

The world in your hands: GeoHealth then and now

Robert Bergquist,¹ Samuel Manda^{2,3}

¹Ingerod, Brastad, Sweden; ²Biostatistics Unit, South African Medical Research Council, Pretoria, South Africa; ³Department of Statistics, University of Pretoria, Pretoria, South Africa

Abstract

Infectious diseases transmitted by vectors/intermediate hosts constitute a major part of the economic burden related to public health in the endemic countries of the tropics, which challenges local welfare and hinders development. The World Health Organization, in partnership with pharmaceutical companies, major donors, endemic countries and non-governmental organizations, aims to eliminate the majority of these infections in the near future. To succeed, the ecological requirements and real-time distributions of the causative agents (bacteria, parasites and viruses) and their vectors must not only be known to a high degree of accuracy, but the data must also be updated more rapidly than has so far been the case. Current approaches include data collection through terrestrial capture on site and satellite-generated information. This article provides an update of currently available sources of remotely-sensed data, including specific information on satellite-borne sensors, and how such data can be handled by Geographical Information Systems (GIS). Computers, when equipped with GIS software based on common spatial denominators, can connect remotely-sensed environmental records with terrestrial-captured data and apply spatial statistics in ways uniquely suited to manage control activities in areas where vector-borne infections dominate.

Correspondence: Robert Bergquist, Ingerod, Brastad, Sweden.

E-mail: editor@geospatialhealth.net

Key words: Geographical Information Systems; Remote sensing; Satellites; Sensors; Environment; BioClim; WorldClim; Spatial statistics.

Contributions: RB conceived and outlined the paper and also produced the tables and figures. RB and SM participated equally in creating the detailed review.

Conflict of interest: the authors declare no potential conflict of interest.

Funding: none.

Received for publication: 9 April 2019.

Accepted for publication: 2 May 2019.

©Copyright R. Bergquist and S. Manda, 2019

Licensee PAGEPress, Italy

Geospatial Health 2019; 14:779

doi:10.4081/gh.2019.779

This article is distributed under the terms of the Creative Commons Attribution Noncommercial License (CC BY-NC 4.0) which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

Introduction

The Romans may or may not have given us the word *malaria*, but the 18th century Italians knew the significance of locations characterized by what they called bad air, *i.e.* malaria. The notion of location was also of interest for the English doctor, Edward Jenner, who noted that smallpox seemed to spare farms with cattle. Through more detailed study, he found that people who had regular contact with cows, *e.g.*, milking them, generally developed a type of skin rash called cowpox. More importantly, those who had had cowpox did not get the dreaded disease smallpox, an observation that seemed to hold for the rest of their lives (Jenner, 1798; Jenson *et al.*, 2016). As the notion of infectious agents had not yet entered medical knowledge, Jenner's approach was predominantly spatial. Indeed, the importance of location became absolutely clear about 50 years later, when the London physician John Snow focused on the spatial variation of cholera cases in a short medical communication (Snow, 1849). However, he was not alone in his thinking as shown by a contemporary colleague who tried to localize the culprit in the Exeter 1932 cholera epidemic by mapping the distribution of cases in the city (Shapter, 1849). Again, both Snow and Shapter lived before the link between infectious diseases and microorganisms had been unravelled and the spread of cholera in England at the time was believed to be due to *bad air*. Snow was eventually credited for solving the quandary by realizing that the problem was not air but rather water. This thought came to him when he noted that most of the homes of people struck with cholera in London in 1854 were concentrated around a specific water pump in Soho (Figure 1). He developed his train of thoughts about the epidemic in a more substantial report (Snow, 1855) with the same title as that of his brief treatise of 1849. Foreshadowing epidemiology as a discipline and, in actual fact also what we now call Geographical Information Systems (GIS), Snow was vindicated when the epidemic promptly disappeared after he managed to convince the authorities to close the Broad Street water pump situated at the geographical centre of the epidemic. The amazing technical progress linked to GIS emerges when we *fast-forward* to the present time. The convergence of Earth sciences, satellite development, computer technology and spatial statistics has not only produced Earth-observing satellites and an ever-increasing stream of remotely-sensed information but has also provided tools needed to manage all these new data. Thanks to GIS, we can now deal rapidly with the proliferation of diseases, in particular the major pandemics spread by mosquitoes, flies, ticks and snails. The GIS approach has contributed to detailed risk maps and Early Warning Systems (EWS) for a large number of vector-borne infections, including reference to birds carrying influenza viruses over large distances. Today, GIS, remote-sensing and Global Positioning Systems (GPS) are seen as part of the epidemiologist's tool box, and few scientists working in the field of public health can manage without them.

The evolving geographical information systems platform

As already stated, the roots of GIS go back to Snow, though in fact they actually go a little further still. Although the *Atlas of Diseases*, published in 1837 by Heinrich Berghaus in Germany (Berghaus, 1837) emphasizes isothermal lines, winds, climatic zones, *etc.* with only an incidental discussion of endemic and epidemic diseases, the book is a landmark in the history of medical cartography as it synthesizes medical mapping of earlier periods and provides an impetus to specific epidemiological mapmaking. That winds and climate information are included is interesting as this builds on Greek thinking in the Antiquity and predates today's environmental focus. The final decade of the 19th century presents firmer ground for the discussion of epidemics, as the role of parasites, bacteria and even viruses had by then become understood. For example, following Snow's idea, Stiles (1903) noted that hookworm infection in the southern parts of the United States (U.S.) was restricted to sandy districts and could not be found where clay is more common. This kind of epidemiological approach, novel then, served to prepare the medical profession for the compelling cause of disease mapping as a way of controlling epidemics. Administrative and cartographic infrastructures made it possible to map the spread of infectious diseases, but map production was still not as straightforward as printing text, which explains why the tabular format was generally used for recording number of cases and areas at risk. Although tables and text can contain the same information as a map, they are less easily exploited for analysis since the reading and interpretation of tables are both laborious and time-consuming. As map-making improved, sets of disease maps enabled rapid visualization of changes in distribution frequency and clusters of infections, but they did not become as useful in practise as they are today when they can instantaneously be called up on the computer screen.

Working on a land inventory project in Canada in the 1960s, Roger Tomlinson coined the term GIS, first seen in print in the published version of his 1968 lecture entitled *A geographic information*

system for regional planning given at a symposium organized by the Commonwealth Scientific and Industrial Research Organization. The idea became better known after he had developed it further in his doctoral thesis presented at the University of London in 1974 (Tomlinson, 1974). Meanwhile, the Environmental Systems Research Institute (ESRI) in Redlands, California (<https://www.esri.com/>) was using topographic maps produced by the U.S. Geological Survey (USGS) in the form of transparent overlays of gridded sheets with separate datasets, such as geology, soil type, topography, *etc.* The computerized version of this approach took off in the mid-1970s when San Diego County in California selected ESRI to develop a system for a land inventory reminiscent of Tomlinson's work in Canada. In a major departure from the previous grid-based overlay approach, ESRI now pioneered the use of vector graphics together with raster-format geographical maps as the basic framework, which led to the 1982 release of ARC/INFO, the first commercial GIS application. Spurred on by the implementation of several successful projects leading to validation of new features, such as GPS access, faster computers, network processing and electronic data publishing, ESRI underwent a period of rapid growth producing a modular, scalable platform. The new desktop product, called ArcGIS, had a major impact on the industry when it was released in 1999 contributing to a wider use of GIS. Moving into the new millennium, many competitors, including open-source GIS, appeared in the shape of Quantum GIS and other applications compatible with Linux, Unix, Mac OS X and Windows resulting in the supply of a mix of commercial and free GIS software. Data collected in the field can be bundled together with environmental information in a GIS visualizing the situation in the form of different overlays against a background canvas, commonly a geographical map. GIS is a highly versatile system that captures, integrates, stores, manages, analyzes and displays data so that models can be constructed and tested under various scenarios. Buffers and different spatial parameters supporting analysis of single overlays, neighbourhoods or larger geographical areas are routinely included in a GIS and data can either be presented as vectors or rasters, where the former are composed of paths (lines) and polygons, and the latter of pixels (Figure 2). The vector format is

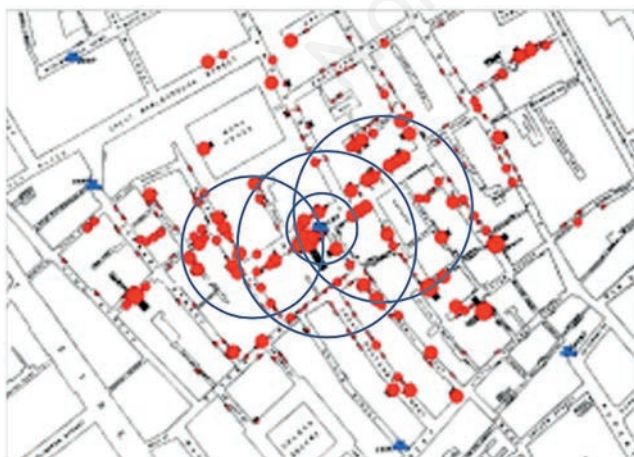


Figure 1. Snow's approach: finding the circle with the highest ratio of cases in relation to its diameter. The figure shows Broad Street pump in London's Soho area at the corner of Broad Street and Lexington Street in the map's centre.

