 24 hours before the sensor tests. For *Cx. pipiens* females, this was done 48 hours before, to improve their affinity for the attractant used in the trap.

Sensor and trap description

The prototype sensor was designed and produced by Irideon S.L. (Barcelona, Spain) and was coupled to the entrance of a commercial BG-Mosquitaire suction trap from Biogents AG (Regensburg, Germany), as shown in **Fig 1a**.

The trap coupled to the sensor was placed in an insect rearing cage (BugDorm-4S4590 W47.5 x D47.5 x H93.0 cm, MegaView Science, Talchung, Taiwan). The trap was fitted with a sachet of BG-Sweetscent chemical attractant from Biogents AG. The air flow generated by the fan was approximately three metres per second in the downward direction. When a mosquito flies close to the entrance funnel of the sensor, it is sucked in by the fan, detected by the sensor and trapped in the catch bag inside the body of the trap.

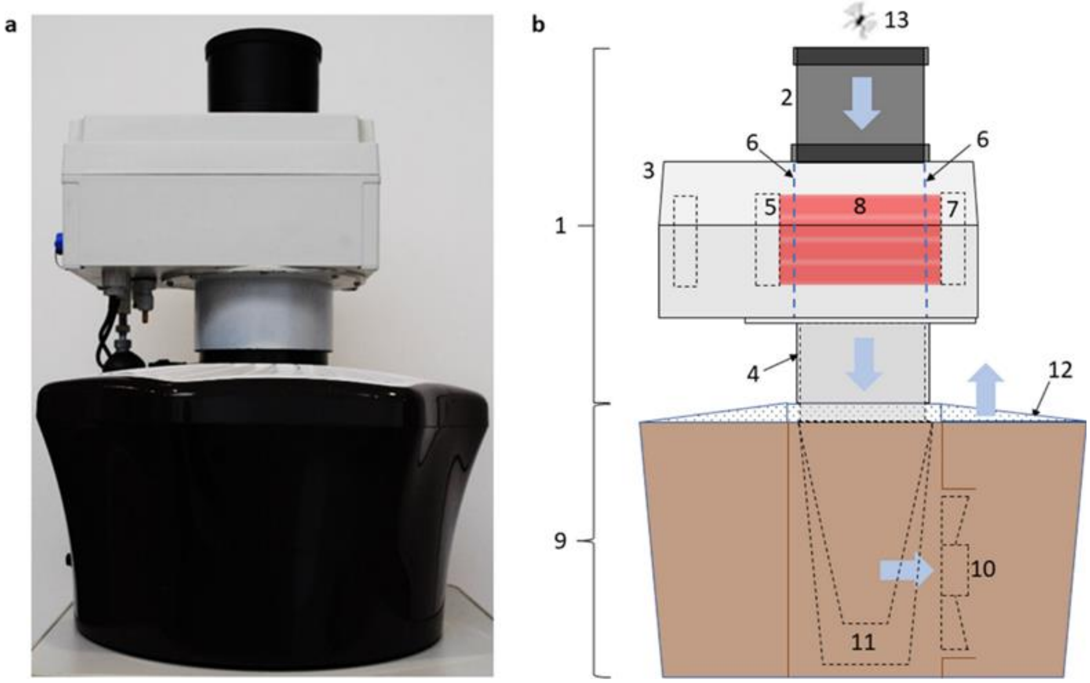


Fig 1. a. Prototype sensor (top) fitted to a BG-Mosquitaire trap (bottom). **b.** Side view diagram of sensor and trap to illustrate operation. The exterior of the sensor unit (1) is formed by an inlet tube with a diameter of approximately 100 mm (2), sensor housing (3) and outlet tube (4). The housing contains an optical emitter (5), which projects collimated beams of light through the transparent flight tube (6) and onto an optical receiver (7) to create a sensing zone (8) within the flight tube. The trap (9) contains a suction fan (10), a removable catch bag (11) made of textile mesh and a perforated lid (12). The fan produces a flow of air downward through the inlet tube, flight tube and catch bag and upward through the perforated lid as indicated by the blue arrows. An insect (13) which flies close to the entrance of the inlet tube may then be sucked downwards through the sensing zone where it will be recorded and then trapped in the catch bag. As the mosquito passes through the sensing zone it casts a shadow upon the optical receiver according to the so-called optical extinction mode of operation. As the insect flaps its wings within the sensing zone, the light falling on the optical receiver is modulated, giving rise to changes in the amplitude in the recorded waveform.

The sensor contains an optical emitter panel and an optical receiver panel which face each other through a transparent flight tube with a diameter of 105 mm. The optical emitter comprises a two-dimensional (2D) array of 940 nm wavelength infrared light emitting diodes (LEDs); and the optical receiver comprises a 2D array of 940 nm photodiodes. The optical sensor has an active length of 70 mm in the downward direction. The basic operating principle of the optical sensor is illustrated in **Fig 1b**.

The output of the optical sensor is amplified and acquired by an analog to digital converter (ADC) with a sampling frequency of 9603 samples per second. When a mosquito enters the sensing volume, it automatically triggers a recording of up to 1024 samples, i.e., of up to 107 milliseconds duration. The duration of a typical mosquito flight is around 50 milliseconds. The sensor automatically adds a timestamp to each recording, along with the measured ambient temperature.

Data acquisition process

Mosquitoes from Aedes and Culex genera were anesthetized with carbon dioxide 48 and 72 h respectively before each experiment. They were separated into groups by species (*Cx. pipiens*, *Ae. albopictus* and *Ae. aegypti*) and sex (male, female).

Culex pipiens and *Ae. albopictus* were introduced into the insect rearing cage in batches of 20 individuals to reduce the chance of multiple mosquitoes passing through the sensor simultaneously. Batches of ten individuals were used for *Ae. aegypti* because of their greater affinity to the attractant. All mosquitoes were introduced at a distance of 20 to 30 cm from the entrance of the sensor to ensure that they could fly freely until they approached it and were sucked in, to approximate field conditions.

Each recording corresponds to a different mosquito, i.e. trapped mosquitoes were not re-used to generate more recordings. Wingbeat files were tagged with species and sex class by the operator. After each experiment, the wingbeat recordings were downloaded from the sensor and processed using a Python script to produce playable and viewable audio files, as depicted in **Fig 2a**. Wingbeat recordings were examined manually, and those deemed to be invalid, such as recordings containing more than one mosquito or where a mosquito may have hit the wall of the flight tube, were excluded from the dataset. The excluded recordings represented 2.3% of the data.

The resulting dataset contained 4335 wingbeat recordings, comprising 2472 of Aedes genus (882 Ae. aegypti and 1590 Ae. albopictus) and 1863 of Culex genus (all Cx. pipiens). There were 1211 Aedes females, 1261 Aedes males, 964 Culex females and 899 Culex males. Females were in an age range of 2 to 16 days old and males were in an age range of 2 to 9 days old. These age ranges provide a representative variety in the dataset.

All recordings took place with the sensor and trap located in the laboratory facilities of IRTA-CReSA during daylight hours. The average ambient temperature measured was 25.8 (standard deviation = 1.2 °C).

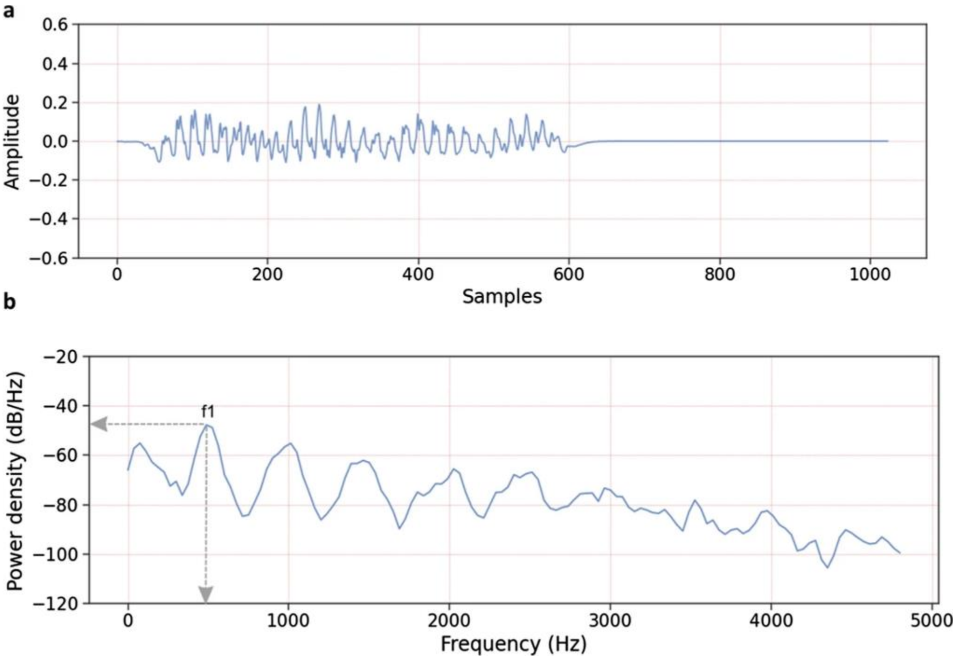


Fig 2. a. Example of a recorded mosquito flight with ADC sample number (0 to 1023) on the *x*-axis and amplitude on the *y*-axis, scaled to a range of [− 1, 1], which equates to the full-scale range of the ADC. A high pass filter in the optical receiver attenuates frequencies < 300 Hz to remove electronic offsets and low-frequency noise, which also attenuates the signal due to the body of the insect. Baseline correction has been applied by subtracting the average value of the recording from each data point in the recording. **B.** Power spectral density (PSD) plot of a typical mosquito flight. The wingbeat fundamental peak is labelled as f1. The fundamental frequency is indicated by the vertical arrow and the fundamental peak power by the horizontal arrow. The various peaks to the right of f1 are harmonics of f1, i.e. at frequencies of 2*f1, 3*f1, etc. The power density has units of (units^2/Hz) on a logarithmic (dB) scale. A level of 0 dB/Hz corresponds to a white noise signal time domain signal with a power density of 1.0 unit^2/Hz. The fundamental peak power density levels in this study are typically < − 40 dB/Hz, i.e. < 1×10^{-4} units^2/Hz. The noise floor of the system, i.e. with sensor active but with no insect in the sensing zone, is < − 85 dB/Hz from 0 to 300 Hz and < − 90 dB/Hz from 300 Hz.

Feature extraction

The following five features were extracted from each wingbeat recording via the application of digital signal processing methods:

- The power spectral density (PSD) shows the power of the signal at different frequencies. It is calculated using Welch's method (Villwock & Pacas, 2008), in which the wingbeat recording is divided into several overlapping segments. A windowing function is applied to each of the segments and a series of periodograms is obtained by calculating the power spectrum of each windowed segment. Finally, the periodograms are averaged to give the PSD (Bisina & Azeez, 2017). A PSD plot of a typical mosquito recording is shown in **Fig 2b**.
- Wingbeat fundamental frequency in Hertz [Hz] is determined from the PSD using a peak search method. The wingbeat fundamental frequency is the frequency at which a mosquito flaps its wings. It is characteristic of mosquito taxonomy and sex, and varies depending on intrinsic variables of mosquito biology (size, age, parity status, mating behaviour) (Genoud et al., 2019, 2020; Gibson et al., 2010; Staunton et al., 2019) and environmental variables such as temperature (Villarreal et al., 2017). The typical range of mosquito wingbeat fundamental frequencies is 300 to 900 Hz (Kim et al., 2021).
- The fundamental peak power density [dB/Hz] (hereafter referred to as fundamental peak power) is also determined from the PSD as shown in **Fig 2b** and represents the peak power density of the sensor output at the wingbeat fundamental frequency. It is equivalent to the intensity of the sound

produced by a flying mosquito, typically ranging from 40 to 80 dB (Dou et al., 2021; Menda et al., 2019).

- The spectrogram is a series of spectra calculated from multiple overlapping segments of the wingbeat recording. Each spectrum is generated by applying a Fourier transform to the segment to provide information about the amplitude of the various frequency components in the segment. The spectrogram represents the variations of the frequency content of the signal over time, rather than an average for the whole signal as given by the PSD (Oppenheim, 1970).
- Mel Frequency Cepstral Coefficients (MFCCs) are calculated by converting the frequencies of a spectrogram to the Mel scale and applying overlapping triangular filter banks before calculating the cepstrum, by transforming the spectra to a logarithmic scale, and then applying an inverse Fourier transform (Zhu, 2011). Please refer to **Additional file 1: Text S1** and **Fig S1** for further details.

The PSDs have 257 values, generated using a window length of 512 samples. The spectrograms and MFCCs are obtained using nine segments of 512 samples; then, 16 Mel filter banks are applied to each spectrum to give a total of 144 values. All the MFCC coefficients are used. Each individual feature and one combined feature (fundamental frequency and fundamental peak power) were used for the machine learning models.

A scatter plot of the wingbeat fundamental frequency and peak power features is shown in **Fig 3a** for the entire dataset, in **Fig 3b** for all *Aedes* samples and in **Fig 3c** for all *Culex* samples. In **Fig 3a**, which is coloured by genus, a high degree of overlap between the genera is observed. In **Fig 3b** and **Fig 3c**, which are coloured by sex, two clearly separated clusters are observed. The distributions of the two single-value features, fundamental frequency and fundamental peak power, for the three classifications are shown in **Additional file 1: Fig S2**.

Machine learning

The goal of the machine learning process was to compare the performance of five selected machine learning algorithms using the features described above, in classifying mosquito genus and sex. A labelled dataset consisting of the feature set was used to train, evaluate, and compare the classification models. The following five machine learning algorithms were used: logistic regression (LR), gradient boosting (GB), random forests (RF), support vector machines (SVM), and a fully connected deep neural network (DNN). These algorithms were chosen due to their widespread usage and good performance (Schmidhuber, 2014). A brief overview of each algorithm is given in Additional file 1: Text S2. Of these algorithms, the more complex ones, such as DNN or RF, were also used with the single-value features (fundamental frequency and fundamental peak power) because they can model nonlinearities, unlike LR.

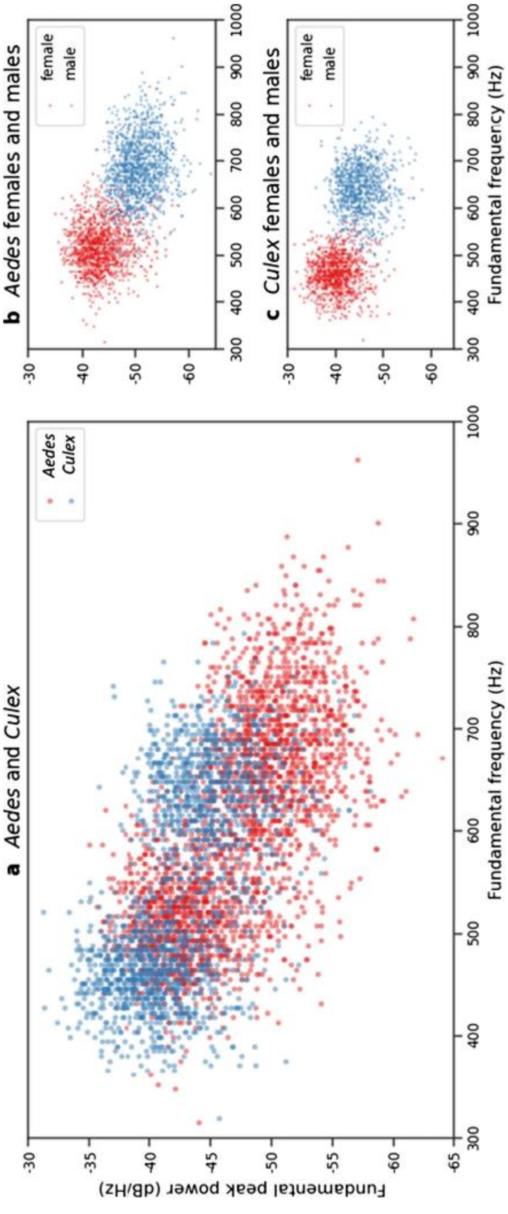


Fig 3. a. Scatterplot of wingbeat fundamental frequency and peak power for the full dataset showing *Aedes* genus in red and *Culex* in blue. **b** Scatter plot of wingbeat fundamental frequency and peak power for *Aedes* genus showing females in red and males in blue. **c** Scatter plot of wingbeat fundamental frequency and peak power for *Culex* genus showing females in red and males in blue.

Three classification tasks were performed: one genus classification (*Aedes/Culex*) and two sex classifications (male/female), one for each genus (sex of *Aedes*, sex of *Culex*). The logic of the classification process is shown in **Additional file 1: Fig S3**.

Balanced datasets, i.e., datasets that contained an equal number of samples in each class were used to make an unbiased assessment. They were obtained by randomly undersampling the classes which had a higher number of available samples.

Model performance was assessed using the accuracy metric, which is calculated by dividing the number of correct predictions by the total number of predictions. The accuracy metric is a simple evaluation metric, which makes it easy to interpret, and is appropriate when using balanced datasets.

The typical machine learning process consists of training, validation and testing. In the training phase, the model is fitted to the data with different configurations of the algorithm determined by hyperparameters, which can have a significant impact on performance. In the validation phase, the performances of the models trained with the different configurations are compared and the best one is selected. The testing phase assesses how well the model generalizes on previously unused data. A schematic overview of the training, validation and testing approach employed in this work is shown in **Additional file 1: Fig S4**.

Seventy-five percent of the recordings in each dataset were chosen randomly to create a training set for use in the training and validation phase. Training and validation were done using fourfold cross-validation, in which the training set is split into four parts of equal size and the model being optimized is trained on three of the four parts and validated on the fourth part. This process is done four times using a different part of the training set for the cross-validation in each iteration. The final cross-validation score was obtained by averaging the four cross-validation results. The model with the best crossvalidation score was then selected for testing. The remaining 25% of each dataset, i.e. that part which was not allocated to training and validation, was used to test the performance of the trained model. Since the data in the test set are completely new to the model, accuracy results for the test set are an indication of how well the model generalizes on new data, and good results cannot be attributed to overfitting of the model.

Error analysis consists of analysing the training and validation accuracies obtained during the training and validation phase. If the training accuracy is considerably higher than the validation accuracy, it indicates overfitting, so more samples could help to improve the model. If, on the other hand, training and validation accuracies have a similar low score, it indicates that the model is too simple and that more training data would probably not help. In this case, the model could possibly be improved by using a different algorithm which is able to learn more complex relationships or to use more features.

Programming was done in *Python* (Phyton software foundation, n.d.).For model generation, *scikit-learn* (Scikit-learn developers, n.d.),*TensorFlow* (TensorFlow developers, n.d.)and *XGBoost* (XGBoost developers, n.d.)were used. Regarding execution times, training of the models took days to weeks, but once done, each new sample was classified in under 1 s.

Results

Genus classification

In the genus classification, mosquitoes were classified into *Aedes* and *Culex* genus. A total of 2688 samples were used comprising: 1344 *Aedes* (672 *Ae. albopictus* and 672 *Ae.* aegypti) and 1344 *Culex* (all *Cx. pipiens*) with an equal number of males and females was used. The dataset was split 75%/25% into the training data set (2016 samples) and the test set (672 samples). The accuracy results for genus classification on the test set are shown in Table 1, with the best performing algorithm for each feature shown in bold. The best result for genus classification was obtained for the DNN algorithm trained on the spectrogram feature, with an accuracy of 94.2%.

Table 1: Genus classification accuracy results, with best results per feature in bold.

	Algorithm				
Feature	LR	GB	RF	SVM	DNN
Fundamental frequency	55.2%	**67.3%**	65.9%	65.5%	66.1%
Fundamental peak power	68.9%	**70.1%**	69.6%	69.8%	70.0%
Fundamental freq. & peak power	70.1%	77.7%	77.2%	77.2%	**77.8%**
PSD	84.8%	**92.3%**	89.0%	90.5%	90.3%
Spectrogram	90.5%	93.2%	91.2%	93.4%	**94.2%**
MFCC	89.3%	**93.2%**	90.2%	93.0%	**93.2%**

Sex classification of Aedes

In this classification, mosquitoes of the *Aedes* genus were classified into males and females. A total of 1344 samples were used, comprising 672 females and 672 males, with each sex group comprising 336 *Ae. aegypti* and 336 *Ae. albopictus*. The dataset was split 75%/25% into the training data set (1008 samples) and the test set (336 samples).The results for this classification on the test set are shown in Table 2. The best performing algorithms for sex classification of *Aedes* were logistic regression trained on spectrogram and MFCC, and gradient boosting trained on MFCC, with an accuracy of 99.4% in each case.

Table 2: Results of sex classification for *Aedes,* with best results per feature in bold.

	Algorithm				
Feature	LR	GB	RF	SVM	DNN
Fundamental frequency	95.5%	95.5%	95.5%	95.5%	95.5%
Fundamental peak power	86.9%	**89.5%**	**89.5%**	89.2%	89.3%
Fundamental freq. & peak power	98.2%	96.7%	97.0%	98.5%	**97.9%**
PSD	97.0%	**98.8%**	97.9%	**98.8%**	98.2%
Spectrogram	**99.4%**	98.8%	98.8%	99.1%	98.8%
MFCC	**99.4%**	**99.4%**	98.8%	98.8%	98.8%

Sex classification of Culex

In this classification, mosquitoes of the *Culex* genus (all *Cx. pipiens*) were separated into males and females. A total of 1560 samples were used comprising 780 females and 780 males. The dataset was split 75%/25% into the training data set (1170 samples) and the test set (390 samples). The results for this classification on the test set are shown in Table 3**.** For *Culex* sex classification, an accuracy of 100% was achieved by all five algorithms trained on MFCC; by logistic regression, SVM and DNN trained on spectrogram; and by SVM trained on PSD.

Table 3: Results of sex classification for *Culex*, with best results per feature in bold.

	Algorithm				
Feature	LR	GB	RF	SVM	DNN
Fundamental frequency	98.0%	98.0%	98.0%	98.0%	98.0%
Fundamental peak power	83.4%	81.3%	81.5%	83.1%	**83.6%**
Fundamental freq. & peak power	**98.7%**	**98.7%**	98.5%	**98.7%**	**98.7%**
PSD	99.7%	99.2%	99.2%	**100%**	99.7%
Spectrogram	**100%**	99.7%	99.7%	**100%**	**100%**
MFCC	**100%**	**100%**	**100 %**	**100%**	**100%**

Summary of the best model performance results

A summary of the classification results, which includes the best performing algorithms and features for each classification, is given in Table 4 in which training and validation accuracies are also listed, with an indication of how the results might be improved. The corresponding hyperparameters are listed in Additional file 1: Table S1.

The best accuracy results were 94.2% for genus classification, 99.4% for sex classification of *Aedes* and 100% for sex classification of *Culex*.

For genus, the training accuracy was 100% and the cross-validation accuracy was significantly lower (95%), which indicates that the model overfits slightly and its performance could possibly be improved with more training samples.

For *Aedes* sex classification, although the best models gave a near perfect accuracy, the training accuracy and cross-validation accuracy are similar (99.5%), which indicates that the model could possibly be improved with a more complex algorithms and/or features, rather than with more training samples. In case of *Culex* sex classification the accuracy was 100%, so no error analysis was necessary.

Table 4: Summary of machine learning classification results.

Classification task	**Best accuracy**	**Best feature**	**Best algorithm**	**No. of samples**	**Training accuracy**	**Cross-validation accuracy**	**Error analysis indication**
Genus	94.2%	Spectrogram	DNN	2688	100%	95%	more training samples
Sex *Aedes*	99.4%	Spectrogram MFCC	LR LR, GB	1344	99.5%	99.5%	more complex features or algorithm
Sex *Culex*	100%	PSD Spectrogram MFCC	SVM LR, SVM, DNN All	1560	100%	100%	No error.

Discussion

In the present study, 4335 mosquito flights were recorded using a novel optical sensor. The sensor was attached to the entrance of a commercial mosquito suction trap inside an insect rearing cage, with mosquitoes flying freely within the cage until they were sucked in by the trap, through the sensor and into the catch bag within the trap. Each flight recording made by the sensor corresponded to a different mosquito. Five features were extracted from each recording and used with five different machine learning algorithms for classification of mosquito genus and sex.

One of the features used was wingbeat fundamental frequency, which has been used in many studies for insect characterization and classification (Cator et al., 2011; Genoud et al., 2018, 2019; Mukundarajan et al., 2017; Potamitis & Rigakis, 2016c; Santos et al., 2019; Staunton et al., 2019). Differences in reported values of wingbeat frequency between studies can be due to intrinsic and/or extrinsic variables such as size, parity status, age and ambient conditions (Dou et al., 2021; Genoud et al., 2018; Gibson et al., 2010; Potamitis et al., 2017; Silva et al., 2015). In this study, the wingbeat fundamental frequency feature gave a high accuracy in sex classification in both Aedes (95.5%) and Culex (98%), but it scored lower (67.3%) in genus classification. These results are consistent with the fundamental frequency histograms in **Additional file 1: Fig S2**, which show very little overlap between the distributions of males and females, especially for Culex (**Additional file 1: Fig S2c**) and considerable overlap between genera (**Additional file 1: Fig S2a**). In the fundamental peak power histograms of **Additional file 1: Fig S2b, c**, a higher degree of overlap is observed between the distributions of males and females, especially for Culex, which helps explain why the accuracy for sex using this feature alone (89.5% for Aedes and 83.6% for Culex) was lower than that of fundamental frequency alone.

As other studies have indicated (Chen et al., 2014; Genoud et al., 2018; Kim et al., 2021; Wang et al., 2020), the use of the wingbeat frequency alone as a feature to differentiate between taxonomical classes or other attributes of mosquito biology can be challenging because of overlap in wingbeat frequency distributions. . To address this, other authors have used additional features (i.e., depolarization ratio) (Genoud et al., 2018)or metadata (i.e., localization, environmental variables, and circadian rhythm) (Chen et al., 2014) in combination with fundamental frequency to improve their classification methods. In the present work, we have tested several features apart from or in combination with the fundamental frequency to better classify mosquito genus and sex.

The use of both fundamental frequency and fundamental peak power, yielded better performance in sex and genus classification than fundamental frequency alone. Although the effect of signal intensity or power has been investigated in mosquito mating and courtship behavioural experiments (Dou et al., 2021; Menda et al., 2019), to the best of our knowledge, fundamental peak power has not been used as a feature in mosquito classification studies. In other sensor systems, the reported signal intensity or power may depend on the position and orientation of the flying mosquito with respect to the sensor (Arthur et al., 2014), whilst our optical setup was designed to measure wingbeat power relatively independently of the position and orientation of the mosquito within the sensing volume.

Despite the better results obtained in this work using the fundamental frequency and power features compared with fundamental frequency alone, the more complex spectrogram and MFCC features provided the best performance for genus and sex classification. MFCCs are normally used in applications such as speech recognition (Ganchev et al., 2005) or music information retrieval (Müller, 2007) and although MFCCs are based on human perception of pitch, they have given good results in sound recognition studies with mosquitoes and other insects (Lukman et al., 2017; Noda et al., 2019; Silva et al., 2015; Zhu, 2011).

The classification of mosquito genus achieved a high accuracy of 94.2% while the classification of sex achieved 99.4% and 100% for Aedes and Culex respectively. The training and validation accuracies indicate that genus classification could possibly be improved with more training samples.

In this study, the best performing machine learning algorithm depended on the classification task. For genus classification, DNN showed the best performance, with an accuracy of 94.2%, trained on the spectrogram feature. In another work (Fanioudakis et al., 2018), DNN also gave the best performance for genus classification between *Aedes* and *Culex*. For sex classification, the best performing algorithms and features were: LR with spectrogram or MFCC and GB with MFCC. Different machine learning algorithms were also compared for mosquito classification in a previous study (Genoud et al., 2020) and it was concluded that the best algorithm for complex classification tasks was SVM. In our study, SVM had an accuracy of 93.4% for genus, although DNN, which was not studied in (Genoud et al., 2020), performed slightly better (94.2%).The classification of mosquito genus achieved a high accuracy of 94.2% while the classification of sex achieved 99.4% and 100% for *Aedes* and *Culex* respectively. Learning and validation accuracies in this work indicate that genus classification could possibly be improved with more training samples.

Other studies have successfully achieved automatic classification of genus (Potamitis & Rigakis, 2016b) and sex (Genoud et al., 2018; Ouyang et al., 2015) using machine learning with relatively large datasets (Silva et al., 2015) and placing emphasis on class balance (Genoud et al., 2020). However, only a small number of sensor studies have been done using a mosquito suction trap, either without an automatic classification system (J. Wang et al., 2020) or with only mosquito and non-mosquito counting and without differentiating mosquito genus and sex (Day et al., 2020).

To our knowledge, we present the first sensor system for use with a commercial mosquito suction trap, which provides automatic classification of genus and sex with high performance, based on a large number of training samples, with class balance. further work includes the study of species classification, study of age groups, training of models with more features and feature combinations, and testing of the system in the field.

Conclusions

In this work, we have presented the results of a novel sensor system for genus and sex classification of *Aedes* and *Culex* mosquitoes captured by a commercial suction trap in laboratory conditions. The obtained results are encouraging for the use of the sensor with standard suction traps in the field, for the remote surveillance and classification of genus and sex of *Aedes* and *Culex* mosquitoes.

Acknowledgements

The authors are grateful to BSL2 facilities staff of IRTA-CReSA for their technical contribution.

Funding

This research was supported by the project VECTRACK. This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No. 853758. The contents of this publication are solely the responsibility of the authors, and the funding agency is not responsible for any use that may be made of it.

Availability of data and materials

The datasets generated during and/or analysed during the current study are not publicly available due to the protection of intellectual property defined under the H2020 agreement No. 853758, but are available from the corresponding author on reasonable request.

Authors' contributions

JE, ST and NB conceived and designed the study. JE, MW and PV designed and built the sensor. ST, NB and CA designed the entomological part of the study. JB and MG conducted the flight assays with the sensor. MG, NP and MV conducted mosquito rearing and maintenance and gave support on flight assays. BF conducted the feature extraction of flight assays and implemented the machine learning process to build the classification models. MG and BF analysed the results and drafted the manuscript. NB, MW, CA, ST and JE revised the manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Competing interests

Irideon S.L. is currently applying for a patent relating to the content of the manuscript.

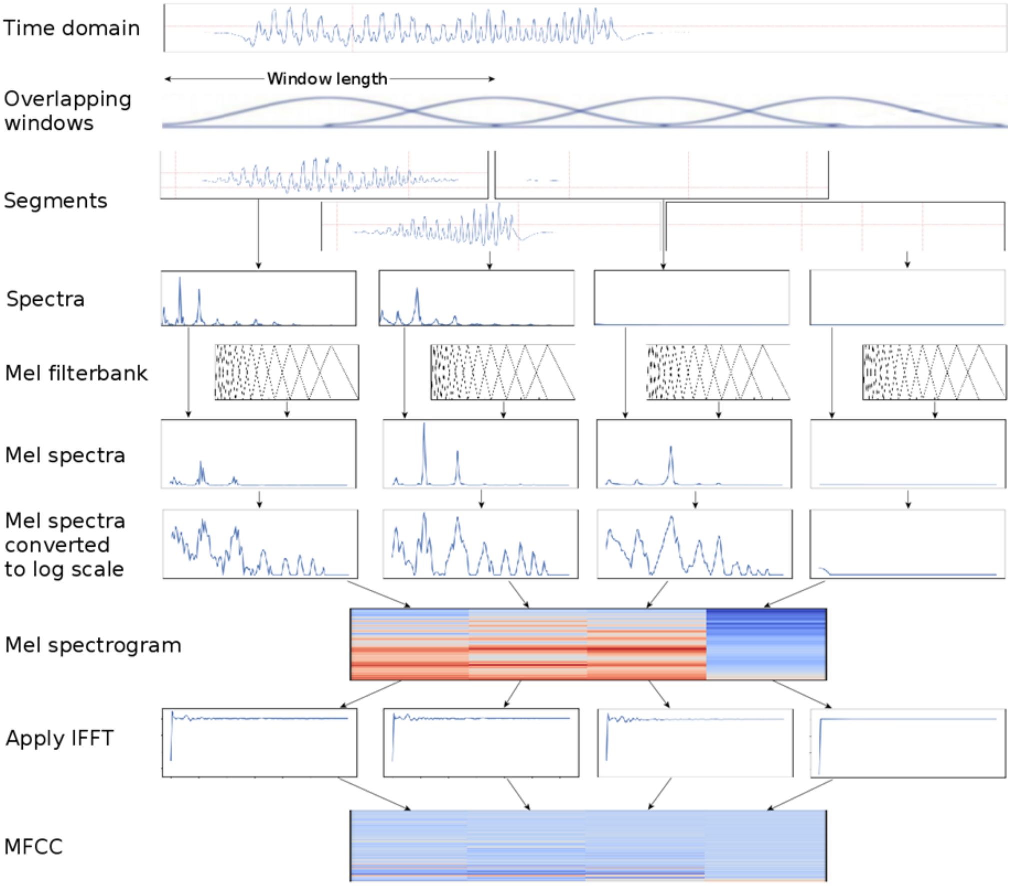
Supplementary information

Text S1. Mel spectrogram and MFCC generation process.

The Mel scale is a perceptual scale which relates sound frequency in Hertz to the pitch perceived by the human auditory system (Stevens et al., 1937). For the filter banks, the minimum frequency is taken as 0 Hz and the maximum frequency is taken as the Nyquist frequency of the sampling system, i.e., $\frac{9603\ \mathrm{Hz}}{2} = 4801.5\ \mathrm{Hz}$, and these are converted from Hertz (f) to Mel (m) using the following formula (Arias-Vergara et al., 2020): $m(f) = 1125 ln\left(1 + \frac{f}{700}\right)$. This span (in Mels) is divided into “n” bands of equal width where “n” is the number of filters applied, and the end points of each filter are converted to Hertz using the following formula: $f = 700\left(10^{\mathrm{m}/2595} - 1\right)$. The resulting values, form the base of each triangular filter in the frequency domain in which the base of each triangle is at zero and the peak is at one, giving values between zero and one along the legs of each triangle. Within each filter band, the values of the frequency spectra are multiplied by the corresponding value of the leg of the triangle and the results are then summed to give one value of the Mel spectrum. As such, the Mel spectrum consists of “n” values. Each Mel spectrum is converted to a logarithmic scale and joined to form a Mel spectrogram (Virtanen et al., 2018). An inverse Fourier transform is then applied to generate the MFCC.

The process of generating the MFCC feature is represented in **Fig S1**.

Fig S1. Diagram to illustrate MFCC generation



Text S2. Description of the machine learning algorithms used in this work.

The following provides a brief description of the machine learning algorithms employed in this work.

- LR fits a sigmoid function to the observations (predictor variables and labels) to generate a classification model (Kleinbaum & Klein, 2010; Peng et al., 2002).
- GB uses the decision trees ensemble model which consists of different classification and regression trees (CART). In GB, the ensemble building is done incrementally so that each new tree helps correct the error of the previous tree (Chen & Guestrin, 2016; Friedman, 2001). The optimized *XGBoost* library was used (XGBoost developers, n.d.).
- RF also uses the decision tree ensemble method, but it grows the different trees simultaneously from a randomly chosen subset of the predictor variables and then they vote on the outcome (Breiman, 2001). Another method of randomization used in RF is bootstrapping where a subset of the samples of the training set is randomly chosen for each tree (Lee et al., 2020)which allows for non-linear models when using one or more predictor variables.
- SVM works by plotting the training samples in an n-dimensional space (where n is the number feature values used) and generating a hyperplane or hyperline that separates the different classes. The best hyperplane is chosen by maximizing

the distance of the hyperplane to the closest samples of either class. These samples form the support vectors. For non-linear models, different kernel functions exist which transform non-linear spaces into linear-spaces (Burges, 1998).

