

**Environmental statistical modelling
of mosquito vectors
at different geographical scales**

Daniela Cianci

Cover design by Ron Ariëns (Grond ontwerp)

Printed and lay-out by Ridderprint

ISBN 978-90-393-6411-6

**Environmental statistical modelling
of mosquito vectors
at different geographical scales**

Statistische modellen van door muggen overgedragen
ziekten, op verschillende ruimtelijke schalen

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht
op gezag van de rector magnificus, prof.dr. G.J. van der Zwaan,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen
op donderdag 22 oktober 2015 des middags te 12.45 uur

door

Daniela Cianci

geboren op 4 april 1985 te Lanciano (Italië)

Promotor: Prof.dr. J.A.P. Heesterbeek

Copromotor: Dr. N.A. Hartemink

This thesis was accomplished with financial support from the EU grant FP7-261504 EDENext and is catalogued by the EDENext Steering Committee as EDENext 366 (<http://www.edenext.eu>).

CONTENTS

Chapter 1	Introduction	7
Chapter 2	Modelling the potential spatial distribution of mosquito species using three different techniques	89
Chapter 3	Modelling the spatial distribution of the nuisance mosquito species <i>Anopheles plumbeus</i> (Diptera: Culicidae) in the Netherlands	37
Chapter 4	High resolution spatial analysis of habitat preference of <i>Aedes albopictus</i> (Diptera:Culicidae) in an urban environment	53
Chapter 5	Estimating mosquito population size from mark–release–recapture data	71
Chapter 6	Characterizing mosquito breeding sites in an urban setting: a case study in Amsterdam	93
Chapter 7	The “Auto-Dissemination” approach: A novel concept to fight <i>Aedes albopictus</i> in urban areas	109
Chapter 8	Summarising discussion	125
	Samenvatting	135
	Summary	137
	Riassunto	139
	Dankwoord	141
	About the author	143
	List of publications	145

CHAPTER 1

Introduction

In this thesis, several statistical methods are applied to study the distribution of mosquito species that are potential vectors of pathogens threatening both human and animal health. The analyses are conducted at different geographical scales. In the first part of the thesis, the distribution is investigated of indigenous mosquito species in the Netherlands. In the second part, environmental preferences are explored of *Aedes albopictus* and *Culex pipiens*, in Rome and Amsterdam respectively. This chapter provides a general introduction to vector-borne diseases and some examples of mosquito-borne diseases and their introduction in Europe. The importance of gaining more knowledge on the spatial distribution of vectors is here explained and various existing modelling techniques used for this purpose are presented.

VECTOR-BORNE DISEASES IN EUROPE

Vector-borne diseases are infections transmitted by the bite of infected arthropods, such as mosquitoes, ticks, fleas, midges and flies. Within the past two decades, many important vector-borne diseases have (re-)emerged in humans and animals or expanded their geographical range. Often considered as primarily a problem for countries with tropical climate (but see Reiter 2001, 2008), vector-borne diseases pose an increasingly wider threat to global public health, both in terms of people affected and their geographical spread (WHO 2014). Mosquitoes are considered as prime candidate for transmitting (re-)emerging vector-borne diseases in Europe (Versteirt et al. 2009).

The invasive mosquito *Aedes albopictus*, commonly known as the 'Asian tiger mosquito', is a good example to illustrate the above. It is now present in several European countries (figure 1). It originated in Southeast Asia, but has spread in temperate Asia, Europe, North America, as well as in South Africa and in several locations in the Pacific and Indian Ocean (Knudsen 1995, Benedict et al. 2007, Paupy et al. 2009). In Europe, *Ae. albopictus* was reported for the first time in Albania in 1979 (Adhami and Murati 1987) and in Italy in 1990, in Genoa, and from there it has gradually spread to several Italian regions (Sabatini et al. 1990, Dalla Pozza and Majori 1992, Della Torre et al. 1992, Romi and Majori 2008). This species is a known vector of chikungunya virus, dengue virus, and a number of other viruses affecting human health have also been isolated from

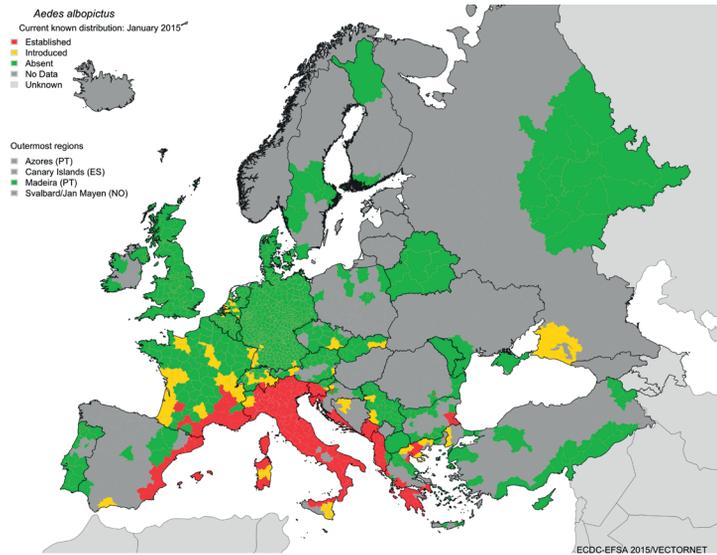


Figure 1. Current known distribution of *Aedes albopictus* in Europe at 'regional' administrative level (NUTS3). The map is based only on confirmed data (published and unpublished) provided by experts from the respective countries as part of the VBORNET project (VBORNET). Source: ECDC (ECDC 2015).

field-collected *Ae. albopictus* in different countries. Chikungunya was reported for the first time in Europe in 2007, in Emilia-Romagna province in Italy (Angelini et al. 2007, Rezza et al. 2007). Autochthonous chikungunya fever cases occurred also in south eastern France in 2010 (Gould et al. 2010). Although generally considered a secondary vector of dengue to *Aedes aegypti*, *Ae. albopictus* has also been associated with dengue virus transmission (Gerhardt et al. 2001). Dengue was made notifiable in France in 2006 (Ledrans and Dejour Salamanca 2008) and the first autochthonous cases of dengue were reported in France during September 2010 (La Ruche et al. 2010) followed by others in Croatia at around the same time (Gjenero-Margan et al. 2011). In late 2012, a dengue outbreak occurred on Madeira, associated with *Ae. aegypti* (Sousa et al. 2012). *Ae. albopictus* is also known to cause significant biting nuisance (Scholte et al. 2007, Vazeille et al. 2008, Genchi et al. 2009); in areas where it is well-established its bites reduce the quality of life of individuals affected (Aranda et al. 2006, Abramides et al. 2011).

Another important group of potential vectors are the members of the *Culex* family that are generally considered the principal vectors of West Nile virus (WNV). In numerous European countries the virus has been isolated in mosquitoes, wild rodents, migrating birds, hard ticks, horses and human beings (Hubálek et al. 2006). In 2010, human cases of West Nile fever were detected in several south eastern EU countries. There have been recent outbreaks since then. In 2014, WNV infection has officially been reported in Austria, Greece, Hungary, Italy, Romania, Bosnia and Herzegovina, Russian Federation and Serbia, as it is shown in figure 2 (ECDC 2015).

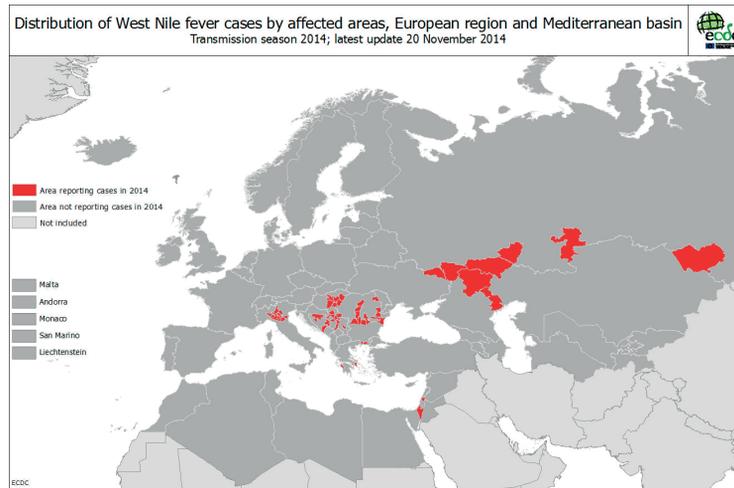


Figure 2. Human cases of West Nile fever reported in the EU and in neighbouring countries in 2014. Source: ECDC (ECDC 2014).

Anopheles spp mosquitoes are the third important class presented in this thesis. Species of this genus can transmit the *Plasmodium* parasite that can cause malaria. Historically malaria was endemic in Europe but it was eventually eliminated in 1975 through a number of factors related to socioeconomic development (Kuhn et al. 2002, Kuhn 2006). *Anopheles* mosquitoes are present in several European countries.

Environmental preferences of *Ae. albopictus* and some *Culex* and *Anopheles* species are investigated in this thesis. In the following paragraphs, it is explained why it is important to gain a better understanding of the spatial distribution of potential vector species and several modelling techniques used for this goal are described.

SPATIAL DISTRIBUTION OF VECTORS

Arthropod vectors are cold-blooded and thus especially sensitive to climatic factors. Weather influences survival and reproduction rates of vectors (Rogers and Randolph 2006), habitat suitability, vector distribution and abundance, intensity and temporal pattern of vector activity (particularly biting rates) throughout the year, and rates of development, survival and reproduction of pathogens within vectors (ECDC 2015). However, climate is only one of many factors influencing vector distribution; pesticide application, changes in land cover and the increased movement of people and goods (Wilson 1995, Tatem et al. 2006) are important for the introduction and spread of vectors and pathogens. Even if a pathogen is introduced, this does not necessary mean that the pathogen can be established.

A measure to study the risk of establishing of infectious diseases is the basic reproduction number, R_0 (Anderson and May 1990, Diekmann et al. 2013). R_0 is a quantity that integrates all factors that determine whether a pathogen can establish or not, in a properly weighted way. It is defined as the average number of secondary cases caused by one typical infectious individual placed in a population consisting entirely of susceptibles. R_0 is a measure of the success of invasion into a population; if R_0 is higher than 1, an outbreak is possible if the pathogen is introduced, whereas if R_0 is smaller than 1, introduction will fail. In order to develop basic reproduction number (R_0) models and construct risk maps that indicate the risk for an outbreak after an introduction, abundance data of vectors are an essential parameter (Hartemink et al. 2009, 2011).

Although vector presence and abundance are not the only factors determining whether or not a pathogen can spread in an area, determining the distribution of the vector is an essential step in studying the risk of transmission of a pathogen. An understanding of the spatial extent of potential vector species, their abundance and seasonal activity, is important for estimating levels of risks of vector-borne diseases, enabling better targeting for surveillance and helps in designing control measures.

SPECIES DISTRIBUTION MODELS

Species distribution models (SDMs) are numerical tools that combine observations of species occurrence or abundance with environmental characteristics. The environment is in most cases represented by climate data (such as temperature, and precipitation), but other variables such as soil type, water depth, and land cover can also be used. Species distribution models quantitatively describe areas that support the presence of a given species, based on known occurrence or abundance data and the associated environmental conditions (Elith and Leathwick 2009). The general idea behind SDMs is to identify, through a statistical model, relationships between known occurrence or abundance of a species and environmental data and to use these relationships to make predictions for all un-sampled areas in the study region (figure 3). These models are used in several research areas in conservation biology, ecology, evolution, biogeography, evolution, and climate change research (Guisan and Thuiller 2005).

A general and recurrent problem in SDM building is identifying the appropriate scale for modelling (Wiens 2002). Small scale maps show less detail and cover larger areas such as a whole country. Large scale maps show more detail and cover smaller areas. The larger the map scale the higher the possible resolution (grid size). As scale decreases, resolution diminishes. The choice of scale is closely related to the type of species considered. Understanding the processes driving the observed distribution patterns is also essential to avoid a mismatch between the scale used for modelling and the one at which key processes occur. Patterns observed on one scale may not be apparent on another scale (Guisan and Thuiller 2005, Zeimes 2015) and the use of different geographical extents may provide contradictory answers to the same ecological question (Thuiller et

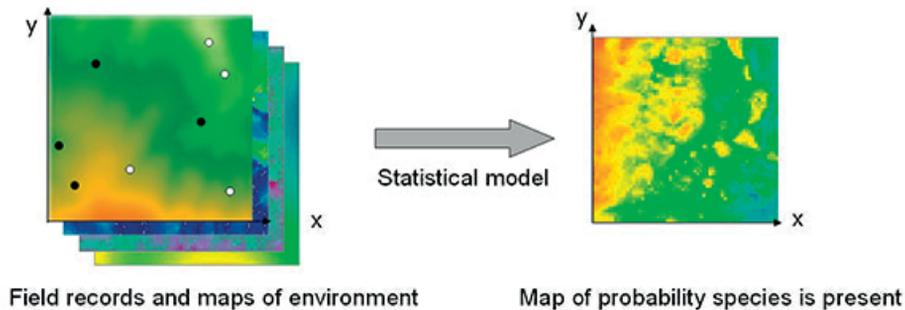


Figure 3. Distribution modelling is used to bring together field data and maps of environmental variables to create habitat suitability maps (map giving the probabilities of species' presence) as results of models (figure taken from <http://www.biodiversityscience.com/2011/04/27/species-distribution-modelling/>).

al. 2003). Hence, the selection of resolution and extent is a critical step in SDM building and an inappropriate selection can yield misleading results. Every question in the field of spatial ecology should be answered referring explicitly to resolution and extent of the study area at which data are measured or analysed (Wiens 2002).

CONTROL MEASURES

A wide range of interventions to control and prevent vector-borne diseases exist. Depending on factors such as local ecology and behaviour of the vector species, human activity, socioeconomic conditions, and environmental change, some will be more appropriate than others. Examples of interventions currently used include long-lasting insecticidal nets, indoor and outdoor spraying, addition of chemicals to water and insect repellents (WHO 2014).

Vector control can also be done adopting environmental management strategies. The breeding habitats can be reduced also through the waste management (e.g., empty tins, plastic bottles, car tyres can serve as breeding site). The vectors can be controlled through the introduction of parasites, predators or other living organisms or through genetic control strategies, which aim to suppress target population or to introduce a harm-reducing trait (WHO 2014).

To face the growing risk of MBD epidemics, the European Centre for Disease Prevention and Control (ECDC) has established a network of medical entomologists and public health experts (VBORNET) and produced guidelines to support the implementation of invasive mosquito species surveillance in Europe (Medlock et al. 2012, Deblauwe et al. 2014). Recently, the control methods available or in development against invasive mosquitoes with a particular focus on those which can be implemented in Europe have been reviewed. These control methods have been divided into five categories: environmental (source reduction), mechanical (trapping), biological (e.g.

copepods, Bti, Wolbachia), chemical (insect growth regulators, pyrethroids) and genetic (sterile insect technique and genetically modified mosquitoes). Effectiveness, ecological impact, sustainability and stage of development of each control method are discussed in the review (Baldacchino et al. 2015).

In this thesis we describe a particularly innovative approach, i.e., the auto-dissemination method. This method uses the natural behaviour of female mosquitoes to go from one breeding site to another to spread a larvacide. An assessment of the effectiveness of this approach is described in this thesis.

OUTLINE OF THIS THESIS

This thesis is part of the research program EDENext [www.edenext.eu], a European project that investigates and models the processes leading to the introduction, establishment and spread of vectors and vector-borne disease and aims at assessing the effect of possible control strategies on breaking their epidemiological cycles. Biological, ecological and epidemiological components of vector-borne disease are studied in this project.

In this thesis, the ecology of mosquito-borne diseases was studied with the aim to improve the knowledge on the spatial distribution and the abundance of vectors. We describe tools to gain better knowledge on the distribution of mosquitoes that are potential vectors of diseases, using surveillance and experimental data. New knowledge on vectors is necessary to predict the (re-) emergence and spread of vector-borne diseases and to develop new interventions to interrupt or limit the spread of vector-borne diseases.

In **chapter 2** three different modelling techniques, i.e., non-linear discriminant analysis, random forest and generalised linear model, were used to explore the potential spatial distribution in the Netherlands for three indigenous mosquito species (*Culiseta annulata*, *Anopheles claviger* and *Ochlerotatus punctor*). The models were used to find a relationship between the occurrence of the species and environmental features and to produce environmental suitability maps in the Netherlands. A second aim of this study was to compare the results produced with the different techniques.

A more specific application of one of the techniques described in the previous study is presented in **chapter 3**, where the potential spatial distribution of one of the most frequently observed species in the Netherlands, *Anopheles plumbeus*, is investigated using random forest models. This species is a nuisance mosquito species and a potential malaria vector. The models were used to find a relationship between the occurrence and the abundance of this species, and environmental features, in order to produce distribution maps for *An. plumbeus* in the Netherlands.

In **chapters 4-6** the spatial distribution of mosquito species is investigated in urban areas, in Rome (Italy) and Amsterdam (the Netherlands). Because the study areas are smaller than in the

studies presented in the previous chapters, it was possible to analyse the environment in more detail, at a higher resolution.

The habitat preference of *Ae. albopictus* is investigated in **chapter 4**. This species is a nuisance and a potential vector of several human pathogens and it has a great capacity to adapt to different environments. Although it has also colonised urban areas in Europe, the habitat preference in urban environments has not been studied in Southern European cities. In this chapter, spatial statistical models were used to evaluate the relationship between egg abundances and land cover types on the campus of Sapienza University in Rome, which is taken as an example of a European urban habitat.

In the same study area, an estimate of the population abundance of *Ae. albopictus* is provided (**chapter 5**). In this chapter, a new method specifically developed to estimate the population size including features of mark-release-recapture (MRR) experiments carried out with mosquitoes is presented. The method is a logistic regression model based on the principle of the well-known Fisher-Ford method, which is based on the assumption that the ratio of recaptured individuals to the total captures approximates the ratio of marked individuals released to the total population. In addition, the model presented here takes into consideration the distance between release points and traps, the time between release and recapture, and the loss of marked mosquitoes to death or dispersal. The performance and accuracy of the logistic regression model has been assessed using simulated data from known population sizes and a real dataset containing data from MRR experiments with *Ae. albopictus* in the campus of Sapienza University in Rome.

An additional study that aims at giving insights in urban mosquito ecology is presented in **chapter 6**. Here the breeding sites of the indigenous mosquito species in Amsterdam are investigated with the aim of identifying the factors that determine a good larval habitat for these species.

A better understanding of the spatial distribution of vectors is needed for help in successful prevention and control activities. In **chapter 7**, the feasibility of the “auto-dissemination” approach as a possible alternative to traditional control tools against *Ae. albopictus* in urban areas is assessed. The approach is based on the possibility that wild adult females, exposed to artificial resting sites contaminated with pyriproxyfen, can disseminate this juvenile hormone analogue to larval habitats, thus interfering with adult emergence.

In **chapter 8** the main issues encountered in **chapters 2-7** are described. Approaches to handle them are proposed and the results are discussed.

REFERENCES

- Abramides, G. C., D. Roiz, R. Guitart, S. Quintana, I. Guerrero, and N. Giménez. 2011. Effectiveness of a multiple intervention strategy for the control of the tiger mosquito (*Aedes albopictus*) in Spain. *Trans. R. Soc. Trop. Med. Hyg.* 105: 281–288.
- Anderson, R. M., and R. M. May. 1990. Immunisation and herd immunity. *Lancet.* 335: 641–645.
- Angelini, R., A. C. Finarelli, P. Angelini, C. Po, K. Petropulacos, G. Silvi, P. Macini, et al. 2007. Chikungunya in north-eastern Italy: a summing up of the outbreak. *Euro Surveill.* 12: E071122.
- Aranda, C., R. Eritja, and D. Roiz. 2006. First record and establishment of the mosquito *Aedes albopictus* in Spain. *Med. Vet. Entomol.* 20: 150–152.
- Baldacchino, F., B. Caputo, F. Chandre, A. Drago, A. della Torre, F. Montarsi, et al. 2015. Control methods against invasive *Aedes* mosquitoes in Europe: a review. *Pest Manag. Sci.*
- Deblauwe, I., C. Sohler, F. Schaffner, L. M. Rakotoarivony, and M. Coosemans. 2014. Implementation of surveillance of invasive mosquitoes in Belgium according to the ECDC guidelines. *Parasit. Vectors.* 7: 1–11.
- Diekmann, O., and J. A. P. Heesterbeek. 2000. *Mathematical epidemiology of infectious diseases: model building, analysis, and interpretation*, Wiley Seri. ed. John Wiley & Sons LTD, New York.
- ECDC. 2014. West Nile fever maps. [http://ecdc.europa.eu/en/healthtopics/west_nile_fever/west-nile-fever-maps/pages/index.aspx#sthash.cdOF4DLX.dpuf].
- ECDC. 2015. Mosquito maps. [http://ecdc.europa.eu/en/healthtopics/vectors/vector-maps/Pages/VBOR-NET_maps.aspx].
- ECDC. 2015. European Centre for Disease Prevention and Control. [ecdc.europa.eu].
- EDENext. n.d. EDENext. [www.edenext.eu].
- Elith, J., and J. R. Leathwick. 2009. Species distribution models: ecological explanation and prediction across space and time. *Annu. Rev. Ecol. Syst.* 40: 677–697.
- Genchi, C., L. Rinaldi, M. Mortarino, M. Genchi, and G. Cringoli. 2009. Climate and *Dirofilaria* infection in Europe. *Vet. Parasitol.* 163: 286–292.
- Gerhardt, R. R., K. L. Gottfried, C. S. Apperson, B. S. Davis, P. C. Erwin, A. B. Smith, et al. 2001. First isolation of La Crosse virus from naturally infected *Aedes albopictus*. *Emerg. Infect. Dis.* 7: 807–811.
- Gjenero-Margan, I., B. Aleraj, D. Krajcar, V. Lesnikar, A. Klobucar, I. Pem-Novosel, et al. 2011. Autochthonous dengue fever in Croatia, August-September 2010. *Euro Surveill.* 16: 19805.
- Gould, E. A., P. Gallian, X. De Lamballerie, and R. N. Charrel. 2010. First cases of autochthonous dengue fever and chikungunya fever in France: from bad dream to reality! *Clin. Microbiol. Infect.* 16: 1702–1704.
- Guisan, A., and W. Thuiller. 2005. Predicting species distribution: offering more than simple habitat models. *Ecol. Lett.* 8: 993–1009.
- Hartemink, N. A., B. V Purse, R. Meiswinkel, H. E. Brown, A. De Koeijer, A. R. W. Elbers, et al. 2009. Mapping the basic reproduction number (R0) for vector-borne diseases: A case study on bluetongue virus. *Epidemics.* 1: 153–161.
- Hartemink, N., S. O. Vanwambeke, H. Heesterbeek, D. Rogers, D. Morley, B. Pesson, et al. 2011. Integrated Mapping of Establishment Risk for Emerging Vector-Borne Infections: A Case Study of Canine Leishmaniasis in Southwest France. *PLoS One.* 6: e20817.

- Hubálek, Z., B. Kříž, and B. Menne. 2006. West Nile virus: ecology, epidemiology and prevention.
- Kuhn, K. G. 2006. Malaria. In *Climate Change and Adaptation Strategies for Human Health*, WHO, Stein. ed. Darmstadt.
- Kuhn, K. G., D. H. Campbell-Lendrum, and C. R. Davies. 2002. A continental risk map for malaria mosquito (Diptera: Culicidae) vectors in Europe. *J. Med. Entomol.* 39: 621–630.
- Ledrans, M., and D. Dejour Salamanca. 2008. Cas importés de chikungunya et de dengue en France métropolitaine. *Saint-Maurice Inst. Veill. Sanit.*
- Medlock, J. M., F. Hansford, Kayleigh M. Schaffner, G. Versteirt, Veerle Hendrickx, et al. 2012. A review of the Invasive Mosquitoes in Europe: Ecology, Public Health Risks, and Control Options. *Vector-Borne Zoonotic Dis.* 12: 435–447.
- Reiter, P. 2001. Climate change and mosquito-borne disease. *Environ. Health Perspect.* 109: 141.
- Reiter, P. 2008. Global warming and malaria: knowing the horse before hitching the cart. *Malar. J.* 7: S3.
- Rezza, G., L. Nicoletti, R. Angelini, R. Romi, A. C. Finarelli, M. Panning, et al. 2007. Infection with chikungunya virus in Italy: an outbreak in a temperate region. *Lancet.* 370: 1840–1846.
- Rogers, D. J., and S. E. Randolph. 2006. Climate change and vector-borne diseases. *Adv. Parasitol.* 62: 345–381.
- La Ruche, G., Y. Souarès, A. Armengaud, F. Peloux-Petiot, P. Delaunay, P. Desprès, et al. 2010. First two autochthonous dengue virus infections in metropolitan France, September 2010. *Euro Surveill.* 15: 19676.
- Scholte, E., E. Dijkstra, H. Ruijs, F. Jacobs, W. Takken, A. Hofhuis, et al. 2007. The Asian tiger mosquito (*Aedes albopictus*) in the Netherlands: should we worry?, p. 131. *In Proc. Sect. Exp. Appl. Entomol. Entomol. Soc.*
- Sousa, C. A., M. Clairouin, G. Seixas, B. Viveiros, M. T. Novo, A. C. Silva, et al. 2012. Ongoing outbreak of dengue type 1 in the Autonomous Region of Madeira, Portugal: preliminary report. *Euro Surveill.* 17: 20333.
- Tatem, A. J., S. I. Hay, and D. J. Rogers. 2006. Global traffic and disease vector dispersal. *Proc. Natl. Acad. Sci.* 103: 6242–6247.
- Thuiller, W., M. B. Araújo, and S. Lavorel. 2003. Generalized models vs. classification tree analysis: predicting spatial distributions of plant species at different scales. *J. Veg. Sci.* 14: 669–680.
- Vazeille, M., C. Jeannin, E. Martin, F. Schaffner, and A. Failloux. 2008. Chikungunya: a risk for Mediterranean countries. *Acta Trop.* 105: 200.
- VBORNET. n.d. VBORNET. [www.vbornet.eu].
- Versteirt, V., E. De Clercq, W. Dekoninck, D. Damiens, A. Ayrinhac, F. Jacobs, et al. 2009. Mosquito vectors of disease: spatial biodiversity, drivers of change, and risk. Brussels.
- WHO. 2014. A global brief on vector-borne diseases.
- Wiens, J. A. 2002. Predicting species occurrences: progress, problems, and prospects. *Predict. species Occur. issues accuracy scale.* 739–749.
- Wilson, M. E. 1995. Travel and the emergence of infectious diseases. *Emerg. Infect. Dis.* 1: 39.
- Zeimes, C. 2015. Combining hazard and exposure to model the spatial distribution of two zoonoses, based on human case records.

CHAPTER 2

Modelling the potential spatial distribution of mosquito species using three different techniques

Daniela Cianci¹, Nienke Hartemink¹, and Adolfo Ibáñez-Justicia²

¹ Faculty of Veterinary Medicine, Utrecht University, Utrecht, The Netherland

² Centre for Monitoring of Vectors, Food and Consumer Product Safety Authority, Wageningen, The Netherlands

International Journal of Health Geographics, 14:10 (2015);

DOI: 10.1186/s12942-015-0001-0

ABSTRACT

Models for the spatial distribution of vector species are important tools in the assessment of the risk of establishment and subsequent spread of vector-borne diseases. The aims of this study are to define the environmental conditions suitable for several mosquito species through species distribution modelling techniques, and to compare the results produced with the different techniques. Three different modelling techniques, i.e., non-linear discriminant analysis, random forest and generalised linear model, were used to investigate the environmental suitability in the Netherlands for three indigenous mosquito species (*Culiseta annulata*, *Anopheles claviger* and *Ochlerotatus punctor*). Results obtained with the three statistical models were compared with regard to: (i) environmental suitability maps, (ii) environmental variables associated with occurrence, (iii) model evaluation. The models indicated that precipitation, temperature and population density were associated with the occurrence of *Cs. annulata* and *An. claviger*, whereas land surface temperature and vegetation indices were associated with the presence of *Oc. punctor*. The maps produced with the three different modelling techniques showed consistent spatial patterns for each species, but differences in the ranges of the predictions. Non-linear discriminant analysis had lower predictions than other methods. The model with the best classification skills for all the species was the random forest model, with specificity values ranging from 0.89 to 0.91, and sensitivity values ranging from 0.64 to 0.95. We mapped the environmental suitability for three mosquito species with three different modelling techniques. For each species, the maps showed consistent spatial patterns, but the level of predicted environmental suitability differed; NLDA gave lower predicted probabilities of presence than the other two methods. The variables selected as important in the models were in agreement with the existing knowledge about these species. All model predictions had a satisfactory to excellent accuracy; best accuracy was obtained with random forest. The insights obtained can be used to gain more knowledge on vector and non-vector mosquito species. The output of this type of distribution modelling methods can, for example, be used as input for epidemiological models of vector-borne diseases.

Keywords: species distribution modelling, non-linear discriminant analysis, random forest, generalised linear model, indigenous mosquito species, vector-borne diseases

BACKGROUND

Mosquitoes (Diptera:Culicidae) are known to be vectors of a large number of pathogens around the globe. Blood-feeding females of several mosquito species are involved in transmission of protozoa (e.g. *Plasmodium*), nematodes and viruses. Mosquitoes are considered as prime candidates for transmitting (re-)emerging vector-borne diseases in Europe (Versteirt et al. 2009).

Accurate information on the spatial distribution of mosquito species is essential for our understanding of the current risk of diseases transmitted by mosquitoes and for preparing for future threats (Versteirt et al. 2013). For the modelling of the spatial distribution of species, several techniques exist (Elith et al. 2006, Rogers 2006), differing in assumptions and predictive performance. The general idea behind species distribution modelling is to identify relationships between known occurrence of a species (presence/absence) and environmental data (e.g. meteorological data, land use covers, remote sensing data) and to use these relationships to make predictions for all un-sampled areas in the study region.

Here, we compare non-linear discriminant analysis (NLDA), random forest (RF) and generalised linear models (GLM), three techniques that have not been compared before, by applying them to a new dataset consisting of systematically collected data on three indigenous mosquito species in the Netherlands. The three mosquito species are *Culiseta annulata* (Schrank, 1776), *Anopheles claviger* (Meigen, 1804) and *Ochlerotatus punctor* (Kirby, 1837). For each of these species, the resulting environmental suitability maps and the most important environmental variables selected by the models are discussed and the techniques are compared in terms of model performance.

RESULTS

The environmental suitability for *Culiseta annulata*, *Anopheles claviger* and *Ochlerotatus punctor* was investigated in the Netherlands using three statistical models, i.e., NLDA, RF and GLM. Through these modelling techniques, occurrence data collected in 766 locations were linked to environmental factors.

The maps in Figure 1 show the observed occurrence data (i.e., the model input) and the predicted environmental suitability (i.e., the model outcomes). The ten most important variables for each model and each species are reported in Table 1. Accuracy measures are given in Table 2. For each species, the environmental suitability maps, the most important environmental variables and the model performances are reported below.

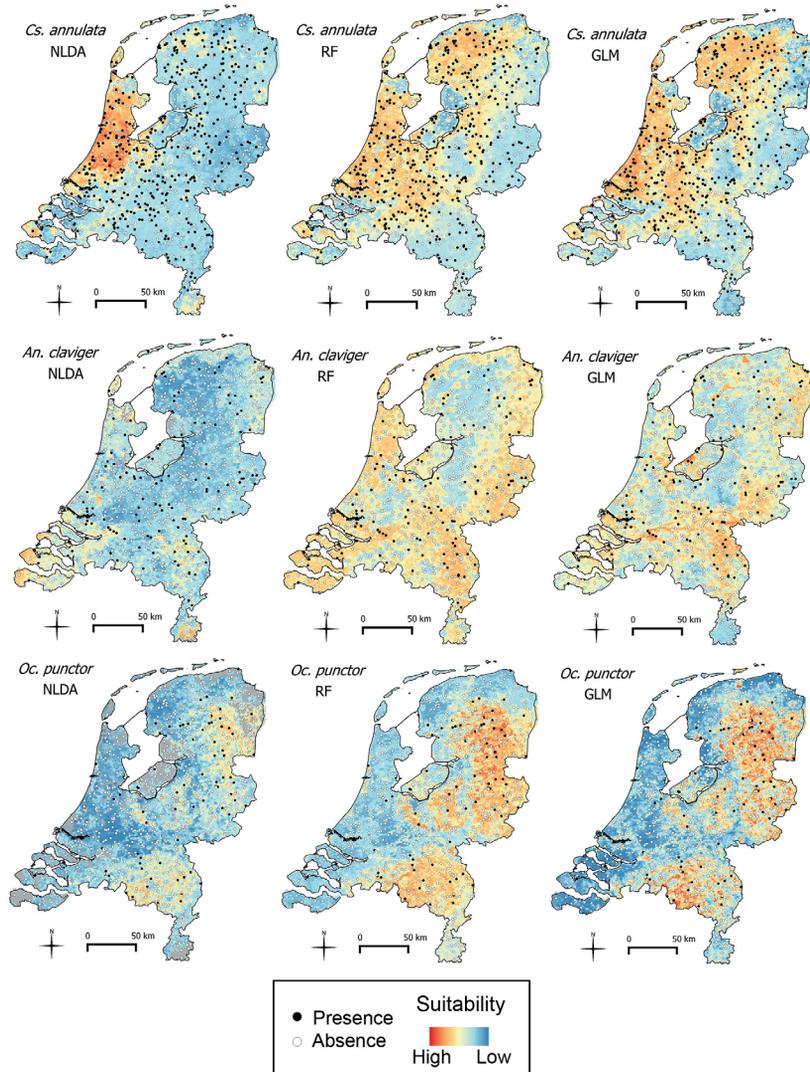


Figure 1. Environmental suitability maps for *Cs. annulata*, *An. claviger* and *Oc. punctor*, produced using non-linear discriminant analysis (NLDA), random forest (RF) and generalised linear model (GLM). Black dots indicate that the species was captured on the sampled locations and white dots indicate that the species was not captured. Environmental suitability is depicted using a gradient fill: blue indicates low environmental suitability, red indicates high suitability. NLDA and GLM bootstrapping was based on 150 presence points and 150 absence points for *Cs. annulata* and 100 presence points and 100 absence points for *An. claviger* and *Oc. punctor*.

Table 1 Most important variables per species and per model

Species	NLDA	RF	GLM
<i>Cs. annulata</i>	•Population density	•NLST P2	•EVI VR
	•WORLDCLIM precipitation P2	•Population density MIR A2	•DEM
	•WORLDCLIM precipitation A0	•DLST A2	•DLST A2
	•WORLDCLIM precipitation D1	•MIR MX	•NLST P3
	•WORLDCLIM precipitation DA	•NDVI A2	•CMORPH precipitation VR
	•CMORPH precipitation A1	•MIR P1	•CMORPH precipitation A3
	•DLST DA	•EVI MN	•DLST D1
	•DLST P1	•CMORPH precipitation P1	•DLST D3
	•DLST A0	•WORLDCLIM precipitation P1	•MIR A2
	•DLST P2	•DLST A3	•MIR 03
<i>An. claviger</i>	•WORLDCLIM precipitation P2	•NLST MX	•EVI P2
	•WORLDCLIM precipitation A0	•MIR MN	•DEM
	•Population density	•NLST A0	•NLST MN
	•MIR A3	•WORLDCLIM precipitation P3	•NLST A2
	•WORLDCLIM precipitation DA	•NLST MN	•CMORPH precipitation A 1
	•EVI D2	•DLST A0	•MIR D3
	•NLST P3	•DLST A1	•WORLDCLIM precipitation D3
	•EVI P2	•DLST MX	•CMORPH precipitation A2
	•NLST A3	•NDVI A2	•NLST A0
	•DLST A0	•NDVI VR	•Population density
<i>Oc. punctor</i>	•Population density	•Population density	•NDVI D1
	•MIR P1	•MIR P1	•MIR P1
	•EVI P3	•EVI P3	•DLST P2
	•NDVI P3	•NDVI P3	•EVI P2
	•NDVI P2	•NDVI P2	•MIR A3
	•DLST MN	•DLST MN	•WORLDCLIM precipitation A3
	•DEM	•DEM	•NDVI A3
	•CMORPH precipitation A2	•CMORPH precipitation A2	•WORLDCLIM P3
	•CMORPH precipitation A1	•CMORPH precipitation A1	•EVI MN
	•WORLDCLIM precipitation P3	•WORLDCLIM precipitation P3	•CMORPH precipitation A2

For non-linear discriminant analysis (NLDA) and generalised linear model (GLM) the top 10 variables average ranks are reported; for random forest (RF) the most important variables are expressed by the mean decrease in Gini index.

Table 2 Accuracy measures for the environmental suitability per species and per model

Species		NLDA	RF	GLM
<i>Cs. annulata</i>	Specificity (CI)	0.805 (0.596-0.884)	0.892 (0.808-0.936)	0.576 (0.442-0.811)
	Sensitivity (CI)	0.639 (0.541-0.829)	0.637 (0.696-0.779)	0.753 (0.498-0.868)
<i>An. claviger</i>	Specificity (CI)	0.670 (0.559-0.850)	0.908 (0.875-0.944)	0.652 (0.452-0.820)
	Sensitivity (CI)	0.772 (0.567-0.866)	0.890 (0.827-0.945)	0.709 (0.512-0.890)
<i>Oc. punctor</i>	Specificity (CI)	0.828 (0.735-0.954)	0.910 (0.825-0.944)	0.765 (0.574-0.828)
	Sensitivity (CI)	0.932 (0.795-1.00)	0.945 (0.890-1.00)	0.808 (0.685-0.945)

The confidence intervals (CI) are based on 2000 stratified bootstrap replicates. The best values for sensitivity and specificity for each species are printed in bold.

Culiseta annulata

Cs. annulata has the highest number of observed presences (438, Table 3) in the study. The presence points -indicated by black dots in Figure 1- show that this species was found almost all over the country. The environmental suitability maps show greater suitability in the western part of the Netherlands and lower suitability more inland. Although the areas identified as suitable (and unsuitable) are similar in all maps (Figure 1), the environmental suitability indicator for NLDA has a wider range of values (0.07-0.91) than RF (0.12-0.79) and GLM (0.10-0.86). This is visible in the different intensities of blue and red in the maps. RF and GLM identify also the northern and central part of the country as suitable environment. In all three models, precipitation and land surface temperature are important variables (Table 1). Population density comes up as an important variable in both the NLDA model and the RF model (first and second-most important variable, respectively). Highly forested areas (e.g., National Park Hoge Veluwe situated in the centre of the country) are not identified as suitable for this species in any of the three models. RF is the most specific method (specificity = 0.89), while GLM is the most sensitive (sensitivity = 0.75) (Table 2).

