Int J Appl Earth Obs Geoinformation 77 (2019) 100–107

Contents lists available at ScienceDirect

Int J Appl Earth Obs Geoinformation

journal homepage: www.elsevier.com/locate/jag

Satellite scatterometer estimation of urban built-up volume: Validation with airborne lidar data

Adam J. Mathews[a,*], Amy E. Frazier[b], Son V. Nghiem[c], Gregory Neumann[c], Yun Zhao[d]

[a] Department of Geography, Western Michigan University, Kalamazoo, Michigan, USA
[b] School of Geographical Sciences and Urban Planning, Arizona State University, Tempe, Arizona, USA
[c] NASA Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, USA
[d] Department of Environmental Studies, University of Illinois at Springfield, Springfield, Illinois, USA

ARTICLE INFO

Keywords:
Radar
Dense sampling method
Lidar
Built-up volume
Land use
Land cover

ABSTRACT

Accurately mapping urban infrastructure and extent is a high priority for resource management and service allocation as well as for addressing environmental, socioeconomic, and geopolitical concerns. Most available data products only document surficial (two-dimensional) land use and land cover (LULC), yet a substantial component of urban growth occurs in the vertical dimension. Light detection and ranging (lidar) data offer the potential for monitoring three-dimensional (3D) change, but the extreme lack of systematic lidar coverage worldwide inflicts considerable gaps in both spatial and temporal coverage. Satellite scatterometer (radar) data may serve as an alternative data source for characterizing urban growth and development in both the horizontal and vertical directions. The accuracy of these radar-based datasets for estimating building volumes remains to be validated quantitatively. For nine U.S. cities, we test whether scatterometer data can be used to estimate 3D urban built-up volume. We found strong, linear correlations between the lidar-derived and radar-derived building volume estimates for all cities with r^2 values as high as 0.98 when using spatial trend analysis. Given the high expense that limits lidar data acquisition to small areas at sporadic points in time, satellite scatterometer data provide a breakthrough method for monitoring both vertical growth and horizontal expansion of cities across the world with a continuous decadal time scale.

1. Introduction

Between 1950 and 2014, the world's urban population grew from 746 million to 3.9 billion (United Nations, 2014), and projections suggest that this growth will continue well into the future (Alig et al., 2004). Accurately mapping urban infrastructure and extent is a high priority for appropriately allocating resources and services as well as responding to environmental, socioeconomic, and geopolitical issues (Lwin and Murayama, 2009). Understanding urban change is also critical for decision-makers to plan for sustainable urban development in the future (Seto et al., 2014). Yet, despite the extensive availability of global remotely sensed datasets, developing accurate methods for mapping urban extents across the world has proven difficult (Nguyen et al., 2018). A major limitation of most imagery is that it only accounts for surficial (two-dimensional, or 2D) land use and land cover (LULC), while much of the development occurring in urban areas includes a significant vertical component. For example, parking lots exhibit very similar spectral land cover characteristics as buildings, but they support very different land uses and population and contribute very differently to phenomenon such as urban heat island effects.

More recently, researchers have explored the use of light detection and ranging (lidar) data for characterizing urban structures and extents (Dong et al., 2010; Qiu et al., 2010; Zhao et al., 2017). Lidar measures distances between the sensor and objects on Earth through light pulses and can provide very high spatial resolution (< 1 m) three-dimensional (3D) information on the heights of structures. These data can then be processed to provide volumetric measurements of objects such as buildings or trees. A major benefit of lidar-derived volumetric estimates is their use in population estimation (Dong et al., 2010; Lu et al., 2011; Zhao et al., 2017), which has proven challenging worldwide due to the lack of population censuses in many countries and infrequent surveys in others. However, despite encouraging findings on the relationships between lidar-derived building volumes and population estimates (Zhao et al., 2017), these approaches are limited by the severe lack of lidar coverage globally, resulting in considerable gaps in the spatial and temporal coverage of these datasets and thereby precluding their

* Corresponding author at: 3219 Wood Hall, Kalamazoo, MI, 49008, USA.
E-mail address: adam.mathews@wmich.edu (A.J. Mathews).

https://doi.org/10.1016/j.jag.2019.01.004
Received 30 October 2018; Received in revised form 30 December 2018; Accepted 7 January 2019

systematic use in change detection studies. Moreover, when lidar data are available, they have often been preprocessed to represent only the bare-Earth digital terrain model, which hinders analyses of surface features such as urban infrastructure. Thus, a different method is required to monitor 3D urban change for population estimation and other urban science research and applications.

Spaceborne scatterometer data have been investigated as an innovative method for characterizing urban growth and development as these data can capture features in both the horizontal and vertical directions (Nghiem et al., 2009; Nguyen et al., 2018). A scatterometer operates by transmitting pulses of microwave energy, typically from onboard an orbiting satellite platform, and measuring the reflected or backscattered power. The U.S. National Aeronautics and Space Administration (NASA) launched several scatterometer missions, including the Seasat Scatterometer (SASS) in 1978 (Jones et al., 1982), the NASA Scatterometer (NSCAT) in 1997 (Naderi et al., 1991), and the SeaWinds scatterometer aboard the QuikSCAT satellite in 1999 (Tsai et al., 2000) and the ADEOS II satellite in 2002 (Freilich et al., 2003). The primary function of these scatterometers was to monitor ocean wind speed and direction, which have played a key role in understanding and predicting complex global weather patterns and climate systems.

