INTERNATIONAL JOURNAL OF SUSTAINABLE CONSTRUCTION ENGINEERING AND TECHNOLOGY
VOL. 14 NO. 5 (2023) 16-36

IJSCET

http://publisher.uthm.edu.my/ojs/index.php/ijscet

ISSN : 2180-3242 e-ISSN : 2600-7959

International Journal of Sustainable Construction Engineering and Technology

The Development of the Vulnerability Index (VI) using Principal Component Analysis (PCA)

Siti Aekbal Salleh[1,2*], Nurul Amirah Isa[3], Nurul Aida Siman[4], Nur Hidayah Zakaria[1], Lynlei L. Pintor[6], Rostam Yaman[2], Nazri Che Dom[2,5]

[1]School of Geomatic Science and Natural Resources, College of Built Environment, Universiti Teknologi MARA, 40450, Shah Alam, Selangor, MALAYSIA

[2]Institute of Biodiversity and Sustainable Development, Universiti Teknologi MARA, 40450, Shah Alam, Selangor, MALAYSIA

[3]Faculty of Asia Built Environment, University of Geomatika Malaysia, Setiawangsa, 54200, Kuala Lumpur, MALAYSIA

[4]As White Global Sdn. Bhd, Suite, 10.1 Level 10, Centrepoint North Mid Valley City, 58000, Kuala Lumpur, MALAYSIA

[5]Faculty of Health Science, Universiti Teknologi MARA, 40450, Shah Alam, Selangor, MALAYSIA

[6]Ecosystems Research and Development Bureau (ERDB), Department of Environment and Natural Resources, College, Laguna, PHILIPPINES

*Corresponding Author

DOI: https://doi.org/10.30880/ijscet.2023.14.05.002
Received 04 April 2023; Accepted 03 July 2023; Available online 31 October 2023

Abstract: Climate change elevates the rate of emergence of urban heat islands (UHIs), especially in the tropics. UHIs severely affect human comfort and health. Many studies have suggested that urban areas should be properly mitigated or planned. To cope with this, it is best to present the issue using easy-to-understand approaches to allow for better decision-making, especially during urban planning. Based on the information, adaptations and mitigation strategies can be suggested in order to reduce the impact. Hence, this research was aimed at determining the heat vulnerability index (HVI) of urban areas. This study was conducted in Malaysia in the Klang Valley, a tropical city with a complex urban morphology. Remote sensing techniques were employed to extract and derive the spatial index values for exposure, sensitivity, and adaptive capacity. A principal component analysis (PCA) was used to estimate the vulnerability as well as to generate the HVI. The most vulnerable districts were found to be Petaling (1.00), Kuala Lumpur (0.99), and Putrajaya (0.95). Kuala Lumpur had a level of exposure that was high (0.56), a level of sensitivity that was high (0.84), and capacity to adapt that was low (0.54), while Petaling had a high exposure value (0.56), very high sensitivity (1), and high adaptive capacity (0.72). A Pearson's correlation (*r*) test also revealed that the variables used were highly correlated. From the preliminary findings, the vulnerability of the population to high temperatures in the Klang Valley can be identified to help develop adaptative plans that are targeted as a response to rapid warming in the future in Malaysia.

Keywords: UHI, GIS, Heat Vulnerability Index (HVI), remote sensing, urban climate

Corresponding author: aekbal@uitm.edu.my

1. Introduction

Climate change has a significant impact on ecosystems, and over the last 130 years there has been a continuous rise in global temperatures with significant implications for a wide range of climate-related factors (Nwankwoala, 2015). Although climate change can be caused by various factors (Isa *et al*., 2020), urbanisation has been identified as the driving factor behind this event in this era of modernisation. Urbanisation can be defined as the development of populations and cities, and the process refers to much more than simple population growth; rather, it involves an analysis of the related economic, social, and political transformations (Shuid, 2004). The population increase in urban areas is mainly due to job opportunities, educational factors, and economic opportunities.

Urban heat islands (UHIs) are created as a result of the urbanisation process. UHIs can be defined as urban areas that experience an unduly hot climate compared to their surrounding areas (Ooi *et al*., 2018; Isa *et al*., 2018). The situation of a UHI arises when a particular urban area is significantly warmer than the surrounding suburban and rural areas as a result of the use of concrete and asphalt, which are able to retain heat and take longer to cool down, in urban sites and buildings (Nayak *et al*., 2018; Zeeshan & Ali, 2022).

Urbanisation also involves the cutting down of trees to make way for commercial development, route designs, industrial sectors, and urban development (Mölders, 2011). Since the role of green areas is to reduce heat, thus, the surrounding temperature has a significant impact on the processes of development and urbanisation. Weather conditions can be influenced by even minor changes in land cover, and these can have a significant impact on the urban climate (Isa *et al*., 2021). The attributes of people (such as their health status, socio-demographics, etc.) and certain elements of their local communities have been linked to their susceptibility to heat (environment, community demographics). These traits, or "heat vulnerability factors", may have a significant impact on a person's capacity to endure heat (New York State Department of Health, 2018). A common approach to developing and mapping vulnerability indices is to apply the characterisation of vulnerability to the process. These indices serve to draw attention to vulnerable locations so that precise mitigation and adaptation strategies can be developed to reduce the likelihood of incident-related effects, including death, disease, loss of livelihood, or damage to property and infrastructure (Reid *et al*., 2012).

Reid *et al*. (2009) created a national heat vulnerability index (HVI) to trace populations that are vulnerable to heat at the sub-metropolitan level via variables associated with vulnerability across the United States. The study found that the HVI was linked to more hospitalisations and deaths in every state, both on days that were normal and very hot. But on extremely hot days, the correlations between thermal illnesses, electrolyte imbalance, acute renal failure, and nephritis in the state of California, death due to any cause in the state of New Mexico, thermal illnesses in the state of Washington, and hospitalisations due to respiratory illnesses in the state of Massachusetts were clearer (interaction p-value of 0.05).

Due to climate change, heat waves are likely to get worse, last longer, and happen more often in many parts of the world (Isa *et al*., 2018; 2020). A lot of evidence shows that heat waves and extremely hot weather lead to more deaths, and there is a growing amount of evidence that they lead to more illnesses (Basu 2009; Basu and Samet 2002). To help guide public health efforts, whether ahead of, during, or in the aftermath of such an event, an early identification of vulnerability to extreme heat events should be developed (Nayak *et al*., 2018). This situation can save many lives, and urban planning can be carried out in the most effective way (Niu *et al*., 2021).

Rural-to-urban migrations frequently lead in environmental problems such as overpopulation, pollution, and poor sanitation. Unfortunately, bringing people out of poverty and into more developed countries frequently comes at the expense of the local ecosystem. A massive urban sprawl, if poorly designed, can increase deforestation, habitat degradation, and greenhouse gas (GHG) or carbon emissions. According to Lee Poh Onn, Senior Fellow of the ISEAS Yusof Ishak Institute, urban growth feeds commercialisation and industrialisation, which will increase the use of fossil fuels that will later contribute to global warming and climate change. In fact, the effects of heat waves caused by UHIs on the environment and health of the people are a major concern today (Kosatsky, 2005). This is because the number and length of extremely hot spells are getting more frequent and intense (Robin *et al*., 2008).

Therefore, the intensity and frequency of bouts of extremely hot weather is forecasted to increase due to the effects of climate change which will, in turn, significantly affect human health (Confalonieri and Menne, 2007; IPCC, 2012), many large cities around the world should be concerned about the effects of UHIs in a potentially warmer world (Hien, 2016). Heat activities, whether due to human or machine activities, contribute to the UHI phenomenon. Urban heat islands (UHIs) are commonly lower in the day and higher at night (Lemonsu, Viguié, Daniel, & Masson, 2015). The main factors with regard to adapting to extreme heat events are mostly land cover and land use. The heat retained by the concrete and asphalt used in urban settings and buildings in urban areas takes longer to be released, thus creating UHIs that are significantly hotter than the surrounding suburban and rural areas. So, during hot seasons, people who live in cities tend to experience higher daytime temperatures, less cooling at night, and more and longer extreme heat events (Santamouris, 2015).

With the increasing frequency, intensity, and period of extreme hot weather, it is no surprise that heat-related morbidity and mortality among vulnerable populations may rise, as human health is strongly related to the extreme heat. If one's person is exposed to heat for a long period of time, it can have a detrimental effect on the body's system, even if the heat is not extreme. Thus, with a higher intensity of temperature for a long period of time, the chances of an increased rate of mortality per area are said to be high. Many studies have examined the effects of heat waves on health as it is a developing concern in the realm of environmental health (Kovacs, Belusko, Pockett & Boland, 2016; Campbell, Remenyi, White & Johnston, 2018; Arsad, Hod, Ahmad, Ismail, Mohamed, Baharom, Osman, Mohd Radi & Tangang, 2022). The heat waves that occurred in previous years, like the one in Europe in 2003 that killed up to 80,000 people (Robin *et al*., 2008), and the one in Russia in 2010 that killed an estimated 54,000 people (Revich, 2011), have drawn attention to this problem all over Europe (Kosatsky, 2005).

As a result, the advancement of a HVI is thought to play an important role in addressing this issue. In this study, the HVI in the study area could provide useful information to the government in developing a heat warning plan or for any related parties, such as planners, decisionmakers, and other government institutions. In some cases, interventions by the relevant authorities in specific areas to prevent heat-related deaths have resulted in a lower mortality rate in subsequent heat events (Ebi *et al*., 2004). The relationship between the HVI and urbanisation can also help to improve future urban planning. An HVI map of an urban area will provide information on different vulnerability levels in different regions of the urban area based on certain parameters such as exposure, sensitivity, and adaptability. Based on the HVI map, the populations and areas within a city that are most vulnerable to heat can be identified and analysed for further investigation. The benefit of this action is that it can help local governments to distribute resources to the areas that are in greatest need (O'Neill *et al*. 2009).