- DNNs mimic the structure of the human brain and consist of multiple layers of relatively simple processing units that receive input and give output. The use of multiple layers and non-linear activation functions allows for non-linearity between predictor variables and outputs (Schmidhuber, 2014; Zhang, 2016).

The logic of the machine learning classification process is shown in **Fig S3**.

Fig S3. Representation of the machine learning classifications (in bold text), with their respective classes immediately below and indicated by the arrow heads.

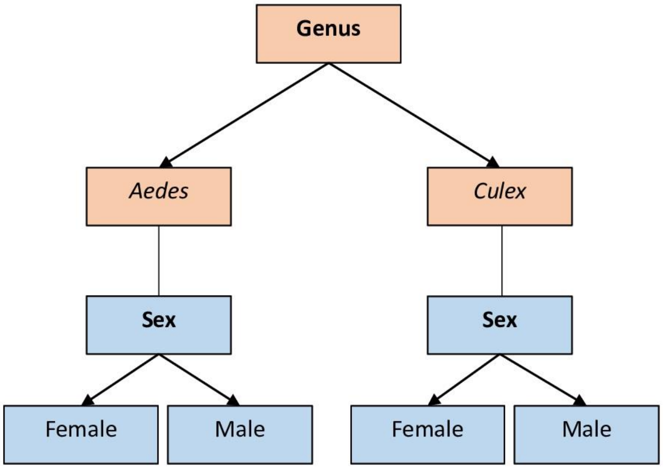


Fig S4. Schematic overview of the training, validation and testing approach. 1 Dataset is randomly separated into training and test sets, accounting for 75% and 25% of the whole dataset respectively. 2 Training set is separated using 4-fold cross-validation into four folds with an equal number of samples in each fold. 3 Four iterations of training and validation take place using a different fold for validation in each iteration. 4 Model with best average validation score, obtained by averaging the four cross-validation results, is selected. 5 Model is evaluated using test set (containing data which was previously unused) to obtain test score.

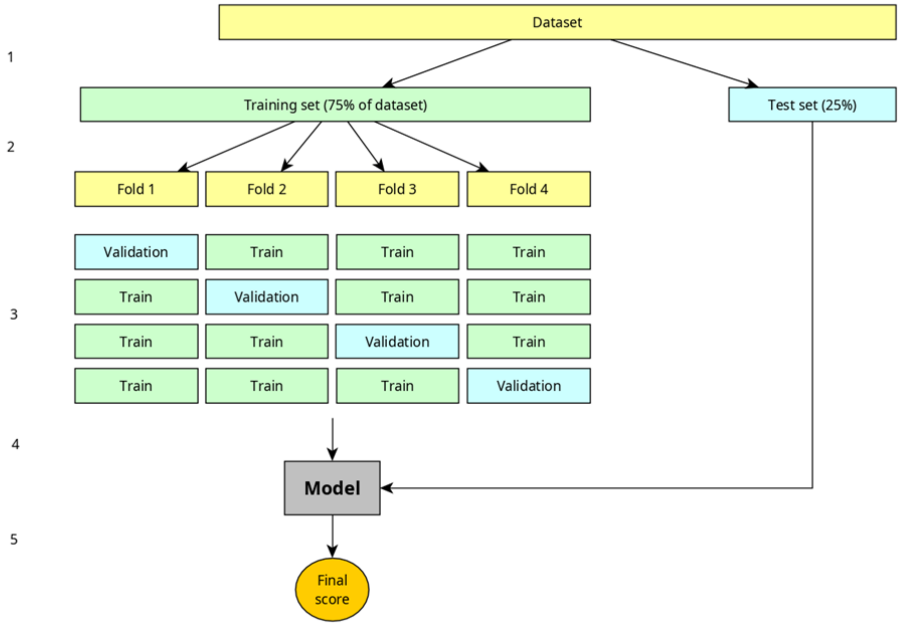


Table S1. Hyperparameters of the trained models which achieved the highest accuracies.

Classification task	**Best feature**	**Best algorithm**	**Hyperparameters**
Genus	Spectrogram	DNN	layers: 4 epochs: 1400
Sex *Aedes*	Spectrogram	LR	penalty: L2 C: 2.8
	MFCC	LR	penalty: L2, C: 1.0
		GB	estimators: 200
Sex *Culex*	PSD	SVM	kernel: RBF C: 2.8
	Spectrogram	LR	penalty: L2 C: 1.0
		SVM	kernel: RBF C: 1.0
		DNN	Layers: 3 epochs: 1500
	MFCC	LR	penalty: L2 C: 1.0
		GB	estimators: 450
		RF	estimators: 100
		SVM	kernel: RBF C: 1.0
		DNN	layers: 3 epochs:1500

4.3 CHAPTER 3: Evaluation of the sensor performance in field conditions.

Field evaluation of an automated mosquito surveillance system which classifies Aedes and Culex mosquitoes by genus and sex (submitted).

Field evaluation of an automated mosquito surveillance system which classifies *Aedes and Culex* mosquitoes by genus and sex.

María I. Gonzalez-Perez[1,2], Bastian Faulhaber[3], Carles Aranda[1,2,4], Mark Williams[3], Pancraç Villalonga[3], Manuel Silva[5], Hugo Costa Osório[5,6], Joao Encarnaçao[3], Sandra Talavera[1,2]†, Núria Busquets[1,2]†*

[1] IRTA. Programa de Sanitat Animal. Centre de Recerca en Sanitat Animal (CReSA). Campus de la Universitat Autònoma de Barcelona (UAB), Bellaterra, Spain.

[2] Unitat mixta d'Investigació IRTA-UAB en Sanitat Animal. Centre de Recerca en Sanitat Animal (CReSA). Campus de la Universitat Autònoma de Barcelona (UAB), Bellaterra, Spain.

[3] Irideon S.L., Barcelona, Spain.

[4] Servei de Control de Mosquits del Consell Comarcal del Baix Llobregat, El Prat de Llobregat, Spain.

[5] National Institute of Health/ Centre for Vectors and Infectious Diseases Research, Águas de Moura, Portugal.

[6] Instituto de Saúde Ambiental, Faculdade de Medicina, Universidade de Lisboa, Lisboa, Portugal.

*Correspondence: nuria.busquets@irta.cat

†Sandra Talavera and Núria Busquets contributed equally to this work.

Abstract

Background

Mosquito-borne diseases are a major concern for public and veterinary health authorities, highlighting the importance of effective vector surveillance and control programs. Traditional surveillance methods are labour-intensive, and do not provide high temporal resolution which may hinder a full assessment of the risk of mosquito-borne pathogen transmission. Emerging technologies for automated remote mosquito monitoring have the potential to address these limitations, however, very few studies have tested the performance of such systems in the field.

Methods

In the present work, an entomological optical sensor coupled to the entrance of a standard mosquito suction trap was used to record 14,067 mosquito flights of *Aedes and Culex* genus at four temperature regimes in the laboratory, and the resulting dataset was used to train a machine learning (ML) model. The trap, sensor, and ML model, which form the core of an automated mosquito surveillance system, was tested in the field for two classification purposes: to discriminate *Aedes* and *Culex* mosquitoes from other insects that enter the trap and to classify the target mosquitoes by genus and sex. The field performance of the system was assessed using balanced accuracy and regression metrics by comparing the classifications made by the system with those made by the manual inspection of the trap.

Results

The field system discriminated the target mosquitoes (*Aedes* and *Culex* genus) with a balanced accuracy of 95.5% and classified the genus and sex of those mosquitoes with a balanced accuracy of 88.8%. An analysis of the daily and seasonal temporal dynamics of *Aedes* and *Culex* mosquito populations was also performed using the time-stamped classifications from the system.

Conclusions

This study report results for automated mosquito genus and sex classification using an entomological sensor coupled to a mosquito trap in the field with high balance accuracy. The compatibility of the sensor with commercial mosquito traps enables the sensor to be integrated into conventional mosquito surveillance methods to provide accurate, high temporal resolution, automatic monitoring of *Aedes* and *Culex* mosquitoes, two of the most concerning genera in terms of arbovirus transmission.

Keywords

Aedes; Automated classification; *Culex*; Field study; Machine learning; Mosquito surveillance; Optical sensor.

Background

Mosquitoes (Diptera: Culicidae) act as vectors of several pathogens such as malaria parasite, dengue (DENV), Zika (ZIKV), yellow fever (YFV), chikungunya (CHIKV) and West Nile (WNV) viruses that cause diseases which result in hundreds of thousands of human deaths per year worldwide, primarily in tropical countries of Africa, East and South-East Asia, and South America (WHO, 2014). In Europe, autochthonous vector species such as *Culex pipiens,* and invasive vector species such as *Aedes albopictus,* are responsible for the transmission of endemic (e.g., WNV, Usutu, Sindbis, Tahyna and Batai viruses, lymphatic filariasis and avian malaria), and imported (CHIKV, DENV and ZIKV) pathogens respectively, and pose a threat to public and veterinary health in the continent (Calzolari, 2016). To mitigate the impact of mosquito borne diseases (MBD), surveillance programs for both native and invasive species are used by public health organizations worldwide to monitor trends in vector populations and to assess the effectiveness of control programs(ECDC, 2012; Schaffner et al., 2014). The availability of high-quality surveillance data is essential for these tasks and to model the risk of MBD (Caputo & Manica, 2020).

Traditional entomological methods for mosquito monitoring generally entail the use of physical traps, which primarily target adult mosquitoes as a proxy for pathogen transmission risk (Silver, 2008). These methods are very costly in terms of the human resources involved in the tasks of sample collection in the field, taxonomical identification of the samples, and data processing. Furthermore, the time lag between the time of capture and the analysis of the samples and processing of the results may hinder a full understanding of the real-time dynamics of mosquito populations. This delay can limit the proper assessment of disease transmission risk and the timely application of control measures. Consequently, the application of new technologies, including machine learning (ML), to the automated and remote real-time characterization of mosquito populations may have a positive impact in the state of the art in entomological surveillance (Joshi & Miller, 2021; Santos et al., 2019).

Over recent years, there has been an increasing number of studies aimed at taxonomically classifying mosquitoes and other attributes of mosquito biology using either acoustic (Mukundarajan et al., 2017; Sinka et al., 2021; Su Yin et al., n.d.; Vasconcelos et al., 2020) or optical sensors (Fanioudakis et al., 2018; Genoud et al., 2018, 2019, 2020; González-Pérez et al., 2022; Potamitis & Rigakis, 2016a; Silva et al., 2015) which take advantage of insect bioacoustic properties. The study of these properties, especially the mosquito flight tone or wing beat frequency, has been used for mosquito characterization and classification purposes since the 1940's (Kahn et al., 1945; Offenhauser & Kahn, 1949). However, the existence of overlapping frequency distributions among different mosquito species (Chen et al., 2014; Genoud et al., 2018) led to the exploration of other predictor variables such as spectrograms, power spectral density, Mel frequency cepstral coefficients or optical depolarization ratio, which provide better classification results (Fanioudakis et al., 2018; Genoud et al., 2020; González-Pérez et al., 2022). In addition to the choice of features, the choice of ML algorithm and its configuration parameters has been shown to contribute to the overall classification accuracy (Fanioudakis et al., 2018; Genoud et al., 2020; González-Pérez et al., 2022).

Despite the growth of research in automated remote mosquito surveillance (Joshi & Miller, 2021), there are very few published papers which describe the evaluation of solutions in the field (Day et al., 2020; Lai et al., 2022; Mukundarajan et al., 2017). Technical constraints such as interference from ambient noise in the case of acoustic sensors (Chen et al., 2014); the presence of heavy rain during the sampling period (Geier, Weber, Rose, Obermayr, et al., 2016); the proportion of mosquitoes relative to other flying insects in the capture (Day et al., 2020); the capture efficiency of the sampling devices (Lai et al., 2022) or the ambient environmental temperature which is known to affect mosquito flight tone (Villarreal et al., 2017)may limit the usage of these systems for field monitoring of mosquito populations. The only example of a commercial mosquito sensor, with reported results, is the BG-Counter (Biogents, Regensberg, Germany) (Geier, Weber, Rose, Obermayr, et al., 2016) which is claimed to distinguish mosquitoes from other insect species, and whose performance was shown to have a high rate of misclassifications when the proportion of non-mosquitoes was significant (Day et al., 2020).

In this contribution, we present the results of a field study of an automated mosquito surveillance system, in which an entomological optical sensor coupled to the entrance of a standard mosquito suction trap automatically differentiates target mosquitoes (*Aedes* and *Culex*) from other insects that enter the trap; and identifies the genus and sex of these target mosquitoes. We previously reported high levels of accuracy for genus and sex classification of *Aedes* and *Culex* mosquitoes in the laboratory using the same technology (González-Pérez et al., 2022). In the current study, a new ML dataset was built with 14,067 mosquito flights in the laboratory, corresponding to a wider range of larval density and ambient temperature conditions to cover the morphological variability and ambient temperature range of the target genera in nature. A new ML model was trained using this dataset. The sensor and trap were deployed and assessed in the field during periods of mosquito activity at two different locations in a Mediterranean climate area with a predominance of *Cx. pipiens* and *Ae, albopictu*s, potential vectors of imported and endemic arboviruses.

Material and methods

Entomological optical sensor

The optical sensor (Irideon, Barcelona, Spain) comprises a light grey waterproof enclosure (W25.5 x D18 x H13 cm), with a black inlet tube of 10 cm diameter at the top of the unit, and a light grey outlet tube on the underside. The sensor contains an optical emitter, comprising a rectangular array of 940 nm wavelength light emitting diodes (LEDs) which together, emit a collimated beam (W10.5 x H7 cm) of near infrared light towards an optical receiver formed by a corresponding array of photodiodes. The emitter and receiver face each other through a transparent circular tube of 10.5 cm diameter which traverses the enclosure from top to bottom, to create a sensing zone with a volume of 600 cm^3. The sensor was placed on the entrance of a BG-Mosquitaire mosquito trap (Biogents, Regensburg, Germany). The trap contains a suction fan, a removable catch bag, and a flap valve which automatically opens when the fan is powered. The suction fan creates a downward flow of air through the inlet tube of the sensor and into the trap. When an insect flies close to opening of the inlet tube, it is likely to be sucked into the tube, and down through the sensing zone, and through the flap valve, and into the catch bag. As the flying insect passes through the sensing zone, it casts a fast-changing shadow upon the optical receiver due to the modulation of the light beam by the wing flap of the insect, and this signal is recorded by the sensor. Two cables exit the sensor: one is connected to a 12 VDC power supply, such as the supply included with the BG Mosquitaire trap, and the other is connected to the trap to power the fan.

Two variants of the sensor were used in the present work: a laboratory version , which was used to record mosquito flights in the laboratory to build the ML dataset; and a field version, which was used to record mosquito flights in the field for automated mosquito classification using the ML model previously trained with the laboratory data. The two variants differ only by their method of data communication, which was via USB to a laptop computer for lab use, or wireless communication via the mobile phone network to a server for field use. Further detail about the sensor was reported in our previous work (González-Pérez et al., 2022).

Mosquito rearing conditions for the creation of the dataset

Two populations of *Ae. albopictus* and one population of *Cx. pipiens* were reared in the laboratory from immature stages (eggs and larvae respectively) collected in the field: *Ae. Albopictus*, population of Rubí (2020), Barcelona, Spain (41.50674, 2.00778); *Ae. albopictus*, population of Vilamoura (2020-2022), Algarve, Portugal (37.08546, 8.11929); and *Cx. pipiens*, population of Bellaterra (2020, 2022), Cerdanyola del Vallés, Barcelona, Spain (41.49903, 2.10872). The mosquito strains obtained in Barcelona were reared in the insectarium facilities of IRTA-CReSA (Campus of the Autonomous University of Barcelona, Cerdanyola del Vallès, Barcelona, Spain). The *Ae. albopictus* strain obtained in Portugal was reared in the insectarium facilities of CEVDI/INSA (Águas de Moura, Setúbal, Portugal).

Larvae were maintained in plastic trays with two larval density regimes (50 and 250 larvae/tray) in 750 mL of dechlorinated tap water, renewed three times per week, and fed with fish food pellets (Goldfish Sticks-TETRA, Melle, Germany). Pupae were placed in plastic cups inside insect rearing cages with dimensions of 30 x 30 x 30 cm (BugDorm-1 Insect Rearing Cage, MegaView Science, Talchung, Taiwan). Adults were fed with 10% sucrose solution *ad libitum,* which was removed 24 hours before the flight assays of the females to increase their appetite and host seeking activity and likelihood of entering the trap. All females used in the experiment were nulliparous and their age ranged from 2-16 days. The age of the males ranged from 2-9 days.

Each development stage of the mosquito life cycle took place inside a climatic chamber at controlled environmental conditions of: 28°C temperature; 80% relative humidity; and a light:dark photoperiod of 12:12 hours for *Ae. albopictus* and 11:11 hours (plus 1 hour of dusk and 1 hour of dawn) for *Cx. pipiens*. All mosquito colonies were maintained until a maximum of 15 generations to minimize any changes to flight characteristics due to the adaptation of wild populations to prolonged confinement.

Flight assays in the laboratory and training of the machine learning model

The sensor and trap were placed in an insect cage (BugDorm-4S4590 W47.5 × D47.5 × H93.0 cm, MegaView Science, Talchung, Taiwan) inside a climatic chamber. The trap was fitted with a sachet of BG-Sweetscent chemical attractant (Biogents, Regensberg, Germany) to attract mosquitoes towards the sensor and trap. Flight assays were performed at different temperatures to cover the range of temperature at which the assayed mosquito species are known to have flight activity: ~15-35°C for *Aedes* (Reinhold et al., 2018)and ~15-30°C for *Culex* (Ruybal et al., 2016).

Before the flight assays, mosquitoes were anesthetized using carbon dioxide and separated in small cardboard boxes sorted by genus (*Aedes or Culex*) and sex (female or male). All mosquitoes were held in the climatic chamber at the designated ambient temperature for 24 hours prior to the start of the assay to acclimatize them. They were then released into the insect cage containing the sensor in batches of 25 individuals every 15 minutes. *Ae. albopictus* and *Cx. pipiens* were assayed at 18°C, 23°C and 28°C in the facilities of IRTA-CReSA (in a climatic chamber: CCK-0/5930m, Dycometal, Barcelona, Spain). *Ae. albopictus* was also assayed at 33°C in the facilities of CEVDI/INSA (in a climatic chamber: FITOCLIMA S600, Aralab, Rio de Mouro, Portugal). They were released at 30 cm from the entrance of the sensor to ensure that they could fly freely before being sucked into the sensor, and to minimize the possibility of multiple insects passing through the sensor at the same time. Mosquitoes that did not enter the trap during the assay were removed from the insect cage with an electronic entomological aspirator (IA-INSECT02USB, Infoagro Systems, Madrid, Spain). After each flight assay, the catch bag was collected and the specimens inside were frozen and then counted.

After each laboratory assay, the recordings were downloaded from the sensor to a laptop computer and then processed using a Python script to produce playable and viewable audio files. Each recording was examined manually, and those considered to be invalid were excluded from the dataset, e.g., recordings containing double flights or those where the mosquito was deemed to have hit the wall of the flight tube inside the sensor. A machine learning model was generated using the methodology described in our previous work (González-Pérez et al., 2022).The gradient boosting algorithm using the XGBoost library (XGBoost developers, n.d.) was trained with 4-fold cross-validation on the extracted spectrograms of a balanced sub-dataset. A test set was previously separated from the dataset to evaluate the trained model on unused data.

Field trial of the automated mosquito surveillance system

The sensor and trap were deployed in the municipalities of El Prat de Llobregat in 2021 (field trial 1) and Rubí in 2022 (field trial 2), in the province of Barcelona (Catalunya, Spain). These locations have a Mediterranean climate, typified by hot dry summers, mild rainy winters and variable temperatures in autumn and spring. The specific location of the sensor and trap at each site (**Fig 1**) was selected to provide shade, nearby vegetation, shelter from rain and wind, and access to electrical power, in a place where mosquitoes are known to be present.

The field trials were performed in the months of peak mosquito activity. Field trial 1 ran from July to October 2021 and used collection cycles (the time between catch bag placement and sample collection) of 24 hours. Field trial 2 ran from June to September 2022 and used collection cycles of 24 to 72 hours. In each trial, dry ice was used as a source of carbon dioxide to attract mosquitoes to the trap. Samples were frozen at -20°C shortly after collection and inspected by a trained entomologist to taxonomically classify and count the content.

The flying insects that entered the trap were automatically detected and recorded by the sensor. Each recording includes the sensor GPS coordinates; and the date and time (time stamp), measured ambient temperature and relative humidity at the time of capture. The field sensor sent batches of sensor recordings via the mobile phone network to the server every 30 minutes, where classification was done by the ML pipeline. The mosquito classification results, with the associated capture time stamps and environmental data were downloaded from the server as .csv file and used for the analysis in this work.

Data analysis of sensor classification in the field

Two main functions of the mosquito surveillance system were assessed: i) target mosquito detection, i.e., the ability of the system to discriminate *Aedes* and *Culex* target mosquitoes from non-target insects which also enter the trap; and ii) mosquito genus and sex classification, i.e., the ability of the system to correctly classify the *Aedes* female, *Aedes* male, *Culex* female and *Culex* male classes.

The relationship between the sensor count (mosquitoes counted by the sensor) and the manual count (mosquitoes counted by manual inspection) was assessed by correlation analysis and linear regression analysis; and was visualized using a time series plot and a scatter plot of manual count versus sensor count per collection cycle. Pearson correlation coefficient (r) and p-value for significance were obtained to analyze how both variables were related. Regression coefficients, i.e., the R^2 coefficient of determination and the linear slope and intercept, were calculated to indicate how well the regression predictions based on sensor count approximated the manual count. A regression slope of greater/less than one, would indicate that overall, sensor counts were greater/lower than the manual counts.

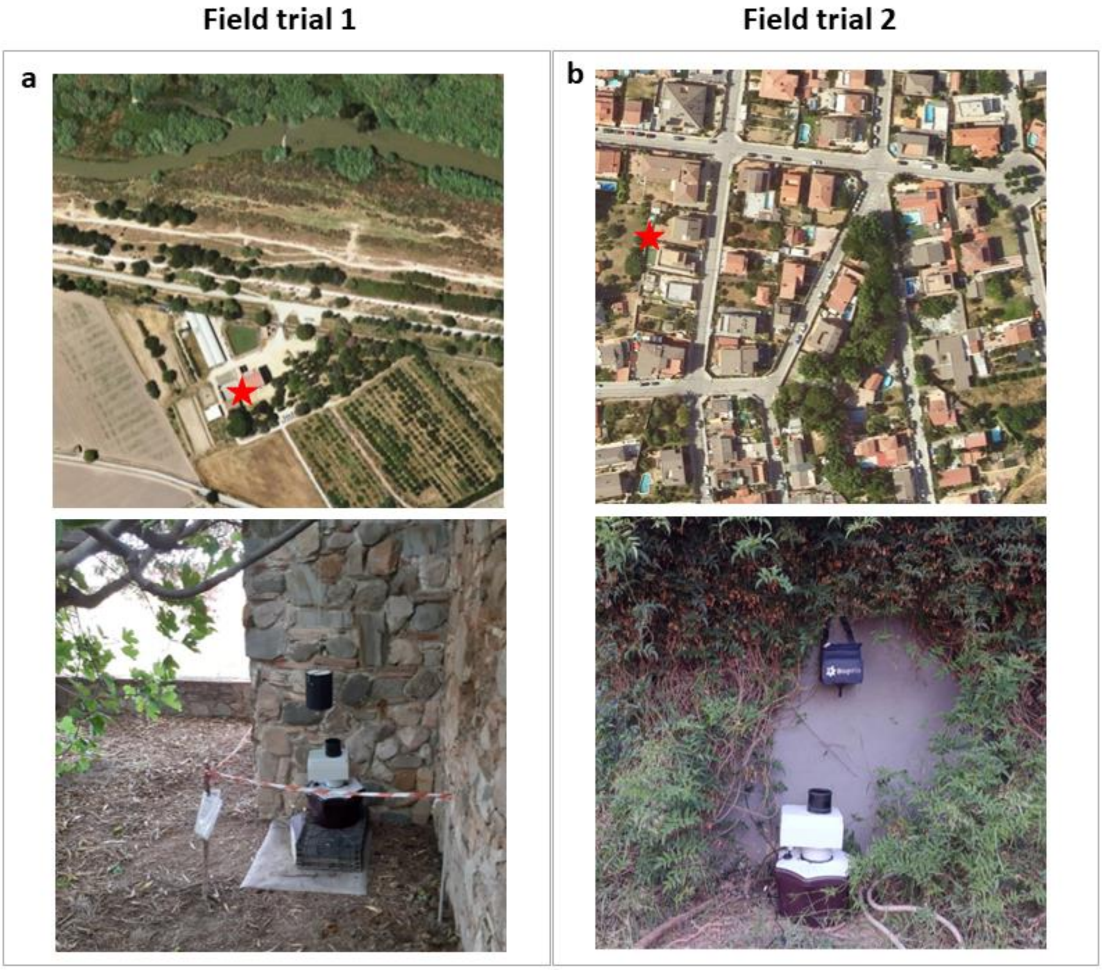


Fig 1. a. Field trial 1: The sensor and trap were installed near *Can Comas*, a 19th century country house located in the Baix Llobregat Agrarian Park in the municipality of El Prat de Llobregat, Barcelona, Spain (41.341286, 2.078259). The Park is a protected natural and rural space located in the alluvial plains of the delta and the lower basin of the Llobregat river. Land use in the area includes rainfed and irrigated agricultural crops (mainly fruit and vegetables); livestock (primarily sheep); and the Barcelona-El Prat international airport. **b.** Field trial 2: The sensor and trap were installed in the backyard of a private house in a residential area of Rubí, Barcelona, Spain (41.472816, 2.032258). The area is a typical peri-urban area, comprising detached houses with a garden or small sparsely planted orchard, some green areas with playgrounds and sport zones, roads, and services such as petrol stations and supermarkets. The neighborhood is bounded by two creeks which are tributaries of the Llobregat river.

A further evaluation metric used in this work was balanced accuracy (BA). This metric was determined by using the number of True Positives (TP), True Negatives (TN), False Positives (FP) and False Negatives (FN), using the manual classification as the reference. TP and TN are the numbers of positive and negative cases respectively that the system classified correctly. FP is the number of negative cases that the system incorrectly classified as positive, and FN is the number of positive cases that the system incorrectly classified as negative. To calculate TP, TN, FP and FN for a particular class, this class was defined as the positive class and the other class(es) were defined as the negatives. For the positive class, TP equals the minimum common value of the sensor and manual count. If the sensor count was greater than the manual count then the difference was taken as FP, otherwise FP equals zero. If the sensor count was less than the manual count, then the difference was taken as FN. TN is calculated by subtracting FP from the manual counts for the negatives. BA gives equal weighting to the proportion of positives and negatives that are correctly classified, and it is appropriate when classes are imbalanced (Brodersen et al., 2010), as is the case for the field captures in this work. The equation for BA is $BA = \frac{Se+Sp}{2}$ (where *Se* refers to sensitivity and *Sp* refers to specificity). Sensitivity, also known as true positive rate or recall, indicates the proportion of positives that are correctly classified by the system ($Se = \frac{TP}{TP+FN}$). Specificity, also known as true negative rate, indicates the proportion of negatives that are correctly classified by the system ($Sp = \frac{TN}{TN+FP}$).

The daily and seasonal temporal dynamics of *Aedes* and *Culex* mosquitoes were also analysed by descriptive statistics using the time-stamped classification results in which the sensor genus counts per hour were averaged for each month over the length of each trial.

Results

Performance of the machine learning model using the laboratory data

A total of 15,208 mosquito flights were recorded in the laboratory, of which 7.5% were rejected in the pre-processing to yield 14,067 valid flights. The valid flights were randomly under-sampled to obtain a balanced dataset from which 1,000 flights were set aside as the test set. The trained ML model achieved an average balanced accuracy of 93.9% for the classification of *Aedes* female, *Aedes* male, *Culex* female and *Culex* male flights in the test set. The BA results per class were: 91.0% for Aedes female, 93.4% for Aedes male, 96.7% for Culex female and 94.4% for Culex male. The same ML model was then used to classify the recordings from the sensor in the field.

Manual classification of the field samples

A total of 53 samples (catch bags) were collected from the traps and underwent manual inspection: 32 in field trial 1 and 21 in field trial 2. Of these, seven showed signs of significant mosquito depredation and/or degradation and were excluded from the analysis due to the impact it would have on the manual count. A further two samples were also excluded due to mobile network connectivity issues during those collection cycles. As a result, a total of 44 samples were inspected and used in the analysis: 29 from field trial 1 and 15 from field trial 2.

In total, 3,634 mosquitoes were classified manually (1,665 in field trial 1 and 1,969 in field trial 2) comprising the following species in decreasing order of number: *Cx. pipiens*, *Ae. albopictus*, *Culiseta longiareolata, Aedes caspius* and *Coquillettidia richiardii* (Table 1). *Cx. pipiens*, *Ae. albopictus* and *Cs. longiareolata* were found in both trials, while *Ae. caspius* was only found in field trial 1 during September and early October. Only one specimen of *Cq. richardii* was found, in field trial 1

Table 1: Total number of mosquitoes and other insects by manual inspection of the samples in each field trial.

Sample composition				**Field trial 1**	**Field trial 2**
Mosquitoes	Target	*Culex*	*Culex pipiens*	1261	1387
		Aedes	*Aedes albopictus*	270	543
			Aedes caspius	39	0
	Non-target	Other genus	*Culiseta longiareolata*	94	39
			Coquilletidia richiardii	1	0
Other insects	Non-target	-	Non-culicidae insects	3188	2125

The mean proportion of target mosquitoes (*Aedes* and *Culex*) compared to total insects in the samples was 32.4% in field trial 1 and 47.1% in field trial 2 with the non-target group mostly comprising Phlebotominae, Chironomidae and a wide variety of small dipterians. The proportions of each genus and sex class within the target mosquitoes, from highest to lowest were: *Culex* female (76.6% in field trial 1 and 67.1% in field trial 2), *Aedes* female (15.0% in field trial 1 and 18.9% in field trial 2), *Aedes* male (4.6% in field trial 1 and 9.2% in field trial 2), and *Culex* male (3.8% in field trial 1 and 4.8% in field trial 2).

Automated target mosquito detection in the field

There was a strong positive correlation between the number of target mosquitoes counted by the sensor and by manual inspection in both field trials (r = 0.983, p-value = 0.000 in field trial 1 and r = 0.915, p-value = 0.000 in field trial 2). This good agreement between manual and sensor counts is shown in **Fig 2**, even when the manual count changed significantly from one collection cycle to the next due to natural conditions. Linear regression analysis indicated a good fit of the linear regression line to manual count versus sensor count ($R^2 = 0.984$, p-value = 0.000), as shown in **Fig 3a**. The linear regression equation ($y = 0.924x + 3.219$) indicated that sensor count was typically 7.6% lower than the manual count.

The average BA for target mosquito detection per collection cycle was 95.9% in field trial 1 and 94.8% in field trial 2. Since the distribution of BA was skewed towards high values, the median, interquartile range

(IQR) and first and third quartiles (Q1, Q3) are given, in Table 2. The correlation between the BA of target mosquito detection, and the proportion of target mosquitoes in each collection cycle was not significant (r= 0.009, p-value = 0.956) and this is also apparent in **Fig 3b**, i.e., the BA of target detection was not dependent on the proportion of target mosquitoes in the samples. Furthermore, the correlation between the BA of target mosquito detection and the duration of each collection cycle (24 hours in field trial 1 and 24-72 hours in field trial 2) was not significant (r = 0.033, p-value = 0.834).

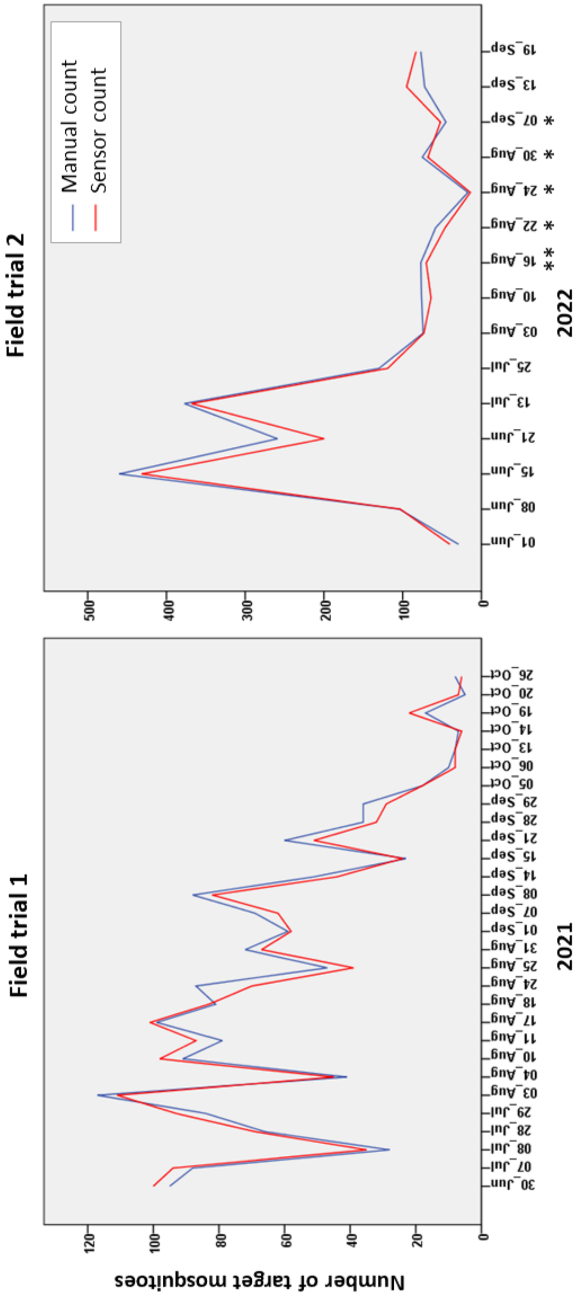


Fig 2. Time series plots representing the number of target mosquitoes (sensor count and manual count), per collection cycle for each field trial. The x-axis indicates the date of start of each collection cycle. Collection cycles lasted 24 hours in field trial 1, and 48 hours in field trial 2 except those marked with * (= 24 h) or ** (= 72 h).