Nevertheless, scientists also recognize the utility of these scatterometer datasets for research in LULC change and impacts (Nghiem et al., 2009, 2014; Jacobson et al., 2015; Masetti et al., 2015; Nghiem, 2015; Stevenazzi et al., 2015; Groisman et al., 2017; Balk et al., 2018; Nguyen et al., 2018). Specifically, the active microwave scatterometer sensors are sensitive to the geometry of the target, and scientists have developed techniques to process these signals into a volumetric measurement (Nghiem et al., 2009) that theoretically represents the volume of the built environment. However, despite demonstrated usages of scatterometer data in urban areas, these data have not yet been validated quantitatively using high spatial resolution lidar data to determine how accurately they estimate building volume. The objective of this study is to validate the use of scatterometer data for deriving urban building volume so that these data may serve as a reliable and consistent global data source for monitoring 3D urban changes continuously over space and time. Using a novel method for processing scatterometer data into volumetric measurements (Nghiem et al., 2009), we compare contemporaneous scatterometer and lidar datasets collected for nine U.S. cities between 2003 and 2008 to determine the utility and accuracy of scatterometer data for urban volume estimates. By validating these estimations using high resolution lidar data, we offer a breakthrough in advancing LULC science from the confines of 2D analyses to incorporate the three dimensions. This validation also has the potential to impact demographic and socioeconomic studies that account for dynamic 3D urban processes worldwide. In the sections that follow, we describe our input data and methodological approach, the results of our statistical analyses, a discussion of the findings, and concluding statements.

2. Data processing and analytical methodology

The data workflow includes acquisition of raw scatterometer and lidar data that are processed and transformed to the same spatial resolution for comparison. The two datasets are then statistically analyzed using a series of correlation metrics (see Fig. 1). Further details on each of the datasets, the intermediary processing steps, and the analytical comparisons are provided in the subsections to follow.

2.1. Dense sampling method for radar scatterometer processing

Launched in June 1999, the QuikSCAT satellite carried aboard a SeaWinds scatterometer, which was an accurate and stable radar operated at the Ku-band frequency of 13.4 GHz. QuikSCAT measured global radar backscatter in the full scanning mode in 2000–2009 with

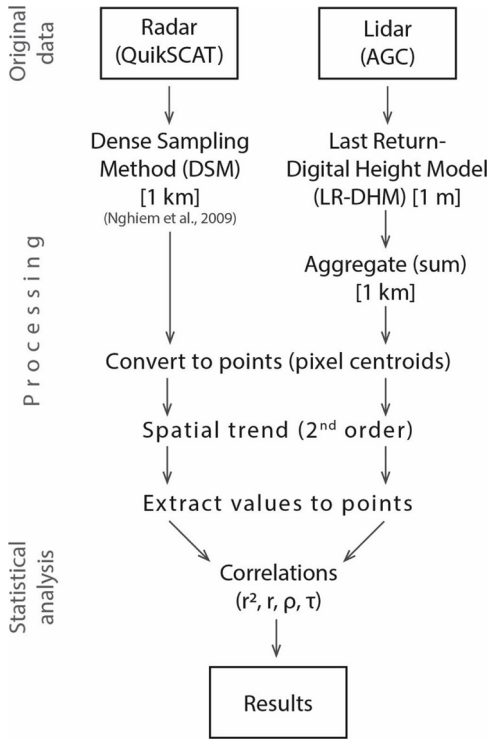

Fig. 1. Data processing and analytical workflow.

an accuracy of 0.2 dB (3σ) with a swath of 1800 km for the vertical polarization and 1400 km for the horizontal polarization. QuikSCAT acquired global data daily around 6:20 am and pm local time along ascending and descending orbits. The sensor footprint was approximately 25 km in azimuth and 37 km in range, and the sensor imaged 90% of the Earth's surface (QuikSCAT Mission, 2014). The QuikSCAT data used in this study were processed with the dense sampling method (DSM), which is a novel processing method developed to ingest a dense ensemble of scatterometer data samples (e.g., data collected every day for a year) and generate a high-resolution data array representing ground features with a grid size of approximately 1 km (Nghiem et al., 2009). The trade-off of the DSM method for converting scatterometer data into the high resolution map array is that the daily or near daily acquisitions of the QuikSCAT data must be aggregated to take advantage of the multiple view angles, leading to a temporal resolution of one year for the resulting data products. However, with respect to urban development, the annual time scale is appropriate for detecting change, so the DSM method is highly relevant for urban studies. A benefit of the DSM method is that the multiple azimuthal

views captured from the QuikSCAT platform are integrated together to allow all facets of buildings in a city to be observed regardless of the arrangements of building geometries (Nghiem et al., 2009).

The specific processing steps involved in the DSM are briefly described here, but readers are referred to Nghiem et al. (2009) for complete details. First, multi-azimuth backscatter slices, obtained by Doppler-compensate processing, are collocated into each 1/120° pixel (˜1 km at the equator) in latitude and longitude (World Geodetic System 84; WGS84). This process is repeated for backscatter ensembles at all locations in the study area. The thin slices at different azimuth angles, when taken together, resemble a rosette diagram, and thus, the transformation to calculate an ensemble average backscatter (in decibels, or dB) is referred to as the Rosette Transform. The equation (Equation 1) to compute an ensemble average of measured backscatter values, $\bar{\sigma}_0$ (dB), is given as

$$\bar{\sigma}_0 = \frac{1}{N\Gamma_A}\sum_{i=1}^{N}\iint_A dx\ dy\ G(\phi_i, x, y)\,\bar{\sigma}_0(x, y) + \frac{1}{N\Gamma_A}\sum_{i=1}^{N}\iint_A dx\ dy\ G(\phi_i, x, y)\varepsilon(\phi_i, t_i, x, y) \quad (1)$$

where N is the total number of scatterometer footprints centered within each pixel of the data grid, G is the radar gain, Γ_A is the integrated radar gain-area product, t_i is the time, ϕ_i is the azimuth angle, $\bar{\sigma}_0(x, y)$ at location (x, y) is the mean backscatter overall ϕ_i and all t_i, and $\varepsilon(\phi_i, t_i, x, y)$ is the zero-mean fluctuation part of the backscatter ensemble in each pixel. In Equation 1, the first term represents the mean part obtained by the Rosette Transform with the equivalent power gain (Equation 2) defined by

$$G(x, y) = \sum_{i=1}^{N}\frac{G(\phi_i, x, y)}{N} \quad (2)$$

and the second term represents the residual from the zero-mean fluctuating part of the radar backscatter.

