Atmosphere, Earth, Ocean & Space

Tzai-Hung Wen
Ting-Wu Chuang
Mathuros Tipayamongkholgul *Editors*

Earth Data Analytics for Planetary Health

Springer

Atmosphere, Earth, Ocean & Space

The series Atmosphere, Earth, Ocean & Space (AEONS) publishes state-of-art studies spanning all areas of Earth and Space Sciences. It aims to provide the academic communities with new theories, observations, analytical and experimental methods, and technical advances in related fields. The series includes monographs, edited volumes, lecture notes and professional books with high quality. The key topics in AEONS include but are not limited to: Aeronomy and ionospheric physics, Atmospheric sciences, Biogeosciences, Cryosphere sciences, Geochemistry, Geodesy, Geomagnetism, Environmental informatics, Hydrological sciences, Magnetospheric physics, Mineral physics, Natural hazards, Nonlinear geophysics, Ocean sciences, Seismology, Solar-terrestrial sciences, Tectonics and Volcanology.

Tzai-Hung Wen · Ting-Wu Chuang ·
Mathuros Tipayamongkholgul
Editors

Earth Data Analytics for Planetary Health

Springer

Editors
Tzai-Hung Wen
Department of Geography
College of Science
National Taiwan University
Taipei, Taiwan

Mathuros Tipayamongkholgul
Faculty of Public Health
Mahidol University
Bangkok, Thailand

Ting-Wu Chuang
Department of Molecular Parasitology
and Tropical Diseases
School of Medicine
College of Medicine
Taipei Medical University
Taipei, Taiwan

ISSN 2524-440X ISSN 2524-4418 (electronic)
Atmosphere, Earth, Ocean & Space
ISBN 978-981-19-8764-9 ISBN 978-981-19-8765-6 (eBook)
https://doi.org/10.1007/978-981-19-8765-6

Contents

Editors and Contributors

About the Editors

Dr. Tzai-Hung Wen is a professor of Geographic Information Science (GISc) and Epidemiology at National Taiwan University (NTU). He currently also serves as the Chairman of the Chinese Cartographic Association, and the Editor-in-Chief of the *Journal of Population Studies*. His research concentrates on developing spatial-temporal statistical and computational methods for examining the associations between urban environmental changes and human health risks.

Dr. Ting-Wu Chuang is a spatial epidemiologist and his research focuses on evaluating environmental influences on transmission dynamics of infectious diseases. He has been using remotely sensed satellite images and geographic information system (GIS) to undertaken spatial and temporal analyses and developing forecasting models for malaria and dengue. He is an associate professor of Molecular Parasitology and Tropical Diseases at Taipei Medical University in Taiwan.

Dr. Mathuros Tipayamongkholgul is an associate professor in Department of Epidemiology at the Faculty of Public Health, Mahidol University, Thailand. She is an epidemiologist by training and her research focus on spatial epidemiology, and time-series analysis in public health.

Contributors

Aji Kusumaning Asri Department of Geomatics, National Cheng Kung University, Tainan City, Taiwan

Luke R. Bergmann Department of Geography, University of British Columbia, Vancouver, BC, Canada

Arika Bridhikitti Earth Science Research Cluster, Mahidol University Kanchanaburi Campus, Kanchanaburi, Thailand;
Environmental Engineering and Disaster Management Program, School of Interdisciplinary Studies, Mahidol University Kanchanaburi Campus, Kanchanaburi, Thailand

Chun-Hsiang Chan Undergraduate Program in Intelligent Computing and Big Data, Chung Yuan Christian University, Taoyuan, Taiwan;
Department of Geography, National Taiwan University, Taipei, Taiwan;
Artificial Intelligence Analytics, Taiwan Cybersecurity Foundry Company, Taipei, Taiwan

Day-Yu Chao Graduate Institute of Microbiology and Public Health, College of Veterinary Medicine, National Chung Hsing University, Taichung, Taiwan

Luis Fernando Chaves Instituto Conmemorativo Gorgas de Estudios de La Salud, Ciudad de Panamá, Panamá;
Department of Environmental and Occupational Health, School of Public Health, Indiana University, Bloomington, IN, USA

Tzu-Hsin Karen Chen Yale Institute for Biospheric Studies, Yale University, New Haven, CT, USA;
Danish Big Data Centre for Environment and Health (BERTHA), Aarhus University, Roskilde, Denmark

Wei Chien Benny Chin Department of Geography, National University of Singapore, Singapore, Singapore

Tzu-How Chu Department of Geography, National Taiwan University, Taipei, Taiwan

Mariel D. Friberg Earth System Science Interdisciplinary Center (ESSIC), University of Maryland, College Park, MD, USA

Nicole L. Gottdenker Department of Veterinary Pathology, University of Georgia, Athens, GA, USA;
Center for the Ecology of Infectious Diseases, University of Georgia, Athens, GA, USA

Mei-Liang Huang Graduate Institute of Microbiology and Public Health, College of Veterinary Medicine, National Chung Hsing University, Taichung, Taiwan

Shin-Ying Huang National Center for Cyber Security Technology, Institute for Information Industry, Taipei, Taiwan

Lisbeth A. Hurtado Instituto Conmemorativo Gorgas de Estudios de La Salud, Ciudad de Panamá, Panamá

Jehn-Yih Juang Department of Geography, National Taiwan University, Taipei, Taiwan

Htet Yamin Ko Ko Department of Mathematics, Faculty of Science, King Mongkut's University of Technology Thonburi (KMUTT), Bangkok, Thailand

Ching-Hao Mao Artificial Intelligence Analytics, Taiwan Cybersecurity Foundry Company, Taipei, Taiwan

Piyawan Nakto Department of Mathematics, Faculty of Science, King Mongkut's University of Technology Thonburi (KMUTT), Bangkok, Thailand

Thanet Natisri Department of Mathematics, Faculty of Science, King Mongkut's University of Technology Thonburi (KMUTT), Bangkok, Thailand

Milixa Perea Instituto Conmemorativo Gorgas de Estudios de La Salud, Ciudad de Panamá, Panamá

Arthit Phosri Department of Environmental Health Sciences, Faculty of Public Health, Mahidol University, Bangkok, Thailand

Alexander V. Prishchepov Department of Geosciences and Natural Resource Management (IGN), University of Copenhagen, København K, Denmark

Chystrie A. Rigg Instituto Conmemorativo Gorgas de Estudios de La Salud, Ciudad de Panamá, Panamá

Nawin Rinrat Department of Mathematics, Faculty of Science, King Mongkut's University of Technology Thonburi (KMUTT), Bangkok, Thailand

Clive E. Sabel Danish Big Data Centre for Environment and Health (BERTHA), Aarhus University, Roskilde, Denmark;
Department of Public Health, Aarhus University, Aarhus, Denmark;
Health Research Institute, University of Canberra, Canberra, Australia

Mathuros Tipayamongkholgul Department of Epidemiology, Faculty of Public Health, Mahidol University, Bangkok, Thailand

Pariwate Varnakovida Department of Mathematics, Faculty of Science, King Mongkut's University of Technology Thonburi (KMUTT), Bangkok, Thailand

Michael C. Wimberly Department of Geography and Environmental Sustainability, University of Oklahoma, Norman, OK, USA

Chih-Da Wu Department of Geomatics, National Cheng Kung University, Tainan City, Taiwan;
National Institute of Environmental Health Sciences, National Health Research Institutes, Miaoli, Taiwan

Hong-Dar Isaac Wu Department of Applied Mathematics and Institute of Statistics, Taichung, Taiwan

Part I
Environmental Quality Monitoring

Chapter 1
Applications of Remote Sensing for Air Pollution Monitoring in Thailand: An Early Warning for Public Health

Arika Bridhikitti

Abstract There are also consistent findings on the adverse effects of air pollution on public health in Thailand. Small size particulate matter, or $PM_{2.5}$, is the most pronounced air pollutant during the haze crisis. $PM_{2.5}$ often comes along with other polluted gases, including carbon monoxide (CO), oxides of nitrogen ($NO_x = NO + NO_2$), sulfur dioxide (SO_2), ozone (O_3), and volatile organic compounds (VOCs). This chapter presents various applications of remote sensing technology for air pollution monitoring, warning, and forecasting. These applications can help assess human exposure to air pollution and determine health risks associated with air pollution. The presentation is divided into four sections. The first section provides an overview of Earth Observing Satellites and current remote sensing technology for air pollution observations. The second section is on assessing the magnitude of atmospheric pollutants and human exposure levels from remote sensing. The third section is on air pollution source identification using remote sensing technology. Finally, the fourth section discusses the possibility of employing satellite information for forecasting haze episodes as the early warning tool. The presentation is based on the recent deployment of remote sensing technology for air pollution monitoring, especially reported for the cases of Thailand and the Southeast Asian region.

Keywords Satellite · Remote sensing · Air quality forecast · Air pollution · Particulate matter · Southeast Asia

A. Bridhikitti (✉)
Earth Science Research Cluster, Mahidol University Kanchanaburi Campus, 199 Moo 9, Lumsum Sub-District, Saiyok District, Kanchanaburi 71150, Thailand
e-mail: arika.bri@mahidol.edu

Environmental Engineering and Disaster Management Program, School of Interdisciplinary Studies, Mahidol University Kanchanaburi Campus, 199 Moo 9, Lumsum Sub-District, Saiyok District, Kanchanaburi 71150, Thailand

T.-H. Wen et al. (eds.), *Earth Data Analytics for Planetary Health*, Atmosphere, Earth, Ocean & Space, https://doi.org/10.1007/978-981-19-8765-6_1

1.1 Introduction

Air pollution in Thailand has become public attention nowadays. In the dry season from December to March, the pollution is often reported at harmful levels, exceeding the national air quality standard. Pinichka et al. [1] studied the burden of disease attributed to air pollution and found that NO_2 and $PM_{2.5}$ could account for 10% and 7.5%, respectively, of the disease burden for all mortality. The $PM_{2.5}$ could also contribute to 16.8% of lung cancer cases and 14.6% of cardiovascular cases, whereas the NO_2 was responsible for 7.8% of respiratory mortality [1]. Jenwitheesuk et al. [2] also showed evidence on $PM_{2.5}$-induced risk of colon cancer, with 15% risk increased for every ten micrograms $PM_{2.5}$ m^{-3} increased. Furthermore, US researchers have strong evidence showing that short-term exposure to $PM_{2.5}$ could significantly increase the risk of COVID-19 cases and death [3]. Not only the small-size aerosol, but the coarse-size particulate matter (PM_{10}) could also suddenly increase in hospital admissions as found a strong association between the PM_{10} level and the number of cardiovascular and respiratory admission in Bangkok, the most populated city in Thailand [4].

The Thai Government set up several policies and solutions to tackle haze pollution. The haze mitigation policy includes inspection and maintenance of vehicles, mandating higher quality fuel, and balancing productivity and environmental conservation in agricultural production [5]. During the haze episode, the policy relied on single command-and-control, framed by the central and provincial government [6]. The policies include prohibiting biomass burning, applying water sprays in public areas, roadside inspection on vehicle exhaust emissions, etc. [7]. Moran et al. [5] criticized that key issues of unsuccessful haze abatement in Thailand are low public participation and poor enforcement of laws or regulations. Scientists recommended policy outlines to minimize health effects from air pollution in the short term and eliminate the haze in the long run. The policy outlines included improvement of capacities to monitor, assess source inventory, and forecast air pollution, probably by incorporating applications of satellite retrievals in combination with ground measurements to fill spatial monitoring gaps [7, 8].

Satellite technology is widely applied for environmental monitoring since it provides spatial advantages for understanding the atmosphere and the land surface at the corresponding timeframe. Since air pollution is mainly released from surface activities, satellite observations could be helpful to provide strong connections between hot spots and pollution plumes or between urban land cover and heat island zone. The satellites could have multiple sensors onboarded, and each sensor was designed to capture specific radiative bands with certain viewing angles. Many Earth Observing Satellites detect solar reflectance and long-wave radiation from the Earth. These electromagnetic spectrums are appropriate for observing Ozone, Aerosol, Hydrocarbons, Greenhouse Gases, and Water vapor in the atmosphere. Some satellite sensors, such as Multiangle Imaging Spectroradiometer (MISR) and Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP), provide aerosol optical properties, beneficial for aerosol source identification. Furthermore, ground-based remote

sensing, such as Aerosol Robotic Network or AERONET, is typically used to validate satellite products. It can provide scattering and extinction properties of the atmospheric aerosols with a high temporal resolution by tracking direct sun and sky radiances.

1.2 Overview Earth Observing Satellites and Current Remote Sensing Technology for Air Pollution Observations

1.2.1 Earth Observing Satellites

By observing the Earth at the top of the atmosphere, satellite remote sensing can provide aerosol and gaseous compositions in the total atmospheric column basis. With approximately 1 μm and smaller, the aerosol highly scatters the energy spectrum in the visible to the near-infrared band from 300 to 1,000 nm. In addition, some types of aerosol absorb energy. The absorbing aerosols include black carbon (absorb both solar and thermal radiation) and mineral dust (scatter sunlight but absorb thermal infrared) [9]. The spectrophotometer is designed to observe solar radiation, and it can measure the extinction (both scattering and absorption) of the solar beam attributed to aerosol, called aerosol optical depth (AOD). Several satellite sensors have been previously employed for the studies of air pollution monitoring and assessment in Thailand. The summary of those sensors is detailed in Table 1.1.

Among the satellites, A-train (afternoon train) satellite constellation can be advantageous for atmospheric observation due to its combining multiple satellite remote sensors to better understand atmospheric and land dynamics. The constellation recently consisted of four satellites, orbiting in sequences, which are OCO-2 (launched in 2014), GCOM-W1 (since 2012), AQUA (since 2002), and AURA (since 2004) [10]. The satellites cross the equator around 1:30 PM local time with a 16-day repeating cycle.

The OCO-2 stands for Orbiting Carbon Observatory-2. The primary mission of OCO-2 is to quantify atmospheric carbon dioxide by its absorption of the visible band [11]. The GCOM-W1 stands for the Global Change Observation Mission-Water Satellite 1. It can explain the water cycle and climate change by observing the atmosphere and water bodies [12]. The GCOM-W1 is onboard with the Advanced Microwave Scanning Radiometer 2 (AMSR2), detecting microwave radiated from the ground, the water surface, and atmospheric gases, primarily greenhouse gases and water [12]. AQUA is designed to provide information about the Earth's hydrosphere. Furthermore, AQUA also enables the observation of aerosol via Moderate Resolution Imaging Spectroradiometer (MODIS) instrument and greenhouse gases via the Atmospheric Infrared Sounder (AIRS) along with the Advanced Microwave Sounding Unit (AMSU-A). AURA consists of four instruments to provide comprehensive studies on stratospheric and tropospheric compositions, including ozone,

Table 1.1 Summary of satellite sensors typically used for air pollution monitoring in Thailand

Sensors	Satellite	Standard detecting air pollution products	Temporal resolution	Spatial resolution, m	Reference
Moderate Resolution Imaging Spectrophotometer (MODIS)	TERRA since 2000, AQUA since 2002	Aerosol optical depth with a combined land and ocean algorithms	Level 2: Daily Level 3: Daily, 8-day, Monthly	Level 2: 10 km × 10 km (at nadir) Level 3: 1° × 1°	NASA Atmosphere Discipline Team Imager Products, https://modis-images.gsfc.nasa.gov/products.html
		Aerosol optical depth with dark target and deep blue algorithm	Level 2: Daily Level 3: Daily, 8-day, Monthly	Level 2: 10 km × 10 km (at nadir) and 3 km × 3 km (at nadir) Level 3: 1° × 1°	
		Aerosol optical depth with Multi-Angle Implementation of Atmospheric Correction, or MAIAC, algorithm	Daily	Level 2: 1 km × 1 km (at nadir)	
Measurements of Pollution in the Troposphere (MOPITT)	TERRA since 2000	CO Total Column	Daily, Monthly	Level 3: 1° × 1° horizontally, 10-level vertical (surface, 900 hPa, 800 hPa, …, 100 hPa)	NASA TERRA, https://terra.nasa.gov/about/terra-instruments/mopitt
		CO Mixing Ratio for the layer above each pressure level			
		CO Mixing Ratio at the surface			

(continued)

Table 1.1 (continued)

Sensors	Satellite	Standard detecting air pollution products	Temporal resolution	Spatial resolution, m	Reference
Multi-angle Imaging Spectroradiometer (MISR)	TERRA since 2000	Aerosol Optical Depth at 550 nm (also *including absorption, non-spherical, small-mode, medium-mode, large-mode*) Ångström Exponent at 550 and 860 nm	Level 3: Daily, Monthly, Quarterly, and Yearly	Level 2: 4.4 km × 4.4 km Level 3: 0.5° × 0.5°	NASA Jet Propulsion Laboratory, California Institute of Technology, https://misr.jpl.nasa.gov/
Ozone Monitoring Instrument (OMI)	AURA since 2004	Total column ozone: DOAS technique	Level 3: daily, monthly	Level 2: 13 km × 48 km at nadir Level 3d: 1° × 1° Level 3e: 0.25° × 0.25°	[16]
		Total column ozone: TOMS version 8 method			
		Aerosol: near UV algorithm			
		Aerosol: multi-wavelength algorithm			
		Ozone profile			
		Total column SO_2, HCHO, BrO, OClO, NO_2			

(continued)

Table 1.1 (continued)

Sensors	Satellite	Standard detecting air pollution products	Temporal resolution	Spatial resolution, m	Reference
Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP)	CALIPSO since 2006	Aerosol Profile	Level 2: ~ 2 to 3 times per month	Level 2: 5 km × 5 km horizontal, 60 m vertical in the troposphere and 180 m vertical in the stratosphere	NASA, https://www-calipso.larc.nasa.gov/
		Vertical Feature Mask (including aerosol subtype)		Level 2: 5 km × 5 km horizontal, 1 km vertical (from 0 to 30 km)	
		Tropospheric Aerosol Profile (similar to Vertical Feature Mask)	Level 3: Monthly	Level 3: 2° latitude × 5° longitude, 60 m vertical from 0.5 to 12 km	

(continued)

Table 1.1 (continued)

Sensors	Satellite	Standard detecting air pollution products	Temporal resolution	Spatial resolution, m	Reference
Visible Infrared Imaging Radiometer Suite (VIIRS)	Suomi NPP since 2011	Level 2 (or Environmental Data Record, EDR): aerosol optical thickness (AOT), aerosol particle size parameter (include Ångström exponent at 445 and 672 nm over land and at 865 and 1610 nm over the ocean), and suspended matter (include aerosol type and smoke concentration) Note: *available for Deep Blue and Dark Target algorithms as well*	Level 2: Daily Level 3: Daily, Monthly	Level 2: ~ 6 km × 6 km at nadir Level 3: $1^{o} \times 1^{o}$	[17–19]
		Aerosol data: Deep Blue Algorithm	Level 3: Daily, Monthly		

Note Level 1 is calibrated, and geolocated product, Level 2 is derived geophysical variables at the same resolution and location as Level 1 source data, or swath products, Level 3 is gridded variables in derived spatial and/or temporal resolutions

water, greenhouse gases, halogen compounds, oxides of nitrogen, carbon monoxide, and aerosols [13]. The AURA sensors include the High Resolution Dynamics Limb Sounder (HIRDLS), the Microwave Limb Sounder (MLS), the Ozone Monitoring Instrument (OMI), and the Tropospheric Emission Spectrometer (TES) [13].

Besides the AQUA satellite, the MODIS instrument is also onboard the TERRA satellite to provide a higher temporal resolution of the atmospheric aerosol. The TERRA passes to tropic at around 10.30 AM (ascending) and 10.30 PM (descending), whereas the AQUA is around 1.30 PM (ascending) and 1.30 AM (descending). Thus, their combined product can be provided four times daily. Even though aerosol observation can only perform in the daytime due to the requirement of light scattering, the nighttime satellite imagery can provide helpful information on socioeconomic parameters, such as population density and gross domestic product, and greenhouse gas emissions [14]. Thus MODIS product quite temporally advantages over other instruments and is universally employed for ground-level air pollution monitoring.

1.2.2 Ground-Based Remote Sensing

With currently more than 1,000 stations (see Fig. 1.1) and the number is growing, the AERONET (Aerosol Robotic Network, https://aeronet.gsfc.nasa.gov/) provided good spatial coverage of aerosol and cloud in the atmosphere. It is widely used as ground truth measurement for satellite retrievals of aerosol around the world. Currently, a total of 21 AERONET sites are based in Thailand. The AERONET program was established by the National Aeronautics and Space Administration (NASA), and LOA-PHOTONS, the French National Observatory for Aerosol, has been operated for more than 25 years. The AERONET aerosol products are measured by sun and sky photometers to measure direct and diffuse radiation. Details on the AERONET products are in Table 1.2. The aerosol products include spectral aerosol optical depth (AOD) and aerosol inversions, which provide aerosol microphysical and radiative properties. In parallel, the NASA Micro-Pulse Lidar Network (MPLNET: https://mplnet.gsfc.nasa.gov/) operates in conjunction with the AERONET to provide vertical integration structures of aerosol and cloud. The MPLNET project started with full operation in 2000. The MPLNET has three active sites in Thailand—Princess Sirindhorn Astro Park in the North, Silpakorn University in the Central, and Songkhla Regional Observatory in the South. Nonetheless, the current aerosol model cannot sufficiently describe vertical aerosol extinction measured during field campaigns over the Indian Ocean [15].

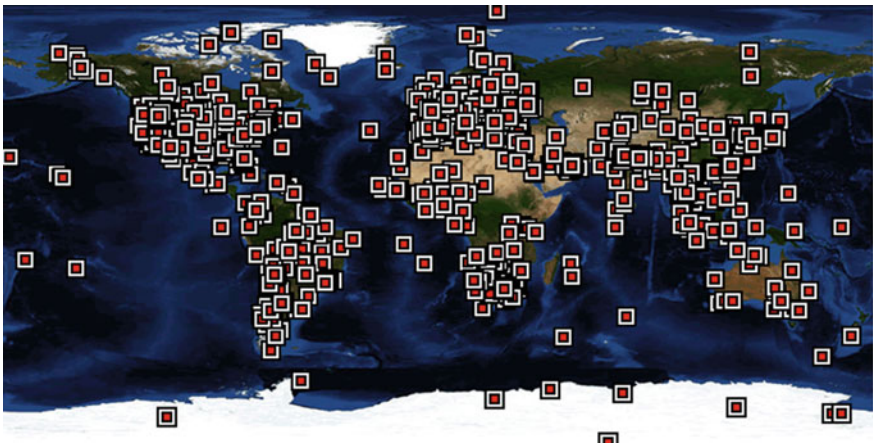


Fig. 1.1 AERONET sites (AERONET, 28 August 2021, retrieved from https://aeronet.gsfc.nasa.gov/)

Table 1.2 Products from direct sun measurements and aerosol optical properties from sun-sky photometers inversion products acquired from AERONET (https://aeronet.gsfc.nasa.gov/new_web/data.html)

AERONET products	Unit	Range
Direct sum measurement products		
Aerosol optical depth (AOD) Levels 1.0, 1.5, and 2.0, AOD (λ)	unitless	Level 1.0 unscreened Level 1.5 cloud screened Level 2.0 quality assured
Angstrom parameter, AE at 440–870 nm	Unitless	
Total water vapor at 940 nm	cm^3/cm^2 or g cm^{-2}	
Aerosol inversion products		
Volume particle size distribution, dV(r)/dlnr	$\mu m^3/\mu m^2$	$0.05\ \mu m \leq r \leq 15\ \mu m$
Volume concentration, C_v	$\mu m^3/\mu m^2$	For total, fine, and coarse aerosol modes Note: the fine and coarse separation point is the minimum within 0.439 to 0.992 μm
Volume median radius, r_v (mean logarithm of radius)	μm	
Standard deviation from r_v	μm	
Effective radius, r_{eff}	μm	
Real-part refractive index, n(λ)	Unitless	λ = 440, 675, 870, 1,020 nm 83 scattering angles, ~$7^o \leq \Theta \leq$ ~170^o
Imaginary-part refractive index, k(λ)	Unitless	
Single scattering albedo, SSA(λ)	Unitless	
Phase function for each scattering angle, P(Θ, λ)	Unitless	
Asymmetry parameter for each phase function, cos(Θ)	Unitless	

1.3 Assessing Magnitude and Extents of Atmospheric Pollutants and Human Exposure Level from Remote Sensing

1.3.1 Tropospheric Ozone O_3 and Its Precursors

Once the tropospheric O_3 is generated, it can disintegrate into OH-radical. Though both the O_3 and OH-radical have short-lived in the environment, they have high oxidation potential, which can harm the soft tissues of plants, animals, and humans. Several previous studies confirmed that the O_3 could significantly increase human mortality [20–22]. The ozone is not directly emitted from the sources. It is the product of the photochemical reaction, in which the precursors mainly are NO_x and non-methane volatile organic compounds.

In Southeast Asia, biomass burning is the key source of O_3 and its precursors [23, 24]. Mekaumnuaychai et al. [25] and Yimlamaid et al. [26] employed satellite retrievals to assess the ozone precursors across Thailand. They found that the formaldehyde (HCHO) retrievals from the OMI/AURA satellite may not well reach the ground level, but the NO_2 retrievals from OMI/AURA and other satellites can represent the ground levels well in dry seasons [25, 26]. Furthermore, the UV radiation product of OMI can infer surface UV radiation with aerosol correction [27, 28]. The UV radiation provides energy driving photochemical reaction for O_3 generation in the troposphere. However, since the OMI provides total-column-based O_3, it may not sufficiently represent the ground levels. Tropospheric O_3 can be estimated by subtracting the OMI total column ozone with the MLS stratospheric ozone, which both OMI and MLS sensors are onboard the AURA satellite [24]. Thus, the data acquired from the OMI/MLS is recommended for assessing variation in tropospheric ozone and suggesting mechanisms for surface ozone control.

1.3.2 Aerosol Loading

Aerosol optical depth (AOD) is one of the aerosol optical products retrievable from satellites with visible to near-infrared band detectors. Satellite remote sensing is advantageous for studying long-term aerosol trends in regional to global scales [29]. Mehta et al. [29] employ satellite retrievals (both MODIS and MISR-Multiangle Imaging Spectroradiometer) for studying regional trends of aerosol magnitude. Their finding showed a noticeable increase in annual AOD over Southeast Asia from September to May, covering post- and pre-monsoon season [29]. The CSR, United States, attempts to predict air quality from MODIS AOD observation and to isolate local sources from the long-range transport sources using ground observation networks [30]. Though AOD is strongly affected by the amount of aerosol scattering and absorption, the AOD may not reach the ground level. The performance of the

new MODIS collection (6.1) aerosol products is spatially and temporally dependent [31]. The performance over the Asia continent is low, with less than 62% of the data sampling falling within the confidence intervals [31]. It could be likely due to the complex geography and various aerosol mixtures from natural and anthropogenic sources [31].

In northern Thailand, MODIS AOD was moderately-to-poorly correlated with ground-observed PM_{10}—$R^2 = 0.42$ during forest fire episodes [32] and $R^2 = 0.21$ for a normal episode [33]. Scientists and engineers have attempted to improve ground-level $PM_{2.5}$ prediction algorithm from satellite observation by developing empirical formula using multiple predictors, which is satellite-based AOD, land use, and meteorological variables, etc. With ground-based meteorological correction, the MODIS AOD at about 10:30 AM correlates better with the hourly ground-level $PM_{2.5}$ and PM_{10} using multiple linear regression ($R^2 = 0.74$), based on the study in Chiangmai, Thailand [33].

Since satellite retrievals are based on the pollutants' electromagnetic scattering and extinction properties in the atmospheric layer, they may not promise with the ground-based measurements using gravitational methods or chemical analytical methods. Nonetheless, satellite retrievals could be well responding to the ground during the high aerosol episode. Sayer et al. [34] studied corresponding between the MODIS collection 6 AOD and the AOD from ground-based sunphotometer measurements from Aerosol Robotic Network (AERONET) in the study at Doi Ang Khang, remote forest area in NW Thailand during the biomass burning episode. They found a strong correlation of R^2 0.93−0.94 [34]. The AERONET sun photometer employs the same approach like that of the satellites, considering aerosol optical properties. The correlations for board arrays of land covers and aerosol loadings were slightly declined as found to be R^2 of 0.81 for MODIS AOD and AERONET AOD and 0.68 for VIIRS/Suomi NPP AOD and AERONET AOD across the Southeast Asia region [35].

1.3.3 Biomass Burning Smoke

Biomass burning releases smoke plumes containing carbonaceous aerosol, CO_2, CO, NO_x, and other uncompleted burning products to the atmosphere [36]. The emission of biomass burning smoke in the Southeast Asian region is globally concerned in terms of its magnitude and its contribution to climate change [37]. The biomass burning contributes to 49% of total PM_{10} in Southeast Asian mainland, and the highest contribution was from Laos (73%), followed by Myanmar (69%), Cambodia (59%), Thailand (45%), China (33%) and Vietnam (31%) [37]. In Thailand, a total of 117.7 Mt of rice residue was left on the field after being harvested, and 15% of the amount was later burnt. The burning emitted approximately 2.19 Mt of combined air pollutants, and 30% of the amount was from the lower Northern Thailand, followed by central Thailand (26%) and western Thailand (17%) [36].

Satellite observations can provide reasonable estimates of smoke magnitude as found in the studies in northern Thailand. Sukitpaneenit and Oanh [32] monitored forest fire smoke using MODIS and Measurement of Pollution in the Troposphere (MOPITT) onboard TERRA satellite. They found that the MODIS AOD can capture variation in ground-measured PM_{10} by ~42% using a linear regression model, whereas the MOPITT CO can explain the ground-level CO by ~35%. Lalitaporn and Boonmee [38] employed tropospheric NO_2 column from SCIMACHY, OMI, GOME-2A, and GOME-2B sensors. They found that high levels of the NO_2 corresponded well with ground-level NO_2 ($R^2 = 0.29$ to 0.66) and a number of biomass burning hot spots.

1.4 Air Pollution Source Identification Using Remote Sensing Technology

1.4.1 Aerosol Optical Properties

Radiative sensors can sense atmospheric aerosols through their scattering and absorption of incoming solar radiation. With these optical properties, the aerosols can influence the Earth's radiative balance, either negative effect (cooling) or positive effect (warming), depending upon their composition, shape, size, and mixing state [9]. Absorbing aerosols include black carbon. The aerosol optical properties can be expressed from the single scattering albedo (SSA), aerosol optical depth (AOD), and Angström exponent (AE), etc. The SSA is the ratio of scattering coefficient to the extinction coefficient (combine scattering and adsorption effects). The AOD refers to the magnitude of aerosol by representing the depth of the direct solar beam lost when traveling through the aerosol layer [9]. AE is the log-slope exponent of the spectral aerosol optical depth between two wavelengths [39].

Recently, scientists have tried to simulate aerosol properties using various climate and aerosol models. Takemura et al. [39] employed an aerosol transport model coupled with an atmospheric general circulation model to simulate the SSA of primary aerosols—carbonaceous (organic and black carbons), sulfate, and soil dust and sea salt aerosols. Age aerosol with internal mixing (among the same type of aerosols) or external mixing (among different types) could exhibit lower SSA since its increasing adsorption efficiency. In contrast, the scattering efficiency is less affected by the mixing [39]. Different aerosol types can be distinguished by knowing the aerosol optical properties, as shown in Tables 1.3 and 1.4. Clean oceanic sea salt exhibits high light scattering (SSA of ~0.99, (Mallet et al. (2004) cited in [40]) and low light absorption (imaginary-part reflective index <0.001) [41]. Long-range transport urban/industrial aerosol from East Asia to Southeast Asia also exhibited high SSA of >0.95 (at 440 nm) and coarser fine mode size of >0.2 μm [42]. The high SSA could be due to the aerosol's hygroscopic growth and secondary aerosol formation along the air trajectory. Nonetheless, long-range transport of the polluted air masses

from NE Asia and SE Asia to the remote island, Mauna Loa in HAWAII, did not exhibit high scattering properties, suggesting its hydrophilic properties and insignificant secondary aerosol formation [43]. From January to April, the aerosol exhibited a lower SSA of approximately 0.90, a higher refractive index, and smaller fine-mode size, suggesting biomass burning smoke [42]. The urban aerosol exhibited a mean SSA of approximately 0.90 (440 nm) or lower. The coarse-mode aerosol, possibly road dust or soil dust, played a role from October to January when seasonal winds were strongest [42].

Scientists try to understand the aerosol optical properties and employ the knowledge to indicate the potential origins of the aerosols. There are many previous studies employing satellite retrievals to identify sources of aerosols, as the detail was given in the following sections.

1.4.2 Local Biomass Burning

Biomass burning is the key source of haze pollution in the Southeast Asia continent, including Thailand. Agricultural residue burning for land preparation occurs regularly before starting cropping season in May [5, 57]. In northern Thailand, forest fires can often be found with agricultural burning resulting in poor air quality in remote areas, such as Doi Ang Khang [34].

Remote sensing can be used to identify biomass burning sources of air pollution by its association among fire hotspots, air pollutant loading, and land cover. Kamthonkiat et al. [58] developed an empirical model using Landsat 8 imagery to estimate ground-measured PM_{10}. Their finding showed that the model is highly correlated with land use/cover, primarily agricultural land and forests, and fire hotspots. They conclude significant contributions of biomass burning in the air pollution in Nan, Thailand [58].

Consistently, the BASE-ASIA in 2006 and the 7-SEAS in 2010–13 campaigns were conducted over northern Southeast Asia by collaboration among the Southeast Asia START region center, NASA, and the University of Hawaii to understand aerosol and cloud properties in this region using satellite and *in-situ* measurements [59]. Based on the BASE-ASIA deployment over Thailand, Myanmar, and Laos, the result shows a good correlation ($R^2 = 0.66$) between the MODIS fire activity and MODIS AOD, suggesting a significant smoke from forest fires and agricultural crop burning [59]. Furthermore, Li et al. [52] presented comprehensive biomass-burning aerosol properties measured at a remote site of Phimai, in NE Thailand. The aerosol had dominant carbonaceous components and substantial loadings of SO_4^{2-}, NH_4^+, NO_3^-. The aerosol also had hygroscopic properties, enhancing aerosol light scattering and hygroscopic growth [59].

Table 1.3 Aerosol optical properties from the literature

Studied areas	Aerosol chemical compositions	Potential sources	Size	Angstrom exponent	SSA	AOD	Reference
The remote island, Mauna Loa in HAWAII	–	Long-range transport of dust and pollution aerosols from NE Asia in spring from March to May	Coarse	–	0.87 (550 nm)	–	[43]
		Biomass burning from SE Asia in spring from March to May	Fine	–	0.86 (550 nm)	–	
		Pacific Ocean	Fine	–	0.86 (550 nm)	–	
East Coast of US	Sulfate	Oceanic	0.18–0.21 μm 1.43–2.05 μm Geometric median diameter	–	0.96 ± 0.03 (550 nm)	0.13 ± 0.10	[44]
Thar Desert, Jodhpur India	High hematite content Moorthy et al. (2007) cited in Gogoi et al. (2013)	Desert dust With high absorption in Blue, UV, and IR	Coarse	0.71 ± 0.2	–	0.66 ± 0.14 (500 nm)	[45]

(continued)

Table 1.3 (continued)

Studied areas	Aerosol chemical compositions	Potential sources	Size	Angstrom exponent	SSA	AOD	Reference
Urban central China	–	Coal combustion (both local and transported) in winter Gas to particle conversion and hygroscopic growth of Nitrates, sulfates, and ammonium sulfates particles	Fine mode decreases from summer (0.23 μm) to spring (0.17 μm)	1.16–1.24 (440–870)	0.87 in winter 0.93 in summer Mean 0.89 (440 nm)	1.16–1.28	[46]
Kanpur, N India	–	Industries, coal thermal power plants, automobile exhausts, bio-fuel combustions, biomass burning in post-monsoon	0.2 μm, 2 μm	1.30 ± 0.11 (440–870 nm)	–	>0.93	[47]
		Local mineral dust in monsoon	2 μm	0.36 ± 0.41 (440 – 870 nm)	–	–	
Delhi, India in winter	–	Secondary (traffic) in the winter morning	$PM_{2.5}$	–	0.93 ± 0.03 (500 nm)	0.95 ± 0.32	[48]
Boreal forest, N Europe	–	Soil dust, sea salt	PM_{10}	1.80 ± 0.55 (450–700 nm)	0.87 ± 0.07 (550 nm)	–	[49]

(continued)

Table 1.3 (continued)

Studied areas	Aerosol chemical compositions	Potential sources	Size	Angstrom exponent	SSA	AOD	Reference
		Soil dust, sea salt, brown carbon (BC + OC) from wood burning in winter, and secondary reaction in summer	PM1	2.22 ± 0.44 (450–700 nm)	0.85 ± 0.08 (550 nm)	–	
Urban Beijing	–	Secondary reaction, cooking, traffic (day), diesel truck transport (night)	–	–	0.82, 0.77, 0.78 (670 nm)	–	[50]
Amazon Basin	–	Biomass burning	1.5 μm cut point	1.8 ± 0.2 (440–670 nm)	0.92 ± 0.02 (545 nm)	3.0	[51]
Laboratory	Sea salt	–	–	–	0.99	–	Mallet et al. (2004) cited in [40]
Coastal industrial areas, France	Sea salt + internally mixed with Black Carbon	–	–	–	0.75	–	
Coastal industrial areas, France	Sea salt + externally mixed with Black Carbon	–	–	–	0.85	–	
Remote Agricultural site, Phimai, NE Thailand	Black Carbon, Organic Carbon, K^+, SO_4^{2-}, NH_4^+, NO_3^-	Biomass burning	–	–	0.86 ± 0.04 to 0.92 ± 0.02 (550 nm)	–	[52]

(continued)

Table 1.3 (continued)

Studied areas	Aerosol chemical compositions	Potential sources	Size	Angstrom exponent	SSA	AOD	Reference
Indochina peninsula	–	Biomass burning	–	–	0.91 ± 0.02 (440–1020 nm) with the lowest over the northern location of Chiang Mai (~0.85)	–	[53]
Boreal forests or peatlands	–	Boreal forest fires Peat burning	Larger fine mode	–	0.95–0.97 (440–1020 nm)	–	[54]
Grasslands, Scrublands, Croplands	–	Biomass burning	Smaller fine mode	–	0.88–0.90 (440–1020 nm)	–	
Tropical forest, Alta Floresta, Brazil	–	Forest fires	Fine mode size, $\mu m = 0.151 + 0.0099\log(2.0 \times AOD_{fine}$ at 440 nm) Coarse mode size, $\mu m = 3.2\ \mu m$	1.95 for 440–870 nm	~ 0.92 (440–1020 nm)	3.78 (max at 440 nm)	
African Savannah, Mongu, Zambia	–	Savannah fires	Fine mode size, $\mu m = 0.175 + 0.012\log(0.12 \times AOD_{fine}$ at 440 nm) Coarse mode size, $\mu m = 3.34$	1.89 for 440–870 nm	~ 0.85 (440–1020 nm)	2.29 (max at 440 nm)	

(continued)

Table 1.3 (continued)

Studied areas	Aerosol chemical compositions	Potential sources	Size	Angstrom exponent	SSA	AOD	Reference
Bac Giang, Vietnam	–	Long-range transport urban/industrial aerosol	r_{eff}-F/C, μm = 0.226/2.465	1.268 for 440–870 nm	0.951/0.952/0.951/0.950 (440/675/870/1020 nm)	–	[42]
	–	Biomass burning smoke	r_{eff}-F/C, μm = 0.167/2.264	1.409 for 440–870 nm	0.900/0.899/0.899/0.897 (440/675/870/1020 nm)	–	
Mukdahan, NE Thailand	–	Biomass burning smoke	r_{eff}-F/C, μm = 0.172/2.453	1.544 for 440–870 nm	0.911/0.896/0.885/0.872 (440/675/870/1020 nm)	–	
	–	Long-range transport urban/industrial aerosol	r_{eff}-F/C, μm = 0.216/2.789	1.328 for 440–870 nm	0.950/0.943/0.937/0.931 (440/675/870/1020 nm)	–	
Pimai, NE Thailand	–	Biomass burning smoke	r_{eff}-F/C, μm = 0.168/2.472	1.541 for 440–870 nm	0.905/0.891/0.886/0.878 (440/675/870/1020 nm)	–	
	–	Long-range transport urban/industrial aerosol	r_{eff}-F/C, μm = 0.203/2.794	1.365 for 440–870 nm	0.950/0.946/0.945/0.942 (440/675/870/1020 nm)	–	
Silpakorn U., Central Thailand	–	Local urban/industrial aerosol	r_{eff}-F/C, μm = 0.156/2.386	1.560 for 440–870 nm	0.884/0.873/0.850/0.830 (440/675/870/1020 nm)	–	
	–	Biomass burning smoke	r_{eff}-F/C, μm = 0.170/2.257	1.536 for 440–870 nm	0.917/0.906/0.886/0.870 (440/675/870/1020 nm)	–	

(continued)

Table 1.3 (continued)

Studied areas	Aerosol chemical compositions	Potential sources	Size	Angstrom exponent	SSA	AOD	Reference
	–	Local urban/industrial aerosol with high fraction of soil dust	r_{eff}-F/C, μm = 0.192/2.990	1.274 for 870–440 nm	0.908/0.905/0.889/0.874 (440/675/870/1020 nm)	–	

Note Internal mixing = mixing among the same type of particles, External mixing = mixing with different types of particles

Table 1.4 Complex refractive indices of the aerosols reported in the literature

Aerosol type	Real (550 nm)	Imaginary (550 nm)	Reference
Water soluble	1.53	0.006	[41]
Dustlike	1.53	0.008	
Soot	1.75	0.44	
Oceanic	1.381	4.26×10^{-9}	
Clean continental (water soluble 48% + dustlike 52%)	1.53	0.00704	
Urban (water soluble 97.5% + dustlike 0.3% + soot 2.2%)	1.535	0.0156	
Maritime (water soluble 95% + Oceanic 5%)	1.388	0.0003	
Sea salt and aged sea salt (Cl + Na + S > 85%)	1.53	–	[55]
Ammonium sulfate (S > 85%)	1.53	–	
Silicates (Al + Si > 60%)	1.53	–	
Metal oxides/hydroxides (Al, Ti, Mn, Fe, Cu, Ni, Zn, Pb > 80%)	2.5	0.05	
Calcium sulfate (Ca > 30%, S > 40%)	1.53	–	
Carbonates (Ca > 60%, S < 20%)	1.53	–	
Soot	1.5	0.47	
Biological (minor elements: Na, Mg, P, S, Cl, K, Ca)	1.4	–	
Carbon/sulfate mixed particles (S > 15%)	1.5	0.05	
Rest of the carbon-rich particles	1.53	–	
Soil dust, sea salt in Boreal forest, N Europe	1.518 ± 0.067	0.02 ± 0.018	[49]
Soil dust, sea salt, brown carbon (BC + OC) from wood burning in winter and secondary reaction in summer in Boreal forest, N Europe	1.484 ± 0.054	0.025 ± 0.018	

(continued)

Table 1.4 (continued)

Aerosol type	Real (550 nm)	Imaginary (550 nm)	Reference
Biomass burning aerosol (oxalic acid 0.6 $\mu g\ m^{-3}$, KNO_3 0.8 $\mu g\ m^{-3}$, K_2SO_4 1.3 $\mu g\ m^{-3}$, Ammonium sulphate 2.3 $\mu g\ m^{-3}$, Levoglucocsan 1.4 $\mu g\ m^{-3}$, EC 5.6 $\mu g\ m^{-3}$, OM 24.7 $\mu g\ m^{-3}$)	1.54	0.013	[56]
Organic matters	1.4	–	
Elemental carbon	1.87	0.22	
Long-range transport urban/industrial aerosol (at 440/675/870/1020 nm)	1.403/1.416/1.423/1.418 at Bac Giang, Vietnam 1.402/1.416/1.423/1.420 at Mukdahan, NE Thailand 1.409/1.423/1.432/1.434 at Pimai, NE Thailand	0.007/0.006/ 0.005/0.005 at Bac Giang, Vietnam 0.007/0.007/ 0.007/0.007 at Mukdahan, NE Thailand 0.006/0.006/ 0.005/0.005 at Pimai, NE Thailand	[42]
Biomass burning smoke (at 440/675/870/1020 nm)	1.428/1.448/1.462/1.458 at Bac Giang, Vietnam 1.433/1.448/1.454/1.451 at Mukdahan, NE Thailand 1.423/1.437/1.450/1.445 at Pimai, NE Thailand 1.413/1.426/1.437/1.433 at Silpakorn U., C Thailand	0.014/0.011/ 0.010/0.009 at Bac Giang, Vietnam 0.014/0.013/ 0.012/0.013 at Mukdahan, NE Thailand 0.014/0.013/ 0.011/0.011 at Pimai, NE Thailand 0.012/0.011/ 0.011/0.012 at Silpakorn U., C Thailand	
Local urban/industrial aerosol (at 440/675/870/1020 nm)	1.469/1.480/ 1.487/1.480 at Silpakorn U., C Thailand	0.019/0.017/ 0.017/0.017 at Silpakorn U., C Thailand	

1.4.3 Urban Pollution

In urban areas, the source contributions cannot be easily identified due to multiple sources, complex ground reactions facilitated by urban heat island characteristics. The Economic and Social Commission for Asia and the Pacific (ESCAP), the United Nations conducted the studies to understand key sources of air pollution in the major cities in Thailand—Bangkok, Chiang Mai, and Nakhon Srithammarat. Their work incorporated both ground-based measurements (meteorological parameters and air quality index) and satellite retrievals (land cover and hotspots) into machine learning models [60]. The findings showed that biomass burning and forest fires are the primary sources in all cities, with minor contributions from urban activities [60].

1.4.4 Long-Range Transport Air Mass

Both ground-based sensing and satellite remote sensing are advantageous for estimating and tracking the extent of biomass burning smoke plumes. Duc et al. [57] employed the integrated data from AERONET, MODIS, and CALIPSO satellites to assess the extent of biomass burning smoke from the Southeast Asian mainland. They found that the smoke plumes could be transported to southern China, Taiwan, and farther. Smoke haze from the Southeast Asian Region can be evident and contribute to the air pollution problem on a global scale. NASA initiated the framework of biomass-burning Aerosols in Southeast Asia: Smoke Impact Assessment, or BASE-ASIA, in 2006 to evaluate the impact of the aerosols [61]. The BASE-ASIA employs various resources from numerical simulation, the analysis of AERONET aerosol optical properties from multiple sites, and satellite observation. From 2008 to 2013, Seven Southeast Asian Studies, or 7-SEAS, campaign were implemented [59]. Under the 7-SEAS campaign, the NASA scientists employed mobile laboratories and ground-based networks (both AERONET and MPLNET) [59]. From BASE-ASIA to 7-SEAS, scientists observed long-range transport of the smoke aerosols and understood dynamics of the aerosol optical properties along with trajectories, affecting climate warming or cooling.

Since polluted air masses can aloft and transport across the national boundaries. In such a case, the previous models developed using local parameters may not well explain these long-range transport air masses. Scientists and engineers use the advantages of CALIOP air products to observe vertical profiles of aerosol types and later employ backward trajectory models to track potential sources of aerosol clouds for individual cases. Bridhikitti [62] employed aerosol layer scenes taken from CALIOP over the Bangkok Metropolitan Region (BMR) and found that the long-range transport smoke was the most often found at the elevation of approximately 0.7–1.5 km (Fig. 1.2). The CALIOP aerosol profile also suggests long-range transport of polluted dust to the BMR at the elevated level of 3.0–3.2 km during monsoon season (June to September).

Backward trajectory analysis also assists in indicating sources of the long-range transport aerosol. For example, as seen from Fig. 1.3 in the work of [62], the smoke (in the left figure) aloft over the BMR could be transported from the E-to-SE Asian continent, and the polluted dust (in the right-hand figure) could be from S Asia. The HYSPLIT trajectory model from NOAA (https://www.ready.noaa.gov/HYSPLIT.php) is widely used interactively for computing the backward air parcel trajectories [62, 63].

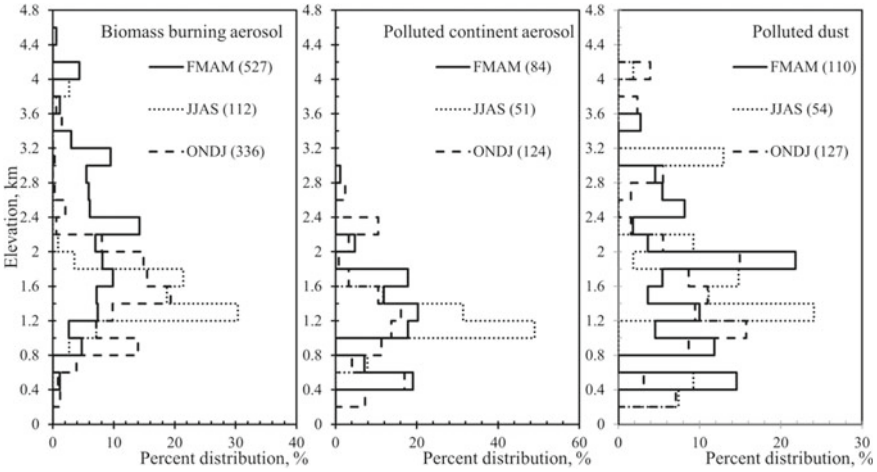


Fig. 1.2 Frequency distribution of aerosol top layer over Bangkok Metropolitan Region from CALIPSO Observations. *Note* FMAM = February to May, JJAS = June to September, ONDJ = October to January [62]

1.5 Satellite Retrievals for Forecasting Haze Episode as the Early Warning Tool

Current approaches for air quality forecasting are applications of chemical transport models, statistical methods, ground and satellite observation networks, and deep learning techniques [64]. The U.S. Environmental Protection Agency's AirNow program employs surface observations of PM2.5 for more than 500 cities to forecast Air Quality Index (AQI). The forecast also combines the satellite observations of aerosol from the Interagency Monitoring of Protected Visual Environments (IMPROVE) network to fill the gaps in the rural areas of the continental U.S. [65] Along with the forecasted AQI, the Chemical Speciation Network (CSN) has been operated to measure the chemical composition of ground-level PM in the urban areas for PM sources identifications [65]. Zhang et al. [65] also proposed the improvement of PM2.5 forecasts in the continental U.S. through the integrating uses of satellite retrievals (including AOD products from TERRA. AQUA, and VIIRS), multi-chemical transport models (including GEOS-Chem, WRF-Chem, and CMAQ), and ground observations. The integration was claimed to provide the best performance for today and next-day forecast [65]. Nonetheless, the study of air pollution episodes in Poland showed highly seasonal dependence and cloud effects on the forecasting performance [66]. From 2005 to 2009, European Commission funded the Global and Regional Earth-System Monitoring Using Satellite and In situ Data, or GEMS project. The project collaborates among researchers from eighteen research institutes in Europe to develop a pilot system for enhancing the forecasting capacity for

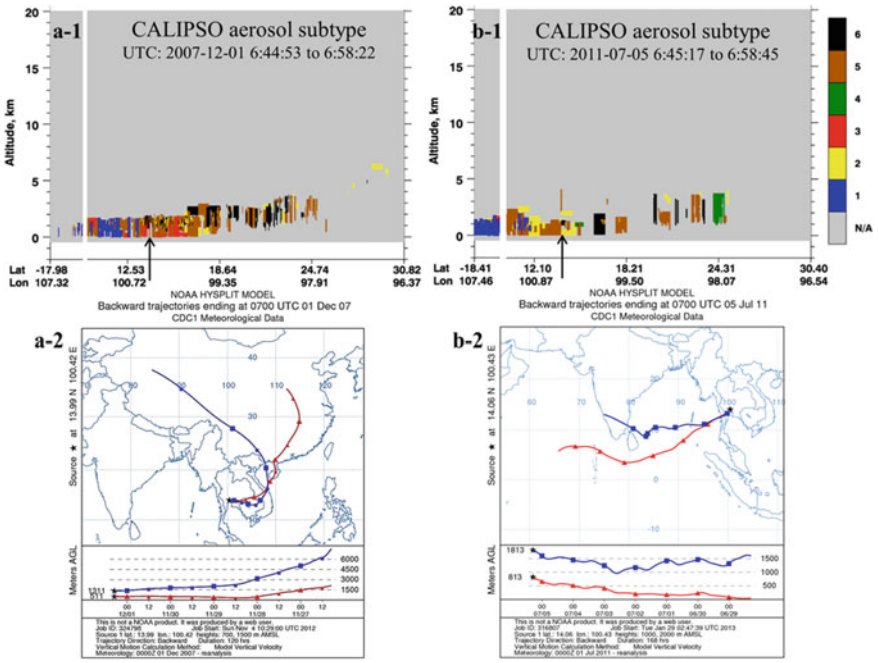


Fig. 1.3 CALIPSO aerosol subtypes (1 = clean marine, 2 = dust, 3 = polluted continental, 4 = clean continental, 5 = polluted dust, 6 = smoke, N/A = not applicable) for the daytime on December 1st, 2007 (**a-1**) and the daytime on July 5th, 2011 (**b-1**) and their corresponding HYSPLIT back trajectories at 13.99°N100.42°E (**a-2**, at 700 m above mean sea level for red line and 1,500 m for blue line) and at 14.06°N 100.43°E (**b-2**, at 1,000 m above mean sea level for red line and 2,000 m for blue line), respectively [62]

aerosols, greenhouse gases, and reactive gases in the atmosphere and improving forecasts for regional air quality [67]. The GEMS production required multi-satellite and remote sensing datasets for developing the best-fitted forecast model. The continuation of GEMS after 2009 was the Monitoring of Atmospheric Composition and Climate (MACC) project. One of the objectives of the MACC project was to provide information services on short-term forecast air quality for Europe [67].

Nowadays, the deep learning model gains more scientific attention for air quality forecast since it can extract complex interconnections using a large pool of ground-measured and satellite-based datasets. Sharma et al. [68] show that the deep learning model can be used to forecast satellite-based PM_{10}. The technique can be beneficial for early warning of hazards associated with air pollution from hotspots in Australia [68]. This method, however, is still in the early stage of development and its lack of interpretability needs to be accomplished to make the technique more useful [64].

In case of Thailand and the Southeast Asian Region, the application of satellite retrievals for air quality forecasting is recommended for remote areas with a sparse ground monitoring network [69]. Li et al. [69] developed a linear regression model

using MODIS AOD and ground-measured meteorological factors to predict PM_{10} levels in the next day. The model provided a percentage error of less than 30% for hindcasts for the cities in China and Thailand [69]. In 2020, SERVIR-Mekong, a partnership between the US Agency for International Development (USAID) and the US National Aeronautics and Space Administration (NASA), developed new web-based tool, which is "Mekong Air Quality Explorer Tool" to improve air quality monitoring and forecasting in Thailand and other countries in the Lower Mekong River Basin. Users can access and visualize 3-day $PM_{2.5}$ forecast derived from NASA GEOS global forecast model [70]. The forecast employs MODIS AOD from TERRA and AQUA satellites and fire products (fire radiative power and number of fires) from the MODIS and VIIRS/Suomi NPP. The forecast simulation for Thailand is bias corrected using machine learning algorithm [70]. The developer aims that the tools can facilitate decision-making on air quality management, especially in the haze season.

1.6 Summary

Satellite remote sensing has been widely used for air pollution monitoring nowadays primarily because it provides spatial advantages over ground-based monitoring. Previous studies in Thailand assessed the magnitude and extents of atmospheric pollutants from remote sensing. The studies had been focused on tropospheric O_3 and its precursors, aerosol loading, and pollutants from biomass burning. The observations, however, are primarily reported as the total atmospheric column, and in some cases, the retrievals may not reach the ground. Therefore, scientists are developing various algorithms using land surface properties and meteorological parameters to assess magnitudes of ground-level air pollutants from satellite retrievals. The temporal resolution of the satellite retrievals is another limitation. The limitation can be accomplished by employing the retrievals from multi-satellite having similar functioning sensors, such as MODIS/TERRA, MODIS/AQUA, and VIIRS/Suomi NPP. The integrated satellite retrievals also yield a better estimate of air pollutant magnitudes with different atmospheric heights. The retrievals from A-Train satellites can be superior since several satellites were orbiting in line, with the near corresponding timeframe. Furthermore, ground-based remote sensing network, such as AERONET for aerosol retrieval, is widely used to understand the dynamics of aerosol properties at particular locations and provide the data used for validating the satellite retrievals.

Using remote sensing technology to identify air pollution sources was successfully implemented in Thailand for biomass burning/wildfires, urban pollution, and long-range transport air masses. Algorithms for the source identification have been relied upon either multivariate analyses of the remote-sensing aerosol optical properties and/or justified from the co-presence of air pollutants as fingerprints for specific sources. The recent development of satellite data analysis for forecasting haze is ongoing. The development requires integrations among long-term satellite observations, surface measurements, and chemical transport models.

References

1. Pinichka C, Makka N, Sukkumnoed D, Chariyalertsak S, Inchai P, Bundhamcharoen K (2017) Burden of disease attributed to ambient air pollution in Thailand: a GIS-based approach. PLoS ONE 12(12):e0189909
2. Jenwitheesuk K, Peansukwech U, Jenwitheesuk K (2020) Accumulated ambient air pollution and colon cancer incidence in Thailand. Sci Rep 10(1):1–12
3. Zhou X, Josey K, Kamareddine L, Caine MC, Liu T, Mickley LJ, Cooper M, Dominici F (2021) Excess of COVID-19 cases and deaths due to fine particulate matter exposure during the 2020 wildfires in the United States. Sci Adv 7(33):eabi8789
4. Phosri A, Sihabut T, Jaikanlaya C (2020) Temporal variations of short-term effects of particulate matter on hospital admissions in the most densely populated city in Thailand. Sci Total Environ 742:140651
5. Moran J, NaSuwan C, Poocharoen OO (2019) The haze problem in northern Thailand and policies to combat it: a review. Environ Sci Policy 97:1–15
6. Tipayarat Y, Sajor EE (2012) State simplification, heterogeneous cause of vegetation fires and implication on local haze management: case study in Thailand. Environ Dev Sustain 14:1047–1064
7. Wongwatcharapaiboon J (2020) Toward future particulate matter situations in Thailand from supporting policy, network and economy. Future Cities Environ 6(1)
8. Shi H, Wang Y, Chen J, Huisingh D (2016) Preventing smog crises in China and globally. J Clean Prod 112:1261–1271
9. Liu Y, Jia R, Dai T, Xie Y, Shi G (2014) A review of aerosol optical properties and radiative effects. J Meteorol Res 28(6):1003–1028
10. Ferrier C, Boschetti M, Fourcade J, Gamet P, Gamet P (2010) Exit the A-Train constellation: PARASOL and CALIPSO different strategies. In: SpaceOps 2010 conference delivering on the dream hosted by NASA marshall space flight center and organized by AIAA, p 2275
11. Murphy R (2014) OCO-2 takes the A-Train to study earth's atmosphere. National Aeronautics and Space Administration. https://www.nasa.gov/jpl/oco2/a-train-atmosphere-20140703. Accessed on 17 October 2021
12. Kawaguchi M, Yoshida T (2013) Regular observation by global change observation mission 1st-water GCOM-W1 (Shizuku). NEC Tech J 8(1):32–35
13. Schoeberl MR, Douglass AR, Hilsenrath E, Bhartia PK, Beer R, Waters JW, Gunson MR, Froidevaux L, Gille JC, Barnett JJ, Levelt PF (2006) Overview of the EOS Aura mission'. IEEE Trans Geosci Remote Sens 44(5):1066–1074
14. Doll CN, Muller JP, Elvidge CD (2000) Nighttime imagery as a tool for global mapping of socioeconomic parameters and greenhouse gas emissions. Ambio 157−162
15. Welton EJ, Voss KJ, Quinn PK, Flatau PJ, Markowicz K, Campbell JR, Spinhirne JD, Gordon HR, Johnson JE (2002) Measurements of aerosol vertical profiles and optical properties during INDOEX 1999 using micropulse lidars. J Geophys Res Atmos 107(D19):INX2–18
16. OMI team (2012) Ozone monitoring instrument (OMI) data user's guide. NASA, Washington DC
17. Jackson JM, Liu H, Laszlo I, Kondragunta S, Remer LA, Huang J, Huang HC (2013) Suomi-NPP VIIRS aerosol algorithms and data products. J Geophys Res Atmos 118(22):12–673
18. Hsu NC, Lee J, Sayer AM, Kim W, Bettenhausen C, Tsay SC (2019) VIIRS Deep Blue aerosol products over land: extending the EOS long-term aerosol data records. J Geophys Res Atmos 124(7):4026–4053
19. Sawyer V, Levy RC, Mattoo S, Cureton G, Shi Y, Remer LA (2020) Continuing the MODIS dark target aerosol time series with VIIRS. Remote Sens 12(2):308
20. Bell ML, McDermott A, Zeger SL, Samet JM, Dominici F (2004) Ozone and short-term mortality in 95 US urban communities, 1987–2000. JAMA 292(19):2372–2378
21. Bell ML, Dominici F (2008) Effect modification by community characteristics on the short-term effects of ozone exposure and mortality in 98 US communities. Am J Epidemiol 167(8):986–997

22. Lelieveld J, Barlas C, Giannadaki D, Pozzer AJACP (2013) Model calculated global, regional and megacity premature mortality due to air pollution. Atmos Chem Phy 13(14):7023–7037
23. Amnuaylojaroen T, Macatangay RC, Khodmanee S (2019) Modeling the effect of VOCs from biomass burning emissions on ozone pollution in upper Southeast Asia. Heliyon 5(10):e02661
24. Sonkaew T, Macatangay R (2015) Determining relationships and mechanisms between tropospheric ozone column concentrations and tropical biomass burning in Thailand and its surrounding regions. Environ Res Lett 10(6):065009
25. Mekaumnuaychai T, Suranowarath K, Kanabkaew T, Lalitaporn P (2020) Observations of atmospheric carbon monoxide and formaldehyde in Thailand using satellites. EnvironmentAsia 13
26. Yimlamaid A, Suranowarath K, Kanabkaew T, Lalitaporn P (2020) Long-term assessment of daily atmospheric nitrogen dioxide in Thailand using satellite observed data. EnvironmentAsia 13
27. Buntoung S, Webb AR (2010) Comparison of erythemal UV irradiances from ozone monitoring instrument (OMI) and ground-based data at four Thai stations. J Geophys Res Atmos 115(D18)
28. Lee J, Choi WJ, Kim DR, Kim SY, Song CK, Hong JS, Hong Y, Lee S (2013) The effect of ozone and aerosols on the surface erythemal UV radiation estimated from OMI measurements. Asia Pac J Atmos Sci 49(3):271–278
29. Mehta M, Singh R, Singh A, Singh N (2016) Recent global aerosol optical depth variations and trends—a comparative study using MODIS and MISR level 3 datasets. Remote Sens Environ 181:137–150
30. Hutchison KD, Smith S, Faruqui S (2004) The use of MODIS data and aerosol products for air quality prediction. Atmos Environ 38(30):5057–5070
31. Wei J, Peng Y, Mahmood R, Sun L, Guo J (2019) Intercomparison in spatial distributions and temporal trends derived from multi-source satellite aerosol products. Atmos Chem Phys 19(10):7183–7207
32. Sukitpaneenit M, Oanh NTK (2014) Satellite monitoring for carbon monoxide and particulate matter during forest fire episodes in Northern Thailand. Environ Monit Assess 186(4):2495–2504
33. Kanabkaew T (2013) Prediction of hourly particulate matter concentrations in Chiangmai, Thailand using MODIS aerosol optical depth and ground-based meteorological data. EnvironmentAsia 6(2)
34. Sayer AM, Hsu NC, Hsiao TC, Pantina P, Kuo F, Ou-Yang CF, Holben BN, Janjai S, Chantara S., Wang SH, Loftus AM (2016) In-situ and remotely-sensed observations of biomass burning aerosols at Doi Ang Khang, Thailand during 7-SEAS/BASELInE 2015. Aerosol Air Qual Res 16(11):2786−2801
35. Nguyen TT, Pham HV, Lasko K, Bui MT, Laffly D, Jourdan A, Bui HQ (2019) Spatiotemporal analysis of ground and satellite-based aerosol for air quality assessment in the Southeast Asia region. Environ Pollut 255:113106
36. Cheewaphongphan P, Garivait S (2013) Bottom up approach to estimate air pollution of rice residue open burning in Thailand. Asia Pac J Atmos Sci 49(2):139–149
37. Vongruang P, Pimonsree S (2020) Biomass burning sources and their contributions to PM10 concentrations over countries in mainland Southeast Asia during a smog episode. Atmos Environ 228:117414
38. Lalitaporn P, Boonmee T (2019) Analysis of tropospheric nitrogen dioxide using satellite and ground based data over Northern Thailand. Eng J 23(6):19–35
39. Takemura T, Nakajima T, Dubovik O, Holben BN, Kinne S (2002) Single-scattering albedo and radiative forcing of various aerosol species with a global three-dimensional model. J Clim 15(4):333–352
40. Wang J, Virkkula A, Gao Y, Lee S, Shen Y, Chi X, Nie W, Liu Q, Xu Z, Huang X, Wang T (2017) Observations of aerosol optical properties at a coastal site in Hong Kong, South China. Atmos Chem Phys 17(4):2653–2671
41. Levoni C, Cervino M, Guzzi R, Torricella F (1997) Atmospheric aerosol optical properties: a database of radiative characteristics for different components and classes. Appl Opt 36(30):8031–8041

42. Bridhikitti A, Overcamp TJ (2011) Optical characteristics of southeast Asia's regional aerosols and their sources. J Air Waste Manag Assoc 61(7):747–754
43. Park JU, Kim SW, Sheridan PJ, Williams A, Chambers SD (2020) Long-term variability of aerosol optical properties at Mauna Loa. Aerosol Air Qual Res 20(7):1700–1711
44. Reidmiller DR, Hobbs PV, Kahn R (2006) Aerosol optical properties and particle size distributions on the east coast of the United States derived from airborne in situ and remote sensing measurements. J Atmos Sci 63(3):785–814
45. Bhaskar VV, Safai PD, Raju MP (2015) Long term characterization of aerosol optical properties: implications for radiative forcing over the desert region of Jodhpur, India. Atmos Environ 114:66–74
46. Zhang M, Ma Y, Gong W, Liu B, Shi Y, Chen Z (2018) Aerosol optical properties and radiative effects: assessment of urban aerosols in central China using 10-year observations. Atmos Environ 182:275–285
47. Kaskaoutis DG, Sinha PR, Vinoj V, Kosmopoulos PG, Tripathi SN, Misra A, Sharma M, Singh RP (2013) Aerosol properties and radiative forcing over Kanpur during severe aerosol loading conditions. Atmos Environ 79:7–19
48. Tiwari S, Pandithurai G, Attri SD, Srivastava AK, Soni VK, Bisht DS, Kumar VA, Srivastava MK (2015) Aerosol optical properties and their relationship with meteorological parameters during wintertime in Delhi, India. Atmos Res 153:465–479
49. Luoma K, Virkkula A, Aalto P, Petäjä T, Kulmala M (2019) Over a 10-year record of aerosol optical properties at SMEAR II. Atmos Chem Phys 19(17):11363–11382
50. Wang T, Du Z, Tan T, Xu N, Hu M, Hu J, Guo S (2019) Measurement of aerosol optical properties and their potential source origin in urban Beijing from 2013–2017. Atmos Environ 206:293–302
51. Chand D, Guyon P, Artaxo P, Schmid O, Frank GP, Rizzo LV, Mayol-Bracero OL, Gatti LV, Andreae MO (2006) Optical and physical properties of aerosols in the boundary layer and free troposphere over the Amazon Basin during the biomass burning season. Atmos Chem Phys 6(10):2911–2925
52. Li C, Tsay SC, Hsu NC, Kim JY, Howell SG, Huebert BJ, Ji Q, Jeong MJ, Wang SH, Hansell RA, Bell SW (2013) Characteristics and composition of atmospheric aerosols in Phimai, central Thailand during BASE-ASIA. Atmos Environ 78:60–71
53. Gautam R, Hsu NC, Eck TF, Holben BN, Janjai S, Jantarach T, Tsay SC, Lau WK (2013) Characterization of aerosols over the Indochina peninsula from satellite-surface observations during biomass burning pre-monsoon season. Atmos Environ 78:51–59
54. Sayer AM, Hsu NC, Eck TF, Smirnov A, Holben BN (2014) AERONET-based models of smoke-dominated aerosol near source regions and transported over oceans, and implications for satellite retrievals of aerosol optical depth. Atmos Chem Phys 14(20):11493–11523
55. Ebert M, Weinbruch S, Rausch A, Gorzawski G, Helas G, Hoffmann P, Wex H (2002) Complex refractive index of aerosols during LACE 98 as derived from the analysis of individual particles. J Geophys Res Atmos 107(D21):8121. https://doi.org/10.1029/2000JD000195
56. Schkolnik G, Chand D, Hoffer A, Andreae MO, Erlick C, Swietlicki E, Rudich Y (2007) Constraining the density and complex refractive index of elemental and organic carbon in biomass burning aerosol using optical and chemical measurements. Atmos Environ 41(5):1107–1118
57. Duc HN, Bang HQ, Quan NH, Quang NX (2021) Impact of biomass burnings in Southeast Asia on air quality and pollutant transport during the end of the 2019 dry season. Environ Monit Assess 193(9):1–28
58. Kamthonkiat D, Thanyapraneedkul J, Nuengjumnong N, Ninsawat S, Unapumnuk K, Vu TT (2021) Identifying priority air pollution management areas during the burning season in Nan Province, Northern Thailand. Environ Dev Sustain 23(4):5865–5884
59. Tsay SC, Hsu NC, Lau WKM, Li C, Gabriel PM, Ji Q, Holben BN, Welton EJ, Nguyen AX, Janjai S, Lin NH (2013) From BASE-ASIA toward 7-SEAS: a satellite-surface perspective of boreal spring biomass-burning aerosols and clouds in Southeast Asia. Atmos Environ 78:20–34

60. UN Economic and Social Commission for Asia and the Pacific, UN ESCAP (2021) Deciphering the black box of air pollution data in Thailand. https://www.unescap.org/blog/deciphering-black-box-air-pollution-data-thailand. Accessed on 26 January 2021
61. Huang K, Fu JS, Hsu NC, Gao Y, Dong X, Tsay SC, Lam YF (2013) Impact assessment of biomass burning on air quality in Southeast and East Asia during BASE-ASIA. Atmos Environ 78:291–302
62. Bridhikitti A (2013) Atmospheric aerosol layers over Bangkok Metropolitan Region from CALIPSO observations. Atmos Res 127:1–7
63. Kanniah KD, Lim HQ, Kaskaoutis DG, Cracknell AP (2014) Investigating aerosol properties in Peninsular Malaysia via the synergy of satellite remote sensing and ground-based measurements. Atmos Res 138:223–239
64. Liao Q, Zhu M, Wu L, Pan X, Tang X, Wang Z (2020) Deep learning for air quality forecasts: a review. Curr Pollut Rep 6(4):399–409
65. Zhang H, Wang J, García LC, Ge C, Plessel T, Szykman J, Murphy B, Spero TL (2020) Improving surface PM2.5 forecasts in the United States using an ensemble of chemical transport model outputs: 1. bias correction with surface observations in nonrural areas. J Geophys Res Atmos 125(14):e2019JD032293
66. Werner M, Kryza M, Guzikowski J (2019) Can data assimilation of surface PM2.5 and Satellite AOD improve WRF-Chem forecasting? A case study for two scenarios of particulate air pollution episodes in Poland. Remote Sens 11(20):2364
67. Hollingsworth A, Engelen RJ, Textor C, Benedetti A, Boucher O, Chevallier F, Dethof A, Elbern H, Eskes H, Flemming J, Granier C (2008) Toward a monitoring and forecasting system for atmospheric composition: the GEMS project. Bull Am Meteorol Soc 89(8):1147–1164
68. Sharma E, Deo RC, Soar J, Prasad R, Parisi AV, Raj N (2022) Novel hybrid deep learning model for satellite based PM10 forecasting in the most polluted Australian hotspots. Atmos Environ 279:119111
69. Li C, Hsu NC, Tsay SC (2011) A study on the potential applications of satellite data in air quality monitoring and forecasting. Atmos Environ 45(22):3663–3675
70. SERVIR-Mekong (2021) Mekong air quality explorer. https://aqatmekong-servir.adpc.net/en/home/. Accessed on 22 August 2021

Dr. Arika Bridhikitti is an assistant professor in the Environmental Engineering and Disaster Management Program at Mahidol University Kanchanaburi Campus, Thailand. Her research includes remote sensing applications for monitoring air pollution and land use changes. She also worked on the hydrogeological processes of small to large-scale watersheds via field monitoring and modeling. Her recent research has focused on environmental monitoring in tropical forests and agroecosystems.