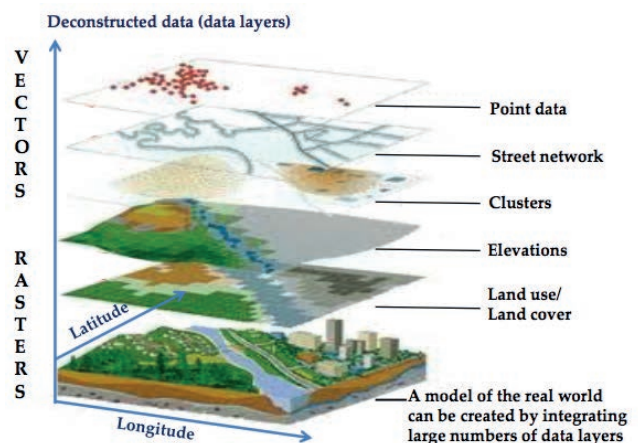


Figure 2. The Geographical Information Systems approach.

based on tables, stored separately in the computer, containing the coordinates of points together with instructions, *e.g.*, how to find points separating those belonging to common sets from those that are independent (Yang *et al.*, 2005; Rinaldi *et al.*, 2006). Geographical base maps as well as continuous phenomena, such as land cover, digital elevation models and climatological information are generally in the raster format but can also be produced by carpets of adjacent polygon cells providing a virtual cover of a given part of a territory. Spatial and temporal points of interest relate to one another in a 3-dimensional arena with time added as a fourth dimension. That means that locations in space may be recorded by coordinates representing longitude, latitude and elevation, while change of location invokes spatio-temporal interpretations. Using location as the key index, data may be visualized in ways that potentially can reveal hidden relationships and patterns, a concept that can open new avenues of scientific inquiry (Rinaldi *et al.*, 2006, 2011). Although GIS can refer to a number of different environments and processes, the overview presented here focuses only on health applications. The reason for the particular strength of GIS with respect to work involving vector-borne diseases is that the environmental requirements of the vectors limit these infections to highly specific ecological niches (Malone *et al.*, 2016). Additionally, the distribution of vector-borne diseases is restricted, not only by the vectors' requirements but also by the ambient temperature inside them as temperature governs the speed of maturation of the parasite stages inside (Yang *et al.*, 2007; Zhou *et al.*, 2008).

Virtual globes, available from Google and other companies, are superior to conventional maps as they allow importation and annotation of datasets of specific interest, *e.g.* environmental parameters, demographic information, *etc.* Although the basic analytical capabilities of virtual globe applications are limited, Google

Earth and its web-based pendant Google Maps are conducive to the combination of results from literature and findings from the field and have therefore a strong potential for data sharing and dissemination of scientific research results as shown by many authors (Stensgaard *et al.*, 2009; Yang *et al.*, 2012; Twardzik *et al.*, 2018).

The concept of *time geography*, advanced by the Swedish geographer Hägerstrand (1970) and chronicled by Ellegård and Svedin (2012), culminated in the creation of the space-time cube (STC), where the base represents a 2-D geographical map with height, normally used for expressing altitude, is instead treated as the time axis. This produces a volume representing merged space-time in which internal objects, *e.g.*, study subjects, move along *world lines* and events are shown as points representing both space and time. This amazing spacetime concept where the time component emerges as a series of spacetime slices, first introduced by Einstein in his relativity theory, is indeed the way Hägerstrand's STC operates. As proposed by Gatalisky *et al.* (2004), the fusion of space and time and the use of appropriate spatial analytical visualization techniques, such as GIS, facilitate understanding. Despite its obvious advantages, this great idea was sidelined due to the immense graphic preparations required at the time when Hägerstrand was developing the technique. However, this drawback disappeared when computers arrived and graphic presentations could easily be created by rapid switches between data packages and geographical perspectives. The revived interest in Hägerstrand's innovative idea (Hedley *et al.*, 1999) eventually led to an ESRI application building on the STC approach (<http://desktop.arcgis.com/en/arcmap/10.3/tools/space-time-pattern-mining-toolbox/visualizing-cube-data.htm>), which is strongly recommended.

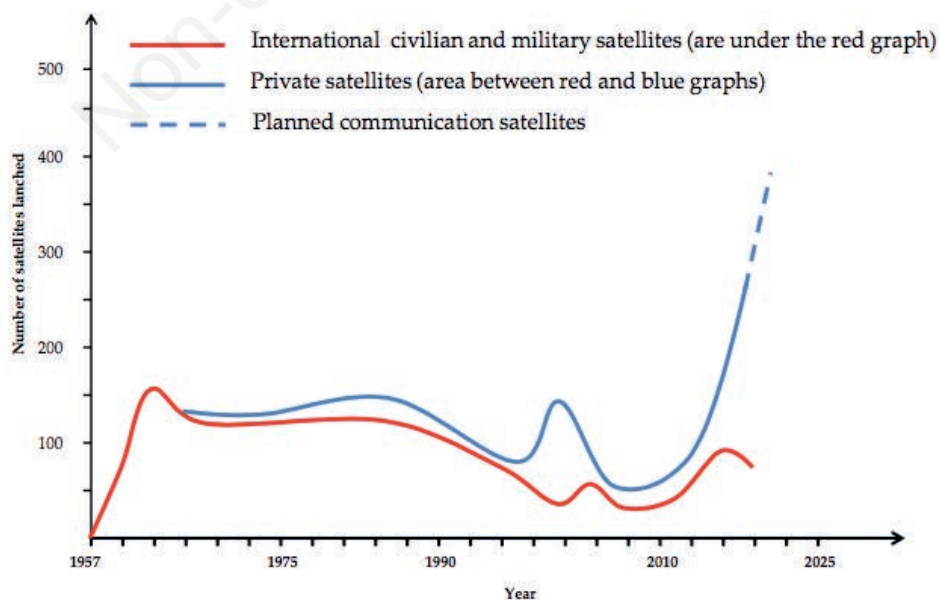


Figure 3. Schematic view of annual satellite launches 1957-2018.



The view from above

The launch of Sputnik on 4 October 1957 initiated a race in the sky where the U.S. the following year launched an ambitious programme that included satellite missions jointly managed by the new National Aeronautics and Space Administration (NASA) and the USGS, while the National Oceanic and Atmospheric Administration (NOAA) engaged in weather studies sparking off the longest-lived series of Earth-observing satellites started by the launch of the Television Infrared Observation Satellite (TIROS) on 1 April 1960. The easily counted satellites of the first *space decade* have sixty years later mushroomed into a cloud of approximately 5,000. The online satellite index kept by the United Nations Office for Outer Space Affairs (UNOOSA) (http://www.unoosa.org/oosa/osoindex/index.jsp?lf_id=) provides the exact number at any given date. The GPS network and the new communication network, currently being planned and launched by several international companies, constitute essential infrastructure building blocks but are not further discussed here; also the military satellites fall outside the scope of this article. What is of interest, however, is the recent growth of privately launched satellites that is expected to double the current number of airborne satellites by 2027 (Figure 3).

Optical Earth-observation matured with Landsat, a joint NASA/USGS programme with a mission to establish and execute a strategy of repetitive observations of Earth's total land mass,

including islands, coastal boundaries and coral reefs. The current Landsat satellites have a near-polar, sun-synchronized orbit at an altitude of 705 km, which allows them to scan all areas of the globe with a spatial resolution of 1.1 km. The unbroken Landsat record since 1972 makes this programme the world's longest running system of moderate-resolution, remotely-sensed imagery satellites (<https://landsat.gsfc.nasa.gov/>). Landsat 1-3 were equipped with the Multi Spectral Scanner, replaced with the Thematic Mapper onboard Landsat 4-5. Landsat-6 (lost at launch) and Landsat 7 were equipped with the Extended Thematic Mapper (ETM) and the ETM+ together with a thermal single-channel sensor. These instruments were finally replaced with the Operational Land Imager (OLI) and the dual Thermal Infrared Sensor (TIRS) onboard Landsat-8, while a still more sophisticated Landsat 9 is planned for launch in 2020.

NOAA records weather changes and has at least two Polar-Orbiting Environmental Satellites (POES) in orbit at all times thereby achieving twice-daily global coverage ensuring that no region has data older than six hours. They orbit Earth at an altitude of around 850 km carrying onboard the 6-channel Advanced Very High Resolution Radiometer (AVHRR/3), a much used data collection system with applications in meteorology, climatology, oceanography and land cover with emphasis on regional and global environmental issues (<https://poes.gsfc.nasa.gov/avhrr3.html>). MetOp, a group of three polar-orbiting meteorological satellites

Table 1. High-resolution imagery satellites.

Satellite	Launch date	Panchrom. - multispec. resolutions	Owner/collaborator as of December 2018	Comment
Dove*	2016-2018	3-m multispectral	Planet Labs, USA	Low-flying (400-km altitude) mini-satellites consisting of 87 satellites
IKONOS	24 Sept. 1999	0.82 m-3.28 m.	Maxar Technologies, Canada	Fifteen-year lifetime (1999-2015). Imagery records remain available
Quickbird	18 Oct. 2001	0.61 m-2.4 m	Maxar Technologies, Canada	Five-year lifetime (2001-2014). Retired on Jan. 27 2015. Imagery records remain available
GeoEye	6 Sept. 2008	0.41 m-1.65 m 0.46 m-1.84 m	Maxar Technologies, Canada	In 2013 the orbit altitude was raised to 770 km reducing the resolution slightly
RapidEye	29 Aug. 2008	6.5 m (panchromatic)	Planet Labs, Inc, USA	Rapid daily revisit suits vegetation and disaster applications due to high sensitivity for chlorophyll
SPOT-5	4 May 2002	2.5-5 m-10 m	Created by CNES Operated by SPOT Image, France	Explores Earth's resources, climate, oceanography and monitors natural phenomena
SPOT-6	9 Sept. 2012	1.5 m-6 m	Created by CNES Operated by SPOT Image, France	Spot-6 and 7 are phased into the same orbit as the two Pléiades satellites to improve revisit time
SPOT-7	30 June 2014	1.5 m-6 m	Space Agency of Azerbaijan	Sold to Azerbaijan's Space Agency in December 2014 and renamed AzerSky. Main use in agriculture
Pléiades HR 1A Pléiades HR 1B	17 Dec. 2011 2 Dec. 2012	0.46 m-1.84 m	CNES SPOT Image distribution of data	SPOT/Pléiades constellation images can be requested less than six hours beforehand
WorldView-1	18 Sept. 2007	Panchromatic: 0.46 m	Maxar Technologies, Canada	Strong geo-location capabilities with targeting and efficient in-track stereo collection
WorldView-2	8 Oct. 2009	0.46 m-1.84 m	Maxar Technologies, Canada	Capability to collect optical data through clouds and smoke
WorldView-3	13 Aug. 2014	0.31 m-1.24 m	Maxar Technologies, Canada	First satellite delivering close to 30-cm resolution imagery
WorldView-4	11 Nov. 2016	0.31 m-1.24 m	Maxar Technologies, Canada	Previously named GeoEye-2. Malfunction due to the gyro system in Jan. 2019 stops imagery production