As noted earlier, the tradeoff in generating the increased spatial resolution gained by the DSM processing technique that allows analyses in urban areas is that multiple daily acquisitions from the QuikSCAT sensor at different azimuthal angles are needed to create an annual 1 km product with a stable mean backscatter value while the fluctuation residual is minimized in each pixel (Nghiem et al., 2009). The resulting data product represents integrated intensity signal measurements, where higher intensities theoretically result from larger features such as skyscrapers and lower signals will result from smaller features such as residential houses. These theoretical links between the DSM transformed scatterometer data and actual urban features have not yet been validated though, and so comparison and validation with actual 3D data of ground objects are needed.

2.2. Lidar data acquisition and processing

Discrete-return, airborne lidar data provided the benchmark to compare and validate the volumetric discrimination capabilities of the scatterometer data. The U.S. Army Geospatial Center (AGC) produces several lidar-derived raster products at 1 m spatial resolution including a digital terrain model (DTM) representing the bare-Earth surface and two digital height models (DHM) created using the first-return of the lidar signal (FR-DHM) and the last return of the signal (LR-DHM), respectively. These DHMs capture the bare-Earth terrain as well as any surface features (e.g., buildings, vegetation, highway overpasses, etc.).

It should be noted that the term digital surface model and the commonly-used acronym 'DSM' are synonymous with digital height model and the acronym DHM (see Dong and Chen, 2018). However, since the acronym DSM is already used in this paper to refer to the Dense Sampling Method for processing the scatterometer data, we refer to the lidar models of surface features as DHMs to avoid confusion. Hereafter, lidar DHM references will be prefixed with FR or LR to denote whether they were created with the first or last return, respectively. The FR-DHM and LR-DHM lidar products vary slightly due to the sensitivity of lidar in capturing multiple returns, but both products have average horizontal and vertical accuracies of 0.5 m and 0.3 m, respectively. The LR-DHM typically contains less vegetation noise than the FR-DHM while still capturing building structures with equivalent accuracy; therefore, the LR-DHM provides the best option for computing building volume (Alharthy and Bethel, 2002; Zhao et al., 2017). As mentioned previously, lidar data availability can be spotty, and the AGC data availability varied for the different cities both in terms of the spatial and temporal extents.

We identified nine cities distributed across the U.S. for which lidar data existed that matched the time period of the QuikSCAT data collection for comparison (Table 1). While the spatial extent of lidar coverage for the different cities varied considerably, the scatterometer data is spatially continuous (over 90% of the globe), allowing us to exactly match the analysis extents constrained by the lidar data. The study cities vary in size both in terms of population and analysis extent (due to lidar data availability) and are located in diverse ecoregions. Additionally, since various modes of urban development may manifest differently in the scatterometer signal, we have included cities that formed and evolved under different urban growth processes. Both the scatterometer data and the lidar data were processed for the years and spatial extents for each city (Table 1).

Table 1
Study cities and data coverage.

City	Year	Analysis extent (km^2)	Population (2010)*
Atlanta, GA	2003	79	420,003
Austin, TX	2006	390	720,390
Buffalo, NY	2004	342	261,310
Detroit, MI	2004	347	713,777
Los Angeles, CA	2007	64	3,792,621
New Orleans, LA	2008	346	343,829
San Antonio, TX	2003	640	1,327,407
Tulsa, OK	2008	1,329	391,906
Washington, DC	2008	8297	601,723

* Population values given are for the entire metropolitan area and do not necessarily conform to the data analysis extent.

Following the methods described in Zhao et al. (2017), we calculated per-pixel volume of surface features from the lidar data by subtracting the DTM from the LR-DHM to create a normalized (relative) surface of features. Using building footprints provided by the AGC (see Cheuk and Yuan, 2009 and Zhao et al., 2017), we extracted only pixels corresponding to buildings, and in this manner were able to ensure the normalized surface features corresponded only to building heights within each city (Fig. 2). While methods are available for filtering out surface features from lidar returns (Meng et al., 2010), we specifically aim to test the capabilities of the QuikSCAT product for measuring building volume change in space and time, as urban growth in the built environment is a pressing social and ecological issue. Therefore, we employed building footprints to ensure precise measurements of urban buildings and reduce noise from non-building surface features such as vegetation.

Since the resolution of the lidar data (1 m) does not match the 1 km posting of the scatterometer data, the DHMs were aggregated to 1 km using a summation function to provide the total building volume within each 1 km pixel. Again, while the DSM processing of the scatterometer data does reduces the temporal resolution of the original acquisition, the benefits of having a globally-available product from which to measure urban volume change, to develop methods that can be implemented worldwide, and to foster replicable and repeatable analyses is a critical data need.

2.3. Analytical comparisons of radar scatterometer and lidar data

When using the DSM, the effective radar gain in a given pixel is not a square function, which results in signal contributions from neighboring pixels that are not fully suppressed (Nghiem et al., 2009). Within the 1 km pixels representing the DSM results, the total intensity value includes backscatter not only from manmade structures (e.g., buildings, bridges, freeway overpasses, walls, fences, power lines, towers) within the pixel footprint but also from pixels in the vicinity. Given the intricacies of the radar DSM product as well as the differences with lidar data, a direct pixel-by-pixel comparison with the aggregated lidar pixels would not be expected to result in consistent comparisons.