Table 3 Number of presence and absence points per species

Species	Presence	Absence
<i>Cs. annulata</i>	438	344
<i>An. claviger</i>	127	655
<i>Oc. punctor</i>	73	709

Anopheles claviger

The number of presence locations for *An. claviger* was 127 (Table 3). This species does not show a particular pattern in the distribution over the country (black dots in Figure 1). All maps (Figure 1) show lower environmental suitability for this species in the northern and central part of the country whereas the environment is identified as more suitable in the eastern part and in the coastal area (especially for RF and GLM). NLDA predicts much lower suitability values than the other techniques (minimum values: NLDA = 0.07, RF = 0.22, GLM = 0.17), only a few values are larger than 0.5 and as a consequence the average value is very low (average values: NLDA = 0.27, RF and GLM \approx 0.50). The GLM map indicates wetlands and floodplains as suitable environments. The most important variables are precipitation, land surface temperature, vegetation indices and middle infra-red, which is a vegetation related index (Table 1). Highly forested areas are not identified as suitable for this species. RF has excellent classification capabilities, the highest when compared to the other techniques (specificity = 0.91, sensitivity = 0.89) (Table 2).

Ochlerotatus punctor

Among the three species presented here, *Oc. punctor* is the least present (73 presence locations, Table 3) and it is the only species showing a clear pattern in the observations (Figure 1); there are more presence points in the east of the country (inland). All three models indicated higher environmental suitability in this part of the country (Figure 1). Comparing the suitability values obtained with the different techniques, the values are higher for RF and GLM and lower for NLDA (maximum values: NLDA = 0.81, RF = 0.94, GLM = 0.92). In the top 10 variables, middle-infra red, vegetation indices, precipitation and day land surface temperature are reported (Table 1). Population density is recorded as the most influential variable in NLDA and RF models. Highly forested areas are identified as suitable for this species in all three models. RF showed excellent classification capabilities (specificity = 0.91, sensitivity = 0.95), and it is the modelling technique with the best accuracy (Table 2).

DISCUSSION

The environmental suitability for *Cs. annulata*, *An. claviger* and *Oc. punctor* has been investigated using field and environmental data and applying three different modelling approaches, i.e., NLDA, RF and GLM. When comparing the maps for each species produced with the three different modelling techniques, we see consistent spatial patterns, but different levels of predicted environmental suitability. The average predicted environmental suitability was lower for NLDA than for the other methods. This is visible in the predominance of blue colours in the NLDA maps.

Most of the variables highlighted by the models as important are in agreement with field experience, existing biological knowledge, and known habitat preference of these species in Belgium (MODIRISK) (Versteirt et al. 2009). Precipitation and temperature for *Cs. annulata* and *An. claviger*

are important in both our study and MODIRISK. For *An. claviger* the population density is also reported as important, both in MODIRISK and in our study (for NLDA and GLM). Both studies show a preference of the latter species for the coastal area. The GLM map for *An. claviger* shows wetland and the floodplains of the big rivers as suitable environments, in accordance with field knowledge. *Oc. punctor* occurrence is related to land surface temperature and particularly with vegetation indices, in the Netherlands as well as in Belgium. This species is generally found in forests and natural areas. Population density is recorded in our study as one of the most influential variables in NLDA and RF models for *Oc. punctor* and *Cs. annulata*. *Cs. annulata* is known to breed in a wide variety of habitats and to be associated with areas with human activity (Becker et al. 2010), whereas *Oc. punctor* prefers swampy forest with boggy waters and seldom flies out of the forest (Becker et al. 2010), characteristics that suggest a negative relationship with human presence.

In terms of model performance, RF shows the best discrimination skills. Also in other studies, this technique was consistently reported to outperform other traditional modelling techniques (Cutler et al. 2007, Peters et al. 2007). Only the GLM for *Cs. annulata* has a higher sensitivity than RF. Random forest sensitivity and specificity are excellent, often equal to or larger than 0.9. However, if we want to interpret these values, we have to consider that the training data are also used to evaluate the model, meaning that the accuracy measures will be overestimated (Harrell et al. 1996, Hastie et al. 2005). Although most modellers consider that external validation is preferable to internal, there are cases where internal validation (i.e., the model ability to fit the training data) is sufficient. If the goal is to describe a pattern, overestimating the accuracy is not a problem. This is the case for models seeking to convert the observed records of a species into a suitability score (Araujo and Guisan 2006), as in our study. Overall, when making predictions based on occurrence data, presence data are more reliable than absence data (Dicko et al. 2014). Absence points may represent areas where the trap failed to catch a mosquito despite these mosquitoes being present in the area, or areas that are in principle suitable, but which have not yet been invaded. Therefore, in these cases, it is recommended to prioritize the sensitivity over the specificity (De Clercq et al. 2015).

Sensitivity and specificity measures were used to compare techniques for the same species and in the same geographical area. It was not possible to compare model performance between species because the traditional methods are highly influenced by the relative areas of occurrence of different species and by the geographical extent: increasing the geographical extent outside the presence environmental domain leads to a larger score for the area under the curve (Lobo et al. 2008). In fact, it has been shown that the relative occurrence area of the species influences the results of the evaluation scores, implying that models of rare species with high environmental specificity will yield to higher discrimination values (McPherson and Jetz 2007, Jiménez-Valverde et al. 2008) and that species with restricted environmental tolerance and/or distributions are usually reported to be well predicted (Tsoar et al. 2007). This is indeed what we would observe if we

compared sensitivity and specificity between species: *Oc. punctor* is mainly observed in the east of the country and the random forest has the highest discrimination skills, compared to other species.

To create a reliable model, it is generally considered necessary to have the same number of presence and absence points as input. This is because having a different number will create a bias in the model prediction towards the more prevalent category (presence or absence) (Fielding and Bell 1997). For NLDA and GLM, this balancing, i.e., considering the same number of presences and absences, is accomplished at the bootstrapping stage. However, for techniques such as random forest it is necessary to select a 'balanced' subset of the data.

Some areas were excluded from the sampling scheme, because they were deemed unsuitable for mosquitoes. At the modelling stage, this had to be corrected by adding absence points in these un-sampled regions (for details, see methods section). In our study, this adjustment was possible because it was known that the un-sampled regions would be negative for presence of mosquitoes. Generally speaking, avoiding bias in sampling strategies is more advisable than correcting for lack of data at the modelling stage.

The aim of the study was to investigate the spatial distribution of mosquito species and to compare the performance of three statistical models. In recent years, predictive modelling of species distribution has become an increasingly important tool to address various issues in ecology, biogeography, evolution, conservation biology and climate change research (Guisan and Thuiller 2005). Beyond describing species distributions, these models have become an important and widely used decision making tool for a variety of biogeographical applications, such as mapping risk of vector-borne disease spread, and determining locations that are potentially susceptible to invasion (Miller 2010). Species distribution modelling using flexible machine learning approaches has been successfully applied to quantify and to map the global distribution of hosts (Alexander et al. 2014), disease vectors (Sinka et al. 2012), pathogens (Bhatt et al. 2013), and infection and outbreak risk (Fischer et al. 2013).

CONCLUSIONS

In this study we mapped the environmental suitability for three mosquito species with three different modelling techniques. For each species, the models produced consistent spatial patterns, but different levels of prediction ranges. The average predicted environmental suitability was lower for NLDA than for the other methods. The variables selected as important in the models were consistent with field experience and the existing knowledge about these species. All the modelling techniques showed a satisfactory to excellent accuracy; the best accuracy was obtained with the random forest model.

The insights obtained in this study can be used to improve future predictions for vector and non-vector species. The output of this type of distribution modelling methods can be used as

input for epidemiological models and can be helpful to identify suitable areas for a given species, at risk of successful invasion if the species is still absent. Such areas may therefore need particular attention in terms of measures of prevention.

METHODS

The input for spatial distribution modelling consists of mosquito field data and environmental, often satellite, data. Here we describe the mosquito data collection, the environmental data used and the statistical methods applied.

Mosquito data

Data were collected by the Dutch Centre for Monitoring of Vectors during the National Mosquito Survey program, from April to October 2010–2013 (Ibañez-Justicia et al. 2015). These consisted of mosquito abundance data, sampled at 778 locations. Each of the locations was sampled only once and each trap was active for one week. At the sampled locations, mosquitoes were captured by means of CO₂-baited Mosquito Magnet Liberty Plus MM3100 traps (Woodstream® Co., Litzitz, USA). These traps have been evaluated successfully for trapping and surveillance against a variety of mosquito genus and species (Dennett et al. 2004, McKenzie and Bedard 2004), and have been used successfully in the national inventory of mosquitoes in Belgium, MODIRISK (Versteirt et al. 2013), also to capture *Culiseta annulata*, *Anopheles claviger* and *Ochlerotatus punctor*. For our survey, the traps were randomly located in the Netherlands, following the study design described in Ibañez-Justicia et al. (2015). Of the traps, 40% were placed in urban areas, 40% in agricultural areas and 20% in natural areas. Natural areas were sampled to a lesser extent because of their presumed lower involvement in human and veterinary health risks.

As described in Ibañez-Justicia et al. (2015), high productivity agricultural areas, such as arable land or permanent crops, were not sampled as they are considered to be unsuitable due to a lack of mosquito breeding sites. Therefore, areas with beet, grain, maize, potatoes and other agricultural crops, bulb flowers, productivity orchards and greenhouses were excluded from the sampling. When the goal is to estimate the potential distribution of a species, it is important that absence data come from environmental conditions that are known to be unsuitable for the species (Chefaoui and Lobo 2008). If information on absence is not available, absences can be generated outside the environmental domain where the species is present (Jiménez-Valverde et al. 2008). This has been done, for example, in Jiménez-Valverde and Lobo (Jiménez-Valverde and Lobo 2007), where probable absences were randomly selected in the areas having environmental values outside the range of observed presences. In a similar way, forty-three absence points in our study were generated in the land cover types that are deemed unsuitable for mosquitoes. Omitting this information, would have led to unrealistic predictions, as is for example shown in Figure 2. In Figure 2a, the presence and absence points for *An. claviger* are shown as black and

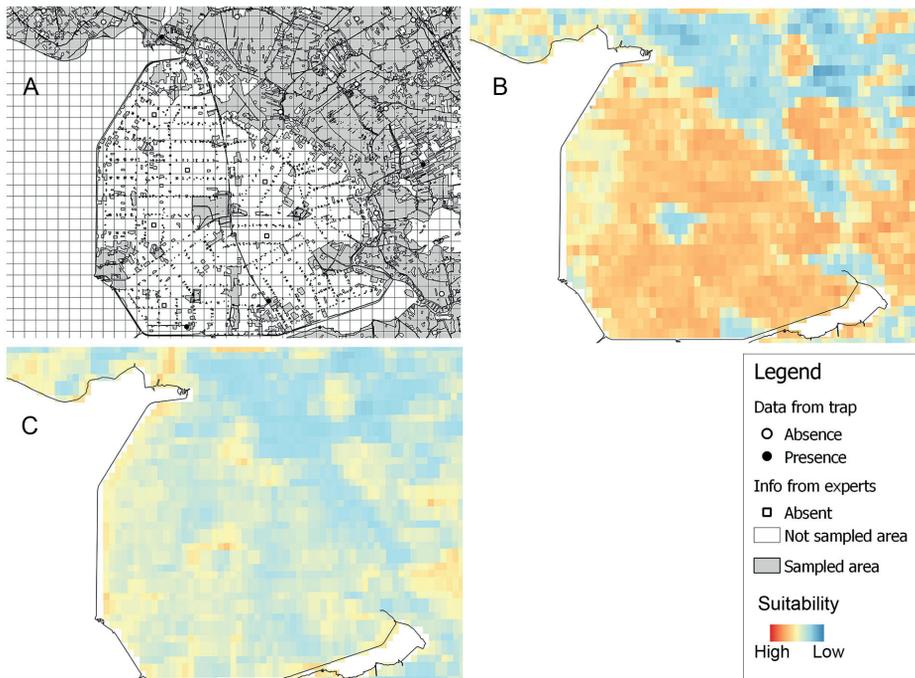


Figure 2. Probable absences added in unsuitable un-sampled areas. **A** – Absence points added to the *An. claviger* data in part of Flevoland province. The grey area was sampled and the white area was excluded from the sampling because it was considered unsuitable for mosquitoes. White and black circles indicate negative and positive traps, respectively. White squares indicate the probable absences added in unsuitable un-sampled areas. **B** – Random forest predictions for *An. claviger* without pseudo-absences. Environmental suitability is depicted using a gradient fill: blue indicates low environmental suitability, red indicates high suitability. **C** – Random forest predictions for *An. claviger* with pseudo-absences.

white circles in an area of Flevoland province (north-west). The grey area was sampled and the white area was excluded from the sampling because it was considered unsuitable for mosquitoes. Only two traps were located in this part of the region, in two fragments that were part of the sampled area, and they were both positive for the presence of mosquitoes. Both positive traps were located in small areas enclaved in pixels that were mainly considered to be unsuitable (intensively used agricultural fields). However, these possibly unsuitable areas were not sampled and therefore there are no data informing the model of their unsuitability. In a model without probable absence added, the two pixels are identified as suitable and consequently the whole area will be predicted as being suitable (Figure 2b), potentially incorrectly. If we introduce also absence points, the model is provided with more complete information and gives more realistic predictions (Figure 2c).

The abundance data were reclassified into data of presence (when at least one mosquito was found in the trap) and absence (when no mosquitoes were found in the trap), because the number of mosquitoes in each location was measured in different weeks and the mosquito abundance is expected to vary seasonally. The resolution used for the maps is 1 km². When a presence and an absence point were in the same square kilometre only the presence point was selected because presences inform about the places that are environmentally suitable for a species, but absences do not necessarily indicate the opposite (Lobo et al. 2010).

Environmental variables

The environmental data included in the analysis as predictor variables are 1 km² resolution satellite images and meteorological data in raster file format, commonly used for mosquito distribution modelling (ECDC 2013). The images were obtained from the MODIS sensor on NASA's Terra and Aqua satellites (Justice et al. 2000, modis.gsfc.nasa.gov) for 2000–2012 and subjected to temporal Fourier transformation (Rogers and Williams 1994, Rogers 2000) to summarise the images and to produce sets of data that capture characteristics of the annual seasonality: the mean, the annual, bi-annual and tri-annual amplitudes and phases, the maxima, minima and variances of variances of the middle infra-red (MIR), the daytime land surface temperature (dLST), the night-time land surface temperature (nLST), the enhanced vegetation index (EVI) and the normalized difference vegetation index (NDVI) signals (Scharlemann et al. 2008). Other environmental data used in this study are precipitation (WorldClim [www.worldclim.org]) and CMORPH [cpc.ncep.

Table 4 Fourier components from temporal Fourier analysis of an imagery time series

Component	Description
A0	Fourier mean for entire time series
MN	Minimum value
MX	Maximum value
A1	Amplitude of annual cycle
A2	Amplitude of bi-annual cycle
A3	Amplitude of tri-annual cycle
VR	Total variance
P1	Phase of annual cycle
P2	Phase of bi-annual cycle
P3	Phase of tri-annual cycle
D1	Proportion of total variance due to annual cycle
D2	Proportion of total variance due to bi-annual cycle
D3	Proportion of total variance due to tri-annual and cycle
DA	Proportion of total variance due to all three cycles

Component is the name used in the software Vecmap.

Table 5 Environmental predictor variables

Source	Variable
MODIS	Middle infra-red (MIR)
MODIS	Day-time land surface temperature (DLST)
MODIS	Night-time land surface temperature (NLST)
MODIS	Enhanced vegetation index (EVI)
MODIS	Normalised difference vegetation index (NDVI)
CMORPH	Precipitation
WorldClim	Precipitation
MODIS	Digital elevation model (DEM)
Gridded Population of the World	Human population density
European Environment Agency	Corine land cover

noaa.gov] 1950–2000), population density (compiled from the Gridded Population of the World Dataset 2000 [sedac.ciesin.columbia.edu]), the digital elevation model (MODIS 2012) and land cover (Corine land cover map of 2006 [www.eea.europa.eu]). A list of the Fourier components is provided in Table 4 and the environmental data are listed in Table 5. Predictor variables were organized as raster type files and for each trap location the pixel values of the environmental variables were extracted.

Statistical analysis

Species distribution models quantitatively describe areas that support the presence of a given species, based on known occurrence data and the associated environmental conditions (Elith and Leathwick 2009). Here, three methods suitable for occurrence data have been applied, i.e., non-linear discriminant analysis, random forest analysis and a generalised linear model, aimed at describing the relationship between response and predictor variables. For all three modelling techniques, the output was an environmental suitability indicator for each species, expressed as a value between 0 (low suitability) and 1 (high suitability). The predicted environmental suitability is visualised in maps.

Non-linear discriminant analysis

Models created using NLDA (Rogers 2000) require presence and absence data to be grouped into clusters based on attribute data. In this way, a discriminant function can be created and predictive maps based on these clusters can be made. The main advantage of the clustering is that it handles spatial heterogeneity of habitat niches and zones. The data were clustered using the k-means clustering algorithm. Since the important variables for the species were unknown, generic variables were used for clustering. These variables were DEM and the means, amplitudes, maxima, minima, variances (of the entire signal) and phases of MIR, LST and NDVI. NLDA models

were bootstrapped (Efron and Tibshirani 1993), meaning that 100 models were run and that for each model a sample of an equal number of presence and absence points was taken from the training set with replacement. The final predictions are based on the average of the 100 models.

Random forest

A random forest (Breiman 2001) method consists of an ensemble of classification (and regression) trees (CART; Breiman et al. 1984) constructed using a random subset of both the available samples and the attributes recorded for each data point. A CART tree is a hierarchical structure that allows a data point to be assigned to a particular class based on its attribute values. For the random forest method it is necessary to have the same number of presence and absence points as input, in order to obtain unbiased model predictions (McPherson et al. 2004). For techniques such as NLDA and GLM, this balancing is guaranteed at the bootstrapping stage, but this is not the case for RF. Therefore, before the model was run, five 'balanced' subsets of the complete dataset were randomly created. For each species, if there were more absence than presence points, all the presence points were used and a random subset of the absences was selected. If the presence points outnumbered the absence points, the procedure was inverted. The RF models were not bootstrapped because inherently RF uses a rationale similar to the bootstrapping approach, being based on several CART trees. The final predictions are the average of the five sub-models.

Logistic regression

For the GLM analysis, a logistic regression model was used because the response is a binary variable (presence/absence). GLM models can account for spatial autocorrelation by using an autoregressive term or mixture model. The effect of spatial autocorrelation on the training set was checked in the correlograms, i.e., plots of distance between points and the Moran's I index of their correlation. Since the correlation effect was not strong (with the exception at extreme distances where it is known that values for Moran's I may be erratic due to fewer points that can be compared (Jong et al. 1984) there was no need to account for spatial autocorrelation. As for NLDA, the GLM models were bootstrapped 100 times with a sample of an equal number of presence and absence points after which the 100 models were averaged to produce the final predictions.

Model evaluation

The choice of evaluation strategy needs to be explicitly related to the subject and goals of modelling. Here, the aim is to describe a given pattern and get a suitability score. In this context, simple forms of verifications, e.g. the number of false negatives, is appropriate to check whether models are performing as intended (Araujo and Guisan 2006). For each model, sensitivity and specificity were calculated, where sensitivity is the ability of a model to correctly identify known positive sites and specificity is the ability of a model to correctly identify known negative sites. Sensitivity and specificity were reported together with the values of their confidence intervals. The confidence

intervals were calculated based on 2000 stratified bootstrap replicates at 95% level. Sensitivity and specificity were used to compare the results produced with NLDA, RF and GLM for the same species. A list of the most important variables used in the models is provided. For NLDA and GLM the top 10 ranked variables are listed and for RF variable importance is given as mean decrease in Gini index (Breiman et al. 1984, Breiman 2001).

The analysis has been performed with the software Vecmap demo version (Avia-GIS 2014). The accuracy measures have been calculated with R 3.0.2 statistical language environment (R Development Core Team 2013), using of the R-packages pROC, ROCR, OptimalCutpoints.

ABBREVIATION

NLDA, Non-linear discriminant analysis; RF, Random forest; GLM, Generalised linear model; *Cs.*, *Culiseta*; *An.*, *Anopheles*; *Oc.*, *Ochlerotatus*; MIR, Middle infra-red; dLST, Daytime Land Surface Temperature; nLST, Night-time Land Surface Temperature; EVI, Enhanced Vegetation Index; NDVI, Normalized Difference Vegetation Index; DEM, Digital Elevation Model.

ACKNOWLEDGEMENTS

The authors would like to thank David Morley, Els Ducheyne and Veerle Versteirt for their help with Vecmap™ and Hans Heesterbeek and the FAH writing club for their helpful comments on the manuscript. The authors would also like to thank their recently deceased beloved colleague Dr. Ernst-Jan Scholte for promoting the idea of modelling the CMV mosquito surveillance data in the Netherlands. This study was funded by EU grant FP7-261504 EDENext and is catalogued by the EDENext Steering Committee as EDENext243 (<http://www.edenext.eu>). The contents of this publication are the sole responsibility of the authors and do not necessarily reflect the views of the European Commission.

REFERENCES

- Alexander, N. S., D. Morley, J. Medlock, K. Searle, and W. Wint. 2014. A First Attempt at Modelling Roe Deer (*Capreolus capreolus*) Distributions Over Europe. *Open Heal. Data*. 2: e2.
- Araujo, M. B., and A. Guisan. 2006. Five (or so) challenges for species distribution modelling. *J. Biogeogr.* 33: 1677–1688.
- Avia-GIS. 2014. Vecmap™: A one-stop-shop for vector mapping, Zoersel, Belgium. <http://www.avia-gis.com/vecmap>.
- Becker, N., D. Petric, M. Zgomba, C. Boase, M. Madon, C. Dahl, et al. 2010. *Mosquitoes and Their Control*. Springer, Heidelberg, Dordrecht, New York.
- Bhatt, S., P. W. Gething, O. J. Brady, J. P. Messina, A. W. Farlow, C. L. Moyes, et al. 2013. The global distribution and burden of dengue. *Nature*. 496: 504–507.
- Breiman, L. 2001. Random forests. *Mach. Learn.* 45: 5–32.
- Breiman, L., J. H. Friedman, R. A. Olshen, and C. J. Stone. 1984. *Classification and regression trees*, Nucl. Instrum. Meth. A. Wadsworth international group, Belmont, California, USA.
- Chefaoui, R. M., and J. M. Lobo. 2008. Assessing the effects of pseudo-absences on predictive distribution model performance. *Ecol. Modell.* 210: 478–486.
- De Clercq, E. M., S. Leta, A. Estrada-Peña, M. Madder, S. Adehan, and S. O. Vanwambeke. 2015. Species distribution modelling for *Rhipicephalus microplus* (Acari: Ixodidae) in Benin, West Africa: Comparing datasets and modelling algorithms. *Prev. Vet. Med.* 118: 8–21.
- Cutler, D. R., T. C. Edwards Jr, K. H. Beard, A. Cutler, K. T. Hess, J. Gibson, et al. 2007. Random forests for classification in ecology. *Ecology*. 88: 2783–2792.
- Dennett, J. A., N. Y. Vessey, and R. E. Parsons. 2004. A comparison of seven traps used for collection of *Aedes albopictus* and *Aedes aegypti* originating from a large tire repository in Harris County (Houston), Texas. *J. Am. Mosq. Control Assoc.* 20: 342–349.
- Dicko, A. H., R. Lancelot, M. T. Seck, L. Guerrini, B. Sall, M. Lo, et al. 2014. Using species distribution models to optimize vector control in the framework of the tsetse eradication campaign in Senegal. *Proc. Natl. Acad. Sci.* 111: 10149–10154.
- European Centre for Disease Prevention and Control (ECDC). 2013. Environmental risk mapping: *Aedes albopictus* in Europe. Stockholm; ECDC. Available from: <http://www.ecdc.europa.eu/en/publications/Publications/climate-change-environmental-risk-mapping-aedes.pdf>
- Efron, B., and R. J. Tibshirani. 1993. *An Introduction to the Bootstrap*. Chapman and Hall, London, UK.
- Elith, J., C. H. Graham, R. P. Anderson, M. Dudík, S. Ferrier, A. Guisan, et al. 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography (Cop.)*. 29: 129–151.
- Elith, J., and J. R. Leathwick. 2009. Species distribution models: ecological explanation and prediction across space and time. *Annu. Rev. Ecol. Evol. Syst.* 40: 677–697.
- European Environment Agency [www.eea.europa.eu].
- Fielding, A. H., and J. F. Bell. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environ. Conserv.* 24: 38–49.

- Fischer, E. A. J., G.-J. Boender, G. Nodelijk, A. A. de Koeijer, and H. J. Van Roermund. 2013. The transmission potential of Rift Valley fever virus among livestock in the Netherlands: a modelling study. *Vet. Res.* 44: 58.
- Guisan, A., and W. Thuiller. 2005. Predicting species distribution: offering more than simple habitat models. *Ecol. Lett.* 8: 993–1009.
- Harrell, F. E., K. L. Lee, and D. B. Mark. 1996. Tutorial in biostatistics multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat. Med.* 15: 361–387.
- Hastie, T., R. Tibshirani, J. Friedman, and J. Franklin. 2005. The elements of statistical learning: data mining, inference and prediction. *Math. Intell.* 27: 83–85.
- Ibañez-Justicia, A., A. Stroo, M. Dik, J. Beeuwkes, and E. J. Scholte. 2015. National Mosquito (Diptera, Culicidae) Survey in the Netherlands 2010–2013. *J. Med. Entomol.* DOI: 10.1093/jme/tju058
- Jiménez-Valverde, A., and J. M. Lobo. 2007. Potential distribution of the endangered spider *Macrothele calpeiana* (Walckenaer, 1805)(Araneae, Hexathelidae) and the impact of climate warming. *Acta Zool. Sin.* 53: 865–876.
- Jiménez-Valverde, A., J. M. Lobo, and J. Hortal. 2008. Not as good as they seem: the importance of concepts in species distribution modelling. *Divers. Distrib.* 14: 885–890.
- Jong, P. de, C. Sprenger, and F. Veen. 1984. On extreme values of Moran's I and Geary's c. *Geogr. Anal.* 16: 17–24.
- Justice, C., J. Townshend, E. Vermote, R. Sohlberg, J. Descloitres, D. Roy, et al. 2000. Preliminary land surface products from the NASA Moderate Resolution Imaging Spectroradiometer (MODIS), pp. 1157–1162. In *Geosci. Remote Sens. Symp. 2000. Proceedings. IGARSS 2000. IEEE 2000 Int.*
- Lobo, J. M., A. Jiménez-Valverde, and J. Hortal. 2010. The uncertain nature of absences and their importance in species distribution modelling. *Ecography (Cop.)*. 33: 103–114.
- Lobo, J. M., A. Jiménez-Valverde, and R. Real. 2008. AUC: a misleading measure of the performance of predictive distribution models. *Glob. Ecol. Biogeogr.* 17: 145–151.
- McPherson, J., and W. Jetz. 2007. Effects of species' ecology on the accuracy of distribution models. *Ecography (Cop.)*. 30: 135–151.
- McKenzie, K. E., and S. D. Bedard. 2004. A Comparison of the Efficacy of Several Models of Mosquito Magnet Traps Baited with Lurex 3 versus Commercially Available Traps Supplied with Octenol on the Collection of Mosquitoes in Oahu, Hawaii, Primarily *Aedes albopictus*.
- McPherson, J., W. Jetz, and D. J. Rogers. 2004. The effects of species' range sizes on the accuracy of distribution models: ecological phenomenon or statistical artefact? *J. Appl. Ecol.* 41: 811–823.
- Miller, J. 2010. Species distribution modeling. *Geogr. Compass.* 4: 490–509.
- MODIS [modis.gsfc.nasa.gov].
- National Weather Service Climate Prediction Center [www.cpc.ncep.noaa.gov].
- Peters, J., B. De Baets, N. E. C. Verhoest, R. Samson, S. Degroeve, P. De Becker, et al. 2007. Random forests as a tool for ecohydrological distribution modelling. *Ecol. Modell.* 207: 304–318.
- R Development Core Team. 2013. R: A Language and Environment for Statistical Computing.
- Rogers, D. J. 2000. Satellites, space, time and the African trypanosomiasis. *Adv. Parasitol.* 47: 129–171.
- Rogers, D. J. 2006. Models for vectors and vector-borne diseases. *Adv. Parasitol.* 62: 1–35.

- Rogers, D. J., and B. G. Williams. 1994. Tsetse distribution in Africa: seeing the wood and the trees.
- Scharlemann, J. P. W., D. Benz, S. I. Hay, B. V Purse, A. J. Tatem, G. R. W. Wint, et al. 2008. Global data for ecology and epidemiology: a novel algorithm for temporal Fourier processing MODIS data. *PLoS One*. 3: e1408.
- Sinka, M. E., M. J. Bangs, S. Manguin, Y. Rubio-Palis, T. Chareonviriyaphap, M. Coetzee, et al. 2012. A global map of dominant malaria vectors. *Parasit Vectors*. 5: 69.
- Socioeconomic Data and Applications Center [sedac.ciesin.columbia.edu].
- Tsoar, A., O. Allouche, O. Steinitz, D. Rotem, and R. Kadmon. 2007. A comparative evaluation of presence-only methods for modelling species distribution. *Divers. Distrib.* 13: 397–405.
- Versteirt, V., S. Boyer, D. Damiens, E. M. De Clercq, W. Dekoninck, E. Ducheyne, et al. 2013. Nationwide inventory of mosquito biodiversity (Diptera: Culicidae) in Belgium, Europe. *Bull. Entomol. Res.* 103: 193–203.
- Versteirt, V., E. De Clercq, W. Dekoninck, D. Damiens, A. Ayrinhac, F. Jacobs, and W. Van Bortel. 2009. Mosquito vectors of disease: spatial biodiversity, drivers of change, and risk. *Final Report*. Belgian Science Policy: Brussels.
- WorldClim - Global Climate Data [www.worldclim.org].

CHAPTER 3

Modelling the spatial distribution of the nuisance mosquito species *Anopheles plumbeus* (Diptera: Culicidae) in the Netherlands

Adolfo Ibañez-Justicia^{1*} and Daniela Cianci^{2*}

¹ Centre for Monitoring of Vectors, Food and Consumer Product Safety Authority, Wageningen, The Netherlands

² Faculty of Veterinary Medicine, Utrecht University, Utrecht, The Netherlands

* Equal contributors

Parasites & Vectors, 8(1):258;
DOI: 10.1186/s13071-015-0865-7

ABSTRACT

Landscape modifications, urbanization or changes of use of rural-agricultural areas can create more favourable conditions for certain mosquito species and therefore indirectly cause nuisance problems for humans. This could potentially result in mosquito-borne disease outbreaks when the nuisance is caused by mosquito species that can transmit pathogens. *Anopheles plumbeus* is a nuisance mosquito species and a potential malaria vector. It is one of the most frequently observed species in the Netherlands. Information on the distribution of this species is essential for risk assessments. The purpose of the study was to investigate the potential spatial distribution of *An. plumbeus* in the Netherlands. Random forest models were used to link the occurrence and the abundance of *An. plumbeus* with environmental features and to produce distribution maps in the Netherlands. Mosquito data were collected using a cross-sectional study design in the Netherlands, from April to October 2010–2013. The environmental data were obtained from satellite imagery and weather stations. Statistical measures (accuracy for the occurrence model and mean squared error for the abundance model) were used to evaluate the models performance. The models were externally validated. The maps show that forested areas (centre of the Netherlands) and the east of the country were predicted as suitable for *An. plumbeus*. In particular high suitability and high abundance was predicted in the south-eastern provinces Limburg and North Brabant. Elevation, precipitation, day and night temperature and vegetation indices were important predictors for calculating the probability of occurrence for *An. plumbeus*. The probability of occurrence, vegetation indices and precipitation were important for predicting its abundance. The AUC value was 0.73 and the error in the validation was 0.29; the mean squared error value was 0.12. The areas identified by the model as suitable and with high abundance of *An. plumbeus*, are consistent with the areas from which nuisance was reported. Our results can be helpful in the assessment of vector-borne disease risk.

Keywords: species distribution modelling, *Anopheles plumbeus*, mosquito nuisance, vector-borne diseases, random forest

BACKGROUND

Mosquitoes (Diptera:Culicidae) are known to be vectors of a large number of pathogens around the globe and are considered as prime candidates for transmitting (re-)emerging vector-borne diseases (VBDs) in Europe (Versteirt et al. 2009). The increased mobility of humans, that has also increased the mobility of livestock and pathogens, as well as environmental modifications and climate changes can contribute to the (re-)emergence of vector-borne diseases (Randolph and Rogers 2010). Furthermore, mosquito bites can cause a considerable nuisance for humans and mammals. Severe nuisance can have negative economic consequences (e.g., in tourism, work productivity outdoors, meat and dairy production) (Connelly et al. 2009). These nuisance situations can eventually lead to autochthonous VBD cases, when in non-endemic areas infectious reservoirs, either humans (travellers, temporary workers) or animals (livestock, migrating animals) come in contact with high density of mosquito vectors.

In 2010, in the Netherlands, the Centre for Monitoring of Vectors (CMV) started a nationwide inventory of indigenous mosquitoes to acquire basic information on the composition, geographical distribution, biodiversity and environmental preferences of mosquito species. In this survey, the nuisance mosquito species and potential malaria vector *Anopheles plumbeus* (Stephens, 1828) (Figure 1) was one of the most frequently collected mosquito species (Ibañez-Justicia et al. 2015). This species has been proven to be able to transmit *Plasmodium falciparum* (Welch, 1897), the causative agent of malaria tropica (Schaffner et al. 2012). Circumstantial evidence for local



Figure 1. *Anopheles plumbeus* female (source: A. Ibañez-Justicia).

transmission of *P. falciparum* malaria by *An. plumbeus* has been reported for Germany (Krüger et al. 2001). *Anopheles plumbeus* has also been incriminated as a vector of *Plasmodium vivax* (Grassi & Feletti, 1890) (Shute and Maryon 1974, Sokolova and Snow KR. 2002) and has been proven to be a laboratory vector of West Nile virus (Schaffner et al. 2001). Even though the health care system is likely to rapidly identify malaria patients and thereby prevent the building up of an infectious human reservoir of *Plasmodium* parasites (Takken et al. 2007), it is important to gain information on the spatial distribution of *An. plumbeus*, in order to inform the health care system on the areas at risk.

Anopheles plumbeus is a mosquito species commonly found in forests, where larvae are usually found in water in rot-holes of trees with high salinity and deficiency of oxygen (Bradshaw and Holzapfel 1991). They can also be found in containers with stagnant rain water and groundwater, such as tyres, rainwater casks and cemetery vases (Schaffner et al. 2012, Marshall 1938). In the last decade, this species has also been associated with abandoned stables where it breeds in the rain water collected in the manure cellars (Dekoninck et al. 2011). This species is known to be a particularly aggressive biter, feeding at any time of the day on different mammalian hosts (including humans), and to a lesser degree on birds and reptiles (Bueno Marí and Jiménez Peydró 2012). In June 2006, nuisance caused by *An. plumbeus* was reported for the first time in the Netherlands, near the city of Nijmegen (Versteirt et al. 2009). Since then, *An. plumbeus* nuisance have been reported every year in the Netherlands, mostly in proximity to abandoned pig stables (Ibanez-Justicia, unpublished).

An understanding of the spatial extent of potential vector species, their abundance and seasonal activity, is important for estimating levels of risks of VBDs and enabling better targeting for surveillance and control. In order to develop basic reproduction number (R_0) models and construct risk maps that indicate the risk for an outbreak after an introduction, abundance data of vectors are an essential parameter (Hartemink et al. 2009, 2011). Although vector presence and abundance are not the only factors determining whether or not a pathogen can spread in an area, determining the distribution of the vector is an essential step in studying the risk of transmission of a pathogen. Given the nuisance and potential risk for the human health, such information on *An. plumbeus* is needed. Currently, no information on the potential spatial distribution of this species is available for the Netherlands.

In this study, we modelled the potential spatial distribution, expressed in occurrence (predicted probability of presence) and abundance of *An. plumbeus* in the Netherlands, based on data collected during the National Mosquito Survey and environmental data. The occurrence was modelled to predict the environmental suitability of the species using a random classification forest model. The abundance was modelled using a random regression forest model with the aim to identify areas where mosquito peaks could be expected. Random (classification and regression) forest models allow external validation through a bootstrapping procedure. The occurrence model was validated also with an external dataset. The resulting maps are in agreement with the

reported nuisance for this species and the predictions show a good matching with an external dataset used to validate the model.

METHODS

Species distribution modelling links the occurrence or the abundance of species with environmental data and estimates the similarity of the conditions at any site based on the conditions at the locations of known occurrence/abundance of a species. Here we describe the mosquito data collection, the environmental data used and the statistical methods applied in this study.

Mosquito data

Mosquito data used for the modelling were obtained from the national mosquitoes survey that was carried out from April to October 2010–2013 by the Dutch Centre for Monitoring of Vectors. Mosquitoes were captured using CO₂ baited Mosquito Magnet Liberty Plus MM3100 (Woodstream® Co., Lititz, USA). Traps were randomly distributed in the country following a cross-sectional study design, with the following constraint: 40% of the traps were placed in urban areas, 40% in rural-agricultural areas and 20% in natural areas (Ibañez-Justicia et al. 2015). Urban and agricultural areas were sampled more intensively, because of the potential higher human and veterinary health risk in those areas due to higher exposure.

Data consisted of mosquito abundance data, sampled at 778 locations. For this study the abundance data were also reclassified into data of presence (when at least one mosquito was found in the trap) and absence (when no mosquitoes were found in the trap). Each of the locations was sampled only once and each trap was active for one week. The content of the traps was collected weekly and sent to the CMV laboratory. In the laboratory, mosquitoes were morphologically identified using, among others, the Culicidae key specifically designed for rapid field-identification of Dutch adult Culicidae (modified key after Snow 1990, Schaffner et al. 2001, Verdonshot 2001, Becker et al. 2010). Twenty-seven mosquito species were found in the National Mosquito Survey and *An. plumbeus* was the 7th mosquito species most commonly found in the Netherlands. This species was active in the whole period of the survey, from April until October (Ibañez-Justicia et al. 2015). When a presence and an absence point were in the same square kilometre only the presence point was used because presences inform about the places that are environmental suitable for a species, but absences do not necessarily indicate the opposite (Lobo et al. 2010).

For the validation of the occurrence model, data on *An. plumbeus* presence from confirmed nuisance notifications and data from other mosquito surveys carried out by the CMV in the Netherlands during the years 2010–2014 were used (Table 1). The mosquito data from these surveys were collected with various trapping methods: dippers, pooters, CDC miniature light traps Model 512 (John W. Hock Company, Gainesville, USA), BG Sentinel traps (Biogents AG, Regensburg,

Table 1 Surveys used for the validation

Survey name	Year	Sampling strategy	Capturing device	Total nr locations
EMS-Used tires	2010-2014	Target longitudinal sampling	Larval sampling, manual aspirator, BG-Sentinel trap	16
EMS-Lelystad	2013	Target sampling	MM-Liberty Plus trap	3
NVS-Limburg	2009	Cross-sectional	MM-Liberty Plus trap	14
NVS-Mosquitoes longitudinal	2011	Target longitudinal sampling	MM-Liberty Plus trap	1
Projects	2011, 2012	Target longitudinal sampling	CDC light trap, manual aspirator	3
Nuisance	2010, 2011, 2013, 2014	Check at locations of reported nuisance	Larval sampling, manual aspirator	6
West-Nile-Virus Wetlands	2010	Target longitudinal sampling	MM-Liberty Plus trap, CDC light trap	2

The predictions obtained with the occurrence model that used National Mosquito Survey data were validated with data from these surveys.

EMS: Exotic Mosquito Survey.

