

Review

Satellite Earth Observation Data in Epidemiological Modeling of Malaria, Dengue and West Nile Virus: A Scoping Review

Elisavet Parselia ^{1,*} , Charalampos Kontoes ¹, Alexia Tsouni ¹ , Christos Hadjichristodoulou ², Ioannis Kioutsioukis ³, Gkikas Magiorkinis ⁴ and Nikolaos I. Stilianakis ^{5,6}

¹ Institute for Space Applications and Remote Sensing, National Observatory of Athens, 15236 Athens, Greece

² Department of Hygiene and Epidemiology, Faculty of Medicine, University of Thessaly, 41500 Larissa, Greece

³ Department of Physics, University of Patras, 26504 Rio, Greece

⁴ Department of Hygiene, Epidemiology and Medical Statistics, Medical School, National and Kapodistrian University of Athens, 11527 Athens, Greece

⁵ Joint Research Centre (JRC), European Commission, 21027 Ispra VA, Italy

⁶ Department of Biometry and Epidemiology, University of Erlangen-Nuremberg, D-91054 Erlangen, Germany

* Correspondence: eparselia@noa.gr

Received: 29 May 2019; Accepted: 4 August 2019; Published: 9 August 2019



Abstract: Earth Observation (EO) data can be leveraged to estimate environmental variables that influence the transmission cycle of the pathogens that lead to mosquito-borne diseases (MBDs). The aim of this scoping review is to examine the state-of-the-art and identify knowledge gaps on the latest methods that used satellite EO data in their epidemiological models focusing on malaria, dengue and West Nile Virus (WNV). In total, 43 scientific papers met the inclusion criteria and were considered in this review. Researchers have examined a wide variety of methodologies ranging from statistical to machine learning algorithms. A number of studies used models and EO data that seemed promising and claimed to be easily replicated in different geographic contexts, enabling the realization of systems on regional and national scales. The need has emerged to leverage furthermore new powerful modeling approaches, like artificial intelligence and ensemble modeling and explore new and enhanced EO sensors towards the analysis of big satellite data, in order to develop accurate epidemiological models and contribute to the reduction of the burden of MBDs.

Keywords: mosquito-borne infectious diseases; Satellite Earth Observation data; epidemiological modeling; entomological data; vector-borne diseases; Earth Observation for health; malaria; dengue; West Nile Virus; scoping review

1. Introduction

Mosquito-Borne Diseases (MBDs) infect almost 700 million people every year and are recognized in over 100 countries affecting all continents apart from Antarctica and causing millions of deaths annually [1]. The burden of MBDs is estimated to be higher in tropical and subtropical areas, affecting disproportionately the poorest populations. Despite the fact that there have been global campaigns to eradicate MBDs [2], these diseases are re-emerging and even more emerging in countries where they were previously unknown. The reason for this may be manifold. The changing climatic and ecological conditions, global travel and trade, human behavior [3], as well as the rapid and unplanned urbanization [4], are key factors that influence the seasonal and geographic distribution of vectors' population and therefore the transmission of the pathogens.

Most of the environmental variables (geographical, climatological, and hydrological) that influence the transmission cycle of MBDs between pathogenic agents, vectors and intermediate hosts can be monitored efficiently from satellites that carry specific instruments capable of capturing these parameters frequently and on a global scale [5–7]. Kazansky et al. listed the various satellite sensors that can provide environmental data and could contribute as an input to a Malaria Early Warning System [8].

It is worth noting that between the years of 2014 until 2018, there has been a remarkable growth in EO satellites, counting in total approximately 700 satellites in space [9], operating advanced sensor payload providing enhanced spectral and spatial resolution with shorter revisit times and larger coverage, enabling improved earth monitoring at global level [10]. This growth of observations was followed by a substantial increase in the number of studies that exploit EO data in order to better understand the geographic distribution, abundance and dynamics of MBDs and the associated vectors and pathogens [11].

The scope of this paper is to review recent literature for identifying studies that utilized satellite EO data for epidemiological modeling of malaria, dengue and West Nile Virus (WNV). Epidemiological models and Early Warning Systems (EWSs) that utilize EO data have been used as tools for helping decision-makers to improve health system responses, take preventive measures in order to curtail the spread of MBDs and address the relevant priorities of the Sustainable Development Goals (SDGs) such as good health and well-being (SDG 3) and climate action (SDG 13) [12].

In this scoping review we solely focused on three MBDs, namely malaria, dengue and WNV. The exclusion of other MBDs serves the economy of the paper only. We believe that these choices are representative of a wide range of MBDs; malaria is most commonly spread by the *Anopheles* mosquito genus that is also responsible for the transmission of Lymphatic Filariasis. The *Aedes* genus can transmit Dengue fever, Chikungunya, Lymphatic Filariasis, Rift Valley Fever, Yellow Fever and Zika, while the *Culex* genus is responsible for transmitting Japanese Encephalitis, Lymphatic Filariasis and WNV. Therefore, the use of EO data in the epidemiology of malaria, dengue and WNV do not differ substantially from those we did not consider.

1.1. Malaria

Malaria is the most prevalent, life-threatening and costly parasitic infection worldwide, affecting over 100 countries and territories which live under the risk of malaria transmission [2]. The global campaign to eradicate malaria rolled out by the World Health Organization (WHO) led to important reductions in new malaria cases in endemic countries during 2000–2015. Despite this achievement there were still 219 million cases of malaria in 90 countries reaching 435,000 deaths in 2017 [13]. Research has been conducted to examine the effect of climatic conditions on malaria transmission [14,15] with ambient ground temperature and moisture affecting *Anopheles* mosquitoes population and the incubation period [16]. Furthermore, intense rainfall can reduce the larvae density by flushing first stage larvae [17].

1.2. Dengue

Dengue is a mosquito-borne viral infection and is endemic in many tropical and subtropical regions in the world [18]. Since 1970 dengue has rapidly spread and can be found in more than 100 countries in regions of the world including the Americas, Eastern Mediterranean, South-East Asian and Western Pacific [19]. Temperature affects the extrinsic incubation period of the *Aedes* mosquito [20] and rainfall is also one of the most important environmental factors that affects the vector's reproduction cycle. The vector abundance can be influenced by rainfall events by increasing the availability of mosquito juvenile habitats (e.g., containers in the patio with standing water) [21] and drought conditions can increase the larval habitat by increasing household water storage [22].

1.3. West Nile Virus

WNV was first identified in the West Nile district of Uganda in 1937 and was considered a low risk disease for humans and livestock species until the 1990s [23]. Since then WNV has rapidly spread across all continents except for Antarctica [24]. In nature, the WNV cycles between birds, which act as the principal hosts and mosquito vectors and transmit the virus to other birds. Humans, equines and other mammals act as incidental or dead-end hosts and are not involved in the transmission cycle. Most human infections are asymptomatic (around 80% of infected people) or can lead to mild symptoms like fever, headache, tiredness and body aches. More severe cases can cause neuroinvasive disease including meningitis, encephalitis, acute flaccid paralysis and death in people [23]. The transmission and geographic distribution of WNV is associated with the existence of both the avian reservoir host and mosquito-vector, which is affected by environmental (abiotic and biotic) and socio-economic conditions [25]. The involvement of birds (in addition to mosquitoes) and the potential of EO techniques that may contribute to understanding of movement of migratory birds make the disease of particular interest in this context. WNV transmission can thrive under favorable environmental conditions; *Culex pipiens* can transmit WNV efficiently at a temperature of 30° [26].

2. Material and Methods

2.1. Literature Search Strategy

This scoping review included epidemiological and entomological studies that utilized EO (climatic and environmental) data in mapping, modeling and forecasting of malaria, dengue and WNV. Whereas a systematic review uses systematic methods to critically appraise a focused research question, a scoping review comprehensively maps evidence across a broader research question using diverse sources [27]. Accordingly, in this scoping review we have sought to review recent literature for identifying the current state-of-the-art in epidemiological modeling of the MBDs using satellite EO data. The search was limited to peer-reviewed literature in English that was conducted during the period 1 January 2012 to 31 December 2018. This search phase was selected due to the fact that from 2012 onward the annual growth rate of publications relating to health and dealing with remote sensing has been steadily increasing, with malaria and dengue being the most frequent disease-specific keywords [11]. The Web of Science, PubMed and the Scopus databases were searched electronically to retrieve relevant literature and articles. Boolean operators combining multiple keywords salient to the research topic were queried in the abovementioned databases. The keywords were “Earth observation”, “Remote sensing”, “Satellite data”, “vector-borne disease*”, “mosquito-borne disease*”, “modeling”, (“NDVI” OR “NDWI” OR “EVI”; AND “malaria” OR “Dengue” OR “WNV”), “temperature”, “precipitation”, “malaria”, “West Nile Virus” and “dengue”. The results were combined using the Mendeley software, and duplicates were removed. The titles and abstracts were initially examined to determine the relevance of the articles. Thereafter, full texts were screened to ascertain if the selection criteria were met. Finally, the reference lists of the reviewed papers were scanned to gain additional literature. All the authors listed in this paper participated in each step of the selection procedure.

2.2. Inclusion and Exclusion Criteria

The scoping review was conducted adapting the Arksey and O’ Malley [28] and Levac et al. [29] methodological framework. This framework includes a transparent method for linking the purpose and the area of research [28] and uses an iterative team approach for the selection of the studies, while it includes a numerical summary and qualitative thematic analysis [29]. The selection criteria involve post hoc inclusion and exclusion criteria. To ensure consistency and eliminate studies that were out of the scope of this paper the authors discussed and agreed on the initial criteria at the beginning of the selection process with further refinements until the final selection.

The articles finally selected were:

1. Peer-reviewed articles published in English between 1 January 2012 and 31 December 2018.
2. Publications that integrated satellite EO derived climatic and environmental predictors for analyzing mosquito-borne epidemics. Studies that did not use satellite EO data or used solely in situ data were excluded from the review.
3. Studies on models that included disease incidence, prevalence and cases as variables, as well as studies that used entomological data as response variables.
4. Articles referring to the impact of (inter-annual) climate variability on pathogen transmission, excluding the ones using climatic scenarios. By climatic scenario we refer to studies that used future projections under different climate change scenarios. Therefore, we only focused on studies that utilized historical data and built knowledge from the past events.
5. Studies that used epidemiological models, making reference to the achieved level of accuracy rates. In contrast studies that did not bring any evidence or information on the accuracy of the used models were excluded.

3. State-of-the-Art Review

A total of 576 relevant articles were initially identified by the electronic search on Web of Science, Scopus, and PubMed during the period January 2012 to December 2018 (Figure 1). The 112 articles in Figure 1 refer to records found through the review of the references of the selected articles as well as articles that authors recommended. A total of 43 articles were finally selected as meeting the eligibility criteria for this scoping review; they are briefly described in Tables A2 and A3 in the Appendix A.

The majority of the 43 studies have analysed cases of malaria ($n = 20$), followed by dengue ($n = 15$), and WNV ($n = 8$). Figure 2 illustrates the number and geographic reference of the selected studies together with the disease of study.

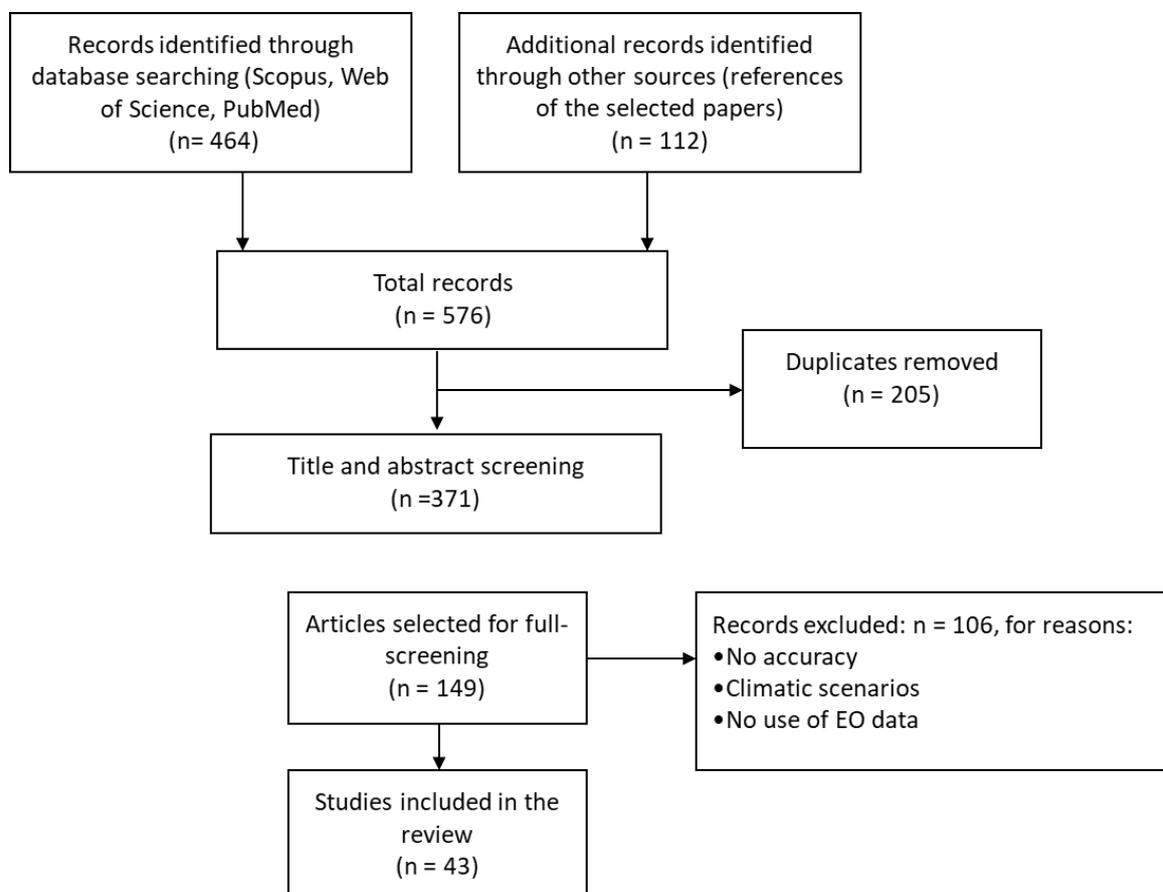


Figure 1. Flow Diagram of article selection (inclusion/exclusion) process.

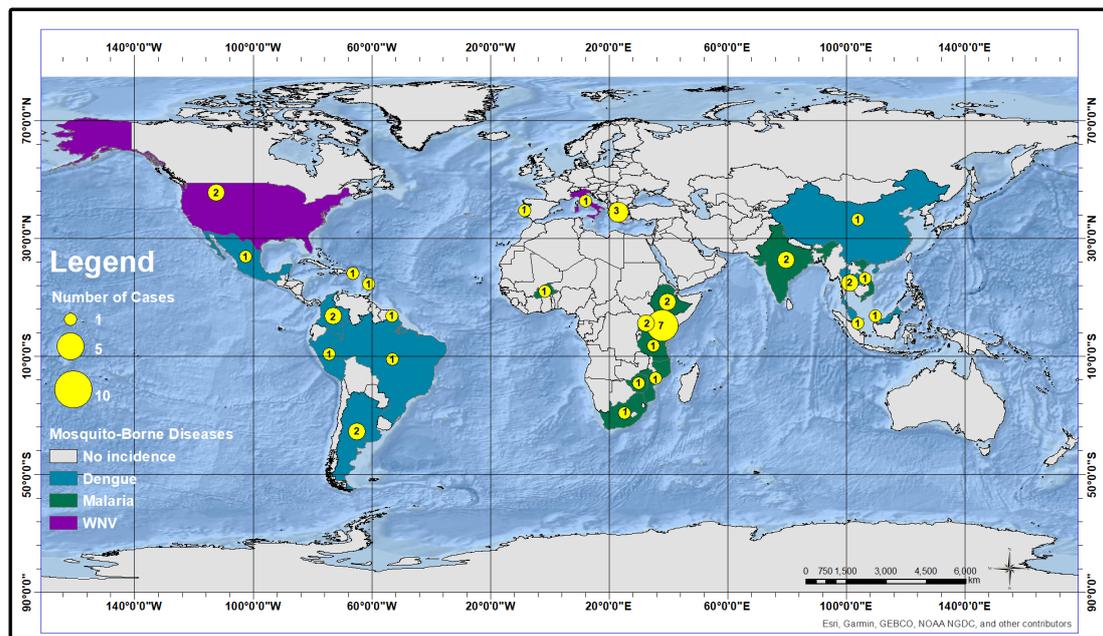


Figure 2. Distribution Map of the MBDs that were included in this review.