This study attempted to identify the HVIs for urban areas using certain variables representing exposure (land surface temperature (LST)), sensitivity (population density, minority, elderly, very young, and differently-abled people), and adaptivity (road density, geographical elevation, and normalised difference vegetation index (NDVI)). Both remote sensing and geographic information system (GIS) technologies were used in this study. To cut the number of correlated variables into fewer uncorrelated components, a principal component analysis (PCA) was used to identify the HVI values in urban areas. A principal component analysis (PCA), which is usually used to search for patterns in high-dimension data, has been regularly used in heat vulnerability studies to create a composite index (Bai *et al*., 2016).

2. Study Area

This present study focused on the Klang Valley region, which encompasses the Federal Territory of Kuala Lumpur as well as towns and cities in the adjacent state of Selangor. This is region is also referred to as "Greater Kuala Lumpur." The Titiwangsa Mountains to the north and east, and the Malacca Strait to the west, form the geographical boundaries of the Klang Valley, which stretches northwest to Rawang, southeast to Semenyih, and southwest to Klang and Port Klang.

Despite the fact that there are no officially designated boundaries, the Klang Valley, with an area of 8318 square kilometres, can be considered to encompass the Federal Territories of Kuala Lumpur and Putrajaya, as well as several other districts in its vicinity, namely the districts of Petaling, Klang, Gombak and Hulu Langat in Selangor.

The valley gets its name from the Klang River, which runs through it. It starts in Port Klang and ends in Hulu Klang, both of which were linked to the development of the area in the late nineteenth century as a cluster of tin mining settlements. Most of the development occurred in the east-west direction (between Gombak and Port Klang), but the urban areas near Kuala Lumpur have since expanded to the south, towards the border with Negeri Sembilan, and to the north, towards Rawang. This study, however, only focused on eleven districts: Kuala Lumpur, Putrajaya, Gombak, Hulu Selangor, Hulu Langat, Petaling, Sepang, Sabak Bernam, Kuala Selangor, Klang, and Kuala Langat. Figure 1 depicts images of the research area in the Klang Valley, Malaysia.

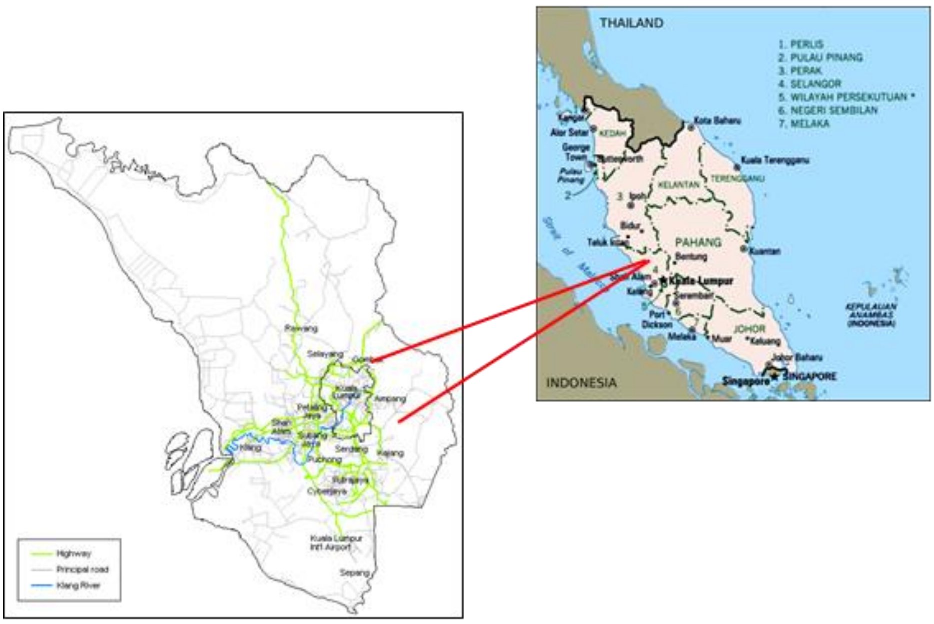

Fig. 1 - Study area

3. Materials

The majority of the data used in this present study comprised remotely-sensed data acquired by the Thermal Infrared Sensor (TIRS) and the Landsat 8 Operational Land Imager (OLI), which is jointly operated by the National Aeronautics and Space Administration (NASA) and the United States Geological Survey (USGS). The data, dated 26 March 2016, consisted of eleven (11) spectral bands with spatial resolutions of 30 to 60 metres. Thermal bands 10 and 11 were bands that specialised in detecting the thermal condition, which was useful for providing accurate surface temperature extractions (Li *et al*., 2013).

In this study, the most vital dataset was the statistical data. Without these data, the variables for the HVI could not have been calculated. The socioeconomic and demographic data for the Klang Valley area were obtained from the Department of Statistics Malaysia (DOSM). The DOSM is the premier government agency that serves to collect data and information related to the economic and social aspects of the country. The data collected involved descriptions of the density, age, citizenship, and physical abilities of the population. Besides the DOSM, some of the data were also obtained from the Ministry of Women, Family and Community Development (KPWKM) through its open-source data website.

Table 1 - The variables, descriptions, and sources of the indices

Index	Variables	Data Description	Data Source
Exposure	LST	LST values	Landsat 8 OLI/USGS 07.03.2016
Sensitivity	Population Density	Inhabitant per hectare of population	DOSM (2016)
	Minorities	Inhabitant per hectare of non-citizen	DOSM (2016)
	Very Young People	Inhabitant per hectare below 5 years old	DOSM (2016)/KKLW
	Elderly People	Inhabitant per hectare above 65 years old	DOSM (2016)/JPBD
	Differently-abled People	Inhabitant per hectare of handicapped people	KPWKM (2016)
Adaptive Capacity	NDVI	NDVI values	Landsat 8 OLI/USGS 26.03.2016
	Geographical Elevation	Means value	USGS
	Road Density	km/km^2 of roads per area	HERE Map Data

4. Methodology

4.1 Extraction of Remotely Sensed Data: Top-of-Atmosphere (TOA) Temperature

The main objective of deriving an LST from the Landsat 8 satellite imagery was to estimate continuous air temperature and identify suburban hot spots. Only Band 10 was used for this study due to large uncertainties and errors in the TIRS band 11 issued by the USGS. A mono-windowed algorithm was used to retrieve the LST from the TIRS data of the Landsat 8. Three steps were required to derive the LST, namely, 1) calculating the top-of-atmosphere (TOA) or at-sensor brightness temperature; 2) estimating the land surface emissivity (LSE); and 3) using the mono-window algorithm to determine the LST.

A series of formulae were needed to calculate the LST value. The NDVI composite layer and the thermal layers of Landsat 8 OLI, which were in Band 10, were imported to the table of contents in the ArcMap. To determine the LST, the digital number data at the TOA were converted to radiance using the Landsat 8 Band 10-specific gain and bias parameters (USGS, 2015). The digital number was converted into radiance using Equation 1. The radiance number for Band 10 was filled, as provided in the metadata, by using the raster calculator function.

$$RADIANCE\_MULTI\_BAND * (BAND\ LAYER) + RADIANCE\_ADD\_BAND \qquad (1)$$

The next step was to transform the spectral radiance to temperature. The K1_CONSTANT_BAND and K2_CONSTANT_BAND for the thermal layer could be found in the metadata document. The value of 273.15 was used to convert the temperature unit from Kelvin to Celsius. The process was performed for the Band 10 thermal layer. Equation 2 was used for the conversion of the TOA brightness temperature, as shown below.

$$\boldsymbol{T = \frac{K2}{ln\,(\frac{K1}{L\lambda})+1}} \qquad (2)$$

$\boldsymbol{T = \frac{K2}{ln\,(\frac{K1}{L\lambda})+1}}$T - TOA brightness temperature (K)

Lλ - TOA spectral radiance (Watts/m2 * srad * μm))

K1 - Band specific thermal conversion constant from the metadata (K1_CONSTANT_BAND_x, where x is the thermal band number)

K2 -Band specific thermal conversion constant from the metadata (K2_CONSTANT_BAND_x, where x is the thermal band number)

Satellite temperature data of the specific time when the satellite image was taken could be viewed by inserting both the atmosphere brightness temperature data of both the thermal layers into the cell statistics tool to obtain the output. This output was not used to create the LST layer.

4.2 Normalised Difference Vegetation Index (NDVI)

The NDVI parameter is often used in investigations into the LST as it is less sensitive to changes in atmospheric conditions compared to other parameters. Thus, it is has become the most important parameter, especially in vegetation monitoring. In calculating the proportion of vegetation, the NDVI value must be inserted into the equation for the *Pv* as its value is related to the NDVI. The *Pv* value is needed in the next calculation for emissivity. Thus, the determination of the NDVI is a very crucial step in the calculation of the LST as they are related to one another. The equation for the NDVI is shown as follows, where *R* is the red band (Band 4) and NIR is the near-infrared band (Band 5).

$$NDVI = (Band\ 5 - Band\ 4) / (Band\ 5 + Band\ 4) \qquad (3)$$

4.3 Proportion of Vegetation

The formula in Equation 4 was used to calculate the proportion of vegetation. The NDVI for vegetation (NDVIv = 0.5) and the NDVI for soil (NDVIs = 0.2) can be used to determine the *Pv* and apply it to global conditions. Although at-surface reflectance can be used to calculate the global NDVI, the TOA reflectivity cannot be used for this purpose as atmospheric conditions affect the NDVIv and NDVIs.

$$Pv = Square\ ((NDVI - NDVIs) / (NDVIv - NDVIs)) \qquad (4)$$

4.4 Emissivity of the Land Surface (LSE)

The LSE is a proportional aspect that gauges black-body radiance, which is based on Planck's law, to forecast the amount of radiance emitted. It is the efficiency of transporting thermal energy from the ground surface into the atmosphere (Sobrino, 2004). The determination of the LSE is based on the conditional Equation 5.

$$\boldsymbol{\varepsilon_{\lambda} = \varepsilon_{\gamma\lambda} P\gamma + \varepsilon_{\varsigma\lambda}\,(1 - P\gamma) + C\lambda} \qquad (5)$$

where, *εs* is the emissivity of the soil, *εv* is the emissivity of the vegetation, and *C* is the roughness of the surface. Rough surfaces are assigned a value that is constant; namely 0.005; while surfaces that are flat and homogenous are assigned 0. If NDVI < 0, it is assigned an emissivity of 0.991 and classified as water. Normalised difference vegetation indices (NDVIs) ranging from 0.2 and 0.5 are considered a blend of soil and vegetation cover and are used to calculate emissivity. However, if NDVI > 0.5, it is considered vegetation and assigned an emissivity of 0.973.