SPOT, Satellite Pour l'Observation de la Terre; CNES, Centre National d'Etudes Spatiales, France. *The Dove satellites are part of satellite network.

developed by the European Space Agency (ESA) and operated by the European Organization for the Exploitation of Meteorological Satellites (EUMETSAT), forms a component of the overall EUMETSAT Polar System (<https://www.eumetsat.int/website/home/Satellites/CurrentSatellites/Metop/index.html>). The MetOP satellites constitute the European contribution to the EUMETSAT/NOAA Initial Joint Polar System (IJPS) with a payload comprising 11 scientific instruments plus two Search and Rescue services. The NOAA Suomi NPP satellite (https://www.nasa.gov/mission_pages/NPP/main/index.html), launched in October 2011 as a stop-gap between the POES and the new IJPS successor has a sun-synchronized orbit at an altitude of 824 km equipped with the Visible Infrared Imaging Radiometer Suite providing measurements related to the climate that was initiated by the Earth-Observing System (EOS) operated by NASA.

NOAA also operates high-altitude weather satellites which originate from NASA's Synchronous Meteorological Satellite programme that launched two satellites providing imagery through a Visible and Infrared Spin Scan Radiometer in 1974/1975. They were replaced by NOAA's own Geostationary Operational Environmental Satellites (GOES), the first of which was launched in October 1975. To stay fixed relative to Earth, the GOES must match Earth's rotation which demands the specific altitude of 35,786 km, high enough to cover half the planet. Two satellites, *GOES East* and *GOES West*, are placed at the equatorial plane allowing global coverage (<https://www.goes.noaa.gov/>). Although the imagery cannot be detailed at that level, a major advantage is that the revisit problem of the near-polar orbits disappears thanks to the continuous view of the same area. The two GOES cover a large section of the electromagnetic spectrum through the Advanced Baseline Sounder, which is a multi-channel radiometer focused on atmospheric cloud temperature- and moisture-profiles and ozone distribution, together with the main optical sensor ABI, the Advanced Baseline Imager.

High-resolution imagery

The first satellites produced only optical views of the land they flew over whereas the next generation also had sensors for other parts of the electromagnetic spectrum. Quite a few satellites, however, are specialized in delivering high-quality imagery (Table 1) with full global coverage due to the rotation of the Earth under their near polar orbits. When launched in 1999, IKONOS was the first publicly available, high-resolution satellite. It was designed at Lockheed Martin Corporation but developed and launched in partnership with Space Imaging, a firm that took the name GeoEye after merging with another company, Orbital Imaging. When GeoEye-1 was launched in 2008, it had the world's highest spatial resolution, while an even sharper-eyed satellite, GeoEye-2 was already on the drawing board at Lockheed Martin. However, five years later, GeoEye was acquired by DigitalGlobe, an aspiring U.S. satellite company which launched Quickbird in 2001, followed by WorldView-1 in 2007 and WorldView-2 in 2009. Following this merger, GeoEye-2 was renamed WorldView-4, while GeoEye-1 became referred to as just GeoEye. When WorldView-3 had been put into orbit in 2014 followed by WorldView-4 at the end of 2016, the scene was set for the final merger in the high-resolution satellite business. MacDonald Dettwiler and Associates (MDA), a Canadian global communications & information company, purchased DigitalGlobe forming the

combined company Maxar Technologies, which is now the dominant force in the commercial market for this kind of satellite.

The latest trend is to use smaller satellites, typically weighing less than five kg, and launching a large number of them together. Planet Lab's 150 mini-satellites provide spatial resolutions that vary between 3, 5 and 0.72 m (<https://www.planet.com/products/planet-imagery/>) and even if their best resolution cannot rival those from Geoeye and WorldView, revisits can quickly be achieved because of the plurality of satellites in space and their higher speed. Importantly, Planet Lab is pioneering machine-learning for analysis of its imagery feeds that improves the speed at which vital information can be extracted.

As seen in Table 1, there is an astonishing intermission of more than six years between the early American high-resolution satellites and the follow-up by GeoEye and the series of WorldView satellites. However, the gap is filled by the *Satellite Pour l'Observation de la Terre* (SPOT) programme initiated by the French space agency *Centre National d'Etudes Spatiales* (CNES). The system includes ground control resources for satellite control and programming, image production and distribution by SPOT Image in Toulouse, France. The SPOT programme not only filled the gap between 2001 and 2007 but also predated the first U.S. high-resolution satellites. Between 1986 and 1998, four SPOT satellites with 10-m panchromatic and 20-m multispectral spatial resolutions were launched. While the first three were identical, the Vegetation Monitoring Instrument with a 1-km resolution for agricultural monitoring was added to SPOT-4. The next in this series, SPOT 5, was equipped with two high-resolution geometrical instruments offering 2.5 to 5 m resolution in panchromatic mode and 10 m in multispectral mode plus a sensor for stereo-pair images. The latest satellites, SPOT-6 and Spot-7, are phased into the same orbit (at 694-km altitude) as Pléiades 1A and Pléiades 1B (see below) making revisits possible on a twice-daily basis. However, the continued role of SPOT-7 in this constellation is unclear since it has recently been sold to Azerbaijan to monitor crop maturation in that country.

The Pléiades satellites (<https://directory.eoportal.org/web/eoportal/satellite-missions/p/pleiades>) operate as a pair offset at 180° to offer a daily revisit capability over any point on the globe. They share the same orbital plane as the two latest SPOT satellites, forming a constellation of four, 90° apart from one another. Equipped with fibre-optic and control-moment gyros, Pléiades-HR 1A and 1B offer exceptional capabilities enabling the system to maximize the number of acquisitions above a given area. This agility, coupled with dynamic image acquisition programming, makes the Pléiades system time-wise extraordinarily responsive to user requirements.

RapidEye (<https://earth.esa.int/web/guest/missions/3rd-party-missions/current-missions/rapideye>) consists of a constellation of five identical, calibrated mini-satellites originally contracted to MDA and put into orbit in 2008. Each of the five satellites travel at the same orbital plane at 630-km altitude, collecting 5-m spatial resolution imagery. The RapidEye concept originated in Germany but eventually ended up in the hands of Planet Labs, which currently operates the largest fleet of Earth-observing satellites. RapidEye includes the first commercial sensor with the Red-Edge band, which has a particular sensitivity to changes in chlorophyll content and thus improves vegetation monitoring (Deleu *et al.*, 2015).

In the period 2014-2018, the ESA entered the fray, launching six satellites in a constellation of three pairs under the Copernicus programme (https://www.esa.int/Our_Activities/Observing_the_



Table 2. Essential sensors for environmental studies.

Sensor/ Mission	Satellite(s)	Resolution		Responsible agency/country	Comment
		Spatial	Temporal		
ABI	GOES-16 (East) GOES-17 (West)	500 m and 1-2 km depending on channel	Scene updated every 5- 5 min	NOAA	GOES are primarily weather satellites, which show the eastern hemisphere from GOES East (since Dec. 2017) and the western hemisphere (since Feb. 2019) while GOES West
ASTER	Terra	15-90 m depending on channel	16-days	Japanese sensor onboard NASA satellite	A 14-band multispectral instrument including NIR and SWIR delivering data on land cover, soils, vegetation and ecosystems, hydrology, climatology with DEM measurements creating detailed stereo imagery
AVHRR	NOAA 18-19 MetOp A-C	1.1 km	Twice daily	NOAA (in collaboration with ESA since 2006)	AVHRR/3 is a 6-channel American instrument providing day and night imaging of land, water and clouds. It also measures sea surface temperature, ice, snow and vegetation
C-SAR	Sentinel-1 A+B	5 m	12-dayst	ESA through EUMETSAT	Day and night radar imaging mission for land coastal water with respect to natural disasters, environment and climate. Sentinel-1 works in conjunction with SMAP (see below) to achieve accurate measures of soil moisture estimates. The Sentinel satellites are part of ESA's Copernicus programme
DPR	Both sensors onboard the Core GPM satellite	125-245 km	Every 2-3 hours	Initiated by JAXA and NASA operating with wide international collaboration	Consists of a core satellite and a constellation of eight other satellites with the aim of measuring rainfall, storms and water/energy cycle. GPM succeeds the TRMM
GMI		885 km			
HyspIRI- TIR	Installed externally on the ISS	38×57 m	3-4 days	NASA	Although results are publicly available, this investigation, entitled ECOSTRESS, is a specific mission on vegetation and its response to stress, such as lack of water. ECOSTRESS does not deliver data for general use
MODIS	Terra/Aqua	250-500 or 1,000 m depending on channel	1-2 days	NASA	A sensor with 36 discrete spectral bands for the study of global dynamics occurring on land, in the oceans and in the lower atmosphere. MODIS contributes to accurate environmental predictions regarding global change. Its relatively low spatial resolution is compensated by a high temporal resolution
MSI	Sentinel-2 A+B	10, 20 or 60 m depending on channel	-	ESA through EUMETSAT	Multispectral land imagery based on 13 spectral channels that include NIR and SWIR for the study of vegetation, soil and water, inland waterways and coastal areas
RMI	SMAP satellite	1-3 km	2-3 hours	NASA	Measures plant evapotranspiration and the water content in the top 5 cm of the soil producing soil moisture estimates in collaboration with Sentinel 1 A+B
OLI	Both sensors onboard Landsat-8	15-30 m	16 days	USGS/NASA	Seven of the 9-channel imager's band are the same as earlier Landsat imagers to assure compatibility with historical data, an additional deep blue and a shortwave-infrared bands allow measuring water quality and improve thin cloud detection
TIRS		100 m			Data for two long wavelength infrared bands will be collected with TIRS. This provides data continuity with Landsat 7's single thermal IR band and adds a second one