NVS: National Vector Survey.

Germany) and Mosquito Magnet traps. These data were extracted from VecBase, a tailor-made application built for CMV in 2010 for vector surveillance data.

Environmental variables

The environmental data included in the analysis as predictor variables are 1 km² resolution satellite images and meteorological data in raster file format, commonly used for mosquito distribution modelling (ECDC 2013). The images were obtained from the MODIS sensor on NASA's Terra and Aqua satellites (Justice et al. 2000, modis.gsfc.nasa.gov) for 2000–2012 and subjected to temporal Fourier transformation (Rogers and Williams 1994, Rogers 2000) to summarise the images and to produce sets of data that capture characteristics of the annual seasonality: the mean, the annual, bi-annual and tri-annual amplitudes and phases, the maxima, minima and variances of the middle infra-red, day and night-time land surface temperature, the enhanced vegetation index and the normalized difference vegetation index signals (Scharlemann et al. 2008). Other environmental data used in this study are precipitation (WorldClim [www.worldclim.org] and CMORPH [www.cpc.ncep.noaa.gov] 1950–2000), population density (compiled from the Gridded Population of the World Dataset [sedac.ciesin.columbia.edu] 2000), the digital elevation model (MODIS 2012) and land cover (Corine land cover map [www.eea.europa.eu] of 2006). The Fourier components used are provided in Table 2 and the environmental data in Table 3. For each trap location the pixel values of the environmental variables were extracted.

Table 2 Fourier components from temporal Fourier analysis of an imagery time series

Component	Description
A0	Fourier mean for entire time series
MN	Minimum value
MX	Maximum value
A1	Amplitude of annual cycle
A2	Amplitude of bi-annual cycle
A3	Amplitude of tri-annual cycle
VR	Total variance
P1	Phase of annual cycle
P2	Phase of bi-annual cycle
P3	Phase of tri-annual cycle
D1	Proportion of total variance due to annual cycle
D2	Proportion of total variance due to bi-annual cycle
D3	Proportion of total variance due to tri-annual and cycle
DA	Proportion of total variance due to all three cycles

Component is the name used in Vecmap.

Table 3 Environmental predictor variables

Source	Variables
MODIS	Middle Infra-red (MIR)
MODIS	Day-time land surface temperature (DLST)
MODIS	Night-time land surface temperature (NLST)
MODIS	Enhanced vegetation index (EVI)
MODIS	Normalised difference vegetation index (NDVI)
CMORPH	Precipitation
WorldClim	Precipitation
MODIS	Digital elevation model (DEM)
Gridded Population of the World	Human population density
European Environment Agency	Corine land cover

Statistical analysis

Occurrence model

Three distribution modelling techniques suitable for occurrence data were applied: non-linear discriminant analysis (Rogers 2000), random classification forest (Breiman 2001) and generalised linear model (McCullagh and Nelder 1989). For each model, the accuracy was assessed using (i) sensitivity, i.e., the ability of a model to correctly identify known positive sites; (ii) specificity, i.e.,

the ability of a model to correctly identify known negative sites; (iii) the area under the curve (AUC) that can be roughly interpreted as the probability that a model will correctly predict positive and negative sites (Fielding and Bell 1997). Of the three techniques, random forest provided the best accuracy and therefore the results of this model are presented.

A random classification forest model consists of an ensemble of trees. To create a reliable model, it is generally considered necessary to have the same number of presence and absence points as input. This is because having a different number will create a bias in the model prediction towards the more prevalent category (presence or absence) (Fielding and Bell 1997). For this reason, a 'balanced' subset of the data, i.e., a dataset with the same number of presences and absences, was selected. The output produced by the model is an environmental suitability indicator, expressed as a value between 0 (low suitability) and 1 (high suitability). The predictions are visualised in a map with colours ranging from red (high suitability) to blue (low suitability). A list of the most important variables used in the model is given based on the mean decrease in Gini index (Breiman 2001, Breiman et al. 1984). Random forest allows external validation through a bootstrapping procedure: for each tree, a random subset of the full dataset is sampled with replacement. The model validation is carried out for each tree using the points not used from the full dataset. This validation method is referred to as external, because the model is validated using data that are not used to build the tree. The comparison of the observed and predicted results enables us to calculate accuracy statistics, such as sensitivity and specificity. These measures are calculated for each tree and then averaged to give the overall values.

The predictions produced by the random classification forest were also externally validated against 45 observations from other surveys (Table 1) that reported only the presence of *An. plumbeus*. Comparing the observations obtained with the other surveys and the predictions made by the model using National Mosquito Survey data, the error rate was calculated as the proportion of incorrectly predicted pixels to the total number of points used in the validation.

Abundance model

The abundance of the species was modelled using a random regression forest model. The abundance data were transformed according to the formula $\log_{10}(\text{abundance} + 1)$ (Avia-GIS 2012). Because the aim was to identify areas where mosquito peaks could be expected, only the data collected in months in which peaks were observed were selected (June-September). The predicted environmental suitability obtained with the occurrence model described above, was included as one of the predictor variables for modelling the abundance of the species, as it is frequently done in this type of analysis (Osborne et al. 2001, Boyce et al. 2002, Gibson et al. 2004, Chefaoui et al. 2005, Calveteet al. 2008). The predicted abundance is interpreted as the expected maximum number of mosquitoes caught in a trap in a certain pixel. The predictions are visualised in a map with colours ranging from light green (low abundance) to dark green (high abundance). The importance of the predictors was assessed using the increase in node purity (INP). The dif-

ference between observed and predicted values was expressed as the mean squared error. The analysis has been performed with the software Vecmap demo version (Avia-GIS 2014). The maps have been produced with Quantum GIS (QGIS Development Team 2012).

RESULTS AND DISCUSSION

The probability of occurrence (environmental suitability) and the abundance of *An. plumbeus* have been predicted using mosquito field data and environmental data. The estimated environmental suitability and abundance are shown in maps. The important environmental variables used in the models and the accuracy of the models are discussed. The fact that out of three different modelling techniques for occurrence data random forest model was selected based upon its higher classification accuracy is consistent with earlier findings; random forest has been reported to outperform other traditional modelling techniques (Cutler et al. 2007, Peters et al. 2007, Cianci et al. 2015).

An. plumbeus was found in 100 locations and it was observed in particular in the eastern part of the Netherlands (Figure 2a). The percentage of traps containing *An. plumbeus* per week is shown in Figure 3. For the modelling, 97 presence points and 97 absence points were selected. Using the

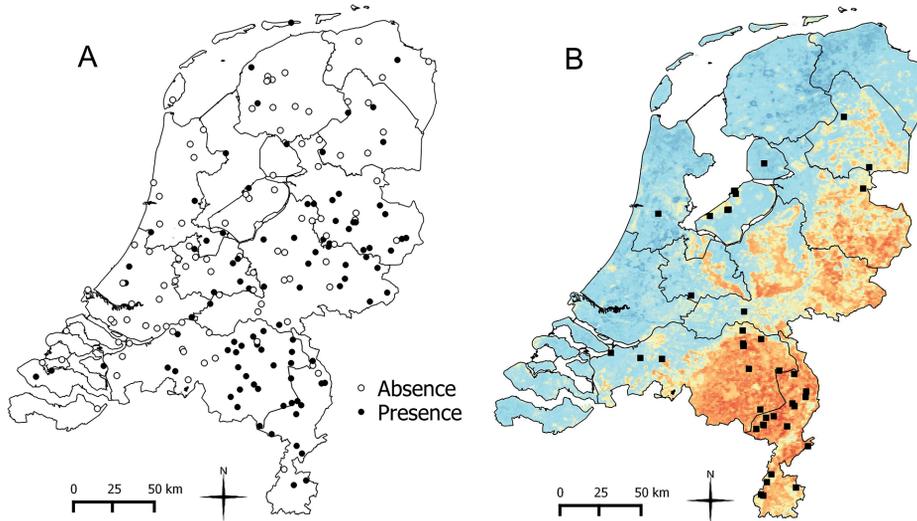


Figure 2. Observed presence and absence points and map of the estimated environmental suitability for *An. plumbeus*. **A-** Presence and absence observed during the National Mosquito Survey program carried out from April to October 2010–2013. **B-** Environmental suitability map of *An. plumbeus* produced using classification random forest. Environmental suitability is depicted using a gradient fill: blue indicates low environmental suitability, red indicates high suitability. Locations where other surveys took place are also shown on the map (black squares).

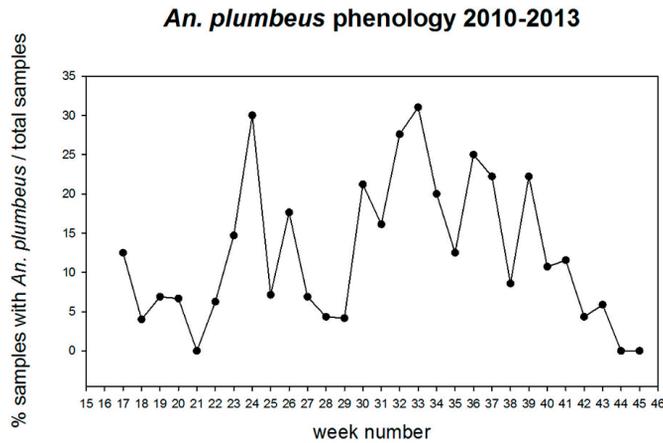


Figure 3. Percentage of positive sites of *An. plumbeus* per week in 2010–2013.

random forest model, forest-rich areas in the centre of the Netherlands (e.g. National Park Hoge Veluwe and National Park Utrechtse Heuvelrug) are predicted as suitable for *An. plumbeus* (Figure 2b). Also the eastern parts of the country and in particular the southeastern provinces (Limburg and North Brabant) are predicted to be suitable. In these two provinces nuisance is often reported, especially close to abandoned and un-cleaned pig stables, where mosquitoes breed in manure pools (Dekoninck et al. 2011). Based on the environmental characteristics included in the analysis, the model was capable of identifying areas where *An. plumbeus* is truly present, meaning that these characteristics can be a good proxy for abandoned and un-cleaned stables.

Elevation, precipitation, day and night temperature, vegetation indices and middle infra-red (index sensitive to changes in the vegetation) were found to be important predictors for environmental suitability for *An. plumbeus* (Table 4). Precipitation and vegetation are likely to be biologically relevant, since this species typically breeds in water-filled tree holes with high organic material content (Becker et al. 2010). Presence of tree-holes is related to the age of the tree or to the tree species. Eggs of this mosquito species are not laid on the water surface but on the sides of potential breeding sites, just above the waterline, so the number of generations produced each year are often dependent upon hydrological conditions (Becker et al. 2010). The occurrence model predicts environmental suitability for *An. plumbeus* in areas where old trees with tree-holes are known to occur (e.g., National Park Hoge Veluwe). Even though the species is considered to be a tree-hole breeding species, results obtained using the random forest occurrence model depict environmental suitability for this species in areas without forests in the Netherlands. The results indicate the potential successful use of unforested environments for *An. plumbeus* populations, and imply a similar trend to that seen in continental Europe and UK, where *An. plumbeus* is shifting habitats from almost exclusively breeding in tree-holes to exploiting a wider array of novel

Table 4 List of the top 10 most important variables in the occurrence model

Rank	Variables
1	DEM
2	CMORPH precipitation, phase of bi-annual cycle
3	CMORPH precipitation, phase of annual cycle
4	Worldclim precipitation, phase of annual cycle
5	Worldclim precipitation, proportion of total variance due to annual cycle
6	MIR, phase of annual cycle
7	NTLS temperature, minimum value
8	DTLS temperature, amplitude of annual cycle
9	NDVI mean
10	CMORPH precipitation, maximum value

The lowest ranking number indicates the most important variable (e.g., rank = 1 is the most important variable).

man-made larval breeding sites (Ibañez-Justicia et al. 2015, Townroe and Callaghan 2014). Day and night temperature were found to be related to environmental suitability also in another study in Belgium (Versteirt et al. 2009).

Fair accuracy was obtained with the model (AUC = 0.73), which showed a better ability in identifying suitable environments (sensitivity 0.71) than unsuitable environments (specificity 0.66). The accuracy is improved as compared with a first attempt of predicting the environmental suitability for *An. plumbeus* in the Netherlands, where the environmental suitability was extrapolated from Belgium to the Netherlands (sensitivity = 0.50, specificity = 0.49) (Versteirt et al. 2009). The error rate calculated to compare the predicted values to data of other surveys was low (0.29); 71% of the pixels were correctly predicted, meaning that the model could make good predictions in non-sampled areas. However, this is only a partial validation because it considers only presence points and does not give information about the performance of the model in predicting absence points.

The observed and estimated abundance are shown in Figure 4. In the summer, the observed field abundance was low; rarely more than 10 mosquitoes per trap were found (Table 5). The maximum value observed was 1701 mosquitoes followed by 62 mosquitoes per trap. Given the huge difference between the maximum value and the numbers of mosquitoes observed in the other traps, the maximum value was considered as outlier and excluded from the analysis. However, the reason of this high abundance was investigated and it turned out that the trap was located in a rural area where cattle farms with abandoned stables, not in use anymore, are present. In total, 505 mosquitoes were used in the analysis and they were captured in 80 traps/weeks from 2010 to 2013. The predicted abundance, produced with the random forest technique, was also low (with a maximum of 15 individuals per trap) and the highest abundance was predicted in the eastern part of the country and especially in the south-eastern provinces Limburg and Brabant. These findings are in agreement with the suitable areas identified by the occurrence model. This is not surprising

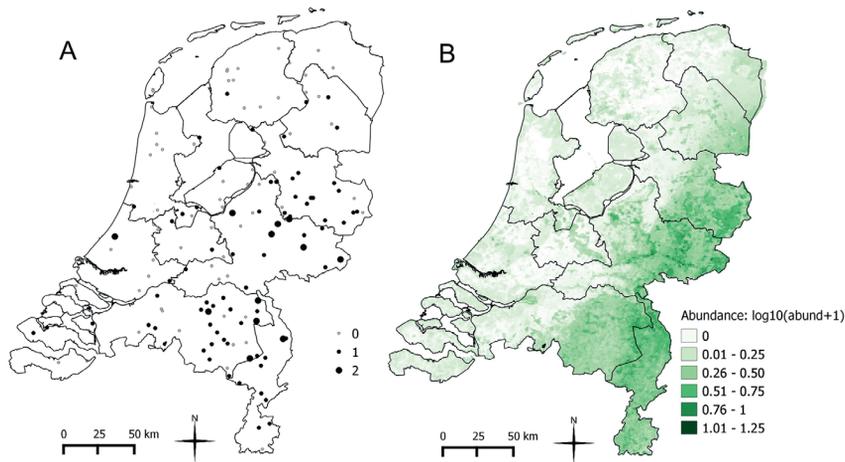


Figure 4. Observed and estimated abundance of *An. plumbeus*. **A** – Observed abundance represented as $\log_{10}(\text{abundance} + 1)$. **B** – Map of the estimated abundance produced using a regression random forest. A darker colour indicates higher abundance.

Table 5 Observed abundance used in the model

Count	Frequency
0	66
1-10	70
11-20	4
21-30	1
31-40	2
41-50	2
51-60	0
61-70	1

because the probability of occurrence was the most important factor among the environmental factors included in the abundance model (Table 6). Similar results, where the probability of occurrence appears to be the most important factor when using this technique, were already observed in a study conducted on other mosquito and biting midges species in the Netherlands (Avia-GIS 2012). Other important predictors for abundance of *An. plumbeus* were precipitation and vegetation, in accordance with the biology of the species, as it is explained above. The difference between the observed abundance and the predicted abundance was small, with a mean squared error value of 0.12, meaning that the model predictions matched the observation.

Table 6 List of the top 10 most important variables in the abundance model

Rank	Variables
1	Occurrence
2	Worldclim precipitation, phase of annual cycle
3	Worldclim precipitation, proportion of total variance due to bi-annual cycle
4	NDVI, amplitude of annual cycle
5	Worldclim precipitation, amplitude of bi-annual cycle
6	MIR, amplitude of annual cycle
7	DEM
8	NTLS temperature, phase of bi-annual cycle
9	Worldclim precipitation, total variance
10	CMORPH precipitation, phase of bi-annual cycle

The lowest ranking number indicates the most important variable (e.g., rank = 1 is the most important variable).

CONCLUSIONS

The aim of this study was to investigate the potential spatial distribution of *An. plumbeus* in the Netherlands. Using random (classification and regression) forest models, we identified areas with high environmental suitability and high abundance of this species in south-eastern provinces of Limburg and Brabant. These areas coincide with the areas where in recent years most nuisances have been reported. The predictions of the occurrence model were accurate and matched the external dataset used for validation. The abundance model predictions also matched the observations.

The output of species distribution modelling method can be used as an input for risk assessment of establishment and spread of vector-borne diseases (Miller 2010, Fischer et al. 2013). Understanding and depicting the potential spatial distribution of mosquito species with modelling techniques is of increasing importance, especially for nuisance mosquito species that can cause economic implications or impact on human health.

ACKNOWLEDGEMENTS

The authors would like to thank Nienke Hartemink and Marieta Braks for a critical reading of the manuscript and useful comments. The authors would also like to thank Els Ducheyne and Wesley Tack for helpful discussion and suggestions. This study was funded by EU grant FP7-261504 EDENext and is catalogued by the EDENext Steering Committee as EDENext311 (<http://www.edenext.eu>). The contents of this publication are the sole responsibility of the authors and do not necessarily reflect the views of the European Commission.

REFERENCES

- Avia-GIS. 2012. Abundance modeling of mosquito and biting midge species in the Netherlands.
- Avia-GIS. 2014. VecmapTM. A one-stop-shop for vector mapping, Zoersel, Belgium. http://www.avia-gis.com/vecmap_
- Becker, N., D., Petric D, Zgomba M, Boase C, Madon M, Dahl C, et al. 2010. Mosquitoes and Their Control. Heidelberg, Dordrecht, New York: Springer; p. 577.
- Boyce, M. S., P. R. Vernier, S. E. Nielsen, and F. K. A. Schmiegelow. 2002. Evaluating resource selection functions. *Ecol. Modell.* 157: 281–300.
- Bradshaw, W. E., and C. M. Holzapfel. 1991. Fitness and habitat segregation of British tree-hole mosquitoes. *Ecol. Entomol.* 16: 133–144.
- Breiman, L. 2001. Random forests. *Mach. Learn.* 45: 5–32.
- Breiman, L., J. H. Friedman, R. A. Olshen, and C. J. Stone. 1984. Classification and regression trees, Wadsworth international group, Belmont, California, USA.
- Bueno Marí, R., R. Jiménez Peydró, and others. 2011. Study of the malariogenic potential of Eastern Spain. *Trop Biomed.* 29:39-50.
- Calvete, C., R. Estrada, M. A. Miranda, D. Borrás, J. H. Calvo, and J. Lucientes. 2008. Modelling the distributions and spatial coincidence of bluetongue vectors *Culicoides imicola* and the *Culicoides obsoletus* group throughout the Iberian peninsula. *Med. Vet. Entomol.* 22: 124–134.
- Chefaoui, R. M., J. Hortal, and J. M. Lobo. 2005. Potential distribution modelling, niche characterization and conservation status assessment using GIS tools: a case study of Iberian *Copris* species. *Biol. Conserv.* 122: 327–338.
- Cianci, D., N. Hartemink, and A. Ibáñez-Justicia. 2015. Modelling the potential spatial distribution of mosquito species in the Netherlands using three different techniques. *Int. J. Health Geogr.* 14:10.
- Connelly, C. R., and D. B. Carlson. 2009. Florida Coordinating Council on mosquito control. Florida mosquito control: The state of the mission as defined by mosquito controllers, regulators, and environmental managers. Vero Beach, Florida: University of Florida, Institute of Food and Agricultural.
- Cutler, D. R., T. C. Edwards Jr, K. H. Beard, A. Cutler, K. T. Hess, J. Gibson, et al. 2007. Random forests for classification in ecology. *Ecology.* 88: 2783–2792.
- Dekoninck, W., F. Hendrickx, W. Van Bortel, V. Versteirt, M. Coosemans, D. Damiens, et al. 2011. Human-induced expanded distribution of *Anopheles plumbeus*, experimental vector of West Nile virus and a potential vector of human malaria in Belgium. *J. Med. Entomol.* 48: 924–928.
- European Centre for Disease Prevention and Control (ECDC). 2013. Environmental risk mapping: *Aedes albopictus* in Europe. Stockholm. Available from: <http://www.ecdc.europa.eu/en/publications/Publications/climate-change-environmental-risk-mapping-aedes.pdf>
- European Environment Agency [www.eea.europa.eu].
- Fielding, A. H., and J. F. Bell. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environ. Conserv.* 24: 38–49.
- Fischer, E. A. J., G.-J. Boender, G. Nodelijk, A. A. de Koeijer, and H. J. Van Roermund. 2013. The transmission potential of Rift Valley fever virus among livestock in the Netherlands: a modelling study. *Vet. Res.* 44: 58.

- Gibson, L. A., B. A. Wilson, D. M. Cahill, and J. Hill. 2004. Spatial prediction of rufous bristlebird habitat in a coastal heathland: a GIS-based approach. *J. Appl. Ecol.* 41: 213–223.
- Hartemink, N., S. O. Vanwambeke, H. Heesterbeek, D. Rogers, D. Morley, B. Pesson, et al. 2011. Integrated Mapping of Establishment Risk for Emerging Vector-Borne Infections: A Case Study of Canine Leishmaniasis in Southwest France. *PLoS One.* 6: e20817.
- Hartemink, N. A., B. V Purse, R. Meiswinkel, H. E. Brown, A. De Koeijer, A. R. W. Elbers, et al. 2009. Mapping the basic reproduction number (R0) for vector-borne diseases: A case study on bluetongue virus. *Epidemics.* 1: 153–161.
- Ibañez-Justicia, A., A. Stroo, M. Dik, J. Beeuwkes, and E. J. Scholte. 2015. National Mosquito (Diptera, Culicidae) Survey in the Netherlands 2010–2013. *J. Med. Entomol.* doi: 10.1093/jme/tju058
- Justice, C., J. Townshend, E. Vermote, R. Sohlberg, J. Descloitres, D. Roy, et al. 2000. Preliminary land surface products from the NASA Moderate Resolution Imaging Spectroradiometer (MODIS), pp. 1157–1162. In *Geosci. Remote Sens. Symp. 2000. Proceedings. IGARSS 2000. IEEE 2000 Int.*
- Krüger, A., A. Rech, X.-Z. Su, and E. Tannich. 2001. Two cases of autochthonous *Plasmodium falciparum* malaria in Germany with evidence for local transmission by indigenous *Anopheles plumbeus*. *Trop. Med. Int. Heal.* 6: 983–985.
- Lobo, J. M., A. Jiménez-Valverde, and J. Hortal. 2010. The uncertain nature of absences and their importance in species distribution modelling. *Ecography (Cop.)*. 33: 103–114.
- Marshall, J. F. 1938. The British mosquitoes. *Brit. Mus. (Nat. Hist.)*, London. 341.
- McCullagh, P., and J. A. Nelder. 1989. *Generalized linear models*. Chapman & Hall/CRC, London.
- Miller, J. 2010. Species distribution modeling. *Geogr. Compass.* 4: 490–509.
- MODIS [modis.gsfc.nasa.gov].
- National Weather Service Climate Prediction Center [www.cpc.ncep.noaa.gov].
- Osborne, P. E., J. C. Alonso, and R. G. Bryant. 2001. Modelling landscape-scale habitat use using GIS and remote sensing: a case study with great bustards. *J. Appl. Ecol.* 38: 458–471.
- Peters, J., B. De Baets, N. E. C. Verhoest, R. Samson, S. Degroeve, P. De Becker, et al. 2007. Random forests as a tool for ecohydrological distribution modelling. *Ecol. Modell.* 207: 304–318.
- QGIS Development Team. 2012. QGIS Geographic Information System.
- Randolph, S. E., and D. J. Rogers. 2010. The arrival, establishment and spread of exotic diseases: patterns and predictions. *Nat. Rev. Microbiol.* 8: 361–371.
- Rogers, D. J., and B. G. Williams. 1994. Tsetse distribution in Africa: seeing the wood and the trees.
- Rogers, D. J. 2000. Satellites, space, time and the African trypanosomiases. *Adv. Parasitol.* 47: 129–171.
- Schaffner, F., G. Angel, B. Geoffroy, J. P. Hervy, A. Rhaiem, and J. Brunhes. 2001. The mosquitoes of Europe, an identification and training programme. Montpellier, Fr. IRD.
- Schaffner, F., I. Thiéry, C. Kaufmann, A. Zettor, C. Lengeler, A. Mathis, and C. Bourguin. 2012. *Anopheles plumbeus* (Diptera: Culicidae) in Europe: a mere nuisance mosquito or potential malaria vector? *Malar. J.* 11: 393.
- Scharlemann, J. P. W., D. Benz, S. I. Hay, B. V Purse, A. J. Tatem, G. R. W. Wint, et al. 2008. Global data for ecology and epidemiology: a novel algorithm for temporal Fourier processing MODIS data. *PLoS One.* 3: e1408.

- Shute, P. G., and M. Maryon. 1974. Malaria in England past, present and future. *J. R. Soc. Promot. Health.* 94: 23–29.
- Snow, K. R. 1990. Mosquitoes., *Naturalist' Handbooks* 14. Richmond Publishing Co. Ltd., Slough, England; 1990:66.
- Socioeconomic Data and Applications Center [sedac.ciesin.columbia.edu].
- Sokolova, M. I., and K. R. Snow. 2002. Malaria vectors in European Russia. *Eur Mosq Bull.* 12: 1–6.
- Takken, W., P. A. Kager, and J. P. Verhave. 2007. Will malaria return to North-west Europe? In *Emerg. Pests Vector-Borne Dis.* Wageningen Academic Publishers, Wageningen. pp. 23–34.
- Townroe, S., and A. Callaghan. 2014. British Container Breeding Mosquitoes: The Impact of Urbanisation and Climate Change on Community Composition and Phenology. *PLoS One.* 9: e95325.
- Verdonschot, P. 2002. Family Culicidae. In: Beuk PLT, editor. *Checklist of the Diptera in the Netherlands.* Utrecht: KNNV Publishers; p. 98.
- Versteirt, V., E. De Clercq, W. Dekoninck, D. Damiens, A. Ayrinhac, F. Jacobs, and W. Van Bortel. 2009. Mosquito vectors of disease: spatial biodiversity, drivers of change, and risk. Final Report. Belgian Science Policy: Brussels.
- WorldClim - Global Climate Data [www.worldclim.org].

CHAPTER 4

High resolution spatial analysis of habitat preference of *Aedes albopictus* (Diptera:Culicidae) in an urban environment

Daniela Cianci¹, Nienke Hartemink¹, Caroline B Zeimes², Sophie O Vanwambeke²,
Annamaria Ienco³, and Beniamino Caputo³

¹ Faculty of Veterinary Medicine, Utrecht University, Utrecht, The Netherlands

² Georges Lemaitre Centre for Earth and Climate Research Earth and Life Institute, Université Catholique de Louvain, Louvain-la-Neuve, Belgium

³ Dipartimento di Sanità Pubblica e Malattie Infettive, Università di Roma "Sapienza", Roma, Italy

Journal of Medical Entomology. 1–7 (2015);

DOI: <http://dx.doi.org/10.1093/jme/tjv026>

ABSTRACT

Over the past decades, the Asian tiger mosquito (*Aedes albopictus* (Skuse, 1895)) has emerged in many countries, and it has colonized new environments, including urban areas. The species gives nuisance and is a potential vector of several human pathogens, and a better understanding of the habitat preferences of the species is needed for help in successful prevention and control. So far, the habitat preference in urban environments has not been studied in Southern European cities. In this paper, spatial statistical models were used to evaluate the relationship between egg abundances and land cover types on the campus of Sapienza University in Rome, which is taken as an example of a European urban habitat. Predictor variables included land cover types, classified in detail on a high resolution image, as well as solar radiation and month of capture. The models account for repeated measures in the same trap and are adjusted for meteorological circumstances. Vegetation and solar radiation were found to be positively related to the number of eggs. More specifically, trees were positively related to the number of eggs and the relationship with grass was negative. These findings are consistent with the species' known preference for shaded areas. The unexpected positive relationship with solar radiation is amply discussed in the paper. This study represents a first step toward a better understanding of the spatial distribution of *Ae. albopictus* in urban environments.

Keywords : Asian tiger mosquito, urban habitat preference, ovitrap, spatial analysis, vector-borne disease

BACKGROUND

The mosquito *Aedes albopictus* (Skuse, 1895), commonly known as the “Asian tiger mosquito,” originated in Southeast Asia, but has spread in temperate Asia, Europe, North America, as well as in South Africa and in several locations in the Pacific and Indian Ocean regions (Knudsen 1995, Benedict et al. 2007, Paupy et al. 2009). The geographical spread of *Ae. albopictus* has mostly occurred during the past few decades, largely through the international trade of used tires (Reiter and Sprenger 1987, Lounibos 2002, Tatem et al. 2006), but also via transport of Lucky Bamboo plants from China (Madon et al. 2002). In Europe, *Ae. albopictus* was reported for the first time in Albania in 1979 (Adhami and Murati 1987) and in Italy in 1990, in Genoa, and from there it has gradually spread to several Italian regions (Sabatini et al. 1990, Dalla Pozza and Majori 1992, Della Torre et al. 1992, Romi and Majori 2008). In Rome, *Ae. albopictus* was detected in 1997 (Romi et al. 1999), and this led to the first example of complete colonization of an urban area in Italy (Di Luca et al. 2001, Toma et al. 2003, Severini et al. 2008). The efficient spread of *Ae. albopictus* to urban areas in temperate regions is associated to two major characteristics of this species: the capacity to produce hibernating eggs (Hawley 1988) and the ability to shift from natural breeding sites (e.g., bamboo stumps, tree holes) to anthropogenic ones (e.g., manholes, water storage containers, used tires, flower pots, cemetery urns; Hawley 1988, Knudsen 1995, Tsuda et al. 2006, Vazeille et al. 2008, Paupy et al. 2009). *Ae. albopictus* is a vector for many arboviral infections, including dengue and chikungunya (Knudsen 1995, Gratz 2004, Angelini et al. 2007, Rezza et al. 2007, Thenmozhi et al. 2007, de Lamballerie et al. 2008) and was responsible for a large chikungunya epidemics in the Indian Ocean in 2005–2006 (Enserink 2006), for a small outbreak in Italy (Angelini et al. 2007, Rezza et al. 2007), and for two human cases of dengue in South of France in 2010 (La Ruche et al. 2010). These latest cases in Europe highlight the potential risk of arbovirus outbreaks in the European cities where the species is now well established. Habitats of *Ae. albopictus* in urban environment have been investigated in the United States (Barker et al. 2003a,b, Braks et al. 2003; Bartlett-Healy et al. 2012, Unlu et al. 2013), South America (Braks et al. 2003), and Asia (Gilotra et al. 1967). However, there is only little information on the habitat preferences of *Ae. albopictus* in urban areas in Europe. Knowledge on the abundance of *Ae. albopictus* (Cianci et al. 2013) and on its spatial distribution may help in designing control measures, such as removing favourable breeding sites or making favourable habitats less attractive or accessible. Also, identifying areas of higher *Ae. albopictus* abundance is important for constructing risk maps that indicate the risk for an outbreak after an introduction (Hartemink et al. 2009, 2011). Although vector presence and abundance are not the only factors determining whether or not a pathogen can spread in an area, determining the distribution of the vector is an essential step in studying the risk of transmission of a pathogen.

In this paper, we model the relationship between the abundance of eggs and different types of land cover, solar radiation and the month of capture, adjusting for the repeated measures in the same trap and the meteorological circumstances. Data were collected in a small area within

Rome, the campus of Sapienza University, which is taken as an example of a European urban habitat. This paper aims at improving the understanding of habitat preferences of *Ae. albopictus* in urban areas, in order to help in the prevention and control of diseases that this mosquito could spread.

MATERIALS AND METHODS

Study area

The campus of Sapienza University in Rome is situated in an urbanized area in the center of Rome (Supp Figure 1) and is mainly characterized by buildings, green areas with grass, trees, bushes and hedges, a botanical garden, car parks and roads. With its mixture of built-up areas and vegetation, the campus may be considered as representative of many areas of the city. A high resolution classification of land cover was produced by hand digitization of an aerial picture ("*Rome, Sapienza.*" 41°54'10.51"N and 12°30'53.24"E. Google Earth. July 2007; 25 April 2012, Supp Figure 2) in Quantum GIS (QGIS Development Team 2012; Figure 1).

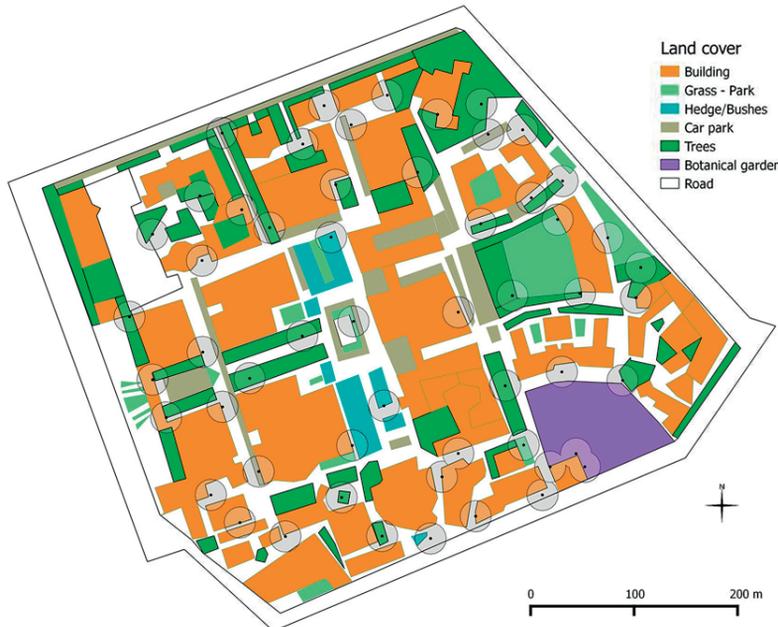


Figure 1. Distribution of the ovitraps (black dots) in the campus of Sapienza University in Rome (Italy) with circular buffers of 15-m radius around each trap. Different colors indicate different land cover classes.

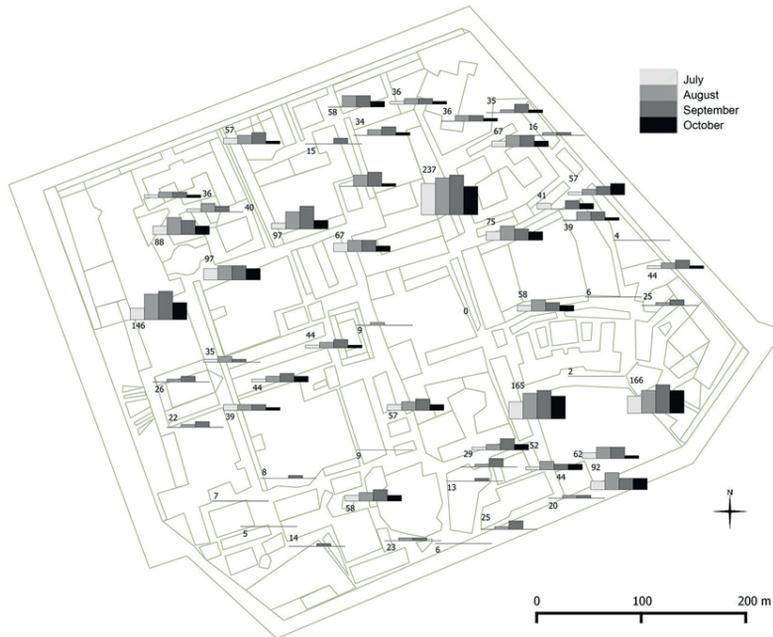


Figure 2. Map of the campus of Sapienza University in Rome (Italy) with bar charts showing the average per month of the weekly counts of eggs for each ovitrap. The numbers indicate the total number of eggs found in each trap.

Data collection

The number of eggs was monitored weekly using ovitraps, starting on the 8th of July until the 21st of October 2011. Fifty-five ovitraps were located at ground level in shaded sites (e.g., under bushes or trees) and distributed over an area of approximately 22 hectares (Figure 1), as measured by Google Earth (Google Earth Beta 5.2.1, 2010). Ovitrap consisted of black acrylonitrile butadiene styrene (ABS) truncated cones (diameter at base 8.5 cm, diameter at top 12 cm, height 13.5 cm) filled with 500 ml of tap water and equipped with a hole about 2 cm from the top edge to prevent complete filling in the case of rain. The ovitraps were lined with germination paper for egg laying. The germination paper and water were replaced and ovitraps were cleaned to remove possible remaining eggs on a weekly basis. Only the eggs adhering to the germination paper (even if already hatched) were taken into consideration, whereas larvae found occasionally in the ovitrap were killed and discarded. Each germination paper was kept in a numbered plastic bag and brought to the lab for egg identification and counting under a stereomicroscope.

Statistical analysis

Regression models for count data were applied to model the relationship between the number of eggs in the traps and the characteristics of the area surrounding the traps. A Poisson model

and a negative binomial model were compared, and the best model was chosen based on the Akaike information criterion (AIC). The best model was defined according to the commonly used rule-of-thumb that a model is better than another model when its AIC is at least 2 units smaller (Burnham and Anderson 2004).

The dependent variable was the weekly number of eggs collected in the ovitraps. The independent variables were the land cover categories, the solar radiation, and the month of capture. A random effect for the trap was included in the model to account for the fact that the observations taken from the same trap are correlated (Verbeke and Molenberghs 2009). Meteorological variables were included to adjust for the weather effect. Models with all possible combinations of variables were compared and the best model was selected based on the AIC values. The variables selection was done with the R package MuMIn (R Development Core Team 2013).

The land cover was classified into buildings, car parks, roads, grass, trees, hedges and botanical garden. The vegetation data were grouped in two different ways: in model 1 all vegetation classes were merged into a single vegetation class, whereas in model 2 all land cover classes were used as separate variables. The land cover classes were incorporated into the model using a circular buffer around each ovitrap location, as described and applied in Vanwambeke et al. (2011). The area (expressed in m²) of each land cover class was calculated within the buffers. Buffers with radius of 10, 15, 20 m were tested, in order to consider the surrounding landscape characteristics at different distances from the traps and the model with buffer size of 15 m was selected because it had the smallest AIC (Table S1). Further discussions about the buffer size can be found in the Supplementary Material (Supp Figure 3). Solar radiation was used as a proxy for the amount of sunshine and the local temperature in the areas around the traps, which were put in shaded sites and never exposed directly to sunlight. Solar radiation represents the expected amount of sunlight in a specific season, based on factors such as altitude, latitude, aspect of a slope, and day length. It is a very general measure, basically reflecting whether a trap was placed at the sunny or shadowed side of a building or a wall. This variable was calculated as the sum of the direct, diffuse, and global insolation calculated at every location on an elevation map (solar radiation tool of ArcGIS version 10.0, Environmental Systems Resource Institute [ESRI] 2011, Rich et al. 1994), and meteorological conditions were not taken into account. The elevation map combined information on the height of the buildings, to account for the shadow they produced, with a map of the ground elevation (Google Earth). The ground elevation map was interpolated by kriging from a set of altitude points collected in Google Earth. For each buffer, the value of the solar radiation was calculated excluding areas of the buffers covered by buildings. This ensured that the solar radiations were calculated only at the ground level, where the traps were located, and not on the roof of the buildings. The solar radiation is calculated for the area around the traps, meaning that it reflects the amount of sunshine and warmth in the proximity of the trap and not necessarily at the site of the trap itself.