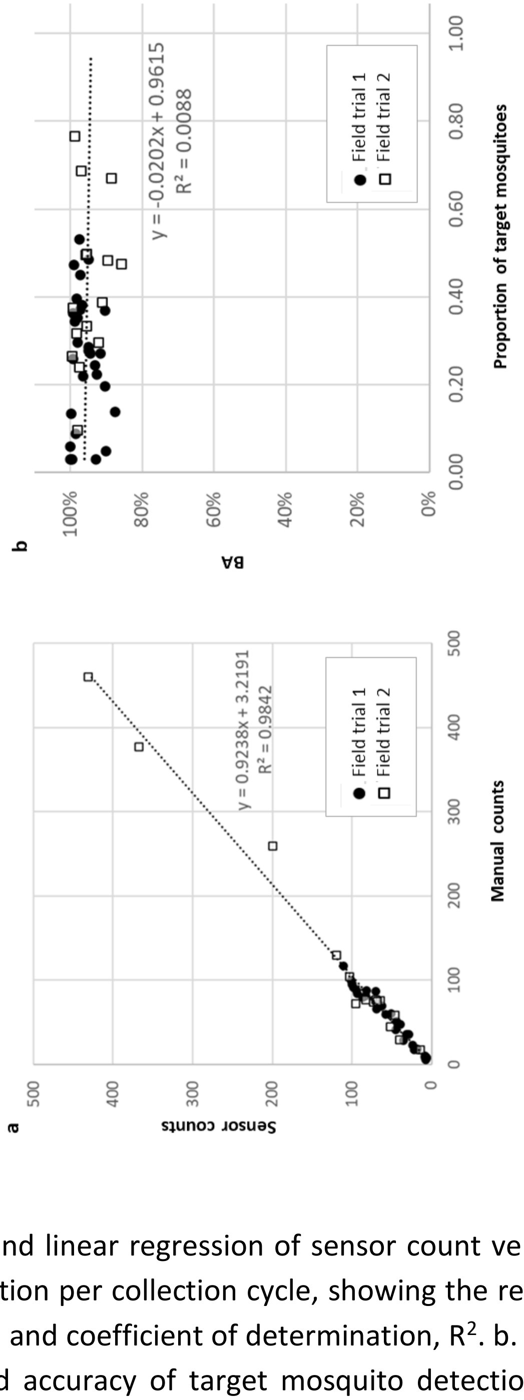


Fig 3. a. Scatter plot and linear regression of sensor count versus manual count for target mosquito detection per collection cycle, showing the regression line equation (slope and y-intercept) and coefficient of determination, R^2. b. Scatter plot and linear regression of balanced accuracy of target mosquito detection per collection cycle versus the proportion of target mosquitoes in the catch, showing the regression line equation (slope and y-intercept) and coefficient of determination, R^2.

Automated mosquito genus and sex classification in the field

There was a strong positive correlation (r = 0.846, p-value = 0.000 in field trial 1 and r = 0.903, p-value = 0.000 in field trial 2) between sensor counts and manual counts for the four mosquito classes (*Aedes* female, *Aedes* male, *Culex* female and *Culex* male) in both field trials, and this agreement can be observed in **Fig 4**. Linear regression analysis indicated a good fit linear regression line to the data points of manual counts versus sensor counts per collection cycle ($R^2 = 0.972$, p-value = 0.000) as shown in **Fig 5**. The regression equation ($y = 0.856x + 2.142$), indicates that the sensor count for genus and sex was typically 14.4 % lower than the manual count.

The BA, calculated over the collection cycles, for each genus and sex class were calculated for both field trials. In field trial 1, the BA results per class were: 88.8% for *Aedes* female, 93.7% for *Aedes* male, 88.9% for *Culex* female and 80.5% for *Culex* male. In field trial 2, the BA results were: 93.3% for *Aedes* female, 95.0% for *Aedes* male, 87.8% for *Culex* female and 85.7% for *Culex* male. The average BA of the four genus and sex classes was 88.0% in field trial 1 and 90.5% in field trial 2. Since the distribution of BA was skewed towards high values, the median, IQR, Q1 and Q3 are given, in Table 2. The BA for genus and sex was not correlated with the proportion of target mosquitoes in the samples (r = 0.153, p-value = 0.321), nor to the proportion of *Aedes* (r = 0.262, p-value = 0.086), *Culex* (r = 0.146, p-value = 0.345), females (r = -0.048, p-value = 0.756), or males (r = 0.037, p-value = 0.810) among the target mosquitoes

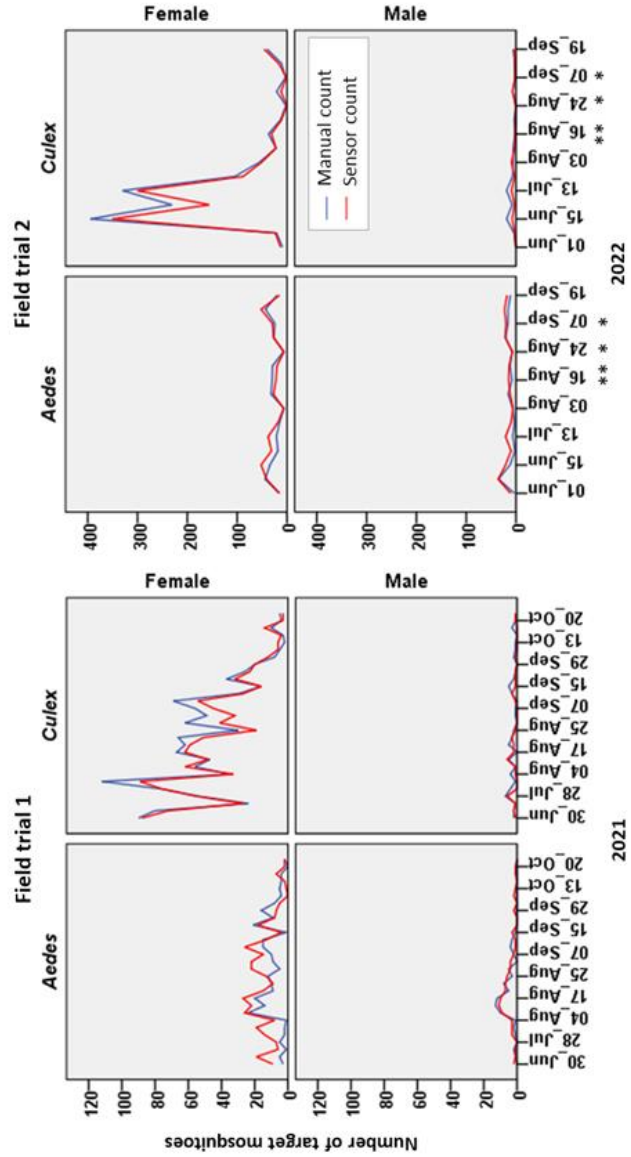


Fig 4. Time series plots representing the number of target mosquitoes (sensor count and manual count) for each genus and sex class, per collection cycle for each field trial. The x-axis indicates the date of start of each collection cycle. Collection cycles lasted 24 hours in field trial 1, and 48 hours in field trial 2 except those marked with * (= 24 h) or ** (= 72 h).

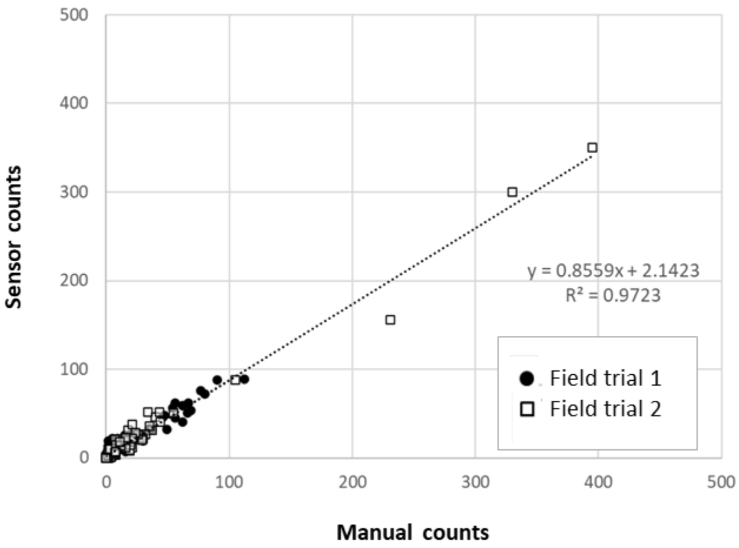


Fig 5. Scatter plot and linear regression of manual count versus sensor count for genus and sex detection per collection cycle, showing the regression line equation (slope and y-intercept) and coefficient of determination, R^2.

Table 2: Overall balanced accuracy results for target mosquito detection and for genus and sex classification for each field trial and for both field trials combined.

	Target mosquito detection			**Genus and sex classification**		
	Field trial 1	**Field trial 2**	**Field trial 1 & 2 combined**	**Field trial 1**	**Field trial 2**	**Field trial 1 & 2 combined**
Average	95.9%	94.8%	95.5%	88.0%	90.5%	88.8%
Median	97.1%	95.8%	96.7%	89.4%	90.6%	90.0%
IQR	5.7%	6.4%	5.9%	10.3%	5.4%	7.9%
Q1	93.1%	91.6%	92.8%	84.1%	87.9%	85.7%
Q3	98.8%	98.1%	98.6%	94.3%	93.3%	93.6%

Time resolution of the automated mosquito surveillance system

The daily and monthly/seasonal activity of *Aedes* and *Culex* mosquitoes for each field trial are represented in **Fig 6**. In both field trials, the peak hourly counts for *Culex* mosquitoes (*Cx. pipiens*) are higher than that of *Aedes* (mostly *Ae. albopictus* but also *Ae. caspius* during September and October in field trial 1).

There is a noticeable difference in the activity of *Culex* between the two sites. Regarding the daily activity of *Culex*, in field trial 1 there was a high and pronounced peak of activity in the morning, after sunrise, and a lower and less pronounced peak in the evening, at sunset. The morning peak was at 06:00 – 08:00 in July, which shifted to 07:00 – 08:30 in August, 07:30 – 09:00 in September, and 08:00 – 09:30 in October, but with very low counts. The evening peak was at 21:30 – 23:00 in July and August and at 20:00 – 21:00 in September. In field trial 2, *Culex* activity was apparent only during the dark photoperiod, starting around 21:00, just before sunset in the first and second months of summer (June and July) and continuing overnight until sunrise at around 07:00 although the overnight counts were much lower in late summer and early Autumn (August and September) than in June and July.

Regarding the daily activity of *Aedes*, there were generally two peaks of activity per day in both field trials: one in the early morning and one in the evening before the sunset. In field trial 1, the morning peak was at 06:00 – 07:00 in July, and 07:30 – 8:30 in August and September, and the evening peak was at 18:00 – 21:00 in August and at 19:00 in September. In field trial 2, the morning peak was at 06:00 – 07:00 in

June and July (being more pronounced in July) and at 07:00 – 08:00 and 10:00 – 11:00 in September; with an evening peak between 18:00 and 20:00 from June to September.

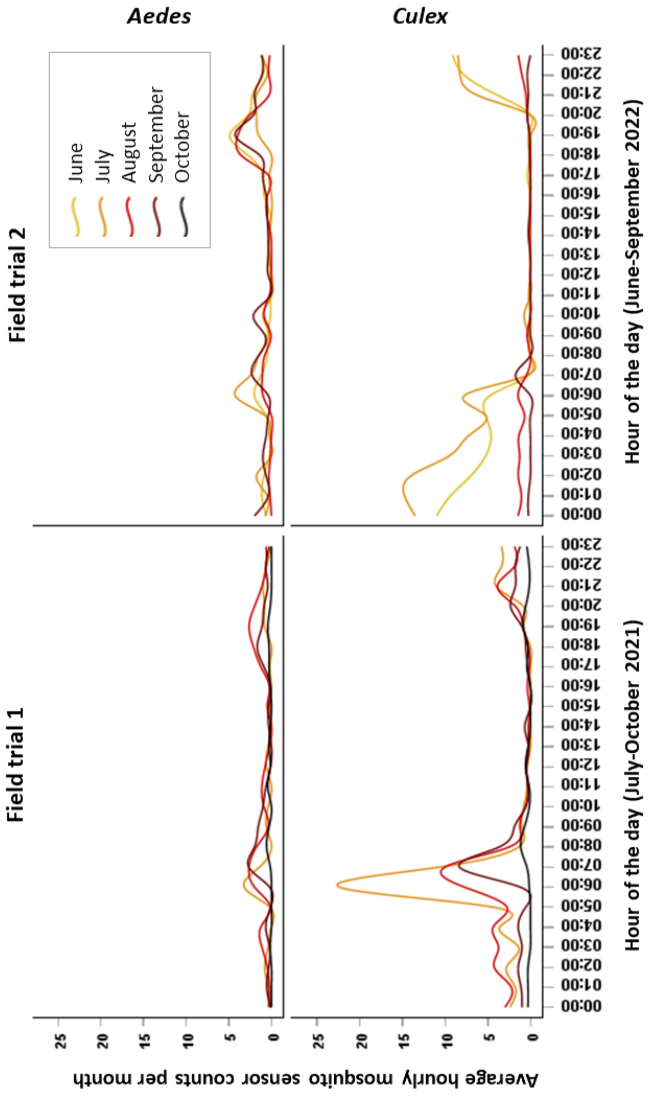


Fig 6. Time series plots of average hourly sensor count per month for the target mosquito genera.

Discussion

The present work tested the performance of an automated mosquito classification system in the field. The system comprises a commercial mosquito suction trap, an entomological optical sensor, and an ML pipeline, enabling target mosquitoes (*Aedes* and *Culex*) to be discriminated from other insects which entered the trap, and the genus and sex of the target mosquitoes to be classified. The data provided by the system also enables the real-time dynamics of the target mosquito populations to be determined with a time resolution as fine as one second. The ML model was trained using recordings from thousands of mosquitoes raised in the laboratory which flew through the sensor under different ambient temperature regimes.

The system distinguished the designated target mosquitoes (*Aedes* and *Culex*) from other flying insects that entered the trap with a BA of 95.5% for the two field trials combined, meaning that the total number of target mosquitoes counted by the sensor was very similar to the number of mosquitoes counted manually. This suggests that the system would be suitable for a range of mosquito surveillance and control activities such as: i) identifying the start and end of a mosquito activity season; ii) monitoring seasonal tendencies of endemic vectors to prioritize geographical areas of intervention; iii) doing quality control checks of control measures aimed at reducing mosquito populations; or iv) identifying the arrival of invasive mosquitoes due to changes in land use, the effects of globalization, or due to climate change.

The BA results for mosquito target detection were not correlated with the proportion of target mosquitoes in the catch. This is in contrasts with the results presented in Day et al (2020), where the mean daily accuracy of the BG-Counter sensor ranged from 9.4% to 80.1% across sites, and was highly dependent on the proportion of mosquitoes in the catch, giving high levels of accuracy only in one site when the mean daily proportion was high (89%), and leading the authors to conclude that the accuracy of the BG-Counter was “context-dependent”. In our work, the overall proportion of target mosquitoes was 39.8% and the BA for target mosquito detection was high (from 92.9% to 100%) for all collection cycles, even with a proportion of target mosquitoes as low as 3%. These results indicate that the present system performs target mosquito detection with a low rate of false positives. This result can be advantageous in routine mosquito surveillance programs in which carbon dioxide is usually substituted by a more cost-effective attractant such as a chemical lure, which leads to a lower proportion of mosquitoes in the catch (Claudel et al., 2022).

In this work, we also describe the automated classification of target mosquitoes by genus and sex in the field, which represents an advance in the state of the art. The system classified *Aedes* females, *Aedes* males, *Culex* females and *Culex* males with an average BA of 88.8%. The genus and sex classification feature can be very useful for public health agencies and biological research in order to: i) detect possible introductions of *Aedes* invasive mosquitoes in new areas; ii) evaluate the effectiveness of control strategies based on mosquito modification such as the Sterile Male Technique or *Wolbachia* infection, that target a specific mosquito genus; or iii) monitor population dynamics of *Aedes* and *Culex* mosquitoes as an indicator of their vectorial capacity for arboviruses. The fact that the BA for genus and sex in the field (88.8%) is only slightly below the result obtained in the laboratory (93.9%) indicates that the ML model, which was developed under controlled laboratory conditions, has generalized well to mosquitoes in the field and validates the methodology developed to train the model.

The daily activity patterns of *Aedes* and *Culex* mosquitoes were monitored in this study taking advantage of the high time resolution of the mosquito surveillance system. A bimodal activity coinciding with the daylight hours was identified for *Aedes* mosquitoes in both field trials as previously reported for this species [24]. *Culex* exhibit a typical endogenous night activity in field trial 2 but showed an unexpected peak just after sunrise in field trial 1. This plasticity in behaviour could be explained by factors such as host availability, environmental conditions, predator inactivity or bioform type (Fritz et al., 2014).Moreover, seasonality can have a considerable impact on vector feeding preferences which may drive the transmission of zoonotic pathogens to humans amplifying the scope of an epidemic (Kilpatrick et al., 2006). In our study, we observed a progressive shift of the early morning activity peak of *Aedes* and *Culex* to later hours over the duration of each trial from early summer to early autumn coinciding with the sunrise. Since most pathogen transmission by vectors occurs during the peaks of feeding activity, a proper characterization of daily dynamics of local vector species may be of value for risk assessment and control programs.

Overall, the system presented here provides several advantages with respect to conventional manual surveillance methods: i) it significantly reduces the manual effort to gather and inspect each catch bag, especially when target mosquitoes must be sorted from a large number of non-target insects, and to manually record the results; ii) it provides classification results much earlier than what is possible in routine monitoring programs with collection cycles of 7 to 15 days, enabling a faster epidemiological response when needed; iii) it is not subject to the effects of predation and degradation of the sample; iv) it associates a time of capture stamp to each classification result, enabling the activity dynamics of the target insects to be determined with time resolutions down to one second; v) the server provides automated results in the form of tables and graphs which may be downloaded or visualized on the server itself and may feed risks maps, via the application programming interface.

The system also provides the following advantages compared to alternative automated mosquito classification systems: i) it provides classification of target mosquito genus and sex in the field, which has not been reported in the scientific literature to date; ii) it provides good classification accuracy results for the target *Aedes* and *Culex* mosquito species over the range of ambient temperatures in which these species are known to be active, independently of the proportion of mosquitoes in the catch; and iii) it may be used with existing commercial mosquito traps used in routine entomological surveillance, allowing manual collection and inspection if needed.

Further work to improve the possibilities of the developed system in the field should include the use of additional environmental variables and biological traits when training the system, as well as addressing other vector species of interest such as *Aedes aegypti* or *Anopheles gambiae* complex to enable the system to be applied in a broader range of geographic regions with these and other vector species.

Conclusions

This work describes the field evaluation of a new entomological optical sensor which operates with commercial mosquito traps routinely used in entomological surveillance. This approach enables the sensor to be integrated into conventional mosquito surveillance methods to provide automatic, high temporal resolution monitoring of populations of *Aedes* and *Culex* mosquitoes; two of the most concerning genera in terms of arbovirus transmission. The system automatically discriminates these target mosquitoes from non-target insects in the catch and classifies the target mosquitoes according to genus and sex, which overcomes the manual effort associated with conventional methods to periodically visit the trap and to manually classify the contents of the catch. The system evaluated in the field in the present work, therefore represents a significant improvement in the state of the art of mosquito surveillance.

List of abreviatons

BA: Balanced accuracy; CHIKV: Chikungunya virus; DENV: Dengue virus; FN: False negatives; FP: False positives; MBD: Mosquito-borne disease; ML: Machine learning; Se: Sensitivity; Sp: Specificity; TN: True negatives; TP: True positives; WNV: West Nile virus; YFV: Yellow fever virus; ZIKV: Zika virus.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

This research was supported by the project VECTRACK. This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No. 853758.

Authors' contributions

MG, CA, ST and NB conceived and designed the study. MG and MS conducted mosquito rearing and maintenance and performed the flight assays in the laboratory with the support of ST and HO. BF conducted the feature extraction of flight assays and implemented the machine learning process to build the classification model. PV, MW, JE and BF developed the hardware, firmware and software components of the system. MG, CA and NB were in charge of the field studies. MG performed the taxonomical identification of samples with the support of CA and ST. MG, JE, BF and MW analysed the results. MG drafted the manuscript. MW, BF, JE, CA, ST and NB revised the manuscript. All authors read and approved the final manuscript.

Acknowledgements

The authors are grateful to BSL2 facilities staff of IRTA-CReSA for their technical support, specially to Marta Verdún and to the Service for mosquito control of Baix Llobregat for their help in trap logistics.

4.4 CHAPTER 4: Determination of the age of *Culex pipiens* mosquitoes.

Age grading of female Culex pipiens s.s mosquitoes by gene profiling and by an optical sensor system (Submitted).

Age grading of female *Culex pipiens* mosquitoes by gene profiling and by an optical sensor system.

María I. Gonzalez-Perez[1,2], Bastian Faulhaber[3], Mark Williams[3], Joao Encarnaçao[3], Pancraç Villalonga[3], Sandra Talavera[1,2], Núria Busquets[1,2*]

[1] IRTA. Programa de Sanitat Animal. Centre de Recerca en Sanitat Animal (CReSA). Campus de la Universitat Autònoma de Barcelona (UAB), Bellaterra, Catalonia, Spain.

[2] Unitat mixta d'Investigació IRTA-UAB en Sanitat Animal. Centre de Recerca en Sanitat Animal (CReSA). Campus de la Universitat Autònoma de Barcelona (UAB), Bellaterra, Catalonia, Spain.

[3] Irideon S.L., Barcelona, Spain.

*Corresponding author: nuria.busquets@irta.cat

Abstract

The age of a mosquito population is a major determinant of its vectorial capacity. To contribute to disease transmission, a competent mosquito vector, carrying a pathogen, must live longer than the extrinsic incubation period of that pathogen to be able to transmit it to a new host. As such, determining the age of wild-caught female mosquitoes is of significant interest for research in vector biology, and for vector surveillance and control programs. In this contribution, two methods were developed to classify the age of *Culex pipiens*, the primary vector of West Nile virus and other pathogens of medical and veterinary importance. The methods are: i) a transcriptional age-grading technique, based on the differential expression of age-responsive genes; and ii) an automated system comprising an optical wingbeat sensor, whose output signal is applied to a machine learning model to classify mosquitoes by age. Models for age classification were developed for each method. The gene profiling method had a mean absolute error (MAE) of 1 day for mosquitoes ≤ 4 days old and a MAE of 5 days for mosquitoes ≥ 5 days. The sensor-based method classified mosquitoes into two groups (≤ 4 days old, and ≥ 7 days old) with an accuracy of 74.7%. When integrated into vector surveillance programs, each of these methods should allow young mosquitoes (≤ 4 days old) to be discriminated from older mosquitoes which may act as vectors of arboviruses. In this work, the gene profiling method was applied to *Culex* mosquitoes for the first time, using a simpler and more cost-effective methodology than previously reported for other mosquito species. The sensor-based method is designed to enable automated age classification of flying mosquitoes in the field, to complement the mosquito genus and sex classification capability of the same system, reported in our previous work.

Author summary

Mosquitoes are a global health concern because they act as vectors of pathogens which cause serious diseases in humans and animals. The age of a female mosquito vector is a determining factor in disease transmission, since to transmit the pathogen to a host, the mosquito

must live longer than the period required for the pathogen to replicate in its gut and reach the saliva. In this paper, we assessed two methods for age grading of female *Culex pipiens* mosquitoes, which are one of the main vector species worldwide: i) a method based on the analysis of genes with age-related expression, and ii) an automated fieldable method, which includes an optical sensor system to detect the wingbeat properties of a flying mosquito, used to classify its age using machine learning. Each of these methods successfully differentiated between mosquitoes which are too young for the pathogen transmission, and older mosquitoes which may spread the pathogen. Our aim is that with further work, the methods developed in this work will be applied to vector surveillance and control programs to reduce the incidence of mosquito-borne diseases.

Background

Culex (Culex) pipiens Linnaeus, 1758, also known as the "Common house mosquito" or "Northern house mosquito" is a polytypic species that is member of a species complex with a worldwide distribution. It is native to Europe where it is the primary vector of important human and animal pathogens such as West Nile virus (WNV), Usutu virus, Sindbis virus, Tahyna virus, lympathic filariasis and avian malaria (ECDC, 2020). Females of *Cx. pipiens* feed on a variety of vertebrate hosts, contributing to the amplification of the enzootic cycle of arboviruses, such as WNV, which is mainly sustained among wild birds, and to the occasional spill-over to humans (Figuerola et al., 2022). Preventing the spread of mosquito-borne diseases (MBD) and responding to outbreaks is a priority for public health authorities, which are developing coordinated strategies to strengthen vector control programs worldwide (WHO, 2017).

The age of a mosquito population is a key determinant of its vectorial capacity (V), which indicates the efficiency of a vector to transmit a pathogen (Kramer & Ciota, 2015). To become a vector, an adult female mosquito must live longer than the combined period of non-feeding activity (around two first days after emergence) and the extrinsic incubation period (EIP) of the pathogen, which ranges from several days to a couple of weeks depending on factors such as vector population, pathogen, pathogen load and ambient temperature. For example, the EIP of WNV in *Cx. pipiens* is typically greater than 7 days at 18 to 27ºC (Anderson et al., 2008; Kilpatrick et al., 2008; Vogels, Göertz, et al., 2017a, 2017b) As such, the impact of MBD can be reduced by vector control strategies which target adult mosquitoes and adult mosquito lifespan (Cook et al.,2008).

The importance of vector survivorship as a key component of the epidemiology and control of MBD was first described in the 1960's for malaria reproduction rate and Anophelinae (Macdonald, 1957). Following this, the concept of V was introduced, incorporating vector longevity as one of the major contributors to pathogen transmission (Garrett-Jones & Grab, 1964). The V is calculated as $V = \frac{ma^2\, bp^n}{-\ln{(p)}}$ where 'm' is the vector density relative to host density; 'a' is the probability a vector feeds on a host in one day; 'b' is the competence of the vector for a particular virus, 'p' is the daily probability of the vector survival, and 'n' is the EIP in days. A linear reduction in 'p' leads to an exponential reduction in V, which highlights the impact of vector longevity on V (Johnson et al., 2020).

Several methods have been described in the literature to estimate mosquitoes' age (Johnson et al., 2020). One of the oldest and most established methods is based on the examination of changes in the ovarian morphology of female mosquitoes according to their reproductive status, such as the ovary tracheation method or the determination of sequential egg laying events (Hugo et al., 2008). More recent analytical methods rely on age-related changes in biochemical signals such as the pteridines or the cuticular hydrocarbons present in different mosquito tissues, quantified by chromatography (L. E. Hugo et al., 2006). Other recent approaches include transcriptional profiling, based on differences in the expression levels of age-responsive genes

(Cook et al., 2007), and near-infrared spectroscopy, based on changes in the absorption spectra of organic compounds in the exoskeleton (Lambert et al., 2018). Despite decades of research, only a few methods based on ovarian dissection and cuticular hydrocarbons have been developed for *Culex* (de Meillon & Khan, 1967; Kardos & Bellamy, 1961; Samarawickrema, 1967). These current available methods are either labour-intensive, require high level of expertise, high equipment and processing, or do not directly provide an estimation of chronological age (calendar days) (Johnson et al., 2020).

To address these limitations, two approaches to the age-grading of female *Cx. pipiens* mosquitoes were developed and assessed in the present work: i) a transcriptional age grading technique, based on the differential expression of age-related genes; and ii) a system, comprising an optical wingbeat sensor with machine learning (ML) to classify mosquitoes by age. Both methods were developed and assessed with a view to their future application vector biology research and in mosquito surveillance and control programs.

Material and methods

Age classification by transcriptional age-grading technique

Mosquito sampling

An established laboratory colony of *Culex pipiens*, population of Bellaterra (2020), Cerdanyola del Vallés, Barcelona, Spain (41.499035, 2.108717) was used for the experiment. Mosquitoes were maintained in insect rearing cages (BugDorm-1 Insect Rearing Cage W30 × D30 × H30 cm, MegaView Science, Talchung, Taiwan) inside a climatic chamber (Telewig, Barcelona, Spain) at 22-26ºC, 80% relative humidity and a light:dark photoperiod of 14:10 hours. They were fed with 10% of sucrose solution, with no blood meals.

Eleven age categories were used in this study: 1, 2, 3, 4, 5, 9, 13, 17, 21, 25 and 29 days old. At each time-point, female mosquitoes were aspirated from the insect rearing cage with an electronic aspirator (IA-INSECT02USB, Infoagro Systems, Madrid, Spain) and anesthetized with carbon dioxide on a plate (Flowbuddy flow regulator, 59-122BC, Flystuff, California, USA). Six females from each age category were collected and preserved individually in an Eppendorf tube containing 500 µL of TRIzol Reagent (Invitrogen, Thermo Fisher Scientific, Massachusetts, USA) with soda-lime glass beads (2 mm of diameter). Samples were homogenized at 30 Hz for 1 minute (Tissuelyser II Bead Mill Sample Disruption Preparation Unit, Qiagen, Denmark) and stored at -80ºC until analysis.

RNA extraction and cDNA synthesis

RNA extraction was performed in a laboratory fume cupboard (Köttermann, Uetze, Germany) using TRIzol Reagent and following the manufacturer's instructions. The RNA of 66 mosquito samples was extracted (6 samples of 1 individual mosquito each per time point) and resuspended in 30 µL of nuclease free water. Total RNA yield of each sample was quantified by spectrophotometric absorbance readings (BioDrop, Biochrom Ltd, Cambridge, United Kingdom). Five hundred ng of RNA were treated with 1 µL of HL-dsDNAsa (ArticZymes technologies, Tromsø, Norway) to eliminate dsDNA. Samples were incubated for 10 minutes at 37ºC and 5 minutes at 58ºC in a thermocycler (GeneAmp PCR System 9700, Applied Biosystems, Thermo Fisher Scientific, Massachusetts, USA). A mix of 1 µL of oligo (dT) (Invitrogen, Massachusetts, USA) and 1 µL of dNTPs 10 mM (BioTools, Madrid, Spain) was added to each DNase treated sample and they were first incubated for 5 minutes at 65ºC and later put in ice for at least 1 minute. Samples were then mixed with SuperScript Reverse Transcriptase III (Invitrogen, Massachusetts, USA) to generate cDNA by following the manufacturer's protocol and incubating at 50ºC for 50 minutes and 85ºC for 5 minutes. To remove RNA, the final cDNA volume of 20 µL was treated with 1 µL of RNase H (Invitrogen, Massachusetts, USA) and incubated at 37ºC for 20 minutes.

Primer design for age-related and reference genes

The reference genome for primer design was a partially sequenced genome from *Culex pipiens spp. pallens,* now completely sequenced (GenBank Ref: GCA_016801865.2). Three candidate genes were selected based on showing strong age-related expression trends in *Ae. aegypti* (Cook et al., 2006)and *Ae. albopictus* (Chathurika Weeraratne et al., 2021) In *Cx. pipiens,* these genes were found by similarity, performing BLAST searches in NCBI databases. The three candidate genes were: i) endocuticle structural glycoprotein ABD-4-like (GenBank Ref: XM_039582691.1), hereafter referred to as CUT; ii) sarcoplasmic calcium-binding protein 1 (GenBank Ref: XM_039575963.1), hereafter referred to as SCP; and cell division cycle protein 20 homolog (GenBank Ref: XM_039574874.1), hereafter referred to as CELL. One reference gene, 40S ribosomal protein S17 (RpS17) (GenBank Ref: XM_039582846.1), was used to normalize the gene expression of age-related genes, according to existing protocols for *Ae. aegypti* (Cook et al., 2007). Gene-specific primer sets (Table 1) were designed using PrimerExpress version 3 (Applied Biosystems).

Table 1. List of reference and age-related candidate genes used for age prediction.

Gene	NCBI Reference sequence (GenBank)	Putative function (Gene Ontology)	Primer sequence	Amplicon size
RpS17	XM_039582846.1	Structural ribosomal protein S17 (GO:0003735)	FW 5' CGAGAAGTACTACACCCGGC RV 5' ACGCTTCATCAGATGCGTCA	130 bp
CUT	XM_039582691.1	Endocuticle structural glycoprotein (GO:0008010)	FW 5' TTGACAAGACGAGCGGTGAA RV 5' GCGGTCACCAGAAGTAGCTC	129 bp
SCP	XM_039575963.1	Sarcoplasmic calcium-binding protein (GO:0005509)	FW 5' GTCTTCGACGGGAGTTGGTT RV 5' TTGTCCAGGAAGCCGTTGTT	147 bp
CELL	XM_039574874.1	Cell cycle/physiology (GO:0010997, GO:1990757)	FW 5' GCGGYAGAACGACCAGAACA RV 5' TTCGCCATCCAGCGTMATGA	144 bp

Determination of transcript abundance using qPCR

Transcript abundance of the three candidate genes (CUT, SCP and CELL) and the RpS17 housekeeping gene was determined for individual mosquitoes. Quantitative real-time polymerase chain reaction (qPCR) was performed on the Applied Biosystems™ 7500 Fast Real-Time PCR Systems (Applied Biosystems, Thermo Fisher Scientific, Massachusetts, USA) by using SYBR Green (Thermofisher) following the manufacturer's protocols. Each reaction mixture (20 µl) was prepared using 2.5 µl cDNA (diluted to 1/10), 7.5 µl Fast SyBr Green master mix, 2.5 µl of each primer (10 mM), and 5 µl nuclease-free water. The thermal cycling conditions of the 7500 Fast system were 95ºC for 20 sec for enzyme activation, followed by 40 cycles of 3 s at 95°C for melting and 30 s at 60°C, for annealing and extension. Reaction specificity was assessed by melting curve analysis and later confirmed by primer sequencing. Cycle threshold (Ct) values were obtained from amplification curves as a measure of transcript abundance. The efficacy of the different qPCRs assayed was assessed by performing five 10-fold serial dilutions.

Age prediction model

Fold change (FC) of age-responsive genes at each time point was calculated, relative to the 1-day-old mosquitoes using the $2^{-\Delta\Delta CT}$ method (Livak & Schmittgen, 2001). Relative changes in gene expression for each individual mosquito were calculated as log contrasts (LogC) following current validated protocols (Cook et al., 2007). To determine any significant variation in the expression (LogC) of the candidate genes across age categories, Kruskal-Wallis tests and pairwise Wilcoxon tests with Benjamini-Hochberg procedure were performed.

To calibrate age prediction models, regression analyses were performed with age as explicative variable and LogC of the candidate genes as the response variable. As a measure of goodness of fit, the coefficient of determination was calculated: using adjusted R^2 for linear regression; and pseudo R^2 for non-linear regression (*R: Pseudo R2 Statistics*, n.d.).

To estimate mosquito age from LogC values of age-responsive genes, a predictive model was developed by reverse regression, in this case with LogC as the predictor variable and age as the response. First, a polynomial regression was performed to determine the linear and/or polynomial components of age-responsive genes that are significant for age prediction. Then, a regression tree was built on the selected candidate genes to set up significant cut-off points for age classification. To assess the fit of the segmented regression, a generalised least square model with a variance weighting function was fitted. Finally, a leave one out cross-validation procedure was used to estimate the performance of the model on the dataset, i.e., how well predictions made by the model (predicted age) match the observed data (actual age). Age estimations were represented with 95% confidence intervals and the mean point was presented as the likely estimated age for a given actual age category. The difference between the predicted and the actual age, i.e., the mean absolute error (MAE), was used to assess the accuracy of the model.

All analyses were performed using RStudio (R version 4.3.1).

Age classification by the optical sensor system

Mosquito rearing conditions

One colony of *Cx. pipiens*, population of Bellaterra (2020, 2022), Cerdanyola del Vallés, Barcelona, Spain (41.49903, 2.10871), was reared in the laboratory from larvae collected in the field. Larvae were maintained in plastic trays with two larval density regimes (50 and 250 larvae/tray) in 750 mL of dechlorinated tap water, renewed three times per week, and fed with fish food pellets (Goldfish Sticks-TETRA, Melle, Germany). Pupae were placed in plastic cups inside insect rearing cages (BugDorm-1 Insect Rearing Cage W30 × D30 × H30 cm, MegaView Science, Talchung, Taiwan) until adult emergence. Adult female mosquitoes were anesthetised with carbon dioxide in a plate (Flowbuddy flow regulator, 59-122BC, Flystuff, California, USA) and sorted by age in three groups: 2-4 days, 7-9 days, and 14-16 days of age. They were fed with 10% sucrose solution *ad libitum,* which was removed 24 hours before the flight assays.