Moreover, calculating a gradient of built-up volume from the urban core to the periphery by taking the derivative of different discretized pixels is an ill-conditioned problem (Knowles and Renka, 2012) that will result in an excessively noisy and inconsistent spatial rate of change. To overcome these limitations and enable testing of the utility of the scatterometer data for capturing the trends across the urban

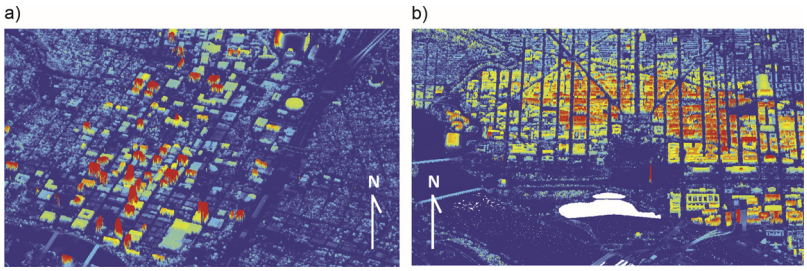

Fig. 2. 3D visualizations of lidar data for (a) Austin, TX and (b) Washington, DC.

extent, we transform both datasets into a comparable format using a second-order polynomial regression as a continuum rather than a pixel-by-pixel discrete comparison. This approach is founded on the concept of spatial trend (Şen, 2017), defined rigorously based on time-variable morphology, which overcomes the issue of taking derivatives of noisy discrete space and time data for quantitative assessment of spatial gradient and temporal rate of change of urban transformation. A key advantage of 3D spatial trend analysis of building volume is that it allows a continual and consistent assessment of the building gradient throughout the city. Spatial trend surfaces were generated using the Trend function in ArcGIS Pro (Esri, Redlands, CA).

In the spatial trend analysis in this paper, a single polynomial is fit to an array of data points (in this case, the pixel centroids extracted from the respective radar and lidar rasters) to model 3D patterns in the data. In most urban contexts, building heights exhibit a clear second-order trend that peaks in the urban core and dissipates as one moves away from the city center. The exact nature of this dissipating trend will vary from city to city according to the directional processes of urban development, and these nuances in the patterns of spatial development can be captured through the polynomial coefficients. It is important to note that cities rarely exhibit an abrupt cutoff of development except in certain cases where an urban growth boundary or other environmental constraints exist (e.g., a water body, which can be accounted for by existing water data products). Therefore, characterizing the spatial distribution of building volumes via a spatial trend is more appropriate and consistent for assessing the utility of the scatterometer data for capturing typical urban growth trends.

When utilizing the spatial trend approach, effects of over-interpolation—a situation in which a greater range of point values is estimated than expected—may cause extraneous negative values in the outer peripherals around a city. Thus, pixels with intensity (i.e. volume) values less than or equal to zero were excluded from the analysis to prevent these non-physical, extraneous effects in our analysis. The spatial trend method here used second order polynomials, and after the regressions were computed, the new pixel values were reassigned to the pixel centroids of the respective rasters for correlation analyses. For completeness, to verify the expected inconsistency in direct comparisons of the difference between DSM and lidar datasets, we also tested the relationship between the untransformed (untrended) DSM data and the untransformed and transformed lidar-derived volumes.

3. Results

Results with accompanying graphics are presented for two example cities: Austin, TX and Washington, DC (Fig. 3). These two cities were chosen because they have similar population sizes (720,390 for Austin and 610,723 for Washington, DC) but vary considerably in terms of their analysis extent (390 km^2 for Austin versus 8297 km^2 for Washington, DC) and thus capture different levels of suburban areas. Statistical results from the comparisons between the radar and lidar data are presented for all nine cities.

The graphical results for Austin and Washington, DC show that while the overall spatial distributions of volumes between the raw, 1 km radar DSM product (Fig. 3c) and the raw, 1 km lidar data (Fig. 3d) are similar, there are discrepancies between them, particularly at the smaller extent for cities such as Austin, that warrant the use of the spatial trend analysis. As expected, linear correlations between the raw, 1 km radar DSM product and the raw, 1 km aggregated lidar data (graphical examples in Fig. 3c-d) were not particularly strong (Table 2). Coefficient of determination (r^2) values for Pearson correlations varied [0.04, 0.32], with the highest value in Washington, DC and the lowest in New Orleans and Los Angeles, CA. Los Angeles was the only city for which the correlation was not statistically significant ($p < 0.05$). These low goodness of fit values are not unexpected due to the spatial aberrancies of the radar data acquisition and the differences between the lidar and radar acquisitions discussed above.

In addition to Pearson correlation and coefficient values, we also computed Spearman (ρ) and Kendall (τ) rank correlation coefficients to show the strength of the ordinal association between ranked values. These two rank-based correlation coefficients provide an indication of how well the radar dataset is capturing the general shape of the urban 3D form while the Pearson correlation coefficient (r) and coefficient of determination (r^2) provide a direct comparison between the two forms of urban volumetric values. Rank correlations (ρ, τ) show that the raw radar DSM product is moderately well suited for capturing general urban volume patterns—values ranged [0.2, 0.67] for ρ and [0.13, 0.5] for τ —particularly for mid-sized cities such as Austin and Washington, DC (Table 2).

Polynomial regressions transformed the data in a manner that allowed the general patterns in the density and directionality of development to emerge (Fig. 3e-f). In Austin, the highest densities are located in the northwest corner of the data extent, and the general directional trend of development occurs in the NNE-SSW direction. In Washington, DC, the highest densities are located in the center of the data extent, and the general direction for development occurs in a NE-SW trending direction. These regression-based visualizations are useful for highlighting broad trends in development patterns across cities. Direct comparisons between the trend surfaces also reveal strong relationships between the radar DSM product and the aggregated lidar data (examples in Fig. 3e–f). Specifically, Pearson correlations improved for all cities (Table 3), with r^2 values ranging [0.33, 0.98]. Highest correlations were again in Austin and Washington, DC as well as San Antonio, TX. Spearman (ρ) and Kendall (τ) coefficients also increased, ranging [0.61, 0.99] and [0.44, 0.91], respectively.