The meteorological data were included in the model in order to adjust for the weather effect, which is known to affect the dynamics of mosquitoes and eggs (Bentley and Day 1989, Deichmeister and Telang 2011). Further investigations on the weather effect were not possible, as no micro climate data at the ovitrap sites were available. Meteorological variables were maximum temperature, average humidity and wind speed, and the cumulative value of the precipitation; all were measured for a period of 5 d before each collection of eggs (Supp Figure 4). The meteorological data were recorded at the nearest weather station, i.e., “Prenestina – Malatesta” in Rome (approximately 2.5 km from the campus; source: www.wunderground.com (accessed 6 March 2015)). Since the ovitraps were located within a small area and checked repeatedly (weekly), the data could be spatially and temporally autocorrelated. The presence of autocorrelation was graphically assessed by looking at the residuals of the model: the autocorrelation function (ACF) plot of residuals was used for the dependence over time (R package stats, R Core Team 2013) and the variogram was used for dependence in space (Ribeiro Jr and Diggle 2001). It was also tested whether including a random effect for the time of the capture would improve the model.

RESULTS

An overview of the spatial distribution of number of eggs collected in the campus is presented in Figure 2, where the average per month of the weekly counts of eggs has been calculated for each trap and plotted in the map. The mean number of eggs per trap with the standard error and the temporal trend of the number of eggs collected in the campus are presented respectively in Supp Figure 5 and 6).

The negative binomial regression model was preferred to the Poisson regression model, because it had a lower AIC (Supp Table 1). The negative binomial model coefficients are interpreted as incidence rate ratios (IRR), which means that when $IRR > 1$ the association between the variable and the number of eggs is positive and when $IRR < 1$ the association is negative. No spatial and temporal autocorrelation was found.

In model 1, where the vegetated land cover classes (i.e., grass, trees, hedges, and botanical garden) were grouped, vegetation, solar radiation, and months of collections were statistically significant at 0.05 level (Table 1). Vegetation and solar radiation were positively associated with the number of eggs.

In model 2, where the vegetation was subdivided into different classes, grass, trees, solar radiation, and months were statistically significant at 0.05 level, with trees and solar radiation having a positive effect and grass having a negative effect (Table 2). The best model selected included the botanical garden, but this variable was not statistically significant at 0.05 level.

Table 1. Model 1: Negative binomial regression model with number of eggs as outcome and the best subset of predictor variables, chosen among vegetation, buildings, car parks, roads, solar radiation, month, maximum temperature, average humidity, wind speed, and precipitation

VARIABLE	IRR ^a	P value
Vegetation^b	1.153	0.0440
Solar radiation^c	1.047	<0.0001
August^d	1.954	<0.0001
September^d	2.801	<0.0001
October^d	1.811	0.0008
Temperature Max	0.980	0.0887
Humidity Max	0.948	<0.0001
Precipitation	0.999	0.774

^aIRR = incident rate ratio; IRR>1 positive effect; IRR<1 negative effect.

^bFor the land cover classes, the IRR is calculated for an increase of 100m².

^cFor the solar radiation, the IRR is calculated for an increase of 1000 WH/km².

^dJuly is the reference category for the month.

In bold variables significant at 0.05 level.

Table 2. Model 2: Negative binomial regression model with number of eggs as outcome and the best subset of predictor variables, chosen among grass, trees, hedges, botanical garden, buildings, car parks, roads, solar radiation, month, maximum temperature, average humidity, wind speed, and precipitation

VARIABLE	IRR ^a	P value
Botanical garden ^b	1.221	0.1003
Grass^b	0.726	0.0396
Trees^b	1.241	0.0017
Solar radiation^c	1.049	<0.0001
August^d	1.992	<0.0001
September^d	2.858	<0.0001
October^d	1.885	<0.0001
Temperature Max	0.982	0.0480
Humidity Max	0.946	0.0285

^aIRR = incident rate ratio; IRR>1 positive effect; IRR<1 negative effect.

^bFor the land cover classes, the IRR is calculated for an increase of 100m².

^cFor the solar radiation, the IRR is calculated for an increase of 1000 WH/km².

^dJuly is the reference category for the month.

In bold variables significant at 0.05 level.

DISCUSSION

Spatial statistical models were applied to ovitrap data to identify ecological factors shaping the distribution of *Ae. albopictus* eggs in the campus of Sapienza University in Rome. Vegetation, solar radiation and month of capture turned out to be significant to explain the differences in egg abundance. Areas with trees were statistically significantly associated with higher numbers of eggs as already shown in its original range (Hawley 1988) and in Thailand and Hawaii (Vanwambeke et al. 2007, 2011), probably because their shadow provides a favourable habitat for mosquitoes. Grass was negatively associated with the number of eggs indicating that an open area without high vegetation and shelter is less attractive for mosquitoes, as already shown in Rey et al. (2006) and Honório et al. (2009). The positive effect of solar radiation suggests that areas with more exposure to sunshine are favourable for oviposition. It should be noted that the calculated solar radiation basically reflects whether the area around the trap is on the sunny or shadowed side of a building or a wall and, thus, may be taken as a proxy of the amount of sunshine and warmth in the proximity of the ovitrap. Also, it is important to stress that the ovitraps were put in shaded sites and never exposed directly to sunlight and that the landscape analysis could not take in specific consideration the determinants of the shade (e.g. small bushes, benches, low walls, etc.). Therefore, the results obtained suggest that solar radiation influences the attractivity of a location for gravid mosquitoes, either by (indirectly) creating a better temperature at the times at which the mosquitoes search for an oviposition site or, for instance, by increasing the water evaporation from the ovitrap and consequently the humidity around the traps, which is known to attract the mosquitoes (Wan-Norafikah et al. 2009, Rohani et al. 2011). The number of eggs increased in August and September, in agreement with the known population dynamics of *Ae. albopictus* in Rome (Toma et al. 2003).

It should be kept in mind that all collection methods have a bias related to the fraction of the population they target. Strictly speaking results based on ovitrap data indicate suitability of a habitat for oviposition, not necessarily for other activities, such as resting or host-seeking. However, correlation was found between adult mosquitoes collected by sticky traps and number of eggs collected by ovitraps (Facchinelli et al. 2007) and ovitraps are widely used to monitor and survey the species distribution and relative densities (Toma et al. 2003, Severini et al. 2008, Becker et al. 2010) as well as to study habitat preference (Barker, Paulson, et al. 2003, Rey et al. 2006). The presence of alternative potential breeding sites close to the traps could affect the results, in the sense that they would attract gravid female mosquitoes that would otherwise have laid their eggs in the traps. In our study area, alternative breeding sites could be expected in the botanical garden, where water could accumulate in plant pots and dishes, after watering of the plants. This means that the positive effect (close to significance at the 0.10 level) of the botanical garden may actually have been underestimated by the model and that in reality, the effect is even stronger.

The aim of the study was to identify favorable habitats for *Ae. albopictus* in a small urban area. If we want to generalize the results, we have to consider that, although the campus is quite rep-

representative of an urban area (with a mixture of buildings, streets and some green areas), it is still a small area, which does not include all possible land cover classes. In further studies it would be interesting to include also information that was not available in this study, such as other land cover types (e.g., rivers, ponds, trees species, flowerbeds, distinguish between footpaths and busy roads), the presence of water (e.g., monitoring the water in the manholes and in other containers), and human density. Also, the effect of the coexistence of different land cover categories could be tested.

To our knowledge, this study is the first to investigate the habitat preferences of *Ae. albopictus* in a European urban area. In Europe, and particularly in Italy, the tiger mosquito is established especially in urban areas. Given the nuisance caused by this species and especially the risk of transmission of pathogens if these pathogens would be introduced, it is important to have information on habitat preferences to plan carefully monitoring surveys and control measures. This study represents a first step toward a better understanding of the spatial distribution of *Ae. albopictus* in urban environments. Although this is not the only factor determining whether or not a pathogen can spread in an area, determining the distribution of the vector is an essential step in studying the risk of *Ae. albopictus*-transmitted pathogens.

ACKNOWLEDGMENTS

We would like to thank Hans Heesterbeek and Alessandra della Torre for their constructive comments on the manuscript, Jan van den Broek for useful discussion on the statistics, and Neil Alexander for help with the digitization of the Google Earth map. This study was funded by EU grant FP7-261504 EDENext and is catalogued by the EDENext Steering Committee as EDENext173 (www.edenext.eu) (accessed 5 March 2015). The contents of this publication are the sole responsibility of the authors and do not necessarily reflect the views of the European Commission.

REFERENCES

- Adhami, J., and N. Murati. 1987. Prani e mushkonjes *Aedes albopictus* ne shqiperi. Rev. Mjekesore. 1: 13–16.
- Angelini, R., A. C. Finarelli, P. Angelini, C. Po, K. Petropulacos, G. Silvi, P. Macini, C. Fortuna, G. Venturi, F. Magurano, et al. 2007. Chikungunya in north-eastern Italy: a summing up of the outbreak. Euro Surveill. 12: E071122.
- Barker, C. M., C. C. Brewster, and S. L. Paulson. 2003. Spatiotemporal oviposition and habitat preferences of *Ochlerotatus triseriatus* and *Aedes albopictus* in an emerging focus of La Crosse virus. J. Am. Mosq. Control Assoc. 19: 382–391.
- Barker, C. M., S. L. Paulson, S. Cantrell, and B. S. Davis. 2003. Habitat preferences and phenology of *Ochlerotatus triseriatus* and *Aedes albopictus* (Diptera: Culicidae) in southwestern Virginia. J. Med. Entomol. 40: 403–410.
- Bartlett-Healy, K., I. Unlu, P. Obenauer, T. Hughes, S. Healy, T. Crepeau, A. Farajollahi, B. Kesavaraju, D. Fonseca, G. Schoeler, et al. 2012. Larval mosquito habitat utilization and community dynamics of *Aedes albopictus* and *Aedes japonicus* (Diptera: Culicidae). J. Med. Entomol. 49: 813–824.
- Becker, N., D. Petric, M. Zgomba, C. Boase, M. Madon, C. Dahl, and K. A. 2010. Mosquitoes and Their Control. Springer, Heidelberg, Dordrecht, New York.
- Benedict, M. Q., R. S. Levine, W. A. Hawley, and L. P. Lounibos. 2007. Spread of the tiger: global risk of invasion by the mosquito *Aedes albopictus*. Vector-Borne Zoonotic Dis. 7: 76–85.
- Bentley, M. D., and J. F. Day. 1989. Chemical ecology and behavioral aspects of mosquito oviposition. Annu. Rev. Entomol. 34: 401–421.
- Braks, M. A. H., N. A. Honório, R. Lourenço-De-Oliveira, S. A. Juliano, and L. P. Lounibos. 2003. Convergent habitat segregation of *Aedes aegypti* and *Aedes albopictus* (Diptera: Culicidae) in southeastern Brazil and Florida. J. Med. Entomol. 40: 785–794.
- Burnham, K. P., and D. R. Anderson. 2004. Multimodel inference understanding AIC and BIC in model selection. Sociol. Methods Res. 33: 261–304.
- Cianci, D., J. Van Den Broek, B. Caputo, F. Marini, A. Della Torre, H. Heesterbeek, and N. Hartemink. 2013. Estimating Mosquito Population Size From Mark-Release-Recapture Data. J. Med. Entomol. 50: 533–542.
- Dalla Pozza, G., and G. Majori. 1992. First record of *Aedes albopictus* establishment in Italy. J. Am. Mosq. Control Assoc. 8: 318–320.
- De Lamballerie, X., E. Leroy, R. Charrel, K. Tsetsarkin, S. Higgs, and E. Gould. 2008. Chikungunya virus adapts to tiger mosquito via evolutionary convergence: a sign of things to come? Virol. J. 5: 33.
- Deichmeister, J. M., and A. Telang. 2011. Abundance of West Nile virus mosquito vectors in relation to climate and landscape variables. J. Vector Ecol. 36: 75–85.
- Della Torre, A., C. Bomboi, G. Cancrini. 1992. Extension of the area of *Aedes albopictus* in Italy. First report of the species in Central Italy. Parassitol. 34: 143–146.
- Di Luca, M., L. Toma, F. Severini, F. D'Ancona, and R. Romi. 2001. *Aedes albopictus* a Roma: monitoraggio nel triennio 1998-2000. Ann. Ist. Super. Sanita. 37: 249–254.
- Enserink, M. 2006. Massive outbreak draws fresh attention to little-known virus. Science. 311: 1085.
- (ESRI) Environmental Systems Resource Institute. 2011. ArcGIS 10.0. Redlands, CA.

- Facchinelli, L., L. Valerio, M. Pombi, P. Reiter, C. Costantini, and A. della Torre. 2007. Development of a novel sticky trap for container-breeding mosquitoes and evaluation of its sampling properties to monitor urban populations of *Aedes albopictus*. *Med. Vet. Entomol.* 21: 183–195.
- Gilotra, S. K., L. E. Rozeboom, and N. C. Bhattacharya. 1967. Observations on possible competitive displacement between populations of *Aedes aegypti* Linnaeus and *Aedes albopictus* Skuse in Calcutta. *Bull. World Health Organ.* 37: 437.
- Google Earth Beta. 2010. Google Earth Beta, version 5.2.1. (<http://www.google.com/earth/index.html>) (accessed 5 March 2015).
- Gratz, N. G. 2004. Critical review of the vector status of *Aedes albopictus*. *Med. Vet. Entomol.* 18: 215–227.
- Hartemink, N. A., B. V Purse, R. Meiswinkel, H. E. Brown, A. De Koeijer, A. R. W. Elbers, G. J. Boender, D. J. Rogers, and J. A. P. Heesterbeek. 2009. Mapping the basic reproduction number (R0) for vector-borne diseases: A case study on bluetongue virus. *Epidemics.* 1: 153–161.
- Hartemink, N., S. O. Vanwambeke, H. Heesterbeek, D. Rogers, D. Morley, B. Pesson, C. Davies, S. Mahamdallie, and P. Ready. 2011. Integrated Mapping of Establishment Risk for Emerging Vector-Borne Infections: A Case Study of Canine Leishmaniasis in Southwest France. *PLoS One.* 6: e20817.
- Hawley, W. A. 1988. The biology of *Aedes albopictus*. *J. Am. Mosq. Control Assoc.* 1: 1–39.
- Honório, N. A., M. G. Castro, F. S. M. de Barros, M. de A. F. M. Magalhães, and P. C. Sabroza. 2009. The spatial distribution of *Aedes aegypti* and *Aedes albopictus* in a transition zone, Rio de Janeiro, Brazil. *Cad. Saude Publica.* 25: 1203–1214.
- Knudsen, B. A. 1995. Geographic spread of *Aedes albopictus* in Europe and the concern among public health authorities. *Eur. J. Epidemiol.* 11: 345–348.
- La Ruche, G., Y. Souarès, A. Armengaud, F. Peloux-Petiot, P. Delaunay, P. Desprès, A. Lenglet, F. Jourdain, I. Leparç-Goffart, F. Charlet, et al. 2010. First two autochthonous dengue virus infections in metropolitan France, September 2010. *Euro Surveill.* 15: 19676.
- Lounibos, L. P. 2002. Invasions by insect vectors of human disease. *Annu. Rev. Entomol.* 47: 233–266.
- Madon, M., M. Mulla, M. Shaw, S. Kluh, and J. Hazelrigg. 2002. Introduction of *Aedes albopictus* (Skuse) in southern California and potential for its establishment. *J. vector Ecol.* 27: 149–154.
- Paupy, C., H. Delatte, L. Bagny, V. Corbel, and D. Fontenille. 2009. *Aedes albopictus*, an arbovirus vector: From the darkness to the light. *Microbes Infect.* 11: 1177–1185.
- QGIS Development Team. 2012. QGIS Geographic Information System
- R Core Team. 2013. R: A Language and Environment for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/> (accessed 5 March 2015)..
- Reiter, P., and D. Sprenger. 1987. The used tire trade: a mechanism for the worldwide dispersal of container breeding mosquitoes. *J. Am. Mosq. Control Assoc.* 3: 494–501.
- Rey, J. R., N. Nishimura, B. Wagner, M. A. H. Braks, S. M. O'Connell, and L. P. Lounibos. 2006. Habitat segregation of mosquito arbovirus vectors in south Florida. *J. Med. Entomol.* 43: 1134–1141.
- Rezza, G., L. Nicoletti, R. Angelini, R. Romi, A. C. Finarelli, M. Panning, P. Cordioli, C. Fortuna, S. Boros, F. Magurano, et al. 2007. Infection with chikungunya virus in Italy: an outbreak in a temperate region. *Lancet.* 370: 1840–1846.
- Ribeiro Jr, P. J., and P. J. Diggle. 2001. geOR: A package for geostatistical analysis. *R news.* 1: 14–18.

- Rich, P. M., R. Dubayah, W. A. Hetrick, and S. C. Saving. 1994. Using viewshed models to calculate intercepted solar radiation: applications in ecology. American Society for Photogrammetry and Remote Sensing Technical Papers, pp. 524–529.
- Rohani, A., I. Suzilah, M. Malinda, I. Anuar, I. M. Mazlan, M. S. Maszaitun, O. Topek, Y. Tanrang, S. C. Ooi, H. Rozilawati, et al. 2011. *Aedes* larval population dynamics and risk for dengue epidemics in Malaysia. *Trop. Biomed.* 28: 237–248.
- Romi, R., and G. Majori. 2008. An overview of the lesson learned in almost 20 years of fight against the tiger mosquito. *Parassitologia.* 50: 117–119.
- Romi, R., M. Di Luca, and G. Majori. 1999. Current status of *Aedes albopictus* and *Aedes atropalpus* in Italy. *J. Am. Mosq. Control Assoc.* 15: 425–427.
- Sabatini, A., V. Raineri, G. Trovato, and M. Coluzzi. 1990. *Aedes albopictus* in Italy and the possible spread of this species in the Mediterranean area. *Parassitol.* 32: 301–304.
- Severini, F., M. Di Luca, L. Toma, and R. Romi. 2008. *Aedes albopictus* in Rome: results and perspectives after 10 years of monitoring. *Parassitologia.* 50: 121–123.
- Tatem, A. J., S. I. Hay, and D. J. Rogers. 2006. Global traffic and disease vector dispersal. *Proc. Natl. Acad. Sci.* 103: 6242–6247.
- Thenmozhi, V., J. G. Hiriyan, S. C. Tewari, P. P. Samuel, R. Paramasivan, R. Rajendran, T. R. Mani, and B. K. Tyagi. 2007. Natural vertical transmission of dengue virus in *Aedes albopictus* (Diptera: Culicidae) in Kerala, a southern Indian state. *Jpn. J. Infect. Dis.* 60: 245–249.
- Toma, L., F. Severini, M. Di Luca, A. Bella, and R. Romi. 2003. Seasonal patterns of oviposition and egg hatching rate of *Aedes albopictus* in Rome. *J. Am. Mosq. Control Assoc.* 19: 19–22.
- Tsuda, Y., W. Suwonkerd, S. Chawprom, S. Prajakwong, and M. Takagi. 2006. Different spatial distribution of *Aedes aegypti* and *Aedes albopictus* along an urban-rural gradient and the relating environmental factors examined in three villages in northern Thailand. *J. Am. Mosq. Control Assoc.* 22: 222–228.
- Unlu, I., A. Farajollahi, D. Strickman, and D. M. Fonseca. 2013. Crouching Tiger, Hidden Trouble: Urban Sources of *Aedes albopictus* (Diptera: Culicidae) Refractory to Source-Reduction. *PLoS One.* 8: e77999.
- Vanwambeke, S. O., P. Somboon, R. E. Harbach, M. Isenstadt, E. F. Lambin, C. Walton, and R. K. Butlin. 2007. Landscape and land cover factors influence the presence of *Aedes* and *Anopheles* larvae. *J. Med. Entomol.* 44: 133–144.
- Vanwambeke, S. O., S. N. Bennett, and D. D. Kapan. 2011. Spatially disaggregated disease transmission risk: land cover, land use and risk of dengue transmission on the island of Oahu. *Trop. Med. Int. Heal.* 16: 174–185.
- Vazeille, M., C. Jeannin, E. Martin, F. Schaffner, and A. Failloux. 2008. Chikungunya: a risk for Mediterranean countries. *Acta Trop.* 105: 200.
- Verbeke, G., and G. Molenberghs. 2009. Linear mixed models for longitudinal data. Springer-Verlag, New York, NY.
- Wan-Norafikah, O., C. D. Chen, H. N. Soh, H. L. Lee, W. A. Nazni, and M. Sofian-Azirun. 2009. Surveillance of *Aedes* mosquitoes in a university campus in Kuala Lumpur, Malaysia. *Trop. Biomed.* 26: 206–215.

SUPPLEMENTARY MATERIAL

Table S1. Comparison of the AIC of Poisson and negative binomial models for different buffer sizes. Negative binomial models have a lower AIC than the Poisson models. The negative binomial model with the lowest AIC is the model with a buffer of 15m. The results of the negative binomial model with a buffer of 15m are presented in the main text.

Model	Distribution	AIC buffer 10	AIC buffer 15	AIC buffer 20
1	Poisson	4672.96	4659.82	4661.14
	Negative binomial	4429.84	4429.62	4430.94
2	Poisson	5121.86	4655.04	4656.94
	Negative binomial	4426.44	4424.74	4426.7

Model 1: Model with number of eggs as outcome and vegetation, buildings, car parks, roads, solar radiation, month as explanatory variables; adjustment for the meteorological circumstances; random effect for the trap

Model 2: Model with number of eggs as outcome and buildings, car parks, roads, grass, trees, hedges, botanical garden, solar radiation, month as explanatory variables; adjustment for the meteorological circumstances; random effect for the trap

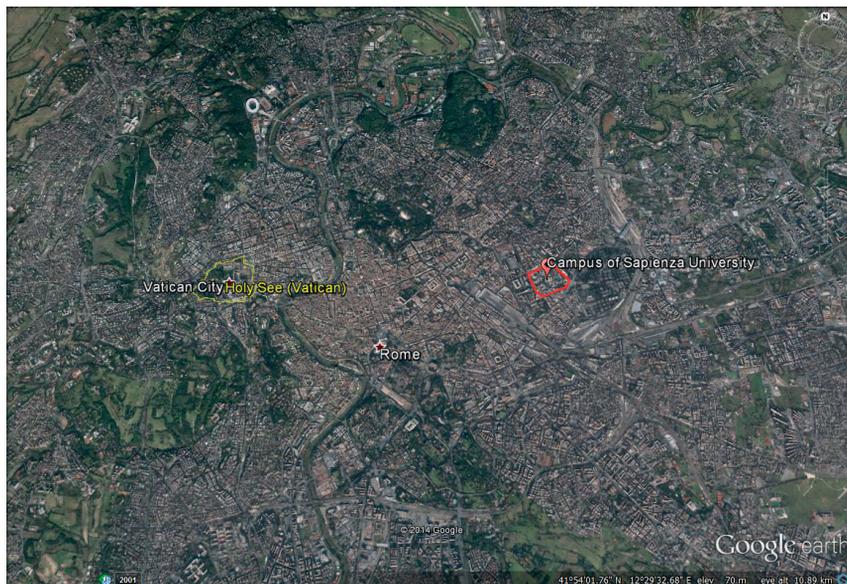


Figure 1. Google Earth image of Rome; the red polygon indicates the campus of Sapienza University in Rome. (source: "Rome, Sapienza." 41°54'10.51N and 12°30'53.24"E. Google Earth 5.0.11733.9347. July 2007; 05 September 2014).



Figure 2. Google Earth image of the campus of Sapienza University in Rome. (source: "Rome, Sapienza." 41°54'10.51N and 12°30'53.24"E. Google Earth 5.0.11733.9347. July 2007; 25 April 2012).

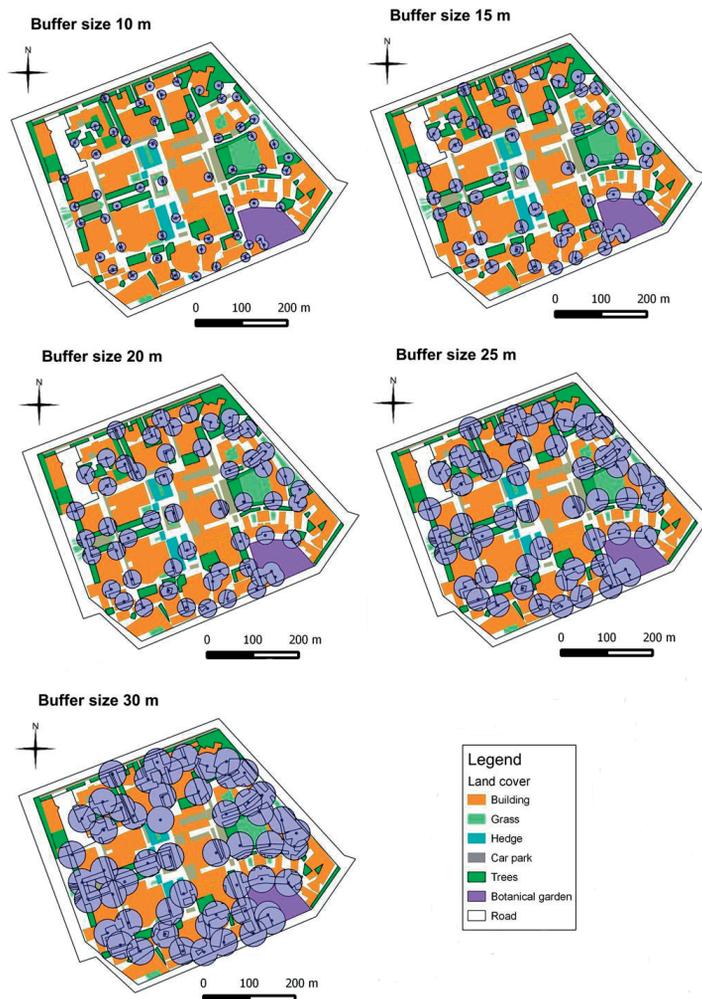


Figure 3. Maps of the campus with different buffer sizes (radius is 10, 15, 20, 25 or 30 meters). In this study, we tested three buffer sizes (10, 15 and 20 meters) to study the impact of landscape factors at different distances from the trap and we selected the buffer sizes that generated the best model in terms of AIC. The choice of buffer size relies on the aim of the study, the biology of the vector and the characteristic of the study area. The buffer size had to be smaller than the flight range, since we are interested in the importance of landscape factors and land covers in determining the suitability of the area for oviposition at a very local scale. This means that mosquitoes must be able to reach the different habitats within the campus, so that the number of eggs in a trap can be assumed to be a reflection of the preference of the mosquitoes. Given the distance between the traps, we could not use a buffer size bigger than 20m, because this would create a lot of overlap between buffers, which would lead to using the same value of land cover for several traps. This would violate the assumption of independent observations that is required for generalized linear models (McCullagh and Nelder 1989).

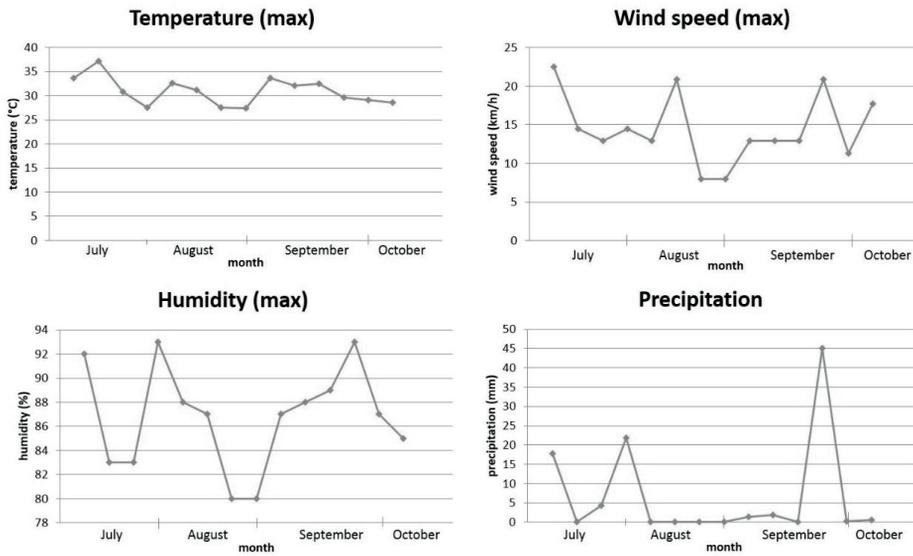


Figure 4. Trend of meteorological variables over the months.

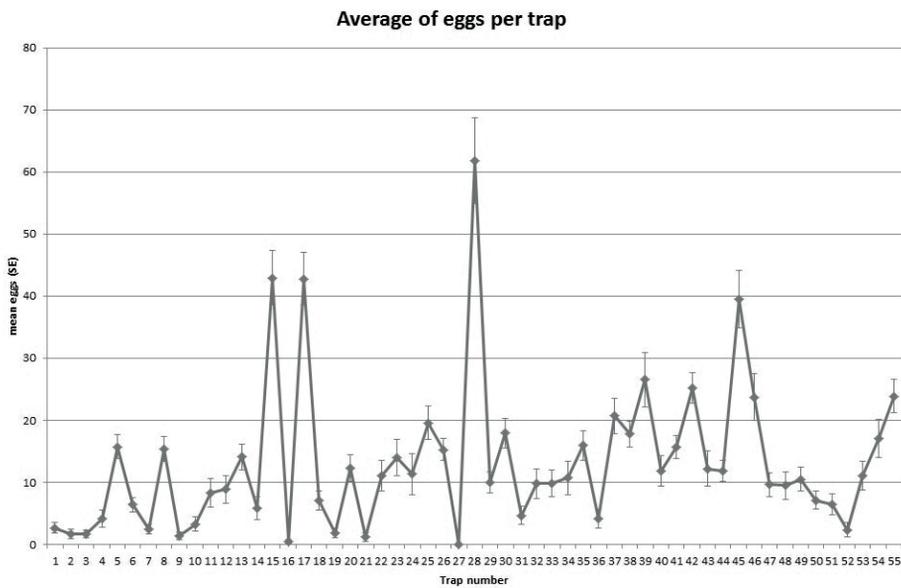


Figure 5. Average number of the weekly counts of eggs per trap with standard error (SE).

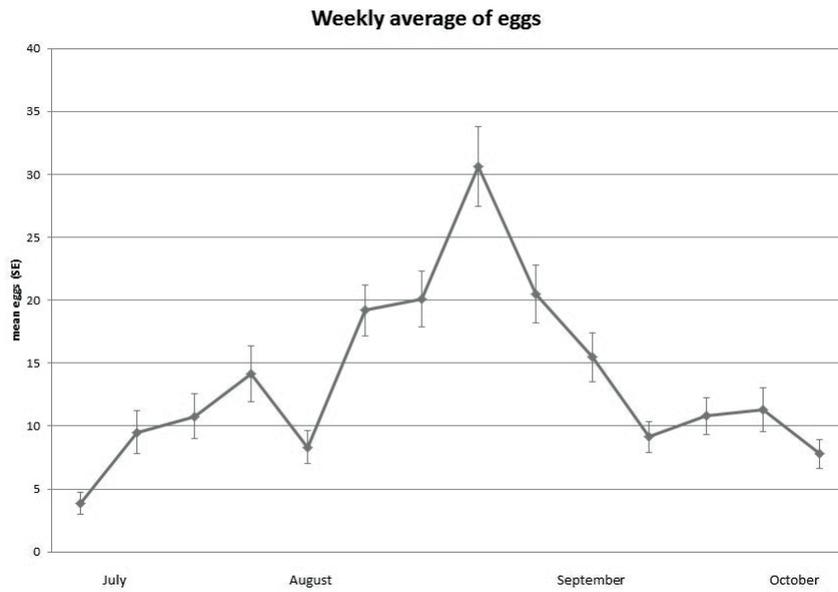


Figure 6. Temporal distribution of *Aedes albopictus* eggs collected in the campus of Sapienza University in Rome (Italy) in July – October 2011. The average of eggs collected weekly and the standard error (SE) are shown.

REFERENCES

McCullagh, P and Nelder, JA 1989. Generalized Linear Model. Volume 37. Chapman & Hall/CRC.

CHAPTER 5

Estimating mosquito population size from mark–release–recapture data

Daniela Cianci¹, Jan Van den Broek¹, Beniamino Caputo², Francesca Marini², Alessandra della Torre², Hans Heesterbeek¹, and Nienke Hartemink³

¹ Faculty of Veterinary Medicine, Utrecht University, The Netherlands.

² Dipartimento di Sanità Pubblica e Malattie Infettive, Università di Roma "Sapienza", Roma, Italy.

³ Georges Lemaître Centre for Earth and Climate Research Earth and Life Institute, Université Catholique de Louvain, Louvain-la-Neuve, Belgium

Journal of Medical Entomology. 50(3): 533-542 (2013);

DOI: <http://dx.doi.org/10.1603/ME12126>

ABSTRACT

Accurate estimation of population size is key to understanding the ecology of disease vectors, as well as the epidemiology of the pathogens they carry and to plan effective control activities. Population size can be estimated through mark-release-recapture (MRR) experiments that are based on the assumption that the ratio of recaptured individuals to the total captures approximates the ratio of marked individuals released to the total population. However, methods to obtain population size estimates usually consider pooled data and are often based on the total number of marked and unmarked captures. We here present a logistic regression model, based on the principle of the well-known Fisher-Ford method, specific for MRR experiments where the information available is the number of marked mosquitoes released, the number of marked and unmarked mosquitoes caught in each trap and on each day, and the geographic coordinates of the traps. The model estimates population size, taking into consideration the distance between release points and traps, the time between release and recapture, and the loss of marked mosquitoes to death or dispersal. The performance and accuracy of the logistic regression model has been assessed using simulated data from known population sizes. We then applied the model to data from MRR experiments with *Aedes albopictus* Skuse performed on the campus of "Sapienza" University in Rome (Italy).

Keywords: *Aedes albopictus*, vector population, logistic regression model, vector-borne disease control, mark-release-recapture

BACKGROUND

Knowledge of the population size of vectors in endemic or epidemic areas is valuable for understanding disease transmission dynamics and for determining the extent of control measures necessary to interrupt transmission. A measure developed in the field of infectious disease epidemiology, the value of which is indicative of control effort required, is the basic reproduction number, R_0 . R_0 combines, in a weighted way, the factors that determine whether a pathogen can become established in an area where it is introduced. One of the factors included in the R_0 formula, to which the numerical value is very sensitive, is the ratio of vectors to hosts (Hartemink et al 2009, 2011). The more precise the estimate of the population size, the better the estimate of R_0 .

Population size and other parameters of animal populations, such as dispersal, survival and flight range are estimated by mark-release-recapture (MRR) experiments (Pollock et al. 1990). MRR methods have a long history in ecology (Le Cren 1965) and a large literature on the modeling of capture-recapture methods exists (Begon 1979; Seber 1982, 1986; White et al. 1982).

Peterson was the first to estimate population size as the number of marked animals released divided by the proportion of marked in the group of captures (Lincoln 1930, Le Cren 1965, Seber 1982, Krebs 1999). The same idea was used afterwards in other models, such as Fisher-Ford's method, which can take into account a death rate from literature (Fisher 1947). Jackson (1937, 1939, 1940) obtained the estimate of population size by fitting a curve to a series of recapture rates, obtained from the change in recapture in time. The program MARK (White 1999) provides parameter estimates from marked animals, including population size. Capture and recapture probability also can be modeled as a function of time, but the software is based on the history of individual animals and therefore may not be useful for most mosquito experiments. Extensive overviews of MRR techniques to estimate animal population abundance are provided in the literature (Pollock 1980, Seber 1986, Schwarz and Seber 1999, Gratz 2004, Efford 2009). Not all the information collected with MRR experiments has been taken into account simultaneously in one model. For instance, it may be important to account for the animal loss rate estimated from the data, because of mortality or dispersal, that recapture occurred on more than one day and to consider whether there is a biological limit to their dispersal distance.

We present a specific method to estimate the population abundance of mosquitoes by using data based on MRR experiments, where recaptures take place on different dates (time effect) and at different distances (distance effect). The method is based on a logistic regression that models the fraction of marked and unmarked mosquitoes as a function of population size, loss rate, distance between the release site and the traps, and number of mosquitoes released. This model is specific for data that do not distinguish individuals. We illustrate the principle by applying the method to data collected during MRR experiments with *Aedes albopictus* Skuse mosquitoes on the campus of "Sapienza" University in Rome (Italy). We also apply the model to simulated data to show how the model performs in different scenarios, including different population size, different number of marked mosquitoes released, and different effect of including data on distance.

MATERIALS AND METHODS

Study area and mark–release–recapture experiments

Three mark-release-recapture experiments were carried out, using the experimental design described in detail in Marini et al. (2010), on the campus of “Sapienza” University in Rome. The campus is situated in an urbanized area in the center of Rome; the buildings -typically 10-20 m high- are separated by lanes and green spaces. For the experiments, 55 sticky traps (STs), specifically designed to collect *Ae. albopictus* females searching for oviposition and resting sites (Faccinelli et al. 2007), were located at ground level in sheltered positions, distributed over an area of ≈ 22 ha (Figure 1a) measured with Google Earth, version 5.2.1 (Google Earth Beta 2010). The Euclidean distances between the STs and the release points were calculated from their georeferenced locations by using ArcGIS (version 9.2, ESRI 2009). The number of sticky traps increases with the distance from the release site to keep the density of sticky-traps constant over space (Marini et al. 2010). *Aedes albopictus* adults used for MRR experiments were obtained from eggs collected by ovitraps within the study area. Eggs were hatched and larvae reared to the adult stage in plastic basins. The pupae were collected and transferred to emergence cages, where adults were maintained with 10% sugar solution for 2–9 d. On the morning of the release, females blood-fed by membrane feeders were marked with orange fluorescent dust and released from the center of the campus. The number of marked mosquitoes released was 464 in MRR1 (August of 2008), 566 in MRR2 (September of 2008), and 552 in MRR3 (October of 2008). Marked and unmarked mosquitoes in each ST were counted at days 5, 9, 13, 17, and 21 after release in MRR1 and at days 2-5, 7, 9, 13, 17, and 21 after release in MRR2 and MRR3. The recapture percentage in the campus, i.e., the number of marked recaptures divided by the number of mosquitoes released, was 4.5% for MRR1 (21 recaptures), 5.1% for MRR2 (29 recaptures), and 3.3% for MRR3 (18 recaptures) (Marini et al. 2010).

Data used in the analysis

Figure 1b indicates, by the size of the pie chart, the total number of mosquitoes captured in the three experiments, with the fraction of recaptured mosquitoes in black. As shown in the figure, few released mosquitoes have been recaptured in the eastern part of the campus. This is probably because of the fact that the western part of the campus (where the mosquitoes were released) is to a large extent separated from the eastern part by an almost continuous line of high buildings, interrupted by few relatively small openings. Because it is likely that these buildings could have represented a physical barrier for the mosquito dispersion eastwards, we decided to use only the trapping results from the area situated on the western side of the buildings crossing the study area for our analyses.