4.5 Land Surface Temperature (LST)

Equation 6 was used to calculate the LST retrieval or the emissivity-corrected LST (Ts).

$$T\varsigma = \frac{BT}{\{\, 1 + [(\lambda BT/\rho)\ ln\ \varepsilon_{\lambda}]} \qquad (6)$$

where, *Ts* is the LST in Celsius (°C); λ is the emitted radiance's wavelength, where a peak response and limiting wavelength average (λ = 10.895) was used; BT is at-sensor BT (°C); and $\varepsilon\lambda$ is the emissivity, which was calculated using Equation 7.

$$\rho = h\frac{c}{\sigma} = 1.438\ x\ 10^{-2}\ m\ K \qquad (7)$$

where, σ is Boltzmann's constant, namely, 1.38×10^{-23} J/K; *h* is the Planck constant, namely, 6.626×10^{-34} J s; and *c* is the light's velocity, namely, 2.998×10^{8} m/s (Weng, 2014).

5. Heat Vulnerability Index (HVI) Development

The creation of an HVI at the block group level, which would enable more spatially-tailored adaptive mitigation approaches, was another area of attention for this work. The general methodology by Reid *et al*. (2009) was used, with very minor alterations, to accomplish this. A PCA was done at the block group level on the variables of total population, age, race, income, education, language ability, household type, and land cover in order to obtain a structure that would lessen the complexity of the variables. To produce an extensive HVI, variables were selected to represent various heat vulnerabilities and statistical components, and the spatial analysis technique was utilised.

By integrating the GIS and remote sensing data comprised of satellite imagery of Landsat 8 OLI, the HVI for each district in the Klang Valley was developed using the suitable method. The three-dimensional patterns of vulnerability to heat were charted across the urban areas according to the districts located all around the Klang Valley, including Kuala Lumpur, Petaling, Putrajaya, Gombak, Hulu Selangor, Hulu Langat, Sabak Bernam, Kuala Selangor, Sepang, Kuala Langat, and Klang. Different types of data to represent the variables of each category of exposure, sensitivity and adaptive capacity were processed and analysed by using a statistical method. The data obtained were limited to the year 2016 only due to the lack of available data, especially at the district level, as most of the data were at the state level. Then, the remote-sense-based indices were combined with the statistical data in order to calculate the indices. The unit of data for sensitivity was uniformly transformed into unit per surface, where this study used unit per hectare.

In order to acquire more accurate indices, most extant studies on vulnerability indices (VIs) used a plethora of multiple combinations of sensitivity, exposure, and adaptive capacity (Inostroza *et al*., 2016). A model was created by compiling a number of factors, that combine the risks of sensitivity and exposure, while also taking into account the capacity for adaptation. This present study viewed VIs as a function of capability to adapt (A); which were respectively stated as sensitivity (S) and exposure (E); and impact components (I). The individual indices derived from the distinct measurements of adaptive capacity, sensitivity, and exposure were combined to identify the vulnerability variations of the model. The use of this technique, as opposed to the accumulation of vulnerability quantifications, can provide decision makers with additional information. As a direct consequence of this, the HVI value, denoted by the letter V, was determined by employing a summary model that characterised this equation:

$$V = f\ (I, A) \qquad (8)$$

$$I = (E, S) \qquad (9)$$

Or

$$V_j = Ej + Sj - A_j \qquad (10)$$

where, Ej is the level of exposure at the j census tract, Sj is the level of sensitivity of the census tract, and A is the capacity to adapt of the j census tract.

5.1 Variables of Exposure

The LST, which is a measurement of heat exposure, was calculated from the thermal LSE of the Klang Valley captured through remote sensing images. It was determined using standard methods and a Landsat 8 thermal band image. The pictures were cropped to solely depict the research zone. ArcGIS 10.6 software was used to calculate the average LST of each district using the statistics of each zone. The estimated exposure of each census tract was determined using the LST based on pixels. For 2016, the exposure level of each census tract was estimated as its mean LST and the addition of one standard deviation (SD) in temperature. This calculation corresponded to a conservative estimate of exposure.

5.2 Variables of Sensitivity

Five variables; namely, 1) minority populations, 2), differently-abled populations 3) very young populations, 4) elderly populations, and 5) population density; were selected from the census database to determine the sensitivity. The population density was defined as the overall number of individuals of age in every household. The minority population was non-Malaysian citizens from other countries; such as Bangladesh, Pakistan, Thailand or Cambodia; living in the Klang Valley. The very young population was the total number of individuals below the age of 5 while the elderly population was senior citizens aged 65 and above.

Table 2 displays the data sources and descriptions. The relative spatial density variable was calculated by adding the overall instances of every category then dividing it by each census tract's net built-up area. The census tract and population data from 2016 were obtained from the DOSM as the most recent census; i.e., 2019; was not available at the time of writing.

5.3 Variables of Adaptive Capacity

Three variables were used to calculate the adaptive capacity: 1) the NDVI, 2) geographical elevation, and 3) roads (Table 2). The NDVI directly measured the extent and composition of the vegetative cover. Every census tract's average NDVI plus one SD was used to estimate higher values. The NDVI values were calculated from the Landsat 8 satellite images (Table 2) using ArcGIS software. Then, the geographical elevation was derived from the Shuttle Radar Topography Mission (SRTM) images that were downloaded from the USGS website. The variable of the road was calculated as the density of paved roads located in every square kilometre in a census tract. ArcGIS 10.6 was used to determine all the spatial statistics.

5.4 Bias Controls

To avoid bias caused by low or high levels of variance in the variables, the data was pre-processed using centring and scaling to guarantee that the potential of the PCA was at its maximum. The main reason for performing these procedures was to ensure that all the variables contributed equally to the vulnerability model so as to have equal weighting in the data analysis. Only the scaling and centring procedures were used for the sensitivity data because the variables shared the same unit (n/hectares). The mean centre was calculated by computing every variable's average value and subtracting it from the data. Unit variance scaling was achieved by multiplying the variables by the inverse of their SD. To standardise the variables of adaptive capacity, which were calculated in diverse units, the average was subtracted, and the result divided by the SD.

5.5 Selection of Indicators

To determine which district would be impacted the most by possible developments, an evaluation of the existing situation was required. An HVI for the human population was designed and customised based on the circumstances and situation so as to identify the vulnerability status in the Klang Valley district. Firstly, the decision to pick a broad or narrow selection of indicators depended heavily on the best available data and representative indicators from Malaysia based on previous literature and studies. Part of the indicator selection involved choosing variables to represent these vulnerability indicators (Tate, 2012). Data accessibility, validity, the intended number of indicators, and statistical properties all influence the choice of representative variables. Vulnerability cannot be measured by a single established collection of indicators. Thus, the decision of which factors to include is left to the researcher, and there is no validation as to whether ta specific indicator should be used to determine the HVI. The use of different kinds of variables from different researchers might produce different results, which can be compared to see the difference between them.

A PCA is used in an inductive method to condense a large number of variables into a small number of uncorrelated factors (Jolliffe & Cadima, 2016). This method includes the process of variable reduction and selection, rather than having researchers choose the variables on their own, thus, making it a more objective than a deductive method (Praene, Damour, Radanielina, Fontaine, & Reviere, 2019). However, it is crucial to note that this logic condenses a large number of variables into fewer ones based on the spatial variation they account for. A few factors with various spatial patterns are produced as a consequence of combining variables with comparable spatial patterns into one factor.

A PCA was used to prevent co-linearity and reduce variable complexity by decreasing the number of principal components (PCs), that cause most of the variations in the detected variables, in the original variables set. The variance-weighted method was used to assign weights to the variables by summing the explained variances of each component to generate a collective PC score (z-score) (Schmidtlein *et al.*, 2008). Eigenvalues associated with the vector for each PC indicated the significance ranking of the PCs, depending on how much data variability they captured (Thurstone, 1947). Kaiser's recommendations and the Pearson correlation (*r*) matrix were used to ensure that the remaining factors solely contained those with PCs that had eigenvalues > 1.0. The eigenvectors were orthogonally rotated using Varimax rotation to optimise the clarifications that this present study obtained and maximise loading distributions across the PCs to produce a collection of factors that were interpretable and depict its straightforward construction (Thurstone, 1947). The matrix of the z-score depicts a fresh set of unrelated variables that mathematics can further manipulate.

The PCs obtained for adaptive capacity and sensitivity were translated into a z-score vector while maintaining the structure of the initial data, with 0 in the vector's middle, and preserving z-scores' signs and weights. The z-scores were calculated for each of the eleven districts, ensuring that each of them contained n PC scores, to determine the n retaining PCs. Typically, a vulnerability index is derived from multiple sets of indicators in order to conduct a quantitative evaluation of the susceptibility of a system. A customised vulnerability index was created by compiling data from numerous sources. A multivariate index was utilised to make comparisons between the different regions. This research analysed three factors that were particularly relevant to this study, and as a result, this study was aimed at investigating the following topics:

a) exposure to the environment pertains to the magnitude and rate of change in climatic variables, such as temperature, which are known to affect the human population. This can be thought of as the vulnerability of the human population to the environment.

b) sensitivity to its effects – the extent to which a community is affected by climate variability or change.

c) adaptive capacity to survive with the effects – a measure of the resources and capabilities of a society to counteract the adverse consequences of change in the environment or capitalise on potential benefits.

5.6 Indicator Normalisation

Equation 11 was used to normalise the partial results of the sensitivity, exposure, and adaptive capacity to a 0 to 1 scale.

$$\beta = [(x - x\_min)/(x\_max - x\_min)] \qquad (11)$$

where, β is the normalised value, χ is the original value, and χmin and χmax are the minimal and maximum values of the dataset, respectively. Using equal intervals, the normalised values were divided into five categories to illustrate the three-dimensional dispersals at the level of the census tract. After normalising the incomplete scores, the equation in Section 5.0 was used to calculate the partial vulnerability value, which was then normalised to obtain the final HVI.