The selected articles were organized into two main categories (Figure 3) with respect to the data used as dependent variables for the prevalence of the diseases: (a) epidemiological data (disease incidence, prevalence or case, mortality data) ($n = 31$) and (b) entomological data ($n = 11$), while Stilianakis et al. has examined both (a) and (b) [30], and Valiakos et al. has additionally used wild bird data in complement to the epidemiological data [31]. The first category (a) used clinical records from the general human population as the main data source. In this case the majority of the studies ($n = 23$) referred to the clinical data as “confirmed cases”, meaning that the patients were confirmed through laboratory testing. Buczak et al. [32] and Arboleda et al. [33] included also cases that were considered as “possible”, meaning that the patients exhibited some of the symptoms of the infection. Refs. [31,34–39] explicitly used laboratory confirmed cases (microscopy/Rapid Diagnostic Tests (RDT)), while Sewe et al. utilized the number of deaths caused by malaria [40]. The second category (b) used entomological data providing information on the vectors’ density, that is highly dependent on the ambient climatic and environmental conditions and significantly influences the transmission of the pathogen. The mosquitoes’ collection was implemented by ovitraps, classical dipping techniques [41,42] or by recording indices, like the Breteau Index (BI) [33,43], House Index (HI) or Container Index (CI) [33,43,44], and the Entomological Inoculation Rate (EIR) [45,46] indices, the latter being a commonly used measure that estimates the number of infected bites per person and per unit time (usually year) [47]. Different kinds of traps (e.g., light traps, magnet traps baited with octenol, CO₂ baited traps, odour-baited MM-X) were used by the various studies [41,45,46,48–52] as a method to capture mosquitoes, and produce the vectors’ density assessments and its status as infected or not. It is worth noting that there is only one study [53] that used crowdsourced data, providing information on both the vector (% of mosquito bites, % of mosquito Larvae) and human cases (% of known human dengue cases) in the area. The share of the applied methodology with respect to the epidemiological or entomological data is illustrated in Figure 3.

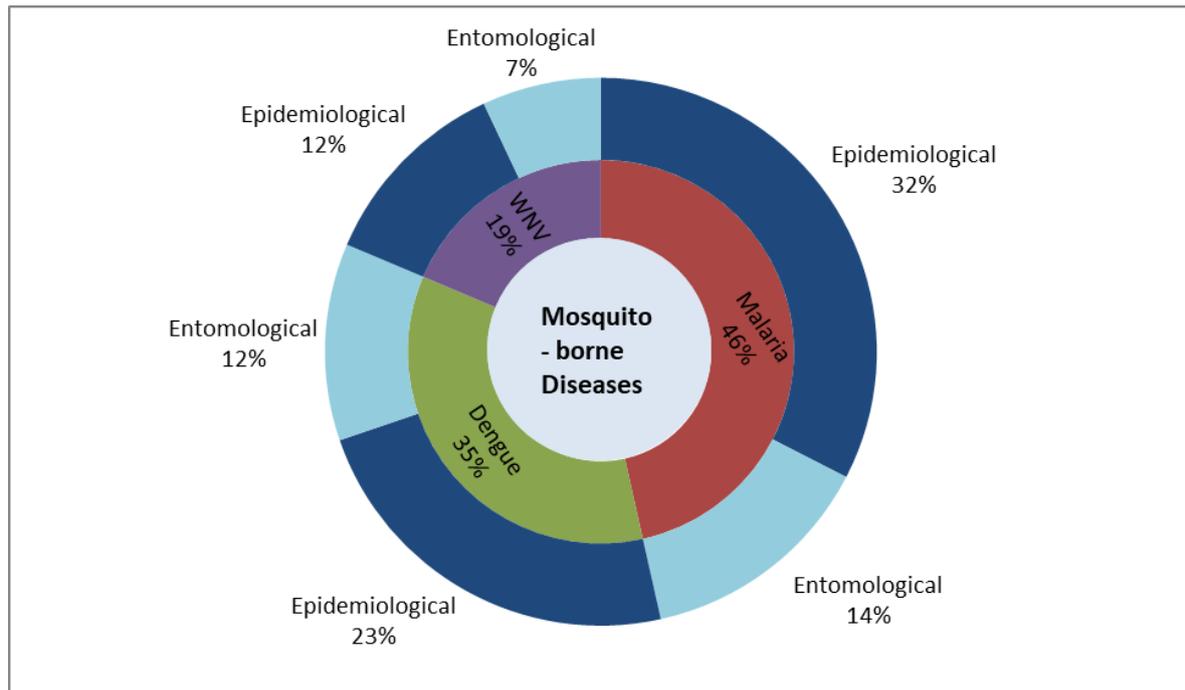


Figure 3. Share of the selected papers as per epidemiological/entomological data and MBDs.

Through our database search, mainly data-driven and statistical approaches were returned. Increased computational power as well as the number of open-source datasets over the past years has given rise to these data-driven models that can relate environmental variables with the species occurrence or abundance [54]. These models represent input-output relationships built upon available datasets and do not require detailed knowledge about the complex interactions of climate, vector, host and pathogen. Mechanistic models on the other hand aim at capturing the biological and environmental mechanisms using dynamic equations [55] in order to define causality. In order to capture the full dynamics of the system, in their majority these models, when the study area has small spatial extent, utilize weather data from ground stations; in situ measurements generally represent the input fluctuations with smaller error compared to the satellite-derived indirect products. This scoping review will focus predominately on data-driven and statistical models.

3.1. Environmental EO Predictors

The environmental EO based predictors that were leveraged by the studies examined in this review are listed in Figure 4. Among the various climatic and environmental variables that were examined as possible predictors in these models, the vast majority of the studies used air, land and soil temperature data ($n = 45$), precipitation ($n = 34$), and vegetation indices ($n = 42$) as listed in details in Figure 4. Many processes that are associated with mosquitoes are strongly influenced by temperature, as the rate of development of the virus inside the vector is linked to warmer temperatures [26]. Air temperature estimates were either indirectly linked to the remotely sensed Land Surface Temperature (LST), which is widely used as a proxy, or by collection of in situ observations. LST is the radiative skin temperature of the land surface and is an important climate variable that is estimated from Top-of-Atmosphere brightness temperatures from the infrared bands of the satellite's sensors [56]. Associating LST and air temperature is a complex task as proved in [57] since it is highly dependent on the geographic location of the study area. Only one study [30] reported the association of WNV infections to soil temperatures obtained from the ECMWF's (European Centre for Medium-Range Weather Forecasts) Re-Analysis (ERA-Interim) datasets. Soil temperature data refer to ground based observation at several depths [58]. Mendez-Lazaro et al. [59] and Laureano-Rosario et al. [60] have examined the influence of the Sea Surface Temperature (SST) instead, because of the vicinity to the coastal areas of San Juan in Puerto

Rico and Yucatan state of Mexico respectively. Likewise for the LST parameter, the results from these studies showed that also the SST was significantly associated with the reported dengue cases. Bhatt et al. developed a dengue-specific temperature suitability index based on a biological model with temperature as an input [61]. This index included two temperature-dependent values affecting the dengue transmission cycle: (i) the life duration of *Aedes* vector and (ii) the Extrinsic Incubation Period (EIP).

Heavy rainfalls have a negative effect on the development of *Anopheles gambiae*; flooding and flushing at the early stage larvae leads to high levels of larval mortality [17]. The same holds for the *Aedes* genus, as their reproductive cycle can be disrupted by extensive rainfall through flushing out the aquatic stages from breeding sites [62]. Precipitation satellite sensor derived data were mainly acquired from the Tropical Rainfall Measuring Mission (TRMM) ($n = 10$ [32,35,40,43,50,51,63–67]) ($n = 10$), while some studies used the WorldClim (<https://www.worldclim.org/>) ($n = 3$ [31,41,61]), ERA-Interim (<https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era-interim>) ($n = 1$ [30]) and Meteosat-7 ($n = 1$ [45]) and other sources such as local ground-stations.

Vegetation and vegetation indices are another important parameter that showed strong correlations with the vectors' behavior and their biological cycle. Most of the studies ($n = 26$) used the Normalized Difference Vegetation Index (NDVI), which is a proxy index of vegetation density and distribution due to the fact that is chlorophyll sensitive. NDVI is not only restricted to studies of plants; various studies have coupled vegetation dynamics with biodiversity, animal species distributions [68], movement patterns of animals (e.g., migratory birds) and the performance of animal populations (reproduction or survival). NDVI data can be used in combination with other data to model the temporal and spatial dynamics of vectors [69]. The Enhanced Vegetation Index (EVI) that is relevant to the canopy's structural variations [70] was also used ($n = 8$), followed by the Green Index (GI), the Soil Adjusted Vegetation Index (SAVI) ($n = 2$) and the quasi-yellowness index (p-YI) [34]. Ruangudomsakul et al. used a multitude of indices derived from the satellite based Global Vegetation Health System of NOAA such as (a) the Smoothed NDVI (SMN) to estimate the vegetation growing and senescence phases, (b) the Smoothed Brightness Temperature index (SMT), (c) the Temperature Condition index (TCI) for assessing the ambient thermal conditions, (d) the Vegetation Condition Index (VCI) that was used as a proxy for assessing the ambient moisture, and (e) the Vegetation Health Index (VHI) [71]. Vegetation information could also be derived from Land Use/Land Cover (LU/LC) maps in order to identify suitable vector breeding sites and was used by several studies ($n = 14$). LU/LC maps were also utilized for identifying other factors that might influence the transmission of MBDs like urban areas, health facilities, proximity to water bodies, etc.

Relative humidity plays an important role in the survival rate of the vectors, affecting differently various species [72]. Relative humidity estimates were derived from EO sensors systems onboard of satellites such as the Indian National Satellite System (INSAT)-3D imager [63], the Atmospheric Infrared Sounder (AIRS) instrument onboard the NASA's Aqua satellite [43] and the ERA-Interim reanalysis approach by Stilianakis et al. [30]. Adde et al. estimated the minimum and maximum relative humidity [48], while in Machault et al. the relative humidity was calculated using only in situ observations [44].

The evapotranspiration (ET) encompasses the amount of water that is removed from the land surface and returns to the atmosphere through the process of evaporation and transpiration [73]. Remote sensing techniques have been used by [74] in order to estimate the actual ET (ET_a), that is the quantity of the water actually removed from the surface of the earth. Positive ET_a is often associated with high levels of surface water and soil moisture availability, both of which indicate suitable vectors' breeding site conditions. In [65,75] the ET_a parameter was derived from MODIS sensor data products, in [66] it was used in the global ET from MOD16 product, while [48] used in situ data to assess the ET.

In the coastal city of San Juan, reference [59] has associated the dengue cases with the Sea Level Pressure (SLP) and the Mean Sea Level (MSL). MSL was one of the variables that was significantly associated with the human dengue cases. Higher dengue incidences were related to the MSL maxima

due to the fact that coastal areas are more prone to flooding during seasonal peak. Stilianakis et al. examined the soil water content as a parameter [30], which did not show an association with the presence of infected mosquitoes. Most of the studies ($n = 16$) referred to the role of the water bodies in the life circle of the mosquitoes because they serve as breeding sites for larval development. In order to depict the water bodies extent, several studies ($n = 9$) used the Normalized Difference Water Index (NDWI), which represents changes in liquid water content, while other studies ($n = 7$) additionally included as a parameter to the model the proximity to the different kinds of water bodies. Studies such as Diboulo et al. [46] and Giardina et al. [39] used the permanent or semi-permanent waters taking into account both natural and man-made containers. Other studies considered proximity to the running water, like rivers [76] or streams [31,77] and stagnant waters and lakes [45,78].

Topography is a significant factor in the transmission of MBDs as it affects the living conditions of the *Anopheles* and *Aedes aegypti* mosquitoes and indicates the best suited breeding sites [79,80]. Fourteen studies in this review [31–33,39,42,44,45,49,67,76–78,81,82] utilized Digital Elevation Models (DEMs) to extract the topographic parameters of elevation, aspect and slope. Moreover, DEMs were used for calculating the Topographic Wetness Index (TWI) [41,42,78], an index that gives information about the wetness of an area, taking into account the topographic slope and the upstream area. Nmor et al. specifically focused on topographic variables for the prediction of malaria vector breeding sites [42]. It was the only study that considered the association of topographic position index (TPI), curvature and Convergence index (CI) with malaria vector habitats.

Two studies [30,59] investigated the association of the MBDs with alternative climatic factors such as the wind speed. According to Stilianakis et al. [30] low wind speed was associated with the existence of infected mosquitoes, while high speed reduced the chances for blood meals, and consequently human infections.