Given the strong correlations between the trended radar DSM product and the trended lidar data, we lastly tested whether the raw, 1 km DSM product is able to capture general trends of urban volume by comparing it to the trended lidar data (graphical examples of the two datasets are shown in Fig. 3d–e). Correlations remained statistically significant ($p < 0.01$) for all cities (Table 4), but correlation coefficients of determination (r^2) decreased in all cases, with values ranging [0.21, 0.75]. The highest r^2 value was for San Antonio, while the lowest was again for New Orleans. Spearman rank correlation coefficients (ρ) remained generally high though, with values ranging [0.5, 0.86]. These values again indicate that the DSM product is able to capture general urban volume trends.

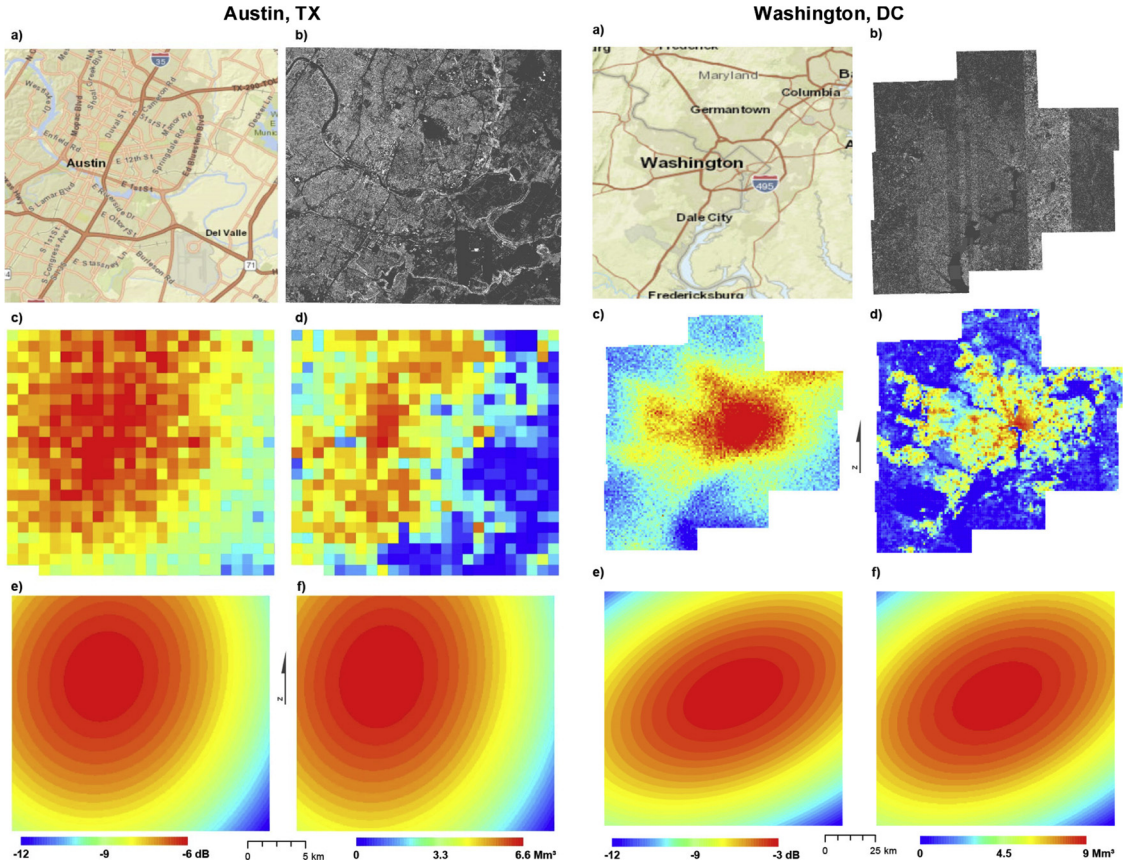

Fig. 3. City and data extents along with raw and processed data and polynomial trend visualizations for Austin, TX (left) and Washington, DC (right). Figure includes: (a) reference map, (b) 1 m lidar LR-DHM, (c) 1 km radar DSM, (d) 1 km aggregated lidar data (buildings only), (e) radar trend surface, and (f) lidar trend surface.

Table 2
Coefficients of association strength between the raw DSM radar data and raw lidar data correlations.

City	r^2	r	ρ	τ
Atlanta, GA	0.13	0.36	0.54	0.38
Austin, TX	0.21	0.45	0.64	0.50
Buffalo, NY	0.14	0.38	0.51	0.35
Detroit, MI	0.10	0.32	0.43	0.30
Los Angeles, CA	0.04*	0.20*	0.44	0.30
New Orleans, LA	0.04	0.19	0.20	0.13
San Antonio, TX	0.20	0.45	0.57	0.40
Tulsa, OK	0.26	0.51	0.59	0.40
Washington, DC	0.32	0.56	0.67	0.48

r^2: coefficient of determination in linear model; r: Pearson correlation coefficient; ρ: Spearman rank correlation coefficient; τ: Kendall rank correlation coefficient. All correlations significant with p-values < 0.01 unless otherwise noted (< 0.05*).

Table 3
Coefficients of association strength between the trended DSM radar data and the trended lidar data.

City	r^2	r	ρ	τ
Atlanta, GA	0.77	0.88	0.90	0.73
Austin, TX	0.98	0.99	0.99	0.91
Buffalo, NY	0.69	0.83	0.86	0.67
Detroit, MI	0.81	0.90	0.93	0.78
Los Angeles, CA	0.64	0.80	0.73	0.55
New Orleans, LA	0.33	0.57	0.61	0.44
San Antonio, TX	0.97	0.98	0.97	0.87
Tulsa, OK	0.84	0.92	0.93	0.77
Washington, DC	0.98	0.99	0.99	0.91

r^2: coefficient of determination in linear model; r: Pearson correlation coefficient; ρ: Spearman rank correlation coefficient; τ: Kendall rank correlation coefficient. All correlations significant with p-values < 0.01.