The mosquito life cycle took place inside a climatic chamber (Telewig, Barcelona, Spain) at 28ºC; 80% relative humidity; and with a light:dark photoperiod of 11:11 hours (plus 1 hour of dusk and 1 hour of dawn). Colonies were maintained until a maximum of 15 generations to minimize any changes in their flight characteristics due to prolonged confinement. All female mosquitoes used in the experiment were nulliparous.

Flight assays using the optical sensor and trap

An optical wingbeat sensor (Irideon, Barcelona, Spain) was coupled to the entrance of a BG-Mosquitaire trap (Biogents, Regensburg, Germany) which contains a suction fan. The sensor comprises an optical emitter formed by a two-dimensional array of light emitting diodes which emits a collimated beam of light (940 nm) towards an optical receiver formed by a two-dimensional array of photodiodes, with a sensing zone formed between them. When a mosquito flies close to the entrance of the sensor, it is likely to be sucked into the sensor by the airflow of the fan and pass through the sensing zone where it casts a fast-changing shadow upon the optical receiver due to the modulation of the light beam by the wingbeat of the mosquito in flight. A detailed description of the sensor is provided in our previous work (González-Pérez et al., 2022). The sensor and trap were placed in an insect cage (BugDorm-4S4590 W47.5 × D47.5 × H93.0 cm, MegaView Science, Talchung, Taiwan) inside a climatic chamber (CCK-0/5930m, Dycometal, Barcelona, Spain), where the flight assays took place. The trap was fitted with a sachet of BG-Sweetscent (Biogents, Regensberg, Germany) to attract mosquitoes towards the sensor. During each flight assay, female mosquitoes belonging to a particular age group were released into the insect cage. Flight assays were performed at ambient temperatures of 18ºC, 23 ºC and 28ºC, and mosquitoes were acclimatized in the climatic chamber for 24 hours prior to the start of the assay.

Machine learning model for age classification

The sensor recordings were downloaded to a laptop computer after each flight assay and were then processed using a Python script to produce playable and viewable audio files. Pre-processing of data included the manual examination of each recording and the exclusion of those considered to be unrepresentative, e.g., single recordings with two mosquitoes or recordings where the mosquito was deemed to have hit the wall of the transparent flight tube inside the sensor. The resulting labelled data was randomly under-sampled to obtain a balanced dataset which was split, with 75% used to train the supervised ML model, with 4-fold cross validation; and the remaining 25% used to test the classification performance of the model. Spectrogram, power spectral density (PSD) and Mel Frequency Cepstral Coefficients (MFCC) features were extracted from the sensor recordings and were used to train the ML model with the XGBoost gradient boosting algorithm(Chen & Guestrin, 2016) using the methodology described in our previous work (González-Pérez et al., 2022). The performance of the ML model was assessed using the accuracy metric, which was calculated by dividing the number of correct predictions by the total number of predictions.

Results

Transcriptional age-grading technique

The FC in the expression of age-responsive genes related to 1-day-old mosquitoes are represented in **Fig 1**. Generally, there was an underexpression in CUT and SCP with age, however, the expression levels of CUT showed a slight increase in days 2 and 3. FCs for CUT and SCP were more pronounced from day 5 onwards in comparison to 1-day old mosquitoes, being the highest FC in days 13 and 17 respectively. The lowest FCs for CUT and SCP were in days 2 and 3 in both cases. In case of CELL, there was a discrete increase in its levels of expression with age, being the FC values quite similar among age categories.

The LogC of Ct values for the three age-responsive genes are represented in **Fig 2**. LogC are inversely proportional to the expression levels of the mentioned genes meaning that an increase in LogC values reflect an increase in Ct values and a decrease in the levels of expression of these genes. In CUT and SCP, the distribution of LogC seemed to fit a sigmoid function or S-shaped curve with an initial lag phase of growth in the first 3 days, and exponential phase between days 3 and 5 (which may be extended to day 9 or 13 in SCP), and a plateau from day 5 onwards. In case of CELL, the distribution of LogC values seemed to fit a constant function with no pronounced differences across age categories.

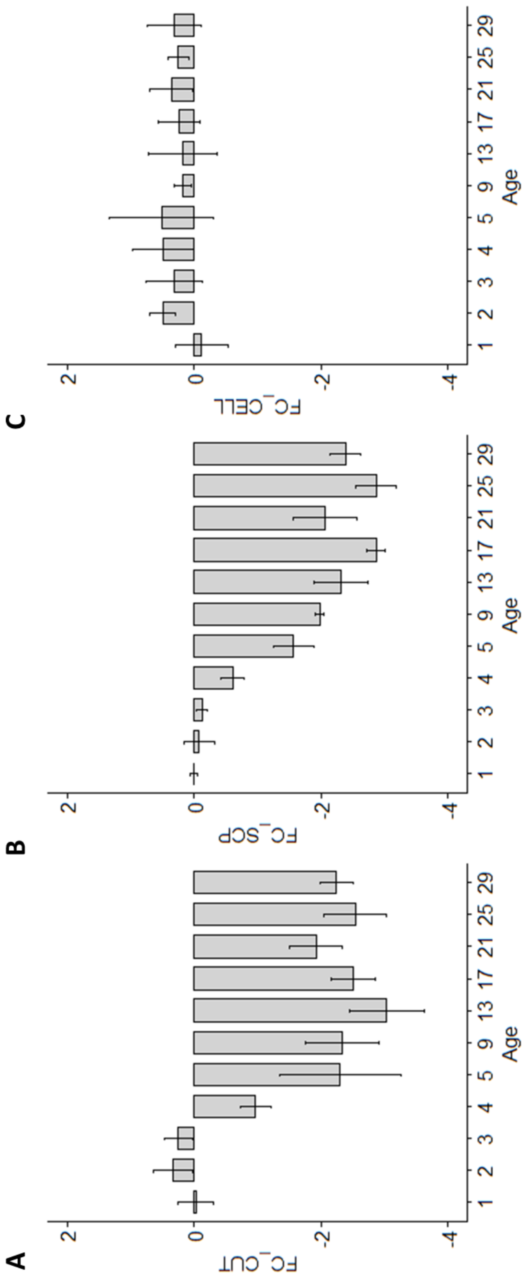


Fig 1. Fold-change in gene expression relative to day 1 mosquitoes calculated by the 2^ ΔΔCt method.

Kruskal-Wallis tests showed that only CUT and SCP genes exhibited a significant change in LogC values between age categories (p-value < 0.05) and were therefore suitable candidates for age prediction. On the contrary, CELL did not exhibit any significant change in their levels of expression with age (p-value > 0.05) and was therefore excluded from latter analysis. Post-hoc comparisons in CUT and SCP revealed differences between age classes (S1 Table). From 1 to 3 days of age, there were no statistically significant differences between age classes for both CUT and SCP, thus constituting a homogeneous group which was different from the rest of ages. LogC values of CUT and SCP in 4-day-old mosquitoes were also significantly different from the other age classes. In case of CUT, from 5 days onwards, mosquitoes did not show any significant differences in the distribution of LogC, except between 9 and 13 age classes. In case of SCP, 5-day-old mosquitoes were significantly different from the other age classes (except the 9-day class); and the 9-day class was different from the rest, except 5- and 21-day classes. From day 13 onwards some specific differences between age classes could be identified for SCP as shown in S1 Table.

Regression analyses were performed for those genes that showed age-related expression (i.e., CUT and SCP) with age as the explicative variable for calibration. Linear regression R^2 values were low: 0.44 for CUT, and 0.66 for SCP (**S1 Fig**). Since data suggested a non-linear fit and showed a sigmoidal shape curve, a logistic regression analysis was used, with $R^2 = 0.91$ for CUT and 0.93 for SCP (**Fig 3**).

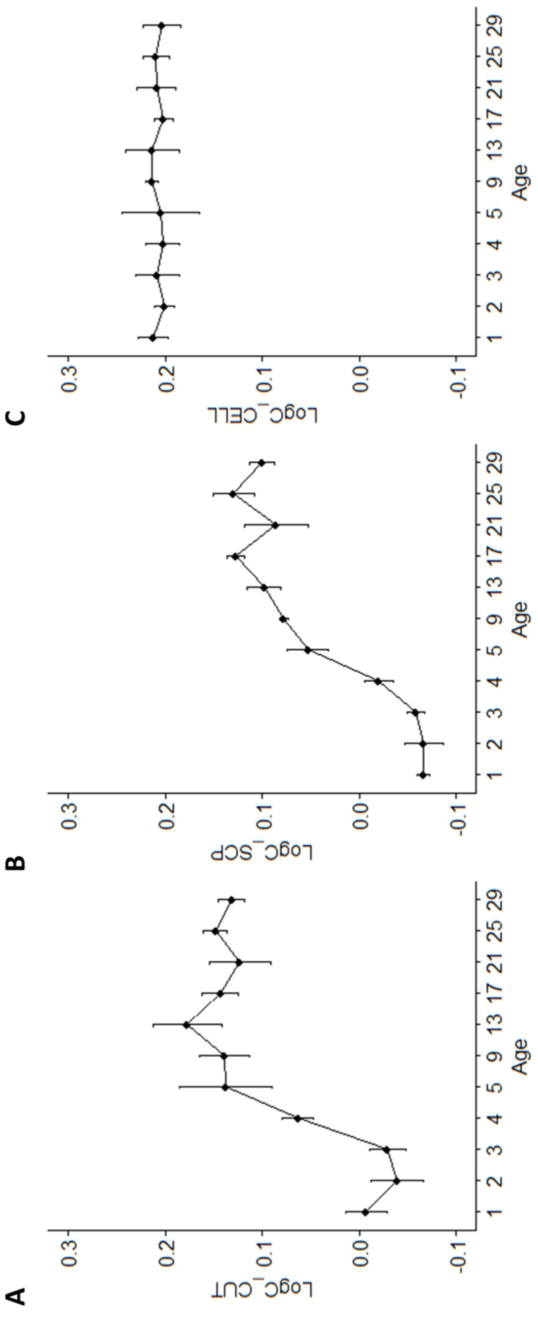


Fig 2. Relative changes in LogC values of age-related genes with age.

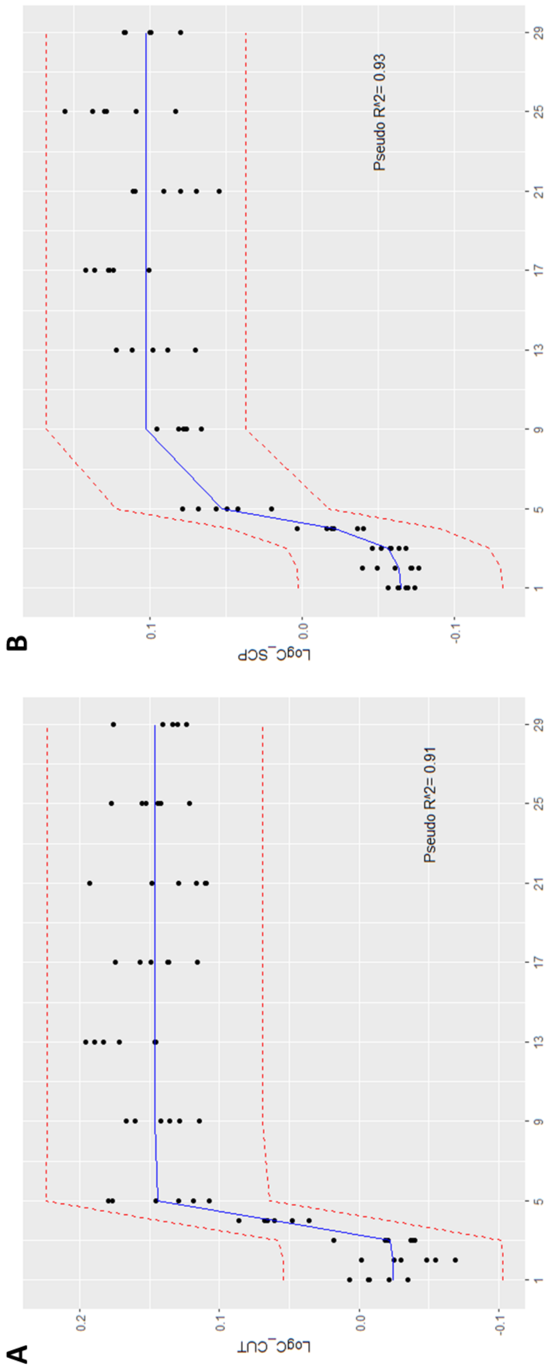


Fig 3. Logistic regression between age and LogC of CUT (**A**) and SCP (**B**).

Reverse regression was then performed to predict age from LogC values of age-responsive genes, with age as the response variable for prediction. A polynomial regression model was tested for CUT and SCP together, revealing that only the linear and polynomial (quadratic and cubic) components of SCP were significant to explain the variance in age (p-value < 0.05). Therefore, SCP was selected as the sole predictor variable for the age prediction model.

Since the LogC values of SCP seemed to be clustered into different groups, a regression tree was applied to the dataset to determine significant thresholds for a posterior segmented regression. Two significant break points in the distribution of SCP and three relevant groups of observations were identified (**S2 Fig**): i) LogCs of SCP < 0.052, with 41% of the observations (N=27) and an average age of 2.8 days old; ii) LogCs of SCP > 0.052 and < 0.072, with 15% of the observations (N=10) and an average age of 11 days old; and iii) LogCs of SCP > 0.079, with 44% of the observations (N=29) and an average age of 20 days old. The R^2 of the regression tree was 0.74. Once the break points were determined, the dataset was partitioned into three intervals and a generalised least squares model with weighted variances was fitted ($R^2 = 0.84$). Then, a leave-one-out cross validation was performed on the dataset, yielding an R^2 of 0.71, a MAE of 3.8 days in age prediction and a root-mean-square error (RMSE) of 5.1 days in age prediction. The model is plotted in **Fig 4**, showing the mean value and the 95% confidence intervals for age prediction.

The MAE between the actual and the predicted age classes was not homogeneously distributed and depended on the age class. For mosquitoes with ages ≤ 4 days old, the MAE was 1 day, and the confidence intervals were between 0 and 5.6 days old. For mosquitoes with ages ≥ 5 days old, the MAE for age prediction was 5.0 and the confidence intervals ranged from -3.7 to 24.9 in the intermediate age classes and from 7.0 to 33.9 in the older age classes. The mean predicted age values and the lower and upper limits of confidence intervals at each actual age class are presented in **S2 Table**.

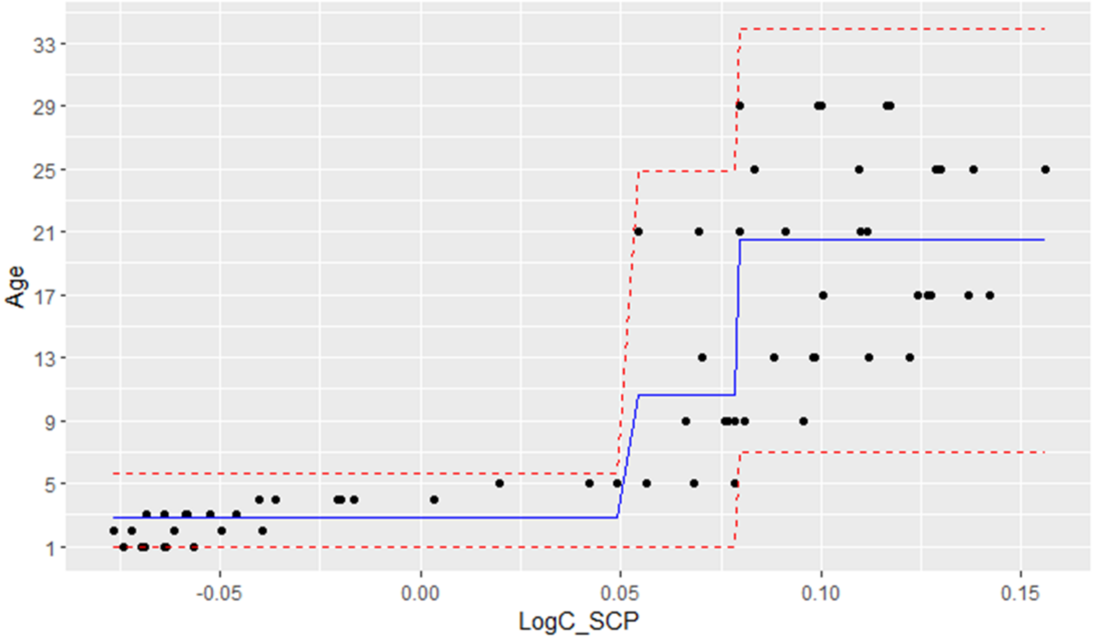


Fig 4. Age prediction model for *Cx. pipiens* based on LogC values of SCP gene. Black dots indicate the actual observations. The mean value for predicted age (blue line) is shown with 95% confidence intervals (red dashed lines). In the plot, the lower limit for age prediction was set to 1 day old.

Automated age classification system

The balanced dataset used for training and test contained a total of 2088 recordings of female *Cx. pipiens* from the three age groups (2-4, 7-9 and 14-16 days).

An exploratory analysis of the data was performed using descriptive statistics for the wingbeat frequency (WBF) which showed a significant difference in the mean of the WBF between the younger (2-4 day) and the old (14-16 day) classes. At 18ºC and 23ºC, the young and the middle (7-9 day) classes were significantly different from each other, but they were not different at 28ºC. The middle and the old classes only differ significantly from each other at 28ºC. Generally, there was a high overlap in the frequency distributions among the different age classes (**Fig 5**).

A first ML model was trained to classify mosquitoes into the three age classes: young (2-4 days), middle (7-9 days) and old (14-16 days). The best performing feature was MFCC, which gave an accuracy of 60.1%. The spectrogram feature and the PSD feature gave an accuracy of 57.5% and 58.6% respectively. The confusion matrix indicated higher confusion between the middle and old age classes (**Fig 6A**).

Considering these results, with the confusion between the middle and old age classes, a second training was performed in which the middle (7-9 days) and old (14-16 days) age groups were combined into a single “older” class (7-16 days). In this case, the balanced dataset contained a total of 1392 recordings and the best performing feature was the PSD, with a classification accuracy of 74.7%. In this case, the spectrogram and the MFCC features gave an accuracy of 72.1% and 70.9% respectively. The proportion of correctly classified age groups was superior to the previous model according to the confusion matrix (**Fig 6B**). A training score of 100% indicated that results could improve using more training data.

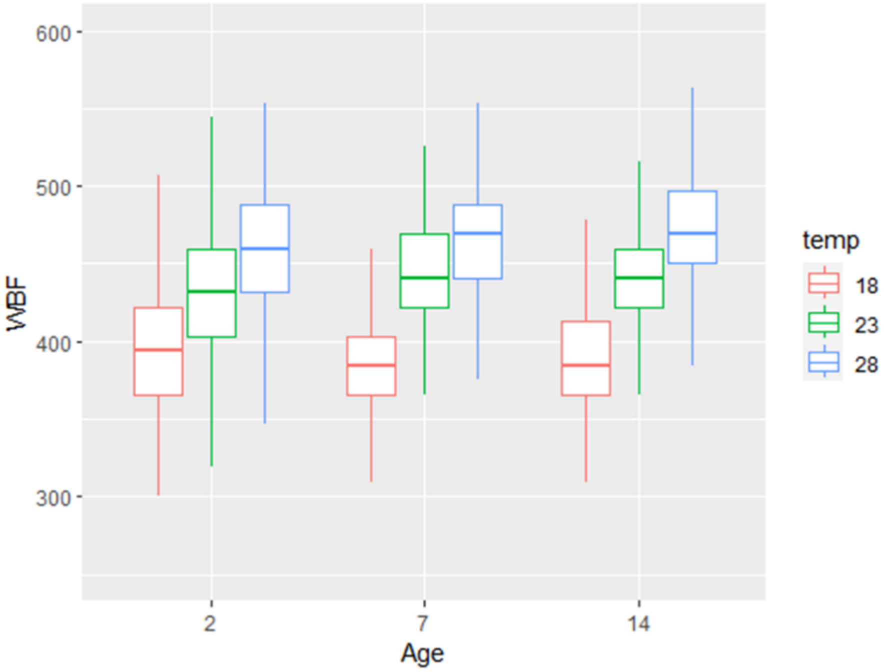


Fig 5. Distribution of the WBF (in Hz) across age categories at different temperature regimes. Significant differences in the distribution of WBF between age classes according to the Wilcoxon test is indicated with an asterisk.

A

Predicted age	2-4 days	7-9 days	14-16 days
Actual age			
2-4 days	121	23	30
7-9 days	36	97	41
14-16 days	25	53	96

B

Predicted age	Young (2-4 days)	Old (7-16 days)
Actual age		
Young (2-4 days)	127	47
Old (7-16 days)	41	133

Fig 6. Confusion matrix showing model performance in **A.** Model based on three age classes (2-4, 7-9 and 14-16 days old); and **B.** Model based on “young” (2-4 days old) and “old” (7-16 days old) age classes.

Discussion

In this contribution, two methods were described to predict the age of *Cx. pipiens* mosquitoes, one of the major vector species worldwide. The first method was a transcriptional age-grading technique, a molecular method based on the analysis of the differential expression of age-responsive genes across age categories, which has been applied to *Cx. pipiens* for the first time in this work. The second method was the novel use of an optical wingbeat sensor with ML to discriminate age classes based on mosquito wingbeat properties. To the best of our knowledge this is the first time, age classification results have been published for an automated entomological sensor.

Previous transcriptional profiling studies in *An. gambiae* (Cook & Sinkins, 2010), *Ae. aegypti* (Cook et al., 2006) and *Ae. albopictus* (Chathurika Weeraratne et al., 2021), demonstrated that the expression of several genes displayed large variation across age categories in adult mosquitoes and identified CUT, SCP and CELL as being the most informative for age prediction (Cook et al., 2006). In the present study, these three genes were tested for *Cx. pipiens*, although only CUT and SCP had a significant change with age.

The strong drop in the expression of CUT and SCP observed from day 1 to day 5 led to the inclusion of intermediate age categories (2, 3 and 4 days), to allow more accurate characterization of gene expression at these early ages. This contrasts to previously described protocols which did not consider these intermediate age classes (Cook et al., 2007). Post-hoc comparisons, which were also not assessed in previous studies, showed that differences in gene expression only occurred between certain age classes. In fact, from day 5 onwards no differences between age classes were observed for CUT, and from day 13 onwards, there was no consistent trend observed for SCP.

The general downregulation of CUT and SCP with age relative to 1-day-old *Cx. pipiens*, which was previously reported for *Aedes* mosquitoes (Chathurika Weeraratne et al., 2021; Cook et al., 2006a), is probably related to the putative function of their codifying proteins (endocuticle structural protein and sarcoplasmic calcium-binding protein, respectively), which are known to be downregulated in post-moulting stages in other invertebrate species (Gao et al., 2006; Zhao et al., 2018)As indicated in Hugo et al. (L. Hugo et al., 2010), which introduced the idea of age assessment based on total RNA yield, high transcription levels in newly emerged adults are probably a residual effect of high transcription rates during metamorphosis and these levels later stabilize at older ages.

When plotting LogC values of CUT and SCP for the different age categories, the calibration model fitted a sigmoid function, with an exponential growth trend from days 1 to 5, flattening from days 5 to 29 in CUT and from day 9 to 29 in SCP. This contrasts to the linear regression fit of other multivariate calibration models for *Ae. aegypti* and *Ae. albopictus* (Chathurika Weeraratne et al., 2021; Cook et al., 2006) where the regressions were performed with the redundancy variate (linear combination) of the three candidate genes, and did not include the 2-, 3- and 4-day age categories. SCP was the most informative gene to explain the variance in age of *Cx. pipiens* and was used as the sole age predictor variable, which simplifies previous approaches based on multivariate analysis (Cook et al., 2007). The final cross-validated model showed a MAE of 3.8 days in age prediction. This error was not equally distributed between age classes (RMSE = 5.1 days) and higher error values were associated with the prediction of older age classes. The accuracy was similar to other gene profiling studies based on multivariate calibration, such as (Cook et al., 2006), which reported a discrepancy of 5 days in age estimations for *Ae. aegypti,* and Weeraratne et al. (Chathurika Weeraratne et al., 2021) which reported a MAE of 2.19 (± 1.66) and 2.58 (± 2.06) days for *Ae. aegypti* and *Ae. albopictus* respectively. When applied to the field, the transcriptional age grading technique reported, in 77% of cases, a difference of 6 days in age predictions for *Ae. aegypti* (L. Hugo et al., 2010). Differences in results between studies may be due to differences in methodology, including mosquito rearing conditions; the type of samples used for RNA extraction (individual mosquitoes in the present study versus individual heads and thoraces (Cook et al., 2006; L. Hugo et al., 2010)or pools of mosquitoes (Chathurika Weeraratne et al., 2021)); or the design of the predictive model. In Weeraratne et al. (Chathurika Weeraratne et al., 2021), the ambient temperature at rearing was shown to influence age prediction, however in Hugo et al. (L. Hugo et al., 2010) no effects on age prediction were attributed to feeding status (presence of blood in the midgut), reproductive status (ovary development) or body size. Potential sources of variability should be investigated to calibrate models for field applications in surveillance and control programs.

Regarding the optical sensor-based approach to age grading developed in the present work, we note that the same system was used in a previous study of genus and sex classification of *Aedes* and *Culex* mosquitoes under laboratory conditions, with high accuracy results (González-Pérez et al., 2022). While other wingbeat sensor systems were used to taxonomically classify flying insects (by genus, species, , sex or parity status) (G. Batista et al., 2011; Fanioudakis et al., 2018; Genoud et al., 2018, 2019, 2020; Potamitis & Rigakis, 2016a; Silva et al., 2015)age classification has not been assessed previously using such sensors.

In many mosquito classification studies, the wingbeat fundamental frequency has been a good predictor variable (Santos et al., 2019), however, it was demonstrated to be insufficient for certain classification tasks (e.g. species classification) where there are overlapping frequency distributions (Genoud et al., 2018). In *Ae. aegypti*, the wingbeat fundamental frequency was reported to increase significantly with age (Staunton et al., 2019), although this increase only took place at young ages and plateaued in older age categories (D. Park et al., 2023). In the present study, due to the high overlap observed in in the WBF distributions between age classes richer features extracted from each sensor recording were used, including the spectrogram, PSD and MFCC. These features have been used in mosquito classification studies based on bioacoustic sensing (Fanioudakis et al., 2018; Su Yin et al., n.d.; Vasconcelos et al., 2020; Wei et al., 2022) and were also used in our own prior work on mosquito genus and sex classification using the same sensor (González-Pérez et al., 2022). In the present study, the widely used ML algorithm XGBoost was used because it has given good accuracy results for mosquito species classification in other works (Fanioudakis et al., 2018), and in our own prior work on mosquito genus and sex [34] and generally provides a good compromise between model performance and execution speed [34]. In the current contribution, the best age classification result for the sensor system was 74.7%, using the MFCC feature to distinguish young (2-4 day) age from older (7-16 day) mosquitoes. As indicated by the training score metric, the accuracy could possibly be improved using more training samples.

It is important to note that the two different age-grading methods for female *Cx. pipiens* assessed in this work (i.e., the transcriptional age grading technique, and the automated classification system) gave similar results. They both gave a reasonable level of accuracy for a binary classification between young mosquitoes (≤ 4 days old) and older mosquitoes. However, they were less accurate when differentiating between middle and old age categories due to the lack of change in the predictor variables between these classes. In line with vector competence studies of *Cx. pipiens* for WNV (Vogels, Göertz, et al., 2017a), this binary classification may serve to discriminate between two functional groups of non-vector and potential vector mosquitoes which are old enough to have overcome the EIP of the virus. The EIP of WNV in *Cx. pipiens* has been mainly reported superior to 7 days at temperatures that ranged from 18 to 27ºC (Anderson et al., 2008; Kilpatrick et al., 2008; Vogels, Göertz, et al., 2017a, 2017b). Therefore, the capacity to differentiate between two populations of mosquitoes of different ages may be useful to assess WNV control interventions targeting vector longevity.

In the case of the transcriptional age grading technique, a new method for age prediction is proposed for *Cx. pipiens* s.s based on a single candidate gene with individual mosquitoes. The one gene approach is simpler and more cost-effective to implement than previous methodologies developed for other mosquito species based on multivariate calibration. Furthermore, the use of whole individual mosquitoes avoids the time required to dissect specimens in other methods. In the case of the sensor-based system, its characteristics should enable automated age-grading of wild female *Culex* mosquitoes in the field without the need for specialised laboratory material, installations, or expert staff, and with potential application to other mosquito vectors. Both methodologies used in the present study represent an advance in the state of the art in mosquito age grading and vector biology, with potential future application in mosquito surveillance and control programs.

Acknowledgements

The authors are grateful to BSL2 facilities staff of IRTA-CReSA for their technical support, especially to Raquel Rivas.

Supporting Information

S1 Table. Pairwise comparisons in the distribution of LogC values between age classes using Wilcoxon rank sum exact test for CUT (**A**) and SCP (**B**). Blue cells indicate significant differences between age categories (p-value < 0.05).

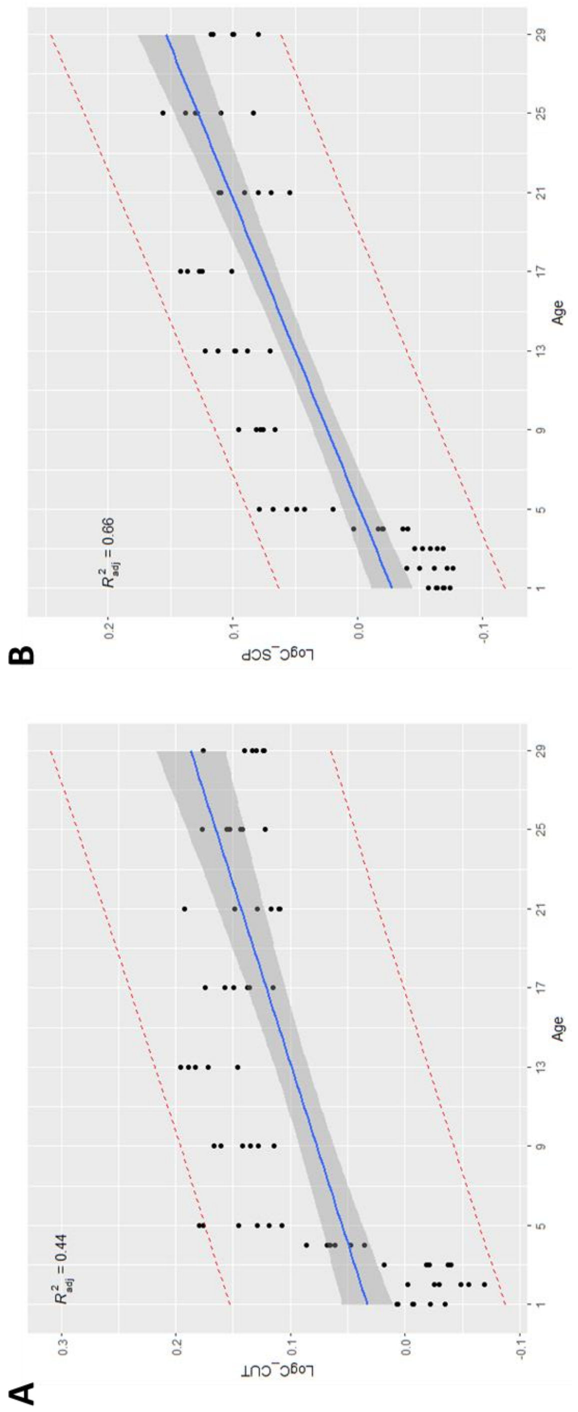
A. CUT

Age classes	1	2	3	4	5	9	13	17	21	25	29
1	-	-	-	-	-	-	-	-	-	-	-
2	0.102	-	-	-	-	-	-	-	-	-	-
3	0.3223	0.3223	-	-	-	-	-	-	-	-	-
4	0.0038	0.0038	0.0038	-	-	-	-	-	-	-	-
5	0.0038	0.0038	0.0038	0.0038	-	-	-	-	-	-	-
9	0.0038	0.0038	0.0038	0.0038	0,9372	-	-	-	-	-	-
13	0.0038	0.0038	0.0038	0.0038	0,1020	0.0433	-	-	-	-	-
17	0.0038	0.0038	0.0038	0.0038	0,9372	0.7395	0.1347	-	-	-	-
21	0.0038	0.0038	0.0038	0.0038	0.7395	0.5797	0.1347	0,3959	-	-	-
25	0.0038	0.0038	0.0038	0.0038	0.9372	0.689	0.1347	0.7395	0.3223	-	-
29	0.0038	0.0038	0.0038	0.0038	0.9372	0.7395	0.026	0.5797	0.5797	0.32	-

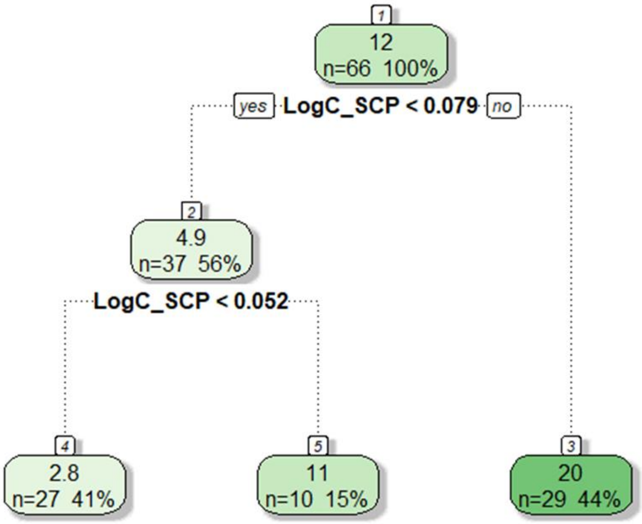
B. SCP

Age classes	1	2	3	4	5	9	13	17	21	25	29
1	-	-	-	-	-	-	-	-	-	-	-
2	0.9372	-	-	-	-	-	-	-	-	-	-
3	0.1513	0.5128	-	-	-	-	-	-	-	-	-
4	0.0035	0.0064	0.0035	-	-	-	-	-	-	-	-
5	0.0035	0.0035	0.0035	0.0035	-	-	-	-	-	-	-
9	0.0035	0.0035	0.0035	0.0035	0.0514	-	-	-	-	-	-
13	0.0035	0.0035	0.0035	0.0035	0.0064	0.0794	-	-	-	-	-
17	0.0035	0.0035	0.0035	0.0035	0.0035	0.0035	0.0116	-	-	-	-
21	0.0035	0.0035	0.0035	0.0035	0,0340	0.611	0.3405	0.0116	-	-	-
25	0.0035	0.0035	0.0035	0.0035	0.0035	0.0064	0.1089	0.8333	0.0514	-	-
29	0.0035	0.0035	0.0035	0.0035	0.0035	0.0116	0.5128	0.0116	0.2697	0.1	-

S1 Fig. Linear regression between age and LogC values for CUT (**A**) and SCP (**B**).



