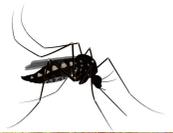


REMOTE SENSING AND **SPATIAL** MODELLING

Applications to the surveillance and control
of mosquito-borne diseases

Annelise Tran, Éric Daudé, Thibault Catry, coord.



Remote Sensing and Spatial Modelling

Applications to the surveillance
and control of mosquito-borne diseases

Annelise Tran, Éric Daudé, Thibault Catry, coordinators

Éditions Quæ

Collection Update Sciences & Technologies

*Transformative Participation
for Socio-ecological Sustainability
Around the CoOPLAGE Pathways*
E. Hassenforder, N., Ferrand, coord.
2024, 270 p.

*Studying Food and Eaters:
A Cocktail of Perspectives and Methods*
O. Lepiller, T. Fournier, N. Bricas,
M. Figuié, coord.
2023, 218 p.

*Public Policies and Food Systems
in Latin America*
J.-F. Le Coq, C. Grisa, S. Guéneau,
P. Niederle, coord.
2022, 446 p.

*Eating in the City
Socio-anthropological Perspectives
from Africa, Latin America and Asia*
A. Soula, C. Yount-André, O. Lepiller,
N. Bricas, J-P. Hassoun, coord.
2020, 158 p.

This publication received support from CIRAD, IRD and CNRS.

To cite this book:

Tran A., Daudé É., Catry T., coord., 2025. *Remote Sensing and Spatial Modelling. Applications to the Surveillance and Control of Mosquito-borne Diseases*, Versailles, Quæ, 148 p., <https://doi.org/10.35690/978-2-7592-4102-6>

This book is the English translation of:

Tran A., Daudé É., Catry T., coord., 2022. *Télé-détection et modélisation spatiale. Applications à la surveillance et au contrôle des maladies liées aux moustiques*, Versailles, Quæ, 148 p.

The digital versions of this book are distributed under the CC-by-NC-ND 4.0 license (<https://creativecommons.org/licenses/by-nc-nd/4.0/>).



Éditions Quæ
RD 10, 78026 Versailles Cedex
www.quae.com – www.quae-open.com

© Éditions Quæ, 2025

ISBN paperback: 978-2-7592-4101-9
ISBN ePub: 978-2-7592-4103-3

ISBN PDF: 978-2-7592-4102-6
ISSN: 1773-7923

Table of contents

Foreword	7
<i>Didier Fontenille</i>	
General introduction	11
<i>Thibault Catry, Éric Daudé, Nadine Dessay, Annelise Tran</i>	
Remote sensing concepts.....	12
Introduction to GIS.....	15

PART 1

SPATIAL DATA FOR VECTOR MOSQUITO SURVEILLANCE AND ASSOCIATED DISEASES

Chapter 1. Relationships between vector mosquitoes and the environment: the role of satellite remote sensing methods	19
<i>Renaud Marti, Claire Teillet, Hobiniaina Anthonio Rakotoarison, Florence Fournet</i>	
Relationships between vector mosquitoes and the environment.....	22
Description of the environment using satellite remote sensing methods	27
Références	34
Chapter 2. Spectral indices and classifications of multispectral images for vector risk mapping	43
<i>Annelise Tran, Renaud Marti, Vincent Herbreteau</i>	
Mapping land cover using remotely sensed data in order to model the distribution of <i>Anopheles</i> mosquitoes in Camargue	45
Spectral indices derived from remote sensing images employed as environmental factors in the analysis of human cases of West Nile fever in Europe	47
Automated production of spectral indices: example of the Sen2Extract tool.....	49
References	51
Chapter 3. Estimation of air temperatures from satellite images and weather stations	53
<i>Barbara Boufhal, Alexandre Cebeillac, Éric Daudé</i>	
Data to measure temperatures.....	53
Air temperature estimation: different methods	54
Applications to Bangkok.....	56
Conclusion	61
References	61

Chapter 4. From census to buildings: generating synthetic populations	63
<i>Alexandre Cebeillac, Olivier Gillet, Éric Daudé</i>	
Population disaggregation and redistribution	64
Synthetic populations, a methodology which supports the fine-scale analysis of health-related issues	69
Conclusion	71
References	71
Chapter 5. Satellite image texture and characterisation of urban environments favourable to vector mosquitoes	73
<i>Claire Teillet, Ophélie Hoarau, Nausicaa Habchi-Hanriot, Benjamin Pillot, Thibault Catry, Annelise Tran</i>	
Different methods to characterise image texture	74
Study of the relationships between urban variables and the distribution of dengue cases in Brasília using a texture-based approach	76
Map of potential larval habitat distribution of the Asian tiger mosquito on Reunion island	79
Conclusion	82
References	83

PART 2

ANALYSING AND PREDICTING THE EFFECT OF ENVIRONMENTAL VARIABLES
ON THE DISTRIBUTION AND DYNAMICS OF VECTOR MOSQUITOES

Chapter 6. Data-driven models: mapping the spatial distribution of vectors	87
<i>Yi Moua, Emmanuel Roux</i>	
Species distribution models	87
Maxent model	89
Sampling bias and minimising its impact on modelling	89
Application to the primary malaria vector in French Guiana	90
Conclusion	95
References	96
Chapter 7. Knowledge-based models: example of a multi-criteria evaluation tool for public health	99
<i>Fanjasoa Rakotomanana, Hobiniaina Anthonio Rakotoarison</i>	
GIS-based multi-criteria analysis, a knowledge-based approach	100
Example of this method used to map the risk of malaria in the Malagasy Central Highlands	102
Conclusion	107
References	107

Chapter 8. Arbocarto: a mechanistic model based on the life cycle of <i>Aedes</i> mosquitoes.....	109
<i>Renaud Marti, Marie Demarchi, Mathieu Castets, Annelise Tran</i>	
A generic model built around the mosquito life cycle.....	110
Adaptation of the model to the species <i>Aedes albopictus</i> and <i>Aedes aegypti</i> and spatialization.....	111
Implementation, initialisation and simulation of <i>Aedes</i> mosquito abundance.....	113
Arbocarto: a specialised interface for vector control measures.....	115
References.....	118
Chapter 9. Spatial simulation of the risk of dengue transmission using vector and host behavioural models.....	119
<i>Éric Daudé, Sébastien Rey-Coyrehourcq, Alexandre Cebeillac</i>	
Individual-based and spatially explicit models.....	120
Application to dengue in Bangkok: MO ³ , methods and data.....	121
MO ³ model simulations.....	126
Conclusion.....	129
References.....	129
General conclusion and perspectives.....	131
<i>Thierry Baldet, H�el�ene Guis</i>	
References.....	135
Acknowledgements.....	137
Glossary.....	139
List of acronyms.....	141
Authors.....	145

Foreword

Mosquitoes. The term is generic in nature, yet it is eminently practical for encompassing a vast array of biological, ecological, health, social, economic and historical diversity. Mosquitoes, or *Culicidae* to use their scientific name, refers to the 3,700 species described here on earth, in addition to an unknown number yet to be described. If nature abhors a vacuum, then so do mosquitoes. They are ubiquitous, occurring on every continent and across all ecosystems, and have been around for far longer than humans.

This is something which our human readers often forget. You might only think about mosquitoes if they're keeping you awake at night, stopping you from enjoying an evening outdoors, or, depending on where you live, for causing an illness in the family or to one of your livestock or pets.

Emergence. This has become a trendy word in recent years. Humans have become aware that diseases can emerge. The French microbiologist Charles Nicolle, in his 1933 work *Destin des maladies infectieuses*, already predicted that "there will be new [infectious] diseases. It's a fatal fact". Some of these diseases, emerging from the wild as a consequence of environmental, climatic, demographic, societal, cultural, health, and economic changes, among other factors, are vectorial diseases, and sometimes mosquitoes are responsible for this transmission through the inoculation of viruses and parasites.

Mosquitoes are insects, but their study and control go well beyond entomology (from *entoma*, meaning insect in ancient Greek). A multitude of complementary disciplines are involved, ranging from taxonomy to public health. Remote sensing and spatial modelling are counted among these, and they have become indispensable tools in medical and veterinary entomology, as well as agricultural entomology.

By the 5th century BC, the Greek philosopher and physician Hippocrates had already established the link between environmental factors and the aetiology of disease. He described fevers with the same set of symptoms as malaria, and noted a connection between the wetlands and these fevers in his treatise *On Airs, Waters, and Places*. Of course, at the time, even though people likely complained about mosquitoes, it was not feasible to form a causal relationship with malaria. In the not-so-distant past and closer to home, in France the inhabitants of the regions now known as Vendée, Sologne, Dombes and Camargue were invaded by mosquitoes, and fevers were commonplace in these areas until the beginning of the 20th century. The construction site of the Palace of Versailles was the site of numerous fatalities, likely attributable to malaria, before the surface water was channelled.

It is only in our recent past, following the formulation of modern germ theory by Louis Pasteur, that causal relationships have been established between the environment, climate, mosquitoes, microbes and diseases. Over the past two decades, significant

progress in our understanding of these relationships have been made, thanks in part to novel genomic techniques, but also due to the emergence of sophisticated remote sensing technologies, spatial analysis tools for biological phenomena (mosquitoes included) and advances in health risk modelling.

The biological diversity of mosquitoes is extraordinary. These 3,700 species are particularly well adapted to specific environments and biotopes. Some mosquito larvae are only found in very specific larval habitats, such as small, water-filled cavities in trees, known as phytotelmata, or the pitchers of carnivorous plants like *Nepenthes*. Others are less picky and are able to thrive in lakes, marshes or on riverbanks; yet others are almost exclusively found in areas where water has collected due to human activities. Certain species are endemic to a single region (*Aedes pia* on the island of Mayotte), whereas others, which have adapted to urban environments, can be found on every continent (*Aedes albopictus*). Some of these can take blood meals from many different animals, including humans (*Anopheles arabiensis*), whereas others have very strict diets (ant regurgitate for *Malaya* sp.). Some species can survive periods of drought or cold by their eggs entering diapause (*Aedes*), or their adult form resting in sheltered sites such as houses and stables. It is, however, essential that they have access to water in order to lay their eggs and for the development of larvae and pupae. Water plays a vital role in mosquito biology, exerting influence through its presence, quality, physical and chemical properties, as well as biotic factors (plants, food, predators). Any approach that seeks to ascertain, analyse and correlate water-related parameters (rainfall, development, vegetation, etc.) is capable of more accurately estimating, or even predicting, the presence or abundance of different mosquito species and populations, as well as the associated risks.

These risks are not trivial. History is replete with examples of fates being decided by mosquitoes, from the death of Alexander the Great attributed to malaria (*Anopheles*) or West Nile disease (*Culex*), and the excavation of the Panama canal being halted by malaria and yellow fever (*Aedes*), to the more recent example of the “vertical forest” buildings in China being abandoned by their inhabitants due to an invasion by Asian tiger mosquitoes. The list of infectious diseases transmitted to humans by different mosquito species is impressive. Nearly 100 human diseases can be attributed to mosquitoes. Some are still rare, such as Mayaro fever in South America. Yet others are much more common, such as malaria, which kills nearly 400,000 children every year in Africa, or dengue fever, which affects more than 300 million people each year and is present on every continent according to the World Health Organization (WHO).

The health, social and economic challenges associated with mosquitoes are therefore immense, without considering the ecological challenges. Although mosquitoes play an important role in the food chain and contribute to biodiversity, it is nevertheless essential to control species that are responsible for major human and animal diseases. This control must be rationalised, integrated, adapted, sustainable accepted and generate the least environmental impact. The era of the intensive use of insecticides is coming to an end. Other more targeted methods, including geographically targeted campaigns, are currently being developed. Approaches such as remote sensing, spatial analysis and modelling have become indispensable tools for achieving these goals, yet they remain underutilised in the decision-making process.

The examples provided in this publication—*Anopheles* and malaria risk in Camargue, French Guiana, Asia and Madagascar; *Aedes* and dengue risk in Thailand, Brazil and the Indian Ocean—show that remote sensing and spatial modelling applied to mosquitoes and mosquito-borne diseases play a crucial role in these efforts. They also show that interdisciplinary collaboration is required. Models based on inadequate documentation of biological data are not only devoid of meaning, but they can also foster false expectations among those that use them. Conversely, rigorous sampling in the field cannot be used to its full potential without good spatial modelling.

Each scientific community has its own concepts and language. Attending a specialist symposium can be a real ordeal if one is unable to decipher the code. Set an entomologist loose in a remote sensing convention, or a geomatics expert in the annual conference of the Society for Vector Ecology, and they may be unable to correctly interpret the words or acronyms being used, such as reflectance, scanning swath, exophilic, sternite, univoltine, raster mode, diapause, spectral signature, gonotrophic, Normalized Difference Vegetation Index (NDVI) and Normalized Difference Water Index (NDWI). Only the word vector can be understood by all, but with two very different meanings: one taken from the field of biology and the other from the field of geomatics. The idea behind this publication, authored by specialists who have worked with or even belong to both communities, is to make these concepts accessible to all with the help of well-documented and concrete examples. My sincerest thanks and best wishes go out to all the contributors. This publication will serve as an invaluable reference for those who recognise the need to adopt a global, spatial and environmental approach for the study of mosquitoes (and other vectors) and the documentation of their biology, distribution, impact and control. It will also prove beneficial to those seeking examples of the application of remote sensing and spatial modelling.

This book acts as a bridge between communities, inviting entomologists to more abstractions and macroscopic perspectives, and those working in the field of remote sensing and geomatics to discover the fascinating world of mosquitoes.

Didier Fontenille
Medical Entomologist, Research Director
IRD, UMR MIVEGEC (University of Montpellier, IRD, CNRS)

General introduction

Thibault Catry, Éric Daudé, Nadine Dessay, Annelise Tran

Remote sensing provides Earth observation data which can be particularly useful for modelling and mapping in public health. The World Health Organization (WHO) considers the identification, monitoring and control of arthropod vectors to be a priority for the surveillance of vector-borne diseases (VBDs). In this regard, over a period of more than two decades, a substantial body of research has demonstrated that satellite images and other spatially explicit data can assist in identifying the environmental and climatic variables that influence the spatio-temporal dynamics of VBDs, with a particular focus on mosquito-borne diseases. The current range of satellite sensors allows data to be acquired at high enough spatial and temporal resolutions to (i) characterise different environmental and climate variables (land cover, precipitation, temperature, humidity, etc.) associated with the presence of favourable habitats, and with vector presence and abundance, (ii) to develop predictive tools and methods to model, at different scales, the risks associated with these vectors and the pathogens they transmit, and (iii) to help monitor the evolution of this risk. These efforts are based on advanced satellite image processing techniques (pixel or object classification, use of time series, artificial intelligence algorithms, see Part 1) using multiple sensors (optical, radar, lidar, etc.) and multiple resolutions (medium, high, very high spatial resolutions), as well as a combination of these remote sensing variables with other types of spatial information in different modelling approaches (based on knowledge, data, processes or behaviours, see Part 2). These models can incorporate a large number of variables (in particular environmental and climate variables) into complex and dynamic systems, thereby enhancing our understanding of the epidemiology of mosquito-borne diseases and their transmission mechanisms, which represent a major public health concern.

The results of these studies, conducted as part of research programmes, have led to the development of operational methods based on remote sensing and modelling, which have proliferated in the field of public health in recent years. Some products (risk maps, processing chains) have thus been made available by organisations such as the land surface data and services hub, Theia¹, and in particular by its “Risks associated with Infectious Diseases” scientific expertise centre (SEC)². Such initiatives

1. www.theia-land.fr/en/

2. www.theia-land.fr/en/ceslist/risks-associated-with-infectious-diseases-sec/

have allowed research communities, whether from the field of geomatics, entomology or epidemiology, as well as other stakeholders in public health, to collaborate in pursuit of common goal: the enhancement of knowledge and tools to control mosquito-borne diseases. In particular, the ANISETTE project³ (Inter-Site Analysis: Evaluation of Remote Sensing as a predictive tool for the surveillance and control of diseases caused by mosquito), funded by the French national space agency, the Centre National d'Etudes Spatiales (CNES), between 2018 and 2022, aimed to assess the interoperability of methods combining remote sensing and spatial modelling to predict the dynamics of vector mosquitoes and their associated diseases. This project is based on the results of various other research projects, both concluded or ongoing, led by teams from different joint research units (ASTRE, Espace-Dev, IDEES and TETIS) who are engaged in entomological modelling in close collaboration with other organisations, such as institutes of the Pasteur Network, and in particular the Institut Pasteur de Madagascar (IPM, Pasteur Institute of Madagascar). The interoperability of the methods developed in the framework of these projects was tested across several sites in South America (Brazil, Antilles, French Guiana), in Europe (France), in the Indian Ocean (Madagascar, Mauritius, Réunion), in South and Southeast Asia (India, Thailand, Cambodia) and in Oceania (New Caledonia).

This publication is the culmination of efforts and reflections undertaken as part of the ANISETTE project. It presents a summary of the theoretical concepts, methodological approaches, tools and main results achieved by the project team, which is primarily composed of geographers, geomatic scientists and modellers. The aim is to introduce the concepts of remote sensing and spatial modelling and apply them to the study of mosquito-borne diseases. It is intended for laypersons who want a better grasp of these notions and their applications to public health⁴. This work is split into two separate parts: Part one covers remote sensing methods for the identification and characterisation of environmental and climate determinants of vector mosquito populations. Part two focuses on the integration of these variables into different modelling approaches in order to implement operational monitoring tools for VBDs caused by certain mosquito species. In order to facilitate comprehension, a glossary is provided at the end of the text, defining several technical terms related to the different fields under discussion, namely entomology, epidemiology, remote sensing, geomatics and mathematics. A list of acronyms and their respective definitions is also provided at the end of this publication.

► Remote sensing concepts

Remote sensing is defined as the set of techniques used to collect information on objects at a distance. In particular, Earth observation uses an instrument (a sensor) on board a platform (satellite, aircraft, drone, etc.) to characterise the Earth's surface (land surface, oceans or atmosphere). Typical examples of remote sensing include the use of satellite imagery or aerial photography.

3. <https://anisetite.cirad.fr/>

4. For more detailed information on vector mosquito biology and ecology, readers can refer to the publication *Le moustique, ennemi public n° 1 ?*, coordinated by S. Lecollinet, D. Fontenille, N. Pagès and A.-B. Failloux, published by Éditions Quæ (2022).

Main characteristics of remote sensors

Different types of remote sensors exist. Passive sensors measure the natural radiation emitted or reflected by the surface being observed, as in the case of optical sensors which rely on an external energy source (sunlight). As for active sensors, these measure the reflected radiation which they themselves emit. This is the case for radar, which emits its own energy source and measures surface roughness and humidity.

The signals measured by remote sensors are referred to as “electromagnetic radiation” and possess properties which can be quantified and described. These properties include the wavelength (which represents the spatial period of a wave, i.e., the distance between two successive maxima), the amplitude (or intensity, which corresponds to the maximum value of the oscillation) and polarisation (relationship between the amplitude and the direction of travel of the wave). Sensors measure the quantity of energy carried by the electromagnetic radiation emitted or reflected by the surfaces. In particular, this includes the albedo or directional-hemispherical reflectance, which is defined as the ratio between the energy emitted and the energy received. Panchromatic images, in black and white, are obtained from recording this radiation in a single band of wavelength. “Multispectral” imaging refers to when these measurements are taken across different wavelengths. Remote sensors can measure signals in the visible spectrum (optical remote sensing), infrared spectrum or microwave spectrum (radar remote sensing), thereby providing supplementary data. Sensors can be ground-based, on board aircraft or drones (airborne remote sensing), or on satellites. Most Earth observation sensors capture information in the form of digital images characterised by pixel size and the width of the snapshot (swath width).

Remote sensors are mainly categorised by three resolutions:

- spatial resolution is the smallest size of observable objects, directly related to the elementary size of the image pixel. We refer to low resolution when images have a spatial resolution greater than 1 km, medium resolution when this is between 100 m and 1 km, high resolution between 10 and 100 m, and very high resolution for values ranging from a few dozen centimetres to several metres;
- temporal resolution, which corresponds to time taken by a satellite to revisit a given location, i.e., the time between taking two identical shots;
- spectral resolution which characterises the ability of the sensor to distinguish between signals of different wavelengths.

Spectral signature concept

Each surface type is characterised by its spectral signature, defined as the variation in reflectance as a function of wavelength (Figure I.1). Spectral signature depends on the nature of the surface, its physical properties and the interaction it has with the incoming electromagnetic waves.

Extracting information from satellite images

There are different ways to make use of the spectral information contained in satellite images. The simplest method consists of a visual interpretation of the image, or photointerpretation. More complex analysis methods are used to classify the spectral content of images based on the information contained in pixels (pixel-oriented approach), or in objects whereby images are segmented into homogeneous regions

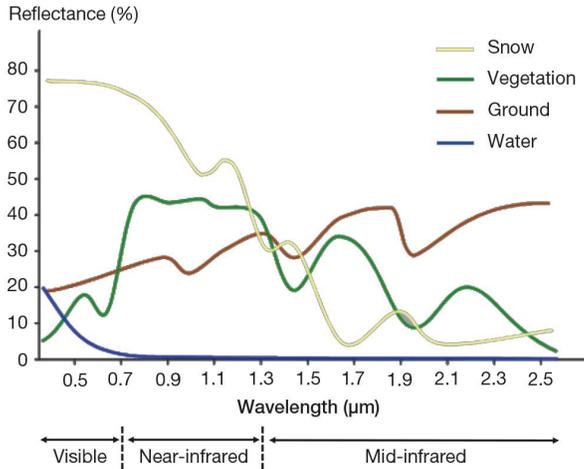


Figure I.1. Spectral signatures of natural surfaces across different wavelength bands. Adapted from <https://e-cours.univ-paris1.fr/>.

of pixels (object-oriented approach). These classification approaches can be unsupervised (without *a priori* knowledge of the image to be classified) or supervised (when prior knowledge is available) [Figure I.2]. Classification algorithms (K-means, Random Forest, Support-Vector Machine, etc.) group the information contained in each pixel or object into a cluster which describes the image in question.

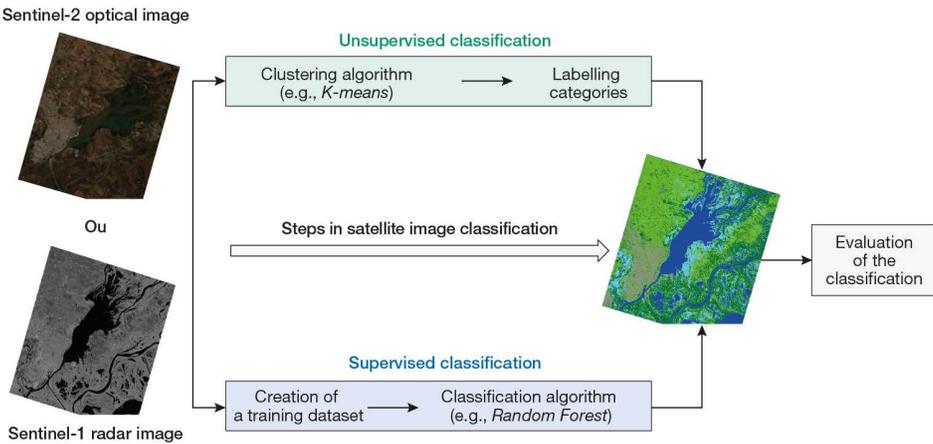


Figure I.2. General principles of supervised and unsupervised approaches to optical and radar satellite image classification.

Satellite image broadcasting

Today, a plethora of platforms exist to provide optical and radar satellite images. Such is the case for the Sentinel-1 and 2 sensors of the European Space Agency in the framework of the Copernicus programme⁵.

5. www.copernicus.eu/en/about-copernicus/infrastructure-overview/discover-our-satellites

A wide range of free software, tools and satellite image processing algorithms are also available which help contribute to the popularisation of this technology. The information extracted from remote sensing images can then be combined and analysed with other spatially explicit data using a Geographic Information System (GIS).

► Introduction to GIS

GIS are computational tools for the acquisition, storage, updating, integration, analysis, visualisation and recovery of georeferenced digital data (i.e., data which can be associated with a specific location through its geographic coordinates). They allow different types and sources of spatially explicit data to be handled and processed.

Georeferenced (or spatially referenced) data is organised in a GIS based on the following principle: each type of object (vegetation, water bodies, towns or mosquito trapping results) is represented by a different data layer (Figure I.3). Overlaying layers according to their spatial references enable each data layer to be visualised and analysed separately (horizontal query within the same layer, e.g., which mosquito species were observed and in what abundance?). Additionally, the relationships between different data layers can also be investigated (vertical query between different layers, e.g., what type of land cover do we observe in places with the highest abundance of mosquitoes?).

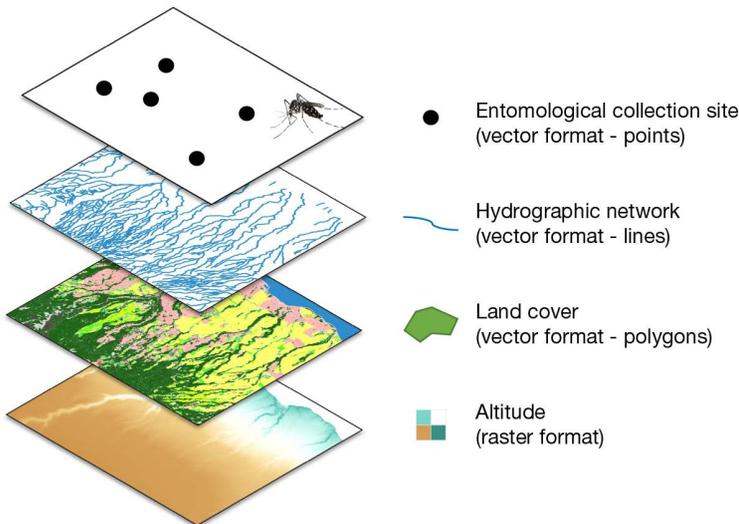


Figure I.3. Principles of data organisation in a GIS—thematic layers and modes of representation.

Within a single spatial data layer, objects are of the same kind and represented in two different modes:

– the vector mode or “vector”: in this mode, each object is represented in the form of polygons (e.g., a plot of vegetation), lines (e.g., a road or river) or points (e.g., location of a trapping site) [Figure I.3]. The most commonly used vector file format is “.shp” (shapefile);

– the matrix mode or “raster”: in this mode, spatial data is represented in the form of an image (or grid) composed of cells of the same size, called pixels (such as in a satellite image). The most commonly used raster file format is “.tif”.

In both cases, geographical data is combined with thematic data, providing information on the properties of the object. In vector mode, this thematic data is stored in an associated attribute table (e.g., for results of entomological traps, represented in form of points, the associated table will list: sample data, captured species, abundance, etc.). In raster mode, the pixel value contains the information represented (for a multispectral satellite image, the value of the pixels will be the reflectance value measured by the sensor).

Part 1

Spatial data for vector mosquito surveillance and associated diseases

The first part of this publication addresses the identification of different environmental, climate and demographic variables that exert an influence on the presence and dynamics of mosquito populations, with a particular focus on satellite images and their contribution to the study of vector-borne diseases.

The first chapter describes these different variables, in addition to the satellite remote sensing data and methodologies that facilitate access to this information. The following chapters present different approaches based on satellite imagery for the extraction of these variables: the use of spectral indices for water and vegetation (Chapter 2), the study of air temperatures (Chapter 3), the characterisation of human populations (Chapter 4) and finally the use of image texture analysis to characterise urban environments (Chapter 5).

Chapter 1

Relationships between vector mosquitoes and the environment: the role of satellite remote sensing methods

*Renaud Marti, Claire Teillet, Hobiniaina Anthonio Rakotoarison,
Florence Fournet*

Certain biological traits are shared among the nearly 3,600 species of mosquitoes (*Diptera: Culicidae*) described throughout the world, and these shape a significant part of their relationships with the environment and different meteorological factors. As ectotherms, the internal temperature of mosquitoes is controlled by the temperature of their environment, which consequently affects their physiology, behaviour, ecology and, more broadly, their survival (Reinhold *et al.*, 2018). The life cycle of mosquitoes is characterised by four distinct stages (Figure 1.1). The first three have an aquatic form: egg, larva (with 4 larval stages called instars each interrupted by a moult) and pupa. The final stage is that of the mature adult, which is associated with a period of reproduction and dispersal, and is distinguished by an aerial form. The aquatic stages develop in areas of water called larval habitats, the characteristics of which vary depending on the species of mosquito in question. Meteorological factors, such as temperature, precipitation or air humidity, are key determinants of the distribution and dynamics of mosquito populations by influencing the development of each of these stages, their transitions and the associated mortality rates.

Only female mosquitoes are haematophagous, and it is during this blood meal that they can be infected with a pathogen (virus, bacteria or parasite). Following the multiplication of this pathogen, the infected mosquito can then transmit it to a new host: the mosquito is then called a “vector”. In addition to exerting an influence on mosquito ecology, meteorological variations impact the entire chain of infection and development of pathogens, thereby influencing the relationships between human populations and vector mosquitoes (Morin *et al.*, 2013; Stresman, 2010). A reduction in the development period for mosquitoes, coupled with enhanced survival, results in accelerated reproductive cycles and a corresponding increase in the frequency of blood meals. An increase in ambient temperature is linked to a more rapid replication of the virus in the vector, or an accelerated parasite life cycle, which consequently results in a shorter extrinsic incubation period. A modest increase of a few degrees, which is compatible with current mosquito biology, has the effect of promoting

pathogen transmission, as evidenced by the aforementioned phenomena. The ongoing effects of climate change, however, remain challenging to assess due to their impact on a multitude of factors and processes across varying scales. They are accompanied by other significant collective changes, such as coevolutionary processes between arboviruses and parasites, vector arthropods and their vertebrate hosts. Furthermore, mosquitoes are subjected to local microclimate conditions that may diverge significantly from macroclimate measurements, particularly in heterogeneous urban environments (Wimberly *et al.*, 2020).

It is in tropical and subtropical environments that arboviruses and parasites transmitted by mosquitoes cause the greatest harm. However, countries situated in temperate latitudes are also affected by certain vector-borne diseases (VBDs). With the increase in international travel and trade, combined with the effects of environmental and climate change, the potential and proven risks of transmission or re-emergence mean that surveillance strategies need to be adapted. Species belonging to the three genera of mosquitoes *Anopheles*, *Aedes* and *Culex* are of major interest in the study of VBDs. *Anopheles* is highly prevalent across the globe, with the exception of the polar regions, and an accurate understanding of the geographical distribution of malaria transmission can only be achieved through a clear delineation between vector and non-vector anopheline species (Duvallat *et al.*, 2017). This distinction can sometimes be tricky due to the presence of so called “twin” species in the same ecological complex. These are morphologically indistinguishable (impossible to tell apart with the naked eye), but possess quite distinct biological, ecological and genetic traits. *Anopheles gambiae* is considered the primary malaria vector worldwide. However, it is a species complex, and therefore the use of molecular methods is required to distinguish between the various species (Coetzee *et al.*, 2013; Davidson, 1964). Out of the 540 species of *Anopheles* recorded (Duvallat *et al.*, 2017), around sixty have been identified as being malaria vectors, i.e., capable of transmitting parasites of the genus *Plasmodium* (Manguin *et al.*, 2008). Sub-Saharan Africa accounts for 85 to 95 % of malaria cases. The principal vector mosquitoes responsible for the transmission of this disease in this region are *An. gambiae*, *Anopheles arabiensis* and *Anopheles coluzzii* of the species complex *An. gambiae* and *Anopheles funestus*, of the species group *An. funestus* (Dahan-Moss *et al.*, 2020). Besides the transmission of parasites from the genus *Plasmodium*, anopheles are also vectors of filaria and o'nyong-nyong virus, an arbovirus of interest to public health, notably transmitted by *An. gambiae* and *An. funestus*, and described for the first time in Uganda in 1959.

The species *Aedes aegypti* and *Aedes albopictus*, belonging to the genus *Aedes* in the historical classification and classified as *Stegomyia* in the updated taxonomic system⁶ proposed by Harbach and Howard (2007), are originally associated with the hot and humid climates of tropical regions. These species have been identified as vectors for several flaviviruses, which are responsible for a range of diseases including dengue, Zika virus disease and yellow fever (tropics of Africa and America), as well as the alphavirus that causes chikungunya. Although *Ae. aegypti*, which exhibits superior vector competence, is the primary vector mosquito of these arboviruses globally, *Ae. albopictus* is an important secondary vector, notably in Asia and islands of the

6. <https://mosquito-taxonomic-inventory.myspecies.info/>

Indian Ocean. In recent years, the distribution range of this species has expanded to encompass regions situated at higher latitudes, including North America and southern Europe (Kraemer *et al.*, 2019).

Among the *Culex pipiens* species assemblage, *Cx. pipiens pipiens* and *Cx. quinquefasciatus* possess the widest distribution globally and have been observed in all urban and suburban temperate and tropical regions. In these areas, they are often identified as the primary vectors of arboviruses of interest to animal and public health. The genus *Culex* is notably associated with the transmission of two flaviviruses. One is responsible for Japanese encephalitis, which is observed in Asia (with *Culex tritaeniorhynchus* as the primary vector). The other is responsible for West Nile fever which is present in Europe, Africa, Asia and North and South America (with *Culex pipiens* in particular as a vector). The genera *Aedes* (particularly the species *Aedes* [*Aedimorphus*] *vexans*) and *Culex* (particularly the species

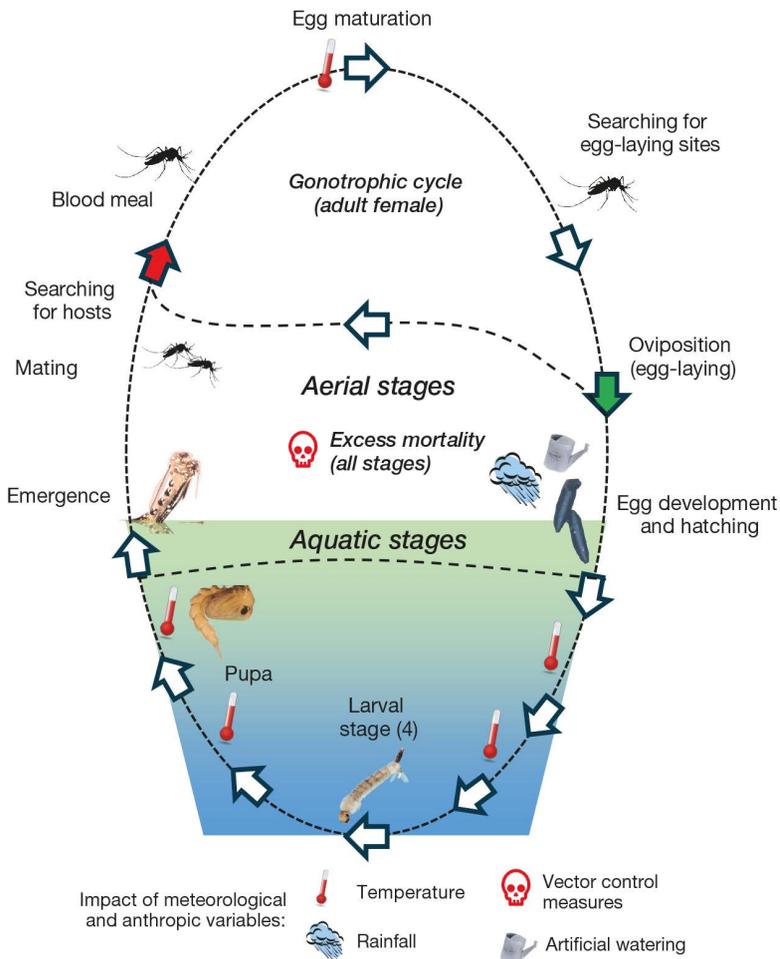


Figure 1.1. Life cycle of female mosquitoes illustrated with reference to the different aerial and aquatic stages of the species *Aedes albopictus*. Photographs of the larval stages: © Nicolas Henon—2022 TIGER (Tri-national Initiative Group of Entomology in Upper Rhine valley).

Culex poicilipes) can be linked to the transmission of Rift Valley fever, a viral zoonotic disease found in Africa and caused by a phlebovirus. Lastly, the three genera *Anopheles*, *Aedes* and *Culex* have also been identified as potential vectors for the transmission of lymphatic filariasis, a tropical parasitic disease caused by nematodes of the family *Filariidae*.

►► Relationships between vector mosquitoes and the environment

The vector system, which comprises infectious agents, vector mosquitoes and vertebrate hosts, is part of an environment. Defined by both biotic factors (vegetation cover, presence of natural larval habitats, food sources, predators, etc.) and abiotic factors (rainfall, temperature, humidity, sunlight, etc.), the environment influences the different population densities that make up the vector system as well as the frequency of their interactions.

Influence of meteorological factors on life cycle

General considerations

Temperature has a non-linear effect on the different developmental stages of mosquitoes, with threshold effects being observed. If the water temperature in the larval habitats is too low, development during aquatic stages is slowed or even stopped, whereas an increase in temperature up to a threshold value tends to accelerate development. An increase in air temperature, provided it remains within an optimal range, also promotes survival and activity during the adult stage, including the consumption of blood meals.

Each species of mosquito lays its eggs in specific larval habitats, whether natural (pools, tree hollows, rock pools, shells, standing water, watercourses, etc.) or artificial (of various shapes and sizes and containing a variety of matter, crawl spaces, etc.), filled with fresh water, brackish water or salt water and containing little or a lot of organic matter. In general, the availability of larval habitats is determined by precipitation levels (rainfall), which may follow a seasonal cycle depending on the latitude and the climate. Depending on the context and human practices, the flow of water into these habitats may also come from artificial sources: watering of indoor plants, outdoor gardens, irrigation or flooding of agricultural crops (e.g., rice fields). Air temperature affects how long the larval habitats contain water through the process of evapotranspiration. Lastly, precipitation promotes vegetation growth, which is both a source of sugar and a resting place for adult mosquitoes (Sallam *et al.*, 2017).

Focus on species of medical interest

Mosquitoes of the genus *Anopheles* lay their eggs one at a time on the water surface. For *An. gambiae*, an optimal temperature for eggs seems to be between 24 and 30 °C, with development periods of less than 7 days, whereas lower temperatures (12 °C) and higher temperatures (48 °C) considerably reduce egg viability (Impoinvil *et al.*, 2007). *Anopheles* eggs have very poor resistance to desiccation, with a markedly lower hatching rate on dry substrates (Duvallat *et al.*, 2017). Conversely, in humid conditions, eggs of

An. gambiae and *An. funestus* were able to hatch up to 12 and 10 days respectively after being laid. Larval growth of *An. gambiae* is also enhanced by increasing temperature, with threshold effects and non-linearities being observed: non-linear growth from an initial threshold of 16°C, a linear increase in growth between 22 and 28°C (optimum), non-linear decrease up to a maximum threshold of 34°C, and detrimental effects observed above these temperatures (Bayoh et Lindsay, 2003). Most anopheles species must endure unfavourable seasons—cold in temperate regions or drought in tropical regions—, which, depending on the species and biogeographical context, can lead to different survival strategies in adult females: diapause, quiescence, migration, aestivation or gonotrophic dissociation (cessation in egg production but continuing to take blood meals) [Duvallet *et al.*, 2017].

Aedes mosquitoes lay their eggs one by one on a substrate near water. During prolonged periods of drought, they are remarkably resistant to desiccation (Duvallet *et al.*, 2017). Furthermore, prolonged drought may in some contexts lead to more water being stored in the form of household containers or tanks, which in turn may provide new egg-laying sites (Pontes *et al.*, 2000). Conversely, excessive rainfall over a short period of time may result in egg destruction and leaching of larvae, leading to excess mortality during aquatic stages (Dieng *et al.*, 2012). The combination of temperature and humidity levels strongly determines the reproductive activity and survival of *Aedes*, which are particularly sensitive to low values of these two mutually influential parameters. At 80% humidity, the females of *Ae. aegypti* appear to survive longer and produce more eggs at 25°C than at 35°C, whereas at 60% humidity, oviposition appears to be drastically reduced or even inhibited (Costa *et al.*, 2010). In tropical environments, marked by alternating periods of high rainfall and dry periods, variations in temperature and humidity, even if small, can lead to seasonal increases or decreases in *Aedes* population density (Nasir *et al.*, 2017). In temperate environments, following the development of an adaptive mechanism by certain strains, *Ae. albopictus* females can lay diapause eggs which are resistant to unfavourable climate conditions (low temperature), resuming development and hatching once more favourable conditions return. Although the eggs of *Ae. aegypti* have been shown to have good cold tolerance in a laboratory setting, and even in the field (Kramer *et al.*, 2020), this diapause mechanism has yet to be observed in a natural environment, which may explain why they have struggled to establish themselves on the European continent in the long term. However, recent experimental results obtained for certain strains of *Ae. aegypti* associated with a temperate region (Buenos Aires, Argentina) appear to show adult forms which are sensitive to a reduction in photoperiod, resulting in a lower rate of egg hatching (Fischer *et al.*, 2019). If confirmed, this ability of *Ae. aegypti* to pause egg hatching in response to shorter photoperiods may facilitate its expansion into regions with colder winters.

Culex pipiens and *Cx. quinquefasciatus* are present in most inhabited areas of the world and are often directly linked to human activities, giving rise to their name of the northern and southern house mosquito (Farajollahi *et al.*, 2011). The females lay their eggs in batches, forming a raft that floats on the surface of the water in latrines, cesspools and drainage systems. The positive correlation between temperature and developmental rate in these species appears to be less pronounced above 24°C, with adult mortality rates increasing at higher temperatures (Ciota *et al.*, 2014). The worldwide distribution of *Cx. pipiens* complex mosquitoes can be explained by their

adaptation to environments that have been altered or modified by humans, which promotes their spread, and by their mixed feeding habits on birds and mammals (including humans). High rainfall, crop irrigation and poor maintenance of urban sewage systems, as well as elevated temperatures, favour the rapid and prolific growth of *Culex* populations: these mosquitoes thrive in polluted water rich in organic matter (Darriet, 2014). In temperate climates, adult females of *Cx. p. pipiens* overwinter in basements and sewage systems (Darriet, 2014).

Interactions between vector mosquitoes and their biotope

General considerations

In favourable environments, mosquitoes make the best use of available resources, with a high birth rate, short generation time and good dispersal ability (Duvallet *et al.*, 2017). Spatial distribution is thus determined by the structure of the environment, most importantly the availability of favourable water bodies for gravid females to lay eggs and for larvae to develop, and the presence of hosts (animals and humans) for blood meals. Due to their remarkable phenotypic plasticity, some mosquito species of the genera *Anopheles*, *Aedes* and *Culex* are perfectly adapted to anthropic environments and exhibit anthropophilic behaviour, which in turn promotes the transmission of pathogens (Cohuet *et al.*, 2010). In particular, these species thrive in urban and peri-urban environments, taking advantage of human activities to breed, lay eggs and find shelter. They target humans either as their primary source of blood meals or more opportunistically.

Environment of aquatic stages

Although all three juvenile stages—egg, larva and pupa—are aquatic, the larva is the only one that feeds, and most of the effects influencing its development and/or mortality rate are related to its population density (Beck-Johnson *et al.*, 2013). There are four larval stages (L1 to L4), and this is the only time when mosquitoes continuously increase in size. The period of larval development varies greatly (from a few days to several months, or more if the larvae enter diapause) is highly dependent on climatic, ecological and environmental conditions. Larval development conditions are determined by the quality and quantity of food (yeast, bacteria, micro-plankton, micro-algae, pollen grains, etc.), temperature, intra- and interspecific competition, and the presence of biotic or abiotic stress factors. Depending on the species in question, the availability of larval habitats is a key factor in population dynamics. The pupal stage is relatively short (in the order of a few days) and dependant on temperature; during this time the pupa does not feed.

Most anopheles species require freshwater larval habitats, although some have a high tolerance for salinity. It is worth noting that an adaptation to polluted water has been observed in some species in urban environments, such as *An. coluzzii* which has developed a tolerance to ammonia, one of the main pollutants found in urban larval habitats (Duvallet *et al.*, 2017). For *Ae. albopictus* and *Ae. aegypti*, all kinds of small containers prone to filling with water, both natural and artificial, are potential egg-laying sites (Estallo *et al.*, 2008). Mosquitoes of the genus *Aedes* lay their eggs one at a time on moist substrates near water. The larval habitats of *Culex* are numerous

and varied (cesspools, drainage systems, etc.), mainly found in the polluted waters and effluents of urban environments, where the larvae can feed on abundant organic matter (Darriet, 2014).

Environment of aerial stages and the gonotrophic cycle

The emergence of adult mosquitoes represents a short period of time (ten to fifteen minutes) during which they are extremely vulnerable to predation and environmental changes. Some species, termed “eurygamous”, require open spaces for mating (e.g., *Ae. caspius*, *Ae. detritus*, *Cx. pipiens pipiens*), while others, stenogamous species, can mate in restricted spaces (*Ae. albopictus*, *Cx. pipiens molestus*). After mating, the nulliparous females (those that have not yet laid eggs) search for a host to take their first blood meal, which is necessary to produce eggs in the case of so-called “anautozenous” species. During their few weeks of existence (between 3 and 5 weeks for most species in tropical regions), female behaviour is dominated by the completion of around a ten gonotrophic cycles (*gono-* relating to sexual reproduction, *-trophic* meaning having nutritional habits or requirements). Each cycle can be broadly broken down into three main stages, with a total duration of 3 to 5 days depending on the species and climate conditions (Duvallet *et al.*, 2017 ; Figures 1.1 and 1.2):

- search and selection of vertebrate host by the fasted female;
- rest for blood digestion and egg maturation;
- search for an egg-laying site and oviposition by the gravid female.

For *Anopheles*, blood meals are taken at dusk or during the night, and humans are either the primary or an opportunistic target depending on the species. Meals are taken either indoors (endophagic) or outdoors (exophagic). For *Aedes*, blood feeding occurs during the day. *Ae. aegypti* prefers to feed indoors on humans, whereas *Ae. albopictus* generally feeds outdoors and targets the mammals present in the environment. *Culex* bites at night and generally feeds on mammals (including humans), with *Cx. pipiens pipiens* preferring to target birds (*bird biting mosquito*) in rural (temperate) environments and humans in urban environments, with a high degree of endophily (distinct preference for human dwellings) for *Cx. pipiens molestus* (temperate environments) and *Cx. quinquefasciatus* (tropical environments) [Duvallet *et al.*, 2017; Farajollahi *et al.*, 2011].