5.7 Statistical Test on Vulnerability Indices

The final step was to subject the vulnerability indices to a statistical test. Pearson's *r* method was used to test the vulnerability indices as to whether all the variables were significant to each other. Pearson's *r* is a statistical formula that measures the strength between variables and relationships. A value of 1 denotes a strong positive association, -1 a strong negative relationship, and 0, no relationship. Every positive rise in one variable causes a fixed proportional increase in the other. A correlation coefficient of -1 means that for every positive increase in one variable, there is a negative reduction in the others, and zero means there is no positive or negative association for every increase. The Pearson's *r* is depicted in Equation 12.

$$r = (n\ (\Sigma xy) - (\Sigma x)\ (\Sigma y))/\sqrt{([n\Sigma x^2 - (\Sigma x)^2]\ [n\Sigma y^2 - (\Sigma y)^2)} \qquad (12)$$

where $\sum xy$ is the total value of variable.

6. Result and Analysis

6.1 Exposure

The temperature data were derived by calculating the LST using the Landsat 8 OLI satellite imagery. The mean temperature for each district was derived and plotted as a linear graph (Temperature vs District). Figure 2 shows the trend of the average LST for each district in the Klang Valley. The graph shows the predicted surface temperatures for the Klang Valley area during March 2016. The highest LST value was found in Kuala Lumpur (32.97ºC), where its vegetation cover levels were lower based on the NDVI value for that area. Petaling had the second highest LST value (32.48ºC), followed by Putrajaya (30.45ºC), Klang (30.15ºC), Kuala Langat (29.65ºC), Sepang (29.57ºC), Kuala Selangor (29.23ºC), Gombak (28.87ºC), Hulu Langat (28.62ºC), Sabak Bernam (28.51ºC), and Hulu Selangor (27.57ºC).

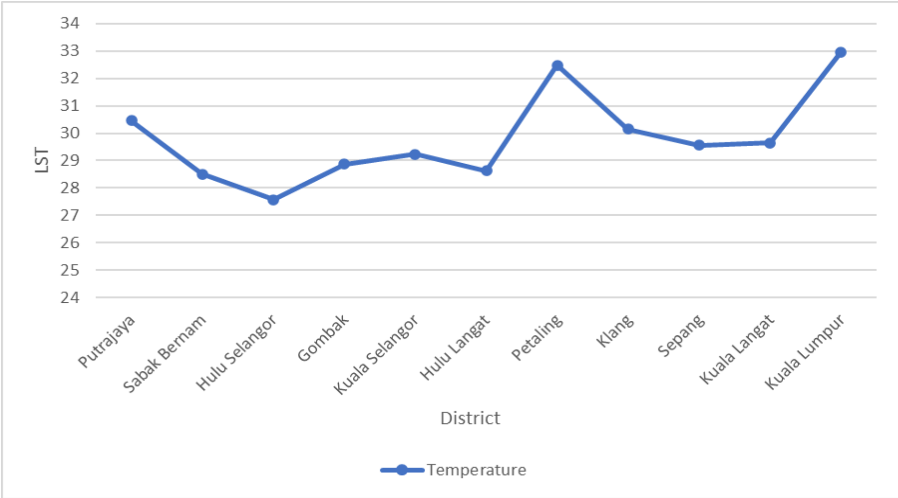

Fig. 2 - Linear graph of the LST of every district

Most of the areas in north-eastern Klang Valley; such as Hulu Selangor and Gombak; had the lowest LSTs. This could be because these areas have low building densities and high vegetated areas or vegetation cover. After all, cooling is prevalent is areas that are large and open. Cooling islands have also been found to correlate with larger quantities of vegetation as well as consolidated green parks, infrastructure, and general areas (Mansor & Harun, 2014). Figure 3 shows the average surface temperature map for the Klang Valley.

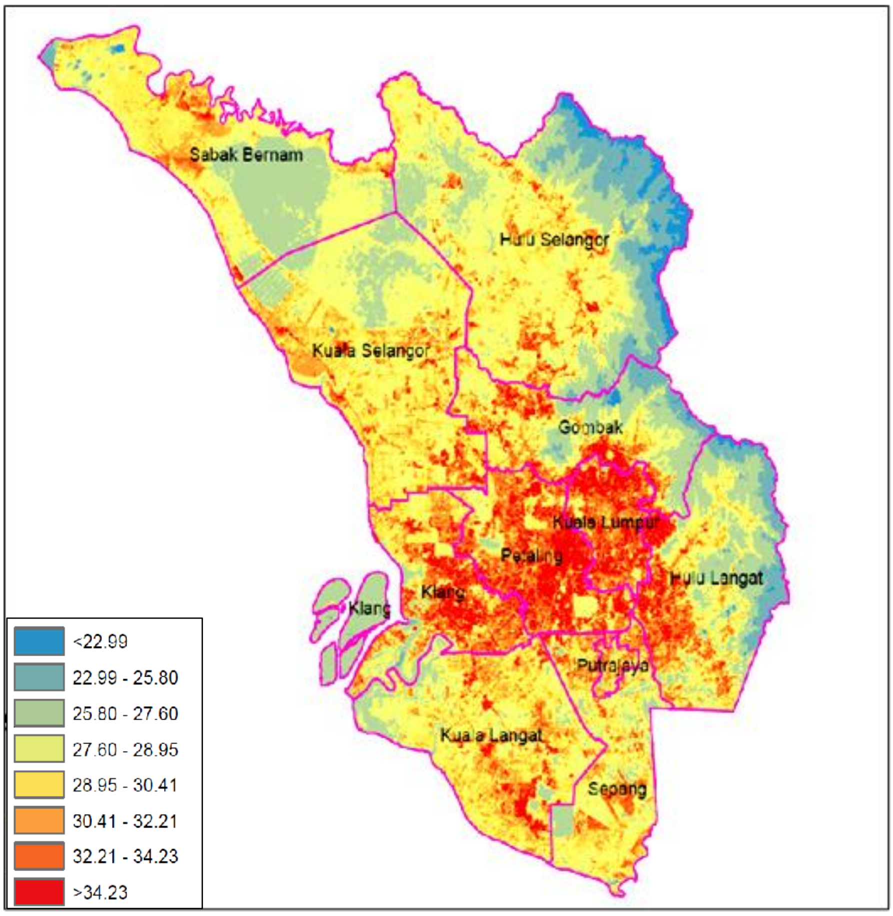

Fig. 3 - Average LST map of the Klang Valley

The average exposure value was 0.48, which is considered moderate on a five-point gauge, while the SD was 0.06. The level of exposure across the Klang Valley, according to their respective exposure values, is shown in Table 2 and Figure 4. Petaling, Kuala Lumpur, and Sabak Bernam had the highest exposure values.

Table 2 - Level of exposure in the Klang Valley

District	Area (Sq Km)	Mean Temperature (Celsius)	Exposure Value
Hulu Langat	841	28.62	0.4037
Petaling	499	32.48	0.5626
Kuala Selangor	1195	29.23	0.4037
Sepang	558	29.57	0.5023
Kuala Langat	859	29.65	0.3708
Klang	633	30.15	0.4881
Gombak	629	28.87	0.5376
Hulu Selangor	1758	27.57	0.5045
Sabak Bernam	993	28.51	0.5516
Kuala Lumpur	241	32.97	0.5571
Putrajaya	44	30.45	0.4470

6.2 Sensitivity

6.2.1 Population Density

Table 3 presents the demographic breakdown of the residents in each district along with an evaluation of the level of sensitivity of that district. The DOSM estimated that there were 8.2234 million people living in the Klang Valley region as at 2015. Putrajaya recorded the lowest population of any district in the Klang Valley with only 88,000 people, while Petaling recorded the highest population of any district in the Klang Valley at 2.085 million people. Putrajaya encompassed the smallest area of any district in Malaysia, at only 4900 hectares, compared to Hulu Selangor's total area of 1740,000 hectares, making it the biggest district in terms of size. The population density, on the other hand, indicated an entirely different scenario than the distribution of population. Putrajaya came in at number three in terms of population density, with 17,959 individuals packed into each hectare.

Among the most densely populated districts were Kuala Lumpur (75.984 people/hectare) and Petaling (43.069 people/hectare). This was followed by Putrajaya (17.959 people/hectare), Klang (15.819 people/hectare), Hulu Langat (15.767 people/hectare), Sepang (12.500 people/hectare), Gombak (12.097 people/hectare), Kuala Langat (3.036 people/hectare), Kuala Selangor (2.025 people/hectare), Hulu Selangor (1.317 people/hectare), and Sabak Bernam, which had the lowest population density (1.220 people/hectare).

Table 3 - Population density of the Klang Valley

District	Area (Hectare)	Population (2016)	Population Density (per hectare)
Hulu Langat	84000	1324400	15.767
Petaling	48432	2085900	43.069
Kuala Selangor	119500	242000	2.025
Sepang	19808	247600	12.500
Kuala Langat	85800	260500	3.036
Klang	62678	991500	15.819
Gombak	65008	786400	12.097
Hulu Selangor	174000	229100	1.317
Sabak Bernam	99710	121600	1.220
Kuala Lumpur	24300	1846400	75.984
Putrajaya	4900	88000	17.959

6.2.2 Minority

The data on the minority variable was based on the non-citizen population for the Klang Valley. The ethnicity of the population can be different from the existing ethnic groups in Malaysia. This group of people usually come to Malaysia from other countries such as Indonesia, Bangladesh, and Cambodia to work in this country or for other purposes. For some reason, the difference in ethnicity adds to the social vulnerability through the lack of resources, quality of housing, cultural differences, and social, economic and political marginalisation (Cutter *et al*., 2003).