S2 Fig. Regression tree for LogC values of SCP.



S2 Table. Age predicted values (mean and 95% confidence intervals) obtained from the cross-validated model.

Actual age	**Predicted Age**		
	Mean	**95% confidence intervals**	
		Lower limit	**Upper limit**
1	2.8	0.0	5.6
1	2.8	0.0	5.6
1	2.8	0.0	5.6
1	2.8	0.0	5.6
1	2.8	0.0	5.6
1	2.8	0.0	5.6
2	2.8	0.0	5.6
2	2.8	0.0	5.6
2	2.8	0.0	5.6
2	2.8	0.0	5.6
2	2.8	0.0	5.6
2	2.8	0.0	5.6
3	2.8	0.0	5.6
3	2.8	0.0	5.6
3	2.8	0.0	5.6
3	2.8	0.0	5.6
3	2.8	0.0	5.6
3	2.8	0.0	5.6
4	2.8	0.0	5.6
4	2.8	0.0	5.6
4	2.8	0.0	5.6
4	2.8	0.0	5.6
4	2.8	0.0	5.6
4	2.8	0.0	5.6
5	2.8	0.0	5.6
5	2.8	0.0	5.6
5	2.8	0.0	5.6
5	10.6	-3.7	24.9
5	10.6	-3.7	24.9
5	10.6	-3.7	24.9
9	10.6	-3.7	24.9
9	10.6	-3.7	24.9
9	10.6	-3.7	24.9
9	10.6	-3.7	24.9
9	20.4	7.0	33.9
9	20.4	7.0	33.9
13	10.6	-3.7	24.9
13	20.4	7.0	33.9
13	20.4	7.0	33.9
13	20.4	7.0	33.9
13	20.4	7.0	33.9
13	20.4	7.0	33.9
17	20.4	7.0	33.9
17	20.4	7.0	33.9
17	20.4	7.0	33.9
17	20.4	7.0	33.9
17	20.4	7.0	33.9
17	20.4	7.0	33.9
21	10.6	-3.7	24.9
21	10.6	-3.7	24.9
21	20.4	7.0	33.9
21	20.4	7.0	33.9
21	20.4	7.0	33.9
21	20.4	7.0	33.9
25	20.4	7.0	33.9
25	20.4	7.0	33.9
25	20.4	7.0	33.9
25	20.4	7.0	33.9
25	20.4	7.0	33.9
25	20.4	7.0	33.9
29	20.4	7.0	33.9
29	20.4	7.0	33.9
29	20.4	7.0	33.9
29	20.4	7.0	33.9
29	20.4	7.0	33.9
29	20.4	7.0	33.9

(continuation)

4.5 CHAPTER 5: Assessment of ZIKV infection in *Aedes albopictus* though its wing-beat pattern.

Automated identification of ZIKV infected Aedes albopictus by the use of an entomological optical sensor trained with machine learning (In preparation).

Automated identification of ZIKV infected *Aedes albopictus* by the use of an entomological optical sensor trained with machine learning

María I. Gonzalez-Perez[1,2], Jaume Gardela[1,2], Bastian Faulhaber[3], Mark Williams[3], Joao Encarnaçao[3], Sandra Talavera[1,2], Núria Busquets[1,2*].

[1] IRTA. Programa de Sanitat Animal. Centre de Recerca en Sanitat Animal (CReSA). Campus de la Universitat Autònoma de Barcelona (UAB), Bellaterra, 08193, Catalonia. Spain.

[2] Unitat mixta d'Investigació IRTA-UAB en Sanitat Animal. Centre de Recerca en Sanitat Animal (CReSA). Campus de la Universitat Autònoma de Barcelona (UAB), Bellaterra, 08193, Catalonia. Spain.

[3] Irideon S.L., Barcelona, Spain.

*Correspondence: nuria.busquets@irta.cat

Abstract

Zika virus (ZIKV) is an emergent mosquito-borne flavivirus that can cause severe malformations and diverse neurological disorders in neonates. After several outbreaks in South America, in 2016 it was declared a public health emergency of international concern by the WHO. Although the epidemic cycle of ZIKV is mainly sustained by *Ae. aegypti*, *Ae. albopictus* is also a competent vector that can have a primary role in its transmission in temperate areas. It has been reported that ZIKV infection is neurotropic in mosquitoes, leading to behavioural changes in the vector, which may influence its vectorial

capacity. Here, we hypothesize that ZIKV infection in *Ae. albopictus* may cause an alteration in the mosquito flight pattern due to the ZIKV replication in the mosquito's nervous system; and that the potential alteration in the flight performance could be detected by an optical sensor, which in combination with machine learning techniques, could serve to differentiate infected from non-infected mosquitoes. To test this hypothesis, we performed a laboratory assay with a smart-trap system formed by an optical sensor coupled to a standard mosquito trap, which reported a 63.0% classification accuracy with the power spectral density (PSD) feature and the deep neural network (DNN) algorithm. This result, although improvable with more training samples, is encouraging for future vector biology research and potential applications of the smart-trap technology in mosquito and arbovirus surveillance and control programs.

Introduction

Zika virus (ZIKV) is an emergent mosquito-borne virus that is a member of the *Flaviviridae* family, which includes other mosquito-borne viruses of public health concern such as dengue (DENV), West Nile (WNV), Japanese encephalitis or yellow fever (Plourde & Bloch, 2016) viruses. As with other flaviviruses, ZIKV is a small spherical particle of ~50 nm containing a positive-sense single-stranded RNA molecule of ~11 kb that encodes for three structural proteins and seven non-structural proteins (Lindenbach & Rice, 2003). This virus was first isolated in 1947 from a febrile rhesus monkey (*Macaca mulatta*) in the Zika forest (Uganda) and a few months later from a pool of *Aedes africanus* mosquitoes in the same forest (GW et al., 1952). According to phylogenetic analyses, there are at least two major lineages of ZIKV: the African and the Asian (Haddow et al., 2012). Reports of human infections were constrained to these continents with very few diagnosed cases (Gubler et al., 2017). It was not until 2007 when the global ZIKV awareness began with the first large outbreak beyond its known distribution range, in the Yap State (Micronesia) (Duffy et al., 2009). From 2013 onwards, ZIKV outbreaks emerged in the Pacific Islands, Africa, Asia, and the Americas (Gubler et al., 2017)5. In 2016, the occurrence of clinical cases of microcephaly in neonates, congenital nervous system malformations and Guillain-Barré syndrome associated with ZIKV infection led the World Health Organization to declare ZIKV as a public health emergency of international concern (Wilder-Smith & Osman, 2020).

Zika virus is maintained in nature primarily through a sylvatic cycle which includes forest-dwelling *Aedes* mosquitoes and non-human primates, although many other non-primate animals have been considered as potential reservoirs (Bueno et al., 2016). Infections in humans occur when there is a spillover from the sylvatic cycle and humans become incidental hosts or when an epidemic transmission cycle becomes established in urban sites with anthropophilic vectors feeding preferentially on humans (Gutiérrez-Bugallo et al., 2019). Sexual transmission of ZIKV also plays a role in the virus spreading among humans (Counotte et al., 2018). Additionally, the confirmation of vertical and venereal transmission of ZIKV for several mosquito species has been suggested as a mechanism that may ensure the maintenance of ZIKV in vector populations during adverse conditions for horizontal transmission among vertebrate hosts (Gutiérrez-Bugallo et al., 2019).

Despite that many mosquito species have been found to be naturally infected by ZIKV, *Aedes aegypti* has been pointed out to be the major vector in urban areas driving recent epidemics, while other *Aedes* species may contribute to the sylvatic transmission cycle (Boyer et al., 2018; Epelboin et al., 2017). In temperate areas where *Ae. aegypti* is absent, other human-adapted species such as *Aedes albopictus*, which are also competent to carry and transmit ZIKV (Mckenzieid et al., 2019), can contribute to disease epidemics even though *Ae. albopictus* transmission rates for ZIKV are low compared to *Ae. aegypti* (Obadia et al., 2022). An example of this was the autochthonous transmission of ZIKV declared in France in 2019 which was attributed to *Ae. albopictus*

(Giron et al., 2019). The highly invasive nature and extensive geographic distribution of *Ae. albopictus*, which tends to predominate over *Ae. aegypti* in habitats where both species are sympatric (M. A. H. Braks et al., 2004; O'Meara et al., 1995; Tedjou ID et al., 2019; Zhou et al., 2022), suggests that *Ae. albopictus* has the potential to become a major vector of ZIKV globally (Gutiérrez-Bugallo et al., 2019). In addition, its ecological plasticity in habitat and host choice, makes it a potential bridge vector, thus increasing the risk of spill-over and spill-back events (Pereira-Dos-Santos et al., 2020).

When a mosquito feeds on a viraemic host, the virus is ingested through the digestive tract triggering an innate immune response in the vector (Samuel et al., 2018). To replicate and disseminate inside the vector, the virus needs to overcome this immune response and cross the different vector barriers in the midgut and the salivary glands. These biological constraints pose an evolutionary pressure on the virus through the effect of sequential bottlenecks which challenges viruses to maintain diversity and fitness (Forrester et al., 2014). There is a consensus assuming that arboviruses infect arthropods without significant damage, supporting a persistent infection in their cells without the cytopathic effect that characterizes vertebrate cell infections (Y. G. Li et al., 2012). However, the effects of arboviral infection on mosquitoes are diverse and show a vector-pathogen specificity, which may be a result of co-evolutionary processes (Ciota et al., 2011).

Some flaviviruses exhibit a remarkable tropism for mosquito nervous tissues, which appeared to be the primary site of amplification for DENV in *Ae. aegypti* (Linthicum et al., 1996). Neurotropism has also been demonstrated for ZIKV with replication in the central nervous system and peripherical sensory organs of *Ae. aegypti* (Gaburro, et al., 2018a). The costs of infection of different flaviviruses on their competent vectors have been published in several studies (Gaburro et al., 2018a; Gaburro, Paradkar, et al., 2018; Lima-Camara et al., 2011; Maciel-de-Freitas et al., 2011; Pedreira Padilha et al., 2018; Platt et al., 1997; Vogels, Fros, et al., 2017). Most of them are based on the effects of DENV in *Ae. aegypti*, thus reporting: a decrease in vector survival, longevity, and fecundity (Maciel-de-Freitas et al., 2011); alterations in the feeding behaviour with longer probing and feeding times (Platt et al., 1997); alterations in oviposition olfactory preferences (Gaburro et al., 2018); or increases in the locomotor activity (Lima-Camara et al., 2011. For WNV and *Culex pipiens*, a decrease in the host-seeking response was found in infected mosquitoes but no effects on host preferences, blood-feeding propensity, olfactory responses, survival, or flight activity were identified (Vogels, Fros, et al., 2017). In the case of ZIKV infecting *Ae. aegypti*, one study found a decrease in their locomotor activity (Pedreira Padilha et al., 2018); while another reported a hyperexcitation of infected neurons and a significant increase in the diurnal locomotion activity (Gaburro et al., 2018a). Although there is growing evidence that infected mosquitoes express a modified behavioural response, data on the actual flight performance of ZIKV-infected mosquitoes are still scarce.

Here, we hypothesize that the infection with ZIKV of the competent vector *Ae. albopictus* may cause an alteration in the mosquito flight pattern due to the ZIKV replication in the mosquito’s nervous system;

and this alteration in the flight performance could be registered by an optical sensor, which in combination with machine learning (ML) techniques, could serve to differentiate infected from non-infected mosquitoes. The sensor prototype, which is coupled to a standard mosquito suction trap, previously reported high accuracy results for genus and sex classification of *Aedes* and *Culex* mosquitoes in laboratory conditions (González-Pérez et al., 2022). In the present contribution, the system is challenged to automatedly differentiate ZIKV-infected from non-infected *Ae. albopictus* through supervised learning trained on mosquito flight features. To the best of our knowledge, the present study is the first time that an automated classification system coupled to a mosquito trap has assessed the classification of mosquitoes by their vector status. This opens a whole new perspective for the future of mosquito and arbovirus surveillance and vector biology research.

Material and methods

Mosquito colony rearing

An established laboratory colony of *Ae. albopictus*, population of Sant Cugat del Vallès (2005), Barcelona, Spain (41.4667°, 2.0833°) was used to run the experiment. Eggs were hatched with autoclaved water and larvae were maintained in plastic trays (density of 250 larvae/tray) with dechlorinated tap water, renewed three times per week, and fed with fish food pellets (Goldfish Sticks-TETRA, Melle, Germany). Pupae were placed in plastic cups inside insect-rearing cages (BugDorm-1 Insect

Rearing Cage W30 × D30 × H30 cm, MegaView Science, Talchung, Taiwan) until adult emergence. Adult female mosquitoes were anesthetised with CO_2 (Flowbuddy flow regulator, 59-122BC, Flystuff, California, USA) and sorted in groups of 15 individuals in cardboard boxes and supplied with 10% sucrose solution administered on soaked cotton pledgets. Mosquito life cycle took place inside a climatic chamber (Telewig, Barcelona, Spain) at 28ºC; 80% relative humidity (RH); and a light:dark photoperiod of 12:12 hours.

Virus production and titration

The ZIKV Martinique strain from Asian lineage (MRS_OPY_Martinique_PaRi_2015, Asian lineage, passage 3, Genbank reference: KU647676), which was originally isolated from a human serum in 2015, was used in the present study to infect *Ae. albopictus* mosquitoes. Viral production was performed from an original stock and was titrated in Vero cells (ATCC, ref. CCL-81) to obtain the plaque forming units per mL (PFU/mL).

Experimental infection assay

Adult female *Ae. albopictus* of 3-5 days of age were transferred to the Biosafety Level 3 facility of IRTA-CReSA, where the experimental infection of ZIKV took place. Sucrose supply was removed 24 h before the infection assay. Then, females were anesthetised with CO_2 in a plate (Flowbuddy flow regulator, 59-122BC, Flystuff, California, USA) and were inoculated intrathoracically (IIT) with a microinjector (XenoWorks®, Sutter Instrument, Novato, USA). A total of 843 females of *Ae. albopictus* were inoculated: 433 with 1 µL of ZIKV from a viral stock of 6.74 $\log_{10}$PFU/mL; and 410 with 1 µL of Dulbecco's modified Eagle's medium (DMEM) (Lonza Group AG, Basel, Switzerland) as a control group.

After inoculation, 6 individual mosquitoes from the IIT ZIKV group and 6 from the control group were preserved in separate tubes containing 500 µL of DMEM (with 2% of Fetal Bovine Serum (FBS) and 1% of antibiotic and antimycotic solution), as 0 days-post-infection (dpi) samples for later virus detection through RT-qPCR. The rest of the mosquitoes were maintained at 28ºC and 80% of RH and fed with sucrose *ad libitum* until 14-16 days of age. At this time point, 7 females from the IIT ZIKV group and 6 from the control group were separated as a 11-12 dpi samples. Females from the IIT ZIKV group were previously anesthetised with CO_2 in the FlyStuff plate and the wings and legs were removed from the body. Saliva was extracted from each mosquito following a capillary technique as described in a previous work (Dubrulle et al., 2009)34. The extraction was carried out during 30 min with a P20 pipette tip filled with 7 µL of FBS and saliva samples were preserved in 193 µL of the same medium described above. Individual bodies from these same individuals were preserved in 500 µL of the same medium. Females from the control group at 11-12 dpi were not dissected and were preserved individually in 500 µL of the same medium. Individual bodies were homogenised at 30 Hz for 1 min using

TissueLyser (Qiagen GmbH, Stockach, Germany) and then stored at -75ºC until its processing.

For immunohistochemistry (IHQ) assays, 6 other females of each group (IIT with ZIKV and ITT with DMEM) at 11-12 dpi were preserved in formaldehyde.

The rest of individuals that survived the period of 11-12 dpi (259 from the ZIKV IIT group and 271 from the control group) were used for the flight assays. The mortality of inoculated mosquitoes was quantified.

Virus detection through real time RT-qPCR

Viral RNA was extracted from body and saliva samples using NucleoSpin® RNA Virus commercial kit (Macherey-Nagel GmbH & Co. KG, Düren, Germany) according to the manufacturer’s protocol. Reverse transcription quantitative polymerase chain reaction (RT-qPCR) for ZIKV was carried out by using the primers and TaqMan probe designed in a previous study (Lanciotti et al., 2008)35 and the AgPath-ID™ One-Step RT-PCR reagents (Applied Biosystems, Thermo Fisher Scientific, Massachusetts, USA). The thermal cycling conditions were 45^{o}C for 10 min; 95^{o}C for 10 min; and 45 cycles of 95^{o}C for 10 s and 60^{o}C for 45 s.

Virus detection and semi-quantification through IHQ

The formaldehyde preserved samples of IIT ZIKV and DMEM mosquitoes at 11-12 dpi were used for IHQ analyses following standard paraffin-embedding histological procedures and haematoxylin and eosin staining. A microtome (HistoCore MULTICUT, Leica, Wetzlar, Germany) was used to cut 3 µm of mosquito paraffin sections, which were then deparaffinized, rehydrated, and incubated with 3% hydrogen peroxide in methanol for endogenous peroxidase inhibition. The sections were first incubated with Target retrieval solution (DAKO, Ref: S1699) and blocked for unspecific staining with bovine albumin (2%) in PBS - Tween20. After that, the sections were incubated with the primary antibody, a mouse monoclonal [D/2/D6/B7] antibody (MoAb) to flavivirus NS1 D2/D6/D7 (Abcam, Ref: ab214337) (1/500). The samples were later incubated with a peroxidase labelled polymer conjugated goat anti-mouse immunoglobulins (DAKO, Ref: AK4001) secondary antibody. Finally, 3’-Diaminobenzidine tetrahydrochloride hydrate (DAB, Sigma-Aldrich, Ref: D5637) substrate was added for staining reaction followed by counter-staining with Mayer’s haematoxylin. For a no primary antibody control, duplicate sections were incubated with the antibody diluent without primary antibody, followed by incubation with the previous secondary antibody and detection reagents. For the isotype control, duplicate sections were incubated with a non-immune antibody of the same isotype and at the same concentration as the primary antibody, followed by incubation with the secondary antibody and detection reagents. Samples were prepared with DPX mounting medium using a fully automated glass coverslipper (Leica, Ref: CV5030) and were then examined under the microscope (Motic, Ref: BA410E).

Mosquito tissues stained in brown revealed the detection of the non-structural protein 1 (NS1) of ZIKV by the primary antibody (MoAb [D/2/D6/B7]) and thus identified as positive for ZIKV infection. Negative controls (non-infected mosquitoes) were analysed to determine unspecific staining which may lead to misinterpretations, e.g., because of chitin content or the presence of artifacts (Chatterjee, 2014)36. A semi-quantification of the ZIKV antigen presence in mosquito tissues was performed based on a scoring scheme from 0 to 3, in which 0 means no cells stained; 1 means few cells stained; 2 means many cells stained; and 3 means all cells stained. The central nervous system (head, thoracic and abdominal ganglia) so as peripherical structures (ommatidia and Johston's organ) were analysed because of the potential affectation of virus neurotropism on mosquito flight behaviour (Gaburro et al., 2018a). Scores for each anatomical structure were averaged for the set of 6 IIT ZIKV samples.

Flight assays with the sensor

Female mosquitoes from the IIT ZIKV group and DMEM control group were used to perform separate flight assays with an optical sensor prototype (Irideon, Barcelona, Spain), which was coupled to the entrance of a BG-Mosquitaire trap (Biogents, Regensburg, Germany). The inner operational principal of the sensor was described in detail in a previous work (González-Pérez et al., 2022). For the flight assays, the sensor and trap system were placed inside an insect-rearing cage (BugDorm-4S4590 W47.5 × D47.5 × H93.0 cm, MegaView Science, Talchung, Taiwan) in a climatic chamber (Dycometal, CCK-0/5930m, Barcelona, Spain) at controlled environmental conditions of 23ºC and 80% of RH. Mosquitoes were released in batches of 15-30 individuals inside the insect-rearing cage where the sensor was placed. Mosquitoes flew freely until they got sucked by the effect of the fan inside the trap, passing through the sensing zone of the sensor where their flight got recorded. All mosquitoes were held in the climatic chamber at the designated ambient temperature for 24 h prior to the start of the assay for acclimatization.

Machine learning model

After the flight assays, the sensor recordings were downloaded to a laptop computer and processed using a Python script to produce playable and viewable audio files. Pre-processing of data included the manual examination of each recording and the exclusion of those considered to be invalid, e.g., recordings containing double flights or those where the mosquito was deemed to have hit the wall of the flight tube inside the sensor. The resulting dataset was split into 75% for training the supervised ML model and 25% for testing its performance. The power spectral density (PSD) feature was extracted from the flight signal and was trained with a deep neural network (DNN) algorithm in order to differentiate ZIKV infected from non-infected *Ae. albopictus*. Four-fold cross-validation was performed on the training set in search of the most optimized model which was then used for testing. Model performance was assessed using the accuracy metric, which is calculated by dividing the number of correct predictions by the total number of predictions. Since the data in the test set are completely new to the model, accuracy results for the test set are an indication of how well the model generalizes on new data.

Results

ZIKV detection by RT-qPCR

The presence of ZIKV was confirmed in all mosquitoes (N=6) at 0 dpi, which were IIT with ZIKV with a mean Ct value of 28.4. At 11-12 dpi, 100% of the inoculated mosquitoes presented infection with a mean Ct value of 19.7, indicating virus replication within all mosquitoes exposed to ZIKV. Among the mosquitoes with infection, three tested positive in saliva meaning that 42.8% of the samples (N=7) were transmitting ZIKV at 11-12 dpi.

ZIKV detection by IHQ

The nervous system of *Ae. albopictus* female mosquitoes supported ZIKV replication at 11-12 dpi in both the ganglionic system (head and thoracic ganglia) and in peripherical structures (mainly the ommatidia), as depicted in **Fig 1**. According to the developed scoring scheme for virus semi-quantification through IQH (Table 1), the most affected organs were the compound eyes (ommatidia), whose average score for antigen-staining was 2.7, showing high ZIKV infectivity in the photoreceptor cells. The second most affected anatomical structure was the ganglionic system of the head and thorax with two clearly differentiated regions: the neuropile, formed by non-medullated axons located in the inwards of the ganglia; and the cortical layer, formed by the nucleus of the neurons (soma) delineating an outline of the ganglionic mass. While the neuropile was barely stained (score of 0.5), the cortical layer showed high infectivity with an average score of 1.8 and 2.0 in the head and thoracic ganglia respectively. The microtome sections of the sampled mosquitoes did not allow to visualise the abdominal ganglia, so scoring data on this structure was not available. The Johston’s organ located in the antennae did not support high positivity neither scoring for ZIKV, showing an average score of 0.5.

Table 1. Summary of the ZIKV antigen semi-quantification in the anatomical structures of the nervous system of *Ae. albopictus* females at 11-12 dpi. Averages and standard deviations (SD) of the scores for all sampled mosquitoes are presented. Abbreviations: Data not available (NA).

ZIKV infected mosquitoes								
	1	**2**	**3**	**4**	**5**	**6**	**Average**	**SD**
Ommatidia	3	3	3	1	3	3	2.7	0.7
Johston's organ	1	NA	1	0	0	NA	0.5	0.5
Head ganglia								
Neuropile	0	1	1	0	0	1	0.5	0.5
Cortical layer	2	2	2	2	2	1	1.8	0.4
Thoracic ganglia								
Neuropile	NA	0	1	0	1	NA	0.5	0.5
Cortical layer	NA	2	2	2	2	NA	2	0
Abdominal ganglia								
Neuropile	NA	NA	NA	NA	NA	NA	NA	NA
Cortical layer	NA	NA	NA	NA	NA	NA	NA	NA

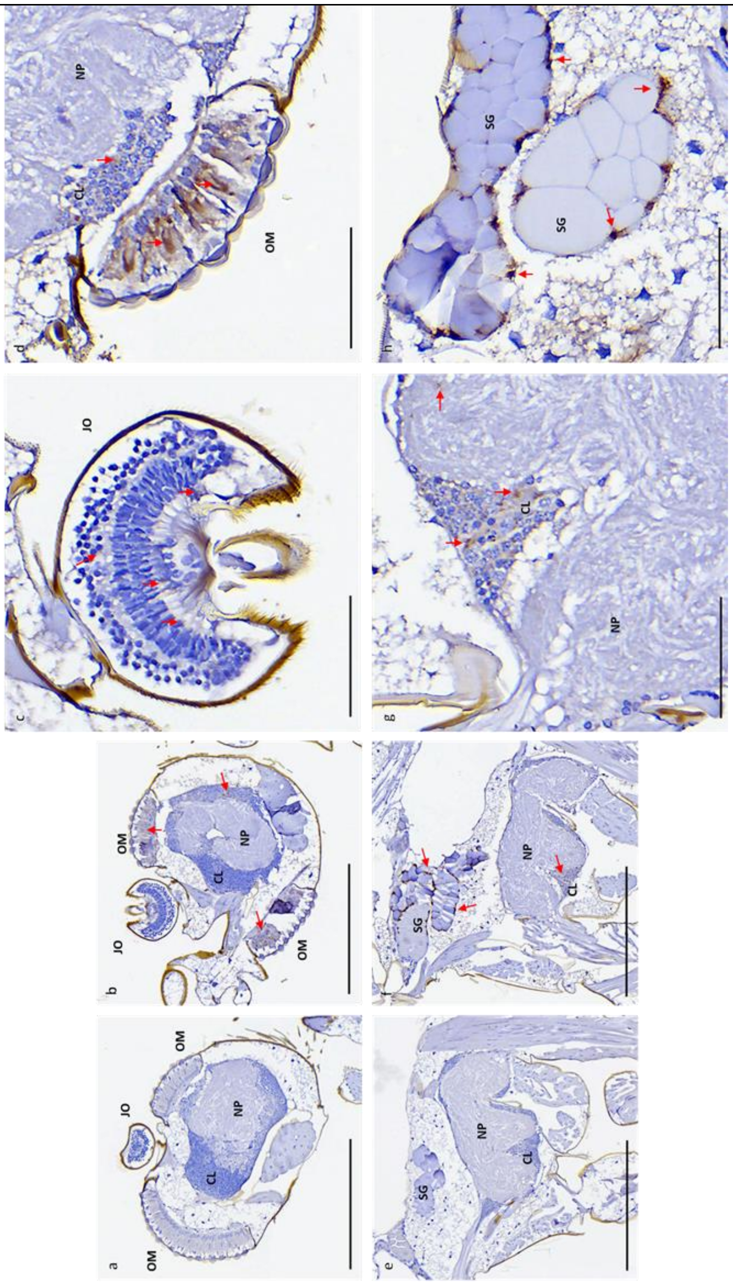


Fig 1. Anatomical structures in head (**a-d**) and thorax (**e-h**) of non-infected (**a, e**) and ZIKV-infected (**b-d**, **f-h**) *Ae. albopictus* females. Pictures a, b, e, f at 10x (scale bar 300 µm) and c, d, g, h at 40x (scale bar 60 µm). Cortical layer (CL), Johnston's organ (JO), neuropile (NP), ommatidia (OM), salivary gland (SG). Red arrows indicate points of antigen-staining.

Classification accuracy of the ML model

From the total of 843 females that were inoculated for the study, 35.8% of the IIT ZIKV group and 29.5% of the control group died before 11-12 dpi, leaving a total of 567 individuals alive. Discounting those that were preserved for virus detection, a total of 530 females were finally available and were released during the flight assays (259 from the IIT ZIKV group and 271 from the control group). Some of the mosquitoes that were released in the flight assays showed certain mobility problems and were inactive on the floor of the insect cage. In the end, 212 ZIKV-infected mosquitoes and 209 non-infected mosquitoes entered the trap and were counted by the sensor, representing respectively 81.8% and 77.1% of the mosquitoes that were originally released. During the preprocessing of data, 5.2% of flights were rejected leading to a final dataset composed of 399 valid flight recordings (197 for the IIT ZIKV group and 202 for de control group). The final dataset was randomly under-sampled to obtain a balanced dataset which was split into a training set and a test set. The training set was used to develop the ML classification model with the extracted PSD feature trained with DNN, reporting an overall accuracy of 63.0% in classifying ZIKV-infected and non-infected *Ae. albopictus*. A training score of 68.6% indicated that with more training data the results could improve.

Discussion

The current study assessed for the first time the automated identification of arbovirus-infected mosquitoes captured by a mosquito trap by the analysis of their wing-beat flight pattern through supervised ML. The developed automated classification system composed of an optical sensor coupled to a mosquito trap, appears to be a potentially useful tool for the future of entomological and arbovirus surveillance. The current classification accuracy of the model is still improvable, and a full development of its capabilities would require more extensive work that includes more training samples.

In recent years, the interest in ML approaches for mosquito surveillance has skyrocketed with an upward trend of published papers aimed at identifying mosquitoes through a variety of sensing devices (Joshi & Miller, 2021). Particularly, the use of optical (also called optoacoustical) sensors to record the mosquito flying signals has experienced considerable growth, attempting different classification tasks regarding mosquito taxonomy (e.g., genus and species classification) and other attributes of mosquito biology (e.g., gender or parity status) (Fanioudakis et al., 2018; Genoud et al., 2018, 2019, 2020; Ouyang et al., 2015; Silva et al., 2015) with free-flying mosquitoes. Other proposals include the use of sensors coupled to mosquito traps which record mosquitoes under a forced flight while they are being sucked by the fan inside the trap and have been designed for mosquito counting (Geier, Weber, Rose, Obermayr, et al., 2016; J. Wang et al., 2020), genus and sex classification(González-Pérez et al., 2022) or for the recognition of certain mosquito species (Lai et al., 2022). However, the automated recognition of mosquitoes based on their infection status through optical sensors had not been attempted before, until the present work.

Our initial hypothesis began on the basis that a disseminated infection of ZIKV in the competent vector *Ae. albopictus* may generate an affectation in the nervous system that would lead to an alteration in the mosquito behaviour, as previously reported for ZIKV-infected *Ae. aegypti* (Gaburro et al., 2018a; Pedreira Padilha et al., 2018). The IHQ analysis confirmed the tropism of ZIKV for *Ae. albopictus* nervous tissues, showing virus replication at 11-12 dpi in the cortical layer of the head and thoracic ganglia and in peripherical sensorial structures, especially in the ommatidia, as indicated by positive staining. These results are in concordance with previous IHQ analysis in *Ae. aegypti* at two weeks post-infection with ZIKV (Gaburro et al., 2018a). The attributed scoring values indicated that ommatidia supported higher virus replication in comparison with other tissues, suggesting that the vision of *Ae. albopictus* may be affected when infected with ZIKV, thus possibly compromising mosquito behaviours mediated by visual stimuli such as the optomotor response, landscape navigation, swarming, and host-seeking behaviours (Hawkes et al., 2022). These infection-driven changes in mosquito rhythmic patterns may have an impact on its vectorial capacity, and consequently, on arbovirus transmission (Vieira Bruno et al., 2020).

All the females IIT with ZIKV sustained replication at 11-12 dpi, however, not all of them were actually transmitting the virus at that time point, which agrees with the transmission rates reported for ZIKV-infected *Ae. albopictus* in other competence studies (Di Luca et al., 2016; Jupille et al., 2016). The high mortality reported at 11-12 days post-inoculation in both infected and uninfected groups is likely associated with the intrathoracic inoculation technique (Sanchez-Vargas et al., 2021), so is the presence of mosquitoes with scarce mobility during the flight assays. A shift towards oral infection may solve these problems but would generate others possibly related to low feeding and infection rates (Yang et al., 2022), hence making it unpractical to apply for this kind of assays that require a lot of data.

Overall, the automated classification system was able to discriminate *Ae. albopictus* mosquitoes with a disseminated ZIKV infection from non-infected ones with 63.0% of classification accuracy. This result points out that ZIKV infection in *Ae. albopictus* may generate an alteration in the mosquito flight pattern that could be identified with an optical sensor coupled to a mosquito trap, thus enabling an automated classification through an ML model. According to the training score, the accuracy of the model could be improved with more training samples which encourages further testing in search of better performance. This study provides preliminary evidence of how new technologies have the chance to enhance the capabilities of mosquito vector surveillance thus allowing a timelier disease risk assessment.

Conclusions

The infection with ZIKV of the competent vector *Ae. albopictus* modifies the mosquito flight pattern possibly due to the amplification of ZIKV in its nervous system. This change in flight performance can be registered by means of optical sensors, which in combination with ML, can provide classification models for the automated differentiation of infected from non-infected mosquitoes. The automated classification system presented here, consisting of a sensor coupled to a mosquito trap, was able to differentiate ZIKV-infected *Ae. albopictus* with 63% of accuracy. Although further assays are needed to improve the accuracy of the model, this work provides preliminary evidence of how this kind of technology has the potential to boost the current state of mosquito vector monitoring with effortless and timelier surveillance methods.

5. GENERAL DISCUSSION

A great fraction of human's world population lives at risk of VBD, accounting for almost 700,000 human deaths per year and being MBD the largest contributors to disease burden (WHO, 2020). In a context of global change, the unprecedented increase in travel, trade and tourism has promoted the circulation of both pathogens and vectors beyond their origin sites (Baker et al., 2022). Moreover, massive changes in land uses from natural to human-dominated landscapes with growing deforestation, agriculture and urbanization have enhanced disease risk by altering the interactions between humans, pathogens, vectors, and vertebrate hosts (Franklinos et al., 2019). Together with environmental change, socioeconomic factors such as poverty index have also been correlated with an increased risk of neglected tropical diseases (e.g., dengue and malaria) in low-income countries (Magalhães et al., 2023). Altogether, there is increasing tendency in the incidence and geographical distribution of MBD which are emerging in new areas and re-emerging in regions from which they had previously been eradicated, thus constituting a major global health issue (WHO, 2014).

Reducing the incidence of MBD is on the focus of the World Health Organization, which in 2017 adopted the Global Vector Control Response aimed at reducing the disease burden of VBD by reinforcing vector surveillance and control programs though a perspective of integrated vector management (WHO, 2017). On one side, mosquito control has been demonstrated highly effective against MBD when comprehensively applied and sustained (Wilson Id et al., 2020). On the other side, mosquito surveillance gathers important entomological information about the vector which allows to stablish intervention thresholds for control actions and serves to evaluate the efficacy of control measures (Caputo & Manica, 2020).

In the last decades, Europe has experienced an increasing number of imported and indigenous cases of MBD (Calzolari, 2016). Native mosquitoes like *Cx. pipiens* s.s. are behind the transmission of endemic pathogens of mandatory declaration for public and animal health such as WNV; but are also responsible of the emergence of other neglected pathogens that remain present in the continent such as USUV, Tahyna or Sindbis virus. Furthermore, native mosquitoes (e.g., *Anopheles* genus) could act as local vectors of already eradicated pathogens (e.g., malaria plasmodia) if they are reintroduced. Parallelly, the expansion of *Aedes* invasive species in Europe has led to the circulation of exotic pathogens within autochthonous transmission cycles thus leading to sporadic disease outbreaks. A proof of this is the increase in the number of CHIKV and DENV outbreaks transmitted by *Ae. albopictus* in the Mediterranean area in the last years (ECDC, 2023b, 2023c).

Appropriate surveillance methods according with the current epidemiological contexts are crucial to obtain quality and timely information for public health institutions for a rapid disease risk assessment and response (M. Braks et al., 2011). The European Centre for Disease and Control is the reference institution in this regard and promotes the strengthen of preparedness and response nets in front of infectious diseases in Europe. One of its goals was the creation of a harmonised framework for vector surveillance among the European countries by elaborating reference guidelines for native and invasive mosquito species (ECDC, 2012; Schaffner et al., 2014).

