

Article

Combining Measurements of Built-up Area, Nighttime Light, and Travel Time Distance for Detecting Changes in Urban Boundaries: Introducing the BUNTUS Algorithm

Muhammad Luqman^{1,*}, Peter J. Rayner¹ and Kevin R. Gurney²

- ¹ School of Earth Sciences, University of Melbourne, 3000 Melbourne, Australia; prayner@unimelb.edu.au
- ² School of Informatics, Computing and Cyber Systems, Northern Arizona University, Flagstaff, AZ 86011, USA; kevin.gurney@nau.edu
- * Correspondence: mluqman@student.unimelb.edu.au

Received: 27 October 2019; Accepted: 6 December 2019; Published: 11 December 2019



Abstract: This paper introduces a new algorithm (BUNTUS—Built-up, Nighttime Light, and Travel time for Urban Size) using remote sensing techniques to delineate urban boundaries. The paper is part of a larger study of the role of urbanisation in changing fossil fuel emissions. The method combines estimates of land cover, nighttime lights, and travel times to classify contiguous urban areas. The method is automatic, global and uses data sets with enough duration to establish trends. Validation using ground truth from Landsat-8 OLI images revealed an overall accuracy ranging from 60% to 95%. Thus, this approach is capable of describing spatial distributions and giving detailed information of urban extents. We demonstrate the method with examples from Brisbane, Australia, Melbourne, Australia, and Beijing, China. The new method meets the criteria for studying overall trends in urban emissions.

Keywords: remote sensing; urban areas; climate; google earth engine; machine learning; land cover; nightlight

1. Introduction

The world is urbanizing rapidly. The United Nations Department of Economic and Social Affairs [1] reported that 55.3% of the world population now lives in urban areas. They projected that the world's urban population will reach 60% by 2030 and 67% by 2050. Urban areas are the hub of economic activities and, consequently, of fossil fuel emissions [2]. The relationship between urban size as well as emissions is, however, complex and contested. Bettencourt and West, (2010) [3] found a relationship between urbanisation and the economy. They stated that it requires only 85% more infrastructure when one doubles the population in an urban area. The authors of reference [4] also found the decreasing trend of carbon emissions intensity in highly dense urban areas (100 metropolitan cities) of the United States. The Intergovernmental Panel on Climate Change (IPCC) Fifth Assessment Report concluded that urban areas generate the majority of carbon emissions from final energy use [5,6]. However, it is not well understood how carbon footprints are distributed among cities. Any global study of such relationships requires a consistent definition of urban extent, which is the task of this paper.

No single definition of urban extent will suit all administrative, policy or research needs. It is recommended that the methodology used for delineating cities should be tailored to its application.

Even within the literature, different disciplines require different definitions. In urban ecology, cities are defined on a qualitative basis—for instance, the areas under the influence of humans [7,8].



In social science, the areas with high population densities are regarded as urban or city areas [8]. Economics defines the city by its political, cultural, and economic characteristics [9].

This work forms the first part of a study of the global distribution of trends in urban fossil fuel emissions and related quantities. That study requires a definition of urban extent so we can attribute emissions as urban or not. The needs of the overall study provide some important requirements for the algorithm we use for the urban extent. We will use these so often throughout the paper that we label them here:

- **R1** The study is global, so we may only choose globally consistent and available datasets.
- **R2** We wish to study enough cities to establish patterns, so the algorithm must be computationally efficient enough for large-scale use.
- **R3** Changes in urban extent can be small, so the algorithm must work at high resolution, no more than 1 km.
- **R4** The study must be long enough to establish trends. We estimate this requires two decades of data.

In this paper, we propose a novel and multiple-step approach which satisfies these requirements. We define an urban area as a contiguous (i.e., simply connected) and compact region including a pre-defined urban center and which satisfies several criteria relating to density and surface properties. The urban boundary is the bounding polygon for this region. The compactness requirement means that gaps such as green belts (but not water) should be included in the urban area. The criteria which define the urban area must be deducible from datasets satisfying R1, R3, and R4.

The outline of the paper is as follows. In Section 2, we review existing methods for determining urban extent, focusing on their utility for our task. In Section 3, we describe our methodology and its underlying data sets. In Section 3, we also present available validation for the method and three case studies of different cities. In Section 4, we validate and discuss the results with a particular focus on the uncertainties of the method and summary of the main results. Section 5 comprises of discussion and Section 6 comprises of conclusion.

2. Review of Existing Methods

We use two classes of remotely sensed data to delineate cities: built-up area and, nighttime lights. We also use travel time based on openly available street maps. Several authors have used one of these. Jaeger et al. (2010), Kasanko et al. (2006), Poelmans and Van Rompaey (2009), and Zhang and Seto (2011) [10–13] used measures of built-up area to study urban expansion, while Schneider and Woodcock (2008), and Schwarz (2010) [14,15] also analysed urban form. Goldblatt et al. (2016), Parés-Ramos et al. (2013), and Small et al. (2005) [16–18] used nighttime lights to delineate cities. Finally, Mindali et al. (2004), Niu et al, (2018), Sun et al., (2016), and Xie and Levinson (2007) [19–22] used travel time with other datasets as a measure of urban extent.

Before describing BUNTUS in detail, we will briefly review some previous work that approaches similar problems. This will establish the need for a new approach, motivate our choice of validation datasets and provide background on our choice of algorithm.

We have noted 12 urban area products with global coverage, using built-up area or human settlements, based on remotely sensed data and freely available. These are summarised in Table 1 and described below. In particular, we note the requirements listed above which each dataset does not meet.

Product Name	Map	Group	Temporal Resolution	Spatial Resolution	Availability
HYDE3.2 [23]	History database of the global environment v3.2		10,000 BCE to 2015 CE	5 arc min	Open Source
GMIS [24]	Global Man-made Impervious Surface (GMIS) Dataset from Landsat, v1 (2010)	Human settlement or built-up oriented products	2010	30 m	Open Source
HBASE [25]	Global Human Built-up and Settlement Extent	_	2010	30 m	Open Source
IMPSA [26]	Global impervious surface area		2000–2001	30 arcsecond	Open Source
GUF [27]	Global Urban Footprint		2011	0.4 arcsec (~12 m)	Free for scientific study
GHSL [28]	Global Human Settlement Layer		1975, 1990, 2000, 2014	38 m to 1 km	Open Source
GRUMP [29]	Global rural-urban mapping project (alpha version)		1995	30 arcsecond (vector)	Open Source
VMAP0 [30]	Vector Map Level Zero		2000	1:1,000,000 scale (vector)	Open Source
MODIS [31] 500 m	MODIS urban land cover (2001v4)	Land cover products consist of built-up and other classes –	2001–2017	500 m	Open Source
CCI [32]	Climate Change Initiative (CCI) Land Cover		1992–2015	300 m	Open Source
GLC30 [33]	Global Land Cover 30 m		2000, 2010	30 m	Open Source
GLC [34]	Global Land Cover	_	2015	100 m	Open Source

 Table 1. Global land cover and urban area products.