The most prominent problem with regard to people of different ethnicities is that they are not fluent in the local language and are unable to understand it. They are often misunderstood or are completely clueless about the instructions that are given with regard to the use of household appliances, including electric fans or air-conditioners. They may also find it difficult to comprehend official data sheets and guidelines about how to respond to heat, which makes them more vulnerable. All the researches that discovered a substantial link between ethnicity and the effects of heat on health were carried out in the US. This was consistent with the findings of the study by Hansen *et al*. (2013), which found that among the many factors influencing the sensitivity of minority ethnic groups are economic and social inequalities, language obstacles, and living situations. However, there is a knowledge deficit about these sociocultural variations and how they affect susceptibility in various countries.

Table 4 shows the total minority population in the Klang Valley using statistical data provided by the DOSM 2016. Kuala Selangor (242 thousand) and Kuala Lumpur (223.5 thousand) had the highest minority population in the Klang Valley. This was because both areas were the main centres for career opportunities. Kuala Lumpur is now called Greater Kuala Lumpur by the DBKL Mayor, and it seems all the opportunities for career development and growth are in this area.

Table 4 - Minority populations in the Klang Valley

District	Area (Hectare)	Minority Population (2016)	Minority (Per Hectare)
Hulu Langat	84000	128900	1.535
Petaling	48432	222500	4.594
Kuala Selangor	119500	242000	0.126
Sepang	19808	30400	1.535
Kuala Langat	85800	16400	0.191
Klang	62678	112700	1.798
Gombak	65008	75600	1.163
Hulu Selangor	174000	12500	0.072
Sabak Bernam	99710	4000	0.040
Kuala Lumpur	24300	223500	9.198
Putrajaya	4900	2800	0.571

6.2.3 Very Young People

The very young population consisted of people aged 5 years and below in the Klang Valley. In the Dutch and worldwide literature, it has been discovered that the demographic categories of the elderly, the very young, and those with pre-existing health impairments are disproportionately susceptible to temperature variations (Huynen *et al*., 2001). Babies are at risk because they still lack the capacity to control their body temperature and are dependent on their carer. Parents may overdress their young children or fail to give them enough shade.

The age distribution is a crucial characteristic for understanding population trends. Table 5 shows the age distribution of the extremely youthful population in the Klang Valley. The DOSM has divided the whole population into sixteen categories separated by five years. Those under 15 years of age are classified as children, those aged 15 to 64 years as working adults, and those older than 65 years as the elderly. Sabak Bernam had the smallest population of very young individuals among the states, at 51424 people, while Petaling had the greatest population (195861), followed by Hulu Langat (1262698).

Table 5 - Very young populations in the Klang Valley

District	Area (Hectare)	Very Young population (2016)	Very Young Population (per hectare)
Hulu Langat	84000	1262698	15.032
Petaling	48432	1958611	40.440
Kuala Selangor	119500	227708	1.906
Sepang	19808	230035	11.613
Kuala Langat	85800	244301	2.847
Klang	62678	934262	14.906
Gombak	65008	189520	2.915
Hulu Selangor	174000	215650	1.239
Sabak Bernam	99710	51424	0.516

Kuala Lumpur	24300	136184	5.604
Putrajaya	4900	75838	15.477

6.2.4 Elderly Population

Age was included in thirty-three researches in the meta-analysis by Romero-Lankao *et al*. (2012), while twenty-three studies have concluded that the elderly are the most vulnerable demographic. A reduced ability to control body temperature, detect thirst, function of the kidneys, produce perspiration, and maintain heart and lung reserves are all symptoms of ageing (RIVM, 2015). Therefore, the elderly are more vulnerable to heat as these bodily functions have higher at their age. Table 6 presents the elderly population by district in the Klang Valley, where Putrajaya had the lowest population of elderly people (1031) amongst the states. The population was the highest for Kuala Lumpur (134703), followed by Petaling (120322).

Table 6 - Elderly populations in the Klang Valley

District	Area (Hectare)	Elderly Population (2016)	Elder Population (per hectare)
Hulu Langat	84000	69878	0.832
Petaling	48432	120322	2.484
Kuala Selangor	119500	16216	0.136
Sepang	19808	9484	0.479
Kuala Langat	85800	17214	0.201
Klang	62678	56181	0.896
Gombak	65008	46911	0.722
Hulu Selangor	174000	13705	0.079
Sabak Bernam	99710	5055	0.051
Kuala Lumpur	24300	134703	5.543
Putrajaya	4900	1031	0.210

6.2.5 Differently-abled People

The data on differently-abled people were based on the number of handicapped persons or people with a pre-existing medical condition per hectare in the Klang Valley district. The latter are more sensitive to heat because of their prescribed medication, limited mobility, and limited awareness of hot environments. There is also a high degree of agreement in the literature that people with pre-existing health impairments are sensitive to heat.

The differently-abled population in the Klang Valley area is shown in Table 7, where it can be seen that Kuala Lumpur had the highest population of differently-abled people (955), followed by Klang (302), and Gombak (203). Meanwhile, Putrajaya had the lowest differently-abled population in the Klang Valley, where the data from DOSM showed no record of any differently-abled person. This might have been due to the failure of households to register the status of their differently-abled family members with any related authority. Another assumption was that there was no differently-abled people in Putrajaya as most of the population was comprised mainly of government servants.

Table 7 - Differently-abled populations in Klang Valley

District	Area (Hectare)	Differently-abled Population (2016)	Differently-abled Population (per hectare)
Hulu Langat	84000	197	0.002
Petaling	48432	152	0.003
Kuala Selangor	119500	67	0.001
Sepang	19808	36	0.002
Kuala Langat	85800	75	0.001
Klang	62678	302	0.004
Gombak	65008	203	0.003
Hulu Selangor	174000	69	0.000
Sabak Bernam	99710	36	0.000

Kuala Lumpur	24300	955	0.039
Putrajaya	4900	0	0.000

6.3 Statistical Analysis of Variables for Sensitivity

According to the Bartlett's test of sphericity shown in Table 8, the level of significance was much lower than the 0.05 cut-off value for alpha. As such, the null hypothesis was rejected as a minimum of one of the intervariable associations differed significantly from 0.

Table 8 - Kaiser-Meyer-Olkin (KMO) and Bartlett's Test of Sphericity

Sampling Adequacy Measured using KMO		0.655
Bartlett's Test	Approx. Chi-Square	103.401
	Df	10
	Sig.	0.000

As seen in Table 9, PCs 1 and 2 were retained based on their eigenvalues. More specifically, PC1 had an extremely high eigenvalue of 3.919 that accounted for 78.4% of the overall variations while that of PC2 was 1.057, which accounted for 21.1% of the overall variations. These two components captured most of the data, as they were able to explain 99.5% of the overall variance in the data.

Table 9 - Overall variations explained for sensitivity

PC	Original Eigenvalue			Extracted Total of Loadings Squared		
	Total	Var (%)	Cum (%)	Total	Var (%)	Cum (%)
1	3.919	78.379	78.379	3.919	78.379	78.257
2	1.057	21.140	99.518	1.057	21.140	99.440
3	0.015	0.290	99.809			
4	0.007	0.144	99.952			
5	0.002	0.048	100.000			

Table 10 shows the construction of the PCs. PC1 was dominated by the following variables: population density (0.948), minority (0.995), elderly (0.985), and differently-abled individuals (0.982). PC1 could therefore be understood as "social isolation." When combined, these elements generated a weak social network comprised primarily of persons who had difficulty seeking assistance in time of need. Only the extremely young population (0.994) was the major variable in PC2, which might be taken as "dependence." Dependency refers to individuals who rely on others owing to age (children) or social circumstances. The average normalised sensitivity was 0.29 (low on the five-point gauge), while the SD was 0.22.

Table 10 - Rotated component matrix for sensitivity

	Component	
	1	2
PopDens	0.948	0.303
Very Young	0.108	0.994
Minority	0.972	0.224
Elderly	0.985	0.158
Differently-abled	0.982	-0.177

This research provided a quantitative assessment of the sensitivity values for the eleven districts located within the Klang Valley based on the heat sensitivity indicators. These regions were broken up into divisions ranging from extremely poor to extremely wealthy regions. The use of natural breaks, also known as Jenks, is a classification technique that emphasises the similarities and differences between groups. This was to ensure that the differences between neighbourhoods that made adjacent classes; such as between high and very high; were also relatively high. The average sensitivity was 0.29, which is low based on the five-point scale, and a SD of 0.33. The map in Figure 4 shows the sensitivity across the Klang Valley, according to their respective exposure values, as shown in Table 11.

Table 11 - Sensitivity values in the Klang Valley

District	Area (Hectare)	Sensitivity Value
Hulu Langat	84000	0.3532
Petaling	48432	1.0000
Kuala Selangor	119500	0.0303
Sepang	19808	0.2722
Kuala Langat	85800	0.0534
Klang	62678	0.2787
Gombak	65008	0.1355
Hulu Selangor	174000	0.0134
Sabak Bernam	99710	0.0000
Kuala Lumpur	24300	0.8195
Putrajaya	4900	0.3253

Two areas, namely, Kuala Lumpur and Petaling, were ranked with a very high index of sensitivity. These two areas also had a relatively very high exposure value compared to the other areas.

6.4 Adaptivity

The geophysical infrastructure of this investigation was comprised of the road density, NDVI, and geographical elevation. Table 13 displays the geophysical infrastructure data that were gathered from multiple agencies and processed. The average elevation of each state above the mean sea level was acquired using the SRTM data, which can be downloaded from the USGS website. According to the data, the Hulu Selangor district is located 302.9 metres above the mean sea-level. Hulu Selangor is located close to the geographically demarcated Titiwangsa Mountains, known as the Main Range (Banjaran Besar), to the north, which is the major mountain range that forms the backbone of Peninsular Malaysia. In contrast, Klang, located just 8.3 metres above the mean sea-level, is the lowest among all the districts.

The road density factor was generated using the data on roads provided by HERE's maps. The road density is defined as the road length over the total area of a state. According to Table 12, Putrajaya had the densest road network in contrast to the size of Putrajaya itself. Putrajaya is the main centre of organisation for Malaysian government departments. Sepang had the second densest road network, probably due to the location of Putrajaya in the Sepang area. Moreover, the rapid expansion of freeways and expressways have led to a rise in the total length of roadways. In contrast, the road networks in Hulu Selangor were the least crowded. The third factor was the NDVI. The purpose of the NDVI was to determine the vegetation distribution index of the area. It was also an easier way to differentiate between land and water. Hulu Selangor had the highest mean NDVI value of 0.416834, while Kuala Lumpur had the lowest NDVI value of 0.217223. Due to the geographical structure of Hulu Selangor, it was clear that there were fewer land use activities there and that its population density was the lowest among the districts.