Scatterplots show that the relationships between the trended radar DSM product and the trended lidar product are linear (Fig. 4), and there is very little variation (i.e., point spread) at larger volumes, indicating that the radar DSM product is able to sufficiently capture heavy urban development. The increase in variation at lower volumes is likely due to a reduction in signal-to-noise ratio expected for pixels with weak radar backscatter (see Fig. 4a for Austin, TX and Fig. 4b for Washington, DC). These results provide evidence that the trended DSM product is appropriate for relative measures of urban volume. For completeness, we also examined the relation between the raw, radar DSM product and the trended lidar data. Results indicate a non-linear relationship with a large data spread for both Austin, TX and Washington, DC. These results confirm the difference between the raw radar DSM and the trended lidar data, and thus, as expected, it is inconsistent to compare these different data types.

4. Discussion

Having globally available datasets that are obtained with consistent temporal regularity is critical for assessing urban change and addressing the pressing social and environmental issues that often accompany urban growth. While lidar data provide an excellent resource for computing building volumes in pursuit of these goals, they are not currently captured systematically on a global scale at regularly repeated time intervals to permit the types of replicable and large-scale analyses without biases in space and in time. Even if regular lidar acquisitions do become reality (e.g., ICESat-2 currently operating and the Global

Table 4
Coefficients of association strength between the raw DSM radar scatterometer data and the trended lidar data.

City	r^2	r	ρ	τ
Atlanta, GA	0.33	0.57	0.58	0.42
Austin, TX	0.72	0.85	0.86	0.67
Buffalo, NY	0.38	0.61	0.64	0.45
Detroit, MI	0.52	0.72	0.74	0.54
Los Angeles, CA	0.26	0.51	0.50	0.35
New Orleans, LA	0.21	0.46	0.47	0.32
San Antonio, TX	0.75	0.87	0.83	0.64
Tulsa, OK	0.63	0.80	0.82	0.61
Washington, DC	0.66	0.81	0.86	0.66

r^2: coefficient of determination in linear model; r: Pearson correlation coefficient; ρ: Spearman rank correlation coefficient; τ: Kendall rank correlation coefficient. All correlations significant with p-values < 0.01.

Ecosystem Dynamics Investigation, or GEDI, instrument that launched on 5 December 2018), the lack of historical datasets make it difficult to monitor growth dynamics.

Validated with lidar data for nine cities of differing size and urban character, spaceborne scatterometer data processed using the DSM (Nghiem et al., 2009) provide an innovative data source for this purpose. We compared the DSM product derived from radar scatterometer data to high resolution lidar data acquired for nine cities in the U.S. that had been filtered to include only buildings. Our results show that the DSM procedure is successful in estimating building volume trends across multiple cities as intensity values from the trended DSM product correlate strongly and linearly with the trended lidar-derived building volumes. It is particularly important that the validation was carried out in terms of spatial trends, which are required to avoid noise amplification in taking derivatives of noisy discrete space and time data and to allow a consistent evaluation of the gradient in space and rate of change in time. Results are encouraging, but the knowledge and technical skills (e.g., DSM and lidar processing) required in addition to the computational cost likely demands the involvement of specialists. However, once these data are processed into high-level products such as urban building volumes, interdisciplinary research communities and application practitioners such as city developers and management agencies can make significant or breakthrough advances as the 3D global results will liberate urban observations from the confines of urban extent in 2D.

Another key finding is the linear relationship between radar-derived and lidar-derived building volume, which holds true for both low and high building volume values. In other words, the radar DSM product is equally able to estimate areas of low building volume as it is able to estimate areas of high volumes within a city extent regardless of whether the total building volume consists of a number of small buildings (e.g., in residential areas with small, wood frame houses surrounded by vegetation such as trees, bushes and grasses on yards, etc.) or a few large buildings (e.g., in an urban core with steel-frame skyscrapers and shopping centers surrounded by impervious surfaces like roads and parking lots). Thus, it is expected that this method remains valid over urban areas with various combinations of building types and vegetation conditions, not only in U.S. but also in other countries. Moreover, because we observed very little variation and no saturation at the top end, we expect this relationship will be transferable to large megacities.

The linear relationship also supports the use of the DSM method for identifying the physical boundaries of cities (Jacobson et al., 2015), and provides further evidence that the DSM method does not suffer from the same 'blooming' effects as other remotely sensed datasets (e.g., nighttime lights data). Identifying the point where the relationship between the trended radar and trended lidar data becomes invalid may provide another application for urban boundary delineation. However, in order to formalize exact thresholds, it is necessary to test multiple urban contexts including peri-, ex-, and suburban areas.

In the correlation plots (Fig. 4), we identified data separation at the low-end tails where slightly higher lidar values were found for relatively smaller DSM values. We found that these tails correspond spatially to the outermost extents of the fitted trend surfaces and are likely an artifact of the second-order polynomial. For instance, we uncovered these outliers for the Washington, DC region and noted that the discrepancies occurred mainly in the northeast corner of the result surface. This area is a region of transition between Washington, DC and the metropolitan region of Baltimore, MD, which is located approximately 75 km to the northeast. Instead of urban infrastructure decreasing monotonically moving away from Washington, DC, it begins to increase again moving toward Baltimore. The second-order polynomial constrained the range of possible values, contributing to the separation in the relationships in this region. This finding suggests that certain city patterns may require the use of higher-order trend models to capture the multimodal form, such as the Dallas–Fort Worth metroplex, which similarly encompasses two urban 'cores' (Nghiem et al., 2009), each with distinct clusters of high rise buildings separated in between by an area of lower density mixed land uses. While we were not able to test higher-order spatial trends given the geographic constraints of the lidar data availability, this is a promising area for future investigations, particularly those involving large, continental-scale investigations.