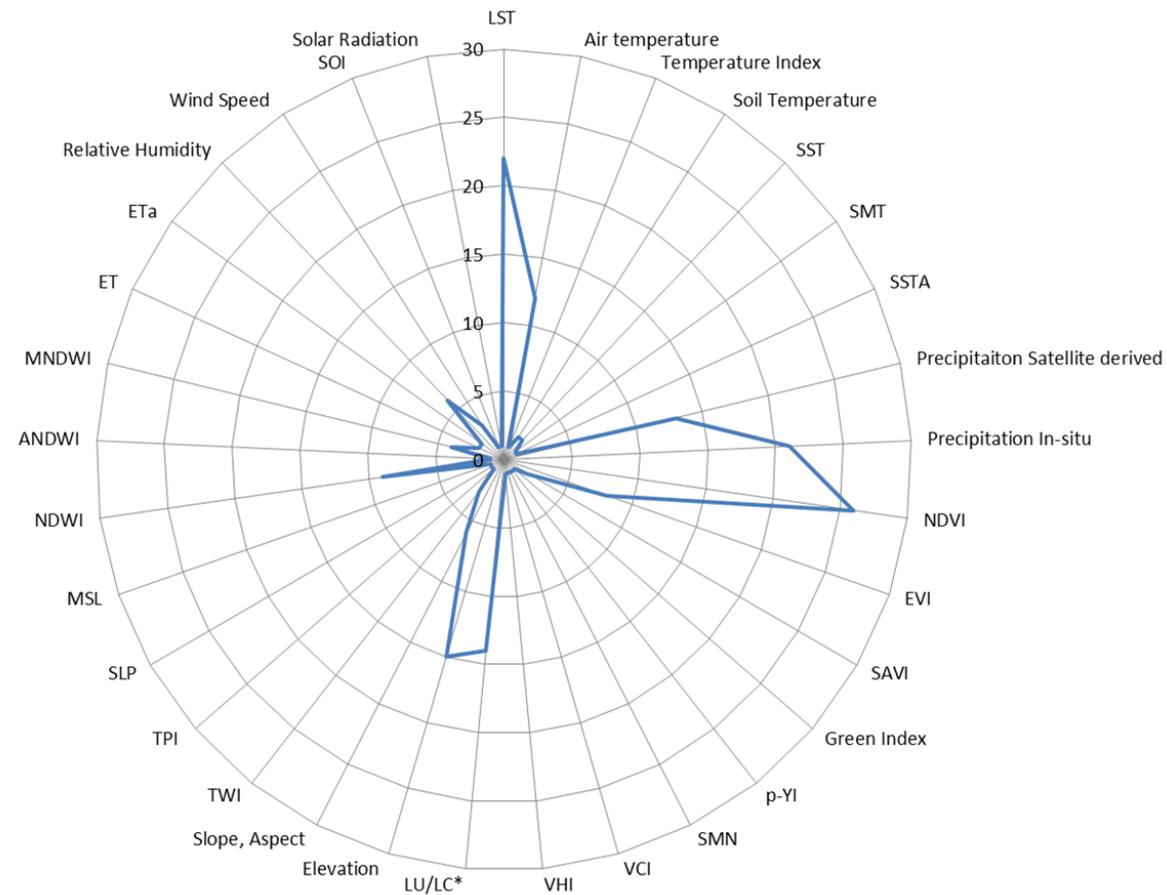


Figure 4. Overview of studies that utilized EO derived environmental/climatic variables:

LST [32,35,37–40,43,45,46,49–52,63–67,75,83–85], Air temperature [31,41,44,48,53,59,60,76,81,82,86,87], Temperature Index [61], Soil Temperature [30], SST [59,60], SMT [71], SSTA [32], Precipitation (Sat) [32,35,40,43,45,50,51,63–67,82], Precipitation In-situ [30,31,37–39,41,44,46,48,49,53,59–61,64,76,81,83,84,87], NDVI [31–34,37,38,40,44–46,50–52,61,65–67,75–78,81,83,84,88], EVI [32,35,41,52,63–65,67], SAVI [34,67], Green Index [34], p-YI [34], SMN [71], VCI [71], VHI [71], LU/LC* [31,37,39,43,44,48,49,52,53,64,76,77,83,85], Elevation [31–34,37–39,43–45,64,76,77,81,82], Slope & Aspect [31,33,34,42,76,77], TWI [41,42,78], TPI [42], SLP [59], MSL [59], NDWI [38,44,49–52,67,83,85], ANDWI [44], MNDWI [34,86], ET [48,66], ETa [65,75], Relative Humidity [30,43,44,48,63], Wind Speed [30,59], SOI [32], Solar Radiation [48]. * = vegetation types, proximity to water, urban areas, etc.

3.2. Other Non-Environmental Predictors

Although climatic parameters are highly influential in the transmission of MBDs, non-climatic factors such as social, economic and demographic parameters listed in Table 1 can affect the magnitude and the spatial extent of the MBDs' transmission [55]. The studies of Quintero et al. [89,90] examined cases of poor water and sanitation conditions that compel inhabitants to store water in open containers, thus building breeding habitats for the mosquitoes. In this review the study that was conducted by Buczak et al. also included variables related to socio-economic parameters, namely running water, hygienic services and electric lighting into the final model [32]. Moreover, Homan et al. highlighted the impact of socio-economic risk factors in malaria spread and constructed a socio-economic status index (SES) [78]. The SES was constructed using Principal Component Analysis (PCA) and was based on six variables: (a) rented or owned dwelling, (b) owned agricultural area, (c) highest education level of household, (d) location of the kitchen, (e) the wall structure and (f) the floor cover. Furthermore, Bhatt et al. included multiple socio-economic variables to generate dengue risk maps [61]; the urban accessibility data that define the travel time of people to a city using land and water—based mass transit mechanisms, relative poverty and demarcation of urban and peri-urban areas.

Furthermore, demographic data were used to estimate the rate of vulnerability of the population and identify the geographical areas where the risk of a disease outbreak is higher [91]. Areas with higher population density are at higher risk of transmitting a pathogen [90]. Several studies ($n = 5$) have used population data as independent variables [41,64,78,82,86], while others ($n = 2$) used the population data for estimating the vulnerability and the level of the risk [38,85]. In general the population density originated from National Administrative Departments. Marcantonio et al. that used Europe as its AOI, additionally used satellite imagery to extract the intensity of light at night that was used as a proxy for human population density [83].

Two studies that investigated the WNV risk factors, used data related to birds, since birds act as the principal host of the WNV transmission cycle. Tran et al. digitized and categorized the birds' migratory routes based on their fly way direction (western and eastern) [86], while Valiakos et al. utilized as a predictor, cases of birds that were positive to WNV antibodies [31].

It is believed that moon light affects the vectors' behavior, increasing its activity during full moon and third quarter moon phase. According to Mokraoui et al., there is a correlation between moon phases and dengue outbreaks [53]. Moon light affects the vectors' behavior, increasing its activity during full moon and third quarter moon phase. This was the only study that examined if there is any correlation between the moon light and dengue outbreaks.

Table 1. Overview of studies that utilized Non- EO Environmental Predictors.

Non-EO Environmental Predictors	Number of Studies (Reference)
Demographic data	
Population density	8 [38,41,64,78,82,83,85,86]
Socio-economic conditions	
Running water	1 [32]
Hygienic services	1 [32]
Electric lighting	1 [32]
Socio-economic status index (SES)	1 [78]
Other non-environmental data	
Birds	2 [31,86]
Moon phase	1 [53]

3.3. Satellite EO Systems Used for Assessing the Environmental Predictors

Satellite EO data were the main sources for assessing the environmental predictors. All studies used remotely sensed data to estimate environmental or climatic parameters, while some others incorporated in situ data additionally. In terms of data sources, the low resolution satellites

(spatial resolution of 300 m and less) were used as the major optical data sources. The MODIS based information products that were mostly used for the prediction of mosquito borne diseases refer to the LST, NDVI, EVI, and NDWI. Moreover, the TRMM mission was the most important source to timely acquire precipitation data, while Meteosat-7 data were also used by Amek et al. to estimate the rainfall at a 8 km wide cell [45]. Landsat-7 and Landsat-8, were both used to derive vegetation indices [33,76,77] as well as to generate LU/LC maps. Correspondingly Adde et al. [48] and Yue et al. [85] generated detailed LU/LC maps from using SPOT-5 and GF-1 data at the spatial resolution of 10 m and 16 m respectively. Machault et al. utilized very high resolution satellite imagery (GeoEye-1, 0.41 m) in order to create detailed LU/LC maps and downscale the study at the level of the household [44]. Similarly, Homan et al. used QuickBird imagery at the spatial resolution of 0.61 m for deriving precise information on household's proximity to lakes, and to the nearest clinic, but also for producing detailed estimations of the NDVI and TWI indices in the neighborhood of the households [78].

4. Results and Discussion

EO data were found to elicit environmental and climatic variables, that could significantly contribute to epidemiological modeling of MBDs referring to predictive mapping, geographic distribution and abundance of the pathogen and vectors, health risk assessment, understanding the transmission dynamics, identification, implementation of appropriate control strategies and their assessment. The associations between predictors and MBDs were either positive or negative as listed in Table 2. Furthermore, the study by Merkord et al. has successfully incorporated satellite EO data into the early warning EPIDEMIA System [67], while Lowe et al. highlighted the potential for integrating EO data into a EWS for Southern Brazil [82]. All studies have used satellite remotely sensed data, while some incorporated in situ data. Although the latter tended to be accurate sources of information, they have been of limited use in the literature, most likely because they were single point measurements and sparsely distributed observations, while the interpolation between points, necessary in case of larger study areas, added more uncertainty in the prediction. Contrariwise, according to the majority of the studies the satellite based observations provided large areas of coverage and uninterrupted acquisitions of series environmental data needed for the predictions. The primary satellite data sources that were exploited extensively from most of the studies in this review were medium to high resolution and freely available. It is worth noting that besides the unforeseen long lifespan of the two most commonly used satellites Terra and Aqua, they are expected to switch off operations in a few years, and there is a need for replacing them with similar satellite/sensor systems e.g., SUOMI-NPP/VIRS, JPSS/VIRS, Sentinel-2, Sentinel-3, which have not yet been fully exploited. From the analysis it is obvious that three MODIS based vegetation indices have been widely used; these are the NDVI, the EVI and the water index NDWI with 500 m spatial and 16 days temporal resolution. The day and night LST products based on MODIS were also extensively exploited providing continuous daily information in 1 km wide cells. Furthermore, remotely sensed precipitation data in the spatial resolution of 0.25 to 5.0 degrees have been derived using the TRMM mission with a revisit time of 23 days at the equator and 46 days at the highest latitudes. The TRMM mission provided precipitation data for 17 years over the tropical and subtropical areas; however this mission is no longer available as it was turned off in 2015.

4.1. Predictors for Malaria

Most of the studies in this review took place in tropical, subtropical, temperate and Sub-Saharan climatic zones as shown in Figure 2. Temperature was one of the most influential parameters affecting the malaria occurrence in tropical regions as shown in [35,37,39,45,46,48,84,88] and only one study [38] located in Tanzania, claimed that temperature performed poorly as a predictor. Lag times and degrees of temperature varied between the studies, which could be explained due to the fact that the studies were located in different climatic zones; Diboulo et al. that is located in Burkina Faso, and as such

is characterized by a Sub-Saharan climate, has found that the density of the vector *An. gambiae* was positively associated with the day temperature during the two previous months counting from the date that the collection of the mosquitoes has occurred, and it was negatively associated with the night temperature during the current and two previous months [46]. In studies that run in the tropical zone, as for example [45] that geographically refers to Western Kenya has shown that there was a three month lag pattern between temperature and peaks of malaria admissions and that the *An. gambiae* mosquito's density was negatively associated with the mean day temperature of 29 °C. The study of Adde et al., which was located in French Guiana, concluded that a minimum temperature of 20 °C proved to be always beneficial for mosquito *An. darlingi* breeding [48], while the study conducted by Ssempiira et al. that was located in Uganda, observed that the incidence of malaria was increased with day temperature, however very high temperatures above 29 °C resulted in a decline of malaria incidences [37]. This result was concordant with the study of Amadi et al., which was located at Baringo in Kenya and found that average monthly minimum temperatures between 16.2–21 °C (lag 1-month) made favorable conditions for the increase of the malaria risk [84].

Additionally, precipitation proved to be another significant predictor in the tropical zone highly associated with malaria occurrence [35,37,40,84]. Contrariwise, the study of Kabaria et al. that was conducted in Tanzania (tropical region) was the only one study that claimed precipitation performed poorly as a predictor [38]. Lag periods ranged significantly in the case of precipitation as well; Sewe et al. found out that all three study regions across the Area of Interest (AOI) located in Western Kenya resulted in different lag periods (0 to 12 weeks) [40] and Amadi et al. found positive associations between rainfall and malaria at a 2-month lag time [84]. Kanyangarara et al. [77] and Midekisa et al. [65] took place in humid subtropical climates; Kanyangarara et al. [77] found higher malaria risk during the rainy season at a total monthly rainfall between 94–181 mm, while Midekisa et al. found positive associations between rainfall and malaria cases at a lag time of one to three months [65].

Different vegetation indices (NDVI & EVI) were also tested in the literature and have been associated with malaria occurrences in tropical areas [35,37,40,76,84]. NDVI values between 0.3–0.4 have showed an increased correlation with the malaria risk as stated in [40,84]. EVI was also positively associated with reported malaria cases in humid subtropical areas like Midekisa et al. [65] and dry Mediterranean climates like Portugal in the study of Benali et al. [52]. The SAVI was positively associated with malaria distribution, meaning that the malaria vector prefers greener vegetation according to Malahlela et al. [34], which examined an AOI that is located in Vhembe District in South Africa and is characterized by varying topography; the north-western part is in the semi-arid climatic zone, while the southern-eastern part lies on the subtropical zone.

Proximity to water bodies however was negatively associated with malaria incidence in tropical areas like Uganda [37] and vector' density in Western Kenya [45], while water indices performed poorly as predictors both in temperate climates like South Africa [34] and tropical like Tanzania [38]. The percentage of dense/riverine vegetation was the most significant predictor for Kabaria et al. [38].

Socioeconomic factors were proved to be significant covariates [78], with the outdoor occupation being the most significant risk factor, followed by the SES and population density. This result is concordant with [92] that associated outdoor occupation with higher malaria risk, since people working outdoors are more exposed to receiving infective mosquito bites.

4.2. Predictors for Dengue

Temperature was an important explanatory variable related to the prediction of dengue cases in the tropical zone [43,59,64,71]. Ashby et al. claimed that day and night temperature played an important role in the determination of the dengue fever niche, with the day temperature limiting the reproduction rate of the main vector of dengue fever, *Aedes aegypti* [64], while the study of Sarfraz et al., which was located in Thailand, found that a temperature range between 30–35 °C had a high impact on *Aedes* vector breeding [43]. Hii et al. mentioned a consistent and stable association between mean temperature and dengue incidence [87] and Mokraoui et al. highlighted the importance of

temperature for estimating the dengue index, as longer dry seasons create more suitable sites for dengue outbreaks [53]. Ssempiira et al. found that both temperature indices of SMT and TCI with a lag time of 8–9 weeks were related to dengue incidence rate [37]. Two studies were located in subtropical areas; Yue et al. was conducted in a coastal area in China and found that day temperature and night temperature were significantly correlated with dengue fever outbreaks [85], while German et al. found that low temperatures had a negative association with the oviposition activity [50].

NDWI is an indirect proxy for precipitation and humidity, and it was associated in many studies conducted in subtropical climates with dengue occurrences; Scavuzzo et al. [51] and German et al. [50] examined areas that were located in the subtropical city of Tartagal in Argentina and claimed that NDWI was positively associated with the oviposition activity of *Aedes aegypti* vector, while [85], found that NDWI was significantly positively correlated with dengue outbreaks.

Other variables that proved to be significant were various LU/LC classes as well as socioeconomic factors; Yue et al. mentioned that the land type was significantly correlated with the dengue fever outbreak [85] and the study conducted by Machault et al. took place in Tartane in French Antilles, which is characterized by a tropical climate and found that the “sparsely vegetated soil” land use class was associated with the presence of water filled containers, while the “asphalt” land use class was negatively associated with the presence of *Aedes* larvae-positive containers [44]. Yue et al. claimed that human population density was one of the most significant predictors [85].

4.3. Predictors for WNV

Temperature was one of the most significant predictors for the WNV incidence, which appeared mainly in temperate [30,31] and continental climates [75,81,83,86], meaning large seasonal temperature differences, with warm to hot summers and cold winters. Temperature was positively associated with the WNV incidence [75,81,83,86], while Stilianakis et al. located in Greece mentioned that soil and air temperature were between the most significant predictors for WNV disease outbreak [30]. Marcantonio et al. [83] and Young et al. [81] which examined Europe and the US Great Plains respectively, both characterized by warm and humid continental climate, suggested precipitation as one significant predictor for WNV incidence. Moreover, and according to [75] located in the US Great Plains, cumulative ETa has shown positive association with WNV relative risk. Elevation played an important role in the prediction of WNV incidence [81], while low elevation was positively associated with both human and wild bird cases [31].

The vegetation index NDVI showed positive associations with the WNV risk according to Chuang et al. [75] and Young et al. [81] both located in the same region of the US Great Plains, while the study of Conley et al. that was conducted in arid and semi-arid areas mentioned that the seasonality of EVI was a significant predictor of the vector’s habitat [41]. Of the land use predictors, the irrigated croplands and the populated forest were the most significant predictors that were positively associated with the WNV incidence [75]. Kanyangarara et al. identified the WNF outbreak of the previous year as a risk factor [77].