In accordance with Kaiser's recommendations, only PC1 (eigenvalue = 1.748, overall variations explained = 58.3%) and PC2 (variance = 1.001, overall variations explained = 33.4%) were maintained a they both preserved most of the structure of the data and accounted for 91.7% of the overall variations. As the other three PCs did not meet Kaiser's criteria, they were not analysed any further. Table 13 shows the total variance explained by the PCs. Following the Varimax rotation in Table 14, the eigenvectors of the geographical elevation (0.94) alone had very high scores in PC1. The NDVI also was the only relevant variable in PC2. The road density variable of (-0.85) and (-0.38) presented lower loads in the first and second PCs, respectively. Thus, the road density variable could not be used to define the class of the factor.

Table 12 - Data adaptive capacity derived from various sources

District	Area (SQ_KM)	Geographical Elevation[1]	Road Density (km/km^2) [2]	Mean NDVI[3]
Hulu Langat	845	225.149	0.4130	0.345376
Petaling	506	42.139	0.8735	0.237051
Kuala Selangor	1205	15.833	0.6689	0.41395
Sepang	559	25.715	1.0429	0.347053
Kuala Langat	861	11.112	0.6423	0.3735

District	Area (SQ_KM)	Geographical Elevation[1]	Road Density (km/km²) [2]	Mean NDVI[3]
Klang	644	8.332	0.7112	0.288294
Gombak	632	220.201	0.5237	0.35468
Hulu Selangor	1761	302.914	0.3021	0.42
Sabak Bernam	1032	11.191	0.8130	0.39
Kuala Lumpur	229	64.556	0.7598	0.217223
Putrajaya	44	45.921	1.4091	0.25

Sources: [1] SRTM data downloaded from USGS; [2] HERE 2010 map data; [3] Landsat 8 OLI downloaded from USGS calculation.

Table 13 - Overall variations explained

PC	Original Eigenvalues			Extracted Total of Loadings Squared			Rotation Total of Loadings Squared		
	Total	Var (%)	Cum (%)	Total	Var (%)	Cum (%)	Total	Var (%)	Cum (%)
1	1.748	58.277	58.277	1.748	58.277	58.277	1.624	54.133	54.133
2	1.001	33.358	91.635	1.001	33.358	91.635	1.125	37.502	91.635
3	.251	8.365	100.000						

Table 14 - Component matrix for adaptivity

	Component	
	1	2
NDVI	0.063	0.983
Road_Dens	-0.854	-0.382
Geo_Elev	0.944	-0.116

The average adaptivity was 0.68, which is moderate based on the five-point gauge, and a SD of 0.25. The map in Figure **4** shows the adaptivity across the Klang Valley according to their respective adaptivity values, as shown in Table 15.

Table 15 - Adaptivity values in the Klang Valley

District	Area (Sq Km)	Adaptivity Value
Hulu Langat	841	0.9591
Petaling	499	0.7230
Kuala Selangor	1195	0.7245
Sepang	558	0.6779
Kuala Langat	859	0.9010
Klang	633	0.5147
Gombak	629	0.9807
Hulu Selangor	1758	1.0000
Sabak Bernam	993	0.4891
Kuala Lumpur	241	0.5381
Putrajaya	44	0.0000

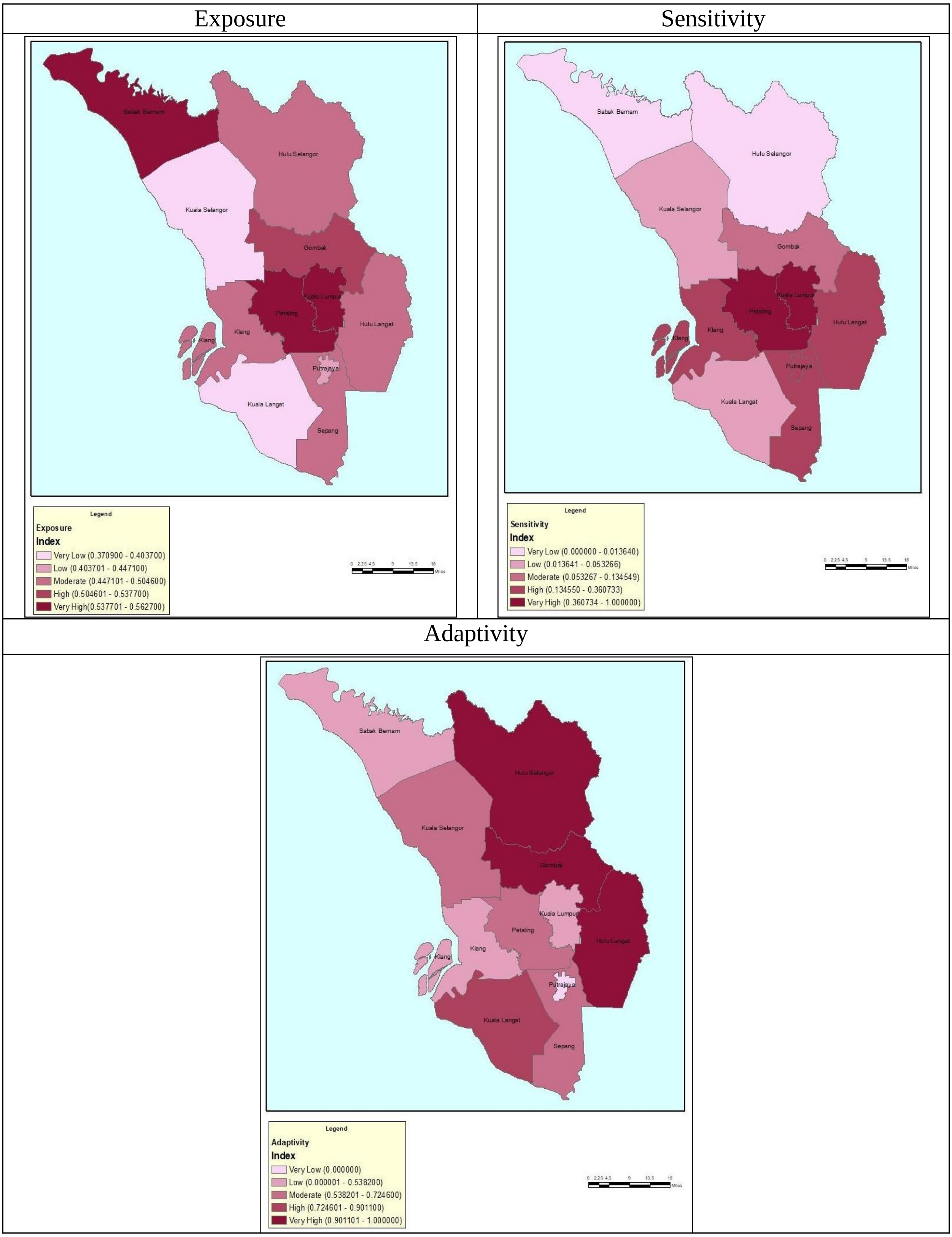

Fig. 4 - Exposure, sensitivity and adaptivity of the study area

6.5 Heat Vulnerability Index (HVI) of the Klang Valley

Iyengar and Sudarshan (1982) developed a technique for constructing a composite HVI from multivariate data, which was used to rank the districts based on their economic performance. This approach is statistically and appropriately suited for the creation of a composite HVI (Hiremath & Shiyani, 2013). Several recent scholarly works have embraced this technique (Chakraborty *et al*., 2019); Hiremath & Shiyani, 2013). The vulnerability index values range between 0 and 1, with a value of 1 indicating that the HVI for the population in that area is very high, and 0 indicating that the HVI for the population in that area is very low.

The average value of the normalised HVI was 0.45, which is considered moderate, with a SD of 0.39. Out of all the eleven areas that were examined, three had high HVI that exceeded 0.6 and two had very high HVI. Kuala Lumpur, which also presented a high level of exposure (0.56), had a very high level of sensitivity (0.84) and a low level of adaptive capacity (0.54). Meanwhile, Petaling had a high level of exposure (0.56) with very high level of sensitivity, and a high level of adaptive capacity (0.72). The ranking of each district according to the HVI is presented in Table 16 and Figure 5. Kuala Lumpur and Petaling were the most vulnerable districts in the Klang Valley with HVI scores of 0.99 and 1.0, respectively, out of 1.0. This was followed by Putrajaya (0.95), Klang (0.56), Sepang (0.44), Sabak Bernam (0.41), Hulu Langat (0.29), Kuala Selangor (0.14), Gombak (0.13), Hulu Selangor (0.00), and Kuala Langat (0.00).