GRUMP and VMAP0 are provided as vector files, while the rest are rasters. The highest resolution GUF and GHSL are based on radar measurements while the rest are based on optical imagery, with a resolution of 0.4 arcseconds. VMAP0, GLC30, MODIS 500, GLC, and CCI are multi-class land cover maps that include an urban class. HYDE3 and IMPSA portray urban land as a persistent variable, as the part of the impenetrable surface and the division of urban land, individually. Each of these products has to solve a common set of problems: the ambiguity of the observable attributes of cities, sporadic coverage (usually due to cloud obscuration), and arbitrary thresholds for quantities in finite-sized pixels. Many products use strategies combining satellite imagery with ground-based results and Geographic Information System (GIS) information layers.

GRUMP urban areas [29] are polygons bounding contiguous regions of nighttime lights. These are combined with estimated metropolitan boundaries extracted from a pre-existing settlements map to account for lack of electrification. GRUMP data approximate the city size by buffering on the points of the settlements. The "bloom" of lights in deserts is the major reason for the overestimation of urban areas [35]. GRUMP has an effective resolution of 1km and is tied to the year 1995. This makes it difficult to use for trend analysis.

MODIS 500 draws on coarse-resolution multispectral MODIS images from the Terra and Aqua platforms for the whole year. It yields the highest accuracy in land cover classification for the year 2009 [36] among all its years. Its effective resolution is 500 m, and it is referenced to 2001.

In 2010, China launched a worldwide mapping project for land cover to generate GlobeLand30 land-cover maps for two decadal epochs (2000 and 2010). The 2010 land cover maps of the Earth were provided to the United Nations as support for worldwide practical improvements and fighting environmental change. GlobeLand30 2010 has the same 30 m spatial resolution as the Landsat sensor [33]. GlobeLand30 has sufficient spatial resolution to see small changes in urban extent from year to year but its temporal resolution is too coarse to establish trends.

TerraSAR-X is an imaging radar launched by the German Aerospace Agency (DLR) in 2007 [27]. In 2016 TerraSAR-X was used to produce the Global Urban Footprint (GUF) [37], a worldwide map of human settlements at 12 m resolution for the reference year 2011. GUF depends on the examination of 182,249 single look complex pictures procured with 3 m ground resolution mostly during 2011 and 2012. GUF has the highest spatial resolution of any current global product but is limited to a single reference year.

The Global Human Settlement Layer (GHSL) project of the Joint Research Center of the European Commission is a general system for completely automated analysis of optical or radar imagery. It has so far been applied to optical imagery from 0.5–80 m resolution and 10–20 m resolution for radar measurements. These can be combined to map built-up surfaces and are produced using a completely computerized image processing technique [28] namely Symbolic Machine Learning (SML) directed classifier [28]. The GHSL methods were also applied to Landsat imagery to produce the built-up zone network made available in 2016. It is generated by re-processing the archive of Landsat imagery with the GHS Sentinel-1 built-up framework and Globe-Land30 2010 as the training data set. The GHSL is limited to epochs 1975, 1990, 2000, and 2014.

ESA CCI Land Cover (v2.0.7) is a project of the European Space Agency (ESA). It generated Annual maps of land cover (including an urban class) from 1992 to 2015 with a spatial resolution of 300-m [32]. The 'urban area' class which is a part of the chosen map for land cover for the year of 2015, has been generated by following the classification standards of the GUF [37] and GHSL [38].

The Global Human Built-Up and Settlement Extent (HBASE) dataset [25] was generated from Landsat imagery for 2010 as the reference year. HBASE has been generated by post-processing the Global Manmade Impervious Surface (GMIS) dataset which approximates fractional impermeable cover worldwide extracted using the Global Land Survey 2010 [25]. HBASE maps the occurrence of human settlements and developed regions.

The above products have a range of strengths and weaknesses. None, however, combined the temporal and spatial resolution we need for our future study of emissions trends. This provides the

motivation for the methodology described and tested in the rest of this paper. Many of the previous products are probably more refined for their chosen reference years than our method. These will provide data for comparison and testing.

We also note that most of the above products do not solve all of our problems. They classify, pixel by pixel, the land cover type at some spatial and temporal resolution. Solving the problem of which of these pixels to assign to a given city is different from the original classification. As we will see, this makes up the majority of the effort and always involves combinations of data sets.

3. Methodology

The flow chart of the adopted methodology is depicted in Figure 1. The empirical framework for delineating urban areas contains three steps—land cover generation, nightlight classification, and travel time calculation. The steps consist of processing an individual dataset. Each of these datasets has some advantages and disadvantages. For example, Landsat provides the free optical remote sensing imagery of medium resolution (30 m). Surface classification using Landsat is limited by cloud obscuration, the similarity of radiance signatures among surface types and incomplete training data [39]. Another problem with land cover is the occurrence of gaps when we focus on one class (built up in the present study). Unlike Landsat, the Nighttime Light (NTL) products do not show any gaps over a city area. However, due to their large pixel size, NTL products provide ambiguous information at the edges of the cities and on coastlines. Travel time provides evidence on the connectivity of a city but not whether this is associated with population so it cannot define urban extent alone. If combined judiciously, the three datasets can provide a more reliable urban boundary estimate than any of them alone. The algorithm is named BUNTUS (Built-up, Nighttime Light, and Travel time for Urban Size). In what follows we show the study area and explain each step separately and their fusion into a final product.



Figure 1. Methodology flow chart of fusion of Landsat, Nighttime light product of Defense Meteorological Satellite Program (DMSP) and Visible Infrared Imaging Radiometer Suite (VIIRS) and travel time.

3.1. Case Studies

In this pilot study, we chose three cities (see Figure 2) that test different aspects of BUNTUS. Beijing has been chosen due to its rapid growth, including agglomeration, Melbourne due to rapid growth, and Brisbane due to its unusual urban form of a long coastal strip.



Figure 2. Study area map.

3.2. Land Cover Classification

We face three significant problems when establishing a land cover classification for the duration of 21 years at global scale.

- 1. Dealing with the inevitable instabilities arising from different instruments and conditions;
- 2. A lack of trustworthy and independent data points for training the classification;
- 3. The computational demands of classification at this resolution.

We are, of course, not the first to face these challenges—the authors of references [16,40–42] have all suggested solutions.

For point 1, we take advantage of the multiple well-calibrated images available for the same location, reducing both calibration and cloud contamination problems. A more refined option would be to combine the same set of Landsat scenes by computing a "cloud score" extracted using the F-mask (Zhu and Woodcock, 2012) cloud detection algorithm.