Table 2. Association between predictors and diseases; the green color shows positive association between the disease and the predictive parameter, while the red color indicates negative association in the different climatic zone. (T = Temperature, P = Precipitation, ET = Evapotranspiration, Veg = Vegetation, El = Elevation, TWI = Topographic Wetness Index, WS= Wind Speed, Hum = Humidity, Pop = Population)

Parameter Association	Malaria		Dengue		WNV	
	(+) association	(-) association	(+) association	(-) association	(+) association	(-) association
T _{Tropical}	[35,37,39,45,48,84,88]	[48]	[43,59,64,71]			
T _{Semi-arid}	[63]					
T _{Subtropical}			[50,85]		[49]	
T _{Sub-Saharan}		[65]				
T _{Continental}					[75,81,83,86]	
T _{Mediterranean}					[30]	
P _{Tropical}	[35,39,40,84]	[48]	[93]			
P _{Semi-arid}	[63]				[81]	
P _{Subtropical}	[65]			[50]	[49]	
P _{Sub-Saharan}		[65]				
P _{Continental}					[83]	
ET _{Tropical}	[48]	[48]				
ET _{Semi-arid}	[63]					
ET _{Continental}					[75]	
Veg _{Tropical}	[38,40,45,84]		[44,71]			
Veg _{Semi-arid}	[34]				[41]	
Veg _{Mediterranean}	[52,65]					
Veg _{Continental}					[75]	
El _{Continental}						[31]
El _{Subtropical}	[77]			[93]		
TWI _{Tropical}	[42]					
NDWI _{Subtropical}			[50,51]	[85]	[86]	[83]
NDWI _{Continental}					[86]	[83]
WS _{Mediterranean}						[30]
Hum _{Mediterranean}						[30]
Pop _{Tropical}					[64]	
Pop _{Semi-arid}					[41]	
Pop _{Subtropical}			[85]			

4.4. Data Driven Uncertainties and Limitations

The quality of the input data has been the major limiting factor in regard to the sensitivity of the models and the accuracy of the predicted variables and risks. Given the dependence of MBDs prediction algorithms on temporal data, data availability and reliability were the major concerns of many studies.

A lack of systematic epidemiological and entomological data collection [34,37,81], uncertainty of the ingested dataset due to under-diagnosis and underreporting [35,38] and a confined number of cases [30] were reported as the main limitation reasons. Furthermore, the computation of entomological indices such as the CI, HI and BI [43] may have constrained the estimations due to the fact that these indices are highly dependent on the samples of the vectors in the containers that use the immature forms of the vectors. Sarfraz et al. suggests that pupal survey may have been more suitable for investigating the risk because it collects the mature form of the vectors, which may be more indicative for revealing the real trend of the risk [43]. The use of the statistical method of Inter-VA Autopsy from Sewe et al. might have under or over estimated the number of deaths [40].

Some of the studies mentioned additional limiting factors that were related to the satellite imagery. Quality problems relevant to image blurring and image stripe was mentioned by Yue et al. [85]. In addition studies that examined areas located in subtropical and tropical regions [34,48] and used optical data faced issues due to the existence of thick clouds. Fusion of optical and Synthetic Aperture Radar (SAR) data seemed to have resolved the problem to some extent.

4.5. Modeling Approaches and Evaluation

There has been a wide variety of approaches used for modeling/correlating MBDs with the EO derived predictors. The various approaches ranged from simple regression to advanced machine-learning techniques as shown in Figure 5. Most of the studies ($n = 27$) used statistical models including regression models [30,31,33,34,42,44,48–50,59,60,66,75,77,83–86,88], AutoRegressive Integrated Moving Average (ARIMA) models [35,65], spatial statistics [37,39,45,46,78], and probabilistic graphical models (Bayesian Networks) [71], while some studies ($n = 13$) used data mining [32,93], machine learning [63,81,93] and ensemble approaches [33,38,41,61,64,66,76], and [51] examined both. Some of the most important findings are presented below.

From this review it comes to our attention that the employment of regression methods has been relatively easy to apply and automate. These approaches included autoregressive terms and functions for seasonality in order to model the serial correlation. A Generalized Linear Mixed model (GLM) with Poisson distribution was utilized by Amadi et al. [84], while Sewe et al. [66] used the Generalized Additive Model (GAM) that is an extension of the GLM, in which the predictor variable is estimated using unspecific (non-parametric) functions [94]. Sewe et al. [66] compared the GAM with the Boosting (GAMBOOST) ensemble model, where GAMBOOST performed better since it reduced the over-fitting of the model and could handle the data that were non-stationary. Logistic regression was employed to model binary response variables for example the presence or absence of the infection by [30,34,44,77,86]; Malahlela et al. used a Stepwise Logistic Regression (SLR) [34], which is one of the most commonly used methods for relating remotely sensed data with disease distribution [95,96] to analyze the spatial distribution of malaria.

The ARIMA models have been used for analyzing and forecasting time series data and have been performing well in cases where the data appear to be non-stationary [97]. Furthermore, ARIMA models seemed well suited for representing temporal patterns, such as seasonality and serial correlation. Extensions of the ARIMA were utilized by two studies; Kanya et al. used the ARIMAX model [35], which is a multivariate Autoregressive Integrated Moving Average Model, extending the ARIMA model by including multiple predictors using current and past values of the independent variables [98] and [65] used a Seasonal Autoregressive Integrated Moving Average (SARIMA) modeling approach, including a seasonal component to relate the lagged association of environmental variables with malaria cases. SARIMA models performed well and could be used in cases where the time series of the

dependent variables exhibit a seasonal variation. However, the SARIMA models might fail to provide accurate prediction, if the preceding sequence of the time series exhibit abnormal variations [99]. Furthermore SARIMA were strongly data-driven, requiring a sufficient time series set of historical data for the model's parameterization.

Spatial statistics were used to analyze and predict the values associated with spatial or spatiotemporal phenomena by studies [37,39,45,46,85]. Yue et al. analyzed spatial patterns of dengue fever by conducting the following spatial analysis methods [85]: point density, average nearest neighbor, spatial autocorrelation and hot spot analysis, while Bayesian binomial models were utilized by [37,39,45,85].

Only Ruangdomsakul et al. estimated the dengue outbreak level by utilizing Bayesian Network (BN) [71]. BN is a probabilistic graphical model that uses Bayesian inference in order to perform probability computations. The goal of the BN is to model conditional dependence of the variables using a Directed Acyclic Graph (DAG) [100]. In this study, three BN models were tested, that included expert knowledge, the Greedy Thick Thinning algorithm (GTT), and a combination of both. After assessing the performance of the three different models, the model that combined the GTT and expert knowledge was suggested for forecasting dengue at the different outbreak levels.

Boosted Regression Trees (BRT) could handle a big variety of predictors, complex nonlinear relationships and missing data and were utilized by [38,41,61,64]; reference [38] used BRT to relate the high resolution urban LC classes, as well as other satellite derived environmental variables with the malaria prevalence, while Ashby et al. used BRT to quantify the risk of dengue incidence comparing the Poisson and the Bernoulli family models [64]. The BRT analysis that was conducted used disease presence/absence data for Bernoulli family and the actual case counted for the Poisson family. The results showed that the Poisson family returned a better model fit compared to the Bernoulli one, with a lower Root Mean Square Error (RMSE) and higher correlation.

The ANNs approach has been used for time series prediction and has been capable of reproducing and modeling nonlinear processes. ANNs had also the advantage of detecting every possible interaction between explanatory variables. On the other hand, the ANNs could not explicitly identify the causal relationships due to the unexplained behavior of the network and have been more prone to overfit the models in case of inadequate input datasets. Bui et al. tested multiple machine learning classifiers and ensemble techniques for relating malaria cases with socio-physical parameters and creating malaria vulnerability maps [76]. ANNs, Support Vector Machine (SVM), J48 and ensemble techniques using the J48 as a base classifier and Adaboost, Bagging and Random Subspace were used; the Random Subspace ensemble model performed the best. In the study conducted by Scavuzzo et al. multiple ML algorithms were examined to model temporal variations of the oviposition in both urban and rural areas [51]. SVM, ANN multi-layer Perceptron, Decision trees and K-Nearest Neighbor (KNN) were compared with two linear regression models. KNN performed better than the rest of the methods. Furthermore, two studies [33,41] utilized the Maximum Entropy (MaxEnt) approach for predicting the distribution of the vectors' population. Both studies compared the MaxEnt with other models; Conley et al. used the BRT method [41] that proved to have a strong agreement in results, and Arboleda et al. showed that the combination of BRT and Genetic Algorithm for Rule-set Prediction (GARP) yields the best models [33].

Although various evaluation metrics were used (Appendix A—Tables A1–A3), we tried to summarize the accuracy of the most commonly used methods in this review; regression methods that predicted malaria incidence varied widely year by year and showed big spatial heterogeneity $R^2 \sim 0.4\text{--}0.9$. However, classification methods in vector population prediction yielded a mean model accuracy of $\sim 80\%$ with a Confidence Interval (CI) of 95% and an AUC ~ 0.8 . Methods for prediction of dengue incidence in humans resulted in a mean $R^2 \sim 0.35$, while mosquito population dynamics were modeled with $R^2 \sim 0.7$. Regression models utilized WNV incidence in humans with the dependent variable showing a high temporal variability over the year ($R^2 \sim 0.1\text{--}0.7$), while early warning vector population dynamics and their potential transmission risk were modeled with $R^2 \sim 0.5$.

Studies used human incidence data that ranged significantly in terms of accuracy, mainly due to data gaps, non-systematic and non-standardized collection of the input data. The assumption that the infections were locally acquired might be biased due to population travel from one region to another. Studies that used entomological data seemed to yield slightly better accuracy. A possible explanation for this result is that disease incidence is highly correlated with the mosquito density and therefore models that took the vectors densities as response variables, seemed to result in more accurate predictions.

4.6. Scalability and Transferability

It is essential to know which methodologies seemed promising and claimed to be easily replicated in different geographic regions, enabling the realization of predictive systems on regional and national scales. The review shows that some of the methods used are scalable and transferable in other areas with similar climatic and MBDs conditions. The automated FARM method used by Buczak et al. [32], was described as generic and extendable in any geographical area. Rosa et al. [49] applied linear mixed models and stated that they are transferable and applicable in other areas with similar climate and land cover conditions, while the principles used in the models' design could be applied in any area. Machault et al. [44] used logistic regression analysis in a two-step approach and mentioned that the equations derived from the final model could be applied in regions with similar morphological and LU/LC conditions, while Scavuzzo et al. utilized different ML algorithms (SVM, ANNs, K-NN, decision trees), allowing transferability of the methodology in other regions as well [51]. In contrast, Sarfraz et al. mentioned that the fuzzy approach could not be extrapolated to a regional scale, since the environmental and social factors affecting the dengue vector density would vary significantly [43].

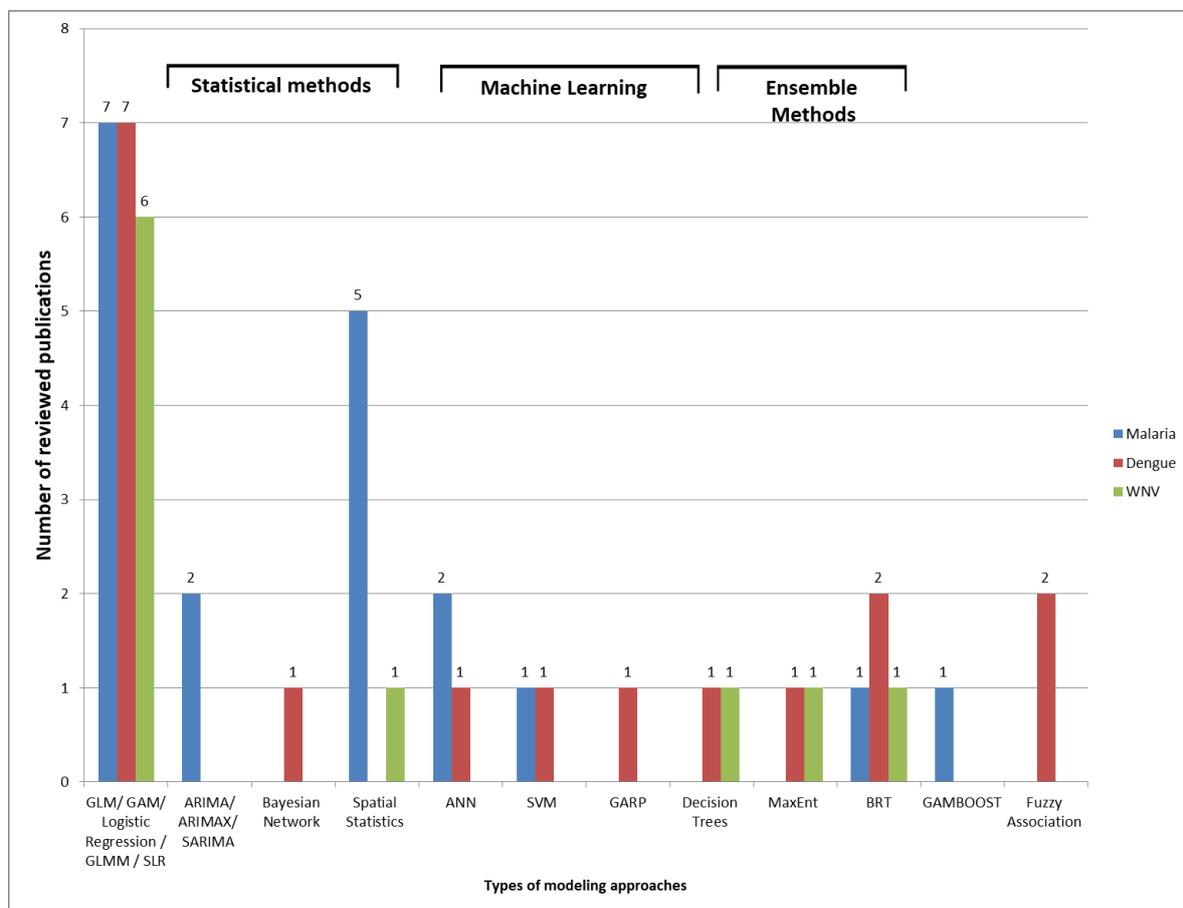


Figure 5. Overview of the methods that were considered in this review.

5. Conclusions

A wide range of both predictors and modeling approaches were found in the literature to forecast epidemic diseases like malaria, dengue and WNV. Researchers have examined different methods and have utilized different data sets in order to model the MBDs. Barriers until recently have been the temporal and the spatial resolution of the data and the data accuracy, as shown in this review; most of the studies have used data from satellite missions that are at the end of their operation or are no longer available. Because of this, it is strongly believed that state-of-the-art EO sensors and satellite systems need to be envisaged for the prediction of MBDs. Actually, with the advent of the Sentinel data, offered freely by the Copernicus EU program, a new challenge has arisen for the analysis of big satellite data and the employment of data science approaches. Their enhanced capabilities for multipurpose environmental monitoring at various scales along with the higher temporal and spectral resolutions will significantly increase the level of information on predictors such as soil moisture, vegetation and water bodies and therefore the accuracy of the models. There is a great advantage to using Sentinel-1 SAR images because of their enhanced azimuth spatial resolution (5 m) and mainly the ability to be used frequently every 6 days during day and night independently of the atmospheric and cloud conditions. Li et al. [101] and Catry et al. [102] have successfully used the fusion of optical and SAR data for generating LU/LC maps to better address the challenge of malaria elimination, while Catry et al. leveraged SAR data for estimating the extent of wetlands in the Amazon river basin [103]. Moreover, the Sentinel-2 images (10 m GSD, 6-days revisit time) offer a unique continuity and high accuracy assessments of indices such as NDVI, EVI, SAVI, NDWI, in complement to the SPOT and Landsat missions which have been widely used so far. In addition, the need to scale up predictions and move from the local to regional or continental level, can be ideally addressed if medium resolution (500m–1km GSD) data from Sentinel-3 are adequately combined with Sentinel-1 and Sentinel-2 and other existing operational HR and VHR satellite missions (SPOT, IKONOS, WorldView, etc.). In this regard, it is worth mentioning the benefits of using the Global Precipitation Measurement (GPM) mission of NASA as it offers an enhanced continuity for the TRMM mission, and provides precipitation data more frequently, with increased accuracy in the spatial resolution of 250 m every 3 h. Last but not least, the Soil Moisture Active/Passive (SMAP) mission, that has not yet been fully exploited, exhibits a high potential as it provides global soil moisture assessments at a spatial resolution of 3 km within 2–3 days revisit time.