Except when they emerge, adult mosquitoes of both sexes do not drink water, but they regularly feed on the sugar-bearing liquids excreted by plants, including flower nectar. Male mosquitoes rely exclusively on these sources of nutrition. Plant preference varies depending on the species of mosquito, geographical habitat and seasonal availability. When sugar sources are scarce, the females of certain species, such as *An. gambiae* and *Ae. aegypti*, may compensate for this by taking larger and more frequent blood meals (Barredo and Degennaro, 2020). Furthermore, some mosquito species are able to feed on sugary solutions and decomposing fruits, which can be found in high quantities in the waste produced by human activities.

In cities, the resources needed to sustain the life cycle of mosquitoes that have adapted to artificial habitats are available within a short distance of each other (mating partners, egg-laying sites, resting places, and blood meal hosts). At the same time, the active dispersal ability of *Ae. aegypti* and *Ae. albopictus* is considered low, with distances generally less than 1 km. Passive dispersal can however occur over great distances,

either in the form of desiccation-resistant eggs (transported by boats carrying used tyres or some plants such as *lucky bamboo* for example), or in their adult form being transported by road or railway, with the heavy traffic increasing the likelihood of new species becoming established (Eritja *et al.*, 2017). The active dispersal distance of *An. gambiae* is considered to be much greater, in the order of 2 to 3 km per generation (Duvallet *et al.*, 2017). The passive dispersal of this species is usually by means of human transport (plane, boat, car, etc.) carrying the adult form.

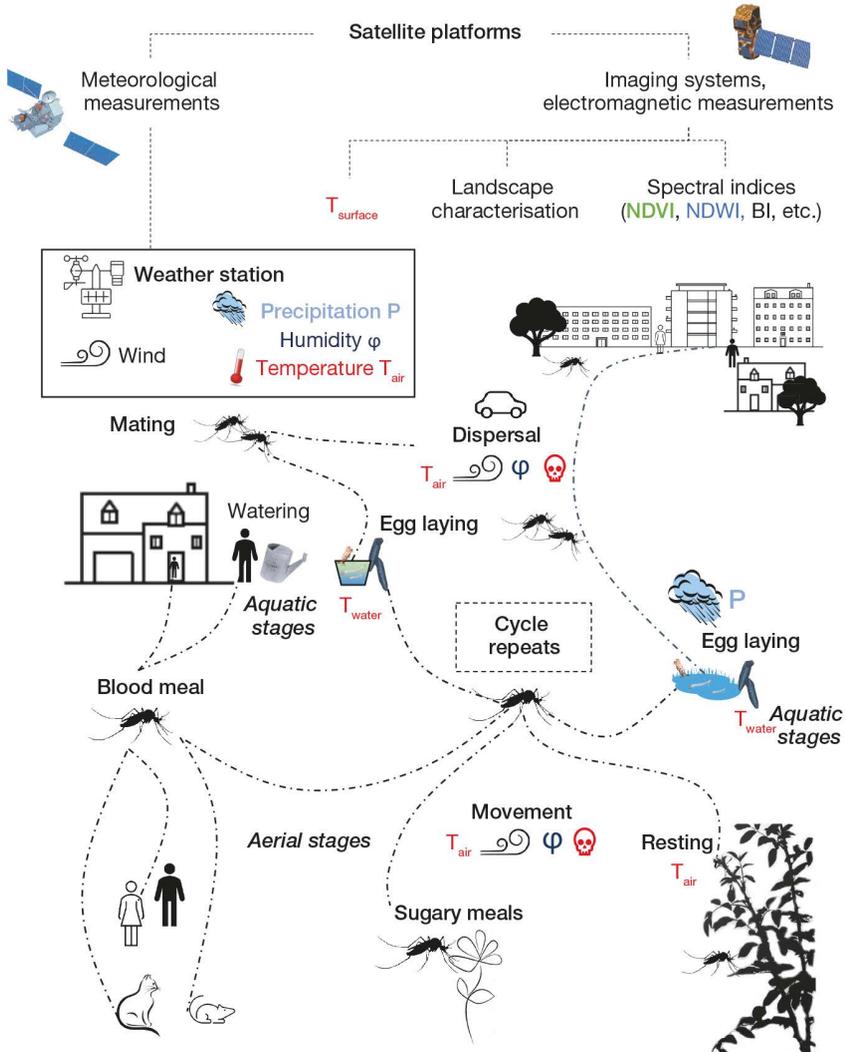


Figure 1.2. Life cycle of *Aedes albopictus* mosquitoes in an urban or peri-urban biotope.

In addition to an active dispersal ability, which allows them to complete their development cycle, passive dispersal by means of human transport (private vehicles, public transport, etc.) can help them to spread over greater distances.

Female mosquitoes are subject to a higher mortality rate (symbolised by a small red skull) when they move (in search of blood meal hosts or egg-laying sites).

The main sources of environmental data from remote sensing platforms are shown at the top of the figure.

► Description of the environment using satellite remote sensing methods

Benefits of remote sensing

Satellite remote sensing has global coverage and the methods used can be dynamically reproduced, in different locations and at different spatial and temporal resolutions. They have proven effective in estimating critical environmental variables such as certain meteorological parameters (temperature, precipitation, water vapour distribution, wind, etc.) or in providing a first level of description of the biophysical environment using spectral indices or more advanced products such as land use or land cover (Parselia *et al.*, 2019; see General Introduction and Chapters 2 to 5). The advantage of using these types of techniques in health ecology is that they are constantly improving thanks to technological innovation and the availability of new sensors and related products (Goetz *et al.*, 2000; Herbreteau *et al.*, 2007). The recent launch of the Copernicus programme⁷ and Sentinel satellites is particularly noteworthy in this regard, with the availability of multi-temporal images (weekly revisit rate) and derived products with decametric spatial resolution.

Cross-referencing remote sensing data with other socio-economic or land use parameters in a geographic information system (GIS) provides a better picture of the environment in a given region in terms of mosquito ecology (larval or adult stages), for example, by taking into account water management or urban waste (Kolimenakis *et al.*, 2021). Products derived from remote sensing data can serve as inputs to various modelling methods, allowing the spatial structure of the environment to be incorporated into models based on research and observation of statistical relationships (Chapter 6), knowledge-based models (Chapter 7), and even models highlighting the processes and biological traits of mosquitoes (Chapters 8 and 9).

In the context of human VBDs such as dengue fever, the detection and characterisation of urban environments should be a priority (Chapters 4 and 5), given the adaptation of certain medically important vectors to urban environments and the global increase in the urban population, which overtook the rural population in 2008 (Prasad *et al.*, 2016). Between 1950 and 2050, the proportion of people living in urban areas will increase from 2 in 10 to 7 in 10, with particularly strong urban growth in Asia and Africa in the coming decades.

Estimating meteorological variables using remote sensing

Weather stations on the planet's surface can measure temperature, humidity, hours of sunshine and wind, all of which can affect the development and spread of mosquitoes (see previous section). For sites not covered by these stations, some of these variables can be indirectly estimated using satellite remote sensing (Tables 1.1 and 1.2). Unlike geostationary satellites (orbiting at a distance of 36,000 km), sun-synchronous satellites (orbiting over the same point on the Earth at the same local mean solar time) orbit much closer to the Earth (usually between 500 and 1,000 km), providing spatial resolutions that are more suitable for studying ecological phenomena.

7. www.copernicus.eu/

Although the sensors on board meteorological satellites allow vertical temperature profiles to be constructed, environmental satellites can take direct measurements of land surface brightness temperature (Table 1.1). The instruments on board measure the emission spectrum in the thermal infrared range (3-14 μm). By using specific algorithms (e.g., Split Window), land surface emissivity (LSE) and land surface temperature (LST) can be decoupled and estimated as separate variables. The spatial structure of urban heat islands can be properly identified and delineated using land surface temperature products. These islands can influence mosquito ecology by creating more favourable local conditions in urban areas (Huraux *et al.*, 2017). However, when taking a physical approach, relating a surface temperature to an air temperature remains difficult and is generally modelled using a statistical relationship with *in situ* observations (Boser *et al.*, 2021; Weiss *et al.*, 2014; Chapter 3). The spatial resolution of sensors used to estimate surface temperatures is currently in the order of a hundred metres at best, or even a kilometre, making it impossible to work at finer scales, such as modelling urban neighbourhoods. The planned joint French-Indian satellite mission, TRISHNA (Thermal infraRed Imaging Satellite for High-resolution Natural resource Assessment), should provide unprecedented spatial resolution (57 m), overcoming some of the current operational constraints of thematic mapping (Lagouarde *et al.*, 2018). In some regional contexts, the Sentinel-3 SLSTR Level-2 LST, with its 290 km footprint, 1 km spatial resolution and daily revisit rate, may also prove useful (Shumilo *et al.*, 2019).

Table 1.1. Primary satellite remote sensing instruments currently in operation associated with the “surface temperature” variable.

Satellite platforms and associated products	Acquisition characteristics			
	Image footprint	Spatial resolution	Revisit rate	Examples of use in the context of VBDs
ECOSTRESS (Ecosystem Spaceborne Thermal Radiometer Experiment)	384 km	70 m	1 day	(Boser <i>et al.</i> , 2021)
Landsat 8/9 TIRS (Thermal Infrared Sensor)	185 km	100 m	16 days	(Ogashawara <i>et al.</i> , 2019)
MODIS MOD11A2 (Moderate Resolution Imaging Spectroradiometer)	1100 km	1 km	8 days	(Yue <i>et al.</i> , 2018)

In terms of precipitation measurement, products derived from the Global Precipitation Measurement (GPM, since 2010) mission provide continuity with the historical data collected by the Tropical Rainfall Measuring Mission (TRMM, 1997-2015). GSMaP (Global Satellite Mapping of Precipitation) products estimate precipitation every 3 days at a resolution of 10 km (Guilloteau *et al.*, 2014), and the Integrated Multi-satellite Retrievals for GPM (IMERG) product combines TRMM and GPM data to estimate precipitation between 2010 and the present day at a spatial resolution of 10 km and a temporal resolution ranging from almost real time to every day or month (Tsantalidou *et al.*, 2021).

In addition, several national, regional and global databases provide climate data (temperature, humidity, precipitation, wind, etc.) in the form of reanalyses generated by

climate models incorporating satellite observational data and *in situ* data. Bioclimatic variables can also be derived from monthly temperature and rainfall data in order to generate more biologically meaningful variables (mean diurnal range, isotherms, mean annual temperature and precipitation, etc.). With a resolution of nearly 1 km², the “WorldClim” database provides gridded weather and climate datasets for past (near-current) and future conditions. Such data has already been used in the context of studying the global distribution range of some vector mosquitoes (Kraemer *et al.*, 2015) and the transmission of certain VBDs (Tsheten *et al.*, 2021).

Table 1.2. Main global precipitation products incorporating satellite remote sensing data.

Meteorological products	Acquisition characteristics			
	Acquisition period	Spatial resolution	Acquisition frequency	Examples of use in the context of VBDs
Product 3B43 v6 Tropical Rainfall Measuring Mission TRMM	1997 -2015	30 km	Monthly average	(Scavuzzo <i>et al.</i> , 2018)
GSMaP/MVK v6 Global Precipitation Measurement Mission	2014 - present	10 km	3-day average	(Guilloteau <i>et al.</i> , 2014)
GPM IMERG Fusion of TRMM (2010-2015) and GPM (2014-present)	2010 - present	10 km	Near-real time, daily and monthly average	(Tsantalidou <i>et al.</i> , 2021)
Worldclim Climatic and bioclimatic variables	Past, present, future	1 km	Monthly average or variable dependent	(Tsheten <i>et al.</i> , 2021)

Describing the biophysical environment with remote sensing

Useful optical and radar sensors in the context of VBDs

Depending on their characteristics, remote sensing images can identify objects as small as a few centimetres (e.g., drone images) or images of the Earth’s surface taken daily (e.g., by the MODIS satellite). In the framework of the Copernicus programme, the European Space Agency (ESA) launched the “Sentinel” satellite constellation. The Sentinel-1A and Sentinel-1B satellites, which were put into orbit on 3 April 2014 and 25 April 2016 respectively, operate day and night and provide high-resolution (10 m) C-band radar images. Launched on 22 June 2015 and 7 March 2017 respectively, Sentinel-2A and Sentinel-2B produce high resolution optical images (primarily 10 and 20 m) which are particularly useful for monitoring vegetation, aquatic and agricultural surfaces, or even urban environments (Table 1.3). Several articles provide a reasonably comprehensive list of the sensors used in the study of vector mosquito ecology and VBD epidemiology, depending on the surfaces being identified (Catry *et al.*, 2018b; Herbreteau *et al.*, 2018). New missions should allow finer levels of detail in the description of the environment, with resolutions approaching those of airborne remote sensing: 30 cm for Pléiades Neo for example (Soubirane, 2019).

Table 1.3. Examples of optical and radar satellite remote sensors used in the study of vector mosquito ecology and VBD transmission.

Mode of acquisition and satellite platform	Acquisition characteristics			
	Image footprint	Spatial resolution (at nadir)	Revisit rate	Examples of use in the context of VBDs
(Time series) Sentinel-1	80 km	10 m	5 days	(Hardy <i>et al.</i> , 2019)
(Time series) Sentinel-2	290 km	10 or 20 m	5 to 10 days	(Tran <i>et al.</i> , 2019)
(Time series) Landsat 7 ETM + Landsat 8	185 km	15 m (P) 30 m (XS)	16 days	(Kofidou <i>et al.</i> , 2021)
(Time series) MODIS	2,230 km	250 m 500 m	1 day 7-day delay	(Troyo <i>et al.</i> , 2009)
(Programming) SPOT 6/7	60 km	1.5 m (P) 8 m (XS)	Programming on request	(Orta-Pineda <i>et al.</i> , 2021)
(Programming) Pléiades	20 km	0.7 m (P) 2.8 m (XS)	Programming on request	(Georganos <i>et al.</i> , 2020)

The letters (P) and (XS) in the “Spatial resolution” column refer respectively to panchromatic and multispectral images (a single wide-spectrum band or specific bands at certain wavelengths, e.g., red, green, blue, near-infrared, etc.).

Calculation of spectral indices

In the field of applied remote sensing, the scientific community has developed several indices using the combination of different spectral bands to characterise the biophysical properties of imaged surfaces (Table 1.4). The Normalized Difference Vegetation Index (NDVI) is a measure of chlorophyll content based on the spectral behaviour of vegetation. It is calculated by measuring the difference between electromagnetic radiation recorded in the near-infrared band (highly reflected by vegetation) and the red band (absorbed). NDVI is an index which ranges in value from -1 (no vegetation) to $+1$ (dense vegetation with high chlorophyll content). Additionally, other vegetation indices can be extracted from the data, such as EVI (Enhanced Vegetation Index), which estimates structural variations of the canopy, or SAVI (Soil Adjusted Vegetation Index) which incorporates soil brightness. The MNDWI (Modified Normalized Difference Water Index) and NDWI (Normalized Difference Water Index) are indices that facilitate the assessment and monitoring of surface water content. In contrast, the NDBI (*Normalized Difference Built-up Index*) and BI (*Brightness Index*) are used to detect the presence of “built-up” surfaces. Thresholding can be applied to each index to determine whether a surface contains vegetation (NDVI), water (MNDWI), or is a built-up area (NDBI) [Chapter 2].

Global and regional mapping products available

National space agencies and major scientific centres (such as NASA, ESA, DLR, JRC and the Theia data and services centre⁸) regularly provide products that utilise global or

8. www.theia-land.fr/en/theia-data-and-services-center/

regional data (Table 1.5). Despite certain inherent limitations of using this data (inconsistent quality, sometimes outdated), these products have proven to be particularly useful in instances where *in situ* observational data is lacking.

Table 1.4. Primary indices derived from remote sensing images to characterise the presence of vegetation, water and buildings.

Spectral indices	Product characteristics			
	Satellite imaging source	Surfaces characterised	Calculation method ⁽¹⁾	Examples of use in the context of VBDs
NDVI (Normalized Difference Vegetation Index)	Multispectral	Vegetated surfaces	Normalized difference of NIR and Red bands	(Richman <i>et al.</i> , 2018)
MNDWI (Modified Normalized Difference Water Index)	Multispectral	Surface water	Normalized difference between bands: Green or NIR and SWIR or MIR	(Malahlela <i>et al.</i> , 2018)
NDBI (Normalized Difference Built-up Index)	Multispectral	Built-up areas (buildings)	Normalized difference between SWIR and NIR bands	(Demets <i>et al.</i> , 2020)

NIR: near-infrared; SWIR: Short Wave Infrared; MIR: mid-infrared.

Table 1.5. Mapping products relevant to the ecology of certain vectors: building footprint, topography, land cover.

Product name	Product characteristics			
	Type of environment	Acquisition platform	Methods	Associated publication
GHSL	Building footprint	Sentinel-1 and 2	MASADA processing chain	(Pesaresi <i>et al.</i> , 2013)
GUF	Building footprint	TerraSAR-X	Texture and Support Vector Data Description	(Esch <i>et al.</i> , 2017)
SRTM	Topography	Space Shuttle Endeavour	Radar interferometry	(Slater <i>et al.</i> , 2006)
OSO (Metropolitan France)	Land cover	Sentinel-2	iota ² processing chain	(Inglada <i>et al.</i> , 2017)

Ad hoc maps and classifications

In many cases, global or regional land cover products may prove inadequate for representing the ecology of vectors and the transmission of VBDs. This is due to factors such as inappropriate or incomplete definitions of categories, unsuitable spatial resolutions, and the use of obsolete products. If users possess the technical skills, they can

create their own mapping products, either with the help of open-source processing chains (Table 1.6), or by building their own tools with the help of a toolbox specially designed for image processing (such as Orfeo Toolbox⁹ or SNAP¹⁰).

Table 1.6. Examples of open-source processing chains for satellite images which characterise the environmental features relevant to the ecological study of vector mosquitoes and VBDs.

Name of processing chain	Production characteristics			
	Satellite imaging source	Type of environment	Methods	Associated publication
Iota ²	Sentinel-1 and 2	Generalist	Pixel-oriented approach	(Inglada <i>et al.</i> , 2017)
Moringa	Sentinel-2 and VHRS (e.g., SPOT 6/7, Pléiades)	Generalist	Object-oriented approach	(Dupuy <i>et al.</i> , 2020)
<i>Urba-Opt</i>	Sentinel-2	Urban	Object-oriented approach	(Puissant <i>et al.</i> , 2019)
Fototex	Pléiades, Sentinel-2	Primarily urban, but can also be used for vegetation	Short-time Fourier transform (texture index)	(Teillet <i>et al.</i> , 2021)
WaterDetect	Sentinel-2	Surface water	Clustering algorithm Naive Bayes classifier	(Cordeiro <i>et al.</i> , 2021)

Users convert the biophysical environment to a digital format, compatible with a GIS, according to the vector's ecology and its interaction with the environment (see previous sections). This representation is achieved through the extraction of geographical objects from the images, or alternatively, through the utilisation of land cover categories pertaining to biophysical coverage (forest, agricultural land, etc.) and/or land use categories indicative of the manner in which the space is being used (e.g., logging or rice farming). This classification or “extraction” must be carried out in accordance with the available images (resolution, footprint, date), the method employed (object-based or pixel-based classification, supervised or unsupervised), the selection of an appropriate set of parameters, the availability of high-quality training data for supervised methods, and an understanding of the vector's ecology (species) within the biogeographical context of the study (climate, seasonality, habitats).

A number of large geographical typologies (which should be adapted to the specific species and transmission cycle in question) facilitate the structuring and characterisation of landscapes within the broader context of vector mosquito and vector-borne disease research:

- building footprint, which indicates human presence and has a significant impact on the surrounding environment (Troyo *et al.*, 2009). The surface area and height of

9. www.orfeo-toolbox.org/

10. <https://step.esa.int/main/download/snap-download/>

buildings (number of floors) can be used as an indicator of human population density, which in turn can be used to determine exposure to mosquito bites (see Chapter 4; Georganos *et al.*, 2020);

- a functional typology associated with building footprints, which may be indicative of the habits and behaviours of human populations and influence the risk of transmission. This includes areas of detached houses with gardens, multi-storey blocks of flats, industrial or commercial areas (e.g., markets), or even informal housing and housing under construction (Flamand, 2015). The type of human activity associated with a given building allows for the identification of the ecological characteristics relevant to vector mosquitoes (e.g., the density of larval habitats for oviposition) or the potential interactions between vector mosquitoes and humans (blood meal availability depending on the occupation rate of the location over the course of the day);

- transportation routes: roads of different sizes, indicative of human mobility, may be synonymous with ecological barriers. The suitability of roads and concrete pavements for egg-laying is contingent upon the type of surface, which may be conducive to egg-laying in the event of water accumulation and the presence of puddles and storm drains (Montalvo *et al.*, 2022);

- surface water, which, depending on its size and content (fresh water, salt water or brackish water), and the specific larval ecology of the species in question, may be considered a potential egg-laying site (e.g., *An. gambiae* and freshwater pools) or an ecological barrier (e.g., watercourses and lakes for *Ae. albopictus*);

- forested land and its evolution over time, taking into account the anthropogenic pressure exerted in the form of deforestation, urbanisation and the expansion of agricultural land, which disturbs balanced ecosystems and promotes human-vector contact (Orta-Pineda *et al.*, 2021; Stefani *et al.*, 2013);

- agricultural areas, which can serve as larval habitats, and notably rice fields which have been identified as a crucial breeding ground for anopheles (Diuk-Wasser *et al.*, 2006). Additionally, these areas act as mosquito bite exposure zones for individuals residing in close proximity.

- bare ground, which can play an important role at the start and end of the wet season (forming pools of stagnant water). Wetlands are also of particular importance in the case of *Anopheles* (Giardina *et al.*, 2015);

- vegetation, which plays an important role in urban environments, by maintaining a favourable microclimate for vector mosquitoes adapted to urban areas in the form of resting places (shaded, cool and wet, or a more temperate environment in winter), as well as increasing the presence of potential egg-laying sites or food sources for these mosquitoes. Depending on the vegetation in question, mosquitoes can feed on the sap and nectar of flowers and trees (Honorio *et al.*, 2009). The layout, height and type of vegetation, as well as the distance to a vegetated surface, are all important explanatory variables for mosquito abundance in urban environments, often considered in aggregate form (Manica *et al.*, 2016). A number of categories can thus be formed, such as “trees”, “sugar cane”, “stubble field”, “lawn”, or “sparsely vegetated land” (Machault *et al.*, 2014);

- environmental topography, which may be associated with a temperature gradient (altitude, exposure), a type of vegetation, or hydrological factors through a derived variable (slope) or calculation of a hydrological index (e.g., Topographic Wetness Index calculated by Homan *et al.* [2016] in the context of a malaria study). Topographical

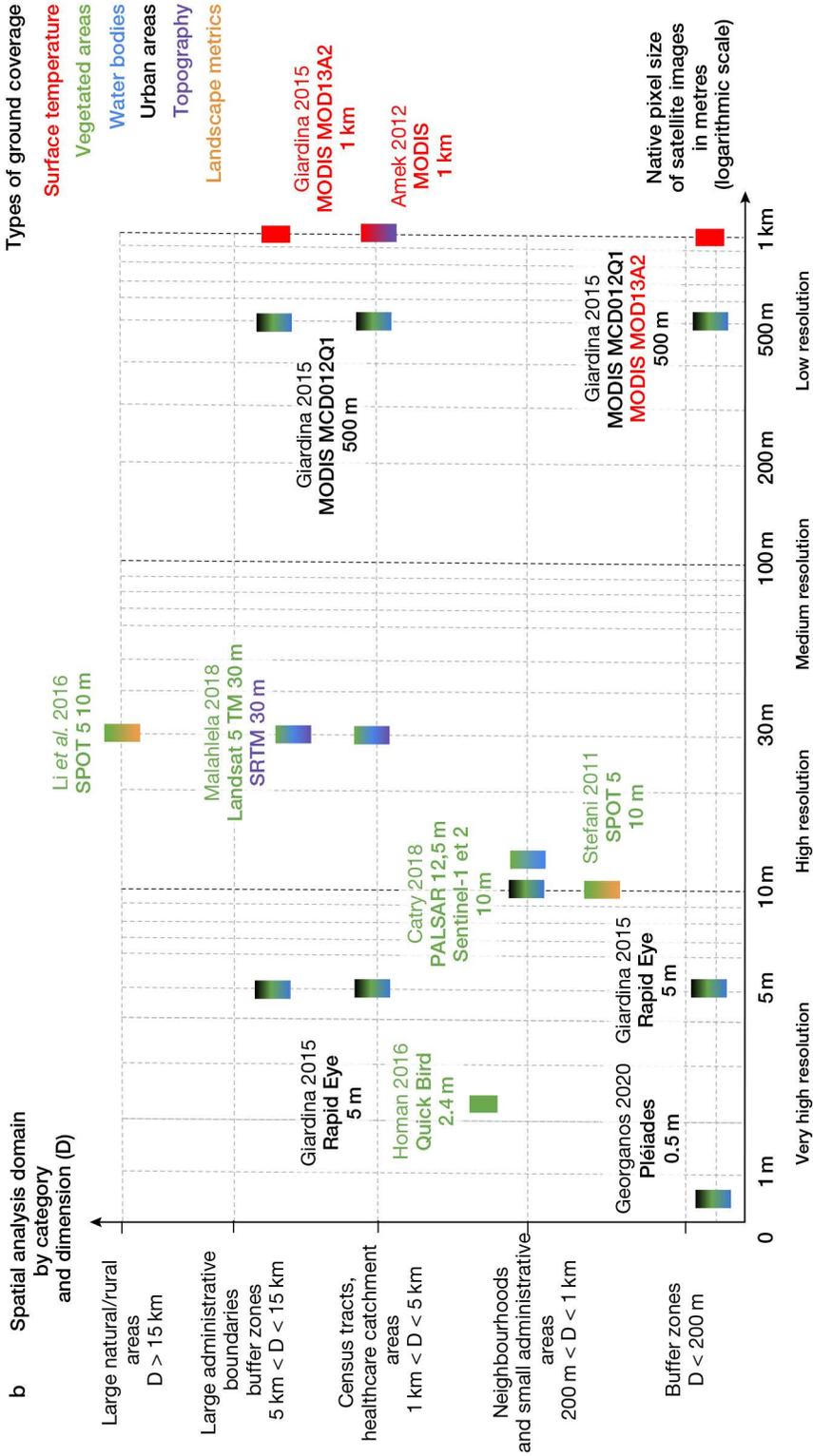


Figure 1.3. Examples of publications using remote sensing images to map the biophysical environment in the context of vector-borne disease transmission (Figure 1.3a: dengue; Figure 1.3b: malaria). Adapted from Marti *et al.* (2020).

The x-axis uses a logarithmic scale to represent native pixel size. The y-axis uses a qualitative dimension (categories) to represent the aggregated geographical unit in the spatial analysis.

data may be available, at medium spatial resolution, in the form of a global layer (such as SRTM, Table 1.5) or generated at high resolution by the user (e.g., digital model of terrain derived from radar interferometry, lidar data or a stereopair).

The literature offers a variety of landscape metrics that can be employed in the analysis of spectral index maps and for the characterisation of land cover and land use. These metrics facilitate the assessment of heterogeneity, fragmentation, composition (such as the degree of anthropisation, see Orta-Pineda *et al.* [2021]) or even the openness of a landscape (presence of forest clearings for *Anopheles*, see Stefani *et al.* [2013]). The interaction between humans and vectors is often considered to be merely a matter of distance from specific elements within a given landscape related to vector ecology, such as the proximity to a tyre pile, or a graveyard in the case of *Ae. albopictus* (Estallo *et al.*, 2008), or the border of the Amazon rainforest for *An. darlingi* (Li *et al.*, 2016). Alternatively, more complex geospatial statistics have been employed to examine human-vector interactions (Giardina *et al.*, 2015). It is also necessary to consider the potential role of habitats that are conducive to the survival and development of species that can act as intermediate hosts in the zoonotic cycle.

A number of studies on the ecology of vectors implicated in disease transmission have employed remote sensing data at different scales and dimensions derived from the various sensors cited in this chapter (optical and, less commonly, radar) in a range of biogeographical contexts (Figure 1.3). The majority of studies on dengue (*Ae. albopictus* et *Ae. aegypti*) and West Nile fever (*Cx. pipiens*, *Cx. tarsalis*) are primarily concerned with determining the spatial organisation of urban environments (Figure 1.3a), while studies on malaria (*An. gambiae* in particular) also include agroforestry environments at the interface of human habitats (Figure 1.3b). The risk maps derived from this data frequently diverge from the native pixel resolutions in order to aggregate data within administrative regions or other types of delineated areas, integrating additional data sources that characterise socio-economic factors and/or the practices of populations. In rare cases, satellite data is disaggregated to be cross-referenced with more finely-resolved data (Richman *et al.*, 2018) or is integrated into risk models with different resolutions (Demets *et al.*, 2020; Giardina *et al.*, 2015).

►► Références

- Andreo V., Cuervo P.F., Porcasi X., Lopez L., Guzman C., Scavuzzo C.M., 2021. Towards a workflow for operational mapping of *Aedes aegypti* at urban scale based on remote sensing. *Remote Sensing Applications: Society and Environment*, 23, 100554.
- Barredo E., Degennaro M., 2020. Not Just from Blood: Mosquito Nutrient Acquisition from Nectar Sources. *Trends Parasitol*, 36 (5), 473-484.
- Bayoh M.N., Lindsay S.W., 2003. Effect of temperature on the development of the aquatic stages of *Anopheles gambiae* sensu stricto (Diptera: Culicidae). *Bulletin of Entomological Research*, 93 (5), 375-381.
- Beck-Johnson L.M., Nelson W.A., Paaijmans K.P., Read A.F., Thomas M.B., Bjørnstad O.N., 2013. The effect of temperature on *Anopheles* mosquito population dynamics and the potential for malaria transmission. *PLoS ONE*, 8 (11), e79276.
- Boser A., Pascolini-Campbell M., Reager J., 2021. ECOSTRESS detects a 20% increase in evapotranspiration over agricultural lands in California, in *AGU Fall Meeting Abstracts*, 13-17 December 2021, (H35T-1270), AGU, New Orleans, USA.
- Catry T., Li Z., Roux E., Herbretreau V., Gurgel H., Mangeas M., Seyler F., Dessay N., 2018a. Wetlands and Malaria in the Amazon: Guidelines for the Use of Synthetic Aperture Radar Remote-Sensing. *International Journal of Environmental Research and Public Health*, 15 (3).

- Catry T., Pottier A., Marti R., Li Z., Roux E., Herbreteau V., Mangeas M., Demagistri L., Gurgel H., Dessay N., 2018b. Apports de la combinaison d'images satellites optique et RADAR dans l'étude des maladies à transmission vectorielle : cas du paludisme à la frontière Guyane française – Brésil. *Confins*, 37.
- Ciota A.T., Maticchiero A.C., Kilpatrick A.M., Kramer L.D., 2014. The effect of temperature on life history traits of *Culex* mosquitoes. *Journal of Medical Entomology*, 51 (1), 55-62.
- Coetzee M., Hunt R.H., Wilkerson R., Della Torre A., Coulibaly M.B., Besansky N.J., 2013. *Anopheles coluzzii* and *Anopheles amharicus*, new members of the *Anopheles gambiae* complex. *Zootaxa*, 3619, 246-274.
- Cohuet A., Harris C., Robert V., Fontenille D., 2010. Evolutionary forces on *Anopheles*: what makes a malaria vector? *Trends Parasitol*, 26 (3), 130-136.
- Cordeiro M.C.R., Martinez J.-M., Peña-Luque S., 2021. Automatic water detection from multidimensional hierarchical clustering for Sentinel-2 images and a comparison with Level 2A processors. *Remote Sensing of Environment*, 253, 112209.
- Costa E., Santos E., Correia J., Albuquerque C., 2010. Impact of small variations in temperature and humidity on the reproductive activity and survival of *Aedes aegypti* (Diptera, Culicidae). *Revista Brasileira de Entomologia*, 54, 488-493.
- Dahan-Moss Y., Hendershot A., Dhoogra M., Julius H., Zawada J., Kaiser M., Lobo N.F., Brooke B.D., Koekemoer L.L., 2020. Member species of the *Anopheles gambiae* complex can be misidentified as *Anopheles lesoni*. *Malaria Journal*, 19 (1), 89.
- Darriet F., 2014. *Des moustiques et des hommes : chronique d'une pullulation annoncée*, Marseille, IRD, 136 p. (coll. Didactiques).
- Davidson G., 1964. *Anopheles gambiae*, a complex of species. *Bulletin of the World Health Organization*, 31 (5), 625-634.
- Demets S., Ziemann A., Manore C., Russell C., 2020. Too big, too small, or just right? The influence of multispectral image size on mosquito population predictions in the greater Toronto area, in *Algorithms, Technologies, and Applications for Multispectral and Hyperspectral Imagery XXVI*, edited by Miguel Velez-Reyes, David W. Messinger, Proceedings of SPIE, 27 April-9 May 2020, 11392, Washington, USA. .
- Dieng H., Rahman G.M.S., Abu Hassan A., Che Salmah M.R., Satho T., Miake F., Boots M., Szalay A., 2012. The effects of simulated rainfall on immature population dynamics of *Aedes albopictus* and female oviposition. *International Journal of Biometeorology*, 56 (1), 113-120.
- Diuk-Wasser M.A., Dolo G., Bagayoko M., Sogoba N., Toure M.B., Moghaddam M., Manoukis N., Rian S., Traore S.F., Taylor C.E., 2006. Patterns of irrigated rice growth and malaria vector breeding in Mali using multi-temporal ERS-2 synthetic aperture radar. *International Journal of Remote Sensing*, 27 (3), 535-548.
- Dupuy S., Defrise L., Gaetano G., Andriamanga V., Rasoamalala E., 2020. Land cover maps of Antananarivo (capital of Madagascar) produced by processing multisource satellite imagery and geospatial reference data. *Data in Brief*, 31, 105952.
- Duvallet G., Fontenille D., Robert V., 2017. *Entomologie médicale et vétérinaire*, Versailles – Marseille, Quæ – IRD, 687 p.
- Eritja R., Palmer J.R.B., Roiz D., Sanpera-Calbet I., Bartumeus F., 2017. Direct Evidence of Adult *Aedes albopictus* Dispersal by Car. *Scientific Reports*, 7 (1), 14399.
- Esch T., Heldens W., Hirner A., Keil M., Marconcini M., Roth A., Zeidler J., Dech S., Strano E., 2017. Breaking new ground in mapping human settlements from space – The Global Urban Footprint. *ISPRS Journal of Photogrammetry and Remote Sensing*, 134, 30-42.
- Estallo E.L., Lamfri M.A., Scavuzzo C.M., Almeida F.F., Introiño M.V., Zaidenberg M., Almiron W.R., 2008. Models for predicting *Aedes aegypti* larval indices based on satellite images and climatic variables. *Journal of the American Mosquito Control Association*, 24 (3), 368-376.
- Farajollahi A., Fonseca D.M., Kramer L.D., Marm Kilpatrick A., 2011. "Bird biting" mosquitoes and human disease: a review of the role of *Culex pipiens* complex mosquitoes in epidemiology. *Infection, Genetics and Evolution, Journal of Molecular Epidemiology and Evolutionary Genetics of Infectious Diseases*, 11 (7), 1577-1585.

Fischer S., De Majo M.S., Di Battista C.M., Montini P., Loetti V., Campos R.E., 2019. Adaptation to temperate climates: Evidence of photoperiod-induced embryonic dormancy in *Aedes aegypti* in South America. *Journal of Insect Physiology*, 117, 103887.

Flamand C., 2015. *Étude des déterminants climatiques et environnementaux de la dengue en Guyane française*, thèse de doctorat en Santé publique – épidémiologie, Université Paris-Saclay (ComUE), 244 p.

Georganos S., Brousse O., Dujardin S., Linard C., Casey D., Milliones M., Parmentier B., Van Lipzig N.P.M., Demuzere M., Grippa T., Vanhuysse S., Mboga N., Andreo V., Snow R.W., Lennert M., 2020. Modelling and mapping the intra-urban spatial distribution of *Plasmodium falciparum* parasite rate using very-high-resolution satellite derived indicators. *International Journal of Health Geographics*, 19 (1), 38.

German A., Espinosa M.O., Abril M., Scavuzzo C.M., 2018. Exploring satellite based temporal forecast modelling of *Aedes aegypti* oviposition from an operational perspective. *Remote Sensing Applications: Society and Environment*, 11, 231-240.

Giardina F., Franke J., Vounatsou P., 2015. Geostatistical modelling of the malaria risk in Mozambique: effect of the spatial resolution when using remotely-sensed imagery. *Geospatial Health*, 10 (2), 333.

Goetz S.J., Prince S.D., Small J., 2000. Advances in satellite remote sensing of environmental variables for epidemiological applications. *Advances in Parasitology*, 47, 289-307.

Guilloteau C., Gosset M., Vignolles C., Alcoba M., Tourre Y.M., Lacaux J.-P., 2014. Impacts of Satellite-Based Rainfall Products on Predicting Spatial Patterns of Rift Valley Fever Vectors. *Journal of Hydrometeorology*, 15 (4), 1624-1635.

Harbach R.E., Howard T.M., 2007. Index of currently recognized mosquito species (Diptera: Culicidae). *European Mosquito Bulletin*, 23, 2-66.

Hardy A., Ettritch G., Cross D.E., Bunting P., Liywalii F., Sakala J., Silumesii A., Singini D., Smith M., Willis T., Thomas C.J., 2019. Automatic Detection of Open and Vegetated Water Bodies Using Sentinel 1 to Map African Malaria Vector Mosquito Breeding Habitats. *Remote Sensing*, 11 (5), 593.

Herbreteau V., Salem G., Souris M., Hugot J.-P., Gonzalez J.-P., 2007. Thirty years of use and improvement of remote sensing, applied to epidemiology: From early promises to lasting frustration. *Health & Place*, 13 (2), 400-403.

Herbreteau V., Kassié D., Roux E., Marti R., Catry T., Attoumane A., Révillion C., Commins J., Dessay N., Mangeas M., Tran A., 2018. Observer la Terre pour appréhender spatialement les inégalités de santé : regard historique et prospectif sur l'utilisation de la télédétection dans le domaine de la santé. *Confins*, 37.

Homan T., Maire N., Hiscox A., Di Pasquale A., Kiche I., Onoka K., Mweresa C., Mukabana W.R., Ross A., Smith T.A., Takken W., 2016. Spatially variable risk factors for malaria in a geographically heterogeneous landscape, western Kenya: an explorative study. *Malaria Journal*, 15, 1.

Honorio N.A., Nogueira R.M., Codeco C.T., Carvalho M.S., Cruz O.G., Magalhaes Mde A., De Araujo J.M., De Araujo E.S., Gomes M.Q., Pinheiro L.S., Da Silva Pinel C., Lourenco-De-Oliveira R., 2009. Spatial evaluation and modeling of Dengue seroprevalence and vector density in Rio de Janeiro, Brazil. *PLOS Neglected Tropical Diseases*, 3 (11), e545

Huraux T., Misslin R., Cebeillac A., Daudé É., Vaguet A., 2017. Modélisation de l'impact des îlots de chaleur urbains sur les dynamiques de population d'*Aedes aegypti*, vecteur de la dengue et du virus Zika, in *Spatial Analysis and GEomatics*, 6-9 novembre 2017, EA LITIS; UMR IDEES, Rouen, France.

Impoinvil D.E., Cardenas G.A., Gihture J.I., Mboga C.M., Beier J.C., 2007. Constant temperature and time period effects on *Anopheles gambiae* egg hatching. *Journal of the American Mosquito Control Association*, 23 (2), 124-130.

Inglada J., Vincent A., Arias M., Tardy B., Morin D., Rodes I., 2017. Operational High Resolution Land Cover Map Production at the Country Scale Using Satellite Image Time Series. *Remote Sensing*, 9 (1), 95.

Kofidou M., De Courcy Williams M., Nearchou A., Veletza S., Gemtzi A., Karakasiliotis I., 2021. Applying Remotely Sensed Environmental Information to Model Mosquito Populations. *Sustainability*, 13 (14), 7655.

- Kolimenakis A., Heinz S., Wilson M.L., Winkler V., Yakob L., Michaelakis A., Papachristos D., Richardson C., Horstick O., 2021. The role of urbanisation in the spread of *Aedes* mosquitoes and the diseases they transmit-A systematic review. *PLOS Neglected Tropical Diseases*, 15 (9), e0009631.
- Kraemer M.U., Sinka M.E., Duda K.A., Mylne A.Q., Shearer F.M., Barker C.M., Moore C.G., Carvalho R.G., Coelho G.E., Van Bortel W., Hendrickx G., Schaffner F., Elyazar I.R., Teng H.J., Brady O.J., Messina J.P., Pigott D.M., Scott T.W., Smith D.L., Wint G.R., Golding N., Hay S.I., 2015. The global distribution of the arbovirus vectors *Aedes aegypti* and *Ae. albopictus*. *Elife*, 4, e08347.
- Kraemer M.U.G., Reiner R.C. Jr., Brady O.J., Messina J.P., Gilbert M., Pigott D.M., Yi D., Johnson K., Earl L., Marczak L.B., Shirude S., Davis Weaver N., Bisanzio D., Perkins T.A., Lai S., Lu X., Jones P., Coelho G.E., Carvalho R.G., Van Bortel W., Marsboom C., Hendrickx G., Schaffner F., Moore C.G., Nax H.H., Bengtsson L., Wetter E., Tatem A.J., Brownstein J.S., Smith D.L., Lambrechts L., Cauchemez S., Linard C., Faria N.R., Pybus O.G., Scott T.W., Liu Q., Yu H., Wint G.R.W., Hay S.I., Golding N., 2019. Past and future spread of the arbovirus vectors *Aedes aegypti* and *Aedes albopictus*. *Nature Microbiology*, 4 (5), 854-863.
- Kramer I.M., Kress A., Klingelhofer D., Scherer C., Phuyal P., Kuch U., Ahrens B., Groneberg D.A., Dhimal M., Muller R., 2020. Does winter cold really limit the dengue vector *Aedes aegypti* in Europe? *Parasit Vectors*, 13 (1), 178.
- Lagouarde J., Bhattacharya B., Crébassol P., Gamet P., Babu S., Boulet G., Briottet X., Buddhiraju K., Cherchali S., Dadou I., Dedieu G., Gouhier M., Hagolle O., Irvine M., Jacob F., Kumar A., Kumar K., Laignel B., Mallick K., Murthy C., Olioso A., Otlé C., Pandya M., Raju P., Roujean J., Sekhar M., Shukla M., Singh S., Sobrino J., Ramakrishnan R., 2018. The Indian-French Trishna Mission: Earth Observation in the Thermal Infrared with High Spatio-Temporal Resolution, in *IGARSS 2018 – 2018 IEEE International Geoscience and Remote Sensing Symposium*, 22-27 July 2018, Valencia, Spain, GRSS – IEEE, 4078-4081.
- Li Z., Roux E., Dessay N., Girod R., Stefani A., Nacher M., Moiret A., Seyler F., 2016. Mapping a Knowledge-Based Malaria Hazard Index Related to Landscape Using Remote Sensing: Application to the Cross-Border Area between French Guiana and Brazil. *Remote Sensing*, 8 (4), 319.
- Lorenz C., Castro M.C., Trindade P.M.P., Nogueira M.L., De Oliveira Lage M., Quintanilha J.A., Parra M.C., Dibo M.R., Favaro E.A., Guirado M.M., Chiaravalloti-Neto F., 2020. Predicting *Aedes aegypti* infestation using landscape and thermal features. *Scientific Reports*, 10 (1), 21688.
- Machault V., Yébakima A., Etienne M., Vignolles C., Palany P., Tourre Y.M., Guérécheau M., Lacaux J.-P., 2014. Mapping Entomological Dengue Risk Levels in Martinique Using High-Resolution Remote-Sensing Environmental Data. *ISPRS International Journal of Geo-Information*, 3 (4), 1352-1371.
- Mala S., Jat M.K., 2019. Implications of meteorological and physiographical parameters on dengue fever occurrences in Delhi. *Science of The Total Environment*, 650, 2267-2283.
- Malahlela O.E., Olwoch J.M., Adjorlolo C., 2018. Evaluating Efficacy of Landsat-Derived Environmental Covariates for Predicting Malaria Distribution in Rural Villages of Vhembe District, South Africa. *Ecohealth*, 15 (1), 23-40.
- Manguin S., Carnevale P., Mouchet J., 2008. *Biodiversité du paludisme dans le monde*, Montrouge, John Libbey Eurotext, 428 p.
- Manica M., Filippini F., D'Alessandro A., Screti A., Neteler M., Rosa R., Solimini A., Della Torre A., Caputo B., 2016. Spatial and Temporal Hot Spots of *Aedes albopictus* Abundance inside and outside a South European Metropolitan Area. *PLOS Neglected Tropical Diseases*, 10 (6), e0004758.
- Marti R., Li Z., Catry T., Roux E., Mangeas M., Handschumacher P., Gaudart J., Tran A., Demagistri L., Faure J.-E., Carvajal J.J., Drumond B., Xu L., Herbreteau V., Gurgel H., Dessay N., Gong P., 2020. A Mapping Review on Urban Landscape Factors of Dengue Retrieved from Earth Observation Data, GIS Techniques, and Survey Questionnaires. *Remote Sensing*, 12 (6), 932.
- Mena N., Troyo A., Bonilla-Carrión R., Calderón-Arguedas O., 2011. Factores asociados con la incidencia de dengue en Costa Rica. *Revista Panamericana de Salud Pública*, 29 (4), 234-242.
- Montalvo T., Higueros A., Valsecchi A., Realp E., Vila C., Ortiz A., Peracho V., Figuerola J., 2022. Effectiveness of the Modification of Sewers to Reduce the Reproduction of *Culex pipiens* and *Aedes albopictus* in Barcelona, Spain. *Pathogens*, 11 (4).