ABI, Advanced Baseline Imager; ASTER, Advanced Space-borne Thermal Emission & Reflection Radiometer; AVHRR, Advanced Very High Resolution Radiometer; C-SAR, C-band synthetic-aperture radar; DPR, Dual-Frequency Precipitation Radar; GMI, Global Precipitation Measurement Microwave Imager; HyspIRI-TIR, HyspIRI Thermal Infrared Radiometer; MODIS, Moderate Resolution Imaging Spectroradiometer; MSI, Multispectral Instrument; RMI, Radar Microwave Imager; OLI, Operational Land Imager; TIRS, Thermal Infrared Sensor; GOES, Geostationary Operational Environmental Satellites; NOAA, National Oceanic and Atmospheric Administration; NASA, National Aeronautics and Space Administration; NIR, near infrared; SWIR, shortwave infrared; DEM, digital elevation model; ESA, European Space Agency; EUMETSAT, European Organization for the Exploitation of Meteorological Satellites; SMAP, Soil Moisture Active Passive; GPM, Global Precipitation Measurement; JAXA, Japan Aerospace Exploration Agency USGS, United States Geological Survey; TRMM, Tropical Rainfall Measuring Mission; ISS, International Space Station; ECOSTRESS, Ecosystem Space-borne Thermal Radiometer Experiment on Space Station; IR, infrared.



Earth/Copernicus). These multi-sensor satellites, named Sentinel (<https://sentinel.esa.int/web/sentinel/home>), are focused on vegetation, soil and water and are thus concerned with temperature and colour of land, inland waterways, coastal water and the oceans. Only Sentinel 2 and 3 include optical scanners. The spatial resolutions of these satellites (10-60 m for Sentinel 2 and 300-1,000 m for Sentinel 3), depending on the wavelength chosen, are moderate; however, their unique combination of broad spectral coverage and wide field of view represents a major step forward in the study of environment, weather and climate, providing seasonal forecasting as well as surveillance of natural disasters.

Great expectations

The first large-scale photographs of Earth's surface were taken by cameras onboard piloted flights, with satellite-generated imagery becoming the norm after the launch of Landsat-1 in 1972 (<https://landsat.gsfc.nasa.gov/about/history/>). Although Cline (1970) realized the opportunities even earlier, papers in the field of public health appeared only sporadically and not until the last two decades of the last century, *e.g.*, Cross *et al.* (1984), Hugh-Jones (1989, 1991), Pope *et al.* (1994), Roberts and Rodriguez (1994), Washino and Wood (1994), Kitron (1998), Hay and Lennon (1999), Kristensen *et al.*, 2001; Huh and Malone (2001). Although satellite-generated imagery immediately impress both laymen and scientists, environmental datasets offer additional more useful information to the epidemiologist as pointed out in reviews by Hay (2000) and Goetz *et al.* (2000). Importantly, the openness of most governments in allowing free access to all sorts of environmental data and imagery proved crucial. Cartographers and meteorologists were the first to take advantage of the new tools, while epidemiologists were slower to react.

Following a brief mention below of the better known satellites and their sensors that were developed for measurements outside the visual electromagnetic range, Table 2 summarizes all information given and adds also some core data.

The Terra/Aqua duo and constellations

The Moderate Resolution Imaging Spectroradiometer (MODIS) is the key instrument onboard both the Terra and Aqua satellites (Tatem *et al.*, 2004), whose orbits are chosen so that Terra passes from north to south across the equator in the morning with Aqua moving in the opposite direction in the afternoon. In this way, a higher temporal resolution is obtained without loss of image quality. Terra was the first in NASA's EOS programme which represents a coordinated series of polar-orbiting satellites observing oceans, land cover, biosphere and atmosphere. It was launched in 2000 and began collecting what has become an 18-year, global, wide-range set of environmental data. In order to produce better information than would be possible from the sum of independent observations, some satellites are placed in close proximity to each other in carefully planned constellations with only seconds or minutes between them. One such group is the A-train, also called the Afternoon Constellation (of satellites), that all cross the equator in the early afternoon travelling north. Aqua, launched in 2002, was the first satellite in this constellation while a smaller group containing Landsat-7, Terra and a few other satellites moving south constitute the Morning train as they cross the equator in the morning. A full account of the sensors onboard Terra and Aqua is given

by Bergquist (2011).

The optical information from Sentinel 2 and 3 mentioned above is joined by other data, particularly from Sentinel 1 which lacks an optical sensor. The Sentinel satellites' main contribution consists of multi-instrument measurements of sea-surface topography as well as sea- and land-surface temperatures.

The Global Precipitation Measurement mission

Measuring Earth's water/energy cycle, the Global Precipitation Measurement (GPM) project continues the Tropical Rainfall Measuring Mission (TRMM) collaboration between NASA and the Japan Aerospace Exploration Agency (JAXA). GPM is based on an international group consisting of a core GPM satellite equipped with the Dual-Frequency Precipitation Radar (DPR) and the GPM Microwave Imager (GMI), plus eight other satellites to provide global coverage.

The DPR provides three-dimensional maps of storms and surface precipitation and the GMI observes the microwave energy emitted by Earth and its atmosphere, while the Integrated Multi-satellite Retrievals for GPM merges and calibrates all satellite microwave precipitation estimates (https://www.nasa.gov/mission_pages/GPM/main/index.html). The DPR has two radar frequencies: the Ku-band radar, which is similar to the PR band of the TRMM, and the Ka-band, the differential attenuation of which enables distinction between rain and snow.

GPM is part of NASA's Earth Systematic Missions program, managed by the Goddard Space Flight Center in Maryland, USA and represents an unprecedented international collaboration to improve precipitation estimates from space for research and application. The original JAXA and NASA collaboration has been joined by CNES, ESA and other national space agencies unifying global precipitation observations.

The Soil Moisture Active Passive satellite

The Soil Moisture Active Passive (SMAP) mission (<https://smap.jpl.nasa.gov/>) is an international network of satellites that provides the next-generation global observations of rain and snow, aiming to advance understanding of the processes that link the terrestrial water, energy and carbon cycles. It features a near-global revisit coverage in 2-3 days improving forecasting of extreme climate events and providing accurate and timely information on land surface soil moisture and freeze-thaw status. The SMAP surface measurements are coupled with hydrologic models to infer soil moisture conditions in the root zone rather than on the surface. These measurements also provide estimates of global water and energy fluxes over land and quantify the net carbon flux in boreal landscapes. It further produces improved flood predictions and drought monitoring capability. All data are publicly available through the NASA archive centres.

The Ecosystem Space-borne Thermal Radiometer Experiment on Space Station mission

The Ecosystem Space-borne Thermal Radiometer Experiment on Space Station (ECOSTRESS) is a radiometer mounted on the International Space Station measuring the temperature of plants growing in specific locations on Earth over the course of a year (Hook and the HypIRI and ECOSTRESS Teams, 2015). Rather than being a general-purpose sensor, ECOSTRESS represents a specific investigation based on a 5-channel prototype, the HypIRI Thermal Infrared Radiometer. ECOSTRESS collects data on the temperature of the surface itself rather than the surface air temper-



ature testing how the terrestrial biosphere responds to changes in water availability; how changes in daily water access impact vegetation and the global carbon cycle; and whether agricultural vulnerability can be reduced through improved drought estimation. These measurements provide an insight into the effects on crops of events, such as heat-waves and droughts on crops. The data collected are publicly available (<https://ecostress.jpl.nasa.gov/early-adopters>).

From BioClim to PaleoClim

Efforts to predict climate suitability for agriculture began in the 18th century. The relationship between plant development and the ambient temperature was first mentioned by Réaumur in France, who discovered what he called the *degree-day unit* (described by Bonhomme, 2000). This unit, a measure of crop maturation defined as the amount of heat an organism needs to accumulate to reach full development, is now referred to as the growing degree-day (GDD) unit and is no longer used only in agriculture but also for predicting the development of parasites (Bernal, 1993; Malone, 2005; Yang *et al.*, 2007). The GDD approach constitutes a major step forward with respect to the parasite/vector relationship and is still applicable (Genchi *et al.*, 2011; Valencia-López *et al.*, 2012).

