Nighttime lights are more strongly related to urban building volume than to urban area

Lingfei Shi[a,b,c]*, Giles M Foody[c], Doreen S Boyd[c], Renoy Girindran[c], Lihui Wang[a], Yun Du[a] and Feng Ling[a]

[a]*Key Laboratory for Environment and Disaster Monitoring and Evaluation, Institute of Geodesy and Geophysics, Chinese Academy of Sciences, Wuhan, 430077, China;*

[b]*The University of Chinese Academy of Sciences, Beijing, 100049, China;*

[c]*School of Geography, University of Nottingham, University Park, Nottingham, NG7 2RD, UK;*

shilingfei14@mails.ucas.ac.cn

Nighttime lights are more strongly related to urban building volume than to urban area

A strong relationship between nighttime light (NTL) data and the areal extent of urbanised regions has been observed frequently. As urban regions have an important vertical dimension, it is hypothesised that the strength of the relationship with NTL can be increased by consideration of the volume rather than simply the area of urbanised land. Relationships between NTL and the area and volume of urbanised land were determined for a set of towns and cities in the UK, the conterminous states of the USA and countries of the European Union. Strong relationships between NTL and the area urbanised were observed, with correlation coefficients ranging from 0.9282-0.9446. Higher correlation coefficients were observed for the relationship between NTL and urban building volume, ranging from 0.9548-0.9604; The difference in the correlations obtained with volume and with area was statistically significant at the 95% level of confidence. Studies using NTL data may be strengthened by consideration of the volume rather than just area of urbanised land.

1 Introduction

Nighttime light (NTL) sensors record visible light emerging from the Earth at night and their imagery are often used as a proxy for a number of variables, including urban density, population size and economic status (Mellander et al. 2015; Wu et al. 2014). The use of NTL data as a proxy variable is based upon there being a direct or an indirect link between the sources of NTL and the variable of interest. Considerable attention has focused on artificial lights and the human population associated with them, especially in relation to urban environments. Generally, there is a strong positive relationship between urban population size and NTL as artificial lighting is commonly used to illuminate homes, workplaces and other urban features such as advertising boards, sporting venues, and streetlights (Sutton et al. 1997; Zhang et al. 2018). Growth of the urban population is typically associated with an increase in the number of dwelling units which in turn is

associated with an increase in the size of the urbanised region. In many instances, the growth involves an increase in the areal extent of the urbanised land and such lateral expansion has been commonly monitored via remote sensing (Weng 2012), including use of NTL sensors (Zhou et al. 2015; Li and Zhou 2017).

Urban growth often involves more than simply an increase in the spatial extent of settlements by lateral expansion (Handayani et al. 2018). Settlements may, for example, become more intensively built-up, with vegetated terrain converted to built-up land cover. This intensification may be studied by, for example, monitoring changes in the impervious cover (Shi et al. 2017). Urban growth may also occur in the vertical domain, often associated with the increasing construction of high-rise buildings (Giridharan, Ganesan, and Lau 2004). The latter may be expected to increase the total source of NTL as, for example, there would be more dwelling units providing light through their windows (Li et al. 2019). Growth in the vertical domain could be studied by monitoring changes in elevation. Increases in the height of urban areas, especially in residential areas, suggest that urban building volume rather than area may more appropriately reflect properties of the settlement, its population and the total source of NTL. By including information on height in studies it may be possible to strengthen the relationships between urban properties and NTL in order to facilitate the enhanced study of urban features and their populations remotely.

Recent developments in remote sensing, notably lidar and radar systems, offer the potential for height estimation over large areas which would allow the estimation of urban building volume. Airborne lidar systems, for example, offer the ability for highly accurate local scale estimates of height (Kada and McKinley 2009). For large areas, height may be estimated, albeit at a relatively coarse scale, as the difference between a digital surface model (DSM) and a digital elevation model (DEM) (Zhang et al. 2018). Critically, it is

now possible to estimate the height of urbanised land as well as its areal extent and hence estimate urban building volume. Here, the aim is to test the hypothesis that a stronger relationship exists between NTL and the volume of urbanised land than with its area.