- Morin C.W., Comrie A.C., Ernst K., 2013. Climate and dengue transmission: evidence and implications. *Environmental Health Perspectives*, 121 (11-12), 1264-1272.
- Nasir S., Jabeen F., Abbas S., Nasir I., Debboun M., 2017. Effect of Climatic Conditions and Water Bodies on Population Dynamics of the Dengue Vector, *Aedes aegypti* (Diptera: Culicidae). *Journal of Arthropod-Borne Diseases*, 11 (1), 50-59.
- Ogashawara I., Li L., Moreno-Madrinan M.J., 2019. Spatial-Temporal Assessment of Environmental Factors Related to Dengue Outbreaks in Sao Paulo, Brazil. *Geohealth*, 3 (8), 202-217.
- Orta-Pineda G., Abella-Medrano C.A., Suzan G., Serrano-Villagrana A., Ojeda-Flores R., 2021. Effects of landscape anthropization on sylvatic mosquito assemblages in a rainforest in Chiapas, Mexico. *Acta Tropica*, 216, 105849.
- Parselia E., Kontoes C., Tsouni A., Hadjichristodoulou C., Kioutsoukis I., Magiorkinis G., Stilianakis N.I., 2019. Satellite Earth Observation Data in Epidemiological Modeling of Malaria, Dengue and West Nile Virus: A Scoping Review. *Remote Sensing*, 11 (16), 1862.
- Pesaresi M., G. Huadong G., Blaes X., Ehrlich D., Ferri S., Gueguen L., Halkia M., Kauffmann M., Kemper T., Lu L., Marin-Herrera M., Ouzounis G., Scavazon M., Soille P., Syrris V., Zanchetta L., 2013. A Global Human Settlement Layer From Optical HR/VHR RS Data: Concept and First Results. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 6 (5), 2102-2131.
- Pontes R.J., Freeman J., Oliveira-Lima J.W., Hodgson J.C., Spielman A., 2000. Vector densities that potentiate dengue outbreaks in a Brazilian city. *American Journal of Tropical Medicine and Hygiene*, 62 (3), 378-383.
- Prasad A., Gray C.B., Ross A., Kano M., 2016. Metrics in Urban Health: Current Developments and Future Prospects. *Annual Review of Public Health*, 37, 113-133.
- Puissant A., Sellé A., Baghdadi N., Thierion V., Le Bris A., Roujean J., 2019. The 'urban' component of the French Land Data and Services Centre (Theia), in *2019 Joint Urban Remote Sensing Event (JURSE)*, 22-24 May 2019, Vannes, France, IEEE, 1-4.
- Reinhold J.M., Lazzari C.R., Lahondere C., 2018. Effects of the Environmental Temperature on *Aedes aegypti* and *Aedes albopictus* Mosquitoes: A Review. *Insects*, 9 (4).
- Richman R., Diallo D., Diallo M., Sall A.A., Faye O., Diagne C.T., Dia I., Weaver S.C., Hanley K.A., Buenemann M., 2018. Ecological niche modeling of *Aedes* mosquito vectors of chikungunya virus in southeastern Senegal. *Parasit Vectors*, 11 (1), 255.
- Sallam M.F., Fizer C., Pilant A.N., Whung P.Y., 2017. Systematic Review: Land Cover, Meteorological, and Socioeconomic Determinants of *Aedes* Mosquito Habitat for Risk Mapping. *International Journal of Environmental Research and Public Health*, 14 (10).
- Scavuzzo J.M., Trucco F., Espinosa M., Tauro C.B., Abril M., Scavuzzo C.M., Frery A.C., 2018. Modeling Dengue vector population using remotely sensed data and machine learning. *Acta Tropica*, 185, 167-175.
- Shumilo L., Kussul N., Shelestov A., Korsunskaya Y., Yailymov B., 2019. Sentinel-3 Urban Heat Island Monitoring and analysis for Kyiv Based on Vector Data, in *2019 10th International Conference on Dependable Systems, Services and Technologies (DESSERT)*, 5-7 June 2019, Leeds, United Kingdom, IEEE, 131-135.
- Slater J.A., Garvey G., Johnston C., Haase J., Heady B., Kroenung G., Little J., 2006. The SRTM Data "Finishing" Process and Products. *Photogrammetric Engineering & Remote Sensing*, 72 (3), 237-247.
- Soubirane J., 2019. Shaping the Future of Earth Observation with Pléiades Neo, in *2019 9th International Conference on Recent Advances in Space Technologies (RAST)*, 11-14 June 2019, Istanbul, Turkey, 399-401.
- Stefani A., Roux E., Fotsing J.M., Carme B., 2011. Studying relationships between environment and malaria incidence in Camopi (French Guiana) through the objective selection of buffer-based landscape characterisations. *International Journal of Health Geographics*, 10, 65.
- Stefani A., Dusfour I., Correa A.P., Cruz M.C., Dessay N., Galardo A.K., Galardo C.D., Girod R., Gomes M.S., Gurgel H., Lima A.C., Moreno E.S., Musset L., Nacher M., Soares A.C., Carme B., Roux E., 2013. Land cover, land use and malaria in the Amazon: a systematic literature review of studies using remotely sensed data. *Malaria Journal*, 12, 192.

- Stresman G.H., 2010. Beyond temperature and precipitation: ecological risk factors that modify malaria transmission. *Acta Tropica*, 116 (3), 167-172.
- Teillet C., Pillot B., Catry T., Demagistri L., Lyszczarz D., Lang M., Couteron P., Barbier N., Adou Kouassi A., Gunther Q., Dessay N., 2021. Fast Unsupervised Multi-Scale Characterization of Urban Landscapes Based on Earth Observation Data. *Remote Sensing*, 13 (12), 2398.
- Tran A., Fall A.G., Biteye B., Ciss M., Gimonneau G., Castets M., Seck M.T., Chevalier V., 2019. Spatial Modeling of Mosquito Vectors for Rift Valley Fever Virus in Northern Senegal: Integrating Satellite-Derived Meteorological Estimates in Population Dynamics Models. *Remote Sensing*, 11 (9), 1024.
- Troyo A., Fuller D.O., Calderón-Arguedas O., Solano M.E., Beier J.C., 2009. Urban structure and dengue incidence in Puntarenas, Costa Rica. *Singapore Journal of Tropical Geography*, 30 (2), 265-282.
- Tsantalidou A., Parselia E., Arvanitakis G., Kyratzi K., Gewehr S., Vakali A., Kontoes C., 2021. MAMOTH: An Earth Observational Data-Driven Model for Mosquitoes Abundance Prediction. *Remote Sensing*, 13 (13), 2557.
- Tsheten T., Clements A.C.A., Gray D.J., Wangdi K., 2021. Dengue risk assessment using multicriteria decision analysis: A case study of Bhutan. *PLOS Neglected Tropical Diseases*, 15 (2), e0009021.
- Weiss D.J., Bhatt S., Mappin B., Van Boeckel T.P., Smith D.L., Hay S.I., Gething P.W., 2014. Air temperature suitability for *Plasmodium falciparum* malaria transmission in Africa 2000-2012: a high-resolution spatiotemporal prediction. *Malaria Journal*, 13, 171.
- Wimberly M.C., Davis J.K., Evans M.V., Hess A., Newberry P.M., Solano-Asamoah N., Murdock C.C., 2020. Land cover affects microclimate and temperature suitability for arbovirus transmission in an urban landscape. *PLOS Neglected Tropical Diseases*, 14 (9), e0008614.
- Yue Y., Sun J., Liu X., Ren D., Liu Q., Xiao X., Lu L., 2018. Spatial analysis of dengue fever and exploration of its environmental and socio-economic risk factors using ordinary least squares: A case study in five districts of Guangzhou City, China, 2014. *International Journal of Infectious Diseases*, 75, 39-48.

Chapter 2

Spectral indices and classifications of multispectral images for vector risk mapping

Annelise Tran, Renaud Marti, Vincent Herbreteau

Traditional methods of analysing remotely sensed multispectral images¹¹ include image classification and the calculation of spectral indices. These two approaches exploit the pre-existing correlation between the attributes of the surfaces under observation and their respective spectral signatures.

Classification methods enable regions of an image with similar spectral responses to be grouped into the same class. Two types of classification methods can be used to produce land cover maps from a multispectral image. Supervised classification methods employ algorithms (maximum likelihood, Random Forest, etc.) to distinguish between the various land cover categories present in the image (water, urban areas, vegetation, bare ground, etc.) which are defined in advance by the user (Rodriguez-Galiano *et al.*, 2012). It is thus necessary to create a training database for each category to be mapped prior to commencing the actual classification process. Unsupervised classification methods (e.g., K-means, Isodata) are more automated processes which do not require prior knowledge of the categories (Duda and Canty, 2002). Both of these methods can be applied to classify each pixel of the image (pixel-oriented approach) or objects (object-based approach) by grouping neighbouring pixels with similar reflectance values before classification by using an image segmentation process (Blaschke *et al.*, 2014). In both instances, the collection of validation data on the ground is essential for the evaluation of classification quality.

Spectral indices exploit the combination of different bands of a multispectral image to highlight certain properties of the surfaces observed (see Chapter 1). For example, the spectral signature of vegetation exhibits high reflectance in the near-infrared band and, conversely, low reflectance in the red band (Figure 1.1). To highlight the areas of vegetation in a multispectral image (Figure 2.1a), vegetation indices use the difference (such as NDVI, Normalized Difference Vegetation Index, see Chapter 1) or the ratio between reflectance values in the near-infrared (NIR) and red (R) bands [Figure 2.1b]:

$$NDVI = \frac{PIR - R}{PIR + R} \text{ (Equation 1)}$$

11. See general introduction.

For water bodies, whose spectral signatures exhibit a decrease in reflectance from the blue to infrared wavelength ranges, water indices (Figure 2.1c) exploit the differences between reflectance values in the green (G) and the near- or mid-infrared (MIR) bands, such is the case for the MNDWI (Modified Normalized Difference Water Index):

$$MNDWI = \frac{V - MIR}{V + MIR} \text{ (Equation 2)}$$

The application of a threshold to the spectral indices facilitates the identification and mapping of areas of ecological interest for mosquitoes (Figure 2.1d).

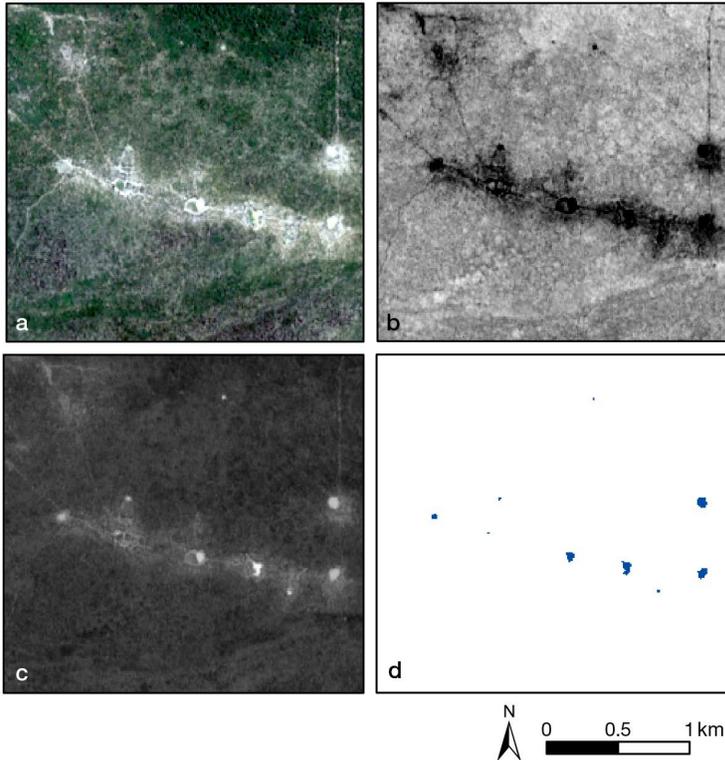


Figure 2.1. Examples of vegetation and water index calculations from a multispectral image.

(a) Extract of a Sentinel-2 image, natural colour composite image. (b) Vegetation index. (c) Water index – from black, areas with low index values, to white, areas with the highest index values. (d) Result of applying image thresholding to a water index – surface water appears in blue.

This chapter presents different examples of the application of traditional multispectral image interpretation methods. The first uses a land cover map derived from a high spatial resolution multispectral image to map the local distribution of *Anopheles* mosquitoes in Camargue. In the second example, a water index is calculated from medium spatial resolution images and used in a predictive risk model for the occurrence of human cases of West Nile fever in Europe at a continental level. Finally, we present an illustrative example of an online tool, Sen2Extract, which facilitates the straightforward extraction of time series of spectral indices.

► Mapping land cover using remotely sensed data in order to model the distribution of *Anopheles* mosquitoes in Camargue

Context

The Camargue is located in the Rhône delta in southern France (Figure 2.2a), with a Mediterranean climate characterised by hot and dry summers, mild and wet winters, and high rainfall in autumn and spring. The region is sparsely populated and comprises a mosaic of wetland landscapes, under the influence of saltwater to the south and with agricultural land to the north. The latter is primarily composed of rice fields and meadows.

The diverse wetland habitats of Camargue provide optimal conditions for the proliferation of mosquitoes, which are prevalent in significant numbers, and the transmission of associated diseases. Malaria was endemic to this region until the beginning of the 20th century (Jetten, Takken, 1994), and subsequently eliminated in the 1950's through the use of insecticides (essentially DDT) and chemical treatments/quinine-based prophylaxis. More recently, research projects have been conducted with the objective of analysing the potential resurgence of malaria in this region in the context of climate change (Ponçon *et al.*, 2008). These studies identified *Anopheles* (*Anopheles*) *hyrcanus* (Pallas), which has rice fields as its primary larval habitat, as a potential vector (Ponçon *et al.*, 2007).

In the example shown below, the relationships between land cover and entomological sampling data on the presence of *An. hyrcanus* presence were analysed in order to map the distribution of this species across Camargue (Tran *et al.*, 2008).

Data

Entomological data

A total of 80 potential larval habitats (different types of wetlands, rice fields) were visited on a monthly basis during the period of mosquito activity, from April to October 2006, by teams from the French National Research Institute for Sustainable Development (IRD; Institut de recherche pour le développement). Mosquito larvae were collected according to standard entomological protocol and identified at a species level (Ponçon *et al.*, 2007).

Remote sensing data

Two Landsat 7 Enhanced Thematic Mapper (ETM+, spatial resolution of 30 m) images were acquired during two contrasting periods (July and October) in order to map the main types of land cover in the area (Figure 2.2a).

Methods

Land cover classification

An object-based supervised classification (eCognition software¹²) was employed, utilising a training dataset derived from field observations and developed in conjunction with entomologists, with the objective of mapping the habitats of several species

12. <https://geospatial.trimble.com/en/products/software/trimble-ecognition>

of interest across five distinct ecotypes (Figure 2.2b). This typology includes all potential larval habitats for *Anopheles* mosquitoes in the region: open water, rice fields, reedbeds, bulrush and rush marshes.

Statistical analyses

Correlations between the presence or absence of *An. hyrcanus* larvae (response variable) and the type of larval habitat and distance to rice fields (explanatory variable) were studied using a logistic regression model.

Key results

Rice fields, bulrush marshes and reedbeds were identified by the logistic regression model as being highly correlated with the presence of *An. hyrcanus* larvae. These results confirm the potential of biotopes other than rice fields to serve as larval habitats for *An. hyrcanus*, which can explain the presence of the species at the end of summer and in autumn once the rice fields have been drained. Additionally, a map showing the probability of the species being present was generated across the entire study area using the land cover map (Figure 2.2c). This map was able to be validated by field data on the distribution of adult *An. hyrcanus* mosquitoes. It is important to note that this same approach was applied to data on other species of *Anopheles* and *Culex* mosquitoes. The results demonstrated that the distribution of larval habitats in wetlands influenced the spatial distribution of adult member of the species in cases where these species were highly dependent on a certain type of larval habitat (Cailly *et al.*, 2011).

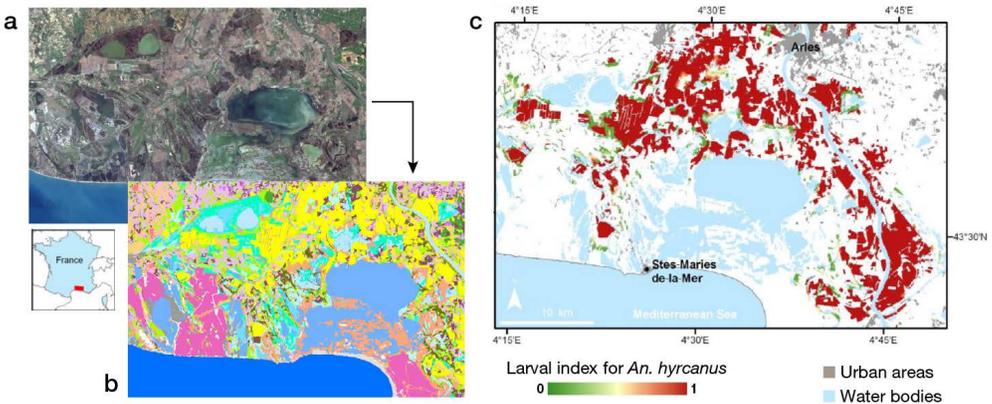


Figure 2.2. Map of the larval index in Camargue, derived from optical satellite imagery, highlighting the areas that are favourable to the *Anopheles hyrcanus* mosquito. Adapted from Tran *et al.*, 2008.

(a) Background: Landsat ETM+ image. (b) Results of classification into different types of land cover. (c) Larval index for the species *An. hyrcanus*, ranging from 0 (highly unfavourable area) to 1 (highly favourable).

► Spectral indices derived from remote sensing images employed as environmental factors in the analysis of human cases of West Nile fever in Europe

Context

West Nile virus (WNV), the causative agent of West Nile fever (WNF), is transmitted by bites from mosquitoes of the genus *Culex*. Wild and domestic birds represent the primary hosts, with occasional transmission to humans or horses, which are considered accidental hosts and epidemiological dead ends. In Europe, the circulation of WNV has been periodically confirmed in recent decades in several countries of the Mediterranean basin. However, during the particularly warm summer of 2010, a significant number of human cases of WNF were reported in areas which had previously been unscathed (Paz *et al.*, 2013). The study presented below sought to identify the principal environmental factors that contribute to the risk of human cases of WNF in Europe. To this end, it employed the use of spectral indices derived from remote sensing data (Tran *et al.*, 2014).

Data

Epidemiological data

Data on confirmed human cases of WNF was catalogued at the district level in Europe from 2002 to 2013. For each year, a district ($n = 1113$) was considered as “infected” if at least once human case of WNF was reported in this district. If no cases were reported then it was considered “not infected”.

Remote sensing data

A time series of MODIS images was downloaded in order to cover the whole of Europe, with a frequency of 8 days throughout the period in question.

Other spatially explicit data

Other georeferenced data was used to characterise, at a district level, potential risk factors for the incidence of human cases of WNF: temperature, population, bird migration routes and wetlands.

Methods

Calculation of vegetation and water indices

The vegetation index (NDVI) and water index (MNDWI) presented at the beginning of this chapter (Equations 1 and 2) were calculated from a time series of MODIS images.

Statistical analyses

Correlations between WNV infection status (response variable) and population, presence of wetlands, presence of bird migration routes, temperature anomaly values, NDVI and MNDWI (explanatory variables) were studied using univariate analysis (testing variables one by one), then using multivariate analysis, combining the most

significant variables using a logistic regression model. The model was constructed using epidemiological data from the period 2001 to 2011, while the predictive capacity was tested using data from the period 2012 to 2013.

Results

The most effective model for predicting the infected/uninfected status of human cases of WNV at the district level incorporates temperature anomalies during July, the MNDWI at the beginning of June (9-16 June), the prevalence of WNV in the previous year, the presence of wetlands, the type of bird migration routes, and human population levels (Figure 2.3a) as explanatory variables. These variables were found to be significantly and positively correlated with the probability of infection (Tran *et al.*, 2014). Using this logistic regression model, maps predicting the probability of WNV infection per district can be produced on an annual basis (Figure 2.3b). The results demonstrate the significance of water bodies in the transmission of WNV in Europe and the utility of a straightforward water index (MNDWI derived from MODIS images) in identifying surface water levels above the seasonal average in June, which can foster the proliferation of mosquitoes during the summer months and the transmission of the virus to humans at the end of summer/start of autumn.

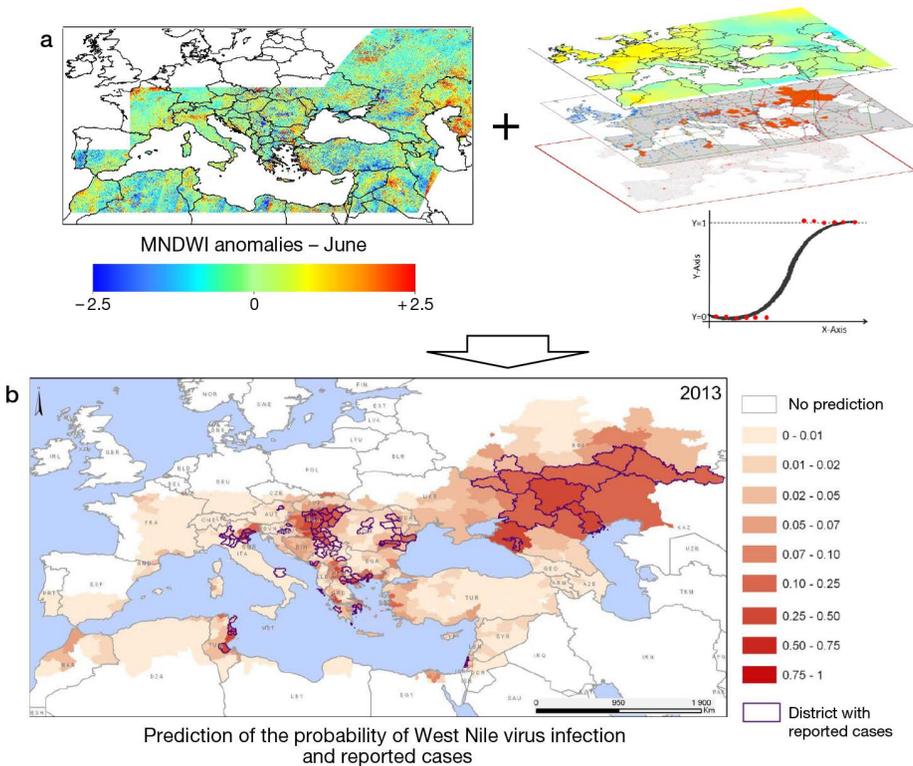


Figure 2.3. Risk map for the transmission of WNV in Europe based on environmental and meteorological indicators. Adapted from Tran *et al.* (2014).

(a) Analysis of a water index and various environmental predictors of WNV using a logistic regression model. (b) Infection risk map.

► Automated production of spectral indices: example of the Sen2Extract tool

The field of remote sensing remains a technically demanding area of study, requiring a comprehensive understanding of the methodologies, software, and algorithms employed in image processing, as well as the wide array of spatial data available. It is therefore unsurprising that the majority of satellite imagery utilised in the field of health is analysed indirectly with pre-calculated products, such as spectral indices, which are readily accessible and free to download. These products have enabled a community of users, including epidemiologists and statisticians with no prior knowledge of remote sensing, to use and incorporate this technology in their studies. While these indices may not be best suited to the characteristics of the area of study or the problem at hand, in numerous instances their consistent availability across space and time has demonstrated their utility as invaluable indicators of seasonal fluctuations.

It was from this observation that tools were developed to automate the production of spectral indices from Sentinel-2A and 2B satellite images of the European Space Agency's (ESA) Copernicus programme. As their name suggests, these two satellites were designed for the purpose of environmental monitoring. Consequently, they may prove to be a valuable resource for the monitoring of environmental diseases. They provide optical images covering the majority of the globe with a spatial resolution of 10 m, and are made available free of charge every five days, since mid-2015. Among the tools developed is Sen2Extract, an online application¹³ that facilitates the straightforward extraction of time series of spectral indices for a specified period and enables the selection of sites according to the user's requirements. A simple and intuitive interface (Figure 2.4) prompts the user to select their sites of interest by uploading a vector data file (points or polygons, in the zipped shapefile format), and then to choose the date range and indices to be calculated (NDVI and different water indices). Once the calculation is complete, an email is sent containing a download link for the results of the query. These results are presented in the form of a data table in CSV format. For each spatial entity (polygon or point) and each date, descriptive statistics are calculated for the values of the chosen index (mean, minimum, maximum, standard deviation, median, 25th and 75th percentiles, number of cloudy or 'no data' pixels). The final file is therefore very small (several kilobytes), in comparison to the size of a Sentinel-2 image and its post-processing, which can reach almost one gigabyte. This enables these images to be used in countries with limited internet access.

Sen2Extract represents the user interface for a set of algorithms developed as part of the S2-Malaria project (Tosca call for research proposals, funded by CNES from 2017 to 2020). The objective was to produce useful satellite observational data for monitoring diseases. These algorithms were then optimised to assess the impacts of cyclones (RenovRisk Interreg project from 2018 to 2020). The Sen2Chain utility, written in Python, is situated upstream of Sen2Extract and constitutes the core of the production process for these indices. It allows Sentinel-2 images to be downloaded from two main catalogues, the ESA's Copernicus Open Access Hub¹⁴ and the CNES's Sentinel Products Exploitation Platform¹⁵ (PEPS). It integrates the ESA's Sen2Cor

13. <https://web.seas-oi.org/sen2extract/>

14. <https://scihub.copernicus.eu/>

15. <https://peps.cnes.fr>

algorithm, which is used to perform atmospheric, terrain and cirrus corrections of raw images (level 1A), with the objective of obtaining level 2A data which is required for comparisons between different dates and scenes. This utility enables the generation of different spectral indices, and the integration of new indices if necessary. It then allows for the extraction of pixel values, which are used to calculate descriptive statistics for defined spatial entities on each date. The code of Sen2Chain is open source¹⁶, thereby enabling it to be freely reused and refined. Sen2Extract is implemented downstream to provide a query interface and perform calculations of index time series in accordance with the user's requirements. This tool is written in R¹⁷ and the online interface is based on the R Shiny¹⁸ package. Furthermore, the code is open source, thereby facilitating its reimplementaion¹⁹. It enables the submission of queries to Sen2Chain, and permits users to define extraction parameters and statistical calculations via a link sent by email. The online interface is installed on the servers of the SEAS-OI satellite ground station on Reunion island. Time series requests may be made primarily for image footprints acquired over the Indian Ocean. For other areas, it is recommended that these tools be installed on new computers or servers.

The set of tasks performed by these processing chains can be fully automated. This allows time series to be routinely generated for use by monitoring tools. The Lepto Yangon application²⁰ is an example of a tool that can map environments conducive to the transmission of leptospirosis in the Yangon region of Myanmar. It can do so as soon as a new image is available (every 5 days). This automation of the production of spectral indices over short time periods is a potential avenue for further study, particularly in relation to favourable environments for mosquito species for which links between these indices and their presence or abundance have been established.

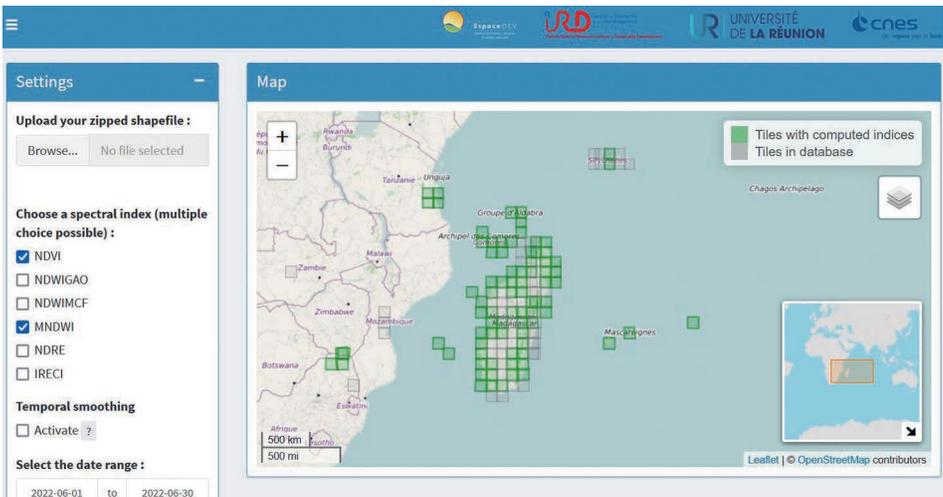


Figure 2.4. Interface of the online tool Sen2Extract (<https://web.seas-oi.org/sen2extract/>).

16. <https://framagit.org/espace-dev/sen2chain>
17. www.R-project.org/
18. <https://shiny.rstudio.com/>
19. <https://framagit.org/espace-dev/sen2extract>
20. <https://leptoyangon.geohealthresearch.org/>

►► References

- Blaschke T., Hay G.J., Kelly M., Lang S., Hofmann P., Addink E., Queiroz Feitosa R., Van der Meer F., Van der Werff H., Van Coillie F., Tiede D., 2014. Geographic Object-Based Image Analysis – Towards a new paradigm. *ISPRS Journal of Photogrammetry and Remote Sensing*, 87, 180-191.
- Cailly P., Balenghien T., Ezanno P., Fontenille D., Toty C., Tran A., 2011. Role of the repartition of wetland breeding sites on the spatial distribution of *Anopheles* and *Culex*, human disease vectors in southern France. *Parasit Vectors*, 4, 65.
- Duda T., Canty M., 2002. Unsupervised classification of satellite imagery: Choosing a good algorithm. *International Journal of Remote Sensing*, 23 (11), 2193-2212.
- Jetten T.H., Takken W., 1994. Anophelism without malaria in Europe, A review of the ecology and distribution of the genus *Anopheles* in Europe, *Wageningen Agricultural University, papers*, 94-5, 69 p.
- Paz S., Malkinson D., Green M.S., Tsioni G., Papa A., Danis K., Sirbu A., Ceianu C., Katalin K., Ferenczi E., Zeller H., Semenza J.C., 2013. Permissive summer temperatures of the 2010 European West Nile fever upsurge. *PLoS One*, 8 (2), e56398.
- Ponçon N., Toty C., L'Ambert G., Le Goff G., Brengues C., Schaffner F., Fontenille D., 2007. Biology and dynamics of potential malaria vectors in Southern France. *Malaria Journal*, 6, 18.
- Ponçon N., Tran A., Toty C., Luty A.J., Fontenille D., 2008. A quantitative risk assessment approach for mosquito-borne diseases: malaria re-emergence in southern France. *Malaria Journal*, 7, 147.
- Rodriguez-Galiano V.F., Ghimire B., Rogan J., Chica-Olmo M., Rigol-Sanchez J.P., 2012. An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, 67, 93-104.
- Tran A., Sudre B., Paz S., Rossi M., Desbrosse A., Chevalier V., Semenza J.C., 2014. Environmental predictors of West Nile fever risk in Europe. *International Journal of Health Geographics*, 13, 26.
- Tran A., Ponçon N., Toty C., Linard C., Guis H., Ferre J.B., Lo Seen D., Roger F., La Rocque S. de, Fontenille D., Baldet T., 2008. Using remote sensing to map larval and adult populations of *Anopheles hyrcanus* (Diptera: Culicidae) a potential malaria vector in Southern France. *International Journal of Health Geographics*, 7, 9.

Chapter 3

Estimation of air temperatures from satellite images and weather stations

Barbara Boufhal, Alexandre Cebeillac, Éric Daudé

Mosquitoes of the genus *Aedes* are sensitive to temperature, i.e., their level of activity (searching for blood meal, movement, etc.) and larval development rate are conditioned by temperature. The comfort zone for these mosquitoes is between 21 °C and 32 °C, with a reduced probability of survival beyond these limits (Brady *et al.*, 2013). At the global level, the prevalence of diseases transmitted by these mosquitoes is predominantly concentrated in tropical and subtropical regions (Bhatt *et al.*, 2013). However, the distribution range of these vectors has now extended to North America and, notably, southern Europe, as a consequence of climate change. (Messina *et al.*, 2019). *Aedes aegypti* mosquitoes are also synanthropic, meaning their life cycle is directly linked to the presence of human populations. *Ae. aegypti* females prefer small artificial larval habitats (flowerpots, water tanks, litter) to lay their eggs and they require a blood meal for egg maturation. The concentration of heat and humans in urban areas creates an ideal environment for vector populations to thrive.

This chapter is dedicated to the estimation of temperatures in urban environments and their potential impact on vector population dynamics. In essence, the temperature gradients present at the micro-scale can, depending on the context, either facilitate or impede the biological development of mosquitoes, and thus their proclivity to move and take blood meals, as well as influencing their reproductive cycle and survival rate. Consequently, accurately gauging temperatures at a fine-scale is an indispensable preliminary step in identifying local environments that are favourable or unfavourable to the mosquitoes. This chapter presents an array of data types for characterising ground and air temperatures, then a set of methods for estimating air temperature. It concludes with a comparative analysis of these methods as applied to Bangkok (Thailand).

►► Data to measure temperatures

An understanding of temperature fluctuations, or at the very least, their fine-scale variations, allows for a more precise characterisation of the micro-environments that can either sustain or limit the growth of mosquito populations. Urban heat islands are characteristic of this temperature differential, which can be quantified in inner-city areas. The presence of a heat island is correlated to the concentration of built-up

areas, which absorb heat during the day and release it overnight. This contributes to an increase in minimum temperatures in densely populated areas. This can further promote the stability of *Aedes* populations, particularly during the coldest periods in urban areas of tropical regions, such as Delhi, India (Misslin *et al.*, 2018), where dengue is endemic. It may, however, reduce survival rates during the hottest periods. In the conventional approach to the identification of heat islands, two temperature classes are typically employed: land surface temperatures and air temperatures.

Land surface temperature (LST) measurements utilise satellite imagery captured by different sensors. These images are available at various spatial and temporal resolutions. It is pertinent to mention the Thermal Infrared Sensor (TIRS) on board Landsat 8, with which it is theoretically possible to estimate surface temperatures on a fortnightly basis with a spatial resolution of 100 m. For shorter time periods, and therefore those more suitable for the mosquito development cycle, the MODIS sensor on board the Aqua and Terra satellites provides land surface temperature data at a resolution of 1 km, four times per day when meteorological conditions are optimal. One of the biggest constraints for satellite images, particularly in tropical and subtropical regions, is the frequent occurrence of a cloud ceiling, which significantly impairs their efficacy.

In addition to exhibiting spatial variability, land surface temperatures also display variation when compared to air temperatures (Figure 3.1). These differences are explained by a combination of many factors acting at the local level, including soil moisture, surface roughness, and wind speed (Stisen *et al.*, 2007). These temperature differences have a direct impact on vector development (Misslin *et al.*, 2016).

The measurement of air temperature is conducted at weather stations. The location and positioning of these stations is subject to strict guidelines, with their location confined to shaded areas. The management of these facilities is the responsibility of local authorities. Data is collected at an altitude of 2 m from the ground, and the contrasting values recorded between different stations in the same town can be attributed to local variations in surface energy balance. One of the challenges when using this data is generally the density of territorial coverage, which often results in an insufficient number of stations to enable the reliable interpolation of data for large cities, given their diverse architectural and structural characteristics.

This chapter proposes combining these two data sources: satellite images (LST), which are incomplete over time, and weather station data, which is incomplete over space, in order to compare different methods of estimating air temperature in the Bangkok metropolitan area (Figure 3.1). For this study, we used the MODIS sensor, given its higher revisit rate and straightforward access to the images (USGS site²¹).

►► Air temperature estimation: different methods

Several methods exist to estimate air temperature based on land surface temperatures. One simple method is based on the linear regression between air temperatures recorded by ground stations and land surface temperatures obtained from satellite imagery. In the event that the number of data points is considerable (obtained over an extended period from a sufficient number of stations) and the correlation

21. <https://lpdaac.usgs.gov/data/get-started-data/collection-overview/missions/modis-overview/>

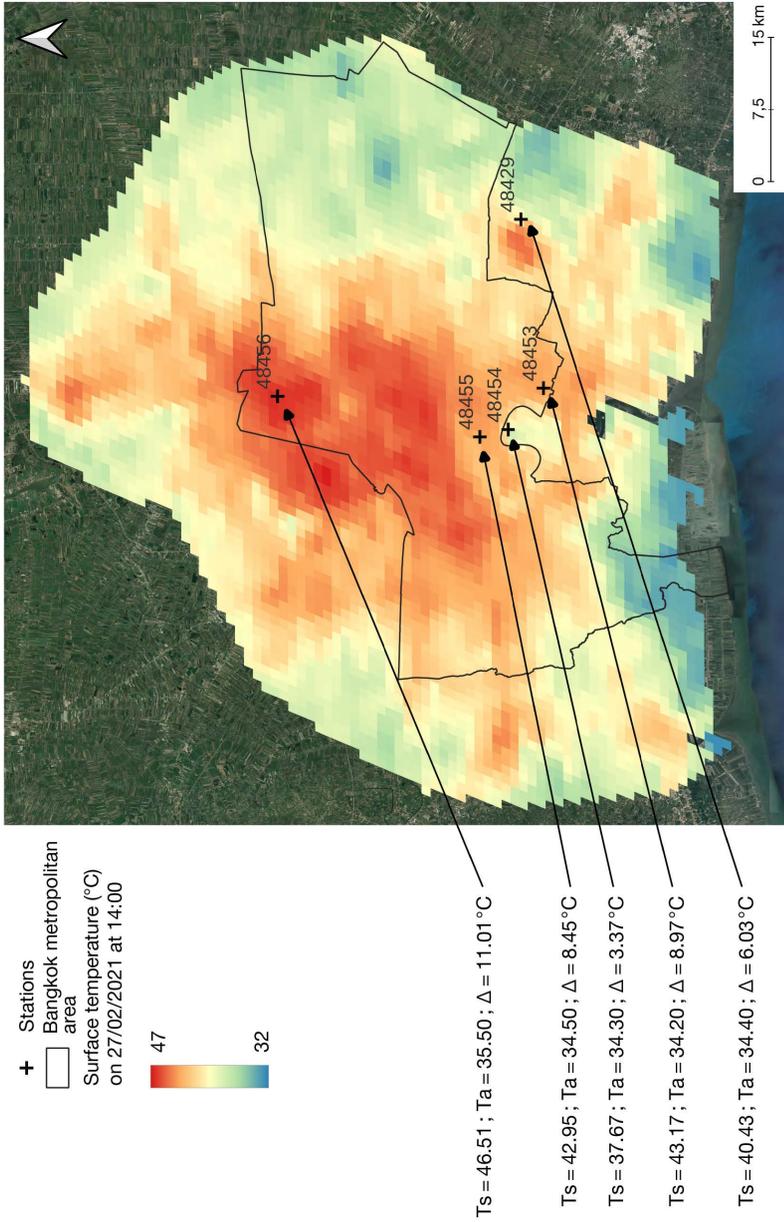


Figure 3.1. Temperatures across the extensive Bangkok metropolitan area.

Stations are identified by numbers displayed to the upper right of the crosses (e.g., “48429” refers to the airport station). The image of land surface temperatures (LSTs) was captured by MODIS on 27 February 2021 at 14:00. In the lower left, the surface temperature (Ts) is given in degrees Celsius, estimated using MODIS satellite imagery, alongside the air temperature (Ta) recorded by the station and the difference (Δ) between the two.

coefficient is strong, regression enables the estimation of air temperatures across the entire region based on land surface temperature data (Zaksek and Schroedter-Homscheidt, 2009). An alternative approach employs the energy balance of the observed surfaces and incorporates concepts from thermodynamics (Sun *et al.*, 2005). Finally, regression analyses (linear or otherwise) between land surface temperature and other supplementary variables, such as land cover or altitude, can be employed (Benali *et al.*, 2012; Oyler *et al.*, 2014). The TVX (*Temperature-Vegetation Index*) method is based on this principle.

The TVX method is based on the assumption that the land surface temperature is in thermal equilibrium with the air temperature at the canopy level (Vancutsem *et al.*, 2010). Furthermore, it assumes that there is a negative correlation between temperature and vegetative activity, as measured by the Normalized Differenced Vegetation Index (NDVI), and that atmospheric conditions are locally uniform (Stisen *et al.*, 2007). Based on these principles, the TVX method consists of calculating linear regression between the LST values (T_s) and those of the NDVI in a moving window of a given size:

$$T_s \sim \alpha NDVI + \beta \text{ (Equation 1)}$$

By assuming the thermal equilibrium at the canopy, it is possible to retain the parameters of Equation 1 in order to obtain the following:

$$T_a \sim \alpha NDVImax + \beta \text{ (Equation 2)}$$

where $NDVImax$ can either be chosen from ground station values of NDVI such that that it minimises the mean absolute error in the relationship defined by Equation 2 (Method 1 or “traditional TVX”), or the maximum NDVI in the local window can be used (Method 2 or “local TVX”). The TVX method has the advantage of requiring minimal supplementary data input, in contrast to the statistical approach. Furthermore, it is not contingent on dense ground station coverage for the generation of accurate results (Misslin *et al.*, 2018).

Subsequently, the TVX method is applied to the Bangkok metropolitan area (Method 1, “traditional TVX”) by proposing different variants, such as the use of local $NDVImax$ (Method 2, “local TVX”) and temperature corrections related to the amount of built-up area (Method 3, “local and built-up TVX”). Indeed, if buildings influence temperature by generating a heat island effect, then incorporating them into the calculations may potentially enhance the precision of air temperature estimates.

► Applications to Bangkok

For this case study, we relied on LST and NDVI data collected by the MODIS sensor on board the Terra and Aqua satellites from 30 June 2015 to 30 April 2022. A total of 1,885 images were identified as potentially suitable for use, with a minimum of 50% of the area covered. However, due to significant cloud cover from May to November, these months only represent 17% of the sample (Table 3.1).

We also use a geographical layer containing more than 1.4 million buildings in Bangkok (see Chapter 4 for details about the method). Linear regression is applied to each of the images in a moving window (between land surface temperature and NDVI for Methods 1 and 2, with the inclusion of the building layer for Method 3). Subsequently, local regression coefficients and intercepts are obtained, as illustrated in Figure 3.2.

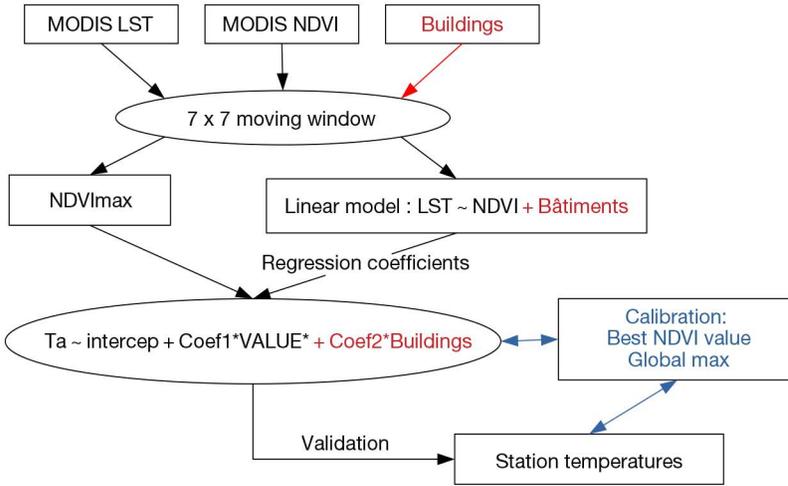


Figure 3.2. Illustration of the methods used.

Method 2 (“local TVX”) follows the path drawn by the black arrows, with Method 3 (“local & built-up TVX”) including the elements in red, and Method 1 (“traditional TVX” approach) including those in blue. *Value* minimises the mean error for all stations in Method 1, whereas it corresponds to the local NDVImax for Methods 2 and 3.

Seven weather stations collect meteorological data across the study area. However, information on built-up surfaces is only available for five of these (those closest to the administrative boundaries of Bangkok, Figure 3.1). These stations provide air temperatures data every three hours, but only the temperatures taken during the closest approaches of the Aqua and Terra satellites are retained for the purposes of calibration and validation. We subsequently put forward three approaches, derived from the TVX method, for comparison.

The first method (Method 1) uses the approach proposed by Misslin *et al.* (2018). The objective is to identify the NDVImax value at which the discrepancy between the air temperatures recorded at the stations and the estimated air temperatures using Equation 2 is at a minimum. The method therefore consists of using the set of NDVI values recorded at each station and selecting the one which results in the lowest mean absolute error (MAE):

$$MAE = \frac{\sum |Ta' - Ta|}{n}$$

with Ta' the estimated air temperature, Ta the recorded air temperature and n the number of observations.

The NDVImax value, representing the entire area, is subsequently employed in Equation 2 for the estimation of air temperatures. Seasonal updates to NDVImax enables the estimation of air temperature fluctuations across the study area (Table 3.1).

It is important to note that, in addition to the NDVImax values varying between months, they also demonstrate variation over the course of the day (Table 3.1). The extremely low values of NDVImax observed during the night highlight the importance

of the intercept (the coefficient β) in estimating nighttime temperatures. Indeed, as NDVImax approaches zero, T_a' becomes very close to β (Equation 2). This observation must be taken into account when comparing the quality of these methods, particularly when estimating nighttime temperatures.

Table 3.1. NDVImax values per period (per month).

	January	February	March	April	May to November	December
NDVImax (daytime)	0.56	0.49	0.47	0.45	0.62	0.58
NDVImax (nighttime)	0.08	0.02	0.08	0.06	0	0.05
Number of images	333	388	334	200	328	302

The other two methods tested (Methods 2 and 3) calculate a local NDVImax using a moving window. In comparison with the global NDVImax calculated using Method 1, it is hypothesised that there is a significant variation in energy balance across urban areas due to the diversity of land use. This justifies the need to account for local variations. Method 2 therefore emphasises the link between land surface temperatures and local NDVImax, and Method 3 (TVX & built-up) incorporates the extent of built-up area into the calculation of regression coefficients by modifying Equation 1 as follows:

$$T_s \sim \alpha_1 NDVI + \alpha_2 \text{built-up} + \beta \text{ (modified Equation 1)}$$

Figure 3.3 shows the distribution of correlation coefficients calculated for each MODIS image in a moving window Figure 3.3a) and the local averages obtained by Method 2 (Figure 3.3b) and Method 3 (Figure 3.3c).

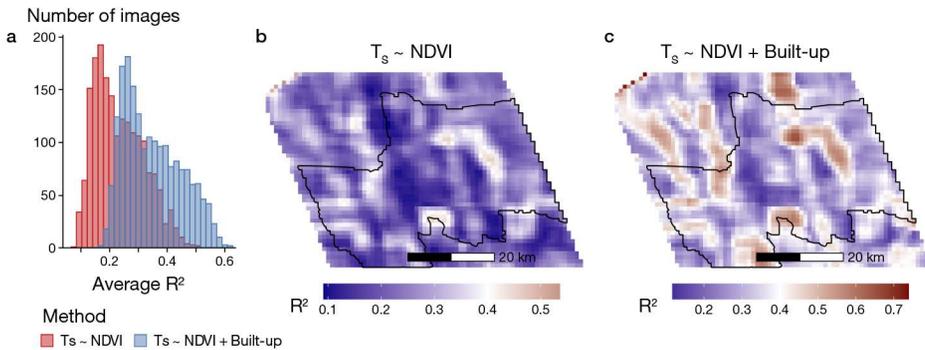


Figure 3.3. A comparison of correlation coefficients between surface temperatures and NDVI for Method 2, and surface temperatures and NDVI with the incorporation of building density for Method 3.

(a) Histogram of global correlation coefficients R^2 calculated for each MODIS image. (b) Average local R^2 (7×7 window) across all images for $T_s \sim NDVI$ relationship. (c) $T_s \sim NDVI + \text{built-up}$.