Chapter 2
A Novel Evaluation of Air Pollution Impact from Stationary Emission Sources to Ambient Air Quality via Time-Series Granger Causality

Chun-Hsiang Chan, Jehn-Yih Juang, Tzu-How Chu, Ching-Hao Mao, and Shin-Ying Huang

Abstract Many heavy industrial cities in the world are suffered from serious air pollution problems from stationary emission sources and the spatial patterns between the sources and the receptors is an important environmental issue. Some existed studies adopted numerical models, machine learning or deep learning to characterize the spatial patterns and impacts of air pollution sources in urban areas. Due to the complexity of the air circulation system and consideration of several factors, the relationship between stationary emission sources and ambient air quality is hard to estimate; as a result, limited studies discussed and gave quantitative evidence. This study aimed to quantify and verify the relative impacts from stationary emission sources to each ambient air quality station via applying time-series Granger causality. The study is conducted in Kaohsiung, the largest industrial metropolitan area in Taiwan. The results from the analysis on the role of transboundary pollutants show that the estimated relative impact does not significantly increase during transboundary-dominating seasons winter and spring in the study area. We found that the spatial characteristics of the estimated relative impacts in seasonal and diurnal variation are strongly related to the geographical factors and wind field, respectively. The major stationary emission source is attributed to the category "Smelting and Refining of Iron and Steel". Moreover, the emission amount of different industrial categories is highly consistent with the estimated relative impacts. This method could efficiently reveal the spatial relationship between stationary emission sources and ambient air quality with limited data; hence, the results could provide as suggestions to local residents,

C.-H. Chan (✉)
Undergraduate Program in Intelligent Computing and Big Data, Chung Yuan Christian University, Taoyuan, Taiwan
e-mail: d04228002@ntu.edu.tw

C.-H. Chan · J.-Y. Juang · T.-H. Chu
Department of Geography, National Taiwan University, Taipei, Taiwan

C.-H. Chan · C.-H. Mao
Artificial Intelligence Analytics, Taiwan Cybersecurity Foundry Company, Taipei, Taiwan

S.-Y. Huang
National Center for Cyber Security Technology, Institute for Information Industry, Taipei, Taiwan

T.-H. Wen et al. (eds.), *Earth Data Analytics for Planetary Health*, Atmosphere, Earth, Ocean & Space, https://doi.org/10.1007/978-981-19-8765-6_2

administrations of government and non-government organizations for policy planning. Nevertheless, this concept could be utilized even in low-infrastructure cities, regions or countries for monitoring or realizing how stationary emission sources affect to ambient air quality.

Keywords Relative impact · Source-receptor relationship · Urban air quality · Causal analysis · Spatial–temporal pattern

2.1 Introduction

The air pollution of stationary emission sources is one of the most serious environmental issues in many industrial cities and their vicinity area [1, 2]. Especially in the developing and developed countries, these stationary emission sources, including different types of industries for energy supply and for manufacturing process of steel, petroleum, chemical, paper, or many other industrial materials [3], usually produce numerous gaseous or particulate air pollutants, and many kinds of toxic chemicals, to affect the ambient air quality. In recently years, many epidemiological studies have reported that these compounds could induce premature mortality and reduce the life expectancy [4–7].

In consequence, the spatial linkage between the emission sources and the ambient air quality is becoming an important research concern in the disciplines of environmental science and urban planning [8, 9]. To characterize the impact of stationary industrial sources on the ambient air quality over different spatial scales, many previous studies analyze the source-receptor relationship through different methodologies by considering different controlling factors [10–13], which include ambient concentration of air pollutants, topographic properties, built-environment factors, and meteorological conditions [14–17]. Most of these studies adopted numerical model, which integrates the emission inventory, meteorological variables and chemical mechanism, to simulate ambient air quality [18, 19]. Meanwhile, some studies attempted to use spatial analysis and stochastic model to evaluate the impact of air pollution [14, 20]. However, these methods usually require complicate setting and parameterizations, and need many computational resources for data processing and calculation [21, 22].

To resolve these problems, some studies applied data-driven approaches to bridge the connection between the observation and the prediction of the ambient air quality by using the newly-developing machine learning [23–25] and deep learning techniques [26–28]. Although these approaches could solve issues of low data resolution and low data quantity, they still need powerful computation equipment for data training and model fitting [29–32].

Granger causality was proposed by Granger (1969), and was used to know the causal relationship between dependent and independent variables by identifying statistical significance of the corresponding coefficients. It could test a series of lagging values to identify which values have causal relationship with the dependent

variables via joint hypotheses test (F-test) [33, 34]. This method has been widely applied in medical, economic or tourism studies to realize whether the certain factors affect the specific variables [35–37]. In environmental issues, some studies quantify the causal relationship between air pollutants and the health expenditures, diseases and even mortality [38–43]. Similar concept was adopted in revealing the relationship between emission sources and controlling factors (e.g. meteorology, terrain, land use and wind field) to ambient air quality [23, 27, 44–46]. Jiang and Bai [46] utilized time-series Granger causality and Pearson correlation to examine the characteristics of air quality in Beijing and its surrounding cities. They explained the interactions of air pollution among these cities and concluded that using Granger causality was resources-efficient for air quality analysis in Beijing area. Wang [23] characterize the relationship between air pollution and environmental factors by using Granger causality, and they mentioned that self-aggregation and self-diffusion process were the major impacts on ambient air quality even under different meteorological conditions.

As a newly developed country, Taiwan has been suffered the problem of air pollution from domestic emission (stationary and mobile sources) and the monsoon-driven transboundary pollutants in the past couple decades [47–50]. Among the different types of emission sources, the stationary sources play significant role to influence the ambient air quality in many cities in Taiwan [51], especially in the southwest coastal plain [52].

The objective of this study is to reveal the impact of stationary emission sources on ambient air quality in spatial and temporal distribution via time-series Granger causality and impact evaluation. Based on the consideration of the emission source characteristics and low chemical conditions, we selected sulfuric dioxide (SO_2) as the air pollution indicator for analysis. Compared to the previous studies [23], this research proposes to utilize Granger causality analysis to quantify the seasonal and diurnal variation of relative casual impact on ambient air quality. In addition, this study aims to identify the contributions of different sources. There are three tasks in this study: the first is to identify the influence of transboundary pollutants on the causality via the seasonal comparison, the second is to characterize spatial–temporal features of the causality over different periods, and the third is to quantify the relative contributions of different industrial categories on ambient air quality in the study area.

Therefore, a data-driven method conducting time-series Granger causality is adopted to measure the impact of SO_2 of each stationary emission source on ambient air quality in a heavy-industrial city, Kaohsiung [52, 53] in southern Taiwan. The spatial–temporal data of stationary sources and ambient air quality observation in Kaohsiung are used in the causality analysis, and the results from this analysis could help us to quickly validate the spatial–temporal characteristics of stationary sources and their relative impact on the ambient air quality.

2.2 Characteristics of the Study Area

Kaohsiung is the biggest metropolis and the most important heavy industrial city in southern Taiwan. The total area of Kaohsiung is 2,952 km^2, and the population is approximately 2.77 million in 2017. According to Taiwan's Ministry of Economic Affairs, there are 7,045 registered factories in this city in 2017. Therefore, the air quality in Kaohsiung has become one of the most important environmental issues in the past decades. To monitor the stationary sources in this major industrial city, Taiwan's Environmental Protection Administration (EPA) established the country's largest number of the continuous emission monitoring systems (CEMS) stations at each registered stationary source in Kaohsiung (105 stations in total, which is approximately 32.2% of total CEMS in Taiwan). The proportions of different industrial categories of the registered CEMS stations in Kaohsiung in 2017 is shown in Fig. 2.1. From the distribution, we found that the biggest one is the category "Smelting and Refining of Iron and Steel" (30.5%), and is followed by "Electricity Supply" (18.1%), "Petrochemical Manufacturing" (13.3%), and "Treatment and Disposal of Waste" (12.4%). The top 4 categories of CEMS contribute roughly about three fourths of the total registered CEMS in Kaohsiung. In addition, there are 899,024 vehicles and 1,999,902 motorcycles in Kaohsiung in 2017. To monitor the ambient air quality, Taiwan's EPA set 12 ambient air quality monitoring system (AQMS) stations in different locations in this metropolis. The spatial distribution of the CEMS and AQMS stations in Kaohsiung is shown in Figs. 2.2 and 2.3.

In this study, we use the data from CEMS and AQMS stations in Kaohsiung to conduct the analysis of causal relationship. Because the spatial causality analysis sensitively relies on source-receptor distance, it is important to estimate the contributions of the domestic sources and transboundary sources. To verify the performance of this model and distinguish the contribution from transboundary pollutions, we used dataset over different seasons to characterize the influence of transboundary

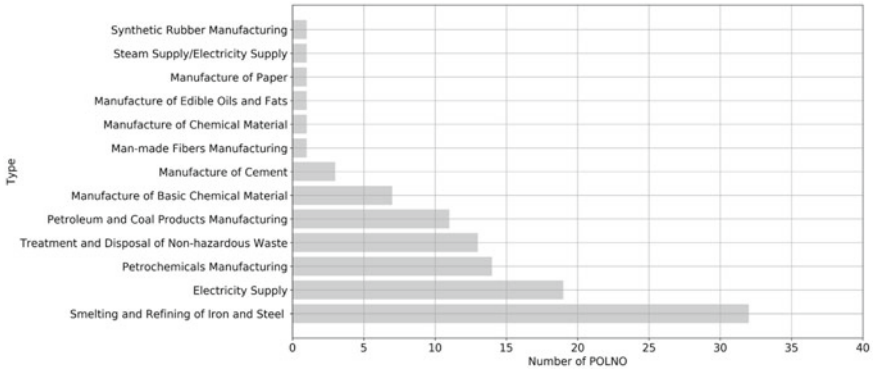


Fig. 2.1 Distribution of registered continuous emission monitoring systems (CEMS) stations in Kaohsiung, 2017

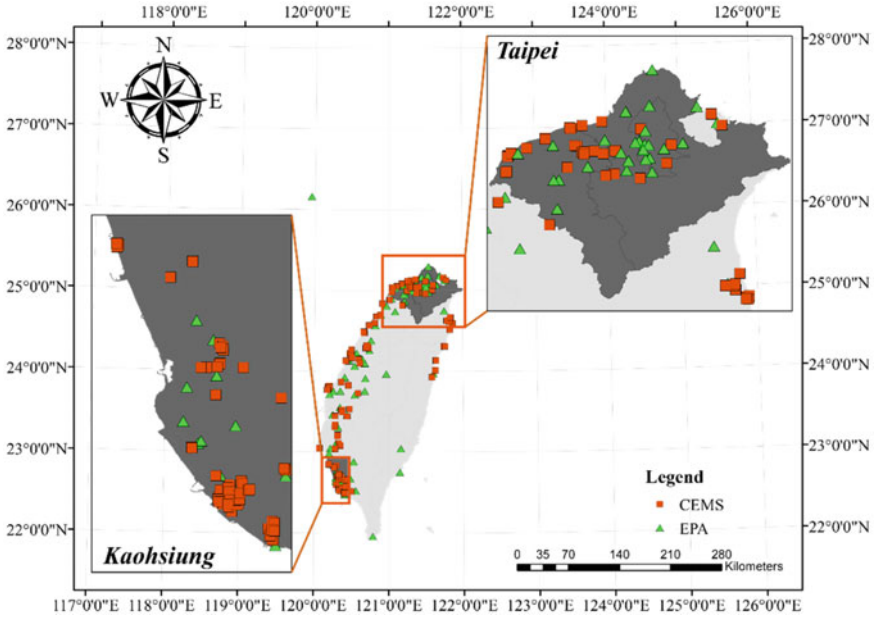


Fig. 2.2 Study Area (Our study selected two regions in Taiwan; the upper right submap is the Taipei metropolitan area, validation region for the influence from transboundary pollutants, including Taipei City, New Taipei City and Taoyuan City, suffered from the domestic pollutants and transboundary pollutants as well; the down left submap is the Kaohsiung City, major region for analyzing the causal relationship between stationary sources and surrounding ambient air quality, existing the largest heavy industrial factories in Taiwan)

pollutants. In addition, we also used the dataset from the Taipei metropolitan area as a test case for model validation.

In contrast to the patterns of air pollution in Kaohsiung, the Taipei metropolitan area (total population: 8,857,963; area: 3,545 km^2; 2,596,830 vehicles and 4,295,661 motorcycles in 2017) has relative fewer stationary emission sources. There are 25 CEMS stations (~7.7% of the entire country) in the Taipei metropolitan area (Fig. 2.2). This is because the major economic activities in Taipei is international business and service industry. Although the total number of heavy industrial factories in Taipei is much less compared to Kaohsiung, the major air quality issues in Taipei is mainly attributed to domestic mobile sources, and the transboundary pollutants due to the seasonal-prevailing monsoon during the winter and spring seasons. To monitor the ambient air quality, there are 19 EPA's AQMS stations spread in this area.

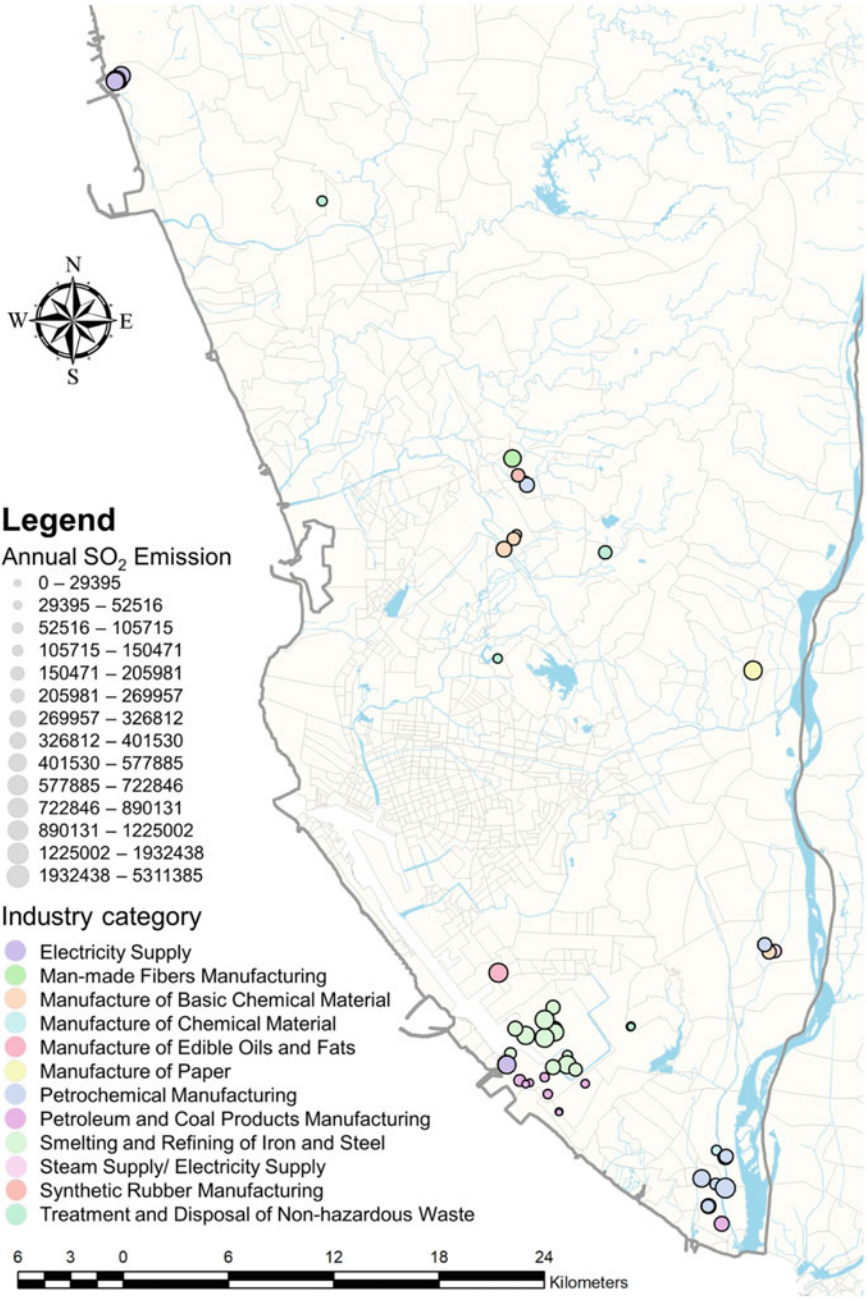


Fig. 2.3 Annual SO_2 emission amount of each CEMS stations in Kaohsiung. Different colors indicate industry categories and maker size demonstrated annual SO_2 emission amount of the specific CEMS stations

2.3 Data and Methodology

2.3.1 Data Collection

In the past couple decades, SO_2 has become one of the most major air quality issue in many cities in Taiwan and in Asia [54–56]. Therefore, in this study we focus on the spatial–temporal patterns of SO_2 by quantifying the Granger causality between the emission sources and the ambient air quality in the study area. In the data analysis, we used SO_2 observation from EPA's AQMS stations and the emission inventory from CEMS for the year of 2017. In addition, in order to consider the effect of the terrain and meteorological conditions, this study utilized the digital elevation model (DEM) data from Taiwan's Ministry of the Interior, and the observation of wind field, including wind direction and wind speed, from Taiwan's Central Weather Bureau (CWB) for analysis.

2.3.2 Data Processing

The time interval of EPA observation data is one hour, and sampling frequency of SO_2 data in CEMS is 15 min. In the stage of data analysis, we averaged every 15-min data into hourly interval to ensure the consistence of both datasets.

In this study, both CEMS data and AQMS data were extracted during the daytime from 08:00 to 17:00 LT due to the significant difference in meteorological properties (wind field and mixed-layer height) between daytime and nighttime in the study area. As for the wind patterns, the wind field during the daytime in Kaohsiung is dominated by sea-breeze system while the wind field during the nighttime is relatively calm. Furthermore, the growth of mixed-layer height after the sunrise brings the pollutants from the aloft atmosphere to the ground surface and causes the changes in near-surface atmospheric stability [57, 58]. Furthermore, every day in the morning each CEMS is regulated by EPA to examine the emission system to ensure the detectors of CEMS are stable. Therefore, to avoid the discontinue emission data in the analysis, we conducted the Granger causality analysis from 08:00 to 17:00 LT during the daytime.

To conduct the source-receptor relationship in this study, we collected data from all AQMS stations which locate within the circular region of 10-km radius of CEMS stations for Granger causality estimation. The impact of mobile sources is a daily and regular periodic influence, so the mobile source was regarded as a background signal or baseline of ambient air quality. In addition, the influence of transboundary pollutants in Taiwan usually happens in winter and spring [47, 48]. Hence, we set a series of analysis to identify the influence of transboundary pollutants via seasonal comparison (summer-autumn versus winter-spring) and spatial comparison (Kaohsiung versus Taipei).

2.3.3 Time-Series Granger Causality

Granger causality is a vector autoregression modeling method to evaluate whether a n consecutive time series of variables $x_{ts} = \{x_t, x_{t-1}, ..., x_{t-n}\}$ can affect a variable y_t at a specific time t [59]:

$$y_t = \sum_{a=1}^{n} \beta_a x_{t-a} + \varepsilon_t,$$

where β_a is a set of corresponding fitting coefficients and ε_t is the error term of the vector autoregression model.

The significance of each coefficient β_a of vector autoregression model could represent whether Granger causal relationship exists between y_t and $\{x_{ts}\}$. The null hypothesis of Granger causality is that $\{x_{ts}\}$ does not affect the y_t, when the p-value of coefficient in F-statistic is larger than the significance level ($\alpha = 0.05$ under 95% confidence level); otherwise, the alternative hypothesis is accepted, as well as, the Granger causality exists.

In the analysis, we utilized a python package, "*grangercausalitytests*" in "*statsmodels*", to compute the Granger causality between AQMS observation and CEMS data. For a given concentration of SO_2 of a specific AQMS station $E_t^{(i,j)}$ at timestamp t, and a set of lagging emission values obtained from the latest consecutive six hours of emission inventory of surrounded CEMS stations, $C_{ts}^{(j)} = \{C_t^{(j)}, C_{t-1}^{(j)}, \cdots, C_{t-5}^{(j)}\}$, within the area of 10-km radius. Therefore, the model for the j-th CEMS station to the i-th AQMS station could be depicted as

$$E_t^{(i,j)} = \sum_{a=0}^{5} \beta_a^{(j)} C_{t-a}^{(j)} + \varepsilon_t^{(i,j)},$$

where $\beta_a^{(j)}$ is the coefficient and $\varepsilon_t^{(i,j)}$ is the error term of the vector autoregression model.

2.3.4 The Evaluation of CEMS Impact

Through the time-series Granger causality, we could obtain a set of Granger causal relationships $GC_{C_{t'}^{(j')} \leftarrow E_{t'}^{(i',j')}}$ between the AQMS station i' and a specific CEMS j' at specific time t' as

$$GC_{C_{t'}^{(j')} \leftarrow E_{t'}^{(i',j')}} = \left\{ C_{t'}^{(j')} \right\},$$

where for all p-values of $\beta_a^{(j')}$ are less than 0.05 under 95% confidence level and all subscript primes of i, j and t are the set of AQMS, CEMS and timestamps which represent there exist Granger causal relationship. Therefore, the impact of one CEMS $P_{C^{(j')}}$ to ambient air quality could be given as the total count of the statistically significant Granger causal relationships expressed as follows,

$$P_{C^{(j')}} = \sum_{t'=1}^{T'} \sum_{i'=1}^{I'} \text{Count}\left(\{ GC_{C_{t'}^{(j')} \to E_{t'}^{(i',j')}} \} \right),$$

where $GC_{C_{t'}^{(j')} \to E_{t'}^{(i',j')}}$ is the set of Granger causal relationship between one CEMS $C_{t'}^{(j')}$ to all surrounded EPA $E_{t'}^{(i',j')}$; I' and T' are the number of AQMS stations and timestamps, corresponding Granger causal relationship between CEMS station j' and AQMS station i' at timestamp t'.

To calculate the overall impact of each industry category, the impact of a specific industry category d in the specific year is aggregated by the impact of all CEMS $P_{C^{(j")}}$. Finally, the overall impact of category d in that year could be defined as

$$Y_d = \sum_{j"=1}^{J"} P_{C^{(j")}}, \forall P_{C^{(j")}} \in \left\{ P_{C^{(j')}} \right\},$$

where $j"$ is the subset of CEMS j' and the $j"$ represents the number of CEMS station, which exists Granger causal relationship.

2.4 Results and Discussions

2.4.1 Seasonal Comparison

Because the compositions of emission sources are different in the study area (Kaohsiung) and test area (Taipei), in this section, we focus on the difference in source impact, and number of GC events in these two areas over different seasons, which are shown in Fig. 2.4. The results indicate that the numbers of GC events in Kaohsiung during winter-spring seasons are about 15% more than the numbers in summer-autumn seasons, but the seasonal variations in Taipei is not significant.

In Taiwan, it is well known that during the winter and spring seasons, the significant portion of air pollution issues could be attributed to transboundary pollution brought by the seasonal monsoon. Meanwhile, during the summer and autumn seasons, the contribution from the transboundary pollutant is pretty minor [47, 60–62]. To quantify whether the transboundary pollution affect the GC events, we conducted ANOVA test and Post-Hoc test on results of the GC analysis. The results show that there is no statistical significance between these two season periods

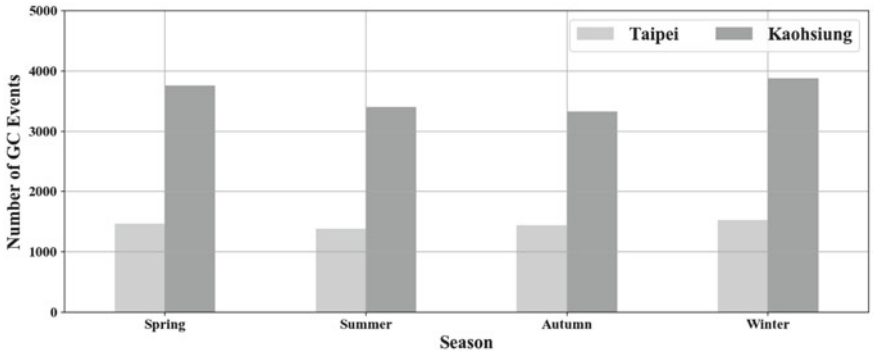


Fig. 2.4 The number of GC events of each season in 2017

(winter-spring and summer-autumn) in both areas in ANOVA (Levene = 0.037, *p*-value = 0.990 > 0.05), and season-by-season comparison in Post-Hoc test also show no significant difference between two season periods (Levene = 0.093, *p*-value = 0.964 > 0.05). This result confirms that the GC events were not affected by trans-boundary air pollutants in both Taipei and Kaohsiung. These results also proved that all changes in large scale, daily variation or background pollutants (mobile sources) do not significantly influence the impact estimation.

Moreover, the spatial distribution of GC events over different seasons has spatial variations (Fig. 2.5). The figure shows that high GC event area is concentrated in inland area for all seasons. However, the high GC event area extended from inland to costal area in summer and autumn. Overall, spring and winter have relatively more GC events than the other two seasons.

As for the contribution of source categories on the GC events in the study area, the category "Smelting and Refining of Iron and Steel" is the largest for the entire year, and "Waste Treatment and Disposal" is the second largest (Fig. 2.6). The top 5 industrial categories have relatively high GC events in spring and winter seasons, while the GC events is relatively lower in summer. It is because during the winter-spring seasons in Kaohsiung, the mixed-layer height is lower for pollutants to accumulate, and the meteorological condition during the summer season is better for diffusion.

2.4.2 Diurnal Variation of Spatial Characteristics of GC Event

In the study, we considered only the daytime data from 08:00 to 17:00 LT. Therefore, in the outcome of the GC analysis for each day, there are five consecutive periods with 6 one-hour timestamps in each period (e.g. 08:00 ~ 13:00, 09:00 ~ 14:00, …, and 12:00 ~ 17:00). The proportions of GC events of the five periods are shown in Table 2.1. From the results, we could find that the proportion gradually increases from

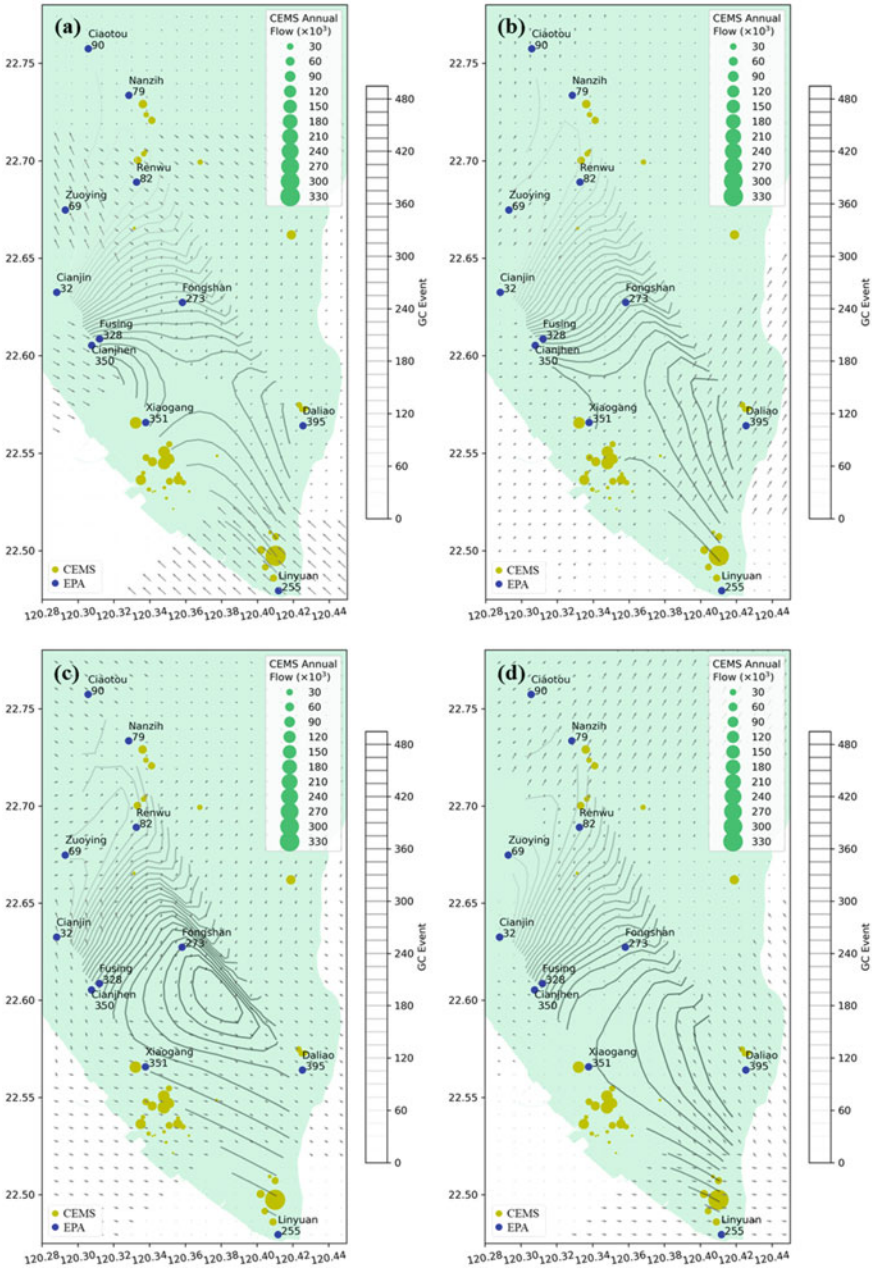


Fig. 2.5 The spatial distribution of GC event in Kaohsiung during different seasons (**a** Spring; **b** Summer; **c** Autumn; **d** Winter)

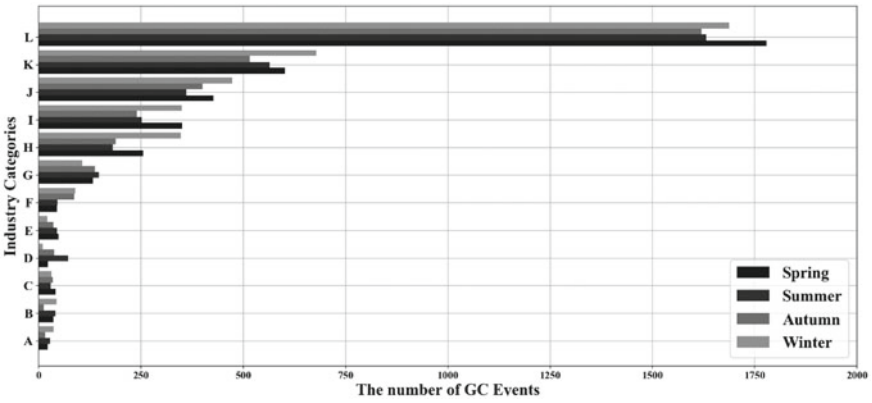


Fig. 2.6 The number of GC events of each season by industry in Kaohsiung. (A: Manufacture of Paper; B: Synthetic Rubber Manufacturing; C: Steam Supply/Electricity Supply; D: Electricity Supply; E: Manufacture of Chemical Material; F: Man-made Fibers Manufacturing; G: Manufacture of Edible Oils and Fats; H: Petroleum and Coal Products Manufacturing; I: Manufacture of Basic Chemical Material; J: Petrochemicals Manufacturing; K: Treatment and Disposal of Non-hazardous Waste; L: Smelting and Refining of Iron and Steel)

16.04% to 21.81% with time, and the pattern is highly correlated to the diurnal wind field shifting in Kaohsiung. Every day in the morning, the wind filed in Kaohsiung starts to change direction from the nocturnal land-breeze system to the sea-breeze system in this area. After around 10:00 LT, the sea-breeze dominate the wind filed in the vicinity. Therefore, the GC events change gradually in the morning and is becoming stable after the late morning (Fig. 2.7). From Fig. 2.7, the spatial distribution of high EPA GC areas during the first period (08:00–13:00) locate in the coastal area (between Siaogang and Cianjhen). Later, the high EPA GC area moves to the inland area (between Fongshan and Daliao), and this phenomenon prolongs to the fourth period (11:00–16:00) and reaches the highest EPA GC events. Moreover, the fifth period (12:00–17:00) shows that the high EPA GC events area started to move back to coastal area due to the decrease of sea breeze.

Table 2.1 The GC event distribution of Kaohsiung in five consecutive periods

Consecutive period	Timestamp period	GC event proportion (%)
1	08:00–13:00	16.04
2	09:00–14:00	18.73
3	10:00–15:00	21.82
4	11:00–16:00	21.58
5	12:00–17:00	21.82

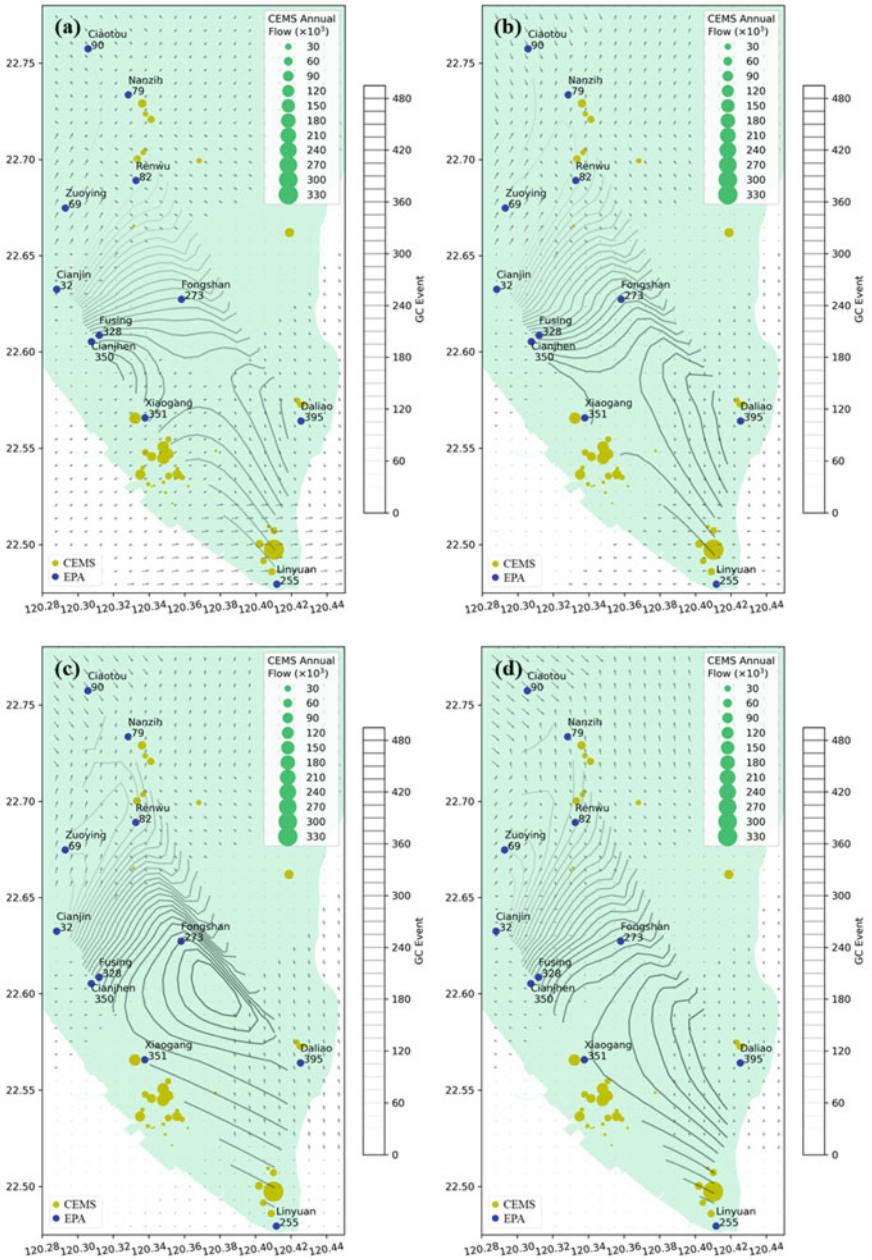
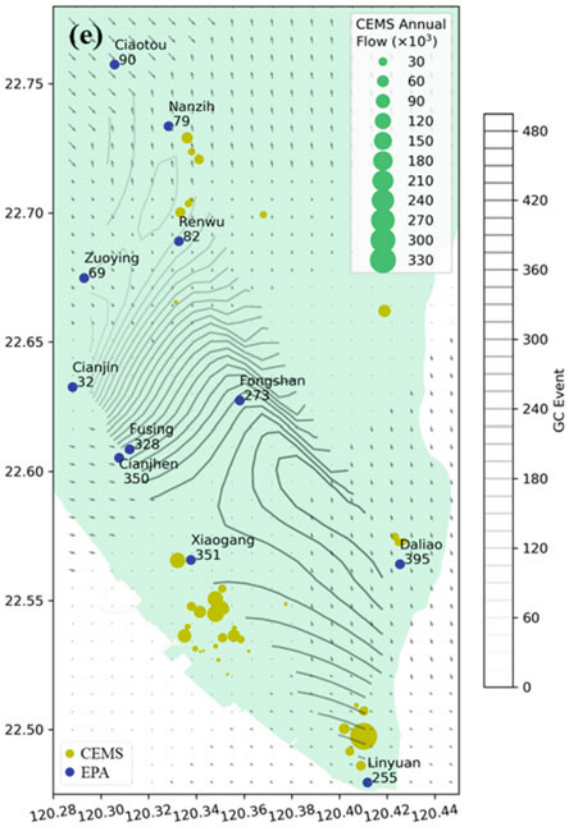


Fig. 2.7 The spatial distribution of GC event in different consecutive periods (**a** The first consecutive period (8:00–13:00); **b** The second consecutive period (9:00–14:00); **c** The third consecutive period (10:00–15:00); **d** The fourth consecutive period (11:00–16:00); **e** The fifth consecutive period (12:00–17:00))

Fig. 2.7 (continued)



2.4.3 Industrial Comparison

One advantage of applying GC estimation in the spatial analysis is to identify the relative contribution of the large number of emission sources. In this section, the relative local emission impact and statistic of the GC events for each industry category are shown in Fig. 2.8. The total number of GC events for the entire year is 14,365, and the biggest one is the category "Smelting and Refining of Iron and Steel", which accounts for roughly about 46.77% (6,718 in total) of the total GC events. The biggest four industry categories account for the vast majority, about 85.6%, of the total GC events in Kaohsiung. The results also show that steel-related, petroleum-related or chemistry-related categories make major contributions on GC events.

In this analysis, the observed power is 1.000 and the partial eta squared is 0.993, so the performance of this GC estimation has good power of test. In addition, on the Levene test for equality of error variances (Levene = 3.812, p-value = 0.001), it rejects the null hypothesis which states that error variance of the dependent variable is not equal across different industry categories. According to descriptive statistics and

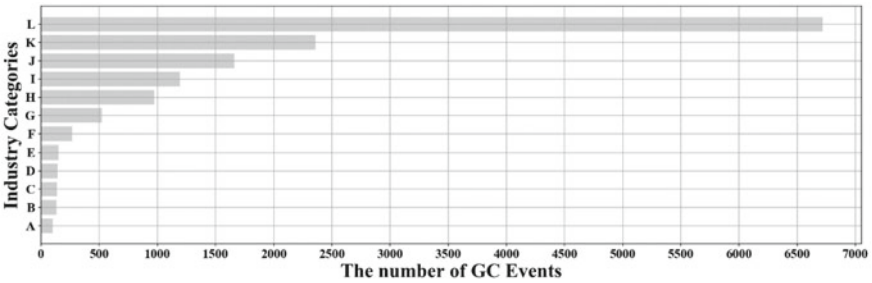


Fig. 2.8 The number of GC events of each industry in Kaohsiung (A: Manufacture of Paper; B: Synthetic Rubber Manufacturing; C: Steam Supply/Electricity Supply; D: Electricity Supply; E: Manufacture of Chemical Material; F: Man-made Fibers Manufacturing; G: Manufacture of Edible Oils and Fats; H: Petroleum and Coal Products Manufacturing; I: Manufacture of Basic Chemical Material; J: Petrochemicals Manufacturing; K: Treatment and Disposal of Non-hazardous Waste; L: Smelting and Refining of Iron and Steel)

homogeneous subsets (Table 2.2), there are 6 subsets and it shows that the category "Smelting and Refining of Iron and Steel" is the highest, and followed by the category "Treatment and Disposal of Non-hazardous Waste".

From the analysis across different industry categories, we found that the waste-related or disposal-related industries are the major emission sources in Kaohsiung. It is because the waste incineration facilities locate close to the city center, therefore the pollutants from the sources could directly affect the high-population areas. In

Table 2.2 Descriptive statistics of monthly total GC events by different industries in Kaohsiung

Entry	Industry	Values
0	Manufacture of Paper	25.75 ± 8.96^{a}
1	Synthetic Rubber Manufacturing	33.50 ± 14.06^{a}
2	Steam Supply/Electricity Supply	34.50 ± 5.45^{a}
3	Electricity Supply	35.75 ± 26.74^{a}
4	Manufacture of Chemical Material	37.75 ± 12.42^{a}
5	Man-made Fibers Manufacturing	67.00 ± 24.86^{a}
6	Manufacture of Edible Oils and Fats	130.75 ± 17.02^{a}
7	Petroleum and Coal Products Manufacturing	$244.25 \pm 76.39^{a, b}$
8	Manufacture of Basic Chemical Material	$298.25 \pm 60.53^{b, c}$
9	Petrochemicals Manufacturing	$415.75 \pm 46.31^{c, d}$
10	Treatment and Disposal of Non-hazardous Waste	590.75 ± 68.05^{e}
11	Smelting and Refining of Iron and Steel	1680.25 ± 71.96^{f}

All data are presented as mean ± standard deviation

[a–f] Different superscript letters in the same column indicate significant statistical differences based on Scheffé post hoc test (p-value < 0.05)

Kaohsiung, most of the factories distribute along the coastal area within few kilometers from the city center. In addition, most of the AQMS stations are set in the populated area. Therefore, the neighbor influence is expected to amplify the GC event impact within the short distance. Furthermore, more AQMS stations surrounded also strengthen the magnitude of GC impacts on ambient air quality in this area.

2.4.4 Comparison with Emissions and Impacts

To quantify the relationship between the estimated emission impact and real emission volumes (Fig. 2.9), this study calculated the impact total emission volumes of SO_2 on the AQMS stations within 10-km radius of CEMS stations. Overall, the emission volumes are highly consistent with the magnitudes of the GC impacts, but there still exist some exceptions. It is because that the locations of these stationary emission sources are far away from the AQMS stations; therefore the GC impact is not significant.

The results show that the high emission source does not directly represent the higher impacts on the ambient air quality, because of the meteorological factors (wind field and atmospheric stability). Moreover, the spatial distributions of AQMS and CEMS stations directly affected the estimation of emission impacts. These results suggest that the proposed GC analysis in this study could sufficiently characterize the spatial relative contribution of different emission sources with the consideration of their distance to the receptors and the meteorological conditions.

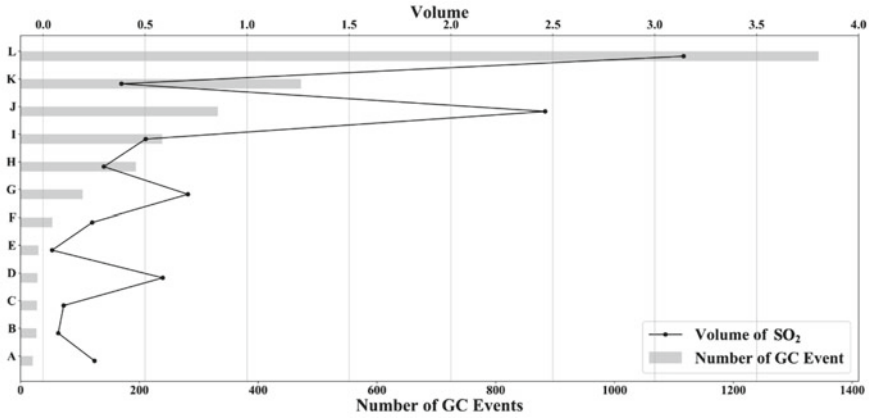


Fig. 2.9 The number of GC events of each industry versus SO_2 emission (A: Manufacture of Paper; B: Synthetic Rubber Manufacturing; C: Steam Supply/Electricity Supply; D: Electricity Supply; E: Manufacture of Chemical Material; F: Man-made Fibers Manufacturing; G: Manufacture of Edible Oils and Fats; H: Petroleum and Coal Products Manufacturing; I: Manufacture of Basic Chemical Material; J: Petrochemicals Manufacturing; K: Treatment and Disposal of Non-hazardous Waste; L: Smelting and Refining of Iron and Steel)

Based on the results abovementioned, for short-term, we could list all CEMS stations, which have high potential to influence a specific AQMS station in a specific study period. For long-term monitoring, the total GC events could sufficiently indicate how a CEMS station influences its ambient air quality. Compared to the previous studies [44, 46], the approach proposed in this study is relatively simple, fast, low cost and low complexity on computing and is capable to support for quickly overviewing the relative contributions of different air pollution sources.

Although the GC analysis in this study is relatively simple and accurate with low requirements on data quality and resolution, its performance still relies on the sampling frequency and spatial allocation of the emission sources and ambient air quality measurement. If the sampling frequency of the data is too low or too discrete, some small sources might be difficult to detect in this analysis. Furthermore, the spatial distribution of the ambient air quality stations in the study area directly affect the accuracy of analysis because of several geographical and meteorological factors could play important roles to affect the estimation.

2.5 Conclusion

This study proposed a modified Granger Causality analysis, a simple but novel data-driven method, to estimate the impact of air pollution from stationary sources on the ambient air quality in the spatial scale. This method provides an efficient way with lower computational resources to construct the causal relationship between the emission sources and the air quality measurement even if the numbers of sources and receptors are abundant. The statistical results of seasonal comparison gave a concrete evidence that this method is not affected by transboundary air pollutants, therefore the corresponding GC impact of each AQMS could be sufficiently derived from the emission sources.

In addition, this model could sufficiently characterize the temporal variation and spatial distribution of GC events over diurnal and seasonal cycles and it is capable to identify the contribution of meteorological condition (wind field). The high GC event area from coastal to inland area before the noon time, and in the afternoon the pattern of hot spots shifts from inland back to costal area due to the shift of sea-land-breeze system. Furthermore, the wind field in different seasons affects the GC event distribution. The highest number of GC events centralizes in inland area in spring season.

As for the analysis on industry category comparison, it showed that this model could sufficiently identify the contributions of different industry category. Furthermore, the trend of total emission volume of SO_2 is highly consistent with the GC event ranking of industry categories. Although the spatial inhomogeneity of emission sources and ambient air quality station might affect the estimation of GC events, this proposed method is still able to quantify the spatial causality between the sources and receptors. The theoretical contribution is to simplify and accelerate the construction of air pollution model for evaluating the relative impact from stationary sources

to ambient air quality. The relative impact could represent the air pollution from the specific factory, company or industry category on ambient air quality in specific period. In the practical contribution, this method provides an innovative and efficient way for citizens, administrations of government and non-government organizations to quickly overview the ambient air quality and the major contributors of local air pollution. As a result, this method could be utilized more efficiently even in cities, regions or countries with low environmental infrastructure.

Acknowledgements This study was supported in part by the Ministry of Science and Technology (MOST) of Taiwan (MOST 108-2111-M-002-013) and by National Taiwan University (NTU-108L901004). In addition, the authors appreciate the assistance from Taiwan's Environmental Protection Administration and CPC Corporation, Taiwan.

References

1. Zhou C, Li S, Wang S (2018) Examining the impacts of urban form on air pollution in developing countries: a case study of China's megacities. Int J Environ Res Public Health 15(8)
2. WHO (2019) Air pollution. [cited 2019 7 August]. https://www.who.int/airpollution/en/
3. Wang Q et al (2019) Impacts of residential energy consumption on the health burden of household air pollution: evidence from 135 countries. Energy Policy 128:284–295
4. Brook RD et al (2004) Air pollution and cardiovascular disease. Circulation 109(21):2655–2671
5. Mirabelli MC et al (2018) Air quality awareness among U.S. adults with respiratory and heart disease. Am J Prev Med 54(5):679–687
6. Khaniabadi, YO et al (2017) Acute myocardial infarction and COPD attributed to ambient SO2 in Iran. Environ Res 156:683–687
7. Amsalu E et al (2019) Short-term effect of ambient sulfur dioxide (SO2) on cause-specific cardiovascular hospital admission in Beijing, China: a time series study. Atmos Environ 208:74–81
8. Yuan Z et al (2006) Identification and spatiotemporal variations of dominant PM10 sources over Hong Kong. Atmos Environ 40(10):1803–1815
9. Yang JC et al (2019) Application of factor and cluster analyses to determine source–receptor relationships of industrial volatile organic odor species in a dual-optical sensing system. Atmos Meas Tech 12(10):5347–5362
10. Liu Y et al (2017) Source-receptor relationships for PM2.5 during typical pollution episodes in the Pearl River Delta city cluster, China. Sci Total Environ 596–597:194–206
11. Snyder MG et al (2013) RLINE: a line source dispersion model for near-surface releases. Atmos Environ 77:748–756
12. Burr MJ, Zhang Y (2011) Source apportionment of fine particulate matter over the Eastern U.S. Part I: source sensitivity simulations using CMAQ with the Brute Force method. Atmos Pollut Res 2(3):300–317
13. Stojić A et al (2015) Characterization of VOC sources in an urban area based on PTR-MS measurements and receptor modelling. Environ Sci Pollut Res 22(17):13137–13152
14. Mercer LD et al (2011) Comparing universal kriging and land-use regression for predicting concentrations of gaseous oxides of nitrogen (NOx) for the Multi-Ethnic study of Atherosclerosis and Air Pollution (MESA Air). Atmos Environ 45(26):4412–4420
15. Vienneau D et al (2010) Comparison of land-use regression models between Great Britain and the Netherlands. Atmos Environ 44(5):688–696
16. Battista G, de Lieto Vollaro R (2017) Correlation between air pollution and weather data in urban areas: assessment of the city of Rome (Italy) as spatially and temporally independent regarding pollutants. Atmos Environ 165:240–247

17. Lee P, Saylor R, McQueen J (2018) Air quality monitoring and forecasting. Atmosphere 9(3)
18. Monk K et al (2019) Evaluation of regional air quality models over Sydney and Australia: part 1—meteorological model comparison. Atmosphere 10(7)
19. Dennis RL et al (1996) The next generation of integrated air quality modeling: EPA's models-3. Atmos Environ 30(12):1925–1938
20. Basagaña X et al (2012) Effect of the number of measurement sites on land use regression models in estimating local air pollution. Atmos Environ 54:634–642
21. Dimov I et al (2004) Computational challenges in the numerical treatment of large air pollution models. Ecol Model 179(2):187–203
22. El-Harbawi M (2013) Air quality modelling, simulation, and computational methods: a review. Environ Rev 21(3):149–179
23. Wang B (2019) Applying machine-learning methods based on causality analysis to determine air quality in China. Pol J Environ Stud 28(5):3877–3885
24. Kang GK et al (2018) Air quality prediction: big data and machine learning approaches. Int. J. Environ. Sci. Dev. 9(1):8–16
25. Zhou Y et al (2018) Data-driven air quality characterization for Urban environments: a case study. IEEE Access 6:77996–78006
26. Zhao J et al (2019) Long short-term memory-Fully connected (LSTM-FC) neural network for PM2.5 concentration prediction. Chemosphere 220:486–492
27. Wang J, Song G (2018) A deep spatial-temporal ensemble model for air quality prediction. Neurocomputing 314:198–206
28. Feng X et al (2015) Artificial neural networks forecasting of PM2.5 pollution using air mass trajectory based geographic model and wavelet transformation. Atmos Environ 107:118–128
29. Zhou L et al (2019) Air big data outlier detection based on infinite Gauss Bayesian and CNN. In: Proceedings of the 2019 11th international conference on machine learning and computing. ACM, Zhuhai, China, pp 317–321
30. Niska H et al (2004) Evolving the neural network model for forecasting air pollution time series. Eng Appl Artif Intell 17(2):159–167
31. Wang W, Men C, Lu W (2008) Online prediction model based on support vector machine. Neurocomputing 71(4):550–558
32. Liu Y, Xu Z, Li C (2018) Online semi-supervised support vector machine. Inf Sci 439–440:125–141
33. Ozturk I, Acaravci A (2013) The long-run and causal analysis of energy, growth, openness and financial development on carbon emissions in Turkey. Energy Econ 36:262–267
34. Acaravci A, Ozturk I (2010) On the relationship between energy consumption, CO2 emissions and economic growth in Europe. Energy 35(12):5412–5420
35. Sahoo D, Honnorat N, Davatzikos C (2019) Sparse low-dimensional causal modeling for the analysis of brain function, vol 10949. SPIE Medical Imaging. SPIE
36. Akadiri SS et al (2018) Examining the causal impacts of tourism, globalization, economic growth and carbon emissions in tourism island territories: bootstrap panel Granger causality analysis. In: Current issues in tourism, pp 1–15
37. Xiong Y, Cui Y, Zhang X (2019) Pharmaceutical expenditure and total health-care expenditure in OECD countries and China: bidirectional Granger causality on the basis of health level. Exp Rev Pharmacoeconomics Outcomes Res 1–8
38. Schwartz J, Bind M-A, Koutrakis P (2017) Estimating causal effects of local air pollution on daily deaths: effect of low levels. Environ Health Perspect 125(1):23–29
39. Chen CWS et al (2018) Causality test of ambient fine particles and human influenza in Taiwan: age group-specific disparity and geographic heterogeneity. Environ Int 111:354–361
40. Bind M-A (2019) Causal modeling in environmental health. Annu Rev Public Health 40(1):23–43
41. Chaabouni S, Zghidi N, Ben Mbarek M (2016) On the causal dynamics between CO2 emissions, health expenditures and economic growth. Sustain Cities Soc 22:184–191
42. Pope CA et al (2019) Mortality risk and fine particulate air pollution in a large, representative cohort of U.S. adults. Environ Health Perspect 127(7):077007

43. Vodonos A, Awad YA, Schwartz J (2018) The concentration-response between long-term PM2.5 exposure and mortality; a meta-regression approach. Environ Res 166:677–689
44. Zhu JY, Sun C, Li VOK (2017) An extended spatio-temporal granger causality model for air quality estimation with heterogeneous urban big data. IEEE Trans Big Data 3(3):307–319
45. Zhang Y (2019) Dynamic effect analysis of meteorological conditions on air pollution: a case study from Beijing. Sci Total Environ 684:178–185
46. Jiang L, Bai L (2018) Spatio-temporal characteristics of urban air pollutions and their causal relationships: evidence from Beijing and its neighboring cities. Sci Rep 8(1):1279
47. Lai IC, Lee C-L, Huang H-C (2016) A new conceptual model for quantifying transboundary contribution of atmospheric pollutants in the East Asian Pacific rim region. Environ Int 88:160–168
48. EPA (2001) The impact on Taiwan's air quality from long-distance delivery of sandstorm. Executive Yuan, Taiwan, p 270
49. Chen T-F, Chang K-H, Tsai C-Y (2014) Modeling direct and indirect effect of long range transport on atmospheric PM2.5 levels. Atmos Environ 89:1–9
50. Li T-C et al (2017) Clustered long-range transport routes and potential sources of PM2.5 and their chemical characteristics around the Taiwan Strait. Atmos Environ 148:152–166
51. EPA (2001) Motorcycle type and emission coefficient in urban area. Executive Yuan, Taiwan: Taiwan, p 244
52. Lee T-C et al (2018) Bottom-up approach for downscaling CO2 emissions in Taiwan: robustness analysis and policy implications. J Environ Plan Manag 61(4):656–676
53. Wang H-K et al (2010) Seasonal variation and source apportionment of atmospheric carbonyl compounds in urban Kaohsiung, Taiwan. Aerosol Air Qual Res 10(6):559–570
54. Kurokawa J et al (2013) Emissions of air pollutants and greenhouse gases over Asian regions during 2000–2008: regional emission inventory in Asia (REAS) version 2. Atmos Chem Phys 13(21):11019–11058
55. Simon S et al (2016) Chemical composition of fog water at four sites in Taiwan. Aerosol Air Qual Res 16(3):618–631
56. Chang C-T et al (2017) Trends of two decadal precipitation chemistry in a subtropical rainforest in East Asia. Sci Total Environ 605–606:88–98
57. Stutz J et al (2004) Vertical profiles of NO3, N2O5, O3, and NOx in the nocturnal boundary layer: 1. Observations during the Texas air quality study 2000. J Geophys Res Atmos 109(D12)
58. Fernando HJS et al (2010) Flow, turbulence, and pollutant dispersion in urban atmospheres. Phys Fluids 22(5):051301
59. Granger CWJ (1969) Investigating causal relations by econometric models and cross-spectral methods. Econometrica 37(3):424–438
60. Tseng C-Y et al (2016) Characteristics of atmospheric PM2.5 in a densely populated city with multi-emission sources, vol 16
61. Lin CY et al (2007) Long-range transport of Asian dust and air pollutants to Taiwan: observed evidence and model simulation. Atmos Chem Phys 7(2):423–434
62. Yeh C-F, Lee C-L, Brimblecombe P (2017) Effects of seasonality and transport route on chemical characteristics of PM2.5 and PM2.5–10 in the East Asian Pacific Rim Region. Aerosol Air Qual Res 17(12):2988–3005

Dr. Chun-Hsiang Chan is an assistant professor at Undergraduate Program in Intelligent Computing and Big Data, Chung Yuan Christian University. He received his Ph.D. from Department of Geography, National Taiwan University. He has experienced spatiotemporal big data analysis with a demonstrated history of working in the academic and information industry.

Dr. Jehn-Yih Juang is currently an associate professor in the Department of Geography at National Taiwan University. His research is focused on surface-atmosphere interactions over different land cover types, urban environmental sciences, and the impacts of land-use change on environment form micro to regional scales.

Dr. Tzu-How Chu is an emeritus professor at Department of Geography, National Taiwan University. He received his Ph.D. from Department of Geography, University of Kansas. His research focused on spatial cognition, spatial information error, spatial information integration and automatic interpretation of remote sensing data.

Dr. Ching-Hao Mao is the Chief Executive Director, Taiwan Cyber Security Foundry Company in Taiwan. He earned his Ph.D. degree in the Department of Computer Science and Information Engineering at the National Taiwan University of Science and Technology and was a member of the Intelligent System Laboratory. His research interests are network security and machine learning.

Dr. Shin-Ying Huang is a senior engineer in the Center for Cybersecurity Service, Institute for Information Industry in Taiwan since 2016. She received her Ph.D. in the Department of Management Information Systems, National Chengchi University, Taiwan. Her research interests are machine learning, artificial intelligence, social media analysis, and cyber security.

Chapter 3
Groundwater Recharge, Monitoring and Finding Suitable Areas for Groundwater Recharge in Northeast Thailand

Pariwate Varnakovida, Htet Yamin Ko Ko, Thanet Natisri, Nawin Rinrat, and Piyawan Nakto

Abstract Groundwater is one of the most important natural resources but unsustainably uses for agriculture in Thailand. Groundwater accounts for available freshwater bodies and it becomes a major freshwater resource. It can be defined as precipitation that has infiltrated the soil surface and collected in soil and aquifer underground. It is the principal water source of domestic, industrial, and agricultural purposes. Serious groundwater shortage is experiencing especially in Northeast of Thailand because groundwater usage is higher than the recharged amount. Currently, groundwater resources are facing two major challenges such as overexploitation and contamination. Overexploitation is caused by the overdraft usage of groundwater resources and contamination is caused by the pollution caused by industrial, agricultural, and other human activities and the salinization and waterlogging caused by improper way to discharge wastewater, inadequate drainage system and insufficient conjunctive use. Groundwater recharge becomes an important matter not only because of previously stated facing major issues but also it is the primary resource of fresh water for drinking and irrigating crops, also in Northeastern of Thailand. This chapter will discuss about the basic knowledge of groundwater, groundwater recharge process, and groundwater monitoring and current methodologies applied in groundwater recharge in Northeast of Thailand.

Keywords Groundwater monitoring · Groundwater recharge · Groundwater replenishment

P. Varnakovida (✉) · H. Y. Ko Ko · T. Natisri · N. Rinrat · P. Nakto
Department of Mathematics, Faculty of Science, King Mongkut's University of Technology Thonburi (KMUTT), 126 Pracha-Uthit Road, Bang mod, Thung khru, Bangkok 10140, Thailand
e-mail: pariwate@gmail.com

T.-H. Wen et al. (eds.), *Earth Data Analytics for Planetary Health*, Atmosphere, Earth, Ocean & Space, https://doi.org/10.1007/978-981-19-8765-6_3

3.1 Introduction

Water is the essential part of life on Earth, and it is also a major part of the climate system such as clouds, air, lakes, vegetation, the ocean, glaciers, and snowpack [1]. Water also dominants climate change and plays a key factor in drought and floods. To meet the requirements of society and ecosystems, it is crucial to monitor the balance of available water resources. Clean water is used for domestic, industrial, agricultural, energy, waste disposal and recreation areas. Therefore, it is important to protect and preserve the clean water resources for human usage and ecosystem health. Fresh water can constitute only 3.5% of all water which covers 71% of Earth and the main source of freshwater comes from groundwater, glacial, streams and lakes. Currently, population, intensive usage of herbicides and pesticides, sprays on croplands, septic tank leakages and landfills are the major causes of water supply depletion [1]. Before learning about groundwater, understanding about Earth's water cycle can provide better understanding of how groundwater forms.

The steady and constant movement of water within atmosphere and earth in three stages (liquid, solid and gas) is defined as the water cycle. Besides other activities such as transpiration, evaporation, precipitation, condensation, and runoff included in the water cycle are explained in Fig. 3.1. Evaporation is the transition of earth surface's water in the liquid state to the vapor state and then entered the atmosphere [2]. Transpiration is the evaporation of water from the leaves of plants and evapotranspiration, or total evaporation is the combination of transpiration and evaporation from soils, ice, snow, vegetation, and other surfaces. Condensation is known as the process of vapor-to-liquid state transition, and it takes place when air contains more water vapor than it can receive. Precipitation is formed by releasing the water vapor in the atmosphere by condensation. Some precipitations that fall to earth can transit into four phases: evaporation, transpiration, percolation into soil by infiltration and flow directly via surface runoff into the sea [3]. Percolation into soil by infiltration penetrates streams as groundwater runoff and some portion of precipitation penetrate the ground to replenish Earth's groundwater.

Various research has been taken place to study about different aspects of groundwater. Water diversion research to induce the groundwater, groundwater quality monitoring research, groundwater recharge estimation research, groundwater recharge balance monitoring research, finding the possibilities of potential groundwater recharge zones research, finding the suitable areas for finding the groundwater recharge process and other research have been carried out and received a tremendous attention over last decade.

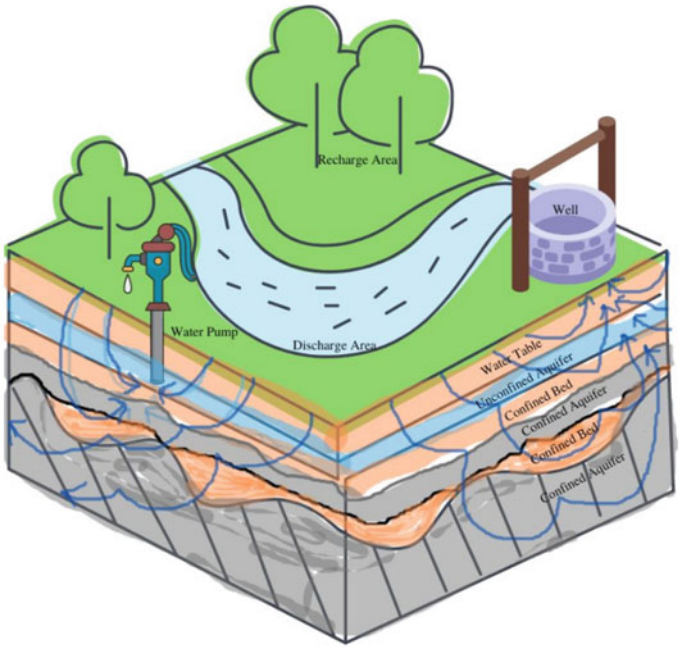


Fig. 3.1 Conceptual groundwater-flow diagram (Adopted from [3])

3.2 Groundwater

Groundwater is amongst the most prime native resources but mostly unrecognized resources for surface-water. Groundwater held 30% of accessible freshwater bodies and it becomes one of major freshwater fund. It can be defined as precipitation that has invaded the soil surface and collected in empty space [4]. It is the principal water source of domestic, industrial, and agricultural purposes. Groundwater can be discovered underground in the gaps and cracks in soil, rock, sand, gravel, and fractured rock. It can be replenished by rainfall and snow melts that seeps down during winter. It can be retrieved naturally by spring or recharged into lakes and streams by default. Groundwater can be found inside an aquifer which is an underground layer of porous rock or sediment. Water soaks into the ground, enters aquifer, rotate around the aquifer and reappear through wells and springs [4].