2 Study areas and data

This article is focused on studies undertaken at two scales using data for three regions (Figure 1). First, a local scale for sites for which highly accurate building height and footprint data were available. This allowed accurate estimation of building areal extent and volume but this does not include other key sources of urban NTL such as streetlights. Consequently, the second set of studies were undertaken at a coarser scale in which the total area of urbanised land cover and its height, estimated from regional scale DSM and DEM data, were used.
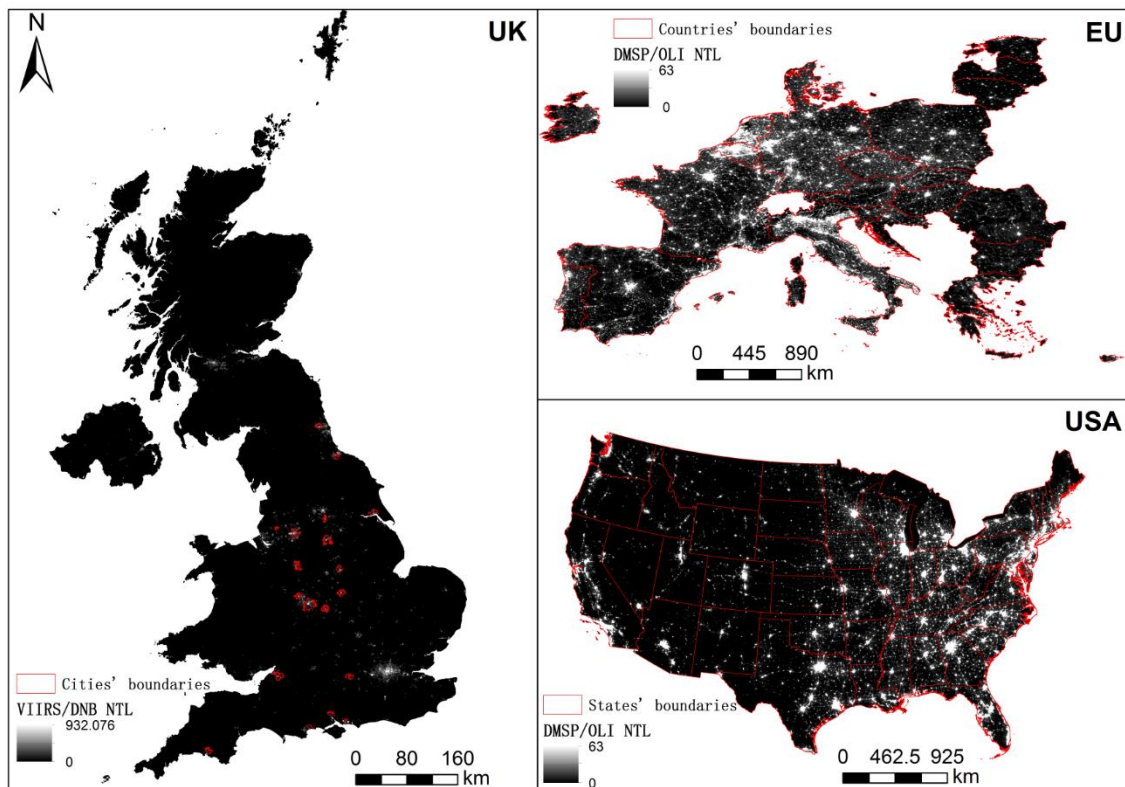

Figure 1. Three study areas with NTL images.

2.1 Local scale

An open building height data for the United Kingdom (UK) provided by Emu Analytics was used to test the relationships between NTL and building area and volume (http://www.emu-analytics.com/products/datapacks.php). This data has detailed information on building footprints and heights for a set of towns and cities in the UK. The height information was generated with lidar data that were acquired by the UK's Environment Agency with an absolute height error of less than ± 15cm in the period 1998-2014. These data combined with building footprint data obtained from Ordnance Survey (OS) Open Map allowed estimation of both building area and volume in each settlement as defined in the national Major Towns and Cities (December 2015) Boundaries data set. Data were available for a total of 25 settlements but those for 6 were excluded from the analyses. The excluded settlements were those for which preliminary assessments highlighted potential problems (e.g. for 5 cities there were buildings that had negative heights) and London which, as so much larger than the others, could be an outlier. The 19 towns and cities included in the analysis were: Birmingham, Bournemouth, Bristol, Coventry, Hull, Leicester, Manchester, Middlesbrough, Newcastle, Nottingham, Plymouth, Portsmouth, Reading, Sheffield, Southampton, Stoke, Wakefield, Wigan, and Wolverhampton.

Visible Infrared Imaging Radiometer Suite Day/Night Band (VIIRS/DNB) data were obtained for the 19 towns and cities in the UK. The VIIRS/DNB data were provided by the National Oceanic and Atmospheric Administration (NOAA). These data are available from 2012 in two forms: monthly and annual composites. Attention here was focused on annual composite data since their pre-processing involves removal of transient lights arising from features such as fires and ships. The annual composite data set closest to the dates of the lidar data, acquired in 2015, was used.