Figure 1a shows the 28 STs selected for the analysis that were situated in an area of 10.6 ha, indicated by the red line. A breakdown of the number of captured and recaptured mosquitoes in

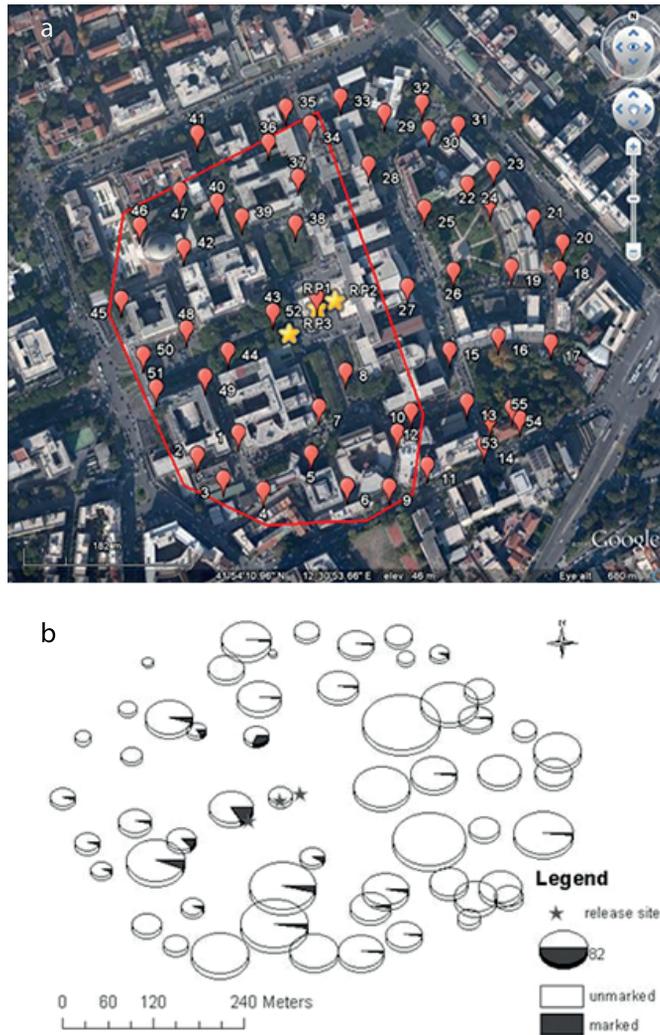


Figure 1. (a) Distribution of the sticky traps and release points on the campus of the University of Rome, “Sapienza.” The red markers indicate the 55 STs and the stars the release sites. The image is taken from Google Earth (source: “Sapienza,” Rome; 41° 53’58.48” N, 12° 31’13.59” E. GOOGLE EARTH. July 2007; 25 April 2012). The red line indicates the area selected for the analysis. **(b)** Distribution of captured female *Ae. albopictus* in the 55 sticky traps on the campus of “Sapienza” University of Rome. The stars indicate the release sites, the black slices represent the number of marked captured mosquitoes, and the white slices represent the number of the unmarked. The size of the pies is proportional to the total number of mosquitoes trapped in the three MRR experiments.

Table 1. Number of marked female *Aedes albopictus* recaptured (unmarked captured) per day in the three mark-release-recapture (MRR) experiments in the selected area (Fig 2.a) of the campus of “Sapienza” University in Rome (Italy).

MRR	Days								
	2	3	4	5	7	9	13	17	21
1	NS (NS)	NS (NS)	NS (NS)	15 (87)	NS (NS)	4 (87)	0 (88)	0 (109)	0 (127)
2	5 (32)	11 (37)	2 (14)	3 (17)	2 (38)	1(17)	0 (36)	0 (36)	0 (45)
3	0 (41)	1 (56)	1 (42)	9 (33)	4 (35)	0 (30)	0 (56)	0 (75)	0 (33)

The number of marked mosquitoes released is 464 (MRR1), 566 (MRR2), 552 (MRR3).
NS= not sampled.

the selected area during each experiment is presented in Table 1. Because no marked mosquitoes were captured more than 9 d after the release in each of the three experiments (Marini et al. 2010), the analysis was restricted to the first 9 d. No marked mosquitoes were found farther than 210 m in the selected area. Therefore, the analysis was restricted to the traps in the range of 210 m from the release site in this area.

To gauge the performance of the model and to compare this in different scenarios, we simulated data for the same experimental set-up. We did this for an existing population of 21,000 unmarked mosquitoes and with the number of mosquitoes released being 150, 510, and 1,005, respectively. The loss rate per day was fixed at 0.05. The distance was considered in the simulated data as three pooled areas; in one scenario marked and unmarked mosquitoes were equally distributed in the three areas, in another scenario only the unmarked mosquitoes were equally distributed and the marked decreased when the distance from the release point increased. The former scenario is unrealistic in nature, but we used it to investigate whether adding a measure of distance improves the performance of the model. In the latter scenario the marked mosquitoes were drawn from a binomial distribution of sample size 100, which is the number of mosquitoes caught in each trap, with a probability that changed according to time and distance. Details on how the data were simulated, including the R-code we used, are given in the supplementary material.

Logistic regression model

The formula derived by Petersen and Lincoln (1930) (Seber 1982) to estimate population size is the simplest estimation obtained from a single marking occasion in a closed population. It was the starting point to develop more intricate and refined methods. We present how it was modified to include a mortality effect and a distance effect, to illustrate the similarities with the logistic regression model we developed, which is based on the same principle.

The principle behind the method is that a known number, M , of individuals is caught, marked and then returned to the environment. After a suitable interval to allow mixing with the population, a second population sample, of size n , is taken and the numbers of marked and unmarked individuals recorded. When the assumption of a closed population is met and the captured animals

(marked and unmarked) constitute a simple random sample, the conditional distribution of the recaptured, m , given M and n , is hypergeometric (Chapman 1951, Seber 1982) (a list of the parameters used is provided in the Table 2). Here, the hypergeometric distribution describes the random variable that counts, for n elements randomly drawn from the population of size N , the number of marked individuals recaptured from the subgroup of the marked individuals released. Applying the technique to the experiment on mosquitoes, the probability to get m marked mosquitoes is:

$$f(m|M, n) = \frac{\binom{M}{m} \binom{N-M}{n-m}}{\binom{N}{n}} \quad [1]$$

Using a binomial approximation to the hypergeometric distribution, we have (Bailey 1951, 1952):

$$f(m|M, n) \approx \binom{n}{m} \left(\frac{M}{N}\right)^m \left(1 - \frac{M}{N}\right)^{n-m} \quad [2]$$

and the maximum likelihood estimate of \hat{N} is the Petersen estimate (Seber 1982). The estimator is based on the assumption that the ratio of marked individuals to the total number in the second sample (both marked and unmarked) reflects the same ratio in the total population, so that:

$$\frac{m}{n} = \frac{M}{N} \quad [3]$$

which provides the estimator:

$$\hat{N} = n \frac{M}{m} \quad [4]$$

(Seber 1982; Pollock 1990). A further extension of this formula, includes a mortality factor:

$$\hat{N} = \phi^t n \frac{M}{m} \quad [5]$$

where ϕ is the survival rate (actually more a "retention rate") and t is the period between the release and the recapture (Conway et al. 1974). This is known as Fisher-Ford's method (Dowdeswell et al. 1940, Fisher 1947). The mortality can be based on a laboratory population (MacLeod 1958)

or obtained by independent assessment in the field. Because M marked adults, reared from eggs or pupa, are added to the natural population of which the size is being estimated, M needs to be subtracted from the estimate (Nelson 1978; Reisen et al. 1979, 1980):

$$\hat{N} = \left[\phi' n \frac{M}{m} \right] - M \quad [6]$$

Using the correction provided by Fisher and Ford for situations when the number recaptured is small (<20) (Bailey 1952), the formula becomes:

$$\hat{N} = \left[\frac{\phi' (n+1)(M+1)}{m+1} \right] - 1 - M \quad [7]$$

We applied these formulas (6 and 7) to the real and simulated data. The confidence intervals related to the estimates are calculated with the method of percentile bootstrap (Efron and Tibshirani 1993) based on 500 bootstrap replicates at 95% level. The survival rate was a value taken from Alto and Juliano (2001), who investigated the relationship between temperature and duration of life of *Ae. albopictus*, through laboratory experiments. They estimated a daily mortality rate at 24°C of 0.0114. This value provides an estimate made in conditions similar to the temperature experienced in Rome during the MRR experiments. However, because the other environmental conditions are different there may nevertheless be a difference between the mortality rate as estimated in the laboratory and that for the experiment is carried out in the field. Since, because of the low recaptures, it was not possible to estimate the mortality rate from the data, we also analyzed the case of a mortality rate of 0.15, which is considered more likely under field conditions.

We now describe our logistic regression approach, based on the same principle as the Fisher-Ford method. Although for the Fisher-Ford method we need to set the value of the mortality rate, the survival rate present in the logistic regression model does not need preset values from the literature, because it can be estimated simultaneously from the data. We expect to recapture more mosquitoes in the first days, because after that they may have spread to a bigger area and some of them will have died. To control for this, the model includes a time effect, which means the effect of mortality and emigration. Furthermore, the logistic regression allows the inclusion of the distance between release point and STs, which may be important, because of the limited flight range of *Ae. albopictus*. Usually in an MRR experiment with insects, traps are placed at different distances from the release point and therefore each trap is expected to have a different recapture probability. This can be controlled for by introducing the distance effect in the model.

The starting point for the logistic regression model is the same as that of Fisher-Ford's method: the fraction of recaptured mosquitoes. If we assume a zero death rate for the logistic regression model, we can see the similarity between the methods. The model aims to estimate the number of mosquitoes in the period of interest and first we assume this quantity is constant and there is no mortality in the entire period. Then, the population fraction of marked mosquitoes is:

$$\pi = \frac{M}{M + N} \quad [8]$$

where the denominator considers that we introduced M mosquitoes in the total population (cf. equation 7). The ratio of the probability to capture a marked mosquito to the probability to capture an unmarked mosquito is:

$$\frac{\pi}{1 - \pi} = \frac{M}{N} \quad [9]$$

π is estimated as the ratio of marked mosquitoes recaptured to the total number of mosquitoes captured ($\pi = m/n$). We can then express the above formula as:

$$\frac{m}{n - m} = \frac{M}{N} \quad [10]$$

and it follows that:

$$N = \frac{M(n - m)}{m} \quad [11]$$

which is similar to the Lincoln index.

In the following, we add first survival and then distance to the model. We assume that N is the number of mosquitoes in the study area per day and that the number of marked mosquitoes released at time point $t=0$ is M . The population of mosquitoes can increase through emergences or decrease through deaths, whereas the population of marked mosquitoes can only decrease. At time t , $MS_m(t)$ is the number of marked mosquitoes that can still be trapped at that time. For the wild population, we are assuming a closed population, which means that when a mosquito is lost,

it is replaced by another one. That is, the number of emergences equals the number of deaths. Then the number of mosquitoes N will be constant over the time. We are also assuming a stable age distribution. Considering these assumptions, the population fraction of marked mosquitoes at time t is:

$$\pi(t) = \frac{MS_m(t)}{N + MS_m(t)} \quad [12]$$

and the ratio of the probability to capture a marked mosquito to the probability to capture an unmarked mosquito is:

$$\frac{\pi(t)}{1 - \pi(t)} = \frac{MS_m(t)}{N} \quad [13]$$

A sample of n mosquitoes at time t is taken and each is checked for marks. The number of marked mosquitoes in the sample has a binomial distribution with sample size n and population fraction π . From this, the logistic regression model is:

$$\log\left(\frac{\pi(t)}{1 - \pi(t)}\right) = -\log(N) + \log[S_m(t)] + \log(M) \quad [14]$$

Assuming an exponential distribution for the survival function, i.e., $S_m = e^{-\lambda t}$ (Macdonald 1952), and a log-linear model for the rate, i.e., $\lambda = e^{\beta_0}$, the logistic regression model with time effect is:

$$\log\left(\frac{\pi(t)}{1 - \pi(t)}\right) = -\log(N) - e^{\beta_0} t + \log(M) \quad [15]$$

The survival function in the model has a broader meaning than usual: it includes both a mortality factor and dispersal to outside the capture area. N can be estimated from the intercept of the model, i.e., $\alpha = -\log(N)$:

$$\hat{N} = \exp(-\hat{\alpha}) \quad [16]$$

and the approximate 95% confidence interval is given by:

$$\exp[-\hat{\alpha} - zSE(\hat{\alpha})] < N < \exp[-\hat{\alpha} + zSE(\hat{\alpha})] \quad [17]$$

where z is a standard normal distribution.

Besides the time effect, we incorporate a distance effect, to check whether it plays an important role in the sense that a model with distance included provides markedly different estimates, or improved estimate-confidence. The distance can be included in the model as a continuous variable, to consider the individual trap-distances, or divided in zones. Given the spatial distribution of the recaptures of our data, we decided to consider three different pooled distance areas from the release site. We consider that a certain fraction p of marked mosquitoes moves from an area to the adjacent one, starting from the release point, $p_i MS(t)$. If we consider three distances, we expect to find the mosquito population N uniformly distributed over the areas ($N/3$ in each area) and the fraction p_2 moving to $area_2$, the fraction p_3 moving to $area_3$, the fraction $p_1 = 1 - p_2 - p_3$ staying in $area_1$. Then $p_1 MS(t)$ mosquitoes will be in $area_1$, $p_2 MS(t)$ mosquitoes will be in $area_2$, $p_3 MS(t)$ mosquitoes will be in $area_3$. The population fraction of marked mosquitoes at time t in area i is then:

$$\pi_i(t) = \frac{p_i MS_m(t)}{\frac{1}{3}N + p_i MS_m(t)} = \frac{3p_i MS_m(t)}{N + 3p_i MS_m(t)} \quad [18]$$

and the ratio of the probability to capture a marked mosquito to the probability to capture an unmarked mosquito becomes:

$$\frac{\pi_i}{1 - \pi_{i1}} = \frac{3p_i MS_m(t)}{N} \quad [19]$$

The logistic regression model with time and distance effect is:

$$\log\left(\frac{\pi_i(t)}{1 - \pi_i(t)}\right) = -\log(N) - e^{\beta_0} t + \log(M) + \log(3p_i) \quad [20]$$

In the application of the model to the real data, the distance was included in the model as a factor and the total experimental area around the release point was divided in three regions of equal surface area (distance smaller than 121 m, between 121 and 171.5 m, between 171.5 and 210 m), with an exception for MRR2 in which we recognized only two areas (distance smaller than 148.5 m, between 148.5 and 210 m), because no marked mosquitoes were found in the third area. Also for the simulated data, the distance was included in the model as a factor and the area was divided in three regions with the same surface area.

The parameters are estimated using the maximum likelihood approach with the package `bbmle` in R (Bolker 2012). The parameters involved in the models are summarized in the Table 2.

Table 2. Parameters involved in the formulas

Parameters	
N	Population size per day
M	Number of marked mosquitoes released
n	Number of mosquitoes captured
m	Number of marked mosquitoes recaptured
ϕ	Survival rate
t	Time between release and capture
S_m	Fraction of marked mosquitoes that can still be trapped at time t
λ	Parameter of the survival function
α	Intercept of the logistic regression, $-\log(N)$
β_0	$\ln(\lambda)$
p_i	Fraction that moves from an area to the one farther

RESULTS

We applied to the data the logistic regression model with only the survival function (equation 15), the logistic regression model with survival function and distance effect (equation 20), and the Fisher-Ford method (equation 7). This is used as a possible comparison with one of the already existing methods; the choice is because of the fact that it is based on the same principle as our method. The same analysis also was applied to simulated data to show how the models perform in different scenarios, when the number of marked mosquitoes released, the population size and the distance effect change.

Analysis of *Ae. albopictus* data

Applying the model with the survival function, the estimated population size per day was 516 (95% CI 392-679) in MRR1, 1,435 (95% CI 1,314-1,567) in MRR2, and 8,723 (95% CI 8,369-9,093) in MMR3. The parameter of the survival function, which describes the amount of mosquitoes lost per time period, was estimated as 0.330 (SE 0.069) in MRR1, 0.223 (SE 0.047) in MRR2, 1.35E-07 (SE 68.792) in MRR3 (Table 3).

Table 3. Estimates of female *Ae. albopictus* population size in the three mark-release-recapture (MRR) experiments, day 1-9, selected part of the campus of "Sapienza" University in Rome (Italy), maximum distance 210 m

MRR	M	Method	Density per ha	N per day	95% CI for N	Width CI	λ	SE(β_0)	AIC
1	464	Model1 ^a	49	516	392 – 679	287	0.330	0.069	2549.53
		Model 2 ^b	23	244	37 – 1611	1,574	0.307	0.490	61.07
		Fisher-Ford ^c	340	3,603	2,245 – 7,219	4,974	-	-	-
		Fisher-Ford ^d	55	580	204 – 1,505	1,301	-	-	-
2	566	Model1 ^a	135	1,435	1,314 – 1,567	253	0.223	0.047	11,328.06
		Model 2 ^b	63	673	232 – 1,953	1,721	0.254	0.680	95.08
		Fisher-Ford ^c	294	3,115	1,932 – 5,451	3,519	-	-	-
		Fisher-Ford ^d	36	379	103 – 919	816	-	-	-
3	552	Model1 ^a	823	8,723	8,369 – 9,093	724	1.35E-07	68.792	13,224.78
		Model 2 ^b	625	6,625	2,880 – 15,238	12,358	3.76E-06	174.820	89.53
		Fisher-Ford ^c	692	7,334	4,661 – 13,282	8,621	-	-	-
		Fisher-Ford ^d	139	1,472	827 – 2,988	2,161	-	-	-

Density is expressed as the number of female per hectare, N is the estimate of the population size per day, λ is the parameter of the survival function estimated with the logistic regression. The daily survival rate used for the method of Fisher-Ford is 0.9886 (Alto and Juliano 2001) and a more likely value for the field of 0.85; the confidence interval is based on 500 bootstrap replicates with the method of percentile bootstrap at 95% level. M=marked mosquitoes released.

$$^a \text{Equation 15: } \log\left(\frac{\pi(t)}{1 - \pi(t)}\right) = -\log(N) - e^{\beta_0} t + \log(M).$$

$$^b \text{Equation 20: } \log\left(\frac{\pi_i(t)}{1 - \pi_i(t)}\right) = -\log(N) - e^{\beta_0} t + \log(M) + \log(3p_i).$$

^c Fisher and Ford with daily survival rate 0.9886 (Alto and Juliano 2001) and correction for low recapture.

^d Fisher and Ford with daily survival rate 0.85 and correction for small recaptures.

AIC = Akaike Information Criterion.

Applying the model with the survival function and the distance effect, the estimated population size per day was 244 (95% CI 37-1,611) in MRR1, 673 (95% CI 232-1,953) in MRR2, and 6,625 (95% CI 2,880-15,238) in MRR3. The parameter of the survival function was estimated as 0.307 (SE 0.490) in MRR1, 0.254 (SE 0.680) in MRR2, 3.76E-06 (SE 174.820) in MRR3 (Table 3).

It is important to note that the estimates of λ s for MRR3 are very small but are also based on very limited data with very low recapture in the experiment. This limits the capability of the model to estimate the parameters, as we can see from the huge standard error (for both models).

The loss survival functions for each experiment and model are shown in Figure 2.

The population size estimated with Fisher-Ford method is very sensitive for the parameter that we choose for the mortality rate. In MRR1 is 3,603 with $\phi=0.9886$ (95% CI 2,245-7,219) and 580 with $\phi=0.85$ (95% CI 204-1,505); in MRR2 is 3,115 with $\phi=0.9886$ (95% CI 1,932-5,451) and 379 with $\phi=0.85$ (95% CI 103-919); in MRR3 is 7,334 with $\phi=0.9886$ (95% CI 4,661-13,282), 1,472 with $\phi=0.85$ (95% CI 827 – 2,988).

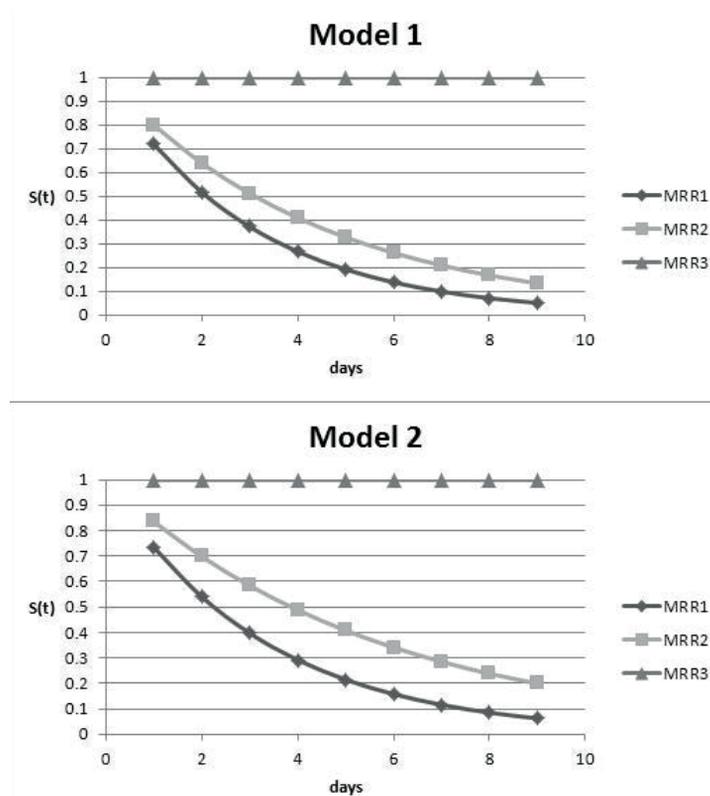


Figure 2. Loss function in models 1 and 2 for each experiment.

Analysis on the simulated data

In all the simulated scenarios and with both logistic regression models (equation 15 and 20), we obtained an estimate of N very close to the predefined size of our simulated population. We present the median value of the estimate (Table 4), because the distribution of the estimates is skewed (Supp Figure 1). The performance of both models is similar when the marked mosquitoes

Table 4. Application of the model on simulated data: $N=21,000$; survival rate=0.95; number of simulation=1,000; except for scenario1 where the number of simulations is 516

Scenario	M in area 1	M in area 2	M in area 3	Model	Median N	First Q	Third Q	IQR	Median AIC
1	50	50	50	1	20095	15401	26854	11453	55.46
				2	20080	15370	26820	11450	57.16
				FF ^a	15914	14129	19622	5493	-
2	105	30	15	1	20597	14890	26874	11984	58.26
				2	20500	14780	26820	12040	50.61
				FF ^a	15914	13942	19622	5680	-
3	170	170	170	1	20439	17022	24879	7857	91.75
				2	20430	16980	24830	7850	93.67
				FF	16507	14987	17955	2968	-
4	357	102	51	1	20675	17075	24411	7336	113.1
				2	20360	16670	23990	7320	84.08
				FF	16507	14987	17955	2968	-
5	335	335	335	1	19903	17370	22704	5334	110.97
				2	19900	17340	22690	5350	113.26
				FF	15927	14830	16997	2167	-
6	704	201	100	1	19990	17605	22920	5315	162.4
				2	19410	17060	22300	5240	102.68
				FF	15761	14830	17188	2358	-

M=marked mosquitoes released.

$$\text{Model 1: } \log\left(\frac{\pi(t)}{1-\pi(t)}\right) = -\log(N) - e^{\beta_0}t + \log(M).$$

$$\text{Model 2: } \log\left(\frac{\pi_i(t)}{1-\pi_i(t)}\right) = -\log(N) - e^{\beta_0}t + \log(M) + \log(3p_i).$$

First Q and third Q are the first and third quartile.

IQR is the interquartile range.

AIC = Akaike Information Criterion.

^a FF: Fisher-Ford method with correction for low recapture (equation 7).

are equally distributed over the areas. In scenarios 1, 3, and 5 the AICs (we report the median value because the AIC has also a skewed distribution) differ $< 2 U$. The model with the distance performs better in the scenarios 2, 4, and 6, when there is a relationship between the distance from the release site and the number of marked mosquitoes caught. In all the scenarios, the precision of the estimates, indicated by the interquartile range, does not substantially change.

In all the simulated scenarios, the true value of the population size is never present in the interquartile range of the estimates produced by Fisher-Ford method. The estimates are more precise, but they lack accuracy.

DISCUSSION

We applied a logistic regression model to estimate the population size of *Ae. albopictus* on the campus of “Sapienza” University in Rome, using data from mark-release-recapture experiments (Marini et al. 2010). We also tested the performances of the model under different simulated scenarios.

The logistic regression model we present is built to accommodate most mosquito MRR protocols where population size is estimated from a single release, captures are carried out on several days after the release and traps are distributed in an area at different distances from the release site. Besides providing an estimate of the population size, the model allowed us to estimate a combined survival and dispersal (or “retention”) factor from the data and to consider the time needed to reach traps situated farther from the release point. This means that it is not necessary to provide a preset value for survival taken from the literature (laboratory experiments or other field experiments). Such values will usually not be a good match with the situation in one’s own experiment. For example, the value we chose for the Fisher-Ford model, based on the only laboratory experimental evidence we could find, was markedly different from the values we estimated directly from the data of the MMR experiments. In our design, we do not have a preset value, but rather a preset *distribution* for the survival. In our analyses we chose an exponential distribution, but this can be changed to another distribution, should more information about the survival function of mosquitoes be available for the field situation one is studying. Taking into account the distance between release site and trap may be important in an experimental setup. We would expect to find more mosquitoes in the traps near the release point in the first days and then a spread to a wider area. For traps at the borders of the natural flight range, one would expect that too few marked individuals are caught to satisfy the (implicit) assumption in our approach that the naturally occurring (unmarked) population and the marked population are well-mixed. If we do not consider the distance effect, we could base our estimate on a too low recapture rate, because of the difficulty the mosquitoes had to reach the traps and because of the chance effects deviating from a well-mixed situation.

We evaluated whether the distance correction really improved the estimates, comparing a model with only time effect (model 1) and a model with time effect and distance effect (model 2) for each experiment. We observed that the introduction of the distance effect improved the model substantially, as shown by the AIC. Second, the width of the confidence intervals shows how the estimates became less precise, with a substantially larger variability. To draw more robust conclusions about the performance of the models, we examined the simulated data, because the real data did not have many observations. The estimates related to the simulated scenarios, appeared very similar for models 1 and 2. They were always close to the real value of 21,000 mosquitoes. The interquartile range assumed similar values in both models. The AIC shows a small improvement in the scenarios in which the number of mosquitoes decreases when the distance increases. The difference was negligible when the marked mosquitoes were distributed equally among the areas.

Examining the number of mosquitoes estimated by the models from the real data, one may be surprised by the relatively low estimates, considering what is known about the abundance of *Ae. albopictus* in Rome (Severini et al. 2008), and Sapienza campus area, in particular (B. C., A. D. T., F. M., personal communication). It is, however, important to remark that the estimates are based on the number of marked mosquitoes recaptured, which also depends on the characteristics of the method used to trap the mosquitoes in the experimental setup. Because the sticky trap attracts mainly gravid female mosquitoes, and all marked mosquitoes were released at the fed stage and presumably recaptured during oviposition (see Marini et al. 2010), it may be more appropriate to consider that our estimates refer mainly to this fraction of the total population. This could then explain why in the third experiment, carried out in October, we found a larger number of mosquitoes in the traps: in fact, at the end of the reproductive season it may be expected that no freshly emerged mosquito are produced and that, a higher fraction of the surviving female mosquitoes is looking for a place to lay eggs.

The logistic regression model proposed is able to include specific features of MRR experiments carried out with mosquitoes. It can be applied whenever there is interest in monitoring the population size of mosquitoes or other insects with mark-recapture experiments. To improve and further validate the model, additional experimental data from the field are needed. This would help to understand which variables are most important and the kind of relationship they have with the outcome. One could then also use the model in combination with simulated data to explore experimental set-ups, comparing different numbers of traps, at different distances and spatial distributions from a release point, and the influence of releasing fewer or more marked individuals. We have used simulated data to at least provide some validity of the approach, showing that the model is able to estimate preset population size with sufficient accuracy to serve as a basis to estimate vector to host ratio's accurately enough for (local) R_0 estimates.

ACKNOWLEDGMENTS

This study was funded by EU grants FP7-261504 EDENext and is catalogued by the EDENext Steering Committee as EDENext040 (<http://www.edenext.eu>). The contents of this publication are the sole responsibility of the authors and do not necessarily reflect the views of the European Commission.

REFERENCES

- Alto, B. W., and S. A. Juliano. 2001. Temperature effects on the dynamics of *Aedes albopictus* (Diptera: Culicidae) populations in the laboratory. *J. Med. Entomol.* 38: 548 -556.
- Bailey, N.T.J. 1951. On estimating the size of mobile populations from recapture data. *Biometrika* 38: 293-306.
- Bailey, N.T.J. 1952. Improvements in the interpretation of recapture data. *J. Anim. Ecol.* 21: 120 -127.
- Begon, M. 1979. Investigating animal abundance: capture-recapture for biologists. Arnold, Edward, London, United Kingdom.
- Bolker, B., and RDC Team. 2012. bbmle: tools for general maximum likelihood estimation, <http://cran.r-project.org/web/packages/bbmle/index.html>.
- Chapman, D. G. 1951. Some properties of the hypergeometric distribution with applications to zoological sample censuses. University of California Publications in Statistics, Berkeley, CA.
- Conway, G. R., M. Trpis, and G.A.H. McClelland. 1974. Population parameters of the mosquito *Aedes aegypti* (L.) estimated by mark-release-recapture in a suburban habitat in Tanzania. *J. Anim. Ecol.* 43: 289 -304.
- Dowdeswell, W. H., R. A. Fisher, and E. B. Ford. 1940. The quantitative study of populations in the lepidoptera I. *Polyommatus icarus* rott. *Ann. Hum. Genet.* 10: 123-136.
- Efford, M. G. 2009. Density estimation by spatially explicit capture-recapture: likelihood-based methods. Springer, New York.
- Efron, B., and R. J. Tibshirani. 1993. An introduction to the bootstrap. Chapman & Hall, London, United Kingdom.
- [ESRI] Environmental Systems Resource Institute. 2009. ArcGIS 9.2. Redlands, CA.
- Facchinelli, L., L. Valerio, M. Pombi, P. Reiter, C. Costantini, and A. Della Torre. 2007. Development of a novel sticky trap for container - Breeding mosquitoes and evaluation of its sampling properties to monitor urban populations of *Aedes albopictus*. *Med. Vet. Entomol.* 21: 183-195.
- Fisher, R. A. 1947. The spread of a gene in natural conditions in a colony of the moth *Panaxia dominula* L. *Heredity* 1: 143-174.
- Google Earth Beta. 2010. Google Earth Beta, version 5.2.1.
- Gratz, N. G. 2004. Critical review of the vector status of *Aedes albopictus*. *Med. Vet. Entomol.* 18: 215-227.
- Hartemink, N. A., B. V. Purse, R. Meiswinkel, H. E. Brown, A. De Koeijer, A.R.W. Elbers, et al. 2009. Mapping the basic reproduction number (R0) for vector-borne diseases: a case study on bluetongue virus. *Epidemics* 1: 153-161.
- Hartemink, N. A., S. O. Vanwambeke, H. Heesterbeek, D. Rogers, D. Morley, B. Pesson, et al. 2011. Integrated mapping of establishment risk for emerging vector-borne infections: a case study of canine leishmaniasis in southwest France. *PLoS ONE* 6: e20817.
- Jackson, C.H.N. 1937. Some new methods in the study of *Glossina morsitans*. *Proc. Zool. Soc.* 106: 811- 896.
- Jackson, C.H.N. 1939. The analysis of an animal population. *J. Anim. Ecol.* 8: 238 -246.
- Jackson, C.H.N. 1940. The analysis of a tsetse fly population. *Ann. Eugen.* 10: 332-369.
- Krebs, C. J. 1999. Ecological methodology. H and R Publishers, New York.
- Le Cren, E. D. 1965. A note on the history of mark-recapture population estimates. *J. Anim. Ecol.* 34: 453-454.
- Lincoln, F. C. 1930. Calculating waterfowl abundance on the basis of banding returns. *Circ. U.S. Dep. Agric.* 118.

- Macdonald, G. 1952. The analysis of the sporozoite rate. *Trop. Dis. Bull.* 49: 569-586.
- MacLeod, J. 1958. The estimation of numbers of mobile insects from low incidence recapture data. *Ecol. Entomol.* 110: 363-392.
- Marini, F., B. Caputo, M. Pombi, G. Tarsitani, and A. Della Torre. 2010. Study of *Aedes albopictus* dispersal in Rome, Italy, using sticky traps in mark-release-recapture experiments. *Med. Vet. Entomol.* 24: 361-368.
- Nelson, R. L. 1978. Estimates of survival, population size, and emergence of *Culex tarsalis* at an isolated site. *Ann. Entomol. Soc. Am.* 71: 801-808.
- Pollock, K. H. 1980. Capture-recapture models: a review of current methods, assumptions and experimental design. North Carolina State University, Raleigh, NC.
- Pollock, K. H., J. D. Nichols, C. Brownie, and J. E. Hines. 1990. Statistical inference for capture-recapture experiments. *Wildl. Monogr.* 107: 1-97.
- Reisen W. K., F. Mahmood, and T. Parveen. 1979. *Anopheles subpictus* Grassi: observations on survivorship and population size using mark-release-recapture and dissection methods. *Res. Popul. Ecol.* 21: 12-29.
- Reisen, W. K., F. Mahmood, and T. Parveen. 1980. *Anopheles culicifacies* Giles: a release-recapture experiment with cohorts of known age with implications for malaria epidemiology and genetical control in Pakistan. *Trans. R. Soc. Trop. Med. Hyg.* 74: 307-317.
- Schwarz, C. J., and G.A.F. Seber. 1999. Estimating animal abundance: review III. *Stat. Sci.* 14: 427-456.
- Seber, G.A.F. 1982. The estimation of animal abundance and related parameters. Griffin, London, United Kingdom.
- Seber, G.A.F. 1986. A review of estimating animal abundance. *Biometrics* 42: 267-292.
- Severini, F., M. Di Luca, L. Toma, and R. Romi. 2008. *Aedes albopictus* in Rome: results and perspectives after 10 years of monitoring. *Parassitologia* 50: 121-123.
- White, G. C., D. R. Anderson, K. P. Burnham, and D. L. Otis. 1982. Capture-recapture and removal methods for sampling closed populations. Los Alamos National Laboratory, Los Alamos, NM.
- White, G. C. 1999. Program MARK: survival estimation from populations of marked animals. *Bird Study* 46: 120-139.

SUPPLEMENTARY MATERIAL

R code used to generate the simulated data and parameters involved. We present the scenarios 3 and 4 as an example.

List of parameters

a0: Intercept of the model
 b0: Parameter involved in the survival function
 c2, c3: Parameter involved in the distance effect
 ti: Time in days
 d1, d2, d3: Areas
 x: Number of marked mosquitoes in the trap
 n: Number of mosquitoes in the trap
 M: Number of mosquitoes released
 N: Population size

```
library(bbmle)

#####
# Scenario 3
#
# M=510, N=21000
# 33.3% in area 1 => N=7000, M=170
# 33.3% moves to area 2 => N=7000, M=170
# 33.3% moves to area 3 => N=7000, M=170
#####

minloglik <- function(a0,b0,c2,c3,ti,d1,d2,d3,x,n,M) {
  p2 <- exp(c2)/(1+exp(c2))
  p3 <- exp(c3)/(1+exp(c3))
  p1 <- 1-p2-p3
  logit<-a0-ti*exp(b0)+log(M)+d1*log(3*p1)+d2*log(3*p2)+d3*log(3*p3)
  popfrac <- exp(logit)/(1+exp(logit))
  -sum(dbinom(x, size=n, prob=popfrac, log = TRUE))
}

day <- rep(1:9,3)
dis1 <- c(rep(1,9),rep(0,9),rep(0,9))
dis2 <- c(rep(0,9),rep(1,9),rep(0,9))
dis3 <- c(rep(0,9),rep(0,9),rep(1,9))

frac <- c((170/7000)*(.95^(1:9)), (170/7000)*(.95^(1:9)),
         (170/7000)*(.95^(1:9)))

for (i in 1:1000){
  marked <- rbinom(n=rep(1,27),size=rep(100,27),prob=frac)
  fit.mod <- mle2(minuslogl=minloglik,
                 start=list(a0=7,b0=.1,c2=-1,c3=1),
                 data=list(ti=day,d1=dis1,d2=dis2,d3=dis3,x=marked,
                           n=rep(100,27),M=510),
                 method="Nelder-Mead",control=list(maxit=50000,
                                                    type=1,trace=1,reltol=10^(8)*.Machine$double.eps))
}
#####
# Scenario 4
#
# M=510, N=21000
# 70% in area 1 => N=7000, M=357
# 20% moves to area 2 => N=7000, M=102
# 10% moves to area 3 => N=7000, M=51
#####

minloglik <- function(a0,b0,c2,c3,ti,d1,d2,d3,x,n,M) {
  p2 <- exp(c2)/(1+exp(c2))
  p3 <- exp(c3)/(1+exp(c3))
  p1 <- 1-p2-p3
  logit<-a0-ti*exp(b0)+log(M)+d1*log(3*p1)+d2*log(3*p2)+d3*log(3*p3)
  popfrac <- exp(logit)/(1+exp(logit))
  -sum(dbinom(x, size=n, prob=popfrac, log = TRUE))
}

day <- rep(1:9,3)
dis1 <- c(rep(1,9),rep(0,9),rep(0,9))
dis2 <- c(rep(0,9),rep(1,9),rep(0,9))
```

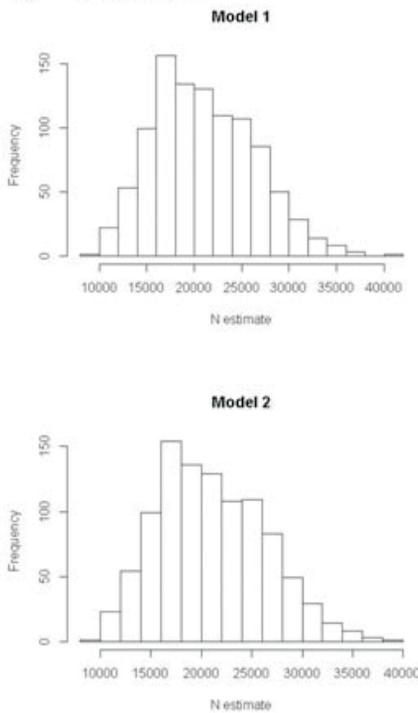
```

dis3 <- c(rep(0,9),rep(0,9),rep(1,9))

frac <- c((357/7000)*(.95^(1:9)),(102/7000)*(.95^(1:9)),
(51/7000)*(.95^(1:9)))
for (i in 1:1000){
marked <- rbinom(n=rep(1,27),size=rep(100,27),prob=frac)
fit.mod <- mle2(minuslogl=minloglik,
start=list(a0=7,b0=.1,c2=-1,c3=-1),
data=list(ti=day,d1=dis1,d2=dis2,d3=dis3,
x=marked,n=rep(100,27),M=510),
method="Nelder-Mead",control=list(maxit=50000,
type=1,trace=1, reltol=10^(8)*.Machine$double.eps))
}
#####

```

a) Scenario 3



b) Scenario 4

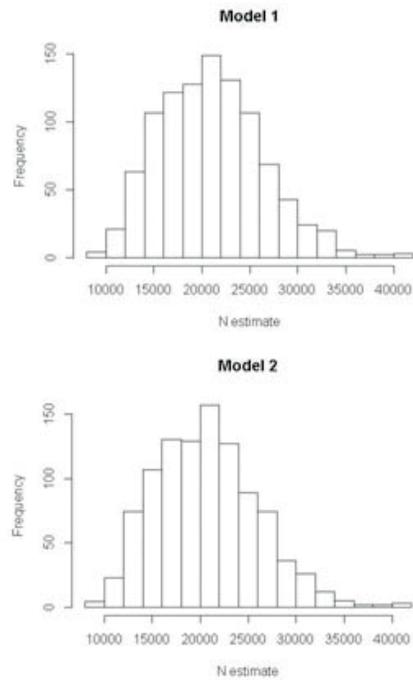


Figure 1. Histograms of the population size estimated in scenarios 3 and 4 with both methods (models 1 and 2).