We note that the addition of the built-up layer in the application of local linear regressions increases global R^2 values (average of 0.18 for Method 2 and 0.25 for Method 3, Figure 3.3a). A similar observation can be made at the local level (Figures 3.3b and c).

However, if a significant number of areas exhibit high correlation coefficients, it becomes challenging to draw conclusions about the influence of urban structures on the relationship $T_s \sim NDVI + \text{built-up}$. While a high level of correlation is not a prerequisite for proceeding to the subsequent stages of the TVX method (Misslin *et al.*, 2018), it is evident that incorporating a built-up layer in conjunction with NDVI markedly strengthens the relationship and may potentially enhance air temperature estimates.

Figure 3.4 shows the results obtained for each method alongside the LST of MODIS images for comparison. Figure 3.4a demonstrates a tendency for daytime air temperatures to be overestimated, while nighttime temperatures are, on the contrary, underestimated. However, during the daytime, the modes of estimated temperatures are closer to the recorded temperatures than the mode of LSTs. It is important to highlight that, despite the increased amplitude of deviations, Method 1 yields results

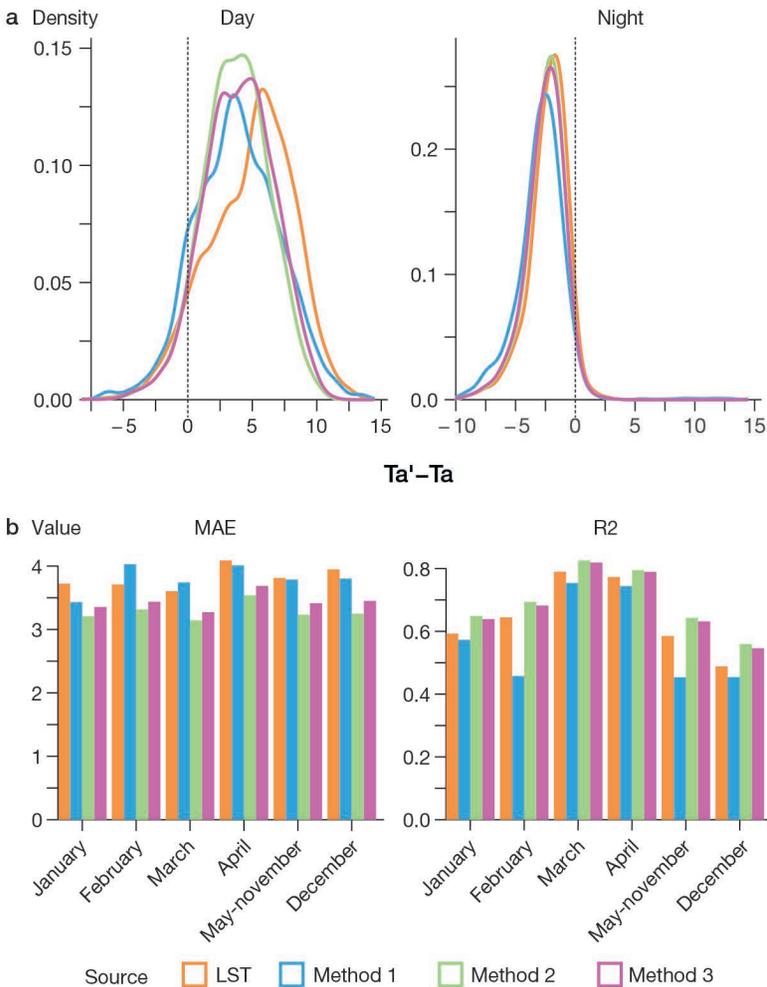


Figure 3.4. Discrepancies between estimated and observed air temperatures

For each method: (a) day (left) and night (right); (b) mean absolute error (left) and correlation coefficient (right) across 6 periods of the year.

that are more closely aligned with the temperatures recorded by ground stations. Conversely, at night, the methods employed to estimate air temperature do not mitigate the discrepancies between estimated and observed temperatures, particularly in the case of Method 1. Figure 3.4b illustrates the discrepancy in estimates across different periods of the year, quantified in terms of MAE and R^2 . With regard to the latter indicator, the air temperatures recorded for the months of March and April are the most reliable, particularly when employing Method 2, which demonstrates marginally superior results in comparison to Method 3 (TVX + built-up). Method 1 is an effective means of reducing the mean absolute error in comparison to the LSTs. However, it remains less effective than Methods 2 and 3, with deviations oscillating between 3.2 and 3.7°C depending on the month in question.

The validation method, which involves a comparison of the results with data from ground-based weather stations, suffer from several drawbacks. Firstly, how can meso-scale data be effectively compared to micro-scale data? The values recorded by these stations depend on their immediate environment, whereas the estimated air temperatures are averages across significantly larger areas. Furthermore, the number of stations is relatively low (5 in Bangkok at the time of the study), and their coverage is not sufficiently homogeneous to adequately represent the diverse urban structures present in Bangkok, which can influence temperatures (densely populated areas, sparsely populated, hybrid, etc.).

A comparison of the mean absolute differences between the various methods of estimated air temperature and the LSTs (Figure 3.5) reveals the existence of areas where air temperatures exceed land surface temperatures as measured by MODIS, particularly in the city centre. The inverse can also be noted, particularly to the east of the city in the vicinity of the airport (station 48429) [Figure 3.1]. The observed deviations can

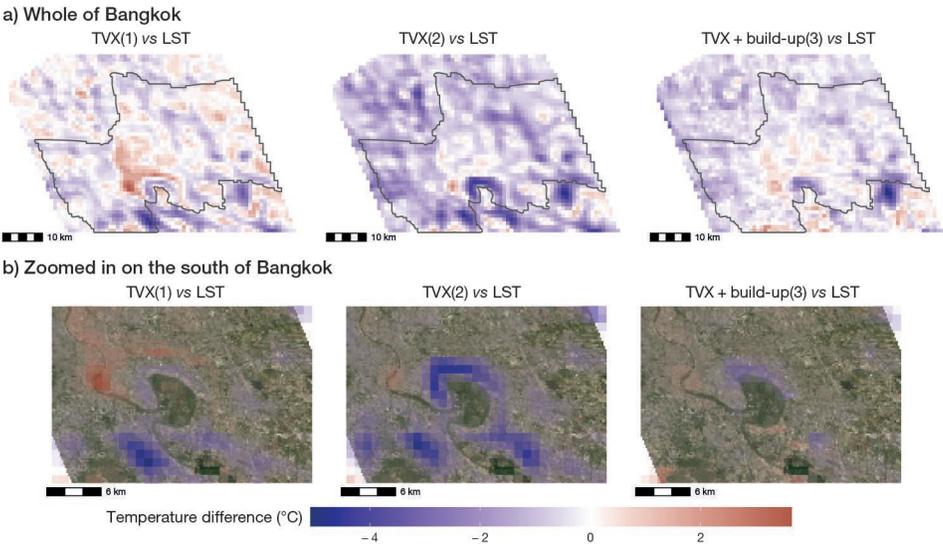


Figure 3.5. Differences between the sum of estimated daytime temperatures (from left to right: Methods 1, 2 and 3) and the LSTs measured by MODIS, weighted by the number of pixels containing data over the period in question: (a) across all of Bangkok, (b) across a smaller and more heterogeneous area to the south of the city.

be explained by a high degree of heterogeneity in land cover at a local scale, particularly in the area surrounding the meander to the south of Bangkok (Figure 3.5b), where vegetation-dominated areas meet densely populated areas. This local heterogeneity exerts an influence on the local regression coefficients and local NDVImax.

►► Conclusion

In this chapter, a comparative analysis was presented of the various methods employed for the estimation of air temperatures based on land surface temperatures derived from MODIS images. While the methodologies presented do mitigate discrepancies between observed and estimated air temperatures, these differences remain considerable, varying between 3 to 5 °C depending on the time of year and the method used. Nevertheless, it is challenging to compare temperatures estimated at the meso-scale and data measured at the micro-scale. It is indeed feasible that the mean effective air temperature in a 1 km² grid may diverge from the temperature recorded at the local level. This difference can be attributed to the placement of the station and its surrounding environment on the one hand, and the impact of smoothing across the grid on the other. Furthermore, the quality of the estimate is contingent upon the density of weather stations utilised and their local environment, which may not be uniform across the grid in question.

it is also important to acknowledge that the methodologies employed have the potential to introduce edge effects associated with the utilisation of a moving window in heterogeneous areas. The calculation of local regression coefficients in homogeneous areas from a land cover perspective would facilitate a more accurate comparison of the different methods.

►► References

- Benali A., Carvalho A.C., Nunes J.P., Carvalhais N., Santos A., 2012. Estimating air surface temperature in Portugal using MODIS LST data. *Remote Sensing of Environment*, 124, 108-121.
- Bhatt S., Gething P.W., Brady O.J., Messina J.P., Farlow A.W., Moyes C.L., Drake J.M., Brownstein J.S., Hoen A.G., Sankoh O., Myers M.F., George D.B., Jaenisch T., Wint G.R., Simmons C.P., Scott T.W., Farrar J.J., Hay S.I., 2013. The global distribution and burden of dengue. *Nature*, 496 (7 446), 504-507.
- Brady O.J., Johansson M.A., Guerra C.A., Bhatt S., Golding N., Pigott D.M., Delatte H., Grech M.G., Leishman P.T., Maciel-de-Freitas R., Styer L.M., Smith D.L., Scott T.W., Gething P.W., Hay S.I., 2013. Modelling adult *Aedes aegypti* and *Aedes albopictus* survival at different temperatures in laboratory and field settings. *Parasit Vectors*, 6, 351.
- Messina J.P., Brady O.J., Golding N., Kraemer M.U.G., Wint G.R.W., Ray S.E., Pigott D.M., Shearer F.M., Johnson K., Earl L., Marczak L.B., Shirude S., Davis Weaver N., Gilbert M., Velayudhan R., Jones P., Jaenisch T., Scott T.W., Reiner R.C. Jr., Hay S.I., 2019. The current and future global distribution and population at risk of dengue. *Nature Microbiology*, 4 (9), 1508-1515.
- Misslin R., Vaguet Y., Vaguet A., Daudé É., 2018. Estimating air temperature using MODIS surface temperature images for assessing *Aedes aegypti* thermal niche in Bangkok, Thailand. *Environmental Monitoring and Assessment*, 190 (9), 537.
- Misslin R., Telle O., Daudé É., Vaguet A., Paul R.E., 2016. Urban climate versus global climate change-what makes the difference for dengue? *Annals of the New York Academy of Sciences*, 1382 (1), 56-72.
- Oyler J., Ballantyne A., Jencso K., Sweet M., Running S., 2014. Creating a topoclimatic daily air temperature dataset for the conterminous United States using homogenized station data and remotely sensed land skin temperature. *International Journal of Climatology*, 35 (9), 2258-2279.
- Stisen S., Sandholt I., Nørgaard A., Fensholt R., Eklundh L., 2007. Estimation of diurnal air temperature using MSG SEVIRI data in West Africa. *Remote Sensing of Environment*, 110 (2), 262-274.

Sun Y.J., Wang J.F., Zhang R.H., Gillies R.R., Xue Y., Bo Y.C., 2005. Air temperature retrieval from remote sensing data based on thermodynamics. *Theoretical and Applied Climatology*, 80 (1), 37-48.

Vancutsem C., Ceccato P., Dinku T., Connor S.J., 2010. Evaluation of MODIS land surface temperature data to estimate air temperature in different ecosystems over Africa. *Remote Sensing of Environment*, 114 (2), 449-465.

Zaksek K., Schroedter-Homscheidt M., 2009. Parameterization of Air Temperature in High Temporal and Spatial Resolution from a Combination of the SEVIRI and MODIS Instruments. *ISPRS Journal of Photogrammetry and Remote Sensing*, 64 (4), 414-421.

Chapter 4

From census to buildings: generating synthetic populations

Alexandre Cebeillac, Olivier Gillet, Éric Daudé

A major challenge for the study of vector-borne diseases is the need for detailed knowledge of human population distribution at a high spatial resolution. In the case of dengue, clusters have been observed to vary in size between a radius of 200 and 800 m (Salje *et al.*, 2012), which may result in epidemic explosions in densely populated areas. This has resulted in the development of a proven strategy for vector control, comprising targeted interventions in the vicinity of the index case, with the objective of reducing the prevalence and size of clusters. In the event of multiple outbreaks, this strategy must be refined to prioritise interventions in areas most at risk, which may include the most populated areas where vector presence has been verified. The integration of epidemiological data (e.g., via sentinel hospitals) and entomological data (such as mosquito risk maps) with population distribution can enhance the accuracy of predictions regarding the risk of incidence and spread of outbreaks.

Unfortunately, a significant proportion of metropolitan areas in tropical regions, which are severely affected by dengue, lack the necessary tools and population maps at very fine spatial resolutions to enable effective intervention. Indeed, institutional data is typically aggregated in accordance with administrative boundaries which do not apply to epidemiological mechanisms. Furthermore, the data is frequently outdated due to the rapid growth of some highly populated urban areas, and in some cases, it is not available at all. While these two aspects, related to the capacity of states and/or municipalities to conduct censuses, is not within the scope of this chapter, it is nevertheless feasible to employ disaggregation algorithms on the available coarse data to disaggregate populations into more granular spatial resolutions (Viel et Tran, 2009). Projects such as Landscan (Dobson *et al.*, 2000), Worldpop (Linard et Tatem, 2012) and the Global Human Settlement Layer (GHSL; Florczyk *et al.*, 2019) are heading in this direction by proposing global population distribution maps with maximum spatial resolutions in the order of one kilometre and down to 250 m.

Below, we present a case study on Bangkok (Thailand) in order to facilitate a more detailed examination of the spatial disaggregation of population data. For this metropolitan area, we obtained census data on populations at the district level, in addition to a layer representing buildings in two dimensions. We then describe a second use case for the disaggregation of population data at another site (Rouen, France).

We demonstrate the application of algorithms that estimate population distribution within buildings, as well as a realistic representation of the composition of each household. This final point is particularly noteworthy in the context of dengue, for which there is a strong link between the seroprevalence and the age distribution of the population. The results of these algorithms can also represent interesting data for simulation models (Chapter 9).

►► Population disaggregation and redistribution

Traditional dasymetric mapping

Dasymetric mapping, which literally means “measuring density”, was first used in 1911 by Semenov-Tian-Shansky (Petrov, 2012) and subsequently popularised by Wright in 1936 (Li and Weng, 2005). It involves the distribution of population counts from a source zone to a set of target zones based on thematic characteristics. Historically, this entailed the redistribution of populations from administrative areas to inhabited areas within the same geographical entity. This principle has been gradually refined by the use of multiple linear regression between the gridded population data and land use/land cover within the same grid (Equation 1).

$$P_s = \beta_0 + \beta_1 U_1 + \beta_2 U_2 + \dots + \beta_n U_n \text{ (Equation 1)}$$

This method proposes a description of the population distribution P_s in a source zone using multiple regression, where each land use category U_n is associated with a coefficient β_n . These coefficients, which represent population densities associated with different categories of land use, are subsequently employed for the purpose of weighting, thereby facilitating the disaggregation of the source population into each target zone in accordance with its surface type. In an ideal scenario, the correlation coefficient would approach a value of 1.

This method was subsequently employed to estimate fine-scale population data for Bangkok (Misslin and Daudé, 2017) based on OLI and TIRS (Landsat 8) satellite images. The proportion of loosely, densely and very densely built-up areas within subdistricts of Thailand’s capital city was estimated using unsupervised classification. Dasymetric mapping in this context is based on the assumption that a pixel of densely built-up area will accumulate more population than a sparsely built-up area. Each pixel is therefore assigned a coefficient (β), obtained by using multiple regression across the entire zone corresponding to its category (U), which is multiplied by the population of the geographic zone in which it is located (Figure 4.1).

Since Landsat 8 satellite images are available for the whole Earth, and this is a well-proven methodology (Hallot *et al.*, 2019), this approach has the advantage of being reproducible in disparate locations, provided census data is available for the area of interest. Other thematic data can also be used to characterise the link with population distribution, such as soil water repellency and its relative surface area (Hallot *et al.*, 2019), or even data on artificial light at night (Briggs *et al.*, 2007).

Nevertheless, any unsupervised classification is inherently imperfect. The advent of new services, such as Google Maps and ESRI²², which provide data on the footprint of buildings and their use, offers promising avenues for further research. This consists

22. <https://www.esrifrance.fr/>

of counting the inhabited buildings and taking into account their footprint within each census tract, thus enabling the application of the same linear regression techniques. For example, OpenStreetMap (OSM) data is freely available and provides information on land cover and land use categories for a specified geographic area, as well as, on occasion, the footprint of buildings. However, the accuracy and reliability of this crowd-sourced data greatly depends on the location. While highly accurate in France, this information is not as precise in other parts of the world, especially in Bangkok. However, by downscaling geospatial data, it is possible to overcome this obstacle.

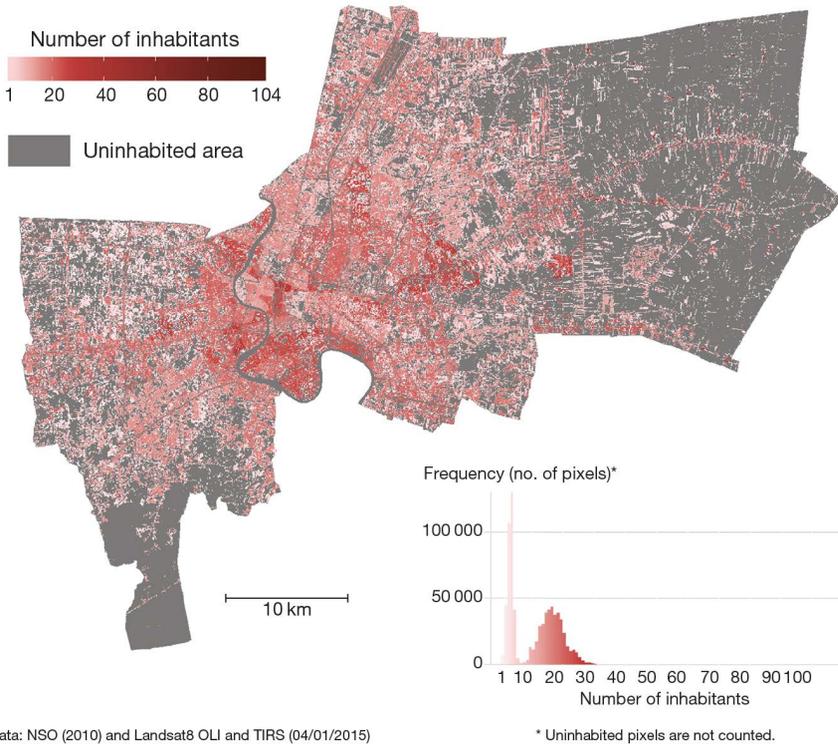


Figure 4.1. Dasymeric map of population distribution in Bangkok on a 30×30 m grid. Source: Misslin and Daudé (2017).

Refining thematic data to improve dasymeric mapping

Web scraping methods allow online databases (private or public) to be accessed, such as those of the ESRI's Web Map Service (WMS). These databases facilitate the generation of thematic maps that delineate the spatial distribution of buildings (Figure 4.2a). A visual comparison between satellite images and data representing buildings from the ESRI's WMS in Bangkok demonstrates that the buildings are accurately represented. Code was developed to extract all of this information²³, firstly in raster form (Figure 4.2b), then using vectors²⁴ (Figure 4.2c and d), generating a database of 1.4 million polygons.

23. It should be noted that web scraping is now legal according to a ruling by a U.S. appeals court (<https://techcrunch.com/2022/04/18/web-scraping-legal-court/>).

24. <https://drive.univ-rouen.fr/f/a20e1dc3baad49caadec/?dl=1>



Figure 4.2. Process of extracting buildings from an ESRI WMS image.

(a) Raw data, (b) colour intensity change, (c) extraction of built-up polygons, and (d) their projection on to a satellite image for visual comparison.

This 2D representation of buildings in Bangkok²⁵ allows for the calculation of the number of buildings, the cumulative perimeter of these buildings, and the built-up surface area for each zone, which in this case are the subdistricts. Additionally, Bangkok is a city where winding neighbourhoods comprising one- and two-storey buildings are interspersed with large blocks of flats. The inclusion of height estimates can thus provide additional useful information for the disaggregation of population data. For this, we collected data from the SRTM (Shuttle Radar Topography Mission) to obtain digital elevation models (DEM) with a surface resolution of around 30 m at the equator²⁶. A digital terrain model (DTM) was then constructed by interpolating the average altitude of canals and rivers in the city to determine a baseline, which was then subtracted from the DEM to obtain local elevation²⁷. Finally, the differentiation between inhabited and uninhabited areas was made possible by the incorporation of a land use layer from OSM. This allowed buildings located in non-residential areas to be removed (temples, industrial zones, shopping centres, recreational facilities, etc.), as well as removing vegetated areas and water bodies.

Using this data, several linear regressions were calculated between the population of Bangkok per subdistrict and land use data: the number of buildings, built-up surface area and boundaries, and the average altitude in each of these subdistricts (Figure 4.3a). We then selected the combination of factors from Equation 1 which gave the lowest possible mean absolute error and the best correlation coefficient. The latter is obtained by using all indicators of a built-up area (Figure 4.3b). However, the combination of at least two parameters (regardless of which) results in R^2 correlation coefficients greater than 0.81. Indeed, with regard to Figure 4.3a, the majority of the parameters that define built-up areas are highly correlated with population ($R^2 > 0.75$).

We then proceeded with three steps. Initially, a grid of the desired size was created (e.g., 50 m) and from this, information on buildings was inferred (number, length, height, and built-up surface area in each grid square). Subsequently, the linear regression coefficients linking population to these indicators at the subdistrict level were applied to each grid square. Finally, the population within each grid square was adjusted according to subdistrict, in alignment with the census population data for that area. A breakdown of the population is thus obtained for a grid square of a given resolution (Figure 4.4b).

25. <https://drive.univ-rouen.fr/f/d71c73ffd3504ae28068/?dl=1>

26. <https://earthexplorer.usgs.gov/>

27. <https://drive.univ-rouen.fr/f/f282bacb588e43818c75/?dl=1>

Nevertheless, as shown in Figure 4.4c, depending on the multiple regression models used at the beginning, significant local variations may be present. These are primarily located in areas with tall buildings, which implies that the regressions taking into account elevation will tend to concentrate the populations within these grid squares, unlike other approaches which will distribute the population more homogeneously.

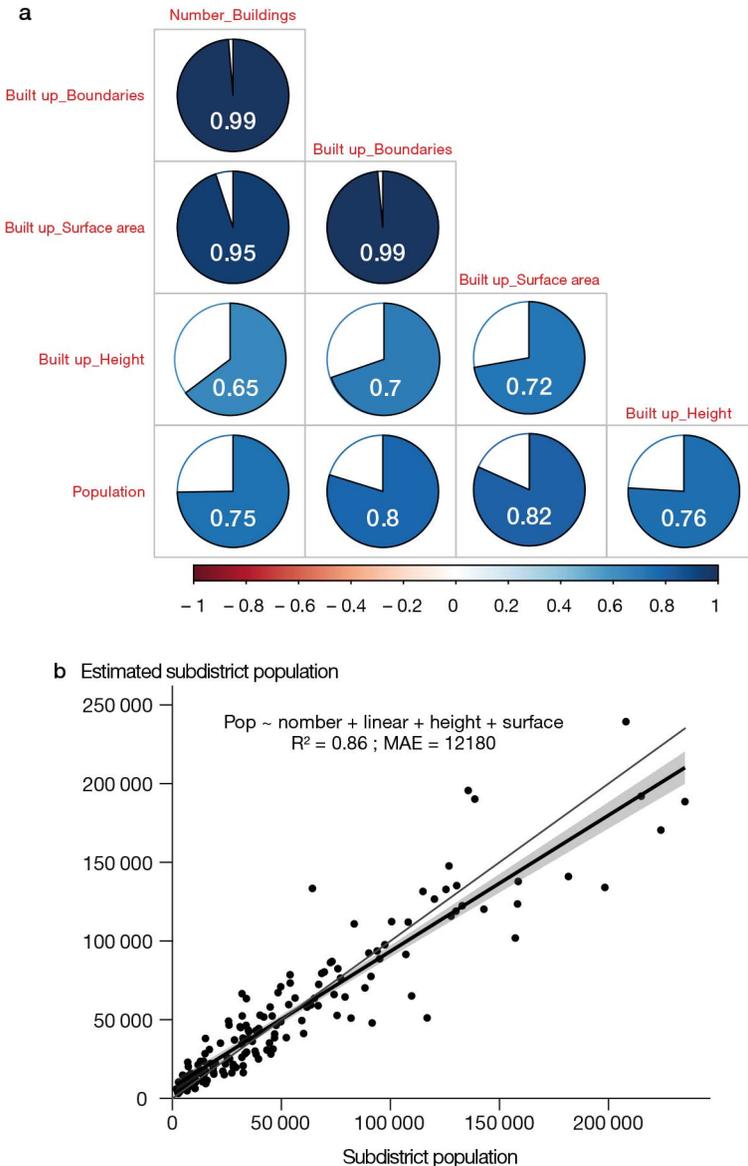


Figure 4.3. Relationships between building characteristics in a given geographical area and the associated population.

(a) Correlations (R^2) between pairs of parameters, and (b) regression between known populations at the subdistrict level and those estimated using a set of land cover indicators (black line, blue line corresponds to the relationship $y = x$).

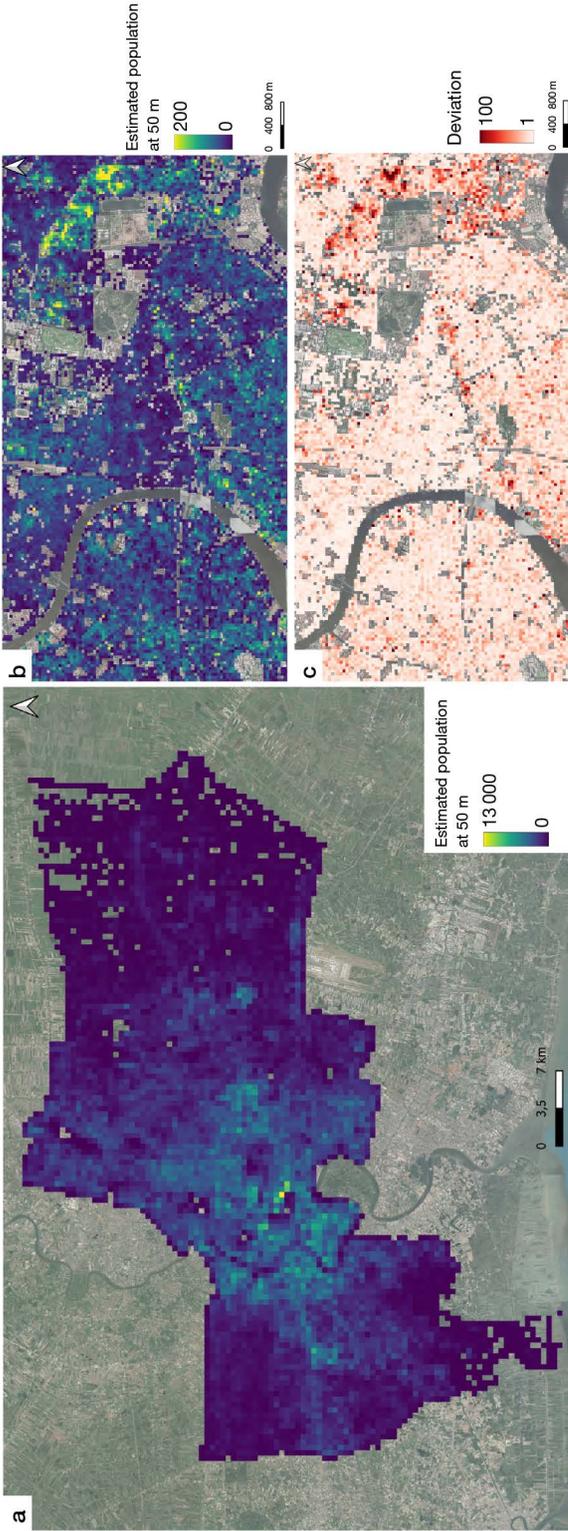


Figure 4.4. Estimating the population of Bangkok using dasymetric mapping (a) 500 m resolution, and (b) 50 m with all of the building parameters taken into account; (c) standard deviation over 10 estimates.

High resolution spatial mapping of population density can therefore aid in decision-making (increased vector control in the most densely populated areas) or be used as inputs for different models (see Chapter 9). In this regard, recent advances have led to new innovations in qualifying high spatial resolution population data. This process entails not only counting the individuals within a given area, but also the characterisation of each individual present. We will illustrate this approach in the city of Rouen, Normandy (France).

► Synthetic populations, a methodology which supports the fine-scale analysis of health-related issues

Computing a synthetic population consists of creating an exhaustive and statistically representative set of individuals and households for a given territory (Lovlace and Dumont, 2016). Several methods have been developed in recent years (synthetic reconstruction, combinatorial optimisation, statistical learning) to generate sets of agents with realistic demographic attributes. The availability and precision of the data will inform the choice of methodology to be employed: with or without sampling, deterministic or stochastic.

Principle of generating a synthetic population

The most commonly used algorithm for generating a synthetic population is IPF, for Iterative Proportional Fitting (Hörl and Balac, 2021). This algorithm uses two types of data: a representative sample of individuals described using multiple terms and the total number of individuals in the population per term. The algorithm will then generate a population of individuals equivalent to the overall population, while respecting the relative distribution of the sample per cross-referenced term. However, IPF is not equipped to handle nested data structures, for example, a number of households composed of n individuals. This algorithm is aimed at finding optimal distribution either at the household or at the individual level (Khachman *et al.*, 2021). The loss of information may prove problematic, particularly when considering individual mobility, which is often contingent upon the individual's characteristics and those of the household. Other methods have been developed to compensate for this loss, such as Iterative Proportional Update (Ye *et al.*, 2009) and the Hierarchical Iterative Proportional Fitting (Yameogo *et al.*, 2021). These algorithms therefore consider not only information at the individual level, but also at the household level in their processes of reconstructing synthetic populations.

From populations to households and from households to individuals

In France, thanks to population censuses, highly accurate information is available on population characteristics: age, occupations, modes of transport used, housing conditions, etc. A synthetic population of Rouen was generated from 2017 data using an IPF algorithm applied to all of the census tracts, known as IRIS (*îlots regroupés pour l'information statistique*). A total of 62,121 households and 110,142 individuals were generated using all 42 IRIS census tracts from the study area. The second stage in this generation process consisted of breaking down the individuals and households located in the buildings of each IRIS. First, it was necessary to estimate the number

of inhabitants using the characteristics of residential buildings taken from the BD TOPO²⁸ database of France’s National Geographic Institute (IGN). The number of inhabitants is thus contingent upon the number of floors, as well as the surface area of each floor in relation to the total surface area of residential buildings. Households were then broken down into buildings according to the number of inhabitants and the theoretical occupancy of the accommodation.

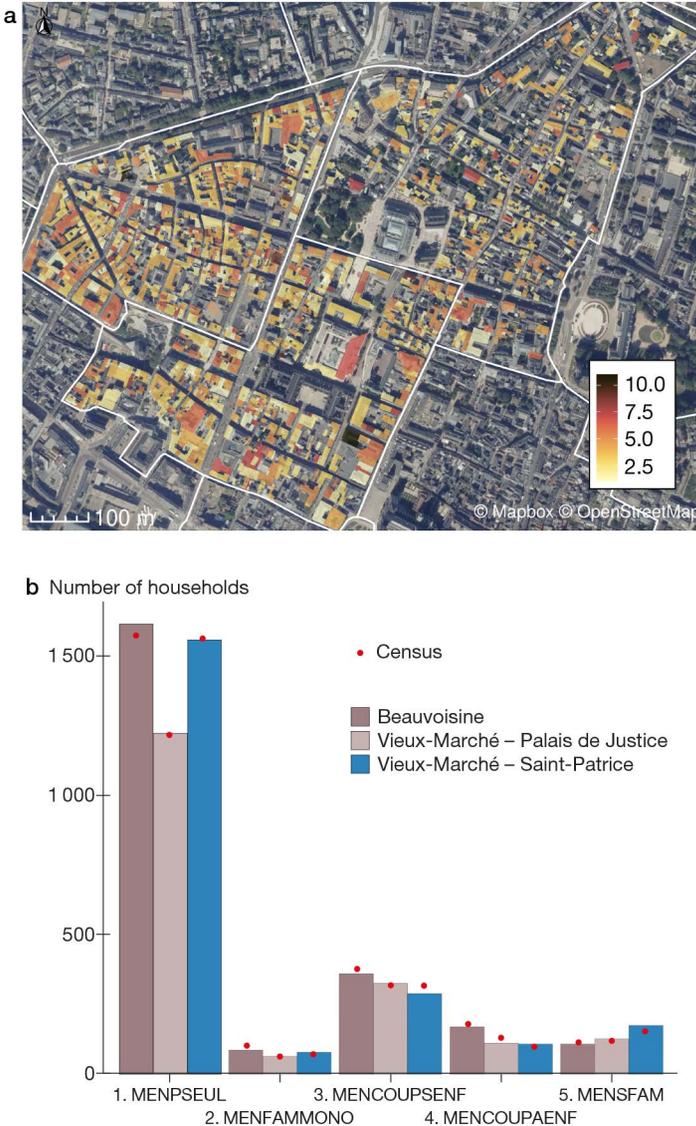


Figure 4.5. Distribution of households across three IRIS census tracts in Rouen. (a) Spatial distribution. (b) Structure of social groups: (1) one-person household, (2) single-parent household, (3) childless couple, (4) couple with child, and (5) complex households.

28. <https://geoservices.ign.fr/documentation/donnees/vecteur/bdtopo>

Figure 4.5 illustrates the outcomes of this synthetic population generation process across three IRIS census tracts from Rouen (Vieux-Marché – Palais de Justice, Vieux-Marché – Saint-Patrice, Beauvoisine). The map represents the number of households in each building and the histogram shows the distribution of these types of households in the three IRIS. The same method can then be used to further break down the population into age categories.

►► Conclusion

Knowledge on population distribution can limit the risks of human exposure to vector-borne diseases. The epidemic risk is a consequence of the simultaneous spatial and temporal coexistence of hosts and vectors. This chapter presents a number of techniques for disaggregating census data from disparate data sources. Census data is frequently presented as gridded data constructed from administrative units, which is of limited relevance to the epidemiological issues under consideration. While the spatialization of individuals is now feasible at high spatial resolutions, one of the key challenges to be addressed is how these locations evolve over time. The socio-demographic data presented here are based on censuses taken from the place of residence. As a result, it is not possible to evaluate individual exposure risk when travelling to different locations.

►► References

- Briggs D.J., Gulliver J.S., Fecht D., Vienneau D., 2007. Dasymetric modelling of small-area population distribution using land cover and light emissions data. *Remote Sensing of Environment*, 108, 451-466.
- Dobson J., Bright E., Coleman P., Durfee R., Worley B., 2000. LandScan: A Global Population Database for Estimating Populations at Risk. *Photogrammetric Engineering & Remote Sensing*, 66 (7), 849-857.
- Florczyk A., Corban C., Ehrlich D., Freire S., Kemper T., Maffeni L., Melchiorri M., Pesaresi M., Politis P., Schiavina M., Sabo F., Zanchetta L., 2019. *GHSL Data Package*, JRC Technical Report, Publications Office of the European Union, Luxembourg.
- Hallot E., Tais G., Stephenne N., Wolff E., 2019. Cartographie détaillée de la densité de population : comparaison de méthodes dasymétriques. *Dynamiques régionales*, 8 (2), 35–56.
- Hörl S., Balac M., 2021. Synthetic population and travel demand for Paris and Île-de-France based on open and publicly available data. *Transportation Research Part C: Emerging Technologies*, 130, 103291.
- Khachman M., Morency C., Ciari F., 2021. Impact of the Geographic Resolution on Population Synthesis Quality. *ISPRS International Journal of Geo-Information*, 10 (11), 790.
- Li G., Weng Q., 2005. Using Landsat ETM+ Imagery to Measure Population Density in Indianapolis, Indiana, États-Unis. *Photogrammetric Engineering & Remote Sensing*, 71 (8), 947–958.
- Linard C., Tatem A.J., 2012. Large-scale spatial population databases in infectious disease research. *International Journal of Health Geographics*, 11, 7.
- Lovelace R., Dumont M., 2016. *Spatial Microsimulation with R*, CRC Press, 260 p. (Chapman & Hall/CRC The R Series).
- Misslin R., Daudé É., 2017. An environmental suitability index based on the ecological constraints of *Aedes aegypti*, vector of dengue. *Revue internationale de géomatique*, 27, 481-502.
- Petrov A., 2012. One Hundred Years of Dasymetric Mapping : Back to the Origin. *The Cartographic Journal*, 49, 256-264.
- Salje H., Lessler J., Endy T.P., Curriero F.C., Gibbons R.V., Nisalak A., Nimmannitya S., Kalayanarooj S., Jarman R.G., Thomas S.J., Burke D.S., Cummings D.A., 2012. Revealing the microscale spatial signature of dengue transmission and immunity in an urban population. *Proceedings of the National Academy of Sciences of the United States of America*, 109 (24), 9535-9538.

Viel J.-F., Tran A., 2009. Estimating denominators: satellite-based population estimates at a fine spatial resolution in a European urban area. *Epidemiology*, 20 (2), 214-222.

Yameogo B.F., Vandanjon P.-O., Gastineau P., Hankach P., 2021. Generating a Two-Layered Synthetic Population for French Municipalities: Results and Evaluation of Four Synthetic Reconstruction Methods. *Journal of Artificial Societies and Social Simulation*, 24 (2), 5.

Ye X., Konduri K.C., Pendyala R.M., Sana B., Waddell P., 2009. Methodology to Match Distributions of Both Household and Person Attributes in Generation of Synthetic Populations, in *Transportation Research Board 88th Annual Meeting Compendium of Papers DVD*, 11-15 January 2009, Washington DC, USA, 24 p.

Chapter 5

Satellite image texture and characterisation of urban environments favourable to vector mosquitoes

*Claire Teillet, Ophélie Hoarau, Nausicaa Habchi-Hanriot,
Benjamin Pillot, Thibault Catry, Annelise Tran*

Texture is an important concept in image processing. It is used during classification, segmentation and image generation (Busch *et al.*, 2004). Texture data is frequently employed in conjunction with spectral information to describe an image (Pacifi *et al.*, 2009). This is achieved by decomposing the image into frequency and orientation components. Haralick (1979) demonstrated that a texture can be represented by two dimensions. The first dimension corresponds to the description of the frequency and orientation of the texture's primitives, while the second dimension corresponds to the spatial organisation of these primitives (Haralick, 1979). The concept of texture is highly dependent on the specific context and challenging to generalise. It is fundamentally rooted in the study of homogeneity within grayscale levels through the identification of primitives within the image (Caloz and Collet, 2001).

The main concepts in texture analysis are illustrated in Figure 5.1. The regular repetition of a primitive in one or more directions corresponds to an anisotropic image, whereas the random distribution of pixels in all directions is considered an isotropic texture, which is characterised by a certain level of homogeneity in its structure. An anisotropic texture is therefore characterised by direction-dependent properties, whereas an isotropic texture has properties that are independent of direction. The terms microtexture and macrottexture are used to describe the respective levels of detail in a texture, from the very fine to the coarser (Figure 5.1b). In addition, texture varies with the scale of observation. An analysis window of an object must be large enough to include the primitive defining the texture, but sufficiently small to preserve spatial dependency and enable identification of the texture that corresponds to the object (Figure 5.1c). During changes in scale, and thus the analysis window, one primitive may be lost whereas others may be emphasised. This concept of texture analysis scaling is analogous to the spatial resolutions of Earth observation images employed as inputs in textural information extraction methods (Figure 5.1d).

Texture analysis has been a widely employed to process and extract information from satellite images, particularly in the context of studies on forest and urban

environments (Lefebvre, 2011; Pandey *et al.*, 2021; Proisy *et al.*, 2007; Puissant *et al.*, 2006; Ruiz Hernandez and Shi, 2018). In urban environments, the density and proportion of vegetation and built-up areas, as well as the spatial layout between the two, are the primary components of texture. To date, texture data has been very rarely used in the study of vector-borne diseases caused by certain species of mosquitoes (see Chapter 1). However, it can serve to augment the environmental, climatic and demographic indicators presented in preceding chapters, particularly in the case of dengue vector mosquitoes *Ae. aegypti* and *Ae. albopictus*, which are predominantly found in urban settings.

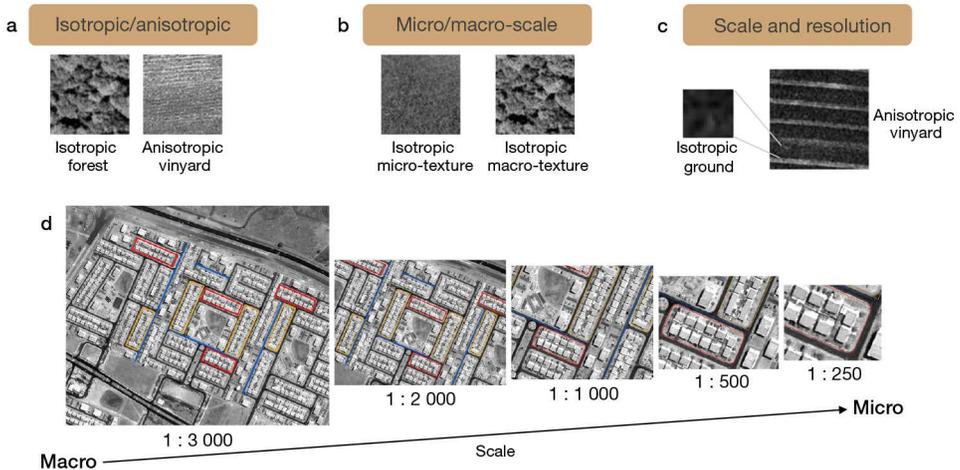


Figure 5.1. Key concepts for the analysis of satellite image textures.

This chapter presents a number of methods for the analysis of textures. In particular, the implementation of the FOTO approach (Couteron *et al.*, 2006) to very high spatial resolution Pléiades images has facilitated the study of the relationships between urban variables, the distribution of dengue cases and the availability of larval habitats in Brazil and La Réunion.

► Different methods to characterise image texture

A number of methods exist to analyse textures, including statistical analysis and frequency-domain analysis.

An example of statistical analysis: the grey level co-occurrence matrix and Haralick features

This approach is based on the statistical representation of grey levels, which consists of calculating a co-occurrence matrix for said levels (GLCM, Grey Level Co-occurrence Matrix). This matrix is the graphical representation of the number of occurrences of grey level pairs separated by a certain distance in a given direction (Figure 5.2a). Haralick (1979) proposed the calculation of features based on this matrix, with those most relevant to the study of urban environments being energy, entropy, contrast, correlation and homogeneity (Maillard, 2003; Pacifici *et al.*, 2009).

An example of frequency-domain analysis: the Fourier transform

In the nineteenth century, Fourier demonstrated that periodic functions can be expressed as a sum of sine and cosine functions. A discrete signal can thus be represented by a function composed of a fundamental frequency (the lowest), harmonics (multiples) and Fourier series coefficients. It is these coefficients that facilitate an understanding of the contribution of each frequency to signal formation (Caloz and Collet, 2001).

The frequency-domain representation of a Fourier transform is an amplitude spectrum, which corresponds to energy values for each frequency (Figure 5.2b). The spectrum considers the distribution of energy within the image while preserving information on the texture's frequency and orientation. A greater distance from the centre results in a higher observed frequency. A smooth texture will exhibit a spectrum that corresponds to low-frequency values at the centre. In contrast, a coarse texture will display values that are more spread out across the spectrum corresponding to higher frequencies (Regniers, 2014). Fourier transforms facilitate a transition from the spatial domain to the frequency domain. This is achieved by the decomposition of the signals that constitute the texture into their constituent frequencies.

The FOTOTEX algorithm is based on this frequency-domain approach utilising Fourier transforms. The following section presents a detailed account of the method and principle of implementing this algorithm.

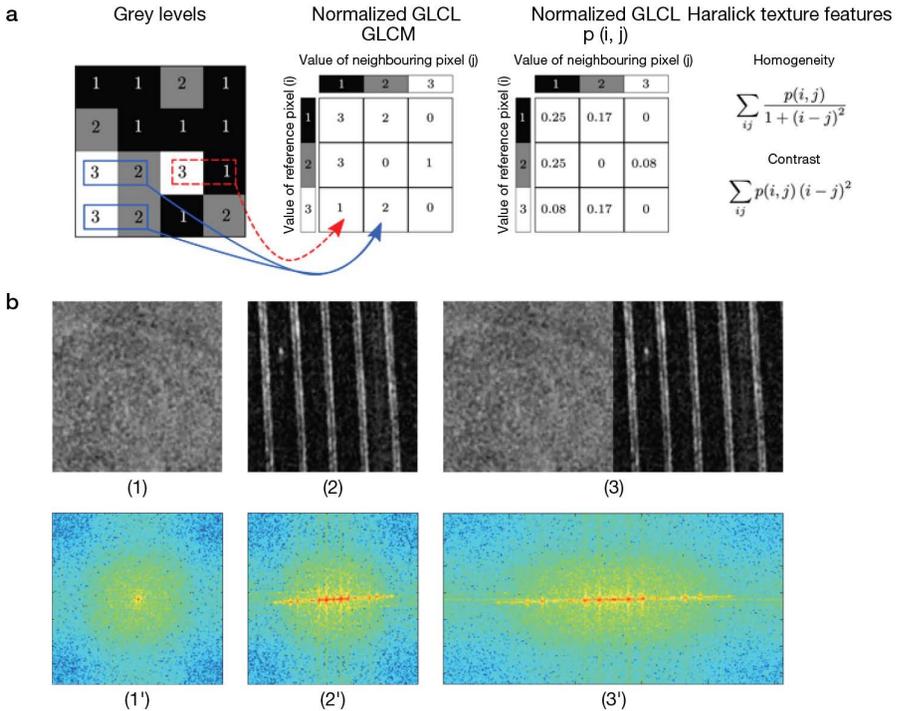


Figure 5.2. Example of texture analysis using (a) statistical analysis via grey level co-occurrence matrix and Haralick features, and (b) frequency-domain analysis via 2D Fourier transform spectra for: (1-1') grasslands, (2-2') cereals, (3-3') grasslands + cereals. Sources: (a) Brynolfsson *et al.* (2017); (b) Lefebvre (2011).