2.2 Regional scale

Additional studies at a coarser, regional, scale that included the full set of urban NTL sources were undertaken. These studies focused on the States of the conterminous United States of America (USA) and countries of the European Union (EU).

For studies focused on the USA, the relationship between NTL and urbanised area and volume was determined at the State level. The impervious surface class depicted in the 2006 National Land Cover Database (NLCD) (Xian, Homer, and Fry 2009) was taken to represent urbanised land cover. These data have a 30 m spatial resolution and allowed estimation of the areal extent of urbanised land cover in each State. The latter was used to estimate the volume of urbanised land together with height information obtained as the difference between a DSM and DEM. The DSM was the Advanced Land Observing Satellite (ALOS) World 3D-30 m (AW3D30) data set (Tadono et al. 2016). This data set has a 30 m spatial resolution and was co-registered to the NLCD data. The DSM was generated from data acquired by the PRISM panchromatic stereo mapping sensor carried on the ALOS that operated between 2006 and 2011. The DEM used was the Shuttle Radar Topography Mission (SRTM) data set generated in 2000. These data also have a 30 m spatial resolution and were co-registered to the NLCD data. Finally, the Defence Meteorological Satellite Program/Operational Linescan System (DMSP/OLS) stable light data was used as the source of NTL data. This data source was selected because the ephemeral events, such as fires and background noise were identified and removed. To reduce temporal differences in the data sets, the F152006 data set of DMSP/OLS stable light data for the year 2006 was used. The total NTL observed over the urban land cover was calculated for each of the 48 contiguous states of the USA.

As with the research focused on the USA, the AW3D30 DSM, SRTM DEM and

DMSP/OLS data were used in the study focused on EU countries. The extent of urbanised land cover was extracted from the GHS BUILT-UP GRID data produced by the EU's Joint Research Centre (Pesaresi et al. 2015). This data set contains a multi-temporal information layer on built-up presence as derived from Landsat image data that was used to represent the urban class. These data have a 38 m spatial resolution and were resampled to 30 m resolution with the nearest neighbour method for integration with the other data sets. The data for the year 2000 was used. To fit with this, the F152000 data set of DMSP/OLS stable light data was used as the source of NTL data for the year 2000. Initial assessments of the data indicated problems with the AW3D30 data set for Finland and Sweden, notably missing data, and hence the data for these countries were excluded from further study. The remaining 25 countries of the EU were included in the analysis.

3 Methods

For each study, the area and volume of the urbanised land cover were calculated. For the 19 towns and cities in the UK, the height of each building in the lidar data set was used to determine its volume and the total volume of all buildings was calculated. At the regional scale, the urban area estimates were the total impervious cover in each State of the USA and the total urban cover in each country in the GHS BUILT-UP GRID for the EU. Volume estimates for USA states and EU countries were obtained by multiplying the urban area by the height information generated by subtraction of the DEM from the DSM.

The NTL data acquired for each study were used to determine the total light intensity measured for urbanised region. The strength of the relationship between NTL and the area urbanised and between NTL and the volume of the urban cover was assessed using Pearson's correlation coefficient (r) (Lee Rodgers and Nicewander 1988). The statistical significance of the difference in the correlation obtained using area and volume

was assessed using Hotelling's (1940) *t* statistic (Hotelling 1940) which was specifically designed for the comparison of two correlation coefficients calculated from overlapping dependent samples (May and Hittner 1997). The Hotelling's (1940) *t* statistic is based on:

$$t = \frac{(r_{12} - r_{13})\sqrt{(N-3)(1+r_{23})}}{\sqrt{2(1 - r_{12}^2 - r_{13}^2 - r_{23}^2 + 2r_{12}r_{13}r_{23})}} \tag{1}$$

where N is the sample size. r_{12} is the correlation coefficient between NTL and volume, r_{13} is the correlation coefficient between NTL and area. r_{23} is the correlation coefficient between area and volume. Recognising the directional nature of the hypothesis being tested, a one-sided significance test was performed at the 95% level of confidence.

4 Result and discussion

Strong and statistically significant relationships for NTL with both the area and the volume of the urbanised land cover were observed (Figure 2). The correlation coefficients observed for the relationship between NTL and urban area in the UK, USA and EU based studies were 0.9350, 0.9282 and 0.9446 respectively, confirming results of other studies that the relationship between NTL and urban area is strong. In contrast, stronger relationships were evident between NTL and urbanised volume. The correlation coefficients observed for the relationship between NTL and urbanised volume in the UK, USA and EU based studies were 0.9548, 0.9563 and 0.9604 respectively. Furthermore, the differences in the magnitude of the correlation coefficients obtained with urbanised area and those with volume were all statistically significant at the 95% level of confidence (Table 1).