Although the GDD concept has no direct link to the Bioclim datasets, it has encouraged thinking on the environment as a set of contiguous niches with diverse ecology as first outlined by Pavlovskii (1945) and later developed further by Hutchinson (1957). Bioclim grew out of the type of niche modelling developed in Australia by Nix (1986) and comprises a package of environmental records for specific sites where a species had been seen that can be used to map out other suitable locations where this species should be able to exist. In this way, habitat suitability modelling would indicate niches suitable for species under past, present or future conditions. Bioclim is used extensively for species surveys, assessments of the risk for invasive plants and animals, studies of historical distributions and climate change evaluations, but the lack of reliable estimates for climatic conditions at sites remote from meteorological stations remains a major limitation (Booth *et al.*, 2014). Bioclim can also be used to predict how existing, established species can cope with climate change (Booth and Williams, 2012). The Anusplin package an interpolation programme, made it possible to develop applications in widely different biotopes (Zuo *et al.*, 1996; Hong *et al.*, 2005). Significantly, it led to the WorldClim database, a set of 1 km² monthly gridded global climate layers aggregated across a temporal range from 1960 to 1990 developed by Hijmans *et al.* (2005). These data can be used for mapping and spatial modelling for all land areas of the globe except Antarctica and have recently been followed up by the publication of WorldClim 2.0 covering 1970–2000 (Fick and Hijmans, 2017). The 19 bioclimatic variables presented in WorldClim are derived from temperatures and rainfall values to generate biologically meaningful variables to be used for species distribution and ecological modelling. Data from weather stations were interpolated using thin-plate splines with elevations and satellite-derived covariates, such as maximum and minimum land surface temperature as well as cloud cover and other parameters.

The biotope concept has also received archaeological interest. Paleoparasitology, defined as the study of parasites in ancient material and their interactions with hosts and vectors, has led to the

creation of PaleoClim, a paleoclimate database (Brown *et al.*, 2018), that should be of interest for evolutionary and ecological studies. PaleoClim is a free database of down-scaled paleoclimate outputs at a resolution of about 5 km² (at the equator) that include surface temperature and precipitation estimates from snapshot-style climate model simulations based on HadCM3, a version of the UK Met Office Hadley Centre General Circulation Model that can be utilized for modelling past and future climates. This database contains climatic data for three key time periods spanning from 3.300 to 0.787 million years ago (Ma): the Marine Isotope Stage 19 (MIS19) in the Pleistocene (~0.787 Ma), the mid-Pliocene Warm Period (~3.264–3.025 Ma), and MIS M2 in the late Pliocene (~3.3 Ma). The availability of climatic simulations representing conditions of the late Pleistocene and Holocene epochs has revolutionized the study of species responses to the late Quaternary climate change. A recent excavation of ancient areas in the Middle East has revealed the incidence of human and animal parasitic infections dating back to 8100 BC (Khodkar *et al.*, 2018), while Banks (2017) used eco-cultural niche modelling as a way to better understand the dynamics of European culture changes as far back as 40,000 years ago.

Points to ponder

In response to calls for action by the 2002 World Summit on Sustainable Development and the G8 leading industrialized countries, the Third Earth Observation Summit in Brussels, Belgium established the Group on Earth Observations (GEO) in 2005. As of 2018, GEO's membership includes 105 governments and 127 organizations with a mandate in Earth observation. GEO's main activity is to set up a Global Earth Observation System of Systems (GEOSS) to connect producers of environmental data with the end users of these data (<https://www.earthobservations.org/geoss.php>) on the basis of a 10-Year Strategic Plan (2016–2025). The plan defines eight areas, which includes public health, biodiversity and ecosystem sustainability. NASA's Earth Science area is moving towards the study of climate and environmental change and the potential impact on public health issues (Luvall, 2013). NASA also promotes linkages with GEO to support collaboration and the use of a standardized, open-source global resource data portal where the Group on Earth observations health network (GEOHealth) within the GEOSS framework would play an important role (Malone *et al.*, 2014). The GEOHealth mandate is outlined in a recent paper by Malone *et al.* (2019).

Earth-observing satellites play a vital role in data collection as they monitor the entire globe within rapidly repeated time periods. More than 30 years of archived data used in epidemiological studies of infectious diseases are currently available. GEO uses satellite-generated observations, including public-health surveillance, to provide alerts regarding air and weather extremes, vector-borne diseases, *etc.* (Igarashi *et al.*, 2014; Gebreslasie *et al.*, 2015; Takane *et al.*, 2016). In order to offer reliable and user-friendly datasets, agencies have validated them by comparison with corresponding ones from other sensors (Martin *et al.*, 2019) as well as with ground-based measurements (Colston *et al.*, 2018). Users should select their products carefully as time series of observations usually are affected by the density of measurement and non-homogeneity caused by non-climatic factors such as relocation of weather stations, changes of instruments, *etc.* One way to avoid this is by reanalysis based on assimilation of combined observations from

many different sources (terrestrial measurements and various satellite-generated datasets) in order to provide a better representation of past and current weather conditions (Toreti *et al.*, 2019).

Environmental data provide useful insights into the different factors related to transmission levels and disease distributions, while mapping and modelling facilitates interpretation, synthesis, and recognition of outbreak frequencies (Rinaldi *et al.*, 2015). Although small-scale maps based on raw data can be difficult to interpret because of differences in disease occurrence between areas due to summation of true, structural differences, remotely sensed information has now reached a level of sophistication allowing almost complete niche specification, and optical scenes that can be zoomed to give information at the household level (Malone *et al.*, 2019).

AVHRR, carried by NOAA's polar-orbiting satellites, and the optical instruments onboard the various Landsat satellites have been in use for close to 40 and 50 years, respectively, while MODIS has been delivering environmental data since the start of the new millennium. The continuing instrumentation upgrade has been highly beneficial to a wide range of scientific inquiry and during the last few years a plethora of exceptional, new sensors have emerged. They include the OLI optical sensor onboard Landsat-8, the GPM for rainfall and the SMAP for soil moisture, while scientists interested in agriculture and vegetation can enjoy access to a daily stream of new data from ECOSTRESS. However, when choosing satellites and sensors for research, it is necessary to consider the four different modes of resolution available (spatial, temporal, spectral and radiometric). As an example, Table 3 shows this choice with respect to the four most common satellites and the six most useful instruments. Acquiring data in 36 spectral groups of wavelengths (bands) from the entire surface of the Earth, MODIS has probably been, and will continue to be, the most useful sensor for biological research.

Spatial statistics

There is growing need to leverage existing and emerging

sources of remotely sensed data and to investigate their viability for improving human health. Satellite imagery and environmental data have significant potential to provide opportune statistical outputs at a more disaggregated level for informed human health decision-making. It is becoming evident that the human health sector is directly affected by climate change (*e.g.* seasonal droughts, intense rainfall, riverine floods and flash floods). Climatic factors are linked to infant malnutrition and chronic ailments associated with malaria, cholera and diarrhoea as a result of droughts and floods. In a report by the World Health Organization (WHO, 2006), it was estimated that 24% of the global burden of disease and 23% of all deaths can be attributed to environmental factors and the United Nations (2011) surmised that *health is becoming a key goal of climate policies and a priority in climate mitigation and adaptation actions*. At the core, there is the aim to not only identify and describe observed associations but also to appreciate the processes resulting in the pattern(s) observed. Thus, statistical methodologies are essential to describe the geographic distribution of disease, to understand the intensity of transmission and risk, to characterize populations at risk and to inform about potential interventions and responses.

Spatial statistics are specific methods that use distance, space, and spatial relationships to describe and analyse patterns in the geographical distribution of human health. This analysis often involves a description of spatial occurrences (usually represented by data on a map); exploratory analysis (to search for patterns and possible explanations); statistical analysis and hypothesis testing of expected phenomena under an assumed statistical model; and modelling and prediction that involve construction of regression models to predict health outcomes at certain locations. Spatial statistics consequently constitute an area much too large to be discussed here in any detail, but some commonly used approaches are expanded on below to emphasize that statistics cannot simply be added onto a dataset to make sense of the information collected. On the contrary, statistical thinking must be involved already from the moment a project is being drafted. Indeed, it will help in advising just which data and how much information on each aspect would be needed.

Table 3. Channel characteristics of some commonly used sensors.

Resolution	NOAA/MetOP AVHRR/3	Terra/Aqua MODIS	Landsat OLI and TIRS	Global Precipitation Measurement DPR	Global Precipitation Measurement GMI
Spatial	All six bands: 1.1 km	Bands 1-2: 250 m Bands 3-7: 500 m Bands 8-36: 1 km	Bands 1-7: 30 m Band 8: 15 m Band 9: 30 m Band 10: 100 m Band 11: 100 m	Both bands: 5 km	Bands 1-2: 19.4×32.2 km Bands 3-4: 11.2×18.3 km Band 5: 9.2×15.0 km Bands 6-7: 8.6×14.4 km Bands 8-13: 4.4×7.3 km
Temporal	Twice daily	2 days	16 days	2 days	2 days
Spectral	Band 1: visible Band 2: NIR Band 3A: SWIR Band 3B: SWIR Band 4: TIR Band 5: TIR	10 bands: visible 9 bands: IR 17 bands: TIR	Bands 1-4: multispectral Band 5: NIR Bands 6-7: SWIR Band 8: pancromatic Band 9: 1370 ± 10 nm Bands 10-11: TIR	The KuPR band: 12-18 GHz The KaPR band: 26-40 GHz	Bands 1-2: 10.65 GHz Bands 3-4: 18.70 GHz Band 5: 23.80 GHz Bands 6-7: 36.64 GHz Bands 8-9: 89.00 GHz Bands 10-11: 165.50 GHz Bands 12-13: 183.31 GHz
Radiometric	10 bits	12 bits	12 bits	Variable	Variable

NOAA, National Oceanic and Atmospheric Administration; AVHRR, Advanced very high resolution radiometer; MODIS, Moderate resolution imaging spectroradiometer; OLI, Operational land imager; TIRS, Thermal infrared sensor; DPR, Dual-frequency precipitation radar; GMI, Global Precipitation Measurement Microwave Imager; NIR, Near infrared; IR, Infrared; SWIR, Shortwave infrared; TIR, Thermal infrared.