Insights into the appropriateness of the DSM product for building volume investigations can also be gained through examining the performances of the individual cities. In particular, New Orleans, LA and Buffalo, NY did not perform as well as the other cities. These two cities are both located adjacent to large bodies of water. New Orleans sits on the Mississippi River near the mouth of the Mississippi Delta and is also adjacent to Lake Pontchartrain. Buffalo is located on the eastern shore of Lake Erie at the head of the Niagara River, which travels north across Niagara Falls and empties into Lake Ontario just 26 km to the north. The lower correlations for these two cities indicate that water may have spillover effects on the DSM product. Therefore, we caution the use of

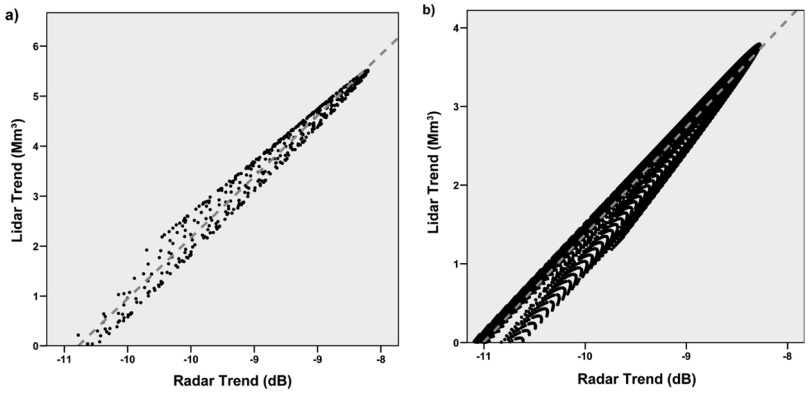

Fig. 4. DSM radar trend vs. lidar trend scatterplots for Austin, TX (a) and Washington, DC (b).

the DSM product for urban volume estimations in urban sectors located adjacent to large water bodies or containing considerable open water bodies, such as the Minneapolis-St. Paul area (Nguyen et al., 2018).

Atlanta, GA and Los Angeles, CA represent the two smallest urban extents we analyzed. Both performed moderately well, but these findings indicate that the method may perform better when the data covers a larger extent of the city. For Los Angeles in particular, the weaker correlations may also be due to the discrepancies introduced from highway overpasses, bridges, walls, and other infrastructure or superstructure signals that are received by the scatterometer instrument but are filtered out of the lidar data. These structures contribute a volumetric signal representing the total manmade structures, including both infrastructure and superstructures, which can be useful in estimating ambient population (including people on the move, at work, and at home) as opposed to census population (registered at residential addresses regardless of the true distribution of such population).

The implications of using the radar DSM product to perform large-scale, continuous monitoring of urban volume growth are considerable despite some limitations. The world's cities are growing at the fastest rate in recorded history, and the raw DSM values can be used in parallel with the trended data to identify intra-urban development hot spots or pockets of built-up development (Masetti et al., 2015). The DSM product can also be used to assess urban change at the district level within a city (Balk et al., 2018). Asia in particular is home to many of the world's megacities—defined as cities with population over 10 million—including Tokyo, Shanghai, Jakarta, Delhi, Seoul, Guangzhou, Beijing, Manila, Mumbai, and Shenzhen. A strong increase in the number of megacities is expected in this region by 2025.

The growth rate and development patterns occurring in these Asian megacities are unlike the more traditional gradient patterns seen in North American and European cities where dense urban cores dissipate into less dense sub- and ex-urban areas. Instead, in the rapidly growing megacities across Asia, high rise buildings are often placed immediately adjacent to what was previously rural farmland, creating a sharp gradient in LULC. Such rapid growth is often accompanied by serious socioeconomic, environmental, and geopolitical impacts. The results from this validation study show that the DSM product is a viable and innovative tool for measuring these types of urban growth dynamics.

A key consideration when relying on satellite data for LULC change analyses is mission continuation to ensure data consistency and accessibility. While the SeaWinds scatterometer aboard the QuikSCAT satellite, which provided the radar dataset used in this study, is no longer operational after a decade of global data acquisition, other scatterometers that may provide continual data are currently operating or planned for the future. For instance, NASA has access to the Oceansat-2 and Scatterometer Satellite-1 (SCATSAT-1) data through an agreement with the Indian Space Agency, and potentially similar scatterometer datasets are available such as from the Haiyang-2 satellite if released by the Chinese government in the future. The European Space Agency has an Advanced Scatterometer (ASCAT) with a C-band (5.255 GHz) on the Meteorological Operational (MetOp) satellite series (MetOp-A, MetOp-B, and MetOp-C), which, while different from the Ku band (13.45 GHz) sensing capabilities of SeaWinds, has been collecting global data since 2006.

The decadal time series of 3D building volume data derived from QuikSCAT radar backscatter can serve as a reference to calibrate other radar observations of 3D urban change using X-band synthetic aperture radar (SAR) systems with an added advantage of finer spatial resolution (10–100 m; Nghiem et al., 2018). X-band SAR systems such as the current TerraSAR-X launched in 2007, TanDEM-X in 2010, COSMO-SkyMed during 2007–2010, ASNARO-2 in 2018, and the future LOTUSat-1 and LOTUSat-2 missions (Pham, 2017) to be launched in 2021 and 2025 respectively, acquire radar data for 3D urban change observation worldwide and for long-term monitoring. As radar data come from different satellites though, data quality, consistency, and cross-calibration will need to be addressed carefully.

5. Conclusions

Through a comprehensive analysis of nine urban areas in the United States, this study showed through simple correlation analyses that the DSM-processed spaceborne radar data effectively correlates with airborne lidar data. The strong linear correlations between the trended DSM product and the trended lidar product indicate that the DSM method is practical and accurate for estimating urban volumes in the absence of large-scale lidar acquisitions. For all but one study city, r^2 values between the trended radar DSM data and the trended lidar DHM data were greater than 0.6 (with five cities greater than 0.8). These results establish the utility of incorporating satellite-based radar backscatter to quantify urban growth in 3D for cities across the world in the long term. This development is notable because of the difficulty and high costs associated with obtaining lidar data to conduct research requiring 4D data sources (3D building volume and 1D time change) to advance LULC science. Future work should further examine the utility of DSM radar data to model the urban environment and, more importantly, evaluate the sensitivity of the data product to differing building materials, its utility in modeling different types of urban areas (e.g., multi-core cities, larger continuous urban complexes containing multiple cities), and the potential to monitor global urbanization in 3D across an interdecadal time scale when data from international satellite missions are made available.