The results show that the strength of the relationship between NTL and urban area can be increased by including information on the height of the urbanised region. As NTL may be used as a proxy for a number of variables, including urbanization, population

density, and economic growth, the increase in correlation coefficient is mainly because that urban building volume rather than area more appropriately reflects properties of the settlement, its population and the total source of NTL. Although there are some limitations to this study, notably in the time gap between some data sets used, the trend throughout suggests that greater attention should be given to the volume rather than area of urbanised environments. With height data becoming easier to acquire and increasingly available this may allow studies using NTL to relate more closely to key variables of urbanised areas and their populations.
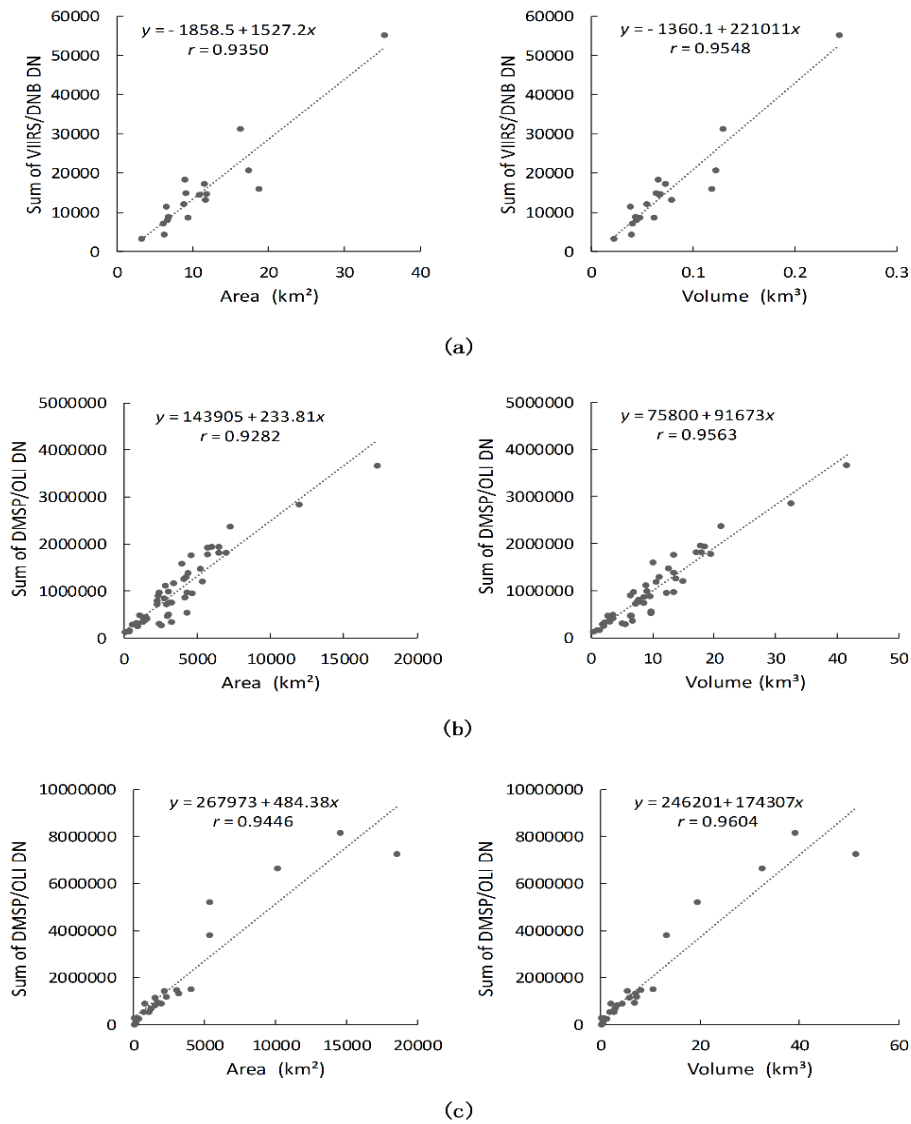

Figure 2. Relationships between NTL and urban extent defined as urban area and urban building volume. (a) UK, (b) USA and (c) EU.

Table 1. The results of the significance test

Study area	UK	USA	EU
N	19	48	25
t	2.1847	3.4964	2.3813
t-critical value	1.7459	1.6794	1.7171
p-value	0.0221	0.0005	0.0132
Significant?	Yes	Yes	Yes

5 Conclusion

The results confirmed the previously observed strong relationship between NTL and the extent of urbanised land and, critically, demonstrated that this relationship could be strengthened by including information on height. The stronger relationships between NTL and the volume rather than area of urbanised land may help enhance studies based on NTL data.