Choropleth maps *vis-à-vis* heat maps

One of the most common methods for describing spatial data is the use of choropleth maps, which are thematic maps where different areas are shaded in proportion to exposure and outcomes (e.g., with respect to health) dealing with parameters such as population density, amount of humidity and proportion of cases, per-capita income, *etc.* They are based on data aggregated over previously defined regions and provide an easy way to measure and visualize how occurrences vary across a geographic area. However, when using choropleth maps, a number of issues need to be considered, e.g., the modifiable areal unit problem (MAUP) where results may depend on the specific geographic unit used (Fotheringham and Wong, 1991), or the choice of a province, county or city which can lead to different results and thus contribute to major misinterpretations. Misuse of these maps can also lead to the so-called ecological fallacy where the results obtained from aggregated data are assumed to apply to individual people. While the use of smaller and more specific regions can decrease the risk of the ecological fallacy and MAUP, it can also make more complicated maps. Raw-count data are often used when constructing choropleth maps as if the population or geographic area are the same size for all geographic units and this, too, can be problematic. When using choropleth maps, the plotted data must always be *normalized* by total population or geographic area.

There is a slight, but important difference between choropleth maps and heat maps: although similar, the former show changes across a landscape with geographic boundaries, e.g., units such as countries, states, census blocks, *etc.*, while the latter, normally generated from point data representing a type of density, show changes across a rasterized dataset, usually conforming to a small grid size. *i.e.* a landscape where boundaries play no role. In this way, these diverse presentations reflect the difference between vectors and rasters as discussed above in relation to GIS presentations. Heat maps are usually used when tracking natural phenomena, such as environmental niches or weather, where established borders and boundaries do not play a role.

Cluster analysis

Perhaps one of the most important concepts in geography and GIS is Tobler's (1970) first law of geography, namely *All places are related but nearby places are more related than distant places*. In spatial statistics this law is confirmed by spatial autocorrelation measures accounting for distance between spatial units. There are several measures of spatial correlation: global measures use a single value which applies to the entire data set, which is then seen as an average for the entire area, while local measures produce a value for each subunit of area, which may imply that different patterns or processes occur in different parts of the region or country under study. Well-known works on spatial autocorrelation are Moran (1950), Geary's C (Geary, 1954) and Getis and Ord (1992).

Correlation and regression models

Remotely-sensed environmental data are often useful when considering the interplay between parameters such as poverty levels, socio-economic variables and health. Standard correlation and ordinary (linear) regression can be utilized to reveal relationships between such datasets using a statistical analytical procedure within a GIS that links health outcomes and spatial relationships,

including geographic features. However, in the presence of (expected) spatial correlations, estimated coefficients and coefficients of determination would appear larger, and standard errors smaller than they really are; this could lead to potentially incorrect conclusions regarding the statistical significance of a relationship. Spatial regression models, including spatial autoregressive lag model error models and geographically weighted regressions, can be used to estimate separate regression values for each area and its neighbours.

Bayesian spatial statistics

In Bayesian statistics probability is expressed as a degree of belief in the occurrence of an event, which can change as new information is gathered, rather than a fixed value based on frequency (Bayes, 1763). Bayesian methods generally require an unrealistic amount of computation, so it was not much used before the advent of powerful computers. Even though the Bayesian methodology for statistical inferences has long been used in the various branches of statistics, its application to spatial statistics was initially hindered by analytical intractability in addition to the computational burden. Consequently, spatial modelling, including spatial autoregressive models and conditional autoregressive models, were rarely applied before advancements in computation and simulation methods appeared (Congdon, 2006). In particular, Markov chain Monte Carlo methods, including the Metropolis-Hastings algorithm (Hastings, 1970) and the Gibbs sampler (Gelfand and Smith, 1990), eased the computational burden of Bayesian spatial inference and are now incorporated into several packages, such as WinBUGS (Lunn *et al.*, 2000) and BayeX (Belitz *et al.*, 2003) contributing to the current exponential growth in epidemiological applications of Bayesian spatial modelling (Congdon, 2006). Bayesian spatial modelling considers both the distributions of the data and the unknown coefficient estimates which are assigned priors based on subjective beliefs about their distributions. In this context, Bayesian spatial modelling, Bayesian estimation and inference of parameter estimates, model specifications and comparisons as well as predictions are the main issues. In most applications in health and epidemiology, Bayesian spatial modelling has become more popular than conventional classical spatial modelling (Banerjee *et al.*, 2004). In recent decades, most applications of Bayesian spatial methods have been based either on the *lattice geographical units* model of Besag, York and Mollie (BYM) (Besag *et al.*, 1991) or the so-called *convolution model*, alternatively the *distance-based geostatistics model* advanced by Diggle and Ribeiro (2007). Measuring the statistical associations between remotely-sensed data and human health depends on the data distributions used to model human health outcomes. In most cases, binomial and Poisson Bayesian spatial models are applied, in which the effect sizes the Odds Ratios and Risk Ratio models, respectively. Such models have been widely used to link environmental data and outcome, e.g., to assess the vegetation's suitability for the tsetse fly (Albert *et al.*, 2015; Lord *et al.*, 2018); the climatic changes and induced vulnerability to malaria and cholera (Reyburn *et al.*, 2011; Zhao *et al.*, 2014); the spatial drivers of land use and land cover change (Kamwi *et al.*, 2018); and the environmental effects on allergies and cardiovascular disease (Brook *et al.*, 2004; Takano and Inoue, 2017). Importantly, today's common generation of large amounts of data leads to problems such as uncertainty and inaccuracies in the generated datasets. This may require complex geospatial modelling, to avoid measurement



errors and uncertainties. The rapid generation of these types of datasets and the need for geospatial modelling are now crucial in facilitating the addition of a geographical component to epidemiological research. We therefore need a consensus on the complexity of spatial models used for estimating and predicting spatial surfaces in order to provide robust and reliable spatial mapping at relevant levels for local planning.

Discussion and Conclusions

With regard to cholera and smallpox, neither Shapter and Snow nor Jenner understood how the places they realized were risky, actually produced the disease and could thus only know the *where* not the *why*. However, that was sufficient in practice. Although the hypothesis, based on Hippocrates' treatise *Airs, waters and places* (Hippocrates, around 400 BC), that places could exert a kind of influence through air or water first led Snow to think that the water must be the culprit, he later realized that this could not be the case since other pumps in London were safe (at least in 1854). Instead it seemed to become bad when fetched – in this he was tantalizing near the hidden truth that something was transmitted from the water at the pump to the person who later fell sick (in fact it also had to be transmitted to the water first at the site). The way we understand the situation today is not only that bacterial transfer takes place but also that certain areas can become conducive for bacterial survival and exchange and thus create an ecological niche. The same argument can be made for malaria, which was already a problem in Roman times and not eliminated from central Italy until the 1930s when efforts to drain the Pontine marches south of Rome finally succeeded, robbing the *Anopheles* mosquito of its ecologic niche.

Although ecology is generally seen as a new science, it has in one way or another existed since time immemorial (McIntosh, 1985). It was already an accepted concept in ancient Greece and during the Roman era, with water sources and special locations believed to play a role for health and disease. Hippocrates' contemporary Theophrastus, although not primarily interested in health, philosophizes about the association between animals and their environment (Ramalay, 1940). However, ecology did not emerge as a distinct discipline until the mid-19th century when defined by Haeckel (1866), a German naturalist active in a range of academic disciplines, who not only discovered, described and named thousands of new species and mapped a genealogical tree relating all life forms, but also coined many biological terms. Haeckel admired his compatriot Alexander von Humboldt who, already in the 18th century, had established an ecological perspective not far from Darwin's ideas, though Haeckel noted that Humboldt's ecology lacked the competition concept that was at the heart of *On the origin of species* (Darwin, 1859).

The niche-related thinking based on the study of the very specific requirements different species that resulted in Bioclim (Nix, 1986) is a useful approach that works well even though it does not lend itself to extrapolation taking into account the reasons why certain other environments lack particular species. Here, methods requiring data on absence as well as presence are needed, e.g., Genetic Algorithm for Rule Set Production, a computer program describing environmental conditions (possible to store in a GIS) under which a species should be able to maintain its population (Stockwell and Peters, 1999). Other modelling methods, such as Maxent (Phillips *et al.*, 2006). Use presence data in conjunction

with background data based on the maximum entropy principle that generates the probability distribution best supported by available data. Extrapolation can be done allowing application of Maxent also for combinations of environments unavailable during model training which is an advantage, though overfitting is a risk that should be taken into consideration.

BioClim is useful in its own right and the interpolations leading to the WorldClim database (Hijmans *et al.*, 2005), which is based on 19 global bioclimatic variables aggregated across the 1960-1990 period, can be used for mapping and spatial modelling for all land areas. The follow-up WorldClim 2,0 (Fick and Hijmans, 2017) covering the years 1970-2000 increases accuracy for current data. However, if the Earth's average temperature continues to rise as predicted by the Intergovernmental Panel on Climate Change (IPCC), we might soon need a WorldClim 3.0. Paleoparasitology, created along the same lines to trace and recover the natural development of parasites as well as the origin of infectious diseases, is an interesting continuation along the BioClim trail. While today's access to broad-range remotely-sensed data has made epidemiology more dependent on such data, this discipline has undergone a paradigmatic change thanks to GIS that facilitates data management and processing. The visualization of epidemiological datasets in a geographical context supports risk profiling and the development of EWS, while the growth of the Internet has made it possible to connect with other platforms, such as web map servers, libraries, spatial database management systems, and software development frameworks. The field has thus become multi-participatory, allowing the advantage of cloud-computing opportunities that makes GIS accessible for anyone, anywhere. However, while the development of near real-time surveillance systems, based on GIS, GPS and remote-sensing, facilitates the establishment of accurate, up-to-date EWS, it is important to understand that GIS neither makes the actual field collection of parasites and their vectors easier (an essential task), nor does it by itself assure the quality of the information gained.