CHAPTER 6

Characterizing mosquito breeding sites in an urban setting: a case study in Amsterdam

Jan Buijs ¹, Daniela Cianci ², Sander Koenraad ³, Arjan Stroo, Chantal Vogels ^{3*}

¹ GGD Amsterdam, Amsterdam, The Netherlands

² Faculty of Veterinary Medicine, Utrecht University, Utrecht, The Netherlands

³ Laboratory of Entomology, Wageningen University and Research centre, P.O. box 16, 6700 AA, Wageningen, The Netherlands

⁴ Centre for Monitoring of Vectors, Food and Consumer Product Safety Authority, Wageningen, The Netherlands

* Authors list in alphabetical order

Work in progress

ABSTRACT

Mosquitoes can be vectors of a large number of pathogens and their bites can cause a considerable nuisance. When a sufficient number of competent vectors come in contact with hosts, an outbreak can occur after the introduction of a pathogen. Amsterdam is a densely populated urban area in the Netherlands and a central point for many national and international transport routes. In this paper we present the results of the first city-wide survey in Amsterdam that aims at increasing the knowledge on composition, geographical distribution, biodiversity and environmental preferences of native mosquito species. Also, the characteristics of their breeding sites are explored. Nine mosquito larvae species were identified during this study. The number of potential larval microhabitats found in urban green was lower than the number of larval microhabitats found in any of the other ecotope types. The extent of urban green and polder areas was negatively associated with the number of potential larval microhabitats. The chance of finding larvae (of any species) in accumulation of water after runoff is higher compared to the other larval microhabitats. No significant associations (at the 0.05 level) were found between ornithophilic and mammophilic biotype and larval microhabitat type, ecotope, land use, type of sewage system and month. The reported nuisance was significantly higher in areas where larvae were found. Knowledge on the distribution of potential native vectors that could be responsible for the introduction and the establishment of vector-borne diseases is crucial, in order to minimize the risk of outbreaks.

Keywords: mosquito distribution, larval habitats, vector-borne diseases, mosquito nuisance, city-wide survey

BACKGROUND

Mosquitoes (Diptera: Culicidae) are known to be vectors of a large number of pathogens around the globe and their bites can cause a considerable nuisance for humans. In an area with sufficient numbers of competent vectors and human or animal vertebrate hosts, even if the area is non-endemic, outbreaks can occur when a pathogen is introduced. Due to the increase in international trading, traffic and global warming, the species ranges as well as the range of possible pathogens that they can spread are changing. The potential spread of invasive mosquito species and the potential expansion to wider areas of pathogens that are transmitted by the native species are a matter of concern. Awareness of the spatial distribution of mosquito vectors in urban environments is important because cities can act as hotspots for vector-borne disease transmission given the high human population density and the relatively high temperature (Cianci et al. 2015).

Amsterdam is a densely populated urban area in the Netherlands, surrounded by polder and it is a central point for many national and international transport routes. It represents, therefore, an area where potential vectors and/or pathogens could easily be introduced and where they could come in contact with humans. Only few studies on the distribution of potential vector species in the Netherlands have been conducted (Ibañez-Justicia et al. 2015, Vogels et al., manuscript in preparation). Most of these studies, in addition, were nation-wide surveys, and had only few sampled locations in urban settings, and specifically few in Amsterdam. Reusken et al. (2010) conducted a study in the summer of 2009 in the Amsterdam metro system, but the focus was on only one mosquito species, *Culex pipiens*. No city-wide surveys aiming at increasing the knowledge on composition, geographical distribution, biodiversity and environmental preferences of mosquito species have been carried out in Amsterdam so far.

In 2014, the public health service of Amsterdam (GGD) started a survey in collaboration with the Dutch Centre for Monitoring of Vectors (CMV) and Wageningen University (WUR) in order to determine the mosquito distributions in Amsterdam. In this survey, the target species are native mosquito species (Culicidae). The main objectives were (i) to identify larval habitats suitable as mosquito breeding sites, (ii) to identify ecological factors that determine a suitable microhabitat for these species and (iii) to investigate differences in microhabitat types of different biotypes of the *Culex pipiens* complex.

In this study, characteristics of the breeding sites and their surrounding environmental factors are described. Also, the species found in the study, the differences in larval microhabitats between species and nuisance reports are presented. Identifying and mapping the different mosquito species and biotypes can give more insight in the behaviour and in the distribution of mosquitoes.

MATERIAL AND METHODS

Study area and mosquito sampling

From May until July 2014, the GGD conducted a survey on potential larval development habitats by sampling 88 locations in Amsterdam. These locations were distributed over five different ecotopes: old residential areas (built before 1950), new residential areas (built after 1950), urban green space, suburban rural area and industrial area (figure 1a). They were distinguished according to the definition provided by the urban ecologists of Amsterdam (van der Halm 2001). The old residential areas are located mainly in the centre of Amsterdam. The new residential, suburban rural and industrial areas are situated in the periphery of Amsterdam, while urban green space is scattered throughout the whole city. The sampled locations were randomly selected within the ecotopes. Each ecotope was sampled with different intensity; in old residential areas and urban green space one location every 1 km² was selected, in new residential areas one location every 2 km² was selected and in suburban rural area and industrial area one location every 4 km² was selected. The sampled locations are visualised in figure 1a. A circular area with a radius of 50 meters around the selected locations was inspected for the presence of larval microhabitats (figure 1b). All the visible and accessible larval microhabitats were identified and possible mosquito larvae were collected. Fourth instar larvae were identified to the species level following Becker et al. (2010). Morphological identifications of a part of the samples were supplemented by molecular

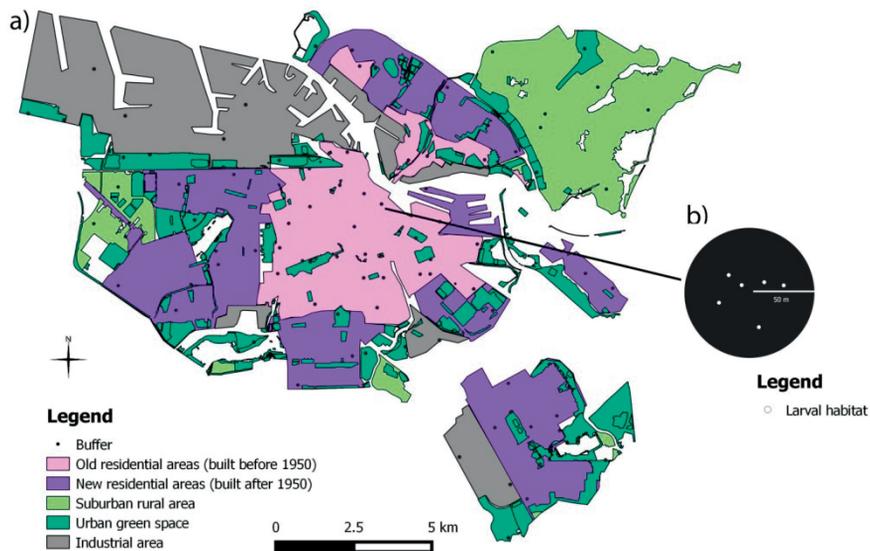


Figure 1. a) Subdivision of Amsterdam into ecotopes and b) a sampled location (buffer) with potential larval microhabitats.

analyses with a real-time PCR assay (Vogels et al., manuscript in preparation), in order to identify *Culex* species and the biotypes within the *Culex pipiens* complex.

Environmental factors

The environmental factors studied in relation with the larval microhabitats were ecotopes, land use classes, type of sewage system and distance to it, human population density and reported nuisance. The five ecotopes were, as indicated above, old residential areas, new residential areas, urban green space, suburban rural area and industrial area. The land use classes were urban green, streets and car parks, water, and buildings. The classes were obtained with the topographic map of Amsterdam (KBK10, geographical service Amsterdam). Three types of sewer systems were considered in the analysis: a separate system that conveys only storm water, a separate system that conveys only wastewater, and a combined system where pipes convey domestic sewage, industrial wastewater and storm water runoff. The distance to each sampled location and the closest sewage system was calculated in QGIS (*v.distance* command in GRASS) (QGIS Development Team 2012). The population density per hectare of each neighbourhood was also included in the analysis. The source of the environmental variables is reported in table 1. Additionally, when possible, up to 5 private addresses per sampling area (buffer) were visited. The inhabitants of the houses were interviewed and it was reported if they experienced mosquito nuisance in the last year. If they had a garden and they gave permission, it was inspected for mosquito larvae.

Table 1. Environmental factors used in the models and their sources

Variable	Source
Larval microhabitat type	Field observations classified by the definitions made by the Centre for Monitoring of Vectors and GGD
Ecotope	Classified according to the city ecologists' definition and validated with building ages of housing blocks (van der Halm 2001)
Land use	KBK10, Geographical department Amsterdam [www.stelsepedia.amsterdam.nl]
Inhabitants	CBS and statistical department of Amsterdam [www.os.amsterdam.nl]
Type of sewage system	Waternet [www.waternet.nl]

Statistical analysis

For each buffer, a nuisance ratio was calculated dividing the number of nuisance cases reported by the total number of interviews. ANOVA was performed to investigate whether there were differences in the mean of nuisance ratio in the buffers with and without larvae.

Regression models for count data were applied to model the relationship between the number of potential breeding sites in each buffer and the ecotope type and the percentage of land use type. The proportion of the buffer that could be inspected is included in the model as offset. A Poisson model and a negative binomial model were compared, and the best model was chosen

based on the Akaike information criterion (AIC). The best model was defined according to the commonly used rule-of-thumb that a model is better than another model when its AIC is at least 2 units smaller (Burnham and Anderson 2004).

An additional analysis was run to investigate the relationship between presence and absence of larvae in the potential larval microhabitat and the characteristics surrounding the breeding sites. For this purpose a mixed effects logistic regression model was used, with larval microhabitat type, ecotope, land use type, type of sewage system, distance to the closest sewage system, population density and sampling month as predictors. The buffer was included in the model as random effect.

In both the regression model for count data and the logistic regression model, models with all possible combinations of variables were compared and the best model was selected based on the AIC values.

According to the biotype, the species were classified in ornithophilic (i.e., *Cx. pipiens pipiens* and *Cx. torrentium*) and mammophilic (i.e., *Cx. pipiens molestus* and hybrids between *Cx. pipiens pipiens* and *molestus*). A Fisher's exact test was performed to investigate the association between biotypes and larval microhabitat type, ecotope, land use, type of sewage system and month.

The statistical analyses were run in R (R Core Team 2013). The variable selection procedure was carried out with the R package MuMIn.

RESULTS

In the residential area 191 potential larval microhabitats were found (103 in the old residential area and 88 in the new residential area). In the urban green spaces 67 potential larval microhabitats were sampled and fewer potential larval microhabitats were found in suburban rural area (33) and industrial area (30). Of the total of 321 potential larval microhabitats, 164 were found in an area that was classified as the land use class 'street', 93 were found in green areas, 52 in proximity to water and 12 close to buildings. The proportion of the buffer that was possible to inspect varied between ecotope: it was easier to sample urban green space and suburban rural areas than residential and industrial areas (figure 2).

Nine mosquito larvae species were identified during this study (table 2). The highest diversity of mosquito species was observed in urban green and suburban rural areas (polder), but the number of larval microhabitats with larvae was lower compared to residential areas. In residential areas a high number of *Culex pipiens* and *Culex torrentium* was found (figure 3).

The breeding sites identified were storm drains, rainwater in various containers, rainwater in boats, storm drains mixed with sewer water, ditches with vegetation, ditches in polder, rainwater in tires, mixed water in tractor tracks, and mixed water in drinking containers for livestock. They were grouped in four categories: accumulation of water after a runoff on a surface (172), direct accumulation of rainwater (56), water in a canal, pond or ditch (64) and mix of rainwater with sew-

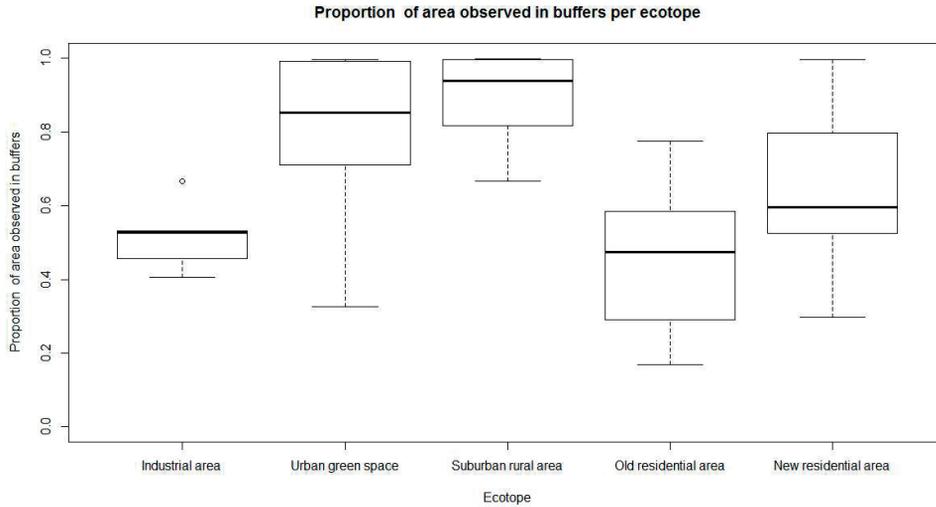


Figure 2. Proportion of area in the buffer for each ecotope that was accessible for inspection.

Table 2. Species identified with a brief description and number of presence locations.

Species	Biotype	Characteristics*	Potential vector of	Presence of larvae
<i>Anopheles claviger</i>		Clean water in shady places	Malaria	1
<i>Anopheles maculipennis</i>		Diversity of standing water habitats with vegetation	Malaria	4
<i>Culiseta annulata</i>		All kinds of standing water with high content of nitrogen		3
<i>Culex pipiens</i>			West Nile Virus	92**
	Pipiens	Rainwater ornithophilic		44
	Molestus	Rainwater and dirty water mammophilic		11
	Hybrid	Rainwater and dirty water mammophilic		13
<i>Culex sp</i> ^b .				22
<i>Culex territans</i>		Ponds, swamps, vegetated streams		1
<i>Culex torrentium</i>		Accumulation in containers with Rainwater ornithophilic	West Nile virus	58

*from Becker et al. 2010

**It was not possible to determine the biotype of all *Cx. pipiens*

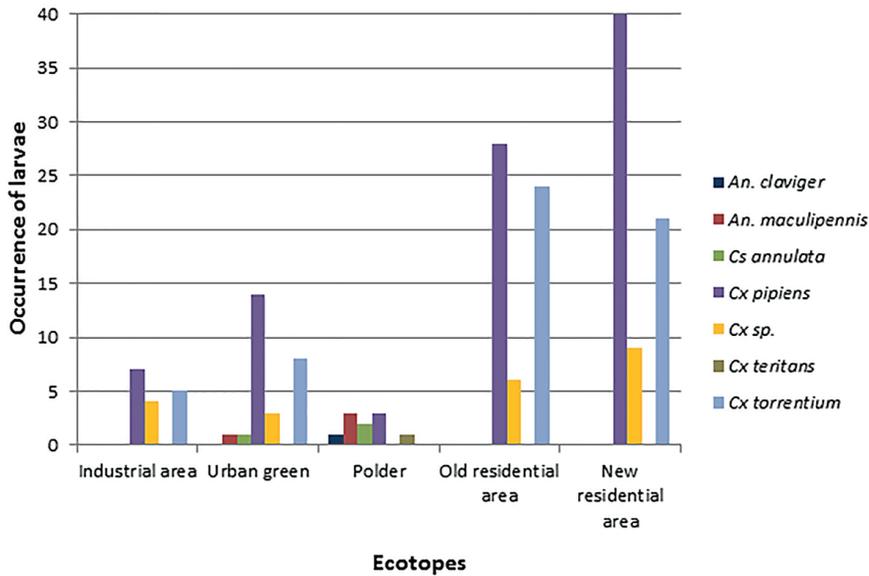


Figure 3. Number of larval microhabitats with presence of larvae of different species per ecotope.

Table 3. Number of presences of larvae per species found in each larval microhabitat type.

Species	Rainwater drainage	Rainwater accumulation	Surface water	Water mix
<i>Anopheles claviger</i>	0	1	0	0
<i>Anopheles maculipennis</i>	0	0	3	1
<i>Culiseta annulata</i>	1	1	1	0
<i>Culex pipiens</i>	69	15	3	5
<i>Culex sp.</i>	13	7	0	2
<i>Culex teritans</i>	0	1	0	0
<i>Culex torrentium</i>	45	11	1	1

age water or water from canals and other possible sources of water (26). In table 3 the occurrence of larvae for each species in each larval microhabitat is reported.

In total 60 people, in all the ecotopes, were interviewed and they were asked whether they experienced nuisance caused by mosquitoes. Actual bites and buzzing throughout the night were considered nuisance. Nuisance was reported by 39 of the interviewed persons. Nuisance was often reported in buffers with many larvae presences, as shown in figure 4. In buffers where larvae were found, 69% of the interviewed persons experienced nuisance, while in buffers where no larvae were found, only 38% of the interviewed persons experienced nuisance. The nuisance

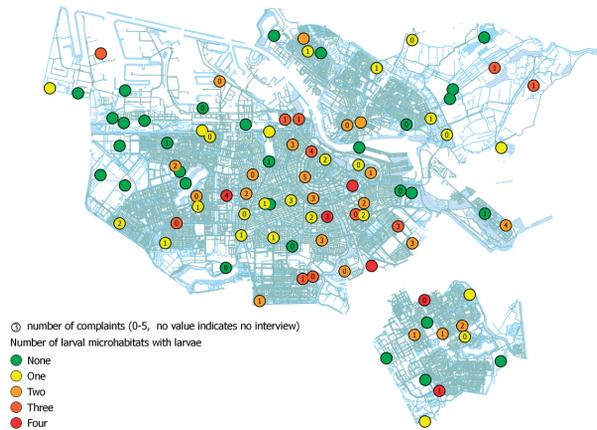


Figure 4. Number of nuisance reports and number of larval microhabitats with larvae within a buffer. The circles are the buffers, the numbers indicate the complaints per buffer, the colour indicate the number of larval microhabitat with larvae per buffer (green corresponds to no larval microhabitat and red to four larval microhabitats).

ratio was significantly different between buffers with and without larvae (one-way ANOVA, p -value=0.02).

Because land use and ecotope provide similar information (Supp figure 1), it was not possible to include both variables in the same model. One model was run to study the effect of land use on the number of potential larval microhabitats and another model was run to study the effect of the ecotope on the number of potential larval microhabitats. The negative binomial regression model was preferred over the Poisson regression model, because it had a lower AIC. The negative binomial model estimates coefficients interpreted as incidence rate ratios (IRR), which means that when $IRR > 1$ the effect of the land use on the number of larval microhabitats is positive (i.e. when a given land use category is present, the number of larval microhabitats is higher compared to the number of larval microhabitats in the land use chosen as reference category), when $IRR < 1$ the effect of the land use on the number of larval microhabitats is negative (i.e. when a given land use category is present, the number of larval microhabitats is lower compared to the number of larval microhabitats in the land use chosen as reference category). The same interpretation holds when the land use is replaced by the ecotope.

After correcting for the area of the buffer that was accessible for inspection, the number of larval microhabitats found in 1 m² in urban green was significantly lower than the number of larval microhabitats per m² found in any of the other ecotopes (table 4). In the model with land

Table 4. Negative binomial regression model with number of breeding sites per buffer as dependent variable and ecotope types as predictors.

Ecotope ^a	IRR ^b	P-value
Industrial area	1.89	0.031*
Polder area	2.20	0.012*
Old residential area	1.73	0.008*
New residential area	1.61	0.025*

^a Urban green area is the reference category for the ecotope.

^b IRR = incident rate ratio; IRR>1 positive effect, IRR<1 negative effect.

* Variables significant at 0.05 level.

Table 5. Negative binomial regression model with number of breeding sites as dependent variable and the best subset of predictor variables chosen among land use classes.

Land use ^b	IRR ^a	P-value
Green	0.978	< 0.0001*
Polder	0.992	0.0355*

^a IRR = incident rate ratio; IRR>1 positive effect, IRR<1 negative effect.

^b For the land use classes, the IRR is calculated for an increase of 100m².

* Variables significant at 0.05 level.

Table 6. Logistic regression model with presence/absence of larvae as dependent variable and the best subset of predictor variables chosen among larval microhabitat type, ecotope, land use class, sewage system types, the distance between each sampled location and the closest sewage system, population density and month.

Variable	OR ^a	P-value
June	46.20	8.3e-06*
July	76.63	2.1e-06*
Rainwater drainage ^b	0.02	2.8e-08*
Rainwater accumulation ^b	0.36	0.02733*
Water mix ^b	0.09	0.00075*

^a OR = incident rate ratio; OR>1 positive effect, OR<1 negative effect.

^b Accumulation of water after runoff is the reference category for larval microhabitat type.

* Variables significant at 0.05 level.

use classes, the number of larval microhabitats slightly decreases when the surface covered by green and polder increases (table 5).

The output of the model to investigate the relationship between the presence of larvae and favourable factors is shown in table 6. The best subset of predictor variables included larval microhabitat type and month of the sampling. The odds of finding larvae in water in canals, ponds or ditches, in direct accumulation of rainwater and in mix of water from different sources are lower than the odds of finding larvae in accumulation of water after runoff. The odds of finding larvae in June and July are higher than the odds in May. Including interaction effects between larval microhabitat types and months did not improve the model.

It was investigated whether there was an association between biotypes and larval microhabitat type, ecotope, and land use. The presence of both biotypes was observed in 18 larval microhabitats, ornithophilic biotypes were present in 24 places and mammophilic biotypes in 2 places. No significant relationships at the 0.05 level were found between ornithophilic and mammophilic biotype and larval microhabitat type, ecotope, land use, type of sewage system and month, individually.

DISCUSSION

In this study, the characteristics of the urban larval habitats of native mosquito species (Culicidae) have been investigated in Amsterdam. Four potential larval microhabitat groups were identified: 1) accumulation of water after a runoff on a surface, 2) direct accumulation of rainwater, 3) water in canal, pond or ditch and 4) mix of rainwater with sewage water or water from canals and other possible sources of water. Larvae of seven mosquito species were found in the studied area and they were not equally distributed among the four larval microhabitat categories. *Culex pipiens* was the most present species and it was mainly found in rainwater drainage; *Culex torrentium* was mainly found in rainwater drainage as well. Rainwater drainage was the category with the highest number of the larvae. Considering all the species together, the chance of finding larvae in accumulation of water after runoff is higher compared to the other larval microhabitats. The positive effect of the storm drains may have been underestimated because, given their high number, it was not possible to sample all of them. Moreover, it was not possible to sample tree holes, roofs and other potential larval microhabitats on high places that were difficult to reach. Consistently with the findings described above, most of the breeding sites were found in the land use class street, where accumulation of rainwater is more likely to be found. This result suggests that the accumulation of rainwater in storm drains and other reservoirs is a crucial factor in the development of mosquitoes in urban environments.

In urban green and suburban rural areas a high species diversity of larvae was observed. The same was observed for adults (J. Buijs, personal communication). However, in these ecotopes, the number of potential larval microhabitats with larvae was lower compared to other ecotopes.

In other words, in urban green and suburban rural areas there was a higher diversity of species, but a lower occurrence. Some species do have specific development microhabitats. For instance, *Anopheles* species are mostly found in standing water with a relatively high amount of vegetation (Becker et al. 2010). *Anopheles plumbeus* is a tree hole breeding species and is also known to live close to abandoned and un-cleaned pig stables with basins containing water with fermented organic matter (Ibañez-Justicia and Cianci 2015). On the other hand, the larvae of *Ochlerotatus caspius* are salt tolerant and are known to be present in coastal marches and fresh water with at least 0.5 g NaCl/l (Becker et al. 2010). These kinds of microhabitats are more available in green and rural areas and in water microhabitats that might have a connection with the brackish water of the IJ, the water connection between the North Sea and IJsselmeer that flows through Amsterdam. *Coquillettidia richardi* larvae develop on the stem of reed or other water vegetation (Becker et al. 2010). No larvae of this species were found, probably because of the sampling method. However, adults of this species were sometimes observed in high numbers in some areas in Amsterdam (J. Buijs, personal communication). Also, not all the private gardens were inspected. These may present a lot of smaller water bodies such as buckets with water, rain water barrels, ponds that are good larval microhabitats for *Culex pipiens*.

Nuisance was often reported in buffers where a high number of larvae was found. This suggests that also the number of adults is high in those areas. In future study it would be interesting to investigate whether a relationship between reported nuisance and number of larvae/adults exists, quantify it and use the reported nuisance to predict the mosquito distribution.

No significant relationships were found between ornithophilic and mammophilic biotype and larval microhabitat type, ecotope, land use, type of sewage system and month, individually. However, this being a pilot study, the number of places with *Culex* larvae was very low (presence of both biotypes 18; presence of only ornithophilic biotypes 24; presence of only mammophilic biotypes 2). This may be the reason why no significant relationships were found.

In this study we only focused on identifying environmental characteristics related to occurrence of larvae. Although we expect that meteorological factors will also affect the presence of larvae and perhaps the species, these factors were not taken into account, because no meteorological data at the buffer level were available.

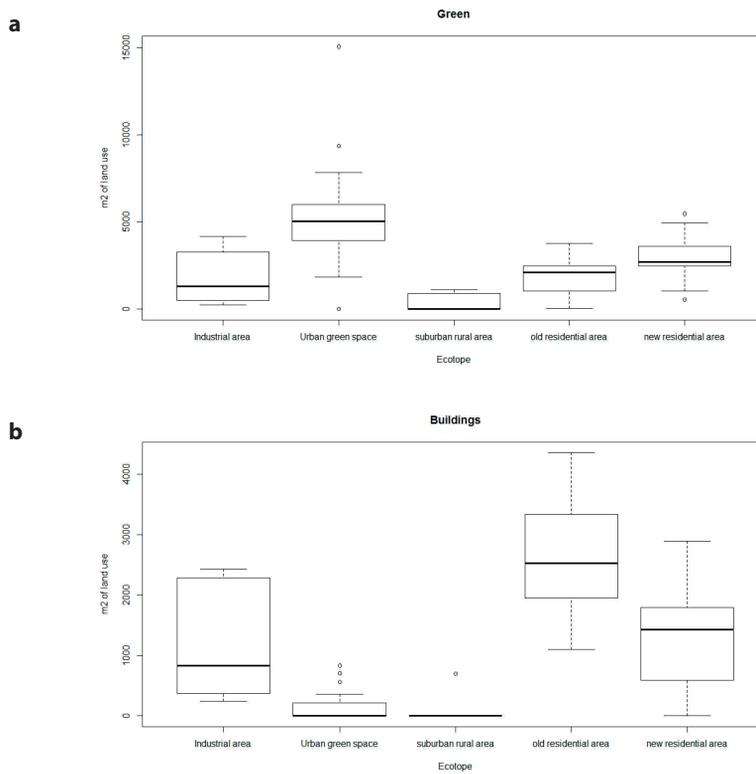
Knowledge on the distribution of native species that are potential vectors of vector-borne diseases is crucial. Identifying key characteristics of the urban mosquito habitats that are suitable for these species will give the possibility to control mosquitoes and the risk of outbreaks of vector-borne disease by reducing the numbers of suitable breeding sites or by targeting control measures specifically to areas with many breeding sites. For example, it might be possible to remove or modify factors that make the habitat suitable for these species.

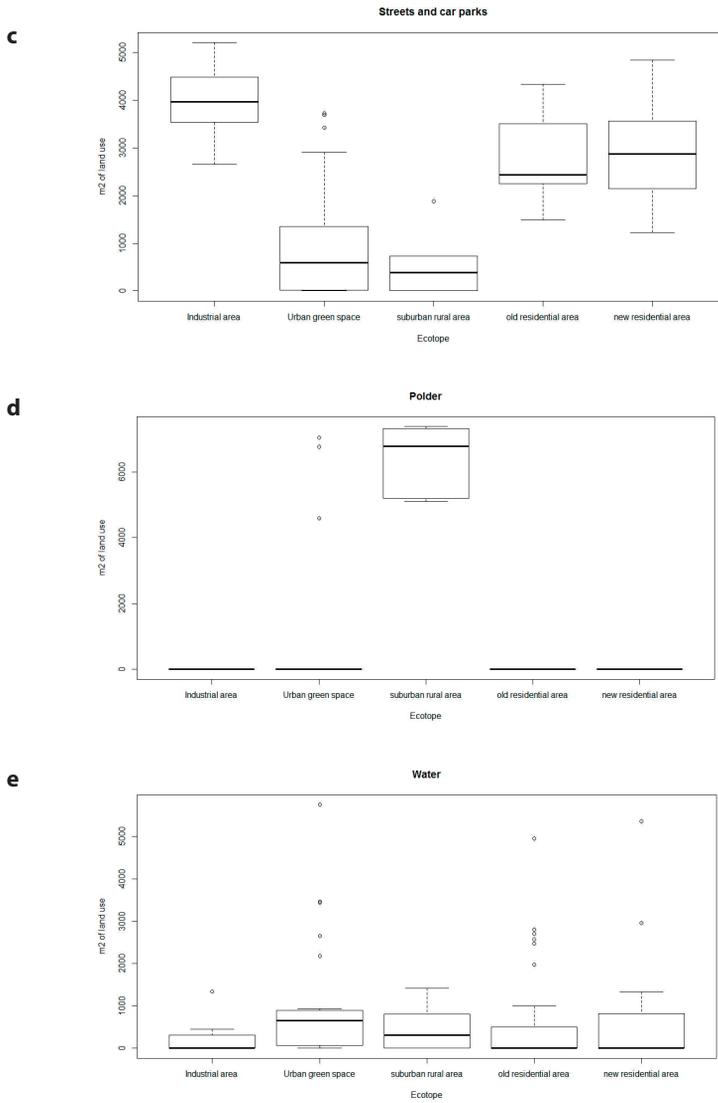
REFERENCES

- Becker, N., D. Petric, M. Zgomba, C. Boase, M. Madon, C. Dahl, et al. 2010. *Mosquitoes and Their Control*. Springer, Heidelberg, Dordrecht, New York.
- Burnham, K. P., and D. R. Anderson. 2004. Multimodel inference understanding AIC and BIC in model selection. *Sociol. Methods Res.* 33: 261–304.
- Cianci, D., N. Hartemink, C. B. Zeimes, S. O. Vanwambeke, A. Ienco, and B. Caputo. 2015. High Resolution Spatial Analysis of Habitat Preference of *Aedes Albopictus* (Diptera: Culicidae) in an Urban Environment. *J. Med. Entomol.* 52: 329–335.
- Van der Halm, H. 2001. *De wilde stad. 100 jaar natuur in Amsterdam*. KNNV Uitgeverij, Utrecht.
- Ibañez-Justicia, A., and D. Cianci. 2015. Modelling the spatial distribution of the nuisance mosquito species *Anopheles plumbeus* (Diptera: Culicidae) in the Netherlands. *Parasit. Vectors.* 8: 258.
- Ibañez-Justicia, A., A. Stroo, M. Dik, J. Beeuwkes, and E. J. Scholte. 2015. National Mosquito (Diptera, Culicidae) Survey in the Netherlands 2010–2013. *J. Med. Entomol.*
- QGIS Development Team. 2012. *QGIS Geographic Information System*.
- R Core Team. 2013. *R: A Language and Environment for Statistical Computing*.
- Reusken, C., A. De Vries, J. Buijs, M. A. H. Braks, W. Den Hartog, and E.-J. Scholte. 2010. First evidence for presence of *Culex pipiens* biotype *molestus* in the Netherlands, and of hybrid biotype *pipiens* and *molestus* in northern Europe. *J. Vector Ecol.* 35: 210–212.

SUPPLEMENTARY MATERIAL

Figure 1. The boxplots show the relationship between land use and ecotope, graphically. Each plot represents a land use class. The amount of m² of each land use class are shown per ecotope. Land use and ecotope provide similar information (land use 'green' mainly in ecotope urban green -fig. 1a; land use 'buildings' mainly present in ecotopes industrial and residential areas -fig. 1b; land use 'streets and car parks' mainly present in ecotopes industrial and residential areas -fig. 1c; land use 'polder' mainly present in ecotope suburban rural area -fig. 1d). For this reason we decided to not include them in the same model, but to run a model to study the effect of land use classes on the number of larval microhabitats and another model to study the effect of the ecotopes on the number of larval microhabitats.





- a) The land use class 'green' is mainly present in the ecotope urban green.
- b) The land use class 'buildings' is mainly present in the ecotopes industrial and residential areas.
- c) The land use class 'streets and car parks' is mainly present in the ecotopes industrial and residential areas.
- d) The land use class 'polder' is mainly present in the ecotope suburban rural area.
- e) The land use class 'water' is pretty equally distribute among the ecotopes.

CHAPTER 7

The “Auto-Dissemination” approach: A novel concept to fight *Aedes albopictus* in urban areas

Beniamino Caputo¹, Annamaria Ienco¹, Daniela Cianci², Marco Pombi¹, Vincenzo Petrarca³, Alberto Baseggio⁴, Gregor J. Devine⁵, Alessandra della Torre¹

¹ Dipartimento di Sanità Pubblica e Malattie Infettive, Università di Roma “Sapienza”, Roma, Italy

² Faculty of Veterinary Medicine, Utrecht University, Utrecht, The Netherlands

³ Dipartimento di Biologia e Biotecnologie “Charles Darwin”, Università di Roma “Sapienza”, Rome, Italy

⁴ I.N.D.I.A. INDUSTRIE CHIMICHE S.p.A, Padua, Italy,

⁵ Cairns Public Health Unit, Tropical Regional Services, Queensland Health, Cairns, Australia

PLoS Neglected Tropical Diseases 6(8): e1793 (2012);

doi: 10.1371/journal.pntd.0001793

ABSTRACT

The main constraint to the fight against container-breeding mosquito vectors of human arboviruses is the difficulty in targeting the multiplicity of larval sources, mostly represented by small man-made water containers. The aim of this work is to assess the feasibility of the “auto-dissemination” approach, already tested for *Aedes aegypti*, as a possible alternative to traditional, inefficient control tools, against *Ae. albopictus* in urban areas. The approach is based on the possibility that wild adult females, exposed to artificial resting sites contaminated with pyriproxyfen, can disseminate this juvenile hormone analogue to larval habitats, thus interfering with adult emergence. We carried out four field experiments in two areas of Rome that are typically highly infested with *Ae. albopictus*, i.e., the main cemetery and a small green area within a highly urbanised neighbourhood. In each area we used 10 pyriproxyfen “dissemination” stations, 10 “sentinel” sites and 10 covered, control sites. The sentinel and control sites each contained 25 *Ae. albopictus* larvae. These were monitored for development and adult emergence. When a 5% pyriproxyfen powder was used to contaminate the dissemination sites, we observed significantly higher mortality at the pupal stage in the sentinel sites (50–70%) than in the controls (<2%), showing that pyriproxyfen was transferred by mosquitoes into sentinel sites and that it had a lethal effect. The results support the potential feasibility of the auto-dissemination approach to control *Ae. albopictus* in urban areas. Further studies will be carried out to optimize the method and provide an effective tool to reduce the biting nuisance caused by this aggressive species and the transmission risk of diseases such as Dengue and Chikungunya. These arboviruses pose an increasing threat in Europe as *Ae. albopictus* expands its range.

BACKGROUND

Aedes albopictus (Skuse, 1894) is native to Southeast Asia. In recent decades this mosquito has invaded and efficiently colonized temperate areas of the US and Europe, thanks to passive transportation of eggs in used tires and the ability to produce diapausing eggs [1], [2]. In its tropical range the species is a secondary vector of human arboviruses, such as Dengue, while in temperate regions its impact on public health is mostly due to its aggressive and diurnal biting behavior. However, in recent years, *Ae. albopictus* has been the sole vector of large epidemics of Chikungunya virus in La Reunion, France [3], [4] and Kerala, India [5]. In August 2007 the species was responsible for a Chikungunya virus outbreak in the Province of Ravenna in north-eastern Italy where more than 250 human cases were confirmed [6] and where the mosquito has repeatedly been identified as infected by Chikungunya [7], [6], [8] and Usutu viruses [9]. In 2010 *Ae. albopictus* was responsible for 2 and 31 cases of endemic transmission of Dengue in France and China, [10], [11]. The report of a human blood index $\geq 80\%$ in Rome raises particular concern for the potential of this species to vector pathogens in urban areas where humans represent the major blood-meal source [12].

In Italy, *Aedes albopictus* was detected for the first time in Genoa in 1990 and the next year in Padua [13], [14]. In the following years it spread to 19 out of 20 Italian regions and over 82 out of 107 Italian provinces [15]. It was first detected in Rome in 1997 [16] and since then it has colonized the whole urban area through three phases “a first massive spread, a following maintenance of infestation, and the colonization of alternative winter breeding sites with favourable climatic conditions” [17]. The successful invasion of Rome and of other urban areas in Italy and worldwide is driven by the ability of this mosquito to exploit a large variety of water containers as larval breeding sites, including the catch basins of storm drains, used tires, domestic containers, vases, etc. The difficulty in identifying and treating all these sources makes control extremely difficult. A survey carried out in Rome’s zoological gardens, which estimated the larval density of mosquitoes breeding in catch basins, showed that a large number of productive alternative larval biotopes exist [18]. Despite this, the most commonly utilized strategy to reduce *Ae. albopictus* densities in Italy is the treatment of catch basins with larvicidal compounds [19]. This only targets an unknown percentage of the overall aquatic habitat. Indoor insecticidal treatments and house screening are not suitable due to the largely exophilic behavior of the species, and outdoor adulticidal treatments are recommended only in case of emergencies (e.g. virus outbreaks or very high nuisance in sensible locations such as outdoor recreational areas) due to their environmental impact and their low cost-efficiency ratio [19].