Groundwater flows underneath the water table which is the first occurrence of the groundwater that is called the saturated zone, in other words, Aquifer. In that zone, all pores and fractures are filled fully with the water. The upper layer of the water table is known as the unsaturated area where only pores are filled partially but not enough to support the enough water to discharge and only moisture the soil (Fig. 3.2).

The artificial groundwater retrieval methods are artesian well and pressure pump. Artesian Well is a place where groundwater is tapping under pressure and some

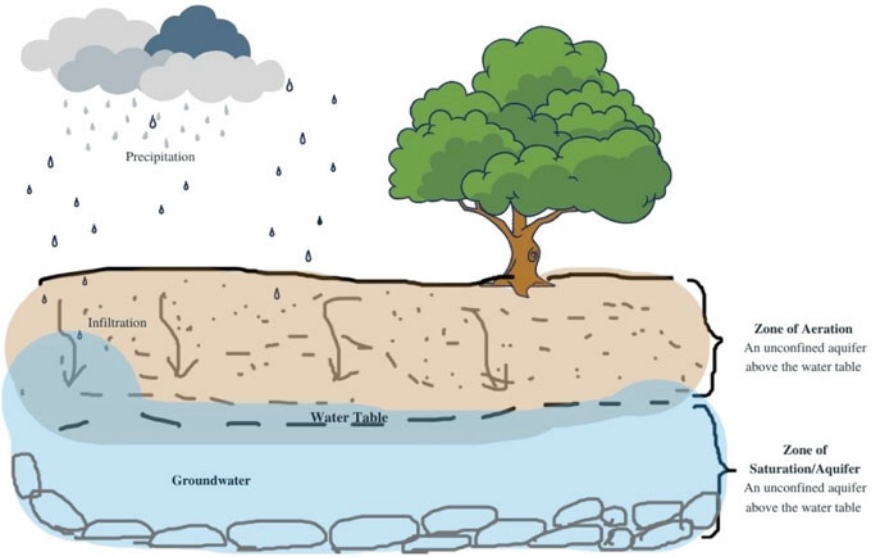


Fig. 3.2 Zonation the groundwater (Adopted from [5])

artesian wells require pump to take out the groundwater while some may have enough pressure to bring water to the surface [4]. Flowing Artesian Wells are the traditional way to get water resources for usage since 1850s. But artesian well can run out if there is reduction in groundwater level than the water intake depth.

Serious groundwater shortage is experiencing because groundwater usage is higher than the recharged amount in and a tremendous effort to monitor the groundwater levels over time to track dock the water-level declination becomes an important matter. Currently, groundwater resource is facing two major challenges such as contamination and overexploitation [6]. Overexploitation is caused by the overdraft usage of groundwater resources and contamination is caused by the pollution caused by industrial, agricultural and other human activities and the salinization and waterlogging caused by improper way to discharge wastewater, inadequate drainage system and insufficient conjunctive use.

Opposite of groundwater shortage in some areas, there are the places which required groundwater induction by using water diversion method. In the incident case of Tham Luang Cave Rescue, the cave was flooded and the level of water inside the cave could harm the divers to reach the boys [3]. In order to reduce the water inside the cave, pumping out the water from the cave was not effective and due to the heavy rain, it was still high. Therefore, with the help of geologists, geohydrologists and hydrologists, water diversion technology was applied to reduce the water level inside the cave by inducting the groundwater level and pumping out the surface water at the same time. Therefore, water level was reduced to the level which divers can dive to rescue the boys.

Fig. 3.3 Artificial groundwater recharge pond

To overcome the two major challenges (contamination and overexploitation) of groundwater resources, one of the possible solutions is groundwater recharge. Groundwater recharge process is groundwater resources replenishing process [7] and this is the main benefit of the groundwater recharge. Overexploitation happens mainly in dry wells, and it can degrade the quality of water and can lead to land subsidence [7]. Groundwater recharge process can be happened naturally, but overexploitation can occur when the draft rate is higher than the recharge rate or the recharge amount also reduce because of the climate change. This can be done by natural and artificial ways. Natural ways of groundwater recharge are to reserve the forest areas, to restore the forests, to increase the plantation in forests and to sustain the groundwater usage in agricultural sectors [7].

Artificial ways of groundwater recharge include collecting rainwater or flooded water and then added to the required/desired/appropriate areas. Later artificial groundwater recharge projects considered to recycle wastewater and agricultural runoff for the groundwater recharge resources [7]. The highly treated wastewater are collected and purified by three-step advanced purification techniques (microfiltration, reverse osmosis and ultraviolet light with hydrogen peroxide). By doing in this way, the contamination also reduces by not discharging the wastewater into the pacific ocean and not replenishing the wastewater blankly into the aquifers (Fig. 3.3).

In order to perform the artificial groundwater recharge process, finding suitable areas for groundwater recharge is one of the importance steps. The detailed explanation about finding potential areas for groundwater recharge will discuss in Groundwater Recharge section.

3.3 Groundwater Recharge

Groundwater Recharge becomes an important matter not only because of previously stated facing major issues but also it is the fundamental root of fresh water for drinking and irrigating crops, especially in island regions [8]. Groundwater Recharge

process is done by water entering an aquifer and it may occur from precipitation, rivers, lakes, canals, urbanization, and irrigation. Two types of recharge process exist, and they are natural process which are naturally recharge through the water cycle and anthropogenic process where reclaimed water and or rainwater is routed to the subsurface [9]. There are different kinds of natural groundwater recharge based on the coming flow of recharge water; these are focused and diffuse recharge. Focused recharge process exhibits the process of water recharge coming from a point source such as depressions, lakes, playas, and streams while diffuse recharge represents distribution of recharge water coming from a wide area by precipitation or irrigation [8, 10].

Due to the climate change and the use of the land have been assumed as some of the reasons for the obstacle of the natural recharge process. Rapid urbanization and agricultural sector demand the water more than it can supply and the groundwater is being cultivated more rapidly than its recharge rate. Artificial recharge methods are being introduced and applied to increase the groundwater supply. Artificial recharge to aquifers is charged by the surface water by using the augmented methods such as construction and spreading water into the underground. Artificial groundwater recharge became popular, and Thailand Authorities also considered to apply artificial recharge technique to reuse the floodwater in the basin area during rainy season to settle the stable groundwater surface and to recharge the depleted groundwater reserves [2].

Groundwater recharge by artificial recharge methods is known as groundwater replenishment and it is crucial to realize the physical characteristics of the study area. Finding the eligible area of groundwater recharge activity can perform by using data analysis methods that will determine the area that is expected to be suitable for groundwater recharge by overlapping the layers of data, for example, Geology. Geomorphology data, soil series data and slope data by assigning score values for each factor and the weight of the geological factors, geomorphology, soil and slope in which each area is assigned the importance of different factors according to the situations and physical characteristics of the study area.

The research of locating eligible areas for groundwater replenishment will output the areas which are very, moderate, low and unsuitable. Mostly four types of data are used in the research for finding the suitable groundwater recharge areas and they are geology, geomorphology, soil types and topographic slopes. In each data type, variety of data also included. For geology data, factors about rock types, weather degree, joints and fractures are considered while nature of earth's crust, faults, fractures, and structural density are considered as features of geomorphology. Type of soil, land use/land cover, spread of area and type of soil cover are factors of soil group and slope gradient of the study area is the feature of topographic slopes (Figs. 3.4 and 3.5).

By analyzing the data to find the eligible locations for groundwater replenishment by overlaying of each factor. The eligible areas for groundwater recharge is the selection of suitable or inappropriate areas which will select a wide area in order to assess the areas that are very suitable, moderate suitable, low suitable and unsuitability. Based on four primary data, namely geology, geomorphology, soil group, topography slope, factor values or criterion for giving different values of each factor

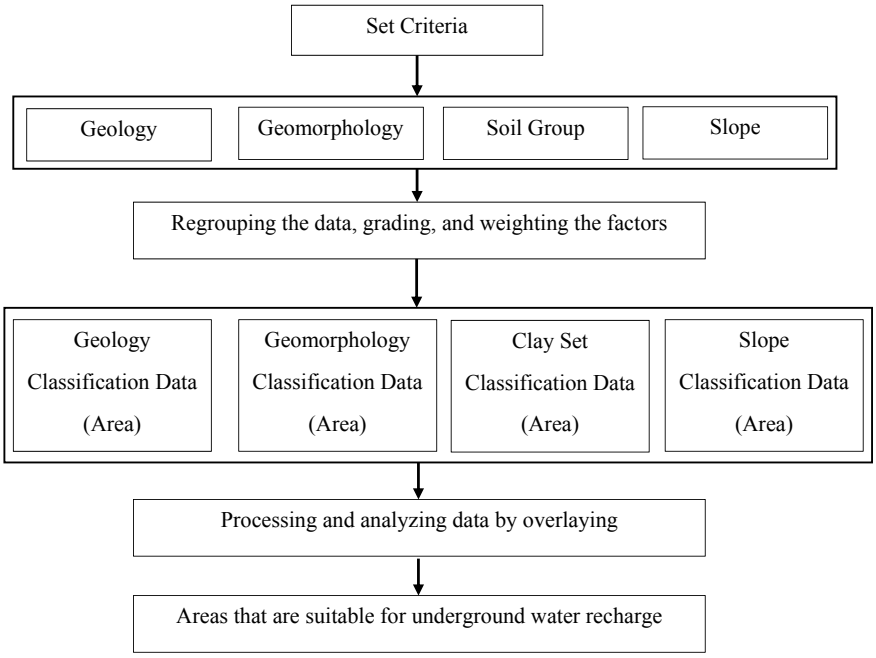


Fig. 3.4 Sample research framework for finding suitable areas for groundwater recharge

are determined by discussing with local geologist experts and also the conditions of the area. The scores for each factor were determined in Table 3.1 and the weights of the geomorphology, geology, soil set, and slope factors were 0.9, 0.8, 0.7 and 0.6, respectively.

The study found that in the northeastern region, the total area in the analysis was 166,676.80 km^2. It is very suitable 4,527.78 km^2 accounted for 2.72% of the total area. It is an area with good physical characteristics for groundwater recharge. Area with a moderate suitability level 52,013.82 km^2 accounted for 31.22% of the area. A detailed study of the groundwater exploration must be carried out before the operation. Low suitable area 98,569.78 km^2 accounted for 59.15% of the total area. Unsuitable area for groundwater recharge contains 115,203.41 km^2 accounted for 6.91% of the total area. If this result is applied in groundwater recharge policy, the water stress in this area might be better for local communities and farmers.

3.4 Groundwater Recharge Monitoring

Levels of groundwater are monitored by the recharge, storage, and discharge balance from an aquifer [11] since its variations in volume and quality take certain amount of times and it is not possible to determine by short time period survey alone [12].

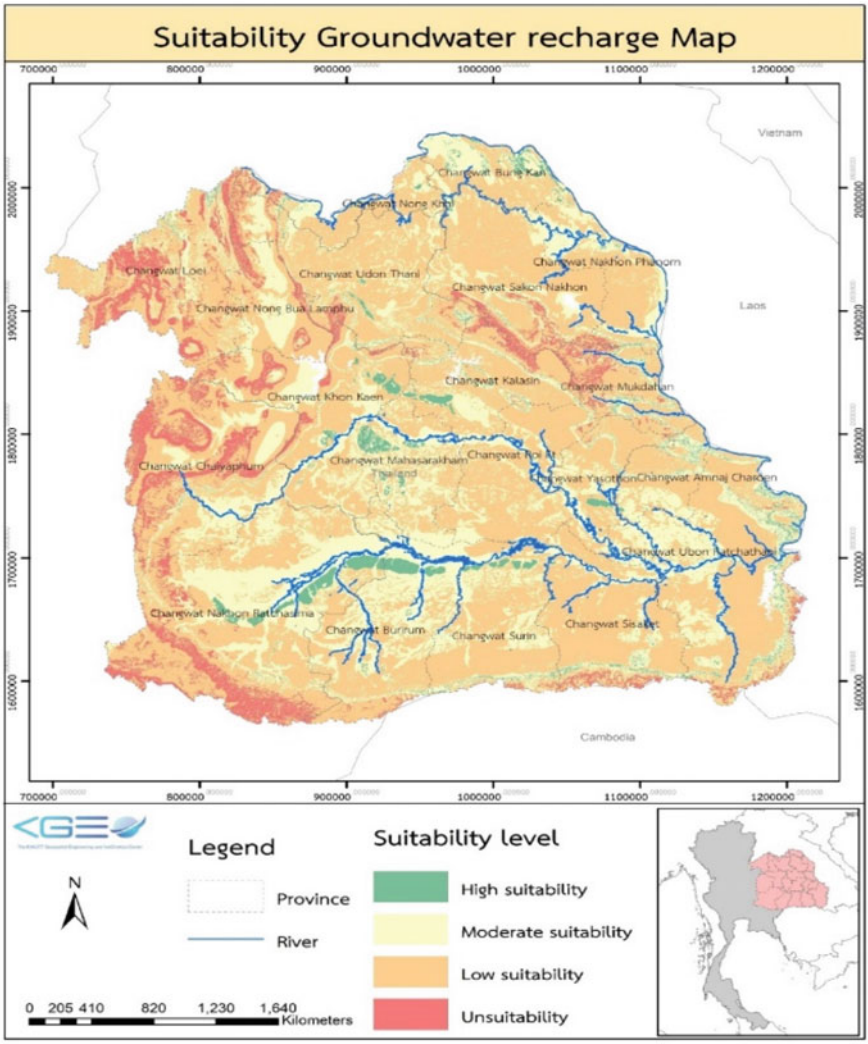


Fig. 3.5 Suitable areas for groundwater recharge in Northeast Thailand

The balance of groundwater is affected by the physical properties of composed aquifer such as porosity, permeability, hydraulic conductivity, transmissivity and rock thickness or sediments, and the climatic and hydrologic factors such as precipitation, evapotranspiration, rainfall, and release from the subsurface to surface-water bodies. For tropical, subtropics and temperate zones, groundwater recharge process rely on rainfall and for cold zones, groundwater recharge process can reply on both snow and rainfall. Water levels are also affected by the seasonal fluctuation, and it raised during spring and winter because of huge amount of precipitation and replenish but

Table 3.1 Factor score for each criteria

Criteria I geology		
No.	Rock type	Factor score
1	Granite, Diorite, Ultra mafic	1
2	Basalt, Rhyolite, Syenite, Andesite, Tuff	2
3	Silt, Mudstone, Shale, Claystone	1
4	Sandstone, Conglomerate	2
5	Limestone	3
6	Gneiss, Schist, Phyllite	1
7	Quartzite	1
8	Marble	1
9	Clay	1
10	Laterite and Rock fragments	4
11	Gravel, Sand, Silt	5
Criteria II geomorphology type		
No.	Geomorphology type	Factor score
1	Mountain	1
2	Foothills	4
3	Stream Terrace	5
4	Alluvial Fan	5
5	Watershed	3
6	Coastal plain	2
Criteria III permeability		
No.	Permeability	Permeation rate (meters per day)
1	Impermeable	<0.03
2	Slow permeation	0.03–0.12
3	Permeate quite slowly	0.12–0.48
4	Moderate permeability	0.48–1.50
5	Intermediate Permeability	1.50–3.00
6	Good Permeability	>3.00
Criteria IV slope		
No.	Slope	Factor score
1	0–2	4
2	2–5	5
3	5–10	3
4	10–20	2
5	>20	1

decreased during fall and summer because of small amount of recharge and greater amount of evapotranspiration.

Beside climate related fluctuations, the rate of underground-water discharge and recharge rate can be also affected by human activities [11]. Deforestation, urban development and draining of wetlands are the main cause of surface runoff and it can reduce the level of recharge while depression, streams, agricultural tillage and artificial wetland creations can increase the rate of groundwater recharge [8, 11]. Estimated amount of groundwater replenishment is an important factor to manage and protect groundwater resource because appropriate quantification of input (recharge) and output (discharge) for sustainable usage of groundwater. Geological factors such as soil type, climate, hydrology, geology, topography, and land use are the factors to consider for the recharge occurrence [13]. Land use and vegetation have effects on groundwater recharge process because evapotranspiration rate is high for vegetated land surface, and it resulted in less water for recharge.

3.5 Current Methodologies for Groundwater Recharge Monitoring

To perform estimation or analysis for groundwater recharge occurrence, deep knowledge about climate zone and hydrogeological conditions about the study area is required. Because evapotranspiration, mean annual rainfall, climate condition and topography consider as significant parts in groundwater replenishment distribution and the methods and the equations can be difference based on the study region [13]. Numerous methods are proposed to calculate the amount of groundwater recharge but selecting the efficient methods for the region of interest is still a challenging task. Important factors to be considered for selecting a research methodology include range, time/space scales, stability of recharge estimates based on different methods [10]. Strengths and weakness of each technique may differ based on various spatial resolutions, time scales for, different time scales over which recharge estimates are averaged and the recharge rate measurement limits [14].

Estimating groundwater recharge can have different study goals such as water resource assessment and aquifer vulnerability evaluation [10]. Groundwater quality evaluation requires recharge information over large spatial scales and decade-long temporal scales while aquifer assessment to pollution requires spatial variability and preferential flow. The purpose of the aquifer management is to regulate and monitor the effectiveness of abstraction on groundwater system [12]. Continuous monitoring of aquifer water balance and groundwater abstraction are the main purpose of groundwater resource management [12]. Therefore, different spatial scales and time periods of recharge estimates are one of the key factors because different research analysis require recharge estimates over different spatial and temporal scales.

The very first stage of groundwater monitoring is data collection about climate factors, geomorphology, hydrology, and geology of the region of interest to develop the conceptual model of recharge system [10]. Two types of data collection techniques to measure and control groundwater abstraction and usage are direction and indirect monitoring of groundwater abstraction. Direct monitoring technique is performed by the water meter reading of individual wells, and it is expensive to operate and required water users' cooperation. Even though it can provide accurate measurements, it is impossible to perform without authority. Invasive regulation of groundwater abstraction includes collection of representative information, estimations of changes and application of remote sensing techniques such as satellite or airborne data. Remote sensing data can reduce the cost of groundwater monitoring and investigation, but its major drawbacks are increase measurement error at low vapor fluxes, inaccurate to determine soil moisture deficit, actual evaporation, and recharge rate of aquifer [12].

Groundwater monitoring techniques can be broadly classified into three types based on three hydrological sources: classical assessment methods for surface water, unsaturated zone, or saturated zone [10, 14]. In surface water methods, stream baseflow, watershed models, tracers, seepage meters and channel water budget measurement methods are included. In unsaturated zone techniques, numerical models, tracers, and direct measurements technology are included whilst tracers, groundwater models and physical measurements are categorized as saturate zone approaches [12] (Table 3.2).

Popular Methods for Estimating Groundwater Recharge are direct measurements, water balance calculations, Darcy approaches, deconvolution of hydrographs, Water table fluctuation method, Tracer techniques and other approaches, mainly empirical

Table 3.2 Types of data required for groundwater management [12]

Type of data	Baseline data	Time-variant data
Ground water occurrence and aquifer properties	Water well records (Hydrological logs, Instantaneous groundwater levels and quality)	Groundwater level monitoring
	Aquifer and well pumping tests	Groundwater quality monitoring
Groundwater use	Water well pump installations	Water well abstraction monitoring (direct or indirect)
	Water-use Inventories	
	Population registers and forecasts	Well groundwater level variations
	Energy consumption for irrigation	
Supporting information	Climatic data	River flow gauging
	Land-use inventories	Meteorological observations
	Geological maps/sections	Satellite land-use surveys

methods. Water Balance Analysis is applied to estimate groundwater recharge. Water Balance Analysis equation may be different based on the climate zone of the area of interest. Gravity Recovery and Climate Experiment (GRACE) data is considered as one of the most effective satellite data available for groundwater resources management [15]. Time-lapse electrical resistivity (ERT) is a geoelectrical survey method used to capture the resistivity differences by water infiltration in the unsaturated area for recharge characterization [8].

Zoomable Object Oriented Distributed Model (ZOODRM) calculate replenishment by implementing soil moisture deficit (SMD) recharge technique through regular time periods. The SMD method determines how the actual evaporation (AE) is associated to the potential evaporation (PE) by the quantity of water collected in the soil at any time. ZOODRM model is appropriate for analysis regions where intend to provide lower data demands, usage of large remotely sensed records and more functionality for directing of overflow water corresponding to topography. The data points requirements of the model are land use, potential evaporation (PE), daily rainfall, topographical aspect directions, digital and geological elevation data.

Water Table Fluctuation (WTF) method is simple and efficient technique in estimation of groundwater recharge process [16]. It can be categorized as one of water-budget method [17]. Sucharit Koontanakulvong [16] integrated WTF method with Chloride Mass Balance method (CMB) and study found out replenishment expected by using CMB process is greater value than WTF method because CMB is local scale estimation and WTF is local scale estimation method.

Chloride Mass Balance (CMB) is founded on the concept of chloride in groundwater is received only from atmosphere via rainfall [8] and it is used to monitor groundwater recharge. The calculation of recharge amount is performed by using the following equation:

$$R = \frac{Cl_p}{Cl_{gw}} P \tag{3.1}$$

where R stands for recharge amount (mm/yr), P is the precipitation amount (mm/yr), Cl_p is weighted average chloride concentration of precipitation (mg/l) and Cl_{gw} is chloride concentration in groundwater (mg/l) respectively. This method provides outstanding results in quantifying recharge in arid and semi-arid regions [8] but later, studies found out that this method is also suitable for tropical climate areas.

Soil–water budget technique is applied to approximate groundwater recharge by conveying the principal of mass maintenance or linked principle where the contrast between inputs and outputs is identical to the variation in the system storage area [13]. Soil control volume is represented by using a simplified water-budget equation, as shown in Eq. 3.2.

$$P - ETR - ES - R = \Delta S \tag{3.2}$$

where P is the rainfall (mm), ES is the surface runoff (mm), ETR is the actual evapotranspiration (mm), R is the groundwater recharge (mm), and ΔS is the soil storage variation (mm). Equation 3.2 considers the major parts of the water balance.

Water-budget are the basic block of hydrologic conceptual model by linking the recharge process and other process in the water cycle [18]. Water-budget equation to assess the groundwater replenishment for an aquifer is shown in Eq. 3.3 where R is the recharge amount, ΔS_{gw} is the change in groundwater storage, Q_{bf} is the baseflow, ET_{gw} is the groundwater evapotranspiration and $(Q_{gw,off} - Q_{gw,on})$ is the net groundwater flow out of the aquifer which includes pumping and inter-aquifer.

$$R = \Delta S_{gw} + Q_{bf} + ET_{gw} + (Q_{gw,off} - Q_{gw,on}) \tag{3.3}$$

Water-budget, hydrograph separation with Eckhardt filter and MGB-IPH hydrological model analysis are performed for watershed groundwater recharge study using large-scale approach [13]. MGB-IPH model is based on the calculation modules of soil–water budget, flow distribution, flow directing through drainage network and evapotranspiration, as shown in Fig. 3.6: Diagrammatic Illustration of MGB-IPH model where HRU stands for Hydrological Response Unit (Adopted from [13]).

MODFLOW is considered for simulating and estimating groundwater/surface water associations and groundwater conditions and it is known as USGS's three-dimensional finite difference groundwater model [16]. MODFLOW considered hydraulic conductivity in x,y and z coordinate axes [0] proposed to study the correlation of groundwater recharge rate with monthly time series data of climate and hydrogeography data in the Nan Basin of the Upper Central Plain of Thailand by applying MODFLOW and using the bias corrected MRI-GCM data.

Xu et al. [19] incorporated a Soil–Water–Atmosphere–Plant (SWAP) with a groundwater flow model (MODFLOW) where SWAP package provides the vertical flux for MODFLOW, and mean water table depth was calculated by MODFLOW to

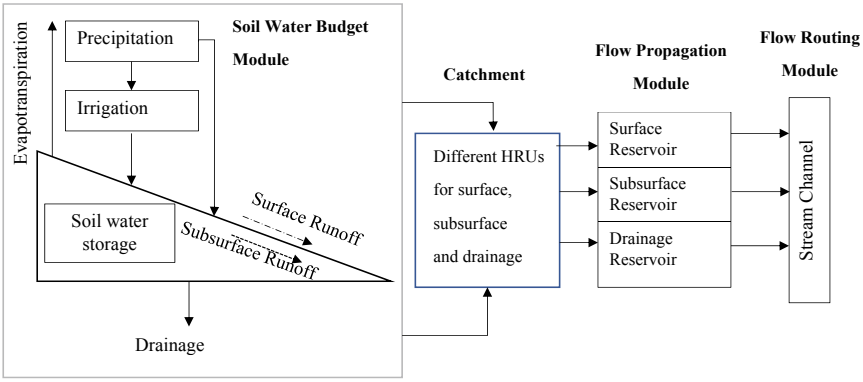


Fig. 3.6 Diagrammatic illustration of MGB-IPH model where HRU stands for hydrological response unit (Adopted from [13])

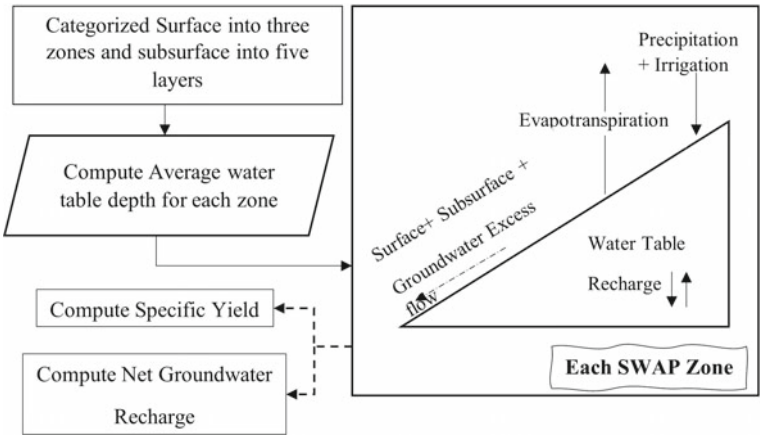


Fig. 3.7 MODFLOW-2000 and its coupling with SWAP Zones (Adapted from [19])

examine the bottom surface condition for SWAP zones. Data about land use, soil type, topology, and water management with geographic information systems (GIS) are also used to derive SWAP zones in MODFLOW. Then the SWAP MODFLOW package was examined using a two-dimensional saturated–unsaturated water table replenishment test. Outcomes revealed that the modeled water table heights equaled with the detected ones not including the initial phase during which they were marginally greater than the detected ones, probably due to ignoring lateral diffusion in the unsaturated zone. Lastly, SWAP package MODFLOW was applied to replicate a regional groundwater flow problem in Hetao Irrigation District, upper Yellow River basin of North China. Authors stated that MODFLOW with SWAP package can be applicable for sensible provincial groundwater shaping (Fig. 3.7).

Annual Groundwater Recharge rate in Thepkasattri watershed, Nothern Thalang District, Phuket Thailand is analyzed by water table fluctuation (WTF) method and chloride mass balance (CMB) method [8]. In this study, seasonal groundwater recharge with time-lapse electrical resistivity analysis is integrated to explore the recharge mechanism of the study region. The study region has 27 observing wells and minimum 42 production wells and authors stated that 28 wells are used for this study. Fifteen rainfall data were collected via 1000 ml and 500 ml polyethylene bottles monthly from April to October to collect average atmospheric wet chloride deposition. PT2X sensor is used to collect hourly groundwater level data from 2012 to 2015 where sensor can measure both pressure and temperature. Study found out that spatial distribution of groundwater recharge is affected by land use and landcover of the study region (Table 3.3).

The usefulness of groundwater monitoring system can be determined by three major components: data interpretation, system implementation and network design

Table 3.3 Summary of literature reviews

References	Study area	Data	Temporal scale	Method
Tesfaldet [8]	Thepkasattri watershed	Satellite Image, Phuket shapefile, Digital Elevation Model	March 2018 to February 2019, Hourly groundwater level data (2012–2015) from Department of Groundwater Resources Thailand	Water Table Fluctuation (WTF), Chloride Mass Balance
Maurício et al. [13]	Two watersheds located over the Serra Geral Aquifer System (Southern South America)	Precipitation (Rainfall Gauges from Brazilian Water Agency), Evapotranspiration and Landuse (MODIS16 with Penman–Monteith method), Meteorological data and emergency balance (Reanalysis), Surface runoff values (HEC-HMS rainfall-runoff model), Soil Moisture Accounting (SMA)model	Yearly	Water Budget Model, Hydrograph Separation, MGB-IPH Hydrological Model
Sucharit Koontanakulvong [16]	Nan Basin of the Upper Central Plain of Thailand	River water level, observation groundwater level, well abstraction, Recharge rate, Climate data, Precipitation, Evaporation, Temperature, Soil Type	Monthly	Linear Regression Method, MODFLOW Groundwater Model, Water Budget Analysis

(continued)

Table 3.3 (continued)

References	Study area	Data	Temporal scale	Method
Xu et al. [19]	Hetao Irrigation District, upper Yellow River basin of North China	Topography, soil type, land use, and water management practice	Yearly	Soil–Water–Atmosphere–Plant, MODFLOW

[12]. In network design component, it is important to define the research objectives clearly, to adapt the program accordingly, to understand groundwater flow system and to select the sampling locations and monitoring parameters accordingly. In system implementation component, joint analysis of groundwater and surface water monitoring should integrate if applicable.

3.6 Conclusion

Like other countries, Thailand also suffers the shortage of water, and it became a huge problem [20]. In 2020, Thailand was faced the worst drought in 40 years, and it exacerbated the current economic pressures by damaging the agricultural sector of sugar, rice, and rubber. Rapid development of Thailand's economic sector increases water demand, but climate change and deforestation caused more floods and droughts. Bangkok is suffering the continuous declination of piezometric levels with the consequences of water quality degradation and land subsidence by the extensive usage of groundwater [2]. Groundwater refueling in Thailand has been initiated for more than 30 years, and there are currently more than 10 projects that have been undertaken across the country for both shallow and deep groundwater recharge.

Different methodologies are applied to calculate the hydrogeological factors of an aquifer to monitor the groundwater balance, but studies find out a specific technique is not suitable to explain the complicated hydrogeological analysis [8]. Therefore, it is encouraged to integrate different methods to increase the reliability of estimations and enhance the information [8]. Groundwater monitoring is considered as expensive and intensive procedure because of capital cost of well monitoring network installation, data collection costs (satellite data, rain gauge), sampling costs for laboratory and data analysis costs [12]. But in return, it can monitor the sustainability of groundwater usage by avoiding the loss of groundwater resources and the costly remediation of aquifer. Therefore, cost-effective, and accurate groundwater monitoring system for the longer run will be the basic requirement of maintaining groundwater resources.

References

1. Administration NOAA (2019) Water cycle. https://www.noaa.gov/education/resource-collections/freshwater/water-cycle#:~:text=The%20water%20cycle%20shows%20the,form%20of%20rain%20and%20snow. Accessed 13 Mar 2022
2. Gupta AD, Babel MS (2005) Challenges for sustainable management of groundwater use in Bangkok, Thailand. Int J Water Resour Dev 453–464. https://doi.org/10.1080/07900620500036570
3. Siripornpibul C (2019) Geoscience and speleological works behind the operation of Tham Luang Cave Rescue, Thailand. http://www.dmr.go.th/download/article/article_20200528161401.pdf. Accessed 28 Mar 2022
4. Geographic N (2022) Aquifers. https://www.nationalgeographic.org/encyclopedia/aquifers/#:~:text=There%20are%20two%20general%20types,a%20permeable%20layer%20of%20soil. Accessed 13 Mar 2022
5. Salako Adebayo O, Adepelumi Abraham A (2018) Aquifer, classification and characterization. IntechOpen, London
6. Survey USG (nd) Conceptual groundwater flow diagram. https://www.usgs.gov/media/images/conceptual-groundwater-flow-diagram. Accessed 31 Mar 2022
7. AQUAOSO (2021) Groundwater recharge its importance and scalability. https://aquaoso.com/how-to-conserve-water-in-agriculture/groundwater-recharge/. Accessed 10 Jun 2022
8. Tesfaldet YT (2019) Groundwater recharge estimation in Thepkasattri Watershed Phuket, Thailand. Prince of Songkla University, Bangkok
9. Cherry AFAJ (1979) Groundwater, 1st edn. Library of Congress Cataloging in Publication Data, New Jersey
10. Scanlon BR, Healy RW, Cook PG (2002) Choosing appropriate techniques for quantifying groundwater recharge. Hydrogeol J 10:18–39. https://doi.org/10.1007/s10040-001-0176-2
11. Taylor CJ, Alley WM (2021) Ground_Water_Level monitoring and the importance of long-term water-level data. U.S. Geological Survey, Colorada
12. Tuinhof A, Dumars C, Foster S, Kemper K, Garduño H, Nanni M (2000) Sustainable groundwater management: concepts and tools. http://web.worldbank.org/archive/website00668/WEB/PDF/BN9.PDF. Accessed 15 Mar 2022
13. Melati MD, Fan FM, Athayde GB (2019) Groundwater recharge study based on hydrological data and hydrological modelling in a South American volcanic aquifer. Comptes Rendus Geosci 351(6):441–450
14. Gilmore TE (2017) Ground water. https://www.groundwater.org/file_download/inline/3e2aa7f2-0bed-4eab-b6a2-29384d0359be. Accessed 15 Mar 2022
15. Masood A, Tariq MAUR, Hashmi MZUR, Waseem M, Sarwar MK, Ali W, Farooq R, Almazroui M, Ng AW (2022) An overview of groundwater monitoring through point-to satellite-based techniques. Water 14. https://doi.org/10.3390/w14040565
16. Sucharit Koontanakulvong CS (2015) The role of groundwater to mitigate the drought and as an adaptation to climate change in the Phitsanulok Irrigation Project, in the Nan Basin Thailand. Jurnal Teknologi 76(15):89–95
17. Tushaar Shah DMRSDS (2000) The global groundwater situation: overview of opportunities and challenges, ed. International Water Management, Colombo, Sri Lanka. ISBN 92-9090-402-X
18. Healy RW (2010) Estimating groundwater recharge, 1st edn. Cambridge University Press, Cambridge
19. Xu X, Huang G, Zhan H, Qu Z, Huang Q (2012). Integration of SWAP and MODFLOW-2000 for modeling groundwater dynamics in shallow water table areas. J Hydrol 412–413: 170–181. https://doi.org/10.1016/j.jhydrol.2011.07.002
20. Manorom K (2020) Thailand's water shortage and inequality crisis. https://www.eastasiaforum.org/2020/03/20/thailands-water-shortage-and-inequality-crisis/#:~:text=Thailand%20is%20currently%20facing%20its,of%20sugar%2C%20rubber%20and%20rice. Accessed 14 Mar 2022

Dr. Pariwate Varnakovida is a director of KMUTT Geospatial Engineering and Innovation Center (KGEO) at the Department of Mathematics, Faculty of Science, King Mongut's University of Technological Thonburi. He is also currently working as Project Manager at Community Water Management by Thai Royal Initiative (CWMRI) for Ministry of Education and Data Scientist at KMUTT Big Data Experience Center. His research interests include Geographic Information System, Climate Change, Groundwater, Geohazards, integrating GIS knowledge into Spatial Decision Support System, Remote Sensing and Digital Satellite Image Processing, Big Data Analysis, Spatial Data Analysis and Sustainability. Dr. Pariwate Varnakovida signed MOU with CWMRI by representing King Mongkutt's University of Technology Thonburi (KMUTT). His contributions to the CWMRI are collaborating with College of Agriculture and Technology (CAT) to quickly establish a learning center through problem solving and mutual support within the water management processes of their respective areas, to use the integrated community-based water management and the King's science to increase awareness for urgent community requirements in water management, and also to create knowledge sharing and better understanding community-based water management to create a good quality of life for the community.

Ms. Htet Yamin Ko Ko is a doctoral student at the Department of Remote Sensing and Geographic Information System, Asian Institute of Technology, Pathum Thani, Thailand and she is currently working as a part-time research assistant at KEGO Centre of King Mongut's University of Technology Thonburi. Her research interests include Web GIS technology, Remote Sensing, Digital Image Processing, Satellite Image Processing, Deep Learning, Machine Learning, Biomedical Signal Processing, Remote Patient Monitoring Systems, Telemedicine and Medical Decision Support System.

Mr. Thanet Natisri is a US groundwater expert and an international team leader of American Groundwater Solutions. He is best known as one of the real heroes behind the 2018 Tham luang Cave Incident, after contributing to the diversion of rainwater inside the mountain and its tunnels. He was also titled the United Nations Human Rights Defender in 2018 and was awarded an Emmy for outstanding research in 2022 after his work on the National Geographic documentary "The Rescue". His research centers on integrating methods and domain knowledge into GIS-based Groundwater and Water Engineering Technology.

Mr. Nawin Rinrat graduated with a bachelor's degree in Marine and Coastal Resources Management Technology. He is currently working as a research assistant at the KEGO Center, King Mongkut's University of Technology Thonburi. His area of research and special interest are Geospatial Analysis, Remote Sensing and Geographic Information Systems, Groundwater Analysis and the integration of application of geographic information systems with environmental resource management.

Ms. Piyawan Nakto graduated with a bachelor's degree in Geographic Information System (GIS) and she is currently working as a research assistant at KEGO Centre of King Mongut's University of Technology Thonburi. Her specialized research and interests are Geospatial Analysis, Remote Sensing and Geographic Information System, Groundwater Bank, Climate change, Image processing, and Analysis of Urban Heat Island Hotspot.

Part II
Environmental Changes and Health

Chapter 4
Health Benefits of Air Pollution Reduction During the COVID-19 Lockdown Period in Thailand Using a Machine Learning Algorithm

Arthit Phosri and Mathuros Tipayamongkholgul

Abstract Due to the COVID-19 pandemic, the Thai government implemented the lockdown measure to restrict the widespread transmission of SARS-CoV-2 between April 3 and May 3, 2020. This measure might subsequently affect ambient air quality and its consequent health effect. Therefore, this study aimed to examine the change of ambient air pollution during the lockdown amid COVID-19 pandemic and its consequent change in mortality. The time-stratified case-crossover design with the conditional Poisson model was utilized to examine province-specific effect of individual air pollutants on non-accidental mortality and the random-effect meta-analysis was used to derive the national estimate of individual air pollutants on mortality during the period 2016–2019. The machine learning algorithm with random forests model was also applied to predict the business-as-usual concentration of air pollutants during lockdown period adjusting for weather variables and long-term trend. Changing in mortality attributed to air pollution change during lockdown period was then calculated. Results indicated that PM_{10}, $PM_{2.5}$, NO_2, and O_3 concentrations during lockdown period were lower than the business-as-usual scenario, leading to a reduction in mortality attributable to air pollutants. Findings from this study imply that stringent implementation of the air pollution control policies could reduce health impacts attributable to air pollution.

Keywords Air pollution · Mortality · Machine learning · Random forest model · COVID-19

A. Phosri (✉)
Department of Environmental Health Sciences, Faculty of Public Health, Mahidol University, 4th Floor, 2nd Building, Rajvithi Road, Bangkok 10400, Thailand
e-mail: arthit.pho@mahidol.ac.th

M. Tipayamongkholgul
Department of Epidemiology, Faculty of Public Health, Mahidol University, 3th Floor, 3nd Building, Rajvithi Road, Bangkok 10400, Thailand
e-mail: mathuros.tip@mahidol.ac.th

T.-H. Wen et al. (eds.), *Earth Data Analytics for Planetary Health*, Atmosphere, Earth, Ocean & Space, https://doi.org/10.1007/978-981-19-8765-6_4

4.1 Introduction

Ambient air pollution is one of the global environmental health issues contributing to a wide range of adverse health consequences with approximately 4.2 million premature deaths per year and 91% of them was observed in low- and middle-income countries [23]. The emergence of COVID-19 pandemic caused by Severe Acute Respiratory Syndrome CoronaVirus 2 (SARS-CoV-2) has imposed many countries to implement the lockdown measure to restrict the widespread transmission of SARS-CoV-2. This measure has been applied in a range of activities, including promoting quarantine measure [14], closing schools as well as implementing social distancing and staying at home to avoid large gathering [4]. These activities have consequently impacted the overall operation of many socioeconomic sectors such as manufacturing processes, logistic and supply chains, resulting in substantial decrease in anthropogenic emission of air pollutants such as particulate matter, O_3, NO_2, SO_2, and CO [7] and plenty improvement of air quality in many countries including China [27], Thailand [22], the United States [20], and Spain [21]. This improvement may subsequently induce public health and economic benefits [1, 26].

Due to the COVID-19 pandemic, the Thai government implemented a partially lockdown measure to restrict the widespread transmission of SARS-CoV-2 between April 3 and May 3, 2020 by means of restricting on the movement of people from their residences between 22.00 and 04.00 local time and limiting on inter-provincial travel. This measure has led to a substantial reduction in the concentration of ambient air pollution, where a previous study has indicated that the 24-h average concentration of $PM_{2.5}$, O_3, and CO in the Bangkok Metropolitan Area during the lockdown period was attenuated by 15.79%, 7.13%, and 8.01% respectively compared to the same period in 2017–2019 [22]. In addition, a growing body of epidemiological study has revealed that short-term exposure to ambient air pollution (PM_{10}, $PM_{2.5}$, NO_2, SO_2, O_3, and CO) was significantly associated with increased risk of human morbidity and mortality in many parts of the word [11, 13, 16–18, 28]. However, the published papers focusing on estimating health benefits associated with air pollution reduction during the COVID-19 lockdown period are limited and there is no such evidence in Thailand to date. Therefore, the objective of this study is to estimate the difference between observed concentrations of PM_{10}, $PM_{2.5}$, NO_2, and O_3 and their business-as-usual concentrations that would have been expected in the absence of lockdown measure during the lockdown periods, predicted through machine learning with meteorological normalization technique, and further estimate the change of mortality burden associated with changing in the concentration of ambient air pollutants.

4.2 The Situation of Air Pollution Concentration in Thailand

As of December 2020, there are 83 fixed-site air quality monitoring stations operated by the Pollution Control Department (PCD) located in 37 provinces throughout Thailand. The hourly concentration of air pollutants (PM_{10}, $PM_{2.5}$, NO_2, and O_3) from January 2016 to December 2020 throughout Thailand was applied in this study and daily concentration of individual air pollutants for each province during such period was calculated by taking the average of hourly concentration. This calculation excluded the day that had missing values more than six hours, where the concentration for that particular days were assigned as missing. The missing data was then imputed using the expectation—maximization (EM) algorithm under the assumptions of multivariate normal distribution [10]. During the study period, the average concentrations ± standard variation of PM_{10}, $PM_{2.5}$, NO_2, and O_3 across 37 provinces was 42.92 ± 27.42, 24.43 ± 18.87, 31.42 ± 31.59, and 74.42 ± 37.26 μg/m^3, respectively. Daily average concentrations of PM_{10}, $PM_{2.5}$, and NO_2 across 37 provinces during the lockdown periods (April 3 and May 3) in 2020 were lower than those in the same period of 2016–2019, but the average concentration of O_3 in 2020 was slightly higher, compared with the same periods of 2016–2019. Specifically, the average concentrations ± standard variations of PM_{10}, $PM_{2.5}$, and NO_2 during the lockdown periods in 2020 were respectively 43.84 ± 26.72, 24.54 ± 21.26, and 20.98 ± 23.35 μg/m^3, whereas those in the same periods of 2016–2019 were respectively 45.27 ± 27.44, 26.23 ± 20.50, and 27.84 ± 26.47 μg/m^3. The daily average concentration of O_3 during the lockdown periods in 2020 was 89.08 ± 35.97 μg/m^3 and that in the same periods of 2016–2019 was 88.76 ± 40.11 μg/m^3 (Fig. 4.1).

4.3 Short-Term Effect of Ambient Air Pollution on Mortality in Thailand

Non-accidental mortality was used as adverse health outcome indicator in the present study, where data was obtained from the Strategy and Planning Division under the Ministry of Public Health of Thailand from 2016 to 2020. Specifically, the mortality data were extracted from death certificate that indicated date of death, age, sex, and primary diagnostic code in accordance with the International Classification of Disease 10th revision (ICD-10), and non-accidental cause of death (ICD-10: A00-R99) was applied in this study. To investigate short-term effect of ambient air pollutants on mortality in Thailand, the statistical analysis was divided into two stages. In the first stage, province-specific effects of individual air pollutants (PM_{10}, $PM_{2.5}$, NO_2, and O_3) on mortality were examined using the Distributed Lag Non-linear Model (DLNM) combined with the case-crossover design and were analyzed using conditional quasi-Poisson model taking into account overdispersion. In particular, the bi-dimensional cross-basis matrix between each air pollutant and its lag was

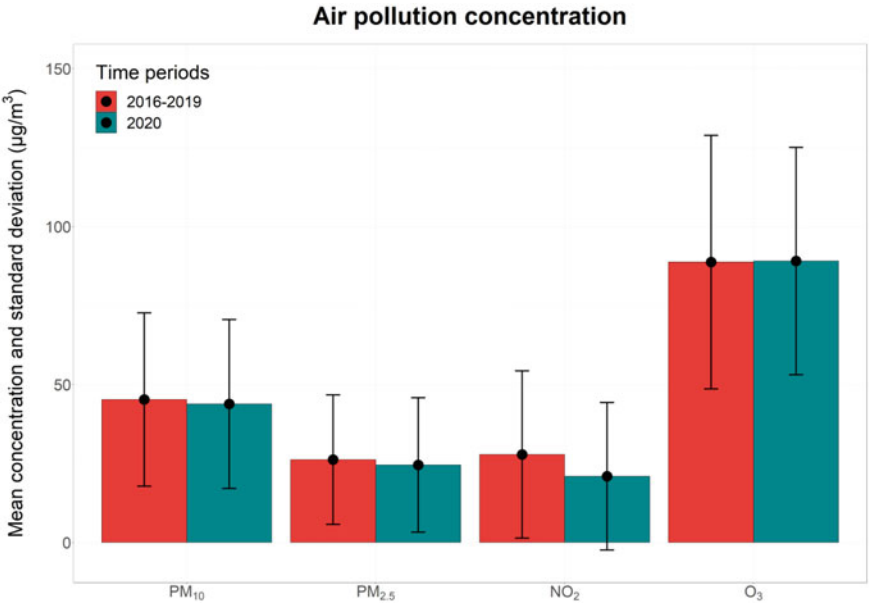


Fig. 4.1 Concentrations (mean ± standard deviations) of air pollutants during COVID-19 lockdown period in 2020 (labeled in green) and the average concentrations during the same period in 2016–2019 (labeled in orange)

constructed in DLNM using linear function for each air pollutant and natural cubic spline function with 3 degrees of freedom for its logarithmic scale of lag over 7 days. This degree of freedom was selected because of minimizing the quasi-Akaike Information Criteria (Q-AIC) [5], and a given cross-basis matrix of each air pollutants was then fitted in the conditional quasi-Poisson regression model designed by case-crossover framework, where three-way interaction terms of year, calendar month, and day of the week of case and its many control days were matched within the same stratum. Specifically, a stratum was defined by matching each case day with many control days that are in the same day of the week within the same calendar month and year as case day. This epidemiological design automatically adjusts for day of the week, seasonality and long-term trend [2]. Moreover, temperature and relative humidity was used as potential confounding factors on the association between air pollutants and mortality by fitting natural cubic spline function with three degrees of freedom each. The algebraic equation applied to analyze province-specific association between air pollutant and mortality is shown below;

$$Log\left[E\left(Y_{t,i}\right)\right] = \alpha + \beta_{p,i} Z_{p,t,l,i} + ns\left(tmean_{t,i},\ 3\right) + ns\left(humid_{t,i},\ 3\right) + stratum_{t,i} \quad (4.1)$$

where $Log[E(Y_{t,i})]$ is the natural logarithm of the expected number of non-accidental mortality at day t and province i; α is model intercept; $\beta_{p,i}$ defines the coefficients of pollutant p associated with mortality at province i; $Z_{p,t,l,i}$ is the cross-basis matrix of

air pollutant p generated by DLNM at day t lag l and province i; $ns(\cdot)$ indicates the natural cubic spline function with predefined degree of freedom; $tmean_{t,i}$ and $hunid_{t,i}$ is daily mean temperature and relative humidity at day t and province i, respectively; $stratum_{t,i}$ is the three-way interaction between case and its many control days at day t and province i.

In the second stage, province-specific estimates obtained from the first stage were pooled to derive the national estimated effects of individual air pollutants on mortality using random-effects meta-analysis with restricted maximum likelihood (REML) estimation. The random-effects model is given by;

$$\beta'_{p,i} = \beta_p + \delta_{p,i} + \theta_{p,i} \tag{4.2}$$

$$\beta'_{p,i} \sim N(\beta_p, D_p + S_{p,i}) \tag{4.3}$$

where $\beta'_{p,i}$ indicates the risk estimate of air pollutant p on mortality for province i; β_p is the national estimate of pollutant p that we would like to obtain; $\delta_{p,i}$ is a vector of between-province random sampling error associated with pollutant p and province i; $\theta_{p,i}$ is a vector of within-province random error for air pollutant p and province i. The province-specific estimates are assumed to be normally distributed, where D_p indicates the covariance matrix of $\delta_{p,i}$ and $S_{p,i}$ is the covariance matrix of $\theta_{p,i}$ obtained from the first stage. The national estimate was shown as relative risk (RR) of non-accidental mortality associated with 10 $\mu g/m^3$ increase of air pollutants. The I^2 statistic and Cochran Q-test was also applied to explore the residual heterogeneity among province-specific estimates [6]. All statistical analyses were performed using the R package for statistical computing (version 4.1.0) with "*dlnm*", "*splines*", and "*gnm*" packages.

The results indicated that exposure to PM_{10}, $PM_{2.5}$, NO_2, and O_3 was significantly associated with an increased risk of mortality in Thailand, where the highest estimates were observed at cumulative lag 0–2 days. Specifically, the pooled RRs of mortality associated with 10 $\mu g/m^3$ increase of PM_{10}, $PM_{2.5}$, NO_2, and O_3 were 1.0097 (95% CI: 1.0064, 1.0131), 1.0171 (95% CI: 1.0108, 1.0234), 1.0391 (95% CI: 1.0225, 1.0560), and 1.0106 (95% CI: 1.0054, 1.0158) at lag 0–2 days, respectively, and these RRs was then applied as concentration–response function to further estimate mortality benefits associated with air pollution reduction during the COVID-19 lockdown period. Findings also revealed that there was a significant heterogeneity among provinces with I^2 statistic and p-value from Cochran Q-test of 43.4% and <0.001 for PM_{10}, 44.9% and <0.001 for $PM_{2.5}$, 67.2% and <0.001 for NO_2, and 32.8% and 0.0016 for O_3 (Fig. 4.2), suggesting that the significant difference on the effect estimates of PM_{10}, $PM_{2.5}$, NO_2, and O_3 on non-accidental mortality among provinces is observed that could be explained by other external predictors.

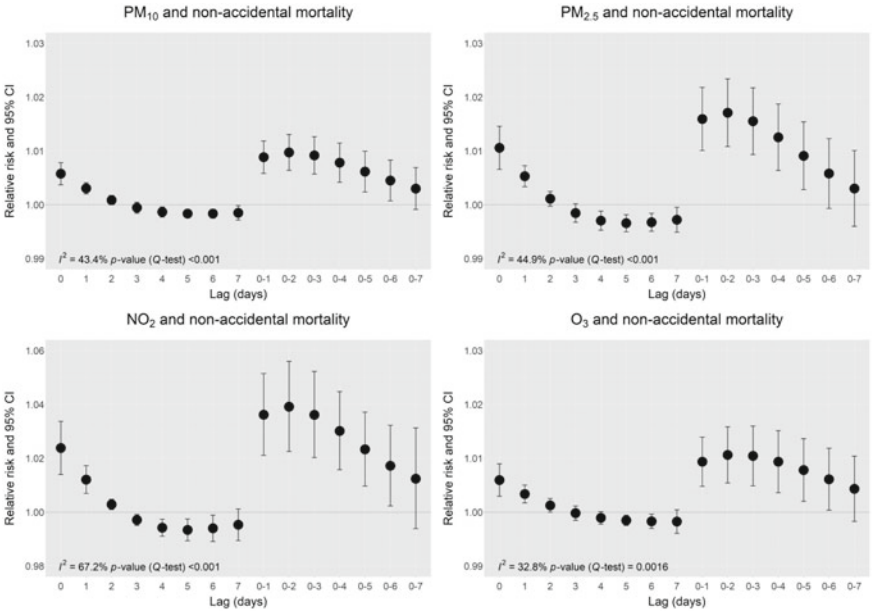


Fig. 4.2 Pooled RR and 95% confidence interval of mortality associated with every 10 $\mu g/m^3$ increase of PM_{10}, $PM_{2.5}$, NO_2, and O_3 at different lag structures

4.4 Business-as-Usual Concentration of Air Pollution During the COVID-19 Lockdown Period

Business-as-usual (BAU) concentration is the concentration of air pollutants that would have been observed when lockdown measure is not implemented. The BAU concentration was predicted using a machine learning algorithms with meteorological normalization technique. The meteorological data, including daily average temperature (°C) and relative humidity (%), as well as monthly air pressure (hPa), and daily cumulative rainfall (mm), were obtained from weather monitoring stations of the Thai Meteorological Department under the Ministry of Digital Economy and Society located in the same provinces as of air pollution data during January 2016 through December 2020. Nevertheless, only 29 out of 37 provinces were finally included to predict the BAU concentrations of PM_{10} and $PM_{2.5}$ during the COVID-19 lockdown period because monthly air pressure and daily cumulative rainfall data were not available in the rest of provinces, in which 28 provinces were ultimately included to estimate the BAU concentrations of NO_2 and O_3 since there was no NO_2 and O_3 data in the rest of province during the study period. A random forests (RF) model is a machine learning algorithm that used to estimate the BAU concentration in this study through meteorological normalization technique. In particular, a number of de-correlated regression trees are basically grown from the original data using bootstrap sampling technique, and further average their predictions from all

regression tress to obtain the final prediction. This method is approximately unbiased, and therefore reduces the variance of an estimated prediction [9]. The RF model based meteorological normalization technique with 50 regression trees was built through development of the decision trees to predict PM_{10}, $PM_{2.5}$, NO_2, and O_3 concentrations separately using ambient temperature, relative humidity, atmospheric pressure, rainfall, as well as day of the year, calendar year, and indicator variable for province as predictor variables. The training set of RF model was performed using the data from 2016 to 2019, and daily province-specific concentration of air pollutants in 2020 was predicted using the estimates obtained from the training set. The predicted concentrations during April 3 through May 3, 2020 were defined as BAU concentration. The RF model used in this study is described below;

$$Z_{p,t} = RF\left(tmean_t,\ humid_t,\ press_t,\ rain_t,\ doy_t,\ year_t,\ province_{i,t}\right) \quad (4.4)$$

where $Z_{p,t}$ is the concentration of air pollutant p on day t; $tmean_t$, $humid_t$, $press_t$, and $rain_t$ is level of temperature, relative humidity, atmospheric pressure, and rainfall at day t, respectively; doy_t is day of the year at day t; $year_t$ is calendar year at day t; $province_{i,t}$ is indicator variable for province i at day t.

Findings revealed that the daily average concentrations of observed PM_{10} and $PM_{2.5}$ over 29 provinces and those of observed NO_2 and O_3 across 28 provinces were lower compared to predicted BAU concentration during the lockdown period in 2020 (Fig. 4.3). In particular, mean ± SD of observed PM_{10}, $PM_{2.5}$, NO_2, and O_3 concentrations were 43.84 ± 26.72, 24.54 ± 21.26, 20.98 ± 23.35, and 89.08 ± 35.97 $\mu g/m^3$, respectively whereas that of predicted BAU concentration of PM_{10}, $PM_{2.5}$, NO_2, and O_3 was 56.13 ± 16.90, 34.51 ± 12.35, 27.86 ± 17.04, and 102.22 ± 27.65 $\mu g/m^3$, respectively. These findings indicate that implementing the COVID-19 lockdown measures lead to a reduction of ambient air pollution.

The validity of RF model was also performed using province-stratified tenfold cross-validation (CV) method using the data from 2016 to 2019. The coefficient of determination (R^2) and root mean square error (RMSE) were calculated as validating parameters. Results revealed that CV R^2 for PM_{10}, $PM_{2.5}$, NO_2, and O_3 was respectively 0.79, 0.79, 0.86, and 0.77, and CV RMSE for PM_{10}, $PM_{2.5}$, NO_2, and O_3 was 11.53, 8.38, 10.28, and 16.70 $\mu g/m^3$, respectively. This finding suggests that RF model based weather normalization technique is accurately predicted BAU concentration of air pollutants in this study, to some extent. The RF model was performed using "*ranger*" and CV was performed using "*caret*" package in R (version 4.1.0).

4.5 Health Benefits of Air Pollution Reduction During the COVID-19 Lockdown Period

The number of mortality attributable to individual air pollutants over n days was computed using the following equation;

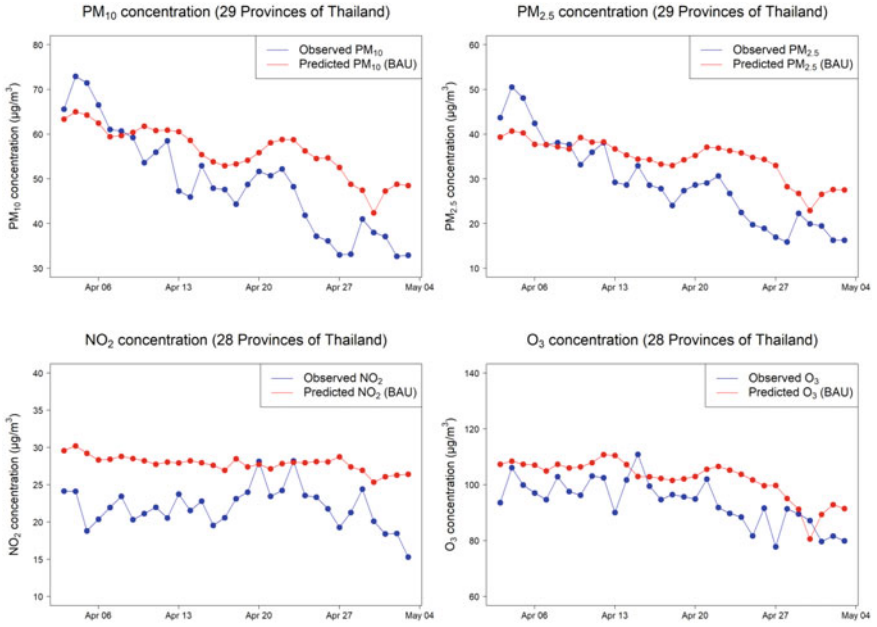


Fig. 4.3 Concentration of PM_{10}, $PM_{2.5}$, NO_2, and O_3 during the COVID-19 lockdown period in 2020 in Thailand, red lines represent the predicted BAU concentration through RF based weather normalization technique and blue lines indicate observed concentration (*Source* Saengsawang and Phosri under review)

$$\sum_{t=1}^{n} \mathrm{MOR_{pt}} = \left[\exp\left(\beta_p * \Delta x_{pt}\right) - 1\right] * N_t \tag{4.5}$$

where MOR_{pt} indicates the number of mortality attributed to air pollutant p on day t; β_p is log relative risk of mortality associated with pollutant p obtained from conditional quasi-Poisson model, therefore β_p in this study is calculated from $\log(RR_p)/10$; Δx_{pt} is concentration of pollutant p on day t exceeding the threshold, where the threshold was set as zero because the association between air pollutants and mortality from preliminary analysis followed the linear pattern with no threshold level, indicating that increased exposure to air pollutants was associated with an increased risk of mortality; N_t defines the total number of mortality at day t. The mortality benefit associated with a reduced concentration of air pollution during the COVID-19 lockdown periods over m provinces was then calculated using equation described below;

$$\sum_{i=1}^{m}\sum_{t=1}^{n} \Delta\mathrm{MOR_{pti}} = \sum_{i=1}^{m}\sum_{t=1}^{n} \mathrm{MOR_{pti(observed)}} - \sum_{i=1}^{m}\sum_{t=1}^{n} \mathrm{MOR_{pti(BAU)}} \tag{4.6}$$

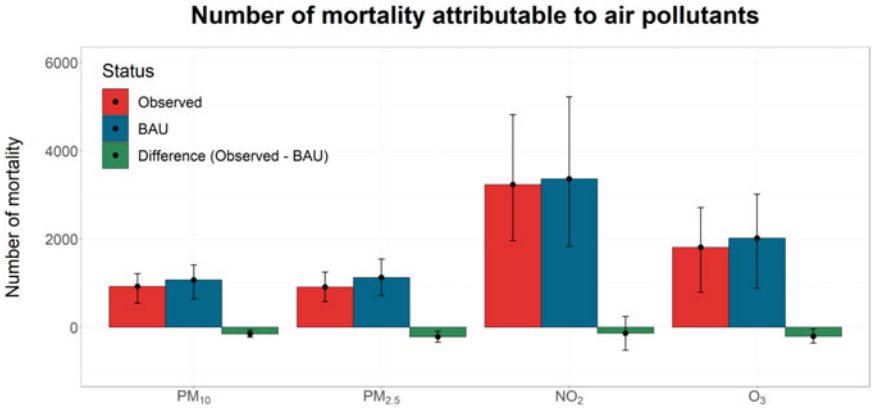


Fig. 4.4 Number of mortality attributable to observed (orange) and predicted BAU (blue) concentration of air pollutants and the difference between the two (green)

where ΔMOR_{pti} indicates the difference in the number of mortality attributed to air pollutant p between observed and BAU concentration at day t and province i; $MOR_{pti(observed)}$ and $MOR_{pti(BAU)}$ is the number of mortality attributed to the observed and predicted BAU concentration of air pollutant p on day t at province i, respectively.

Findings revealed that implementation of the COVID-19 lockdown measures in Thailand leads to a reduction in number of mortality attributable to a reduced concentration of PM_{10}, $PM_{2.5}$, NO_2, and O_3 over 28–29 provinces. Specifically, a reduction of PM_{10}, $PM_{2.5}$, NO_2, and O_3 concentrations over the lockdown periods (i.e., From April 3 to May 3, 2020) was associated with a reduction of −150 (95% CI: −225, −59) deaths, −215 (95% CI: −338, −91) deaths, −134 (95% CI: −517, 245) deaths, and −205 (95% CI: −359, −37) deaths, respectively (Fig. 4.4).

4.6 Discussion and Conclusions

RF model based meteorological normalization technique was used to predict the BAU concentration of PM_{10}, $PM_{2.5}$, NO_2, and O_3 during the COVID-19 lockdown periods, and the concentration–response function obtained from the epidemiological model was subsequently applied to estimate the change in number of mortality attributable to air pollution reduction over 29 provinces for PM_{10} and $PM_{2.5}$, and across 28 provinces for NO_2 and O_3 in Thailand.

The results of RF model, a machine learning algorithm based weather normalization technique, show that the predicted BAU concentration of PM_{10}, $PM_{2.5}$, NO_2, and O_3 was relatively accurate as indicated by CV R^2 and RMSE, although many other potential factors that influence the concentration of air pollutants were not taken into consideration, such as wind speed, wind direction, sunshine duration, and planetary boundary layer height (PBLH) [24, 25]. This finding is similar to previous

studies in China [26] and Spain [1], suggesting that RF model based meteorological normalization technique has accurately performed to predict the BAU concentration of air pollutants. The results also reveal that the average of observed concentration of PM_{10} and $PM_{2.5}$ over 29 provinces and NO_2 and O_3 across 28 provinces was lower than their predicted BAU concentration during the COVID-19 lockdown periods, especially that of NO_2. These results can be explained by the fact that the COVID-19 lockdown measures substantially restrict socioeconomic activities, including industrial processes and logistic and supply chains that may lead to a substantial reduction in anthropogenic emission and subsequent decrease in ambient air pollution concentration [22].

Findings from epidemiologic model indicate that the highest RR of non-accidental mortality associated with air pollutants was observed at lag 0–2 days, where RR of mortality associated with exposure to every 10 $\mu g/m^3$ increase of PM_{10}, $PM_{2.5}$, NO_2, and O_3 was 1.0097 (95% CI: 1.0064, 1.0131), 1.0171 (95% CI: 1.0108, 1.0234), 1.0391 (95% CI: 1.0225, 1.0560), and 1.0106 (95% CI: 1.0054, 1.0158), respectively. The magnitude of the estimates observed in this study was slightly higher compared to a recent study in Thailand [8]. This difference might be due in part to different modeling frameworks, study periods, and a limited number of monitoring stations and provinces included in a previous study, leading to bias the effect toward the null hypothesis. However, the association and lagged patterns observed in this study were similar to previous studies in many parts of the world, showing that the effect of PM_{10}, $PM_{2.5}$, NO_2, and O_3 on non-accidental mortality seem to be linear (no threshold level) with the highest magnitude observed at between lag 0 and lag 7 days [3, 8, 15].

During the COVID-19 lockdown period in Thailand, hundreds of deaths were avoided due to reduction in air pollution concentrations, especially in $PM_{2.5}$ and O_3. This finding is similar to a previous study in China, showing that reduction of PM_{10}, $PM_{2.5}$, NO_2, and O_3 during the lockdown period was associated with a reduction in the number of deaths by 2777 (95% CI: 1565, 3995), 1239 (95% CI: 844, 1578), 4711 (95% CI: 3649, 5781), and 215 (95% CI: 116, 314), respectively throughout 367 Chinese cities [26]. However, the magnitude of mortality benefits attributable to air pollutants reduction in this study was different from that of a previous study. This different finding could be explained by the number of target populations included in the study. Specifically, previous study applied mortality rate and number of populations in 367 cities to estimate the number of death, but the current study used the total number of death obtained directly from the Ministry of Public Heath in 28 or 29 provinces, and further quantifies the decreased number of deaths attributable to air pollution reduction during the lockdown period. In this study, some limitations had to be acknowledged. First, RF model based weather normalization technique was applied to estimate the BAU concentration during the lockdown period, but many sources of uncertainties that influence the concentration of air pollutants were not taken into account. However, the model performance seems to be satisfied, to some extent, with CV R^2 for PM_{10}, $PM_{2.5}$, NO_2, and O_3 of 0.79, 0.79, 0.86, and 0.77, respectively. Second, people spend most of their time indoor during the lockdown period, where indoor air pollutants could mainly influence

adverse health effects. Nevertheless, indoor air quality data was not accounted in this study because previous studies in Thailand has revealed that outdoor air pollutants were well correlated with indoor concentration and outdoor air pollutants are the main contributor to indoor concentration [12, 19]. Third, health benefits attributed to air quality improvement in this study were only estimated over 29 provinces for PM_{10}, $PM_{2.5}$ and 28 provinces for NO_2, and O_3. Moreover, other adverse health outcomes beside mortality, such as emergency room visits and hospital admission, are also associated with air pollutants. Therefore, the magnitude of health benefits seems to be underestimated, but finding from this study implied that improvement in air quality contributes to a reduction in the number of deaths associated with air quality improvement. In conclusion, this study highlighted that air quality was improved during the COVID-19 lockdown period over 29 provinces in Thailand, and the number of deaths was consequently avoided. Although, the COVID-19 situation is inappropriate for alleviating public health problems, in turn increasing public health crisis, this study can aid policy makers to make decision regarding air quality management and related interventions to solve the public health burden attributable to air pollution in Thailand.