Acknowledgements

References

Giridharan, R., S Ganesan, and S.S.Y Lau. 2004. "Daytime urban heat island effect in high-rise and high-density residential developments in Hong Kong." *Energy & Buildings* 36 (6):525-34.

Handayani, Hepi, Yuji Murayama, Manjula Ranagalage, Fei Liu, and DMSLB Dissanayake. 2018. "Geospatial analysis of horizontal and vertical urban expansion using multi-spatial resolution data: A case study of Surabaya, Indonesia." *Remote Sensing* 10 (10):1599.

Hotelling, Harold. 1940. "The teaching of statistics." *The Annals of Mathematical Statistics* 11 (4):457-70.

Kada, Martin, and Laurence McKinley. 2009. "3D building reconstruction from LiDAR based on a cell decomposition approach." *International Archives of*

Photogrammetry, Remote Sensing and Spatial Information Sciences 38 (Part 3):W4.

Lee Rodgers, Joseph, and W Alan Nicewander. 1988. "Thirteen ways to look at the correlation coefficient." *The American Statistician* 42 (1):59-66.

Li, Xi, Ruiqi Ma, Qingling Zhang, Deren Li, Shanshan Liu, Tao He, and Lixian Zhao. 2019. "Anisotropic characteristic of artificial light at night–Systematic investigation with VIIRS DNB multi-temporal observations." *Remote Sensing of Environment* 233:111357.

Li, Xuecao, and Yuyu Zhou. 2017. "Urban mapping using DMSP/OLS stable night-time light: a review." *International Journal of Remote Sensing* 38 (21):6030-46.

May, Kim, and James B. Hittner. 1997. "Tests for Comparing Dependent Correlations Revisited: A Monte Carlo Study." *Journal of Experimental Education* 65 (3):257-69.

Mellander, Charlotta, José Lobo, Kevin Stolarick, and Zara Matheson. 2015. "Night-time light data: A good proxy measure for economic activity?" *PloS One* 10 (10):e0139779.

Pesaresi, Martino, Daniele Ehrilch, Aneta J Florczyk, Sergio Freire, Andreea Julea, Thomas Kemper, Pierre Soille, and Vasileios Syrris. 2015. "GHS built-up grid, derived from Landsat, multitemporal (1975, 1990, 2000, 2014)." *European Commission, Joint Research Centre, JRC Data Catalogue*.

Shi, Lingfei, Feng Ling, Yong Ge, Giles Foody, Xiaodong Li, Lihui Wang, Yihang Zhang, and Yun Du. 2017. "Impervious surface change mapping with an uncertainty-based spatial-temporal consistency model: A case study in wuhan city using landsat time-series datasets from 1987 to 2016." *Remote Sensing* 9 (11):1148.

Sutton, Paul, Dar Roberts, Chris Elvidge, and Henk Meij. 1997. "A comparison of nighttime satellite imagery and population density for the continental United States." *Photogrammetric Engineering and Remote Sensing* 63 (11):1303-13.

Tadono, T., H. Nagai, H. Ishida, F. Oda, and H. Iwamoto. 2016. "Generation of the 30 m-mesh global digital surface model by ALOS PRISM." *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XLI-B4:157-62.

Weng, Qihao. 2012. "Remote sensing of impervious surfaces in the urban areas: Requirements, methods, and trends." *Remote Sensing of Environment* 117:34-49.

Wu, Jiansheng, Lin Ma, Weifeng Li, Jian Peng, and Hao Liu. 2014. "Dynamics of urban density in china: Estimations based on DMSP/OLS nighttime light data." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 7 (10):4266-75.

Xian, George, Collin Homer, and Joyce Fry. 2009. "Updating the 2001 National Land Cover Database land cover classification to 2006 by using Landsat imagery change detection methods." *Remote Sensing of Environment* 113 (6):1133-47.

Zhang, Junlin, Wei Xu, Lianjie Qin, and Yugang Tian. 2018. "Spatial Distribution Estimates of the Urban Population Using DSM and DEM Data in China." *ISPRS International Journal of Geo-Information* 7 (11):435.

Zhou, Yuyu, Steven J Smith, Kaiguang Zhao, Marc Imhoff, Allison Thomson, Ben Bond-Lamberty, Ghassem R Asrar, Xuesong Zhang, Chunyang He, and Christopher D Elvidge. 2015. "A global map of urban extent from nightlights." *Environmental Research Letters* 10 (5):054011.