The second challenge is best addressed by exploiting local existing or global and coarse resolution maps (e.g., the already cited Globcover map [43] to automatically select significant locations for the training data. Additionally, classification by ensemble learning, i.e., by means of the joint use of multiple classifiers, is one approach to improve robustness.

For point 3, we are indebted to the Google Earth Engine (GEE). It provides both an interface to a parallel architecture for large-scale processing and to the entire LANDSAT archive.

For the land cover mapping, Landsat Thematic Mapper (TM), Landsat Enhancement Thematic Mapper (ETM), and Landsat Operational Line Imager (OLI)'s Top of Atmosphere (TOA) products have been used. TOA products are radiometrically calibrated by following the radiometric calibration coefficients defined by Chander et al. (2009) [44] and available at GEE. So, instead of generating our own radiometrically calibrated Landsat product, we utilized the already available dataset. As the aim is to extract almost cloud-free collections for a specific period of interest, we adopted the Hu and Hu, (2019)'s [45] methodology and applied *Algorithms.Landsat.simpleCloudScore* function to get the cloud distribution probability score (0–100) of selected images for each year. Then we applied the *image.updateMask* function to remove the cloudy regions with cloud score greater than 40. After the removal of clouds, we applied a median *ee.Reducer* function to the collection of images having unmasked pixels to reduce the image collection to a single output image representing the median values. This generated cloud free composites of images from 1998 to 2018.

Before constructing a classifier, we need to choose a training data set consistent with our goals of resolution, time-span, and global coverage. We used the Moderate Resolution Imaging Spectroradiometer (MODIS) Terra and Aqua reflectance monthly land cover product (known as MCD12Q1.006). MCD12Q1 version 6 land cover data has been derived from six classification schemes and available at an annual interval from 2001 to 2017 [31]. For the years 1998–2000, we used the 2001 values for training while for 2018, we used 2017 values. Using different years for training does not affect much the accuracy of classification as landcover varies much more in space than one year to the next.

After the selection of a Landsat TOA product and training sites, the next step was the selection and application of a classifier. Achieving high accuracy in land cover classification is challenging, particularly in urban areas because of heterogeneity in landscape and mixed pixels [46,47]. For example, bare soil and sand are often confused with the concrete pixels because of similar reflectance [48,49]. To deal with such problems, many land cover classification algorithms such as Decision Trees, Artificial Neural Networks, Support Vector Machines, and Random Forest are used in remote sensing (Akar and Güngör, 2012; Deilmai et al., 2014). We used the Random forest [50] which has been used recently in a large number of studies [51–53] due to its ability to handle high data multicollinearity and dimensionality efficiently [54]. The model works by randomly picking a sample of observations and a sample of variables (a few hundred to several thousand depending on the size and nature of the training set) with replacement to produce a large number of individual decision trees [55]. The classified image consisted of 16 classes (the same as the classes of MCD12Q1) of land cover at 30 m spatial resolution. The code for the land cover classification is available at https://code.earthengine.google.com/f5974440d4a895f6f97edf9c6af4b780.

3.2.1. Accuracy Assessment

A confusion matrix [56] was employed to assess the accuracy of land cover generated through classification. A confusion matrix is a summary of prediction results on a classification problem. The element i, j in a confusion matrix, shows the proportion of true class I that is classified as class j. The identity matrix indicates perfect classification. We applied the *validated.errorMatrix* function at the Google Earth Engine platform to get the accuracy of land cover. A commonly acceptable accuracy limit for land cover maps derived from the classification of satellite data is 85% [57,58]. The accuracy does not only depend upon the input and classifier, but it also depends upon the quality of the satellite imagery (e.g., radiometric calibration and cloud obscuration). The Landsat imagery in this study had little cloud obscuration (except Melbourne for 2005 and 2010). We tested the land cover classification of 16 classes and the average of minimum accuracy values for all cities were 75%.

3.2.2. Urban Area Generation

The generated land cover had 16 classes. The class type and selection standards are defined by Sulla-Menashe and Friedl, (2018) [31]. There were two reasons for the selection of 16 classes. First, we wished to use the MCD12Q1 as is. Second, a higher number of classes minimizes the effect of mixed pixels. We remapped the (16 classes) into two classes—built-up and non-built-up. The built-up was assigned a value of 1, and everything else was reclassified as non-built up and assigned a value of 0. As our main focus was the built-up area, we again assessed the accuracy of two classes on Google Earth. Our two-class classification was more than 90% successful in all datasets. Next, we needed to deal with small gaps in built-up areas (e.g., parks, water bodies) which arise from the high resolution of our data. A neighborhood circle of one square kilometer was applied to the built-up class to filter-out the contiguous and non-contiguous area by using the focal statistic tool of ArcMap. The tool performed a neighborhood operation to compute the values of each output cell around a specified cell after the application of cells window. The output represented the urban density ranging from 0 to 1, 0 for others and 1 for completely urban area. The output was multiplied with the built-up class to produce built-up and non-built-up class and calculated the new urban area. A connectivity raster was created between the cells of the same value which were connected to each other in any direction. The final output was a raster with minimum gaps to be used in the fusion step. The final output had two classes: built-up (1) and non-built-up (0) at a spatial resolution of 30 m.

3.3. Nighttime Light Data Processing

Nighttime light (NTL) satellite images are a class of remote sensing product, globally available for multiple years. The global open source night light images are available from two sensors; The Operational Linescan System (OLS) instrument mounted on the United States Airforce Defense Meteorological System Program (DMSP) and the Visible Infrared Imaging Radiometer Suite (VIIRS) sensor mounted and flown on NASA/NOAA Suomi-National Polar-Orbiting Operational Environmental Satellite System (SNPP). The OLS images are processed by NOAA's National Geophysical Data Center (NGDC) and are available from 1992 to 2013 as annual composites. The OLS product has two spatial resolution—"fine" data have 0.56 km, and "smooth" data have 2.5 km spatial resolution (Huang et al., 2014). DMSP/OLS suffers from poor intercalibration. It also saturates easily (having only six bits of dynamic range). VIIRS is considered superior to DMSP-OLS product [35] because of high spatial, temporal, and radiometric resolution which are 742 m, monthly composites, and 12 bits, respectively. The VIIRS product for stable lights is already intercalibrated and needs no further processing. We consequently use different strategies for DMSP/OLS and VIIRS data.

DMSP-OLS sensors capture artificial lighting from the earth surface [18,26,59] which can be utilized to determine the boundary of urban territories. The image product of DMSP-OLS sensor stores the reflectance in the form of Digital Numbers (DNs). The DN value is higher in the areas where the night time light is higher and vice versa. This product can help to distinguish between urban and non-urban areas on the basis of DNs. A pixel of this product is viewed as urbanized if its extent surpasses a limit [18]. A cluster of same DNs would represent a land cover over [60–62]. However, the derivation of land cover by using NTL information is sometimes incorrect, particularly in less dense urban territories [13]. DMSP-OLS can likewise misrepresent the extent of urban regions [18,60] while disregarding small or emerging settlements. Also, the extent and intensity of lit regions cannot straightforwardly delimit urban areas due to the "blooming" impact [63,64]. "Blooming" alludes to the appearance of lit territories as bigger than the settlements they are related with [18].