References

1. Achebak H, Petetin H, Quijal-Zamorano M, Bowdalo D, Pérez García-Pando C, Ballester J (2021) Trade-offs between short-term mortality attributable to NO2 and O3 changes during the COVID-19 lockdown across major Spanish cities. Environ Pollut 286:117220
2. Armstrong BG, Gasparrini A, Tobias A (2014) Conditional Poisson models: a flexible alternative to conditional logistic case cross-over analysis. BMC Med Res Methodol 14(1):1–6
3. Chen Y, Jiao Z, Chen P, Fan L, Zhou X, Pu Y et al (2021) Short-term effect of fine particulate matter and ozone on non-accidental mortality and respiratory mortality in Lishui district, China. BMC Public Health 21(1):1–11. https://doi.org/10.1186/s12889-021-11713-9
4. Flaxman S, Mishra S, Gandy A, Unwin HJT, Coupland H, Mellan TA, et al (2020) Report 13: Estimating the number of infections and the impact of non-pharmaceutical interventions on COVID-19 in 11 European countries. United Kingdom: Imperial College COVID-19 Response Team: 1–35. https://doi.org/10.25561/77731
5. Gasparrini A, Armstrong B, Kenward MG (2010) Distributed lag non-linear models. Stat Med 29(21):2224–2234. https://doi.org/10.1002/sim.3940
6. Gasparrini A, Armstrong B, Kenward MG (2012) Multivariate meta-analysis for non-linear and other multi-parameter associations. Stat Med 31(29):3821–3839. https://doi.org/10.1002/sim.5471
7. Guevara M, Jorba O, Soret A, Petetin H, Bowdalo D, Serradell K et al (2021) Time-resolved emission reductions for atmospheric chemistry modelling in Europe during the COVID-19 lockdowns. Atmos Chem Phys 21(2):773–797. https://doi.org/10.5194/acp-21-773-2021
8. Guo Y, Li S, Tawatsupa B, Punnasiri K, Jaakkola JJK, Williams G (2014) The association between air pollution and mortality in Thailand. Sci Rep 4:5509. https://doi.org/10.1038/srep05509
9. Hastie T, Tibshirani R, Friedman J (2009) Random forests. In: The elements of statistical learning: data mining, inference, and prediction. Springer Series in Statistics, 2nd ed. New York, NY, Springer, pp. 587–624. https://doi.org/10.1007/978-0-387-84858-7_15
10. Junger WL, Ponce de Leon A (2015) Imputation of missing data in time series for air pollutants. Atmos Environ 102:96–104. https://doi.org/10.1016/j.atmosenv.2014.11.049

11. Karimi B, Shokrinezhad B, Samadi S (2019) Mortality and hospitalizations due to cardiovascular and respiratory diseases associated with air pollution in Iran: a systematic review and meta-analysis. Atmos Environ 198:438–447. https://doi.org/10.1016/j.atmosenv.2018.10.063
12. Leong ST, Muttamara S, Laortanakul P (2003) Preliminary study of relationship between outdoor and indoor air pollutant concentrations at Bangkok's major streets. Sci Technol Asia 8(3):29–39
13. Liu Y, Xie S, Yu Q, Huo X, Ming X, Wang J et al (2017) Short-term effects of ambient air pollution on pediatric outpatient visits for respiratory diseases in Yichang city, China. Environ Pollut 227:116–124. https://doi.org/10.1016/j.envpol.2017.04.029
14. Mazumder A, Arora M, Bharadiya V, Berry P, Agarwal M, Shewade HD, et al (2020) SARS-CoV-2 epidemic in India: epidemiological features and in silico analysis of the effect of interventions. F1000Res 9: 315.
15. Meng X, Liu C, Chen R, Sera F, Vicedo-Cabrera AM, Milojevic A et al (2021) Short term associations of ambient nitrogen dioxide with daily total, cardiovascular, and respiratory mortality: multilocation analysis in 398 cities. BMJ 372:n534. https://doi.org/10.1136/bmj.n534
16. Niu Z, Liu F, Yu H, Wu S, Xiang H (2021) Association between exposure to ambient air pollution and hospital admission, incidence, and mortality of stroke: an updated systematic review and meta-analysis of more than 23 million participants. Environ Health Prev Med 26(1):1–14
17. Phosri A, Ueda K, Phung VLH, Tawatsupa B, Honda A, Takano H (2019) Effects of ambient air pollution on daily hospital admissions for respiratory and cardiovascular diseases in Bangkok, Thailand. Sci Total Environ 651:1144–1153. https://doi.org/10.1016/j.scitotenv.2018.09.183
18. Qiu H, Pun VC, Tian L (2018) Short-term effects of fine and coarse particles on deaths in Hong Kong elderly population: an analysis of mortality displacement. Environ Pollut 241:148–154. https://doi.org/10.1016/j.envpol.2018.05.056
19. Sompornrattanaphan M, Thongngarm T, Tantilipikorn P, Kreetapirom P, Johnson Foo BE (2018) The contribution of outdoor fine particulate matter to indoor air quality in Bangkok metropolitan region, Thailand-are indoor dwellers safe? Siriraj Med J 70(4):265–271
20. Son JY, Fong KC, Heo S, Kim H, Lim CC, Bell ML (2020) Reductions in mortality resulting from reduced air pollution levels due to COVID-19 mitigation measures. Sci Total Environ 744:141012. https://doi.org/10.1016/j.scitotenv.2020.141012
21. Tobías A, Carnerero C, Reche C, Massagué J, Via M, Minguillón MC et al (2020) Changes in air quality during the lockdown in Barcelona (Spain) one month into the SARS-CoV-2 epidemic. Sci Total Environ 726:138540. https://doi.org/10.1016/j.scitotenv.2020.138540
22. Wetchayont P, Hayasaka T, Khatri P (2021) Air quality improvement during covid-19 lockdown in bangkok metropolitan, Thailand: effect of the long-range transport of air pollutants. Aerosol Air Qual Res 21(7):1–16. https://doi.org/10.4209/aaqr.200662
23. WHO (2016) Ambient air pollution: a global assessment of exposure and burden of disease. Geneva, Switzerland. https://www.who.int/teams/environment-climate-change-and-health/air-quality-and-health/ambient-air-pollution
24. Yang L, Xu H, Yu S (2021a) Estimating PM2.5 concentrations in Contiguous Eastern Coastal Zone of China using MODIS AOD and a two-stage random forest model. J Atmos Ocean Technol 38(12): 2071–2080. https://doi.org/10.1175/JTECH-D-20-0214.1
25. Yang Z, Yang J, Li M, Chen J, Ou CQ (2021b) Nonlinear and lagged meteorological effects on daily levels of ambient PM2.5 and O3: evidence from 284 Chinese cities. J Clean Prod 278: 123931. https://doi.org/10.1016/j.jclepro.2020.123931
26. Ye T, Guo S, Xie Y, Chen Z, Abramson MJ, Heyworth J et al (2021) Health and related economic benefits associated with reduction in air pollution during COVID-19 outbreak in 367 cities in China. Ecotoxicol Environ Saf 222:112481. https://doi.org/10.1016/j.ecoenv.2021.112481
27. Zheng H, Kong S, Chen N, Yan Y, Liu D, Zhu B, et al (2020) Significant changes in the chemical compositions and sources of PM2.5 in Wuhan since the city lockdown as COVID-19. Sci Total Environ 739: 140000. https://doi.org/10.1016/j.scitotenv.2020.140000
28. Zhou H, Geng H, Dong C, Bai T (2021) The short-term harvesting effects of ambient particulate matter on mortality in Taiyuan elderly residents: A time-series analysis with a generalized additive distributed lag model. Ecotoxicol Environ Saf 207:111235

Dr. Arthit Phosri is an Assistant Professor in the Department of Environmental Health Sciences at Faculty of Public Health, Mahidol University, Thailand. His research interest is characterized by an interdisciplinary perspective on exposure sciences and environmental epidemiology driven by Big Data Analytics (e.g. remote sensing, health informatics), machine learning, and advanced statistical modeling.

Dr. Mathuros Tipayamongkholgul is an Associate Professor in Department of Epidemiology at the Faculty of Public Health, Mahidol University, Thailand. She is an epidemiologist by training and her research focus on spatial epidemiology, and time-series analysis in public health.

Chapter 5
Satellite-Derived Vegetation Indices as a Criterion for Assessing Green Exposure that is Related to Human Health Burdens

Chih-Da Wu and Aji Kusumaning Asri

Abstract Remote sensing technology is used practically in epidemiological studies. This chapter demonstrates how green exposure that is related to health burdens is analyzed. Using satellite-derived data, green exposure is represented as a Normalized Difference Vegetation Index (NDVI). Population and individual studies are used, and several green estimation methods are implemented, wherein the spatial data for the study object is linked to the NDVI. The variables that influence the health burden and the effectiveness of the effect of green exposure on health are determined. This study uses adjusted spatial-statistical methods to determine the relationship between green exposure and health burdens. The overall findings show a significant relationship between green exposure and health burden in both regional and global analyses. This study increases knowledge of remote sensing applications to preserve the environment and health.

Keywords Green exposure · Human health · Regional-global analysis · Satellite-derived vegetation indices

5.1 Introduction

Increasing urbanization in various regions means that the natural environment, such as green spaces, has decreased significantly. Significant changes in land-use allow remote sensing technology and Geographic Information System (GIS) to be used for monitoring and determining spatial–temporal changes in land surfaces, such as green spaces.

C.-D. Wu (✉) · A. K. Asri
Department of Geomatics, National Cheng Kung University, Tainan City, Taiwan
e-mail: chidawu@mail.ncku.edu.tw

C.-D. Wu
National Institute of Environmental Health Sciences, National Health Research Institutes, Miaoli, Taiwan

T.-H. Wen et al. (eds.), *Earth Data Analytics for Planetary Health*, Atmosphere, Earth, Ocean & Space, https://doi.org/10.1007/978-981-19-8765-6_5

In assessing exposure to greenness, many studies use various data and technical methods. Hussein and colleagues analyzed green exposure using two satellite-dsserived vegetation indices from a Moderate Resolution Imaging Spectroradiometer (MODIS), a Normalized Difference Vegetation Index (NDVI) and an Enhanced Vegetation Index (EVI) [1]. An NDVI is widely used to estimate greenness because this index is sensitive to the presence of vegetation and its dynamics [2]. Apart from the MODIS-NDVI, the green index is also measured using other satellite imagery, such as high spatial resolution, WorldView-2 [3], Landsat [4], RapidEye [5] or Sentinel-2 [6].

Satellite data gives information that is broad in scope and continuous, so it is widely used to study environmental conditions that are related to human health. Satellite data is used to assess environmental green exposure for epidemiological observations in a study by [7], which uses a multi-temporal MODIS-NDVI to determine the association between green exposure and depressive symptoms in Korea. NDVI satellite data was also used to study the relationship between green space metrics and health and behavioral outcomes in areas with various buffer sizes [8]. A study in the USA assessed greenness using a NDVI to determine the relationship between green exposure and volatile organic compounds, which can trigger the development of cardiovascular diseases [9].

Satellite-derived greenness that is related to health burdens has been studied extensively but studies only give a local-level analysis. This study presents a developmental concept for ecological study for which satellite data is used to determine the effect of green exposure on health burdens globally. A local-level study with individual-based analysis and uses similar method is also performed.

This study gives a better understanding of how the application of remote sensing technology can promote environment preservation and human health. The first section details how this study is presented. In Sect. 5.2, the methodological framework that is used to determine satellite-derived green exposure that is associated with health burdens is proposed. Several application examples of the proposed method for population and individual-based studies are described in Sect. 5.3. Section 5.4 draws conclusions.

5.2 Methodology

This study uses a conceptual framework that has two main stages: estimating the green exposure and assessing its relationship to the health burden. Remote sensing data is used as the primary vegetation index to determine green exposure. The scale factor is adjusted to determine the valid range values for the vegetation index. In addition, negative values of vegetation index are removed to minimize the effect of water exposure in the greenness estimation. GIS is used to determine green exposure for each study area and to establish the spatial distribution pattern for the individuals' address who are confirmed to have a particular health burden.

Statistical association models coupled with multiple sensitivity tests are used to evaluate the association between exposure and outcome. The consistent results of the sensitivity test demonstrate the robustness of the main statistical models. Furthermore, stratified analyses are also examined to determine the effect of greenness on health for various conditions. The following gives a detailed description of each stage,

5.2.1 Definitions of Vegetation Indices

In terms of the global and regional analyzes that are presented, remote sensing data with an adjustable resolution is used to observe the spatial–temporal conditions for green exposure that are represented by the vegetation index. The vegetation index is widely used to estimate green exposure because it is very sensitive to the presence of vegetation and its dynamics. The best-known is the Normalized Difference Vegetation Index (NDVI) [2]. The NDVI, which is derived from satellite data is a dimensionless index that defines the difference between the reflectance of visible and near-infrared vegetation cover and is used to calculate the green density for an area [10]. The NDVI value varies from −1.0 to +1.0. Positive values represent greener areas because healthy vegetation has low red-light reflectance and high near-infrared reflectance. An NDVI value that is close to zero and a decrease to a negative value indicates non-vegetative features [11]. Formally, the formula for the NDVI is expressed as in Eq. 5.1:

$$NDVI = \frac{(NIR - Red)}{(NIR + Red)} \tag{5.1}$$

where *NIR* is Near-Infrared Regions and *Red* is the spectral reflectance measurements that are acquired in red visible radiation.

This study uses an NDVI that is measured using a Terra Moderate Resolution Imaging Spectroradiometer (Terra-MODIS) sensor Version 6 to calculate the presence of green exposure for each study area [7]. This data is provided by the National Aeronautics and Space Administration and includes the monitoring and measuring of vegetation, plants, biomass production, and the components of greenness [12].

5.2.2 Calculating Green Exposure Using Satellite-Based Vegetation Indices

Previous studies calculate green exposure using remote sensing data from GIS [9, 13]. Geospatial processing software is used to assign NDVI grid cell estimates in a particular study area. The vegetation index is calculated using MODIS satellite imagery by adjusting the NDVI value range from −2000 to 10,000 for raw data to

a real NDVI value of −0.2 to 1.0, as shown in Fig. 5.1. The real NDVI value is calculated by multiplying the NDVI range by the scale factor (0.0001), as shown in Eq. 5.2:

$$NDVI_{Adj} = NDVI_0 \times scale\ factor \tag{5.2}$$

where $NDVI_{Adj}$ is the real value of the NDVI (after adjustment), $NDVI_0$ is *the* NDVI values that are identified as raw data from the image in integer values and the NDVI scale factor is 0.0001.

A negative value of NDVI indicates proximity to water [14] so pixels with negative values are excluded to avoid misclassification bias due to another environmental effect. The effect of water is excluded from this green exposure estimate because a previous study confirmed that water or blue space is beneficial to health [15], so this may affect the analysis. By eliminating negative values in the image, the related grid is identified as areas that do not have vegetation data (not 0) and the grids that have a vegetation index value are retained. Figure 5.2 shows valid NDVI values with a range of 0 to 1.

The valid values for the NDVI (0 to 1) are used to calculate exposure to green vegetation in the study area. For a population-based analysis, green exposure is calculated using the regional boundary (Fig. 5.3-left). Using boundaries at various levels (township, district, city/county, province, or country) allows the green exposure to be calculated using the average index value for all pixels that fall within a specific boundary.

As shown in Fig. 5.3 (right), if the estimated green exposure is calculated on a country level, the spatial distribution for the high vegetation index, which is greener exposure that is clustered in the central states, can be calculated. Countries that are marked with dark green are areas have high vegetation index or high green exposure and light green indicates countries with low green exposure.

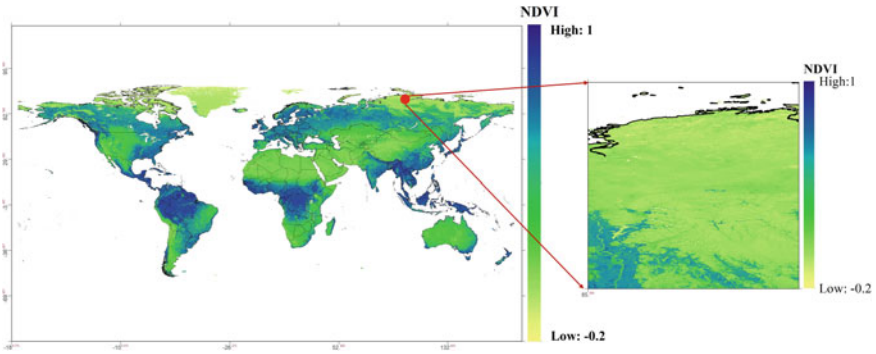


Fig. 5.1 Adjusted NDVI value ranges from −0.2 to 1, October 2012

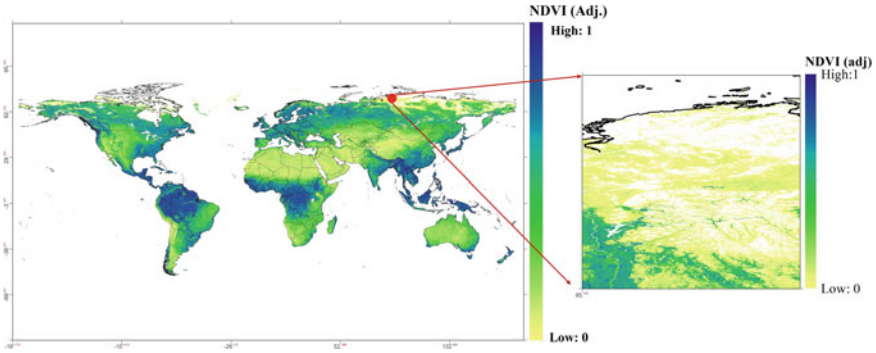


Fig. 5.2 NDVI value for analysis ranges from 0 to 1: October 2012

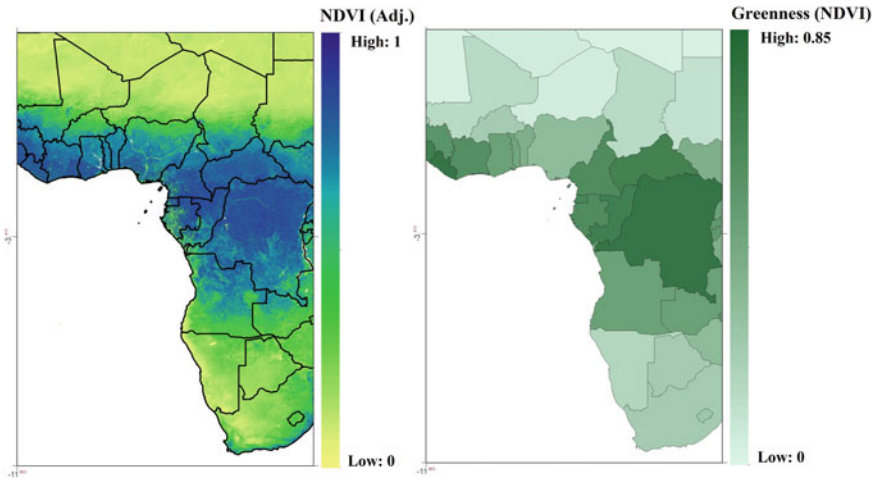


Fig. 5.3 Spatial distribution of estimated green exposure (NDVI): population-based study

As shown in Fig. 5.4, for analysis on an individual basis, estimates of green exposure are based on individual location points (x and y coordinates). In this case, the estimated value for the calculated vegetation index does not use an area boundary: the patient's address or the area with a buffer radius around the individual's site is shown in Fig. 5.4 (left and middle). The buffer rings around each individual's location are used to determine the degree of green exposure to which a subject has access. In the example in Fig. 5.4 (right), someone living in the eastern region has a higher chance of accessing green space than someone in the eastern region.

To calculate the green exposure using the vegetation index, for a population and individual analysis, categorical quartiles of NDVI are used for the main analysis to determine potential nonlinearity and to minimize the effect of outliers in the NDVI distribution.

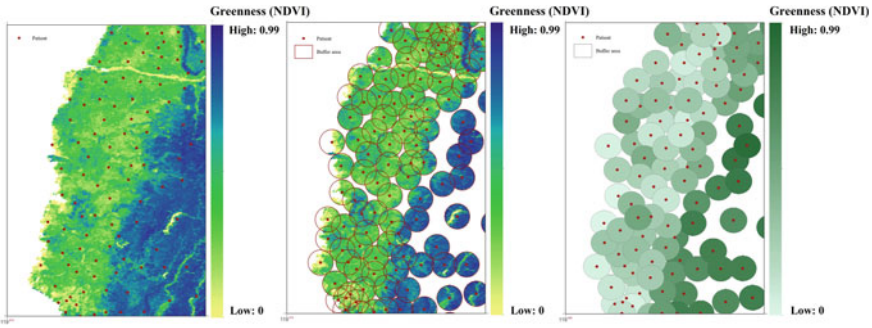


Fig. 5.4 Spatial distribution of estimated green exposure (NDVI): cohort-based study

5.2.3 Assessing Green Exposure in Relation to Health Burden

This section uses the statistical association to determine the relationship between green exposure and health burden. Several potential risk factors that are associated with corresponding health burdens are controlled as fixed-effect variables. Coefficient and risk estimates with 95% confidence intervals are reported and a p-value < 0.05 is considered to indicate statistical significance. The statistical calculations adjusted to the study design are described below.

5.2.3.1 Ecological Study

For this population-based analysis or ecological study, green exposure NDVI is denoted as X (predictor) and health burden as Y (outcome). A linear mixed effect model is used to determine the relationship between exposure and outcome. This model accounts for fixed and random effects and provides a flexible approach for analyzing health outcome [16]. Equation 5.3 shows the general matrix notation for the linear mixed-effects model [17]:

$$\boldsymbol{Y}_{ij} = \boldsymbol{X'}_{ij}\boldsymbol{\beta} + \boldsymbol{Z'}_{ij}\boldsymbol{b}_i + \boldsymbol{\epsilon}_{ij} \tag{5.3}$$

where Y_{ij} is the link function or $(N \times 1)$ column vector, the outcome variables (health burden), X'_{ij} is a $(N \times p)$ matrix for the p predictor variables (green exposure NDVI), β is a $(p \times 1)$ column vector for the fixed-effect coefficients, Z'_{ij} is a $(N \times p)$ matrix for the q random effects, b_i is a $(q \times 1)$ column vector for the random-effect coefficients and $€_{ij}$ is a $(N \times 1)$ column vector for the residuals, which is the part of Y_{ij} that is not explained by the model. The random effects, b_i, are assumed to have a multivariate normal distribution with a mean of zero and a covariance matrix G, so the equation becomes (Eq. 5.4):

$$\boldsymbol{b_i} \sim \boldsymbol{N}(0, \boldsymbol{G}) \tag{5.4}$$

where b_i is distributed as normal with a mean of zero and a covariance matrix G. The mean is zero because the fixed effects, including the fixed effect intercept, are calculated directly so random effect complements are modeled as deviations from the fixed effect. G is a 2×2 variance–covariance matrix for the random effects with unique components. For random intercepts and random gradients, then variance–covariance matrix G is expressed as (Eq. 5.5):

$$\boldsymbol{G} = \begin{bmatrix} \boldsymbol{\sigma}^2_{\boldsymbol{int}} & \boldsymbol{\sigma}^2_{\boldsymbol{int,slope}} \\ \boldsymbol{\sigma}^2_{\boldsymbol{int,slope}} & \boldsymbol{\sigma}^2_{\boldsymbol{slope}} \end{bmatrix} \tag{5.5}$$

In terms of residual errors, ϵ_{ij}, are assumed to have a multivariate normal distribution with a mean of zero and a covariance matrix R_i, so the equation becomes:

$$\boldsymbol{\epsilon_{ij}} \sim \boldsymbol{N}(0, \boldsymbol{R_i}) \tag{5.6}$$

where ϵ_{ij} is distributed as a normal with a mean of zero and a covariance matrix R_i, $\mathrm{R}_i = \sigma^2 I$ and I is the identity matrix. If conditional independence is assumed, $R_i = \sigma^2 I$, the errors ϵ_{ij} within a subject are uncorrelated with homogeneous variance, so a structured model for R_i is assumed, such as an autoregressive model. The "area ID" is a clustered unit that is treated as a random intercept and is used to minimize the temporal correlation due to repeated measurements within the area.

To validate the model, several sensitivity tests are used to determine the robustness of the results. Positive–negative control analyses are also used to determine the strength of causal inference for an exposure-outcome association. Two methods are used: positive–negative outcome control and positive–negative exposure control. A positive–negative outcome control determines whether using the same exposure and replacing studied outcomes with other outcomes yields consistent results. A positive–negative exposure control determines whether using studied outcomes and replacing exposure variables generates a consistent association. Using this validation, causal inference is assumed to be acceptable if the positive control analysis shows an equally significant relationship to that in the main model and the negative control analysis shows an insignificant relationship.

5.2.3.2 Cohort Study

In terms of the analysis at the individual level, Cox proportional hazard regression models are used to determine the association between green exposure and health burden and to calculate the hazard ratio (HR) with 95% confidence intervals (CI). For this model, the proportionality of hazards is determined by testing interactions using the log of time [18]. The formula for this survival time model is:

$$h_t = h_0(t)\mathbf{exp}(b_1 X_1 + b_2 X_2 + \cdots + b_n X_n) \tag{5.7}$$

where $h(t)$ is the expected hazard at time t and $h_0(t)$ is the baseline hazard and represents the hazard when all independent variables X_1, X_2, X_n are equal to 0 (*zero*). The estimated hazard $h(t)$ or the suffering rate for the next event is the product of the baseline hazard $h_0(t)$ and the exponential function for the linear combination of the predictors. Therefore, the predictor (green exposure NDVI) has a multiplicative or proportional effect on the estimated hazard (health burden).

Similar to the ecological study, as part of the validation, sensitivity tests using different covariates are performed to determine the robustness of the developed models.

5.3 Example Cases: Green Exposure Related to Health Burdens

Three examples of studies at the population and individual levels to determine the relationship between green exposure derived from satellite information and the burden on human health are presented. The relationship between greenness and cardiovascular diseases and the all-cause mortality will be described in detail as follows.

5.3.1 The Relationship Between Greenness and IHD and Stroke: Global Ecological Study

Since a global analysis gives stronger scientific evidence, an ecological-based study was undertaken by Asri et al. (2020) to determine the association between greenness and cardiovascular burdens, in terms of ischemic heart disease (IHD) and stroke globally [19]. Cardiovascular diseases are the health burden of concern because the Global Burden of Disease study in 2016 confirmed that IHD and stroke are the leading causes of mortality in the global population. Using country-level datasets, this study collects data from 183 countries for 2000, 2010, 2015, and 2016. Disability-adjusted life year (DALY) loss data from the World Health Organization is used to represent the health burdens due to IHD and stroke. Because the available data cover all age groups, the data of the population aged under 5 years are excluded due to the low burden and ineffectiveness of detection of cardiovascular symptoms in this age group.

This study calculated the ratio for DALY and the spatial distribution of DALY loss due to IHD and stroke and the results are shown in Fig. 5.5. Based on the spatial plotting of DALY loss, for which darker areas indicate a higher risk of burden, the global IHD risk (Fig. 5.5a) is identified as high in Europe and low in African countries.

Stroke (Fig. 5.5b) is identified as a high risk in European and Asian countries and low in Australia and some countries in America and Africa.

The Terra-MODIS NDVI with $1 \times 1\ km^2$ spatial resolution (in 2000–2016) is used to determine the presence of greenness in each country. The monthly NDVI for this study is MOD13A3 Version 6 and pixels with negative values are omitted to avoid bias estimates because of water effect on the image [14]. To determine the relationship between exposure and outcome, one of the mixed-effect algorithms, the generalized linear mixed model with a penalized quasi-likelihood, is used with a sensitivity analysis and validity analysis. In terms of potential risk factors, several covariates such as demographic characteristics, socioeconomic factors, lifestyle behaviors, socioculture, healthcare statuses, comorbidities, air pollution–$PM_{2.5}$, and meteorological factors are controlled for model adjustment.

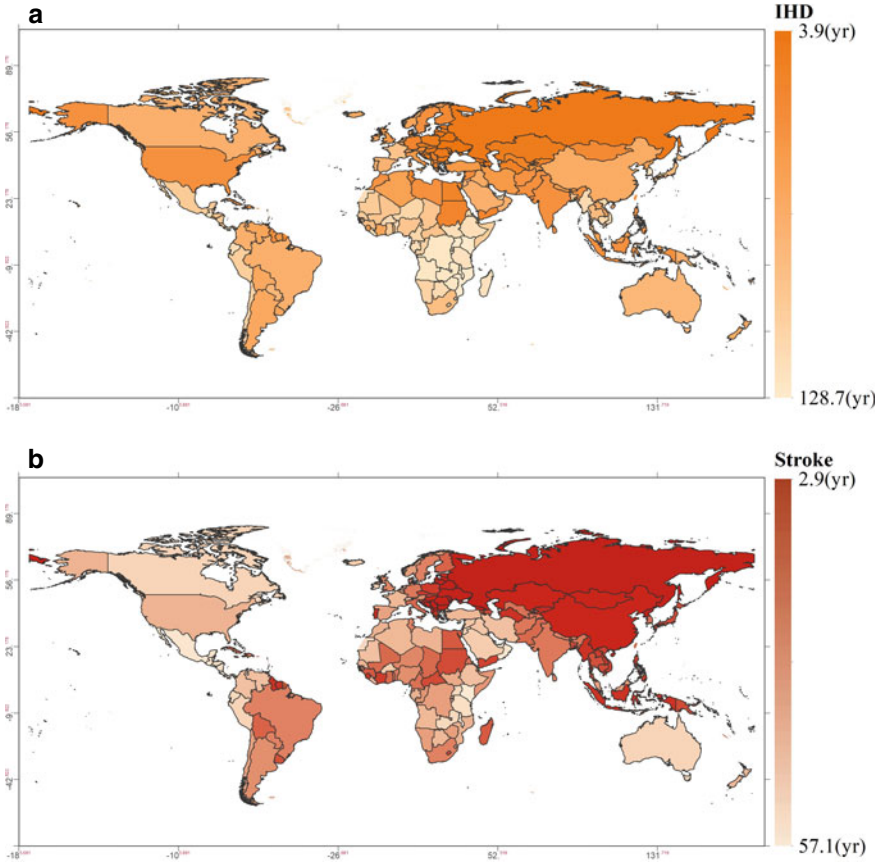


Fig. 5.5 Map of spatial distribution of DALY loss due to **a** ischemic heart disease (IHD) and **b** stroke for all countries worldwide (modified from Asri et al. [19])

The results (Table 5.1) show that the association model predicts a significantly negative association between greenness and DALY loss due to IHD-stroke, with a coefficient estimation per unit increment for the NDVI from 0 to 1 of −11.25 (95% CI: −16.77, −5.72) and −4.387 (95% CI: −7.93, −0.09), respectively. These results show that greenness significantly reduced the health burden due to IHD and stroke. The six sensitivity test models show a consistently significant negative association between greenness and IHD and stroke for different risk factor settings.

As part of the validation tests, the results of the positive–negative control analysis agree with the assumptions for this study (Table 5.2). Positive outcome control shows a significantly negative association between greenness and falls (p-value < 0.05) but does not achieve a statistically significant result for the relationship between greenness and road injury (negative outcome control). Positive exposure control shows a significantly positive association between IHD and $PM_{2.5}$ and an insignificant association between IHD and wind speed for the negative exposure control analysis.

Knowing that greenness varies by country, this study analyzed the four regional groups based on the quartile of NDVI and coded Q1 as the lowest green areas. As shown in Table 5.3, after adjusting for covariates, the areas with the highest quartile of NDVI have a significantly negative association with IHD, compared to the lowest

Table 5.1 Association models between greenness and IHD-stroke in the DALY changes per unit increment of NDVI from 0 to 1 (modified from Asri et al. [19])

	IHD		Stroke	
	β coefficient of NDVI (95% CI)	p-value	β coefficient of NDVI (95% CI)	p-value
Main model[a]	−11.25 (−16.77, −5.72)	<0.001	−4.39 (−7.93, −0.09)	<0.05
Sensitivity test				
Model 1 [b]	−11.81 (−17.73, −5.88)	<0.001	−5.27 (−9.06, −1.48)	<0.01
Model 2 [c]	−11.87 (−17.73, −5.92)	<0.001	−5.45 (−9.27, −1.64)	<0.01
Model 3 [d]	−11.87 (−17.83, −5.92)	<0.001	−5.73 (−9.57, −1.89)	<0.01
Model 4 [e]	−11.53 (−17.49, −5.58)	<0.001	−4.96 (−8.81, −1.10)	<0.01
Model 5 [f]	−10.23 (−15.99, −4.46)	<0.001	−4.10 (−7.71, −0.49)	<0.05
Model 6[g]	−10.72 (−16.29, −5.14)	<0.001	−4.10 (−7.68, −0.52)	<0.05
Model 7[h]	−6.10 (−10.41, −1.80)	<0.01	−1.62 (−5.52, −0.002)	<0.05

CI = confidence interval; IHD = ischemic heart disease
[a]Control covariates include population size, sex (%, female), age, economic status, smoking, alcohol, education, $PM_{2.5}$, BMI, cholesterol, blood pressure and healthcare expenditure
[b]Adjusted for population size, sex (%, female) and age
[c]Adjusted for Model 1 and $PM_{2.5}$
[d]Adjusted for Model 2 and economic status
[e]Adjusted for Model 3 and alcohol
[f]Adjusted for Model 4, smoking, and education
[g]Adjusted for Model 5, BMI, cholesterol and systolic blood pressure
[h]Controlled for all covariates and data from Eastern Europe and Central Asia/Russia countries is eliminated

Table 5.2 Positive–negative control analysis (modified from Asri et al. [19])

Model	β coefficient (95% CI)	*p*-value	β coefficient (95% CI)	*p*-value
	IHD ~ $PM_{2.5}$ (*positive*)		IHD ~ Wind speed (*negative*)	
Exposure control model [a]	0.11 (0.04, 0.17)	<0.01	−1.13 (−0.36, 2.61)	0.15
	Green ~ falls (*positive*)		Green ~ Injury (*negative*)	
Outcome control model [b]	−1.62 (−2.64, −0.60)	<0.001	−0.56 (−1.38, 2.50)	0.58

CI = confidence interval; IHD = ischemic heart disease
[a]Control covariates include population size, sex (%, female), age, economic status, smoking, alcohol, education, BMI, cholesterol, systolic blood pressure and healthcare expenditure
[b]Control covariates include population density, sex (%, female), age, economic status, smoking, alcohol, education, $PM_{2.5}$ and healthcare expenditure

Table 5.3 Coefficient estimations of greenness by quartile attributed to DALY loss due to IHD and stroke in multivariable adjusted models (modified from Asri et al. [19])

NDVI	IHD		Stroke	
	β coefficient of NDVI (95% CI)	*p*-value	β coefficient of NDVI (95% CI)	*p*-value
Quartile 1 (*low*)	*Reference*		*Reference*	
Quartile 2	−3.64 (−8.94, 1.67)	0.19	−0.12 (−0.36, 0.12)	0.34
Quartile 3	−7.80 (−13.17, −2.44)	<0.01	−0.01 (−0.24, 0.21)	0.84
Quartile 4 (*high*)	−9.53 (−15.21, −3.85)	<0.001	−0.07 (−0.32, 0.18)	0.57

CI = confidence interval; IHD = ischemic heart disease
Control covariates include population size, sex (%, female), age, economic status, smoking, alcohol, education, $PM_{2.5}$, BMI, cholesterol, systolic blood pressure and healthcare expenditure

quartile (coef., 95% CI = −9.53, −15.21 to −3.85). There is a linear relationship between NDVI estimates for IHD and the coefficient increases from −9.53 (Q4), −7.80 (Q3), to −3.64 (Q2). This study also determines a negative correlation between greenness and stroke for all classes, although this result is not significant.

This study uses stratified analyses for different levels of demographic factors (Fig. 5.6). In terms of DALY loss based on sex, the effect of greenness on IHD and stroke for both females and males indicate a significant negative association (p-value < 0.05). There is no notable difference in the effect of greenness between the sexes in terms of reducing the health burden due to IHD and stroke.

Stratification by economic status confirms that greenness has a significant negative correlation to IHD in low and middle-income countries and is negatively correlated at all economic levels for stroke, although this result is insignificant. Stratification by age shows a significant negative association for IHD and greenness for all studied populations. A negative association is also identified for all age groups for stroke and this result is significant for ages 15–29, 50–59, and 60–69 years.

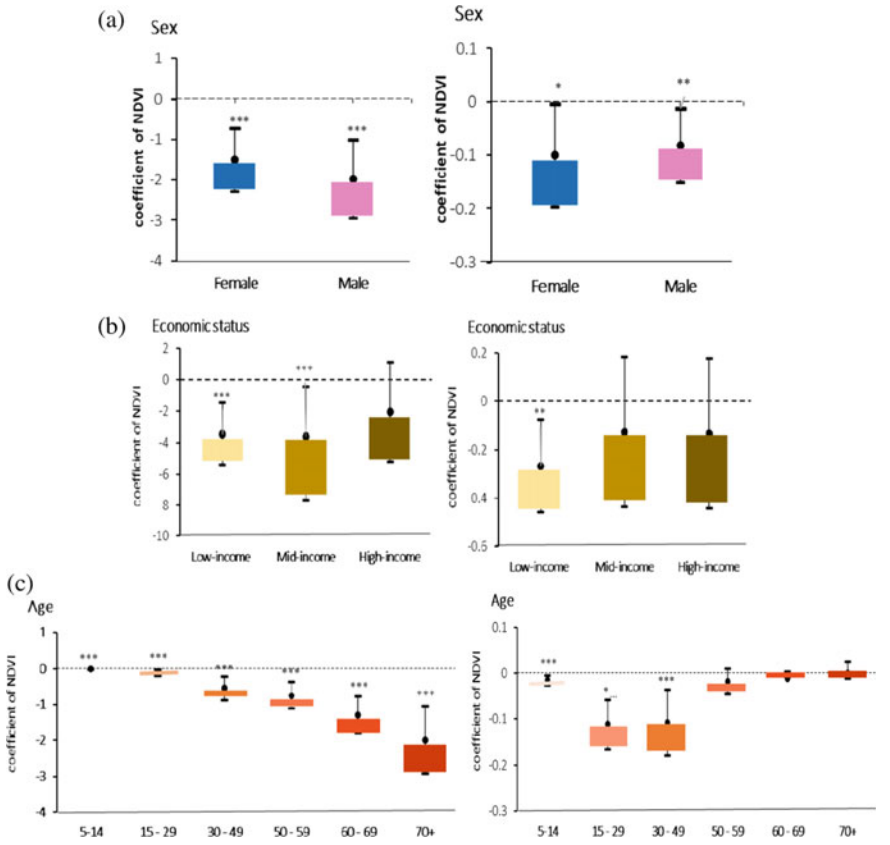


Fig. 5.6 Subgroup analysis, stratified by **a** sex, **b** age group and **c** economic status for greenness in relation to DALY loss due to IHD and stroke (modified from Asri et al. [19])

5.3.2 The Relationship Between Residential Green Space and Ischemic Stroke: A Cohort Study

The proposed method is used for an individual-level analysis of the results of a study by Wilker and colleagues (2014) to determine the link between green space and mortality following ischemic stroke in the greater Boston metropolitan area [18]. 1763 patients of 21 years or older who were admitted to Beth Israel Deacon Medical Center (BIDMC) between April 1999–October 2008 and who resided within 40 km from the hospital are studied. The medical records during hospitalization are used to distinguish medical history, presenting symptoms and demographic information for patients. To determine the outcome, the number of mortalities until June 2012, as registered in the Social Security Death Index (SSDI) (929 deaths among 1645 patients) are used. MODIS NDVI with 250-m spatial resolution as a gridded product vegetation index is used to determine the residential green space (Fig. 5.7). The

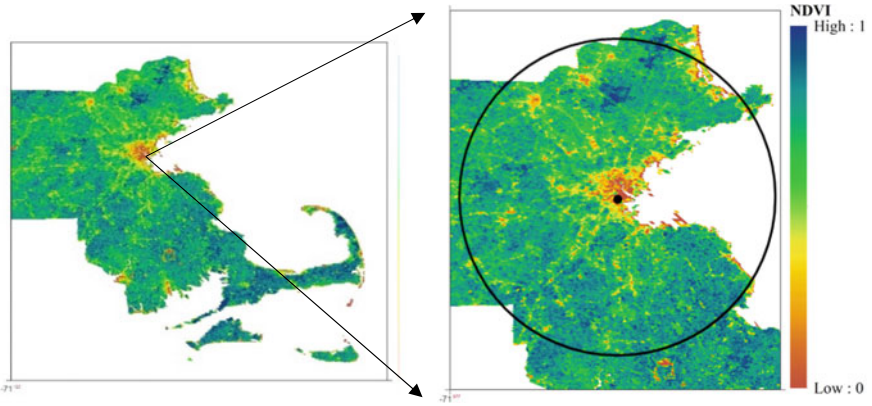


Fig. 5.7 Green space calculated using the NDVI for the studied area: July 2012

average NDVI levels for July from 2000 to 2012 (two July NDVI measures per year) are calculated. This produces 26 images that are used to predict the amount of green space. Grids around patient addresses that are surrounded by water areas are excluded from the analysis.

For the statistical analysis, Cox proportional hazard algorithm is used to calculate the hazard ratio (HR) for the association between green space and mortality following ischemic stroke. Several pertinent covariates, including age, sex, race, Hispanic ethnicity, smoking status, history of coronary artery disease, stroke, atrial fibrillation, heart failure, diabetes, dyslipidemia, hypertension, % of people older than 25 without a high school diploma, median household income and the natural logarithm of the distance to a major roadway are used for the model adjustments. Sensitivity analyses using various covariate settings are used as validity tests to determine the robustness of the models.

The association models for green space and ischemic stroke mortality are shown in Table 5.4. After adjustment for age and sex, Model 1 shows a significant relationship between green space and ischemic stroke (HR = 0.87, 95% CI: 0.78–0.97, p-value < 0.05), but this result is insignificant when it is controlled for all covariates.

The quartile is designated as the primary analysis to evaluate potential nonlinearity and to minimize the outlier effects in the NDVI distribution and Table 5.5 shows the association models for green space and mortality following ischemic stroke, as stratified by quartile of NDVI. After further adjustment for all potential covariates (Model 2), the hazard ratio for patients living in the highest quartile of green space is 0.80 (95% CI: 0.65–0.99, p-value < 0.05). This result shows that patients living in the highest quartile of green space have a 20% lower mortality rate due to ischemic stroke than those living in the lowest green space. Living in the third quartile is also associated with a lower mortality rate (HR = 0.79, 95% CI: 0.65–0.96, p-value < 0.05. The robustness of the association model is confirmed by consistent results for the sensitivity test for Model 1.

Table 5.4 Hazard ratio for the association between green space (NDVI) and mortality following ischemic stroke, with NDVI as a linear continuous term (modified from Wilker et al. [18])

Model	HR (95% CI)	*p*-value
Main model [a]	0.92 (0.81, 1.05)	0.22
Sensitivity test		
Model 1[b]	0.87 (0.78, 0.97)	<0.05
Model 2[c]	0.91 (0.80, 1.03)	0.12

CI = confidence interval; HR = hazard ratios
[a]Controlled for age, sex, race, Hispanic, smoking, history of coronary artery disease, history of stroke, atrial fibrillation, heart failure, diabetes, dyslipidemia, hypertension, % of people older than 25 without a high school diploma, median household income and the log of the distance to a road
[b]Controlled for age and sex
[c]Controlled for age, sex, race, Hispanic, smoking, history of coronary artery disease, history of stroke, atrial fibrillation, heart failure, diabetes, dyslipidemia, hypertension, % of people older than 25 without a high school diploma and median household income
[d]Scaled to an interquartile difference in NDVI (0.22)

Table 5.5 Hazard ratio by quartile for the association between green space (NDVI) and mortality following ischemic stroke (modified from Wilker et al. 2014)

NDVI	Model 1[a]		Model 2[b]	
	HR (95% CI)	*p*-value	HR (95% CI)	*p*-value
Quartile 1 (*low*)	*Reference*		*Reference*	
Quartile 2	0.90 (0.75, 1.08)	–	0.91 (0.76, 1.10)	–
Quartile 3	0.79 (0.65, 0.94)	<0.05	0.79 (0.65, 0.96)	<0.05
Quartile 4 (*high*)	0.77 (0.64, 0.92)	<0.05	0.80 (0.65, 0.99)	<0.05

CI = confidence interval; HR = hazard ratio
[a]Controlled for age and sex
[b]Controlled for age, sex, race, Hispanic, smoking, history of coronary artery disease, history of stroke, atrial fibrillation, heart failure, diabetes, dyslipidemia, hypertension, % of people older than 25 without a high school diploma, median household income and the log of the distance to a road
[c]Scaled to an interquartile difference in NDVI (0.22)

5.3.3 The Relationship Between Residential Greenness and Mortality in the Elderly: A Cohort Study

Using a cohort study design, a peer study by Ji and colleagues conducted an analysis of greenness related to all-cause mortality for the elderly in China [20]. This study uses five pooled waves from the China Longitudinal Healthy Longevity Survey (CLHLS) data from 2000 to 2014 as the determinants of healthy longevity in the older population. 23,754 individuals aged $\geq$ 80 years in 2000 and 65–79 years from 2002 are included in the final sample and these were interviewed to gather personal-health data.

Interviews were also used to identify demographic factors, lifestyle behavior, socioeconomic status, physical activity and social-leisure activity. This study gathered all cause-specific mortality information for 2476 of 18,948 deaths.

To determine exposure, the NDVI from MODIS with a 250-m spatial resolution and 16-day temporal resolution is used to measure residential greenness (Fig. 5.8). Each residential address (longitude and latitude) is plotted and linked to the NDVI imagery and the greenness for the 250 and 1250 m buffer radii is calculated. The 250 m radius is a measure of the green exposure surrounding the residence and the 1250 m buffer represents greenness within walking distance.

Accounting for seasonal variations, this study determines the NDVI values for each of the four seasons from February 2000 to October 2014 (January, April, July, October). As measurement metrics, contemporaneous NDVI (value at the individual's residential address at the time closets to the incident), cumulative NDVI (mean value of all NDVI measured over the follow-up period) and changes in NDVI in the residential area over the course of the follow-up period) are used to calculate green exposure. Quartiles of NDVI are used for the primary analysis.

A Cox proportional hazard model is used to calculate the mortality hazard ratio (HR) and survival time is calculated from the first to the last date of interview (up to 2014). Potential covariates, such as age, sex, marital status, ethnicity, socioeconomic, geographic region, social and leisure activity, smoking status, alcohol consumption and physical activity, are accounted for in the model adjustments. Adjustments for age and full covariates are used as a sensitivity test to determine the robustness of the models.

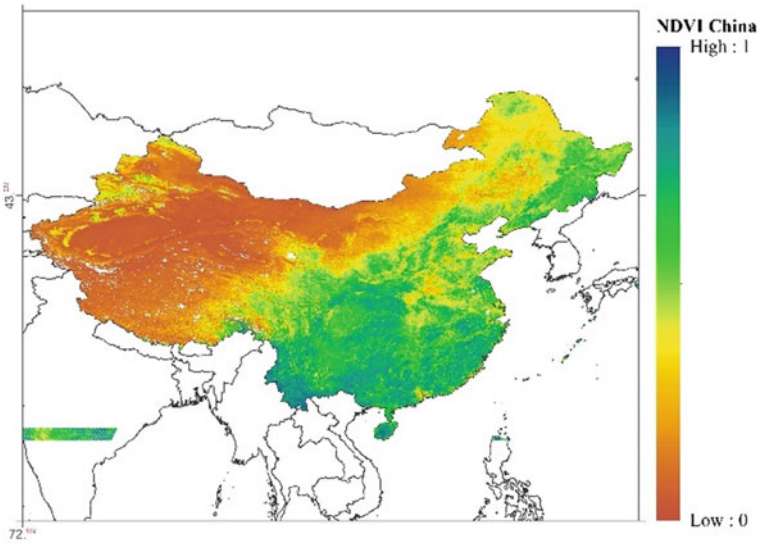


Fig. 5.8 NDVI predictions for the studied area

When covariates are controlled, Table 5.6 shows the hazard ratio with a 95% confidence interval for the association between greenness and mortality for age-adjusted and fully adjusted models. For fully adjusted models, the results show that individuals residing in an area with a 250-m radius that falls within the highest quartile of contemporaneous NDVI (Q4) have 27% (HR = 0.73, 95% CI: 0.70–0.76, *p*-value < 0.001) lower mortality than those in the lowest quartile. Using a radius of 1250 m, the results remain consistent, and 30% lower mortality is confirmed for those living in areas with high greenness, compared to those living in areas with low greenness (HR = 0.70, 95% CI: 0.67–0 0.74, *p*-value < 0.001).

Using cumulative values for the NDVI (Table 5.7), this study shows that there is no significant association between mortality for the highest quartile of greenness for either a 250-m radii (HR = 1.05, 95% CI: 1.01–1.10, *p*-value 0.24) or a 1250-m radii (HR = 1.05, 95% CI: 1.00–1.10, *p*-value 0.17), possibly because economic development that is associated with a reduction in green space, such as infrastructure and roads, may give increased access to healthcare facilities, which indirectly reduces the risk of mortality. Deaths among the elderly can also be caused by acute, rather than long-term exposure, as this group spends more time at home.

This study also constructs association models for all-cause mortality and changes in greenness (NDVI) over time and the results are shown in Table 5.8. For individuals living in areas with a significant increase in NDVI in a 250-m or 1250-m radius, the results show that there is an increase in mortality, compared to those living in areas with decreased NDVI. The effect of modification based on socioeconomic factors is detailed in this study. The results show that greenness has a greater protective effect in females, the financially independent and those who exercise than in males, the financially dependent and those who do not exercise.

Table 5.6 Hazard ratio for the association between residential greenness and all-cause mortality for the CLHLS, using contemporaneous NDVI (modified from Ji et al. [20])

NDVI	HR (95% CI) 250-m buffer		HR (95% CI) 1250-m buffer	
	Model 1[a]	Model 2[b]	Model 1[a]	Model 2[b]
Q1 *(low)*	*Reference*		*Reference*	
Q2	0.92 (0.89, 0.96)	0.90 (0.87, 0.94)	0.91 (0.87, 0.94)	0.89 (0.85, 0.92)
Q3	0.91 (0.87, 0.94)	0.88 (0.84, 0.92)	0.87 (0.83, 0.90)	0.83 (0.80, 0.87)
Q4 *(high)*	0.76 (0.73, 0.79)	0.73 (0.70, 0.76)	0.74 (0.71, 0.77)	0.70 (0.67, 0.74)
p-value	<0.001	<0.001	<0.001	<0.001

CI = confidence interval; HR = hazard ratio; Q = quartile
[a]Age-adjusted hazard ratio
[b]Fully adjusted hazard ratio

Table 5.7 Hazard ratio for the association between residential greenness and all-cause mortality in the CLHLS using cumulative NDVI (modified from Ji et al. [20])

NDVI	HR (95% CI) 250-m buffer		HR (95% CI) 1250-m buffer	
	Model 1[a]	Model 2[b]	Model 1[a]	Model 2[b]
Q1 (*low*)	*Reference*		*Reference*	
Q2	1.06 (1.02,1.10)	1.04 (1.00, 1.09)	1.07 (1.02,1.11)	1.05 (1.01,1.11)
Q3	0.99 (0.95, 1.03)	0.97 (0.93, 1.02)	1.01 (0.97, 1.05)	1.00 (0.96, 1.05)
Q4 (*high*)	1.05 (1.01, 1.10)	1.05 (1.01, 1.10)	1.05 (1.01, 1.09)	1.05 (1.00, 1.10)
p-value	0.10	0.24	0.09	0.17

CI = confidence interval; HR = hazard ratio; Q = quartile
[a]Age-adjusted hazard ratio
[b]Fully adjusted hazard ratio

Table 5.8 Hazard ratio for the association between residential greenness and all-cause mortality in the CLHLS using cumulative NDVI (modified from Ji et al. 2019)

	HR (95% CI) 250-m buffer		HR (95% CI) 1250-m buffer	
	Model 1[a]	Model 2[b]	Model 1[a]	Model 2[b]
Significant decrease (3174 participants)	*Reference*		*Reference*	
Not significant change (14,757 participants)	0.99 (0.94, 1.03)	0.99 (0.95, 1.04)	0.97 (0.93,1.01)	0.99 (0.94, 1.03)
Significant increase (5823 participants)	0.98 (0.94, 1.03)	0.98 (0.94, 1.03)	0.98 (0.93, 1.03)	0.98 (0.93, 1.03)
p-value	>0.10			

CI = confidence interval; HR = hazard ratio
[a]Age-adjusted hazard ratio
[b]Fully adjusted hazard ratio

5.4 Conclusions

Humans and nature are two fundamentally interrelated elements of life and nature greatly affect human welfare. In this era, the use of remote sensing to explore natural exposures such as greenness is a form of application that represents how technological advances can provide information that is useful welfare in terms of health.

This study uses satellite-derived data to determine greenness, in terms of population-based and individual-based exposure. Using the boundary of areas or individual location coordinates, greenness is calculated and its effect on a particular health burden is determined.

This study framework, which uses several case examples, gives an overview of the effect of green exposure on health and shows how spatial and statistical methods are used for epidemiological studies. However, the study has some limitations. Due to the lack of data, this study does not consider specific green exposures, such as

biodiversity, green land cover, or vegetation type (field, grassland, forest, scrub, park, garden, cultivation, etc.). At the individual level, the amount of time that is spent in green spaces is also not known with certainty.

References

1. Hussein SO, Kovács F, Tobak Z (2017) Spatiotemporal assessment of vegetation indices and land cover for Erbil City and its surrounding using MODIS imageries. J Environ Geogr 10(1–2):31–39
2. Zhang X, Friedl MA, Schaaf CB (2006) Global vegetation phenology from moderate resolution imaging spectroradiometer (MODIS): evaluation of global patterns and comparison with in situ measurements. J Geophys Res Biogeosci 111(4)
3. Nouri H, Beecham S, Anderson S, Nagler P (2014) High spatial resolution worldview-2 imagery for mapping NDVI and its relationship to temporal urban landscape evapotranspiration factors. Remote Sens 6(1):580–602
4. Nguyen TH, Jones S, Soto-Berelov M, Haywood A, Hislop S (2020) Landsat time-series for estimating forest aboveground biomass and its dynamics across space and time: a review. Remote Sens 12:98
5. Kross A, McNairn H, Lapen D, Sunohara M, Champagne C (2015) Assessment of RapidEye vegetation indices for estimation of leaf area index and biomass in corn and soybean crops. Int J Appl Earth Obs Geoinf 34(1):235–248
6. Devaux N, Crestey T, Leroux C, Tisseyre B (2019) Potential of Sentinel-2 satellite images to monitor vine fields grown at a territorial scale. OENO One 53(1):51–58
7. Song H, Lane KJ, Kim H, Kim H, Byun G, Le M et al (2019) Association between urban greenness and depressive symptoms: evaluation of greenness using various indicators. Int J Environ Res Public Health 16(2):173
8. Su JG, Dadvand P, Nieuwenhuijsen MJ, Bartoll X, Jerrett M (2019) Associations of green space metrics with health and behavior outcomes at different buffer sizes and remote sensing sensor resolutions. Environ Int 126:162–707
9. Yeager R, Riggs DW, DeJarnett N, Srivastava S, Lorkiewicz P, Xie Z et al (2020) Association between residential greenness and exposure to volatile organic compounds. Sci Total Environ 707:135435
10. Weier J, Herring D (2000) Measuring vegetation (NDVI & EVI). NASA Earth Observatory, Washington DC
11. Wu C-D, Chen Y-C, Pan W-C, Zeng Y-T, Chen M-J, Guo YL et al (2017) Land-use regression with long-term satellite-based greenness index and culture-specific sources to model $PM_{2.5}$ spatial-temporal variability. Environ Pollut 224:148–157
12. Gascon M, Cirach M, Martínez D, Dadvand P, Valentín A, Plasència A et al (2016) Normalized difference vegetation index (NDVI) as a marker of surrounding greenness in epidemiological studies: the case of Barcelona city. Urban Forest Urban Green 19:88–94
13. Kim S, Kim H, Lee JT (2019) Interactions between ambient air particles and greenness on cause-specific mortality in seven Korean metropolitan cities, 2008–2016. Int J Environ Res Public Health 16(10)
14. Wheeler BW, White M, Stahl-Timmins W, Depledge MH (2012) Does living by the coast improve health and wellbeing? Health Place 18(5):1198–1201
15. Britton E, Kindermann G, Domegan C, Carlin C (2020) Blue care: a systematic review of blue space interventions for health and wellbeing. Health Promot Int [Internet] 35(1):50–69
16. Leung WTV, Tam YTT, Pan W-C, Wu C-D, Lung S-CC, Spengler JD (2019) How is environmental greenness related to students' academic performance in English and Mathematics? Landsc Urban Plan 181:118–124

17. Laird NM, Ware JH (1982) Random-effects models for longitudinal data. Biometrics 38(4):963
18. Wilker EH, Wu C-D, McNeely E, Mostofsky E, Spengler JD, Wellenius GA et al (2014) Green space and mortality following ischemic stroke. Environ Res 133:42–48
19. Asri AK, Yu C-P, Pan W-C, Guo YL, Su H-J, Lung S-CC et al (2020) Global greenness in relation to reducing the burden of cardiovascular diseases: ischemic heart disease and stroke. Environ Res Lett 15(12):124003
20. Ji JS, Zhu A, Bai C, Wu CD, Yan L, Tang S et al (2019) Residential greenness and mortality in oldest-old women and men in China: a longitudinal cohort study. Lancet Planet Health 3(1):e17–25

Dr. Chih-Da Wu is an Associate Professor at the Department of Geomatics, National Cheng Kung University. Currently, he is also serving as an Adjunct Associate Research Fellow at the National Health Research Institutes, Taiwan. Pursuing the application of Geo-AIoTs, he has expertise in the field of air pollution modeling, greenspace assessment, and urban heat island assessment which then link to health effects.

Aji Kusumaning Asri is a Ph.D. student at the Department of Geomatics, National Cheng Kung University, Taiwan. Lifting applications of geographic information system and remote sensing technologies on environmental epidemiology studies, her expertise focus on investigating the impact of the environment exposures in terms of greenspace and air pollution on human health.

Chapter 6
Five Common Myths About Land Use Change and Infectious Disease Emergence

Luis Fernando Chaves, Chystrie A. Rigg, Mariel D. Friberg, Milixa Perea, Lisbeth A. Hurtado, Nicole L. Gottdenker, and Luke R. Bergmann

Abstract The literature about emerging infectious diseases is often filled with assumptions that are not fully substantiated or not supported by more relational research. Here we present five common myths in research that has linked land use change with the emergence of infectious diseases. Our intention is to raise awareness about points that deserve special attention when contextualizing observations about land use change and its internal relations to the emergence of new infectious diseases.

Keywords Population · Deforestation · Modeling · Land sparing · Pathogen spillover

6.1 Pervasive Social Constructs in Inferences About Land Use Change and Disease Emergence

The concepts of modern population sciences in the western world are interdisciplinary in their sources, including substantial influence by the development of ecology and evolutionary biology [1, 2]. As such, some abstractions that were useful

L. F. Chaves (✉) · C. A. Rigg · M. Perea · L. A. Hurtado
Instituto Conmemorativo Gorgas de Estudios de La Salud, Apartado Postal, 0816-02593 Ciudad de Panamá, Panamá
e-mail: lfchavs@gmail.com

M. D. Friberg
Earth System Science Interdisciplinary Center (ESSIC), University of Maryland, College Park, MD 20740-3823, USA

N. L. Gottdenker
Department of Veterinary Pathology, University of Georgia, Athens, GA 30602, USA

Center for the Ecology of Infectious Diseases, University of Georgia, Athens, GA 30602, USA

L. R. Bergmann
Department of Geography, University of British Columbia, Vancouver, BC V6T 1Z2, Canada

L. F. Chaves
Department of Environmental and Occupational Health, School of Public Health, Indiana University, Bloomington, IN 47408, USA

T.-H. Wen et al. (eds.), *Earth Data Analytics for Planetary Health*, Atmosphere, Earth, Ocean & Space, https://doi.org/10.1007/978-981-19-8765-6_6

to help build a common ground for population sciences reflect biases and misconceptions that got ingrained as inherent to the field. In ecology and evolutionary biology, these may have been mere assumptions that were open to challenge at the time they were introduced [3]. Yet research around topics relating land use change and infectious disease emergence keep repeating, and amplifying, under-substantiated assumptions that need to be carefully assessed in interdisciplinary context when performing research about land-use change and infectious disease emergence. Here, we elaborate on five common 'myths' (in the sense of narratives that are often accepted but not properly evaluated) we have seen repeatedly mentioned in the literature. For each myth we cite at least one article accepting it and one article refuting it.

6.2 The Five Myths

6.2.1 Everything is Driven by Population Growth

Probably one of the most common myths in population sciences is that population growth is at the root of most current environmental crises [4, 5]. For infectious diseases, this idea has been repeated in several instances [6, 7]. Interestingly, little actual reference is made to whether populations are growing, or the scale at which, if population growth is happening, population growth or density might be a problem for the emergence of new infectious diseases. Currently, we can affirm that at a global level, population growth and fertility rates are declining [8, 9]. Much research does show that ideas about either fixed global "carrying capacity" or limits to population growth as originally suggested by Malthus [10] and think tanks like the *Club of Rome* do not reflect the potential to change the internal relations of labour with food production [11] or to create niches and environments that allow higher population densities [12, 13]. Relationships between population and disease emergence are complex, nonlinear, and confounded by processes often not considered in research, such as multi-layered historic and contemporary economic, social and political forces [14, 15]

6.2.2 Deforestation is Due to Landless Peasant Groups

Deforestation has been often referred to as a major driver for infectious disease emergence [16–20]. Another common affirmation is that landless local, or migrant, populations and indigenous groups constitute a major threat to the integrity of forests. Some studies have argued about this point and made contextualized demographic connections, e.g., referring to population growth in the agricultural frontiers [21, 22], which has been an advance in light of previous beliefs about pressures for

deforestation where population growth was fully decontextualized [23]. However, little is said about factors driving migrations, for example, how land tenure disparities might drive such a focalized demography [24] and how land use policy for land tenure might drive deforestation [25, 26]. As shown by relational research, major pressures for deforestation increasingly are associated with large scale agribusiness involved in broader global circuits of capital accumulation [27–30]. Given the highly contextualized nature of deforestation a major question when assessing its role on disease emergence is inquiring about its causes and the connections with wider phenomena that also make populations more vulnerable to diseases [31, 32]. For example, we can ask: how might deforestation be one among many expressions of modes of production that release new pathogens into human populations?

6.2.3 All Agricultural Land Use Change is Detrimental to Biodiversity—Intensification of Agriculture and Land Sparing are the Solution

Ecological synthesis and meta-analysis have stressed that land use change for agricultural use is detrimental to biodiversity [33, 19]. Instead of conversion of land into more formal agricultural use, there is pressure to intensify production on existing agricultural land, thereby 'sparing' land. There is a prominent lobby for agricultural intensification and land sparing as the ultimate solution to increasing rates of disease emergence [34] and a necessary condition for biodiversity conservation [35]. These are ideas that were instilled early on in ecology, presented in tandem with the myths of uncontrolled population growth [4] and the benefits to privatizing and commodifying common natural spaces [36, 37, 5].

The types of agricultural intensification are more complicated than are often recognized, however, and they likely differ in their effects on ecology and disease emergence. The FAO noted [38], "Agricultural intensification can be technically defined as an increase in agricultural production per unit of inputs (which may be labour, land, time, fertilizer, seed, feed or cash)." Not all studies have suggested all forms of agricultural intensification reduce disease emergence. Some may actually lead to unprecedented rates and types of disease emergence—intensified livestock operations have come into particular question [18, 39]. Others have found that land use change can decrease disease transmission or have variable impacts depending in the context of infectious disease emergence [40]. Agroecological land use can reduce the abundance of medically important disease vector insects such as sand flies, while increasing their overall diversity [41, 42]; these are patterns that extend to most functional groups of species in ecologically managed agricultural systems [43, 44]. Indeed, land sparing can be associated with forms of intensification that define the plantationocene [45], a system of food production that maximizes economic profit and externalizes the stunting of human development, equally exploiting labour from slaves or marginalized populations. The plantationocene as a food production model

is an expression of the need for specialization in agricultural and other economic systems for capital accumulation [46] driving large scale land use change for distant economic growth and benefit [27, 28]. The plantationocene is in conflict with both biodiversity conservation and protection from infectious diseases emergence, considering vulnerabilities to infectious disease are shaped by socio-economic inequities [47].

The pursuit of ecologically- and socially- sound alternatives to land sparing and agribusiness-led intensification of the plantationocene is important. We suggest biodiversity conservation and infectious disease prevention may come from a focus on the agricultural matrix, the ecological space where food is produced and where organisms interact with the environment [48]. Agroecology, encompassing a variety of historical and current practices of many peoples and places, under constant experimentation and exploration [49–51], offers a framework through which the landscape may suppress and reduce instead of catalyze and amplify disease emergence [52].

6.2.4 Spillover Occurs Because of Wet Markets and People that Eat Wildlife

With the emergence of COVID-19 [34, 53], and other zoonotic diseases [54], increased calls for criminalizing traditional food markets and wildlife consumption have been aired. Similarly, interactions between local populations and wildlife tend to be scrutinized from a limited perspective that sees wildlife animals simultaneously as sources of diseases and biodiversity components threatened by people living nearby [55]. The assertions behind these claims tend to be made without reference to the historical, and current, cultural and social contexts where wet markets exist [56]. They tend to generalize and prejudge traditional practices, failing to even try to understand the roles that wildlife meat might play as sustainable protein source in the context of food sovereignty and security [57], and the sustainability of the markets as not posing threats to species conservation in contexts where they are linked with food sovereignty [58]. For example, capybaras are well adapted to the flooding plains of South America, and this giant rodent has historically been an important protein source for local populations and an important element of food sovereignty [59]. Similarly, the implementation of relatively simple hygiene measures such as having one day of market closure, cleaning at regular intervals, and selling or slaughtering all animals by the end of trading each day can significantly reduce the risk of transmission for highly virulent zoonotic pathogens [60–62]. As it happens with most spatial phenomena, the local context is also important to understand the risk of highly virulent zoonotic pathogens. For example, for avian influenza, markets near areas with rivers and other habitats where birds, pathogens, and sales can co-occur may increase transmission risk [63].

6.2.5 *Models Tell "The Truth"*

We want to now focus on a problem that has become pervasive in population sciences, the fetishization of simplistic models and quantitative relationships over the less formalized understanding of patterns and processes in populations. The problem is not unique to population sciences, as it has been well identified and discussed in geography [64–66], where for some, quantification and mathematical modelling too easily ended up taking away the value both of philosophical inquiry, on the one hand, and on the other, of empirical descriptions foregrounding (or at least not devaluing) 'mess' and complication exceeding the grasp of models. In population sciences, the fetishization of models become increasingly problematic with the use and abuse of computationally intensive tools that analyze big datasets [67–70]. This type of exercise, too often foregrounding models and results over assumptions, alternative possible assumptions, inherent limitations**,** and what is empirically not well-captured by models tends to generate research results that unconsciously reflect social constructs and beliefs that partially shape life sciences in general and the analytical methods chosen in particular; numbers do not simply speak but respond to the script used to analyze them. As warned by Box [71] "All models are wrong, but some are useful". Moreover, models are valuable tools when they serve the goal of abstraction of natural phenomena [72], when abstractions can be triangulated or checked for robustness [73], in a process where empirical and conceptual work can lead to false dichotomies being debunked [74]. Confronting the risk of oversimplification with the need of abstraction for the apprehension of complexity requires the development of models and techniques that look for drivers able to explain contradictions in quantitative relationships. It also often requires us to think and analyze more systemically, representing the 'internal relationships' between organisms in which what appears to be a bounded entity is understood as always emerging through its relationships in larger environmental networks. Such approaches are being explored by new forms of geographical information systems where the representation of space can be very different from cartesian coordinates, instead focusing on the relations of objects over space defined by interactions [75]. They are also found in research reconceptualizing relationships between land use change and infectious disease emergence by modeling land ownership dynamics and disease transmission in the historical (and perhaps ongoing) formation of large agricultural estates, a.k.a., latifundia [76]. Thus, inherent to the effort of generating "useful" models, perhaps the most pressing needs become the examination of assumptions and the need for pushing down the walls around what is merely assumed, examining what is taken as granted, questioning the unquestionable. In that struggle, the incorporation of different and diverse viewpoints and personal experiences becomes a necessity.

6.3 Inferences About Land Use Change and Disease Emergence and the Society We Can Build

The five myths we discussed are illustrated in Fig. 6.1. As the figure shows the myths often converge together and can lead to narratives that become mythologies, in the sense that the narrative might be appealing for some, used for the oppression of others, but not well grounded on phenomena occurring in nature. At best these "mythologies" end up reflecting beliefs and doctrines that are necessary for the functioning of the world as we know it and limit the ways in which science could help to solve, or even alleviate, major environmental and health problems.

Pushing the boundaries of what is commonly assumed in science is necessary to gain insight and understanding enabling successful solutions to current problems in society. In that sense demystifying truisms, as the myths we just discussed, is a necessary step to remove barriers for an impactful science whose understandings lend themselves to preventing more health problems, conserving species biodiversity, and improving standards of living, often by demonstrating the positive effects of reducing socio-economic and health inequalities. For the problem of land use change and infectious disease emergence, we consider that it is urgent to reframe research within a 'structural one health' [77] that seeks to understand the role that abstractions about capital and its dynamics have in shaping patterns of disease transmission.

Fig. 6.1 Five common myths about land use change and infectious disease emergence. When taken together the five myths we discussed can lead to narratives that can be appealing for certain groups and stakeholders. However, they can obscure the magnitude and the relation between different factors as well as how we can help society to reduce the emergence of diseases and, more generally, to solve any environmental crisis. Illustration by Nicole L. Gottdenker

Acknowledgements This work was funded by NSF CNH2-1924200, Instituto Conmemorativo Gorgas de Estudios de la Salud, Indiana University and the University of Georgia. This research was also undertaken, in part, thanks to support from the Canada Research Chairs Program.