The aim of this work is to assess the feasibility of a new approach for the control of *Ae. albopictus* in urban areas, inspired by results obtained on the tropical Dengue vector, *Ae. aegypti*, in Thailand [20] and in Peru [21]. This approach, named “auto-dissemination”, is based on the possibility that wild adult mosquitoes exposed to artificial resting sites contaminated with pyriproxyfen (PPF, a juvenile hormone analogue), can disseminate insecticide to larval breeding sites, thus preventing

adult emergence. This strategy is facilitated by the oviposition behaviour of both *Ae. aegypti* and *Ae. albopictus*, that typically scatter the eggs from a single gonothrophic cycle among several temporary sites. This increases the probability of at least some larvae reaching the adult stage [22]. Second, extraordinarily low doses of PPF (*Ae. aegypti*: $LC_{50} = 0.011$ p.p.b., [23], 0.0039 p.p.b. [24]; *Ae. albopictus*: $LC_{50} = 0.11$ p.p.b. [25], are needed to interfere with the metamorphosis of juvenile stages [26], and/or to cause morpho- logical and functional aberrations in emerging adults, such as decreased fertility in males and females [27], [28]. Third, evidence from laboratory experiments shows that females either forced to walk on PPF-treated paper or topically contaminated can contaminate larval sites and significantly inhibit adult emergence [29], [30].

Thus, the “auto-dissemination” approach can be proposed as a ‘pull’ (i.e., attraction of wild mosquitoes to PPF-contaminated sites for contamination) and ‘push’ (i.e., dispersal of contaminated mosquitoes and dissemination of PPF to larval habitats) control strategy with the potential to target the myriad of cryptic larval breeding sites that cannot be reached by traditional larvicidal applications.

We here present the results of four “auto-dissemination” experiments carried out in two areas of Rome with high infestations of *Ae. albopictus*: a cemetery and an enclosed garden.

MATERIALS AND METHODS

Dissemination stations

Dissemination stations (DS, Figure 1) were adapted from modified sticky traps (ST), previously developed by our group [31]. The four sticky surfaces were replaced by four 1268 cm black cotton cloths (Figure 1b). A thick net was placed over the water to prevent mosquitoes from ovipositing (Figure 1c). Before each experiment, each DS was filled with 700 ml of tap water and each cloth was dusted with 1 g of powdered PPF. This was obtained by manually grinding 0.5% or 5% PPF tablets (Proxilar, I.N.D.I.A. Industrie Chimiche S.p.A.) to a granule average size of 40–80 micron.

Sentinel sites and control sites

Larval/pupal mortality in sentinel sites (SS) that were potentially contaminated with PPF by wild mosquitoes was compared to mortality in uncontaminated control sites (CS). Sentinel sites were assembled by inserting a 600 ml Pyrex beaker into a standard ovitrap (i.e., a black vase). This was done to allow easy decontamination after the experiments (PPF adheres to plastics, but not to glass). The border of the beaker was covered with black tape to avoid reflections. Control sites were similar to SS but closed with white nets to prevent mosquitoes from entering the beakers and transferring PPF. Each SS and CS contained 200 ml of tap water, 0.07 g of cat biscuits and 25 third-instar *Ae. albopictus* larvae. These were obtained from eggs collected from the “Sapienza” University campus by ovitraps during the weeks before the experiments. After hatching, the

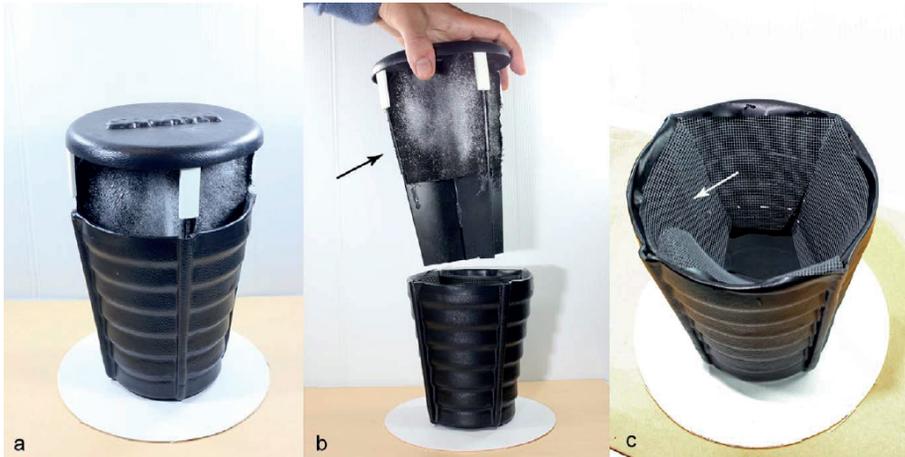


Figure 1. Dissemination station (DS) used for “auto-dissemination” experiments in Rome. **a** – whole DS; **b** – higher and lower parts of DS, where arrow indicates black cotton cloth dusted with powdered pyriproxyfen; **c** – lower part of DS, where arrow indicates the net placed above the water level to prevent mosquitoes contacting with water.

larvae were reared outdoors on the terrace of the Department of Public Health and Infectious Diseases of “Sapienza” University at a density of about 1 larva/ml.

Study areas

The experiments were carried out in the following two sites located in a highly urbanized area of Rome (Italy) adjacent to the University “Sapienza” campus.

Site 1- Verano Cemetery.

Verano is the largest (98 ha) and the oldest cemetery in Rome that is still in use. It is a green area within an urban neighbourhood less than 500 m away from Sapienza” University campus. Very high densities of *Ae. albopictus* have been reported, as is common in cemeteries due to the high number of potential larval breeding sites represented by flower vases [12]. The site chosen for the experiments (41°54’15.43”N and 12°31’34.89”) is a shallow underground crypt formed by eight 22x2 m corridors with walls of tombs on both sides. Ten DS and 10 SS were placed in a single corridor, in pairs, as shown in Figures 2a. These were about 90 cm distant from each other and 4 m from the nearest other pair. Similarly, ten DS and 10 CS were placed in a second corridor located 50 m from the first one. Experiments were carried out from 17 to 24 August 2010, with 0.5% PPF powder (Exp. 1.1) and from 10 to 22 July 2011, with 5% powder (Exp. 1.2).

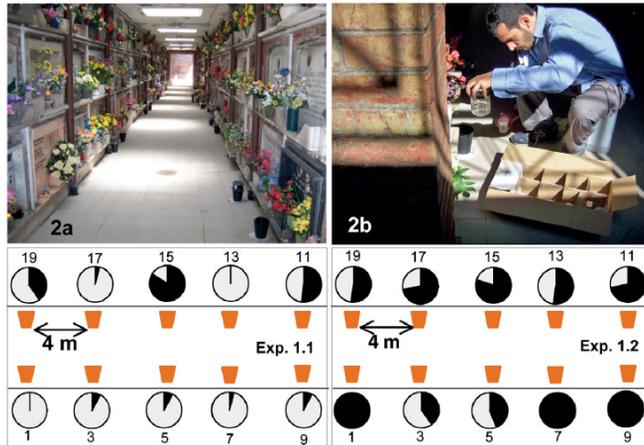


Figure 2. Ecology and results of “auto-dissemination” experiments in the Verano Cemetery in Rome (Italy). **a** - internal view of the burial corridor with dissemination stations and sentinel sites; **b** – monitoring of *Aedes albopictus* larval/pupal mortality; Exp. 1.1 and Exp. 1.2: frequencies (%) of dead larvae/pupae (in black) and of emerged adults (in white) in the first and second experiment carried out in the study area with 0.5% and 5% pyriproxyfen powder, respectively; sentinel sites are numbered accordingly to Figure 4.

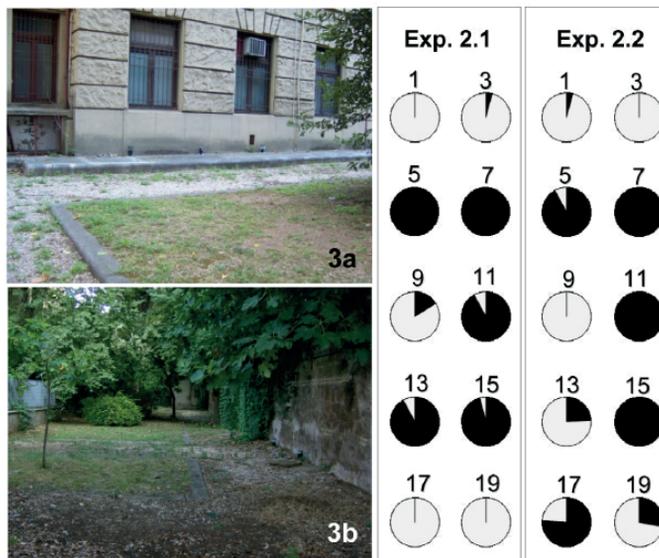


Figure 3. Ecology and results of “auto-dissemination” experiments in an enclosed garden in Rome (Italy). **a** and **b** – views of the study area with dissemination stations and sentinel and control sites; **Exp. 2.1** and **Exp. 2.2:** frequencies (%) of dead *Aedes albopictus* larvae/pupae (in black) and of emerged adults (in white) in the first and second experiment carried out in the study area with 5% pyriproxyfen powder; sentinel sites are numbered accordingly to Figure 4.

Site 2 – Enclosed garden.

The garden is part of the Institute of Anatomy and is less than 300 m from the main “Sapienza” University campus and about 850 m from the Verano cemetery (41°54'23.32"N and 12°30'57.35"E). The experiments were carried out in the 1 ha garden (Figures 3a and 3b), where 10 DSs dusted with 5% pyriproxyfen were placed at a distance of about 2 m right from an equal number of SS; 10 CS were also placed about 2 m from SS. Two replicates were carried out from 28 July to 11 August 2011 (Exp. 2.1) and from 26 August to 15 September 2011 (Exp. 2.2), respectively.

Experimental procedures

During the experiments, SS and CS were monitored every two days. For each larval cohort, we derived cumulative totals of dead larvae, dead pupae and emerged adults, as follows: i) live larvae were left to develop further; ii) dead larvae and pupae were counted and removed (Figure 2b); iii) live pupae were counted and transferred by pipette to a separate disposable cup containing uncontaminated water; these cups were covered with netting and maintained under semi-natural conditions until adult emergence or pupal death; iv) the water level in the SS and CS was maintained at 200 ml volume; v) temperature was recorded using one data logger for each corridor in Site 1, and by individual data loggers located close to each SS in Site 2. Monitoring continued until all the original 25 larvae were either dead or emerged (7–12 days).

Statistical analysis

Mortality in SS and CS was analysed by nonparametric alternatives to the t-test, since the distribution was not normal. We used the Kolmogorov-Smirnov (K-S) two-sample test (two-tailed), which is powerful even when the distributions differ in terms of dispersion. We used Wilcoxon signed-rank tests (two-tailed) to compare the mortality in SS between experiments within same field site: this is a non-parametric paired difference test that compares two related samples or repeated measures on a single sample.

Two mixed-effect logistic regression models (one per study area) were applied to quantify PPF-related mortality. Each model includes, as independent variables, the number of replicates (“replicate effect”), the treatment (i.e., potential contamination in SS by wild mosquitoes vs. no contamination in CS) and the interaction between the replicates and the treatments, which allows an evaluation of the effect on the mortality in the sentinel sites in each replicate. The code identifying each SS and each CS was included as random effect to correct for the repeated measures within sites (“site-effect”).

RESULTS

Table 1 shows the overall mortality observed in SS and CS in the four experiments. Overall mortality in SS was $\geq 50\%$ in the 3 experiments carried out with 5% PPF concentration, and 20.8% in

Table 1. *Aedes albopictus* mortality in control and sentinel sites during the experiments carried out in Rome.

PPF	T	Mortality in CSs			Mortality in SSs		
		Larvae	Pupae	Total	Larvae	Pupae	Total
0.5%	26.8°C	3	3	2.4%	3	49	20.8%
5%	23.8°C	0	3	1.2%	4	174	71.2%
5%	25.5°C	0	3	1.2%	1	124	50.0%
5%	25.8°C	1	3	1.6%	12	119	52.4%

PPF = pyriproxyfen concentration; T = mean daily temperature during experiments; Total = percentage of dead larvae/pupae over 250 third instar larvae in 10 control sites (CS) and 10 sentinel (SS), respectively.

Table 2. Results of mixed effect logistic regression analysis on data obtained in Verano Cemetery, Rome.

Variables	OR	95% CI	
Intercept	0.004	0.002	0.009
Exp 1.1 - SS vs CS	8.942	3.568	22.413
Exp 1.2 - SS vs CS	66.500	15.501	285.283
Treatment - Exp 1.2 vs Exp 1.1	3.179	0.741	13.637
No treatment - Exp 1.2 vs Exp 1.1	0.427	0.103	1.770

The model takes into account the effect of: the treatment (i.e., sentinel, SS, versus control, CS, sites), the replicates (i.e., Exp 1.1 and Exp 1.2) and the interaction between treatment and replicates on pyriproxyfen-related mortality. OR = odds ratio; CI = confidence interval. Statistically significant ($p < 0.05$) odds ratios in bold.

Table 3. Results of mixed effect logistic regression analysis on data obtained in enclosed garden, Rome.

Variables	OR	95% CI	
Intercept	0.001	0.000	0.004
Exp 2.1 - SS vs CS	48.668	8.181	289.515
Exp 2.2 - SS vs CS	37.248	7.177	193.304
Treatment - Exp 2.2 vs Exp 2.1	0.742	0.143	3.851
No treatment - Exp 2.2 vs Exp 2.1	0.970	0.191	4.932

The model takes into account the effect of: the treatment (i.e., sentinel, SS, versus control, CS, sites), the replicates (i.e., Exp 2.1 and Exp 2.2) and the interaction between treatment and replicates on pyriproxyfen-related mortality. OR = odds ratio; CI = confidence interval. Statistically significant ($p < 0.05$) odds ratios in bold.

the first experiment in Site 1, with 0.5% concentration. In all experiments, >90% of the mortality recorded in SS occurred at the pupal stage. In Site 2 however, about 13% of deaths occurred dur-

ing adult emergence. This has been frequently noted in insects treated with PPF [32], but never in mosquitoes. The overall mortality in CS was $\leq 2.5\%$.

In Site 1, a single experiment (1.1) utilised the 0.5% PPF formulation and a subsequent experiment utilised the 5% formulation (1.2). No significant differences in mortality in SS vs. CS were observed using the 0.5% concentration (K-S, $p = 0.164$; Figure 2). However, at 5% concentration mortality in SS was higher than that observed in CS (K-S, $p < 0.001$; Figure 2). Mortality in SS was lower in Exp.1.1 (median = 2, $Q_1 = 1$, $Q_3 = 8$) than in Exp 1.2 (median = 18, $Q_1 = 13$, $Q_3 = 23.75$) (significant difference shown by Wilcoxon test, $p = 0.004$).

In Site 2, both experiments (2.1 and 2.2) utilised the 5% concentration. The difference in mortality in SS vs. CS was significant only in Exp. 2.2 (K-S, $p = 0.015$; Figure 3), although the trend was confirmed in Exp. 2.1 (K-S, $p = 0.055$; Figure 3). No difference in mortality in SS was observed between the two replicates (Wilcoxon test, n.s.; Exp 2.1: median = 13.5, $Q_1 = 0.25$, $Q_3 = 23.75$; Exp. 2.2: median = 13, $Q_1 = 2.25$, $Q_3 = 24.5$).

The results of the mixed-effect logistic regression model showed that mortality was always significantly higher in SS than in CS. In Site 1, mortality was 9- and 66.5-fold higher in SS than in

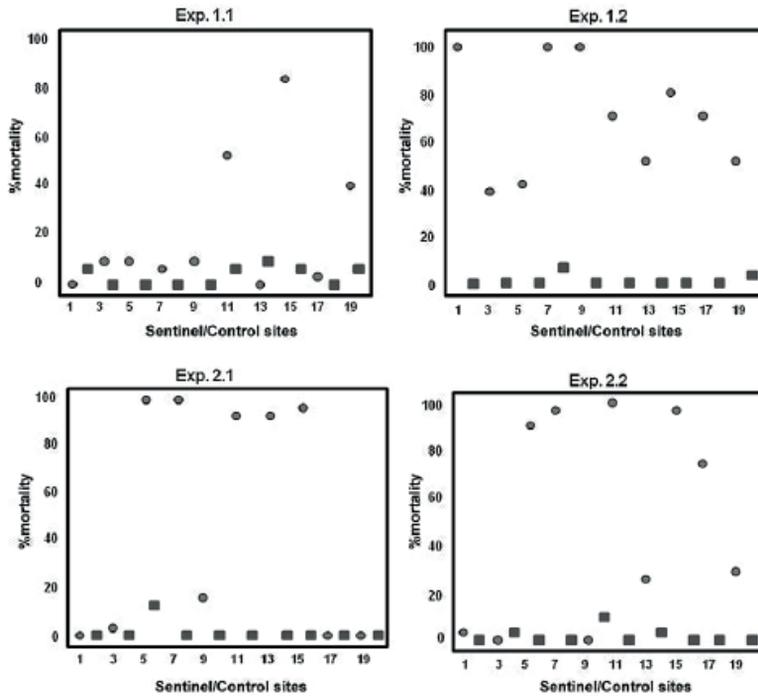


Figure 4. Overall mortality (%) in sentinel and control sites during the four experiments carried out in Rome. Experiment 1.1: 0,5% pyriproxyfen concentration; experiments 1.2, 2.1, 2.2: 5% pyriproxyfen concentration. Circles: sentinel sites numbered accordingly to Figures 2 and 3. Squares: control sites.

CS- in Exp.1.1 and in Exp 1.2, respectively, although mortality observed in SS was approximately 3 times lower in Exp 1.1 than in Exp 1.2 (Table 2). In Site 2, mortality was 49- and 37-fold higher in SS than in CS in Exp.2.1 and Exp 2.2, respectively (Table 3). Mortality was similar in the two replicates and no significant differences were found in the interaction between “replicate” and “treatment”.

Figure 4 shows mortality in SS in the four experiments. In Exp 1.1 mortality was concentrated in 3 out of 10 SS, while in Exp 1.2 it ranged between 40 and 100%. In both Site 2 experiments mortality >76% was observed in 5 out of 10 SS, 4 of which were located in the same position in both replicates.

DISCUSSION

A number of experiments carried out under laboratory conditions have shown that adult *Ae. albopictus* topically contaminated with PPF can transfer enough material to water containers to exert a significant lethal effect on pre-imaginal stages developing therein [23], [20], [29], [30]. For the first time we tested the hypothesis that wild *Ae. albopictus* adults can act as “auto-disseminators” of PPF and inhibit adult emergence from sentinel sites. Although it was not possible to measure PPF concentration in those sites (the low doses of the product are beneath the limits of detection for any published method), the evidence supports the working hypothesis and suggests that auto-dissemination could represent a valid, novel approach for reducing *Ae. albopictus* densities in urban temperate areas.

We observed significantly higher mortality in our sentinel sites than in the controls. This shows that PPF was transferred by mosquitoes into sentinel sites and elicited a lethal effect. Mortality was almost exclusively limited to the pupal stage, i.e., the stage on which PPF is known to have its major effect [26]. Note that control sites were exposed to exactly the same experimental conditions as sentinel sites, except for the fact that contact with potentially contaminated mosquitoes was prevented by a net cover in the controls.

Mortality was not uniformly distributed among sentinel sites. This strongly suggests that PPF contamination occurred at some sites, but not others - presumably as a result of differences in the frequency of visits made by contaminated mosquitoes. This is particularly evident in Site 2, where mortality was observed in 5 out of 10 SSs, 4 of which were located at the same position in both replicates (Figures 3 and 4), leading to hypothesize that these sites were more attractive for the mosquitoes than the remaining ones. The same applies for the first replicate in Site 1. In the second replicate, however, mortality ranging from 40 to 100% was observed in all SSs suggesting that these were all visited and contaminated but with differing frequency.

In the cemetery experiments, a 10-fold increase in PPF concentration resulted in a 3-fold increase in mortality (Table 2).

Median mortality in sentinel-sites, in the three experiments carried out with a 5% PPF concentration, was higher in Site 1 than in Site 2 (18 and 13 deaths/25 larvae, respectively). In the

mixed-effect logistic regression model, the variance due to the “site effect” was 0.12 (Standard deviation = 0.35) and 1.74 (Standard Deviation = 1.32), respectively. This is the likely consequence of the fact that, as mentioned above, in the Site 2 most mortality was observed only in 50% of the sentinel sites, while in Site 1 40–100% mortality was observed in every sentinel site. This presumably reflects the different ecology of the two experimental areas. Site 1, being an underground corridor, is ecologically very homogeneous and it is reasonable to hypothesize that all sentinel sites were equally attractive to the flying wild mosquitoes. The enclosed garden in Site 2 was far more heterogeneous and sentinel sites were located outdoors in a wider area compared to Site 1. It is reasonable to hypothesize that under these conditions, sentinel sites vary in attractiveness depending on contrast against background, shade, humidity, etc. However, even under such heterogeneous conditions the overall mortality in sentinel sites was about 37–49 fold higher than in controls. It is also important to emphasize that the experiments were not designed with the aim of identifying environmental factors that might optimise impact, although the effect of sun exposure has already been hypothesized in auto-dissemination experiments carried out against *Ae. aegypti* in Peru [21]. In fact, it is relevant to highlight that the major/minor attractiveness of sentinel sites to mosquitoes is likely to affect the results of the experiments, but has a relatively low practical relevance. In fact, if the approach is applied to reduce *Ae. albopictus* adult densities, the mosquitoes themselves will disseminate PPF to the most attractive (i.e., most productive) natural breeding sites. Further studies are required to assess the ideal conditions for the location of dissemination stations, as well as to evaluate the persistence of the lethal effect under different environmental (e.g. temperatures, sun exposures) and ecological (e.g. air flow, presence of animals, of humans, other disturbances/attractions) variables.

In the only other auto-dissemination experiment ever carried out in the field outdoors (i.e., in Peru with a 0.5% PPF concentration [21]), the overall reductions in *Ae. aegypti* adult emergence was 49–84%, as opposed to a 7–8% mortality in controls. In the single experiment carried out in Rome with the same PPF concentration, the overall reduction in *Ae. albopictus* adult emergence was 20.8%, as opposed to a 2.4% mortality in controls. This is quite encouraging: in fact a lower overall effect could have been expected in our experiments, since PPF LC50 reported for *Ae. albopictus* (0,11 ppb [25]) is about 10 times higher than that reported for *Ae. aegypti* (0.011 ppb [23]; 0.0039 ppb [24]). It is also worth noting that in Rome control sites were monitored concurrently to sentinel sites and were located at the same distance from dissemination stations, while in Peru they were separated in time. Thus, our results rule out the hypothesis that PPF could be passively transported by wind. If this had occurred, we would have expected higher mortality in at least some control sites.

Overall, the results from these small-scale experiments carried out in Rome strongly encourage further studies to evaluate the feasibility of the exploitation of the auto-dissemination approach to control *Ae. albopictus* densities in urban areas. In fact, the experimental design probably underestimates the overall impact of the approach, as other known effects of PPF on mosquitoes – such

as sterilizing effects on adult females [28] and a decrease in male spermiogenesis [27] - were not taken into consideration. Moreover, the effect of auto-disseminated PPF was only monitored on third instar larvae (or later stages), while in natural breeding sites younger larvae will also be contaminated, presumably increasing the overall impact. The auto-dissemination approach has the potential to effectively counter the main challenge to conventional larviciding approaches by effectively targeting the myriad of cryptic breeding sites that these mosquitoes utilise. Based on results from mark-release-recapture experiments carried out in the campus of Sapienza University in Rome [33], it is possible to hypothesize that gravid, PPF-contaminated *Ae. albopictus* females could contaminate breeding sites in a 200-m radius area around a dissemination station. Other relevant advantages of the auto-dissemination approach are: i) a higher residual activity of PPF (4 months in water [34]) compared to that of other compounds commonly used for larval control and ii) no risk for human health, due to the low-toxicity of the product for vertebrates and the high sensitivity of mosquito larvae/pupae [35]. With regard to this latter point, although PPF is effective against many insects, the proposed approach targets container-breeding species with such tiny amounts of compound, disseminated exclusively to their breeding sites, that impacts on non-target species are likely to be minimal. Finally, the auto-dissemination approach could be a very cost-effective control tool. Once deployed, the dissemination station does not require frequent maintenance, nor frequent toxicant applications, thanks to PPF's outstanding persistence [28]. It is also of note that other Culicidae species (e.g. *Culex pipiens*, a common nuisance species in urban Italy, sharing some *Ae. albopictus* breeding sites) may contribute to the dissemination of the product [36] and be affected by it themselves. Indeed, larger-scale field experiments are needed to evaluate more precisely the worldwide feasibility of the approach and to promote its use against *Ae. albopictus*.

With appropriate modifications to the dissemination stations, auto-dissemination tools may be simple enough to be deployed by members of the public. In the experiments carried out in Rome, PPF dissemination stations were adapted from existing sticky-traps for collecting *Ae. albopictus* and *Ae. aegypti* females [31], [37]. The sticky surfaces were replaced with cloth surfaces dusted with pulverized PPF. In Peru, dissemination stations consisted of 1-liter plastic disposable pots lined with damp black cloth. These dissemination tools could be optimized to increase their attractiveness for mosquitoes and their overall practicability. The fabric of the cloths could be optimized for the release of PPF to mosquito legs, and the method of application of the compound into the cloth could be optimized and standardized. Other methods to attract and contaminate adult mosquitoes could be developed (see, for instance, [30]). Finally, the auto-dissemination approach could be exploited to spread other mosquito toxic compounds, such as other juvenile hormones or fungi, which might have an even greater impact than PPF [38], [39], [40], [41].

ACKNOWLEDGMENTS

We are grateful to the personnel of AMA-Cimiteri Capitolini and of the Unit of Comparative Anatomy of the Department of Biology and Biotechnology of Sapienza University for hosting our experiments and for helping with the logistics. We thank I.N.D.I.A. INDUSTRIE CHIMICHE S.p.A (Padua, Italy) and Dr. Elisabetta Gianecchini for technical support in the preparation of pyriproxyfen talcum powder used in the experiments. This study was funded by EU grant GOCE-2003-010284 EDENext and is catalogued by EDENext Steering Committee as EDENext 035. The contents of this publication are the sole responsibility of the authors and don't necessarily reflect the views of the European Commission. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

REFERENCES

1. Hawley WA (1988) The biology of *Aedes albopictus*. *Journal of the American Mosquito Control Association Supplement* 1: 1.
2. Scholte EJ, Schaffner F (2007) 14. Waiting for the tiger: establishment and spread of the *Aedes albopictus* mosquito in Europe. *Emerging pests and vector-borne diseases in Europe* 1: 241.
3. Renault P, Solet J-L, Sissoko D, Balleydier E, Larrieu S, et al. (2007) A major epidemic of Chikungunya virus infection on Reunion Island, France, 2005–2006. *American Journal of Tropical Medicine & Hygiene* 77: 727–731.
4. De Lamballerie X, Leroy E, Charrel RN, Tsetsarkin K, Higgs S, et al. (2008) Chikungunya virus adapts to tiger mosquito via evolutionary convergence: a sign of things to come. *Virology Journal* 5: 422X–5.
5. Kumar NP, Joseph R, Kamaraj T, Jambulingam P (2008) A226V mutation in virus during the 2007 Chikungunya outbreak in Kerala, India. *Journal of General Virology* 89: 1945–1948. doi:10.1099/vir.0.83628-0.
6. Rezza G, Nicoletti L, Angelini R, Romi R, Finarelli AC, et al. (2007) Infection with Chikungunya virus in Italy: an outbreak in a temperate region. *Lancet* 370: 1840–1846. doi:10.1016/S0140-6736(07)61779-6.
7. Beltrame A, Angheben A, Bisoffi Z, Monteiro G, Marocco S, et al. (2007) Imported Chikungunya infection, Italy. *Emerging Infectious Diseases* 13: 1264.
8. Calzolari M, Bonilauri P, Bellini R, Caimi M, Defilippo F, et al. (2010) Arboviral survey of mosquitoes in two northern Italian regions in 2007 and 2008. *Vector Borne & Zoonotic Diseases* 10: 875–884. doi:10.1089/vbz.2009.0176.
9. Calzolari M, Bonilauri P, Bellini R, Albieri A, Defilippo F, et al. (2010) Evidence of Simultaneous Circulation of West Nile and Usutu Viruses in Mosquitoes Sampled in Emilia-Romagna Region (Italy) in 2009. *PLoS ONE* 5: e14324. doi:10.1371/journal.pone.0014324.
10. La Ruche G, Souare`S Y, Armengaud A, Peloux-Petiot F, Delaunay P, et al. (2010) First two autochthonous Dengue virus infections in metropolitan France, September 2010. *Euro Surveillance* 15: 19676.
11. Peng HJ, Lai HB, Zhang QL, Xu BY, Zhang H, et al. (2012) A local outbreak of Dengue caused by an imported case in Dongguan China. *BMC Public Health* 12: 83.
12. Valerio L, Marini F, Bongiorno G, Facchinelli L, Pombi M, et al. (2010) Host-feeding patterns of *Aedes albopictus* (Diptera: Culicidae) in urban and rural contexts within Rome Province, Italy. *Vector-Borne & Zoonotic Diseases* 10: 291–294.
13. Sabatini A, Raineri V, Trovato G, Coluzzi M, others (1990) [*Aedes albopictus* in Italy and possible diffusion of the species into the Mediterranean area]. *Parassitologia* 32: 301.
14. Dalla Pozza G, Majori G (1992) First record of *Aedes albopictus* establishment in Italy. *Journal of the American Mosquito Control Association* 8: 318–320.
15. Romi R, Majori G (2008) An overview of the lesson learned in almost 20 years of fight against the “tiger” mosquito. *Parassitologia* 50: 117–119.
16. Romi R, Di Luca M, Majori G (1999) Current status of *Aedes albopictus* and *Aedes atropalpus* in Italy. *Journal of the American Mosquito Control Association* 15: 425–427.
17. Toma L, Severini F, Di Luca M, Bella A, Romi R (2003) Seasonal patterns of oviposition and egg hatching rate of *Aedes albopictus* in Rome. *Journal of the American Mosquito Control Association* 19: 19–22.

18. Pombi M, Costantini C, della Torre A (2003) [*Aedes albopictus* (Diptera: Culicidae) in Rome: experimental study of relevant control strategy parameters]. *Parassitologia* 45: 97–102.
19. Romi R, Toma L, Severini F, Luca M, Boccolini D, et al. (2009) Guidelines for control of potential arbovirus mosquito vectors in Italy. *Rapporti ISTISAN- Istituto Superiore di Sanita`*.
20. Itoh T (1994) Utilization of Blood fed Females of *Aedes aegypti* as a Vehicle for the Transfer of the Insect Growth Regulator, Pyriproxyfen, to Larval Habitats. *Tropical Medicine* 36: 243–248.
21. Devine GJ, Perea EZ, Killeen GF, Stancil JD, Clark SJ, et al. (2009) Using adult mosquitoes to transfer insecticides to *Aedes aegypti* larval habitats. *Proceedings of the National Academy of Science USA* 106: 11530–11534. doi:10.1073/pnas.0901369106.
22. Trexler JD, Apperson CS, Schal C (1998) Laboratory and field evaluations of oviposition responses of *Aedes albopictus* and *Aedes triseriatus* (Diptera: Culicidae) to oak leaf infusions. *Journal of Medical Entomology* 35: 967–976.
23. Itoh T, Kawada H, Abe A, Eshita Y, Rongsriyam Y, et al. (1994) Utilization of bloodfed females of *Aedes aegypti* as a vehicle for the transfer of the insect growth regulator pyriproxyfen to larval habitats. *Journal of the American Mosquito Control Association* 10: 344–347.
24. Henrick C (1995) Juvenoids. *Agrochemicals from natural products*: 147–213.
25. Ali A, Nayar JK, Xue RD (1995) Comparative toxicity of selected larvicides and insect growth regulators to a Florida laboratory population of *Aedes albopictus*. *Journal of the American Mosquito Control Association* 11: 72–76.
26. Hirano M, Hatakoshi M, Kawada H, Takimoto Y (1998) Pyriproxyfen and other juvenile hormone analogues. *Reviews in Toxicology* 2: 357–394.
27. Iwanaga K, Kanda T (1988) The Effects of a Juvenile Hormone Active Oxime Ether Compound on the Metamorphosis and Reproduction of an Anopheline Vector, *Anopheles balabacensis* (Diptera: Culicidae). *Applied Entomology and Zoology* 23: 186–193.
28. Sihuincha M, Zamora-Perea E, Orellana-Rios W, Stancil JD, Lo´pez-Sifuentes V, et al. (2005) Potential use of pyriproxyfen for control of *Aedes aegypti* (Diptera: Culicidae) in Iquitos, Peru´. *Journal of Medical Entomology* 42: 620–630.
29. Dell Chism B, Apperson CS (2003) Horizontal transfer of the insect growth regulator pyriproxyfen to larval microcosms by gravid *Aedes albopictus* and *Ochlerotatus triseriatus* mosquitoes in the laboratory. *Medical and Veterinary Entomology* 17: 211–220.
30. Gaugler R, Suman D, Wang Y (2012) An autodissemination station for the transfer of an insect growth regulator to mosquito oviposition sites. *Medical and Veterinary Entomology* 26: 37–45. doi:10.1111/j.1365-2915.2011.00970.x.
31. Facchinelli L, Valerio L, Pombi M, Reiter P, Costantini C, et al. (2007) Development of a novel sticky trap for container-breeding mosquitoes and evaluation of its sampling properties to monitor urban populations of *Aedes albopictus*. *Medical and Veterinary Entomology* 21: 183–195. doi:10.1111/j.1365-2915.2007.00680.x.
32. Bagheri F, Talebi K, Hosseinaveh V (2011) Biological effects of the juvenoid pyriproxyfen on the pistachio green stink bug, *Brachynema germari* Kol. (Hem.:Pentatomidae). *Archives ff Phytopathology and Plant Protection* 44: 1273–1284. doi:10.1080/03235408.2010.493749.

33. Marini F, Caputo B, Pombi M, Tarsitani G, della Torre A (2010) Study of *Aedes albopictus* dispersal in Rome, Italy, using sticky traps in mark-release-recapture experiments. *Medical and Veterinary Entomology* 24: 361–368. doi:10.1111/j.1365-2915.2010.00898.x.
34. Seccacini E, Lucia A, Harburguer L, Zerba E, Licastro S, et al. (2008) Effectiveness of pyriproxyfen and diflubenzuron formulations as larvicides against *Aedes aegypti*. *Journal of American Mosquito Control Association* 24: 398–403.
35. World Health Organization (2007) Pyriproxyfen in Drinking-water: Use for Vector Control in Drinking-water Sources and Containers.
36. Devine GJ, Killeen JF (2010) The potential of a new larviciding method for the control of malaria vectors. *Malaria Journal* 9: 142.
37. Facchinelli L, Koenraadt CJM, Fanello C, Kijchalao U, Valerio L, et al. (2008) Evaluation of a Sticky Trap for Collecting *Aedes* (*Stegomyia*) Adults in a Dengue-Endemic Area in Thailand. *American Journal of Tropical Medicine & Hygiene* 78: 904–909.
38. Becnel JJ, Johnson MA (2000) Impact of *Edhazardia aedis* (Microsporidia: Culicosporidae) on a Seminatural Population of *Aedes aegypti* (Diptera: Culicidae). *Biological Control* 18: 39–48. doi:10.1006/bcon.1999.0805.
39. Scholte E-J, Knols BG, Takken W (2004) Autodissemination of the entomopathogenic fungus *Metarhizium anisopliae* amongst adults of the malaria vector *Anopheles gambiae* s.s. *Malaria Journal* 3: 45. doi:10.1186/1475-2875-3-45.
40. Reyes-Villanueva F, Garza-Hernandez JA, Garcia-Munguia AM, Tamez- Guerra P, Howard AF, Rodriguez-Perez MA (2011) Dissemination of *Metarhizium anisopliae* of low and high virulence by mating behavior in *Aedes aegypti*. *Parasite & Vectors* 4: 171. doi:10.1186/1756-3305-4-171.
41. García-Munguía AM, Garza-Hernández JA, Rebollar-Tellez EA, Rodríguez-Pérez MA, Reyes-Villanueva F (2011) Transmission of *Beauveria bassiana* from male to female *Aedes aegypti* mosquitoes. *Parasite & Vectors* 4: 24. doi:10.1186/1756-3305-4-24.

CHAPTER 8

Summarising discussion

A large number of algorithms have been used in species distribution modelling. In this thesis we applied regression and machine learning methods at different geographical scales. In the first part, the analysis was conducted on indigenous mosquito species (vectors and non-vectors) observed at a national level. In the second part, the focus shifted to urban and suburban areas, and to two potential vector species, i.e., *Aedes albopictus* and *Culex pipiens*. In the seventh chapter a novel concept to control *Ae. albopictus* on a suburban area is studied.

MODELLING THE POTENTIAL SPATIAL DISTRIBUTION OF MOSQUITO SPECIES AT LOW RESOLUTION

Non-linear discriminant analysis, random forest models and generalised linear models are used in **Chapter 2** to study the potential spatial distribution in the Netherlands for three indigenous mosquito species (*Culiseta annulata*, *Anopheles claviger* and *Ochlerotatus punctor*). Mosquito occurrence data and 1 km² resolution environmental and meteorological images were used in the models. Most of the variables highlighted by the models as important are in agreement with field experience and existing biological knowledge. The maps produced with the three different modelling techniques showed consistent spatial patterns for each species, but differences in the ranges of the predictions. Non-linear discriminant analysis gave lower predicted probabilities of presence than the other two methods. A random forest model was the model with the best classification performance for all the species. This result is consistent with earlier findings where random forest models have been reported to outperform other traditional data modelling techniques (Cutler et al. 2007, Peters et al. 2007).

It is important to remember that to create a reliable occurrence model it is generally considered necessary to have the same numbers of presence and absence points as input (Fielding and Bell 1997, McPherson et al. 2004). For techniques such as random forest modelling it is necessary to select a 'balanced' subset of the data in order to have the same number of presence and absence points. In the data preparation it is also important to avoid having presence and absence points in the same pixel, otherwise the model cannot make a prediction in that pixel. For this reason,

when a presence and an absence point were in the same square kilometre only the presence point was selected because presences inform about the places that are environmentally suitable for a species, but absences do not necessarily indicate the opposite (Lobo et al. 2010). In the study in **Chapter 2**, some areas were excluded from the sampling scheme, because they were deemed unsuitable for mosquitoes. At the modelling stage, this had to be corrected by adding absence points in these un-sampled regions. In our study, this adjustment was possible because it was known that in the un-sampled regions there were no mosquitoes. Generally speaking, avoiding bias in sampling strategies is more advisable than correcting for lack of data at the modelling stage.

Based on the results of the **Chapter 2** and on results found in literature (Cutler et al. 2007, Peters et al. 2007), in **Chapter 3**, a random forest model was chosen among several species distribution modelling techniques to investigate the relationship between the occurrence of the nuisance mosquito species and the potential malaria vector *Anopheles plumbeus* and environmental features. A random forest model was used also to investigate the relationship between the abundance of this species with environmental features. As in **Chapter 2**, the environmental and meteorological predictors were 1 km² resolution satellite-derived images. Using these data in the models, occurrence and abundance distribution maps in the Netherlands were produced. The models produced plausible maps of the actual distribution; the areas identified by the model as suitable and with high abundance of *An. plumbeus*, are consistent with the areas from which nuisance was reported.