To overcome the blooming and saturation effect we utilized reference [65]'s intercalibration parameters for DMSP-OLS annual composites (see Figure 3), and to identify the urban areas, we used the multivariate analysis technique of Parés-Ramos et al. (2013) [17] to find the threshold DN for urban regions. For the multivariate analysis, we classified the NTL imagery into five classes to establish the relationship between DN and land cover. We classified the DMSP and VIIRS datasets at their native resolutions of 2.5 km and 742 m, respectively, into five classes and overlayed on Google Earth Imagery and Landsat imagery (see Figure 4). After several tries, we found the threshold as shown in Figure 4. The land cover of each threshold has been shown with the help of arrows in Figure 4. Our focus was on the extraction of urban area from the NTL, so we checked the pixel values of NTL at the urban areas. Using these thresholds, we reclassified the DMSP/OLS data into two classes (1) PV

0-45 non-urban (value 0), (2) and PV 45.01–64, urban (value 1). We tested these thresholds by direct visualisation of Melbourne, Brisbane, and Beijing. Our objective requires a consistent methodology, so we used the same threshold for all cities. We classified the VIIRS data into two classes: (1) PV -0.89-5.2, no development (value 0), (2) PV 5.3–10,000, urban (value 1).



Figure 3. Classification of Nighttime light product of Defense Meteorological Satellite Program Operational Linescan System (DMSP-OLS) intercalibrated image of Brisbane.



Figure 4. Overlay of nighttime light product of Defense Meteorological Satellite Program Operational Linescan System (DMSP-OLS) classification of Brisbane on high-resolution satellite imagery for sensitivity test.

Our classification captures the fine boundaries of built-up areas with high precision (Figure 4). However, during visualization, we noticed mixed pixels at the edges of the urban class. To minimize this problem, we resampled the two-class night time light datasets to a spatial resolution of 30 m. We later overlayed with Landsat's urban area, enforcing the correct surface type at a resolution of 30 m.

3.4. Travel Distance Raster Creation

The road network provides a third view of urban extent since it measures the connectivity of a space. Complete and accurate geospatial road network data is, therefore, a valuable dataset. Open Street Maps (OSM) provides the road network geospatial data at global level at no cost. In the present study, OSM planet road data has been utilized, which was acquired from http://download.geofabrik.de/.

The idea behind the inclusion of the travel time was to check the contiguity and connectivity of the urban areas and to include the emissions economically connected to the city by travel. The motivation for including travel time in the present study follows Day et al. (2017) [66]. Day et al. (2017) [66] calculated the urban footprints of Sri Lanka for the year 2000 by using travel time, and the population in their gravity model. Because of the lack of historical travel time polygons, it was difficult for us to define long-term urban area boundaries by using travel time alone. Rather, we combined it with other datasets.

Road network datasets were downloaded as shapefiles of line features. The extent of the region to download was based on the largest urban image from landcover classifications after applying a 15-km buffer. The road network shapefile was processed in ArcGIS 10.5.1 by using the network analysis toolset. An average speed of 45 km per hour (km/h) was chosen for each road network. The length of each road was calculated in kilometers. Once the road length (in km) and speed (45 km/h) was calculated, we calculated the travel time (TT) for each road in minutes by using the following formula:

$$\Gamma T_{(Minutes)} = 60(Length/Speed).$$
 (1)

The district center point (taken from the ArcGIS Hub [67]) was chosen as the facility point and the travel time from that point was calculated at the interval of 10 min. The output was a service area of equal time. Hence multiple polygons were generated with 10 min interval as shown in Figure 5. The travel time depends on the extent of the city and density of the roads. The travel time output was in the form of polygons, i.e., a vector data. We converted the vector data into raster at a spatial resolution of 30 m, to merge with the other two datasets. Since NTL and Landsat datasets were in classes, we reclassified the travel time rasters by taking the reciprocal of the values and multiplying by 10. The reason for taking the reciprocal of travel time was to give more weight to the area close to the urban core. For example, the 10 min travel time value resulted in 1 (10*1/10 = 1) and 20 min travel time resulted in 0.5 (10*(1/20)) and so on. So, the travel time rasters were float rasters with the values 0-1.



Figure 5. Travel time distance polygons of Beijing, Brisbane, and Melbourne.

3.5. Fusion of Datasets

Once three rasters were produced, we proceeded to merge those rasters. A simple sum of three rasters was generated according to

$$DN_{out} = DN_B + DN_{NTL} + DN_{TT}.$$
(2)

This summed raster takes the values from 0 to 3 (see Figure 6). Our final step is to classify an area as urban using a threshold on this summed raster. By visualizing many rasters on Google Earth Imagery we chose a threshold of 1.5. This thresholding may generate islands of urban areas which were not contiguous to the urban core. We selected the largest contiguous urban area which contained the urban core and converted it into a vector data polygon.

One significant simplification is the final decision rule itself. Inspection of Equation (2) shows it can be replaced with a decision tree. If built-up and nightlights agree they determine the value, otherwise the point is considered urban if travel time is \leq 50 min. This may mean that the travel time information is rarely active. We can see this by considering a map of points in which built-up and nightlights disagree and superimposing the travel-time 50-min polygon. Little of the peri-urban area lies inside the polygon. The travel-time metric does, however, serve to in-fill gaps within the city such as large parks which might be excluded otherwise. It does not remove all gaps within the urban boundary since, for many of these, built-up and nightlights may agree that the pixel is not urban. This can be addressed by defining the smallest polygon bounding all urban points.



Figure 6. Fusion of three datasets (Melbourne in 2018). Plate (**A**) represents the travel time raster (DN_{TT}) , plate (**B**) represents the built-up area (DN_B) , plate (**C**) represents the Nighttime Light classified raster (DN_{NTL}) , plate (**D**) represents the fusion of all three datasets (DN_{out}) , plate (**E**) represents the classified sum with a threshold value 1.5, and plate F represents the largest contiguous urban area boundary.

3.6. Processing Time and Equipment

The processing time of each dataset depends upon the extent of the city. For land cover classification, processing time depends upon the number of training sites and availability of Google Earth Engine. Also, processing time varies with the usage of satellite sensors. For example, it took 8–10 min to process TM and ETM images of Melbourne but took more than 15 min in case of OLI against 1200 training sites for a single year. Similarly, the processing time of the network dataset varies with the size of the city and number of line features. For nighttime light rasters processing and travel time data development, ArcGIS 10.7.1 batch processing has been used on Lenovo X250. The computational cost of combining three datasets is bearable because the BUNTUS algorithm processes the subsets of all three datasets and combine them as a single band raster.