References

1. Levins R, Lewontin RC (1985) The dialectical biologist. Harvard University Press
2. Lewontin RC, Levins R (2007) Biology under the influence: dialectical essays on ecology, agriculture, and health. Monthly Review Press, New York
3. Lewontin RC, Levins R (2000) Let the numbers speak. Int J Health Serv 30:873–877
4. Ehrlich PR, Holdren JP (1971) Impact of population growth. Science 171:1212–1217
5. Hardin G (1968) The tragedy of the commons. Science 162:1243–1248
6. McMichael AJ (2002) Population, environment, disease, and survival: past patterns, uncertain futures. The Lancet 359:1145–1148
7. Pimentel D, Cooperstein S, Randell H, Filiberto D, Sorrentino S, Kaye B, Nicklin C, Yagi J, Brian J, O'Hern J, Habas A, Weinstein C (2007) Ecology of increasing diseases: population growth and environmental degradation. Hum Ecol 35:653–668
8. Cohen JE (1997) Population, economics, environment and culture: an introduction to human carrying capacity. J Appl Ecol 34:1325–1333
9. Cohen JE (2020) Population, population, and population. Bull Ecol Soc Am 101:e01694
10. Malthus TR, Winch D, James P (1992) Malthus: "an essay on the principle of population." Cambridge University Press
11. Boserup, E. (1975) The impact of population growth on agricultural output. Q J Econ 257–270
12. Cohen JE (2005) Human population grows up. Sci Am 293:48–55
13. Neurath P (2017) From Malthus to the Club of Rome and back: problems of limits to growth, population control, and migrations. Routledge
14. Awerbuch T (1994) Evolution of mathematical models of epidemics. Ann N Y Acad Sci 740:232–241
15. Awerbuch T, Kiszewski AE, Levins R (2002) Surprise, nonlinearity and complex behaviour. In: Martens P, McMichael AJ (eds) Environmental change, climate and health. Cambridge University Press, Cambridge, pp 96–119
16. Castro MCD, Monte-Mór RL, Sawyer DO, Singer BH (2006) Malaria risk on the Amazon frontier. Proc Natl Acad Sci 103:2452–2457
17. Herrer A, Christensen HA (1976) Epidemiological patterns of cutaneous leishmaniasis in Panama. I. Epidemics among small groups of settlers. Ann Trop Med Parasitol 70:59–65
18. Jones BA, Grace D, Kock R, Alonso S, Rushton J, Said MY, McKeever D, Mutua F, Young J, McDermott J, Pfeiffer DU (2013) Zoonosis emergence linked to agricultural intensification and environmental change. Proc Natl Acad Sci U S A 110:8399–8404
19. Jones KE, Patel NG, Levy MA, Storeygard A, Balk D, Gittleman JL, Daszak P (2008) Global trends in emerging infectious diseases. Nature 451:990–993
20. Vinson JE, Gottdenker NL, Chaves LF, Kaul RB, Kramer AM, Drake JM, Hall RJ (2022) Land reversion and zoonotic spillover risk. R Soc Open Sci 9:220582
21. López-Carr D (2021) A review of small farmer land use and deforestation in tropical forest frontiers: implications for conservation and sustainable livelihoods. Land 10:1113
22. López-Carr D, Burgdorfer J (2013) Deforestation drivers: population, migration, and tropical land use. Env Sci Policy Sustain Dev 55:3–11
23. Allen JC, Barnes DF (1985) The causes of deforestation in developing countries. Ann Assoc Am Geogr 75:163–184
24. Wright AL, Wolford W (2003) To inherit the earth: the landless movement and the struggle for a new Brazil. Food First Books

25. Rosero-Bixby L, Maldonado-Ulloa T, Bonilla-Carrion R (2002) Forests and population on the Osa Peninsula, Costa Rica. Rev Biol Trop 50:585–598
26. Rosero-Bixby L, Palloni A (1998) Population and deforestation in Costa Rica. Popul Environ 20:149–185
27. Bergmann L, Holmberg M (2016) Land in motion. Ann Am Assoc Geogr 106:932–956
28. Ceddia MG (2020) The super-rich and cropland expansion via direct investments in agriculture. Nat Sustain 3:312–318
29. Curtis PG, Slay CM, Harris NL, Tyukavina A, Hansen MC (2018) Classifying drivers of global forest loss. Science 361:1108–1111
30. Sacchi LV, Gasparri NI (2016) Impacts of the deforestation driven by agribusiness on urban population and economic activity in the Dry Chaco of Argentina. J Land Use Sci 11:523–537
31. Chaves LF, Cohen JM, Pascual M, Wilson ML (2008) Social exclusion modifies climate and deforestation impacts on a vector-borne disease. PLoS Negl Trop Dis 2:e176
32. Yamada K, Valderrama A, Gottdenker N, Cerezo L, Minakawa N, Saldaña A, Calzada JE, Chaves LF (2016) Macroecological patterns of American Cutaneous Leishmaniasis transmission across the health areas of Panamá (1980–2012). Parasite Epidemiol Control 1:42–55
33. Gibb R, Redding DW, Chin KQ, Donnelly CA, Blackburn TM, Newbold T, Jones KE (2020) Zoonotic host diversity increases in human-dominated ecosystems. Nature 584:398–402
34. Daszak P, Amuasi J, das Neves CG, Hayman D, Kuiken T, Roche B, Zambrana-Torrelio C, Buss P, Dundarova H, Feferholtz Y, Földvári G, Igbinosa E, Junglen S, Liu Q, Suzan G, Uhart M, Wannous C, Woolaston K, Mosig Reidl P, O'Brien K, Pascual U, Stoett P, Li H, Ngo HT (2020) Workshop report on biodiversity and pandemics of the intergovernmental platform on biodiversity and ecosystem services, vol 108. Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES), Bonn, Germany
35. Bengochea Paz D, Henderson K, Loreau M (In Press) Habitat percolation transition undermines sustainability in social-ecological agricultural systems. Ecol Lett n/a
36. Ehrlich PR (1982) Human carrying capacity, extinctions, and nature reserves. Bioscience 32:331–333
37. Fischer J, Brosi B, Daily GC, Ehrlich PR, Goldman R, Goldstein J, Lindenmayer DB, Manning AD, Mooney HA, Pejchar L (2008) Should agricultural policies encourage land sparing or wildlife-friendly farming? Front Ecol Environ 6:380–385
38. FAO (2004) The ethics of sustainable agricultural intensification. FAO, Rome, Italy
39. Wallace RG (2009) Breeding influenza: the political virology of offshore farming. Antipode 41:916–951
40. Gottdenker NL, Streicker DG, Faust CL, Carroll CR (2014) Anthropogenic land use change and infectious diseases: a review of the evidence. EcoHealth 11:619–632
41. Alexander B, Agudelo LA, Navarro F, Ruiz F, Molina J, Aguilera G, Quiñones ML (2001) Phlebotomine sandflies and leishmaniasis risks in Colombian coffee plantations under two systems of cultivation. Med Vet Entomol 15:364–373
42. Alexander B, Oliveria EBD, Haigh E, Almeida LLD (2002) Transmission of Leishmania in coffee plantations of Minas Gerais, Brazil. Memórias do Instituto Oswaldo Cruz 97:627–630
43. Perfecto I, Vandermeer J, Philpott SM (2014) Complex ecological interactions in the coffee agroecosystem. Annu Rev Ecol Evol Syst 45:137–158
44. Vandermeer J, Armbrecht I, de la Mora A, Ennis KK, Fitch G, Gonthier DJ, Hajian-Forooshani Z, Hsieh H-Y, Iverson A, Jackson D, Jha S, Jiménez-Soto E, Lopez-Bautista G, Larsen A, Li K, Liere H, MacDonald A, Marin L, Mathis KA, Monagan I, Morris JR, Ong T, Pardee GL, Rivera-Salinas IS, Vaiyda C, Williams-Guillen K, Yitbarek S, Uno S, Zemenick A, Philpott SM, Perfecto I (2019) The community ecology of herbivore regulation in an agroecosystem: lessons from complex systems. Bioscience 69:974–996
45. Wolford W (2021) The Plantationocene: a Lusotropical contribution to the theory. Ann Am Assoc Geogr 111:1622–1639
46. Sheppard E (2012) Trade, globalization and uneven development: Entanglements of geographical political economy. Prog Hum Geogr 36:44–71

47. Galea S (2021) The contagion next time. Oxford University Press
48. Perfecto I, Vandermeer J (2010) The agroecological matrix as alternative to the land-sparing/agriculture intensification model. Proc Natl Acad Sci 107:5786–5791
49. Griffon D, Hernandez M-J (2020) Some theoretical notes on agrobiodiversity: spatial heterogeneity and population interactions. Agroecol Sustain Food Syst 44:795–823
50. Griffon D, Hernandez M-J, Ramírez D (2021) Theoretical clues for agroecological transitions: the conuco legacy and the monoculture trap. Front Sustain Food Syst 5
51. Levins R, Vandermeer JH (1990) The agroecosystem embedded in a complex ecological community. In: Carroll CR, Vandermeer J, Rosset PM (eds) Agroecology. Wiley and Sons, Hoboken, USA, pp 341–362
52. Wallace R, Chaves LF, Bergmann L, Ayres Lopes CFJ, Hogerwerf L, Kock R, Wallace RG (2018) Clear-cutting disease control: capital-led deforestation, public health austerity, and vector-borne infection. Springer, New York
53. Xiao X, Newman C, Buesching CD, Macdonald DW, Zhou Z-M (2021) Animal sales from Wuhan wet markets immediately prior to the COVID-19 pandemic. Sci Rep 11:11898
54. Wolfe ND, Daszak P, Kilpatrick AM, Burke DS (2005) Bushmeat hunting, deforestation, and prediction of zoonoses emergence. Emerg Infect Dis 11:1822–1827
55. Li H, Mendelsohn E, Zong C, Zhang W, Hagan E, Wang N, Li S, Yan H, Huang H, Zhu G, Ross N, Chmura A, Terry P, Fielder M, Miller M, Shi Z, Daszak P (2019) Human-animal interactions and bat coronavirus spillover potential among rural residents in Southern China. Biosaf Health 1:84–90
56. Wallace RG (2016) Big farms make big flu: dispatches on influenza, agribusiness, and the nature of science. Monthly Review Press
57. Hoffman LC, Cawthorn D-M (2012) What is the role and contribution of meat from wildlife in providing high quality protein for consumption? Anim Front 2:40–53
58. Cowlishaw G, Mendelson S, Rowcliffe JM (2005) Evidence for post-depletion sustainability in a mature bushmeat market. J Appl Ecol 42:460–468
59. Nogueira-Filho SLG, da Cunha Nogueira SS (2018) Capybara meat: an extraordinary resource for food security in South America. Meat Sci 145:329–333
60. Leung YH, Lau EH, Zhang LJ, Guan Y, Cowling BJ, Peiris JS (2012) Avian influenza and ban on overnight poultry storage in live poultry markets, Hong Kong. Emerg Infect Dis 18:1339–1341
61. Wang W, Artois J, Wang X, Kucharski AJ, Pei Y, Tong X, Virlogeux V, Wu P, Cowling BJ, Gilbert M, Yu H (2020) Effectiveness of live poultry market interventions on human infection with avian influenza A(H7N9) virus, China. Emerg Infect Dis 26:891–901
62. Wang X, Wang Q, Cheng W, Yu Z, Ling F, Mao H, Chen E (2017) Risk factors for avian influenza virus contamination of live poultry markets in Zhejiang, China during the 2015–2016 human influenza season. Sci Rep 7:42722
63. Mellor KC, Meyer A, Elkholly DA, Fournié G, Long PT, Inui K, Padungtod P, Gilbert M, Newman SH, Vergne T, Pfeiffer DU, Stevens KB (2018) Comparative epidemiology of highly pathogenic avian influenza virus H5N1 and H5N6 in Vietnamese Live Bird Markets: spatiotemporal patterns of distribution and risk factors. Front Vet Sci 5
64. Bunge W (1977) The point of reproduction: a second front. Antipode 9:60–76
65. Harvey D (1990) Between space and time: reflections on the geographical imagination. Ann Assoc Am Geogr 80:418–434
66. O'Sullivan D, Manson SM (2015) Do physicists have geography envy? And what can geographers learn from It? Ann Assoc Am Geogr 105:704–722
67. Han BA, O'Regan SM, Paul Schmidt J, Drake JM (2020) Integrating data mining and transmission theory in the ecology of infectious diseases. Ecol Lett 23:1178–1188
68. Kreuder Johnson C, Hitchens PL, Smiley Evans T, Goldstein T, Thomas K, Clements A, Joly DO, Wolfe ND, Daszak P, Karesh WB, Mazet JK (2015) Spillover and pandemic properties of zoonotic viruses with high host plasticity. Sci Rep 5:14830
69. Olival KJ, Hosseini PR, Zambrana-Torrelio C, Ross N, Bogich TL, Daszak P (2017) Host and viral traits predict zoonotic spillover from mammals. Nature 546:646–650

70. Stephens PR, Gottdenker N, Schatz A, Schmidt J, Drake JM (2021) Characteristics of the 100 largest modern zoonotic disease outbreaks. Philos Trans R Soc B 376:20200535
71. Box GEP (1979) Some problems of statistics and everyday life. J Am Stat Assoc 74:1–4
72. Levins R (2006) Strategies of abstraction. Biol Philos 21:741–755
73. Levins R (1968) Evolution in changing environments. Some theoretical explorations. Princeton University Press, Princeton
74. Levins R (1998) The internal and external in explanatory theories. Sci Cult 7:557–582
75. Bergmann L, O'Sullivan D (2018) Reimagining GIScience for relational spaces. Can Geogr/Le Géographe Canadien 62:7–14
76. Chaves LF (2013) The dynamics of Latifundia formation. PLoS ONE 8:e82863
77. Wallace RG, Bergmann L, Kock R, Gilbert M, Hogerwerf L, Wallace R, Holmberg M (2015) The dawn of Structural One Health: a new science tracking disease emergence along circuits of capital. Soc Sci Med 129:68–77

Dr. Luis Fernando Chaves (Ph.D., M. Sc., Lic.) is a disease ecologist interested in the linkages between uneven development and disease emergence and transmission. Most of his work has been focused on vector-borne diseases. He is currently an Associate Professor at the Department of Environmental and Occupational Health, School of Public Health at Indiana University-Bloomington in the United States.

Ms. Chystrie A. Rigg (M.Sc., Lic.) is a medical entomologist interested in sand flies and mosquitoes and the parasitic pathogens these insects transmit. She has extensive field and laboratory experience working with medically important insects, and often collaborates with the Vector Control Department at Panamá's Ministry of Health. She is currently a tenured assistant health research scientist III at the Instituto Conmemorativo Gorgas de Estudios de la Salud in Panamá.

Dr. Mariel D. Friberg (Ph.D., B. Eng.) is an environmental engineer interested in modeling, big data analytics and remote sensing applied to air quality, wildfires and global health problems. She is an assistant research scientist at the NASA Goddard Space Center Earth Science Division and the University of Maryland, College Park Earth System Science Interdisciplinary Center in the United States.

Ms. Milixa Perea (M.S., Lic) is a parasitologist interested in enteric and vector-transmitted parasites. Her major research topic has been the molecular characterization of *Blastocystis* spp. subtypes. She currently is a tenured assistant health research scientist III at the Instituto Conmemorativo Gorgas de Estudios de la Salud in Panamá.

Ms. Lisbeth Amarilis Hurtado (M.S., Lic.) is a statistician working on infectious disease epidemiology. Her research integrates field studies with advanced statistical modeling tools. She currently works in the Biostatistics Unit at the Instituto Conmemorativo Gorgas de Estudios de la Salud in Panamá.

Dr. Nicole L. Gottdenker (DVM, Dipl. ACVP, Ph.D., M.S., B.S.) works both as a pathologist and disease ecologist at the University of Georgia (USA), where she is an associate professor of veterinary pathology. Most of her research has been centered around the impacts of land use change on multihost pathogens.

Dr. Luke R. Bergmann (Ph.D., M.A., B.S.) is a geographer interested in critical-computational and social-theoretic geographies. He currently holds the Canada Research Chair in GIS, Geospatial Big Data and Digital Geohumanities and is an Associate Professor of Geography at the University of British Columbia in Vancouver, Canada.

Part III
Data and Methodological Issues for Health Studies

Chapter 7
Geospatial Environmental Data for Planetary Health Applications

Michael C. Wimberly

Abstract Planetary health research focused on vector-borne and zoonotic diseases often requires data on the environmental factors that influence vectors, hosts, and pathogens. We summarized major types of geospatial environmental data that are freely available and potentially useful for planetary health applications. There are many relevant geospatial data products that characterize weather, climate, vegetation, land surface temperature, land cover and land use, human population characteristics, and hydrology. However, these datasets differ greatly in their underlying measurement techniques and spatial and temporal resolutions. Although many datasets have global coverage, they vary considerably in their spatial accuracy and suitability for local applications. We recommend that researchers carefully consider the strengths and limitations of alternative data sources with a particular focus on the spatial and temporal scales of the data relative to the specific organisms and processes of interest. Research that addresses the sensitivities of analytical results and model predictions to alternative data sources can provide additional guidance to inform these decisions.

Keywords Climate and weather · Land cover and land use · Geographic information systems · Remote sensing

7.1 Introduction

The field of planetary health addresses myriad interconnections between global environmental change and the health of humans, animals, and the ecosystems they inhabit [1]. It shares this conceptual foundation with related interdisciplinary fields such as One Health [2], EcoHealth [3], and GeoHealth [4]. All are based on a holistic framework that emphasizes the relationships between human health, the social environment, the physical environment, and the non-human organisms that are hosts and vectors for disease-causing pathogens. Because of this breadth, there is a need for diverse sources of data to characterize multiple aspects of human and natural systems.

M. C. Wimberly (✉)
Department of Geography and Environmental Sustainability, University of Oklahoma, Norman, OK 73061, USA
e-mail: mcwimberly@ou.edu

T.-H. Wen et al. (eds.), *Earth Data Analytics for Planetary Health*, Atmosphere, Earth, Ocean & Space, https://doi.org/10.1007/978-981-19-8765-6_7

Geospatial data that map the spatial patterns of relevant phenomena are particularly important for assessing spatial relationships and identifying hot spots with high risk of disease transmission.

Vector-borne and zoonotic diseases are particularly sensitive to features of the physical environment that influence the reproduction, growth, and survival of vectors, hosts, and pathogens. Climate, weather, water, vegetation, and land use influence transmission cycles through their effects on vector and host habitats, pathogen development and transmission, and human exposure to vectors [5–7]. To conduct research on these diseases and translate the results into applications, it is essential to measure the relevant environmental variables. Accurate and timely data are needed to test hypotheses about drivers of disease transmission, develop maps of infectious disease risk based on environmental factors, and forecast future disease risk resulting from changes in weather and climate. Even when the research is focused on other questions, such as the effectiveness of public health interventions, it is still necessary to control for background effects of environmental variation on spatial and temporal patterns of disease transmission [8].

These environmental factors are heterogeneous at multiple spatial and temporal scales. Broad climate gradients vary geographically with latitude and elevation and change gradually over decades. Within a given climate, weather fluctuates continuously and exhibits diurnal, seasonal, and interannual cycles. More localized patterns related to vegetation, topography, and human land use vary at spatial scales from hundreds of meters to hundreds of kilometers and change over time scales from years to decades. These landscape features create microclimates that differ considerably from the broader macroclimate, and these local conditions can facilitate disease transmission even when the broader macroclimate is unsuitable [9, 10]. When selecting the environmental data for a planetary health application, it is essential to understand the scales of environmental measurements and match them with the specific ecological and epidemiological processes of interest.

There are numerous geospatial data products that characterize a variety of environmental characteristics. Many of these products are updated regularly and are available at continental to global extents, providing opportunities for widespread use in planetary health. However, the underlying data are collected over a wide range of spatial and temporal scales. Measurement techniques and the resulting accuracies also vary among data products, as do the techniques used for spatial interpolation and filling of data gaps. All of these factors can affect inferences about environment-disease relationships and the accuracy of predictive models based on these relationships. The goal of this chapter is to summarize the main environmental data sources that have been used in planetary health applications related to vector-borne and zoonotic diseases. Strengths and limitations of various data products are highlighted and emerging trends are discussed.

7.2 Meteorological Data

Ground-based meteorological stations provide in situ observations of weather, and long-term summaries of these data are the basis for measuring climate and tracking climate change. Standard variables monitored at weather stations include air temperature, precipitation, humidity, atmospheric pressure, wind speed, and solar radiation. A critical objective in designing and siting weather stations is ensuring consistent observations that can be compared over time and between different locations [11]. Meteorological stations are therefore located in open areas where measurements are not influenced by buildings or tall vegetation. Instruments are enclosed to protect them from direct solar radiation, condensation, and precipitation while allowing sufficient ventilation to facilitate airflow over the sensors. Because of the expense of installation, equipment maintenance, and data curation, meteorological stations have historically been operated by government agencies [12]. However, volunteer observers are also an important part of the enterprise through programs like the National Weather Service (NWS) Cooperative Observer Program, and the availability of low-cost digital home weather stations has allowed private citizens to provide crowdsourced weather observations [13].

Station data are typically regarded as the gold standard for near-surface observations of weather and climate [14]. However, the types of instruments, frequencies of measurements, and completeness of the resulting data all vary between stations and over time. In general, high-income countries have well developed weather monitoring systems with higher densities of stations and more technologically advanced equipment and data infrastructure than lower-income countries in the Global South [15]. Even in countries with highly resourced weather monitoring infrastructures, most of the places where disease transmission occurs are located relatively far from extant weather stations. Thus, an important issue is determining the degree to which distant weather stations are representative of the environments that directly influence disease transmission cycles.

One way to obtain more spatially precise estimates of local weather and climate is to interpolate the point data collected at meteorological stations (Fig. 7.1). This approach involves predicting meteorological variables at unsampled locations based on the spatial pattern of nearby measurements. In some cases, ancillary variables that are strongly associated with climate gradients, such as elevation, are incorporated to increase the accuracy and precision of local measurements. Commonly used techniques include various types of regression, kriging, self-organizing maps, and thin-plate splines [16–18]. In most cases, data users do not need to carry out this interpolation themselves, as there are many free gridded weather and climate products produced by various institutions. Meteorological variables can also be extracted from reanalysis data sets, which are generated using data assimilation methods that combine multiple sources of historical weather data with numerical weather models [19].

Although many of these products appear similar, there are underlying differences in the methods used to generate the data and the characteristics of the resulting

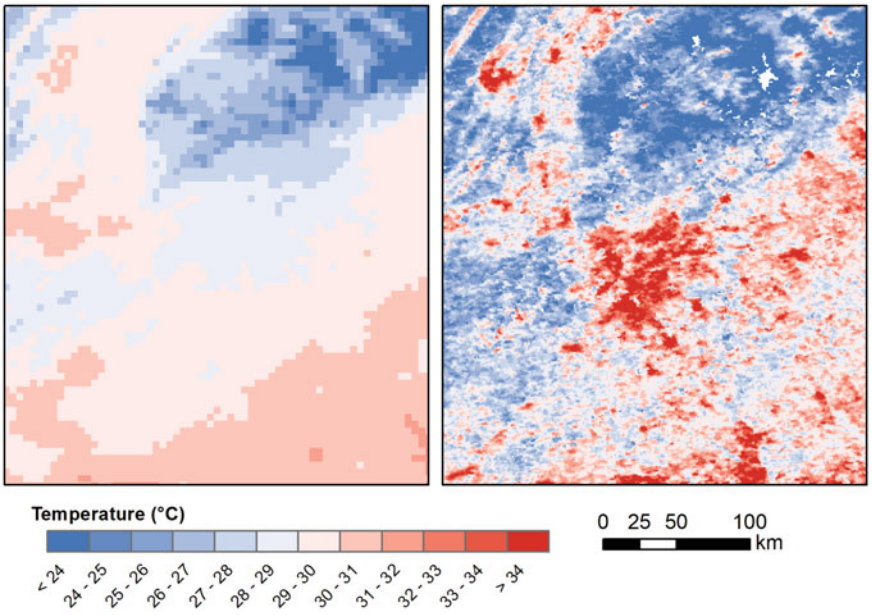


Fig. 7.1 Comparison of two temperature datasets for north Georgia, USA, in June 2020. Left: PRISM interpolated monthly maximum near-surface air temperature (4 km cell size). Right: MODIS Aqua daytime land surface temperature 8-day composite from June 10–17 (1 km cell size)

meteorological grids that can influence results when they are used for planetary health applications [20, 21]. Gridded meteorological and climate data vary considerably in their spatial and temporal scales. For example, the University of East Anglia Climate Research Unit (CRU) datasets provide global historical monthly time series and climatologies at a grid cell size of 0.5° (approximately 55 km) [22]. The Climate Hazards Group Coupled Infrared Precipitation with Stations (CHIRPS) [23] and Temperature with Stations (CHIRTS) [24] datasets combine interpolated station data with satellite estimates of precipitation and land surface temperature to produce daily and monthly estimates at a much smaller grid cell size of 0.05° (approximately 5.5 km). Other downscaled climate data products like WorldClim [25] and Chelsa [26] use high-resolution elevation data to downscale climate grids to a cell size of 30 arc seconds (approximately 1 km).

Differences in the methods used for interpolation and downscaling lead to variations in the meteorological grids that can influence results when they are used for planetary health applications [19, 20]. There are also trade-offs between dataset attributes such as grid cell size, frequency of measurement, and the time required to process the data and make them available. In the United States, the Parameter-Elevation Regressions on Independent Slopes Model (PRISM) climate dataset provides monthly meteorological data at a spatial resolution of 800 m [27]. In contrast, the National Land Data Assimilation System (NLDAS) forcings dataset provides many of the

same meteorological variables on a 30-min time step with a latency of several days, but the grid cell size is 0.125° (approximately 14 km) [28]. The GridMET dataset combines these two data sources to provide gridded meteorological variables at a spatial resolution of 4 km and a daily time step [29].

7.3 Satellite Vegetation Indices

Earth-observing satellites are another source of geospatial environmental data that can be used to predict spatial and temporal patterns of infectious disease transmission [30–32]. Unlike the point-level data obtained from weather stations, satellite images provide spatially continuous measurements over large areas of the Earth's surface and are repeated at intervals ranging from days to weeks. They are fundamentally different from weather station data in that they typically measure conditions on the land surface, not in the near-surface atmosphere. The most commonly used satellite remote sensing data are observations of reflected solar radiation in the visible and infrared wavelengths. These data are measured as radiance or reflectance in one or more spectral bands, where each band encompasses a specific range of wavelengths. These bands are then used to calculate spectral indices that characterize physical properties of the Earth's surface.

The most common spectral index is the normalized difference vegetation index (NDVI, Fig. 7.2), which measures green vegetation using red and near infrared spectral bands [33]. In most cases, vegetation greenness itself is not a proximal driver of disease transmission. However, the NDVI index is highly sensitive to meteorological factors such as temperature and precipitation [34]. In temperate environments, NDVI changes in response to vegetation greenup in the spring and senescence in the fall and can provide information about timing and length of disease transmission seasons [35, 36]. In water-limited environments, NDVI is sensitive to rainfall (Fig. 7.3) and can be an indicator of water availability and drought [37]. Several variations of the NDVI have been developed to improve greenness estimates in particular situations. For example, the enhanced vegetation index (EVI) was developed to mitigate issues with index saturation in dense forests [38], and the soil-adjusted vegetation index (SAVI) was designed to correct for effects of soil brightness in areas with low vegetation cover [39].

A major advantage of NDVI is that the necessary data are widely available over long time periods for nearly every location on Earth. The NDVI can be calculated using data from a variety of satellite sensors, which provide data at different spatial and temporal scales. The earliest applications of satellite imagery for research on vector-borne diseases involved the Advanced Very High Resolution Radiometer (AVHRR), which has been operational on United States National Oceanic and Atmospheric Administration (NOAA) weather satellites since 1981 and provides daily data at pixel sizes of 1000–4000 m [40, 41]. The more recent MODIS sensor, on board the National Aeronautics and Space Administration (NASA) Terra and Aqua Satellites,

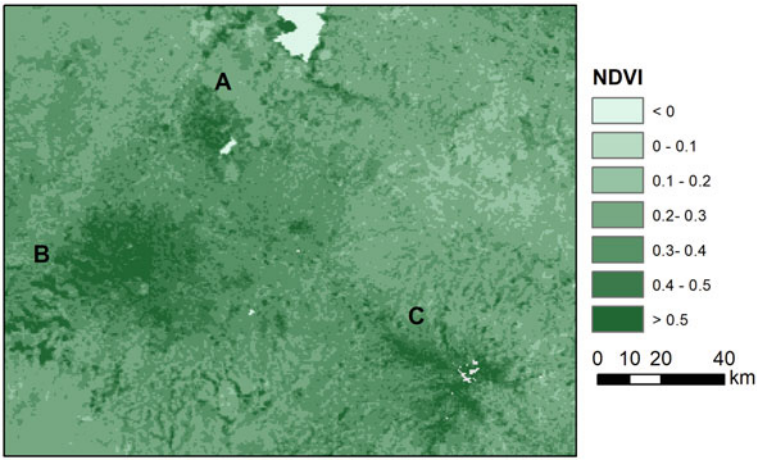


Fig. 7.2 Normalized difference vegetation index (NDVI) for part of the Amhara region of Ethiopia on May 1, 2019. The index was calculated using MODIS BRDF-Adjusted Reflectance Data (500 m cell size). Locations with high vegetation greenness include irrigated agriculture (**a**), areas with high densities of tree cover (**b**), and high elevation zones (**c**)

has provided daily global NDVI data since 2000 at spatial resolutions between 250–1000 m. These data have been widely used to model infectious disease outcomes over relatively large areas when frequent measurement intervals are required [36, 42]. The Visible Infrared Imaging Radiometer Suite (VIIRS) instrument, carried aboard multiple NOAA satellites, also generates daily global estimates of NDVI at spatial resolutions between 500–1000 m and will provide continuity after the end of the MODIS mission. Data from the Landsat and Sentinel missions can be used to derive NDVI at spatial resolutions from 10-30 m with weekly revisit intervals. These data can be applied when higher-resolution environmental measurements are needed for more localized predictions of vector habitats, host habitats, and disease transmission risk [43].

Other advantages of using NDVI and related spectral indices to measure environmental variability include the global availability of satellite imagery and the relatively high spatial resolution of the data compared to the grid size of interpolated meteorological datasets [34]. However, NDVI also has important limitations as an environmental metric for planetary health. NDVI is an indirect environmental measure that is sensitive to multiple environmental factors and the ecological characteristics of the observed vegetation. Therefore, the underlying mechanisms of the relationships between NDVI and disease risk can be obscured, and it is usually not possible to generalize across multiple ecosystems with different landscapes and vegetation. Another major challenge with NDVI is that the underlying visible and infrared imagery is affected by cloud cover [44]. This results in missing data, particularly in cloudy tropical regions, which must be imputed using gap filling techniques or otherwise accounted for in subsequent analyses.

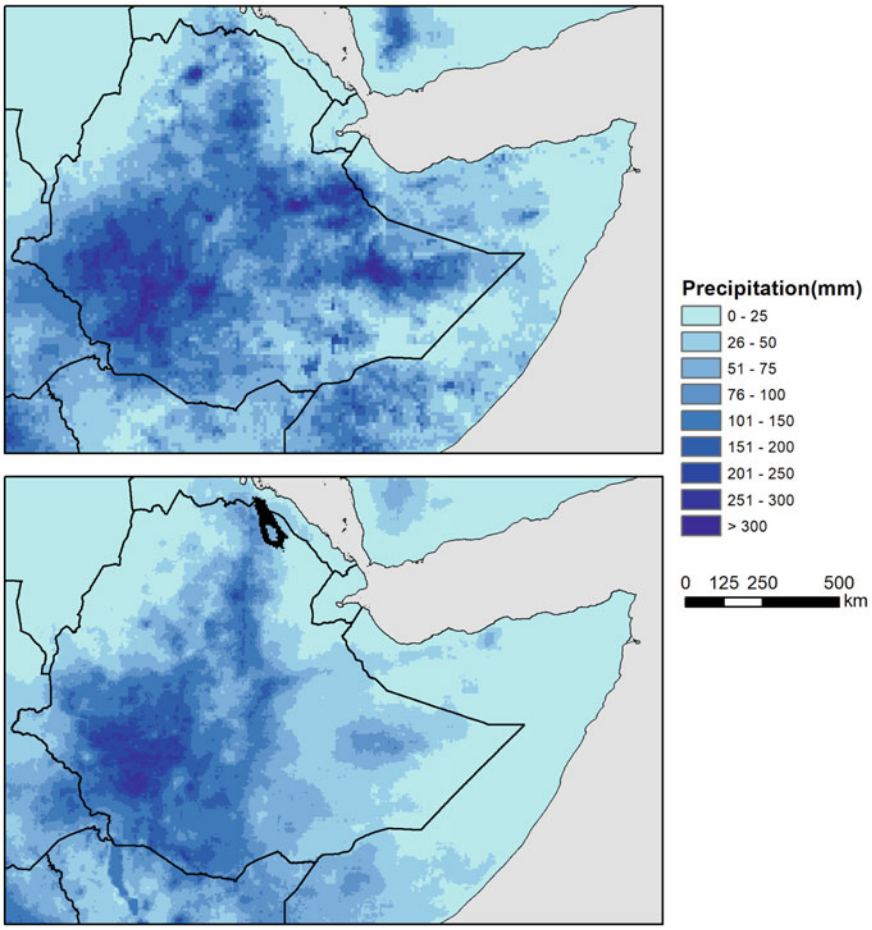


Fig. 7.3 Comparison of two gridded precipitation datasets that combine satellite estimates with ground station data for Ethiopia in March 2019. Top: IMERG (10 km cell size). Bottom: CHIRPS (5.5 km cell size)

7.4 Satellite Land Surface Temperature

Satellite sensors can also measure emitted longwave infrared radiation, which provides information about the temperature of the Earth's surface (Fig. 7.1). Land surface temperature (LST) is a characteristic of the topmost surface layer, which may be vegetation, soil, water, or human-built impervious surfaces depending on the land cover characteristics at a particular location [45]. Importantly, LST measured by satellites is not the same as the near-surface air temperatures (typically 2 m above the

land surface) measured by meteorological stations and represented in gridded meteorological data products. In meteorology, near surface air temperatures are sometimes referred to as simply "surface temperatures", which can lead to confusion.

LST and near-surface air temperature generally exhibit similar patterns of change over time, including diurnal and seasonal cycles as well as long-term trends [46]. At global and regional scales, LST and near-surface air temperature also follow similar spatial gradients with latitude and elevation. At more localized scales, LST and air temperature usually differ because of the effects of solar radiation, wind, and soil moisture [47]. For example, during the day a paved surface will be warmer than the air above it because it absorbs and re-radiates thermal energy, whereas well-watered vegetation will be cooler because of latent heat loss due to evapotranspiration. At night when there is no incoming solar radiation, land surface temperature and air temperature are usually more similar than during the day [48]. Because of these differences, land surface temperature may not be a precise indicator of the air temperature experienced by organisms above the ground surface or underneath a forest canopy. However, LST is often a reliable proxy for relative variation in air temperature over space and time and can be particularly useful in situations where reliable in-situ measurements of localized temperature are not available.

As with greenness indices, LST data are available from multiple sensors over a range of spatial and temporal resolutions. Daily daytime and nighttime LST estimates are available from MODIS [49] and VIIRS at a grid cell size of 1000 m [50]. Biweekly daytime observations at grid cell sizes of 60–120 m are available from the Thematic Mapper (TM), Enhanced Thematic Mapper (ETM+) and Thermal Infrared Sensor (TIRS) on board Landsat 4–5, 7, and 8–9 respectively [51]. The ECOsystem Spaceborne Thermal Radiometer Experiment on Space Station (ECOSTRESS) sensor provides LST measurements every 4–5 days at a grid cell size of 70 m [52]. All of these sources provide standard data products that estimate LST by combining atmospherically corrected observations of emitted radiation in the thermal wavelengths with measurements of emissivity. These methods are complex and are outside the expertise of most end users in planetary health. However, it is important to recognize that LST estimates can vary depending on the specific method used [53]. Although LST measurements are subject to missing data from cloud cover, the thermal wavelengths used to measure LST are less sensitive to clouds than the shorter-wavelength visible and near infrared bands used to compute greenness indices.

7.5 Satellite Precipitation Estimates

In addition to the interpolated meteorological data products discussed previously, gridded precipitation estimates can also be derived from satellite observations [54]. Satellite precipitation estimates are based on visible/infrared data, passive microwave data, and active microwave (radar) data. Because convecting clouds are usually bright and cold, they can be detected indirectly from their reflectance in the visible and near-infrared wavelengths combined with temperature estimates from thermal

infrared observations. Passive and active microwave observations provide more direct estimates because microwaves can penetrate clouds and are scattered by water droplets and ice particles in the atmosphere. The algorithms used to generate satellite precipitation estimates typically integrate satellite data from multiple sensors.

Planetary health researchers can obtain satellite precipitation data from multiple products, each of which uses different input data sources and estimation algorithms (Fig. 7.3). These products often have relatively coarse grid cell sizes, with measurements taken hourly and made available almost immediately. The satellite data can be combined with ground data from weather stations to improve the estimates. For example, the NASA IMERGE product provides global precipitation estimates at a 10 km grid cell size at a time step of 30 min [55]. It includes "Early Run" and "Late Run" datasets that are based only on satellite data and have latencies of less than one day, and a "Final Run" dataset that incorporates station data from the Global Precipitation Climatology Centre but has a latency of >3 months. Other widely-used satellite precipitation data products include Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN) [56], Climate Prediction Center Morphing Technique (CMORPH) [57], and the Global Precipitation Climatology Project (GPCP) [58].

7.6 Land Cover and Land Use Change

Land cover encompasses the biophysical characteristics of the Earth's land surface, including natural and cultivated vegetation, bare soil, human-built impervious surfaces, and water bodies. Land use describes human activities on the land surface, which can range from development and habitation to agricultural practices to nature preservation. Land cover and land use are often related. Locations with a high coverage of impervious surfaces are likely to be residential, commercial, or industrial areas, and agriculture replaces natural vegetative cover with new vegetation consisting of crop plants. However, land cover is not always an indicator of land use. Forest cover, for example, may result from low-density human habitation, forest management for timber production, or land preservation as a park or conservation area. Because satellite remote sensing measures physical characteristics of the land surface, it can be used to generate gridded maps of land cover and monitor changes over time. In some cases, it is also possible to infer information about land use from these land cover characteristics. The resulting data products are often referred to as land cover and land use (LCLU) products.

A large number of LCLU datasets are available at extents ranging from nations or regions to the entire globe. Coarse-grained global LCLU maps with a grid cell size of 500 m and an annual time step have been developed using data from the MOderate Resolution Imaging Spectroradiometer (MODIS) sensors on board NASA's Aqua and Terra spacecraft, which began collecting data in 2000 [59]. The grid cells are classified into relatively broad land cover types such as deciduous and evergreen forests, grasslands, shrublands, croplands, and built-up areas. Because of the relatively coarse

grid size, many cells are not homogeneous, and instead contain mixtures of multiple LCLU types. An alternative approach is to map LCLU as continuous fields, where the proportion of each grid cell containing a particular LCLU is estimated. For example, the MODIS Vegetation Continuous Fields (VCF) product provides global data on the percentage of tree cover, non-tree vegetation cover, and non-vegetated cover in 250 m grid cells [60]. The Copernicus Global Land Cover fractional cover layers (Fig. 7.4) similarly provide annual fractional cover estimates for a variety of LCLU classes such as trees, shrubs, herbaceous vegetation, crops, bare soil, and built-up areas at a grid cell size of 100 m [61].

Although these datasets provide information about general patterns of LCLU over space and time, planetary health applications often require more detailed information at finer spatial resolutions. For example, research on the habitat associations of vector and host species may require maps of land use practices such as irrigated agriculture or details about the sizes, shapes, and connectivity of habitat fragments. Satellite missions with finer grid cell sizes such as Landsat (30 m multispectral), Sentinel-1

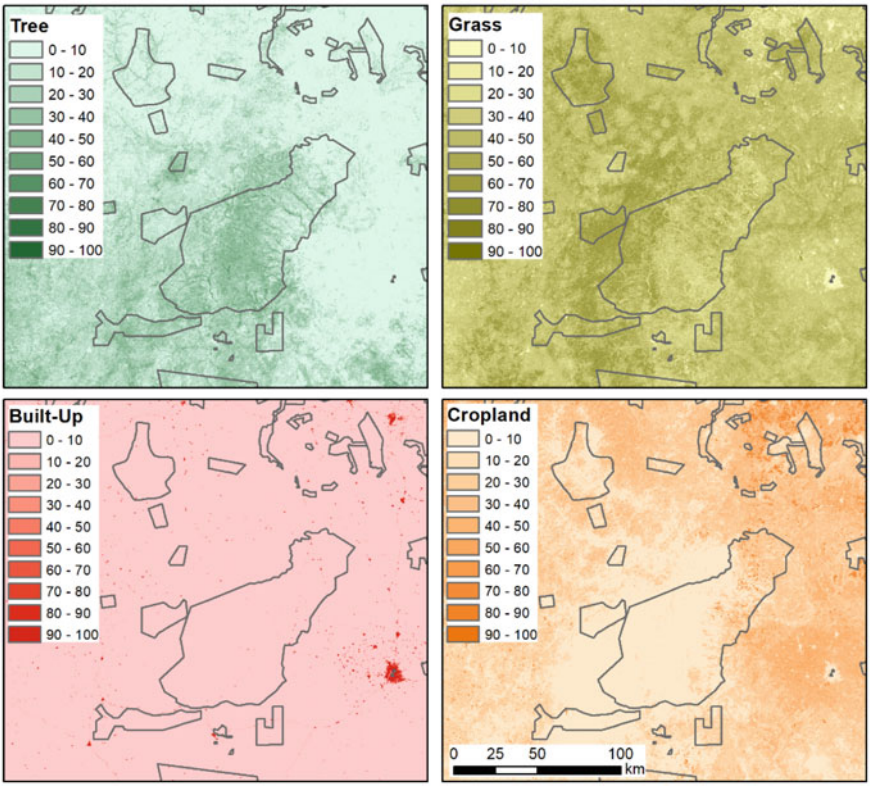


Fig. 7.4 Maps of 2019 land cover in the savanna zone of northern Ghana from the Copernicus Global Land Cover dataset. Land cover of trees, grasses, built-up areas, and croplands is represented as percent cover within 100 m grid cells

(10 m synthetic aperture radar) and Sentinel-2 (10–20 m multispectral) provide data that can be used to generate higher-resolution LCLU products. Many of these datasets are available globally, including data on forest cover and change [62], croplands [63], and cities [64]. However, global availability does not mean that a dataset is well suited for every location across the globe as local accuracy can vary considerably in different locations. In many cases, datasets developed at the regional, national, or local scales may be more accurate and include more relevant LCLU characteristics than global products [65]. These data are often more challenging to discover and access than global products.

7.7 Human Populations

Human population density is a land use characteristic that is particularly important for planetary health research and applications. Data on the human population is needed to calculate the population at risk for epidemiological rates such as incidence and prevalence, and the number of susceptible humans is an important factor influencing the transmission patterns of many infectious diseases. The most common sources of human population data are national censuses, in which people are enumerated within administrative units. Population characteristics can be summarized and mapped within polygons that outline the boundaries of these areas. Although these datasets are produced by individual countries, aggregated global populations datasets such as the Global Rural–Urban Mapping Project (GRUMP) and the Gridded Population of the World (GPW) are also available (Fig. 7.5). These products are published as grids with cell sizes from 1–110 km, but the true spatial resolution of the data is still the administrative unit within which they were aggregated.

It is often desirable to have population data with a finer spatial grain so that urban and rural areas can be distinguished and population density can be estimated for individual settlements or neighborhoods. A common method for generating finer-grained population data is the "top down" approach, in which census data are disaggregated from their administrative units to smaller grid cells using spatial information on land cover, land use, roads, and other factors that are expected to influence population density [66]. These variables are used to calculate a layer of gridded weights that are used to distribute the population within an administrative unit to reflect differences between densely populated urban areas and more sparsely populated rural locations. An alternative is a "bottom up" approach where high-resolution imagery to census individual dwellings are combined with local survey data to estimate population density at a high resolution [67]. The WorldPop project provides a global archive of population and other demographic data products generated using both top-down and bottom-up approaches [68]. LandScan is another widely used gridded population data product that has used multiple sources of satellite imagery and other spatial data to produce annual 1km^2 global population grids from 2000 to the present [69].

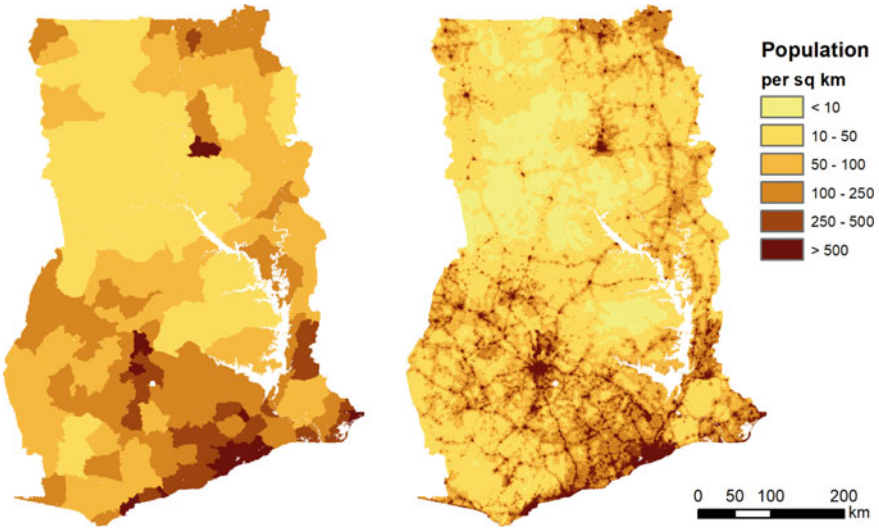


Fig. 7.5 Comparison of two population density datasets for Ghana in 2020. Left: Gridded Population of the World Version 4, which is based on administrative boundaries used for census data collection. Right: WorldPop, which downscales census data based on land cover, roads, and other localized information

7.8 Surface Water and Hydrology

Access to clean water is essential for human health, animal health, and agricultural productivity. However, water also provides habitat for vector and host species and facilitates the transmission of many disease-causing pathogens [70]. Thus, hydrological data are critical for many planetary health applications. Water bodies can be mapped with satellite remote sensing along with other LCLU features, and most LCLU data products include a classification of permanent water bodies such as lakes and large rivers. Understanding how surface water varies over time is also essential. For example, flowing water is not a suitable habitat for vector mosquitoes, but large rivers can provide suitable standing water when their flows decline and leave isolated pools on their floodplains [6]. In flood-prone areas, rising waters are often contaminated by human and animal waste, exposing local populations to a variety of water-borne pathogens [71]. Droughts can also facilitate water-borne pathogen transmission when large groups of people congregate to use the few remaining water sources [72]. More generally, hydrological events like droughts and floods often trigger large-scale human movements and resettlements that facilitate long-distance movement of pathogens and provide novel opportunities for transmission.

Several types of geospatial datasets can provide useful hydrological information for planetary health applications. Gridded elevation datasets are produced for many countries by government mapping agencies, and global elevation products derived from satellite observations are also available. At the most basic level, these data can

be used to identify topographic features such as valley bottoms that are subject to flooding and may serve as locations for water-borne disease transmission or provide larval habitats for mosquitoes [73]. The topographic index, calculated as a function of slope angle and upslope drainage area, is an important input to dynamic hydrological models such as TOPMODEL that can be coupled with mechanistic models of water-associated diseases such as fasciolosis [74]. At coarser grid cell sizes (~10 km or larger) Land Data Assimilations Systems (LDAS) combined gridded meteorological data with other environmental inputs to drive hydrological models that estimate evapotranspiration, soil moisture, and runoff. Various LDAS datasets with different spatial extents, grid cell sizes, and time steps are available, and these data have been used in a variety of disease applications [75–77].

Satellite observations, including passive sensors in the optical and infrared wavelengths and active remote sensing with synthetic aperture radar, can be used to detect and map open water. Surface water is highly variable in locations with pronounced wet and dry seasons, and individual observations are inadequate for characterizing these dynamics. Surface water data products such as the Global Surface Water Explorer [78] and the Global Surface Water Dynamics dataset [79] use time series of Landsat data to map the extent, seasonality, and long-term trends in surface water at 30 m resolution for the entire globe. Global products frequently do not capture smaller water bodies that may serve as larval habitats for mosquitoes or sources of drinking water for humans and livestock. However, they can be used to provide training data for the development of more precise, local maps that include both large and small water bodies [80]. Identifying areas with high seasonal variation in water coverage, including impoundments, wetlands, floodplains, and irrigated areas, is often particularly important in planetary health. These can be identified by analyzing satellite data over multiple seasons and by incorporating topographic variables along with spectral indices [81, 82].

7.9 Synthesis and Conclusions

Researchers and practitioners in planetary health have access to a diverse set of high-quality geospatial data products that characterize environmental factors relevant to human health. Many of these products are global in extent and available at no cost, making them ideal for planetary health assessments and applications in low- and middle-income where locally collected data are sparse. However, it is important to recognize that a global dataset is not necessarily optimal for every location on the Earth [65]. Interpolated meteorological grids and classified LCLU maps have inherent error, and their accuracies can vary considerably among locations. Similarly, the environmental sensitivities of satellite vegetation indices and land surface temperature will vary with the climatic and land surface characteristics in different areas. Before selecting a particular dataset for specific application, potential users should carefully examine the spatial and temporal patterns within their areas of interest to verify that important regional and local features are being captured. If this is not the

case, then it may be necessary to develop bespoke data products that are optimized for the particular region and application [80].

Although planetary health studies frequently incorporate geospatial datasets characterizing climate, LCLU, and human populations, they vary greatly in the specific data used and the manner in which they are applied. A recent systematic review of malaria mapping studies found that the most commonly used covariates were rainfall and temperature [83]. However, the individual studies used a variety of data sources, including ground station measurements, gridded meteorological datasets, satellite vegetation indices, land surface temperature, and satellite precipitation estimates. The degree to which different results are contingent upon differences in the underlying temperature and precipitation data are not well understood. In most cases, the rationale for using a particular source of environmental data is not stated, and decisions are presumably based at least in part on familiarity with particular datasets and ease of data access and use.

This chapter has focused on geospatial environmental data for planetary health applications related to vector-borne and zoonotic infectious diseases. However, geospatial information is also essential for other aspects of planetary health, including natural disasters, food systems and nutrition, and exposure to toxins and pollutants. Timely geospatial data for monitoring meteorological and hydrological variables is essential for monitoring droughts and providing early warning of the risk of food insecurity [84]. Exposure to air pollution is one of the most important global health risks [85], and satellite remote sensing is widely used to obtain spatially explicit measurements of various pollutants [86]. For example, satellite measurements of aerosol optical depth are widely used to estimate ground level concentrations of fine particulate matter generated by combustion of fossil fuels, dust storms and wildfires [87]. There is growing evidence that exposure to greenspace has a variety of health benefits for urban and suburban populations [88], and satellite-based measurements of greenness are commonly used to study these relationships [89]. Although this chapter does not address these topics in detail, many of the data sources that were highlighted in the context of vector-borne and zoonotic diseases are also relevant to these other aspects of planetary health.

Looking forward, more studies on the local accuracies of commonly used geospatial data products would provide evidence to support the choice of geospatial datasets. For example, an accuracy assessment of 20 global precipitation products in Ethiopia found that only three could adequately characterize the spatial extent and severity of historical drought events [90]. An accuracy assessment of multiple gridded climate datasets within the United States found that the most accurate dataset varied by ecoregion [20]. Additionally, comparative analyses of health outcomes based on environmental data from multiple products can help identify the data that are most suitable for specific applications. A West Nile virus risk mapping study compared three predictive models based on land cover and topography data, gridded climate data, and remotely sensed vegetation and moisture indices [91]. Overall accuracy was similar, but the resulting maps based on each dataset exhibited different spatial patterns. A combined model that incorporated variables from all three datasets had the highest overall accuracy. For species range predictions of European tick species, climate

niche models based on an interpolated meteorological dataset had higher accuracies than models based on satellite observations of LST and NDVI from MODIS [92]. Further studies like these will contribute to a broader body of evidence to inform the selection of geospatial environmental datasets for planetary health research and applications.

References

1. Whitmee S, Haines A, Beyrer C, Boltz F, Capon AG, de Souza Dias BF, Ezeh A, Frumkin H, Gong P, Head P (2015) Safeguarding human health in the Anthropocene epoch: report of The Rockefeller Foundation-Lancet Commission on planetary health. Lancet 386(10007):1973–2028
2. Zinsstag J, Schelling E, Waltner-Toews D, Tanner M (2011) From "one medicine" to "one health" and systemic approaches to health and well-being. Prev Vet Med 101(3–4):148–156
3. Charron DF (2012) Ecohealth: origins and approach. In: Charron DF (ed) Ecohealth research in practice. Springer, New York, pp 1–30
4. Gorris ME, Anenberg SC, Goldberg DL, Kerr GH, Stowell JD, Tong D, Zaitchik BF (2021) Shaping the future of science: COVID-19 highlighting the importance of GeoHealth. GeoHealth:e2021GH000412
5. Mordecai EA, Caldwell JM, Grossman MK, Lippi CA, Johnson LR, Neira M, Rohr JR, Ryan SJ, Savage V, Shocket MS (2019) Thermal biology of mosquito-borne disease. Ecol Lett 22(10):1690–1708
6. Smith M, Willis T, Alfieri L, James W, Trigg M, Yamazaki D, Hardy A, Bisselink B, De Roo A, Macklin M (2020) Incorporating hydrology into climate suitability models changes projections of malaria transmission in Africa. Nat Commun 11(1):1–9
7. Parham PE, Waldock J, Christophides GK, Hemming D, Agusto F, Evans KJ, Fefferman N, Gaff H, Gumel A, LaDeau S (2015) Climate, environmental and socio-economic change: weighing up the balance in vector-borne disease transmission. Philos T R Soc B 370(1665):20130551
8. Thomson MC, Ukawuba I, Hershey CL, Bennett A, Ceccato P, Lyon B, Dinku T (2017) Using rainfall and temperature data in the evaluation of national malaria control programs in Africa. Am J Trop Med Hyg 97(Suppl 3):32–45
9. Murdock CC, Evans MV, McClanahan TD, Miazgowicz KL, Tesla B (2017) Fine-scale variation in microclimate across an urban landscape shapes variation in mosquito population dynamics and the potential of *Aedes albopictus* to transmit arboviral disease. PLoS Negl Trop Dis 11(5):e0005640
10. Wimberly MC, Davis JK, Evans MV, Hess A, Newberry PM, Solano-Asamoah N, Murdock CC (2020) Land cover affects microclimate and temperature suitability for arbovirus transmission in an urban landscape. PLoS Negl Trop Dis 14(9):e0008614
11. Snyder RL, Spano D, Duce P (2013) Weather station siting: effects on phenological models. In: Schwartz MD (ed) Phenology: an integrative environmental science. Springer, New York, pp 367–382
12. Fiebrich CA (2009) History of surface weather observations in the United States. Earth-Sci Rev 93(3–4):77–84
13. Venter ZS, Brousse O, Esau I, Meier F (2020) Hyperlocal mapping of urban air temperature using remote sensing and crowdsourced weather data. Remote Sens Environ 242:111791
14. Colston JM, Ahmed T, Mahopo C, Kang G, Kosek M, de Sousa JF, 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
15. Dinku T (2019) Challenges with availability and quality of climate data in Africa. In: Melesse AM, Wossenu A, Senay G (eds) Extreme hydrology and climate variability. Elsevier, pp 71–80

16. Hofstra N, New M, McSweeney C (2010) The influence of interpolation and station network density on the distributions and trends of climate variables in gridded daily data. Clim Dyn 35(5):841–858
17. Bostan P, Heuvelink GB, Akyurek S (2012) Comparison of regression and kriging techniques for mapping the average annual precipitation of Turkey. Int J Appl Earth Obs Geoinf 19:115–126
18. Herrera S, Fernández J, Gutiérrez JM (2016) Update of the Spain02 gridded observational dataset for EURO-CORDEX evaluation: assessing the effect of the interpolation methodology. Int J Climatol 36(2):900–908
19. Fuka DR, Walter MT, MacAlister C, Degaetano AT, Steenhuis TS, Easton ZM (2014) Using the Climate Forecast System Reanalysis as weather input data for watershed models. Hydrol Process 28(22):5613–5623
20. Behnke R, Vavrus S, Allstadt A, Albright T, Thogmartin WE, Radeloff VC (2016) Evaluation of downscaled, gridded climate data for the conterminous United States. Ecol Appl 26(5):1338–1351
21. Zandler H, Haag I, Samimi C (2019) Evaluation needs and temporal performance differences of gridded precipitation products in peripheral mountain regions. Sci Rep 9(1):1–15
22. Harris I, Jones PD, Osborn TJ, Lister DH (2014) Updated high-resolution grids of monthly climatic observations–the CRU TS3. 10 Dataset. Int J Climatol 34(3):623–642
23. Funk C, Peterson P, Landsfeld M, Pedreros D, Verdin J, Shukla S, Husak G, Rowland J, Harrison L, Hoell A (2015) The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. Sci Data 2:150066
24. Verdin A, Funk C, Peterson P, Landsfeld M, Tuholske C, Grace K (2020) Development and validation of the CHIRTS-daily quasi-global high-resolution daily temperature data set. Sci Data 7(1):1–14
25. Fick SE, Hijmans RJ (2017) WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. Int J Climatol 37(12):4302–4315
26. Karger DN, Conrad O, Böhner J, Kawohl T, Kreft H, Soria-Auza RW, Zimmermann NE, Linder HP, Kessler M (2017) Climatologies at high resolution for the earth's land surface areas. Sci Data 4(1):1–20
27. Daly C, Neilson RP, Phillips DL (1994) A statistical topographic model for mapping climatological precipitation over mountainous terrain. J Appl Meteorol 33(2):140–158
28. Xia Y, Mitchell K, Ek M, Sheffield J, Cosgrove B, Wood E, Luo L, Alonge C, Wei H, Meng J, Livneh B, Lettenmaier D, Koren V, Duan Q, Mo K, Fan Y, Mocko D (2012) Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products. J Geophys Res 117(D3)
29. Abatzoglou JT (2013) Development of gridded surface meteorological data for ecological applications and modelling. Int J Climatol 33(1):121–131
30. Ford TE, Colwell RR, Rose JB, Morse SS, Rogers DJ, Yates TL (2009) Using satellite images of environmental changes to predict infectious disease outbreaks. Emerg Infect Dis 15(9):1341–1346
31. Parselia E, Kontoes C, Tsouni A, Hadjichristodoulou C, Kioutsioukis I, Magiorkinis G, Stilianakis NI (2019) Satellite earth observation data in epidemiological modeling of malaria, dengue and West Nile virus: a scoping review. Remote Sens 11(16):1862
32. Wimberly MC, de Beurs KM, Loboda TV, Pan WK (2021) Satellite observations and malaria: new opportunities for research and applications. Trends Parasitol 37(6):525–537
33. Tucker CJ (1979) Red and photographic infrared linear combinations for monitoring vegetation. Remote Sens Environ 8:127–150
34. Pettorelli N, Vik JO, Mysterud A, Gaillard J-M, Tucker CJ, Stenseth NC (2005) Using the satellite-derived NDVI to assess ecological responses to environmental change. Trends Ecol Evol 20(9):503–510
35. Merkle JA, Cross PC, Scurlock BM, Cole EK, Courtemanch AB, Dewey SR, Kauffman MJ (2018) Linking spring phenology with mechanistic models of host movement to predict disease transmission risk. J Appl Ecol 55(2):810–819

36. Chuang TW, Wimberly MC (2012) Remote sensing of climatic anomalies and West Nile virus incidence in the northern Great Plains of the United States. PLoS ONE 7(10):e46882
37. Klisch A, Atzberger C (2016) Operational drought monitoring in Kenya using MODIS NDVI time series. Remote Sens 8(4):267
38. Jiang Z, Huete A, Didan K, Miura T (2008) Development of a two-band enhanced vegetation index without a blue band. Remote Sens Environ 112(10):3833–3845
39. Huete AR (1988) A soil-adjusted vegetation index (SAVI). Remote Sens Environ 25(3):295–309
40. Thomson M, Connor S, Milligan P, Flasse S (1997) Mapping malaria risk in Africa: What can satellite data contribute? Parasitol Today 13(8):313–318
41. Rogers D, Randolph S (1991) Mortality rates and population density of tsetse flies correlated with satellite imagery. Nature 351(6329):739–741
42. Midekisa A, Senay G, Henebry GM, Semuniguse P, Wimberly MC (2012) Remote sensing-based time series models for malaria early warning in the highlands of Ethiopia. Malar J 11:165
43. Brown HE, Diuk-Wasser MA, Guan Y, Caskey S, Fish D (2008) Comparison of three satellite sensors at three spatial scales to predict larval mosquito presence in Connecticut wetlands. Remote Sens Environ 112(5):2301–2308
44. Hilker T, Lyapustin AI, Tucker CJ, Sellers PJ, Hall FG, Wang Y (2012) Remote sensing of tropical ecosystems: atmospheric correction and cloud masking matter. Remote Sens Environ 127:370–384
45. Li Z-L, Tang B-H, Wu H, Ren H, Yan G, Wan Z, Trigo IF, Sobrino JA (2013) Satellite-derived land surface temperature: Current status and perspectives. Remote Sens Environ 131:14–37
46. Liu J, Hagan DFT, Liu Y (2020) Global land surface temperature change (2003–2017) and its relationship with climate drivers: airs, modis, and era5-land based analysis. Remote Sens 13(1):44
47. Cao J, Zhou W, Zheng Z, Ren T, Wang W (2021) Within-city spatial and temporal heterogeneity of air temperature and its relationship with land surface temperature. Landsc Urban Plan 206:103979
48. Vancutsem C, Ceccato P, Dinku T, Connor SJ (2010) Evaluation of MODIS land surface temperature data to estimate air temperature in different ecosystems over Africa. Remote Sens Environ 114(2):449–465
49. Wan Z (2014) New refinements and validation of the collection-6 MODIS land-surface temperature/emissivity product. Remote Sens Environ 140:36–45
50. Hulley GC, Malakar NK, Islam T, Freepartner RJ (2017) NASA's MODIS and VIIRS land surface temperature and emissivity products: a long-term and consistent earth system data record. IEEE J Sel Top Appl Earth Obs Remote Sens 11(2):522–535
51. Malakar NK, Hulley GC, Hook SJ, Laraby K, Cook M, Schott JR (2018) An operational land surface temperature product for Landsat thermal data: methodology and validation. IEEE Trans Geosci Remote Sens 56(10):5717–5735
52. Hulley GC, Göttsche FM, Rivera G, Hook SJ, Freepartner RJ, Martin MA, Cawse-Nicholson K, Johnson WR (2021) Validation and quality assessment of the ECOSTRESS level-2 land surface temperature and emissivity product. IEEE Trans Geosci Remote Sens 60:1–23
53. Chakraborty T, Lee X, Ermida S, Zhan W (2021) On the land emissivity assumption and Landsat-derived surface urban heat islands: a global analysis. Remote Sens Environ 265:112682
54. Sun Q, Miao C, Duan Q, Ashouri H, Sorooshian S, Hsu KL (2018) A review of global precipitation data sets: Data sources, estimation, and intercomparisons. Rev Geophys 56(1):79–107
55. Kirschbaum DB, Huffman GJ, Adler RF, Braun S, Garrett K, Jones E, McNally A, Skofronick-Jackson G, Stocker E, Wu H (2017) NASA's remotely sensed precipitation: a reservoir for applications users. Bull Am Meteorol Soc 98(6):1169–1184
56. Nguyen P, Ombadi M, Sorooshian S, Hsu K, AghaKouchak A, Braithwaite D, Ashouri H, Thorstensen AR (2018) The PERSIANN family of global satellite precipitation data: A review and evaluation of products. Hydrol Earth Syst Sci 22(11):5801–5816

57. Xie P, Joyce R, Wu S, Yoo S-H, Yarosh Y, Sun F, Lin R (2017) Reprocessed, bias-corrected CMORPH global high-resolution precipitation estimates from 1998. J Hydrometeorol 18(6):1617–1641
58. Adler RF, Sapiano MR, Huffman GJ, Wang J-J, Gu G, Bolvin D, Chiu L, Schneider U, Becker A, Nelkin E (2018) The Global Precipitation Climatology Project (GPCP) monthly analysis (new version 2.3) and a review of 2017 global precipitation. Atmosphere 9 (4):138
59. Sulla-Menashe D, Gray JM, Abercrombie SP, Friedl MA (2019) Hierarchical mapping of annual global land cover 2001 to present: the MODIS collection 6 land cover product. Remote Sens Environ 222:183–194
60. DiMiceli C, Townshend J, Carroll M, Sohlberg R (2021) Evolution of the representation of global vegetation by vegetation continuous fields. Remote Sens Environ 254:112271
61. Buchhorn M, Lesiv M, Tsendbazar N-E, Herold M, Bertels L, Smets B (2020) Copernicus global land cover layers—collection 2. Remote Sens 12(6):1044
62. Hansen MC, Potapov PV, Moore R, Hancher M, Turubanova SA, Tyukavina A, Thau D, Stehman SV, Goetz SJ, Loveland TR, Kommareddy A, Egorov A, Chini L, Justice CO, Townshend JR (2013) High-resolution global maps of 21st-century forest cover change. Science 342(6160):850–853
63. Potapov P, Turubanova S, Hansen MC, Tyukavina A, Zalles V, Khan A, Song X-P, Pickens A, Shen Q, Cortez J (2022) Global maps of cropland extent and change show accelerated cropland expansion in the twenty-first century. Nat Food 3(1):19–28
64. Gong P, Li X, Wang J, Bai Y, Chen B, Hu T, Liu X, Xu B, Yang J, Zhang W (2020) Annual maps of global artificial impervious area (GAIA) between 1985 and 2018. Remote Sens Environ 236:111510
65. Tulbure MG, Hostert P, Kuemmerle T, Broich M (2021) Regional matters: on the usefulness of regional land-cover datasets in times of global change. Remote Sens Ecol Conserv 8(3):272–283
66. Stevens FR, Gaughan AE, Linard C, Tatem AJ (2015) Disaggregating census data for population mapping using random forests with remotely-sensed and ancillary data. PLoS ONE 10(2):e0107042
67. Wardrop N, Jochem W, Bird T, Chamberlain H, Clarke D, Kerr D, Bengtsson L, Juran S, Seaman V, Tatem A (2018) Spatially disaggregated population estimates in the absence of national population and housing census data. Proc Natl Acad Sci USA 115(14):3529–3537
68. Tatem AJ (2017) WorldPop, open data for spatial demography. Sci Data 4(1):1–4
69. Dobson JE, Bright EA, Coleman PR, Durfee RC, Worley BA (2000) LandScan: a global population database for estimating populations at risk. Photogramm Eng Remote Sens 66(7):849–857
70. Yang K, LeJeune J, Alsdorf D, Lu B, Shum CK, Liang S (2012) Global distribution of outbreaks of water-associated infectious diseases. PLoS Negl Trop Dis 6(2):e1483
71. Cann K, Thomas DR, Salmon R, Wyn-Jones A, Kay D (2013) Extreme water-related weather events and waterborne disease. Epidemiol Infect 141(4):671–686
72. Stanke C, Kerac M, Prudhomme C, Medlock J, Murray V (2013) Health effects of drought: a systematic review of the evidence. PLoS Curr June 5
73. Cohen JM, Ernst KC, Lindblade KA, Vulule JM, John CC, Wilson ML (2008) Topography-derived wetness indices are associated with household-level malaria risk in two communities in the western Kenyan highlands. Malar J 7:40
74. Beltrame L, Dunne T, Vineer HR, Walker JG, Morgan ER, Vickerman P, McCann CM, Williams DJ, Wagener T (2018) A mechanistic hydro-epidemiological model of liver fluke risk. J R Soc Interface 15(145):20180072
75. Colston JM, Zaitchik B, Kang G, Yori PP, Ahmed T, Lima A, Turab A, Mduma E, Shrestha PS, Bessong P (2019) Use of earth observation-derived hydrometeorological variables to model and predict rotavirus infection (MAL-ED): a multisite cohort study. Lancet Planet Health 3(6):e248–e258
76. Shaman J, Day JF, Komar N (2010) Hydrologic conditions describe West Nile virus risk in Colorado. Int J Environ Res Public Health 7(2):494–508

77. Davis JK, Vincent GP, Hildreth MB, Kightlinger L, Carlson C, Wimberly MC (2018) Improving the prediction of arbovirus outbreaks: a comparison of climate-driven models for West Nile virus in an endemic region of the United States. Acta Trop 185:242–250
78. Pekel J-F, Cottam A, Gorelick N, Belward AS (2016) High-resolution mapping of global surface water and its long-term changes. Nature 540(7633):418–422
79. Pickens AH, Hansen MC, Hancher M, Stehman SV, Tyukavina A, Potapov P, Marroquin B, Sherani Z (2020) Mapping and sampling to characterize global inland water dynamics from 1999 to 2018 with full Landsat time-series. Remote Sens Environ 243:111792
80. Worden J, de Beurs KM (2020) Surface water detection in the Caucasus. Int J Appl Earth Obs Geoinf 91:102159
81. Midekisa A, Senay GB, Wimberly MC (2014) Multisensor earth observations to characterize wetlands and malaria epidemiology in Ethiopia. Water Resour Res 50(11):8791–8806
82. Catry T, Li Z, Roux E, Herbreteau V, Gurgel H, Mangeas M, Seyler F, Dessay N (2018) Wetlands and malaria in the Amazon: guidelines for the use of synthetic aperture radar remote-sensing. Int J Environ Res Public Health 15(3):468
83. Odhiambo JN, Kalinda C, Macharia PM, Snow RW, Sartorius B (2020) Spatial and spatio-temporal methods for mapping malaria risk: a systematic review. BMJ Glob Health 5(10):e002919
84. Senay G, Velpuri NM, Bohms S, Budde M, Young C, Rowland J, Verdin J (2015) Drought monitoring and assessment: remote sensing and modeling approaches for the famine early warning systems network. In: Parron P, Baldassarre GD (eds) Hydro-meteorological hazards, risks and disasters. Elsevier, pp 233–262
85. Lelieveld J, Pozzer A, Pöschl U, Fnais M, Haines A, Münzel T (2020) Loss of life expectancy from air pollution compared to other risk factors: a worldwide perspective. Cardiovasc Res 116(11):1910–1917
86. Anenberg SC, Bindl M, Brauer M, Castillo JJ, Cavalieri S, Duncan BN, Fiore AM, Fuller R, Goldberg DL, Henze DK (2020) Using satellites to track indicators of global air pollution and climate change impacts: lessons learned from a NASA-supported science-stakeholder collaborative. GeoHealth 4(7):e2020GH000270
87. Chu Y, Liu Y, Li X, Liu Z, Lu H, Lu Y, Mao Z, Chen X, Li N, Ren M (2016) A review on predicting ground PM2. 5 concentration using satellite aerosol optical depth. Atmosphere 7(10):129
88. Frumkin H, Bratman GN, Breslow SJ, Cochran B, Kahn PH Jr, Lawler JJ, Levin PS, Tandon PS, Varanasi U, Wolf KL (2017) Nature contact and human health: a research agenda. Environ Health Persp 125(7):075001–075001
89. Labib S, Lindley S, Huck JJ (2020) Spatial dimensions of the influence of urban green-blue spaces on human health: a systematic review. Environ Res 180:108869
90. Degefu MA, Bewket W, Amha Y (2022) Evaluating performance of 20 global and quasi-global precipitation products in representing drought events in Ethiopia I: visual and correlation analysis. Weather Clim Extrem 35:100416
91. Hess A, Davis JK, Wimberly MC (2018) Identifying environmental risk factors and mapping the distribution of West Nile virus in an endemic region of North America. GeoHealth 2(12):395–409
92. Estrada-Peña A, Estrada-Sánchez A, Estrada-Sánchez D, de la Fuente J (2013) Assessing the effects of variables and background selection on the capture of the tick climate niche. Int J Health Geogr 12(1):1–13

Dr. Michael C. Wimberly is an ecologist and geographer who studies environmental change and its effects on ecosystem and human health. He is a professor at the University of Oklahoma, where he conducts research on the influences of climate and land use on malaria and West Nile virus. He applies the results to develop risk maps and early warning systems for mosquito-borne diseases.