►► Study of the relationships between urban variables and the distribution of dengue cases in Brasília using a texture-based approach

Context

In the framework of the Apureza project (Remote sensing analysis of relationships between urban landscape factors of dengue and Zika) financed by the CNES (2017-2020), a multi-scale classification algorithm for the textures of urban areas was developed. It aimed to evaluate the value of a frequency-domain approach to texture analysis in the study of relationships between urban variables and the distribution of dengue cases, in particular in the administrative region of Sao Sebastiao in the city of Brasilia (Teillet *et al.*, 2021).

In recent years, a dramatic rise in cases of dengue has been observed in Brasilia: in 2004-2005, there were 200 confirmed cases, but by 2015, this number had reached 11,722.

Data

In order to characterise the urban footprint, a high resolution (HR) Sentinel-2 image acquired on 31 August 2018 was used, downloaded from the ESA's (European Space Agency) Copernicus²⁹ data space ecosystem for Sentinel products. Sentinel-2 images comprise thirteen spectral bands, with a spatial resolution of between 10 and 60 m.

To characterise inner-city environments, a very high resolution (VHR) Pléiades image, captured on 14 January 2020, was provided by the GEOSUD team in Montpellier³⁰. This image includes multiple bands: a panchromatic band P (480-830 nm) at 0.7 m resolution, resampled to 0.5 m, and multispectral bands (MS) at 2.8 m, resampled to 2 m, including blue (430-550 nm), green (490-610 nm), red (600-720 nm) and near-infrared (750-950 nm) channels.

Method

The methodology developed uses texture information extracted from Earth observation data acquired at different spatial resolutions in order to characterise urban landscapes of the Global South at three different scales, defined by Theia's "Urban" Scientific Expertise Centre (SEC)³¹:

- urban footprint (macro-scale);
- urban cell scale (meso-scale);
- object scale such as buildings (micro-scale).

For this, the FOTOTEX algorithm, which is based on the FOTO (Fourier-based Textural Ordination) algorithm, is employed. The original objective of the FOTO algorithm was to characterise ecosystem vegetation models (Couteron *et al.*, 2006; Proisy *et al.*, 2007). In order to provide a simple, robust and effective tool, this version of the algorithm was developed in Python and made available online³². The algorithm

29. <https://dataspace.copernicus.eu/>

30. <https://ids-dinamis.data-terra.org/>

31. <https://www.theia-land.fr/en/ceslist/urban-sec/>

32. <https://framagit.org/espace-dev/fototex>

has been optimised in terms of process and calculation time. Further details on the utilisation and configuration of parameters for the FOTOTEX algorithm can be found in the article by Teillet *et al.* (2021). The method is summarised below:

Step 1: Image partitioning

The first step in this method is to partition the image (Figure 5.3a). This partitioning corresponds to defining a parsing window's path rule on the image. Two methods for partitioning images are available. The first is the block method (method = “*block*”), which analyses the texture for each image window block by block. The second is the moving method (method = “*moving window*”), which analyses texture for a given window by sliding from west to east and from north to south with a shift of one pixel.

Step 2: Spectral analysis by Fourier transform and “r-spectra” computing

A fast Fourier transform (FFT) is computed for each analysis window of the original image and allows the variance of an image to be expressed as a weighted sum of cosine and sine waveforms of varying direction and spatial frequency (Couteron *et al.*, 2006; Lang *et al.*, 2018; Proisy *et al.*, 2007; Figure 5.3b). The weighting coefficients quantify the contribution of each frequency and each direction to signal formation. A “periodogram” is computed for each window and expresses the proportion of variance explained by each frequency pair (p,q), where p and q are spatial frequencies along the image column and row directions (Proisy *et al.*, 2007). Finally, an isotropic r-spectrum is computed by averaging the periodogram in all directions (from 0° to 360°)

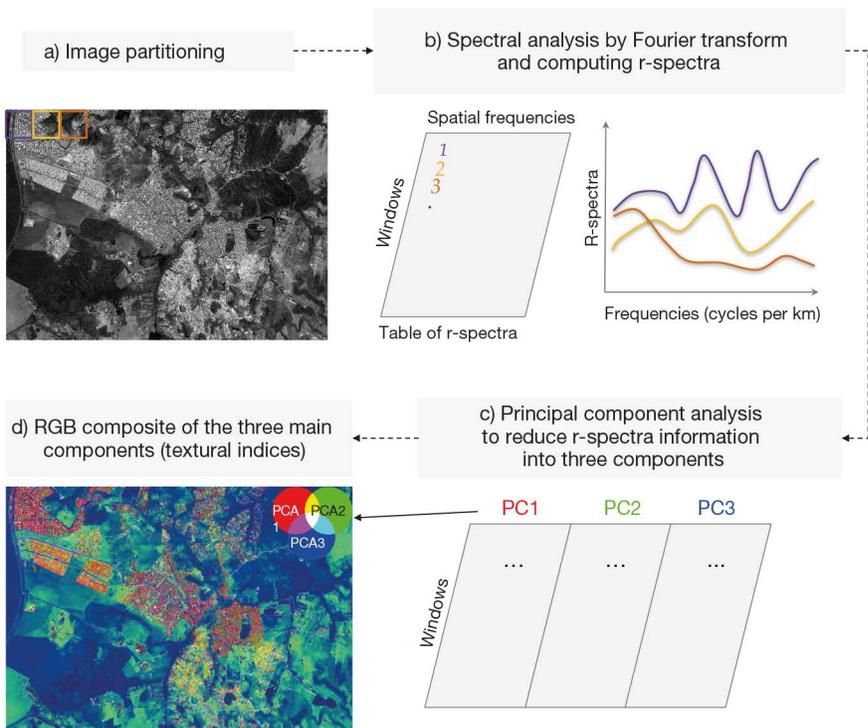


Figure 5.3. Simplified methodological framework of the FOTOTEX algorithm.

to obtain an averaged radial spectrum, called “r-spectrum” and denoted by $I(r)$, for each analysis window. A table compiling the r-spectra is produced. In this table, each row corresponds to the r-spectrum of a window, and each column corresponds to a spatial frequency (Lang *et al.*, 2018).

Step 3: Principal component analysis

Principal component analysis (PCA) is applied to the r-spectra table in order to summarise the information content (Figure 5.3c and d). This statistical method ordiates the correlated variables into new variables, called “principal components” or “principal axes”, which are decorrelated from each other—the information is thereby reduced, since the first principal axes are sufficient to represent a large portion of total variability. The first three main components resulting from PCA are textural indices which can then be visualised in form of a red-green-blue (RGB) composite image (Proisy *et al.*, 2007).

Results

Figure 5.4 shows the results obtained for Sao Sebastiao, a satellite City of Brasilia in Brazil (Figure 5.4a). Texture information obtained with FOTOTEX allows the urban footprint (macro-scale) to be outlined in the Sentinel-2 image, in addition to the identification of neighbourhoods represented by different texture information within this urban area (Figure 5.4b).

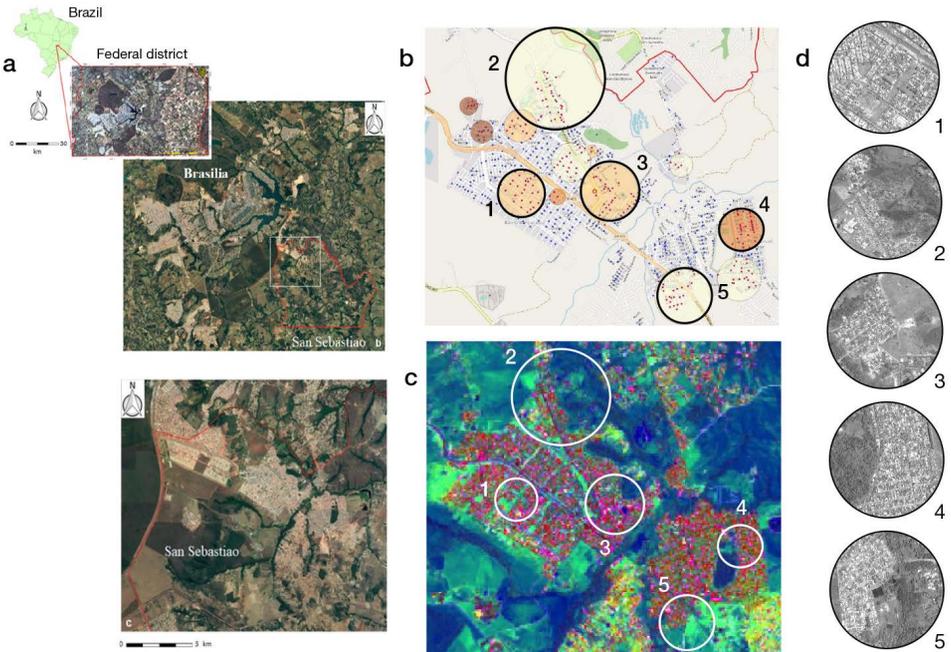


Figure 5.4. Distribution of dengue case clusters in Sao Sebastiao (Brasilia) as a function of neighbourhood type identified by Pléiades image textures.

(a) Location of the area of study, the administrative region of Sao Sebastiao, Brasilia (Brazil). (b) Case clusters of dengue identified by SATSCAN from 2007-2017. (c) RGB colour composite output by FOTOTEX. (d) Organisation of built-up areas and vegetation on a Pléiades image defining the urban typology corresponding to each cluster.

These neighbourhoods are linked to a heterogeneous spatial and temporal distribution of dengue cases (Figure 5.4c). Case clusters were identified using SATSCAN³³ software on geolocation data for dengue cases in the area from 2007-2017 sourced from the SINAN database³⁴. Each cluster is characterised by a distinctive urban landscape configuration, with a specific spatial organisation between built-up areas (structure, density, height) and urban vegetation (Figure 5.4d).

The results of statistical analyses have demonstrated that the density of the built-up area and the proportion of vegetation, which are the primary factors comprising texture information in urban environments, are positively correlated with the number of dengue cases in Sao Sebastiao. Positive correlations also exist with other variables, such as land surface temperature.

►► Map of potential larval habitat distribution of the Asian tiger mosquito on Reunion island

Context

In the context of the ANISETTE project (Inter-Site Analysis: Evaluation of Remote Sensing as a predictive tool for the surveillance and control of diseases caused by mosquito), the FOTOTEX algorithm described above was applied to Reunion island. This French department, situated in the Indian Ocean, has rugged terrain (between 0 and 3070 m), resulting in the concentration of habitats along coastal areas. On this island, the Asian tiger mosquito, *Aedes albopictus*, is the primary vector of arboviruses, such as dengue (epidemics in 1977, 2004 and since 2016) and chikungunya (epidemic in 2005-2006). In this study, the relationships between the textural indices obtained and data collected in the field on the number of larval habitats of Asian tiger mosquitoes were analysed. This was done to map areas conducive to *Ae. albopictus* proliferation at a regional scale (Hoarau, 2021).

Data

Asian tiger mosquito larval habitat data

As part of its vector control (VC) efforts, the regional health agency (ARS) of La Réunion collected data on larval habitats by going door-to-door and prospecting gardens (Figure 5.5). For each area of entomological surveillance (AES), comprised of around 200 houses, data was collected on the number and type of potential breeding sites. For this study, we had access to data on 346 AES (out of a total of 1,203 monitored by the ARS, i.e., one quarter) collected between January 2020 and March 2022.

Environmental data

The Pléiades 2020 image mosaic of La Réunion is a product generated from twelve images provided in the context of the Kalideos programme³⁵. The characteristics of these images have been described in previous sections.

33. <https://www.satscan.org/>

34. <https://portalsinan.saude.gov.br/>

35. <https://www.kalideos.fr>

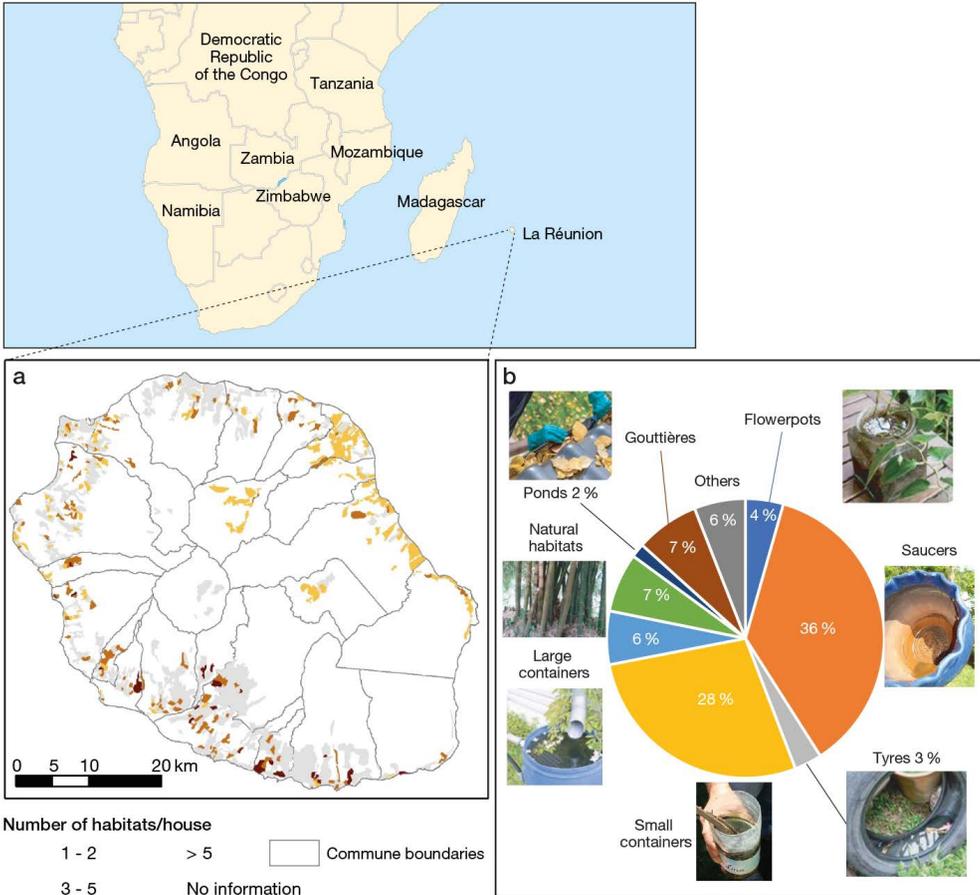


Figure 5.5. Type and distribution of Asian tiger mosquito larval habitats in La Réunion, 2021-2022.

(a) Number of potential habitats observed per neighbourhood. (b) Distribution according to type of habitat observed. Source: (a) ARS La Réunion data.

Data on annual rainfall and mean annual temperature, obtained by interpolating data from the Météo-France/CIRAD monitoring network, was downloaded from AWARE, the Agricultural Web Atlas for Research³⁶.

Methods

Textural index calculation

After performing different tests on the image extracts, the FOTOTEX algorithm was applied following the steps described above to the built-up areas of the Pléiades mosaic (Figure 5.6a), with a square window measuring 67 pixels per side (i.e., 33.5 m) and using the block method of image partitioning (Figure 5.6b).

36. <https://aware.cirad.fr/>

Spectral index calculation

Several indices were calculated from the Pléiades mosaic: vegetation indices (Figure 5.6c) and normalized difference water indices (these indices are described in Chapter 2), in addition to the brightness index (BI) calculated from the red and near-infrared bands which highlight reflective surfaces such as buildings, roads, bare ground, etc. (Figure 5.6d).

Statistical analyses

The correlations between the number of potential larval habitats observed (response variable) and the average textural and spectral indices, as well as mean annual rainfall and temperature for each neighbourhood (explanatory variables), were analysed using univariate analysis (by testing the variables one by one), then using multivariate analysis, combining the most relevant variables with the help of a generalised linear model (GLM).

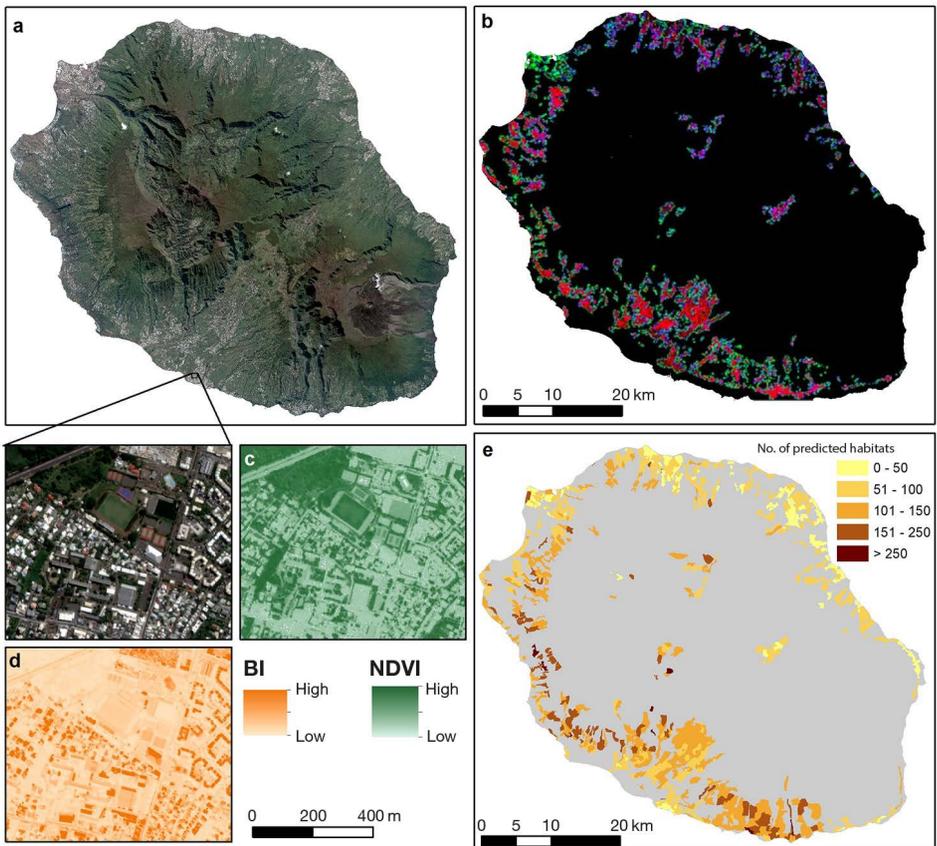


Figure 5.6. Illustration of textural and spectral indices calculated from a Pléiades mosaic, Reunion island, 2020, and predictive map of the number of potential larval habitats of the Asian tiger mosquito per area of entomological surveillance (AES).

(a) Pléiades 2020 mosaic. (b) RGB colour composite output by FOTOTEX. (c) Vegetation index (NDVI). (d) Brightness index (BI) (e) number of predicted potential larval habitats per AES.

Results

As with the study sites in Brazil, the visualisation of textural indices output by the FOTOTEX algorithm provides insight into the typology of neighbourhoods (Figure 5.6b):

- neighbourhoods with very high building density and little vegetation;
- neighbourhoods with more spread-out buildings, separated by vegetation;
- neighbourhoods with very low building density and lots of vegetation;
- industrial zones characterised by the presence of large buildings, large impervious surfaces and little vegetation.

These indices, combined with other environmental variables (spectral indices, rainfall, temperature) have been identified as highly correlated with field data on larval habitats of *Ae. albopictus* (Hoarau, 2021). The resulting relationship allows for the prediction of the number of potential larval habitats for all neighbourhoods on the island, based on the environmental variables (Figure 5.6e).

►► Conclusion

The results presented in this chapter, based on the use of FOTOTEX, demonstrate the value of textural information for the characterisation of urban areas. The addition of texture provides supplementary information to that derived from spectral indices or land use. In particular, it is possible to describe urban environments at different scales, including the ability to describe urban cells characterised by specific typologies (heterogeneous spatial organisation of built-up areas and vegetation). This information has proven to be particularly useful in the study of relationships between urban landscapes and the spatiotemporal dynamics of dengue using very high spatial resolution satellite imagery (Pléiades, SPOT 6/7), as shown in Brazil and La Réunion.

In both of these areas, statistical analyses have highlighted that the texture signal of urban cells is highly correlated with the spatial distribution of dengue cases (Brazil), and it represents a reliable indicator of the number of potential larval habitats in urban areas (La Réunion). The identification of these correlations represents a significant contribution to the study of mosquito population dynamics and the risk of dengue transmission in urban environments. By fairly accurately predicting the number of potential larval habitats, the texture contained in very high spatial resolution satellite images can compensate for the lack of *in situ* data.

These datasets, which are essential for calibrating and validating model outputs for predictive mosquito mapping, such as AlboRun (Tran *et al.*, 2020) or Arbocarto (see Chapter 8), are unfortunately scarce, which often significantly constrains their applications in public health. It is thus feasible to employ texture information as an input variable in models of sites where larval habitat data is lacking. Furthermore, the links between texture and the spatiotemporal dynamics of cases presents a novel avenue for investigating dengue transmission risk based on satellite imagery. However, this remains a very complex mechanism involving a multitude of variables, thereby necessitating the use of additional data to supplement satellite imagery.

►► References

- Brynnolfsson P., Nilsson D., Torheim T., Asklund T., Karlsson C.T., Trygg J., Nyholm T., Garpebring A., 2017. Haralick texture features from apparent diffusion coefficient (ADC) MRI images depend on imaging and pre-processing parameters. *Scientific Reports*, 7 (1), 4041.
- Busch A., Boles W., Sridharan S., 2004. Logarithmic quantisation of wavelet coefficients for improved texture classification performance, in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 17-21 May 2004, Montreal, QC, Canada, IEEE, iii-569.
- Caloz R., Collet C., 2001. *Précis de télédétection – Volume 3 : Traitements numériques d'images de télédétection*, Québec, Presses de l'Université du Québec – Agence Universitaire de la Francophonie, 386 p. (coll. Universités francophones).
- Couteron P., Barbier N., Gautier D., 2006. Textural Ordination Based on Fourier Spectral Decomposition: A Method to Analyze and Compare Landscape Patterns. *Landscape Ecology*, 21, 555–567.
- Haralick R., 1979. Statistical and Structural Approaches to Texture. *Proceedings of the IEEE*, 67 (5), 786-804.
- Hoarau O., 2021. *Télédétection et modélisation de populations de moustiques vecteurs : Extraction d'indicateurs paysagers à partir d'images d'Observation de la Terre pour l'estimation de la distribution des gîtes larvaires en milieu urbain*, mémoire de master 2, spécialité Sciences de la terre, des planètes et de l'environnement, Université de La Réunion, Saint-Denis, 38 p.
- Lang M., Alleaume S., Luque S., Baghdadi N., Féret J.-B., 2018. Monitoring and Characterizing Heterogeneous Mediterranean Landscapes with Continuous Textural Indices Based on VHSR Imagery. *Remote Sensing*, 10 (6), 868.
- Lefebvre A., 2011. *Contribution de la texture pour l'analyse d'images à très haute résolution spatiale : application à la détection de changement en milieu périurbain*, thèse de doctorat, Université Rennes 2, Université Européenne de Bretagne, 284 p.
- Maillard P., 2003. Comparing Texture Analysis Methods through Classification. *Photogrammetric Engineering & Remote Sensing*, 4 (11), 357-367.
- Pacifici F., Chini M., Emery W., 2009. A neural network approach using multi-scale textural metrics from very high-resolution panchromatic imagery for urban land-use classification. *Remote Sensing of Environment*, 113 (6), 1276-1292.
- Pandey G., Sharma V., Chaudhary P., Chowdary V., 2021. Integration of Texture and Spectral Response with AI Techniques for Buildings Footprint Identification Using High-Resolution Satellite Images. *Journal of the Indian Society of Remote Sensing*, 49 (6), 1439-1452.
- Proisy C., Couteron P., Fromard F., 2007. Predicting and mapping mangrove biomass from canopy grain analysis using Fourier-based textural ordination of IKONOS images. *Remote Sensing of Environment*, 109 (3), 379-392.
- Puissant A., Hirsch J., Weber C., 2006. The utility of texture analysis to improve per-pixel classification for high to very high spatial resolution imagery. *International Journal of Remote Sensing*, 26 (4), 733-745.
- Regniers O., 2014. *Méthodes d'analyse de texture pour la cartographie d'occupations du sol par télédétection très haute résolution : Application à la forêt, la vigne et les parcs ostréicoles*, thèse de doctorat, Université de Bordeaux, 165 p. <https://hal.archives-ouvertes.fr/tel-01143739>.
- Ruiz Hernandez I., Shi W., 2018. A Random Forests classification method for urban land-use mapping integrating spatial metrics and texture analysis. *International Journal of Remote Sensing*, 39 (4), 1175-1198.
- Teillet C., Pillot B., Catry T., Demagistri L., Lyszczarz D., Lang M., Couteron P., Barbier N., Adou Kouassi A., Gunther Q., Dessay N., 2021. Fast Unsupervised Multi-Scale Characterization of Urban Landscapes Based on Earth Observation Data. *Remote Sensing*, 13 (12), 2398.
- Tran A., Mangeas M., Demarchi M., Roux E., Degenne P., Haramboure M., Le Goff G., Damiens D., Gouagna L.C., Herbreteau V., Dehecq J.-S., 2020. Complementarity of empirical and process-based approaches to modelling mosquito population dynamics with *Aedes albopictus* as an example - Application to the development of an operational mapping tool of vector populations. *PLoS One*, 15 (1), e0227407.

Part 2

Analysing and predicting the effect of environmental variables on the distribution and dynamics of vector mosquitoes

A model is a simplified representation of a real-world system and therefore always contains a number of simplifications in order to answer a given question. This is why modellers often quote the phrase of the statistician George Box, “*All models are wrong but some are useful*”. With regard to vector-borne diseases, and in particular those involving mosquitoes, the modelling approaches are highly complementary to experimental and observational approaches. Modelling can be used from two different perspectives. This first is about gaining a better understanding of the distribution of vectors, their dynamics, and their links with environmental and climatic variables (such as those outlined in Part 1), in addition to the epidemiological transmission cycle of vectored pathogens. The second allows for the prediction of the presence of vector mosquitoes, their abundance, and the risk of transmission, as well as the variation of these factors in different scenarios, in both space and time. Modelling approaches are varied, and the choice depends on the question needing to be answered, the knowledge about the system, available data, etc.

In Part 2, different modelling approaches are presented, as well as the operational tools for entomological or epidemiological surveillance and vector control developed using these approaches, with examples given for different geographical contexts. The first chapter of Part 2 (Chapter 6) describes data-driven models (models of species distribution or ecological niches). The following chapters present knowledge-based approaches (Chapter 7) and processes at the scale of populations (mechanistic models: Chapter 8) or individuals (behavioural models: Chapter 9).

Chapter 6

Data-driven models: mapping the spatial distribution of vectors

Yi Moua, Emmanuel Roux

The mapping of a species (animal or plant), and *a fortiori* a vector mosquito species, may be done with the help of expert knowledge and the use of the multi-criteria analysis methods discussed in the subsequent chapter (Chapter 7). Alternatively, the mapping process may result from the implementation of so-called “mechanistic” models, i.e., those explicitly based on the knowledge of bioecological processes which govern the life cycle of individual members of a species (see Chapters 8 and 9). However, there may be a lack of knowledge about the species in question, which may impede the construction of such models. An alternative approach is to construct models based on observations made at specific sites, with the aim of developing a model of the habitats of the species. This enables the quality of these habitats (which can be considered a proxy for the probability of species presence) to be predicted for the entire study area. This chapter addresses these types of models, referred to as “species distribution models” or “ecological niche models”. In particular, it describes one of the most commonly used models, called Maxent, as well as its application to mapping the primary vector of malaria in the Amazon region, *Anopheles darlingi*. This chapter also describes the innovative solutions proposed by the authors to minimise the impact of sampling bias on the results of such models.

►► Species distribution models

The construction of species distribution or ecological niche models is based on observational data, with the objective of spatially predicting the habitat quality of one or more species (animals or plants). These models establish a link between the documented distribution of species and the environmental data that characterise the habitats in which they have been observed (Guisan and Thuiller, 2005). The final output is a habitat quality map, which can be converted into a probability map indicating the likelihood of the species in question occurring in a given area.

Since the 1990’s, the use of these models has steadily increased. They are used for different purposes:

- to better characterise the environmental niches of species;
- to predict the distribution and range shifts of invasive species;

- to evaluate the impact of climate change, land use, and land cover on species distribution;
- to predict the distribution of rare species facing extinction, with a particular focus on providing support for conservation and reintroduction efforts for endangered species.

In particular, with regard to the prediction of vector mosquito distribution, Boosted Regression Trees, based on the principle of ensemble modelling and implementing regression trees, have been used to map the distribution of primary malaria vectors on the American continent (Sinka *et al.*, 2010) and at a global scale (Sinka *et al.*, 2012).

The environmental variables used as inputs provide information on the environmental conditions of the area of study. They are referred to as “explanatory variables”, “predictors”, “covariables” or “inputs”. These are often derived from satellite images, aerial photography or data extrapolated from field observations. The environmental conditions that they characterise may exert a direct or indirect influence on habitat quality, either positively or negatively (see Chapter 1).

A number of applications of species distribution models employ a set of variables, termed bioclimatic variables, derived from meteorological variables. These variables are intended to more accurately characterize the biological conditions associated with species habitats. A popular data source is the WorldClim database³⁷. A set of variables such as this is useful for large-scale studies, as well as for predicting the medium- to long-term effects of climate change on the distribution of the species in question, but does not explicitly take into account land cover or land use. High resolution land cover and land use data, along with the resulting landscape variables, allow for an examination of the complex relationships between species presence and the landscape at a local scale—due to the lack of high spatial resolution data available at larger scales—, in addition to the explicit consideration of anthropogenic activities and their impacts. Conversely, the capacity to make predictions according to different global change scenarios is diminished.

In terms of species observation data, there are two main approaches. The first model is based on presence-absence data for species, and can implement a statistical approach for discrimination or supervised learning. The second model is based on presence-only data, involving specific approaches to compensate for the absence of “counter-examples” in the observations. The second approach has proven to be of particular interest, given that information regarding the absence of species is often unavailable or unreliable. The absence of a species in a given location may be due to:

- non-detection of the species even though it is effectively present;
- actual absence of the species in a suitable habitat, for reasons related to dispersion: the species has not yet colonised the environment, but will arrive sooner or later; ecological barriers preventing the species from accessing certain parts of the region;
- actual absence of the species due to unsuitable ecological conditions.

Presence-only data may originate from different sources: historic databases and collections from universities, institutes, or museums; online databases dedicated to biodiversity (e.g., Global Biodiversity Information Facility – GBIF³⁸) or to specific

37. <https://www.worldclim.org/data/index.html>

38. <https://www.gbif.org/>

species, such as VectorBase³⁹, which is dedicated to invertebrate vectors of human pathogens; field observations performed specifically for the study.

One of the most commonly used models which employs presence-only data is Maxent.

►► Maxent model

The Maxent model is based on the maximum entropy principle (Elith *et al.*, 2011). In information science, entropy is a measure of information quantity and uncertainty. High information entropy is associated with varied and rich information content, better able to differentiate between different situations or objects to which the information pertains. Maxent is a method for determining the probability distribution of a species' occurrence across a given geographic region (i.e., all pixels within the area of study), where entropy is at a maximum and observational constraints have been satisfied. These constraints take into account the environmental conditions in which the species was actually observed. For each environmental variable, these constraints are expressed mathematically by the relationship between the expected value of the variable—associated with the probability distribution of occurrence of the species—and the sample mean of the variable in sites where the species was actually observed (sites of occurrence). This model has been shown to give reliable results, even with a low number of observations (Hernandez *et al.*, 2006).

The Maxent algorithm is available in the *dismo* package (Species Distribution Modelling; Hijmans *et al.*, 2017) of the R programming environment⁴⁰.

►► Sampling bias and minimising its impact on modelling

Every analysis or model based on observations will be biased if the observations themselves were made in a biased manner and if no procedure was in place to, at minimum, identify the biases and, at best, minimise their impact.

If the default Maxent parameters are used, there is a random uniform sampling of a large number of “background” sites, assumed to represent the set of environmental conditions of the area of study (in our example, the territory of French Guiana) and their relative frequency. Such a configuration assumes that sampling (species counts) was carried out in a balanced manner, i.e., the various environmental conditions which exist were sampled proportionally relative to their frequency across the area in question. However, this is seldom the case, with counts being taken in areas which are accessible, most often in the vicinity of thoroughfares, and according to other criteria, explicit or otherwise. In the case of a vector mosquito species, sampling is often conducted primarily around confirmed cases of the disease.

We have therefore proposed a method to correct the sampling biases in models using presence-background data, which includes Maxent (Moua *et al.*, 2020). The method is based on an initial estimate of the sampling effort (in this case, based on the counts of all mosquito species which were performed) within the domain of environmental

39. <https://vectorbase.org/vectorbase/app/>

40. R Core Team, 2021. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, <https://www.R-project.org/>.

characteristics⁴¹. A random selection of background sites is then performed in this same domain, weighted by the previously estimated sampling effort. The method was evaluated using synthetic data, and in certain applications, it was found to be more effective than other debiasing approaches, particularly when applied to small datasets on sites of occurrence. The algorithm used to estimate sampling effort is written in pseudocode and evaluated by Yi *et al.* (2020).

► Application to the primary malaria vector in French Guiana

Malaria around the world and in French Guiana

Malaria is a disease caused by a parasite of the genus *Plasmodium* and is transmitted by species of mosquitoes of the genus *Anopheles*. In 2020, the estimated number of malaria cases in the 85 countries where the disease is endemic (among which is France—French Guiana) was 241 million. The number of deaths worldwide was estimated to be 627,000. The positive outcomes achieved in malaria control between 2000 and 2015 have prompted the United Nations to set the objective of eliminating the disease in at least 35 countries by 2030 (SDG Target 3.3 and Global technical strategy for malaria). However, since 2015, the situation has deteriorated, as evidenced by the latest World Malaria Report, which recorded a 12% increase in deaths between 2019 and 2020 (WHO, 2021).

French Guiana recorded an average of around 3,500 cases per year during the 2000's. This number drastically decreased at the start of the 2010's, subsequently stabilising at around a few hundred cases per year after 2013. However, major outbreaks were observed in 2017 and 2018 (Mosnier *et al.*, 2020). Since 2019, the number of cases has once again decreased, and the objective of eradicating the disease within the French territory has been reaffirmed, notably in 2019 by Agnès Buzyn, then Minister for Solidarity and Health, during the 6th replenishment conference of the Global Fund to Fight AIDS, Tuberculosis and Malaria.

The Global Technical Strategy for Malaria incorporates vector control and risk assessment to enhance the precision of intervention strategies and/or anticipate the potential for recurrence in regions where the disease has been eradicated. This is based, in part, on a more comprehensive understanding of the habitats that are conducive to the proliferation of mosquitoes, their spatial distribution, and the probability of occurrence of the vector species.

Anopheles darlingi, the primary malaria vector in French Guiana

Anopheles darlingi can be found in the rural areas of French Guiana. The species exhibits a preference for taking blood meals from humans (so called anthropophilic behaviour), both indoors and outdoors (exo-endophagic), and prefers outdoors resting sites with vegetation. It prioritises larval habitats that receive sunlight, but with enough shade to maintain a temperature between 20 and 28°C (Hiwat and Bretas, 2011). The larvae of this species have been observed in bodies of clean and fresh water, with little to no

41. In other words, it is an estimate of the density of sample sites (whether or not the species was observed) in a space defined by the set of variables describing the area of study's environmental conditions, and not in a geographic space (latitude, longitude).

current and some vegetation, such as riverbanks, creeks, pools formed near riverbeds left behind after flooding, wetlands, and flooded or flood-prone savannahs and forests (Hiwat *et al.*, 2010; Rozendaal, 1987). Larval habitats are rarely located in dense forest due to the acidity of the water and the lack of sunlight under the forest canopy.

An. darlingi is present in low numbers across a range of environments, with an expansive distribution. This species is also extremely difficult to raise in a laboratory setting. The combination of these factors presents a significant challenge in terms of capturing the species, which in turn makes it difficult to map its spatial distribution and study its bioecology. In this context, species distribution modelling—and models such as Maxent in particular—has proven to be particularly useful. The work described below was published in the *Journal of Medical Entomology* (Moua *et al.*, 2017).

***Anopheles darlingi* presence-only data**

A database of historical observations of the presence of *Anopheles* mosquitoes in French Guiana was constructed from the archives of the Institut Pasteur d'Algérie and Institut Pasteur de la Guyane (Pasteur Institutes of Algeria and French Guiana respectively), with activity reports from the DDAS (Directorate of Mosquito Control and Health Action) of the region of French Guiana, replaced by the DDAS of the territorial collectivity of French Guiana, the Service de santé des armées (Defence Health Service), and publications from the ORSTOM (Office of Scientific and Technical Research Overseas, later becoming the French National Research Institute for Sustainable Development, IRD, in 1998). It is available via the Global Biodiversity Information Facility (GBIF) platform (Moua *et al.*, 2019).

From this data, only the sites of occurrence of *An. darlingi* identified since 2000 were taken into account for the model. This approach was adopted in order to avoid considering sites which may have undergone significant change over time and for which information on the presence of the species could be called into question. A total of 48 sites were identified, although some were located within the same 1 km² pixel. Consequently, only 39 pixels were included in the Maxent input data set.

Nearly all of the sites in French Guiana with confirmed *An. darlingi* presence are located in coastal areas or along the two rivers forming the international borders, which are easy to access and are home to the vast majority of the population. This suggests that there is a significant sampling bias and justifies the use of the method described above which aims to minimise the impact of such biases on modelling results.

Environmental data

The selection of environmental data was based on the findings presented in the existing literature and the expertise of entomologists at the Institut Pasteur de la Guyane. From these sources, it has been shown that three major types of environments must be considered:

- natural environments, for which the presence of *An. darlingi* depends on the value or class of the associated environmental variable (NAT_ENV);
- environments associated with anthropic activities which do not permanently alter the local natural environment, and which exert a positive influence on the presence of *An. darlingi* (ANTHROP_NON_PERM);

– highly anthropized environments with long-term effects, corresponding to the presence of human activities which permanently alter the natural environments across a wide area and which limit the presence of *An. darlingi* (ANTHROP).

Each of these environments was characterised by a set of variables (see Table 6.1), most of which come from remotely sensed data with normalised spatial resolutions of 1 km.

Table 6.1. Environmental data selected to model suitable habitats for *Anopheles darlingi* in French Guiana. Source: Moua *et al.* (2017).

Variable	Source data and remotely sensed data employed	Type of variable	Type of environment to which the variable primarily refers
Altitude	Shuttle Radar Topography Mission (SRTM), NASA	Continuous numeric (quantitative)	NAT_ENV
Geomorphological landscapes	Geomorphological landscapes (Guitet <i>et al.</i> , 2013)	Categorical (qualitative)	NAT_ENV
Geomorphological units	derived from the SRTM	Categorical (qualitative)	NAT_ENV
Land cover	Gond <i>et al.</i> (2011), using SPOT4/VEGETATION data	Categorical (qualitative)	NAT_ENV
Human activities and presence which do not permanently affect the local environment (HUMACT)	Human footprint (Thoisy de <i>et al.</i> , 2010)	Ordinal categorical (semi-quantitative)	ANTHROP_NON_PERM
Density of roads and tracks (ROADS)	BD TOPO®, IGN	Continuous numeric (quantitative)	ANTHROP_NON_PERM
Percentage of urbanisation in the surroundings (URBAN)	Human footprint (Thoisy de <i>et al.</i> , 2010)	Continuous numeric (quantitative)	ANTHROP

The full method of implementing Maxent for mapping *An. darlingi* habitats in French Guiana is shown in Figure 6.1.

Results

Figure 6.2 represents the habitat quality for *An. darlingi*. The probability of species presence is linked monotonically with habitat quality (as quality increases, so does the chance of the species being present).

The results show that habitat quality is strongly linked to human activities and presence, characterised by the variables HUMACT and ROADS, in particular along the coast, and in the centre and west of the territory, even though the thoroughfares in the latter two areas are essentially rivers and unpaved trails.

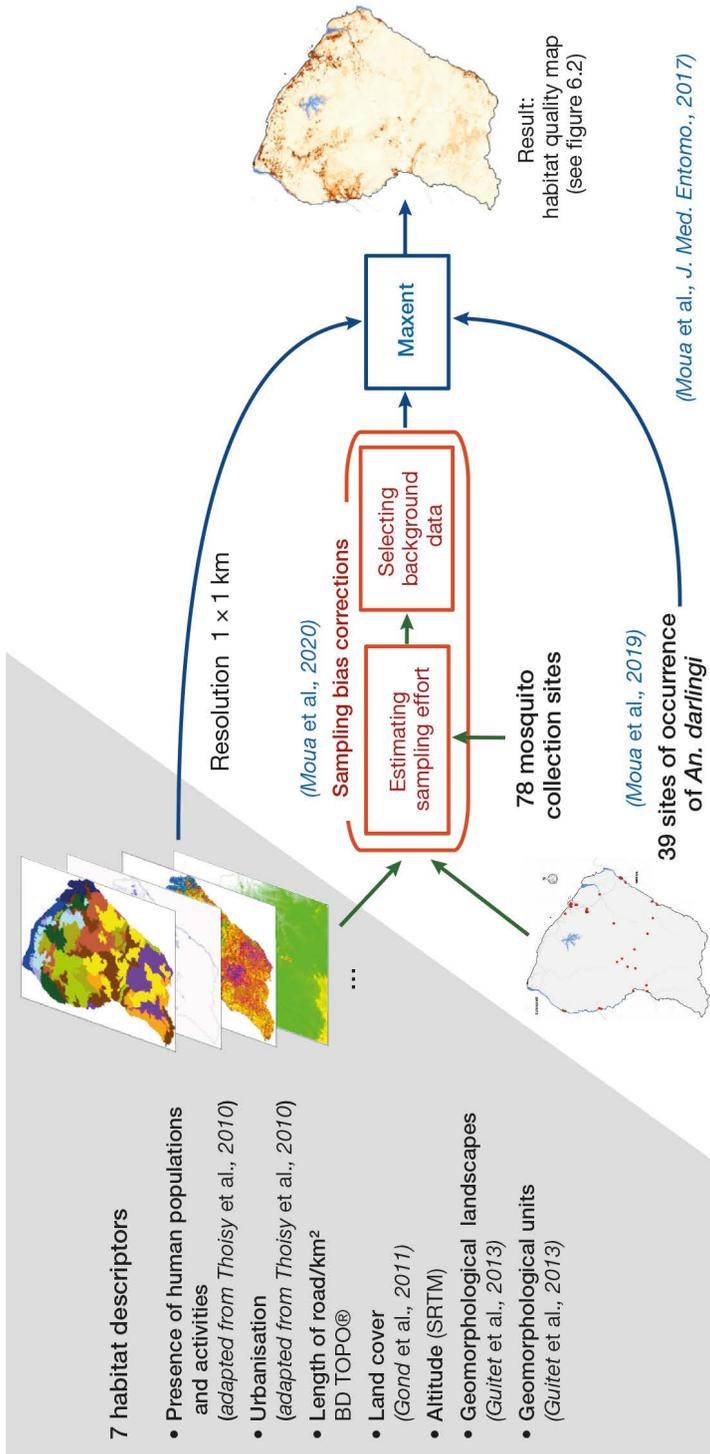


Figure 6.1. General method for qualitative mapping of habitats suitable for *Anopheles darlingi* in French Guiana, minimising the impacts of sampling bias.

The significant contribution of the length of roads and tracks (ROADS) to habitat quality appears to be consistent with the knowledge about the impact of environmental modifications and human activities: the construction of these routes and trails is accompanied by deforestation, potentially leading to new larval habitats (e.g., ditches) [Singer and Castro de, 2001], and bringing with it the sunlight required for larval development, as well as the presence of human hosts. However, it should be noted that above a certain threshold, around 7km of roads/tracks per km², habitat quality stops increasing and begins to decrease, implying that higher densities of road networks (characteristic of urbanisation in an area) lead to unfavourable environmental changes for *An. darlingi* (lack of vegetation for gravid females to rest, pollution of runoff water, etc.). Similarly, the phenomenon of urbanisation (characterised by the URBAN variable) is not favourable to *An. darlingi* (Stefani *et al.*, 2013). This is particularly evident in the areas of Cayenne and Kourou. Conversely, Saint-Laurent-du-Maroni, one of the three largest urban areas in Guiana by size and level of urbanisation, alongside Cayenne and Kourou, appears to have an extremely high habitat quality, notably due to a lower estimated built-up density. However, Saint-Laurent-du-Maroni has undergone sustained demographic growth (+2.3% per year between 2013 and 2019 according to the French National Institute of Statistics and Economic Studies, INSEE⁴²), meaning an update to the distribution map of suitable habitats for malaria vector mosquitoes is necessary, one which takes into account this recent expansion and higher built-up density.

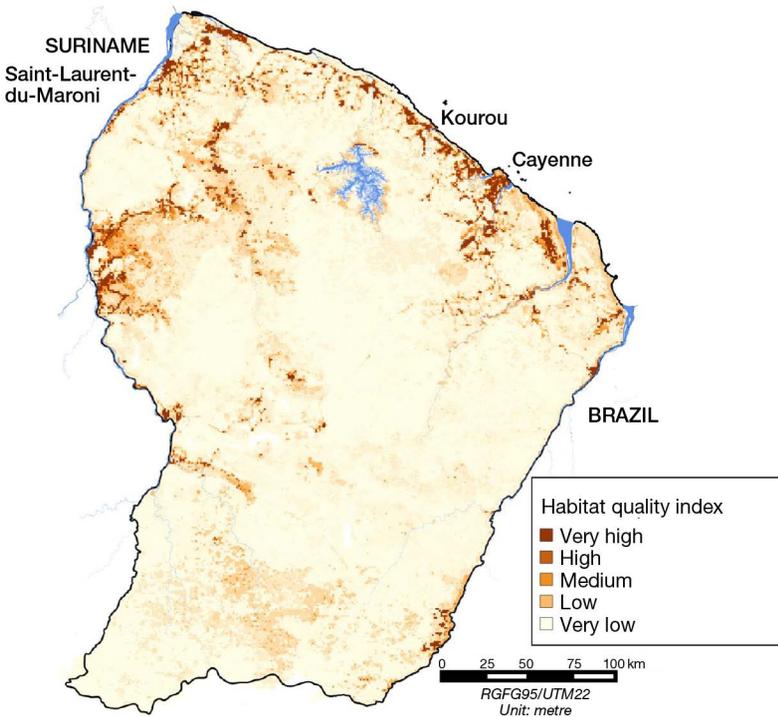


Figure 6.2. *Anopheles darlingi* habitat quality in French Guiana.