4. Results

4.1. Validating BUNTUS

Direct quantitative validation of BUNTUS is difficult. As we pointed out in Section 2, no other dataset has the time and space resolution for complete comparison. More seriously, the definition of urban extent is arbitrary. Our task is to capture urban dynamics. Thus, we can compare magnitudes of rankings of changes from BUNTUS and other datasets.

The most direct, if labour-intensive, comparison is with direct urban imagery. Figure 7 shows BUNTUS boundaries of Melbourne from 1996 to 2016 overlaid on Google Earth imagery. For the year 2017 and 2018 Google Earth did not have imagery at the time of writing. We see agreement in both extent and form between Google Earth imagery and BUNTUS estimates. Comparison for Beijing and Brisbane indicate similar performance.



Figure 7. Built-up, Nighttime Light, and Travel time for Urban Size (BUNTUS) boundaries of Melbourne on Google Earth Imagery from 1998 to 2016.

4.2. Comparison

In Figure 8, we have compared the boundaries of our urban area with some of the world's freely available urban area boundaries. Fine resolution satellite imagery was acquired from Google Earth for the year 2010. The urban area boundaries generated in the present study has been compared (visually) with GRUMP (vector data for urban settlement), MODIS, Global Cover (both Land Cover with multiclass), HYED3.2 (multi-class), GMIS, and HBASE (both urban settlement datasets) for the year 2010. We used GRUMP's 1995 data as it is not available for 2010. It is clear that the GRUMP 1995 urban area is more extended than BUNTUS. Likewise, the MODIS, HYED3.2, and Global Cover are showing a rough urban area boundary due to their coarser spatial resolution. GMIS and HBASE look similar to BUNTUS in their extent, shape and urban pattern. However, both GMIS and HBASE contain gaps within their urban areas, nor do they consist of a contiguous area to the urban core.

From Figure 8, it is clear that the GRUMP, MODIS, HYDE, and Global Cover 300 do not match with the settlements on Google Earth Imagery. We could not calculate a comparable urban area from MODIS, HYDE and Global Cover 300. When we presented their multiple land cover onto a built-up area classification, they generated coarse and scattered pixels and urban agglomeration. We could not distinguish the exact Beijing, Brisbane, and Melbourne extents from these products. GMIS and HBASE match with the BUNTUS (see Figure 9). We extracted the contiguous pixels of GMIS and HBASE and calculated the urban area of Melbourne for 2010. The HBASE showed an annual growth rate of 69% from 2000 (2675 km²) to 2010 (2865 km²) and GMIS showed an annual growth rate of 51% from 2000 (2104 km²) to 2010 (2212 km²). BUNTUS showed an annual growth rate of 1.1% for the same period. We note that Melbourne's population growth during this period was approximately 2% making an area growth-rate of 0.5% in area unlikely.



Beijing

Figure 8. Comparison of Built-up, Nighttime Light, and Travel time for Urban Size (BUNTUS) boundaries with Global Rural-Urban Mapping Project (GRUMP 1995 product), Moderate Resolution Imaging Spectroradiometer (MODIS) 500-m land cover, History Database of the Global Environment (HYDE version 3.2), Global Man-made Impervious Surface (GMIS), and Global Cover and Human Built-up and Settlement Extent (HBASE) global datasets for the year 2010.



Figure 9. A close comparison of Built-up, Nighttime Light, and Travel time for Urban Size (BUNTUS), Global Man-made Impervious Surface (GMIS), and Global Cover and Human Built-up and Settlement Extent (HBASE) of Melbourne.

4.3. Urban Expansion

Figure 10 presents growth in urban boundaries for the three cities for the 20 annual increments. Figure 10 is helpful to understand how urban areas have grown over time. For example, Brisbane has grown more on its northern and southern sides, Beijing in all directions, and Melbourne has grown more on its western and southwestern sides.

Figure 11 summarizes the area of the three cities from 1998 to 2018. The graph, in the case of Beijing, shows continuous growth annually except for some points where some inputs are problematic. For example, the graphs show that Beijing in 2015 (3154 km²) to 2018 (3180 km²) was smaller than in 2014 (3206 km²). This could have happened because of the low accuracy of 2014 image classification. Similarly, Brisbane and Melbourne showed some irregularities but overall the urban growth was smooth. Some apparent discontinuities were explainable. For example, Brisbane's area suddenly increased in 2013 due to its urban agglomeration with the Gold Coast. In the following year Brisbane's area decreased. The reason was that our algorithm generated some gap between Brisbane undoing the agglomeration. Hence there was a decrease in Brisbane's area in 2014 (see Figure 10).



Figure 10. Urban area boundaries of three cities, 1998–2018.



Figure 11. Urban area versus year plot.

Generally, urban area does not shrink except during war or natural disasters. We can take advantage of this fact by making the boundary calculation purely additive, i.e., once a pixel is included in a city it remains there. Thus, the agglomeration of Brisbane in 2013 is not undone in subsequent years. This alternative methodology showed higher accuracy than the default algorithm when compared with high-resolution imagery. The boundaries generated through this method followed the urban settlements edge with high accuracy. Also, the urban area from this algorithm showed a smooth growth rate (see Figure 12).

From 1998 to 2018, the Beijing's area has increased by ~96% from 1723 km² to 3384 km², Brisbane's area has increased by ~76% from 1105 km² to 1948 km², and Melbourne's area has increased by ~31% from 1569 km² to 2049 km².



Figure 12. Post-processing urban area versus year plot.

5. Discussion and Limitations

5.1. Discussion

When assessing the utility of a dataset like BUNTUS, we need to keep in mind its intended application. This is true for any urban data set, since no single definition would suffice for all applications, even if the data to determine it existed. Our task here is to define an algorithm at once robust, consistent, and efficient enough to define global trends in urban extent for many cities over decades. The simplifications necessary to do this should be borne in mind when using BUNTUS, especially for other purposes.

A phenomenon which may confuse results is agglomeration, in which two cities fuse. For example, until 2011, Brisbane comprised a core city and surrounding suburbs. By 2013, the region of the Gold Coast to the south had fused into the Brisbane area. This resulted in a big increase in the Brisbane urban area. This change can be seen in Figure 10. Before 2013, the Gold Coast and Brisbane were separate entities (as per our boundaries). We remain consistent and followed our definition of urban areas, that the urban areas are those areas which were contiguous to the urban core. In general, where area definitions may be unstable, it is safer to calculate intensive quantities like emissions per unit area or per capita than absolute or extensive changes. This will be the focus of future research using BUNTUS.