As in **Chapter 2**, it was necessary to select a 'balanced' subset of the data in order to have the same numbers of presence and absence points and to remove the absence points when they were in the same pixels of presence points. Besides the occurrence model, in **Chapter 3** also an abundance model was run and this required an appropriate data preparation. Since each location was sampled only once and each trap was active for one week, it was difficult to compare abundance values across the sites. Because the aim of the model was to show areas where mosquito peaks could be expected, only the data collected in months in which high number of mosquitoes were observed (June–September) were used in the analysis. In this way the abundance values across the sites were made comparable and the predicted abundance can be interpreted as the expected maximum number of mosquitoes caught in a trap in a certain pixel, in accordance with the aim of the study. Another possible way to collect data for this purpose is to sample all locations in the same period, but this would either severely reduce the number of locations monitored or drastically increase the costs of the survey.

MODELLING THE POTENTIAL SPATIAL DISTRIBUTION OF MOSQUITO SPECIES AT HIGH RESOLUTION

The second part of this thesis aimed at investigating the spatial distribution of mosquito species, but now at a much higher resolution; the study area is moved from the national to the urban and suburban level. In **Chapters 4** and **5**, the distribution in terms of abundance of a nuisance species and potential vector of several human pathogens, *Aedes albopictus*, was studied in a suburban area.

Chapter 4 represents a first step toward a better understanding of the spatial distribution of *Ae. albopictus* in urban environments. Spatial statistical models were used to evaluate the relationship between *Ae. albopictus* egg abundances and land cover types on the campus of Sapienza University in Rome. The study area consisted of a mixture of built-up areas and vegetation and may be considered as representative of many areas of the city. It is here taken as an example of a European urban habitat. The land cover types were classified in detail on a high resolution image. Solar radiation and month of capture were also included in the model. Vegetation and solar radiation were found to be positively related to the number of eggs. In particular, areas with trees were associated with higher numbers of eggs probably because their shadow provides a favourable habitat for mosquitoes. Grass was associated with lower number of eggs indicating that an open area without high vegetation and shelter is less attractive for mosquitoes. The positive effect of solar radiation suggests that warm areas are favourable for oviposition. Here, solar radiation does not represent the amount of sunshine on the trap because the ovitraps were always put in shaded sites and never exposed directly to sunlight.

In **Chapter 4** the campus of Sapienza University in Rome is taken as an example of a European urban habitat. The average of the minimum distance between traps is about 40 meters and this is possible due to the small study area (approximately 22 hectares). However, although the campus is quite representative of an urban area (with a mixture of buildings, streets and some green areas), it is still a small area, which does not include all possible land cover classes present in European cities. It would be interesting to investigate the relationship between *Ae. albopictus* egg abundances and land cover types in a bigger area, including information on more land cover types, presence of water (e.g., in manholes or other containers) and human density. In this small study area, it was possible to have a high trap density. However, the data were not collected for this specific purpose and the study design was taken from another study that had a different aim. The number of traps and their position were not specifically chosen. If the study would be repeated in a larger area, the trap density will probably be reduced; otherwise the costs will also substantially increase. A possible method to calculate the minimum number of ovitraps needed to obtain a prefixed precision level is presented by Carrieri et al. (2011). Furthermore, the location of the traps could be chosen in order to include more land cover types, such as water bodies, different trees species, streets, suburban rural areas.

In another study, in the same area, the abundance of *Ae. albopictus* has been estimated using data from mark-release-recapture (MRR) experiments (**Chapter 5**). A logistic regression model was built to accommodate mosquito MRR protocols where population size is estimated from a single release, captures are carried out on several days after the release and traps are distributed in an area at different distances from the release site. The model allowed estimation of the population size, as well as a combined survival and dispersal factor, obtained from the data. It also takes into account the time needed to reach traps situated farther from the release point, without the need to provide values for survival taken from the literature or laboratory or field experiments. Taking into account the distance between release site and trap is important because we expect to find most mosquitoes in the traps near the release point in the first days and then a spread to a wider area. Neglecting the distance effect could result in estimates biased by the low recapture rate, which is due to the difficulty that the mosquitoes had to reach the traps, and to the violated assumption of well-mixed marked and unmarked population. The importance of having the distance correction in the model was also tested comparing a model with only time effect and a model with time and distance effect. The estimates improved in the model with the distance correction. Because the real data did not have many observations, the performance of the models was assessed comparing different simulated scenarios. The size of the mosquito population was set for different scenarios and it was shown that the models were able to estimate values close to the pre-defined values.

An additional study that aims at giving insight into urban mosquito ecology is presented in **Chapter 6**. Here the characteristics of the breeding sites for native mosquito species (Culicidae) have been investigated in Amsterdam. Larvae of nine species were found. *Culex pipiens* was the most present species and it was mainly found in rainwater drainage. The potential larval habitats were mainly found in the land use class street, where accumulation of rainwater is easy to be found. In urban green and suburban rural areas a high diversity of species was observed, but the numbers of observed larvae were comparatively low. The extent of urban green and polder areas was negative related to the number of potential larval habitats. The number of potential larval habitats found in urban green was lower than the number of larval habitats found in any of the other ecotope type. The chance of finding larvae (of any species) in accumulation of water after runoff was higher compared to the other larval habitats. According to the biotype, the *Culex* species were classified in ornithophilic and mammophilic and it was tested whether there was an association between biotypes and larval habitat type, ecotope, land use, type of sewage system and month. No significant associations were found. The reported nuisance was significantly higher in areas with larvae.

Although the aim of this study was to gain insight into the diversity of breeding sites of all the native mosquito species in Amsterdam, it was not always possible to examine the entire area selected for the sampling. For instance, tree holes, roofs, other potential microhabitats on high places and most of the private properties were not sampled because it was difficult to reach them.

Also, it was not possible to sample all storm drains given their high number. The impossibility to sample specific larval habitats could result in an underestimate of the presence of some species or in not finding species with specific -not sampled- habitat. A possible solution to this issue would be to use another sampling method. The transect method, for example, might be more appropriate in this context, since it will probably provide the possibility to visit a wider range of microhabitats in less time (Krebs 1999).

CONTROL MEASURES

A better understanding of the spatial distribution of vectors is needed for help in successful prevention and control activities. In **Chapter 7**, the feasibility of the “auto-dissemination” approach as a possible alternative to traditional control tools against *Ae. albopictus* in urban areas is assessed. The approach is based on the possibility that wild adult females, exposed to artificial resting sites contaminated with pyriproxyfen, can disseminate this juvenile hormone analogue to larval habitats, thus interfering with adult emergence. The results support the potential feasibility of the auto-dissemination approach to control *Ae. albopictus* in urban temperate areas. It is also known that other Culicidae species may contribute to the dissemination of pyriproxyfen and could also be affected by it. The results from this experiment carried out in a small area strongly encourage further studies to evaluate the feasibility of this approach to bigger areas.

APPLICABILITY

The output of modelling methods presented in this thesis can, for example, be used as input for epidemiological models of vector-borne diseases (Heesterbeek et al. 2015). In recent years, predictive modelling of species distribution has been broadly used to address various issues in ecology, biogeography, evolution, conservation biology and climate change research (Guisan and Thuiller 2005). These models are useful tools to describe species distributions and have become an important and widely used decision making tool for a variety of biogeographical applications, such as mapping risk of vector-borne disease spread, and determining locations that are potentially susceptible to invasion (Miller 2010). Species distribution modelling using flexible machine learning approaches has been successfully applied to quantify and to map the global distribution of hosts (Alexander et al. 2014), disease vectors (Sinka et al. 2012), pathogens (Bhatt et al. 2013), and infection and outbreak risk (Fischer et al. 2013).

Knowledge on the spatial distribution -in terms of occurrence and abundance- of vectors may help in designing control measures, such as removing favourable breeding sites or making favourable habitats less attractive or accessible. Also, identifying areas of high vector abundance is important for constructing risk maps that indicate the risk for an outbreak after an introduction (Hartemink et al. 2009, 2011). Although vector presence and abundance are not the only factors

determining whether or not a pathogen can spread in an area, determining the distribution of the vector is an essential step in studying the risk of transmission of a vector-borne pathogen.

REFERENCES

- Alexander, N. S., D. Morley, J. Medlock, K. Searle, and W. Wint. 2014. A First Attempt at Modelling Roe Deer (*Capreolus capreolus*) Distributions Over Europe. *Open Heal. Data*. 2: e2.
- Bhatt, S., P. W. Gething, O. J. Brady, J. P. Messina, A. W. Farlow, C. L. Moyes, et al. 2013. The global distribution and burden of dengue. *Nature*. 496: 504–507.
- Carrieri, M., A. Albieri, P. Angelini, F. Baldacchini, C. Venturelli, S. M. Zeo, et al. 2011. Surveillance of the chikungunya vector *Aedes albopictus* (Skuse) in Emilia-Romagna (northern Italy): organizational and technical aspects of a large scale monitoring system. *J. Vector Ecol.* 36: 108–116.
- Cutler, D. R., T. C. Edwards Jr, K. H. Beard, A. Cutler, K. T. Hess, J. Gibson, et al. 2007. Random forests for classification in ecology. *Ecology*. 88: 2783–2792.
- Fielding, A. H., and J. F. Bell. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environ. Conserv.* 24: 38–49.
- Fischer, E. A. J., G.-J. Boender, G. Nodelijk, A. A. de Koeijer, and H. J. Van Roermund. 2013. The transmission potential of Rift Valley fever virus among livestock in the Netherlands: a modelling study. *Vet. Res.* 44: 58.
- Guisan, A., and W. Thuiller. 2005. Predicting species distribution: offering more than simple habitat models. *Ecol. Lett.* 8: 993–1009.
- Hartemink, N. A., B. V. Purse, R. Meiswinkel, H. E. Brown, A. De Koeijer, A. R. W. Elbers, et al. 2009. Mapping the basic reproduction number (R0) for vector-borne diseases: A case study on bluetongue virus. *Epidemics*. 1: 153–161.
- Hartemink, N., S. O. Vanwambeke, H. Heesterbeek, D. Rogers, D. Morley, B. Pesson, et al. 2011. Integrated Mapping of Establishment Risk for Emerging Vector-Borne Infections: A Case Study of Canine Leishmaniasis in Southwest France. *PLoS One*. 6: e20817.
- Heesterbeek, H., R. M. Anderson, V. Andreasen, S. Bansal, D. De Angelis, C. Dye, et al. 2015. Modeling infectious disease dynamics in the complex landscape of global health. *Science*. 347: aaa4339.
- Krebs, C. J. 1999. *Ecological methodology*. New York, USA.
- Lobo, J. M., A. Jiménez-Valverde, and J. Hortal. 2010. The uncertain nature of absences and their importance in species distribution modelling. *Ecography (Cop.)*. 33: 103–114.
- McPherson, J., W. Jetz, and D. J. Rogers. 2004. The effects of species' range sizes on the accuracy of distribution models: ecological phenomenon or statistical artefact? *J. Appl. Ecol.* 41: 811–823.
- Miller, J. 2010. Species distribution modeling. *Geogr. Compass*. 4: 490–509.
- Peters, J., B. De Baets, N. E. C. Verhoest, R. Samson, S. Degroeve, P. De Becker, et al. 2007. Random forests as a tool for ecohydrological distribution modelling. *Ecol. Modell.* 207: 304–318.
- Sinka, M. E., M. J. Bangs, S. Manguin, Y. Rubio-Palis, T. Chareonviriyaphap, M. Coetzee, et al. 2012. A global map of dominant malaria vectors. *Parasit Vectors*. 5: 69.

Samenvatting

Vector-overgedragen ziekten zijn infecties overgebracht door de beet van geïnfecteerde vectoren, zoals muggen, teken, vlooiën en vliegen. Een vector is een organisme dat ziekten of parasieten kan overbrengen naar mens, dier of plant. Vector-overgedragen ziekten vormen een bedreiging voor de wereldwijde volksgezondheid, zowel in termen van getroffen mensen als in geografische spreiding. Muggen behoren tot de belangrijkste vectoren in Europa. In dit proefschrift richten we ons op de door muggen-overgedragen ziekten. In het bijzonder zijn de muggensoorten *Anopheles*, *Aedes albopictus* en *Culex* onderzocht.

De verspreiding van de vectoren is hier onderzocht door middel van soort-verspreiding-modellen (SVM). SVM zijn analyse methoden die gegevens van de aanwezigheid of populatie-omvang van soorten met omgevingskenmerken combineren. SVM wordt hier gebruikt om door middel van statistische modellen, relaties te identificeren tussen gemeten verspreiding van een soort en omgevingsgegevens. Deze relaties zijn gebruikt om voorspellingen te maken voor de niet bemonsterde gebieden in de studieregio.

Verschillende ruimtelijke schalen zijn gebruikt in elke studie en zijn gekozen afhankelijk van het doel van de analyses en de beschikbare gegevens. Kaarten met een kleine schaal (lage resolutie) tonen minder details en beschrijven grote gebieden. Kaarten met een grote schaal (hoge resolutie) tonen meer details en beschrijven kleine gebieden.

In het eerste deel van dit proefschrift, is de analyse uitgevoerd op inheemse muggensoorten (vectoren en niet-vectoren) op nationaal niveau. In het tweede deel is de focus verschoven naar stedelijke en voorstedelijke gebieden, en twee potentiële vectorsoorten, *Aedes albopictus* en *Culex pipiens*.

In **hoofdstuk 1**, worden de thema's van dit proefschrift geïntroduceerd. In **hoofdstuk 2** zijn drie modelleringstechnieken gebruikt om de mogelijke verspreiding van drie inheemse muggensoorten in Nederland te verkennen. De modellen zijn gebruikt om een verband tussen de aanwezigheid van de soort en omgevingskenmerken te vinden en milieu geschiktheidskaarten te produceren van Nederland. De mogelijke verspreiding van een van de meest voorkomende soorten in Nederland, *Anopheles plumbeus*, is onderzocht in **hoofdstuk 3**. Deze soort geeft overlast met zijn beet en is een mogelijke malaria vector. Aanwezigheid/afwezigheid en populatie-omvang van *An. plumbeus* in Nederland worden weergegeven op verspreidingskaarten.

In de **hoofdstukken 4-6** is de verspreiding van muggensoorten onderzocht in de stedelijke gebieden, Rome (Italië) en Amsterdam (Nederland). Omdat de studiegebieden kleiner zijn dan de in de voorgaande hoofdstukken gepresenteerde studies, was het mogelijk om de omgevingen meer in detail te analyseren.

De invasieve muggen *Ae. albopictus*, algemeen bekend als de 'tjergemug', is nu aanwezig in verschillende Europese landen. *Aedes albopictus* is een bekende vector van verschillende virussen, onder andere Chikungunya en dengue virus en is ook bekend voor de overlast die zijn beet

veroorzaakt. Het voorkeursleefgebied van de tijgermug is onderzocht in **hoofdstuk 4**. Statistische modellen zijn gebruikt om inzicht te krijgen in de relatie tussen eieren van *Ae. albopictus* en bodembedekking types op de campus van de Universiteit Sapienza in Rome. Gebieden met bomen worden geassocieerd met een hoger aantal eieren en gras wordt geassocieerd met een lager aantal eieren. Dit suggereert dat schaduw een gunstige habitat voor muggen verschaft. Warmte van zonnestraling lijkt een positief effect op het aantal eieren te hebben. Voor hetzelfde studiegebied is een schatting gemaakt van de populatie-omvang van *Ae. albopictus* op basis van mark-release-recapture experimenten (**hoofdstuk 5**). Mark-release-recapture experimenten bestaan uit het uitzetten van gemarkeerde muggen en het tellen, na enkele dagen, van het aantal gemarkeerde en ongemarkeerde muggen die gevonden worden in de vallen in het studiegebied. Ons model is gebaseerd op een in de literatuur bekende methode, maar is door ons specifiek aangepast om de populatie-omvang van muggen te schatten met behulp van kenmerken van mark-release-recapture experimenten.

Resultaten van het onderzoek naar muggen in steden wordt gepresenteerd in **hoofdstuk 6**. Hier zijn de broedplaatsen van inheemse muggen in Amsterdam onderzocht met als doel het identificeren van de factoren die een goede habitat voor larven van deze soorten bepalen. *Culex pipiens* is een belangrijke vector van het West Nijl virus en is in het onderzoek de meest voorkomende soort in Amsterdam. Deze soort wordt vooral gevonden in regenwaterafvoeren. In stadsgroen en polders is sprake van een hoge soortendiversiteit.

Meer kennis van de verspreiding van vectoren is belangrijk voor effectieve preventie en controle. In **hoofdstuk 7** wordt de haalbaarheid beoordeeld van "auto-dissemination" (auto-verspreidingsaanpak) als een mogelijk alternatief voor de traditionele controle-instrumenten tegen *Ae. albopictus* in stedelijke gebieden. De aanpak is gebaseerd op de mogelijkheid dat wilde volwassen vrouwtjes worden blootgesteld aan het juveniel hormoon pyriproxyfen. De wilde vrouwtjes verspreiden het hormoon naar habitats voor larven die zich ten gevolge hiervan niet tot volwassen exemplaren kunnen ontwikkelen. De studieresultaten bevestigen de haalbaarheid van de auto-verspreidingsaanpak. In **hoofdstuk 8** worden de belangrijkste resultaten van de **hoofdstukken 2-7** beschreven en besproken.

Kortom, in dit proefschrift worden verschillende aspecten van muggen vectoren bestudeerd met als doel de kennis van hun ruimtelijke verspreiding te verbeteren. Nieuwe kennis over vectoren is noodzakelijk om (her-)ontstaan en verspreiding van door vectoren overgedragen ziekten te voorspellen en om nieuwe interventiemethoden te ontwikkelen die de verspreiding van door vectoren overgedragen ziekten tegen kunnen gaan.

Summary

Vector-borne diseases are infections transmitted by the bite of infected arthropod vectors, such as mosquitoes, ticks, fleas, midges and flies. Vector-borne diseases pose an increasingly wider threat to global public health, both in terms of people affected and their geographical spread. Mosquitoes are one of the most important vectors in Europe. In this thesis, we focus on mosquito borne-diseases. In particular, the species *Anopheles*, *Aedes albopictus* and *Culex* are investigated.

The distribution of vectors is investigated using species distribution models (SDMs). SDMs are numerical tools that combine observations of species presence or abundance with environmental characteristics. The general idea behind SDMs is to identify, through a statistical model, relationships between the known distribution of a species and the environmental characteristics; these relationships are used to make predictions for all un-sampled areas in the study region.

Different geographical scales are used in each study, according to the objectives of the study and the data available. Small scale maps (i.e., low resolution) show less detail and cover larger areas such as a whole country. Large scale maps (i.e., high resolution) show more detail and cover smaller areas.

In the first part of this thesis, the analysis was conducted on indigenous mosquito species (vectors and non-vectors) observed at a national level at low resolution. In **chapter 1**, the themes covered in this thesis are introduced.

In **chapter 2** three different modelling techniques were used to explore the potential spatial distribution in the Netherlands of three indigenous mosquito species. The models were used to investigate which environmental features were associated with the presence of the species, and to produce environmental suitability maps in the Netherlands. The potential spatial distribution of one of the most frequently observed species in the Netherlands, *Anopheles plumbeus*, is investigated in **chapter 3**. This species is a nuisance mosquito species and a potential malaria vector. Presence/absence and abundance distribution maps in the Netherlands were produced.

In the second part of this thesis, the focus shifted to urban and suburban areas, and to two potential vector species, i.e., *Aedes albopictus* and *Culex pipiens*. In **chapters 4-6** the spatial distribution of mosquito species is investigated in urban areas, in Rome (Italy) and Amsterdam (the Netherlands). Because the study areas are smaller than those presented in the previous chapters, it was possible to analyse the environment in more detail (at a higher resolution).

The invasive mosquito *Ae. albopictus*, commonly known as the 'tiger mosquito', is now present in several European countries. *Aedes albopictus* is a known vector of several viruses, among the others chikungunya and dengue virus and it is also known to cause significant biting nuisance. The habitat preference of the tiger mosquito is investigated in **chapter 4**. Spatial statistical models were used to evaluate the relationship between *Ae. albopictus* egg abundance and land cover types on the campus of Sapienza University in Rome. Areas with trees were associated with higher numbers of eggs and grass was associated with a lower number of eggs, suggesting that shade

provides a favourable habitat for mosquitoes. Solar radiation had a positive effect on the number of eggs, suggesting that warm areas are favourable for oviposition. Here, solar radiation does not represent the amount of sunshine on the trap as the ovitraps were always put in shaded sites and never exposed directly to sunlight.

In the same study area, an estimate of the population abundance of *Ae. albopictus* based on mark-release-recapture experiments is provided in **chapter 5**. Mark-release-recapture experiments consist of releasing marked mosquitoes from a release point and, after several days, counting, the number of marked and unmarked mosquitoes found in sticky traps distributed over the study area. The model presented here is based on a method known in literature but it was specifically developed to estimate the population size by including features of mark-release-recapture experiments carried out with mosquitoes.

An additional study that aims at giving insights into urban mosquito ecology is presented in **chapter 6**. Here the breeding sites of the indigenous mosquito species in Amsterdam are investigated with the aim of identifying the factors that determine a good larval habitat for these species. *Culex pipiens*, a species considered the principal vectors of West Nile virus, was the most common species and it was mainly found in rainwater drainage. In urban green and suburban rural areas a high diversity of species was observed. The chance of finding larvae (of any species) in accumulation of water after runoff was higher compared to the other larval habitats.

A better understanding of the spatial distribution of vectors is needed to help with successful prevention and control. In **chapter 7**, the feasibility of the “auto-dissemination” approach as a possible alternative to traditional control tools against *Ae. albopictus* in urban areas is assessed. The approach is based on the assumption that wild adult females, exposed to artificial resting sites contaminated with pyriproxyfen, can disseminate this juvenile hormone analogue to larval habitats, thus interfering with adult emergence. The results support the potential feasibility of the auto-dissemination approach to control *Ae. albopictus* in urban temperate areas.

In **chapter 8** the main issues encountered in **chapters 2-7** are highlighted. Approaches to handle them are proposed and the results are discussed.

In short, in this thesis, the ecology of mosquito-borne diseases was studied with the aim to improve the knowledge on the spatial distribution and the abundance of mosquito vectors. New knowledge on vectors is necessary to predict the (re-)emergence and spread of vector-borne diseases and to develop new interventions to interrupt or limit the spread of vector-borne diseases.

Riassunto

Le malattie trasmesse da vettori sono infezioni diffuse a seguito della puntura di vettori artropodi infetti, come zanzare, zecche, pulci, moscerini e mosche. Con il termine vettore si intende un organismo che può trasmettere malattie o parassiti a persone, animali o piante. Queste malattie costituiscono una minaccia sempre più ampia alla salute pubblica mondiale, sia in termini di persone colpite sia per la loro diffusione geografica. Le zanzare sono uno dei principali vettori presenti in Europa. Per tale ragione, in questa tesi focalizziamo l'attenzione sulle malattie trasmesse da zanzare. In particolare, sono qui studiate le specie *Anopheles*, *Aedes albopictus* e *Culex*.

La distribuzione delle suddette specie viene qui investigata attraverso l'utilizzo di modelli di distribuzione della specie (MDS). I MDS sono strumenti numerici che combinano osservazioni su presenza o abbondanza di una specie con caratteristiche ambientali. L'idea generale alla base dei MDS è di identificare, attraverso modelli statistici, la relazione tra la distribuzione nota di una specie e le caratteristiche ambientali; questa relazione viene usata per fare previsioni nelle aree in cui la distribuzione è sconosciuta.

In ciascuno studio sono state utilizzate diverse scale geografiche, in base agli obiettivi e ai dati disponibili. In generale, una mappa con piccola scala mostra pochi dettagli e ampie aree, come ad esempio un'intera nazione. In questo caso si parla di bassa risoluzione spaziale. Invece, una mappa con larga scala mostra più dettagli e piccole aree. Allora la risoluzione spaziale è alta.

Nella prima parte della tesi, l'analisi è stata condotta su zanzare indigene (vettori di malattie e non) osservate a livello nazionale. La risoluzione spaziale è bassa.

Nel **capitolo 1** vengono introdotte le tematiche trattate nella tesi. Nel **capitolo 2**, sono usate tre tecniche statistiche per esplorare la potenziale distribuzione geografica di tre specie di zanzare presenti nei Paesi Bassi. I modelli statistici sono usati per trovare una relazione tra presenza delle specie e caratteristiche ambientali circostanti. Questa relazione viene usata per costruire mappe che mostrano l'idoneità ambientale di ciascuna delle tre specie. La potenziale distribuzione spaziale di una delle specie più frequentemente osservate nei Paesi Bassi, *Anopheles plumbeus*, è studiata nel **capitolo 3**, nel quale sono state prodotte mappe di presenza/assenza e abbondanza nei Paesi Bassi. *Anopheles plumbeus* provoca fastidio con le sue punture ed è un potenziale vettore di malaria.

Nella seconda parte della tesi, l'attenzione è spostata su aree urbane e suburbane e su due specie che sono potenziali vettori di malattie, ovvero *Ae. albopictus* e *Culex pipiens*. Nei **capitoli 4-6** la distribuzione spaziale di zanzare è studiata in aree urbane, a Roma (Italia) e Amsterdam (Paesi Bassi). Poiché l'area di studio è più piccola rispetto a quella dei capitoli precedenti, è stato possibile analizzare le caratteristiche ambientali in maggior dettaglio, ad alta risoluzione spaziale.

La specie invasiva *Ae. albopictus*, comunemente conosciuta come 'zanzara tigre', è attualmente presente in diverse nazioni europee. La Febbre chikungunya e la dengue sono tra le malattie acute virali trasmesse da *Ae. albopictus*. Questa specie è particolarmente nota per il fastidio arre-

cato dalle sue punture. Le preferenze di habitat della zanzare tigre sono analizzate nel **capitolo 4**. Modelli statistici spaziali sono usati per studiare la relazione tra l'abbondanza di uova di *Ae. albopictus* e le caratteristiche del territorio nel campus dell'università "La Sapienza" di Roma. Aree con alberi sono risultate associate ad un maggior numero di uova, mentre per i prati si è osservata la relazione opposta. Questi risultati suggeriscono che aree ombreggiate costituiscono un habitat favorevole per le zanzare. Le radiazioni solari sono positivamente associate al numero di uova, indicando che aree calde sono favorevoli per la deposizione di uova.

Nella stessa area di studio, viene fornita (**capitolo 5**) anche una stima di abbondanza della popolazione di *Ae. albopictus*. La stima è basata su un esperimento di marcatura, rilascio e ricattura di zanzare. In questo tipo di esperimenti, zanzare marcate vengono rilasciate (da un punto di rilascio) e dopo diversi giorni le zanzare marcate ricatturate e le zanzare non marcate catturate vengono contate nelle trappole distribuite nell'area di studio. Il modello presentato in questo capitolo è basato su un modello presente in letteratura, ma è specificatamente sviluppato per stimare la popolazione di zanzare a partire da esperimenti di marcatura, rilascio e ricattura e includerne caratteristiche specifiche, quali la distanza tra punto di rilascio e ricattura e il numero di giorni trascorsi tra il rilascio e la ricattura.

Nel **capitolo 6** è presentato uno studio aggiuntivo che mira ad approfondire l'ecologia urbana di zanzare. Qui i siti di riproduzione di zanzare presenti ad Amsterdam sono analizzati per identificare i fattori che determinano un buon habitat larvale per queste specie. *Culex pipiens*, specie considerata tra i principali vettori del virus del Nilo occidentale, è stata la specie trovata con maggiore frequenza e principalmente in acqua drenata dalla pioggia. In aree verdi urbane e suburbane (polder) è stata osservata un'ampia varietà di specie. La possibilità di trovare larve (di qualsiasi specie) in un accumulo di acqua dopo un deflusso su una superficie è risultata maggiore rispetto agli altri habitat larvali.

Una migliore comprensione della distribuzione spaziale di vettori è necessaria ed utile per programmare azioni di prevenzione e controllo. Nel **capitolo 7**, viene valutata la praticabilità dell'approccio "auto-disseminazione" come possibile alternativa ai mezzi di controllo tradizionali contro la zanzara tigre in zone urbane. L'approccio è basato sulla possibilità che femmine adulte, esposte a luoghi di riposo artificiali contaminati con pyriproxyfen, possano disseminare questo ormone in habitat larvali, riducendo lo sviluppo di adulti. I risultati supportano la fattibilità dell'approccio "auto-disseminazione" per controllare *Ae. albopictus* in aree urbane temperate. Nel **capitolo 8** sono descritte le principali questioni emerse nei **capitoli 2-7** e vengono discussi i risultati emersi.

In breve, in questa tesi, l'ecologia delle malattie trasmesse da zanzare è studiata con l'obiettivo di ampliare le conoscenze sulla distribuzione spaziale di zanzare che possono agire da vettori di malattie. Nuove conoscenze sui vettori sono necessarie per predire il (ri)emergere e la diffusione di malattie trasmesse da vettori e sviluppare interventi che ne interrompano o ne limitino la diffusione.

Dankwoord

Het was een lange en uitdagende reis en gelukkig was ik niet alleen. Ik wil al mijn reisgenoten bedanken!

Eerst, Hans, dank je wel voor het vertrouwen en dat je mij de kans hebt gegeven om een PhD te doen. In deze vier jaar heb ik veel geleerd, onder andere dat een PhD student ook zelf kan denken en niet hoeft te wachten op een begeleider die opdrachten geeft...

Nienke, dank je wel voor je professionele, warm menselijke en morele hulp! Ik kon mij geen betere co-promotor dromen!

Jan, het was altijd super interessant om met je over statistiek te praten met op de achtergrond mooie klassieke muziek! Dank je daarvoor en voor je adviezen. Hans, het was leuker als jij meedeed aan onze vergaderingen. Dank je wel, je hebt altijd tijd gevonden als ik dat nodig had en dat was meestal zonder afspraak (ze zeggen dat Nederlanders dat niet kunnen, maar met jou was het wel mogelijk!).

Een grote dank aan Linsey, Gerrit, Lenny, Ellen, Miriam, Tijs, Niels, Liesbeth, Sanja, Kimm, Judith, Saskia, Maite, Helena, Francesca, Don, Francisca, Christian, Mirjam, Elise, Hilde en alle andere collega's voor de gezellige conversaties over Nutella en de rare Italiaanse gewoonten en voor de meer serieuze wetenschappelijke gesprekken. Een speciale dank aan mijn sportieve collega's Jamal, Antonie, Robbert, Eline, Marjolein, Giovanni, Joost, Arjan: het sporten, rennen, schaatsen en naar de sportschool gaan met jullie heeft veel geholpen tegen de gewoonlijke PhD frustraties.

Een heel speciale dank aan mijn paranimfen Marjolein en Claudia, ik ben zo blij dat ik jullie heb ontmoet. Jullie ondersteuning en aanmoediging was essentieel voor het succes van mijn PhD!

Dank dank dank aan...

...de Averechts familie: Sarah, Liza, Sabijn, Joram, Eline, Joost, Hilde, Nancy, Nila, Corine, Sharon en alle andere vrijwilligers

...aan Marianne, Jurgen, Be, de andere vrijwilligers en de dieren van Kinderboerderij Abstede

...aan Ton, Paul, Madeleine, Theo en de vrijwilligers van het Park Bloeyendael

...aan Marjon, voor je fantastische lessen Nederlands

...aan de vrienden van de leesclub voor de gezelligheid

...aan Libreria Bonardi, Pizzeria Tricolore en Roberto Gelato, mijn Italiaanse hoekjes in Nederland!

...aan Henny, Meindert, Mark, Daan, Max en Björn omdat jullie mij liefdevol hebben verwelkomd in de familie

...aan Mayumi, voor je vrolijke inbreng in de familie Cianci deze zomer (helaas slechts voor enkele dagen)

...aan Nikolas, voor het verzorgen van de plantjes, voor je ironie, voor de ijsjes...voor alles!

Of course I would like to thank all the co-authors of the papers included in this thesis for the fruitful collaborations and for trying to explain me something about the mosquitoes' world. Thanks to the EDENext colleagues, in particular to Caroline, Maryam, Neil, David, Els and Marieta.

Eimear, thanks for having welcomed me in Utrecht and for having introduced me to the International Neighbour Group. Thanks for always being there when I need and for having discovered the best restaurants in Utrecht and Amsterdam!

Un grazie va a tutti gli amici, vicini e lontani: Simone, Adele, Dado e Titta, Martina, Carla e Domenico, Chiara, Chiara e Marcello, Grazia, Cecilia, Rosalinda, Antonio, Lorenzo e Vittoria, Tiziana e Ivan, Giovanni e Marie, Charlotte, Silvia e Alessandro, Federica, Lucia, Laura, Angelica.

Grazie a...

...le mie MilanoAmiche Mappi, Masha, Dany e Giulia e a Svevina, Adele, Fabrizio e Ale, con voi la distanza non esiste!

...a Paola, per aver altruisticamente "finto" che, in viaggio verso Detroit, non c'era lettura più interessante da fare se non i miei articoli

...a Roberta, Cristina e Luisa, per aver stimolato il mio interesse verso la ricerca

...agli amici del cammino d'Abruzzo (in siccità!) e di altre avventure: Gloria, Maria, Davide, Livia, grazie per la vostra allegria!

...ai miei nonni, miei grandi maestri di vita.

Mamma e papà, grazie per aver sempre supportato le mie scelte, anche quando mi hanno portata a 1700 km di distanza da voi. Paolo, il miglior allenatore, non solo nella corsa!

Mathijn, grazie per avermi accompagnato con la tua dolce fermezza in questa avventura, sei pronto per la prossima?

About the author

Daniela Cianci was born on 4th April 1985 in Lanciano, Italy. After attending the high school Liceo Scientifico in Lanciano, in 2004 she went to study Statistics and Social Research at the University of Bologna, where she received her BSc. In 2007 she moved to Milan to study for an MSc in Biostatistics. For her master thesis she did an internship at the Laboratory of Genetic Epidemiology and Medical Statistics, at the University of Pavia, where she worked on a project looking at risk factors for acute myocardial infarction, focussing on gender differences. After obtaining her MSc, she won a one-year scholarship in the same group. In that period she conducted research in epidemiology and biomedical statistics and she was also involved in teaching activities. In March 2011, she moved to the Netherlands to start a PhD on modelling of vector-borne diseases in the Theoretical Epidemiology Group at Utrecht University, from which the results are presented in this thesis. Her PhD was part of the EDENext project, a European research project dedicated to investigating several aspects of vector-borne diseases. During her PhD, Daniela collaborated with the Department of Public Health and Infectious Diseases at 'Sapienza' University in Rome (Italy), the Earth and Life Institute at the Université catholique de Louvain (Belgium), the Centrum Monitoring Vectoren in Wageningen (the Netherlands) and the GGD Amsterdam (the Netherlands). She also presented her work at several international conferences. In 2014 she won the poster award with the poster "First attempt to model the spatial distribution of mosquito species in the Netherlands using Vecmap", which was presented at the EDENext annual meeting in Lapland. Daniela is currently working as teacher in statistics at Biometris, Plant Sciences Group, Wageningen University.

List of publications

- A. Ibáñez-Justicia and **D. Cianci**. (2015). Modelling the spatial distribution of the nuisance mosquito species *Anophels plumbeus* (Diptera: Culicidae) in the Netherlands. *Parasites & Vectors*, *8*(1):258 doi:10.1186/s13071-015-0865-7
- D. Cianci**, N. Hartemink, C.B. Zeimes, S.O. Vanwambeke, A. Ienco and B. Caputo. (2015). High resolution satial analysis of habitat peferece of *Aedes albopictus* in an urban environment. *Journal of Medical Entomology*, doi: 10.1093/jme/tjv026.
- D. Cianci**, N. Hartemink, A. Ibáñez-Justicia. (2015). Modelling the potential spatial distribution of mosquito species using three different techniques. *International Journal of Health Geographics*, *14*(1), 10.
- N. Hartemink, **D. Cianci**, P. Reiter. (2015). R0 for Vector-Borne Diseases: Impact of the Assumption for the Duration of the Extrinsic Incubation Period. *Vector-Borne and Zoonotic Diseases*, *15*(3): 215-217.
- D. Cianci**, J. Van Den Broek, B. Caputo, F. Marini, A. Della Torre, H. Heesterbeek and N. Hartemink. (2013). Estimating mosquito population size from mark-release-recapture data. *Journal of Medical Entomology*, *50*(3), pp. 533-542.
- B. Caputo, A. Ienco, **D. Cianci**, M. Pombi, V. Petrarca, A. Baseggio, G.J. Devine and A. della Torre. (2012). The "Auto-Dissemination" Approach: A Novel Concept to Fight *Aedes albopictus* in Urban Areas. *PLoS Neglected Tropical Diseases*, *6*(8) , art. no. e1793.

In preparation

- J. Buijs, **D. Cianci**, S. Koenraadt, A. Stroo, C. Vogels. Characterizing mosquito breeding sites in an urban setting: a case study in Amsterdam

CONTRIBUTIONS TO CONFERENCES AND SYMPOSIA

Oral presentations

- D. Cianci**, N. Hartemink and A. Ibáñez-Justicia. Modelling the spatial distribution of mosquito species using VecmapTM. Abstract Book the 19h International Conference E-Sove 2014. 13-17th October 2014, Thessaloniki, Greece.
- D. Cianci**, N. Hartemink and A. Ibáñez-Justicia. Modelling the spatial distribution of mosquito species in the Netherlands using VecmapTM. Annual meeting of the Dutch Society of Theoretical Biology (NVTB) 22-23 May 2014, Schoorl, The Netherlands.

- D. Cianci**, N. Hartemink, C. B. Zeimes, S. O. Vanwambeke, A. Ienco, A. della Torre and B. Caputo. High resolution spatial analysis of habitat preference of *Aedes albopictus* in an urban environment. 25th Annual meeting of the Dutch Society for Veterinary Epidemiology and Economics (VEEC) 12 November 2013, RIVM, Bilthoven, The Netherlands.
- D. Cianci**, J. Van Den Broek, B. Caputo, F. Marini, A. Della Torre, C. Zeimes, S. Vanwambeke, H. Heesterbeek and N. Hartemink. Models for abundance and habitat preference of *Aedes albopictus*. Abstract Book the 18th International Conference E-Sove 2012. 8-11th October 2012, Montpellier, France.

Poster presentations

- D. Cianci**, N. Hartemink, B. Caputo, A. Ibáñez-Justicia, H. Heesterbeek. Modelling the spatial distribution of mosquitoes at different geographical scales. GERI 2015. 21-23 April 2015, Heraklion, Crete, Greece.
- A. Ibañez-Justicia and **D. Cianci**. Modelling vector surveillance data of the nuisance mosquito species *Anopheles plumbeus* (Diptera: Culicidae) in the Netherlands. Abstract Book the 19th International Conference E-Sove 2014. 13-17th October 2014, Thessaloniki, Greece.
- J. Buijs, C. Vogels, **D. Cianci**, W. den Hartog, S. Koenraadt and A. Stroo. Mosquito surveillance in Amsterdam Urban ecology, mosquito presence & nuisance perception. Abstract Book the 19th International Conference E-Sove 2014. 13-17th October 2014, Thessaloniki, Greece.