Nowadays, new IT technologies allow for high computational performance to perform time-series analysis of big satellite derived data in order to estimate infectious disease trends enabling more accurate predictions for MBDs. Despite the progress made in epidemic forecasting there is still the need to exploit in depth new powerful modeling approaches like artificial intelligence and ensemble modeling ingesting long-lasting EO observations (space/in situ) and EO derived variables that allow the identification of highly complex relationships across data and risk factors influencing the MBDs transmission. However, possessing the ability to unhindered and continuous processing of volumes of data leveraging on High Performance Computing environment and Data Cubes, will assure geographic upscaling and transferability of the predictions in larger geographic areas. This is the primary challenging scientific problem of the days in EO, which in turn, if met, will lead to data driven decisions of high societal benefit.

6. Disclaimer

The views expressed are purely those of the writer (N.I.S.) and may not in any circumstance be regarded as stating an official position of the European Commission.

Author Contributions: Writing—original draft preparation, E.P.; writing—review and editing, C.K., A.T., C.H., I.K., G.M., N.I.S.; supervision, C.K.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ACTs	Artemisinin Combination Therapies
ADDS	Africa Data Disseminating Services
AIC	Akaike's Information Criterion
AIRS	Atmospheric Infrared Sounder
ANDWI	Adapted NDWI Mac Feetters Index
ANNs	Artificial Neural Networks
AOI	Area of Interest
ARIMA	AutoRegressive Integrated Moving Average
ARIMAX	Autoregressive Integrated Moving Average with Explanatory Variable
AST	Air Surface Temperature
AUC	Area under the ROC Curve
AVHRR	Advanced Very High Resolution Radiometer
BI	Breteau Index
BN	Bayesian Network
BRT	Boosted Regression Tree
BT	Brightness Temperature
CDC	Center for Disease Control
CHIRPS	Climate Hazards Group InfraRed Precipitation
CI	Container Index
CLMM	Cumulative Link Mixed Model
DAG	Directed Acyclic Graph
DEMs	Digital Elevation Models
ECMWF	European Centre for Medium-RangeWeather Forecasts
EIP	Extrinsic Incubation Period
EO	Earth Observation
ET	Evapotranspiration
ETa	Actual Evapotranspiration
EVI	Enhanced Vegetation Index
EWEM	Early Warning and Environmental Monitoring Program
EWs	Early Warning Systems
FARM	Fuzzy Association Rule
FSCPE	Federal State Cooperative Program
GAM	Generalized Additive Model
GAMBOOST	Generalized Additive Model with Boosting
GAUL	Global Administrative Unit Layers
GARP	Genetic Algorithm for Rule-set Prediction
GI	Green Index
GLM	General Linear Mixed
GLR	General Linear Regression
GPM	Global Precipitation Measurement
GPW	Gridded Population of the World
GTT	Greedy Thick Thinning
HI	Household Index
HR	High Resolution
IBGE	Brazilian Institute for Geography and Statistics
ICT	Information and Communication Technology
INEI	Peru National Institute of Statistics and Information 2007
INSA	Indian National Satellite System
ITN	Insecticide treated

KNN	K-Nearest Neighbor
LMM	Linear mixed effects models
LST	Land Surface Temperature
LU/LC	Land Use/Land Cover
MAE	Mean Absolute Error
MaxEnt	Maximun Entropy
MBDs	Mosquito-Borne Diseases
MR	Medium Resolution
ML	Machine Learning
MNDWI	Modified Normalized Difference Water Index
MSL	Mean Seal Level
NDVI	Normalized Difference Vegetation Index
NDWI	Normalized Difference Water Index
NOAA	National Oceanic and Atmospheric Administration
NUTS	Nomenclature of Territorial Units for Statistics
ONI	Oceanic Niño Index
PCA	Principle Component Analysis
PPV	Positive Predictive Value
RDT	Rapid Diagnostic Tests
RMSE	Root Mean Square Error
RMSPE	Root Mean Square Percentage Error
p-YI	quasi-Yellowness Index
SAR	Synthetic Aperture Radar
SAVI	Soil-Adjusted Vegetation Index
SES	Socio-Economic Status
SDGs	Sustainable Development Goals
SLP	Sea Level Pressure
SLR	Stepwise Logistic Regression
SMAP	Soil Moisture Active/Passive
SMAPE	Symmetric Mean Absolute Percentage Error
SMN	Smoothed and normalized difference vegetation index
SMT	Smoothed brightness temperature index
SOI	Southern Oscillation Index (SOI)
SRTM	Shuttle Radar Topography Mission
SST	Sea Surface Temperature
SSTA	Sea Surface Temperature Anomaly
SVM	Support Vector Machine
TCI	Temperature Condition Index
TPI	Topographic Position Index
TRMM	Tropical Rainfall Measuring Mission
TWI	Topographical wetness index
USGS	United States Geological Survey
VCi	Vegetation Condition Index
VHI	Vegetation Health Index
VHR	Very High Resolution
WHO	World Health Organization
WNND	West Nile neuro-invasive disease
WNV	West Nile Virus

Appendix A

Table A1. Summary table of malaria predictive studies included in review AI: Artificial Intelligence, SM: Statistical Method.

Reference	Period of Study	Validation	Dependent Variable	Number of Independent Variables/Best	Method	Score
[63]	1995–2015	1995–2014: training set 2015: testing set	Relative Malaria abundances	6 vars (lagged, monthly). Best: temperature and rainfall	ANN/AI	RMSPE ranged from 18% to 117%.
[48]	2012–2014	10-fold cross-validation	Malaria vector densities	22 landscape and 2214 meteorological variables Best: Meteorological variables: rainfall, evapotranspiration, min and max temperature Landscape variable: dense forest surface, built surface	Multivariate analysis-CLMM/SM	AUC of mosquito density classes: “Low” –0.78 “Medium” –0.64 “High” –0.80
[34]	2005	60% training set 40% validation set	Probability of malaria distribution	9 vars. Best: soil-adjusted vegetation index (SAVI)	Stepwise logistic regression model/SM	Classification accuracy of 82% at a threshold of 0.9 (buffer distance of 10 km)
[35]	2006 to 2013	site implementation-31/05/2012: training set 01/06/2012–31/05/2013: testing set	Weekly number of laboratory- confirmed malaria cases	19 vars. Half of the predictor series were lagged, ranging from lags of 1 to 52 weeks Best: Drug treatment, precipitation	ARIMAX/SM	SMAPE ranged from 26% to 128%
[77]	2012–2015	Monte Carlo cross validation 66.6%: training set 33.3%: testing set	Household with at least one member test positive by RDT	11 vars. Best: Distance to the Mozambique border, elevation	Multivariate logistic regression/SM	For the rainy season, the sensitivity and specificity of the model were 61% and 80%, respectively. The model performance during the dry season had better specificity (96%) but far worse sensitivity (37%).

Table A1. Cont.

Reference	Period of Study	Validation	Dependent Variable	Number of Independent Variables/Best	Method	Score
[37]	2013–2017	Markov chain Monte Carlo (MCMC) simulation. A two-chain algorithm for 200,000 iterations with an initial burn-in period of 5000 iterations	Malaria incidence in each age group was estimated by dividing the district aggregated malaria cases by the district age group-specific population	4 vars, (lagged: (i) current and previous month, (ii) current and two previous months, (iii) current and three previous months). Best: LST night, LST day	Bayesian spatio-temporal negative binomial models/Geostatistics	At least one ITN was associated with a decline in malaria incidence in children < years by 73%
[38]	2006–2014	Randomly split 75%: training set 25%: testing set	Malaria parasite prevalence	10 vars. Best: percentage of dense/riverine vegetation	Boosted Regression Tree (BRT) modeling/AI	Model prediction accuracy: AUC = 0.89
[88]	1997–2006	A single year was left out one-by-one from the data set	Malaria incidences	2 vars. Best: TCI	OLS, Principal Component Regression (PCR) PCR performed better	PCR: $R^2 = 0.68$ OLS: $R^2 = 0.43$
[65]	2001–2009	2001–2008: training set 2009: testing set	Monthly time series of malaria cases	4 vars lagged (monthly, yearly). Best: Precipitation (lag one to three months)	Seasonal autoregressive integrated moving average (SARIMA)/SM	Akaike weights greater than 85%
[40]	Asembo and Gem: 2003–2012 Karemo: 2008–2012	-	Number of malaria deaths	3 vars lagged (0–12 weeks) Best: Precipitation	Distributed Lag Non Linear Modeling/SM	Precipitation relative risk: 1.68 in Asembo Vegetation relative risk: 3.4 in Gem

Table A1. Cont.

Reference	Period of Study	Validation	Dependent Variable	Number of Independent Variables/Best	Method	Score
[52]	2001–2010	Statistical association	Monthly values of Vector density	5 vars, Best: Land cover, EVI	(1) Simple Linear relations (2) non-linear multi-variable models (Gaussian, log-logistic, first to third degree polynomials and an inverse second degree polynomial equation)/SM	Coupling static land cover suitability with dynamic vegetation data Nash–Sutcliffe model efficiency (MEF) index MEF = 0.90
[84]	2009–2012	-	Malaria cases	3 vars, lagged (1,2 and 3 months) NDVI (+) and monthly total precipitation (+)	Poisson regression—GLMM/SM	T min and NDVI accounted for 66% (29.9, 36.1 respectively) of the total variation in malaria incidence explained by model
[45]	2002–2004	Randomly selected locations 85%: training set 15%: testing set	Mosquito density	3 vars lagged up to 3 month Best: Distance to water bodies, mean value of NDVI during the month of collection and average day temperature during the current and the previous month of collection	Geostatistical zero inflated binomial and negative binomial models/SM	mosquito densities 66% the zero inflated spatio-temporal negative binomial model and 83% zero inflated spatial negative binomial model respectively(CI:95%)
[78]	2012–2013	Different samples of training and a validation sets were considered to validate predictions in every cluster	Variable malaria positive or negative	6 vars Best: Higher socioeconomic status	geographically-weighted regression (GWR)/SM	R ² values per cluster vary between 32% and 87% with a mean of 63%
[76]	2016–2017	Ten- fold cross validation	malaria incidence	10 vars Best: Land Use, Distance to residence	ANN, SVM, ensemble techniques (J48)/AI	Best model: Random Subspace ensemble model overall accuracy: 94.2%

Table A1. Cont.

Reference	Period of Study	Validation	Dependent Variable	Number of Independent Variables/Best	Method	Score
[66]	2003–2013	5-k cross validation 2003–2012: training set 2013: testing set	malaria admissions	5 vars, (lagged 1 to 3 months) Observed: lag pattern of rainfall and temperature	GAMBOOST, GAM/SM	GAMBOOST: $R^2 = 0.71$ GAM: $R^2 = 0.44$
[46]	2001–2004	Randomly selected locations subset 85% training set 15% testing test	mosquito density and EIR	6 vars (lagged 1 to 3 months) Association: Rainfall (–) and night temperature (–)	Bayesian geostatistical zero-inflated binomial and negative binomial models Bayesian kriging/SM Geostatistics	Mosquito density models 58%: zero in-flated spatio-temporal negative binomial mode 73%: zero inflated spatial negative binomial model(CI:95%)
[39]	2011	Subset of 35 location: testing set	malaria infected individuals	4 vars, Best: rainfall and LST day	Bayesian geostatistical model/Geostatistics	Best model: Log-predictive density of –115.12
[42]	2006	The accuracy of the model predictions was tested using independent breeding site data from Nyamanga in 2010	malaria breeding sites	9 vars: TWI (+)	logistic regression/SM	In the test site AUC for SRTM: 0.829 AUC for ASTER: 0.799

Table A2. Summary table of dengue predictive studies included in review AI: Artificial Intelligence, SM: Statistical Method.

Reference	Period of Study	Validation	Dependent Variable	Number of Independent Variables/Best	Method	Score
[32]	2001–2009	Data set was disjoint into training, validation and test set	Dengue incidence	108 variables (lagged: 3 week, 4 weeks and 4–7 weeks ahead)	Fuzzy Association Rule Mining (FARM)/AI	4–7 weeks from time of prediction yielded a PPV = 0.686, PPN = 0.976, sensitivity: 0.615, specificity: 0.982

Table A2. Cont.

Reference	Period of Study	Validation	Dependent Variable	Number of Independent Variables/Best	Method	Score
[71]	2007–2015	10-fold cross validation procedure	Outbreak level of dengue	5 vars (lagged weekly): Best: SMN, SMT, and TCI	Bayesian Network/SM	Best model: BN model built with an expert and GTT overall Accuracy: 0.906 and AUC 0.954
[50]	2012–2016	2012–2014: training set 2015–2016: testing set	Ovipositor (Nr od Eggs)	42 vars (lagged weekly): Best temperature, humidity and precipitation	linear multivariate method/SM	final model with prediction capacity ($R^2 = 0.7, p < 0.05$) was established.
[44]	2009–2011	Validation was performed on the same dataset as the one used to fit the model	Presence of <i>Aedes aegypti</i> larvae	8 vars: sparsely vegetated soil (+)	Two-step approach, Logistic regression analysis/SM	84% of the experimental units were correctly predicted, the percentage of correctly classified predictions ranged from 67% to 92% depending on sections. The positive predictive value of the two step scenario was 57%, and the negative predictive value was 90%. Sensitivity: 57% Specificity: 90%
[93]	2009–2011	Analytic hierarchy process	Larval density	7 vars. Elevation (–), Temperature 30–35 °C (+), Rainfall 40–81 mm\hr	Fuzzy logic Data mining and the decision tree method/Data Mining/ AI	Overall accuracy of 80%

Table A2. Cont.

Reference	Period of Study	Validation	Dependent Variable	Number of Independent Variables/Best	Method	Score
[51]	2012–2016	Time series cross validation procedure (80% training, 20% test)	Oviposition	5 vars (lagged 3 weeks)	SVM, ANN, K-NN, decision trees, Linear and ridge regression/SM-AI	KNN, correlation between observed and fitted values 90%
[64]	2012–2014	BRT cross validation (75% training set, 25% testing set)	Disease presence/absence	64 vars Best: Population density and daytime LST min	Boosted Regression Trees: (1) Bernoulli Family (presence/absence) (2) Poisson Family (actual case counts)/AI	Poisson family better model fit compared to the Bernoulli, with lower RMSE and higher correlation. 50% of the relative influence was due to the population density and the daytime LST Poisson: Pearson r 0.9 Bernoulli Pearson r 0.25
[85]	2014	-	Dengue fever cases	6 vars. Best: Population density, night and day LST	OLS/SM	$R^2 = 0.320$
[33]	2002–2008	Maxent 50% of the data were used to train the model	Presence of breeding vector populations in	11 vars (lagged 1–3 weeks)	Ecological niche models Maxent and GARP/AI	Two models Maxent + GARP Prediction Maxent: 46–83% GARP: 23–61% Combination: 44.3–76.1%
[59]	1992–2011	-	Dengue incidence	7 vars (Monthly and annual averages, amplitudes, and anomalies) Strongest correlation: SST and AST	Logistic regression/SM	PCA: 4 vars explained 72% of the variance $R = 0.04$ – 0.56 varied over years
[60]	2006–2015	-	Dengue incidence rates	7 vars (lagged weekly) Best: previous dengue cases	Stepwise multiple regression analyses/SM	Best model: $R^2 = 0.42$
[61]	1960–2012	Statistic association	Probability of dengue occurrence	8 vars: Best: rainfall, temperature and the degree of urbanization.	BRT/AI	AUC: 0.81

Table A3. Summary table of WNV predictive studies included in review AI: Artificial Intelligence, SM: Statistical Method.