5.2. Limitations

Finally, BUNTUS is limited by the quality of data. For example, due to the unavailability of cloud-free satellite images, the accuracy of the built-up classification is reduced for some regions and years. Similarly, the unavailability of Landsat product for the year 2012 left a gap in the study. Although the Landsat imagery with scanline error has been available for the year 2012, because of the consideration of high accuracy, we skipped Landsat7 product with scanline error in the present study [68]. Another limitation is the difference in resolution of the different data sets, resulting in mixed pixels. Though we tried to minimize the effect of mixed pixels when dealing with the coarser resolution datasets by resampling, they still play a role at a smaller scale.

6. Conclusions

Results from this exploration demonstrate that, by combining Landsat information with the nighttime lights and travel time information, it was possible to outline regions on a city level with sufficient stability and precision to determine trends. The methodology is fairly robust to variable data quality and availability and the algorithm sufficiently automatic and general to derive trends for many cities. The motivation behind BUNTUS was the development of urban size boundaries for the study of emissions trends. However, considering our definition of the urban area, BUNTUS boundaries can be used for other kinds of urban studies and trends.

7. Future Work

The development of this algorithm was the first step in a study of the relationship between urbanisation and trends in fossil fuel use. Future work will focus on the application of the algorithm to problems like the energy efficiency of cities as a function of size and development, relationships with other pollutants and the possible role of different urban forms in these trajectories [69]. As a first step, we will generate trends for 100 cities and overlay the boundaries on gridded population and emissions estimates. We will also test the trends in urban boundaries for more cities against a wider array of high-quality though temporally sparse datasets.

Author Contributions: Conceptualization, M.L., P.J.R., and K.R.G.; Formal analysis, M.L.; Funding acquisition, P.J.R.; Investigation, M.L., P.J.R., and K.R.G.; Methodology, M.L. and P.J.R.; Project administration, M.L.; Software, M.L.; Supervision, P.J.R.; Validation, M.L. and P.J.R.; Visualization, M.L. and K.R.G.; Writing—original draft, M.L.; Writing—review & editing, M.L., P.J.R., and K.R.G.

Funding: Muhammad Luqman is thankful to the University of Melbourne for a Melbourne Research Scholarship.

Acknowledgments: We are thankful to Google for providing the Google Earth Engine platform. We acknowledge the United States Geological Survey's (USGS) Landsat datasets, Open Street Map's network datasets, DMSP-OLS, and VIIRS products, and Google Earth. We also acknowledge the valuable suggestions of Robyn Schofield, Jason Thompson, Kerry Nice, and Roger Dargaville.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

- 1. United Nations. World Urbanization Prospects 2018—Population Division—United Nations. 2018. Available online: https://esa.un.org/unpd/wup/ (accessed on 13 June 2018).
- 2. Hoornweg, D.; Sugar, L.; Trejos Gómez, C.L. *Cities and Climate Change: An Urgent Agenda*; World Bank: Washington, DC, USA, 2010.
- 3. Bettencourt, L.; West, G. A unified theory of urban living. *Nature* 2010, 467, 912–913. [CrossRef] [PubMed]
- 4. Brown, M.A.; Southworth, F.; Sarzynski, A. The geography of metropolitan carbon footprints. *Policy Soc.* **2009**, *27*, 285–304. [CrossRef]
- Creutzig, F.; Baiocchi, G.; Bierkandt, R.; Pichler, P.-P.; Seto, K.C. Global typology of urban energy use and potentials for an urbanization mitigation wedge. *Proc. Natl. Acad. Sci. USA* 2015, 112, 6283–6288. [CrossRef] [PubMed]
- 6. Edenhofer, O.; Pichs-Madruga, R.; Sokona, Y.; Farahani, E.; Kadner, S.; Seyboth, K.; Adler, A.; Baum, I.; Brunner, S.; Eickemeier, B.; et al. Summary for policymakers climate change 2014, mitigation of climate change, IPCC 2014. In *Climate Change 2014: Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*; Cambridge University Press: Cambridge, UK; New York, NY, USA, 2014.
- 7. Marcotullio, P.J.; Solecki, W. What is a city? An essential definition for sustainability. In *Urbanization and Sustainability*; Springer: Dordrecht, The Netherlands, 2013; pp. 11–25.
- 8. Mcintyre, N.E.; Knowles-Yánez, K.; Hope, D. Urban ecology as an interdisciplinary field: Differences in the use of 'urban' between the social and natural sciences. *Urban Ecosyst.* **2000**, *4*, 5–24. [CrossRef]