A review of the traditional methods versus novel technological solutions for mosquito surveillance was carried out in the **chapter 1** of the present thesis. The assortment of mosquito collection methods is very diverse, and their suitability depend on the targeted species so as on the requested information about the vector (e.g., presence, distribution, status, abundance, seasonal activity, longevity, etc) (ECDC, 2014). Many studies are focused on monitoring adult mosquitoes rather than immature stages, probably because they are easier to monitor, collect and identify. Also, adult sampling has shown to perform better than larval indices for measuring the seasonal variation in mosquito abundance (Codeço et al., 2015) and for predicting disease risk (Leandro et al., 2022). When searching for vector longevity, dispersal, biting behaviour, etc, collection of adult mosquitoes is the only option. Active surveillance of adult mosquitoes is mainly performed by the means of suction traps (e.g., BG-Sentinel, BG-Mosquitaire, CDC trap or EVS trap, among others), normally baited with CO_2 or chemical attractant.

In routine mosquito surveillance programs, traps are typically revised twice a month during mosquito season and the content of the catch bags is periodically sent to the laboratory where it is inspected for mosquito classification. The collected samples are usually composed by a big amount of non-target insects, which must be manually separated from mosquitoes. Mosquitoes are then taxonomically identified until the species level and usually sorted by sex, with the aid of identification keys or stereoscopic microscopes if necessary. As exposed, traditional methods for mosquito surveillance are very labour intensive and require expert entomologists to perform the taxonomical classification of the specimens as well as a minimum of laboratory infrastructure. The great quantity of samples, which need to be timely processed, generates a heavy workload to professionals in charge. Besides, the inevitable time lag between the trap is placed and the data is processed can affect the accurate representation of the true time dynamics of mosquito populations thus compromising a rapid response if needed.

The emergence of new technologies applied to the automated classification of mosquitoes have the chance to profoundly impact the current state of art of mosquito surveillance by overcoming some of the major limitations of traditional methods. Several techniques have been described so far to this purpose based on the analysis of the acoustical properties of mosquitoes using either acoustical or optical sensors and ML models (Santos et al., 2019). Also, new approaches based on image-recognition and deep learning have recently arisen for the classification of vector mosquitoes (Armin Pataki et al., 2021; Kittichai et al., 2023; J. Park et al., 2020).

In the present thesis, a new classification system based on an optical sensor coupled to a mosquito trap and trained with ML was developed for the automated remote monitoring of *Aedes* and *Culex* mosquitoes. In **chapter 2**, a novel optoelectronic sensor prototype was used to record a dataset of 4335 mosquito flights from which five different flight features were extracted and used to train five different ML algorithms in search of the best classification accuracy for mosquito genus and sex in laboratory conditions. The system was able to classify *Aedes* and *Culex* mosquitoes by genus with 94.2% of accuracy and to classify their sex with 99.4% of accuracy in *Aedes* and 100% of accuracy in *Culex.* This is in concordance with other previous studies which have attempted genus, sex or species classification of free-flying mosquitoes by using optical sensors (Batista et al., 2011; Fanioudakis et al., 2018; Genoud et al., 2018, 2020; Ouyang et al., 2015; Potamitis & Rigakis, 2016a; Silva et al., 2015).

The use of mosquito bioacoustics for characterization and classification purposes has been an active field since the first half of the 20th century (Kahn & Offenhauser, 1949), specially by using wing beat fundamental frequency (Santos et al., 2019). However, this feature has been demonstrated to be insufficient as the only predictor variable for complex classification tasks where there are overlapping frequency distributions such as species identification (Chen et al., 2014; Genoud et al., 2018; Kim et al., 2021). In the **chapter 2**, while the wing beat fundamental frequency alone gave high accuracy values for sex classification (95.5% in *Aedes* and 98% in *Culex*), it scored lower for genus classification (67.3%). In this case, the use of more complex features such as the spectrogram provided the maximum performance for genus classification (94.2%) when trained with DNN algorithm. The combination of the spectrogram feature with deep learning architectures have also been employed in the literature for species classification reporting good accuracy results (Fanioudakis et al., 2018;

Fernandes et al., 2021; Khalighifar et al., 2022; Wei et al., 2022). For sex classification, the use of other extracted flight features such as the spectrogram, the MFCC or the PSD in combination with different ML algorithms (LR, GB or SVM, among others) also provided better accuracy results for this task in both genera (99.4% for *Aedes* and 100% for *Culex*).

The main difference of our proposed system with respect to most of previous classification studies (Batista et al., 2011; Fanioudakis et al., 2018; Genoud et al., 2018, 2020; Ouyang et al., 2015; Potamitis & Rigakis, 2016a; Silva et al., 2015) is that, in our case, the sensor was coupled to a standard mosquito trap, and mosquitoes were recorded in a forced flight while they were sucked by the fan inside the trap, meanwhile the previous studies were performed in free-flight. As reported in another study of an automated mosquito counting system coupled to a suction trap (J. Wang et al., 2020), the airflow generated by the fan makes the mosquitoes pass through the sensing zone of the sensor quickly and produces an impact on mosquitoes' flight, which adds certain complexity to the automated classification. Besides, the big number of data samples required for ML models and the fact that in our system each recorded event corresponded to a single mosquito made the data gathering task something labour-intensive. As a counterpart, the system's performance showed its potential implementation in routine field mosquito surveillance and offer the possibility of capturing the recorded mosquitoes, which is very useful for vector surveillance, e.g., to confirm the detected species or for pathogen screening.

Despite the growing research about automated remote sensing in the field of entomology, there are very few available examples of automated mosquito classification systems which have been specifically tested in the field (Day et al., 2020; Lai et al., 2022). In **chapter 3**, the performance of the automated classification system described in **chapter 2** was assessed in the field for genus and sex classification of *Aedes* and *Culex* mosquitoes. In this case, a new dataset composed by 14,067 flights recorded in the laboratory at different temperature regimes and larval density conditions (to simulate the heterogeneity of the field) was used to build a ML model with the extracted spectrogram feature and GB algorithm, reporting an accuracy of 93.9% in laboratory conditions. Then, the classification system composed by the sensor, the trap and the associated ML model, was tested in two field sites during periods of mosquito activity. The accuracy in the field was evaluated by comparing the automated classification results reported by the system with those from the traditional inspection of the samples carried out by a trained entomologist.

The developed system tested in the field distinguished the designated target mosquitoes (from *Aedes* and *Culex* genus) from other flying insects that entered the traps with an average BA of 95.5%. This result suggests that the system would be suitable for a range of mosquito surveillance and control activities such as: i) identifying the start and end of a mosquito activity season; ii) monitoring seasonal tendencies of endemic vectors to prioritize geographical areas of intervention; iii) performing quality control checks of control measures aimed at reducing mosquito populations; or iv) identifying the arrival of invasive mosquitoes due to changes in land use, the effects of globalization, or due to climate change. A previous work, which analysed the performance in the field of an automated mosquito counting system (the BG-Counter), reported that the accuracy of mosquito detection depended on the relative abundance of mosquitoes in the catch with respect to other non-target insects, giving high levels of accuracy only when the percentage of mosquitoes in the sample was high (89%) (Day et al., 2020) .In our field study, the BA results ranged from 92.9% to 100%, even with a proportion of target mosquitos as low as 3%, which indicates that the system also performed accurate target mosquito detection with a low rate of false positives.

When the system identified an entering insect as a mosquito, it automatically classified it by genus and sex in the following categories of *Aedes* female, *Aedes* male, *Culex* female and *Culex* male with an average BA of 88.8%. The fact that this field accuracy was only slightly below the one obtained for the laboratory (93.9%) indicates that the ML model has generalized well to mosquitoes in the field. These results outperformed those from a previous study based on a next-generation trap (the MS-300) containing an infrared window which detected *Ae. albopictus* and *Cx. quinquefasciatus* with a mean accuracy of 64.9% and 79.4% depending on the sampling location (Lai et al., 2022). Our high accuracy results for genus and sex classification of *Aedes* and *Culex* mosquitoes in the field are of great utility for public health agencies and biological research in order to: i) detect possible introductions of *Aedes* invasive mosquitoes in new areas; ii) evaluate the effectiveness of control strategies based on mosquito modification such as the Sterile Male Technique or Wolbachia infection, that target a specific mosquito genus; or iii) monitor population dynamics of *Aedes* and *Culex* mosquitoes as an indicator of their vectorial capacity for arboviruses. Furthermore, the high time resolution of the system allowed the representation of the daily and seasonal activity patterns of *Aedes* and *Culex* mosquitoes in real-time.

Overall, the system provides several advantages with respect to conventional manual surveillance methods described in chapter 1: i) it promises a significant reduction in manual effort to gather and inspect each catch bag, especially when target mosquitoes must be sorted from a large number of non-target insects, and to manually record the results; ii) it provides classification results much earlier than what is possible in routine monitoring programs with collection cycles of 7 to 15 days, enabling a faster epidemiological response when needed; iii) it is not subject to the effects of predation and degradation of the sample; iv) it associates a time of capture stamp to each classification result, enabling the activity dynamics of the target insects to be determined with time resolutions down to one second; v) the server provides automated results in the form of tables and graphs which may be downloaded or visualized on the server itself and may feed risks maps, via the application programming interface. Accurate fine-scale time modelling of vector populations may be crucial for forecasting disease risk (Bartlow et al., 2019) so as for designing targeted precision control actions (Fouet & Kamdem, 2018), thus reinforcing an early warning response in front of MBD.

Since the age of mosquitoes is a critical determinant of their VC to transmit pathogens (Kramer & Ciota, 2015), targeting age through vector control is strategic to reduce the impact of MBD. Then, the ability to characterize the age of wild mosquitoes provide entomological evidence of the epidemiological impacts of these control strategies aimed at reducing vector lifespan (O'Neill et al., 2008). Despite decades of research in age-grading methods (Johnson et al., 2020) and the relevance of *Cx. pipiens* as a vector (ECDC, 2020), there was a lack of updated methods to estimate the age of *Cx. pipiens*. To fulfil this gap, in **chapter 4**, two techniques were described to this purpose: the transcriptional age-grading technique, a molecular method based on the analysis of the differential expression of age-responsive genes across age categories; and an automated classification system which discriminated between age classes based on differences in mosquito flight pattern.

Previous transcriptional profiling studies in *An. gambiae* (Cook & Sinkins, 2010), *Ae. aegypti* (Cook et al., 2006)and *Ae. albopictus* (Weeraratne et al., 2021), demonstrated that the expression of several genes displayed large variation across age categories in adult mosquitoes, signalling three of them (CUT, SCP, and CELL) as the most informative for age prediction (Cook et al., 2006). In the present chapter, these three genes were tested for *Cx. pipiens*, although in this case only two of them (CUT and SCP) reported significant changes with age. Since the SCP was the most informative gene to explain the variance in age of *Cx. pipiens*, it was used as the only predictor variable for age prediction, thus simplifying previous approaches based on multivariate analysis (Cook et al., 2007) and making it easier and cost-effective to be implemented. The cross-validated model generated showed a MAE of 3.8 days in age prediction. This error was not equally distributed between age classes: for mosquitoes ≤ 4 days old, the MAE was 1 day, and the confidence intervals were between 0 and 5.6 days; however, for mosquitoes ≥ 5 days old, the MAE was 5.0 and the confidence intervals ranged from -3.7 to 24.9 in the intermediate age classes and from 7.0 to 33.9 in the older age classes. Higher error values for age prediction in older age classes occurred due to the inexistence of significant differences in the predictor variables between these classes. The accuracy was quite similar to other gene profiling studies based on multivariate calibration such as Cook et al. (2006) which reported a discrepancy of 5 days in age estimations for *Ae. aegypti*; Weeraratne et al. (2021) which reported a MAE of 2.19 (± 1.66) and 2.58 (± 2.06) days in age prediction for *Ae. aegypti* and *Ae. albopictus*, respectively; and Hugo et al. (2010) which reported in 77% of cases a difference of 6 days in age prediction for *Ae. aegypti* in a field study.

Regarding the automated mosquito classification system, while other studies based on the analysis of mosquito wing-beat pattern through optical sensors have demonstrated to be accurate to taxonomically classify mosquitoes and other attributes of their biology in free-flight (G. E. A. P. A. Batista et al., 2011; Fanioudakis et al., 2018; Genoud et al., 2018, 2019, 2020; Potamitis & Rigakis, 2016b; Silva et al., 2015), automated age classification had never been assessed before. In the present chapter, the system composed of an entomological optical sensor coupled to a mosquito trap (previously assessed in **chapters 2** and **3**) was tested for age classification in *Culex pipiens* female mosquitoes. The extracted MFCC feature was trained with XGBoost, providing a mean balanced accuracy of 75% in distinguishing young (2-4 days) from older (7-16 days) mosquitoes. XGBoost was already used in our previous study reporting the maximum accuracy with MFCC for mosquito sex classification (González-Pérez et al., 2022), and also reported good accuracy results for mosquito species classification in other laboratory works (Fanioudakis et al., 2018).

Both methods developed in **chapter 4** to determine the age of *Cx. pipiens* female mosquitoes (the transcriptional age grading technique and the automated classification system) were able to perform a major binary classification between young mosquitoes (≤ 4 days old) and older mosquitoes with reasonable levels of accuracy. According with vector competence studies of *Cx. pipiens* for WNV (Vogels et al., 2017a), this binary classification may serve to discriminate between two functional groups of non-vector and potential vector mosquitoes which are old enough to have overcome the EIP of the virus (mainly superior to 7 days at temperatures from 18 to 27ºC) (Anderson et al., 2008; Kilpatrick et al., 2008; Vogels, Göertz, et al., 2017a, 2017b). Therefore, the capacity to differentiate between two populations of mosquitoes of different ages may be useful to assess WNV control interventions targeting vector longevity. Further steps to consolidate these methodologies should include semi-field studies prior to their application into real field work scenario.

While in recent years the interest in ML approaches for mosquito surveillance has skyrocketed with an upward trend of published papers aimed at identifying mosquitoes through a variety of sensing devices Joshi & Miller, 2021), the automated recognition of mosquitoes based on their infection status through optical sensors had not been attempted before, until the present work. In **chapter 5**, the automated classification of arbovirus-infected mosquitoes captured by a mosquito trap was assessed by the analysis of their wing-beat flight pattern through an optical sensor trained with ML. The initial work hypothesis began on the basis that a disseminated infection of ZIKV in the competent vector *Ae. albopictus* (Mckenzieid et al., 2019) may generate an affectation in the nervous system that would lead to an alteration in the mosquito behaviour, as previously reported for ZIKV-infected *Ae. aegypti* (Gaburro et al., 2018a; Pedreira Padilha et al., 2018)and this alteration in the flight performance could be registered by an optical sensor, which in combination with ML techniques, could serve to differentiate infected from non-infected mosquitoes.

In the experimental infection assay carried out at the BSL3 facilities of IRTA-CReSA, a total of 843 females of *Ae. albopictus* were inoculated (433 with ZIKV and 410 with DMEM as a control group). All the ZIKV IIT *Ae. albopictus* females showed a disseminated ZIKV infection at 11-12 dpi and IHQ analysis confirmed the tropism of ZIKV for *Ae. albopictus* nervous tissues, showing virus replication in the head and thoracic ganglia (mainly in the cortical layer) and in peripherical sensorial structures (especially in the ommatidia). These results were in concordance with previous IHQ analysis in *Ae. aegypti* at two weeks post-infection with ZIKV (Gaburro et al., 2018a). Virus replication found in the central nervous system of *Ae. albopictus* may disrupt neuronal communication and affect mosquito locomotor activity as suggested previously for *Ae. aegypti* (Gaburro et al., 2018). The high virus detection in ommatidia suggests that the vision of *Ae. albopictus* may be affected when infected with ZIKV, thus possibly compromising mosquito behaviours mediated by visual stimuli such as the optomotor response, landscape navigation, swarming, and host-seeking behaviours (Hawkes et al., 2022). These infection-driven changes in mosquito rhythmic patterns may have an impact on its vectorial capacity, and consequently, on arbovirus transmission (Vieira Bruno et al., 2020).

The automated classification system (already described in **chapters 2-4**) recorded 421 mosquitoes that entered the trap during the flight assays (212 infected with ZIKV and 209 from the control group). The system was able to discriminate *Ae. albopictus* mosquitoes with a disseminated ZIKV infection from non-infected ones with 63.0% of classification accuracy. This result points out that ZIKV infection in *Ae. albopictus* may generate an alteration in the mosquito flight pattern that could be identified with an optical sensor thus enabling an automated classification through a ML model. According to the training score, the accuracy of the model could be improved with more training samples which encourages further testing in search of better performance. This study provides preliminary evidence of how new technologies have the chance to enhance the capabilities of mosquito vector surveillance thus allowing a timelier disease risk assessment.

Along the present thesis, an automated classification system consisting of an optical sensor coupled to a mosquito suction trap was trained on a large database of mosquito flights from the vector species *Ae. albopictus* and *Cx. pipiens* at different raising and environmental conditions for different classification tasks: i) the identification of mosquito genus and sex in laboratory conditions (**chapter 2**); ii) the identification of mosquitoes over other non-Culicidae insects and the classification of the genus and sex of those mosquitoes in the field (**chapter 3**); iii) the estimation of mosquito's age (**chapter 4**); and iv) the identification of arbovirus-infected mosquitoes (**chapter 5**). The accuracy of the system was dependent on the classification task pursued, so as in the number of samples used for training the model, and the combination of extracted flight features and ML algorithms used. Some classification tasks such as genus and sex classification, which obtained high accuracy results in laboratory conditions, were tested in the field as a proof of concept of the system's high-performance results in a real field work scenario. Other tasks such as the classification of mosquitoes by age or infection status, which were never assessed before by the means of optical sensors, would still require more training samples to improve their classification accuracy, but promising results were obtained.

Overall, the present thesis has contributed to increase the knowledge about automated remote monitoring methods for mosquito surveillance by developing an automated classification system coupled to a mosquito trap which monitors *Aedes* and *Culex* mosquitoes in near real-time thus enabling a fine-time scale modelling of vector populations, which is crucial for MBD assessment, and has paved the way for future vector biology research.

6. GENERAL CONCLUSIONS

1. The development of new technologies applied for the automated remote monitoring of mosquito populations in real-time represents a great improvement with respect to traditional surveillance methods by eliminating the intensive workload, the need of experts and providing unprecedented time-resolution.
2. The automated classification system assessed in the present thesis, composed by an optical sensor coupled to a mosquito trap, is highly accurate to classify the genus and sex of *Aedes* and *Culex* mosquitoes in laboratory conditions.
3. The spectrogram feature extracted from mosquito flight recordings reported the highest accuracy for genus and sex classification in laboratory conditions when trained with dense neural networks and logistic regression respectively. MFCC provided the best accuracy for sex classification when trained with logistic regression and gradient boosting. This supports the use of these features and algorithms for future mosquito classification tasks.
4. The automated classification system allows to distinguish *Aedes* and *Culex* mosquitoes from other non-targeted insects also captured by the trap, and to automatically classify them by genus and sex in the field with high accuracy, which is a proof of concept that this technology is suitable for field applications of vector surveillance.

5. The accuracy of target mosquito detection in the field is not correlated with the proportion of non-target insects that entered the trap, meaning that the system is efficient to correctly identify mosquitoes over other non-Culicidae insects with a low rate of false positives, even when the proportion of mosquitoes is very low.
6. The accuracy of genus and sex classification in the field is very similar to the one obtained in the laboratory at different environmental and rearing conditions. This fact indicates that the applied ML model generalises well in the field, highlighting the importance of integrating biological and environmental variability in mosquito classification studies.
7. The representation of the daily and seasonal activity patterns of *Aedes* and *Culex* mosquitoes by the automated classification system in the field paves the way to improve the knowledge about mosquito behaviour, which may favour the elaboration of more accurate prediction models.
8. The two methodologies developed to estimate the age of *Cx. pipiens* female mosquitoes are useful to distinguish between young mosquitoes (≤ 4 days of age) and older ones, which may be potential WNV vectors. The capacity to classify mosquitoes by the age may be useful to assess the risk of WNV transmission and the effectiveness of control interventions targeting vector longevity.
9. ZIKV infection in nervous tissues of *Ae. albopictus*, especially in the cortical layer of the thoracic and head ganglia and the ommatidia, is possibly modifying mosquito's flight, which could

be registered by the automated classification system. The moderate accuracy achieved to differentiate ZIKV-infected from non-infected *Ae. albopictus* offers promising perspective for the future of vector surveillance, although further assays are needed to improve the accuracy of the system.

7. REFERENCES

Aardema, M. L., Olatunji, S. K., & Fonseca, D. M. (2022). The enigmatic *Culex pipiens* (Diptera: Culicidae) species complex: phylogenetic challenges and opportunities from a notoriously tricky mosquito group. *Annals of the Entomological Society of America*, *115*(1), 95–104. https://doi.org/10.1093/AESA/SAAB038

Agarwal, A., Parida, M., & Dash, P. K. (2017). Impact of transmission cycles and vector competence on global expansion and emergence of arboviruses. *Reviews in Medical Virology*, 27(5). https://doi.org/10.1002/rmv.1941

Aldersley, A., & Cator, L. J. (2019). Female resistance and harmonic convergence influence male mating success in Aedes aegypti. *Scientific Reports*, *9*(1), 1–12. https://doi.org/10.1038/s41598-019-38599-3

Aldersley, A., Champneys, A., Homer, M., & Robert, D. (2016). Quantitative analysis of harmonic convergence in mosquito auditory interactions. *Journal of the Royal Society Interface*, *13*(117). https://doi.org/10.1098/RSIF.2015.1007

Anderson, J. F., Main, A. J., Delroux, K., & Fikrig, E. (2008). Extrinsic incubation periods for horizontal and vertical transmission of West Nile virus by *Culex pipiens pipiens* (Diptera: Culicidae). *Journal of Medical Entomology*, *45*(3), 445–451. https://doi.org/10.1093/JMEDENT/45.3.445

Aranda, C., Eritja, R., & Roiz, D. (2006). First record and establishment of the mosquito Aedes albopictus in Spain. *Medical and Veterinary Entomology*, *20*(1), 150–152. https://doi.org/10.1111/J.1365-2915.2006.00605.X

Arias-Vergara, T., Klumpp, P., Vasquez-Correa, J. C., Nöth, E., Orozco-Arroyave, J. R., & Schuster, M. (2020). Multi-channel spectrograms for speech processing applications using deep learning methods. *Pattern Analysis and Applications 2020 24:2*, *24*(2), 423–431. https://doi.org/10.1007/S10044-020-00921-5

Armin Pataki, B., Garriga, J., Eritja, R., Palmer, J. R. B., Bartumeus, F., & Csabai, I. (2021). Deep learning identification for citizen science surveillance of tiger mosquitoes. *Scientific Reports* , *11*. https://doi.org/10.1038/s41598-021-83657-4

Arthur, B. J., Emr, K. S., Wyttenbach, R. A., & Hoy, R. R. (2014). Mosquito (*Aedes aegypti*) flight tones: Frequency, harmonicity, spherical spreading, and phase relationships . *The Journal of the Acoustical Society of America*, *135*(2), 933–941. https://doi.org/10.1121/1.4861233

Aryaprema, V. S., Steck, M. R., Peper, S. T., Xue, R. De, & Qualls, W. A. (2023). A systematic review of published literature on mosquito control action thresholds across the world. *PLoS Neglected Tropical Diseases*, *17*(3). https://doi.org/10.1371/journal.pntd.0011173

Azar, S. R., & Weaver, S. C. (2019). Vector Competence: What Has Zika Virus Taught Us? *Viruses*, *11*(9), 867. https://doi.org/10.3390/v11090867

Baker, R. E., Mahmud, A. S., Miller, I. F., Rajeev, M., Rasambainarivo, F., Rice, B. L., Takahashi, S., Tatem, A. J., Wagner, C. E., Wang, L.-F., Wesolowski, A., Jessica, C., & Metcalf, E. (2022). Infectious disease in an era of global change. *Nature Reviews. Microbiology*, *20*, 193–205. https://doi.org/10.1038/s41579-021-00639-z

Bakonyi, T., & Haussig, J. M. (2020). West Nile virus keeps on moving up in Europe. *Euro Surveillance : Bulletin Europeen Sur Les Maladies Transmissibles = European Communicable Disease Bulletin*, *25*(46). https://doi.org/10.2807/1560-7917.ES.2020.25.46.2001938

Barker, C. M., & Reisen, W. K. (2018). Epidemiology of vector-borne diseases. *Medical and Veterinary Entomology* (pp. 33–49). Elsevier. https://doi.org/10.1016/B978-0-12-814043-7.00004-2

Bartlow, A. W., Manore, C., Xu, C., Kaufeld, K. A., Del Valle, S., Ziemann, A., Fairchild, G., & Fair, J. M. (2019). Forecasting Zoonotic Infectious Disease Response to Climate Change: Mosquito Vectors and a Changing Environment. *Veterinary Sciences*, *6*(40). https://doi.org/10.3390/vetsci6020040

Batista, G. E. A. P. A., Hao, Y., Keogh, E., & Mafra-Neto, A. (2011). Towards automatic classification on flying insects using inexpensive sensors. *International Conference on Machine Learning and Applications*. https://ieeexplore-ieee-org.are.uab.cat/stamp/stamp.jsp?tp=&arnumber=6146999

Batista, G., Keogh, E., Neto, A. M., & Rowton, E. (2011). SIGKDD Demo: Sensors and software to allow computational entomology, an emerging application of data mining. *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, 761–764. https://doi.org/10.1145/2020408.2020530

Becker, N., Petrić, D., Zgomba, M., Boase, C., Madon, M. B., Dahl, C., & Kaiser, A. (2020). Biology of Mosquitoes. *Mosquitoes. Identification, Ecology and Control* (3rd ed.). https://doi.org/10.1007/978-3-030-11623-1_2

BiPRO, & EC. (2009). Development of guidance for establishing Integrated Pest Management (IPM) principles.

Bisina, K. V., & Azeez, M. A. (2017). Optimized estimation of power spectral density. *Proceedings of the 2017 International Conference on Intelligent Computing and Control Systems, ICICCS 2017*, *2018-January*, 871–875. https://doi.org/10.1109/ICCONS.2017.8250588

Bonizzoni, M., Gasperi, G., Chen, X., & James, A. A. (2013). The invasive mosquito species *Aedes albopictus*: current knowledge and future perspectives. *Trends in Parasitology*. https://doi.org/10.1016/j.pt.2013.07.003

Boyer, S., Calvez, E., Chouin-Carneiro, T., Diallo, D., & Failloux, A. B. (2018). An overview of mosquito vectors of Zika virus. *Microbes and Infection*, *20*(11–12), 646–660. https://doi.org/10.1016/J.MICINF.2018.01.006

Braks, M. A. H., Honório, N. A., Lounibos, L. P., Lourenço-De-Oliveira, R., & Juliano, S. A. (2004). Interspecific competition between two invasive species of container mosquitoes, *Aedes aegypti* and *Aedes* *albopictus* (Diptera: Culicidae), in Brazil. *Annals of the Entomological Society of America* , *97*(1), 130–139. https://academic.oup.com/aesa/article-abstract/97/1/130/11472?login=false

Braks, M., Van Der Giessen, J., Kretzschmar, M., Van Pelt, W., Scholte, E. J., Reusken, C., Zeller, H., Van Bortel, W., & Sprong, H. (2011). Towards an integrated approach in surveillance of vector-borne diseases in Europe. *Parasites and Vectors*, *4*(1), 1–11. https://doi.org/10.1186/1756-3305-4-192/FIGURES/4

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Brodersen, K. H., Soon Ong, C., Stephan, K. E., & Buhmann, J. M. (2010). The balanced accuracy and its posterior distribution. https://doi.org/10.1109/ICPR.2010.764

Brosa, J. (2018). Evaluation of a novel sensor system integrated into a mosquito trap to determine mosquito species, age and sex [Universitat Autònoma de Barcelona]. https://ddd.uab.cat/record/210874

Brown, H. E., Paladini, M., Cook, R. A., Kline, D., Barnard, D., & Fish, D. (2008). Effectiveness of mosquito traps in measuring species abundance and composition. *Journal of Medical Entomology*, *45*(3), 517–521. https://doi.org/10.1093/jmedent/45.3.517

Brugman, V. A., Hernández-Triana, L. M., Medlock, J. M., Fooks, A. R., Carpenter, S., & Johnson, N. (2018). The role of *Culex pipiens* L. (Diptera: Culicidae) in virus transmission in Europe. *Int. J. Environ. Res. Public Health,* 15(2). https://doi.org/10.3390/ijerph15020389

Brydegaard, M., & Lazzari, C. R. (2015). Towards quantitative optical cross sections in entomological laser radar - Potential of temporal and spherical parameterizations for identifying atmospheric fauna. *PLoS ONE*, *10*(8). https://doi.org/10.1371/journal.pone.0135231

Bueno, M. G., Martinez, N., Abdalla, L., Duarte dos Santos, C. N., & Chame, M. (2016). Animals in the Zika virus life cycle: What to expect from megadiverse Latin American countries. *PLOS*

Neglected Tropical Diseases, *10*(12), e0005073. https://doi.org/10.1371/JOURNAL.PNTD.0005073

Burges, C. J. C. (1998). A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, *2*(2), 121–167. https://doi.org/10.1023/A:1009715923555

Cadar, D., Lühken, R., van der Jeugd, H., Garigliany, M., Ziegler, U., Keller, M., Lahoreau, J., Lachmann, L., Becker, N., Kik, M., Oude Munnink, B. B., Bosch, S., Tannich, E., Linden, A., Schmidt, V., Koopmans, M. P., Rijks, J., Desmecht, D., Groschup, M. H., … Schmidt-Chanasit, J. (2017). Widespread activity of multiple lineages of Usutu virus, western Europe, 2016. *Euro Surveillance : Bulletin Europeen Sur Les Maladies Transmissibles = European Communicable Disease Bulletin*, *22*(4). https://doi.org/10.2807/1560-7917.ES.2017.22.4.30452

Calzolari, M. (2016). Mosquito-borne diseases in Europe: an emerging public health threat. *Reports in Parasitology*, 1. https://doi.org/10.2147/RIP.S56780

Cansado-Utrilla, C., Zhao, S. Y., McCall, P. J., Coon, K. L., & Hughes, G. L. (2021). The microbiome and mosquito vectorial capacity: rich potential for discovery and translation. *Microbiome 2021 9:1*, *9*(1), 1–11. https://doi.org/10.1186/S40168-021-01073-2

Caputo, B., & Manica, M. (2020). Mosquito surveillance and disease outbreak risk models to inform mosquito-control operations in Europe. *Current Opinion in Insect Science*, *2020*, 101–108. https://doi.org/10.1016/j.cois.2020.03.009

Cator, L. J., Arthur, B. J., Harrington, L. C., & Hoy, R. R. (2009). Harmonic convergence in the love songs of the dengue vector mosquito. *Science*, *323*(5917), 1077–1079. https://doi.org/10.1126/science.1166541

Cator, L. J., Arthur, B. J., Ponlawat, A., & Harrington, L. C. (2011). Behavioral observations and sound recordings of free-flight mating swarms of Ae. aegypti (Diptera: Culicidae) in Thailand. *Journal of*

Medical Entomology, *48*(4), 941–946. https://doi.org/10.1603/ME11019

CDC. (2020). *Integrated Mosquito Management*. Centers for Disease Control and Prevention. https://www.cdc.gov/mosquitoes/mosquito-control/professionals/integrated-mosquito-management.html

CDC.(2023). Mosquitoes. Centers for Disease Control and Prevention. https://www.cdc.gov/mosquitoes/about/index.html

Chan, M., & Johansson, M. A. (2012). The Incubation Periods of Dengue Viruses. *PLoS ONE*, *7*(11), 50972. https://doi.org/10.1371/journal.pone.0050972

Charrel, R. N., de Lamballerie, X., & Raoult, D. (2007). Chikungunya outbreaks--the globalization of vectorborne diseases. *The New England Journal of Medicine*, *356*(8), 769–771. https://doi.org/10.1056/NEJMP078013

Chathurika Weeraratne, T., Parakrama Karunaratne, S. H. P., Reimer, L., Priyanka P. de Silva, W. A., & Wondji, C. S. (2021). Use of transcriptional age grading technique to determine the chronological age of Sri Lankan *Aedes aegypti* and *Aedes albopictus* females. *Parasites & Vectors*, *14*(493). https://doi.org/10.1186/s13071-021-04994-x

Chatterjee, S. (2014). Artefacts in histopathology. *Journal of Oral and Maxillofacial Pathology*, *18*(4). https://doi.org/10.4103/0973-029X.141346

Chen, T., & Guestrin, C. (2016). *XGBoost: A Scalable Tree Boosting System*. https://doi.org/10.1145/2939672.2939785

Chen, Y., Why, A., Batista, G., Mafra-Neto, A., & Keogh, E. (2014). Flying Insect Classification with Inexpensive Sensors. *Journal of Insect Behavior*, *27*, 657–677. https://doi.org/10.1007/s10905-014-9454-4

Ciota, A. T., Styer, L. M., Meola, M. A., & Kramer, L. D. (2011). The costs of infection and resistance as determinants of West Nile virus susceptibility in Culex mosquitoes. *BMC Ecology*, *11*(1), 1–8. https://doi.org/10.1186/1472-6785-11-23/TABLES/2

Claudel, I., Brouazin, R., Lancelot, R., Gouagna, L.-C., Dupraz, M., Baldet, T., & Jérémy Bouyer, &. (2022). Optimization of adult mosquito trap settings to monitor populations of Aedes and Culex mosquitoes, vectors of arboviruses in La Reunion. *Scientific Reports 2022 12:1*, *12*(1), 1–10. https://doi.org/10.1038/s41598-022-24191-9

Codeço, C. T., Lima, A. W. S., Araújo, S. C., Lima, J. B. P., Maciel-de-Freitas, R., Honório, N. A., Galardo, A. K. R., Braga, I. A., Coelho, G. E., & Valle, D. (2015). Surveillance of Aedes aegypti: Comparison of House Index with four alternative traps. *PLOS Neglected Tropical Diseases*, *9*(2), e0003475. https://doi.org/10.1371/JOURNAL.PNTD.0003475

Cook, P. E., Hugo, L. E., Aki Iturbe-Ormaetxe, I., Williams, C. R., Chenoweth, S. F., Ritchie, S. A., Ryan, P. A., Kay, B. H., Blows, M. W., & O’neill, S. L. (2006). The use of transcriptional profiles to predict adult mosquito age under field conditions. *PNAS.* http://doi.org/10.1073pnas.0604875103

Cook, P. E., Hugo, L. E., Iturbe-Ormaetxe, I., Williams, C. R., Chenoweth, S. F., Ritchie, S. A., Ryan, P. A., Kay, B. H., Blows, M. W., & O’Neill, S. L. (2007). Predicting the age of mosquitoes using transcriptional profiles. *Nature Protocols*, *2*(11), 2796–2806. https://doi.org/10.1038/nprot.2007.396

Cook, P. E., Mcmeniman, J., & O’neill, S. L. (2008). *Modifying Insect Population Age Structure to Control Vector-Borne Disease*. *Advances in Experimental Medicine and Biology,* 627, 126-140. http://doi.org/ 10.1007/978-0-387-78225-6_11/COVER