Reference	Period of Study	Validation	Dependent Variable	Number of Independent Variables/Best	Method	Score
[81]	2003–2008	80%: training set 20%: testing set	WNV incidence for each US county	5 vars (lagged, monthly). Best X: precipitation for the whole 6y period.	Decision trees/ML	Annual results were highly variable: R = 0–0.84 Tested the whole 6 years period: R = 0.86
[49]	2001–2011	-	Annual Culex population at 44 sites	5 vars (lagged window of 12 consecutive weeks). Best X: (a) days of precipitation at the start of the year (+) and (b) distance to rice fields (-)	Linear mixed models/SM	R ² = 0.46–0.49
[86]	2002–2013	2002–2011: training set 2012–2013: testing set	Probability of WND human infections	12 vars. Best model: Temperature anomalies in July, MNDWI in early June Multivariate	Logistic Regression/ML	AUC: 0.819 for 2012 AUC: 0.853 for 2013.
[83]	2010–2012	Statistical association	Annual total of WNV incidences for each NUTS3 area per 100 K inhabitants	10 vars (lagged into 9 blocks of 4 months). Best model: uses: NDWI (-) in spring-early summer, T (+) in summer, days of precipitation (+), area covered with irrigated croplands (+)	linear mixed-effects models (LMMs)/SM	Best model: (R ² = 0.32)
[41]	2009	Training set: Locations recorded as positive for Cx. pipiens from Egypt (<i>n</i> = 239) from Lebanon (<i>n</i> = 83) Testing set: Independent locations of Cx. pipiens from Israel (<i>n</i> = 23) from Egypt (<i>n</i> = 56)	Habitats of Culex mosquitoes	24 vars. Best X: population density, seasonality of EVI	Maximum entropy and BRT/AI	In regions with high risk: R 0.73–0.77, In regions with low risk: R 0.55–0.66

Table A3. Cont.

Reference	Period of Study	Validation	Dependent Variable	Number of Independent Variables/Best	Method	Score
[30]	2010–2014	Statistical Association	Reported human cases of WNF/ WNND (441 cases), in weekly blocks	6 vars (lagged 1–3 weeks), wind speed (–), relative humidity (–), air temperature (+)	Multiple logistic regression/SM	The odds ratios (OR) OR wind speed: 0.76 (95% CI) OR relative humidity: 0.60 (95% CI), respectively, for lag 0.
[75]	2004–2010	Drop 1 year at a time	The number of county-level WNV neuroinvasive cases and WNV fever cases (3131 cases), in 8day blocks	3vars: The models were sensitive to: the timing of spring green up (measured with NDVI), temperature variability in early spring and summer (measured with LST), and moisture availability from late spring through early summer (measured with ETa)	Non-linear generalized additive models (GAMs)/SM	R ² 0.18 (April)–0.62 (August).
[31]	2010–2012	2010–2011 training set 2012 testing set	Positive human cases	37 vars. Best X: elevation (–) and distance from water (–)	Two step cluster analysis/SM	80% of incidences in 2012 occurred in areas recognized by the model as high-risk

Table A4. Overview of studies that utilized EO derived environmental/climatic variables and epidemiological data * in situ data.

Reference	Mosquito-Borne Disease	Study Area	Climatic Zone	Epidemiological Data	EO Climatic/Environmental Data
[32]	Dengue	Peru, province of Loreto	Tropical	Probable and confirmed cases longer text	Precipitation (TRMM), Temperature (USGS), NDVI (USGS), EVI (USGS), SOI (NCAR), SSTA (GCMD), Elevation (NOAA), running water, sanitation, electric lighting (INEI)
[81]	WNV	Great Plains region of the US	Semi arid	WNV incidence data aggregated to the county level-(I.R.)	NDVI (MODIS), Elevation (SRTM), Land Cover (NLCD2006 Landsat ETM+), temperature *, precipitation * (Oregon State University's PRISM)
[71]	Dengue	Sisaket province, Thailand	Tropical	Dengue confirmed cases (I.R.)	SMN, SMT, VCI, TCI, VHI (AVHRR)
[63]	Malaria	Local Khammam district, Telangana, India	Semi arid	Positive cases without symptoms	precipitation (TRMM), day and night LST (MODIS), EVI (MODIS), relative humidity * (MOSDAC)
[86]	WNV	1113 districts, Europe	Warm humid continental	Number of districts reporting WND cases in humans	Air temperature (NOAA NCEP-NCAR database), MNDWI (MODIS), Wetlands (GLWD)
[35]	Malaria	Uganda	Different climates, but is dominated by tropical savanna climate	Laboratory confirmed malaria cases	day and night LST, EVI (MODIS), precipitation (TRMM)
[77]	Malaria	region Mutasa District, Zimbabwe	Humid subtropical	RDT-positive participants-Household RDT	Elevation, slope, aspect (STRM), NDVI, LU (Landsat-8), distance to streams, distance to main road, distance to health facility, distance to Mozambique border
[83]	WNV	146 NUTS3 regions, Europe	Warm humid continental climate	WNV incidence	daily LST, NDVI, NDWI(MODIS), precipitation (ECA&D), LC, water bodies (OpenStreetMap), protected areas (IUCN and UNEP), Light at night (VIIRS)
[37]	Malaria	Uganda	Tropical savanna climate	Confirmed malaria cases by RDT-district aggregated monthly malaria cases	day and night LST (MODIS), NDVI(MODIS), LC, elevation (SRTM), distance to water, precipitation (EWEM)

Table A4. Cont.

Reference	Mosquito-Borne Disease	Study Area	Climatic Zone	Epidemiological Data	EO Climatic/Environmental Data
[38]	Malaria	City of Dar es Salaam, Tanzania	Tropical	RDT test -standardized parasite prevalence into the 2–10 years age group	LC, NDVI, NDWI (SPOT-6), Distance to inland water, Percentage dense/riverine vegetation, Percentage built-up, Elevation (ASTER GDEM), Compound Topographic Index (CTI) , daily LST(MODIS), precipitation (RFE 2.0)
[88]	Malaria	Tripura state India	Tropical savanna climate	Malaria cases-slidy positive rate and % of malaria cases/total Nr. of patients tested	NDVI, BT, (VCI-TCI) ((AVHRR) (NOAA GVI))
[64]	Dengue	Magdalena River watershed of Colombia	Tropical	Confirmed cases Dengue Fever	day and night LST (MODIS), EVI (MODIS), precipitation (TRMM), LU/LC, elevation (SRTM)
[30]	WNV	Northern Greece	Humid subtropical	Confirmed laboratory cases/mosquito traps	Air temperature, relative humidity, soil temperature, soil water content, wind speed, precipitation (ERA-Interim)
[75]	WNV	66 counties from the Northern Great Plains US	Weather varies throughout the year with cold winters, hot summers, and strong winds	Positive human case–Logarithm of relative risk (LRR)	mean LST (MODIS), NDVI (MODIS), ETa (FEWS-NET)
[65]	Malaria	District Amhara region of Ethiopia	Humid subtropical	Clinically diagnosed malaria cases	precipitation (TRMM), eight-day composite LST (MODIS), NDVI (MODIS), EVI (MODIS), ETa (MODIS)
[40]	Malaria	Western Kenya	Tropical	VA4 method-malaria deaths	day LST, NDVI (MODIS), precipitation (TRMM)
[84]	Malaria	Baringo, Kenya	Tropical savanna	Clinical malaria cases	monthly LST, NDVI (MODIS), precipitation (CHIRPS)
[85]	Dengue	Guangzhou, China	Subtropical coastal area with an oceanic subtropical monsoon climate	Confirmed dengue cases (clinical and laboratory diagnosis)	Land type(LT), NDWI (GF-1 satellite), day and night LST(GF-1 satellite)

Table A4. Cont.

Reference	Mosquito-Borne Disease	Study Area	Climatic Zone	Epidemiological Data	EO Climatic/Environmental Data
[78]	Malaria	Western Kenya	Tropical	Malaria prevalence (positive to antigens) RDT Household information / odour-baited MM-X traps	elevation relative to lake (ASTER), distance to lake, distance to nearest clinic, NDVI (QuickBird), TWI (QuickBird)
[76]	Malaria	Province of Dak Nong, Vietnam	Tropical	Malaria incidence	Elevation, aspect, slope (Aster Global DEM), temperature *, precipitation * (meteorological stations), NDVI (Landsat 8), LU type, distance to road, distance to residential area, distance to river (Landsat 8)
[53]	Dengue	State of Selangor, Malaysia	Tropical rainforest	Dengue cases	Temperature, precipitation (Weather department Malaysia), LU (Town Planning Department Malaysia)
[31]	WNV	Greece	Humid subtropical	laboratory-confirmed human cases, wild birds seroprevalence	temperature (WorldClim), precipitation (WorldClim), NDVI (WorldClim), elevation, slope, aspect (STRM DEM), LU (Corine Land Cover 2000 database), distance to water, distance to nearest village
[66]	Malaria	Western Kenya	Tropical	Confirmed malaria cases	day and night LST(MODIS), NDVI (MODIS), ET (MODIS), precipitation (TRMM)
[87]	Dengue	Singapore	Tropical	Dengue Cases	temperature (NOAA), precipitation (rain gauges)
[82]	Dengue	South Brazil	Tropical	Laboratory and clinically confirmed dengue cases/Dengue Incidence rate	precipitation (GPCP), ONI (NOAA CPC), mean surface air temperature (NCEP/NCAR), elevation (IBGE)
[59]	Dengue	San Juan, Puerto Rico	Tropical island	Confirmed dengue cases	surface air temperature (maximum and minimum), precipitation, sea level pressure (SLP), wind speed (NOAA-National Climatic Data Center), Sea Surface Temperature (SST), precipitation (AVHRR), Mean Sea level (MSL) * Sea Level Pressure SLP) *
[39]	Malaria	Mozambique	Tropical	Number of Infected children from 0 to 5 years old	LST (MODIS), precipitation (ADDS/RFE 2.0), elevation (GDEM/USGS), LU\LC (RapidEye)
[67]	Malaria	Amhara region	Humid subtropical	Malaria cases (clinically diagnosed as well as confirmed)	Precipitation (TRMM), mean daily LST (MODIS), NDVI (MODIS), SAVI (MODIS), EVI (MODIS), NDWI (MODIS)

Table A4. Cont.

Reference	Mosquito-Borne Disease	Study Area	Climatic Zone	Epidemiological Data	EO Climatic/Environmental Data
[60]	Dengue	State of Yucatan, Mexico	Tropical	Confirmed dengue fever cases	Day- and night-time SST (AVHRR), air temperature *, humidity *, and precipitation * (CONAGUA)
[61]	Dengue	Worldwide	-	dengue occurrence records	Precipitation * (WorldClim), temperature index, NDVI (AVHRR)

Table A5. Overview of studies that utilized EO derived environmental/climatic variables and entomological data * in situ data.

Reference	Mosquito-Borne Disease	Study Area	Climatic Zone	Entomological Data	EO Climatic/Environmental Data
[48]	Malaria	Municipality of Saint-Georges de l'Oyapock French Guiana	Tropical	Mosquito Magnet traps baited with octenol-weekly number of specimens	Land cover (SPOT-5), precipitation *, temperature *, relative humidity *, solar radiation *, evapotranspiration * (Meteo-France weather station)
[49]	WNV	Eastern Piemonte, Italy	Humid subtropical	CO ₂ baited traps-start (ON) and end (OFF) of mosquito season (threshold values for population abundance)	daily LST (MODIS), Precipitation * (ECA&D), NDWI (MODIS), LU\LC (Corine Land Cover), proximity to mosquito traps and rice fields
[34]	Malaria	Vhembe District Municipality in Limpopo Province of South Africa	North-western part is semi arid, south-eastern is subtropical	Presence of Malaria agent from patients that were tested positive for P. falciparum-presence/pseudo-absence were generated at buffer distances of 0.5–20 km	NDVI(Landsat 5), NNDWI, GI, SAVI, p-YI, Moisture index (Landsat 5), aspect (ASTER), Elevation (SRTM)
[50]	Dengue	Tartagal City, Argentina	Humid subtropical	Ovitrap–sum of eggs	NDVI (MODIS), NDWI (MODIS), LST (MODIS), precipitation (TRMM)

Table A5. Cont.

Reference	Mosquito-Borne Disease	Study Area	Climatic Zone	Entomological Data	EO Climatic/Environmental Data
[44]	Dengue	Tartane (Martinique, French Antilles)	Tropical	<i>Aedes</i> larvae-positive cases identified at the experimental units (houses visited multiple time)	Temperature *, humidity *, precipitation * (Météo-France stations), NDVI (Geoeye-1), NDWI (Geoeye-1), ANDWI (Geoeye-1), LU\LC (Geoeye-1), elevation (Litto3D)
[93]	Dengue	Phetchabun Province, Thailand	Tropical	Larval density inside and around residential homes-HI, CI, BI	day LST, night LST (MODIS), precipitation (TRMM), relative humidity (Aqua/AIRS), elevation (SRTM), LU\LC (MODIS)
[51]	Dengue	Tartagal city Northwest of Argentina in Salta Province	Humid subtropical	Vector population–Ovitrap Egg	NDVI (MODIS), NDWI (MODIS), day and night LST(MODIS), local precipitation (TRMM)
[41]	WNV	Middle East and North Africa	Arid and semi-arid	(1) Adult sample–CDC light traps (2) Larval samples–classical dipping method form artificial and natural breeding sites	Temperature (WorldClim), precipitation (WorldClim) EVI (MODIS), TWI (GLSDM)
[30]	WNV	Northern Greece	Humid subtropical	Confirmed laboratory cases/mosquito traps	Air temperature, relative humidity, soil temperature, soil water content, wind speed, precipitation (ERA-Interim)
[52]	Malaria	southern Portugal	Mediterranean	CDC light traps/vector density	8 day LST (MODIS), NDVI (MODIS), EVI (MODIS), NDWI (MODIS), LC (Corine Land Cover 2006)
[33]	Dengue	Antioquia, Colombia	Tropical rainforest	The basic unit of sampling was the house, where systematic searches found water containers with larvae, pupae, or exuviae/Breteau index Human cases at least suspected	Landsat band 1–7, NDVI (Landsat 7), elevation, slope, aspect (SRTM)
[45]	Malaria	Western Kenya	Tropical	Mosquito density CDC light traps/EIR	day and light LST (MODIS), NDVI (MODIS), precipitation (Meteosat 7), elevation (USGS EROS Data Center), distance to the nearest water source

Table A5. Cont.