Niche characterization can now be achieved by applying the full force of remote-sensing to a small area. Facilitated by new satellite systems, big data, climatology advances, and the novel sensors, we are close to succeeding in the development of dynamical multi-dimensional models of the major endemic diseases. Importantly, the highly specific measurements of environmental variables together with the sub-meter resolutions achieved by the latest satellite imagery can provide risk assessments for as limited an area as a community or even a household. Although falling outside the scope of this article, the potential of high-altitude aircraft equipped with particular sensors, now promoted by NASA (<https://aviris.jpl.nasa.gov/>), as well as the use of low-flying drones should be seriously contemplated with respect to high-resolution health-related data collection within a user-set agenda.

References

- Albert M, Wardrop NA, Atkinson PM, Torr SJ, Welburn SC, 2015. Tsetse fly (*G. f. fuscipes*) distribution in the Lake Victoria basin of Uganda. PLoS Negl Trop Dis 9:e0003705.
- Banerjee S, Carlin BP, Gelfand AE, 2004. Hierarchical modeling and analysis for spatial data. In: Isham V, Keiding N, Louis T, Reid N, Tibshirani R, Tong H. Monographs on statistics and applied probability. Chapman & Hall/CRC Press, Boca Raton, USA.



- Banks WE, 2017. The application of ecological niche modeling methods to archaeological data in order to examine culture-environment relationships and cultural trajectories. *Quaternaire* 28:271-6.
- Bayes T, 1763. An essay towards solving a problem in the doctrine of chances (communicated by Richard Price in a letter to John Canton). *Phil Trans Roy Soc London* 53:370-418.
- Belitz C, Brezger A, Kneib T, Lang S, 2003. BayesX methodology manual. Available from: <https://pdfs.semanticscholar.org/346d/56bb6aa017a043a8a9e19e619430e68838a6.pdf> Accessed: 6 April 2019.
- Berghaus H, 1837. [Berghaus' Physikalischer Atlas] Available from: <http://www.atlassen.info/atlassen/perthes/berpa.html>
- Bergquist R, 2011. New tools for epidemiology: a space odyssey. *Mem Inst Oswaldo Cruz* 106:892-900.
- Bernal J, Gonzalez D, 1993. Experimental assessment of a degree-day model for predicting the development of parasites in the field. *J Appl Ent* 116:459-66.
- Besag J, York J, Mollié A, 1991. Bayesian image restoration, with two applications in spatial statistics. *Ann Inst Stat Math* 43:1-20.
- Bonhomme R, 2000. Bases and limits to using 'degree.day' units. *Europ J Agron* 13:1-10.
- Booth TH, Nix HA, Busby JR, Hutchinson, MF, 2014. BIOCLIM: the first species distribution modelling package, its early applications and relevance to most current MAXENT studies. *Diversity Distrib* 20:1-9.
- Booth TH, Williams KJ, 2012. Developing biodiverse plantings suitable for changing climatic conditions 1: underpinning scientific methods. *Ecol Manag Restor* 13:267-73.
- Brook RD, Franklin B, Cascio W, Hong Y, Howard G, Lipsett M, Luepker R, Mittleman M, Samet J, Smith SC Jr, Tager I, 2004. Expert Panel on Population and Prevention Science of the American Heart Association. Air pollution and cardiovascular Disease. A statement for healthcare professionals from the Expert Panel on Population and Prevention Science of the American Heart Association. *Circulation* 109:2655-71.
- Brown JL, Hill DJ, Dolan AM, Carnaval AC, Haywood AM, 2018. PaleoClim, high spatial resolution paleoclimate surfaces for global land areas. *Sci Data* 5:180254.
- Cline BL, 1970. New eyes for epidemiologists: aerial photography and other remote sensing techniques. *Am J Epidemiol* 92:85-9.
- Colston JM, Ahmed T, Mahopo C, Kang G, Kosek M, Sousa Junior F, Shrestha PS, Svensen E, Turab A, Zaitchik B, 2018. Evaluating meteorological data from weather stations, and from satellites and global models for a multi-site epidemiological study. *Environ Res* 165:91-109.
- Congdon P, 2006. Bayesian statistical modelling. 2nd ed. Wiley Series in probability and statistics. John Wiley & Sons, Ltd. Chichester, England.
- Cross ER, Perrine R, Sheffield C, 1984. Predicting areas endemic for schistosomiasis using weather variables and a Landsat database. *Mil Med* 149:542-4.
- Darwin CA, 1859. On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life. John Murray, London, UK.
- Deleu J, Franke J, Gebreslasie M, Linard C, 2015. Improving AfriPop dataset with settlement extents extracted from RapidEye for the border region comprising South-Africa, Swaziland and Mozambique. *Geospat Health* 10:48-54.
- Diggle PJ, Ribeiro PJ, 2007. Model-based geostatistics. Springer series in statistics. Springer, New York, USA.
- Ellegård K, Svedin U, 2012. Torsten Hagerstrands time-geography as the cradle of the activity approach in transport geography. *J Transp Geogr* 23:17-25.
- Fick SE, Hijmans RJ, 2017. Worldclim 2: New 1-km spatial resolution climate surfaces for global land areas. *Int J Climatol* 37:4302-15.
- Fotheringham AS, Wong DWS, 1991. The modifiable areal unit problem in multivariate statistical analysis. *Environ Plan* 23:1025-44.
- Gatalsky P, Andrienko N, Andrienko K, 2004. Interactive analysis of event data using space-time cube. In: Banissi E, ed. Proceedings of the 8th International Conference on Information Visualization, London, July 2004. IEEE Computer Society, Los Alamitos, pp 145-152.
- Geary RC, 1954. The contiguity ratio and statistical mapping. *Incorporated Statist* 5:115-45.
- Gebreslasie MT, Bauwens I, 2015. MALAREO: a user-driven project. *Geospat Health* 10:80-7.
- Gelfand AE, Smith AFM, 1990. Sampling-based approaches to calculating marginal densities. *J Am Stat Assoc* 85:398-409.
- Genchi C, Mortarino M, Rinaldi L, Cringoli G, Traldi G, Genchi M, 2011. Changing climate and changing vector-borne disease distribution: the example of *Dirofilaria* in Europe. *Vet Parasitol* 176:295-9.
- Getis A, Ord JK, 1992. The analysis of spatial association by use of distance statistics. *Geogr Anal* 24:189-206.
- Goetz SJ, Prince SD, Small J, 2000. Advances in satellite remote sensing of environmental variables for epidemiological applications. *Adv Parasitol* 47:289-307.
- Haeckel EHPA, 1866. [Generelle Morphologie der Organismen. Volume II: Allgemeine Entwicklungsgeschichte der Organismen]. Georg Reimer, Berlin, Germany. p 286.
- Hägerstrand T, 1970. What about people in regional science? *Reg Sci Assoc Pap* 24:7-21.
- Hastings WK, 1970. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* 57:97-109.
- Hay SI, 2000. An overview of remote sensing and geodesy for epidemiology and public health application. *Adv Parasitol* 47:1-35.
- Hay SI, Lennon JJ, 1999. Deriving meteorological variables across Africa for the study and control of vector-borne disease: a comparison of remote sensing and spatial interpolation of climate. *Trop Med Int Health* 4:58-71.
- Hedley NR, Drew CH, Arfin EA, Lee A, 1999. Hägerstrand revisited: interactive space-time visualizations of complex spatial data. *Informatica Int J Comput Inf* 23:155-68.
- Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A, 2005. Very high resolution interpolated climate surfaces for global land areas. Available from: <http://www.worldclim.org/version1>.
- Hippocrates, ca 400 BC. On airs, waters, and places. Translated by Francis Adams. Available from: <http://classics.mit.edu/Hippocrates/airwatpl.1.1.html>.
- Hong Y, Nix HA, Hutchinson MF, Booth TH, 2005. Spatial interpolation of monthly mean climate data for China. *Int J Climatol* 25:1369-79.
- Hook SJ, HyspIRI and ECOSTRESS Teams, 2015. HyspIRI Thermal & ECOSTRESS Presented at HyspIRI Symposium. NASA/Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA. Available from: https://hyspiri.jpl.nasa.gov/downloads/2015_