Cook, P. E., & Sinkins, S. P. (2010). Transcriptional profiling of Anopheles gambiae mosquitoes for adult age estimation. *Insect Molecular Biology*, *19*(6), 745–751. https://doi.org/10.1111/j.1365-2583.2010.01034.x

Counotte, M. J., Kim, C. R., Wang, J., Bernstein, K., Deal, C. D., Broutet, N. J. N., & Low, N. (2018). Sexual transmission of Zika virus and other flaviviruses: A living systematic review. *PLOS Medicine*, *15*(7), e1002611. https://doi.org/10.1371/JOURNAL.PMED.1002611

Cramer, M. M., Gabel, T. M., & Duvall, L. B. (2023). Characterizing physical interactions between male and female mosquitoes (*Aedes aegypti*) in relation to female receptivity and insemination outcomes using a hydrophobic fluorescent dye. *Integrative and Comparative Biology*, *63*(2), 382–392. https://doi.org/10.1093/icb/icad041

Day, C. A., Richards, S. L., Reiskind, M. H., Doyle, M. S., & Byrd, B. D. (2020). Context-dependent accuracy of the bg-counter remote mosquito surveillance device in north carolina. *Journal of the American Mosquito Control Association*, *36*(2), 74–80. https://doi.org/10.2987/19-6903.1

De Lamballerie, X., Leroy, E., Charrel, R. N., Ttsetsarkin, K., Higgs, S., & Gould, E. A. (2008). Chikungunya virus adapts to tiger mosquito via evolutionary convergence: A sign of things to come? *Virology Journal*, *5*(1), 1–4. https://doi.org/10.1186/1743-422X-5-33/FIGURES/2

de Meillon, B., & Khan, Z. H. (1967). Examples of the use of simple age-grading in the assessment of Culex pipiens fatigans populations. *Bulletin of the World Health Organization*, *36*(1), 169. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2476341/

Delamater, P. L., Street, E. J., Leslie, T. F., Yang, Y. T., & Jacobsen, K. H. (2019). Complexity of the basic reproduction number (R0). *Emerging Infectious Diseases*, *25*(1), 1–4. https://doi.org/10.3201/eid2501.171901

Dente, M. G., Riccardo, F., Bolici, F., Colella, N. A., Jovanovic, V., Drakulovic, M., Vasic, M., Mamlouk, H., Maazaoui, L., Bejaoui, M., Zakhashvili, K., Kalandadze, I., Imnadze, P., Declich, S., Knjeginic, V., Stojkovic, B., Labus, T., Milicevic, V., Veljovic, L., … Kekelidze, A. (2019). Implementation of the One Health approach to fight arbovirus infections in the Mediterranean and Black Sea Region:

Assessing integrated surveillance in Serbia, Tunisia and Georgia. *Zoonoses and Public Health*, *66*(3), 276–287. https://doi.org/10.1111/ZPH.12562

Di Luca, M., Severini, F., Toma, L., Boccolini, D., Romi, R., Remoli, M. E., Sabbatucci, M., Rizzo, C., Venturi, G., Rezza, G., & Fortuna, C. (2016). Experimental studies of susceptibility of Italian Aedes albopictus to Zika virus. *Eurosurveillance*, *21*(18), 30223. https://doi.org/10.2807/1560-7917.ES.2016.21.18.30223/CITE/PLAINTEXT

Diabate, A., & Tripet, F. (2015). Targeting male mosquito mating behaviour for malaria control. In *Parasites and Vectors* (Vol. 8, Issue 1, pp. 1–13). BioMed Central Ltd. https://doi.org/10.1186/s13071-015-0961-8

Dicker, R. C., Coronado, F., Koo, D., & Parrish, R. G. (2006). *Principles of epidemiology in Public Health practice; an introduction to applied epidemiology and biostatistics.*

Dietz, K. (1993). The estimation of the basic reproduction number for infectious diseases. https://doi.org/10.1177/096228029300200103

Dohm, D. J., O’Guinn, M. L., & Turell, M. J. (2002). Effect of environmental temperature on the ability of *Culex pipiens* (Diptera: Culicidae) to transmit West Nile virus. *Journal of Medical Entomology*, *39*(1), 221–225. https://doi.org/10.1603/0022-2585-39.1.221

Dou, Z., Madan, A., Carlson, J. S., Chung, J., Spoleti, T., Dimopoulos, G., Cammarato, A., & Mittal, R. (2021). Acoustotactic response of mosquitoes in untethered flight to incidental sound. *Scientific Reports 2021 11:1*, *11*(1), 1–9. https://doi.org/10.1038/s41598-021-81456-5

Dubrulle, M., Mousson, L., Moutailier, S., Vazeille, M., & Failloux, A. B. (2009). Chikungunya Virus and Aedes Mosquitoes: Saliva Is Infectious as soon as Two Days after Oral Infection. *PLOS ONE*, *4*(6), e5895. https://doi.org/10.1371/JOURNAL.PONE.0005895

Duffy, M. R., Chen, T.-H., Hancock, W. T., Powers, A. M., Kool, J. L., Lanciotti, R. S., Pretrick, M., Marfel, M., Holzbauer, S., Dubray, C., Guillaumot, L., Griggs, A., Bel, M., Lambert, A. J., Laven, J., Kosoy, O., Panella, A., Biggerstaff, B. J., Fischer, M., & Hayes, E. B. (2009). Zika virus outbreak on Yap Island, Federated States of Micronesia. *The New England Journal of Medicine*, *360*(24), 2536–2543. https://doi.org/10.1056/NEJMOA0805715

ECDC. (2012). *Guidelines for the surveillance of invasive mosquitoes in Europe*. https://doi.org/10.2900/61134

ECDC. (2014). *Field sampling methods for mosquitoes, sandflies, biting midges and ticks*. https://doi.org/10.2900/416333

ECDC. (2016). *Aedes albopictus - Factsheet for experts*. European Centre for Disease Prevention and Control. https://www.ecdc.europa.eu/en/disease-vectors/facts/mosquito-factsheets/aedes-albopictus

ECDC. (2019). *West Nile virus infection - Annual Epidemiological Report for 2018*. https://www.ecdc.europa.eu/en/publications-data/west-nile-virus-infection-annual-epidemiological-report-2018

ECDC. (2020). *Culex pipiens - Factsheet for experts*. https://www.ecdc.europa.eu/en/infectious-disease-topics/related-public-health-topics/disease-vectors/facts/mosquito-factsheets

ECDC. (2021). *Zika virus disease - Annual Epidemiological Report for 2019*. https://www.ecdc.europa.eu/en/publications-data/zika-virus-disease-annual-epidemiological-report-2019

ECDC. (2022). *“Reverse” identification key for mosquito species*. https://www.medilabsecure.com/moskeytool.html.

ECDC. (2023a). *Aedes albopictus-current known distribution: August 2023*. Https://Www.Ecdc.Europa.Eu/En/Publications-Data/Aedes-Albopictus-Current-Known-Distribution-August-2023.

ECDC. (2023b). *Autochthonous transmission of chikungunya virus in mainland EU/EEA, 2007–present*. https://www.ecdc.europa.eu/en/infectious-disease-topics/z-disease-list/chikungunya-virus-disease/surveillance-threats-and

ECDC. (2023c). *Autochthonous vectorial transmission of dengue virus in mainland EU/EEA, 2010-present*. https://www.ecdc.europa.eu/en/all-topics-z/dengue/surveillance-and-disease-data/autochthonous-transmission-dengue-virus-eueea

Eliopoulos, P., Potamitis, I., & Kontodimas, D. (2016). Estimation of population density of stored grain pests via bioacoustic detection. *Crop Protection*, *85*, 71–78. https://doi.org/10.1016/j.cropro.2016.04.001

Eliopoulos, P., Tatlas, N. A., Rigakis, I., & Potamitis, I. (2018). A "smart" trap device for detection of crawling insects and other arthropods in urban environments. *Electronics (Switzerland)*, *7*(9). https://doi.org/10.3390/electronics7090161

EMCA. (n.d.). *Traps*. Retrieved October 22, 2023, from https://www.emca-online.eu/traps

Epelboin, Y., Talaga, S., Epelboin, L., & Dusfour, I. (2017). Zika virus: An updated review of competent or naturally infected mosquitoes. *PLOS Neglected Tropical Diseases*, *11*(11), e0005933. https://doi.org/10.1371/JOURNAL.PNTD.0005933

Fanioudakis, E., Geismar, M., & Potamitis, I. (2018). Mosquito wingbeat analysis and classification using deep learning. *26th European Signal Processing Conference (EUSIPCO)*.

Farajollahi, A., Fonseca, D. M., Kramer, L. D., & Kilpatrick, A. M. (2011). *"'Bird biting'" mosquitoes and human disease: A review of the role of Culex pipiens complex mosquitoes in epidemiology*. https://doi.org/10.1016/j.meegid.2011.08.013

Farajollahi, A., Kesavaraju, B., Price, D. C., Williams, G. M., Healy, S. P., Gaugler, R., & Nelder, M. P. (2009). Field Efficacy of BG-Sentinel and Industry-Standard Traps for Aedes albopictus (Diptera:

Culicidae) and West Nile Virus Surveillance. In *J. Med. Entomol* (Vol. 46, Issue 4). https://academic.oup.com/jme/article-abstract/46/4/919/939858

Farlow, R., Russell, T. L., & Burkot, T. R. (2020). Nextgen Vector Surveillance Tools: sensitive, specific, cost-effective and epidemiologically relevant. *Malaria Journal*, *19*(1), 1–13. https://doi.org/10.1186/S12936-020-03494-0/TABLES/7

Fernandes, M. S., Cordeiro, W., & Recamonde-Mendoza, M. (2021). Detecting Aedes aegypti mosquitoes through audio classification with convolutional neural networks. *Computers in Biology and Medicine*, *129*, 104152. https://doi.org/10.1016/j.compbiomed.2020.104152

Figuerola, J., Jiménez-Clavero, M. Á., Ruíz-López, M. J., Llorente, F., Ruiz, S., Hoefer, A., Aguilera-Sepúlveda, P., Jiménez-Peñuela, J., García-Ruiz, O., Herrero, L., Soriguer, R. C., Fernández Delgado, R., Sánchez-Seco, M. P., Martínez-de la Puente, J., & Vázquez, A. (2022). A One Health view of the West Nile virus outbreak in Andalusia (Spain) in 2020. https://doi.org/10.1080/22221751.2022.2134055

Flores, C. (2015). Mosquito Surveillance for Effective Mosquito Population Control International. Vector Disease Control International. http://www.vdci.net/blog/mosquito-surveillance-for-effective-mosquito-population-contro

Focks, D. A. (2003). *A review of entomological sampling methods and indicators for dengue vectors.*

Focks, D. A., & Barrera, R. (2006). Dengue transmission dynamics: assessment and implications for control. In *Report on Dengue*. http://www.who.int/tdr/publications/publications/swg_dengue_2.htm

Forrester, N. L., Coffey, L. L., & Weaver, S. C. (2014). Arboviral Bottlenecks and Challenges to Maintaining Diversity and Fitness during Mosquito Transmission. *Viruses 2014, Vol. 6, Pages 3991-4004*, *6*(10), 3991–4004. https://doi.org/10.3390/V6103991

Foster, W. A., & Walker, E. D. (2019). *Mosquitoes (Culicidae)*. https://doi.org/10.1016/B978-0-12-814043-7.00015-7

Fouet, C., & Kamdem, C. (2018). *Integrated Mosquito Management: Is Precision Control a Luxury or Necessity?* https://doi.org/10.1016/j.pt.2018.10.004

Franklinos, L. H. V, Jones, K. E., Redding, D. W., & Abubakar, I. (2019). The effect of global change on mosquito-borne disease. *Lancet Infect Dis*, *19*, 302–314. https://doi.org/10.1016/S1473-3099(19)30161-6

Freier, B., & Boller, E. F. (2009). Integrated pest management in Europe - History, policy, achievements and implementation. *Integrated Pest Management*, *2*, 435–454. https://doi.org/10.1007/978-1-4020-8990-9_14/COVER

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, *29*(5), 1189–1232. https://doi.org/10.1214/aos/1013203451

Fritz, M. L., Walker, E. D., Yunker, A. J., & Dworkin, I. (2014). Daily blood feeding rhythms of laboratory-reared North American Culex pipiens. *Journal of Circadian Rhythms*, *12*(1), 1–9. https://doi.org/10.1186/1740-3391-12-1/FIGURES/4

Gaburro, J., Bhatti, A., Harper, J., Jeanne, I., Dearnley, M., Green, D., Nahavandi, S., Paradkar, P. N., & Duchemin, J. B. (2018a). Neurotropism and behavioral changes associated with Zika infection in the vector Aedes aegypti article. *Emerging Microbes and Infections*, *7*(1). https://doi.org/10.1038/s41426-018-0069-2

Gaburro, J., Paradkar, P. N., Klein, M., Bhatti, A., Nahavandi, S., & Duchemin, J. B. (2018b). Dengue virus infection changes Aedes aegypti oviposition olfactory preferences. *Scientific Reports 2018 8:1*, *8*(1), 1–11. https://doi.org/10.1038/s41598-018-31608-x

Gallana, M., Ryser-Degiorgis, M.-P., Wahli, T., & Segner, H. (2013). Climate change and infectious diseases of wildlife: Altered interactions between pathogens, vectors and hosts. *Current*

Zoology, *59*(3), 427–437. https://doi.org/10.1093/CZOOLO/59.3.427

Ganchev, T., Fakotakis, N., & Kokkinakis, G. (2005). *Comparative evaluation of various MFCC implementations on the speaker verification task*. https://www.researchgate.net/publication/228756314

Gao, Y., Gillen, C. M., & Wheatly, M. G. (2006). *Molecular characterization of the sarcoplasmic calcium-binding protein (SCP) from crayfish Procambarus clarkii*. https://doi.org/10.1016/j.cbpb.2006.04.007

Garrett-Jones, C., & Grab, & B. (1964). The Assessment of Insecticidal Impact on the Malaria Mosquito's Vectorial Capacity, from Data on the Proportion of Parous Females. *Bull. Org. Mond. Sante*, *31*, 71–86.

Geier, M., Weber, M., Rose, A., Gordon, U., Pruszynski, C., & Dole, M. (2016). A smart Internet of Things (loT) device for monitoring mosquito trap counts in the field while drinking coffee at your desk. In *AMCA 82nd Annual Meeting*. *AMCA.* https://doi.org/10.13140/RG.2.1.5156.5843

Geier, M., Weber, M., Rose, A., Obermayr, U., & Abadam, C. (2016). *A smart Internet of Things (loT) device for monitoring mosquito trap counts in the field while drinking coffee at your desk*. https://doi.org/10.13140/RG.2.1.5156.5843

Genoud, A. P., Basistyy, R., Williams, G. M., & Thomas, B. P. (2018). Optical remote sensing for monitoring flying mosquitoes, gender identification and discussion on species identification. *Applied Physics B*, *124*, 46. https://doi.org/10.1007/s00340-018-6917-x

Genoud, A. P., Gao, Y., Williams, G. M., & Thomas, B. P. (2019). Identification of gravid mosquitoes from changes in spectral and polarimetric backscatter cross sections. *Journal of Biophotonics*, *12*(10). https://doi.org/10.1002/jbio.201900123

Genoud, A. P., Gao, Y., Williams, G. M., & Thomas, B. P. (2020). A comparison of supervised machine learning algorithms for mosquito identification from backscattered optical signals. *Ecological Informatics*, *58*(101090). https://doi.org/10.1016/j.ecoinf.2020.101090

Gibson, G., Warren, B., & Russell, I. J. (2010). Humming in tune: Sex and species recognition by mosquitoes on the wing. *JARO - Journal of the Association for Research in Otolaryngology*, *11*(4), 527–540. https://doi.org/10.1007/s10162-010-0243-2

Giron, S., Franke, F., Decoppet, A., Cadiou, B., Travaglini, T., Thirion, L., Durand, G., Jeannin, C., L'Ambert, G., Grard, G., Noël, H., Fournet, N., Auzet-Caillaud, M., Zandotti, C., Aboukaïs, S., Chaud, P., Guedj, S., Hamouda, L., Naudot, X., … Leparc-Goffart, I. (2019). Vector-borne transmission of Zika virus in Europe, southern France, August 2019. *Eurosurveillance*, *24*(45), 1. https://doi.org/10.2807/1560-7917.ES.2019.24.45.1900655

Gjenero-Margan, I., Aleraj, B., Krajcar, D., Lesnikar, V., Klobučar, A., Pem-Novosel, I., Kurečić-Filipović, S., Komparak, S., Martić, R., Duričić, S., Betica-Radić, L., Okmadžić, J., Vilibić-Čavlek, T., Babić-Erceg, A., Turković, B., Avšić-Županc, T., Radić, I., Ljubić, M., Šarac, K., … Mlinarić-Galinović, G. (2011). Autochthonous dengue fever in Croatia, August- September 2010. *Eurosurveillance*, *16*(9), 1–4. https://doi.org/10.2807/ESE.16.09.19805-EN/CITE/PLAINTEXT

González-Pérez, M. I., Faulhaber, B., Williams, M., Brosa, J., Aranda, C., Pujol, N., Verdún, M., Villalonga, P., Encarnação, J., Busquets, N., & Talavera, S. (2022). A novel optical sensor system for the automatic classification of mosquitoes by genus and sex with high levels of accuracy. *Parasites and Vectors*, *15*(1), 1–11. https://doi.org/10.1186/S13071-022-05324-5/TABLES/4

Gould, E. A., Gallian, P., De Lamballerie, X., & Charrel, R. N. (2010). First cases of autochthonous dengue fever and chikungunya fever in France: from bad dream to reality! *Clinical Microbiology and Infection*, *16*(12), 1702–1704. https://doi.org/10.1111/J.1469-0691.2010.03386.X

Gu, W., Unnasch, T. R., Katholi, C. R., Lampman, R., & Novak, R. J. (2008). Fundamental issues in mosquito surveillance for arboviral transmission. *Transactions of The Royal Society of Tropical Medicine and Hygiene*, *102*(8), 817–822. https://doi.org/10.1016/J.TRSTMH.2008.03.019

Gubler, D. J., Vasilakis, N., & Musso, D. (2017). History and Emergence of Zika Virus. *The Journal of Infectious Diseases*, *216*(Suppl 10), S860. https://doi.org/10.1093/INFDIS/JIX451

Gunay, F., Picard, M., & Robert, V. (2018). *MosKeyTool, an interactive identification key for mosquitoes of Euro-Mediterranean. Version 2.1*. www.medilabsecure.com/moskeytool

Gutiérrez-Bugallo, G., Piedra, L. A., Rodriguez, M., Bisset, J. A., Lourenço-de-Oliveira, R., Weaver, S. C., Vasilakis, N., & Vega-Rúa, A. (2019). Vector-borne transmission and evolution of Zika virus. *Nature Ecology & Evolution*, *3*(4), 561. https://doi.org/10.1038/S41559-019-0836-Z

GW, D., SF, K., & AJ, H. (1952). Zika virus. I. Isolations and serological specificity. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, *46*(5), 509–520. https://doi.org/10.1016/0035-9203(52)90042-4

Haba, Y., & McBride, L. (2022). Origin and status of Culex pipiens mosquito ecotypes. *Current Biology* , *32*(5), 237–246. https://doi.org/10.1016/J.CUB.2022.01.062

Haddow, A. D., Schuh, A. J., Yasuda, C. Y., Kasper, M. R., Heang, V., Huy, R., Guzman, H., Tesh, R. B., & Weaver, S. C. (2012). Genetic characterization of Zika virus strains: geographic expansion of the Asian lineage. *PLOS Neglected Tropical Diseases*, *6*(2), e1477. https://doi.org/10.1371/JOURNAL.PNTD.0001477

Hagstrum, D. W., Phillips, T. W., & Cuperus, G. (2012). *Stored Product Protection*. Kansas State University.

Hamer, G. L., Kitron, U. D., Brawn, J. D., Loss, S. R., Ruiz, M. O., Goldberg, T. L., & Walker, E. D. (2008). *Culex pipiens* (Diptera: Culicidae): a bridge vector of West Nile virus to humans. *Journal of Medical Entomology*, *45*(1), 125–128. https://doi.org/10.1093/JMEDENT/45.1.125

Hamer, W. H., A, M., Cantar, M. D., & P, F. R. C. (1906). Epidemic disease in England-The evidence of variability and the persistence of type. *The Lancet*, *167*(4307), 733–739. https://doi.org/10.1016/S0140-6736(01)80340-8

Hawkes, F. M., Zeil, J., & Gibson, G. (2022). Chapter 19: Vision in mosquitoes. *Sensory Ecology of Disease Vectors*, 511–533. https://doi.org/10.3920/978-90-8686-932-9_19

Huang, Y. J. S., Higgs, S., & Vanlandingham, D. L. (2017). Biological control strategies for mosquito vectors of arboviruses. *Insects*, *8*(1). https://doi.org/10.3390/INSECTS8010021

Hugo, L., Cook, P. E., Johnson, P. H., Rapley, L. P., Kay, B. H., Ryan, P. A., Ritchie, S. A., & O'neill, S. L. (2010). Field validation of a transcriptional assay for the prediction of age of uncaged Aedes aegypti mosquitoes in Northern Australia. *PLoS Neglected Tropical Diseases*, *4*(2). https://doi.org/10.1371/journal.pntd.0000608

Hugo, L. E., Kay, B. H., Eaglesham, G. K., Holling, N., & Ryan, P. A. (2006). Investigation of cuticular hydrocarbons for determining the age and survivorship of Australasian mosquitoes. *American Journal of Tropical Medicine and Hygiene*, *74*(3), 462–474. https://doi.org/10.4269/ajtmh.2006.74.462

Hugo, L. E., Quick-Miles, S., Kay, B. H., & Ryan, P. A. (2008). Evaluations of Mosquito Age Grading Techniques Based on Morphological Changes. *J. Med. Entomol*, *45*(3), 353–369. https://academic.oup.com/jme/article/45/3/353/905817

Iowa State University. (n.d.). *Life Cycle of the Mosquito*. Retrieved October 25, 2023, from https://www.extension.iastate.edu/equine/blog/dr-peggy-m-auwerda/life-cycle-mosquito

Johnson, B. J., Hugo, L. E., Churcher, T. S., Ong, O. T. W., & Devine, G. J. (2020). Mosquito age grading and vector-control programmes. In *Trends in Parasitology* (Vol. 36, Issue 1, pp. 39–51). Elsevier Ltd. https://doi.org/10.1016/j.pt.2019.10.011

Johnson, B. J., & Ritchie, S. A. (2016). The siren's song: exploitation of female flight tones to passively capture male *Aedes aegypti* (Diptera: Culicidae). *Journal of Medical Entomology*, *53*(1), 245–248. https://doi.org/10.1093/jme/tjv165

Johnson, B. J., Rohde, B. B., Zeak, N., Staunton, K. M., Prachar, T., & Ritchie, S. A. (2018). A low-cost, battery-powered acoustic trap for surveilling male Aedes aegypti during rear-and-release operations. *PLoS ONE*, *13*(8). https://doi.org/10.1371/journal.pone.0201709

Joshi, A., & Miller, C. (2021). Review of machine learning techniques for mosquito control in urban environments. *Ecological Informatics*, *61*, 101241. https://doi.org/10.1016/J.ECOINF.2021.101241

Jupille, H., Seixas, G., Mousson, L., Sousa, C. A., & Failloux, A. B. (2016). Zika virus, a new threat for Europe? *PLOS Neglected Tropical Diseases*, *10*(8), e0004901. https://doi.org/10.1371/JOURNAL.PNTD.0004901

Kahn, M. C., Celestin, W., & Offenhauser, W. (1945). Recording of sounds produced by certain disease-carrying mosquitoes. *Science*, *101*(2622), 335–336. https://doi.org/10.1126/SCIENCE.101.2622.335/ASSET/8600DF32-4881-49D2-A32E-6C0EF52DDF8E/ASSETS/SCIENCE.101.2622.335.FP.PNG

Kahn, M. C., & Offenhauser, W. (1949). The first field tests of recorded mosquito sounds used for mosquito destruction. *The American Journal of Tropical Medicine and Hygiene*, *29*(5), 811–825. https://doi.org/10.4269/ajtmh.1949.s1-29.811

Kampen, H., Medlock, J. M., Vaux, A. G. C., Koenraadt, C. J. M., Van Vliet, A. J. H., Bartumeus, F., Oltra, A., Sousa, C. A., Chouin, S., & Werner, D. (2015). Approaches to passive mosquito surveillance in the EU. *Parasites and Vectors*, *8*(1), 1–13. https://doi.org/10.1186/s13071-014-0604-5

Kaptoul, D., Viladrich, P. F., Domingo, C., Niubó, J., Martínez-Yélamos, S., De Ory, F., & Tenorio, A. (2007). West Nile virus in Spain: Report of the first diagnosed case (in Spain) in a human with aseptic meningitis. *Scandinavian Journal of Infectious Diseases*, *39*(1), 70–71. https://doi.org/10.1080/00365540600740553

Kardos, E. H., & Bellamy, R. E. (1961). Distinguishing nulliparous from parous female *Culex tarsalis* by examination of the ovarian tracheation. *Annals of the Entomological Society of America*, *54*(3), 448–451. https://doi.org/10.1093/AESA/54.3.448

Kermack, W. 0, & Mckendrick, A. G. (1927). A Contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society A*. https://royalsocietypublishing.org/

Khalighifar, A., Jiménez-García, D., Campbell, L. P., Mensah Ahadji-Dabla, K., Aboagye-Antwi, F., Ibarra-Juárez, L. A., & Peterson, A. T. (2022). Application of deep learning to community-science-based mosquito monitoring and detection of novel species. *Journal of Medical Entomology*, *59*(1), 355–362. https://doi.org/10.1093/jme/tjab161

Khasnis, A. A., & Nettleman, M. D. (2005). Global warming and infectious disease. In *Archives of Medical Research* (Vol. 36, Issue 6, pp. 689–696). https://doi.org/10.1016/j.arcmed.2005.03.041

Kilpatrick, A. M., Kramer, L. D., Jones, M. J., Marra, P. P., & Daszak, P. (2006). West Nile virus epidemics in North America are driven by shifts in mosquito feeding behavior. *PLoS Biology*, *606*(610), e82. https://doi.org/10.1371/journal.pbio.0040082

Kilpatrick, A. M., Meola, M. A., Moudy, R. M., & Kramer, L. D. (2008). Temperature, viral genetics, and the transmission of West Nile virus by *Culex pipiens* mosquitoes. *PLoS Pathogens*, *4*(6), 1000092. https://doi.org/10.1371/JOURNAL.PPAT.1000092

Kim, D., DeBriere, T. J., Cherukumalli, S., White, G. S., & Burkett-Cadena, N. D. (2021). Infrared light sensors permit rapid recording of wingbeat frequency and bioacoustic species identification of mosquitoes. *Scientific Reports 2021 11:1*, *11*(1), 1–9. https://doi.org/10.1038/s41598-021-89644-z

Kirkeby, C., Wellenreuther, M., & Brydegaard, M. (2016). Observations of movement dynamics of flying insects using high resolution lidar. *Scientific Reports*, *6*(1), 1–11. https://doi.org/10.1038/srep29083

Kiskin, I., Cobb, A. D., Wang, L., & Roberts, S. (2020). Humbug Zooniverse: A Crowd-Sourced Acoustic Mosquito Dataset. *IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, *2020-May*, 916–920. https://doi.org/10.1109/ICASSP40776.2020.9053141

Kittichai, V., Kaewthamasorn, M., Samung, Y., Jomtarak, R., Naing, K. M., Tongloy, T., Chuwongin, S., & Boonsang, S. (2023). Automatic identification of medically important mosquitoes using embedded learning approach-based image-retrieval system. *Scientific Reports 2023 13:1*, *13*(1), 1–19. https://doi.org/10.1038/s41598-023-37574-3

Kleinbaum, D., & Klein, M. (2010). *Logistic regression: a self-learning text* (Springer).

Kramer, L. D., & Ciota, A. T. (2015). Dissecting vectorial capacity for mosquito-borne viruses. *Current Opinion in Virology*, *15*, 112. https://doi.org/10.1016/J.COVIRO.2015.10.003

Krökel, U., Rose, A., Eiras, A., & Geier, M. (2006). New tools for surveillance of adult yellow fever mosquitoes: Comparison of trap catches with human landing rates in an urban environment. *Article in Journal of the American Mosquito Control Association*. https://doi.org/10.2987/8756-971X(2006)22[229:NTFSOA]2.0.CO;2

Kweka, E. J., Baraka, V., Mathias, L., Mwang'onde, B., Baraka, G., Lyaruu, L., Mahande, A. M., Kweka, E. J., Baraka, V., Mathias, L., Mwang'onde, B., Baraka, G., Lyaruu, L., & Mahande, A. M. (2018). Ecology of Aedes mosquitoes, the major vectors of arboviruses in human population. In *Dengue Fever. A resilient threat in the face of innovation*. IntechOpen. https://doi.org/10.5772/INTECHOPEN.81439

Lacour, G., Chanaud, L., L'ambert, G., & Hance, T. (2015). *Seasonal Synchronization of Diapause Phases in Aedes albopictus (Diptera: Culicidae)*. https://doi.org/10.1371/journal.pone.0145311

Lai, Z., Wu, J., Xiao, X., Xie, L., Liu, T., Zhou, J., Xu, Y., Cai, Y., Lin, F., Li, B., Gan, L., James, A. A., & Chenid, X. G. (2022). Development and evaluation of an efficient and real-time monitoring system for the vector mosquitoes, Aedes albopictus and Culex quinquefasciatus. *PLoS Neglected Tropical Diseases*, *16*(9). https://doi.org/10.1371/JOURNAL.PNTD.0010701

Lambert, B., Sikulu-Lord, M. T., Mayagaya, V. S., Devine, G., Dowell, F., & Churcher, T. S. (2018). Monitoring the age of mosquito populations using near-Infrared spectroscopy. *Scientific Reports*, *8*(1). https://doi.org/10.1038/s41598-018-22712-z

Lanciotti, R. S., Kosoy, O. L., Laven, J. J., Velez, J. O., Lambert, A. J., Johnson, A. J., Stanfield, S. M., & Duffy, M. R. (2008). Genetic and serologic properties of Zika virus associated with an epidemic, Yap State, Micronesia, 2007. *Emerging Infectious Diseases*, *14*(8), 1232. https://doi.org/10.3201/EID1408.080287

Leandro, A. S., de Castro, W. A. C., Lopes, R. D., Delai, R. M., Villela, D. A. M., & Maciel-De-Freitas, R. (2022). Citywide integrated aedes aegypti mosquito surveillance as early warning system for arbovirus transmission, Brazil. *Emerging Infectious Diseases*, *28*(4), 707. https://doi.org/10.3201/EID2804.211547

Lee, T. H., Ullah, A., & Wang, R. (2020). Bootstrap aggregating and random Forest. *Advanced Studies in Theoretical and Applied Econometrics*, *52*, 389–429. https://doi.org/10.1007/978-3-030-31150-6_13/COVER

Li, Y. G., Siripanyaphinyo, U., Tumkosit, U., Noranate, N., A-Nuegoonpipat, A., Tao, R., Kurosu, T., Ikuta, K., Takeda, N., & Anantapreecha, S. (2012). Chikungunya virus induces a more moderate cytopathic effect in mosquito cells than in mammalian cells. *Intervirology*, *56*(1), 6–12. https://doi.org/10.1159/000339985

Li, Y., Su, X., Zhou, G., Zhang, H., Puthiyakunnon, S., Shuai, S., Cai, S., Gu, J., Zhou, X., Yan, G., & Chen, X. G. (2016). Comparative evaluation of the efficiency of the BG-Sentinel trap, CDC light trap and Mosquito-oviposition trap for the surveillance of vector mosquitoes. *Parasites and Vectors*, *9*(1), 1–8. https://doi.org/10.1186/s13071-016-1724-x

Li, Y., Zilli, D., Chan, H., Kiskin, I., Sinka, M., Roberts, S., & Willis, K. (2017). Mosquito detection with low-cost smartphones: data acquisition for malaria research. *NIPS Workshop on Machine Learning for the Developing World* .

Lima-Camara, T. N., Bruno, R. V., Luz, P. M., Castro, M. G., Lourenço-de-Oliveira, R., Sorgine, M. H. F., & Peixoto, A. A. (2011). Dengue Infection Increases the Locomotor Activity of Aedes aegypti Females. *PLOS ONE*, *6*(3), e17690. https://doi.org/10.1371/JOURNAL.PONE.0017690

Lindenbach, B. D., & Rice, C. M. (2003). Molecular biology of flaviviruses. *Advances in Virus Research*, *59*, 23–61. https://doi.org/10.1016/S0065-3527(03)59002-9

Linthicum, K. J., Platt, K., Myint, K. S., Lerdthusnee, K., Innis, B. L., & Vaughn, D. W. (1996). Dengue 3 virus distribution in the mosquito *Aedes aegypti*: an immunocytochemical study. *Medical and Veterinary Entomology*, *10*(1), 87–92. https://doi.org/10.1111/J.1365-2915.1996.TB00086.X

Liu, N. (2015). Insecticide resistance in mosquitoes: impact, mechanisms, and research directions. https://doi.org/10.1146/ANNUREV-ENTO-010814-020828

Liu, Q.-M., Gong, Z.-Y., & Wang, Z. (2023). A review of the surveillance techniques for *Aedes albopictus*. https://doi.org/10.4269/ajtmh.20-0781

Livak, K. J., & Schmittgen, T. D. (2001). Analysis of relative gene expression data using Real-Time Quantitative PCR and the 2 C T Method. *METHODS*, *25*, 402–408. https://doi.org/10.1006/meth.2001.1262

Lorenz, C., Almeida, F., Almeida-Lopes, F., Louise, C., Pereira, S. N., Petersen, V., Vidal, P. O., Virginio, F., & Suesdek, L. (2017). Geometric morphometrics in mosquitoes: What has been measured? https://doi.org/10.1016/j.meegid.2017.06.029

Louis, C. (2012). Daily newspaper view of dengue fever epidemic, Athens, Greece, 1927-1931. *Emerging Infectious Diseases*, *18*(1), 78–82. https://doi.org/10.3201/eid1801.110191

Lowe, S., Browne, M., Boudjelas, S., & De Poorter, M. (2000). 100 of the world's worst invasive alien species. A selection from the global invasive species database. www.issg.org/booklet.pdf

Lühken, R., Pfitzner, W. P., Börstler, J., Garms, R., Huber, K., Schork, N., Steinke, S., Kiel, E., Becker, N., Tannich, E., & Krüger, A. (2014). Field evaluation of four widely used mosquito traps in Central Europe. *Parasites and Vectors*, *7*(1), 1–11. https://doi.org/10.1186/1756-3305-7-268

Lukman, A., Harjoko, A., & Yang, C.-K. (2017). MFCC feature classification from Culex and *Aedes aegypti* mosquitoes noise using support vector machine. https://doi.org/10.1109/ICSIIT.2017.28

Macdonald, G. (1957). The Epidemiology and Control of Malaria. *The Epidemiology and Control of Malaria.*

Macdonald, G., G, C. M., D, M., P, F. R. C., H, D. P., & M, D. T. (1956). Epidemiological basis of malaria control. *Bull. Org. Mond. Santé*, *15*, 613–626.