42. <https://www.insee.fr/fr/statistiques/6012651>

Human activities (HUMACT variable) are linked to habitats that are particularly conducive to the presence of *An. darlingi* in the central region of the territory. These habitats are primarily associated with mining activity. In 2014, the number of illegal gold mining operations in French Guiana was estimated to be around 700, with approximately 10,000 to 15,000 workers involved (Douine *et al.*, 2016). This contributes to the maintenance of malaria in the region through a variety of mechanisms. Gold mining activities encourage the proliferation of *An. darlingi* by creating gaps in the dense forest and semi-shaded pools of water which serve as potential larval habitats. These gold mining sites are frequented by a considerable number of people, who are highly exposed to mosquito bites due to their outdoor activities and poor living conditions, leading to situations where there is a high risk of local malaria transmission. Furthermore, this population is highly mobile and receives little medical care, strongly contributing to the circulation of the parasite within populations of the region, encouraging the emergence and maintenance of transmission foci, promoting antimalarial drug resistance, etc. (Thoisy de *et al.*, 2021; Douine *et al.*, 2021).

It should also be noted that *An. darlingi* is not the only malaria vector present at gold mining sites. In particular, *An. marajoara* appears to significantly contribute to transmission in highly anthropised forest environments such as mining sites (Pommier de Santi *et al.*, 2016). All of the above make gold mining, and particularly illegal gold mining, one of the main factors in allowing the autochthonous transmission of malaria to persist, representing a significant obstacle to its eradication in the region.

Certain land cover categories are associated with high habitat quality. This is the case of “wooded savannah / dry forest”, which correspond to areas that are dry but regularly flooded, and are thus associated with the presence of potential larval habitats (Gond *et al.*, 2011; Rosa-Freitas *et al.*, 2007). The results are consistent with those of previous studies conducted along the coastal savannahs (Dusfour *et al.*, 2013; Vezenegho *et al.*, 2015), which confirmed the presence of *An. darlingi* in this type of environment, at times in significant numbers.

► Conclusion

This chapter has demonstrated that it is feasible to construct models based on presence-only species data and the environmental conditions in the observation sites. These models facilitate the prediction of habitat quality and, by extension, the probability of species presence throughout the entire area of study. Among these models, Maxent is one of the most popular, particularly due to its efficacy in the event of limited datasets on species presence. However, as is the case with all models, it is susceptible to sampling bias, particularly when the number of observation sites is limited. However, through the implementation of diverse approaches, including those proposed by the authors of this chapter, the impact of these biases can be minimised. The current trend in species distribution modelling is to use ensemble models, which consists of creating several species distribution models and combining the outputs to achieve the best results. An R library has been developed for this type of approach: biomod2 (Thuiller *et al.*, 2022). However, any modelling approach requires an understanding of the parameters of the model and the capacity to adapt them to the context and objectives of the study, in order to obtain reliable and pertinent results. The use of multiple models may require considerable effort.

In summary, mapping the spatial distribution of vector species is intended to guide vector control actions. However, it remains underutilised in this field, largely due to the absence of sufficiently fine-scale spatial resolution data and the inability to predict vector abundance. Proposed solutions to these shortcomings include the use of predictive, data-driven models on mosquito abundance. For example, Tran *et al.* (2020) present a model based on the support vector machine (SVM) approach using data on *Aedes* mosquito larval density in different municipalities on Reunion island. This model produced promising results, which can supplement those achieved by mechanistic modelling (see Chapter 8). Adde *et al.* (2016) successfully modelled the dynamics of *An. darlingi* density using mixed effects ordinal logistic regression in a municipality of French Guiana. However, such models are difficult to apply to large areas. They are particularly data intensive, and the collection of mosquito abundance data is especially challenging and costly to obtain (in terms of time, personnel, and financial resources), meaning it is rarely available. Furthermore, regularly trapping mosquitoes across vast territories is highly impractical (it should be noted, however, that mosquito trapping methods are becoming increasingly diverse, more efficient and less costly, which could lead to more robust and generalised models being developed). Lastly, one of reasons for the low adoption of species distribution modelling for malaria vector control is that vector distribution cannot be directly translated to transmission risk, let alone epidemic risk. The determination of these risks is contingent upon a number of additional factors pertaining to the vector (population abundance and age), as well as human presence and activities, the state of the health system, and individual human factors (behaviour, genetic predisposition, immunity, etc.).

In order to ensure the effective implementation of species distribution modelling in vector control and, more generally, to create action plans for vector-borne diseases, it is essential to provide a more detailed description of the advantages associated with each modelling approach. This enables a transition towards more integrated and systemic approaches, which allow for the consideration of different risk components through the application of a formalistic approach and the utilisation of various modelling methods (knowledge-based, mechanistic, data-driven) and their respective advantages, contingent on the data and knowledge available.

►► References

- Adde A., Roux E., Mangeas M., Dessay N., Nacher M., Dusfour I., Girod R., Briolan S., 2016. Dynamical Mapping of *Anopheles darlingi* Densities in a Residual Malaria Transmission Area of French Guiana by Using Remote Sensing and Meteorological Data. *PLoS One*, 11 (10), e0164685.
- Douine M., Musset L., Corlin F., Pelleau S., Pasquier J., Mutricy L., Adenis A., Djossou F., Brousse P., Perotti F., Hiwat H., Vreden S., Demar M., Nacher M., 2016. Prevalence of Plasmodium spp. in illegal gold miners in French Guiana in 2015: a hidden but critical malaria reservoir. *Malaria Journal*, 15, 315.
- Douine M., Lambert Y., Galindo M.S., Mutricy L., Sanna A., Peterka C., Marchesini P., Hiwat H., Nacher M., Adenis A., Demar M., Musset L., Lazrek Y., Cairo H., Bordalo Miller J., Vreden S., Suarez-Mutis M., 2021. Self-diagnosis and self-treatment of malaria in hard-to-reach and mobile populations of the Amazon: results of Malakit, an international multicentric intervention research project. *The Lancet Regional Health – Americas*, 4, 100047.
- Dusfour I., Carinci R., Issaly J., Gaborit P., Girod R., 2013. A survey of adult anophelines in French Guiana: enhanced descriptions of species distribution and biting responses. *Journal of Vector Ecology*, 38 (2), 203-209.

- Elith J., Phillips S.J., Hastie T., Dudík M., Chee Y.E., Yates C.J., 2011. A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions*, 17 (1), 43-57.
- Gond V., Freycon V., Molino J.-F., Brunaux O., Ingrassia E., Joubert P., Pekel J.-F., Prévost M.-F., Thierron V., Trombe P.-J., Sabatier D., 2011. Broad-scale spatial pattern of forest landscape types in the Guiana Shield. *International Journal of Applied Earth Observation and Geoinformation*, 13 (3), 357-367.
- Guisan A., Thuiller W., 2005. Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, 8 (9), 993-1009.
- Guitet S., Cornu J.-F., Brunaux O., Betbeder J., Carozza J.-M., Richard-Hansen C., 2013. Landform and landscape mapping, French Guiana (South America). *Journal of Maps*, 9 (3), 325-335.
- Hernandez P.A., Graham C.H., Master L.L., Albert D.L., 2006. The effect of sample size and species characteristics on performance of different species distribution modeling methods. *Ecography*, 29 (5), 773-785.
- Hijmans R., Phillips S., Leathwick J., Elith J., 2017. Dismo: Species Distribution Modeling. R package version 1.1-4.
- Hiwat H., Bretas G., 2011. Ecology of *Anopheles darlingi* Root with respect to vector importance: a review. *Parasit Vectors*, 4, 177.
- Hiwat H., Issaly J., Gaborit P., Somai A., Samjhawan A., Sardjoe P., Soekhoe T., Girod R., 2010. Behavioral heterogeneity of *Anopheles darlingi* (Diptera: Culicidae) and malaria transmission dynamics along the Maroni River, Suriname, French Guiana. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 104 (3), 207-213.
- Mosnier E., Dusfour I., Lacour G., Saldanha R., Guidez A., Gomes M.S., Sanna A., Epelboin Y., Restrepo J., Davy D., Demar M., Djossou F., Douine M., Ardillon V., Nacher M., Musset L., Roux E., 2020. Resurgence risk for malaria, and the characterization of a recent outbreak in an Amazonian border area between French Guiana and Brazil. *BMC Infectious Diseases*, 20 (1), 373.
- Moua Y., Roux E., Seyler F., Briolant S., 2020. Correcting the effect of sampling bias in species distribution modeling – A new method in the case of a low number of presence data. *Ecological Informatics*, 57, 101086.
- Moua Y., Roux E., Girod R., Dusfour I., Thoisy B. de, Seyler F., Briolant S., 2017. Distribution of the Habitat Suitability of the Main Malaria Vector in French Guiana Using Maximum Entropy Modeling. *Journal of Medical Entomology*, 54 (3), 606-621.
- Moua Y., Roux E., Dusfour I., Girod R., Chantilly S., Seyler F., Gomes M., Zanini V., Galardo A., Lima A., 2019. A database of *Anopheles* (Diptera, Culicidae): historical presence records in French Guiana and in the state of Amapá, in South America. Version 1.3, Institut Pasteur de la Guyane, *Occurrence dataset*.
- Pommier de Santi V., Girod R., Mura M., Dia A., Briolant S., Djossou F., Dusfour I., Mendibil A., Simon F., Deparis X., Pages F., 2016. Epidemiological and entomological studies of a malaria outbreak among French armed forces deployed at illegal gold mining sites reveal new aspects of the disease's transmission in French Guiana. *Malaria Journal*, 15, 35.
- Rosa-Freitas M., Tsouris P., Peterson A., Honório N., De Barros F., Aguiar D. de, Gurgel H., Arruda M. de, Vasconcelos S., Luitgards-Moura J., 2007. An ecoregional classification for the state of Roraima, Brazil : the importance of landscape in malaria biology. *Memórias do Instituto Oswaldo Cruz*, 102 (3), 349-358.
- Rozendaal J., 1987. Observations on the biology and behaviour of Anophelines in the Suriname rainforest with special reference to *Anopheles darlingi* Root. *Cahiers ORSTOM Série Entomologie Médicale et Parasitologie*, 25 (1), 33-43.
- Singer B.H., Castro M.C. de, 2001. Agricultural colonization and malaria on the Amazon frontier. *Annals of the New York Academy of Sciences*, 954, 184-222.
- Sinka M.E., Rubio-Palis Y., Manguin S., Patil A.P., Temperley W.H., Gething P.W., Van Boeckel T., Kabaria C.W., Harbach R.E., Hay S.I., 2010. The dominant *Anopheles* vectors of human malaria in the Americas: occurrence data, distribution maps and bionomic précis. *Parasites & Vectors*, 3 (1), 72.

Sinka M.E., Bangs M.J., Manguin S., Rubio-Palis Y., Chareonviriyaphap T., Coetzee M., Mbogo C.M., Hemingway J., Patil A.P., Temperley W.H., Gething P.W., Kabaria C.W., Burkot T.R., Harbach R.E., Hay S.I., 2012. A global map of dominant malaria vectors. *Parasites & Vectors*, 5 (1), 69.

Stefani A., Dusfour I., Correa A.P., Cruz M.C., Dessay N., Galardo A.K., Galardo C.D., Girod R., Gomes M.S., Gurgel H., Lima A.C., Moreno E.S., Musset L., Nacher M., Soares A.C., Carne B., Roux E., 2013. Land cover, land use and malaria in the Amazon: a systematic literature review of studies using remotely sensed data. *Malaria Journal*, 12, 192.

Thoisy B. de, Richard-Hansen C., Goguillon B., Joubert P., Obstancias J., Winterton P., Brosse S., 2010. Rapid evaluation of threats to biodiversity: human footprint score and large vertebrate species responses in French Guiana. *Biodiversity and Conservation*, 19 (6), 1567-1584.

Thoisy B. de, Duron O., Epelboin L., Musset L., Quenel P., Roche B., Binetruy F., Briolant S., Carvalho L., Chavy A., Couppie P., Demar M., Douine M., Dusfour I., Epelboin Y., Flamand C., Franc A., Ginouves M., Gourbiere S., Houel E., Kocher A., Lavergne A., Le Turnier P., Mathieu L., Murienne J., Nacher M., Pelleau S., Prevot G., Rousset D., Roux E., Schaub R., Talaga S., Thill P., Tirera S., Guegan J.-F., 2021. Ecology, evolution, and epidemiology of zoonotic and vector-borne infectious diseases in French Guiana: Transdisciplinarity does matter to tackle new emerging threats. *Infection, Genetics and Evolution, Journal of Molecular Epidemiology and Evolutionary Genetics of Infectious Diseases*, 93, 104 916.

Thuiller W., Georges D., Gueguen M., Engler R., Breiner F., Lafourcade B., 2022. biomod2: Ensemble Platform for Species Distribution Modeling. R package version 4.0.

Tran A., Mangeas M., Demarchi M., Roux E., Degenne P., Haramboure M., Le Goff G., Damiens D., Gouagna L.C., Herbreteau V., Dehecq J.-S., 2020. Complementarity of empirical and process-based approaches to modelling mosquito population dynamics with *Aedes albopictus* as an example – Application to the development of an operational mapping tool of vector populations. *PLoS One*, 15 (1), e0227407.

Vezenegho S.B., Carinci R., Gaborit P., Issaly J., Dusfour I., Briolant S., Girod R., 2015. *Anopheles darlingi* (Diptera: Culicidae) Dynamics in Relation to Meteorological Data in a Cattle Farm Located in the Coastal Region of French Guiana: Advantage of Mosquito Magnet Trap. *Environmental Entomology*, 44 (3), 454-462.

WHO, 2021. *World Malaria Report 2021*. Geneva, World Health Organization, 263 p. <https://www.who.int/teams/global-malaria-programme/reports/world-malaria-report-2021>.

Chapter 7

Knowledge-based models: example of a multi-criteria evaluation tool for public health

Fanjasoa Rakotomanana, Hobiniaina Anthonio Rakotoarison

In order to minimise the impact of a disease on individuals, it is essential that cases are managed effectively. However, the actions of public health stakeholders in both human and animal health, which aim to prevent the risk of disease spread, are crucial in reducing its impact at the population level. Infectious diseases affecting humans and animals emerge as a consequence of exposure to infectious agents present in their immediate or broader environment. These infectious agents, as well as their arthropod vectors (organisms that transmit the infectious agent), are themselves under the influence of environmental determinants whose parameters vary depending on the physical, chemical and climate conditions, in addition to the anthropogenic pressure exerted on their natural habitat. It is therefore essential to understand these factors in the epidemiology of such diseases and to monitor their evolution in order to detect environmental changes and assess the risks of an epidemic occurring.

In the event of an epidemic or pandemic, particularly when an unknown pathogen emerges, the use of models has become standard practice in public health. Descriptive models of a health phenomenon are commonplace, particularly as they allow the prediction of disease incidence and prevalence. However, these models can also be used to understand transmission mechanisms, or to predict the scale of the epidemic or the impact of control measures (Dubois, 2005). Models provide information that helps estimate of the technical and financial resources required to address a potentially critical public health situation. They play an essential role in decision-making. It is also crucial to incorporate a spatial dimension into models to gain insight into the initial stages of an epidemic. The oldest thematic distribution map of cholera cases in London in 1854 demonstrated that integrating spatial data in an explanatory epidemiological approach facilitated the understanding of disease transmission mechanism and enable the epidemic to be controlled (Snow, 1855).

A large number of models are based on data analysis (see Chapter 6). Furthermore, initiatives and data management and collection platforms are in place at the global or continental level, with the creation of reporting systems for epidemic-prone diseases. However, several constraints affect the acquisition of comprehensive and reliable data, both in terms of spatial and temporal coverage, which are essential for

fully addressing the issue. For example, since 2019, Madagascar's Ministry of Public Health has relied on the DHIS2⁴³ platform, a health management information system (HMIS) used by more than 100 countries worldwide. In 2021, the promptness and completeness of data reporting remained below 80% for the malaria-specific integrated national surveillance system, which represents a significant weakness in the management and response to malaria epidemics (NMCP, 2021). Additionally, data on environmental risk factors is often lacking in countries with limited resources, such as Madagascar. This is also the case at the global level during the emergence of new pathogens. This data is vital for implementing strategies to control epidemic-prone diseases such as malaria (WHO, 2001).

This chapter addresses knowledge-based models, which can be employed in the absence of epidemiological data, with geographic information system (GIS)-based multi-criteria analysis given as an example. This method helps overcome the need for large epidemiological datasets required to develop data-driven models (see Chapter 6). This chapters provides an example of a knowledge-based model used to identify areas, which are priorities for the malaria vector control campaign in Madagascar. Finally, this chapter discusses Full MCE for Public Health, a tool developed at the Institut Pasteur de Madagascar (IPM), which enables this method to be implemented via a plug-in for the GIS software QGIS.

► GIS-based multi-criteria analysis, a knowledge-based approach

GIS software is designed to manage multi-source environmental data and provide a decision-making tool. The use of multi-criteria analysis (MCA) in conjunction with a GIS is becoming increasingly prevalent across a range of disciplines. It is based on expert knowledge. Several examples have shown that MCA combined with GIS, also known as spatialised multi-criteria analysis (SMCA), is an highly effective tool for addressing issues in human and animal health. For example, Bell *et al.* (2007) developed deprivation indices, which are indicators related to the health of populations, providing socio-economic gradients to decision-makers based on a SMCA approach. In the absence of precise data on the foci of African swine fever (ASF), SMCA allowed ASF transmission risks to be mapped in Africa and Madagascar, compensating for the lack of quantitative data on the epidemiology of the disease in Africa (Glanville de *et al.*, 2014). Knowledge-based models have also been used to map areas conducive to the amplification and spread of Rift Valley fever (RVF) in four East African countries: Kenya, Tanzania, Uganda and Ethiopia (Tran *et al.*, 2016).

SMCA consists of five main steps, which are considered the building blocks of spatial decision-making:

1. Defining the risk to be mapped (e.g., risk of introduction, amplification, spread, etc.) and the area concerned, depending on the objective, which requires making one or more decisions. A comprehensive understanding of the risks allows the identification of factors to be considered in the subsequent phase.

43. <https://dhis2.org/fr/about/>

2. Identifying the criteria for analysis: the data to be used as factors will be identified with the help of expert opinions and collected from various sources (national geographic institutes or other geoportals). The criteria can be divided into two types: constraints (a mask limiting the areas to be considered) and factors (criteria defining a certain degree of suitability for a given area) [Rakotoarison *et al.*, 2020; Rakotomanana *et al.*, 2007]. The factors (e.g., temperature, which influences mosquito population dynamics and the biological life cycle of plasmodium in mosquitoes) and constraints (e.g., inhabited areas, target areas for vector control intervention: there is no point in predicting risk outside of these areas) are then processed and combined in a different way. Experts such as epidemiologists, entomologists, climatologists, ecologists or specialist in human or animal health are involved in this stage to help identify the criteria. Their opinions are supported by literature reviews resulting from prior bibliographic research. Remote sensing data is a particularly useful source for acquiring spatially explicit data on climate and environmental factors or constraints. The use of satellite data to understand the various types of land cover and their change over time has become the most valuable GIS application for obtaining the up-to-date data on a study area (see Chapter 1).

3. Standardising factors: this consists of transforming the original evaluation factors into comparable units in order to obtain suitability values for an event (e.g., suitability of the occurrence of a disease) expressed on a continuous scale (depending on the choice from 0 to 1, or 0 to 255). A value of zero corresponds to no suitability, whereas a value of 1 or 255 (depending on the choice of scale) represents maximum suitability for the event being studied. Standardisation can be done using thresholds (above or below which, the risk is present or absent) or based on fuzzy logic, the most common approach used in multi-criteria decision analysis. This technique allows for a gradual transition between factors suitable for risk prediction and those unsuitable, depending on the relationship between the factor and the risk (increasing, decreasing, symmetrical, etc.). Figure 7.1 shows two examples of relational functions for the factors and risks to be evaluated.

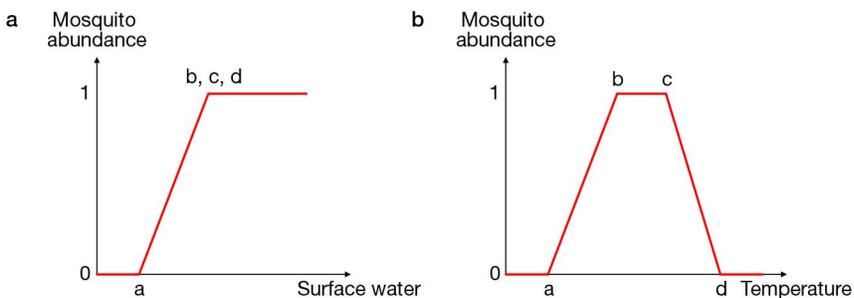


Figure 7.1. Examples of membership functions for standardisation of factors: (a) increasing function; (b) symmetrical function with plateau.

4. Factor weighting: this step consists of assigning a weight (between 0 and 1) to each factor according to its importance relative to the other factors considered. The closer the weight is to 1, the greater the contribution of the factor to the event. Several methods, including pairwise comparison, have been proposed to weight factors

depending on the context. The pairwise comparison method was developed by Saaty as part of the analytic hierarchy process (AHP) [Saaty, 1977]. This method is based on the creation of a matrix wherein the factors are evaluated in pairs through the assignment of scores based on a comparative scale (Table 7.1) Using this matrix, the AHP method allows a weight to be calculated for each factor. However, pairwise comparison may result in inconsistent relationships. These inconsistencies can be measured by calculating a consistency ratio (CR). Pairwise comparisons with $CR < 0.1$ are considered consistent, while a $CR \geq 0.1$ indicates inconsistent judgements, which should be re-evaluated.

Table 7.1. Example of a comparison scale for two risk factors.

Importance of factor A in relation to factor B								
Extreme	Very high	High	Moderate	Equal importance	Moderate	High	Very high	Extreme
Less important					More important			
1/9	1/7	1/5	1/3	1	3	5	7	9

5. The final step involves combining risk indices (standardised and weighted factors) and integrating them into a GIS in order to produce a risk map. The resulting map indicates the level of risk per basic geographic unit. If a continuous scale is used, ranging from 0 to 255, a value of 0 indicates a low risk, whereas a value of 255 indicates a high risk. The varying weights assigned according to the opinions and estimates of experts allow uncertainty to be predicted at the level of the geographic unit in question. An uncertainty map may be combined with a risk map. The model is considered stable if it does not vary significantly, regardless of the variation in weights assigned to the factors by different experts. The risk map can be validated depending on the availability of epidemiological data.

►► Example of this method used to map the risk of malaria in the Malagasy Central Highlands

Context

Malaria control remains a global challenge in public health, with significant economic implications. The 2030 target for malaria control is to reduce the incidence and mortality rates of malaria by at least 90% globally compared to 2015 levels. In 2019, there were an estimated 227 million malaria cases across 85 malaria-endemic countries; this figure increased by 14 million in 2020 to reach 241 million (WHO, 2021). For countries with limited resources, combating this scourge primarily relies on support from financial partners through the global fund to fight major epidemics, such as malaria, AIDS and tuberculosis (NMCP, 2017; WHO/GMP, 2021). In Madagascar, malaria ranks as the fourth leading cause of morbidity in health centres and in-hospital mortality (NMCP, 2017). Malaria is characterised depending by four distinct patterns, which vary depending on the duration and intensity of transmission: East, West, Semi-arid and Malagasy Central Highlands (MCH).

Application of the SMCA method to the *Full MCE for Public Health* tool

Rakotoarison *et al.* (2020) applied the SMCA method to the Malagasy Central Highlands to identify areas, which are priorities for indoor residual spraying (IRS) campaigns. The model was developed with the Full MCE for Public Health tool.

For users who are not specialists in geomatics, the IPM created a plugin for the open-source software QGIS, providing stakeholders in the health sector with an interactive and semi-automated tool. The Full MCE for Public Health plugin was developed using Python 2 for QGIS version 2.x and features a user-friendly interface⁴⁴ (Figure 7.2). Training was provided to human and animal health workers by the IPM team at both the national and international levels, aimed at improving their proficiency with the tool.

The steps of the SMCA method described in the previous paragraph are outlined below, along with illustrations of the Full MCE for Public Health interface.

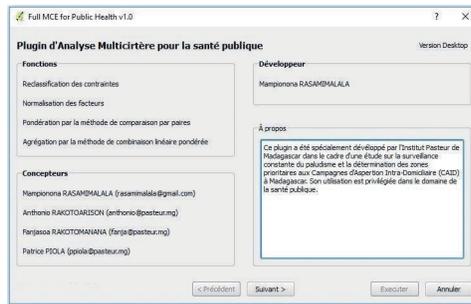


Figure 7.2. Welcome screen of the Full MCE for Public Health plugin in QGIS 2.x.

1. Definition of risks and priority areas

Firstly, the MCHs were identified as a priority area, as this region is prone to epidemics of unstable malaria. Malaria vector control is based on IRS campaigns. Areas of intervention are selected according to the available means of control. By mapping the areas at the highest risk, it is possible to target them with greater precision.

2. Identification of criteria

In collaboration with medical entomology specialists, stakeholders in the national malaria control programme (NMCP), and other local experts, various criteria influencing the risk of malaria transmission in the MCH were identified. This step draws on knowledge of malaria in general, and the MCH in particular.

Chapter 1 discussed the role of the environment in the transmission of mosquito-borne disease and the benefits of using remote sensing. In this case of malaria in the MCH, rice fields represent the larval habitats of choice for *Anopheles funestus* mosquitoes, which are the primary vectors in this region. For various reasons (changes in rainfall, potential for irrigation, insecurity), crop cover may vary from one year to the next. Between crop cycles, a rice field may be drained or remain underwater depending on the irrigation system. Occasionally, these agricultural plots are used for an entirely

44. <https://github.com/SaGEOTeam/FullMCE>

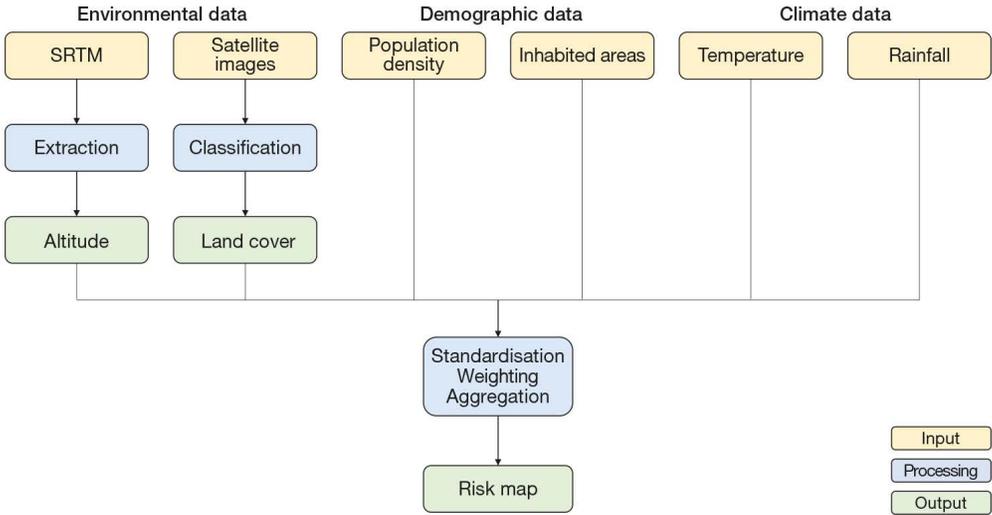


Figure 7.3. Input data and data flow diagram to obtain malaria risk maps for the Malagasy Central Highlands. Adapted from Rakotoarison *et al.* (2020).

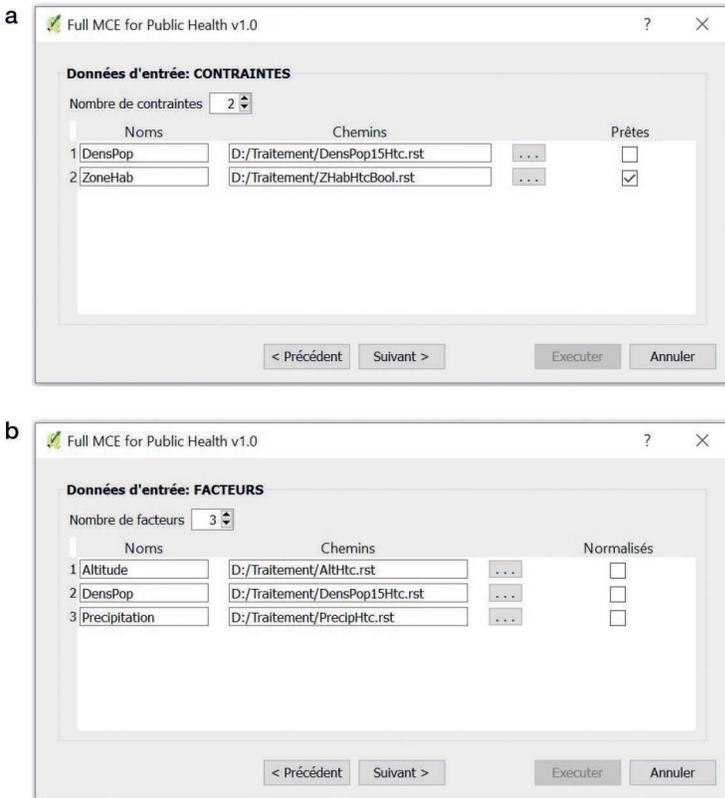


Figure 7.4. Illustration of the Full MCE for Public Health plugin showing the integration of (a) constraints, and (b) factors to be considered in the GIS-based multi-criteria analysis.

different purpose, with the topsoil being harvested for the manufacture of bricks. The water-filled holes left behind after these anthropic activities may also serve as ideal larval habitats for malaria vector mosquitoes. Furthermore, a reduction in rice crop cover encourages the proliferation of other vector species, such as *An. gambiae*, in natural environments. Wetlands represent potential larval habitats for anopheles malaria vectors, and thus one of the main criteria chosen for SMCA. Remote sensing (Landsat images) allows for the annual mapping of rice fields and natural wetlands (Rakotoarison *et al.*, 2020).

Altitude, temperature, precipitation, human population density (which can affect larval habitat distribution), and inhabited areas (targets of IRS) were also selected as criteria (Figure 7.3) (Rakotoarison *et al.*, 2020).

Pre-processing (e.g., satellite image classification) is required to prepare the corresponding data for analysis (Figure 7.3). The interface of the Full MCE for Public Health plugin allows these various criteria to be integrated as either constraints or factors, in raster format (Figure 7.4).

3. Standardisation

The plugin's interface then allows the factors to be standardised. In accordance with expert opinions, the parameters can be defined for the type of standardisation functions to be used (linear, sigmoid, etc.), their rate of change (increasing, decreasing, symmetrical), and the thresholds to be applied (Figure 7.5a).

4. Factor weighting

In the weighting step, the user compares different factors in pairs. Values for the pairwise comparison matrix (Table 7.1) can be easily entered via the interface (Figure 7.5b). The weight assigned to each factor is automatically calculated using the AHP method and a summary is provided on the final screen (Figure 7.6).

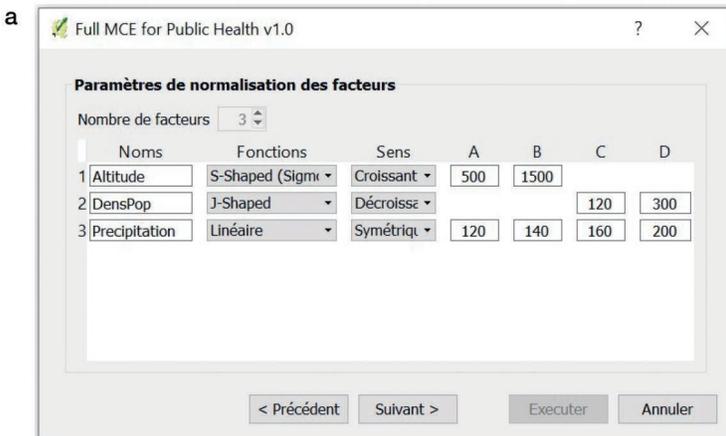


Figure 7.5a. Illustration of the Full MCE for Public Health plugin showing the steps for standardisation.

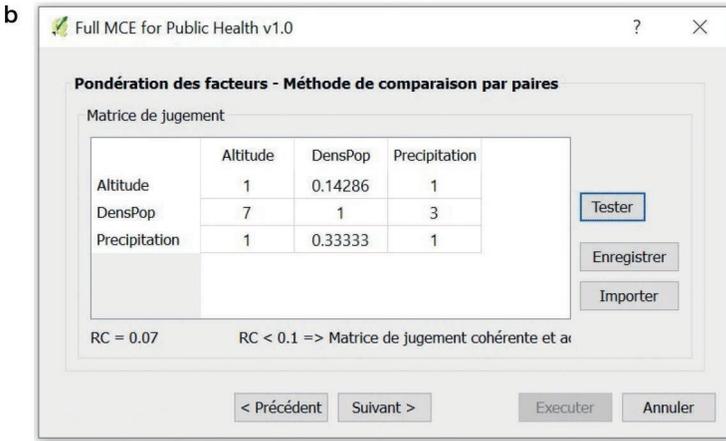


Figure 7.5b. Illustration of the Full MCE for Public Health plugin showing the steps for generation of the pairwise comparison matrix for weighting factors.

In the study conducted by Rakotoarison *et al.* (2020), population density was considered the most important factor by the experts, with the highest weight assigned by the AHP method (weight: 0.50). This was followed by the distance to rice fields and wetlands (as the distance increases, the risk decreases; weight: 0.18), then temperature (weight: 0.17), altitude (weight: 0.09) and precipitation (weight: 0.06).

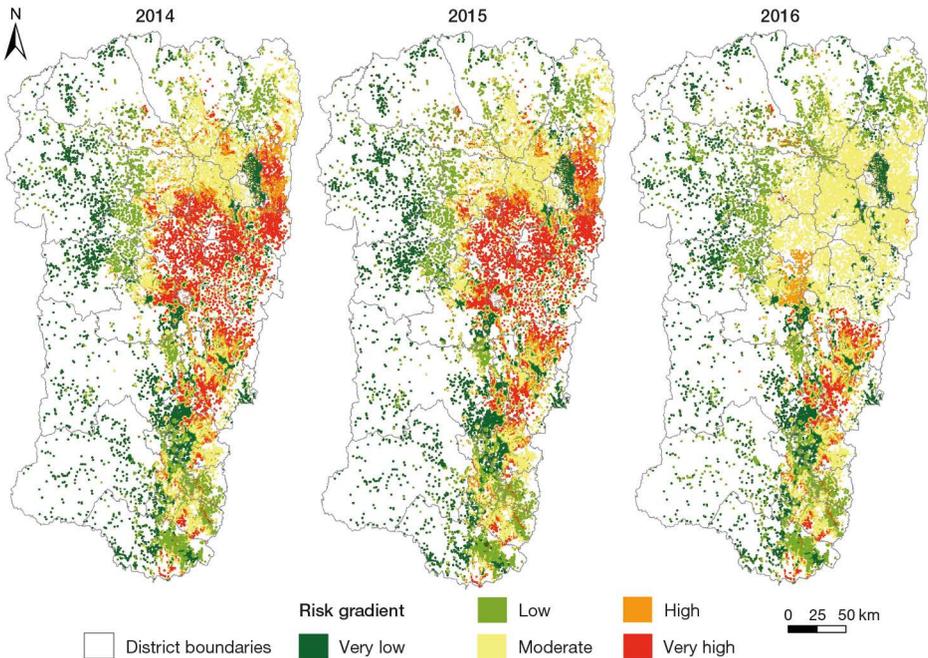


Figure 7.6. Malaria risk maps in the Malagasy Central Highlands Adapted from Rakotoarison *et al.* (2020).

5. Combination of risk indices

Finally, the standardised and weighted factors (population density, distance to wetlands, temperature, altitude, precipitation), as well as the constraints (inhabited areas, altitude, population density), are combined to obtain risk maps for IRS prioritisation for different years (2014, 2015 and 2016) by updating land cover, temperature and precipitation maps on an annual basis. The identified areas of risk were classified according to a gradient ranging from very low to very high, highlighting the significant spatial and interannual variations in transmission risk (Figure 7.6).

►► Conclusion

The influence of geographic location and the role of the environment in disease distribution are long-established concepts in the field of health. The use of satellite data to gain insight into the different types of land cover and their change over time is essential for obtaining up-to-date data and updating risk maps. Knowledge-based models promote a collaborative approach, whereby experts are involved in evaluating potential risks. The tool was developed based on this method and is freely available. It can be used in other fields, such as agriculture, environmental protection, or land use planning.

►► References

- Bell N., Schuurman N., Hayes M.V., 2007. Using GIS-based methods of multicriteria analysis to construct socio-economic deprivation indices. *International Journal of Health Geographics*, 6, 17.
- Dubois M., 2005. Modélisation en épidémiologie : objectifs et méthodes. *Epidémiologie et santé animale*, 47, 1-13.
- Glanville W.A de., Vial L., Costard S., Wieland B., Pfeiffer D.U., 2014. Spatial multi-criteria decision analysis to predict suitability for African swine fever endemicity in Africa. *BMC Veterinary Research*, 10, 9.
- PNLP, 2017. *Plan stratégique nationale de lutte contre le paludisme 2018-2022*. Programme national de lutte contre le paludisme, Antananarivo, Madagascar.
- PNLP, 2021. *Plan de réponse aux épidémies et aux recrudescences de paludisme à Madagascar*. Programme national de lutte contre le paludisme, Antananarivo, Madagascar.
- Rakotoarison H.A., Rasamimalala M., Rakotondramanga J.M., Ramiranirina B., Franchard T., Kapesa L., Razafindrakoto J., Guis H., Tantely L.M., Girod R., Rakotoniaina S., Baril L., Piola P., Rakotomanana F., 2020. Remote Sensing and Multi-Criteria Evaluation for Malaria Risk Mapping to Support Indoor Residual Spraying Prioritization in the Central Highlands of Madagascar. *Remote Sensing*, 12 (10), 1585.
- Rakotomanana F., Randremanana R.V., Rabarijaona L.P., Duchemin J.-B., Ratovonjato J., Arieux F., Rudant J.-P., Jeanne I., 2007. Determining areas that require indoor insecticide spraying using Multi Criteria Evaluation, a decision-support tool for malaria vector control programmes in the Central Highlands of Madagascar. *International Journal of Health Geographics*, 6, 2.
- Saaty T., 1977. A scaling method for priorities in hierarchical structures. *Journal of Mathematical Psychology*, 15, 234–281.
- Snow J., 1855. *On the Mode of Communication of Cholera*, London, John Churchill.
- Tran A., Trevennec C., Lutwama J., Sserugga J., Gely M., Pittiglio C., Pinto J., Chevalier V., 2016. Development and Assessment of a Geographic Knowledge-Based Model for Mapping Suitable Areas for Rift Valley Fever Transmission in Eastern Africa. *PLOS Neglected Tropical Diseases*, 10 (9), e0004999.
- WHO, 2001. *Le Paludisme, Systemes de Pré-Alerte - Cadre pour la recherche de terrain en Afrique*, Genève, Faire reculer le paludisme – Organisation mondiale de la Santé, 88 p. https://apps.who.int/iris/bitstream/handle/10665/67392/WHO_CDS_RBM_2001.32_fre.pdf?sequence=1&isAllowed=y.
- WHO, 2021. *World Malaria Report 2021*. Geneva, World Health Organization, 263 p. <https://www.who.int/teams/global-malaria-programme/reports/world-malaria-report-2021>.
- WHO/GMP, 2021. *2021 African Union Malaria Progress Report*.

Chapter 8

Arbocarto: a mechanistic model based on the life cycle of *Aedes* mosquitoes

Renaud Marti, Marie Demarchi, Mathieu Castets, Annelise Tran

Due to the nature of their design, process-based models, also termed “mechanistic models”, are explanatory models focused on the causality of relationships between inputs and outputs (Craver, 2006). This type of approach requires the explicit delineation of causal relationships (e.g., the effect of temperature on mosquito development) in a modelling framework (see Introduction to Part 2), usually in mathematical form (e.g., an equation expressing development rate as a function of temperature), based on previously established knowledge (from observational or experimental studies) about the system being modelled.

Depending on the goals of the model, this approach also requires a mandatory and sometimes difficult step of simplification. This step, undertaken at the discretion of the modellers and based on their understanding of the significance of the processes to be considered, enables them to develop parcimonious models. In addition, an explanation of the processes through a mechanistic approach enables an *in silico* simulation of scenarios (e.g., impact of insecticide treatment using different protocols to control mosquito populations) and, through their analysis, identification of the control points for the system in question. Experimental studies would be comparatively difficult and costly to perform.

Since the advent of big data, the interest of such approaches combined with the understanding, exploitation and validation of explanatory mechanisms is at times questionable, especially when considering the hegemony of data-driven approaches (see Chapter 2), and the predictive and sometimes spectacular power of machine learning methods (Baker *et al.*, 2018). However, both of these approaches provide complementary information (Tran *et al.*, 2020), and their integration can yield valuable insights (Baker *et al.*, 2018).

In this chapter, we present an example of a mechanistic model transformed into a software tool in the context of entomological surveillance and vector control against *Aedes* mosquitoes, based on an explicit description of the mosquito life cycle. This tool was developed in the framework of the Arbocarto⁴⁵ project, which looks at how to consolidate the surveillance and control strategies used by vector control services in different

45. <https://www.arbocarto.fr/en>

regions of France. The model bearing the same name is a predictive mapping tool for population densities of *Aedes* mosquitoes, designed with a spatial scale suitable for use by surveillance and control organisations, and based on weather data (daily rainfall and temperatures) and high to very high spatial resolution remote sensing products.

► A generic model built around the mosquito life cycle

The life cycle of mosquitoes is uniform across species, commencing with an immature aquatic phase, which is subdivided into three principal developmental stages: egg, larva, and pupa. A rapid emergence phase ending in the adult form marks the transition to the aerial stage, during which mosquitoes will alternate between reproductive, feeding, resting and dispersal behaviours until death (see Chapter 1). This development, which takes the form of a complete metamorphosis (holometabolous insect), allows for the biological cycle of mosquitoes to be interpreted as a compartmental model (Cailly *et al.*, 2012). Each stage of the cycle is assigned a “box” (or “compartment”) into which all of the mosquitoes sharing the same morphological state and similar behaviours are placed (e.g., “Pupae”). The boxes are connected by arrows which mark the transitions between different stages (Figure 8.1).

The demographic changes associated with each compartment can be described by a system of ordinary differential equations (ODE). In the generic model of mosquito population developed by Cailly *et al.* (2012), an ODE system then allows each of the 10 population stages to be expressed over time, in the form of the same number of basic demographic groups: 3 aquatic stages (*E* eggs, *L* larvae, *P* pupae), 1 emerging adult stage (*Aem*), 3 nulliparous stages (*A1h*, *A1g*, *A1o*) and 3 parous stages (*A2h*, *A2g*, *A2o*), with adult females subdivided according to their behaviour during the gonotrophic cycle (*h*: host-seeking, *g*: transition from engorged to gravid, *o*: oviposition site seeking). The variation in the number of mosquitoes in a given stage (e.g., “Larvae”) over a defined period of time (1 day in this example) is then expressed by adding the fraction of the population originating from the previous compartment (e.g., “Eggs”) and subtracting the fraction of the population progressing to the next compartment (e.g., “Pupae”) and assigning a mortality rate (Figure 8.1). To quantify the flow of individuals from one compartment to the next, this model uses parameters which are considered to be stable over time (e.g., the number of eggs laid) and functions which vary according to meteorological conditions such as rainfall or temperature (e.g., mortality rates or transition rates to the next stage) [Cailly *et al.*, 2012]. This approach (Figure 8.1) permits the aggregation of data regarding the influence of meteorological and environmental variables on mosquito population dynamics, facilitating the adaptation of a generic model to a specific species or geographic area (Ezanno *et al.*, 2015).

As previously outlined, certain simplifications have been incorporated into this model, with the objective of prioritising the inclusion of specific biological mechanisms over others that are considered to be of secondary importance. Due to their nuisance capacity and/or their epidemiological significance in relation to blood-feeding (bites), this model prioritises the representation of the abundance of adult (female) mosquitoes capable of taking a blood meal (host-seeking female adults, sum of *A1h* and *A2h*) and transmitting a pathogen after a first infected bite (*A2h*). Furthermore, the principal stages of the mosquito life cycle were retained to facilitate simulations of the impact of a control measure on a specific stage of the cycle.

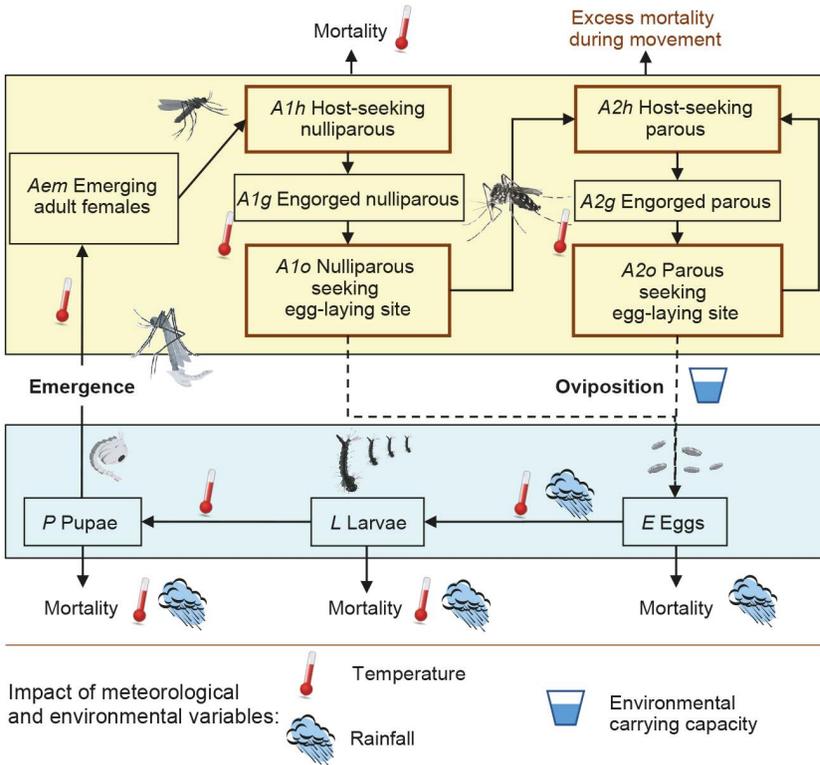


Figure 8.1. Schematic diagram of the life cycle of mosquitoes using a compartmental model. Adapted from Cailly *et al.* (2012).

Conversely, some processes are not represented:

- male mosquito compartments (due to the male not biting); however, this compartment should be added to test the impact of control measures such as the sterile insect technique (SIT), which consists of releasing sterile male mosquitoes for the purposes of population control (Haramboure *et al.*, 2020);
- the movement of mosquitoes (which remains limited to the expression of the mortality rate for females in search of a host or egg-laying site, conveying the risk associated with flight, biting and egg-laying); this model thus does not permit the simulation of mosquito dispersion.