Table 16 - HVI ranks in the Klang Valley

District	HVI	Ranking
Hulu Langat	0.29	7
Petaling	1.00	1
Kuala Selangor	0.14	8
Sepang	0.44	5
Kuala Langat	0.00	10.5
Klang	0.56	4
Gombak	0.13	9
Hulu Selangor	0.00	10.5
Sabak Bernam	0.41	6
Kuala Lumpur	0.99	2
Putrajaya	0.95	3

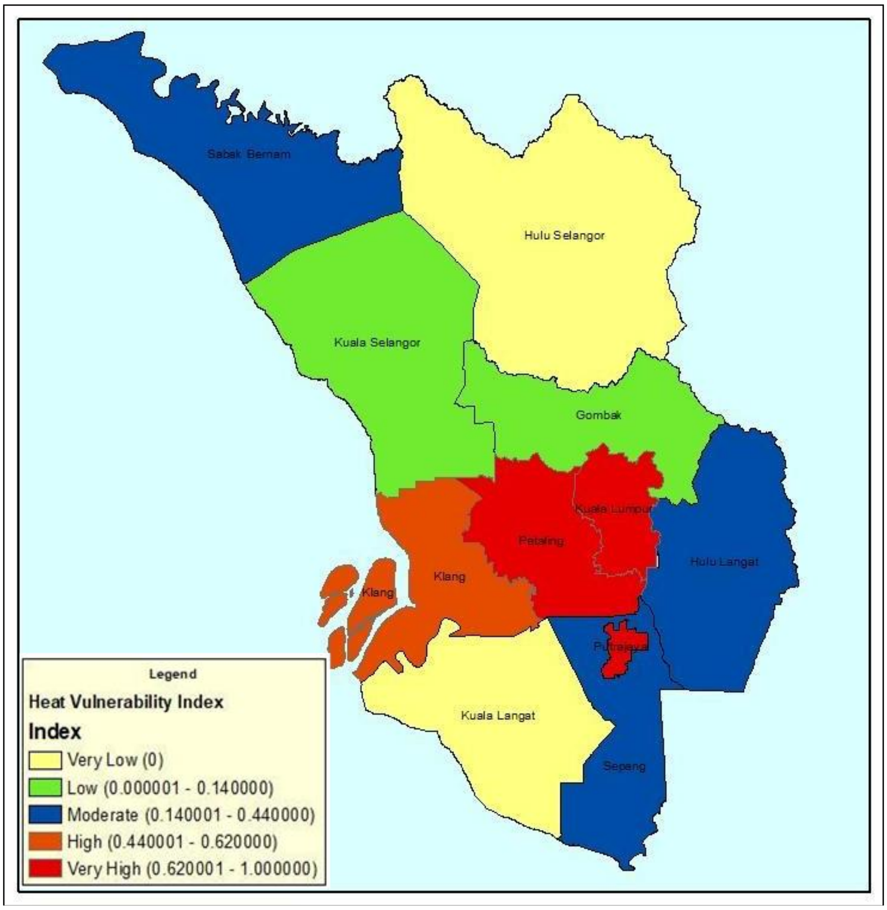

Fig. 5 - HVI map of the Klang Valley

The purpose of developing a vulnerability index from the composite index is to perform a vulnerability assessment. These may be due to the variables of the natural environment, infrastructure, people, society, economy, and adaptive capacity that interact with heat or climate change at the same time. Consequently, it can be adequately reflected by a collection of composite indices. The degree to which each state is vulnerable to climate change is determined using the composite indices, where each state is categorised using the newly-created index based on substantial multivariate data. Vulnerability due to climate change can be very subjective.

The components chosen specifically for this study were climate, natural catastrophes, infrastructure, human, social, economic, and environmental vulnerability. Each component was made up of various sub-indicators.

6.6 Statistical Test of the Developed Indices

The components of the HVI were statistically tested by analysing the Pearson's *r*, and the results of the analysis are shown in Table 17. Not all the variables were significant to each other and had a strong relationship. The Pearson's *r* was more naturally and easily interpreted. It measured the extent to which the components were in agreement with the heat.

Table 17 - Pearson's *r* of the variables

	LST	Population Density	Elderly	Very young	Minority	Differently-abled	NDVI	Road Density (km/km²)²	Geographical Elevation
LST	1.0000								
Population Density	0.4789	1.0000							
Elderly	0.4836	0.9733	1.0000						
Very young	0.2220	0.4116	0.2627	1.0000					
Minority	0.5037	0.9815	0.9950	0.3272	1.0000				
Differently-abled	0.3858	0.8705	0.9349	0.0687	0.9130	1.0000			
NDVI	0.1548	0.1882	0.1028	0.0693	0.0883	0.2118	1.0000		
Road Density (km/km²)²	0.0963	0.2118	0.0386	0.3466	0.0889	0.0060	0.3674	1.0000	
Geographical Elevation	0.0624	0.1284	0.0924	0.1831	0.1175	0.0741	0.0009	0.6524	1.0000

A suitable technique and software were required to obtain the HVI for the Klang Valley. The data from the DOSM and Landsat 8 OLI were used in this research in order to obtain the HVI for all the eleven areas in the Klang Valley. A central goal of this work was to provide a better understanding of the spatial and substantive relationships between temperature and vulnerability across the urban areas. The Klang Valley has been categorised as a developed region in Malaysia. From the findings, it was concluded that the three districts that were most vulnerable in terms of their HVI were Kuala Lumpur, Petaling and Putrajaya. This pattern showed that the populations with a high HVI were usually located in the high-density urban regions, which were characterised by warmer conditions compared to other areas. The districts in urban regions with a low HVI were always characterised by cooler temperatures and a higher vegetation rate.

Nine indicator variables were used in this research to justify the indices, namely, the LST, population density, elderly people, very young population, minority, differently-abled people, NDVI, geographical elevation, and road density. From these nine indicator variables, only two PCs were retained for sensitivity. The first PC consisted of the population density, elderly, minority and differently-abled people. These could be described as a social isolation factor. For the second PC, only the very young population was grouped under the dependency factor. Two PCs were also retained for the adaptive capacity parameter, but road density was not categorised in any of those factors.

In order to develop the HVI, 9 variables were weighted using the PCA, where the sum of the score factors generated was used to calculate the HVI. These variables were obtained from the USGS website relating to the Landsat 8 OLI imagery of the Klang Valley dated 26 March 2016, statistical data from the DOSM, and road density information from the HERE's data. For this study, the main parameters in the HVI were exposure, sensitivity, and adaptive capacity. These three parameters would be related to each other in order to obtain the HVI value. There was a specific indicator variable for each parameter, where the exposure indicator was used for the environmental capacity analysis, the sensitivity parameter indicator was used for the sociodemographic and socioeconomic terms, while the adaptive capacity parameter was used for the human activity and adaptation facilities in their life activities. So, these were important knowledge and concepts that were necessary to conduct the HVI analysis.

All the nine variables were obtained and classed into the exposure, sensitivity, and adaptivity capacity categories. They were then centred and scaled during pre-processing to ensure that the potential of the PCA was at its maximum. A PCA was then used to assign weights to the variables, more specifically, the variance weighted method. From the PCA, the 5 variables representing sensitivity were reduced to 2 factors, while for adaptivity, 3 variables were reduced to 2 factors as well. The sum of the score factors for both sensitivity and adaptivity was normalised using Equation 9, and was further used to calculate the HVI. The average of the LST plus one SD was used for the exposure value.

7. Conclusion and Recommendations

The purpose of this present study was to identify the HVI for urban areas in Malaysia. Through this finding, an investigation of the HVI for people living in urban areas can be carried out so as to finally identify the most vulnerable districts. The target area of this study was the Klang Valley, as it can be categorised as an urban area based on the continuous economic development that goes on there, which is an indication of the urbanisation process. The HVIs for the Klang Valley were developed and mapped to give a clearer image of the levels of vulnerability between the districts spatially. From the HVI map, it could be seen that the most vulnerable district in the Klang Valley was Petaling, followed by Kuala Lumpur and Putrajaya. These three districts can be said to be the busiest developed areas among all the other districts in the Klang Valley, in terms of commercialisation and industrialisation, which attract people mostly for job opportunities and for better settlement. Thus, it can be concluded that with the aid of the GIS and remote sensing data, the HVIs of urban areas can be developed, and these can be a source of information to combat the effects of UHIs.

The following recommendations are given for better understanding and as an aid to improve future studies in relation to the estimations of the HVI and its mapping in the area of interest:

1. This research can be furthered by using another type of satellite imagery from different satellite sensors such as ASTER, SENTINEL and many more since the images for this study were obtained using Landsat 8 OLI satellite images downloaded from the USGS website.
2. Instead of a dimension-reduction PCA, which is an inductive process, other methods, including a deductive process for weightage, can be performed to obtain the HVI for this research to compare the results.
3. Besides the ArcGIS and IBM SPSS Statistics software, further research can be performed using other types of GIS software such as ENVI, ERDAS IMAGINE, and QGIS, and other types of calculation software such as Microsoft Excel XLSTAT, to run the PCA analysis.
4. The health data can be added as a variable in the research to see the association between a person's health and hospitalisation rate with the heat wave.
5. This kind of research can be utilised to predict the future vulnerability of people in a certain area by using the projected data such as population density to help with the early mitigation of incoming disasters.
6. Due to the limited availability of data from the related agencies, the number and types of variables used to represent the vulnerability elements were lacking to describe the association between the variables and the variance of the data. Thus, updated and more detailed data, especially that cover the district, should be provided by the related agencies.
7. This research was only conducted for urban areas. Thus, a study of the HVI can be conducted for rural areas in the future to see the difference in the vulnerability indices between urban and rural areas, and to see whether the UHI phenomenon also occurs in rural areas.

Acknowledgement

The author is gratefully for the financial support received from Universiti Teknologi under the Strategic Research Partnership Grant Scheme (SRP) 100-RMC 5/3/SRP (084/2022) and Universiti Teknologi MARA (UiTM) in collaboration with the Ecosystem Research Development Bureau (ERDB), Philippines for facilitating this research. The Landsat TM data is available on the USGS website.

References

Arsad, F.S.; Hod, R., Ahmad, N.; Ismail, R.; Mohamed, N., Baharom, M.; Osman, Y.; Radi, M.F.M.; Tangang, F. (2022) The Impact of Heatwaves on Mortality and Morbidity and the Associated Vulnerability Factors: A Systematic Review. Int. J. Environ. Res. Public Health, 19, 16356. https://doi.org/10.3390/ijerph192316356

Bai, L., Woodward, A., Cirendunzhu, & Liu, Q. (2016). County-level heat vulnerability of urban and rural residents in Tibet, China. *Environmental Health*, *15*(3), 1-10.

Basu, R. (2009). High ambient temperature and mortality: a review of epidemiologic studies from 2001 to 2008. *Environ Health 8*(40). doi:10.1186/1476-069X-8-40 [Online 16 September 2009]

Basu, R., & Samet, J. M. (2002). Relation between elevated ambient temperature and mortality: a review of the epidemiologic evidence. *Epidemiol Rev*, *24*(2), 190–202.

Campbell, S., Remenyi T. A., White C. J. & Johnston F. H. (2018). Heatwave and Health Impact Research: A Global Review. *Health & Place.* 53, 210-218.