Chapter 8
Delineating Zones of Disease Diffusion from the Amenity-Sharing Network in Peninsular Malaysia

Wei Chien Benny Chin

Abstract Disease diffusion happens when infected and susceptible individuals move around and closely interact with each other. The boundary of human movement can be found by analysing the locations of amenities and people. The human movement boundaries can be considered as the zones of disease diffusion, which is essential for the establishment of disease control measures. During the COVID-19 pandemic, Malaysia had gone through a series of nationwide Movement Control Orders (MCO)—multiple phases of country/city lockdown measures—starting from March 2020. One key order during MCO was to restrict the travel distance to a 10 km radius from household locations. However, this movement restriction can only eliminate/reduce the long-range disease spreading (relocation) but not the disease diffusion within a local area (expansion). The disease can still be transmitted within a neighbourhood and between closely located or densely interacted neighbourhoods. In other words, people who visited the same region of an outbreak cluster would still expose to the disease. This study analyses the boundaries of densely connected neighbourhoods based on the amenity-sharing relationships, i.e., the disease diffusion zones, and identify the vulnerable locations in terms of spreading and contracting diseases, i.e., the centre(s) of zones. Using Peninsular Malaysia as a case study, a four-step framework was established, which utilised the open-data materials and open-source software. The analysis results from the case study showed that while some of the zones resembled the administrative boundaries, a considerable proportion of the zones extended the city area to the neighbouring urbanised area while some zones split a city into separated zones. These identified zones function as a reference for future policymaking on disease control issues.

Keywords Complex network analysis · MapEquation · Epidemic · Neighbourhood · Road network · Accessibility · Movement control orders · Semi-lockdown

W. C. B. Chin (✉)
Department of Geography, National University of Singapore, Singapore 117570, Singapore
e-mail: wcchin@nus.edu.sg

T.-H. Wen et al. (eds.), *Earth Data Analytics for Planetary Health*, Atmosphere, Earth, Ocean & Space, https://doi.org/10.1007/978-981-19-8765-6_8

8.1 Introduction

Contagious diseases spread between people, and people move between spaces. Disease diffusion occurs when infected cases increase and the people move throughout space and time [14, 35]. There are two major types of disease diffusion: expansion and relocation. Expansion diffusion describes the spread of cases from one outbreak location to the spatially neighbouring area, while relocation diffusion captures the jump of cases from one origin to a distant location [24]. The spatial spreading of diseases usually happens as a combination of the two diffusion processes [22], e.g., the hierarchical diffusion indicates the situation when the disease spread outwards from one city centre (expansion) to the neighbouring satellite towns, it also jumps to some other city centres and expands from there. This leads to the discussion of the spatial interaction boundary of people's activities in their daily lives, i.e., the geographical zones of human activity [8, 9, 28]. When two locations belong to the same zone of human activity, it indicates that the interaction of people between the two places is dense, thus providing opportunities for disease to spread from one location to another. Furthermore, as described by the hierarchical diffusion, the centres of zones may be the location that is more susceptible to diseases and which can initiate the local expansion diffusion [5, 15, 17, 26]. Therefore, delineating the zones of human activity and searching the centres of zones is crucial to disease control.

The analysis of geographical zoning is discussed in the literature on activity spaces and functional regions. The activity space of a person is the combination of places where his/her daily routine occurs, e.g., the location of his/her home and other amenities, including workplace or schools and the supermarket or restaurant that he/she usually goes to purchase groceries or foods. The activity space for a set of residents can be delineated by the boundary of the accumulated daily activity spaces of the population [13, 25, 37]. A functional region describes the boundary of geographical spaces that contains both the supplies and demands—the locations of amenities that provide services (e.g., shops, hospitals, parks, schools, banks, etc.) and the servicing area of these amenities (e.g., the residential areas) [19, 21, 32]. Both concepts show a simple idea—people move between their home locations and places where they can access some services or functions. The locations of amenities are the places where people physically interact or closely contact [4]. In other words, these places can be the location where the transmission of disease pathogens occurs, and which can lead to the diffusion of disease through expansion and relocation. This highlights the importance of analysing the activity spaces and functional zones in the context of disease diffusion since the zoning of human activity and interactions can provide insights into the boundary of disease expansion diffusion and the centre point of hierarchical diffusion [8, 21].

COVID-19 appeared in Malaysia during the early stage of the global pandemic. The first incident of COVID-19 cases happened in January 2020 when three Chinese citizens travelled to Malaysia from an international conference event held in Singapore where they were infected [33, 40]. Although these initial cases were quarantined from the beginning and did not cause a super-spreading situation, these cases were

followed by a lot of other imported cases and a mass local gathering and super-spreading event [27] which eventually lead to the nationwide lockdown in March 2020—the Malaysian Movement Control Order (MCO) [3]—that continues with several phases of National Recovery Plan (NRP) and a series of state-level/national-level lockdown until April 2022 [40–42]. During MCO, people were not allowed to cross states and only one person per household was allowed to leave their home for buying groceries or getting essential services within a 10 km radius [1]. The restricted number of persons was relaxed in the various phases of NRP and other extended MCO periods. This 10 km radius moving distance aligns with the regular daily commuting distance by public transit—i.e., on average 7.1 km in a single direction, and 39% of commuters travel over 12 km [29]—and it is about half of the distance for overall regular commuting distance (19.54 km) by cars (70.20% of the total commuters) [30] in Kuala Lumpur, the capital city of Malaysia. In other words, a 10 km moving radius is sufficient for people to access essential amenities for their daily life.

In the Malaysian context, for both the regular situation and the MCO period, people usually (or can only) travel within a 10 km radius from their home to access amenities. During a disease outbreak event, diseases can spread between people who share the amenities when they visit the same location (co-visit). Thus, delineating the zones of disease diffusion is crucial in the decision and policymaking process for setting up disease control measures since it indicates the boundary of the impact area of the high-risk locations [18, 36]. The development of computing techniques and the open data warehouses of Geographical Information System (GIS) provides a workable solution to explore and delineate the boundary of human activity zones from the locations of amenities and people. The OpenStreetMap platform gathered and supplied the global-wide GIS data that contains the road network layout and the point-of-interest (POI) data [12]. POI data indicates the locations of various urban functions and services, e.g., commercial, business, education, civic related, health, religious, outdoors, and recreational locations. Road networks reveal more locations of human activity, including residential, industrial, and transportation areas. Therefore, the two can be considered as surrogate for the location of amenities and people. The remaining research gap would then be the question of how to identify the zones of disease diffusion. This study demonstrates a four-step framework to analyse the ‘zones’ of human interaction activity based on the amenity-sharing network. The study framework focused on the spatial interactions between places—i.e., the sharing of amenities between neighbourhoods—and a series of geospatial network analysis tools for the identification and discussion of the disease diffusion zones and city centres.

8.2 Materials and Methods

8.2.1 Study Site

Located in Southeast Asia, Malaysia is a country composed of two parts: East Malaysia covers 60% (about 198, 172 km^2) of the total land area, and West Malaysia covers about 40% (about 132, 631 km^2). The South China Sea separate the two parts. Located on the island of Borneo, East Malaysia contains two states (Sabah and Sarawak) and one federal territory (Labuan). On the other side, West Malaysia (i.e., Peninsular Malaysia) is a peninsula that shares a land border with Thailand to the North; the island of Singapore lies across the Strait of Johor from the southern tip of the peninsula. According to the 2010 census data [10], the total population living in Peninsular Malaysia is around 22 million, which is about 80% of the total population of Malaysia (27.7 million), leaving around 20% of the population living in East Malaysia. In this study, we focused only the Peninsular Malaysia (West Malaysia) because East Malaysia is geographically separated, and West Malaysia contains a larger proportion of the population and urban area.

Peninsular Malaysia contains eleven states and two federal territories (FTs), which can be divided into the two sides of the peninsula: the west coast (eight states and two FTs) and the east coast (three states). The eight states on the west coast are (from north to south): Perlis, Kedah, Pulau Pinang (or Penang), Perak, Selangor, Negeri Sembilan, Melaka (or Malacca), and Johor. The two FTs—Kuala Lumpur (the capital city) and Putrajaya (a financial district)—locates within the Selangor state area. On the east coast, the other three states are (from north to south) Kelantan, Terengganu, and Pahang. The west coast and the east coast are divided by the Titiwangsa Range which runs from north to south along the peninsula. This geographical division has significant effects on economic development and population settlement—the main economic development corridor runs on the west coast. While the area sizes are similar between the west coast (52% of the Peninsular Malaysia area) and the east coast (48%), the west coast of Peninsular Malaysia is more urbanised and populated than the east coast. Among the population in Peninsular Malaysia (22 million), 82% live on the west coast and only 18% on the east coast [10]. Furthermore, the two FTs have even more extensive urbanisation and development progress. The two FTs together are the home of 7% of the population while the area is only about 0.2% of Peninsular Malaysia. This notable urban–rural hierarchy indicates the need for a separated exploration and discussion among the various parts of Peninsular Malaysia.

8.2.2 Datasets

Two datasets were used in this study: (1) road network and (2) POI. Both datasets were downloaded from Geofabrik—the platform providing the extracted OpenStreetMap archive files [12]. Figure 8.1 shows the size of the area, the density of the POI and road

network data by states/FTs. The states/FTs are sorted by the area sizes in ascending order. The 2 FTs have the smallest area. In addition to the 2 FTs, three of the states also have a small area size (Perlis, Pulau Pinang, and Melaka). While the area is small, the five states/FTs hold a higher density of the three measurements (POI, road features, and junctions) than the other places. In contrast, the largest state (Pahang) has the least density of the three measurements.

Figure 8.2 shows the kernel density estimation for the distribution of POI and junctions. Seven clusters of POIs can be observed (Fig. 8.2a): five on the west coast, one at the southern tip and one at the northern of the east coast. From north to south, the five clusters on the west coast are: (1) Kedah, (2) Pulau Pinang, (3) Ipoh, Perak, (4) Kuala Lumpur–Putrajaya–Selangor, and (5) Melaka. The cluster at the southern tip is (6) Johor Bahru, Johor. The cluster on east coast is (7) Kota Bharu, Kelantan.

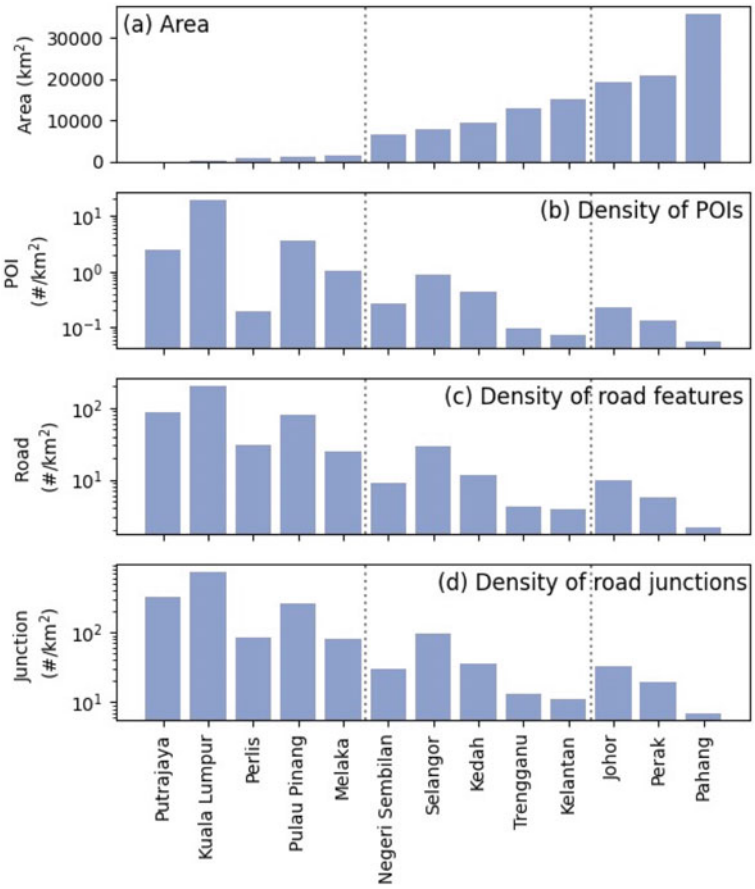


Fig. 8.1 A summary of the data by states/FTs: **a** the area size, **b** density of POI, **c** density of road features (centroids), and **d** density of road junctions

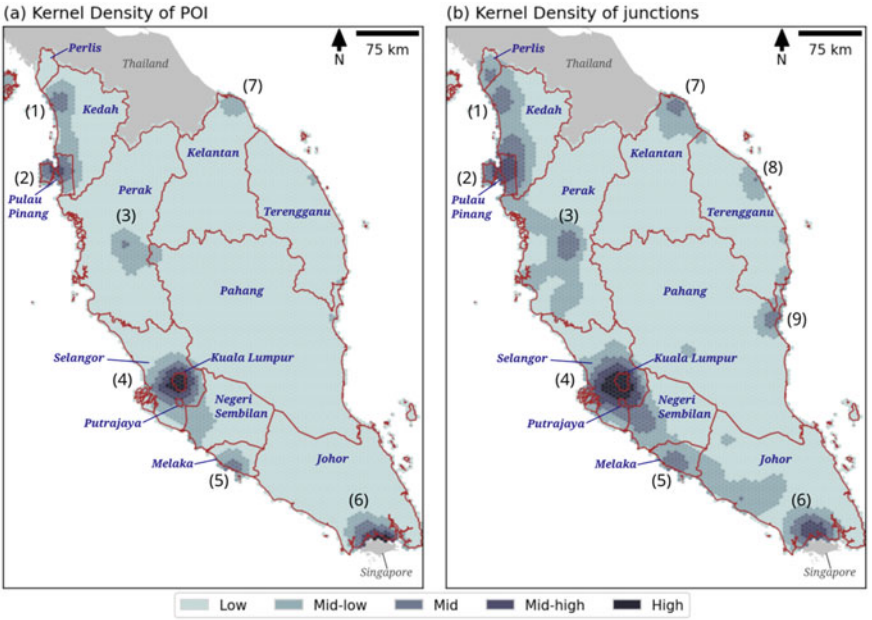


Fig. 8.2 The kernel density estimation of **a** POI and **b** road junctions. The red lines highlight the states/FTs boundaries

The kernel density of road junctions (Fig. 8.2b) is the expansion of the POI clusters to the surrounding areas. Unlike the POI which captures only places of attraction, road junctions could capture other types of human activities, e.g., residential, industrial, or transportation. At the north of the west coast, the Perlis cluster appears and connects to the Kedah–Pulau Pinang clusters. The Ipoh cluster also extends in three directions: one to the north that connects to Pulau Pinang, one to the west that reaches the coastal area, and one to the south that connects to the Selangor boundary. The expansion to the south from the Kuala Lumpur–Putrajaya–Selangor cluster connects to the Melaka cluster and continues extending to the Johor Bahru cluster. On the east coast, two more clusters are found at (8) Kuala Terengganu, Terengganu, and (9) Kuantan, Pahang.

8.2.3 Study Framework and Analyses

The framework of this study is shown in Fig. 8.3. The framework contains two parts of network data processing and two parts of analysis. The two network data processing parts involve two levels of networks: the road network and the amenity-sharing network (the higher-level spatial interactions capturing the human movement structure). Part 1 aims to identify the neighbourhoods from the raw data—road

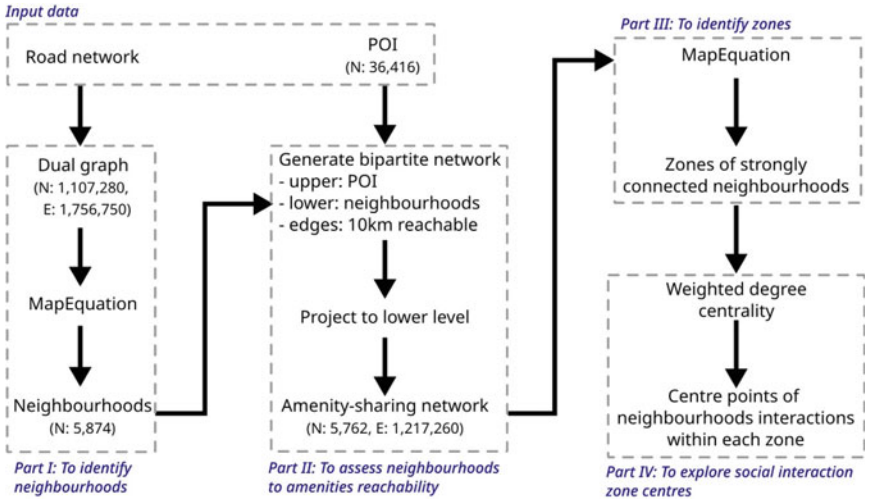


Fig. 8.3 The study framework

network. Part 2 aims to generate the amenity-sharing network from the neighbourhoods (as lower nodes) and POI (as upper nodes) bipartite network. The goals of the third and fourth parts are to identify the zones between neighbourhoods using the amenity-sharing network and to extract important area centres within the zones.

The road network data were prepared and provided on Geofabrik's website,[1] which were extracted from OpenStreetMap (on 26 July 2021) [12]. The data were clipped for the Peninsula Malaysia region.[2] The Peninsular Malaysia administrative boundary data were downloaded and extracted from the database of Global Administrative Areas (GADM).[3] For part 1, the dual graph was generated from the road network data, i.e., road segments were nodes, and intersections (junctions) were used to generate links. In total, 1.107 million nodes (road segments) and 1.757 million links were found in the road network dataset. A community detection method based on MapEquation (Infomap[4]) [34] was used to group densely connected nodes (i.e., community). Each group of densely connected nodes (road segments) was considered a 'neighbourhood', i.e., these road segments were more connected to each other than to the nodes in other groups. These neighbourhoods were then used as the study units in the following analyses.

In part 2, a total of 36, 416 POI points was included. These POI data were downloaded from the Geofabrik website [12]. There are several types of POIs, including commercial, business, education, civic related, health, religious, outdoors, and recreational types. Although the POIs data was strongly diverse in types and weights (e.g.,

[1] Geofabrik's free download server: https://download.geofabrik.de/asia.html.

[2] All spatial data were projected to EPSG:3375 projection from EPSG:4326.

[3] GADM database, version 3.4: https://gadm.org/download_country.html.

[4] Infomap, version 1.7.0: https://www.mapequation.org/infomap/.

capacity, service area), this dataset was used to capture the locations of amenities where people might gather or interact, thus the types and weights were not considered in this study. In the construction of the bipartite network (neighbourhoods as lower nodes and POI as upper nodes), the pairs of neighbourhoods and POI with Manhattan distance below 10 km were identified and converted to links. In other words, the bipartite network indicated the reachable relationship between POIs and neighbourhoods within a 10 km travel.

After the bipartite network was generated, it was projected to the co-occurrence network of the lower nodes—i.e., the network where the nodes were the neighbourhoods (lower nodes) and the links indicated the relationship that the two neighbourhoods shared (co-visited) at least one amenity (POI), and the weight of links indicated the number of shared amenities. In other words, the bipartite network shows the accessible amenities from each neighbourhood, and the amenity-sharing network shows the potential interactions between neighbourhoods. When two neighbourhoods shared a lot of amenities, the frequency of the people visiting the same set of amenities is high, hence the possibility of the people living in the two neighbourhoods having physical interactions or close contact is high. This can lead to disease spreading events.

Using the amenity-sharing network, two analyses (parts 3 and 4) were performed. In part 3, the Infomap community detection algorithm was applied to the amenity-sharing network (nodes were neighbourhoods and links were co-visit frequency). Here, the purpose was to identify the strongly interconnected zones of neighbourhoods, i.e., zones of inseparable neighbourhoods that co-visit to many amenities, and which indicated a high frequency of potential physical interactions. Since the community detection algorithm identifies chains of nodes that are densely connected, the nodes in different parts of a community may not be immediately connected, i.e., not sharing any amenity. Although the people in these neighbourhoods may not directly share any amenity, they are linked to a middle neighbourhood or a set of chained neighbourhoods. These relationships are important in disease spreading analysis because the neighbourhoods in the middle can become the middle place (namely, the zone centre) where the expansion diffusion may have started. From the view of disease control and management, the zones represent clusters of neighbourhoods that share a large number of amenities through a transition chain in the amenity-sharing network structure, i.e., it would be difficult to split the neighbourhoods within the same zone as the residents would go to similar sets of amenities for their daily activities. Therefore, this zoning result could also be understood as disease control zones.

In part 4, a weighted degree analysis was done for several large zones. Since the weights of links in the amenity-sharing network show the frequency of the ‘sharing’ relationship between pairs of neighbourhoods, the weighted degree of a node indicates the aggregated frequencies of interactions for each neighbourhood. In other words, if a node’s weighted degree is high, it means that the people in this neighbourhood have co-visited many POIs with a large number of other neighbourhoods—the residents were exposed to a more dangerous situation in terms of both density and diversity [5]. Because of the strong urban–rural hierarchy between different parts

of Peninsular Malaysia, the comparison of weighted degrees between zones in the country was not useful—all high degree nodes would be found at/near the capital city, causing the other regional centres to be neglected. Therefore, the local comparison is more useful to observe the regional zone centres of neighbourhoods—the local high degree nodes.

In summary, this analysis framework was designed to explore the potential physical interaction structure based on the road network and POI data. The road network reflects the location of human activity, whereas the location of amenities indicates the supplies of daily life activities. We identify the neighbourhoods as an intermediate output and then based on the amenity reachable relationship, the amenity-sharing network is generated. This network could then be used for zone identification and weighted degree centrality analysis.

8.3 Analyses and Results

8.3.1 Part I: Delineate Neighbourhoods from Road Network

The community detection analysis was done using MapEquation on the dual graph of the road network. In summary, the algorithm identified 13 levels of communities; some communities stopped at the 5th level—these communities could not be divided into a lower level. Therefore, we use the 5th level communities to capture the area of neighbourhoods. The centroid of the road segments in each neighbourhood was used to calculate the location of the neighbourhood. The convex hull of the road segments' centroid was identified, and the area was calculated. Figure 8.4 shows the spatial distribution of the convex hulls in Peninsular Malaysia and a focus in Selangor with the two FTs highlighted.

Figure 8.5a shows the number of neighbourhoods in each state/FT. The states/FTs (on the horizontal axis) were arranged in ascending order based on the number of neighbourhoods. The first two states with the least neighbourhoods were small states or FTs (Putrajaya and Perlis were respectively the smallest and third-smallest states/FTs, see Fig. 8.1), whereas the third and fourth states (Kelantan and Terengganu) were the fourth- and fifth-largest states (see Fig. 8.1). The other two small area states—Pulau Pinang and Melaka—contained a moderate number of neighbourhoods. The capital city (Kuala Lumpur) has the third-largest number of neighbourhoods (see Fig. 8.4b). These indicated that the number of identified neighbourhoods was not correlated with the area size.

Figure 8.5b presents the boxplot distribution for the area sizes of each neighbourhood, grouped by the states/FTs. Overall, the median size of the neighbourhood was about 0.95 km^2, and the 25th to 75th percentile range was 0.23 km^2 and 5.35 km^2, respectively. Some extra small neighbourhoods were identified in Kuala Lumpur, Perak, and Selangor—the top three states/FTs with the largest number of neighbourhoods. The small size of the neighbourhood indicated that the urban form was

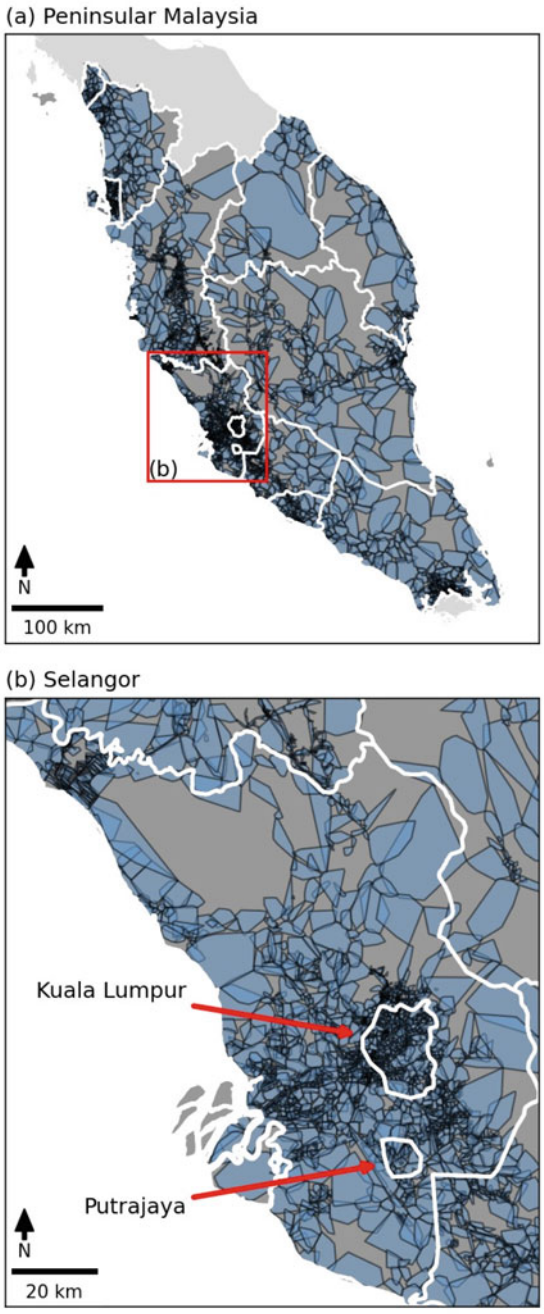


Fig. 8.4 The distribution of the neighbourhoods (convex hulls) in **a** Peninsular Malaysia, **b** Selangor state and the two FTs. The white lines highlight the states/FTs boundaries. The convex hulls in blue indicate neighbourhoods; the grey colour area is an empty land parcel that does not form any neighbourhood, i.e., could be rural, waterbody, forest, or only have a single road segment

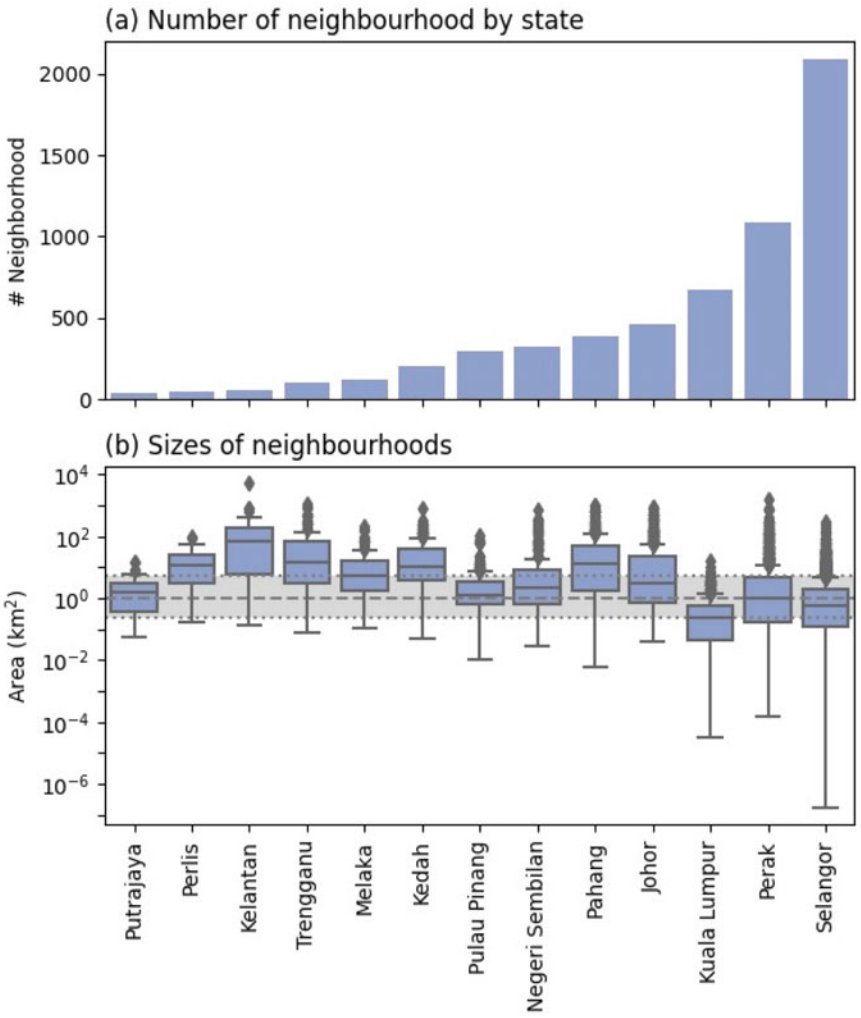


Fig. 8.5 The distribution of **a** the number of neighbourhoods by states/FT and **b** the size of each neighbourhood by states/FT. The dashed line in **b** is the median of all areas, and the grey area indicates the range of 25th to 75th percentiles

denser, and the large numbers indicated a larger proportion of urbanised areas (see Fig. 8.4b for the capital city and surrounding area). On the other hand, the area size of neighbourhoods in five states—Perlis, Kelantan, Terengganu, Kedah, and Pahang—tended to be larger than in other states. Three of these states (Kelantan, Terengganu, and Pahang) are located on the east coast, and the other two (Perlis and Kedah) were at the northern tip of Peninsular Malaysia. The large area size of neighbourhoods indicated a sparse distribution and sprawling type of urban/rural structure (see Fig. 8.4a). This indicated that the activity space comes in diverse sizes for different

states. Thus, some states may need a larger moving distance to access daily services as the neighbourhood size itself is larger than others.

8.3.2 *Part II: Explore Reachability of Neighbourhood Amenities*

A two-mode bipartite network was generated in this study to capture the reachability between neighbourhoods and amenities. In particular, the neighbourhoods were linked to the POIs that fall within a 10 km Manhattan distance range. Figure 8.6 shows the number of linked neighbourhoods for each POI and the number of linked POIs for each neighbourhood in the 13 states/FTs. The number of neighbourhoods each POI is linked to can be viewed as the serving neighbourhood size for each amenity (Fig. 8.6a). The number of linked POIs for each neighbourhood (Fig. 8.6b) can be understood as the number of amenities that can be accessed by each neighbourhood.

Overall, the two boxplots presented a similar pattern. The amenity and neighbourhoods in Kuala Lumpur had the largest size of servicing neighbourhoods and reachable amenities, respectively. In Fig. 8.6a, Selangor and Putrajaya were higher than the global median, showing that the amenity in the two states/FTs would have to serve a larger population than many other places. In Fig. 8.6b, in addition to the two states/FTs, Pulau Pinang and Johor also showed a distribution that was slightly higher than the global median, this indicated that these four states/FTs would have a larger group of reachable amenities. In both box plots, the three east coast states (Kelantan, Terengganu, and Pahang) showed a lower distribution—most of the POIs and neighbourhoods in the three states had fever links. These findings indicated that the differences between states/FTs in Malaysia were large and this heterogeneity should not be neglected in the following analyses. Furthermore, this also indicated that to optimise the effectiveness of movement control orders and to ensure the accessibility to essential services/goods, different moving distance needs to be set up according to the reachability to amenities.

8.3.3 *Part III: Identify Zones of Disease Diffusion*

The amenity-sharing network was constructed after the projection of the neighbourhood-amenity bipartite network to the neighbourhood (lower node) level. Using the MapEquation algorithm, a total of 208 zones were identified based on the amenity-sharing network. Among these zones, 155 were isolated neighbourhoods (single node communities), 8 were composed of only two neighbourhoods (formed a line instead of an area), and the other 45 communities held at least three neighbourhoods. In the following analysis, we focused on the zones that had more than 100 neighbourhoods—the top-11 zones. Figure 8.7 shows the number of neighbourhoods

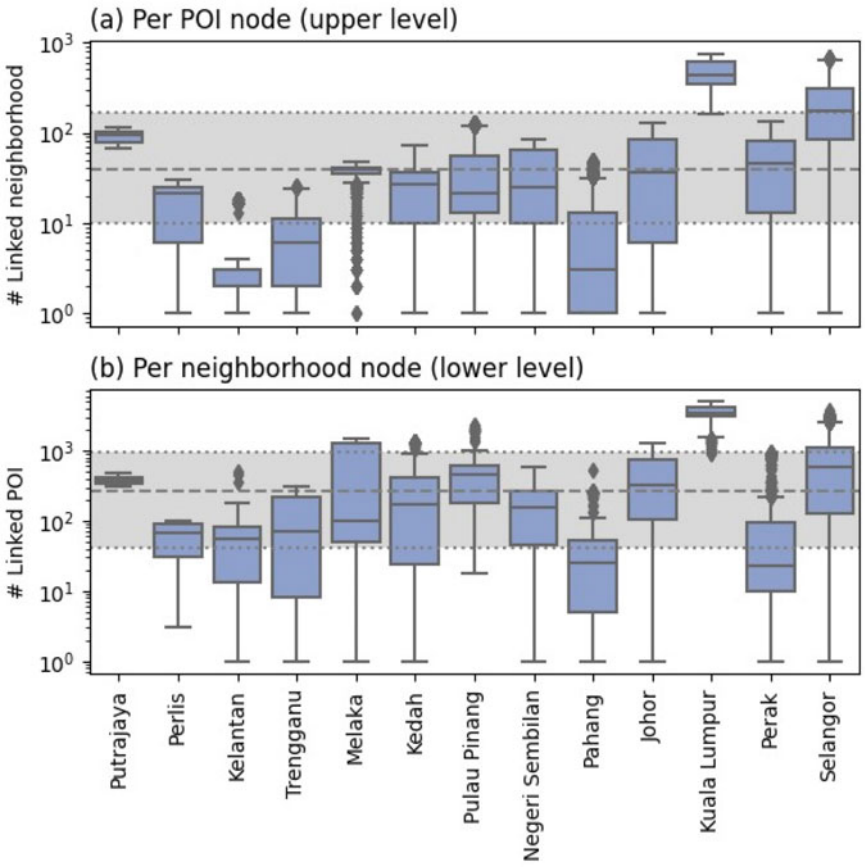


Fig. 8.6 The box plots of the degree by the two levels: **a** the upper level, showing the number of linked neighbourhoods for each POI node; **b** the lower level, showing the number of linked POI for each neighbourhood node. The states/FTs were sorted by the number of neighbourhoods in each state/FT (see Fig. 8.5). The dashed lines indicated the global median and the grey region indicates the global range of 25th to 75th percentiles

and the weighted degree of the top-11 zones, sorted by the number of neighbourhoods in descending order.

In Fig. 8.7a, Zone-A contains a notable large number of neighbourhoods (2437) compared to the other 10 zones (less than 500). Figure 8.7b shows the boxplot of weighted degrees in each zone. The weighted degree distributions of the nodes in Zone-A were mostly higher than 10^5—higher than the other zones. The distribution of the weighted degree for Zone-B and Zone-C (both in Perak) dropped to the range between 10^3 and 10^4, which were lower than Zone-D (somewhere in between 10^4 and 10^5, which was in Pulau Pinang), indicating that the former two zones had a less dense distribution than the latter. The neighbourhoods in Zone-H and Zone-J (both

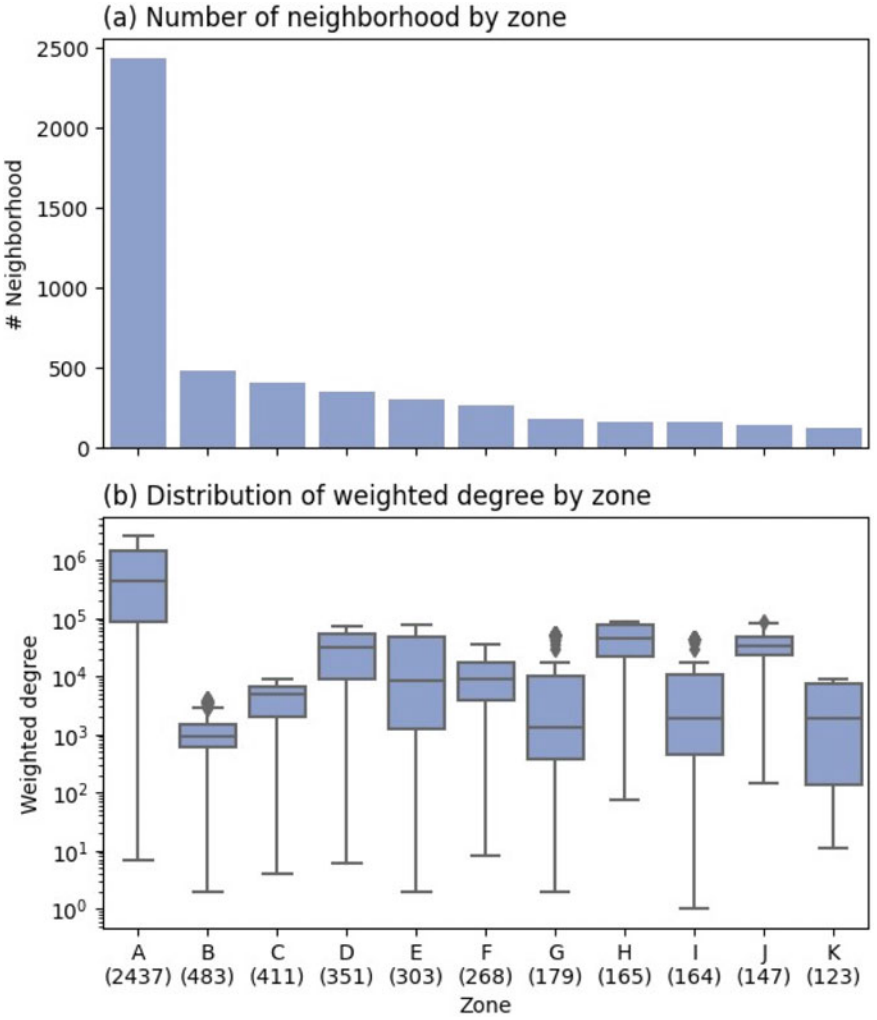


Fig. 8.7 **a** The number of neighbourhoods in each of the top-11 largest zones and **b** the corresponding distribution of weighted degree. Note that the vertical axis of **b** is in log scale

in Johor) had a similar range of weighted degrees to Zone-D, which were higher than the other places, indicating a dense spatial interaction within the zones.

The spatial distributions of the top-11 zones were shown in Fig. 8.8. The 11 zones were shown as red colour convex hulls, and the neighbourhoods (nodes) were coloured based on the categories of their weighted degree values (high, mid-high, mid-low, and low) according to the local (within-zone) distribution. The zone centres (golden 'X' symbols) were identified as the centroids of the neighbourhoods with high weighted degree values (nodes with weighted degrees greater than Q3 of each zone). Except for Zone-K (Kuantan, the capital city of Pahang) which was found

on the east coast, the other ten zones were found on the west coast of Peninsular Malaysia.

Zone-A covered most of the urbanised area in Selangor, including the two FTs. The zone centre was found in Kuala Lumpur. For Zone-B, because there were two spatially separated groups of high weighted degree neighbourhoods (two groups of red nodes), two centroids (zone centres) were identified. Zone-B was found as a narrow-shaped

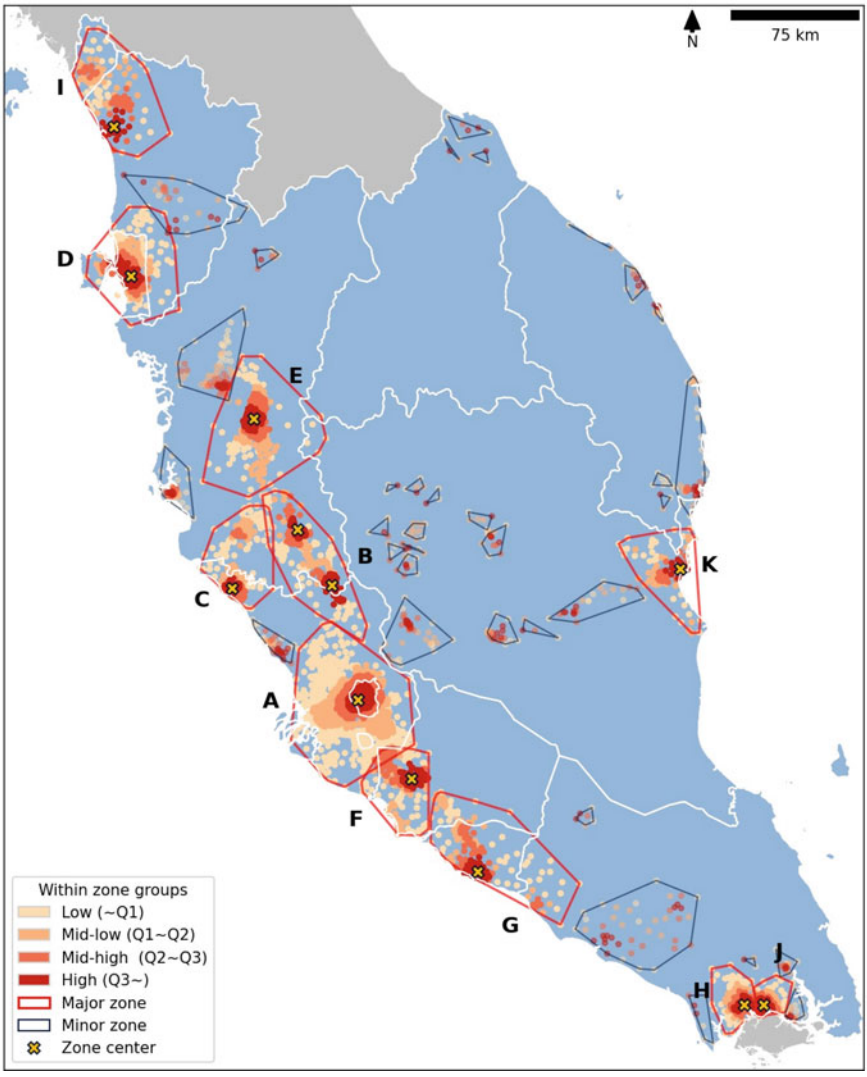


Fig. 8.8 The spatial distribution of the top-11 zones (red polygons). The colours of the neighbourhood (node) show the within zone categorisation by weighted degree. The golden 'X' markers indicate the centre point of each zone. The white lines highlight the states/FTs boundaries

zone that connects Zone-A and Zone-E—an economic corridor development—a lot of amenities located along the economic corridor area that merged the two subzones. Zone-C was located beside Zone-B—the coastal side linkage between the states of Selangor and Perak. While the upper part of Zone-C was geographically located in Perak, the neighbourhoods were more connected to the zone centre in Selangor, which was located at the lower and coastal side of Zone-C. Zone-E was mainly composed of an inland town—Ipoh, the capital city of Perak. Zone-D covered the Pulau Pinang state, including the island side and the peninsula side, and extended to the surrounding area that belongs to Kedah. Zone-I was another zone found in the north of Peninsular Malaysia. This zone covered the whole Perlis area—the most northern state—and Alar Setar, the capital city of Kedah.

As a close neighbour of Zone-A at its southern border, Zone-F was composed of the Seremban city—the capital city of Negeri Sembilan. Although the neighbourhoods were spatially near to each other at the border of the two states, the amenity-sharing relationship successfully captured the administrative border between the two states. Similar to Zone-D, Zone-G has covered the Melaka state and extended to the corridor region connecting to Seremban, Melaka and Muar (a northern district of the state of Johor). In other words, Muar was closer to Melaka state than to the capital city of Johor, which was at the southern tip of Peninsular Malaysia. Johor Bahru—the capital city of Johor state—formed two zones: Zone-H and Zone-J. Zone-H was found at the west of the city while Zone-J covered the east part. Here, the actual spatial distribution of zones is presented, which were identified based on the 10 km moving distance threshold. As discussed in the previous section, these zones can be considered as the spatial boundaries of disease control zones—inseparable neighbourhoods since the residents co-visit a lot of amenities together, which implies complex and dense interactions between neighbourhoods in the same zone that lead to a high possibility of disease transmission.

8.3.4 Part IV: Explore the Zone Centres of the Neighbourhoods Interactions

One or two zone centres were identified for each of the top-11 zones using the high weighted degree neighbourhoods. In Fig. 8.9, the distances from each neighbourhood to the zone centre were calculated and compared with the weighted degree values. The weighted degree indicates the total frequency of a neighbourhood sharing amenities with other neighbourhoods. Overall, the weighted degree of all zones showed a two-step decreasing trend with the distance to the zone centres. The weighted degree dropped with a steep slope as the distances to zone centres increased and reached a low weighted degree. After a turning point, which was approximately 15 km, the decreasing rate of weighted degree became flat, indicating the neighbourhoods that locate beyond this threshold distance had a similar or constant weighted degree value. In other words, the gradients of high neighbourhood interactions to

low neighbourhood interactions were strongly related to the distances to the zone centres for the neighbourhoods located within a 15 km radius of the zone centres. The neighbourhoods beyond this threshold distance usually had low values of neighbourhood interactions, i.e., had a simpler amenity-sharing structure. In addition, this also indicates that the area size of most cities in Malaysia is around a 15 km radius. Neighbourhoods that fall beyond this threshold radius were sparsely distributed, hence the low weighted degree.

In Fig. 8.9, the four groups (high, mid-high, mid-low, and low) were shown using the boxes with the same colour scheme as in Fig. 8.8. The first two groups (high and mid-high) usually had a narrow rectangle shape that was long in the vertical direction. The other two groups (mid-low and low) had a wide rectangle shape that was long in the horizontal direction. The neighbourhoods in the high and mid-high

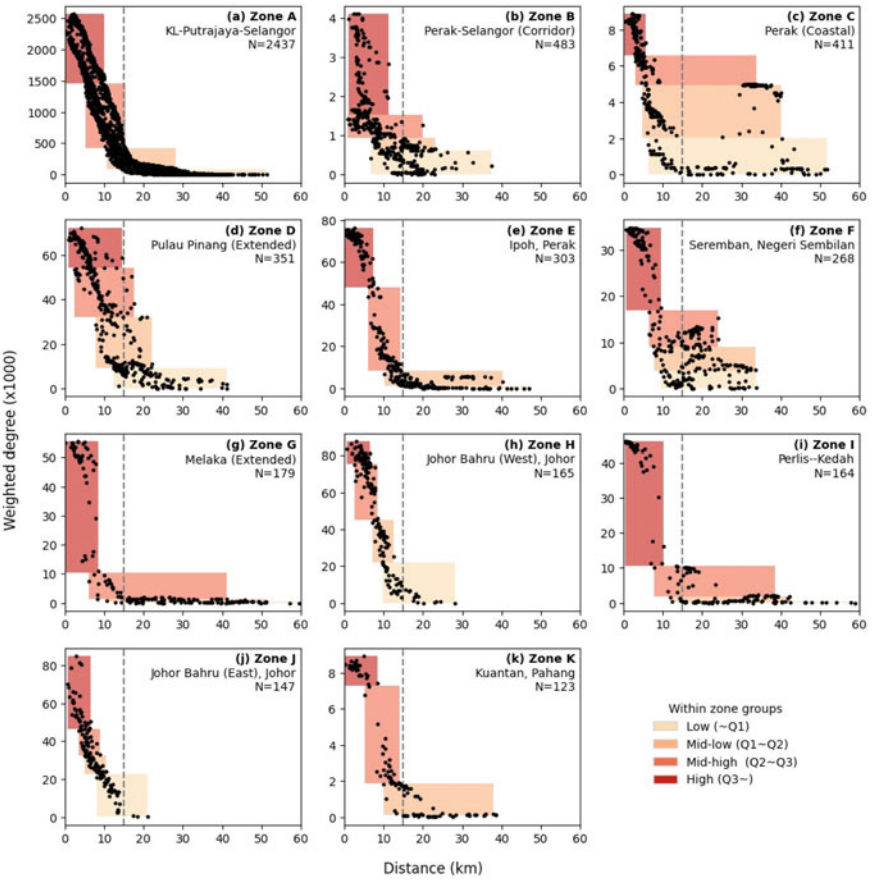


Fig. 8.9 The scatter plot shows the weighted degrees (vertical axis) by the distances to centre points (horizontal axis) for each zone. Boxes in different colours indicate the range of the four groups. The vertical dashed line (15 km) indicates the approximate turning point of the decreasing pattern

groups were usually within a 15 km radius from the zone centres. Zone-C (the coastal part of Perak and Selangor) contained a secondary core region that had only a few mid-high neighbourhoods but none of the neighbourhoods was in the high group, i.e., the north/inner part of the zone. These neighbourhoods were far from the coastal core region—located about 30 km to 40 km from the zone centre—while some of them had moderate values of weighted degree. For Zone-D and Zone-F, some of the neighbourhoods in the mid-high group exceed the 15 km threshold with a moderate high weighted degree, indicating that the urbanised region spread out from the zone centre to a larger radius. For Zone-G and Zone-I, the neighbourhoods in the mid-high group were a short (vertical direction) and wide (horizontal direction) rectangle shape, which indicated that the mid-high group had entered the slow decreasing stage and the weighted degree was constantly low for 75% of the neighbourhoods. In other words, there were fever neighbourhoods (about 25% of the neighbourhoods) located near the core region (with higher values of weighted degree) and only the high group neighbourhoods experienced a quick decreasing slope.

Figure 8.10 focused on the five zones: (a) Zone-A, a large area that was composed of Kuala Lumpur, Putrajaya, and Selangor; (b) Zone-D, a coastal and island state (Pulau Pinang) that included a connection between the peninsular region and the island of Penang; (c) Zone-E, an inland city (Ipoh, Perak); and (d) the two zones (Zone-H and Zone-J) that were closely linked to each other at Johor Bahru. In Fig. 8.10a, the zone centre was in Kuala Lumpur and the weighted degree reduced with the distance to the centre—a typical monocentric city pattern. The second FT (Putrajaya) was mainly covered with a mid-low weighted degree in the amenity-sharing network, indicating a less crucial role in the spatial interaction structure. Zone-D (Fig. 8.10b) formed an area that was larger than the Pulau Pinang State (white border area), that included the Sungai Petani at the north and Kulim at the east; both Sungai Petani and Kulim were part of the neighbouring state (Kedah) but the two towns were more connected to Pulau Pinang than to the capital city of Kedah (Alor Setar). The zone centre was shifted slightly to the south of the central business districts (CBD) of the state. Both George Town and Butterworth (primary and secondary CBD) is located slightly to the north of the zone centre and George Town is on the island. The primary and secondary CBD was developed as a historical product—near the location of harbours and the bridge linking the island and the peninsula. The identified zone centre indicated the hinterland location where more neighbourhoods (especially residential areas) were found, and more spatial interaction happened.

Ipoh (Fig. 8.10c) is the capital city of Perak, which is the largest inland zone in terms of area size among the zones on the west coast of Peninsular Malaysia (the other inland zone is Zone-B). The north-to-south narrow shape of the spatial distribution is caused by the natural landscape: the city is located in the Kinta Valley and surrounded by two mountain ranges—Titiwangsa Range and Kledang Range—on the east and west sides of the city, respectively. The zone centre is located almost at the centre of the city, and the weighted degree decreased from the zone centre along with the north and south directions. In Fig. 8.10d, the two zones were identified in the southern part of the Peninsular Malaysia—Johor Bahru. Neighbouring

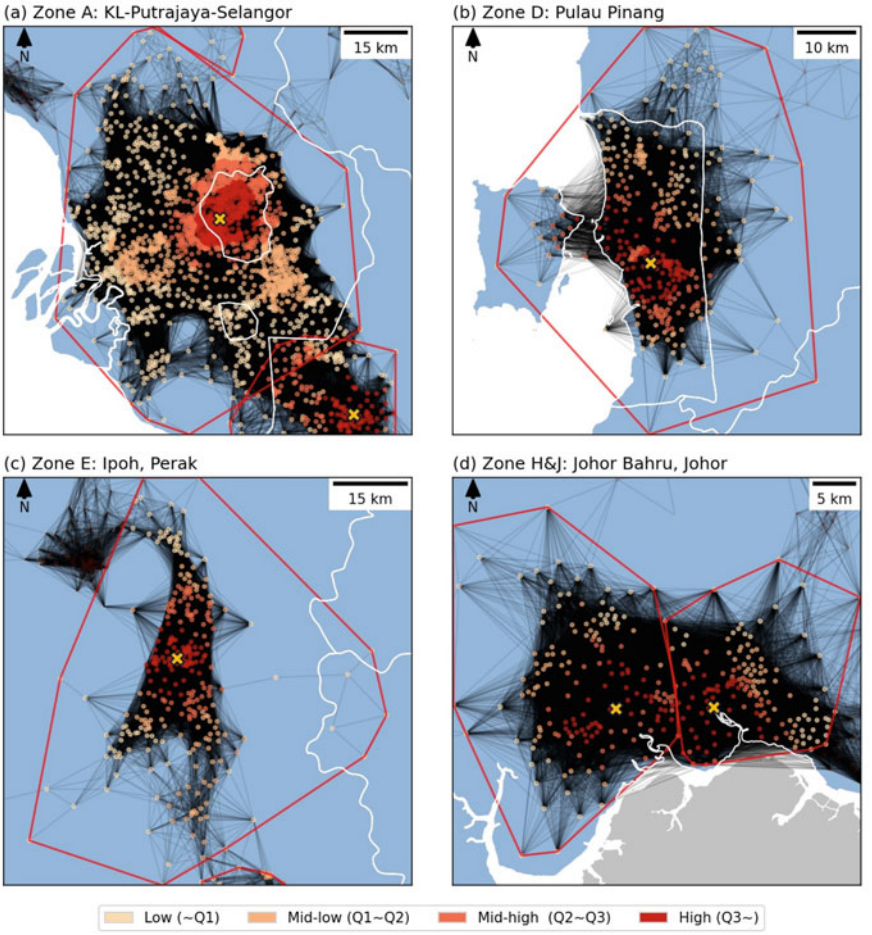


Fig. 8.10 The amenity-sharing networks of four regions: **a** Zone-A, which includes the two FT and Selangor, **b** Zone-D, which covers the entire Pulau Pinang and also extends to the neighbouring state of Kedah, **c** Zone-E: Ipoh, the capital of Perak, and **d** the Zone-H (on the left) and Zone-J (on the right) that covers the Johor Bahru, the capital city of Johor. The colours of neighbourhood nodes show the within zone groupings by the weighted degree. The golden 'X' markers indicate the zone centres. The white lines highlight the states/FTs boundaries

Singapore, Johor Bahru contained the most financial amenities and residential area in Johor. While the two zones belong to the same city, the separation of zones from the community detection algorithm based on the amenity-sharing network indicates that the distinction of communities occurred within the city. In other words, although there were a lot of cross-zone connections, the internal connections within each of the two zones were stronger and denser than the connection between the two zones. This could be observed from the nodes located at the border of the two zones—these neighbourhoods had a slightly lower weighted degree (mid-high group). Similar to

the other three zones, the zone centres of Zone-H and Zone-J were located in the hinterland of the two zones. The decreasing of weighted degree occurred with the distance from the zone centre to the outer area, i.e., toward the west and north for Zone-H and toward the east and north for Zone-J. In summary, based on the four examples shown in Fig. 8.10, the zone centres were usually found in the hinterland or CBD of the city, which also implies the location with a high frequency of spatial interaction within the zone.

8.4 Discussions

This study demonstrated a data-driven framework for delineating zones of disease diffusion. The first part identified neighbourhoods that represent where people live from the micro-scale spatial units—road segments. The relationships of sharing amenities were identified in the second part, which was used to identify the zones of disease diffusion (part 3) and calculate the weighted degrees for identifying city centres (part 4). These analyses captured where people interact with each other heavily and that can lead to disease diffusion. The underlying structure of places and amenities can be analysed by a complex network analysis approach [5, 7, 39]. The identified spatial boundaries of human interaction and city centres can be useful information for the health authorities in the policymaking process, e.g., where the boundary of the lockdown zones should be set or where to allocate the medical resources [20]. While some zones' boundaries resemble the boundary of states, some zones extend to the neighbouring states, and some states contain multiple zones. In other words, the spatial structure of human activity in the cities are not always following the administrative boundaries, which are drawn for political and administrative purposes. Therefore, making policies for disease controls according to administrative boundaries is not ideal. Here we presented a data-driven approach to the identification of the human movement boundaries that can be used as a reference.

The movement of people from one place to another provides an opportunity for disease pathogens to travel and spread spatially [5, 6, 21]. In this study, the concept of disease diffusion and human movement was captured by analysing the locations of neighbourhoods and amenities within a 10 km radius—a situation that corresponds with the MCO (lockdown measure) and regular situation. The co-visit relationship—i.e., the amenity-sharing network—indicates the possible spatial interactions between neighbourhoods. The identified zones using the amenity-sharing network and a community detection method (MapEquation) imply that the places are densely interconnected if they are found in the same zone, i.e., a lot of people visit a similar set of amenities from these neighbourhoods. The distribution of zones uncovers the spatial structure of urban form, i.e., the shape and size of a city/town/settlement. In the context of disease outbreaks, the amenity-sharing relationship indicates the possible spread of disease from one place to another, and the identified zones indicate the possible impact area of a disease cluster. When a new disease cluster is reported, people from other neighbourhoods in the zone should be aware of the situation and

take corresponding actions, e.g., reduce outdoor gatherings and enhance personal hygiene habits to avoid catching the disease.

The spatial distribution of weighted degrees indicates the strength of amenity-sharing relationships. A high weighted degree means more co-visit connections that imply both high diversity (co-visit with various neighbourhoods) and high density (a large frequency of co-visit interactions). Therefore, in the disease diffusion context, the high weighted degree places have a high probability to be spatial super-spreaders or spatial super-susceptibles—the disease can easily spread from or to these places [5, 15, 17]. The results indicate that the locations of high weighted degrees are mainly found in the CBD or hinterland of the areas/states. This is because the CBD usually has more amenities and a more complex road network structure. In other words, when a disease outbreak happens, the CBD should get more attention, e.g., to reduce human activities, or to temporally close some non-essential shops within the area.

The weighted degree categorisation and zone centre identification was done for individual zones (i.e., regional) in this study due to the strong heterogeneity between zones and between states. This notable difference indicates a large urban–rural hierarchy between regions or states occurs within the country. Although the differences between zones/states are large, the distribution of within-zone weighted degree to the distances from the zone centres shows a similar pattern—a rapid decay of weighted degree by distance to around 15 km radius. This suggests that the size of a city/town in Malaysia usually has a 15 km radius size (or less), regardless of the highest value of the weighted degree that was derived from the number of amenities. This also suggests that the places within a 15 km radius from the zone centre are densely interconnected, with a rapid change of weighted degree (strength of spatial interaction) depending on the distance to the zone centre. Therefore, these areas should also be aware of the outbreak situations. In addition, this does not mean that places beyond 15 km are not important in the disease control measures; it only indicates that those places have less complex spatial interactions, hence the risk of disease spreading is relatively low in comparison to the core area. However, these rural and distant places usually have lower accessibility to medical resources. Thus, these places should not be neglected in the policymaking process—some other types of policies or disease control measures may be helpful for these locations, e.g., more medical resources allocation, and more controls on the incoming/outgoing movement restrictions.

This study provided an example of using open data and open-source software to develop a framework to support policymaking on disease control events. The presented framework uses two basic and common types of datasets—a road network and a set of POIs. Both are available without any cost from the OpenStreetMap platform, which covers the global area [12]. In other words, this framework can be easily applied to other cities or countries. For some countries, including Malaysia, a lot of spatial data was not available as open data from the official authorities. The official spatial data is not available because either it does not exist, it is confidential, or it is only available for purchasing. Therefore, it is difficult to run a micro-scale analysis for various purposes, including the situation during a disease outbreak. Open spatial data and open-source software are available for anyone and thus it may be an ideal solution for most countries where the official data is unavailable [2, 48]. This

study provides a way to identify the zones of disease diffusion that can be reproduced with open-source software (the MapEquation tool and some Python packages) and open data.

8.5 Conclusion

The COVID-19 situation is still ongoing in Malaysia [41]. As of this writing, several variants and sub-variants of COVID-19, including Delta, Omicron, Omicron subvariant-BA-5, and Omicron subvariant-BA.2.12.1, have hit Malaysia and the world in the last two years [43, 44]. Understanding the spatial human activity structure can be beneficial for disease spatial modelling and policymaking. Future spatial epidemiology studies on COVID-19, including spatial modelling [8, 22], simulation [16, 36], resource planning and allocation [20], regional lockdown measures assessment [11, 23, 45], risk and vulnerability assessment [18, 24, 46], identification of possible spatial super-spreader/super-susceptible [5, 17], and even sentiment or mental health impact from COVID-19 [31, 47] or building scale human movement and disease spreading analysis [38] etc., can be done on top of the results from this study, e.g., focusing on individual zones or analysing the city centre points. For practical purposes, the result from this study provides the boundaries of zones that can be used as disease control zoning, e.g., for reducing the cross-zones movement to avoid relocation diffusion, to allocate and set up the testing centres, vaccination centres or emergency response teams within each zone, or design different disease control strategies for different zones. In conclusion, the proposed analysis framework can delineate the zones of disease diffusion, and which can be used to support the fight against contagious diseases.

There are several limitations and future research suggestions. First, no validation was done in the analysis because the micro-scale (e.g., neighbourhood or POI level) disease data was not available. In addition, there is no publicly available data for other types of contagious diseases that could be used as surrogate data. While this study aims to identify zones and the neighbourhood-level degree distribution, large scale (state level) data is not suitable for validation. Second, the identification of city centres was slightly arbitrary and not programmatic. The current method is to visually identify the clusters of high degree (greater than Q3) neighbourhoods and then calculate a centroid for each cluster. In the case study, only one zone contains two clusters (Zone-B) among the top-11 zones. The current method can be problematic for cities with multi-core structures—if two cores were close to each other, the inter-core buffer zones area is small and the weighted degrees of the neighbourhoods in the buffer area are high, then only one centroid will be identified, and it would locate at the buffer area instead of the two city centres. To cope with this issue, a high-resolution kernel density estimation may be useful in presenting the distribution and providing some hints on the numbers and locations of zone centres.

Third, while this study focused on human activity and interactions, the situations at periphery locations (e.g., rural, agricultural, and indigenous settlements)

were ignored. In the first step, the neighbourhoods and settlements in the middle area and east coast of Peninsular Malaysia were identified from the road segment data. However, in the amenity-sharing analyses (part-2 to part-4), these areas were fading-out due to the lack of amenity-sharing relationships (links); also, these neighbourhoods/settlements did not connect to the main network component. In other words, the framework can only focus on and analyse the disease diffusion in the urbanised area. Fourth, the framework captures mainly the strong and dense interactions—the 'bond links' in the amenity-sharing network. The long-range connections, e.g., the places that can be reached by a highway within a short amount of time, were ignored. These long-distance 'bridge' links (weak links) extend the movement radius to a far location and provide a possible route for disease spreading (i.e., relocation). Similarly, the domestic airlines and rail lines were not included in the analysis. In future analysis, these long-range connections can be included using a time-based accessibility measurement and threshold, e.g., whether the two neighbourhoods are reachable within 1 hour, to generate the amenity-sharing network.