- Symposium/day1/4_HyspIRIThermalECOSTRESS_Hook.pdf.
- Hugh-Jones M, 1989. Applications of remote sensing to the identification of the habitats of parasites and disease vectors. *Parasitol Today* 5:244-51.
- Hugh-Jones M, 1991. Satellite imaging as a technique for obtaining disease-related data. *Rev Sci Tech* 10:197-204.
- Huh OK, Malone JB, 2001. New tools: potential medical applications of data from new and old environmental satellites. *Acta Trop* 79:35-47.
- Hutchinson GE, 1957. Concluding remarks. *Cold Spring Harb Symp Quant Biol* 22:415-27.
- Igarashi T, Kuze A, Sobue S, Yamamoto A, Yamamoto K, Oyoshi K, Imaoka K, Fukuda T 2014. Japan's efforts to promote global health using satellite remote sensing data from the Japan Aerospace Exploration Agency for prediction of infectious diseases and air quality. *Geospat Health* 8:S603-10.
- Jenner E, 1798. An inquiry into the causes and effects of the *variola vaccinae*, a disease discovered in some of the western counties of England, particularly Gloucestershire, and known by the name of the cow pox. Sampson Low, London, England.
- Jenson AB, Ghim SJ, Sundberg JP, 2016. An inquiry into the causes and effects of the *variola* (or Cow-pox. 1798). *Exp Dermatol* 25:178-80.
- Kamwi JM, Cho MA, Kaetsch C, Manda SOM, Graz FP, Chirwa PW, 2018. Assessing the spatial drivers of land use and land cover change in the protected and communal areas of the Zambezi Region, Namibia. *Land* 7:131.
- Khodkar I, Feizhadad MH, Tavalla M, 2018. Paleoparasitology in Iran: a review. *Infez Med* 26:396-402.
- Kitron U, 1989. Landscape ecology and epidemiology of vector-borne diseases: tools for spatial analysis. *J Med Entomol* 35:435-45.
- Kristensen TK, Malone JB, McCarroll JC, 2001. Use of satellite remote sensing and geographic information systems to model the distribution and abundance of snail intermediate hosts in Africa: a preliminary model for *Biomphalaria pfeifferi* in Ethiopia. *Acta Trop* 79:73-8.
- Lord JS, Torr SJ, Auty HK, Brock PM, Byamungu M, Hargrove JW, Morrison LJ, Mramba F, Vale GA, Stanton MC, 2018. Geostatistical models using remotely-sensed data predict savanna tsetse decline across the interface between protected and unprotected areas in Serengeti, Tanzania. *J Appl Ecol* 55:1997-2007.
- Lunn DJ, Thomas A, Best N, Spiegelhalter D, 2000. WinBUGS – A Bayesian modelling framework: Concepts, structure, and extensibility. *Stat Comput* 10:325-37.
- Luvall JC, 2013. NASA's Public Health Program: how we can use NASA satellite data to study global public health issues. NASA Technical Reports Server. Available from: https://pmm.nasa.gov/sites/default/files/document_files/NASA%20and%20Global%20Health.pdf
- Malone JB, 2005. Biology-based mapping of vector-borne parasites by Geographic Information Systems and Remote Sensing. *Parassitologia* 47:27-50.
- Malone JB, Bergquist R, Martins M, Luvall JC, 2019. Use of geospatial surveillance and response systems for vector-borne diseases in the elimination phase. *Trop Med Infect Dis* 4:pii:E15.
- Malone JB, Bergquist R, Rinaldi L, 2016. Geospatial surveillance and response systems for schistosomiasis. In: James B, ed. *Schistosoma: biology, pathology and control*. CRC Press, Boca Raton, FL, USA, pp 479-497.
- Malone JB, Tourre YM, Faruque F, Luvall JC, Bergquist R, 2014. Towards establishment of GeoHealth, an open-data portal for health mapping and modelling based on Earth observations by remote sensing. *Geospat Health* 8:S599-602.
- Martin MA, Ghent D, Cordeiro Pires A, Göttsche F-M, Cermak J, Remedios JJ, 2019. Comprehensive in situ validation of five satellite land surface temperature data sets over multiple stations and years. *Remote Sens* 11:479.
- McIntosh R, 1985. The background of ecology. Concept and theory. Cambridge University Press, New York.
- Moran PAP, 1950. Notes on continuous stochastic phenomena. *Biometrika* 37:17-23.
- Nix HA, 1986. A biogeographic analysis of Australian elapid snakes. In: Longmore R, ed. *Atlas of elapid snakes of Australia: Australian Flora and Fauna series no. 7*. Bureau of Flora and Fauna, Canberra, Australia, pp 4-15.
- Pavlovskii EN, 1945. The ecological parasitology. *J Gen Biol* 6:65-92.
- Phillips SJ, Anderson RP, Schapire RE, 2006. Maximum entropy modeling of species geographic distributions. *Ecol Modell* 190:231-59.
- Pope KO, Rejmankova E, Savage HM, Arredondo-Jimenez JJ, Rodriguez MH, Roberts DR, 1994. Remote sensing of tropical wetlands for malaria control in Chiapas, Mexico. *Ecol Appl* 4:81-90.
- Ramalay F, 1940. The growth of a science. University of Colorado Studies, USA.
- Reyburn R, Kim DR, Emch M, Khatib A, von Seidlein L, Ali M, 2011. Climate variability and the outbreaks of cholera in Zanzibar East Africa: a time series analysis. *Am J Trop Med Hyg* 84:862-9.
- Rinaldi L, Genchi C, Musella V, Genchi M, Cringoli G, 2011. Geographical information systems as a tool in the control of heartworm infections in dogs and cats. *Vet Parasitol* 176:286-90.
- Rinaldi L, Hendrickx G, Cringoli G, Biggeri A, Ducheyne E, Catelan D, Morgan E, Williams D, Charlier J, von Samson-Himmelstjerna G, Vercruysse J, 2015. Mapping and modelling helminth infections in ruminants in Europe: Experience from GLOWORM. *Geospat Health* 9:257-9.
- Rinaldi L, Musella V, Biggeri A, Cringoli G, 2006. New insights into the application of geographical information systems and remote sensing in veterinary parasitology. *Geospat Health* 1:33-47.
- Roberts DR, Rodriguez MH, 1994. The environment, remote sensing, and malaria control. *Ann N Y Acad Sci* 740:396-402.
- Shapter T, 1849. The history of the cholera in Exeter in 1832. John Churchill, London, England.
- Snow J, 1849. On the mode of communication of cholera. John Churchill, London, England.
- Snow J, 1855. On the mode of communication of cholera. 2nd ed. John Churchill, London, England.
- Stensgaard AS, Saarnak CF, Utzinger J, Vounatsou P, Simoonga C, Mushinge G, Rahbek C, Møhlenberg F, Kristensen TK, 2009. Virtual globes and geospatial health: the potential of new tools in the management and control of vector-borne diseases. *Geospat Health* 3:127-41.
- Stiles CW, 1903. Report upon the prevalence and geographic distribution of hookworm disease. *Hygienic Laboratory Bulletin* No. 10, Gov. Printing Office, Washington DC, USA.



- Stockwell D, Peters D, 1999. The GARP modelling system: problems and solutions to automated spatial prediction. *Int J Geogr Inf Sys* 13:143-58.
- Takane M, Yabe S, Tateshita Y, Kobayashi Y, Hino A, Isono K, Okayasu H, Diop OM, Tadono T, 2016. Satellite imagery technology in public health: analysis of site catchment areas for assessment of poliovirus circulation in Nigeria and Niger. *Geospat Health* 11:355-61.
- Takano H, Inoue KI, 2017. Environmental pollution and allergies. *J Toxicol Pathol* 30:193-9.
- Tatem AJ, Goetz SJ, Hay SI, 2004. Terra and Aqua: new data for epidemiology and public health. *Int J Appl Earth Obs Geoinf* 6:33-46.
- Tobler W, 1970. A computer movie simulating urban growth in the Detroit region. *Econ Geogr* 46:234-40.
- Tomlinson RF, 1974. Geographical information systems, spatial data analysis and decision making in government. Available from: <http://discovery.ucl.ac.uk/id/eprint/1563584>.
- Toreti A, Maiorano A, De Sanctis G, Webber H, Ruane AC, Fumagalli D, Ceglar A, Niemeyer S, Zampieri M, 2019. Using reanalysis in crop monitoring and forecasting systems. *Agric Syst* 168:144-53.
- Twardzik E, Antonakos C, Baiers R, Dubowitz T, Clarke P, Colabianchi N, 2018. Validity of environmental audits using GigaPan® and Google Earth Technology. *Int J Health Geogr* 17:26.
- United Nations, 2011. United Nations Climate Conference in Durban, South Africa (COP17/CMP7). Available from: <https://web.archive.org/web/20111110104913/http://www.cop17-cmp7durban.com/> Accessed: 4 April 2019.
- Valencia-López N, Malone JB, Carmona CG, Velásquez LE, 2012. Climate-based risk models for *Fasciola hepatica* in Colombia. *Geospat Health* 6:S67-85.
- Washino RK, Wood BL, 1994. Application of remote sensing to arthropod vector surveillance and control. *Am J Trop Med Hyg* 50:134-44.
- World Health Organization (WHO), 2006. Preventing disease through healthy environments. World Health Organization, Geneva, Switzerland. Available from https://www.who.int/quantifying_ehimpacts/publications/previdisexecsume.pdf - https://www.who.int/quantifying_ehimpacts/publications/preventingdisease6.pdf Accessed: 4 April 2019.
- Yang GJ, Utzinger J, Sun LP, Hong QB, Vounatsou P, Tanner M, Zhou XN, 2007. Effect of temperature on the development of *Schistosoma japonicum* within *Oncomelania hupensis*, and hibernation of *O. hupensis*. *Parasitol Res* 100:695-700.
- Yang GJ, Vounatsou P, Zhou XN, Utzinger J, Tanner M, 2005. A review of geographic information system and remote sensing with applications to the epidemiology and control of schistosomiasis in China. *Acta Trop* 96:117-29.
- Yang K, Sun LP, Huang YX, Yang GJ, Wu F, Hang DR, Li W, Zhang JF, Liang YS, Zhou XN, 2012. A real-time platform for monitoring schistosomiasis transmission supported by Google Earth and a web-based geographical information system. *Geospat Health* 6:195-203.
- Zhao X, Chen F, Feng Z, Li X, Zhou X-H, 2014. Characterizing the effect of temperature fluctuation on the incidence of malaria: an epidemiological study in southwest China using the varying coefficient distributed lag non-linear model. *Malar J* 13:192.
- Zhou XN, Yang GJ, Yang K, Wang XH, Hong QB, Sun LP, Malone JB, Kristensen TK, Bergquist NR, Utzinger J, 2008. Potential impact of climate change on schistosomiasis transmission in China. *Am J Trop Med Hyg* 78:188-94.
- Zuo H, Hutchinson MF, McMahon JP, Nix HA, 1996. Developing a mean monthly climatic database for China and Southeast Asia. In: Booth TH, ed. *Matching trees and sites*. ACIAR, Canberra, Australia, pp 10-17.