►► Adaptation of the model to the species *Aedes albopictus* and *Aedes aegypti* and spatialization

Modelling the abundance of *Ae. albopictus* and *Ae. aegypti* mosquitoes represents a significant area of interest within the context of vector control. Globally, these two species are the primary vectors of dengue, chikungunya and Zika viruses. In France, *Ae. albopictus* is notably present in the Indian Ocean (La Réunion and Mayotte) and in numerous departments in metropolitan France, while *Ae. aegypti* has been observed in the Antilles, French Guiana and in the Indian Ocean. In the event that a user wishes to simulate the abundance dynamics of one of these species within a

particular biogeographical context, it is necessary to adapt the model presented in the preceding section, which is based on the generic mosquito life cycle. This adaptation is achieved by selecting values for the various parameters of the model and transition functions between the different life cycle stages (Ezanno *et al.*, 2015), in addition to the spatialization of the model.

Choice of parameters and functions for modelling mosquito population dynamics

This step requires a targeted literature review of the species and the environmental context (Ezanno *et al.*, 2015) to ascertain which publications explicitly reference the parameters under consideration (e.g., Delatte *et al.*, 2009, or Lacour *et al.*, 2010, for *Aedes*). Transition functions between stages can typically be derived from the outcomes of *in vitro* experimental studies, with rigorous control of the influential parameters (temperature, pressure, humidity) and by varying a single parameter at a time (Delatte *et al.*, 2009; Lacour *et al.*, 2010). A biostatistical relationship can then be established, thereby characterising the dynamics of the transition. In the absence of data on specific parameters or functions, consultations with medical entomologists and field observations by vector control services can be instrumental in estimating these parameters by taking into account the environmental context of the region (e.g., date on which *Ae. albopictus* exits diapause in temperate regions).

The set of parameters (constants of the model) and functions used to model *Ae. albopictus* (in temperate and tropical environments) and *Ae. aegypti* population dynamics are discussed in the publications of Focks *et al.* (1993) and Tran *et al.* (2013, 2020). For the same species (e.g., *Ae. albopictus* in a tropical environment), these parameters and functions can be reused in a different geographical context. This approach was used for experimental data obtained for the species *Ae. albopictus* on Reunion island (Delatte *et al.*, 2009), which allowed the population dynamics of the Asian tiger mosquito to be modelled on this island (Tran *et al.*, 2020), as well as on the neighbouring island of Mauritius (Iyaloo *et al.*, 2021).

Spatialization of the model: the concept of environmental carrying capacity

In addition to the temporal dynamics that are subject to meteorological variations, the spatial variations in mosquito population abundance between different regions or within the same region are constrained by the distribution of available resources, particularly the density of potential egg-laying and larval development habitats. In the case of *Ae. aegypti* and *Ae. albopictus*, the size of these habitats is typically small (plant pot saucers, tyres, cavities, etc.) and must be filled with water in the short, medium, or even long term. This is due to the high desiccation resistance of *Aedes* eggs, as discussed in Chapter 5. This process of filling with water, which may be attributed to either a natural phenomenon (rainfall) or an artificial intervention (watering), triggers the hatching of mature eggs and the subsequent progression to the larval stage (see Chapter 1). If we consider an area several hectares in size, the density of larval habitats may be aggregated in the form of a single average value for the area, representative of the “environmental carrying capacity”. This definition has its roots in the constant K in foundational ecology, which is linked to the logistic equation model (Verhulst, 1845).

It expresses a term of density-dependent mortality, which is applicable in this case to the larval stage and the emergence of an adult mosquito. This carrying capacity K is an important parameter and represents the maximum number of larvae that the environment can support without affecting the population itself. It therefore directly depends on the number of available larval habitats in a given area and the average production associated with each habitat (e.g., 10 larvae per habitat).

The characteristics and density of these habitats can vary significantly between regions, contingent on the nature of the environment (presence or absence of vegetation, woodland, etc.) and the usage and behaviour of human populations (water storage, plant watering, waste storage, etc.). Accordingly, the model must enable the user to set the carrying capacity of the environment based on field observations (Tran *et al.*, 2020) or estimates derived from information extracted from remotely sensed images, including land cover, vegetation, and the type of built-up area (see Chapter 5), in addition to external data such as the location of abandoned houses, graveyards, etc. Due to the type of larval habitats used by *Ae. albopictus* and *Ae. aegypti* mosquitoes, the environmental carrying capacity is split into two terms, K_{fix} and K_{lvar} , depending on the origin of the water filling the site: K_{fix} for sites watered by humans (considered “fixed”, i.e., does not vary over time) or K_{lvar} for sites watered by rainfall (considered naturally variable).

► Implementation, initialisation and simulation of *Aedes* mosquito abundance

Implementation using the *Ocelet* modelling language

In order to simulate the abundance of mosquitoes in a given area associated with one of the *Aedes* species described, the model is implemented in a computing environment. Implementation consists of defining and writing the code and programming variables that facilitate the automation of all calculations for simulating mosquito population dynamics for each plot under consideration. In spatial modelling, the *Ocelet* platform offers a suitable conceptual and practical framework that can be leveraged to gain valuable insights (Degenne and Lo Seen, 2016). This platform offers a business language that is both generic and precise for the description of dynamic phenomena in a landscape based on the concept of interaction graphs. It also provides a set of tools that enable the identification of relationships between geographic features that are compatible with geographic information systems (GIS).

In *Ocelet*, a system is represented by defining “entities”, which may have different spatial representations (polygons, points, etc.). The system dynamics are simulated using a “scenario”, in which the state of the entities will change over time via the application of functions. These functions may be specific to a given entity type (“services”) or dependent on other entities with which they interact (“relations”) [Degenne and Lo Seen, 2016].

The inputs of the model are as follows (Figure 8.2a):

- the location of weather stations (shapefile geographic format);
- the daily rainfall and maximum and minimum temperatures (text format). As previously indicated (Figure 8.1), these variables will impact the transition rates between stages as well as mortality rates;

– an “environmental file” (shapefile format), which includes the location of the plots for which the densities of each stage will be estimated. These plots are characterised by their altitude and environmental carrying capacity (denoted K , see previous section).

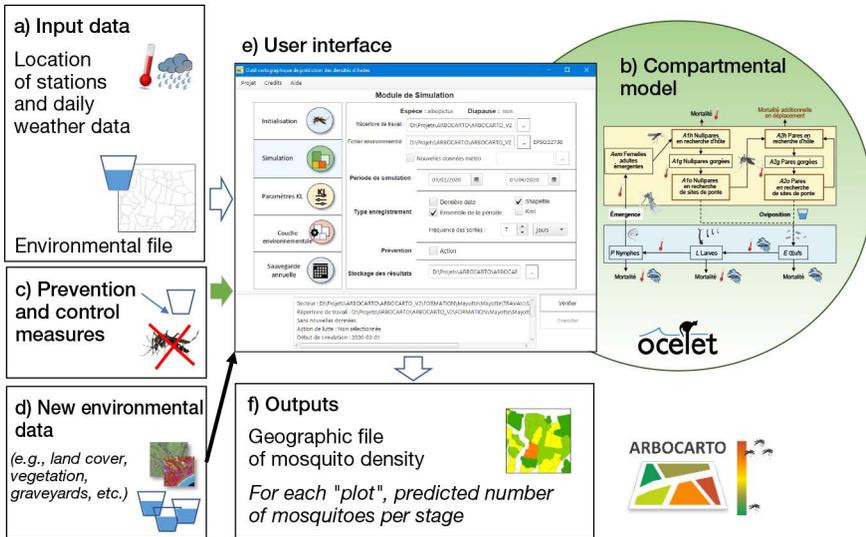


Figure 8.2. Functional diagram of the Arbocarto.

(a) Input data required to begin a simulation, (b) integrated biological model (Figure 8.1), (c, d) optional input data, (e) interface (Figure 8.4), and (f) output generated.

Each plot is considered independent from one another and, for each step in time:

- the temperature and rainfall values for each weather station are read and assigned to each corresponding plot (the closest station);
- the functions of the models are updated for each plot;
- the ODE system (Figure 8.2b) is resolved for each plot allowing for an estimation of number of individuals in each stage ($E, L, P, Aem, A1h, A1g, A1o, A2h, A2g, A2o$).

For each plot of the “environmental file”, the model outputs (Figure 8.2f) the number of individuals in each stage at each step of time. It is important to note that the model predicts mosquito densities for the day corresponding to the last date of meteorological data, and does not predict future densities. However, the prediction of aquatic stage densities (P and L) allows adult densities to be estimated for a few days.

Initialisation and example of simulations

The initialisation phase is carried out as follows: arbitrary assignment of a large number of eggs (e.g., 1,000,000) for each constituent plot of the area in question and null values for all the other stages. An embargo period of at least one year is allowed to pass in the simulation (the first year results of the simulation are not used, but serve to initialise mosquito population dynamics in a more realistic manner). This phase reiterates a crucial point previously highlighted in the simplifications made in Section 2. The model has been constructed to reflect the dynamics of an established mosquito population, once the colonisation process has stabilised at the specific site in question.

The simulation is based on the calculation of abundance for a specific time step, with a minimum duration of one day and a typical duration of one week. The length of the simulated period is dependent on the availability of meteorological data and the implementation or absence of control measures against mosquito populations. The advantages of a mechanistic approach are exemplified by the simplicity with which the majority of vector control measures can be implemented:

- the destruction of habitats is modelled by a decrease in the environmental carrying capacity for the plot(s) in question:
- the use of larvicides or adulticides (fumigation) is equivalent to temporarily increasing the mortality rate of the targeted stage: larvae or adults.

The model permits the execution of multiple consecutive simulations (with or without control measures) and the evaluation of the potential impact of a single measure or an integrated control strategy (combining multiple measures) through *in silico* experimentation (Figure 8.3).

Model validation

A comparison between the simulation data (driven by daily temperature and rainfall values measured by the corresponding weather stations) and entomological data collected in the field allowed the model outputs to be validated for several test sites in temperate and tropical environments (Iyaloo *et al.*, 2021; Tran *et al.*, 2013, 2020).

This step allows the validity of model outputs to be verified prior to their utilisation for predictive purposes or for the evaluation of the impact of different control scenarios.

►► Arbocarto: a specialised interface for vector control measures

This model, based on the mosquito's life cycle, considers both the fluctuations in meteorological data (temperature and rainfall) that drive the temporal dynamics of the model and the suitability of the environment in terms of larval habitat density. This allows for the provision of a spatialization that is consistent at the landscape scale (most often surface areas in the order of several hectares).

The utilisation of the Ocelet platform allows simulations to be quickly launched in a scientific context. However, in order to reach the greatest number of users, it is essential that the model can be utilised by individuals lacking any specialised IT knowledge. The Ocelet tool allows an executable file (JAR format) to be exported in a Java runtime environment (JRE). Nevertheless, it is still feasible to access a comprehensive array of variables for the model, including the choice of species, the period and location of the simulation, as well as the incorporation of black spots observed in the field (high density of larval habitats) or vector control measures. To this end, an interface developed in JavaFx⁴⁶ affords comprehensive control of the model, its initialisation, and the simulation phase, all of which can be accomplished in a few clicks (Figure 8.4).

Additionally, the user is able to modify the initial environmental file in accordance with the specific characteristics of the land cover or vegetation, or by incorporating one or more supplementary layers of geographic information, including the location

46. <https://openjfx.io/>

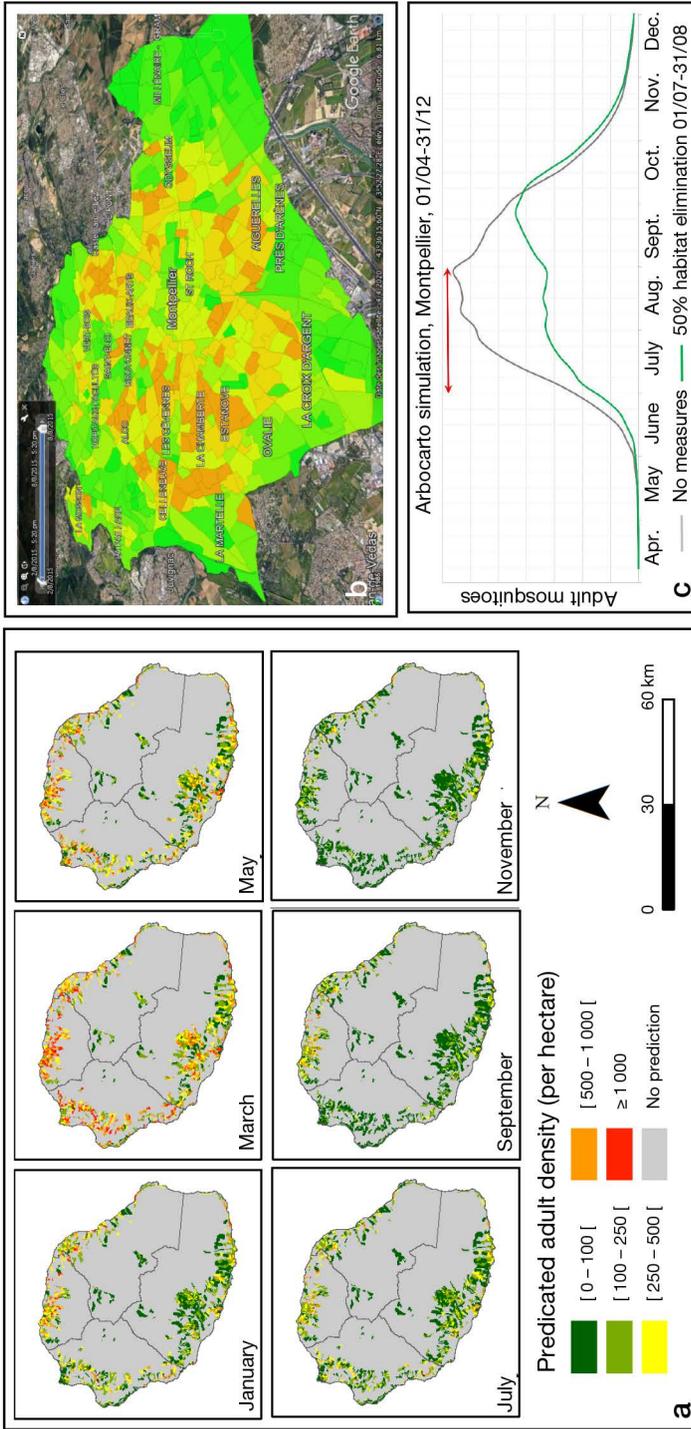


Figure 8.3. Examples of outputs from the Arbocarto tool.

(a) **Multi-day** outputs, in a geographic information system, Reunion island, 2018.
 (b) Population dynamics output in KML format for visualisation with Google Earth, Montpellier, 2015. (c) Temporal outputs enabling the impact of various control measures to be visualised for a plot.

of potential larval habitat sites (graveyards, garages, home gardens, terrace gardens, abandoned houses, etc.; Figure 8.2f); the addition of this information modifies the environmental carrying capacity of each plot in question.

The interface is divided into three main sections (Figure 8.4): on the left is the menu with the various modules to be enabled; in the middle are the parameters to be set for each module; at the bottom, warning messages or information is shown to ensure the simulation proceeds smoothly.

A text file is generated for each output produced by the application, in which the parameters selected by the user to generate this output are listed.

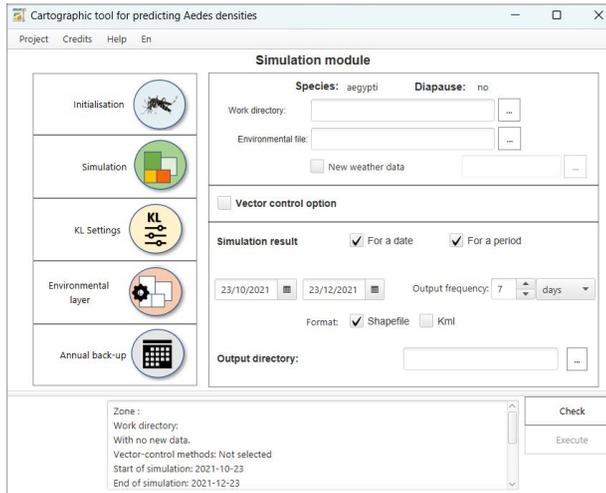


Figure 8.4. Screenshot of the Arbocarto interface for controlling the settings of the population dynamics model for *Aedes albopictus* and *Aedes aegypti* mosquitoes.

The model is controlled via specific modules for initialisation and simulation, thereby enabling vector control measures and entomological observations in the field to be taken into account.

In order to facilitate user proficiency with the tool, training has been provided. The Arbocarto tool is now entering its final phase, namely its operational deployment and routine use in organisations and frameworks responsible for surveillance and vector control, including interdepartmental agreements for mosquito control (EID, *ententes inter-départementales pour la démoustication*), regional health agencies (ARS, *agences régionales de santé*) and local authorities. These structures will employ this tool in addition to their prospecting efforts, awareness campaigns, and vector control measures in the event of an epidemic. The project has a dedicated website⁴⁷ providing information on its current initiatives and ways to access the application under the CeCILL-C free software licence.

47. www.arbocarto.fr/en

►► References

- Baker R.E., Pena J.M., Jayamohan J., Jerusalem A., 2018. Mechanistic models versus machine learning, a fight worth fighting for the biological community? *Biology Letters*, 14 (5).
- Cailly P., Tran A., Balenghien T., L'Ambert G., Toty C., Ezanno P., 2012. A climate-driven abundance model to assess mosquito control strategies. *Ecological Modelling*, 227, 7-17.
- Craver C.F., 2006. When mechanistic models explain. *Synthese*, 153 (3), 355-376.
- Degenne P., Lo Seen D., 2016. Ocelet: Simulating processes of landscape changes using interaction graphs. *SoftwareX*, 5, 89-95.
- Delatte H., Gimonneau G., Triboire A., Fontenille D., 2009. Influence of Temperature on Immature Development, Survival, Longevity, Fecundity, and Gonotrophic Cycles of *Aedes albopictus*, Vector of Chikungunya and Dengue in the Indian Ocean. *Journal of Medical Entomology*, 46 (1), 33-41.
- Ezanno P., Aubry-Kientz M., Arnoux S., Cailly P., L'Ambert G., Toty C., Balenghien T., Tran A., 2015. A generic weather-driven model to predict mosquito population dynamics applied to species of *Anopheles*, *Culex* and *Aedes* genera of southern France. *Preventive Veterinary Medicine*, 120 (1), 39-50.
- Focks D.A., Haile D.G., Daniels E., Mount G.A., 1993. Dynamic Life Table Model for *Aedes aegypti* (Diptera: Culicidae): Analysis of the Literature and Model Development. *Journal of Medical Entomology*, 30 (6), 1003-1017.
- Haramboure M., Labbé P., Baldet T., Damiens D., Gouagna L.C., Bouyer J., Tran A., 2020. Modelling the control of *Aedes albopictus* mosquitoes based on sterile males release techniques in a tropical environment. *Ecological Modelling*, 424, 109002.
- Iyaloo D.P., Degenne P., Elahee K.B., Lo Seen D., Bheecarry A., Tran A., 2021. ALBOMAURICE: A predictive model for mapping *Aedes albopictus* mosquito populations in Mauritius. *SoftwareX*, 13, 100638.
- Lacour G., Carron A., Jeannin C., Delaunay P., Benoit R., Perrin Y., Lagneau C., 2010. *Aedes albopictus* (Skuse, 1894) (Diptera: Culicidae) in France: gonotrophic cycle, fecundity and longevity, in *Proceedings of the International Conference "Emerging Vector-borne Diseases in a Changing European Environment"*, 10-12 May 2010, Montpellier, France.
- Tran A., L'Ambert G., Lacour G., Benoit R., Demarchi M., Cros M., Cailly P., Aubry-Kientz M., Balenghien T., Ezanno P., 2013. A rainfall- and temperature-driven abundance model for *Aedes albopictus* populations. *International Journal of Environmental Research and Public Health*, 10 (5), 1698-1719.
- Tran A., Mangeas M., Demarchi M., Roux E., Degenne P., Haramboure M., Le Goff G., Damiens D., Gouagna L.C., Herbreteau V., Dehecq J.-S., 2020. Complementarity of empirical and process-based approaches to modelling mosquito population dynamics with *Aedes albopictus* as an example - Application to the development of an operational mapping tool of vector populations. *PLoS One*, 15 (1), e0227407.
- Verhulst P., 1845. Recherches mathématiques sur la loi d'accroissement de la population. *Nouveaux mémoires de l'Académie Royale des Sciences et Belles-Lettres de Bruxelles*, 18 (1-45).

Chapter 9

Spatial simulation of the risk of dengue transmission using vector and host behavioural models

Éric Daudé, Sébastien Rey-Coyrehourcq, Alexandre Cebeillac

Arthropod-borne diseases, including dengue, chikungunya and those transmitted by the Zika virus, represent a significant threat to public health when they reach epidemic levels. At the global level, the occurrence of epidemics is linked with climatic parameters (temperature and rainfall) and urbanisation. However, understanding the transmission dynamics of these epidemics remains a significant challenge at the suburban level. In essence, at this scale, the heterogeneity of the environment, spatio-temporal variations in vector density and human population mobility make the geography of the disease and its spread particularly complex. Efforts to control the spread of these diseases primarily target their vectors, predominately the urban mosquito *Aedes aegypti*. However, these efforts are undermined by the vast expanse of the areas in question, the difficulty in reaching the locations where the mosquitoes or their egg-laying sites are potentially located, and the financial and personnel resources that are required. Despite the efficacy of the control strategies developed in the 1960s, which included the intensive use of insecticides such as DDT, these strategies have since become partially ineffective. However, an understanding of the pathogen system and technological advances has enabled the development of targeted control strategies, which are focused on vector mosquitoes, particularly in the absence of a vaccine (Hoffmann *et al.*, 2011; Seixas *et al.*, 2019). However, the identification of target areas for vector control in large metropolitan areas remains a major challenge.

In spatial epidemiology, one of the most frequently proposed hypotheses is that specific locations have the potential to become hyper-localised contamination hotspots, responsible for the multiplication of cases and the amplification of the virus. From these locations, as a result of human mobility, other locations contribute to viral propagation at a larger scale, acting as dispersal nodes that favour the spread of pathogens in urban environments (Daudé *et al.*, 2015). How are these locations identified? Do they have any special socio-environmental characteristics? How do they change over time and what are the impacts on epidemiological dynamics?

At the scale of a large city with several million inhabitants, the identification of these locations could permit the implementation of more targeted and less intrusive

control measures, thereby reducing the number of epidemic foci, reducing the burden on healthcare services in times of an epidemic and reducing the economic burden for individuals and authorities. Exploring the relative weight of these factors—local outbreaks vs spread via human mobility—in the overall dynamics of an epidemic is a challenging undertaking. This is due to the non-linear relationships that exist between these factors and the difficulties encountered in accessing or collecting relevant data to evaluate these different mechanisms. Spatial modelling is particularly well-suited to exploring these two lines of research: firstly, the differential weighting of locations in transmission risk exposure; and secondly, the role of individual mobility in the spread of the virus.

► Individual-based and spatially explicit models

A number of different approaches can be employed to model and characterise an object or domain of study. The choice of one or the other depends on the goals of the model, the extent of knowledge about the phenomenon in question, and the availability of data to describe it.

Different types of modelling

The first type of model is based on the availability of quantitative data which is used to provide a summary or description: these are statistical models. Multinomial logistic regression models and spatial statistics such as spatio-temporal autocorrelation can be used to search for links between disease clusters, the socio-economic and environmental characteristics of neighbourhoods, and types of land cover. Residual analysis of these models allows researchers to look for local explanations for these differences and to understand, for instance, why some neighbourhoods seem to be more exposed to the disease than others (Zellweger *et al.*, 2017).

A second type of model is based on the characterisation of processes, particularly for the estimation of values that are unknown at the local level or for the prediction of system evolution. A gravity model can be used to calculate the flow of commuters between different towns. This is done by taking into account the mass of the towns in question (expressed by the volume of their populations) and the distance between them. This model can then be used to estimate the risk of a virus spreading within an urban system, or even between countries via international flights (Salami *et al.*, 2020).

Finally, a third approach is based on behavioural models, i.e., the way that an entity functions or performs actions in their environment. A pattern of movement can thus describe the daily movement of an individual, which is constrained by the activities that they must perform. The simulation of thousands of these trips can then reproduce the collective mobility dynamics within a region combined with an interpersonal interaction mechanism in order to simulate virus transmission (Karl *et al.*, 2014). A multi-agent system model (Treuil *et al.*, 2008) is particularly well-suited for this type of research question.

Use of multi-agent systems in epidemiology

Multi-agent systems (MAS) use so-called “individual-based” methods. They enable the execution of computer simulations which are then used as “artificial worlds” in order to

perform experiments under conditions similar to those found in a laboratory setting. In essence, the utilisation of these models enables the manipulation and alteration of parameters, as well as the replication of a series of experiments through simulation. Observing and measuring the evolution and change in the agent's state allows for an inference of what would happen in the real world under similar conditions.

MAS come from research in distributed artificial intelligence and robotics, whereby the agents of a MAS interact with one another and with their environment, typically according to cooperative, competitive and coexisting modes. This is done with the objective of resolving problems that are too complex for resolution by a single individual. By explicitly characterising the space, time, interactions, behaviours and levels of organisation, these methods represent a valuable contribution to geographic research, particularly in studies exploring phenomena which propagate through space. A detailed characterisation is proposed for the fundamental mechanisms involved in the dynamics of the phenomenon, which is then employed in a computer simulation to assess its validity.

In areas where the disease is endemic, our initial hypothesis is that urban environments are characterised by a limited number of locations that possess the requisite socio-economic conditions to sustain mosquito populations during inter-epidemic periods. This presence is sufficient to guarantee the local and continuous low-intensity circulation of the viruses. During seasonal changes throughout the year, such as temperature rises and/or monsoons, vector populations explode, thereby increasing the overall risk of viral propagation. Our second hypothesis is that the socio-spatial organisation and discontinuities in urban environments, which structure daily mobility, are the source of pathogen propagation. From these initial focal points, the pathogens then spread throughout the urban area, leading to the occurrence of epidemics of varying sizes. It is therefore imperative to identify potential reservoirs of an infectious agent, understood here to be a set of hyper-localized sites linked to environmental and anthropogenic factors, in order to act efficiently against the spread of an epidemic.

However, the assessment of the epidemiological significance of these locations remains a challenge. In essence, the data collected by surveillance systems records the residence of the infected individual as the site of contamination, although this may not be the original source of contamination. In addition, not all cases are recorded. Epidemiological studies based on this data alone have a tendency to overestimate the local risk (contagion by proximity) in comparison to the overall risk (jump diffusion). In this context, modelling and targeted field research may elucidate the role of local microenvironments and human mobility in the dissemination of pathogens in urban settings. A major obstacle in modelling and data acquisition arises when attempting to consider the impact of urban environmental and socio-economic heterogeneity at high resolutions on epidemiological dynamics.

►► Application to dengue in Bangkok: MO³, methods and data

We developed a set of submodels for the MO³ simulator (Daudé *et al.*, 2015). This enabled a simulation model to be designed, made of different components (Figure 9.1), which integrated biological knowledge of the vector and daily human mobility. The dynamics of these two submodels were linked to socio-environmental data at high resolution. The field of application of this model was Bangkok, Thailand.

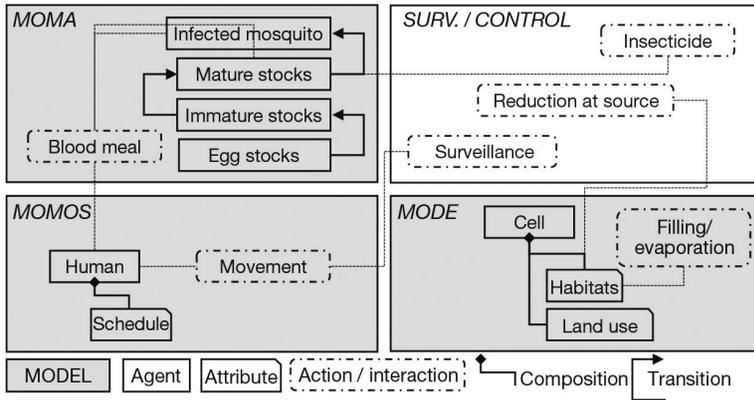


Figure 9.1. Primary components of the MO³ model.

MOMA: Model Of Mosquito Aedes. MOMOS: Model Of Mobility Simulation. MODE: Model Of Dynamical Environment.

The following sections present three models: one for the environment (MODE), one for the vector mosquito (MOMA), and one for the host (MOMOS). Additionally, the transmission mechanisms of infectious agents between vectors and hosts are discussed. The last section of the chapter is dedicated to an MO³ simulation and methods to study this type of model.

Environmental modelling

Environmental dynamics are simulated by the model of dynamical environment, MODE (Misslin and Daudé, 2016). The agents of this model are cells which represent the specific portions of the territory to which the model is being applied. These cells are characterised by different environmental parameters which influence the behaviour of *Ae. aegypti*: classification of land cover (vegetation, water, type of built-up area, etc.), the number of potential larval habitats (receptacles which may contain fresh water), air temperature and precipitation. The data that enables these attributes to be quantified is derived from satellite image processing, weather station records and expert knowledge. Each agent of the environment (cell) also possesses functions which describe the change in the quantity of water available in potential egg-laying sites (filling and draining) and temperature fluctuations. These variations are calculated based on data obtained from MODIS thermal images (Aqua and Terra satellites). In the event of excessive cloud cover, it is possible that some pixels may return null values. In this scenario, MODE assigns a global temperature value to the specified area, which is derived from a weather station in Bangkok⁴⁸. It is these same stations that provide rainfall data on a daily basis. The spatial resolution of an MO³ cell is 30 × 30 m, which corresponds to the resolution of Landsat 8 images (OLI). This resolution is compatible with the average range of dispersal of mosquitoes observed by mark–release–recapture studies (Sheppard *et al.*, 1969).

MODE enables the generation of a substantial quantity of spatially explicit and dynamic data (Figure 9.2), which is used for behavioural models of mosquitoes and humans. Air temperature values can then be calculated, thereby facilitating the identification

48. <https://aqicn.org/city/thailand/bangkok/thai-meteorological-department-bangna/vn/wxqa.com>

of large differentials across Bangkok (Figure 9.2a; see Chapter 3), a vegetation index (Figure 9.2b) and estimates of human population distribution (Figure 9.2c; see Chapter 4). The number of mosquito egg-laying sites is also estimated (Figure 9.2d). In light of the high cost of collecting data in the field, an approach based on knowledge of the terrain and expert assessments was deemed preferable in this case. Given the density of mosquito habitats in urban spaces is highly linked to the presence of humans and their activities (Ooi *et al.*, 2006), an index was created for the abundance of potential egg-laying sites based on the local density of households (Misslin *et al.*, 2016).

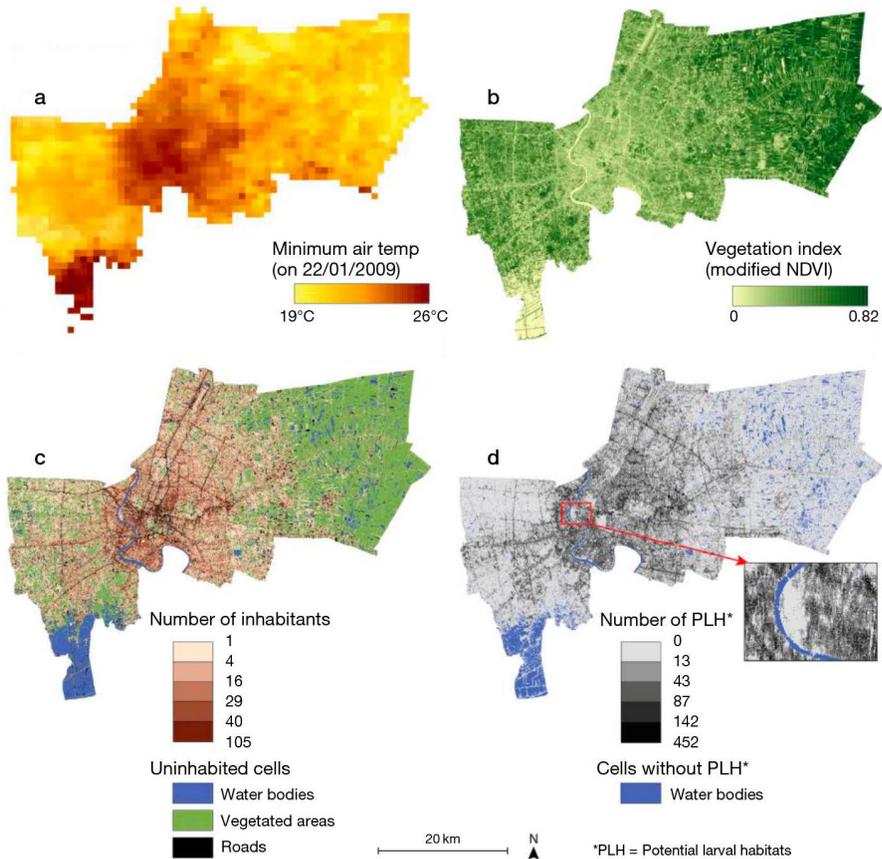


Figure 9.2. Environmental factors constituting an ecological niche for the vector.

(a) Estimated air temperature. (b) Vegetation index. (c) Distribution of human population. (d) Distribution of potential larval habitats.

Some of this spatially explicit data is then input into the mosquito model (Misslin and Daudé, 2017).

Vector mosquito modelling

The model of mosquito aedes (MOMA) is a behavioural submodel for the MO³ simulator. This model permits the simulation of vector populations in relation to environmental characteristics (Maneerat and Daudé, 2016). Figure 9.3 shows the

general concept of the transition between the different states for *Aedes* (aquatic and aerial phases); Figure 9.4 shows the interactions between the adult female mosquito and the environment in the model.

For simulations at a city scale, aquatic development phases are described by stock models at a spatial resolution of 30 × 30 m, whereas the aerial phase is described by an individual-based model (Figure 9.3). *Aedes* females generally lay their eggs in containers where fresh water collects. Egg development begins once it comes into contact with water. This development phase lasts 2 to 3 days depending on water temperature. The mosquito transitions to the larval stage upon hatching. Larvae are capable of moving in water to eat and breathe. At the end of 4 to 6 days, the larva transitions to the pupa stage. This development stage lasts between 2 and 7 days depending on the temperature. Up until this point in time, the development of the mosquito in MO³ has been entirely constrained by the presence of water in its habitat. This reproduces the strong link which exists between the increase in vector density and precipitation. Once the aquatic stage is complete, the mosquito progresses to its aerial stage. In this model, once the pupal stock has reached maturity, it triggers the creation of mosquito agents, with only the females being simulated. Blood meals provide the energy required for the completion of the oviposition phase (O). The females begin their gonotrophic cycle (Gn), during which they are prompted to bite mammals (essentially humans).

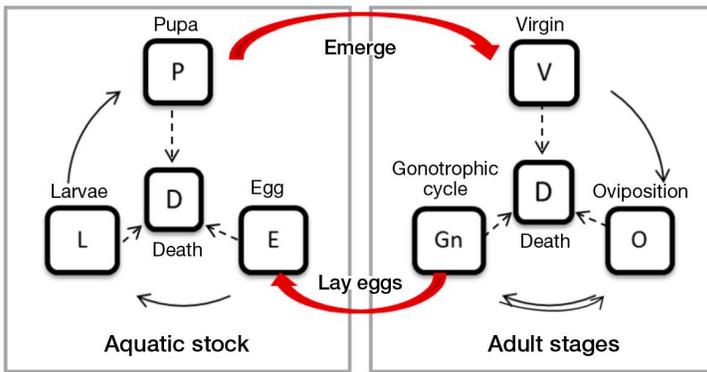


Figure 9.3. Diagram of state transitions for the *Aedes* agent life cycle according to its development stage (immature or aquatic, adult or aerial).

The arrows represent the possibilities of transitioning from one state to another.

MOMA is an individual-based model of *Ae. aegypti* adult female mosquitoes which are responsible for the transmission of viruses during blood meals. The activities which model mosquito behaviour are as follows (Figure 9.4): feeding, reproduction, resting, movement and survival. These activities are related to the needs of the species, which can be satisfied by the presence of blood and nectar sources for feeding, eggs-laying containers for reproduction, shaded areas for resting and circulation areas for movement. As for survival, this essentially depends on temperature, humidity and a stochastic parameter which factors in the dangers to which the mosquito is exposed on a daily basis (in particular predation). *Aedes* therefore acts in accordance with its biological state, as defined by its internal variables, and in response to the resources present in its immediate environment.

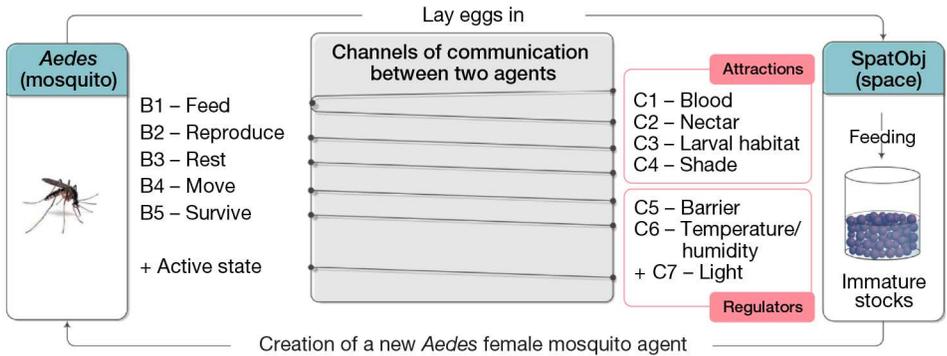


Figure 9.4. Modelling interactions between females mosquitoes and the environment.

Modelling hosts and their mobility

The model of mobility simulation (MOMOS) describes the host component, which essentially corresponds to human mobility behaviour (Cebeillac *et al.*, 2017). A host agent is situated within the environment through which they move in order to complete their activities. The act of moving and changing location exposes humans to a range of risks (Figure 9.5). In the model, movements are made between different

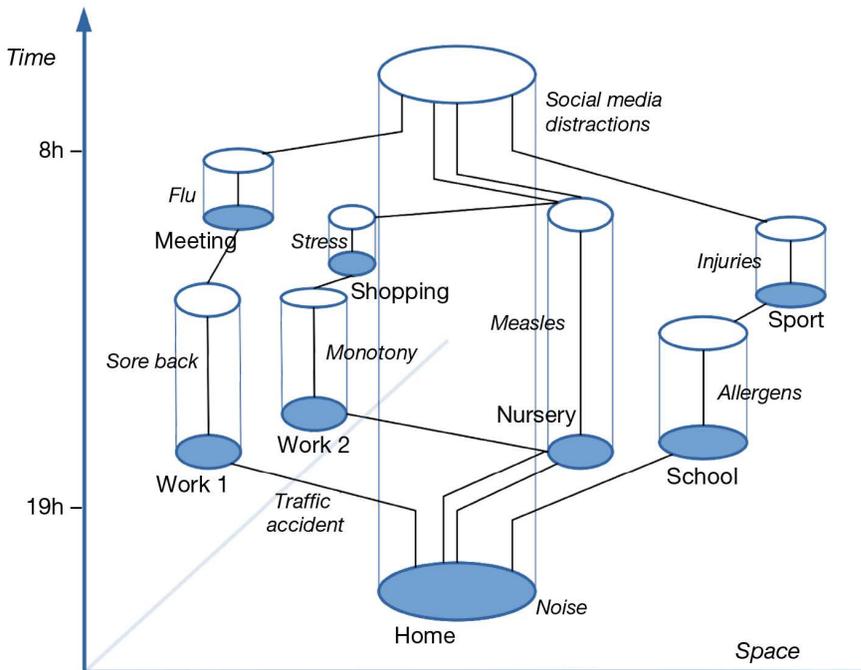


Figure 9.5. Representation of the activity space defined by Hägerstrand (1970) in the context of health. In this example, four people from the same household visit different places throughout the day. Each of these places is associated with a health risk according to the activity performed in this place and the time of day. As defined by Cebeillac and Daudé (2022).

areas of activities (house, school, workplace, recreational activities, etc.) based on a mobility schedule created for each agent. The schedules are generated using data from a variety of sources, including social networks (Twitter and Facebook) and timesheets from surveys conducted (Cebeillac *et al.*, 2018). Agents are generated using a synthetic population generator (see Chapter 4), with each agent being characterised by geographic attributes, such as a place of residence and schedule of daily activities. In the model, the activities performed by an individual, which determine the places visited, depend on their age, place of residence and epidemiological status (an agent presenting symptoms of dengue does not move).

Mechanisms of virus transmission

Mosquito agents (*m*) and host agents (*h*) are defined, according to their exposure to the virus (Figure 9.6), as susceptible (*Sm*, *Sh*), exposed (*Em*, *Eh*) and infected (*Im*, *Ih*) respectively. Only the host agent can be immunised (*Rh*), with the mosquito agent remaining infected, and therefore contagious (*Im*), until its death. The transmission of the virus from an infected host to a healthy mosquito, or from an infected mosquito to a healthy host, occurs when the mosquito takes a blood meal.

These different states may or may not change the behaviour of the agents. For example, it is feasible to restrict or even halt the mobility of the host agent in the event of contamination, in order to account for the appearance of incapacitating symptoms of varying degrees.

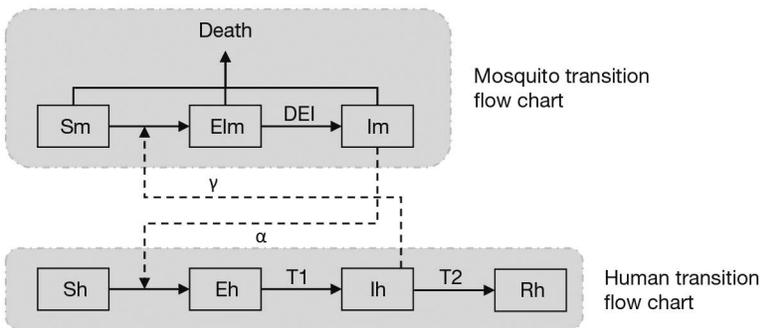


Figure 9.6. Transition diagram between different states for a mosquito agent (*m*) and host agent (*h*).

The host agent, in a susceptible state (*Sh*), may be infected by the dengue virus following a bite from an infected female mosquito (*Im*) with a risk of transmission (α). Once infected, the host develops the virus over a period of time ($T1$), which is referred to as the intrinsic period. Subsequently, the host enters the viraemia phase ($T2$), during which the virus may be transmitted to a healthy mosquito that has bites the host (*Sm*) with a probability (γ). Infected *Aedes* mosquitoes enter a period of extrinsic incubation (*Elm*) before becoming infectious (*Im*).

►► MO³ model simulations

MO³ simulations are multi-temporal, with one time step (iteration) per minute. The simulation time is therefore split into three scales depending on the processes: days, hours and minutes. Aquatic development of the mosquito and environmental variables are updated according to average daily temperatures, rainfall and evaporation.

The mobility of human agents is simulated on an hourly basis, according to their schedule and the activities of adult mosquitoes according to temperature, daylight hours and the need for specific actions on a minute-by-minute basis. The recording of the various parameters during each iteration permits the reconstruction and analysis of the evolution of multiple dynamics. It is thus feasible to monitor the development of mosquito populations in their aquatic stages and calculate the density of adult mosquitoes at the resolution of the cell. Furthermore, it is possible to identify the various points of interaction between hosts and vectors and to calculate the environmental profiles that are at risk at the time of these interactions, which in turn result in contamination loops. The simulation interface of the model permits the visualisation of the evolution of different indicators in the form of graphics and maps (Figure 9.7)

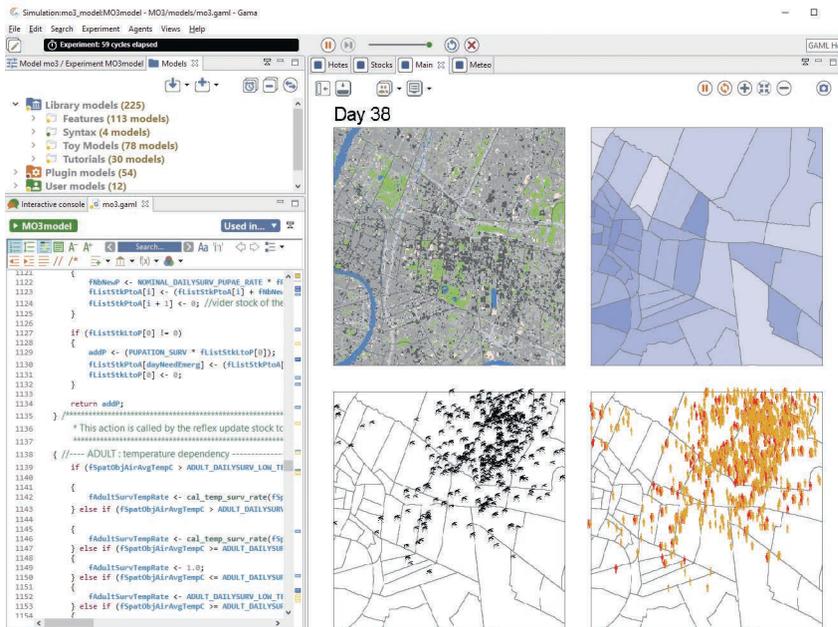


Figure 9.7. MO³ model simulation interface.

Mosquito life cycle dynamics

Over the course of a one-minute iteration, mosquito agents explore their environment and may select a target which responds to their needs and move towards it. If no target is identified, as none satisfy one of the expressed needs, the mosquito moves at random. If a target is selected, the mosquito moves in its direction. In the mechanism of movement, the porosity coefficient of the environment plays a major role. For instance, it will limit the passage between one spatial object, such as a road, to another, such as a house (barrier effect). Conversely, it will facilitate the passage from a road to a green space (corridor effect) [Maneerat and Daudé, 2016]. To illustrate, we conducted a simulation of the population dynamics and dispersal of mosquitoes based on geographic configurations. The objective was to measure the effects of land use categories and human densities on the dispersal of mosquitoes from disparate emergence locations (Figure 9.8).

Firstly, these simulations enabled us to ascertain whether the mean distances travelled by virtual mosquito cohorts in urban areas align with the empirical data and literature-based estimates. Secondly, we were able to establish relationships between these relatively low distances (several dozen metres to several hundred metres at most) and local land use configurations and human densities. Hence, the highly populated and densely built-up area of the urban slum (at the centre of Figure 9.8) showed low dispersal of mosquito cohorts, contrary to the more peripheral areas, which are less dense and characterised by rows of houses and gardens which the mosquitoes can cross, from one to another and depending on their needs, thus travelling over greater distances over the course of their life.