References

1. Bernama (2020) MCO: travel now restricted to 10-km radius. Bernama. Accessed 2022 Jan 03
2. Boeing G, Higgs C, Liu S, Giles-Corti B, Sallis JF, Cerin E, Lowe M, Adlakha D, Hinckson E, Moudon AV, Salvo D, Adams MA, Barrozo LV, Bozovic T, Delclòs-Alió X, Dygrýn J, Ferguson S, Gebel K, Ho TP, Lai P-C, Martori JC, Nitvimol K, Queralt A, Roberts JD, Sambo GH, Schipperijn J, Vale D, Van de Weghe N, Vich G, Arundel J (2022) Using open data and opensource software to develop spatial indicators of urban design and transport features for achieving healthy and sustainable cities. Lancet Glob Health 10(6):e907–e918
3. Bunyan J (2020) PM: Malaysia under movement control order from Wed until March 31, all shops closed except for essential services. Malay Mail. Accessed 2022 Jan 03
4. Chin WCB (2021) Daily life pattern of a city: delineating activity space and time using social media data. SSRN, p 3961269
5. Chin WCB, Bouffanais R (2020) Spatial super-spreaders and super-susceptibles in human movement networks. Sci Rep 10:18642
6. Chin WCB, Huang C-Y (2020) Comments on "EpiRank: modeling bidirectional disease spread in asymmetric commuting networks" for analyzing emerging coronavirus epidemic patterns. MedRxiv
7. Chin WCB, Wen T-H (2015) Geographically modified PageRank algorithms: identifying the spatial concentration of human movement in a geospatial network. PLoS One 10(10):e0139509
8. Chin WCB, Wen T-H, Sabel CE, Wang I-H et al (2017) A geo-computational algorithm for exploring the structure of diffusion progression in time and space. Sci Rep 7(1):1–13
9. Cliff AD, Haggett P, Smallman-Raynor M, Smallman-Raynor MR (2000) Island epidemics. Oxford University Press on Demand
10. Department of Statistics, Malaysia (2011) Population distribution and basic demographic characteristics 2010. http://www.statistics.gov.my/portal/download_Population/files/census2010/Taburan_Penduduk_dan_Ciri-ciri_Asas_Demografi.pdf. Accessed 2021 Nov 16
11. Di Domenico L, Pullano G, Sabbatini CE, Boëlle P-Y, Colizza V (2020) Impact of lockdown on COVID-19 epidemic in Île-de-France and possible exit strategies. BMC Med 18(1):1–13
12. Geofabrik (2021) Download OpenStreetMap data for this region: Malaysia, Singapore, and Brunei. https://download.geofabrik.de/asia/malaysia-singapore-brunei.html. Accessed 2021 Jul 26

13. Golledge RG (1997) Spatial behavior: a geographic perspective. Guilford Press
14. Haggett P (1966) Locational analysis in human geography. St. Martin's
15. Huang C-Y, Chin WCB, Wen T-H, Fu Y-H, Tsai Y-S (2019) Epirank: modeling bidirectional disease spread in asymmetric commuting networks. Sci Rep 9(1):1–15
16. Huang C-Y, Wen T-H, Tsai Y-S (2014) FLUed: a novel four-layer model for simulating epidemic dynamics and assessing intervention policies. J Appl Math 2013:325816
17. Huang J, Kwan M-P, Kan Z (2021) The superspreading places of COVID-19 and the associated built-environment and socio-demographic features: a study using a spatial network framework and individual-level activity data. Health Place 72:102694
18. Kan Z, Kwan M-P, Huang J, Wong MS, Liu D (2021) Comparing the space-time patterns of high-risk areas in different waves of COVID-19 in Hong Kong. Trans GIS 25(6):2982–3001
19. Klapka P, Halas M (2016) Conceptualising patterns of spatial flows: five decades of advances in the definition and use of functional regions. Morav Geogr Rep 24(2):2–11
20. Kuo F-Y, Wen T-H (2021) A mathematical model for evaluating the medical resource availability of COVID-19 in time and space. In: Shaw S-L, Sui D (eds) Mapping COVID-19 in space and time: understanding the spatial and temporal dynamics of a global pandemic, Chapter 15. Springer International Publishing, Cham, pp 295–308
21. Kuo F-Y, Wen T-H (2021) Regionalization for infection control: an algorithm for delineating containment zones considering the regularity of human mobility. Appl Geogr 126:102375
22. Kuo F-Y, Wen T-H, Sabel CE (2018) Characterizing diffusion dynamics of disease clustering: a modified space-time DBSCAN (MST-DBSCAN) algorithm. Ann Am Assoc Geogr 108(4):1168–1186
23. Lau H, Khosrawipour V, Kocbach P, Mikolajczyk A, Schubert J, Bania J, Khosrawipour T (2020) The positive impact of lockdown in Wuhan on containing the COVID-19 outbreak in China. J Travel Med 27(3):1–7
24. Lee J, Lay J-G, Chin WCB, Chi Y-L, Hsueh Y-H (2014) An experiment to model spatial diffusion process with nearest neighbor analysis and regression estimation. Int J Appl Geosp Res 5(1):1–15
25. Lee JH, Davis AW, Yoon SY, Goulias KG (2016) Activity space estimation with longitudinal observations of social media data. Transportation 43(6):955–977
26. Leong C-H, Chin WCB, Feng C-C, Wang Y-C (2021) A socio-ecological perspective on COVID19 spatiotemporal integrated vulnerability in Singapore. In: Shaw S-L, Sui D (eds) Mapping COVID-19 in space and time: understanding the spatial and temporal dynamics of a global pandemic, Chapter 6. Springer International Publishing, Cham, pp 81–111
27. Mat NFC, Edinur HA, Razab MKAA, Safuan S (2020) A single mass gathering resulted in massive transmission of COVID-19 infections in Malaysia with further international spread. J Travel Med 27(3):1–4
28. Meade MS, Emch M (2010) Medical geography. Guilford Press
29. Moovit (2020) Kuala Lumpur public transit statistics. https://moovitapp.com/insights/en/Moovit_Insights_Public_Transit_Index_Malaysia_Kuala_Lumpur-1082. Accessed 2022 Jan 03
30. Numbeo (2020) Traffic in Kuala Lumpur, Malaysia. https://www.numbeo.com/traffic/in/Kuala-Lumpur. Accessed 2022 Jan 03
31. Pfefferbaum B, North CS (2020) Mental health and the Covid-19 pandemic. N Engl J Med 383(6):510–512
32. Philbrick AK (1957) Principles of areal functional organization in regional human geography. Econ Geogr 33(4):299–336
33. Pung R, Chiew CJ, Young BE, Chin S, Chen MI, Clapham HE, Cook AR, Maurer-Stroh S, Toh MP, Poh C et al (2020) Investigation of three clusters of COVID-19 in Singapore: implications for surveillance and response measures. The Lancet 395(10229):1039–1046
34. Rosvall M, Axelsson D, Bergstrom CT (2009) The map equation. Eur Phys J Spec Top 178(1):13–23
35. Sabel CE, Pringle D, Schærström A (2010) Infectious disease diffusion. In: Brown T, McLafferty S, Moon G (eds) Companion to health and medical geography, Chapter 7. WileyBlackwell, Wiley-Blackwell Malden, MA, pp 111–132

36. Salim N, Chan WH, Mansor S, Bazin NEN, Amaran S, Faudzi AAM, Zainal A, Huspi SH, Khoo EJH, Shithil SM (2020) COVID-19 epidemic in Malaysia: impact of lock-down on infection dynamics. MedRxiv
37. Schönfelder S, Axhausen KW (2003) Activity spaces: measures of social exclusion? Transp Policy 10(4):273–286
38. Srikanth ADS, Chin WCB, Bouffanais R, Schröpfer T (2022) Complexity science-based spatial performance analyses of UNStudio/DP Architects' SUTD Campus and WOHA's Kampung Admiralty. In: As I, Basu P, Talwar P (eds) Artificial intelligence in urban planning and design: technologies, implementation, and impacts, Chapter 12. Elsevier, pp 217–244
39. Srikanth ADS, Chin WCB, Bouffanais R, Schröpfer T (2022) Complexity science for urban solutions. In: As I, Basu P, Talwar P (eds) Artificial intelligence in urban planning and design: technologies, implementation, and impacts, Chapter 3. Elsevier, pp 39–58
40. Tang KHD (2020) Movement control as an effective measure against Covid-19 spread in Malaysia: an overview. J Public Health 30:583–586
41. Tang KHD (2021) From movement control to National Recovery Plan: Malaysia's strategy to live with COVID-19. Int J Sci Healthc Res 6(4):286–292
42. Tang KHD, Chin BLF (2021) Correlations between control of COVID-19 transmission and influenza occurrences in Malaysia. Public Health 198:96–101
43. The Star (2022) Covid-19: high possibility of Omicron wave hitting Malaysia, says Khairy. The Star. Accessed: 2022 Jan 04
44. The Star (2022) Covid-19: two Omicron sub-variants detected in Malaysia, says Khairy. The Star. Accessed 2022 Jun 10
45. Wen T-H, Chin WCB (2015) Incorporation of spatial interactions in location networks to identify critical geo-referenced routes for assessing disease control measures on a large-scale campus. Int J Environ Res Public Health 12(4):4170–4184
46. Wen T-H, Tsai C-T, Chin WCB (2016) Evaluating the role of disease importation in the spatiotemporal transmission of indigenous dengue outbreak. Appl Geogr
47. Yan Y, Chin WCB, Leong C-H, Wang Y-C, Feng C-C (2021) Emotional responses through COVID-19 in Singapore. In: Shaw S-L, Sui D (eds) Mapping COVID-19 in space and time: understanding the spatial and temporal dynamics of a global pandemic, Chapter 5. Springer International Publishing, Cham, pp 61–79
48. Yao XA, Huang H, Jiang B, Krisp JM (2019) Representation and analytical models for location-based big data. Int J Geogr Inf Sci 33(4):707–713

Dr. Wei Chien Benny Chin works as a research fellow in the Department of Geography, National University of Singapore. He is a geographical information scientist. His research interests include computational geography, complexity science, and scaling. He works on complex human movement networks and spatial epidemiology projects.

Chapter 9
Approaches for Spatial and Temporal-Spatial Clustering Analysis in Avian Influenza Outbreaks

Mei-Liang Huang, Hong-Dar Isaac Wu, and Day-Yu Chao

Abstract Avian influenza virus (AIV) belongs to the genus Influenza A virus of the family *Orthomyxoviridae*. The virus can infect a variety of avian species, but the low pathogenic AIVs do not usually cause explicit symptoms in poultry. In contrast, the highly pathogenic avian influenza (HPAI) viruses continue to cause outbreaks among poultry, wild birds and occasionally humans in Asia, the Middle East, North America, and Africa. Environmental factors associated with cross-species transmission have been substantially reviewed before. However, acquiring the knowledge of a number of environmental factors with spatial structures, which usually are not randomly distributed, for timely implementation of control measures rely on accurate identification of the spatial clustering in a global or local scale. In this article, we review different approaches in identifying spatial or temporal-spatial clustering in avian influenza outbreaks. In the future perspective, we propose to develop intuitive tools for timely identify the dynamic changes of clustering and viral spreading. Such tools will assist in not just the identifying the environmental factors associated with the clustering or spreading direction, but also timely control measures to prevent further damage.

Keywords Avian influenza virus (AIV) · Cluster analysis · Hotspot analysis · Scan statistics · Space–time permutation model · Knox test · Standard deviational ellipse (SDE) method · Regression modeling

M.-L. Huang · D.-Y. Chao (✉)
Graduate Institute of Microbiology and Public Health, College of Veterinary Medicine, National Chung Hsing University, 145 Xingda Rd., Taichung 402, Taiwan
e-mail: dychao@nchu.edu.tw

H.-D. I. Wu
Department of Applied Mathematics and Institute of Statistics, Taichung 402, Taiwan

T.-H. Wen et al. (eds.), *Earth Data Analytics for Planetary Health*, Atmosphere, Earth, Ocean & Space, https://doi.org/10.1007/978-981-19-8765-6_9

9.1 Introduction

Avian influenza, caused by influenza A virus (IAV), is a zoonotic influenza that affects a wide variety of birds, poultry and occasionally humans. Influenza A virus is the only species of the genus Alpha-influenza virus of the family *Orthomyxoviridae*. The structure of influenza A virus consists of a lipid envelope and a negative-sense single-stranded ribonucleic acid (RNA) genome with eight segments [1]. Influenza A viruses can be classified into subtypes based on the combination of the spike hemagglutinin (HA) attachment protein and the neuraminidase (NA) protein. To date, 18 HA subtypes (H1 to H18) and 11 NA subtypes (N1 to N11) have been identified [2], while only 131 subtypes have been detected in nature [3]. Subtypes of IAV can be further divided into clades and subclades based on the similarity of HA genes [4, 5], and subtypes can also be subdivided into genotypes based on the combination of internal gene segments. The nomenclature of influenza viruses has been standardized, and the name of a new strain consists of a combination of antigen type, original host, geographic origin, strain name, year of isolation, and subtype (HxNy) [6].

The genome segments of IAV encode different viral proteins. The structural proteins express in the envelope containing the surface proteins, which are HA attachment proteins and NA proteins, and the membrane ion channel (M2) proteins. Internal proteins include the nuclear protein (NP), matrix protein (M1), and the polymerase complex consisting of three subunits, namely polymerase basic protein 1 (PB1), polymerase basic protein 2 (PB2), and polymerase acidic protein (PA). Nonstructural protein 1 (NS1) and nonstructural protein 2 (NS2), the nuclear export protein (NEP), are encoded by segment 8. AIVs use host proteases to cleave the HA0 molecule into HA1 and HA2 subunits, which are essential for the uncoating step of viral replication. AIVs can be defined as low pathogenic avian influenza (LPAI) viruses and highly pathogenic avian influenza (HPAI) viruses based on their virulence in chickens. Thus, if mutations result in the insertion of multiple lysine and arginine residues into the HA0 cleavage site of the virus, termed the multilocus cleavage site, which can be recognized by the ubiquitous and extensive proteases in host tissues, it becomes an HPAI virus. As a corollary, HPAI viruses may replicate throughout the host, systematically destroying tissues, leading to multiple organ failure and ultimately to host death. However, LPAI viruses have only one arginine at the cleavage site, which can only be recognized by trypsin-like proteases. Therefore, replication of LPAI virus is restricted to the respiratory and gastrointestinal tracts, where expression of this protease occurs [7–10].

Although avian influenza viruses (AIVs) replicate in wild bird reservoirs, the viruses spread out through the saliva, mucus, and feces of infected birds. Spillover AIVs can be transmitted from infected wild birds to poultry, primarily through direct contact with wild birds or indirect contact through human activities and contaminated water or other media. Most AIVs cause gastrointestinal infections in chickens, while those with no or minimal clinical signs are LPAI viruses, whose distribution in wild birds varies by subtype depending on geographic location, bird abundance,

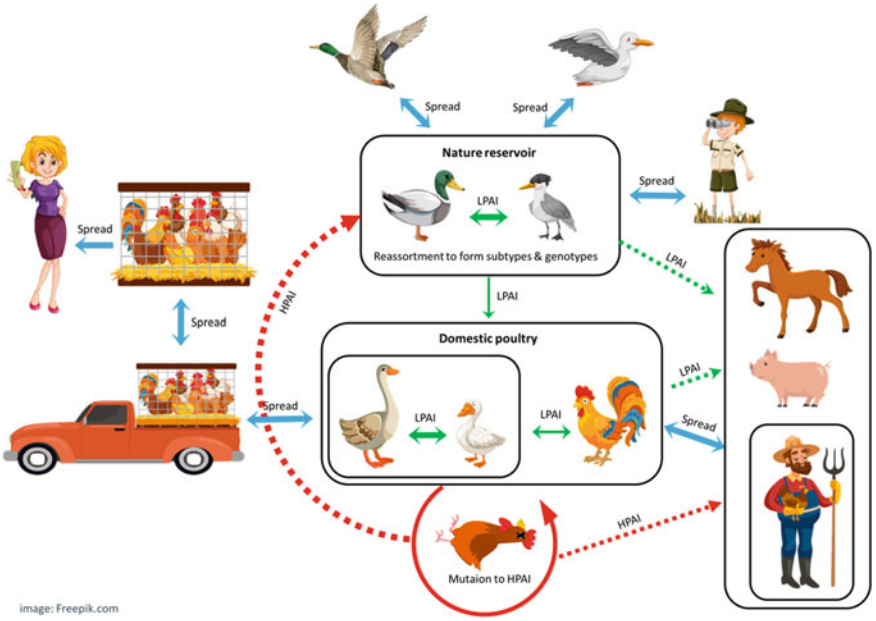


Fig. 9.1 The evolution and transmission routes of avian influenza virus between and among animal species and humans. Green arrows indicate the transmission routes of low pathogenic avian influenza (LPAI), red arrows indicate the transmission routes of high pathogenic avian influenza (HPAI) and blue arrows indicate the additional spread routes of avian influenza. Solid arrows represent frequent transmission events, and dashed arrows represent sporadic or limited transmission events. Once HPAI viruses become introduced into wild bird populations, the spread and maintenance of these viruses in wild birds will be determined by different factors involved the types of host birds, the viruses, and the ecology

and prevalence. HPAI viruses can affect poultry as well as wild birds. Infections in chickens and turkeys induce severe disease with mortality rates as high as 90–100%. So far, only the H5 and H7 subtypes of AIVs have been recorded as causing HPAI outbreaks in poultry, but most of the H5 and H7 subtypes are LPAI viruses. HPAI viruses evolve by mutation, amino acid substitution or recombination after long-term circulation and efficient replication of LPAI viruses in poultry [11] (Fig. 9.1).

9.2 Factors Associated with Zoonotic Transmission of Avian Influenza Virus

Outbreaks of HPAI were first described as "fowl plague" in the 1880s, and subsequent outbreaks in Europe from then onwards were caused exclusively by HPAI H7N7 and H7N1 viruses until the first confirmed outbreak of HPAI H5N1 occurred among

chickens in Scotland in 1959 [12]. HPAI H5Nx (N1-9) and H7Nx (N1, N3, N4, N7-N9) viruses [13, 14] have been reported to cause thousands of outbreaks in domestic poultry and wild birds in more than 60 countries, killing numerous poultry through HPAI virus attacks or mass culling strategies and causing huge economic losses. During HPAI epidemics in poultry, viruses can spill back into wild birds, where they subsequently circulate asymptomatically or cause disease and death [15, 16], even generate reassortants with LPAI or wild bird-adapted strains [17–19]. These viruses can also spill over into mammals, including pigs, horses, whales, seals, and humans [20]. However, it is considered that AIVs do not replicate efficiently enough in humans to sustain human-to-human transmission [21] (Fig. 9.1).

Long-distance migratory birds played an influential role in the global spread of HPAI viruses [22–26], while wild birds may also be involved in local HPAI virus amplification and reassortment [27]. The HPAI found in wild birds was highly associated with the geographical locations of poultry farms [20, 28, 29]. However, such association has been not significant since the emergence in 2014 of the HPAI virus Gs/Gd clade 2.3.4.4 which has dominated in outbreaks in poultry and wild birds with abundant genetic reassortments resulting in H5N1, H5N2, H5N3, H5N4, H5N5, H5N6 and H5N8 subtypes [24, 30, 31].

The most predominant natural reservoirs of HPAI H5 viruses are Anseriformes, which are responsible for the maintenance, rapid transmission, and geographic expansion of these viruses. The other prominent reservoirs, Charadriiformes, are possible reasons for the rapid global spread of HPAI viruses due to their fast-moving, long-distance migration and highly gregarious during migration period [32–34]. The transmission rates of HPAI H5 viruses within Anseriformes and Galliformes are high, but transmission between these orders is limited [18, 35]. Understanding the mechanisms of HPAI virus transmission and maintenance in wild birds can provide a reference for surveillance strategies.

Human infection with AIVs is a rare and sporadic event, however, AIVs subtypes H5, H6, H7, H9, and H10, have been recorded infecting humans to cause clinical disease of varying severity. Exposure histories of human cases and phylogenetic analyses of AIVs isolated from wild birds, poultry, humans, and associated environments suggested that cross-species poultry-to-human transmission of AIVs frequently occurs on poultry farms. In addition, live bird markets are active sites for interspecies dissemination, where AIVs can be transmitted from birds to humans or reassort influenza gene segments in different host species [36, 37].

HPAI H5N1 outbreaks have occurred in a variety of ecological systems with economic, agricultural and environmental differences, which pose the threat to the poultry production sector. Factors affecting the spatial and temporal distribution of the outbreak of AIVs have been investigated in many studies previously. Marius Gilbert and Pfeiffer [38] summarized the risk factors considered for HPAI H5N1 presence in previous studies in nine categories, including(1) farming practice and local biosecurity, (2) poultry and livestock census data with longitude and latitude, (3) anthropogenic variables, (4) socio-economic variables, (5) variables indicative of the presence or abundance of wild birds, (6) variables indicative of the presence or abundance of rivers, lakes or wetlands, (7) eco-climatic variables obtained using

weather station data or remote sensing, (8) land-use and cropping variables, and finally (9) topography. Among these factors, the density of domestic waterfowl, anthropogenic variables (human population density, distance to roads) and indicators of water presence were identified positively correlated with HPAI H5N1 presence across studies and regions.

9.3 Spatial Clustering Analysis of Avian Influenza Viruses Transmission

Commonly used statistical clustering approaches from the literatures to identify spatial distribution patterns and transmission mechanisms can provide additional information for AIV control and prevention strategies.

9.3.1 Cluster Analysis

The spatial distribution pattern of HPAI cases was clustered, dispersed, or randomly distributed, which can be measured by global spatial autocorrelation analysis, such as the global Moran's I statistics [39]. The null hypothesis of global Moran's I is spatial randomness. Global Moran's I index is the correlation coefficient between the eigenvalue and its surrounding values, which can be transformed into z-score and p-value to infer whether the overall spatial distribution has statistically significant clusters. A positive z-score with a statistically significant p-value indicates spatial clustering, and a negative z-score with a statistically significant p-value indicates spatial dispersion. Threshold spatial distances are analyzed by incremental spatial autocorrelation analysis for a series of increasing distances at intervals of interest over a spatial range, and spatial clustering is measured by the z-score of each distance interval. The z-score usually peaks at some distance where the spatial clustering is most salient within the specified spatial extent. The distance associated with the statistically significant peak is selected as the threshold spatial distance for a cluster.

9.3.2 Hotspot Analysis

However, these methods do not account for the location of clusters. The Local Indicator of Spatial Autocorrelation (LISA) with Local Moran's I statistics calculates the eigenvalues of each geographic boundary region and assesses the significance of the region's similarity to its surroundings to identify statistically significant local clusters, such as high-high hot spots and low-low cold spots, or high-low–high local spatial outliers. The high positive z-scores of the test demonstrate the statistically

significant high-high cluster of hotspots [39]. Based on global and local Moran's I analyses, the distribution of H7N9 human cases in Zhejiang Province, China, showed statistically significant spatial autocorrelation in some epidemic waves and identified the statistically significantly high-high clusters and high-low outlier clusters mostly located in the northern part of this province [40]. Shan et al. [41] used global Moran's I analysis to identify that the distribution of H7N9 human cases in Mainland China showed statistically significant spatial autocorrelation during the five epidemic waves.

Liang et al. [42] evaluated environmental factors associated with clusters of outbreaks and multiple subtypes co-circulating of HPAI H5Nx viruses in Taiwan. Global Moran's I analysis was conducted to determine the grid size for covering Taiwan when measuring the clusters of H5Nx outbreak farms and found the optimal distance to be 3 km. Therefore, a 3 km square grid covering Taiwan was used for LISA and local Moran's I statistics, and the results indicated that the hotspots of H5Nx outbreak farms were located on the west coast of Taiwan from 2015 to 2017, where covered more than 75% of outbreaks farms in 2015 and 2017. Multivariate stepwise logistic regressions comparing hotspots and non-hotspots were developed to analyze four categories of variables: farm-related, farm biosecurity-related, wild bird-related, and anthropological. Notably, this study used satellite remote sensing methods to establish unregistered poultry farm data and merged it with the official poultry farm registration database to complete the poultry farm census dataset. A poultry heterogeneity index was also created in this study to describe the heterogeneity of the total number of domesticated waterfowls versus land fowl in each grid. Four risk factors consistently showed a strong association with the spatial clusters of HPAI H5N2 and H5N8 circulations during 2015 and 2017, including high poultry farm density, poultry heterogeneity index, non-registered waterfowl flock density, and a higher percentage of cropping land coverage. Using estimates from the regression models of 2015 and 2017, risk maps were generated to predict high-risk areas and further validated by using outbreaks from the first half of 2018. The results showed that the risk maps for 2015 and 2017 had a prediction rate higher than 55%.

Unlike the local Moran's I statistic, the Getis-Ord Gi* statistic calculates each feature in the context of neighboring features in the dataset to measure the degree of spatial clustering. Features in geographic boundary regions that are similar to adjacent features, and the sum of these features, including the feature itself, that differ significantly from the expected sum, will yield a statistically significant z-core and p-value result. The larger the statistically significant positive z-value, the stronger the aggregation of hot spots, and the smaller the statistically significant negative z-value, the stronger the aggregation of cold spots [43].

In order to investigate the locations of disease clusters, Shan et al. and Huang et al. [41, 44] used the Kernel density estimation to present the clustering areas of human cases caused by infection of AIVs in China in different epidemic waves. Kernel density estimation is a non-parametric statistical method to estimate the probability density function of a random variable. It converts point features into smoothly curved density surfaces by calculating the sum of a kernel function on each data point.

The kernel function calculates the surface value of each point by weighting the distances of all points at each specific location in the distribution. The surface value of each point location is the highest, and it decreases with the distance from the point increases. The density of each point is the sum of all surface values of that point. If more points cluster in one location, the higher the density of that location, as the higher the probability of seeing a point at that location [45]. Based on Kernel density estimation, the Yangtze River Delta region and the Pearl River Delta region had the highest density and the intensity had gradually shifted during the epidemics [41, 46].

9.4 Temporal Spatial Clustering Identification of Avian Influenza Viruses Transmission

Approaches used to early and accurately characterize epidemiologic patterns of disease incidence in a temporal and spatial series are becoming increasingly important. Statistical analysis for detecting spatial–temporal clusters of health-related events is often used for epidemiological and biomedical studies. Timely identification of anomalies of disease or poisoning incidence during ongoing surveillance or an outbreak requires the use of sensitive statistical methods that recognize an incidence pattern at the time of occurrence. The following sections reviewed analytical methods commonly used to study temporal-spatial patterns.

9.4.1 Scan Statistics or Space–Time Permutation Model

Cluster analysis, such as scan statistics, are generally designed for retrospective detection of epidemiologic anomalies in a temporal or space–time series. Spatial scan statistics is a widely-used approach to detect spatiotemporal clustering although several conventional cluster analysis methods such as gap-statistic or K-means have been developed. The scan statistic employs a moving window, possibly with varied shapes, of predetermined radius or geographical unit with fixed population and finds the maximum number of cases revealed through the window as it slides over the entire region [47–49]. The scan test is structured to detect the largest cluster of incidences. The maximum number of events occurring in a window is the test statistic for the scan test. However, calibrating proper spatial and temporal windows in scan statistics is difficult, which requires a process of model tuning. Huang et al. and Dong et al. [44, 46] used the space–time permutation model to analyze the spatial–temporal clustering of H7N9 human cases. Assuming that the population changes are homogeneous, and the spatial extent of the cluster does not change during the scanning process, it only needs the spatial location and time data of the cases. Scan statistics use a varied-size cylindrical moving window with space as the base and time as the height to scan the target area in the time period of interest. Observed and expected numbers of

cases were obtained from the scan of each location and size of the window, and the likelihood ratio or relative ratio statistics were used to evaluate whether there is a cluster in the cylinder. In the space–time permutation model, the spatial and temporal data of the case being studied are used to adjust multiple tests through thousands of random permutations. The cluster with the largest log-likelihood ratio is simulated for each of these permutations of the data set, and the P value for hypothesis testing is used Monte Carlo simulations [50].

According to space–time permutation model scan statistics, the epidemic of H7N9 human cases from 2013 to 2017 showed six statistically significant clusters. In 2017, there were four clusters, with centers located in Beijing, Hubei, Sichuan and Shanghai. One cluster in Xinjiang from July to December 2014, and one cluster in Guangdong from July 2013 to March 2015 [44]. Further analysis of the first two epidemics in 2013–2014 with 5 days as the time unit, in the first and second epidemic waves, two and three statistically significant clusters were identified. In the first wave, the most likely cluster of epidemics was observed in the southeast region centered on Fujian Province from April 27 to May 11, 2013, and the second cluster of epidemics occurred in Jiangsu province and Shanghai from March 13 to April 11, 2013. In the second wave, the earlier cluster of epidemics was in Yangtze River Delta from January 12 to January 31, 2014. The second cluster of epidemics was in Pearl River delta from February 16 to March 2, 2014, and the third cluster of epidemics was in six provinces centered on Anhui Province from April 22 to May 31, 2014.

Zhang et al. [51] also analyze space–time clustering of human infection with H7N9 virus in county level in 2013–2014. The peak z-score indicates that there are obvious spatial clusters at the distance of 30 and 250 km in the incremental spatial autocorrelation analysis, and the distinct temporal clustering at the duration of 14 to 26 days in the temporal autocorrelation analysis. Based on this, 250 km and 14 days are selected as the "Threshold" of distance in space and time for the next space–time hotspot analysis. Getis-Ord Gi* z-score illustrated that there were two statistically significant space–time clustering near Shanghai and Zhejiang in March 26 to April 18, 2013, and near Guangzhou and Shenzhen from February 3 to 4, 2014. Zhang et al. also used a space–time permutation scan statistic model to investigate epidemic pattern of these human cases. The results showed that there were six statistically significant spatiotemporal clusters from 2013 to 2014. The cluster near Shanghai and Zhejiang from March 13, to April 9, 2013, and the cluster near Guangzhou and Shenzhen from February 5 to 25, 2014, were similar to the results of hotspot analysis, indicating the good consistency between these two methods.

9.4.2 Knox Test

Knox proposed a method that allows for statistical testing of the interaction of incidents of infectious disease in space and time that does not use an arbitrary critical value of distance or time for determining local clusters [52]. The Knox statistic is calculated by pairing all possible data points (e.g., location in space and time of the

death of birds) within a clearly defined geographic area and temporal interval and testing them against assigned values of what is "close" in space and time. The number of close space–time data pairs is compared with what would be expected if there were no space–time cluster. Based on Knox settings, Barton and David [53] proposed a "intersection" approach" to obtain spatial–temporal clustering. They suggested to connect the pairs with temporal clustering by line segments to form a temporal map, and then connect the pairs with spatial clustering to form a spatial map. Combining these two maps produces a spatial–temporal clustering [54]. This is reasonable but hard to implement because in a highly dense incidence region, thousands of lines tangled together will make the discrimination among clusters difficult. In addition to the "intersection approach", Openshaw et al. [55] considered a "geographical analysis machine" (GAM) method which draw r-radius circles for the areas with dense incidence when "r" is permitted to varied (say, r = 1, 2, or 4 km). Those corresponding dense circles visually formed bunches of circles, and is decided to be spatially clustered. See also Turnbull [56] for more discussion. In this study published in Scientific Report (2021), Wu et al. [57] showed that Knox-based approach can still display spatiotemporal clusters, in particular when the outbreaks occur in multiple places. When circling the major spatial clusters, each circle has a "diameter" within 3 km, which is the size of the control zone established once HPAI-infected farm identified in Taiwan. When an infected premises (IP) is reported, all poultry from that particular IP will be culled and all farms within 3 km radius of that infected premises will be targeted for intensive surveillance. Therefore, outside the 3 km control zone stands for the spreading of HPAI viruses requiring epidemiological investigation.

9.4.3 Standard Deviational Ellipse (SDE) Method

The standard deviational ellipse (SDE) method was a widely applied approach to displaying geographic distribution of occurrence of some events [58–61], including chronic diseases and infectious diseases, etc. [62–64]. It combines the concern of location, (two-dimensional) dispersion, and orientation (meaning direction plus shape) in a simple optimization calculation. When SDE is used repeatedly over a specified time period (say, every week during the emergent outbreak period), the mean area center is the origin of these two axes [60]. It suggests that one, two, and three standard deviation ellipses will cover approximately 68, 95, and 99% of the points [65]. The orientation of the long axis indicates the direction of the point distribution, therefore, the greater the difference between the long and short axis, the more obvious the direction trend. Connecting the centers of each ellipse offers a clue for disease transmission. This connecting line can be compared with long axes of consecutive ellipses. To sketching the spatial trends of H7N9 human cases over time in China during 2013–2017, SDE analysis was used by Huang et al. and Dong et al. [44, 46]. They analyzed the distribution of cases for each month in the epidemic wave. SDE analysis showed that the first wave of the 2013 epidemic started in the three Yangtze River Delta provinces and spread from Jiangsu to Guangdong. The

second wave occurred in the southeastern coastal provinces from Jiangsu Province to Guangdong Province, and expanded the epidemic area with a coastal orientation, but spread toward the inland in the last two months. The third wave of the epidemic started from the southeast coastal area to Gansu Province, and then gradually narrowed down to the Yangtze River Delta. The fourth wave of the epidemic occurred in the southeastern coastal region and then spread to the northern coastal region. In the fifth wave, the epidemic occurred along the eastern coast and then gradually spread to most of mainland China.

9.4.4 Regression Modeling

The interpretation of the shape of SDE need to be cautious as it might be area-specific. While the connection between ellipses reveals a different story implying the development among sub-areas with dense emergent cases, it shows a temporary geographic pattern or latent mode of spreading of events. Using SDE method to estimate the transmission direction needs mild correction when the concerned infections have become endemic; i.e., the virus tends to be localized and existed there all year round. An alternative approach to estimate the direction of spreading is a regression model proposed in Zinszer et al. [66], hereafter called Zinszer model, which attempted to estimate local transmission directions for Ebola epidemic. It states that for an outbreak event occurred at calendar time T_i and at location (X_i, Y_i) with corresponding explanatory variable (possibly a vector) Z_i, and T_{i+1} is the time of the next (Ebola) outbreak case so that the inter-outbreak "gap time" $\tau_i = T_{i+1} - T_i$ can be modeled as:

$$\tau_i = \beta_0 + \beta_1 X_i + \beta_2 Y_i + \gamma' Z_i + \varepsilon_i.$$

The parameters β_1 and β_2 interpret the inverse of rate of transmission in the direction of X and Y, respectively, usually adopted as the longitude (X) and latitude (Y) of the event spot indexed by "i". Depicting weekly SDEs and connecting consecutive centers to exhibit transmission direction employs parallel idea but roles of time and space interchange: Time interval is now not random; it is fixed to be one week. The magnitude of changes in X and Y are random, implying the velocity (speed plus direction) of transmission.

9.5 Future Perspectives

For infectious diseases such as avian influenza, spatial clustering of outbreaks plays a highly significant role in ecological dynamics and viral spread. However, accurate identifying the spatial cluster and predicting the direction of viral spread requires the knowledge of a number of environmental factors with spatial structures, which are

not only non-randomly distributed across a country but also change through time. Instead of applying complex spatial statistics for clustering tests to detect a series of epidemiological anomalies, development of intuitive tools for timely identification of spatial–temporal clusters will assist control measures to prevent further damage. Wu et al. [57] proposed two visual approaches to identify spatial–temporal cluster with its dynamic change through two-stage methods. In the first stage, they utilized common concepts of Knox test and scan likelihood ratio statistics to determine spatiotemporal cluster. Although there is no universally feasible method to estimate the direction of transmission, the use of SDE in the second stage to visualize the geographical distribution of a series of social, biological or environmental events is still very attractive [64, 67–69]. Geographically, the scale wider than local infections was presented by simply connecting the centroids corresponding to each week's ellipse. If the initial pattern was influenced by local factors, the direction connecting centroids can be exerted by a later "strength" existing among ellipses. Time-varying SDEs are applied to individual spatial clusters, defined by the Knox method, to reveal its local transmission by week. By connecting the consecutive centers of weekly SDEs, the direction of transmission can be easily visualized, which may imply the playing roles of local factors, such as wild bird movement, transportation vehicles, human activities or other meteorological factors acted within the spatial clusters [22, 38, 46, 70–72]. Other non-local factors, such as factors related to poultry market supply networks or the long-distance movement of certain bird species, contributing to the HPAI transmission between spatial clusters can be investigated and differentiated from the local factors [73–75]. Careful identification of influencing factors can help precautionary measures, public health control and prevent further outbreaks. Therefore, a Knox-based combined SDE visualization tool is suggested to identify the spatial-temporal clustering of poultry farm HPAI outbreaks in Taiwan.

On the other hand, AGC (Fig. 9.2)-based second-order aggregation maps based on scan statistics likelihood ratio as two-stage approach in a regular interval provide a quantitative risk in regional level and its dynamic change further indicates the direction of transmission [55]. The likelihood ratio statistic constructed in the first stage considers two "reference populations" to serve as the basis for statistical testing on global and local spatial clustering. A map based on drawing the AGC index, which can capture the aggregation pattern of disease clusters is very useful for displaying hotspots. That is, the aggregation of those sub-regions with higher R_j or AGC index is called hotspots. The identified major clusters are similar in both Knox-based and AGC mapping methods. Although the AGC map inevitably depends on the choice of the critical value of the AGC index, the difference between two results is small. These major spatial clusters or hotspots could share common environmental risk factors contributing to the poultry farm outbreaks by HPAI as we published previously [42]. By monthly depicting the AGC maps, the changes in the hotspot pattern over a period of time also provide clues of the direction of HPAI viral transmission. If the AGC maps of different months remain unchanged, it means that the hotspot is very "stable" in a sense. Note that the formation of AGC map depends on the choice of the cutoff point for the number of clusters. The traditional elbow method based on minimizing

$$AGC \equiv \frac{1}{n}\sum_{i=1}^{k} R_i \text{ , } R_i = \frac{log\lambda_{i(Reference\ 1)} - log\lambda_{i(Reference\ 2)}}{log\lambda_{i(Reference\ 1)} + log\lambda_{i(Reference\ 2)}} \text{ , } i = 1, \ldots, k$$

Fig. 9.2 The aggregation of clustering (AGC) index. It is based on the ratio of the difference between the spatial scan statistics of two reference areas and is used to estimate the clustering of outbreaks in an area in a regular interval

the overall within-cluster variation can be applied, or the more modern gap statistics can be used in the future [76, 77].

In conclusion, various approaches to study spatial or temporal-spatial clustering in infectious diseases have been proposed. However, the knowledge of a number of environmental factors with spatial structures is necessary to accurately identify the spatial clustering in a global or local scale. Development of a visual tool in a webpage will assist in accurate identification of such clustering and predicting the direction of viral spreading.

References

1. Palese P, Shaw M (2007) Orthomyxoviridae: the viruses and their replication, 5th ed. Knipe DMHP, Griffin D, Lamb R, Martin M, Roizman B, Strauss S, editors. Philadelphia: Lippincott Williams and Wilkins: Fields virology
2. Tong S, Zhu X, Li Y, Shi M, Zhang J, Bourgeois M et al (2013) New world bats harbor diverse influenza A viruses. PLoS Pathog 9(10):e1003657
3. (CDC) CfDCaP (1997) Isolation of avian influenza A(H5N1) viruses from humans–Hong Kong, May–December 1997. MMWR Morb Mortal Wkly Rep 46(50):1204–7
4. Group WOFHNEW (2008) Toward a unified nomenclature system for highly pathogenic avian influenza virus (H5N1). Emerg Infect Dis 14(7):e1
5. Anderson TK, Macken CA, Lewis NS, Scheuermann RH, Van Reeth K, Brown IH et al (2016) A phylogeny-based global nomenclature system and automated annotation tool for H1 hemagglutinin genes from swine influenza A viruses. mSphere 1(6)
6. (WHO) WHO (1980) A revision of the system of nomenclature for influenza viruses: a WHO memorandum. Bull World Health Organ 58(4):585–91
7. Klenk HD, Garten W, Bosch FX, Rott R (1982) Viral glycoproteins as determinants of pathogenicity. Med Microbiol Immunol 170(3):145–153
8. Kawaoka Y, Webster RG (1988) Sequence requirements for cleavage activation of influenza virus hemagglutinin expressed in mammalian cells. Proc Natl Acad Sci U S A 85(2):324–328
9. Vey M, Orlich M, Adler S, Klenk HD, Rott R, Garten W (1992) Hemagglutinin activation of pathogenic avian influenza viruses of serotype H7 requires the protease recognition motif R-X-K/R-R. Virology 188(1):408–13
10. Wood GW, McCauley JW, Bashiruddin JB, Alexander DJ (1993) Deduced amino acid sequences at the haemagglutinin cleavage site of avian influenza A viruses of H5 and H7 subtypes. Arch Virol 130(1–2):209–217
11. Alexander DJ (2007) An overview of the epidemiology of avian influenza. Vaccine 25(30):5637–5644
12. Lee DH, Criado MF, Swayne DE (2021) Pathobiological origins and evolutionary history of highly pathogenic avian influenza viruses. Cold Spring Harb Perspect Med 11(2)

13. Naguib MM, Verhagen JH, Mostafa A, Wille M, Li R, Graaf A et al (2019) Global patterns of avian influenza A (H7): virus evolution and zoonotic threats. FEMS Microbiol Rev 43(6):608–621
14. Selleck PW, Arzey G, Kirkland PD, Reece RL, Gould AR, Daniels PW et al (2003) An outbreak of highly pathogenic avian influenza in Australia in 1997 caused by an H7N4 virus. Avian Dis 47(3 Suppl):806–811
15. Chen H, Smith GJD, Zhang SY, Qin K, Wang J, Li KS et al (2005) Avian flu: H5N1 virus outbreak in migratory waterfowl. Nature 436:191–192
16. Liu J, Xiao H, Lei F, Zhu Q, Qin K, Zhang XW et al (2005) Highly pathogenic H5N1 influenza virus infection in migratory birds. Science 309(5738):1206
17. Lycett SJ, Pohlmann A, Staubach C, Caliendo V, Woolhouse M, Beer M, Kuiken T, Global Consortium for HN, Related Influenza V (2020) Genesis and spread of multiple reassortants during the 2016/2017 H5 avian influenza epidemic in Eurasia. Proc Natl Acad Sci USA 117:20814–20825
18. Prosser DJ, Chen J, Ahlstrom CA, Reeves AB, Poulson RL, Sullivan JD et al (2022) Maintenance and dissemination of avianorigin influenza A virus within the northern Atlantic Flyway of North America. PLoS Pathog 18(6):e1010605
19. Cuia Y, Li Y, Lia M, Zhaoa L, Wang D, Tian J, Bai X, Cia Y et al (2020) Evolution and extensive reassortment of H5 influenza viruses isolated from wild birds in China over the past decade. Emerg Microbes Infect 9(1):1793–1803
20. Runstadler J, Hill N, Hussein IT, Puryear W, Keogh M (2013) Connecting the study of wild influenza with the potential for pandemic disease. Infect Genet Evol 17:162–187
21. Taubenberger JK, Morens DM (2009) Pandemic influenza–including a risk assessment of H5N1. Rev Sci Tech 28(1):187–202
22. Global Consortium for H5N8 and Related Influenza Viruses (2016) Role for migratory wild birds in the global spread of avian influenza H5N8. Science 354(6309):213–7
23. Kwon J, Youk S, Lee DH (2022) Role of wild birds in the spread of clade 2.3.4.4e H5N6 highly pathogenic avian influenza virus into South Korea and Japan. Infect Genet Evol 101:105281
24. Verhagen JH, Fouchier RAM, Lewis N (2021) Highly pathogenic avian influenza viruses at the wild-domestic bird interface in Europe: future directions for research and surveillance. Viruses 13(2):212
25. Bevins SN, Shriner SA, Cumbee JC Jr, Dilione KE, Douglass KE, Ellis JW et al (2022) Intercontinental movement of highly pathogenic avian influenza A(H5N1) clade 2.3.4.4 virus to the United States, 2021. Emerg Infect Dis 28(5):1006–1011
26. van der Kolk JH (2019) Role for migratory domestic poultry and/or wild birds in the global spread of avian influenza? Vet Q 39(1):161–167
27. Poen MJ, Bestebroer TM, Vuong O, Scheuer RD, van der Jeugd HP, Kleyheeg E et al (2018) Local amplification of highly pathogenic avian influenza H5N8 viruses in wild birds in the Netherlands, 2016 to 2017. Euro Surveill 23(4):17–00449
28. Kwon YK, Joh SJ, Kim MC, Lee YJ, Choi JG, Lee EK, Wee SH, Sung HW, Kwon JH, Kang MI, Kim JH (2005) Highly pathogenic avian influenza in magpies (Pica pica sericea) in South Korea. J Wildl Dis 41:618–623
29. Lee CW, Suarez DL, Tumpey TM, Sung HW, Kwon YK, Lee YJ, Choi JG et al (2005) Characterization of highly pathogenic H5N1 avian influenza A viruses isolated from South Korea. J Virol 79:3692–3702
30. Caliendo V, Leijten L, van der M Bildt, Germeraad E, Fouchier RAM, Beerens N, Kuiken T (2022) Tropism of highly pathogenic avian influenza H5 viruses from the 2020/2021 epizootic in wild ducks and geese. Viruses 14(2):280
31. Engelsma M, Heutink R, Harders, Germeraad EA, Beerens N (2022) Multiple introductions of reassorted highly pathogenic avian influenza H5Nx viruses clade 2.3.4.4b causing outbreaks in wild birds and poultry in The Netherlands, 2020–2021. Microbiol Spectr 10(2):e0249921
32. Hill NJ, Bishop MA, Trovão NS, Ineson KM, Schaefer AL, Puryear WB, Zhou K et al (2022) Ecological divergence of wild birds drives avian influenza spillover and global spread. PLoS Pathog 18(5):e1010062

33. Tang L, Tang W, Li X, Hu C, Wu D, Wang T, He G (2020) Avian influenza virus prevalence and subtype diversity in wild birds in Shanghai, China, 2016–2018. Viruses 12(9):1031
34. Runstadler J, Hill N, Hussein IT, Puryear W, Keogh M (2013) Connecting the study of wild influenza with the potential for pandemic disease. Infect Genet Evol 17:162–187
35. Hicks JT, Edwards K, QiuI X, Kim DK, Hixson J, Krauss S et al (2022) Host diversity and behavior determine patterns of interspecies transmission and geographic diffusion of avian influenza A subtypes among North American wild reservoir species. PLoS Pathog 18(4):e1009973
36. Chen Y, Liang W, Yang S, Wu N, Gao H, Sheng J et al (2013) Human infections with the emerging avian influenza A H7N9 virus from wet market poultry: clinical analysis and characterisation of viral genome. Lancet 381(9881):1916–1925
37. Zhang T, Bi Y, Tian H, Li X, Liu D, Wu Y et al (2014) Human infection with influenza virus A(H10N8) from live poultry markets, China, 2014. Emerg Infect Dis 20(12):2076–2079
38. Gilbert M, Pfeiffer DU (2012) Risk factor modelling of the spatio-temporal patterns of highly pathogenic avian influenza (HPAIV) H5N1: a review. Spat Spatiotemporal Epidemiol 3(3):173–183
39. Souris M, Bichaud L (2011) Statistical methods for bivariate spatial analysis in marked points. Examples in spatial epidemiology. Spat Spatiotemporal Epidemiol 2(4):227–34
40. Wu H, Wang X, Xue M, Wu C, Lu Q, Ding Z et al (2017) Spatial characteristics and the epidemiology of human infections with avian influenza A(H7N9) virus in five waves from 2013 to 2017 in Zhejiang Province, China. PLoS One 12(7):e0180763
41. Shan X, Wang Y, Song R, Wei W, Liao H, Huang H et al (2020) Spatial and temporal clusters of avian influenza a (H7N9) virus in humans across five epidemics in mainland China: an epidemiological study of laboratory-confirmed cases. BMC Infect Dis 20(1):630
42. Liang WS, He YC, Wu HD, Li YT, Shih TH, Kao GS et al (2020) Ecological factors associated with persistent circulation of multiple highly pathogenic avian influenza viruses among poultry farms in Taiwan during 2015–17. PLoS One 15(8):e0236581
43. ArcGIS Pro 2.8. How Hot Spot Analysis (Getis-Ord Gi*) works. https://pro.arcgis.com/en/pro-app/2.8/tool-reference/spatial-statistics/h-how-hot-spot-analysis-getis-ord-gi-spatial-stati.htm. Accessed 17 Jun 2022
44. Huang D, Dong W, Wang Q (2021) Spatial and temporal analysis of human infection with the avian influenza A (H7N9) virus in China and research on a risk assessment agent-based model. Int J Infect Dis 106:386–394
45. Krisp JM, Špatenková O (2010) Kernel density estimations for visual analysis of emergency response data. Geographic information and cartography for risk and crisis management, pp 395–408
46. Dong W, Yang K, Xu Q, Liu L, Chen J (2017) Spatio-temporal pattern analysis for evaluation of the spread of human infections with avian influenza A(H7N9) virus in China, 2013–2014. BMC Infect Dis 17(1):704
47. Duczmal L, Kulldorff M, Huang L (2006) Evaluation of spatial scan statistics for irregularly shaped clusters. J Comput Gr Stat 428–42
48. Kulldorff M, Huang L, Pickle L, Duczmal L (2006) An elliptic spatial scan statistic. Stat Med 25(22):3929–3943
49. Kulldorff M, Hjalmars U (1999) The Knox method and other tests for space-time interaction. Biometrics 55(2):544–552
50. Kulldorff M, Heffernan R, Hartman J, Assuncao R, Mostashari F (2005) A space-time permutation scan statistic for disease outbreak detection. PLoS Med 2(3):e59
51. Zhang Y, Shen Z, Ma C, Jiang C, Feng C, Shankar N et al (2015) Cluster of human infections with avian influenza A (H7N9) cases: a temporal and spatial analysis. Int J Environ Res Public Health 12(1):816–828
52. Knox EG, Bartlett MS (1964) The detection of space-time interactions. J Roy Stat Soc Ser C (Appl Stat) 13(1):25–30
53. Barton DE, David FN (1966) The random intersection of two graphs. Wiley, London, New York. Research papers in statistics: festschrift for J. Neyman, pp 455–9

54. Mantel N (1967) The detection of disease clustering and a generalized regression approach. Cancer Res 27(2):209–220
55. Openshaw S, Craft AW, Charlton M, Birch JM (1988) Investigation of leukaemia clusters by use of a Geographical Analysis Machine. Lancet 1(8580):272–273
56. Turnbull BW, Iwano EJ, Burnett WS, Howe HL, Clark LC (1990) Monitoring for clusters of disease: application to leukemia incidence in upstate New York. Am J Epidemiol 132(1 Suppl):S136–S143
57. Wu HI, Chao DY (2021) Two-stage algorithms for visually exploring spatio-temporal clustering of avian influenza virus outbreaks in poultry farms. Sci Rep 11(1):22553
58. Lefever DW (1926) Measuring geographic concentration by means of the standard deviational ellipse. Am J Sociol 32(1):88–94
59. Furfey PH (1927) A note on lefever's "Standard Deviational Ellipse." Am J Sociol 33(1):94–98
60. Yuill RS (1971) The standard deviational ellipse; an updated tool for spatial description. Geogr Annal Ser B Hum Geogr 53(1):28–39
61. Wang B, Shi W, Miao Z (2015) Confidence analysis of standard deviational ellipse and its extension into higher dimensional euclidean space. PLoS One 10(3):e0118537
62. Cressie N, Chan NH (1989) Spatial modeling of regional variables. J Am Stat Assoc 84(406):393–401
63. Eryando T, Susanna D, Pratiwi D, Nugraha F (2012) Standard Deviational Ellipse (SDE) models for malaria surveillance, case study: Sukabumi district-Indonesia, in 2012. Malar J 11(1):P130
64. Satoto TBT, Satrisno H, Lazuardi L, Diptyanusa A, Purwaningsih, Rumbiwati et al (2019) Insecticide resistance in Aedes aegypti: an impact from human urbanization? PLoS One 14(6):e0218079
65. Reference APT. How Directional Distribution (Standard Deviational Ellipse) works. https://pro.arcgis.com/en/pro-app/2.7/tool-reference/spatial-statistics/h-how-directional-distribution-standard-deviationa.htm. Accessed 3 Jul 2021
66. Zinszer K, Morrison K, Anema A, Majumder MS, Brownstein JS (2015) The velocity of Ebola spread in parts of west Africa. Lancet Infect Dis 15(9):1005–1007
67. Spumont F, Viti F (2018) The effect of workplace relocation on individuals' activity travel behavior. J Trans Land Use 11(1):985–1002
68. Liu S, Qin Y, Xie Z, Zhang J (2020) The spatio-temporal characteristics and influencing factors of Covid-19 spread in Shenzhen, China-an analysis based on 417 cases. Int J Environ Res Public Health 17(20)
69. Moore TW, McGuire MP (2019) Using the standard deviational ellipse to document changes to the spatial dispersion of seasonal tornado activity in the United States. npj Clim Atmos Sci 2(1):21
70. Artois J, Jiang H, Wang X, Qin Y, Pearcy M, Lai S et al (2018) Changing geographic patterns and risk factors for avian influenza A(H7N9) infections in humans, China. Emerg Infect Dis 24(1):87–94
71. Busani L, Valsecchi MG, Rossi E, Toson M, Ferre N, Pozza MD et al (2009) Risk factors for highly pathogenic H7N1 avian influenza virus infection in poultry during the 1999–2000 epidemic in Italy. Vet J 181(2):171–177
72. Paul M, Tavornpanich S, Abrial D, Gasqui P, Charras-Garrido M, Thanapongtharm W et al (2010) Anthropogenic factors and the risk of highly pathogenic avian influenza H5N1: prospects from a spatial-based model. Vet Res 41(3):28
73. Hogerwerf L, Wallace RG, Ottaviani D, Slingenbergh J, Prosser D, Bergmann L et al (2010) Persistence of highly pathogenic avian influenza H5N1 virus defined by agro-ecological niche. EcoHealth 7(2):213–225
74. Lai PC, Wong CM, Hedley AJ, Lo SV, Leung PY, Kong J et al (2004) Understanding the spatial clustering of severe acute respiratory syndrome (SARS) in Hong Kong. Environ Health Perspect 112(15):1550–1556
75. Leibler JH, Otte J, Roland-Holst D, Pfeiffer DU, Soares Magalhaes R, Rushton J et al (2009) Industrial food animal production and global health risks: exploring the ecosystems and economics of avian influenza. EcoHealth 6(1):58–70

76. Everitt BS, Landau S, Leese M, Stahl D (2011) Cluster analysis, 5th ed. Wiley
77. Tibshirani R, Walther G, Hastie T (2001) Estimating the number of clusters in a data set via the gap statistic. J R Stat Soc B 63:411–423

Dr. Huang's research focuses on the epidemiology of infectious diseases. In addition to exploring the distribution of infectious diseases from different perspectives, she is also interested in molecular diagnostic methods for infectious diseases.

Dr. Wu is an associate professor of statistics at the National Chung-Hsing University. His speciality and research interests include: survival analysis, spatiotemporal data analysis, and epidemiologic study design.

Dr. Chao is a professor at Institute of Microbiology and Public Health, National Chung-Hsing University and specialized in infectious disease epidemiology. Her research interests mainly focus on using different tools, either geospatial or molecular, to describe the transmission of zoonotic infectious diseases. She also serves in editorial boards of Frontiers in Microbiology and Microbiology Spectrum.

Chapter 10
Detecting Urban form Using Remote Sensing: Spatiotemporal Research Gaps for Sustainable Environment and Human Health

Tzu-Hsin Karen Chen, Alexander V. Prishchepov, and Clive E. Sabel

Abstract Remote sensing offers large-scale and longitudinal assessment of the size, density, and function of cities associated with the sustainable environment and human well-being. In this chapter, we synthesize 376 peer-reviewed studies on urban land cover, building density, three-dimensional (3-D) structure, and land use using remote sensing approaches. We evaluated the sources of data, detection methods, as well as the spatiotemporal characteristics (e.g., locations and spatiotemporal scales). Our review identifies three research gaps: (1) Many urbanization studies monitor urban/non-urban change for a long period but not for the patterns of 3-D urban structure; (2) Increasing number of studies use deep learning approaches to detect urban land cover in large scales, especially with Sentinel-2, but there is a lack of time-series analysis and temporal accuracy assessment; (3) most of the urban land change studies focused on North America and East Asia but not in the Global South. A dilemma lies behind these research gaps: newer, high-resolution imagery, able to detect nuanced urban attributes, has a relatively short temporal span. In contrast, older imagery can detect long-term changes but has a lower resolution. This problem has led to a considerable paucity in investigating long-term urban dynamics. For instance, most of the studies investigating 3-D urban form covered less than five years. Dealing with these issues, recent developments in data fusion, temporal accuracy assessment, object-based image analysis, and deep learning methods are showing promise to enhance spatial resolution, extend temporal coverage, and to characterize land use intensity

T.-H. K. Chen (✉)
Yale Institute for Biospheric Studies, Yale University, New Haven, CT 06511, USA
e-mail: karen.t.chen@yale.edu

T.-H. K. Chen · C. E. Sabel
Danish Big Data Centre for Environment and Health (BERTHA), Aarhus University, 4000 Roskilde, Denmark

A. V. Prishchepov
Department of Geosciences and Natural Resource Management (IGN), University of Copenhagen, København K 1350, Denmark

C. E. Sabel
Department of Public Health, Aarhus University, 8002 Aarhus, Denmark

Health Research Institute, University of Canberra, Canberra, Australia

T.-H. Wen et al. (eds.), *Earth Data Analytics for Planetary Health*, Atmosphere, Earth, Ocean & Space, https://doi.org/10.1007/978-981-19-8765-6_10

and 3-D structure, which are important factors affecting temperature, physical activities associated with public health. The increasing availability of computational power such as via Google Earth Engine allows analysis at large spatiotemporal scales such as comparing urban form and sustainable/health outcomes across multiple cities. We foresee increasing importance of remote sensing in providing evidence-based knowledge for policies and science of health cities.

Keywords Urban form · Sustainable cities · Urban remote sensing · Machine learning · Urban dynamics · Time series analysis

10.1 Introduction

Urbanization is amongst the top challenges faced globally, adding another approximately 2.5 billion urban dwellers between 2018 and 2050, which represents an increase in the global population from 56 to 68% [1]. The unprecedented urban land expansion competes with agricultural and natural lands, with the total area of land predicted to be converted to urban areas exceeding 600,000 km^2 between 2015 and 2050 [2], affecting how the Earth physical system functions today, how it supports life, and how conditions might change to alter climate. Not only does urban expansion drive environmental change, but urbanization and the changing built environment also shape *urban form*—the shape, size, density, and layout of cities. These urban form attributes affect dweller's daily experience, salutogenic or harmful to their health and well-being: heat stress [3, 4], air quality [5], food consumption [6, 7], disease outbreaks [8, 9], and biodiversity and access to green space [10], mood and psychological distress [11, 12]. One urban form attribute might be favorable for a sustainable goal but not for another. For instance, denser environments, one of the features of compact city theory, may potentially promote walking reduce the risk of obesity [13, 14].

Therefore, the development of sustainable urban form needs backing up by scientific knowledge and comprehensive information [15]. There have been attempts to evaluate urban form impacts on coupled human-environment systems, from global change [10, 16] to health sciences [9, 17, 18]. A recent review, however, found that the evidence-based knowledge of sustainable urban form does not reach a consensus between locations and times [19]. For instance, how high-density development relates to energy consumption is contested in the literature and varies by study area. Compact cities were found to promote short-distance daily travel and reduce emissions in many US cities [20–23]. At the same time, another study in Norway considering leisure trips at the weekend showed lower urban density could reduce energy consumption [24, 25]. The relationship between air pollution and densification is also disputed in the literature. A US study found that high-density cities were associated with lower air pollution [26], while air quality was negatively correlated with high-density development in India [27] and European large urban zones [28]. Some studies, using the same methodology, within US metropolitan regions, have pointed to opposite

correlations between population density and outdoor concentrations of fine particulate matter (PM2.5) when one used data for a single year [29] and another used data for five years [5]. Going forward with evidence-based solutions, scientists face a significant challenge when there are inconsistencies in the temporal and spatial scales for measuring characteristics of the built environment and desired sustainable outcomes. More harmonized data collection and interpretation are urgent for improving the understanding of our living environment and provide knowledge that is useful for scientists, urban planners, and policymakers, to create a sustainable urban future.

Remote sensing is a valuable source of data, which is complementary to survey data due to its high temporal and spatial resolution. On average, a household in Africa waits for over 1000 years to be surveyed once, and in North America it is 50–100 years [30]. In contrast, the frequency of publicly available earth observations ranges from daily to monthly. With its global availability and low costs, time-series satellite image analysis has been used to identify urban sprawl and shrinkage in many parts of the world. For instance, urban expansion has been investigated with Landsat time series over more than two decades in India [31], the United States [32, 33], Japan [34], and China [35]. Cross-country research also accelerates with the release of global multitemporal urban data, such as Global Annual Urban Dynamics (GAUD) [36], Global Artificial Impervious Area (GAIA) [37], and Global Human Settlement Layer (GHSL) [38], World Settlement Footprint [39], and European Space Agency's Climate Change Initiative (CCI) land cover data.

In contrast to remote sensing approaches, there have also been several studies that used simulation models, such as Markov chains and cellular automata, to quantify urban growth. Remote sensing, nevertheless, remains an indispensable approach because the simulation models are based on historical data often derived from satellite or airborne imagery [40, 41]. Additionally, continuous earth observations are essential to capture nonlinearities in time, because simulation models often fail to capture shifting points in urban systems when policies move in another direction [42]. For instance, satellite imagery allowed to detect the building damage and land-use changes due to environmental disturbances, such as the 2015 Gorkha earthquake in Nepal [43] and the 2011 Fukushima nuclear disaster in Japan [44].

Existing review articles provide a valuable synthesis of how remote sensing data can allow and limit the analysis of urban form spatially and temporally. For example, Weng [45] highlighted the impact of spatial resolution on urban mapping. With a lack of high-resolution images before 2000, there had been scarce research in remote sensing of the built environment due to the requirement of a minimum resolution of ~1 m to detect buildings using pixel-based classification methods. Bhatta et al. [46] underlined the role of spatial resolution in measuring urban sprawl, referring to a problem where low- and high-density developments may not be distinguishable using 30 m-resolution images. Recently, Zhu et al. [47] pinpointed high-frequency analysis as a key research gap in urban remote sensing, undermining the understanding of vital urban processes. Reba and Seto [48] clarified that some studies observed urban land change at high frequencies, but they mostly cover a short time period, thus not providing comprehensive information about the non-linear and long-term

process of evolution of the urban environment. Most of the reviews acknowledged that spatiotemporal characteristics are the major restraining factors of urban form detection research associated with urban sustainability. However, a synthesis of how data and methods affect the spatiotemporal focuses of urban form studies and potential solutions is still lacking.

Our goal in this chapter is to provide a review of the literature on urban form detection to identify spatiotemporal patterns of these studies and their association with adopted data and methods. We define the scope of urban form by the following attributes: urban land cover, building density, three-dimensional (3-D) structure, and land use, which are the fundamental components for deriving other urban form characteristics, such as continuity, nuclearity, and land use mix [15]. These four components can be remotely sensed and reflect a continuum of the ways of thinking, from binary, continuous, to contextual features of the built environment. We aim to answer the following questions: (1) Which data have been used for mapping key urban land attributes at certain scales in space and time? (2) Whether commonly used remote sensing data are suited for characterizing simple and complex urban form features at various spatiotemporal scales? And (3) What are the emerging methods and data that could help to fill the spatiotemporal research gaps? Thus, we identify spatial characteristics in the remote sensing literature for each urban form component, by comparing at which spatial unit comparisons were made, and at what scale the features were analyzed. We also dissect the temporal properties of the research—whether the dynamics of cities are assessed, and if so, over what frequency and time scales changes have been investigated.

We structure the chapter by first describing the background of the four components of urban form and specifying the methods of the systematic review. In the results section, we start with the bibliometric sources of selected papers (Sect. 4.1). Then, we synthesize prevalent remote sensing data sources, data integration (Sect. 4.2), and classification techniques (Sect. 4.3). Lastly, we visualize the global distribution and time scale of studied cities, and present statistics of spatiotemporal characteristics thorough out the four categories of urban form literature (Sect. 4.4). Based on the results, we rethink how commonly used remote sensing data have constrained urban form studies and propose recommendations for future studies from a perspective of environmental health and sustainable urban development.

10.2 Background of Conceptualizing Urban Form

Urban land cover

Urban land cover occupies less than 1% of Earth's land surface area [2], but its extent, distribution, and evolution have enormous impact on environmental and socioeconomic dynamics worldwide [10, 16, 49–51]. Detailed and accurate measures of urban land cover have been fundamental for estimating resource consumption of human activities and the impacts of urban land expansion on arable land, natural habitats,

and environmental degradation. The definitions of urban areas vary by country in United Nations documentation, which considers a range of factors such as population, industry, or urban infrastructure [47]. Many environmental studies use remote sensing approaches to consistently map urban extent by classifying impervious surface, represented by the physical state of the land surface [15, 52, 53].

Building density

Building density defines the amount of built-up area in each geographic area, which is also known as built-up area ratio or fractional built-up areas. The thresholds of building density have been commonly used to define levels of urban development for government data, such as the National Land Cover Database in the United States [54]. Building density is fundamental parameter in urban planning and design [19], related to spacing between buildings [55], and other landscape metrics such as edge density [56]. Often landscape ecologists use urban patch and edge density to evaluate impacts of urban expansion on ecosystem function and biodiversity [57, 58], while urban planning researchers use compactness and sprawling metrics to evaluate urban form's impacts on economic, energy, environmental quality, and public health goals [16, 59].

3-D structure

Three-dimensional structure reflects building height, building volume, canyon geometry, and surface roughness (see Table 10.1). In climate science, the height to floor area ratio has been used to evaluate the thermal mass of the built environment [60] and has been found useful for simulating local climate within cities [61, 62]. Air flow and wind simulations also require surface roughness parameters such as roughness length and porosity [63, 64]. Other three-dimensional structure proxies related to canyon geometry, such as the height to (road) depth ratio and the sky view factor, are associated to crowding stress, thermal comfort, and safety perception people experience in their daily life [12, 65, 66]. Recently, a notable framework Local Climate Zone has mingled the horizontal and vertical dimensions by compactness and height (e.g., *compact high, compact, sparse high, sparse low*) [62]. This framework has been increasingly used in the urban heat island and climatology literature [67, 68].

Land use

Urban land use or functional zoning represent the types of socio-economic functions and human activities in cities [75]. It has been typically categorized, for instance, in the USGS classification system, as residential, commercial, industrial, transportation or other land use [76]. However, such classifications can also depend on the geographical context. While regular the classification of land use is more formal in countries where land use is largely regulated by urban planning laws [77], the diverse characteristics of land use in informal settlements or/and developing countries are found [78–80].

Table 10.1 Common metrics to measure 2-D and 3-D built-up intensity

Attribute	Metric	Definition	Representative studies
Two-dimensional density	Built-up area ratio	Proportion of built-up from the total area	[42]
	Floor area ratio	Ratio of total floor area to lot area	[69]
	Housing density	Number of house/household per land unit	[55]
	Patch density	Number of patches per unit area	[70]
	Road density	Number of roads per unit area	[71]
	Edge density	Total length of edge per unit area	[72]
Three-dimensional structure	Building volume density	Ratio of the volume of building to lot area	[61]
	Height to depth ratio	Ratio of the average height of the buildings along a road to the depth of the road	[65]
	Height to floor area ratio	Ratio of the height of buildings to the total floor area	[73]
	Sky view factor	Ratio of the area of the visible sky to the area of a hemisphere centered at a certain location	[60]
	Roughness length	Height where the wind velocity is equal to zero	[64, 74]
	Urban porosity	Ratio of the open air volume in the urban canopy layer	

10.3 Methods

The systematic literature search was conducted using Scopus databases to search journal articles published after 1990. Very few studies in urban form detection were found in the early 1990s because high-resolution images were not available yet, and medium-resolution images such as Landsat and SPOT were expensive at that time [45]. We designed three criteria and used search terms of publication title, abstract, or keywords to collect relevant literature:

- The publication had to use remote sensing data (e.g., “images”, “street view”).

- The method used for urban form detection had to be an automatic approach (e.g., "classification", "machine learning").
- The publication provides empirical results of one of the urban form components: urban extent, two-dimensional density, three-dimensional morphological attributes, or land use (terms listed in Table 10.2).

The full query to ensure an extensive collection of literature of the three criteria is listed in Table 10.5 (Appendix). The search cutoff was on December 31, 2018. The initial search led to the selection of 1,044 papers. Following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [81], we recorded bibliographic data (e.g., authors and journal), the context of study cases (i.e., place and time), data (i.e., sensor and resolution) and methods enabling urban form detection (Table 10.6). Publications that failed to provide the necessary information, did not conduct urban form detection (i.e., only analyzed urban form products from another study), or did not report empirical results, were excluded. In addition, we found a few studies that covered more than one urban form component. For instance, some studies drawing on 3-D structure and land use also presented the extent of urban land cover. These cases, containing urban form attributes more detailed than the dichotomy, were accounted for in the 3-D structure or land-use groups and removed from the urban land cover group.

We analyzed the trend of academic activities of the urban form detection literature by calculating citation counts, the number of articles for each year, and the annual citation counts per article. For comparison, we additionally present annual citation counts per article for two larger communities: remote sensing and urban sustainability. The general remote sensing community was represented by articles published by journals including "remote sensing" in the journal title. The urban sustainability community was represented by research articles found with "urban

Table 10.2 Summary of reviewed articles on urban form attribute extraction

#	Attribute	Terms	Initial result	Final result
1	Land cover	Impervious surface OR urban extent OR built-up area	489	208
2	Building density	Building density OR urban density OR building cover ratio OR subpixel impervious surface	164	64
2	Three-dimensional structure	Building height OR urban structure OR urban morphology OR building volume OR sky view factor OR roughness length OR urban porosity	167	38
3	Land use	Urban land use mapping (classification) OR functional zone OR urban land use change	224	66
	Total		1,044	376

Table 10.3 Scheme for surveying spatiotemporal characteristics

Characteristic	Definition	Example
Spatial resolution	Clarity of spatial data that are analyzed for attribute extraction, or clarity of the resulted map	30-m resolution images
Spatial unit	Unit of single observations	A county-scale study, comparing observations in 36 counties
Spatial scale	Included area of observations	An international study
Temporal frequency	Density of multi-temporal observations	Every-five-year survey
Time scale	Length of time between the first and the last observation	25-year observation of urban expansion

sustainability", "sustainable urban development", or "sustainable urbanization" in their title, abstract, or keywords.

To present the spatiotemporal characteristics of the urban form literature, we counted the distribution of studies over time and space. While scientists sometimes use intermingled definitions of scale, we adopted overarching definitions to help communication between remote sensing and sustainability sciences (Table 10.3). Five spatiotemporal characteristics of data and studies, including spatial resolution, scale, extent, temporal frequency, and coverage, were analyzed.