Macias, A., Martín, P., Pérez-Olmeda, M., Fernández-Martínez, B., Gómez-Barroso, D., Fernández, E., Ramos, J. M., Herrero, L., Rodríguez, S., Delgado, E., Sánchez-Seco, M. P., Galán, M., Corbacho, A. J., Jimenez, M., Montero-Peña, C., Valle, A., & Vázquez, A. (2023). West Nile virus emergence in humans in Extremadura, Spain 2020. *Frontiers in Cellular and Infection Microbiology*, *13*, 1155867. https://doi.org/10.3389/FCIMB.2023.1155867/BIBTEX

Maciel-de-Freitas, R., Koella, J. C., & Lourenço-de-Oliveira, R. (2011). Lower survival rate, longevity and fecundity of *Aedes aegypti*

(Diptera: Culicidae) females orally challenged with dengue virus serotype 2. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, *105*(8), 452–458. https://doi.org/10.1016/J.TRSTMH.2011.05.006/2/M_105-8-452-FIG001.GIF

Magalhães, A. R., Codeço, C. T., Svenning, J. C., Escobar, L. E., Van de Vuurst, P., & Gonçalves-Souza, T. (2023). Neglected tropical diseases risk correlates with poverty and early ecosystem destruction. *Infectious Diseases of Poverty*, *12*(1), 1–15. https://doi.org/10.1186/S40249-023-01084-1/FIGURES/5

Martínez-De La Puente, J., Ferraguti, M., Ruiz, S., Roiz, D., Llorente, F., Pérez-Ramírez, E., Jiménez-Clavero, M. Á., Soriguer, R., & Figuerola, J. (2018). Mosquito community influences West Nile virus seroprevalence in wild birds: implications for the risk of spillover into human populations. *Scientific Reports 2018 8:1*, *8*(1), 1–7. https://doi.org/10.1038/s41598-018-20825-z

Mckenzieid, B. A., Wilson, A. E., & Zohdyid, S. (2019). *Aedes albopictus* is a competent vector of Zika virus: A meta-analysis. https://doi.org/10.1371/journal.pone.0216794

Medley, G. F., & Vassall, A. (2017). When an emerging disease becomes endemic. *Science*, *357*(6347), 156–158. https://doi.org/10.1126/SCIENCE.AAM8333

Menda, G., Nitzany, E. I., Shamble, P. S., Wells, A., Harrington, L. C., Miles, R. N., & Hoy, R. R. (2019). The long and short of hearing in the mosquito *Aedes aegypti*. *Current Biology*, *29*(4), 709-714.e4. https://doi.org/10.1016/J.CUB.2019.01.026

Mitchell, C. J. (2020). Occurrence, Biology, and Physiology of Diapause in Overwintering Mosquitoes. *The Arboviruses: Epidemiology and Ecology*, 191–218. https://doi.org/10.1201/9780429280221-7

Mordecai, E. A., Caldwell, J. M., Grossman, M. K., Lippi, C. A., Johnson, L. R., Neira, M., Rohr, J. R., Ryan, S. J., Savage, V., Shocket, M. S., Sippy, R., Stewart Ibarra, A. M., Thomas, M. B., & Villena, O. (2019).

Thermal biology of mosquito-borne disease. *Ecology Letters,* 22(10), 1690–1708. https://doi.org/10.1111/ele.13335

Mukundarajan, H., & … R. K. (2019). ABUZZ: A mobile phone-based citizen science platform for crowdsourcing ecological data for mosquito surveillance. *American Society of Tropical Medicine & Hygiene,* 101, 448-448.

Mukundarajan, H., Hol, F. J. H., Castillo, E. A., Newby, C., & Prakash, M. (2017). Using mobile phones as acoustic sensors for high-throughput mosquito surveillance. *ELife*, *6*. https://doi.org/10.7554/eLife.27854

Mullen, E. R., Rutschman, P., Pegram, N., Patt, J. M., Adamczyk, J. J., & Johanson. (2016). Laser system for identification, tracking, and control of flying insects. *Optics Express*, *24*(11), 11828. https://doi.org/10.1364/oe.24.011828

Müller, M. (2007). *Information Retrieval for Music and Motion* . https://doi.org/10.1007/978-3-540-74048-3

Napp, S., Petrić, D., & Busquets, N. (2018). West Nile virus and other mosquito-borne viruses present in Eastern Europe. *Pathog Glob Health*, 112(5), 233-248. https://doi.org/10.1080/20477724.2018.1483567

Nayak, B., Khuntia, · Bonomali, Laxman, ·, Murmu, K., Sahu, · Bijayalaxmi, Rabi, ·, Pandit, S., Tapan, ·, & Barik, K. (2023). Artificial intelligence (AI): a new window to revamp the vector-borne disease control. *Parasitology Research*, *1*, 3. https://doi.org/10.1007/s00436-022-07752-9

Noda, J. J., Travieso-González, C. M., Sánchez-Rodríguez, D., & Alonso-Hernández, J. B. (2019). Acoustic classification of singing insects based on MFCC/LFCC fusion. *Applied Sciences*, *9*(19), 4097. https://doi.org/10.3390/app9194097

Obadia, T., Gutierrez-Bugallo, G., Duong, V., Nuñez, A. I., Fernandes, R. S., Kamgang, B., Hery, L., Gomard, Y., Abbo, S. R., Jiolle, D., Glavinic, U., Dupont-Rouzeyrol, M., Atyame, C. M., Pocquet, N., Boyer, S., Dauga, C., Vazeille, M., Yébakima, A., White, M. T., … Failloux, A. B.

(2022). Zika vector competence data reveals risks of outbreaks: the contribution of the European ZIKAlliance project. *Nature Communications*, *13*(1). https://doi.org/10.1038/S41467-022-32234-Y

Offenhauser, W. H., & Kahn, M. C. (1949). The sounds of disease-carrying mosquitoes. *Journal of the Acoustical Society of America*, *21*(3), 259–263. https://doi.org/10.1121/1.1906505

Ohm, J. R., Baldini, F., Barreaux, P., Lefevre, T., Lynch, P. A., Suh, E., Whitehead, S. A., & Thomas, M. B. (2018). Rethinking the extrinsic incubation period of malaria parasites. *Parasites and Vectors*, *11*(1), 1–9. https://doi.org/10.1186/S13071-018-2761-4/FIGURES/3

O'Meara, G. F., Evans, L. F., Gettman, A. D., & Cuda, J. P. (1995). Spread of *Aedes albopictus* and decline of *Ae. aegypti* (Diptera: Culicidae) in Florida. *Journal of Medical Entomology*, *32*(4), 554–562. https://doi.org/10.1093/JMEDENT/32.4.554

O'Neill, S. L., Cook, P. E., & McMeniman, C. J. (2008). Modifying insect population age structure to control vector-borne disease. *Advances in Experimental Medicine and Biology*, *627*, 126–140. https://doi.org/10.1007/978-0-387-78225-6_11/COVER

Oppenheim, A. V. (1970). Speech spectrograms using the fast Fourier transform. *IEEE Spectrum*, *7*(8), 57–62. https://doi.org/10.1109/MSPEC.1970.5213512

Ouyang, T. H., Yang, E. C., Jiang, J. A., & Lin, T. Te. (2015). Mosquito vector monitoring system based on optical wingbeat classification. *Computers and Electronics in Agriculture*, *118*, 47–55. https://doi.org/10.1016/j.compag.2015.08.021

Palmer, J. R. B., Oltra, A., Collantes, F., Delgado, J. A., Lucientes, J., Delacour, S., Bengoa, M., Eritja, R., & Bartumeus, F. (2017). Citizen science provides a reliable and scalable tool to track disease-carrying mosquitoes. *Nature Communications 2017 8:1*, *8*(1), 1–13. https://doi.org/10.1038/s41467-017-00914-9

Park, D., Bowles, | Joy, Norrid, K., Dobson, | F Stephen, Abebe, | Ash, Haripriya, |, Narayanan, V., Prakash, M., Blagburn, B., Starkey, | Lindsay, & Zohdy, S. (2023). Effect of age on wingbeat frequency of *Aedes aegypti* and potential application for age estimation of mosquitoes. *Medical and Veterinary Entomology*, 37(3), 491-498. https://doi.org/10.1111/mve.12647

Park, J., Kim, D. I., Choi, B., Kang, W., & Kwon, H. W. (2020). classification and morphological analysis of vector mosquitoes using deep convolutional neural networks. *Scientific Reports 2020 10:1*, *10*(1), 1–12. https://doi.org/10.1038/s41598-020-57875-1

Paupy, C., Delatte, H., Bagny, L., Corbel, V., & Fontenille, D. (2009). *Aedes albopictus*, an arbovirus vector: from the darkness to the light. *Microbes and Infection*, *11*, 1177–1185. https://doi.org/10.1016/j.micinf.2009.05.005

Peach, D. A. H., & Gries, G. (2020). Mosquito phytophagy – sources exploited, ecological function, and evolutionary transition to haematophagy. In *Entomologia Experimentalis et Applicata*, 168 (2), 120–136. https://doi.org/10.1111/eea.12852

Pedreira Padilha, K., Resck, M. E. B., Talyuli Da Cunha, O. A., Teles-De-Freitas, R., Campos, S. S., Ferreira Sorgine, M. H., Lourenço-De-Oliveira, R., Farnesi, L. C., & Vieira Bruno, R. (2018). Zika infection decreases Aedes aegypti locomotor activity but does not influence egg production or viability. *Mem Inst Oswaldo Cruz, Rio de Janeiro*, *113*(10), 1–6. https://doi.org/10.1590/0074-02760180290

Peng, C. Y. J., Lee, K. L., & Ingersoll, G. M. (2002). An introduction to logistic regression analysis and reporting. *Journal of Educational Research*, *96*(1), 3–14. https://doi.org/10.1080/00220670209598786

Pereira-Dos-Santos, T., Roiz, D., Lourenço-De-Oliveira, R., & Paupy, C. (2020). A Systematic Review: Is Aedes albopictus an Efficient Bridge Vector for Zoonotic Arboviruses? *Pathogens, 9*(4), 266. https://doi.org/10.3390/PATHOGENS9040266

Petrić, D., Bellini, R., Scholte, E. J., Rakotoarivony, L. M., & Schaffner, F. (2014). Monitoring population and environmental parameters of invasive mosquito species in Europe. *Parasites and Vectors*, *7*(1). https://doi.org/10.1186/1756-3305-7-187

Phyton software foundation. (n.d.). *Phyton Documentation*.

Platt, K. B., Linthicum, K. J., Myint, K. S. A., Innis, B. L., Lerdthusnee, K., & Vaughn, D. W. (1997). Impact of dengue virus infection on feeding behavior of Aedes aegypti. *The American Journal of Tropical Medicine and Hygiene*, *57*(2), 119–125. https://doi.org/10.4269/AJTMH.1997.57.119

Plourde, A. R., & Bloch, E. M. (2016). A Literature Review of Zika Virus. *Emerging Infectious Diseases*, *22*(7), 1185. https://doi.org/10.3201/EID2207.151990

Potamitis, I. (2014). Classifying insects on the fly. *Ecological Informatics*, *21*, 40–49. https://doi.org/10.1016/j.ecoinf.2013.11.005

Potamitis, I., Eliopoulos, P., & Rigakis, I. (2017). Automated remote insect surveillance at a global scale and the internet of things. *Robotics*, *6*(3). https://doi.org/10.3390/robotics6030019

Potamitis, I., Ganchev, T., & Kontodimas, D. (2009). On automatic bioacoustic detection of pests: the cases of *Rhynchophorus ferrugineus* and *Sitophilus oryzae*. *Journal of Economic Entomology*, *102*(4), 1681–1690. https://doi.org/10.1603/029.102.0436

Potamitis, I., & Rigakis, I. (2015). Novel noise-robust optoacoustic sensors to identify insects through wingbeats. *IEEE Sensors Journal*, *15*(8), 4621–4631. https://doi.org/10.1109/JSEN.2015.2424924

Potamitis, I., & Rigakis, I. (2016a). Large aperture optoelectronic devices to record and time-stamp insects' Wingbeats. *IEEE Sensors Journal*, *16*(15), 6053–6061. https://doi.org/10.1109/JSEN.2016.2574762

Potamitis, I., & Rigakis, I. (2016b). Measuring the fundamental frequency and the harmonic properties of the wingbeat of a large number of mosquitoes in flight using 2D optoacoustic sensors. *Applied Acoustics*, *109*, 54–60. https://doi.org/10.1016/j.apacoust.2016.03.005

Potamitis, I., Rigakis, I., & Fysarakis, K. (2015). Insect Biometrics: Optoacoustic Signal Processing and Its Applications to Remote Monitoring of McPhail Type Traps. *PLOS ONE*, *10*(11), e0140474. https://doi.org/10.1371/JOURNAL.PONE.0140474

Potamitis, I., Rigakis, I., Vidakis, N., Petousis, M., & Weber, M. (2018). Affordable Bimodal Optical Sensors to Spread the Use of Automated Insect Monitoring. *Journal of Sensors*. https://doi.org/10.1155/2018/3949415

Qiu, Y., Spitzen, J., Smallegange, R., & Knols, B. (2007). Monitoring systems for adult insect pests and disease vectors. In Emerging pests and vector-borne diseases in Europ*e*. Wageningen Academic. https://doi.org/10.3920/978-90-8686-626-7

R: Pseudo R2 Statistics. (n.d.). https://search.r-project.org/CRAN/refmans/DescTools/html/PseudoR2.html

Reinhold, J. M., Lazzari, C. R., & Lahondère, C. (2018). Effects of the environmental temperature on *Aedes aegypti* and *Aedes albopictus* mosquitoes: A review. *Insects*, *9*(4). https://doi.org/10.3390/insects9040158

Rezza, G., Nicoletti, L., Angelini, R., Romi, R., Finarelli, A., Panning, M., Cordioli, P., Fortuna, C., Boros, S., Magurano, F., Silvi, G., Angelini, P., Dottori, M., Ciufolini, M., Majori, G., & Cassone, A. (2007). Infection with chikungunya virus in Italy: an outbreak in a temperate region. *Lancet (London, England)*, *370*(9602), 1840–1846. https://doi.org/10.1016/S0140-6736(07)61779-6

Richards, I. R. (1955). Photoelectric cell observations of insects in flight. In *Nature* (Vol. 175, Issue 4446, pp. 128–129). Nature Publishing Group. https://doi.org/10.1038/175128b0

Rizzoli, A., Bolzoni, L., Chadwick, E. A., Capelli, G., Montarsi, F., Grisenti, M., De La Puente, J. M., Muñoz, J., Figuerola, J., Soriguer, R., Anfora, G., Di Luca, M., & Rosà, R. (2015). Understanding West Nile virus ecology in Europe: *Culex pipiens* host feeding preference in a hotspot of virus emergence. *Parasites and Vectors*, *8*(1), 1–13. https://doi.org/10.1186/S13071-015-0831-4/COMMENTS

Rizzoli, A., Tagliapietra, V., Cagnacci, F., Marini, G., Arnoldi, D., Rosso, F., & Rosà, R. (2019). Parasites and wildlife in a changing world: The vector-host-pathogen interaction as a learning case. https://doi.org/10.1016/j.ijppaw.2019.05.011

Rohde, B. B., Staunton, K. M., Zeak, N. C., Beebe, N., Snoad, N., Bondarenco, A., Liddington, C., Anderson, J. A., Xiang, W., Mankin, R. W., & Ritchie, S. A. (2019). Waterproof, low-cost, long-battery-life sound trap for surveillance of male Aedes aegypti for rear-and-release mosquito control programmes. *Parasites and Vectors*, *12*(1), 1–7. https://doi.org/10.1186/S13071-019-3647-9/FIGURES/4

Roiz, D., Roussel, M., Munõz, J., Ruiz, S., Soriguer, R., & Figuerola, J. (2012). Efficacy of mosquito traps for collecting potential west nile mosquito vectors in a natural mediterranean wetland. *American Journal of Tropical Medicine and Hygiene*, *86*(4), 642–648. https://doi.org/10.4269/ajtmh.2012.11-0326

Romero-Vivas, C. M. E., & Falconar, A. K. I. (2005). Investigation of relationships between *Aedes aegypty* egg, larvae, pupae and adult density indices where their main breeding sites were located indoors. *Journal of American Mosquito Control Association*, *21 (1)*, 15–21.

Rossati, A. (2017). Global warming and its health impact. *International Journal of Occupational and Environmental Medicine*, 8 (1), 7–20. *NIOC Health Organization*. https://doi.org/10.15171/ijoem.2017.963

Rudolf, I., Betášová, L., Blažejová, H., Venclíková, K., Straková, P., Šebesta, O., Mendel, J., Bakonyi, T., Schaffner, F., Nowotny, N., & Hubálek, Z. (2017). West Nile virus in overwintering mosquitoes, central Europe. *Parasites & Vectors*. https://doi.org/10.1186/s13071-017-2399-7

Ruybal, J. E., Kramer, L. D., & Kilpatrick, A. M. (2016). Geographic variation in the response of *Culex pipiens* life history traits to temperature. *Parasites and Vectors*, *9*(1), 1–9. https://doi.org/10.1186/S13071-016-1402-Z/FIGURES/6

Salim, Z. T., Hashim, U., Arshad, M. K., Fakhri, M. A., & Salim, E. T. (2017). Frequency-based detection of female Aedes mosquito using surface acoustic wave technology: Early prevention of dengue fever. *Microelectronic Engineering*, *179*, 83–90. https://doi.org/10.1016/j.mee.2017.04.016

Samarawickrema, W. A. (1967). A study of the age-composition of natural populations of *Culex pipiens fatigans* in relation to the transmission of filariasis due to *Wuchereria bancrofti* in Ceylon. *Bulletin of the World Health Organization*, *37*(1), 117.

Samuel, G. H., Adelman, Z. N., & Myles, K. M. (2018). Antiviral immunity and virus-mediated antagonism in disease vector mosquitoes. *Trends in Microbiology*, *26*(5), 447–461. https://doi.org/10.1016/J.TIM.2017.12.005

Sanchez-Vargas, I., Williams, A. E., Franz, A. W. E., & Olson, K. E. (2021). Intrathoracic inoculation of Zika virus in *Aedes aegypti*. *Bio-Protocol*, *11*(18), 4165. https://doi.org/10.21769/BioProtoc.4165

Santos, D. A. A., Rodrigues, J. J. P. C., Furtado, V., Saleem, K., & Korotaev, V. (2019a). Automated electronic approaches for detecting disease vectors mosquitoes through the wing-beat frequency. *Journal of Cleaner Production*, *217*, 767–775. https://doi.org/10.1016/j.jclepro.2019.01.187

Santos, D. A. A., Teixeira, L. E., Alberti, A. M., Furtado, V., & Rodrigues, J. J. P. C. (2018). sensitivity and noise evaluation of an optoelectronic sensor for mosquitoes monitoring. *3rd International Conference on Smart and Sustainable Technologies, SpliTech.*

Sauer, F. G., Grave, J., Lühken, R., & Kiel, E. (2021). Habitat and microclimate affect the resting site selection of mosquitoes.

Medical and Veterinary Entomology, *35*(3), 379–388. https://doi.org/10.1111/MVE.12506

Schaffner, F., Angel, G., Geoffroy, B., Hervy, J.-P., Rhaiem, A., & Brunhes, J. (2001). *The mosquitoes of Europe. An identification and training programme*. IRD Editions & EID Méditerranée.

Schaffner, F., Medlock, J. M., & Van Bortel, W. (2013). Public health significance of invasive mosquitoes in Europe. *Clinical Microbiology and Infection,* 19(8), 685–692. https://doi.org/10.1111/1469-0691.12189

Schmidhuber, J. (2014). Deep Learning in Neural Networks: An Overview. *Neural Networks*, *61*, 85–117. https://doi.org/10.1016/j.neunet.2014.09.003

Scikit-learn developers. (n.d.). *Scikit-learn*. https://scikit-learn.org/stable/

Scott, T. W., & Takken, W. (2012). Feeding strategies of anthropophilic mosquitoes result in increased risk of pathogen transmission. *Trends in Parasitology*, *28*(3), 114–121. https://doi.org/10.1016/j.pt.2012.01.001

Sejvar, J. J. (2003). West Nile Virus: An historical overview. *The Ochsner Journal*, *5*(3), 6.

Semenza, J. C. (2013). Prototype early warning systems for vector-borne diseases in Europe. *Int. J. Environ. Res. Public Health*, *12*, 12. https://doi.org/10.3390/ijerph120606333

Service, M. W. (1993). Mosquito Ecology. Field sampling methods. https://doi.org/10.1007/978-94-015-8113-4

Sherpa, S., Blum, M. G. B., Capblancq, T., Cumer, T., Rioux, D., & Després, L. (2019). Unravelling the invasion history of the Asian tiger mosquito in Europe. *Molecular Ecology*, *28*(9), 2360–2377. https://doi.org/10.1111/MEC.15071

Silva, D. F., Vinícius, ·, Souza, M. A., Daniel, ·, Ellis, P. W., Eamonn, ·, Keogh, J., Gustavo, ·, Batista, E. A. P. A., Silva, D. F., Souza, V. M. A., Batista, · G E A P A, Batista, G. E. A. P. A., Ellis, D. P. W., & Keogh, E.

J. (2015). Exploring low cost laser sensors to identify flying insect species evaluation of machine learning and signal processing methods. *J Intell Robot Syst*, *1*, 313–330. https://doi.org/10.1007/s10846-014-0168-9

Silver, J. B. (2008). Mosquito Ecology. Field sampling methods (Springer, Ed.).

Simões, P. M. V., Ingham, R. A., Gibson, G., & Russell, I. J. (2016). A role for acoustic distortion in novel rapid frequency modulation behaviour in free-flying male mosquitoes. *Journal of Experimental Biology*, *219*(13), 2039–2047. https://doi.org/10.1242/jeb.135293

Sinka, M. E., Zilli, D., Li, Y., Kiskin, I., Kirkham, D., Rafique, W., Wang, L., Chan, H., Gutteridge, B., Herreros-Moya, E., Portwood, H., Roberts, S., Willis, K. J., & Marianne Sinka, C. E. (2021). HumBug-An acoustic mosquito monitoring tool for use on budget smartphones. *Methods Ecol Evol*, *12*, 1848–1859. https://doi.org/10.1111/2041-210X.13663

Sivagnaname, N., & Gunasekaran, K. (2012). Need for an efficient adult trap for the surveillance of dengue vectors. *Indian J Med Res*, *136*, 739–749.

Smithburn, K. C., Hughes, T. P., Burke, A. W., & Paul, J. H. (1940). A Neurotropic Virus Isolated from the Blood of a Native of Uganda. *American Journal of Tropical Medicine*, *20*, 471–472.

Song, Z., Zhang, B., Feng, H., Zhu, S., Hu, L., Brydegaard, M., Li, Y., Jansson, S., Malmqvist, E., Svanberg, K., Zhao, G., Bood, J., Svanberg, S., & Li, D. (2019). Application of lidar remote sensing of insects in agricultural entomology on the Chinese scene. *Journal of Applied Entomology*. https://doi.org/10.1111/jen.12714

Staunton, K. M., Crawford, J. E., Liu, J., Townsend, M., Han, Y., Desnoyer, M., Howell, P., Xiang, W., Burkot, T. R., Snoad, N., & Ritchie, S. A. (2020). A Low-Powered and Highly Selective Trap for Male Aedes (Diptera: Culicidae) Surveillance: The Male Aedes Sound Trap. *Journal of Medical Entomology*. https://doi.org/10.1093/jme/tjaa151

Staunton, K. M., Rohde, B. B., Townsend, M., Liu, J., Desnoyer, M., Howell, P., Amos, B., Crawford, J., Snoad, N., & Ritchie, S. A. (2020). Investigating male *Aedes aegypti* (Diptera: Culicidae) attraction to different oviposition containers using various configurations of the Sound Gravid Aedes Trap. *Journal of Medical Entomology*, *57*(3), 957–961. https://doi.org/10.1093/jme/tjz229

Staunton, K. M., Usher, L., Prachar, T., Ritchie, S. A., Snoad, N., & Johnson, B. J. (2019). A novel methodology for recording wing beat frequencies of untethered male and female *Aedes aegypti*. *Journal of the American Mosquito Control Association*, *35*(3), 169–177. https://doi.org/10.2987/18-6799.1

Stevens, S. S., Volkmann, J., & Newman, E. B. (1937). A Scale for the Measurement of the Psychological Magnitude Pitch. *Journal of the Acoustical Society of America*, *8*(3), 185–190. https://doi.org/10.1121/1.1915893

Su Yin, M., Haddawy, P., Ziemer, T., Wetjen, F., Supratak, A., Chiamsakul, K., Siritanakorn, W., Chantanalertvilai, T., Sriwichai, P., & Sa-ngamuang, C. (n.d.). A deep learning-based pipeline for mosquito detection and classification from wingbeat sounds. https://doi.org/10.1007/s11042-022-13367-0

Swan, T., Russell, T. L., Staunton, K. M., Field, M. A., Ritchie, S. A., & Burkot, T. R. (2022). A literature review of dispersal pathways of Aedes albopictus across different spatial scales: implications for vector surveillance. *Parasites & Vectors 2022 15:1*, *15*(1), 1–13. https://doi.org/10.1186/S13071-022-05413-5

Tedjou ID, A. N., Kamgang, B. I., lie Yougang, A. P., Njiokou, F., & Wondji, C. S. (2019). Update on the geographical distribution and prevalence of *Aedes aegypti* and *Aedes albopictus* (Diptera: Culicidae), two major arbovirus vectors in Cameroon. https://doi.org/10.1371/journal.pntd.0007137

TensorFlow developers. (n.d.). *TensorFlow*. Retrieved October 25, 2023, from https://www.tensorflow.org

Thongsripong, P., Hyman, J. M., Kapan, D. D., & Bennett, S. N. (2021). Human-mosquito contact: a missing link in our understanding of mosquito-borne disease transmission dynamics. *Entomological Society of America*, *114*(4), 397–414. https://doi.org/10.1093/aesa/saab011

Troupin, A., & Colpitts, T. M. (2016). Overview of west Nile virus transmission and epidemiology. *Methods in Molecular Biology*, *1435*, 15–18. https://doi.org/10.1007/978-1-4939-3670-0_2/COVER

Vasconcelos, D., Nunes, N. J., & Gomes, J. (2020). An annotated dataset of bioacoustic sensing and features of mosquitoes. *Scientific Data 2020 7:1*, *7*(1), 1–8. https://doi.org/10.1038/s41597-020-00725-6

Vasconcelos, D., Nunes, N., Ribeiro, M., Prandi, C., & Rogers, A. (2019). LOCOMOBIS: a low-cost acoustic-based sensing system to monitor and classify mosquitoes. *16th IEEE Annual Consumer Communications & Networking Conference (CCNC)*.

Verdonschot, P. F. M., & Besse-Lototskaya, A. A. (2014). Flight distance of mosquitoes (Culicidae): A metadata analysis to support the management of barrier zones around rewetted and newly constructed wetlands. *Limnologica*, *45*, 69–79. https://doi.org/10.1016/j.limno.2013.11.002

Vieira Bruno, R., Cristina Farnesi, L., & Ordunha Araripe, L. (2020). The effects of infection on mosquito rhythmic behavior. *Vector-Borne Diseases - Recent Developments in Epidemiology and Control*. https://doi.org/10.5772/INTECHOPEN.89409

Vilibic-Cavlek, T., Petrovic, T., Savic, V., Barbic, L., Tabain, I., Stevanovic, V., Klobucar, A., Mrzljak, A., Ilic, M., Bogdanic, M., Benvin, I., Santini, M., Capak, K., Monaco, F., Listes, E., & Savini, G. (2020). Epidemiology of Usutu Virus: The European Scenario. *Pathogens, 9*(9), 699. https://doi.org/10.3390/PATHOGENS9090699

Villarreal, Susa. M., Winokur, O., & Harrington, L. (2017). The impact of temperature and body size on fundamental flight tone variation in the mosquito vector *Aedes aegypti* (Diptera: Culicidae):

Implications for acoustic lures. *Journal of Medical Entomology*, *54*(5), 1116–1121. https://doi.org/10.1093/jme/tjx079

Villwock, S., & Pacas, M. (2008). Application of the welch-method for the identification of two- and three-mass-systems. *IEEE Transactions on Industrial Electronics*, *55*(1), 457–466. https://doi.org/10.1109/TIE.2007.909753

Virtanen, T., Vincent, E., & Gannot, S. (2018). *Time-frequency processing - Spectral properties*. https://hal.inria.fr/hal-01881426

Vlaskamp, D. R. M., Thijsen, S. F. T., Reimerink, J., Hilkens, P., Bouvy, W. H., Bantjes, S. E., Vlaminckx, B. J. M., Zaaijer, H., van den Kerkhof, H. H. T. C., Raven, S. F. H., & Reusken, C. B. E. M. (2020). First autochthonous human west nile virus infections in the Netherlands, July to August 2020. *Eurosurveillance*, *25*(46), 1–4. https://doi.org/10.2807/1560-7917.ES.2020.25.46.2001904/CITE/PLAINTEXT

Vogels, C. B. F., Fros, J. J., Pijlman, G. P., Van Loon, J. J. A., Gort, G., & Koenraadt, C. J. M. (2017). Virus interferes with host-seeking behaviour of mosquito. *The Journal of Experimental Biology*, *220*(Pt 19), 3598–3603. https://doi.org/10.1242/JEB.164186

Vogels, C. B. F., Göertz, G. P., Pijlman, G. P., & Koenraadt, C. J. M. (2017a). Vector competence of European mosquitoes for West Nile virus. *Emerging Microbes & Infections*, *6*(11). https://doi.org/10.1038/EMI.2017.82

Vogels, C. B. F., Göertz, G. P., Pijlman, G. P., & Koenraadt, C. J. M. (2017b). Vector competence of northern and southern European Culex pipiens pipiens mosquitoes for West Nile virus across a gradient of temperatures. *Medical and Veterinary Entomology*, *31*(4), 358–364. https://doi.org/10.1111/mve.12251

Wahid, I., Sunahara, T., & Mogi, M. (2002). Mouthparts of male Aedes (Stegomyia) mosquitoes. *Journal of the American Mosquito Control Association*, *18*(2), 73–80. https://europepmc.org/article/med/12083358

Wang, G. H., Gamez, S., Raban, R. R., Marshall, J. M., Alphey, L., Li, M., Rasgon, J. L., & Akbari, O. S. (2021). Combating mosquito-borne diseases using genetic control technologies. *Nature Communications 2021 12:1*, *12*(1), 1–12. https://doi.org/10.1038/s41467-021-24654-z

Wang, J., Zhu, S., Lin, Y., Svanberg, S., & Zhao, G. (2020). Mosquito counting system based on optical sensing. *Applied Physics B: Lasers and Optics*, *126*(2), 28. https://doi.org/10.1007/s00340-019-7361-2

Weaver, S. C., Charlier, C., Vasilakis, N., & Lecuit, M. (2018). Zika, chikungunya, and other emerging vector-borne viral diseases. *Https://Doi.Org/10.1146/Annurev-Med-050715-105122*, *69*, 395–408. https://doi.org/10.1146/ANNUREV-MED-050715-105122

Wei, X., Zakir Hossain, M., & Asif Ahmed, K. (2022). A ResNet attention model for classifying mosquitoes from wing-beating sounds. *Scientific Reports*, *12*(10334). https://doi.org/10.1038/s41598-022-14372-x

WHO. (2008).International health regulations (2005)

WHO. (2014). A global brief on vector-borne diseases. www.who.int

WHO. (2016). Technical handbook for dengue surveillance, dengue outbreak prediction/detection and outbreak response.

WHO. (2017). Global Vector Control response 2017 -2030: An integrated approach for the control of vector borne diseases. *World Health Organization*, *March*, 2. http://www.who.int/vector-control/burden_vector-borne_diseases.pdf

WHO. (2020). *Vector borne diseases*. World Health Organization. https://www.who.int/news-room/fact-sheets/detail/vector-borne-diseases

WHO. (2023). *Factsheets: Malaria*. World Health Organization. https://www.who.int/news-room/fact-sheets/detail/malaria

Wilder-Smith, A., & Osman, S. (2020). Public health emergencies of international concern: a historic overview. *Journal of Travel Medicine*, *27*(8), 19. https://doi.org/10.1093/JTM/TAAA227

Wilkerson, R. C., Linton, Y.-M., & Strickman, D. (2021). Mosquito Diversity, Systematics, and Medical Importance. *Mosquitoes of the World*, 352. https://books.google.com/books/about/Mosquitoes_of_the_World.html?hl=es&id=xsoMEAAAQBAJ

Wilson Id, A. L., Courtenay, O., Kelly-Hope, L. A., Scott Id, T. W., Takken, W., Torr, S. J., & Lindsay, S. W. (2020). *The importance of vector control for the control and elimination of vector-borne diseases*. https://doi.org/10.1371/journal.pntd.0007831

Work, T. H., Hurlbut, H. S., & Taylor, R. M. (1955). Indigenous wild birds of the Nile Delta as potential West Nile virus circulating reservoirs. *The American Journal of Tropical Medicine and Hygiene*, *4*(5), 872–888. https://doi.org/10.4269/AJTMH.1955.4.872

XGBoost developers. (n.d.). *XGBoost Documentation*. Retrieved October 25, 2023, from https://xgboost.readthedocs.io/en/stable/

Yan, J., Gangoso, L., Ruiz, S., Soriguer, R., Figuerola, J., & Martínez-De La Puente, J. (2021). Understanding host utilization by mosquitoes: determinants, challenges and future directions. https://doi.org/10.1111/brv.12706

Yang, C., Wang, F., Huang, D., Ma, H., Zhao, L., Zhang, G., Li, H., Han, Q., Bente, D., Salazar, F. V., Yuan, Z., & Xia, H. (2022). Vector competence and immune response of *Aedes aegypti* for Ebinur Lake virus, a newly classified mosquito-borne orthobunyavirus. *PLoS Neglected Tropical Diseases*, *16*(7). https://doi.org/10.1371/JOURNAL.PNTD.0010642

Zhang, Z. (2016). A gentle introduction to artificial neural networks. *Annals of Translational Medicine*, *4*(19). https://doi.org/10.21037/atm.2016.06.20

Zhao, X., Jia, P., Zhang, J., Yang, Y., Liu, W., & Zhang, J. (2018). Structural glycoprotein LmAbd-9 is required for the formation of the endocuticle during locust molting. https://doi.org/10.1016/j.ijbiomac.2018.11.279

Zhou, J., Liu, S., Liu, H., Xie, Z., Liu, L., Lin, L., Jiang, J., Yang, M., Zhou, G., Gu, J., Zhou, X., Yan, G., James, A. A., & Chen, X.-G. (2022). Interspecific mating bias may drive Aedes albopictus displacement of Aedes aegypti during its range expansion. *PNAS Nexus*, *1*(2), 1–10. https://doi.org/10.1093/PNASNEXUS/PGAC041

Zhu, L. Q. (2011). Insect sound recognition based on MFCC and PNN. *Proceedings - 2011 International Conference on Multimedia and Signal Processing, CMSP 2011*, *2*, 42–46. https://doi.org/10.1109/CMSP.2011.100

Ziegler, U., Santos, P. D., Groschup, M. H., Hattendorf, C., Eiden, M., Höper, D., Eisermann, P., Keller, M., Michel, F., Klopfleisch, R., Müller, K., Werner, D., Kampen, H., Beer, M., Frank, C., Lachmann, R., Tews, B. A., Wylezich, C., Rinder, M., … Lühken, R. (2020). West Nile Virus epidemic in Germany triggered by epizootic emergence, 2019. *Viruses*, *12*(4). https://doi.org/10.3390/V12040448