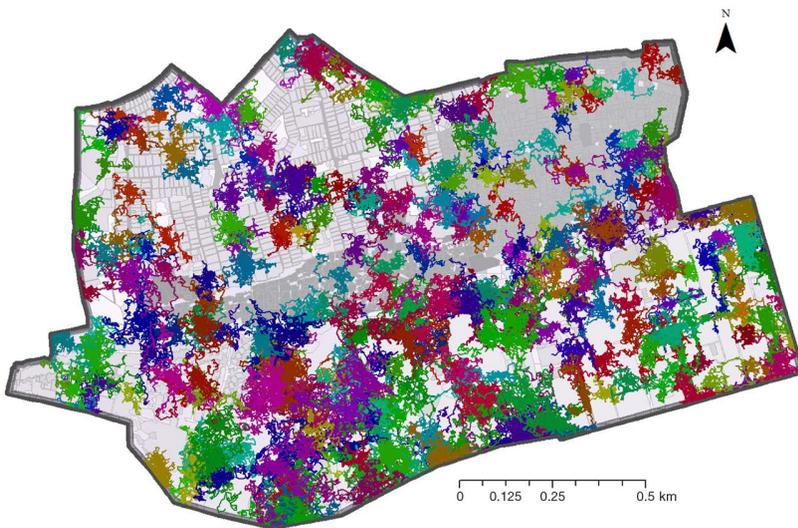


Figure 9.8. Movement patterns of agent mosquitoes in a neighbourhood of Delhi (India). Each line of colour traces the path taken by the mosquito over the course of its life.

More than 50 parameters and variables are employed in the MO³ model, developed using the multi-paradigm programming language Scala⁴⁹. This level of detail results in a particularly complex calibration and validation process. In essence, the number of potential combinations of parameters is so great that it is almost impossible to systematically explore all of their domains. In this example, the validation of the mosquito model was performed in accordance with two criteria: expert knowledge and survey data (Maneerat and Daudé, 2017). Research is ongoing to study the behaviours of this model, and to calibrate and explore different vector control scenarios.

Exploration of model behaviours

The calibration data employed is derived from vector surveillance and epidemiological data. One set of data is used to calibrate and validate the mosquito model, while the other is employed to test the hypothesis of the role of human mobility in the spread of the disease at a city scale. The reconstruction of a virtual world based on a specified

49. <https://www.scala-lang.org/>

number of proxies from the real world and stylised facts produces results that must still be evaluated in order to ascertain their suitability for application in real-world scenarios. A model, in and of itself, is not capable of capturing the full complexity of the phenomena that occur in the real world. It is essential to validate the processes at both micro- and macro-scales in order to guarantee the consistency and robustness of the results produced by this simulation data. The generation of this data is not intended to provide definitive answers to all questions; rather, it is a means of addressing a specific set of questions concerning the selected mechanisms and parameters. MO³ submodels use a multitude of interacting mechanisms and parameters, which generate an extensive range and volume of behaviours and data, rendering them infeasible to evaluate without the use of more targeted methods.

In order to cope with the vast number of possible combinations—for models potentially running for several minutes or even hours—it is necessary to employ exploration algorithms (evolutionary algorithms) and the OpenMOLE platform (Reuillon *et al.*, 2013) capable of exploiting the parallelism offered by HPC (High Performance Computing) environments. MO³ can benefit from the European grid computing infrastructure (4,000 cores) and the high-performance computing power of CRIANN, the computing mesocentre for Normandy (10,000 cores). These resources can be employed for weeks at a time to run simulations and evaluate the consistency of the results obtained for these questions.

►► Conclusion

MO³ is a simulator based on behavioural models at high spatial and temporal resolutions. It describes a complex pathogen system to study the environmental and anthropic factors in the spread of dengue. MO³ is currently in the development phase and has been used throughout the project of the same name⁵⁰ to direct data collection and field experiments. Upon completion, it may be utilised to ascertain the most effective vector control strategies for an *in situ* evaluation based on optimisation algorithms.

►► References

- Cebeillac A., Daudé É., 2022. Mobilité quotidienne et diffusion des épidémies. *in* Valley J. (ed.), *Mobilité quotidienne et santé*, London, ISTE Science publishing LTD.
- Cebeillac A., Daudé É., Huraux T., 2017. Where? When? And How Often? What Can We Learn on Daily Urban Mobilities From Twitter Data and Google Map in Bangkok (Thailand) and Which Perspectives For Dengue Studies? *Netcom, Networks and Communication Studies*, 31 (3-4), 283-308.
- Cebeillac A., Daudé É., Vaguet A., 2018. Spatial discontinuities, health and mobility. What do the Google's POIs and Tweets tell us about Bangkok's (Thailand) Structures and Spatial Dynamics? *International Review of Geomatics*, 28 (4), 389-407.
- Daudé É., Vaguet A., Paul R., 2015. La dengue, maladie complexe. *Natures Sciences Sociétés*, 23 (4), 331-342.
- Hägerstrand T., 1970. What about people in regional science? *Papers of Regional Science Association*, 24, 7-21.
- Hoffmann A.A., Montgomery B.L., Popovici J., Iturbe-Ormaetxe I., Johnson P.H., Muzzi F., Greenfield M., Durkan M., Leong Y.S., Dong Y., Cook H., Axford J., Callahan A.G., Kenny N., Omodei C., Mcgraw E.A., Ryan P.A., Ritchie S.A., Turelli M., O'Neill S.L., 2011. Successful establishment of *Wolbachia* in *Aedes* populations to suppress dengue transmission. *Nature*, 476 (7 361), 454-457.

50. www.MO3.cnrs.fr

Karl S., Halder N., Kelso J.K., Ritchie S.A., Milne G.J., 2014. A spatial simulation model for dengue virus infection in urban areas. *BMC Infectious Diseases*, 14, 447.

Maneerat S., Daudé É., 2016. A spatial agent-based simulation model of the dengue vector *Aedes aegypti* to explore its population dynamics in urban areas. *Ecological Modelling*, 333, 66-78.

Maneerat S., Daudé É., 2017. Étude par simulation à base d'agents des effets des discontinuités intra-urbaines à Delhi sur la dispersion des moustiques *Aedes aegypti*, vecteurs de la dengue, de la fièvre jaune, du chikungunya et du virus Zika. *Cybergeo : European Journal of Geography*, 10.4000/cybergeo.28078.

Misslin R., Daudé É., 2016. Génération d'environnements artificiels pour la simulation spatiale d'arboviroses en milieu urbain : application à la dengue et au virus Zika, in *Actes du colloque SAGéo, Conférence internationale de Géomatique et Analyse Spatiale*, décembre 2016, Nice, 389-402.

Misslin R., Daudé É., 2017. An environmental suitability index based on the ecological constraints of *Aedes aegypti*, vector of dengue. *Revue Internationale de Géomatique*, 27, 481-502.

Misslin R., Telle O., Daudé É., Vaguet A., Paul R.E., 2016. Urban climate versus global climate change-what makes the difference for dengue? *Annals of the New York Academy of Sciences*, 1382 (1), 56-72.

Ooi E.E., Goh K.T., Gubler D.J., 2006. Dengue prevention and 35 years of vector control in Singapore. *Emerging Infectious Diseases*, 12 (6), 887-893.

Reuillon R., Leclaire M., Rey-Coyrehourcq S., 2013. OpenMOLE, a workflow engine specifically tailored for the distributed exploration of simulation models. *Future Generation Computer Systems*, 29 (8), 1981-1990.

Salami D., Capinha C., Martins M., Sousa C.A., 2020. Dengue importation into Europe: A network connectivity-based approach. *PLoS One*, 15 (3), e0230274.

Seixas G., Paul R.E.L., Pires B., Alves G., Jesus A. de, Silva A.C., Devine G.J., Sousa C.A., 2019. An evaluation of efficacy of the auto-dissemination technique as a tool for *Aedes aegypti* control in Madeira, Portugal. *Parasit Vectors*, 12 (1), 202.

Sheppard P., Macdonald W., Tonn R., Grab B., 1969. The dynamics of an adult population of *Aedes aegypti* in relation to dengue haemorrhagic fever in Bangkok. *Journal of Animal Ecology*, 38, 661-702.

Treuil J.-P., Zucker J.-D., Drogoul A., 2008. *Modélisation et simulation à base d'agents - Exemples commentés, outils informatiques et questions théoriques*, Paris, Dunod, 352 p. (coll. Sciences Sup).

Zellweger R.M., Cano J., Mangeas M., Taglioni F., Mercier A., Despinoy M., Menkes C.E., Dupont-Rouzeyrol M., Nikolay B., Teurlai M., 2017. Socioeconomic and environmental determinants of dengue transmission in an urban setting: An ecological study in Noumea, New Caledonia. *PLOS Neglected Tropical Diseases*, 11 (4), e0005471.

General conclusion and perspectives

Thierry Baldet, H el ene Guis

Mosquito-borne diseases, such as malaria, continue to represent a significant challenge for healthcare systems and development in many tropical countries, particularly in Sub-Saharan Africa. Additionally, in recent decades, a multitude of mosquito-borne diseases, including dengue, chikungunya, Zika, West Nile fever and Rift Valley fever, have emerged or re-emerged globally, gaining new ground and causing harm to both animal and human health. The persistence or emergence of these vector-borne diseases (VBDs) depends on numerous climatic, environmental and demographic factors specific to each region. These factors influence the way these diseases operate at varying levels. For two decades now, geomatics has been used to better understand the epidemiological processes of vector-borne diseases and to better predict their evolution in a world undergoing profound change.

This book presents a distinctive and contemporary overview of the potential applications of remote sensing and spatial modelling in the enhanced monitoring and control of infectious diseases whose pathogens are transmitted by mosquitoes. The presence, abundance and behaviour of vector mosquitoes, and by extension the epidemiology of the diseases they transmit, are complex phenomena which depend on a range of geographic, climatic and environmental factors. These include factors related to human demographics and activities, which shape the local expression of these diseases in space and time.

The combination of remote sensing and Geographic Information System (GIS)-based techniques with different modelling approaches offers significant advantages in the study of VBDs:

- mapping the distribution of larval habitats and adult mosquito population densities;
- evaluating the risk of vector exposure at different scales, for diseases already present or those that could be introduced;
- determining locations and seasons at high risk of transmission (hot spots) in order to optimise surveillance and control;
- determining the environmental, demographic and sociocultural factors which lead to a higher risk of infection for certain communities in order to implement preventive measures;

- taking into account anthropogenic changes and the environmental modifications that follow, in particular transformations of agricultural land and vegetation cover, urban sprawl, and industrialisation in order to predict short- and medium-term changes in the risk of exposure to vectors;
- assess the effectiveness and help optimise the various vector control methods, implemented singly or in combination, as part of an integrated control strategy.

Information derived from satellite data can facilitate a more comprehensive understanding of the risk of exposure to vectors, as discussed in Part one. This data is readily accessible to researchers and health service managers at a reduced cost. This information is integrated with other climatic, entomological, demographic and epidemiological data in GIS platforms, enabling the complex processes of vector-borne diseases to be understood. This approach enables all relevant determinants and their spatial dimensions to be taken into account.

The diverse modelling methodologies outlined in Part 2 are subsequently used to improve understanding of the mechanisms underlying VBD dynamics within a specific context. This facilitates more accurate forecasting of the disease's spatial and temporal evolution, incorporating considerations of vector control measures with environmental and climate alterations. The choice of approach is contingent upon the specific objectives, knowledge base, and information available for a given vector and VBD. Part 2 presents these diverse but often complementary approaches (species distribution models, knowledge-based models, mechanistic models, behavioural models) used to study a range of VBDs (*Anopheles*/malaria, *Aedes*/dengue models) in different contexts (French Guiana, Madagascar, La Réunion, Thailand) which enable useful operational tools to be developed for decision-makers and health service managers in order to optimise entomological surveillance and vector control.

In the context of an island territory such as La Réunion, which is subject to a range of vector exposure risks, recent successes demonstrate the potential of these modelling approaches to serve as a toolbox for achieving a desired goal. The spatial and seasonal characterisation of the risk related to flea-borne typhus, bluetongue and epizootic haemorrhagic disease⁵¹ enables the implementation of targeted awareness campaigns for health professionals and livestock farmers (Grimaud *et al.*, 2021, 2019; Tran *et al.*, 2021). Predictive mapping tools for vector densities and the risk of dengue transmission are fitting in the context and have been used for several years by the agency in charge of vector control (ARS La Réunion) in order to optimise entomological surveillance and preventive and control measures (Tran *et al.*, 2020). Modelling tools have been developed and made available to evaluate and optimise vector control methods, either alone or in combination as part of an integrated control strategy (Douchet *et al.*, 2021). The aforementioned successes in La Réunion have been facilitated by the accessibility and depth of expert knowledge, as well as the availability of experimental and observational data. Furthermore, the proximity of researchers from a multitude of disciplines, working in close collaboration with the beneficiaries of this research, has proved invaluable. Such examples include the ARS La Réunion as part of

51. Bluetongue disease and epizootic haemorrhagic disease are two viral animal diseases which regularly afflict Reunion island. These are known locally as "bavites" (derived from the French *baver*, meaning to drool, due to the excessive salivation they cause) and are responsible for major economic losses in the region.

the One Health Network – Indian Ocean⁵², a Platform in partnership for research and training (*Dispositif de recherche et d'enseignement en Partenariat*). The flexibility and adaptability of these approaches also permit their application to other contexts that are less well documented, as evidenced by the work performed on the neighbouring island of Madagascar cited in this publication (Rakotoarison *et al.*, 2020).

As we have seen in the example of Reunion island, the convergence of factors such as the availability of multi-temporal satellite data and entomological and epidemiological data, collaboration between remote sensing experts, modellers and biologists, and the availability of appropriate statistical systems and image processing algorithms creates a fertile research environment. The combination of upstream beneficiaries of research projects and a genuine commitment to transfer the results of these projects through the joint development of operational management tools promote their uptake and effective use in surveillance and vector control programmes. These remote sensing and spatial modelling approaches are relatively generic and can be applied to multiple areas, whether these are vector-borne zoonotic diseases (Rift Valley fever, West Nile fever, flea-borne typhus), human diseases (malaria, dengue) or animal diseases (bluetongue, epizootic haemorrhagic disease), and naturally combine environmental and human sciences. These are the pillars of the “One Health”⁵³ initiative, which include interdisciplinary and intersectoral collaboration.

The application of geospatial technologies to the field of VBDs has seen rapid growth in recent decades. As shown in this publication, a wide range of powerful image processing software, GIS, statistical tools and modelling approaches is now available in an accessible office environment, allowing epidemiologists and biologists to experiment with new spatial analysis techniques. Although the studies discussed in this document demonstrate the effectiveness of remote sensing and other geospatial technologies for VBD control and surveillance, several aspects must be considered in order to see more widespread adoption of these technologies in public health services, particularly in the Global South. Among others, these include the availability of resources for the collection, processing and modelling of geospatial data, staff training in the acquisition and correct interpretation of results, the cost-efficiency of these surveillance techniques, and the continuous and timely availability of remote sensing data. It should also be highlighted that assigning resources to these new technologies should not be to the detriment of other core disease prevention and management activities at the community level.

In order to enhance the capabilities of geospatial technologies applied to the field of VBDs and form interdisciplinary and intersectoral partnerships, the importance of training must be emphasised: postgraduate training (masters, PhDs) for students will help to achieve a critical mass of skills in this field, particularly in countries of the Global South. The incorporation of specific modules on how to use these approaches in the basic training of future stakeholders/decision-makers and the organisation of professional training courses will enable advances in research on Earth observation

52. <https://www.onehealth-oi.org/>

53. The *One Health* initiative is supported by the WHO, the World Organisation for Animal Health (WOAH), the United Nations Food and Agriculture Organization (FAO) and the United Nations Environment Programme (UNEP).

data and models to be shared amongst the fields of human health, animal health and ecology, by highlighting their potential and limitations. The recent growth of online training courses is increasing the uptake of remote sensing and modelling products and methods.

There are a number of opportunities for using spatial information, in particular Earth observation images, to better understand, predict and prevent the transmission of VBDs. Firstly, despite the exponential growth in the application of remote sensing and spatial modelling techniques, these have yet to be applied to many VBDs in a significant number of regions. The accuracy of predictive maps for these diseases must also be verified in the field: there will always be a need for specialists in entomology and epidemiology. In addition to biological approaches, social and behavioural patterns, such as the time spent outdoors by individuals, which increases the risk of exposure to anthropophilic vectors, the types of house construction, the use of mosquito nets and repellents, as well as the availability of basic sanitation facilities and primary health care, which are related to socio-economic conditions, are important for the prevention and control of vector-borne diseases. The characteristics of these diseases have been directly linked to poverty and social inequality. These socio-anthropological and entomological factors cannot be deduced from remote sensing techniques alone.

One interesting line of research opened up by modelling is the possibility of testing future change scenarios and studying their potential impact on health risks. For example, the impact of climate change on the distribution and abundance of vector populations, as well as pathogen transmission risk, can be studied via modelling (Guis *et al.*, 2012; Kraemer *et al.*, 2019). Most of the models presented in Part 2 integrate temperature and precipitation variables. Scenario-based projections of these variables (temperature increases, changes in rainfall patterns) produced by the Intergovernmental Panel on Climate Change (IPCC) can be used as inputs for these models to simulate their effects *in silico*. Furthermore, the impact of vector control measures can also be tested by models (Douchet *et al.*, 2021; Haramboure *et al.*, 2020) and discussed with vector control stakeholders to optimise their efforts, particularly as part of the implementation of integrated control strategies. Indeed, this is one of the main features requested of the Arbocarto tool by vector control agencies in France (see Chapter 8). Lastly, other types of scenarios can be tested, such as the impact of demographic changes (population growth, human mobility), as well as environmental changes. In the latter example, the benefit of modelling is to provide decision-makers with a map-based visualisation of the impact of planning and mitigation measures, which can help in the decision-making process by integrating public health issues, in particular the prevention of epidemics.

Finally, although the spatial data and methods presented in this publication focus on mosquito-borne diseases, the approaches discussed (data extraction from satellite imagery, modelling of species distribution, population dynamics and mobility) are generic in nature. They can therefore be adapted to other questions and challenges, such as biodiversity (and studying their links to health), as well as food security challenges (with common questions related to plant disease epidemiology, as well as the distribution and dynamics of crop pests). These questions are part of the “One Health” approach, which, in addition to addressing human and animal health, must also integrate plant and ecosystem health.

►► References

- Douchet L., Haramboure M., Baldet T., L'Ambert G., Damiens D., Gouagna L.C., Bouyer J., Labbé P., Tran A., 2021. Comparing sterile male releases and other methods for integrated control of the tiger mosquito in temperate and tropical climates. *Scientific Reports*, 11 (1), 7354.
- Grimaud Y., Tran A., Benkimoun S., Boucher F., Esnault O., Cetre-Sossah C., Cardinale E., Garros C., Guis H., 2021. Spatio-temporal modelling of *Culicoides* Latreille (Diptera: Ceratopogonidae) populations on Reunion Island (Indian Ocean). *Parasit Vectors*, 14 (1), 288.
- Grimaud Y., Guis H., Chiroleu F., Boucher F., Tran A., Rakotoarivony I., Duhayon M., Cetre-Sossah C., Esnault O., Cardinale E., Garros C., 2019. Modelling temporal dynamics of *Culicoides* Latreille (Diptera: Ceratopogonidae) populations on Reunion Island (Indian Ocean), vectors of viruses of veterinary importance. *Parasit Vectors*, 12 (1), 562.
- Guis H., Caminade C., Calvete C., Morse A.P., Tran A., Baylis M., 2012. Modelling the effects of past and future climate on the risk of bluetongue emergence in Europe. *Journal of the Royal Society Interface*, 9 (67), 339-350.
- Haramboure M., Labbé P., Baldet T., Damiens D., Gouagna L.C., Bouyer J., Tran A., 2020. Modelling the control of *Aedes albopictus* mosquitoes based on sterile males release techniques in a tropical environment. *Ecological Modelling*, 424, 109002.
- Kraemer M.U.G., Reiner R.C. Jr., Brady O.J., Messina J.P., Gilbert M., Pigott D.M., Yi D., Johnson K., Earl L., Marczak L.B., Shirude S., Davis Weaver N., Bisanzio D., Perkins T.A., Lai S., Lu X., Jones P., Coelho G.E., Carvalho R.G., Van Bortel W., Marsboom C., Hendrickx G., Schaffner F., Moore C.G., Nax H.H., Bengtsson L., Wetter E., Tatem A.J., Brownstein J.S., Smith D.L., Lambrechts L., Cauchemez S., Linard C., Faria N.R., Pybus O.G., Scott T.W., Liu Q., Yu H., Wint G.R.W., Hay S.I., Golding N., 2019. Past and future spread of the arbovirus vectors *Aedes aegypti* and *Aedes albopictus*. *Nature Microbiology*, 4 (5), 854-863.
- Rakotoarison H.A., Rasamimalala M., Rakotondramanga J.M., Ramiranirina B., Franchard T., Kapesa L., Razafindrakoto J., Guis H., Tantely L.M., Girod R., Rakotoniaina S., Baril L., Piola P., Rakotomanana F., 2020. Remote Sensing and Multi-Criteria Evaluation for Malaria Risk Mapping to Support Indoor Residual Spraying Prioritization in the Central Highlands of Madagascar. *Remote Sensing*, 12 (10), 1585.
- Tran A., Mangeas M., Demarchi M., Roux E., Degenne P., Haramboure M., Le Goff G., Damiens D., Gouagna L.C., Herbreteau V., Dehecq J.-S., 2020. Complementarity of empirical and process-based approaches to modelling mosquito population dynamics with *Aedes albopictus* as an example--Application to the development of an operational mapping tool of vector populations. *PLoS One*, 15 (1), e0227407.
- Tran A., Le Minter G., Balleydier E., Etheves A., Laval M., Boucher F., Guernier V., Lagadec E., Mavingui P., Cardinale E., Tortosa P., 2021. Describing fine spatiotemporal dynamics of rat fleas in an insular ecosystem enlightens abiotic drivers of murine typhus incidence in humans. *PLOS Neglected Tropical Diseases*, 15 (2), e0009029.

Acknowledgements

This publication presents a summary of the work carried out in the framework of the ANISETTE project (Inter-Site Analysis: Evaluation of Remote Sensing as a predictive tool for the surveillance and control of diseases caused by mosquito). The coordinators and authors would like to thank CNES (French National Space Agency) which financed the ANISETTE project from 2019-2022 in the framework of the Tosca programme “Land surface” hub).

The coordinators would also like to extend a special thanks to Didier Fontenille for agreeing to write the foreword to this publication, Yvette Vaguet for revision, as well as Thierry Baldet and H el ene Guis for their significant contributions to revision and corrections. We would also like to express our gratitude to all of the authors who participated in this collaborative endeavour and for their promptness in providing feedback on the final manuscript.

The authors would also like to acknowledge the contributions of the Mediterranean Interdepartmental Agreement for Mosquito Control (EID M editerran ee), the Joint Research Unit for Infectious Diseases and Vectors: Ecology, Genetics, Evolution, and Control (UMR MIVEGEC⁵⁴), the Key Initiative for Infectious Risks & Vectors in Occitanie (D efi Cl e RIVOC⁵⁵) and the Key Initiative MUSE Risks & Vectors (KIM RIVE⁵⁶) for their their invaluable input and support in the form of formal and informal discussions, as well as their commitment to continuously improving the overall knowledge on vector mosquitoes, particularly with regard to the content of Chapter 1.

The work presented in Chapter 2 was carried out in the context of different projects: the European Union EDEN project, Emerging Diseases in a changing European Environment (GOCE-CT-2003-010284 EDEN, 2005-2010) and the MoBoD project, Mosquito-borne diseases determinants, funded by the European Centre for Disease Prevention and Control, ECDC (OJ/2012/02/16 – PROC/2 012/015). The tools Sen2Chain and Sen2Extract were developed as part of the S2-Malaria Project thanks to the support of the CNES Tosca program (2017-2020) and the RenovRisk Interreg project (2018-2020). We want to also thank everyone who contributed to these developments at the joint research unit UMR Espace-Dev: Christophe R evillion, Pascal Mouquet, Didier Bouche, J er emy Commins, Sylvaine J ego et Charlotte Wolff.

Chapters 3 and 4 were completed in the framework of the MO³ project, funded by the French National Research Agency, ANR (ANR-19-CE03-0004-01). The authors would like to extend a special thanks to Rick Paul (Institut Pasteur de Paris) and Yvette Vaguet (UMR IDEES, CNRS – University of Rouen).

54. <https://mivegec.fr>

55. <https://rivoc.edu.umontpellier.fr/>

56. <https://www.umontpellier.fr/en/universite/projets-emblematiques/programme-dexcellence-i-site>

Chapter 5 summarises the main results (development of the FOTOTEX processing chain and application to the study of dengue in Brazil) of the Apureza and Deliciosa projects funded by the CNES Tosca programme from 2017-2023. The study on the La Réunion was funded as part of the ANISETTE project and benefited from partnerships with members of the Platform in partnership for research and training, One Health – Indian Ocean.

The work presented in Chapter 6 was funded by the European Social Fund (ESF), the French National Space Agency (CNES), the territorial collectivity of Guiana (CTG), as well as the future investment programme coordinated by the French National Research Agency (ANR) through the Centre for the Study of Biodiversity in Amazonia, CEBA (Ceba, ANR-10-LABX-0025), and the Franco-Brazilian Research and Scientific Cooperation Programme (IRD, CIRAD, Embassy of France in Brazil, territorial collectivity of Guiana, and research foundations from the Brazilian states of Amapá, Amazonas and Maranhão) via the GAPAM-Sentinela project. The authors would like to especially thank the members of the Institut Pasteur de la Guyane, particularly Benoît de Thoisy, Antoine Adde, Romain Girod and Isabelle Dusfour, as well as Sébastien Briolant from the French Armed Forces Biomedical Research Institute (IRBA) and Frédérique Seyler from the French National Research Institute for Sustainable Development (IRD/Espace-Dev) for their decisive contributions.

The work discussed in Chapter 7 was carried out as part of the Surveillance and data management project No. AID-687-G-13-00003, funded by USAID (United States Agency for International Development).

The authors of Chapter 8 would like to acknowledge all of the partners who helped fund, develop and operationally deploy the Arbocarto tool: the French Directorate General for Health (DGS, Albert Godal), the French National Space Agency (CNES) and the Space Climate Observatory (SCO), the National Health Agencies (ARS) of Occitanie, Auvergne Rhône-Alpes, Nouvelle-Aquitaine, Réunion and Mayotte, members of the Mediterranean and Rhône-Alpes Interdepartmental Agreements for Mosquito Control (EID), the Departmental Council of Charente-Maritime and the territorial collectivity of Martinique.

Chapter 9 was completed in the framework of the MO³ project, funded by the National Research Agency's (ANR) MO³ project (ANR-19-CE03-0004-01). The authors would also like to extend their gratitude to Renaud Misslin (UMR LAE, INRAE – University of Lorraine) and Somsakun Maneerat (UMR IDEES, CNRS – University of Rouen) for the invaluable contributions.

Glossary

Pathogen: an agent (virus, bacteria, protozoa, nematode, etc.) capable of causing injury or disease in animals (including humans) or plants.

Arbovirus: contraction of the term arthropod-borne virus, a virus transmitted by arthropods. Arbovirosis is a disease caused by an arbovirus.

Arthropod: group of animals with a segmented body and covered by an exoskeleton, such as ticks and mosquitoes.

Spatio-temporal autocorrelation: statistical measurement which answers the following question: do nearby spatial entities change over time in the same way as entities which are further away?

Bias: The statistical bias of an estimator is the difference between this estimator's expected value and the true value of the parameter being estimated. The term "bias" also encompasses the set of procedures and approaches which lead to statistical bias, i.e., the errors in the results of an analysis or model.

Biotope: ecological space of varying size which provides viable conditions, which are generally fairly homogeneous and constant, to all of the living organisms coexisting in this space.

Index case: in epidemiology, this term refers to the first person of an epidemic (or epidemiological cluster) to be infected by a pathogen.

Cluster (aggregate): group of objects linked in space and time. In epidemiology, a disease cluster refers to a health event connecting individuals.

Vector competence: quantitative estimate of a vector's ability to transmit a pathogen under natural conditions.

Diapause: an organism's physiological state of rest, determined genetically, which decreases metabolic rate. Often occurring on a seasonal basis, this state can be induced by photoperiod (Duvall *et al.*, 2017, see references in Chapter 1).

Dormancy: period of minimal activity without growth or development, minimising risks related to unfavourable environmental conditions and determined by the environment itself (Duvall *et al.*, 2017, see references in Chapter 1).

Expected value: the expected value of a random variable is a numerical value representing the average outcome of a random experience. This is equivalent in probability to the mean of a statistical series in statistics.

Aestivation: period during which certain arthropods reduce their metabolism and draw from fat reserves due to unfavourable environmental conditions (temperature, humidity). Depending on the season in question, this may also refer to hibernation.

Eurygamous: refers to a species where mating occurs during flight over a large area; breeding in a cage is therefore difficult, or even impossible. Conversely, **stenogamous** refers to a species where mating takes place in a confined area.

Larval habitat: area in which the aquatic stages of mosquitoes develop and also associated with the egg-laying sites chosen by females. Larval habitats vary greatly depending on the species and their ecological needs during the aquatic stages: fresh or brackish water, exposed to sunlight or sheltered, permanent or temporary, containing plant matter or clear water, etc.

Gravity model: in geography, models of spatial interactions which aim to describe the relationships between places. Gravity models enables the strength of these relationships to be characterised between geographic cells by considering both their potential (population, resources, etc.) and their distance.

Nulliparous: refers to a female which has never laid eggs, opposite of “parous”.

Oviposition: (from the Latin *ovum*, meaning egg), act of laying eggs at a site chosen by a gravid female (related to larval ecology), which allows them to develop within the environment.

Extrinsic incubation period: time required for the development of the ingested virus in order to render the vector infectious when taking its next blood meal.

Raster: a raster is a grid, also referred to as a matrix, organised into columns and rows. Each cell of this grid is a single pixel with assigned values. The distribution of all of these pixels in the grid (row *i*, column *j*) creates an image (generally in “.tif” format).

Residuals: In a statistical model, the residuals correspond to the observational data that is not explained by the model. In a regression model, the residuals correspond to the difference between the values predicted by the model and those observed.

Vector (biology): refers to an organism which transmits an infectious agent. In particular, these may include a haematophagous arthropod responsible for the active biological (or mechanical) transmission of a pathogen from one vertebrate to another.

Vecteur (geomatics): Vector data is used to represent real-world features in a GIS A vector feature can be defined according to its geometry type, which may be a point, line or polygon. Each vector feature is accompanied by a set of attribute data that describes it. Feature geometry is described in terms of vertices⁵⁷.

57. https://docs.qgis.org/3.34/en/docs/gentle_gis_introduction/vector_data.html#what-have-we-learned

List of acronyms

AES	Entomological surveillance area
AHP	Analytic Hierarchy Process
ARS	Agence régionale de santé (French regional health agency)
ASF	African swine fever
BI	Brightness index
CIRAD	Centre de coopération internationale en recherche agronomique pour le développement (French agricultural research and international cooperation organisation)
CNES	Centre national d'études spatiales (French national space agency)
CR	Consistency ratio
DDT	Dichlorodiphenyltrichloroethane (insecticide)
DEM	Digital elevation model
DGS	Direction générale de la santé (Directorate-general of health)
DLR	Deutsches Zentrum für Luft- und Raumfahrt (German national space agency)
ECOSTRESS	Ecosystem Spaceborne Thermal Radiometer Experiment on Space Station
EID	Entente interdépartementale pour la démoustication (Interdepartmental agreement for mosquito control)
ESA	European Spatial Agency
ESRI	Environmental Systems Research Institute
ETM+	Enhanced Thematic Mapper Plus
EVI	Enhanced Vegetation Index
FAO	Food and Agriculture Organization
FFT	Fast Fourier Transform
FOTO	Fourier-based Textural Ordination
GBIF	Global Biodiversity Information Facility
GHSL	Global Human Settlement Layer
GIS	Geographic information system
GIS-MCA	Geographic Information System-based multi-criteria analysis
GLCM	Grey Level Co-occurrence Matrix
GLM	Generalized linear model
GPM	Global Precipitation Measurement
GSMaP	Global Satellite Mapping of Precipitation
HMIS	Health management information system
HPC	High Performance Computing
HR	High (spatial) resolution
IGN	Institut national de l'information géographique et forestière (French National Geographic Institute)

Remote Sensing and Spatial Modelling

IMERG	Integrated Multi-satellite Retrievals for GPM
INRAE	Institut national de recherche pour l'agriculture, l'alimentation et l'environnement (French National Research Institute for Agriculture, Food and Environment)
INSEE	Institut national de la statistique et des études économiques (French National Institute of Statistics and Economic Studies)
IPCC	Intergovernmental Panel on Climate Change
IPF Iterative	Proportional Fitting
IPM	Institut Pasteur de Madagascar (Pasteur Institute of Madagascar)
IRD	Institut de recherche pour le développement (French National Research Institute for Sustainable Development)
IRIS	Îlots regroupés pour l'information statistique (fundamental unit for dissemination of infra-municipal data used in France)
IRS	Indoor residual spraying
JRC	Joint Research Center
JRE	Java Runtime Environment
LMI	Laboratoire mixte international (International Joint Laboratory of the IRD)
LST	Land Surface Temperature
MAE	Mean Absolute Error
MAS	Multi-agent system
MCA	Multi-criteria analysis
MCH	Malagasy Central Highlands
MIR	Mid-infrared
MNDWI	Modified Normalized Difference Water Index
MODIS	Moderate-Resolution Imaging Spectroradiometer
NASA	National Aeronautics and Space Administration
NDBI	Normalized Difference Built-up Index
NDVI	Normalized Difference Vegetation Index
NDWI	Normalized Difference Water Index
NIR	Near-infrared
NMCP	National malaria control programme
ODE	Ordinary differential equations
OLI	Operational Land Imager
OSM	OpenStreetMap
PCA	Principal component analysis
PEPS	Plateforme d'exploitation des produits Sentinel (Sentinel products exploitation platform)
RGB	Red, green, blue
SAVI	Soil Adjusted Vegetation Index
SCO	Space Climate Observatory
SEAS-OI	Satellite-Assisted Environmental Monitoring of the Indian Ocean (Research centre for data-reception and remote-sensing centre of expertise)
SEC	Scientific Expertise Centre (of the Theia hub)
SIT	Sterile insect technique
SRTM	Shuttle Radar Topography Mission
SVM	Support Vector Machine

SWIR	Short Wave Infrared
TIRS	Thermal Infrared Sensor
TRMM	Tropical Rainfall Measuring Mission
TVX	Temperature-Vegetation Index
UMR	Unité mixte de recherche (Joint research unit)
UNEP	United Nations Environment Programme
USGS	United States Geological Survey
VBD	Vector-borne disease
VC	Vector control
VHSR	Very high (spatial) resolution
WHO	World Health Organization
WHO	World Health Organization
WMS	Web Map Service
WNF	West Nile fever
WNV	West Nile virus

Authors

Thierry Baldet is researcher in medical entomology specialising in emergent vector-borne diseases at the joint research unit UMR ASTRE (Animal, Health, Territories, Risks, Ecosystems) of the French agricultural research and cooperation organisation, CIRAD. He has conducted research on filariasis, malaria and arboviruses in a number of African countries, including Benin, where he also played a role in the coordination of the international master's programme in medical and veterinary entomology. Based in La Réunion, since 2019, he has been engaged in research and training activities pertaining to vector risk prevention and control in the Indian Ocean. In 2020, he assumed the role of coordinating the One Health network in the region (<https://www.onehealth-oi.org/>).

Barbara Boufhal is currently enrolled in her second year of a master's programme in Geography, planning, environment and development at the University of Rouen. Her specialisation is in Territorial analysis in environment and health. She is engaged in an internship at the joint research unit UMR IDEES (Areas and Societies their Identities and Differences), where she is contributing to the detection of urban environments conducive to the development of *Aedes aegypti* and *Aedes albopictus* across two sites of study: Bangkok and the metropolitan area of Rouen.

Mathieu Castets is a researcher in spatial modelling and computing at CIRAD, where he is a member of the joint research unit UMR TETIS (Territories, environment, remote sensing and spatial data). He is currently based on Reunion island. His work focuses on developing methods for modelling spatial dynamics.

Thibault Catry is a remote sensing research engineer working at the IRD, where he is a member of the joint research unit UMR Espace-Dev. He is currently based at the the Remote Sensing Centre in Montpellier. His research is centred on the development of techniques for processing optical and radar satellite images in order to characterise the environment. The potential applications of this research include the fields of health, hydrology, and deforestation.

Alexandre Cebeillac is a postdoctoral researcher at the Institut Pasteur where he is a member of joint research unit UMR IDEES (Normandy University). His research is primarily concerned with the study of human mobility, with a particular focus on its role in dengue epidemics. This work is based on the collection, analysis and processing of large and varied datasets, ranging from geolocated digital traces to satellite images, contributing to the creation of multi-agent models.

Éric Daudé is a geographer and research director at the CNRS assigned to the UMR IDEES (Normandy University). His research is focused on the examination of the vulnerabilities of territories and populations exposed to a range of hazards (environmental and technological), and risk management (vector control) and crisis management strategies (mass evacuation of populations) based on the concept of spatial simulations and the analysis of potential disaster scenarios.

Marie Demarchi is an independent geomatics engineer and the managing director of the company of the same name. She is based at the Remote Sensing Centre in Montpellier and primarily works in partnership with research and training organisations to support them through different projects phases: spatial modelling and analysis; consulting, guiding and coordinating stakeholders and training; application and communication of research findings to territories. Her 25 years of experience in geomatics are applied to a range of fields, including land use planning, environmental management, agriculture, urban development and health. Marie Demarchi co-developed the Arbocarto application alongside CIRAD for the French Ministry of Health.

Nadine Dessay is a remote sensing research engineer at the IRD and a member of the UMR Espace-Dev, based at the Remote Sensing Centre in Montpellier. Her research focuses on the use of satellite images and developing tools to create spatially explicit indicators characterising the links between environmental variables and health.

Didier Fontenille is researcher on vector-borne diseases and vectors from IRD, Montpellier. He spent 22 years in tropical countries (Madagascar, Senegal, Cameroun and Cambodia), working on mosquito-borne diseases, particularly malaria, dengue, chikungunya, yellow fever, Japanese encephalitis, Rift Valley fever. He was director of the MIVEGEC research unit (IRD, CNRS, University of Montpellier) and of the Institut Pasteur du Cambodge.

Florence Fournet is a medical entomology researcher at the IRD and a member of the UMR MIVEGEC (Infectious Diseases and Vectors: Ecology, Genetics, Evolution, and Control). Her research interests lie in the relationships between urban environments and health, with a particular focus on vector-borne diseases. She has worked in West Africa, Côte d'Ivoire and Burkina Faso, developing geographical approaches to study the inequalities in exposure to malaria and dengue in urban environments.

Olivier Gillet is a doctoral student at the UMR IDEES (University of Rouen Normandy). His research focuses on the management of volcano risk in Guadeloupe, with a particular emphasis on mass evacuation strategies for the population. The objective is to explore the potential eruption scenarios using computer simulations in order to provide stakeholders engaged in crisis management with information pertaining to these diverse strategies.

Hélène Guis is a veterinary epidemiologist and researcher at CIRAD within the UMR ASTRE research group. She is currently stationed at the Pasteur Institute of Cambodia. Her research activities are centred on the epidemiology of vector-borne diseases in humans and animals, the distribution of vector arthropods, and the surveillance of arboviruses.

Nausicaa Habchi-Hanriot is a medical entomologist at the vector control service of the Regional Health Agency of La Réunion (ARS La Réunion). Her research activities are focused on the surveillance of local Culicidae and the investigation of innovative vector control approaches. Her research interests include the use of tools to complement traditional vector control and deploying them on Reunion island, specifically in the context of *Aedes albopictus* management.

Vincent Herbreteau is research fellow and health geographer at the IRD within the UMR Espace-Dev research group. He is currently stationed at the Pasteur Institute of Cambodia. His research is centred on the spatial analysis of diseases, ranging from the study of pathogen vector and host ecology, the exposure of human populations to these vectors and pathogens (social models, access to healthcare services, etc.), and the development of strategies disease transmission risk management.

Ophélie Hoarau completed an internship at CIRAD in 2021 as part of the second year of her master's programme in Natural risks and resources of tropical environments at the University of Reunion, specialising in Geomatics and remote sensing. This internship was conducted as part of the ANISETTE project. She worked on the extraction of landscape indicators from Earth observation images to estimate larval habitat distribution of *Aedes albopictus* in La Réunion.

Renaud Marti is a research engineer for INRAE and a member of the UMR TETIS research group at the Remote Sensing Centre in Montpellier. His research aims to produce a digital render of the environment, in the form of land cover and land use maps, landscape metrics and spatially explicit indicators linked to the ecological processes associated with the dynamics of reservoirs, vectors and pathogens in order to integrate this spatial information in spatially explicit models which predict high risk areas and periods of disease.

Yi Moua holds a doctorate from the University of French Guiana (Université de Guyane) and is currently a research engineer at the IRD, UMR Espace-Dev, working on the Regional Cooperation Project: Satellite Observation of the Guianas (Progysat), financed by the European Regional Development Fund (ERDF) of French Guiana and the Interreg Amazon Cooperation Programme (IACP). She is an expert in geomatics, species distribution modelling and information systems.

Benjamin Pillot is a modeller and research fellow at the IRD, UMR Espace-Dev. His research focuses on the development and integration of spatio-temporal models in the field of sustainability science, with particular emphasis on the creation of multidimensional scenarios for the energy transition and sustainable energy access for the Global South. His research interests include the modelling and processing of heterogeneous data in related fields, including health and vector-borne diseases.

Hobiniaina Anthonio Rakotoarison is a research engineer at the Pasteur Institute of Madagascar (IPM), Epidemiology and Clinical Research Unit (EPI-RC), and teaches at the Department of Earth and Environmental Sciences (EES) at the University of Antananarivo (Université d'Antananarivo) in Madagascar. His research interests lie in the application of geomatics in the field of public health, with a particular focus on the spatial modelling of vector-borne diseases.

Fanjasoa Rakotomanana is a research fellow at the Pasteur Institute of Madagascar, in the Epidemiology and Clinical Research Unit (EPI-RC). She is head of the Health and Geomatics group (SaGEO) and her research concerns spatial epidemiology, with a particular focus on vector-borne diseases such as Malaria in Madagascar.

Sébastien Rey-Coyrehourcq is an engineer at the UMR IDEES, specialising in the construction and exploration of multi-agent simulation models. Additionally, he is an expert in the creation and use of innovative tools and exploration methods for these simulations (OpenMOLE, MGO) backed by high-performance computing.

Emmanuel Roux is a researcher at the French National Research Institute for Sustainable Development (IRD) and a member of UMR Espace-Dev. His research involves the use of data sciences and models to study, monitor and control eco-epidemiological systems associated with vector-borne diseases, with a particular emphasis on malaria. Specifically, he is interested in cross-border regions along the international borders of the Brazilian Amazon and is a co-director of the Franco-Brazilian International Joint Laboratory (LMI) "Sentinela" (IRD, France; Oswaldo Cruz Foundation and University of Brasilia, Brazil).

Claire Teillet is a doctoral student in geography and remote sensing at the IRD and a member of UMR Espace-Dev. She is based at the Remote Sensing Centre in Montpellier. Her research is focused on the use of remote sensing to characterise the urban environments of *Aedes* mosquitoes and the identification of the risk of exposure for human populations. The objective of her thesis is to develop an approach to spatialising risk that optimises the use satellite data and which is not influenced by the availability of disease case data and vector data.

Annelise Tran is a geomatics researcher at CIRAD within the UMR TETIS and a research associate at the UMR ASTRE. She is currently based at the Remote Sensing Centre in Montpellier. Her research is focused on the development of remote sensing and spatial modelling methods, with a primary focus on applications in the field of health. She coordinates the Scientific Expertise Centre (SEC) "Risks associated with infectious diseases" in collaboration with Thibault Catry, in addition to directing various research initiatives relating to the use of remote sensing in the field health, including the ANISETTE project.

Cover photo: © frank29052515, stock.adobe.com; © IRD / Contains modified Copernicus Data

Traduction: Atenao
Infographics: H  l  ne Bonnet, Studio 9
Formatting: H  l  ne Bonnet, Studio 9
Legal deposit: June 2025

Mosquitoes are vectors of many disease-causing pathogens, including malaria, dengue, chikungunya, and yellow fever. According to the World Health Organization, these vector-borne diseases account for several hundred thousand deaths annually. They also cause zoonoses, such as Rift Valley fever and West Nile fever.

In this context, the development of operational tools to support surveillance and control strategies is essential—not only in countries of the Global South, where mosquito-borne diseases are most prevalent in tropical and subtropical regions, but also in the countries of the North, where the establishment of invasive species such as the tiger mosquito is increasing the risk of disease emergence. To address these challenges, Earth observation imagery offers valuable potential: the spatial distribution and seasonal dynamics of mosquito populations are closely linked to climatic factors (such as temperatures, rainfall and humidity) and environmental variables (such as the presence of water bodies and vegetation), many of which can be monitored through satellite data.

Numerous recent studies have led to the development of innovative methods that combine remote sensing with spatial modelling to predict the spatial and temporal dynamics of vector mosquitoes and associated diseases. Moving beyond proof-of-concept, some of these approaches have given rise to operational tools and processing chains that are now actively used by public health authorities and vector control agencies.

This book, intended for students, researchers, and public health professionals, offers a synthesis of current research and operational tools in the field.

Annelise Tran is a researcher at CIRAD, within the UMR TETIS in Montpellier. Her research focuses on the development of methods in remote sensing and spatial modelling with applications in the field of health.

Éric Daudé is a research director at the CNRS and deputy director of the UMR IDEES at Normandy University. His research explores the vulnerability of territories to natural, technological and environmental risks, and focuses on developing spatial modelling approaches to support crisis management and informed decision-making.

Thibault Catry is a research engineer in remote sensing at IRD, within the UMR Espace-dev in Montpellier. He develops processing methods for satellite imagery dedicated to the characterization of environmental dynamics, with a particular interest in the relationship between environment and health.

éditions
Quæ

Éditions Cirad, Ifremer, INRAE
www.quae.com



26 €

ISBN : 978-2-7592-4101-9



9 782759 241019

ISSN : 1773-7923

Réf. : 03016