Reference	Mosquito-Borne Disease	Study Area	Climatic Zone	Entomological Data	EO Climatic/Environmental Data
[78]	Malaria	Western Kenya	Tropical	Malaria prevalence (positive to antigens) RDT Household information/odour-baited MM-X traps	elevation relative to lake (ASTER), distance to lake, distance to nearest clinic, NDVI (QuickBird), TWI (QuickBird)
[46]	Malaria	Nouna district in Burkina Faso	Sub-Saharan	CDC light traps/EIR	NDVI, day and night LST (MODIS), precipitation (ADDS), water bodies (Health Mapper)
[42]	Malaria	Rusinga Island in western Kenya	Tropical	Vector density	slope, aspect, plan curvature, profile curvature, convergence index, and wetness index, topographic position index (TPI) (SRTM, ASTER)

References

- World Health Organization. Mosquito-borne diseases. 2018. Available online: https://www.who.int/neglected_diseases/vector_ecology/mosquito-borne-diseases/en/ (accessed on 20 November 2018).
- World Health Organization. Eliminating Malaria. 2016. Available online: https://apps.who.int/iris/bitstream/handle/10665/205565/WHO_HTM_GMP_2016.3_eng.pdf;jsessionid=F61B110C5B2AE747195723077A15AF09?sequence=1. (accessed on 30 November 2018).
- Ford, T.E.; Colwell, R.R.; Rose, J.B.; Morse, S.S.; Rogers, D.J.; Yates, T.L. Using satellite images of environmental changes to predict infectious disease outbreaks. *Emerg. Infect. Dis.* **2009**, *15*, 1341–1346. [[CrossRef](#)] [[PubMed](#)]
- Gubler, D.J. Dengue, Urbanization and Globalization: The Unholy Trinity of the 21(st) Century. *Trop. Med. Health* **2011**, *39*, 3–11. [[CrossRef](#)] [[PubMed](#)]
- Bauwens, I.; Franke, J.; Gebreslasie, M. Malareo-Earth observation to support Malaria Control in Southern Africa. In Proceedings of the 2012 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Munich, Germany, 22–27 July 2012; pp. 7252–7255. [[CrossRef](#)]
- Hay, S.I.; Packer, M.J.; Rogers, D.J. Review article The impact of remote sensing on the study and control of invertebrate intermediate hosts and vectors for disease. *Int. J. Remote. Sens.* **1997**, *18*, 2899–2930. [[CrossRef](#)]
- Kalluri, S.; Gilruth, P.; Rogers, D.; Szczur, M. Surveillance of arthropod vector-borne infectious diseases using remote sensing techniques: A review. *PLoS Pathog.* **2007**, *3*, 1361–1371. [[CrossRef](#)] [[PubMed](#)]
- Kazansky, Y.; Wood, D.; Sutherlun, J. The current and potential role of satellite remote sensing in the campaign against malaria. *Acta Astronaut.* **2016**, *121*, 292–305. [[CrossRef](#)]
- Pixalytics Ltd. 2016. How many Earth observation satellites are in space in 2018?. 2018. Available online: <https://www.pixalytics.com/eo-satellites-in-space-2018/> (accessed on 4 December 2018).
- Ma, Y.; Wu, H.; Wang, L.; Huang, B.; Ranjan, R.; Zomaya, A.; Jie, W. Remote sensing big data computing: Challenges and opportunities. *Future Gener. Comput. Syst.* **2014**. [[CrossRef](#)]
- Viana, J.; Santos, J.V.; Neiva, R.M.; Souza, J.; Duarte, L.; Teodoro, A.C.; Freitas, A. Remote sensing in human health: A 10-year bibliometric analysis. *Remote Sens.* **2017**, *9*, 1225. [[CrossRef](#)]
- SDGs: Sustainable Development Knowledge Platform. Available online: <https://sustainabledevelopment.un.org/sdgs> (accessed on 28 February 2019).
- World Health Organization. Malaria. 2019. Available online: <https://www.who.int/news-room/fact-sheets/detail/malaria> (accessed on 30 April 2019).
- Sadoine, M.L.; Smargiassi, A.; Ridde, V.; Tusting, L.S.; Zinszer, K. The associations between malaria, interventions, and the environment: A systematic review and meta-analysis. *Malar. J.* **2018**, *17*, 73. [[CrossRef](#)]
- Paaijmans, K.P.; Read, A.F.; Thomas, M.B. Understanding the link between malaria risk and climate. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 13844–13849. [[CrossRef](#)]
- Rogers, D.J.; Randolph, S.E.; Snow, R.W.; Hay, S.I. Satellite imagery in the study and forecast of malaria. *Nature* **2002**, *415*, 710–715. [[CrossRef](#)]
- Paaijmans, K.P.; Wandago, M.O.; Githeko, A.K.; Takken, W. Unexpected High Losses of Anopheles gambiae Larvae Due to Rainfall. *PLoS ONE* **2007**, *2*, e1146. [[CrossRef](#)] [[PubMed](#)]
- Whitehorn, J.; Simmons, C.P. The pathogenesis of dengue. *Vaccine* **2011**, *29*, 7221–7228. [[CrossRef](#)] [[PubMed](#)]
- WHO. *What Is Dengue?* WHO: Geneva, Switzerland, 2017.
- Barbazan, P.; Guiserix, M.; Boonyuan, W.; Tuntaprasart, W.; Pontier, D.; Gonzalez, J.P. Modelling the effect of temperature on transmission of dengue. *Med. Vet. Entomol.* **2010**, *24*, 66–73. [[CrossRef](#)] [[PubMed](#)]
- Stewart Ibarra, A.M.; Ryan, S.J.; Beltrán, E.; Mejía, R.; Silva, M.; Muñoz, Á. Dengue Vector Dynamics (*Aedes aegypti*) Influenced by Climate and Social Factors in Ecuador: Implications for Targeted Control. *PLoS ONE* **2013**, *8*, e78263. [[CrossRef](#)] [[PubMed](#)]
- Pontes, R.J.; Freeman, J.; Oliveira-Lima, J.W.; Hodgson, J.C.; Spielman, A. Vector densities that potentiate dengue outbreaks in a Brazilian city. *Am. J. Trop. Med. Hyg.* **2000**, *62*, 378–383. [[CrossRef](#)] [[PubMed](#)]
- World Health Organization. West Nile Virus. 2018. Available online: <http://www.who.int/news-room/fact-sheets/detail/west-nile-virus> (accessed on 3 December 2018).
- Reisen, W.K. Ecology of West Nile virus in North America. *Viruses* **2013**, *5*, 2079–2105. [[CrossRef](#)] [[PubMed](#)]

25. Bertolotti, L.; Kitron, U.D.; Walker, E.D.; Ruiz, M.O.; Brawn, J.D.; Loss, S.R.; Hamer, G.L.; Goldberg, T.L. Fine-scale genetic variation and evolution of West Nile Virus in a transmission “hot spot” in suburban Chicago, USA. *Virology* **2008**, *374*, 381–389. [[CrossRef](#)]
26. Dohm, D.J.; O’guinn, M.L.; Turell, M.J. Effect of Environmental Temperature on the Ability of *Culex pipiens* (Diptera: Culicidae) to Transmit West Nile Virus. *J. Med. Entomol.* **2002**, *39*, 221–225. [[CrossRef](#)]
27. Munn, Z.; Peters, M.D.J.; Stern, C.; Tufanaru, C.; McArthur, A.; Aromataris, E. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC Med. Res. Methodol.* **2018**, *18*, 143. [[CrossRef](#)]
28. Arksey, H.; O’Malley, L. Scoping studies: Towards a methodological framework. *Int. J. Soc. Res. Methodol.* **2005**, *8*, 19–32. [[CrossRef](#)]
29. Levac, D.; Colquhoun, H.; O’Brien, K.K. Scoping studies: Advancing the methodology. *Implement. Sci.* **2010**, *5*, 69. [[CrossRef](#)] [[PubMed](#)]
30. Stilianakis, N.I.; Syrris, V.; Petroliaqkis, T.; Pärt, P.; Gewehr, S.; Kalaitzopoulou, S.; Mourelatos, S.; Baka, A.; Pervanidou, D.; Vontas, J.; et al. Identification of Climatic Factors Affecting the Epidemiology of Human West Nile Virus Infections in Northern Greece. *PLoS ONE* **2016**, *11*, e0161510. [[CrossRef](#)] [[PubMed](#)]
31. Valiakos, G.; Pappaspyropoulos, K.; Giannakopoulos, A.; Birtsas, P.; Tsiodras, S.; Hutchings, M.R.; Spyrou, V.; Pervanidou, D.; Athanasiou, L.V.; Papadopoulos, N.; et al. Use of Wild Bird Surveillance, Human Case Data and GIS Spatial Analysis for Predicting Spatial Distributions of West Nile Virus in Greece. *PLoS ONE* **2014**, *9*, e96935. [[CrossRef](#)] [[PubMed](#)]
32. Buczak, A.L.; Koshute, P.T.; Babin, S.M.; Feighner, B.H.; Lewis, S.H. A data-driven epidemiological prediction method for dengue outbreaks using local and remote sensing data. *BMC Med. Inform. Decis. Mak.* **2012**, *12*. [[CrossRef](#)] [[PubMed](#)]
33. Arboleda, S.; Jaramillo-o, N.; Peterson, A.T. Spatial and temporal dynamics of *Aedes aegypti* larval sites in Bello, Colombia. *J. Vector Ecol.* **2012**, *37*, 37–48. [[CrossRef](#)]
34. Malahlela, O.E.; Olwoch, J.M.; Adjorlolo, C. Evaluating Efficacy of Landsat-Derived Environmental Covariates for Predicting Malaria Distribution in Rural Villages of Vhembe District, South Africa. *EcoHealth* **2018**, *15*, 23–40. [[CrossRef](#)]
35. Kanya, M.R.; Dorsey, G.; Kigozi, R.; Brownstein, J.S.; Charland, K.; Buckeridge, D.L.; Brewer, T.F.; Zinszer, K. Forecasting malaria in a highly endemic country using environmental and clinical predictors. *Malar. J.* **2015**, *14*, 245. [[CrossRef](#)]
36. Kanyangara, M.; Mamini, E.; Mharakurwa, S.; Munyati, S.; Gwanzura, L.; Kobayashi, T.; Shields, T.; Mullany, L.C.; Mutambu, S.; Mason, P.R.; et al. High-resolution plasmodium falciparum malaria risk mapping in Mutasa District, Zimbabwe: Implications for regaining control. *Am. J. Trop. Med. Hyg.* **2016**, *95*, 141–147. [[CrossRef](#)]
37. Ssempiira, J.; Kissa, J.; Nambuusi, B.; Mukooyo, E.; Opigo, J.; Makumbi, F.; Kasasa, S.; Vounatsou, P. Interactions between climatic changes and intervention effects on malaria spatio-temporal dynamics in Uganda. *Parasite Epidemiol. Control* **2018**, *3*, e00070. [[CrossRef](#)]
38. Kabaria, C.W.; Molteni, F.; Mandike, R.; Chacky, F.; Noor, A.M.; Snow, R.W.; Linard, C. Mapping intra-urban malaria risk using high resolution satellite imagery: A case study of Dar es Salaam. *Int. J. Health Geogr.* **2016**, *15*, 26. [[CrossRef](#)]
39. Giardina, F.; Franke, J.; Vounatsou, P. Geostatistical modelling of the malaria risk in Mozambique: Effect of the spatial resolution when using remotely-sensed imagery. *Geospat. Health* **2015**, *10*, 232–238. [[CrossRef](#)] [[PubMed](#)]
40. Sewe, M.O.; Ahlm, C.; Rocklöv, J. Remotely sensed environmental conditions and malaria mortality in three malaria endemic regions in western kenya. *PLoS ONE* **2016**, *11*, e0154204. [[CrossRef](#)] [[PubMed](#)]
41. Conley, A.K.; Fuller, D.O.; Haddad, N.; Hassan, A.N.; Gad, A.M.; Beier, J.C. Modeling the distribution of the West Nile and Rift Valley Fever vector *Culex pipiens* in arid and semi-arid regions of the Middle East and North Africa. *Parasites Vectors* **2014**, *7*, 289. [[CrossRef](#)] [[PubMed](#)]
42. Nmor, J.C.; Sunahara, T.; Goto, K.; Futami, K.; Sonye, G.; Akweywa, P.; Dida, G.; Minakawa, N. Topographic models for predicting malaria vector breeding habitats: Potential tools for vector control managers. *Parasites Vectors* **2013**, *6*, 14. [[CrossRef](#)] [[PubMed](#)]

43. Sarfraz, M.S.; Tripathi, N.K.; Faruque, F.S.; Bajwa, U.I.; Kitamoto, A.; Souris, M. Mapping urban and peri-urban breeding habitats of *Aedes* mosquitoes using a fuzzy analytical hierarchical process based on climatic and physical parameters. *Geospat. Health* **2014**, *8*, S685–S697. [[CrossRef](#)] [[PubMed](#)]
44. Machault, V.; Yébakima, A.; Etienne, M.; Vignolles, C.; Palany, P.; Tourre, Y.; Guérécheau, M.; Lacaux, J.P. Mapping Entomological Dengue Risk Levels in Martinique Using High-Resolution Remote-Sensing Environmental Data. *ISPRS Int. J. -Geo-Inf.* **2014**, *3*, 1352–1371. [[CrossRef](#)]
45. Amek, N.; Bayoh, N.; Hamel, M.; Lindblade, K.A.; Gimnig, J.E.; Odhiambo, F.; Laserson, K.F.; Slutsker, L.; Smith, T.; Vounatsou, P. Spatial and temporal dynamics of malaria transmission in rural Western Kenya. *Parasites Vectors* **2012**, *5*, 86. [[CrossRef](#)] [[PubMed](#)]
46. Diboulo, E.; Sié, A.; Diadier, D.A.; Voules, D.A.K.; Yé, Y.; Vounatsou, P. Bayesian variable selection in modelling geographical heterogeneity in malaria transmission from sparse data: An application to Nouna Health and Demographic Surveillance System (HDSS) data, Burkina Faso. *Parasites Vectors* **2015**, *8*, 118. [[CrossRef](#)] [[PubMed](#)]
47. Shaukat, A.M.; Breman, J.G.; McKenzie, F.E. Using the entomological inoculation rate to assess the impact of vector control on malaria parasite transmission and elimination. *Malar. J.* **2010**, *9*, 122. [[CrossRef](#)]
48. Adde, A.; Roux, E.; Mangeas, M.; Dessay, N.; Nacher, M.; Dusfour, I.; Girod, R.; Briolant, S. Dynamical mapping of anopheles darlingi densities in a residual malaria transmission area of French guiana by using remote sensing and meteorological data. *PLoS ONE* **2016**, *11*, e0164685. [[CrossRef](#)]
49. Rosà, R.; Marini, G.; Bolzoni, L.; Neteler, M.; Metz, M.; Delucchi, L.; Chadwick, E.; Balbo, L.; Mosca, A.; Giacobini, M.; et al. Early warning of West Nile virus mosquito vector: Climate and land use models successfully explain phenology and abundance of *Culex pipiens* mosquitoes in north-western Italy. *Parasites Vectors* **2014**, *7*. [[CrossRef](#)] [[PubMed](#)]
50. German, A.; Espinosa, M.O.; Abril, M.; Scavuzzo, C.M. Exploring satellite based temporal forecast modelling of *Aedes aegypti* oviposition from an operational perspective. *Remote. Sens. Appl. Soc. Environ.* **2018**, *11*, 231–240. [[CrossRef](#)]
51. Scavuzzo, J.M.; Trucco, F.; Espinosa, M.; Tauro, C.B.; Abril, M.; Scavuzzo, C.M.; Frery, A.C. Modeling Dengue vector population using remotely sensed data and machine learning. *Acta Trop.* **2018**, *185*, 167–175. [[CrossRef](#)] [[PubMed](#)]
52. Benali, A.; Nunes, J.P.; Freitas, F.B.; Sousa, C.A.; Novo, M.T.; Lourenço, P.M.; Lima, J.C.; Seixas, J.; Almeida, A.P. Satellite-derived estimation of environmental suitability for malaria vector development in Portugal. *Remote. Sens. Environ.* **2014**, *145*, 116–130. [[CrossRef](#)]
53. Mokraoui, L.; Noor, N.; Abdullah, A. Developing dengue index through the integration of crowdsourcing approach (X-Waba). *IOP Conf. Ser. Earth Environ. Sci.* **2018**, 169. [[CrossRef](#)]
54. Lessler, J.; Azman, A.S.; Grabowski, M.K.; Salje, H.; Rodriguez-Barraquer, I. Trends in the Mechanistic and Dynamic Modeling of Infectious Diseases. *Curr. Epidemiol. Rep.* **2016**, *3*, 212–222. [[CrossRef](#)]
55. Parham, P.E.; Waldock, J.; Christophides, G.K.; Hemming, D.; Agosto, F.; Evans, K.J.; Fefferman, N.; Gaff, H.; Gumel, A.; Ladeau, S.; et al. Climate, environmental and socio-economic change: Weighing up the balance in vector-borne disease transmission. *Philos. Trans. R. Soc. Lond. Ser. Biol. Sci.* **2015**, *370*, 20130551. [[CrossRef](#)] [[PubMed](#)]
56. Freitas, S.C.; Trigo, I.F.; Macedo, J.; Barroso, C.; Silva, R.; Perdigão, R. Land surface temperature from multiple geostationary satellites. *Int. J. Remote Sens.* **2013**, *34*, 3051–3068. [[CrossRef](#)]
57. Weiss, D.J.; Bhatt, S.; Mappin, B.; Van Boeckel, T.P.; Smith, D.L.; Hay, S.I.; Gething, P.W. Air temperature suitability for *Plasmodium falciparum* malaria transmission in Africa 2000–2012: A high-resolution spatiotemporal prediction. *Malar. J.* **2014**, *13*, 171. [[CrossRef](#)]
58. Albergel, C.; Dutra, E.; Muñoz-Sabater, J.; Haiden, T.; Balsamo, G.; Beljaars, A.; Isaksen, L.; de Rosnay, P.; Sandu, I.; Wedi, N. Soil temperature at ECMWF: An assessment using ground-based observations. *J. Geophys. Res.* **2015**, *120*, 1361–1373. [[CrossRef](#)]
59. Méndez-Lázaro, P.; Muller-Karger, F.E.; Otis, D.; McCarthy, M.J.; Peña-Orellana, M.; Méndez-Lázaro, P.; Muller-Karger, F.E.; Otis, D.; McCarthy, M.J.; Peña-Orellana, M. Assessing Climate Variability Effects on Dengue Incidence in San Juan, Puerto Rico. *Int. J. Environ. Res. Public Health* **2014**, *11*, 9409–9428. [[CrossRef](#)] [[PubMed](#)]