10.4 Results

10.4.1 Synopsis of the Reviewed Study Cases

In total, 376 articles that mapped urban form were published between 1990 and 2018. Most reported was urban land cover (208 papers), followed by land use (66 papers), two-dimensional building density (64 papers), and finally three-dimensional structure (38 papers). In total, the publications covered 810 cities globally. An increasing interest in quantifying urban form was observed (Fig. 10.1), which was reflected in the growth of citation counts per article per year. Since 2000, publications have gradually risen in number, associated with when medium-resolution images such as Landsat and SPOT were more readily available. Following the 2007 launch of the urban-focused satellites TerraSAR-X and WorldView-1, followed soon after by many others, the number of urban form studies has substantially increased. Another boost of publications was observed between 2016 and 2017, again, after the launch of Copernicus Sentinel-1 and Sentinel-2 satellites. In contrast with the remote sensing and urban sustainability fields, the overlapping area of urban form detection was very dynamic. The average annual citation among the summarized studies was higher than

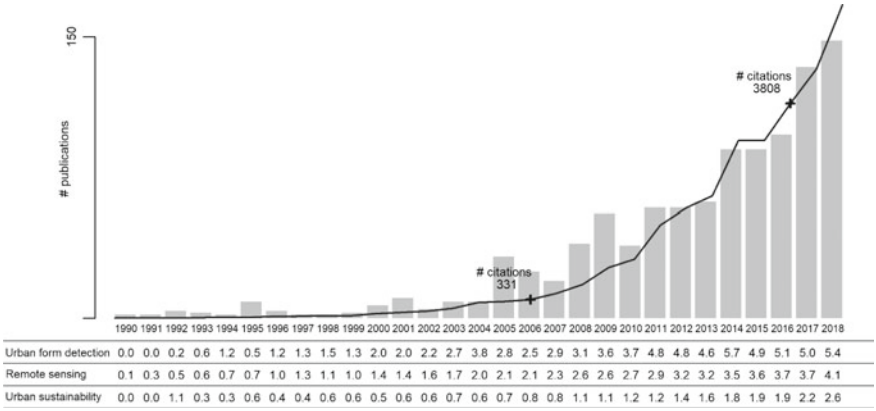


Fig. 10.1 Yearly publications and citations from 1990 to 2018 of the literature on urban form detection. The average annual citation (number of citation per article per year) is presented at the bottom to compare with the remote sensing and the urban sustainability community

4.5 since 2011, while at the same time, it was 2.9 for the whole field of remote sensing and 1.2 for the urban sustainability literature.

10.4.2 Remote Sensing Data Used to Detect Urban Form

Data used for mapping urban land cover

Passive remote sensing was the most commonly used data for mapping urban land cover in the past two decades (Fig. 10.2). Multispectral imagery alone can be used to distinguish the impervious surface from green space without other supporting data sources [82, 83]. Among optical imagery sensors, freely accessible imagery from Landsat satellites, which covers a long time span, was predominantly used (129 case studies), followed by SPOT (18), QuickBird (15), MODIS (13), WorldView (8), IKONOS (8), and Sentinel-2 (4) (Table 10.4). Nighttime light data, such as the Defense Meteorological Satellite Program–Operational Linescan System (DMSP-OLS) data, available since 1992, has been used for mapping global urban land cover before data from MODIS and most of the SAR sensors were released [84]. Freely available data with an increase in spatial resolution, such as Sentinel-2 (2015) and Visible Infrared Imaging Radiometer Suite (VIIRS) nighttime light data (2012), have brought a noticeable improvement of urban mapping accuracy [85, 86], and consequently we observed more applications in the past five years (Fig. 10.2). Fewer papers use coarse resolution (>100 m) imagery in urban studies, but MODIS data is worthy of discussion, as its high temporal resolution (daily), allowing to compensate for cloud cover, and the coarse resolution minimizes computation costs, combine to make global urban mapping feasible [87].

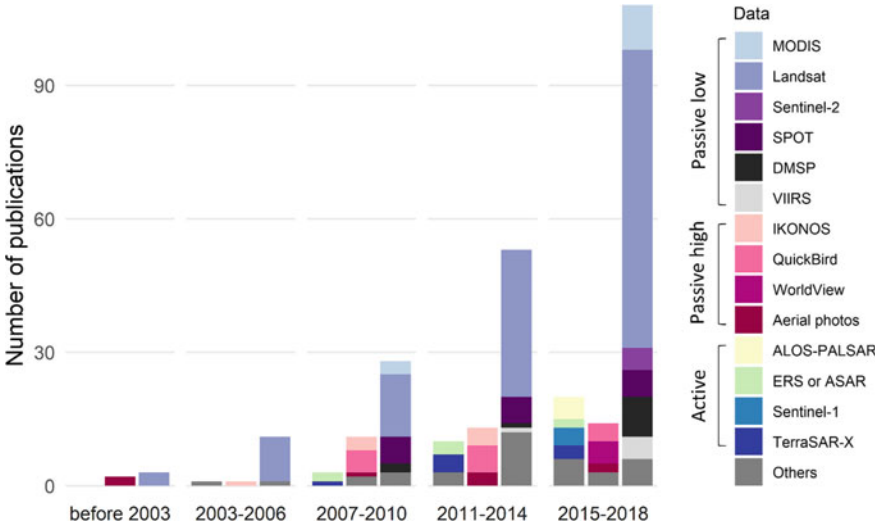


Fig. 10.2 The remote sensing data used, over time, for mapping urban land cover. The data include images collected by active sensors and passive sensors, at low (>=5 m) and high resolution (<5 m)

We use Fig. 10.3 to present an overview of data sources commonly used for urban land cover mapping for different spatiotemporal needs. We noticed that high frequency urban land cover analyses (annual or seasonal) relied on Landsat datasets. Although MODIS has the capacity for frequent assessments, its coarse resolution prohibits detailed delineation of settlements and thus was not a popular data source for urban land cover mapping. In contrast, Landsat imagery is one of the data sources mostly used for studies at longer time scales (more than 15 years), along with SPOT imagery and DMSP nighttime light data. A few studies used QuickBird imagery to map long-term urban change but were limited to local scales.

With the improvement of computer hardware in recent years, many urban land cover products are produced using Landsat (e.g., European commission's Global Human Settlement Layer, Global Artificial Impervious Area, Global Annual Urban Dynamics, and World Settlement Footprint). Others also use Sentinel-1 radar information to improve detailed mapping of building boundaries, such as Global Urban Footprint and World Settlement Footprint created by teams from the German Aerospace Center. But these enhanced global products based on Sentinel satellites are limited to a single time point, at least currently, because the first Sentinel satellite was launched in 2014. QuickBird, IKONO, and WorldView imagery are common sources for urban land cover mapping with a high resolution (<5 m). These data do not have global availability regularly because commercial satellite data usually require pre-orders to make sure an image was taken at the selected location and time. As a result, temporal availability of these data is inconsistent at large scales.

The fusion of multi-sensor data helped identification of specific urban conditions (e.g., functional zones), which may not be possible to accurately detect with only a

Table 10.4 Prevalent data and methods used for detecting urban form. The sources and techniques are presented in the order of the number of publications (with brackets). The sum of proportion used exceeds 100% since some papers used multiple sources

Data type	Major sources	Models/techniques	Proportion in the reviewed literature (%)
Urban land cover			
Passive low (>=5 m)	Landsat(130), SPOT(18), MODIS(13), DMSP(12), VIIRS(6), Sentinel-2 (5)	MLC(32), SVM(19), CART(16)	81
Passive high (<5 m)	QuickBird(15), IKONOS(8), WorldView(8)	Object-based(13), MLC(7), SVM(4)	14
Active	TerraSAR-X(8), ALOS-PALSAR (6)	SVM(5), Object-based(3), RF(4)	12
Building density			
Passive low (>=5 m)	Landsat(36), ASTER(7), MODIS(4), Hyperion(3)	LSMA(22), MESMA(11), CART(6)	78
Passive high (<5 m)	IKONOS(3), WorldView(3)	LSMA(3), OBIA(2)	15
Active	TerraSAR-X(3), LiDAR(3)	OBIA(2)	9
3-D structure			
Passive low (>=5 m)	Landsat(4), SPOT(2)	Shadow analysis(3), CART(2)	21
Passive high (<5 m)	WorldView(7), QuickBird(4), IKONOS(2)	Shadow analysis(5), stereo analysis(5), OBIA(3)	45
Active	TerraSAR-X (4), LiDAR (4)	Stereo analysis(4)	32
Street view	Baidu street view(3), Google street view(1)	CNN(2)	11
Land use			
Passive low (>=5 m)	Landsat(18), SPOT(6)	MLC(11), OBIA(5)	39
Passive high (<5 m)	Aerial photographs(14), QuickBird(7), WorldView(7)	OBIA(16), SVM(8), RF(6), CNN(3)	53
Active	LiDAR(6)	OBIA(4), CART(3)	14
Street view	Google street view(3)	CNN(2)	6

(continued)

Table 10.4 (continued)

Data type	Major sources	Models/techniques	Proportion in the reviewed literature (%)
Other urban data	Street map(6), parcel map(6), point of interest(5), elevation(4), cellphone signal(3)	RF(3), CART(3)	27

Note MLC—maximum likelihood classifier; CART—classification tree or regression; RF—random forest; SVM—support vector machine; CNN—convolutional neural network; LSMA—linear spectral mixture analysis; MESMA—multiple endmember spectral mixture analysis; OBIA—object-based image analysis

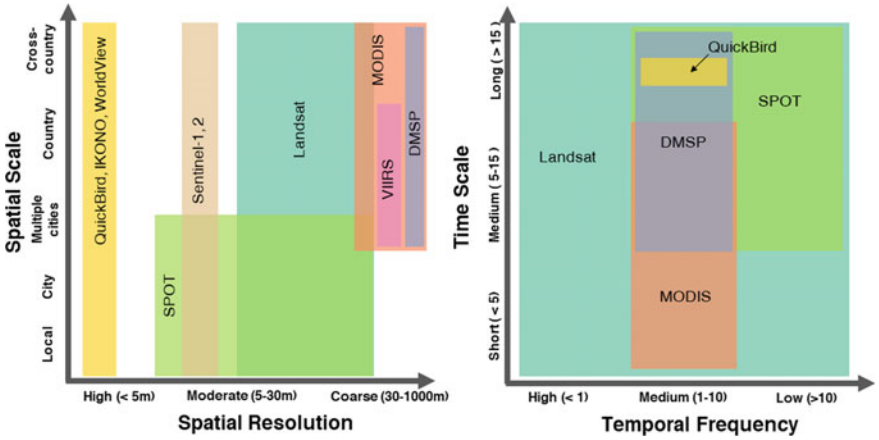


Fig. 10.3 Common remote sensing data used for mapping urban land cover at different spatial scales, resolutions, time scales, and temporal frequencies

single remote sensing product (e.g., optical imagery). Popular approaches for urban form detection included fusion of images from passive and active sensors and images taken in the day and the night. Compared to the spectral signals, SAR signals are more sensitive to roughness and geometric characteristics [88]. Optical and SAR data have been integrated to map urban extent to reduce the confusions between bright impervious surface and bare soil [89], and along fast-developing urban-rural fringes [90]. Recently, features extracted by combining nighttime light and optical imagery have been developed, such as Vegetation Adjusted Nighttime-light Urban Index (VANUI) [91] and the Normalized Difference Spectral Vector (NDSV) [92]. Moreover, with the emerging free service of cloud computing to process planetary-scale Landsat imagery, a method fusing DMSP nighttime light and Landsat imagery has been developed for global urban mapping at 30-m resolution [93].

Mapping building density and 3-D structure

We aggregated the analysis of building density and 3-D structure in this subsection as both of them describe the intensity of built-up areas. A variety of remote sensing data, including images collected by passive and active sensors and at high as well as low resolution, have been used to address urban land use intensity (Fig. 10.4). Spectral mixing analysis, a technique producing fractional estimate of land cover types which can represent building density, usually requires images of two different spatial resolution to provide fractional samples [94–97]. Airborne LiDAR data at very high resolution (<1 m) has been used for modeling the 3-D structure of cities and extracting two-dimensional building footprints [98, 99]. NASA provides one of the publicly available LiDAR data—Geoscience Laser Altimeter System (GLAS) onboard Ice, Cloud, and land Elevation Satellite (ICESat), which can be used to model building height [100]. 3-D morphological variables (e.g., building height and floor area ratio) can also be derived by stereo images that comprise a pair of multi-spectral imagery taken from different angles [101, 102], by SAR images recording the double-bounce reflection of a building [103, 104], by high resolution optical imagery that includes shadows of high-rise buildings [105], or by medium-resolution optical imagery (e.g., Landsat) using spatial configuration of the landscape as predictors [106]. Although optical imagery's spectral features do not direly reflect height information, the classification of high-, mid-, low-rise buildings can be indirectly predicted through the spatial configuration of land uses. The SeaWinds scatterometer onboard the QuikSCAT satellite launched in 1999, was designed to measure the speed and direction of winds that cause ocean waves but also have been used to characterize building volumes and height [107, 108]. Additionally, Fig. 10.4 shows that street view data began to be popular in the last decade for characterizing cities' 3-D space. By using Baidu [109, 110] and Google street view [111], studies have captured street-level sky view factor in high-density urban environments, such as Shanghai, Manhattan Island, and Hong Kong.

It is clear that the selection of data sources affects the spatial resolution and scale of 3-D mapping results (Fig. 10.5). TerraSAR-X and WorldView stereo are few common sources for extracting high-resolution 3-D attributes. While these data are not freely-available, the ICESat-1 and -2 data provided by NASA is free to be used for height extraction at medium-coarse resolution (20–70 m). As the above three datasets only partially cover the land surface, alternative data sources for large-scale 3-D mapping mostly rely on European Space Agency (ESA)'s Sentinel-1and -2 and NASA's Landsat and SeaWinds microwave scatterometer data. SeaWinds microwave scatterometer data covers the Earth fully but has a coarse resolution. For large-scale mapping, however, the ICESat data's partial cover can still serve as validation for 3-D mapping using Sentinel or Landsat data [112].

Few 3-D structure studies consider the dynamics of urban form with frequent mapping as very few free satellite data are useful for this purpose (Fig. 10.5). SeaWinds data have multitemporal information at a one-year interval, but they were only used to estimate building height change at lower temporal frequency (>8 years). Chen et al. used Landsat time-series to predict building height and density at a

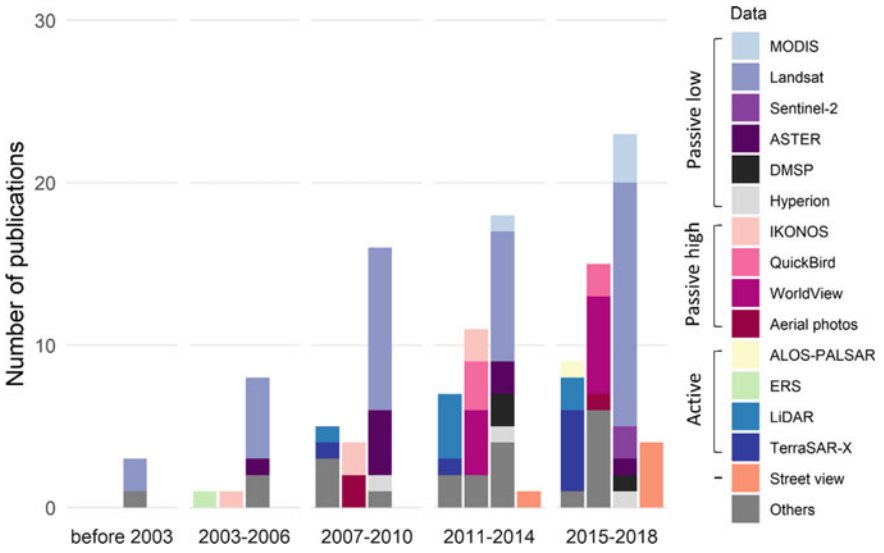


Fig. 10.4 The remote sensing data used, over time, for mapping built-up intensity (i.e., building density or 3-D structure). The data include street view and images collected by active sensors and passive sensors, at low (>=5 m) and high resolution (<5 m)

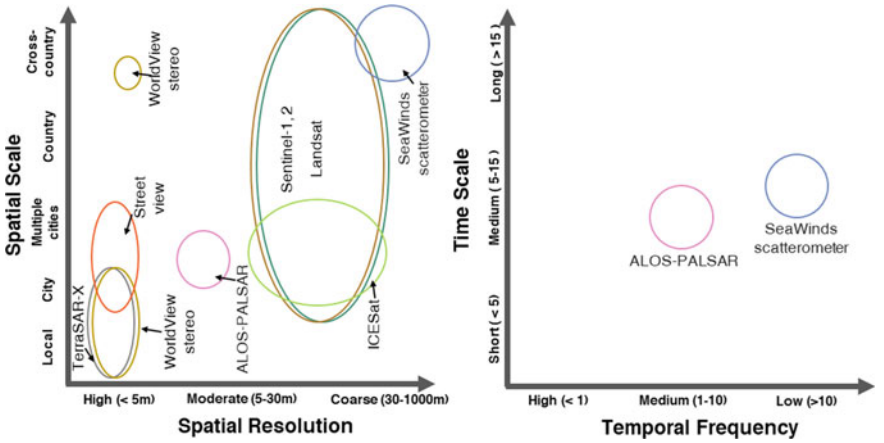


Fig. 10.5 Common remote sensing data used for mapping urban 3-D structure at different spatial scales, resolutions, time scales, and temporal frequencies. Note that the number of the studies analyzing change in 3-D structure shown in the right panel were small (n = 3)

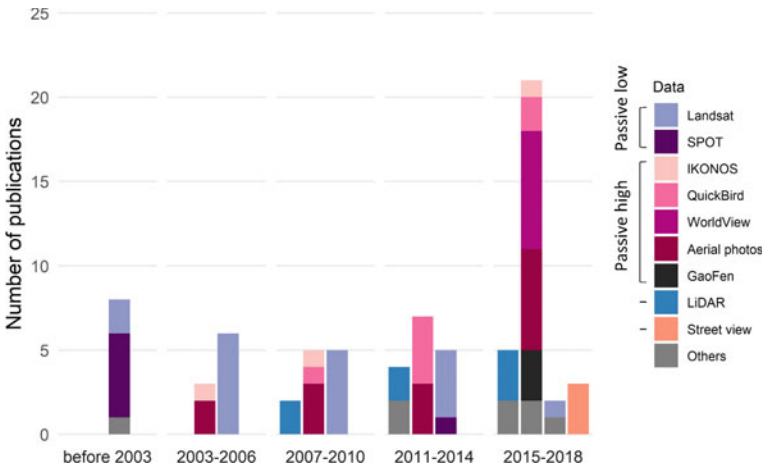


Fig. 10.6 The remote sensing data used, over time, for mapping urban land use, including street view and images collected by active sensors (LiDAR) and passive sensors, at low (>=5 m) and high resolution (<5 m)

one-year interval [106]. While these datasets have potential to be used for high frequency assessments of 3-D structure, temporal accuracy of these assessments remain understudied.

Mapping land use

Studies before 2007 have mostly relied on optical images at resolution lower than 5 m to map urban land use, while studies afterwards shifted to predominantly use high-resolution images (Fig. 10.6). Also, it is challenging to extract nuanced urban land use through a single remote sensing product. Mapping urban land use, such as urban functional zones, places more demands on the quantity of information required. Adding LiDAR elevation [113], street view [114], points of interest [115], or mobile signal density data [116] to passive spectral signals, has helped to distinguish between residential, commercial and other land use.

10.4.3 Methods Used to Detect Urban Form

Common detection methods vary depending on the target attribute of urban form and the properties of imagery data, such as spectral, spatial, and temporal resolution (Table 10.4).

Detecting urban land cover

Supervised machine learning approaches have gained interest recently with the increased accessibility of satellite data and improvement of computation power,

many urban extent studies used maximum likelihood classification (MLC) (e.g., [117]), which is fast and often available on remote sensing software. To overcome the limitation of the normal distribution assumption of MLC that is usually violated in real-world data, support vector machine (SVM), a non-parametric classifier, has demonstrated its ability to incorporate multi-sensor data and achieving high accuracy of mapping [118]. In comparison, other non-parametric classifiers with lower computation cost, such as decision tree [119] and ensemble classifiers (e.g., random forest) [120], have shown their performance in global urban mapping. As the data science and remote sensing communities began to align over the past decade, machine learning using spatial information (e.g., Convolutional Neural Networks) shows significant improvement than models manually incorporate texture indices.

Detecting building density and 3-D structure

To overcome the resolution limitation of freely available satellite imagery, scientists usually incorporated sub-pixel analysis to explore built-up area density. Inspired by the conceptual vegetation-impervious surface-soil (V-I-S) model [121], many studies applied spectral mixture analysis (SMA) by targeting impervious surface mapping at subpixel levels. Several methods for urban mapping have been developed, including linear spectral mixture analysis (LSMA) [122], and multiple endmember spectral mixture analysis (MESMA) that allows various urban compositions [96]. Furthermore, a method combining temporal mixture and the spectral mixture was developed to take advantage of multi-date imagery that was found to successfully reduce noise [123]. Unlike categorical urban extent or land use mapping, predicting urban density usually relied on regression approaches. Besides parametric statistical regression (e.g., linear regression), researchers also utilized machine-learning methods to solve continuous urban density problems (e.g., regression tree [124] and support vector regression [125]).

For 3-D structure mapping, the methods mainly include theoretical and empirical models. Theoretical models rely on the physical mechanism between spectral signal (e.g., SAR) and height to establish a generalizable formula [112], while empirical models (e.g., machine learning) use some real-world examples (i.e., training data) of image features and heights to construct a relationship and therefore are able to predict heights for other areas where heights were unknown [106, 126, 127].

Detecting land use

Deep learning methods using spatial information such as convolutional neural networks (CNNs) are useful to detect contextual urban attributes that are not easy to observe by spectral signal alone. The main advantage of using CNNs over MLC, RF, and NN is that it enables building a hierarchy of local to global features at the spatial dimension [128]. Because of CNN's capabilities in recognizing the spatial patterns of image patches, recent studies applied CNN to street view and aerial photographs for quantifying sky view of street canyons [109], mapping complex urban land use [129, 130] and classifying specific types of building instance (e.g., church and garage) [131].

Object-based image analysis (OBIA) has been largely used for high-resolution urban extent and land-use mapping (Table 10.4). The applications include passive as well as active remote sensing data. The main advantage of object-based analysis in urban form detection is to offer meaningful objects in cities that help to utilize texture features for capturing complex urban classes [132]. Another advantage is to reduce noise (e.g., shadows of buildings, cars on the road) in a high-resolution urban image [133]. Object-based CNN approaches have also been developed to improve irregular urban block mapping [75].

10.4.4 Spatiotemporal Characteristics of the Urban Form Studies

Assessed urban land cover studies—high variety in spatiotemporal scales

Studies that monitored urban form with remote sensing imagery covered all populated continents but unevenly (Fig. 10.7). The studies detecting urban extent were primarily common in East Asia, including Beijing (30), Shenzhen (24), and Guangzhou (23). Similarly, the studies particularly concerned with the temporal aspect (i.e., urban expansion, occupying around 54% of the urban extent studies) were based in East Asia. Among a range of spatial scales of analysis, the majority of the urban extent studies investigated a single city (Fig. 8b). On the other hand, they generally covered higher time depths than other subjects of urban form research. More than a quarter observed urban areas for more than 20 years (Fig. 8a).

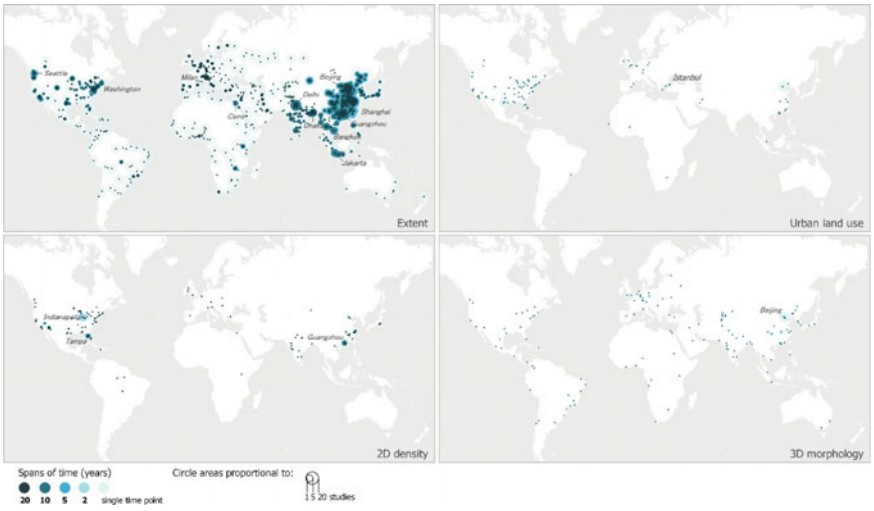


Fig. 10.7 Studied cities, by attributes of urban form, over time

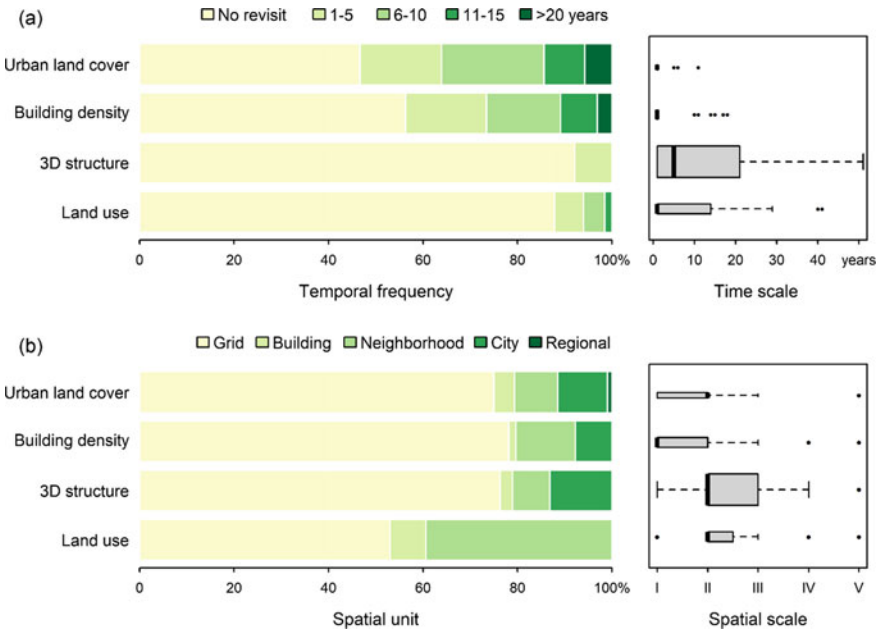


Fig. 10.8 Proportion of papers with analysis related to different **a** temporal frequency and time scale and **b** spatial unit and scale (I = Subset of a city; II = City; III = Multiple cities; IV = Countrywide; V = Cross-country) in the reviewed literature in four categories, urban land cover, building density (2-D), 3-D structure, and land use. The height of the box plots is proportional to the amount of the literature

Studies assessing urban extent cover are composed of various temporal coverages ranging from a single observation to a time series longer than a half century (e.g., [134]). About half of the studies on urban extent had several time steps to trace changes in urban form (Fig. 8a). Most studies on urban expansion monitored changes in urban boundaries every 6–10 years. However, in several rapidly urbanizing regions, changes in the impervious area were detected annually [35] or even seasonally [135].

Assessed studies on building density and 3D structure—primary geographic scopes are neighborhood and individual city

The majority of the papers detecting two-dimensional density investigated only a single city (Fig. 8b). Those cities were mainly located in North America and East Asia, for instance Indianapolis (9) and Guangzhou (9) (Fig. 10.7). Three-dimensional structure studies were primarily conducted on cities in Eurasia (e.g., [136]). Notably, six were in Beijing and three in Hong Kong. These three-dimensional studies often researched within a subset of a city; rarely studies of these interests surveyed the full extent of a city and much less examined multiple cities (Fig. 8b). Only six studies presented between-city comparison to validate their proposed methods (e.g., [136, 137]).

Regarding the time aspect, around 43% of the two-dimensional density studies used multi-temporal observations, and most were every 1–5 years. In contrast, the three-dimensional studies were predominately studied at a single time point (94%). There were only three publications on three-dimensional densification over time. One of them investigated changes of three-dimensional infrastructure through correlating microwave scatterometer data with height across global mega-cities for nearly one decade [107, 108].

Assessed land use studies—least characterization of dynamics and trends

Regarding spatial characteristics, 46% of case studies aggregated pixel-based land-use outcomes to another spatial entity (i.e., block, neighborhood, or an administrative unit) that is more relevant to planning and decision-making. Urban land use was frequently studied in Austin/ Texas (7) and Beijing (6). However, the extent of the study area did not expand as the scale of analysis aggregated. We observed that between-city comparison and between-country comparison were lacking attention in the urban land use literature (Fig. 8b).

While there have been numerous studies on the changes of land cover, the biophysical state of the earth's surface, we found only a few (12%) studies urban land use change, including in Istanbul (2) (e.g., [138]). These studies usually adopted bi-temporal detection over a short period.

10.5 Discussion: Research Gaps and the Way Forward

10.5.1 Dilemma of the Spatial and Temporal Needs

Availability of data has always shaped our scientific questions. Here we have provided an overview of urban form detection using remote sensing, going back 30 years. The spatial resolution and temporal depth of remote sensing data present urban form studies with a dilemma: newer, high-resolution imagery, able to detect nuanced urban attributes, has a relatively short temporal span. Following this, three research gaps have emerged: (1) characterizing long-term urban dynamics, (2) synchronizing within-city variation with between-city diversity, and (3) comparative studies between long-standing research hotspots and underexplored regions.

First, the temporal aspect of urban dynamics remains underexplored. Especially in studies concerning more nuanced features, including three-dimensional structure and urban land use, time depth was usually limited. Although there are around 1-km resolution urban luminosity data observed from the DMSP nighttime satellite since 1992 (e.g., [139]) and 30-m resolution optical imagery from Landsat satellites for more than three decades, these data were mostly used to distinguish urban dichotomy and to measure density by viewing cities as flat surfaces. Extracting three-dimensional structure or activity-based attributes requires high-resolution imagery or other urban information covered from crowdsourced data, for instance, from cellphone signals

or Open Street Map [140]. Those data are usually not retrospectively available and thus prohibit assessing urban form transitions.

Second, the mapping of fine-scale attributes lacks between-city comparison. High-resolution imagery has supported spatially explicit mapping and characterizing within-city variations, such as the distribution of building footprints and different types of land use. However, it is uncertain whether such methods are transferable to other types of cities and whether the empirical findings can be generalized. For instance, is a method detecting fast construction in the coastal regions in China applicable to the small-scale urban regeneration in the Nordic countries? The cost of high-resolution imagery undermines comparisons of methods in diverse cities. Up to now, sensors with freely accessible data were primarily designed to monitor vegetation, but not detailed urban form. The urban landscape is typically composed of features that are smaller than the spatial resolution of commonly utilized data from USGS 30-m Landsat and the up to 10-m ESA Copernicus Sentinel 1,2 missions, namely, a mosaic picture of roads, garden, water, and a wide variety of buildings. Previous studies approach this challenge by using subpixel analysis to explore within-pixel composition [96, 122]. However, to our understanding, the subpixel approaches for urban mapping are mainly used to separate impervious areas from vegetation and soil following the V-I-S model proposed by [121], but tell very little, for instance, about the classification of the built environment features and urban functional zones.

Last, the urban form of fast urbanizing regions has been unevenly explored. For instance, although there were many studies in East Asia, very few were in Africa, which is also experiencing fast urbanization. The research needs to address urban form in a comparable manner, because the quality and performance of training data, the parameters of segmentation and classifiers, and the resulting map accuracy, all depend on the location [93, 141]. The input data, features, and models can better perform in some places of the world, but not in others. However, 39% of studies detecting three-dimensional structure, and 34% detecting two-dimensional building density, were only validated in cities in East Asia. Another 43% of methods for mapping two-dimensional density and 45% for detecting urban land use were developed based on the North American context. Comparative studies that adopt similar models and data on varying types of urban environments across geographies are urgently needed to inform sustainable solutions and to open communication internationally.

10.5.2 Strategies Forward

We foresee that three trends of remote sensing research can deal with the dilemma, providing more comprehensive coverage over spatial and temporal scales: (1) integrating multiple data, (2) enhancing spatial and temporal transferability, (3) shifting from pixels to other meaningful analytic scales, and (4) applying advanced computer vision technologies. Improving the scientific approaches of urban form detection

brings *two critical advancements* in the understanding of sustainable urban development for human health. *First,* causality between public health and the built environment can be better understood when dynamic exposure assessment is built upon a long-term time frame. Researchers would be interested in exploring how, for example, local scale urban form, such as census block architecture, height and density, detected and classified remotely, could be linked to individuals' wellbeing acquired using personal sensors. With long-term urban form measurement, researchers can further explore how change in the built environment are related to change in health outcome. *Second*, transferable methods of urban form detection can help compare urban form features across multiple cities and deepen our understanding of healthy city designs for underexplored regions.

Combinations of street view data and survey/field data, are emerging approaches that help to understand people's living experience, including tree exposure [142], the living quality of neighborhoods [66]. In addition, combinations of images from satellites of a similar design can increase the period of observations. Image fusion using machine learning approaches can predict a high-resolution image by using a low-resolution image as a predictor [143]. We foresee emerging approaches fusing newly available Sentinel-2 at 20-m resolution with the long-standing sensors from the Landsat satellite program back to 1972. Long-term high-resolution information facilitates a deeper understanding of the urbanizing process and causality between urbanization and the environment–whether urbanization drives environmental change or the other way around. On a regional scale, for instance, how long-term patterns of natural hazards (e.g., landslide [144]) and human settlement patterns interacts can be investigated. Other questions at local scales, such as how microclimate, urban land use, and peoples' recreational activities reshape each other, are also crucial for urban planning.

We have seen a progressive development in the spatial transferability of urban form mapping, as models developed using reference data collected in a limited number of cities have been validated by data for other geographic regions [127, 145, 146]. Transferable models can allow for investigation in underexplored regions such as the Global South. These regions are often lack urban form training data, which has made it difficult to map urban form over a large area or a long period. With a transferable model, it would be possible to predict the urban form of a data-deficient region using a model trained on a data-rich region. At the meanwhile, we need to enhance the robustness of our methods in terms of temporal transferability. As collecting reference data with temporal profiles is tremendously labor-intensive, temporal transferability needs to be addressed, namely: Can models developed using data collected at one point in time be applied to map urban attributes at another point in time? Stemming from this core question, two others come up: Can models using training data limited in temporal density accurately predict turning point of urban growth? And, can models trained by data limited for recent years retrospectively map urban form in the previous years? We expect methods for mapping urban form dynamics to mature when strategies for temporal sampling become better developed and commonly applied.

For the spatial dimension, fused images with enhanced spatial resolution can better reflect urban complexity. The distribution of public services and the inner-city variations of building density can be detected at a higher accuracy with fine-resolution imagery. Another way is to enrich urban-relevant information by integrating different data, from active (e.g., SAR) and passive (e.g., optical) sensors, or from satellites and non-satellite sources (e.g., social media). To date, the majority of image fusion approaches developed to enhance the spatial resolution of freely accessible earth observation data are designed for ecological studies, urban sciences have yet to contribute fully.

Wisely selecting a spatial unit for analysis can be another solution to understand urban complexity. Instead of making predictions from pixels, researchers could label, train, and classify features at block, patch, or neighborhood unites. Such unites make it easier to predict three-dimensional features of urban form and human activities because spatial relations and inter-pixel texture are useful predictors. Object-based approaches utilize spatial information besides spectral information, and deep learning techniques enable automated spatial feature extraction [106, 147]. Recently, more advanced deep learning based semantic segmentation approaches, such as fully convolutional networks, UNet, and DeepLab, are gaining popularity in the remote sensing field [148–150]. Unlike traditional CNN-based scene classifications, these semantic segmentation algorithms allow localizing predictions while using the most of contextual information, thus help to detect local urban objects [151, 152]. Object-based deep learning methods, such as mask R-CNN is suitable for extracting building boundaries in complex urban landscapes [153].

By capturing urban complexity, scientific questions can move forward from where urban areas have expanded to more nuanced features, such as how housing and architecture have transformed. Previous literature reviews provided the foundation of the concepts and empirical findings on sustainable urban form. References [15, 19] have summarized the urban form attributes for several sustainable environmental goals. Reviews on resilience, with more focus on disaster risk reduction and adaptation, have addressed resilient urban attributes from a city scale [154] to a neighborhood scale [155]. The strategies we propose can support debates closer to the scale of people's daily life. To join the debate of classical urban planning theories such as 'compact city', it is important to discuss how an individual's experience deviates from a city-scale point of view. This way, we measure how different groups of people can be exposed to different types of neighborhood (e.g., building height, accessibility to open space). The measurement of individual experience can help answer factors of human well-being, such as how individual experience of urban form relates to their commuting behavior and their cardiovascular and stress-related diseases.

Lastly, improvements in computation power and the availability of varying satellite datasets, via Google Earth Engine for instance [146], opens opportunities to study urban transitions in greater spatial and temporal details. The improvement in hard technology also makes deep learning approaches, such as convolutional neural networks, more efficient. The raised capacities in computation bring our questions to large-scale and long-term observations of the Anthropocene. We are on the cusp of exploring impacts of urbanization on human society from the perspective of space.

10.6 Conclusions

This review synthesizes urban form studies on mapping urban land cover, building density, three-dimensional structure, and land use, from a remote sensing perspective. Spatially explicit urban information can support scholars to investigate urban form dynamics and their impacts on human health and well-being. We found it integral to move beyond the global research hotspots, namely East Asia and North America, particularly in areas where urban form has different urban traditions, and thus may have different spatiotemporal features, such as the types of buildings, the spatial configuration of urban land use and transformation paths, particularly in Africa, Europe, and South America. Additionally, our review reveals that there has been progress in urban extent mapping, with much less research focus on "the contexts of the urban environment" in terms of three-dimensional structures and the patterns of varying urban activities. The gap is further enlarged when it comes to a consideration of high frequency and long-term mapping of 3-D structure.

Our review culminates posing a question, whether commonly used remote sensing data are suited for detailed urban form mapping. We found that remote sensing science has a dilemma: newer, high-resolution imagery, able to detect nuanced urban attributes, has a relatively short temporal span. In contrast, older imagery can detect long-term change, but has a lower resolution. Until recently, Landsat imagery was the primary workhorse of urban form studies, but with a limited capability to detect urban form detail due to its resolution. Multisource data fusion and deep learning could be the solutions for urban form research. As big geo-data and artificial intelligence methods become available, urban form studies can better handle the spatiotemporal dilemma, such as the long-term mapping of urban form in two- and three-dimensions, and transferable urban form detection for the Global South. We foresee the possibility of a change in paradigm: instead of formulating research questions by what we can see from pixels, urban scientists can deal with what the needs of people are. For instance, how individuals experience urban redevelopment (e.g., from town housing to high rises) and how does this contribute to their chronic diseases that appear at their later life stages, and how hazard patterns change over time and how does it interact with urbanization. The urban remote sensing community should bring its expertise to collaborate with other urban scholars in health, hazard, climate, and social sciences, and thereby to engage with the sustainable urban development agenda, by upscaling urban form studies in space and time.

Appendix

See Tables 10.5 and 10.6

.

Table 10.5 Full query to search relevant literature on Scopus

#	Attribute	Query
1	Urban land cover	(KEY(“machine learning” OR “classification” OR “deep learning” OR “image analysis” OR “mapping” OR “object-based image analysis” OR “extraction” OR “prediction” OR “detection”) AND TITLE-ABS-KEY(“Satellite data” OR “remote sensing” OR “images” OR “imagery” OR LiDAR OR SAR OR “aerial photos” OR “street view”)AND (TITLE-ABS-KEY (“impervious surface” OR “urban extent” OR “built-up area”)) AND TITLE-ABS-KEY(urban OR city)AND PUBYEAR > 1989 AND PUBYEAR < 2019 AND DOCTYPE(ar)) AND (LIMIT-TO (LANGUAGE,“ English”))
2	Two-dimensional building density	(KEY(“machine learning” OR “classification” OR “deep learning” OR “image analysis” OR “mapping” OR “object-based image analysis” OR “extraction” OR “prediction” OR “detection”) AND TITLE-ABS-KEY(“Satellite data” OR “remote sensing” OR “images” OR “imagery” OR LiDAR OR SAR OR “aerial photos” OR “street view”)AND (TITLE-ABS-KEY (building density OR “urban density” OR “building cover ratio”) OR TITLE-ABS-KEY(“impervious surface” AND “spectral mixture analysis”)) AND TITLE-ABS-KEY(urban OR city)AND PUBYEAR > 1989 AND PUBYEAR < 2019 AND DOCTYPE(ar)) AND (LIMIT-TO (LANGUAGE,“English”))
2	Three-dimensional structure	(KEY(“machine learning” OR “classification” OR “deep learning” OR “image analysis” OR “mapping” OR “object-based image analysis” OR “extraction” OR “prediction” OR “detection”) AND TITLE-ABS-KEY(“Satellite data” OR “remote sensing” OR “images” OR “imagery” OR LiDAR OR SAR OR “aerial photos” OR “street view”) AND TITLE-ABS-KEY(“building height” OR “height of building” OR “building volume” OR “sky view factor” OR “urban structure” OR “urban morphology” OR “roughness length” OR “urban porosity”) AND TITLE-ABS-KEY(urban OR city)AND PUBYEAR > 1989 AND PUBYEAR < 2019 AND DOCTYPE(ar)) AND (LIMIT-TO (LANGUAGE, “English”))

(continued)

Table 10.5 (continued)

#	Attribute	Query
3	Land use	(KEY(“machine learning” OR “classification” OR “deep learning” OR “image analysis” OR “mapping” OR “object-based image analysis” OR “extraction” OR “prediction” OR “detection”) AND TITLE-ABS-KEY(“Satellite data” OR “remote sensing” OR “images” OR “imagery” OR LiDAR OR SAR OR “aerial photos” OR “street view”)AND TITLE-ABS-KEY(urban AND (“land use mapping” OR “land use classification” OR “land use extraction” OR “functional zone”) OR (“urban land use” AND change) OR (“urban landuse” AND change)) AND TITLE-ABS-KEY(urban OR city)AND PUBYEAR > 1989 AND PUBYEAR < 2019 AND DOCTYPE(ar)) AND (LIMIT-TO (LANGUAGE, “English”))

Table 10.6 Checklist of general items (ID #1 to #11) and spatiotemporal characteristics (ID #12 to #24) used when constructing the review database for urban form detection

#	Fields	Definition	Type	Categories
1	Study ID	ID of reviewed study	Numeric	
2	Title	Title of the article	Text	
3	Year	Year of publication	Text	
4	Authors	Author(s)	Text	
5	Research institute	Name and place of research institutes	Text	
6	Source title	Journal name	Text	
7	Citations	No. of citations	Numeric	
8	Category	Category of urban form	Classes	Urban land cover; 2-D building density; 3-D structure; land use
9	Model	The machine learning or image classification method		Threshold-based; MLC; SVM; decision tree; RF; CNN; OBIA
10	Data type	Name of data type(s)	Classes	Multispectral imagery; Hyperspectral imagery; SAR; LiDAR; Nighttime light; DEM; Street view; Others
11	Data	Name of sensor(s) or non-satellite data	Text	
12	Geotag	The original name of the study area	Text	
13	Study city	Cities where the study area is located	Text	

(continued)

Table 10.6 (continued)

#	Fields	Definition	Type	Categories
14	Study country	Countries where the study area is located	Text	
15	Resolution	The spatial resolution of the data	Numeric	Unit: m
16	Passive low	Whether low resolution (>=5 m) data from passive sensor was used	Binary	0;1
17	Passive high	Whether high resolution (<5 m) data from passive sensor was used	Binary	0;1
18	Active	Whether data from active sensor was used	Binary	0;1
19	Street view	Whether street view data was used	Binary	0;1
20	Other data	Whether other ancillary data was used	Binary	0;1
21	Spatial unit	Spatial unit of single observations	Classes	Grid; Building; Neighborhood/district/patch; City; Regional
22	Spatial scale	Total area of observation	Numeric	Subset of a city; City; Multiple cities; Countrywide; Multiple countries
23	Temporal frequency	Density of multi-temporal observations	Classes	No revisit; 1–5 years; 6–10 years; 11–15 years; >15 years
24	Time scale	Length of time between the first and the last observation	Numeric	Unit: years

References

1. United Nations (2018) World urbanization prospects 2018. United Nations Department for Economic and Social Affiars
2. Huang K et al (2019) Projecting global urban land expansion and heat island intensification through 2050. Environ Res Lett 14(11):114037
3. Huang K et al (2021) Persistent increases in nighttime heat stress from urban expansion

despite heat island mitigation. J Geophys Res Atmos, p e2020JD033831
4. Stone B, Hess JJ, Frumkin H (2010) Urban form and extreme heat events: are sprawling cities more vulnerable to climate change than compact cities. Environ Health Perspect 118(10):1425–1428
5. Bereitschaft B, Debbage K (2013) Urban form, air pollution, and CO_2 emissions in large US metropolitan areas. Prof Geogr 65(4):612–635
6. d'Amour CB et al (2017) Future urban land expansion and implications for global croplands. Proc Natl Acad Sci 114(34):8939–8944
7. d'Amour CB et al (2020) Urbanization, processed foods, and eating out in India. Glob Food Secur 25:100361
8. Frank LD, Wali B (2021) Treating two pandemics for the price of one: Chronic and infectious disease impacts of the built and natural environment. Sustain Cities Soc 73:103089
9. Chen T-HK, Chen VY-J, Wen T-H (2018) Revisiting the role of rainfall variability and its interactive effects with the built environment in urban dengue outbreaks. Appl Geogr 101:14–22
10. Seto KC, Güneralp B, Hutyra LR (2012) Global forecasts of urban expansion to 2030 and direct impacts on biodiversity and carbon pools. Proc Natl Acad Sci 109(40):16083–16088
11. Samuelsson K et al (2018) Impact of environment on people's everyday experiences in Stockholm. Landsc Urban Plan 171:7–17
12. Evans GW (2003) The built environment and mental health. J Urban Health 80(4):536–555
13. Kyttä M et al (2016) Urban happiness: context-sensitive study of the social sustainability of urban settings. Environ Plan B Plan Des 43(1):34–57
14. Sallis JF et al (2016) Physical activity in relation to urban environments in 14 cities worldwide: a cross-sectional study. Lancet 387(10034):2207–2217
15. Stokes EC, Seto KC (2019) Characterizing and measuring urban landscapes for sustainability. Environ Res Lett 14(4):045002
16. Güneralp B et al (2017) Global scenarios of urban density and its impacts on building energy use through 2050. Proc Natl Acad Sci, 201606035
17. Mouratidis K (2018) Built environment and social well-being: How does urban form affect social life and personal relationships? Cities 74:7–20
18. Sallis JF et al (2012) Role of built environments in physical activity, obesity, and cardiovascular disease. Circulation 125(5):729–737
19. Gren Å et al (2018) How smart is smart growth? Examining the environmental validation behind city compaction. Ambio, 1–10
20. Ko Y, Radke JD (2014) The effect of urban form and residential cooling energy use in Sacramento, California. Environ Plan B Plan Des 41(4):573–593
21. Lariviere I, Lafrance G (1999) Modelling the electricity consumption of cities: effect of urban density. Energy Econ 21(1):53–66
22. Lee S, Lee B (2014) The influence of urban form on GHG emissions in the US household sector. Energy Policy 68:534–549
23. Seto KC et al (2014) Human settlements, infrastructure and spatial planning
24. Holden E, Norland IT (2005) Three challenges for the compact city as a sustainable urban form: household consumption of energy and transport in eight residential areas in the greater Oslo region. Urban Stud 42(12):2145–2166
25. Holden E, Linnerud K (2011) Troublesome leisure travel: the contradictions of three sustainable transport policies. Urban Stud 48(14):3087–3106
26. Stone B (2008) Urban sprawl and air quality in large US cities. J Environ Manag 86(4):688–698
27. Dave S (2010) High urban densities in developing countries: a sustainable solution? Built Environ 36(1):9–27
28. Rodríguez MC, Dupont-Courtade L, Oueslati W (2016) Air pollution and urban structure linkages: evidence from European cities. Renew Sustain Energy Rev 53:1–9
29. Clark LP, Millet DB, Marshall JD (2011) Air quality and urban form in US urban areas: evidence from regulatory monitors. Environ Sci Technol 45(16):7028–7035

30. Yeh C et al (2020) Using publicly available satellite imagery and deep learning to understand economic well-being in Africa. Nat Commun 11(1):1–11
31. Sharma R, Joshi P (2013) Monitoring urban landscape dynamics over Delhi (India) using remote sensing (1998–2011) inputs. J Indian Soc Remote Sens 41(3):641–650
32. Li X et al (2018) Mapping annual urban dynamics (1985–2015) using time series of Landsat data. Remote Sens Environ 216:674–683
33. Sexton JO et al (2013) Urban growth of the Washington, DC–Baltimore, MD metropolitan region from 1984 to 2010 by annual, Landsat-based estimates of impervious cover. Remote Sens Environ 129:42–53
34. Bagan H, Yamagata Y (2012) Landsat analysis of urban growth: how Tokyo became the world's largest megacity during the last 40 years. Remote Sens Environ 127:210–222
35. Shi L et al (2017) Impervious surface change mapping with an uncertainty-based spatial-temporal consistency model: a case study in Wuhan city using Landsat time-series datasets from 1987 to 2016. Remote Sens 9(11):1148
36. Liu X et al (2020) High-spatiotemporal-resolution mapping of global urban change from 1985 to 2015. Nat Sustain 3(7):564–570
37. Gong P et al (2020) Annual maps of global artificial impervious area (GAIA) between 1985 and 2018. Remote Sens Environ 236:111510
38. Corbane C et al (2019) Automated global delineation of human settlements from 40 years of Landsat satellite data archives. Big Earth Data 3(2):140–169
39. Marconcini M et al (2020) Outlining where humans live, the World Settlement Footprint 2015. Sci Data 7(1):1–14
40. Aburas MM et al (2016) The simulation and prediction of spatio-temporal urban growth trends using cellular automata models: a review. Int J Appl Earth Obs Geoinf 52:380–389
41. Zhang W et al (2018) Analyzing horizontal and vertical urban expansions in three East Asian megacities with the SS-coMCRF model. Landsc Urban Plan 177:114–127
42. Song X-P et al (2016) Characterizing the magnitude, timing and duration of urban growth from time series of Landsat-based estimates of impervious cover. Remote Sens Environ 175:1–13
43. Roback K et al (2018) The size, distribution, and mobility of landslides caused by the 2015 Mw7. 8 Gorkha earthquake, Nepal. Geomorphology 301:121–138
44. Sekizawa R, Ichii K, Kondo M (2015) Satellite-based detection of evacuation-induced land cover changes following the Fukushima Daiichi nuclear disaster. Remote Sens Lett 6(11):824–833
45. Weng Q (2012) Remote sensing of impervious surfaces in the urban areas: requirements, methods, and trends. Remote Sens Environ 117:34–49
46. Bhatta B, Saraswati S, Bandyopadhyay D (2010) Urban sprawl measurement from remote sensing data. Appl Geogr 30(4):731–740
47. Zhu Z et al (2019) Understanding an urbanizing planet: strategic directions for remote sensing. Remote Sens Environ 228:164–182
48. Reba M, Seto KC (2020) A systematic review and assessment of algorithms to detect, characterize, and monitor urban land change. Remote Sens Environ 242:111739
49. Seto KC, Ramankutty N (2016) Hidden linkages between urbanization and food systems. Science 352(6288):943–945
50. Salerno F, Gaetano V, Gianni T (2018) Urbanization and climate change impacts on surface water quality: enhancing the resilience by reducing impervious surfaces. Water Res 144:491–502
51. Cui Y et al (2019) The cost of rapid and haphazard urbanization: lessons learned from the Freetown landslide disaster. Landslides 16(6):1167–1176
52. El Garouani A et al (2017) Analysis of urban growth and sprawl from remote sensing data: case of Fez, Morocco. Int J Sustain Built Environ 6(1):160–169
53. Jat MK, Garg PK, Khare D (2008) Monitoring and modelling of urban sprawl using remote sensing and GIS techniques. Int J Appl Earth Obs Geoinf 10(1):26–43
54. Homer C et al (2004) Development of a 2001 national land-cover database for the United States. Photogramm Eng Remote Sens 70(7):829–840

55. Salvati L et al (2012) Low-density settlements and land use changes in a Mediterranean urban region. Landsc Urban Plan 105(1–2):43–52
56. McGarigal K et al (2002) FRAGSTATS: spatial pattern analysis program for categorical maps
57. Johnson MT, Munshi-South J (2017) Evolution of life in urban environments. Science 358(6363):eaam8327
58. Fahrig L et al (2011) Functional landscape heterogeneity and animal biodiversity in agricultural landscapes. Ecol Lett 14(2):101–112
59. Liu Y et al (2018) The relationship between urban form and air pollution depends on seasonality and city size. Environ Sci Pollut Res 25(16):15554–15567
60. Knowles RL (1974) Energy and form: an ecological approach to urban growth
61. Xu Y et al (2017) Urban morphology detection and computation for urban climate research. Landsc Urban Plan 167:212–224
62. Stewart ID, Oke TR (2012) Local climate zones for urban temperature studies. Bull Am Meteorol Soc 93(12):1879–1900
63. Unger J, Lelovics E, Gál T (2014) Local Climate Zone mapping using GIS methods in Szeged. Hung Geogr Bull 63(1):29–41
64. Adolphe L (2001) A simplified model of urban morphology: application to an analysis of the environmental performance of cities. Environ Plan B Plan Des 28(2):183–200
65. Lau SSY et al (2011) The study of summer-time heat island, built form and fabric in a densely built urban environment in compact Chinese cities: Hong Kong, Guangzhou. Int J Sustain Dev 14(1–2):30–48
66. Naik N et al (2017) Computer vision uncovers predictors of physical urban change. Proc Natl Acad Sci 114(29):7571–7576
67. Kotharkar R, Bagade A (2018) Local Climate Zone classification for Indian cities: a case study of Nagpur. Urban Clim 24:369–392
68. Lelovics E et al (2014) Design of an urban monitoring network based on Local Climate Zone mapping and temperature pattern modelling. Clim Res 60(1):51–62
69. Lin J-J, Yang A-T (2006) Does the compact-city paradigm foster sustainability? An empirical study in Taiwan. Environ Plan B Plan Des 33(3):365–380
70. Tian Y, Jim C, Wang H (2014) Assessing the landscape and ecological quality of urban green spaces in a compact city. Landsc Urban Plan 121:97–108
71. Ou J et al (2013) Quantifying the relationship between urban forms and carbon emissions using panel data analysis. Landsc Ecol 28(10):1889–1907
72. Schindler S, Poirazidis K, Wrbka T (2008) Towards a core set of landscape metrics for biodiversity assessments: a case study from Dadia National Park, Greece. Ecol Indic 8(5):502–514
73. Giridharan R, Ganesan S, Lau S (2004) Daytime urban heat island effect in high-rise and high-density residential developments in Hong Kong. Energy Build 36(6):525–534
74. Gál T, Unger J (2009) Detection of ventilation paths using high-resolution roughness parameter mapping in a large urban area. Build Environ 44(1):198–206
75. Zhang C et al (2018) An object-based convolutional neural network (OCNN) for urban land use classification. Remote Sens Environ 216:57–70
76. Anderson JR (1976) A land use and land cover classification system for use with remote sensor data, vol 964. US Government Printing Office
77. Wentz EA et al (2014) Supporting global environmental change research: a review of trends and knowledge gaps in urban remote sensing. Remote Sens 6(5):3879–3905
78. Heinzel J, Kemper T (2015) Automated metric characterization of urban structure using building decomposition from very high resolution imagery. Int J Appl Earth Obs Geoinf 35:151–160
79. Novack T, Kux H (2010) Urban land cover and land use classification of an informal settlement area using the open-source knowledge-based system InterIMAGE. Health Risk Soc 55(1):23–41
80. Zhang X, Du S, Wang Q (2018) Integrating bottom-up classification and top-down feedback for improving urban land-cover and functional-zone mapping. Remote Sens Environ 212:231–248

81. Moher D et al (2015) Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev 4(1):1
82. Xiao R-B et al (2007) Spatial pattern of impervious surfaces and their impacts on land surface temperature in Beijing, China. J Environ Sci 19(2):250–256
83. Zha Y, Gao J, Ni S (2003) Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. Int J Remote Sens 24(3):583–594
84. Zhou Y et al (2015) A global map of urban extent from nightlights. Environ Res Lett 10(5):054011
85. Pesaresi M et al (2016) Assessment of the added-value of sentinel-2 for detecting built-up areas. Remote Sens 8(4):299
86. Shi K et al (2014) Evaluation of NPP-VIIRS night-time light composite data for extracting built-up urban areas. Remote Sens Lett 5(4):358–366
87. Schneider A, Friedl MA, Potere D (2009) A new map of global urban extent from MODIS satellite data. Environ Res Lett 4(4):044003
88. Salentinig A, Gamba P (2015) Combining SAR-based and multispectral-based extractions to map urban areas at multiple spatial resolutions. IEEE Geosci Remote Sens Mag 3(3):100–112
89. Zhang Y, Zhang H, Lin H (2014) Improving the impervious surface estimation with combined use of optical and SAR remote sensing images. Remote Sens Environ 141:155–167
90. Qin Y et al (2017) Quantifying annual changes in built-up area in complex urban-rural landscapes from analyses of PALSAR and Landsat images. ISPRS J Photogramm Remote Sens 124:89–105
91. Jing W et al (2015) Mapping urban areas with integration of DMSP/OLS nighttime light and MODIS data using machine learning techniques. Remote Sens 7(9):12419–12439
92. Trianni G et al (2015) Scaling up to national/regional urban extent mapping using Landsat data. IEEE J Sel Top Appl Earth Obs Remote Sens 8(7):3710–3719
93. Goldblatt R et al (2018) Using Landsat and nighttime lights for supervised pixel-based image classification of urban land cover. Remote Sens Environ 205:253–275
94. Lu D, Weng Q (2009) Extraction of urban impervious surfaces from an IKONOS image. Int J Remote Sens 30(5):1297–1311
95. Michishita R, Jiang Z, Xu B (2012) Monitoring two decades of urbanization in the Poyang Lake area, China through spectral unmixing. Remote Sens Environ 117:3–18
96. Powell RL et al (2007) Sub-pixel mapping of urban land cover using multiple endmember spectral mixture analysis: Manaus, Brazil. Remote Sens Environ 106(2):253–267
97. Zhang Y et al (2017) Analyzing the impacts of urbanization and seasonal variation on land surface temperature based on subpixel fractional covers using Landsat images. IEEE J Sel Top Appl Earth Obs Remote Sens 10(4):1344–1356
98. Elghazali E (2011) Performance of Quickbird image and Lidar data fusion for 2D/3D city mapping. Aust J Basic Appl Sci 5(11):1588–1600
99. Hung C-LJ, James LA, Hodgson ME (2018) An automated algorithm for mapping building impervious areas from airborne LiDAR point-cloud data for flood hydrology. GISci Remote Sens 55(6):793–816
100. Gong P et al (2010) ICEsat GLAS data for urban environment monitoring. IEEE Trans Geosci Remote Sens 49(3):1158–1172
101. Duan G et al (2018) Establishment of an improved floor area ratio with high-resolution satellite imagery. J Indian Soc Remote Sens 46(2):275–286
102. Peng F et al (2017) A new stereo pair disparity index (SPDI) for detecting built-up areas from high-resolution stereo imagery. Remote Sens 9(6):633
103. Brunner D et al (2010) Building height retrieval from VHR SAR imagery based on an iterative simulation and matching technique. IEEE Trans Geosci Remote Sens 48(3):1487–1504
104. Sauer S et al (2009) Polarimetric dual-baseline InSAR building height estimation at L-band. IEEE Geosci Remote Sens Lett 6(3):408–412
105. Shao Y, Taff GN, Walsh SJ (2011) Shadow detection and building-height estimation using IKONOS data. Int J Remote Sens 32(22):6929–6944

106. Chen T-HK et al (2020) Mapping horizontal and vertical urban densification in Denmark with Landsat time-series from 1985 to 2018: a semantic segmentation solution. Remote Sens Environ 251:112096
107. Frolking S et al (2013) A global fingerprint of macro-scale changes in urban structure from 1999 to 2009. Environ Res Lett 8(2):024004
108. Mahtta R, Mahendra A, Seto KC (2019) Building up or spreading out? Typologies of urban growth across 478 cities of 1 million+. Environ Res Lett 14(12):124077
109. Gong F-Y et al (2018) Mapping sky, tree, and building view factors of street canyons in a high-density urban environment. Build Environ 134:155–167
110. Zeng L et al (2018) A fast approach for large-scale Sky View Factor estimation using street view images. Build Environ 135:74–84
111. Liang J et al (2017) Automatic sky view factor estimation from street view photographs—a big data approach. Remote Sens 9(5):411
112. Li X et al (2020) Developing a method to estimate building height from Sentinel-1 data. Remote Sens Environ 240:111705
113. Hermosilla T et al (2012) Assessing contextual descriptive features for plot-based classification of urban areas. Landsc Urban Plan 106(1):124–137
114. Zhang W et al (2017) Parcel-based urban land use classification in megacity using airborne LiDAR, high resolution orthoimagery, and Google Street View. Comput Environ Urban Syst 64:215–228
115. Song J et al (2018) Mapping urban functional zones by integrating very high spatial resolution remote sensing imagery and points of interest: a case study of xiamen, China. Remote Sens 10(11):1737
116. Jia Y et al (2018) Urban land use mapping by combining remote sensing imagery and mobile phone positioning data. Remote Sens 10(3):446
117. Dewan AM, Yamaguchi Y (2009) Land use and land cover change in Greater Dhaka, Bangladesh: using remote sensing to promote sustainable urbanization. Appl Geogr 29(3):390–401
118. Cao X et al (2009) A SVM-based method to extract urban areas from DMSP-OLS and SPOT VGT data. Remote Sens Environ 113(10):2205–2209
119. Schneider A, Friedl MA, Potere D (2010) Mapping global urban areas using MODIS 500-m data: new methods and datasets based on 'urban ecoregions.' Remote Sens Environ 114(8):1733–1746
120. Esch T et al (2012) TanDEM-X mission-new perspectives for the inventory and monitoring of global settlement patterns. J Appl Remote Sens 6(1):061702
121. Ridd MK (1995) Exploring a VIS (vegetation-impervious surface-soil) model for urban ecosystem analysis through remote sensing: comparative anatomy for cities. Int J Remote Sens 16(12):2165–2185
122. Lu D, Weng Q (2006) Spectral mixture analysis of ASTER images for examining the relationship between urban thermal features and biophysical descriptors in Indianapolis, Indiana, USA. Remote Sens Environ 104(2):157–167
123. Li W, Wu C (2015) Incorporating land use land cover probability information into endmember class selections for temporal mixture analysis. ISPRS J Photogramm Remote Sens 101:163–173
124. Deng C, Wu C (2013) The use of single-date MODIS imagery for estimating large-scale urban impervious surface fraction with spectral mixture analysis and machine learning techniques. ISPRS J Photogramm Remote Sens 86:100–110
125. Zhang T et al (2017) Urban building density estimation from high-resolution imagery using multiple features and support vector regression. IEEE J Sel Top Appl Earth Obs Remote Sens 10(7):3265–3280
126. Yoo C et al (2019) Comparison between convolutional neural networks and random forest for local climate zone classification in mega urban areas using Landsat images. ISPRS J Photogramm Remote Sens 157:155–170

127. Rosentreter J, Hagensieker R, Waske B (2020) Towards large-scale mapping of local climate zones using multitemporal Sentinel 2 data and convolutional neural networks. Remote Sens Environ 237:111472
128. LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521(7553):436
129. Cao R et al (2018) Integrating aerial and street view images for urban land use classification. Remote Sens 10(10):1553
130. Huang B, Zhao B, Song Y (2018) Urban land-use mapping using a deep convolutional neural network with high spatial resolution multispectral remote sensing imagery. Remote Sens Environ 214:73–86
131. Kang J et al (2018) Building instance classification using street view images. ISPRS J Photogramm Remote Sens 145:44–59
132. Zhang X et al (2018) Multiscale geoscene segmentation for extracting urban functional zones from VHR satellite images. Remote Sens 10(2):281
133. Li M et al (2016) Urban land use extraction from Very High Resolution remote sensing imagery using a Bayesian network. ISPRS J Photogramm Remote Sens 122:192–205
134. Modara M, Ait Belaid M, AlJenaid S (2014) Assessment of land-use/land-cover change in Muharraq Island using multi-temporal and multi-source geospatial data. Int J Image Data Fus 5(3):210–225
135. Zhang L, Zhang M, Yao Y (2018) Mapping seasonal impervious surface dynamics in Wuhan urban agglomeration, China from 2000 to 2016. Int J Appl Earth Obs Geoinf 70:51–61
136. Wurm M et al (2014) Investigating the applicability of Cartosat-1 DEMs and topographic maps to localize large-area urban mass concentrations. IEEE J Sel Top Appl Earth Obs Remote Sens 7(10):4138–4152
137. Biljecki F, Ledoux H, Stoter J (2017) Generating 3D city models without elevation data. Comput Environ Urban Syst 64:1–18
138. Erbek FS, Özkan C, Taberner M (2004) Comparison of maximum likelihood classification method with supervised artificial neural network algorithms for land use activities. Int J Remote Sens 25(9):1733–1748
139. Zhou Y et al (2018) A global record of annual urban dynamics (1992–2013) from nighttime lights. Remote Sens Environ 219:206–220
140. Li M et al (2017) Incorporating open source data for Bayesian classification of urban land use from vhr stereo images. IEEE J Sel Top Appl Earth Obs Remote Sens 10(11):4930–4943
141. Mertes CM et al (2015) Detecting change in urban areas at continental scales with MODIS data. Remote Sens Environ 158:331–347
142. Li X et al (2015) Assessing street-level urban greenery using Google Street View and a modified green view index. Urban For Urban Green 14(3):675–685
143. Ke Y et al (2016) Downscaling of MODIS One kilometer evapotranspiration using Landsat-8 data and machine learning approaches. Remote Sens 8(3):215
144. Chen T-HK et al (2019) Detecting and monitoring long-term landslides in urbanized areas with nighttime light data and multi-seasonal Landsat imagery across Taiwan from 1998 to 2017. Remote Sens Environ 225:317–327
145. Lyu H, Lu H (2017) A deep information based transfer learning method to detect annual urban dynamics of Beijing and Newyork from 1984 to 2016. In: 2017 IEEE international geoscience and remote sensing symposium (IGARSS). IEEE
146. Liu X et al (2018) High-resolution multi-temporal mapping of global urban land using Landsat images based on the Google Earth Engine Platform. Remote Sens Environ 209:227–239
147. Burke M et al (2021) Using satellite imagery to understand and promote sustainable development. Science 371(6535):eabe8628
148. Ronneberger O, Fischer P, Brox T (2015) U-net: convolutional networks for biomedical image segmentation. In: International conference on medical image computing and computer-assisted intervention. Springer
149. Long J, Shelhamer E, Darrell T (2015) Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition

150. Chen L-C et al (2017) Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. IEEE Trans Pattern Anal Mach Intell 40(4):834–848
151. Zhu XX et al (2017) Deep learning in remote sensing: a comprehensive review and list of resources. IEEE Geosci Remote Sens Mag 5(4):8–36
152. Qiu C et al (2020) A framework for large-scale mapping of human settlement extent from Sentinel-2 images via fully convolutional neural networks. ISPRS J Photogramm Remote Sens 163:152–170
153. Zhao K et al (2018) Building extraction from satellite images using mask R-CNN with building boundary regularization. In: Proceedings of the IEEE conference on computer vision and pattern recognition workshops
154. Sharifi A (2019) Resilient urban forms: a macro-scale analysis. Cities 85:1–14
155. Sharifi A (2019) Urban form resilience: a meso-scale analysis. Cities 93:238–252

Dr. Tzu-Hsin Karen Chen is a land change and health scientist. Her work focuses on urban environmental change and its relation to physical and mental health inequality. She enjoys integrating machine learning, GIS science and remote sensing methods.

Dr. Alexander V. Prishchepov is an associate professor at the University of Copenhagen. He focuses on monitoring with satellite remote sensing the dynamic of land cover, land use and human footprint. He analyzes the causal drivers of land-use/land-cover change, land potentials. His research addresses many UN Sustainable Development Goals.

Professor Clive E. Sabel is a spatial data scientist, working within the field of Public Health. His work reveals epidemiological relationships to physical and social environmental exposures, using Big Data, machine learning and wearable sensors.