- 9. Tabuchi, T. Agglomeration in World Cities. Procedia Soc. Behav. Sci. 2013, 77, 299–307. [CrossRef]
- 10. Jaeger, J.A.G.; Bertiller, R.; Schwick, C.; Cavens, D.; Kienast, F. Urban permeation of landscapes and sprawl per capita: New measures of urban sprawl. *Ecol. Indic.* **2010**, *10*, 427–441. [CrossRef]
- Kasanko, M.; Barredo, J.I.; Lavalle, C.; McCormick, N.; Demicheli, L.; Sagris, V.; Brezger, A. Are European cities becoming dispersed? A comparative analysis of 15 European urban areas. *Landsc. Urban Plan.* 2006, 77, 111–130. [CrossRef]
- 12. Poelmans, L.; van Rompaey, A. Detecting and modelling spatial patterns of urban sprawl in highly fragmented areas: A case study in the Flanders–Brussels region. *Landsc. Urban Plan.* **2009**, *93*, 10–19. [CrossRef]
- 13. Zhang, Q.; Seto, K.C. Mapping urbanization dynamics at regional and global scales using multi-temporal DMSP/OLS nighttime light data. *Remote Sens. Environ.* **2011**, *115*, 2320–2329. [CrossRef]
- 14. Schneider, A.; Woodcock, C.E. Compact, Dispersed, Fragmented, Extensive? A Comparison of Urban Growth in Twenty-five Global Cities using Remotely Sensed Data, Pattern Metrics and Census Information. *Urban Stud.* **2008**, *45*, 659–692. [CrossRef]
- 15. Schwarz, N. Urban form revisited—Selecting indicators for characterising European cities. *Landsc. Urban Plan.* **2010**, *96*, 29–47. [CrossRef]
- 16. Goldblatt, R.; You, W.; Hanson, G.; Khandelwal, A. Detecting the Boundaries of Urban Areas in India: A Dataset for Pixel-Based Image Classification in Google Earth Engine. *Remote Sens.* **2016**, *8*, 634. [CrossRef]
- 17. Parés-Ramos, I.K.; Álvarez-Berríos, N.L.; Aide, T.M. Mapping Urbanization Dynamics in Major Cities of Colombia, Ecuador, Perú, and Bolivia Using Night-Time Satellite Imagery. *Land* **2013**, *2*, 37–59. [CrossRef]
- 18. Small, C.; Pozzi, F.; Elvidge, C.D. Spatial analysis of global urban extent from DMSP-OLS night lights. *Remote Sens. Environ.* **2005**, *96*, 277–291. [CrossRef]
- 19. Mindali, O.; Raveh, A.; Salomon, I. Urban density and energy consumption: A new look at old statistics. *Transp. Res. Part A Policy Pract.* **2004**, *38*, 143–162. [CrossRef]
- 20. Niu, Q.; Wang, Y.; Xia, Y.; Wu, H.; Tang, X. Detailed Assessment of the Spatial Distribution of Urban Parks According to Day and Travel Mode Based on Web Mapping API: A Case Study of Main Parks in Wuhan. *Int. J. Environ. Res. Public Health* **2018**, *15*, 1725. [CrossRef]
- 21. Sun, Y.; Fan, H.; Li, M.; Zipf, A. Identifying the city center using human travel flows generated from location-based social networking data. *Environ. Plan. B Plan. Des.* **2016**, *43*, 480–498. [CrossRef]
- 22. Xie, F.; Levinson, D. Measuring the Structure of Road Networks. Geogr. Anal. 2007, 39, 336–356. [CrossRef]
- 23. Goldewijk, K.K. Estimating global land use change over the past 300 years: The HYDE database. *Glob. Biogeochem. Cycles* **2001**, *15*, 417–433. [CrossRef]
- 24. De Colstoun, E.C.B.; Huang, C.; Wang, P.; Tilton, J.C.; Tan, B.; Phillips, J.; Niemczura, S.; Ling, P.Y.; Wolfe, R. *Global Man-Made Impervious Surface (GMIS) Dataset from Landsat*; NASA Socioeconomic Data and Applications Center (SEDAC): Palisades, NY, USA, 2017.
- 25. Wang, P.; Huang, C.; de Colstoun, E.C.B.; Tilton, J.C.; Tan, B. *Global Human Built-up and Settlement Extent* (*Hbase) Dataset from Landsat*; NASA Socioeconomic Data and Applications Center (SEDAC): Palisades, NY, USA, 2017.
- 26. Elvidge, C.; Tuttle, B.; Sutton, P.; Baugh, K.; Howard, A.; Milesi, C.; Bhaduri, B.; Nemani, R. Global distribution and density of constructed impervious surfaces. *Sensors* **2007**, *7*, 1962–1979. [CrossRef]
- 27. Esch, T.; Bachofer, F.; Heldens, W.; Hirner, A.; Marconcini, M.; Palacios-Lopez, D.; Roth, A.; Üreyen, S.; Zeidler, J.; Dech, S.; et al. Where we live—A summary of the achievements and planned evolution of the global urban footprint. *Remote Sens.* **2018**, *10*, 895. [CrossRef]
- 28. Pesaresi, M.; Ehrlich, D.; Ferri, S.; Florczyk, A.; Freire, S.; Halkia, M.; Julea, A.; Kemper, T.; Soille, P.; Syrris, V. *Operating Procedure for the Production of the Global Human Settlement Layer from Landsat Data of the Epochs* 1975, 1990, 2000, and 2014; Publications Office of the European Union: Ispara, Italy, 2016; pp. 1–62.
- 29. CIESIN. Global Rural-Urban Mapping Project, Version 1 (GRUMPv1): Urban Extent Polygons, Revision 01; NASA Socioeconomic Data and Applications Center (SEDAC): Palisades, NY, USA, 2017.
- 30. *National Imagery and Mapping Agency, Washington DC, VMap0*; National Imagery and Mapping Agency: Washington, DC, USA, 1997.
- 31. Sulla-Menashe, D.; Friedl, M.A. *User Guide to Collection 6 MODIS Land Cover (MCD12Q1 and MCD12C1) Product;* USGS: Reston, VA, USA, 2018.

- Kirches, G.; Brockmann, C.; Boettcher, M.; Peters, M.; Bontemps, S.; Lamarche, C.; Schlerf, M.; Santoro, M.; Defourny, P. Land Cover CCI-Product User Guide-Version 2. Available online: http://maps.elie.ucl.ac.be/ CCI/viewer/download/ESACCI-LC-PUG-v2.5.pdf (accessed on 25 January 2019).
- Chen, J.; Chen, J.; Liao, A.; Cao, X.; Chen, L.; Chen, X.; He, C.; Han, G.; Peng, S.; Lu, M.; et al. Global land cover mapping at 30 m resolution: A POK-based operational approach. *ISPRS J. Photogramm. Remote Sens.* 2015, *103*, 7–27. [CrossRef]
- 34. Arino, O.; Gross, D.; Ranera, F.; Leroy, M.; Bicheron, P.; Brockman, C.; Defourny, P.; Vancutsem, C.; Achard, F.; Durieux, L.; et al. GlobCover: ESA service for global land cover from MERIS. In Proceedings of the 2007 IEEE International Geoscience and Remote Sensing Symposium, Barcelona, Spain, 23–28 July 2007; pp. 2412–2415.
- 35. Elvidge, C.D.; Baugh, K.E.; Zhizhin, M.; Hsu, F.-C. Why VIIRS data are superior to DMSP for mapping nighttime lights. *Proc. Asia Pac. Adv. Netw.* **2013**, *35*, 62. [CrossRef]
- 36. Potere, D.; Schneider, A.; Angel, S.; Civco, D.L. Mapping urban areas on a global scale: Which of the eight maps now available is more accurate? *Int. J. Remote Sens.* **2009**, *30*, 6531–6558. [CrossRef]
- 37. Esch, T.; Marconcini, M.; Felbier, A.; Roth, A.; Heldens, W.; Huber, M.; Schwinger, M.; Taubenböck, H.; Müller, A.; Dech, S.J.I.G. Urban footprint processor—Fully automated processing chain generating settlement masks from global data of the TanDEM-X mission. *IEEE Geosci. Remote Sens. Lett.* 2013, 10, 1617–1621. [CrossRef]
- 38. Pesaresi, M.; Ehrilch, D.; Florczyk, A.J.; Freire, S.; Julea, A.; Kemper, T.; Soille, P.; Syrris, V. GHS Built-up Grid, Derived from LANDSAT, Multitemporal (1975, 1990, 2000, 2014). European Commission, Joint Research Centre (JRC) [Dataset] PID. Available online: http://data.europa.eu/89h/jrc-ghsl-ghs_built_ldsmt_globe_ r2015b (accessed on 25 January 2019).
- 39. Irish, R.R. Landsat 7 Automatic Cloud Cover Assessment. In Proceedings of the Volume 4049, Algorithms for Multispectral, Hyperspectral, and Ultraspectral Imagery VI, Orlando, FL, USA, 23 August 2000; p. 348.
- 40. Gorelick, N.; Hancher, M.; Dixon, M.; Ilyushchenko, S.; Thau, D.; Moore, R. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* **2017**, *202*, 18–27. [CrossRef]
- 41. Huang, C.; Yang, J.; Jiang, P. Assessing Impacts of Urban Form on Landscape Structure of Urban Green Spaces in China Using Landsat Images Based on Google Earth Engine. *Remote Sensing* **2018**, *10*, 1569. [CrossRef]
- 42. Trianni, G.; Lisini, G.; Angiuli, E.; Moreno, E.A.; Dondi, P.; Gaggia, A.; Gamba, P. Scaling up to National/Regional Urban Extent Mapping Using Landsat Data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 2015, *8*, 3710–3719. [CrossRef]
- 43. Arino, O.; Perez, J.R.; Kalogirou, V.; Defourny, P.; Achard, F. *GlobCover* 2009. 2010. Available online: https://epic.awi.de/id/eprint/31046/1/Arino_et_al_GlobCover2009-a.pdf (accessed on 20 January 2019).
- 44. Chander, G.; Markham, B.L.; Helder, D.L. Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. *Remote Sens. Environ.* **2009**, *113*, 893–903. [CrossRef]
- 45. Hu, Y.; Hu, Y. Land Cover Changes and Their Driving Mechanisms in Central Asia from 2001 to 2017 Supported by Google Earth Engine. *Remote Sens.* **2019**, *11*, 554. [CrossRef]
- 46. Rutherford, G.N.; Bebi, P.; Edwards, P.J.; Zimmermann, N.E. Assessing land-use statistics to model land cover change in a mountainous landscape in the European Alps. *Ecol. Model.* **2008**, *212*, 460–471. [CrossRef]
- Yuan, F.; Sawaya, K.E.; Loeffelholz, B.C.; Bauer, M.E. Land cover classification and change analysis of the Twin Cities (Minnesota) Metropolitan Area by multitemporal Landsat remote sensing. *Remote Sens. Environ.* 2005, *98*, 317–328.
- 48. Deilmai, B.R.; Ahmad, B.B.; Zabihi, H. Comparison of two classification methods (MLC and SVM) to extract land use and land cover in Johor Malaysia. *IOP Conf. Ser. Earth Environ. Sci.* **2014**, *20*, 012052. [CrossRef]
- 49. Xu, X.; Min, X. Quantifying spatiotemporal patterns of urban expansion in China using remote sensing data. *Cities* **2013**, *35*, 104–113. [CrossRef]
- 50. Breiman, L. Random forests. Mach. Learn. 2001, 45, 5-32. [CrossRef]
- Akar, Ö.; Güngör, O. Classification of multispectral images using Random Forest algorithm. J. Geod. Geoinf. 2012, 1, 105–112. [CrossRef]
- 52. Gislason, P.O.; Benediktsson, J.A.; Sveinsson, J.R. Random Forests for land cover classification. *Pattern Recognit. Lett.* **2006**, *27*, 294–300. [CrossRef]
- Rodriguez-Galiano, V.F.; Ghimire, B.; Rogan, J.; Chica-Olmo, M.; Rigol-Sanchez, J.P. An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS J. Photogramm. Remote Sens.* 2012, 67, 93–104. [CrossRef]