60. Laureano-Rosario, A.E.; Garcia-Rejon, J.E.; Gomez-Carro, S.; Farfan-Ale, J.A.; Muller-Karger, F.E. Modelling dengue fever risk in the State of Yucatan, Mexico using regional-scale satellite-derived sea surface temperature. *Acta Trop.* **2017**, *172*, 50–57. [[CrossRef](#)] [[PubMed](#)]
61. 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.; et al. The global distribution and burden of dengue. *Nature* **2013**, *496*, 504–507. [[CrossRef](#)] [[PubMed](#)]
62. Benedum, C.M.; Seidahmed, O.M.E.; Eltahir, E.A.B.; Markuzon, N. Statistical modeling of the effect of rainfall flushing on dengue transmission in Singapore. *PLOS Neglected Trop. Dis.* **2018**, *12*, e0006935. [[CrossRef](#)] [[PubMed](#)]
63. Thakur, S.; Dharavath, R. Artificial neural network based prediction of malaria abundances using big data: A knowledge capturing approach. *Clin. Epidemiol. Glob. Health* **2018**, *7*, 121–126. [[CrossRef](#)]
64. Ashby, J.; Moreno-Madriñán, M.M.J.; Yiannoutsos, C.T.C.; Stanforth, A. Niche modeling of dengue fever using remotely sensed environmental factors and boosted regression trees. *Remote Sens.* **2017**, *9*. [[CrossRef](#)]
65. Midekisa, A.; Senay, G.; Henebry, G.M.; Semuniguse, P.; Wimberly, M.C. Remote sensing-based time series models for malaria early warning in the highlands of Ethiopia. *Malar. J.* **2012**, *11*, 165. [[CrossRef](#)] [[PubMed](#)]
66. Sewe, M.O.; Tozan, Y.; Ahlm, C.; Rocklöv, J. Using remote sensing environmental data to forecast malaria incidence at a rural district hospital in Western Kenya. *Sci. Rep.* **2017**, *7*, 2589. [[CrossRef](#)]
67. Merkord, C.L.; Liu, Y.; Mihretie, A.; Gebrehiwot, T.; Awoke, W.; Bayabil, E.; Henebry, G.M.; Kassa, G.T.; Lake, M.; Wimberly, M.C. Integrating malaria surveillance with climate data for outbreak detection and forecasting: The EPIDEMIA system. *Malar. J.* **2017**, *16*, 1–89. [[CrossRef](#)]
68. Pettorelli, N.; Ryan, S.; Mueller, T.; Bunnefeld, N.; Jedrzejewska, B.; Lima, M.; Kausrud, K. The Normalized Difference Vegetation Index (NDVI): Unforeseen successes in animal ecology. *Clim. Res.* **2011**, *46*, 15–27. [[CrossRef](#)]
69. Liu, J.; Chen, X.P. Relationship of Remote Sensing Normalized Differential Vegetation Index to Anopheles Density and Malaria Incidence Rate. *Biomed. Environ. Sci.* **2006**, *19*, 130–132. [[PubMed](#)]
70. Huete, A.; Didan, K.; Miura, T.; Rodriguez, E.; Gao, X.; Ferreira, L. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sens. Environ.* **2002**, *83*, 195–213. [[CrossRef](#)]
71. Ruangudomsakul, C.; Duangsin, A.; Kerdprasop, K.; Kerdprasop, N. Application of Remote Sensing Data for Dengue Outbreak Estimation Using Bayesian Network. *Int. J. Mach. Learn. Comput.* **2018**, *8*. [[CrossRef](#)]
72. Chabot-Couture, G.; Nigmatulina, K.; Eckhoff, P. An environmental data set for vector-borne disease modeling and epidemiology. *PLoS ONE* **2014**, *9*. [[CrossRef](#)] [[PubMed](#)]
73. Cao, G.; Han, D.; Song, X. Evaluating actual evapotranspiration and impacts of groundwater storage change in the North China Plain. *Hydrol. Process.* **2014**, *28*, 1797–1808. [[CrossRef](#)]
74. Tsouni, A.; Kontoes, C.; Koutsoyiannis, D.; Elias, P.; Mamassis, N. Estimation of Actual Evapotranspiration by Remote Sensing: Application in Thessaly Plain, Greece. *Sensors* **2008**, *8*, 3586–3600. [[CrossRef](#)] [[PubMed](#)]
75. Chuang, T.W.; Wimberly, M. Remote Sensing of Climatic Anomalies and West Nile Virus Incidence in the Northern Great Plains of the United States. *PLoS ONE* **2012**, *7*. [[CrossRef](#)]
76. Bui, Q.T.; Nguyen, Q.H.; Pham, V.M.; Pham, M.H.; Tran, A.T. Understanding spatial variations of malaria in Vietnam using remotely sensed data integrated into GIS and machine learning classifiers. *Geocarto Int.* **2018**, *6049*, 1–15. [[CrossRef](#)]
77. Kanyangarara, M.; Mamini, E.; Mharakurwa, S.; Munyati, S.; Gwanzura, L.; Kobayashi, T.; Shields, T.; Mullany, L.C.; Mutambu, S.; Mason, P.R.; et al. Reduction in malaria incidence following indoor residual spraying with actellic 300 CS in a setting with pyrethroid resistance: Mutasa District, Zimbabwe. *PLoS ONE* **2016**, *11*, e0151971. [[CrossRef](#)]
78. Homan, T.; Maire, N.; Hiscox, A.; Di Pasquale, A.; Kiche, I.; Onoka, K.; Mweresa, C.; Mukabana, W.R.; Ross, A.; Smith, T.A.; et al. Spatially variable risk factors for malaria in a geographically heterogeneous landscape, western Kenya: An explorative study. *Malar. J.* **2016**, *15*, 1. [[CrossRef](#)]
79. Atieli, H.E.; Zhou, G.; Lee, M.C.; Kweka, E.J.; Afrane, Y.; Mwanzo, I.; Githeko, A.K.; Yan, G. Topography as a modifier of breeding habitats and concurrent vulnerability to malaria risk in the western Kenya highlands. *Parasites Vectors* **2011**, *4*, 241. [[CrossRef](#)] [[PubMed](#)]

80. Watts, A.G.; Miniota, J.; Joseph, H.A.; Brady, O.J.; Kraemer, M.U.G.; Grills, A.W.; Morrison, S.; Esposito, D.H.; Nicolucci, A.; German, M.; et al. Elevation as a proxy for mosquito-borne Zika virus transmission in the Americas. *PLoS ONE* **2017**, *12*, e0178211. [[CrossRef](#)]
81. Young, S.G.; Tullis, J.A.; Cothren, J. A remote sensing and GIS-assisted landscape epidemiology approach to West Nile virus. *Appl. Geogr.* **2013**, *45*, 241–249. [[CrossRef](#)]
82. Lowe, R.; Bailey, T.C.; Stephenson, D.B.; Jupp, T.E.; Graham, R.J.; Barcellos, C.; Carvalho, M.S. The development of an early warning system for climate-sensitive disease risk with a focus on dengue epidemics in Southeast Brazil. *Stat. Med.* **2013**, *32*, 864–883. [[CrossRef](#)] [[PubMed](#)]
83. Marcantonio, M.; Rizzoli, A.; Metz, M.; Rosà, R.; Marini, G.; Chadwick, E.; Neteler, M. Identifying the Environmental Conditions Favouring West Nile Virus Outbreaks in Europe. *PLoS ONE* **2015**, *10*, e0121158. [[CrossRef](#)] [[PubMed](#)]
84. Amadi, J.A.; Olago, D.O.; Ong'amo, G.O.; Oriaso, S.O.; Nanyingi, M.; Nyamongo, I.K.; Estambale, B.B. Sensitivity of vegetation to climate variability and its implications for malaria risk in Baringo, Kenya. *PLoS ONE* **2018**, *13*. [[CrossRef](#)]
85. Yue, Y.; Sun, J.; Liu, X.; Ren, D.; Liu, Q.; Xiao, X.; Lu, L. 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. *Int. J. Infect. Dis.* **2018**, *75*, 39–48. [[CrossRef](#)]
86. Tran, A.; Sudre, B.; Paz, S.; Rossi, M.; Desbrosse, A.; Chevalier, V.; Semenza, J. Environmental predictors of West Nile fever risk in Europe. *Int. J. Health Geogr.* **2014**, *13*. [[CrossRef](#)]
87. Hii, Y.L.; Zhu, H.; Ng, N.; Ng, L.C.; Rocklöv, J. Forecast of Dengue Incidence Using Temperature and Rainfall. *PLoS Neglected Trop. Dis.* **2012**, *6*, e1908. [[CrossRef](#)]
88. Nizamuddin, M.; Kogan, F.; Dhiman, R.; Guo, W.; Roytman, L. Modeling and Forecasting Malaria in Tripura, INDIA using NOAA/AVHRR-Based Vegetation Health Indices. *Int. J. Remote. Sens. Appl.* **2013**, *3*, 108–116.
89. Quintero, J.; Carrasquilla, G.; Suárez, R.; González, C.; Olano, V.A. An ecosystemic approach to evaluating ecological, socioeconomic and group dynamics affecting the prevalence of *Aedes aegypti* in two Colombian towns. *Cadernos de Saúde Pública* **2009**, *25*, s93–s103. [[CrossRef](#)] [[PubMed](#)]
90. Quintero, J.; Brochero, H.; Manrique-Saide, P.; Barrera-Pérez, M.; Basso, C.; Romero, S.; Caprara, A.; De Lima Cunha, J.C.; Beltrán - Ayala, E.; Mitchell-Foster, K.; et al. Ecological, biological and social dimensions of dengue vector breeding in five urban settings of Latin America: A multi-country study. *BMC Infect. Dis.* **2014**, *14*, 38. [[CrossRef](#)]
91. Ayala, R.G.; Estrugo, A. *Assessing the Effects of Climate and Socioeconomic Factors on Vulnerability to Vector-Borne Diseases in Latin America*; Inter-American Development Bank: Washington, DC, USA, 2014.
92. Monroe, A.; Asamoah, O.; Lam, Y.; Koenker, H.; Psychas, P.; Lynch, M.; Ricotta, E.; Hornston, S.; Berman, A.; Harvey, S.A. Outdoor-sleeping and other night-time activities in northern Ghana: Implications for residual transmission and malaria prevention. *Malar. J.* **2015**, *14*, 35. [[CrossRef](#)] [[PubMed](#)]
93. Sarfraz, M.S.; Tripathi, N.K.; Kitamoto, A. Near real-time characterisation of urban environments: A holistic approach for monitoring dengue fever risk areas. *Int. J. Digit. Earth* **2014**, *7*, 916–934. [[CrossRef](#)]
94. Generalized Linear Models and Generalized Additive Models 13.1 Generalized Linear Models and Iterative Least Squares. Available online: <https://www.stat.cmu.edu/~cshalizi/uADA/12/lectures/ch13.pdf> (accessed on 13 May 2019).
95. Adimi, F.; Soebiyanto, R.P.; Safi, N.; Kiang, R. Towards malaria risk prediction in Afghanistan using remote sensing. *Malar. J.* **2010**, *9*, 125. [[CrossRef](#)] [[PubMed](#)]
96. De Oliveira, E.C.; Dos Santos, E.S.; Zeilhofer, P.; Souza-Santos, R.; Atanaka-Santos, M. Geographic information systems and logistic regression for high-resolution malaria risk mapping in a rural settlement of the southern Brazilian Amazon. *Malar. J.* **2013**, *12*, 420. [[CrossRef](#)] [[PubMed](#)]
97. Shumway, R.H.; Stoffer, D.S. *Time Series Analysis and Its Applications: With R Examples*, 3rd ed.; Springer: New York, NY, USA, 2011.
98. Zinszer, K.; Verma, A.D.; Charland, K.; Brewer, T.F.; Brownstein, J.S.; Sun, Z.; Buckeridge, D.L. A scoping review of malaria forecasting: Past work and future directions. *BMJ Open* **2012**, *2*, e001992. [[CrossRef](#)]
99. Wang, S.; Feng, J.; Liu, G. Application of seasonal time series model in the precipitation forecast. *Math. Comput. Model.* **2013**, *58*, 677–683. [[CrossRef](#)]
100. Khameneh, N.J. Machine Learning for Disease Outbreak Detection using Probabilistic Models. Ph.D. Thesis, École Polytechnique de Montréal, Montreal, QC, Canada, 2014.

101. Li, Z.; Roux, E.; Dessay, N.; Girod, R.; Stefani, A.; Nacher, M.; Moiret, A.; Seyler, F. 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 Sens.* **2016**, *8*, 319. [[CrossRef](#)]
102. Catry, T.; Li, Z.; Roux, E.; Herbreteau, V.; Révillion, C.; Dessay, N. Fusion of SAR and optical imagery for studying the ecoepidemiology of vector-borne diseases in tropical countries. In Proceedings of the 2016 European Space Agency Living Planet Symposium, Prague, Czech Republic, 9–13 May 2016.
103. Catry, T.; Li, Z.; Roux, E.; Herbreteau, V.; Gurgel, H.; Mangeas, M.; Seyler, F.; Dessay, N.; Catry, T.; Li, Z.; et al. Wetlands and Malaria in the Amazon: Guidelines for the Use of Synthetic Aperture Radar Remote-Sensing. *Int. J. Environ. Res. Public Health* **2018**, *15*, 468. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).