Chakraborty, T., Hsu, A., Manya, D., & Sheriff, G. (2019). Disproportionately higher exposure to urban heat in lower-income neighbourhoods: A multi-city perspective. *Environmental Research Letters*, *14*(10). https://doi.org/10.1088/1748-9326/ab3b99

Confalonieri, U., Menne, B. (Eds.). (2007). Climate Change 2007: impacts, adaptation and vulnerability. Contribution of working group ii to the fourth assessment report of the intergovernmental panel on climate change. Cambridge University Press, Cambridge, UK.

Ebi, K. L., Teisberg, T. J., Kalkstein, L. S., Robinson, L., & Weiher, R. F. (2004). Heat watch/warning systems save lives: estimated costs and benefits for Philadelphia 1995–98. *Bulletin of the American Meteorological Society*, *85*(8), 1067-1074. IPCC (2007) Climate change 2007: impacts, adaptations and vulnerability. Contribution of Working Group II to the Fourth Assessment Report. Cambridge University Press, Cambridge, UK

Hasan, A. R., & Nair, P. L. (2014). Urbanisation and growth of metropolitan centres in Malaysia. Malaysian Journal of Economic Studies, 51(1), 87–101. https://search.informit.org/doi/10.3316/informit.554433876393606

Hansen J, Kharecha P, Sato M, Masson-Delmotte V, Ackerman F, Beerling DJ, et al. (2013) Assessing “Dangerous Climate Change”: Required Reduction of Carbon Emissions to Protect Young People, Future Generations and Nature. PLoS ONE 8(12): e81648. https://doi.org/10.1371/journal.pone.0081648

Hien, W. N. (2016). Urban heat island research: Challenges and potential. *Frontiers of Architectural Research*, *5*, 276–278. https://doi.org/10.1016/j.foar.2016.04.001

Hiremath, D. B., & Shiyani, R. L. (2013). Research Note: Analysis of Vulnerability Indices in Various Agro-Climatic Zones of Gujarat. *Indian Journal of Agricultural Economics*, *68*(902-2016-66825), 122-137.

Inostroza, L., Palme, M., & De La Barrera, F. (2016). A heat vulnerability index: Spatial patterns of exposure, sensitivity and adaptive capacity for Santiago de Chile. *PLoS ONE*, *11*(9), 1–26. https://doi.org/10.1371/journal.pone.0162464

IPCC (2012). Special Report: Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation (SREX). Cambridge University Press, Cambridge, UK and New York, USA.

Isa, N. A., Mohd, W. W., Salleh, S. A., & Ooi, M. C. G. (2018). The effects of green areas on air surface temperature of the Kuala Lumpur city using WRF-ARW modelling and Remote Sensing technique. In *IOP Conference Series: Earth and Environmental Science* (Vol. 117, No. 1, p. 012012). IOP Publishing.

Isa, N. A., Salleh, S. A., WANMOHD, W. M. N., & Chan, A. (2018). Kuala Lumpur city of tomorrow: Integration of geospatial urban climatic information in city planning. *Theoretical and Empirical Researches in Urban Management*, *13*(4), 5-27.

Isa, N. A., Salleh, S. A., Mohd, W. M. N. W., Ooi, M. C. G., Chan, A., & Islam, M. A. (2021). Impacts of Lateral Boundary Condition Resolution in Tropical Urban Climate Modelling for Kuala Lumpur. *Engineering Journal*, *25*(1), 165-175.

Isa, N. A., Wan Mohd, W. M. N., Salleh, S. A., Ooi, M. C. G. & Chan, A. (2020). Land Cover Impacts Towards Thermal Variation in the Kuala Lumpur City. *Journal of Urban and Regional Analysis 12*(1), 91-111.

Iyengar, N. S., & Sudarshan, P. (1982). A method of classifying regions from multivariate data. *Economic and Political Weekly*, 2047-2052.

Jolliffe IT, Cadima J. (2016) Principal component analysis: a review and recent developments. *Phil. Trans. R. Soc. A* **374**: 20150202. http://dx.doi.org/10.1098/rsta.2015.0202

Kovacs, G. H., Belusko, M., Pockett, J. & Boland J. (2016) Assessment of Heatwave Impacts. *Procedia Engineering*, 169, 316-323.

Kosatsky, T. (2005). The 2003 European Heatwaves. European Communicable Desease Journal, *10*(7-9), 148.

Lemonsu, A., Viguie, V., Daniel, M., & Masson, V. (2015). Vulnerability to Heat Waves: Impact of Urban Expansion Scenarios on Urban Heat Island and Heat Stress in Paris (France). *Urban Climate*, *14*(4), 586-605.

Li, Z.-L., Tang, B.-H., Wu, H., Ren, H., Yan, G., Wan, Z., Trigo, I. F., & Sobrino, J. A. (2013). Satellite-derived land surface temperature: Current status and perspectives. *Remote Sensing of Environment*, *131*, 14–37. https://doi.org/10.1016/j.rse.2012.12.008

Mansur, M. & Harun N. Z. (2014) Health Issues and Awareness, and the Significant of Green Space for Health Promotion in Malaysia. Procedia - Social and Behavioural Sciences 153 (2014) 209 – 220.

Ministry of Housing and Local Government. (2006). *National Urbanization Policy, Malaysia*. National Weather Service. (2016). What is the heat index ?, 1–2.

Mölders, N. (2011). Land-Use and Land-Cover Changes: Impact on Climate and Air Quality. *Springer Verlag Journal.* (Vol.44).

Nayak, S. G., Shrestha, S., Kinney, P. L., Ross, Z., Sheridan, S. C., Pantea, C. I., … Hwang, S. A. (2018). Development of a heat vulnerability index for New York State. *Public Health*, *161*, 127–137. https://doi.org/10.1016/j.puhe.2017.09.006

New York State Department of Health. (2014). *Heat Vulnerability Index.* https://www.health.ny.gov/environmetal/weather/vulnerability_index/

Niu, Y., Li, Z., Gao, Y., Liu, X., Xu, L., Vardoulakis, S., Yue, Y., Wang, J., & Liu, Q. (2021). A Systematic Review of the Development and Validation of the Heat Vulnerability Index: Major Factors, Methods, and Spatial Units. *Current Climate Change Reports*, *7*(3), 87–97. https://doi.org/10.1007/s40641-021-00173-3

Nwankwoala, H. N. L. (2015). Causes of Climate and Environmental Changes: The Need for Environmental-Friendly Education Policy in Nigeria. *Journal of Education and Practice*, *6*(30), 224–234.

O'Neill MS, Carter R, Kish JK, Gronlund CJ, White-Newsome JL, Manarolla X, et al. Preventing heat-related morbidity and mortality: new approaches in a changing climate. Maturitas. 2009;64(2):98–103

Ooi, M. C. G., Chan, A., Kumarenthiran, S., Morris, K. I., Oozeer, M. Y., Islam, M. A., & Salleh, S. A. (2018). Comparison of WRF local and nonlocal boundary layer Physics in Greater Kuala Lumpur, Malaysia. In *IOP Conference Series: Earth and Environmental Science* (Vol. 117, No. 1, p. 012015). IOP Publishing.

Praene, J. P., Damour B. M., Radanielina M.H., Fontaine, L., Riviere G. (2019). GIS-based Approach to Identify Climatic Zone: A Hierarchical Clustering on Principal Component Analysis. *Building and Environment.* 164, ISSN 0360-1323

Reid, C. E., O'Neill, M. S., Gronlund, C. J., Brines, S. J., Brown, D. G., Diez-Roux, A. V., & Schwartz, J. (2009). Mapping Community Determinants of Heat Vulnerability. *Environmental Health Perspectives*, *117*(11), 1730–1736. https://doi.org/10.1289/ehp.0900683

Reid, C. E., Mann, J. K., Alfasso, R., English, P. B., King, G. C., Lincoln, R. A., Margolis, H. G., Rubado, D. J., Sabato, J. E., West, N. L., Woods, B., Navarro, K. M., & Balmes, J. R (2012). Evaluation of a Heat Vulnerability Index on Abnormal Hot Days: An Environmental Public Health Tracking Study. *Environmental Health Perspectives*, *120*(5), 715–720. https://doi.org/10.1289/ehp.1103766

Revich, B. A. (2011). Heat-wave, air quality and mortality in European Russia in summer 2010: preliminary assessment. *Yekologiya Cheloveka/Human Ecology*, 7, 3–9.

Robin, J. M., Cheung, S. L. K., Le Roy, S., Van Oyen, H., Griffiths, C., Michel, J. P., & Herrmann, F. R. (2008). Death toll Exceeded 70,000 in Europe During the Summer of 2003. *Comptes Rendus Biologie*, *331*(2), 171–178.

Romero-Lankao, P., Qin, H., & Dickinson, K. (2012). Urban Vulnerability to Temperature-Related Hazards: A Meta-Knowledge Approach. *Global Environmental Change*.

Santamouris, M. (2015). Analysing the heat island magnitude and characteristics in one hundred Asian and Australian cities and regions. *Science of the Total Environment*, *512–513*, 582–598. https://doi.org/10.1016/j.scitotenv.2015.01.060

Schmidtlein, M. C., Deutsch, R. C., Piegorsch, W. W., & Cutter, S. L. (2008). A sensitivity analysis of the social vulnerability index. *Risk Analysis: An International Journal*, *28*(4), 1099-1114.

Shuid, S. (2004). Urbanization and Housing in Kuala Lumpur City Centre : *Earoph World Planning and Housing Congress*, 3–4.

Tate, E. (2012). Social vulnerability indices: a comparative assessment using uncertainty and sensitivity analysis. *Natural Hazards*, *63*(2), 325-347.

Tate, E. (2013). Uncertainty analysis for a social vulnerability index. *Annals of the association of American geographers*, *103*(3), 526-543.

Thurstone, L. L. (1947). Multiple-factor analysis; a development and expansion of The Vectors of Mind.

Zeeshan, M., & Ali, Z. (2022). The Potential of Cool Materials Towards Improving Thermal Comfort Conditions Inside Real-urban Hot-humid Microclimate. *Environment and Urbanization ASIA*, *13*(1), 56-72.