- 54. Belgiu, M.; Drăguţ, L. Random forest in remote sensing: A review of applications and future directions. *ISPRS J. Photogramm. Remote Sens.* **2016**, *114*, 24–31. [CrossRef]
- 55. Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction;* Springer: New York, NY, USA, 2009.
- 56. Stehman, S.V. Selecting and interpreting measures of thematic classification accuracy. *Remote Sens. Environ.* **1997**, *62*, 77–89. [CrossRef]
- 57. Anderson, J.R. *A Land Use and Land Cover Classification System for Use with Remote Sensor Data;* US Government Printing Office: Washington, DC, USA, 1976; Volume 964.
- 58. Lewis, H.G.; Brown, M. A generalized confusion matrix for assessing area estimates from remotely sensed data. *Int. J. Remote Sens.* 2001, *22*, 3223–3235. [CrossRef]
- 59. Sutton, P.; Roberts, D.; Elvidge, C.; Baugh, K. Census from Heaven: An estimate of the global human population using night-time satellite imagery. *Int. J. Remote Sens.* **2001**, *22*, 3061–3076. [CrossRef]
- 60. Henderson, M.; Yeh, E.T.; Gong, P.; Elvidge, C.; Baugh, K. Validation of urban boundaries derived from global night-time satellite imagery. *Int. J. Remote Sens.* **2003**, *24*, 595–609. [CrossRef]
- 61. Liu, F.; Zhang, Z.; Wang, X. Forms of Urban Expansion of Chinese Municipalities and Provincial Capitals, 1970s–2013. *Remote Sens.* **2016**, *8*, 930. [CrossRef]
- 62. Zhou, D.; Zhao, S.; Liu, S.; Zhang, L. Spatiotemporal trends of terrestrial vegetation activity along the urban development intensity gradient in China's 32 major cities. *Sci. Total Environ.* **2014**, *488*, 136–145. [CrossRef]
- 63. Imhoff, M.L.; Lawrence, W.T.; Stutzer, D.C.; Elvidge, C.D. A technique for using composite DMSP/OLS 'city lights' satellite data to map urban area. *Remote Sens. Environ.* **1997**, *61*, 361–370. [CrossRef]
- 64. Hsu, F.-C.; Baugh, K.; Ghosh, T.; Zhizhin, M.; Elvidge, C. DMSP-OLS radiance calibrated nighttime lights time series with intercalibration. *Remote Sens.* **2015**, *7*, 1855–1876. [CrossRef]
- 65. Elvidge, C.D.; Ziskin, D.; Baugh, K.E.; Tuttle, B.T.; Ghosh, T.; Pack, D.W.; Erwin, E.H.; Zhizhin, M. A Fifteen Year Record of Global Natural Gas Flaring Derived from Satellite Data. *Energies* **2009**, *2*, 595–622. [CrossRef]
- 66. Day, J.; Chen, Y.; Ellis, P.; Roberts, M. A free, open-source tool for identifying urban agglomerations using polygon data. *Environ. Syst. Decis.* **2017**, *37*, 68–87. [CrossRef]
- Pulsipher, L.M.; Pulsipher, A.; Johansson, O. World Regional Geography: Global Patterns, Local Lives, 7th ed.;
 W. H. Freeman: New York, NY, USA, 2017.
- 68. USGS. Landsat 7. 2019. Available online: https://www.usgs.gov/land-resources/nli/landsat/landsat-7?qt-science_support_page_related_con (accessed on 27 October 2019).
- 69. Thompson, J.; Stevenson, M.; Wijnands, J.S.; Nice, K.; Aschwanden, G.; Silver, J.; Nieuwenhuijsen, M.; Rayner, P.; Schofield, R.; Hariharan, R.; et al. *Injured by Design: A Global Perspective on Urban Design and Road Transport Injury*; SSRN Scholarly Paper ID 3307635; Social Science Research Network: Rochester, NY, USA, 2018.



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