Acknowledgments

This material is based upon work supported by the National Aeronautics and Space Administration (NASA) under Grant No. NNX15AK42A issued through Oklahoma NASA EPSCoR to A.J. Mathews and A.E. Frazier in collaboration with the NASA Jet Propulsion Laboratory (JPL). The research carried out at the JPL (S.V. Nghiem and G. Neumann), California Institute of Technology, was supported by the NASA Land Cover/Land Use Change (LCLUC) Program and in part by the NASA Earth Science R&A Program. The authors wish to thank the Army Geospatial Center for providing the lidar data used in this study. Additionally, we appreciate the efforts of Emily A. Ellis, K. Colton Flynn, and Gustavo A. Ovando-Montejo in processing the lidar datasets. Availability of the data used in this study: QuikSCAT satellite data from the NASA JPL (https://podaac.jpl.nasa.gov) and lidar data from the Army Geospatial Center with permission (https://www.agc.army.mil).

References

Alharthy, A., Bethel, J., 2002. Heuristic filtering and 3D feature extraction from LIDAR data. Int. Arch. Photogr. Remote Sens. Spatial Inf. Sci. 34 (3/A), 29–34.

Alig, R.J., Kline, J.D., Lichtenstein, M., 2004. Urbanization on the US landscape: looking ahead in the 21st century. Landsc. Urban Plan. 69, 219–234. https://doi.org/10.1016/j.landurbplan.2003.07.004.

Balk, D.L., Nghiem, S.V., Jones, B., Liu, Z., Dunn, G., 2018. Up and out: a multifaceted approach to characterizing urbanization in Greater Saigon, 2000-2009. Landsc. Urban Plan. https://doi.org/10.1016/j.landurbplan.2018.07.009. forthcoming.

Cheuk, M.L., Yuan, M., 2009. Assessing spatial uncertainty of lidar-derived building model. Photogramm. Eng. Remote Sens. 75 (3), 257–269. https://doi.org/10.14358/PERS.75.3.257.

Dong, P., Chen, Q., 2018. LiDAR Remote Sensing and Applications. CRC Press, New York, USA.

Dong, P., Ramesh, S., Nepali, A., 2010. Evaluation of small-area population estimation using LiDAR, Landsat TM and parcel data. Int. J. Remote Sens. 31 (21), 5571–5586. https://doi.org/10.1080/01431161.2010.496804.

Freilich, M.H., Long, D.G., Spencer, M.W., 2003. SeaWinds: a scanning scatterometer for ADEOS-II-science overview. Proceedings of the IEEE International Geoscience and Remote Sensing Symposium. https://doi.org/10.1109/IGARSS.1994.399313.

Groisman, P., Shugert, H., Kicklighter, D., Henebry, G., Tchebakova, N., Maksyutov, S., Monier, E., Gutman, G., Gulev, S., Qi, J., Prishchepov, A., Kukavskaya, E., Porfiriev, B., Shiklomanov, A., Loboda, T., Shiklomanov, N., Nghiem, S., Bergen, K., Albrechtova, J., Chen, J., Shahgedanova, M., Shvidenko, A., Speranskaya, N., Soja, A., de Deurs, K., Bulygina, O., McCarty, J., Zhuang, Q., Zolina, O., 2017. Northern Eurasia Future Initiative (NEFI): facing the challenges and pathways of global change in the 21st Century. Prog. Earth Planet. Sci. 4 (41). https://doi.org/10.1186/s40645-017-0154-5.

Jacobson, M.Z., Nghiem, S.V., Sorichetta, A., Whitney, N., 2015. Ring of impact from the mega-urbanization of Beijing between 2000 and 2009. J. Geophys. Res. Atmos. 120 (12), 5740–5756. https://doi.org/10.1002/2014JD023008.

Jones, W.L., Schroeder, L.C., Boggs, D.H., Bracalente, E.M., Brown, R.A., Dome, G.J., Pierson, W.J., Wentz, F.J., 1982. The SEASAT-A satellite scatterometer: the geophysical evaluation of remotely sensed wind vector over the ocean. J. Geophys. Res. 87 (C5), 3297–3317.

Knowles, I., Renka, R.J., 2012. Methods for numerical differentiation of noisy data, variational and topological methods: theory, applications, numerical simulations, and open problems. Electron. J. Differ. Equ. 235–246 ISSN: 1072-669.

Lu, Z., Im, J., Quackenbush, L., 2011. A volumetric approach to population estimation using LiDAR remote sensing. Photogramm. Eng. Remote Sens. 77 (11), 1145–1156. https://doi.org/10.14358/PERS.77.11.1145.

Lwin, K.K., Murayama, Y., 2009. A GIS approach to estimation of building population for micro-spatial analysis. Trans. Gis 13 (4), 401–414. https://doi.org/10.1111/j.1467-9671.2009.01171.x.

Masetti, M., Nghiem, S.V., Sorichetta, A., Stevennazi, S., Fabbri, P., Pola, M., Filippini, M., Brakenridge, G.R., 2015. Investigating urban changes and environmental impacts in Italy. Eos 96 (21), 13–16.

Meng, X., Currit, N., Zhao, K., 2010. Ground filtering algorithms for airborne LiDAR data: a review of critical issues. Remote Sens. 2 (3), 833–860. https://doi.org/10.3390/rs2030833.

Mission, Quik S.C.A.T., 2014. Jet Propulsion Laboratory, California Institute of Technology (Accessed 10 July 2018) Retrieved from. https://winds.jpl.nasa.gov/missions/quikscat/.

Naderi, F.M., Freilich, M.H., Long, D.G., 1991. Spaceborne radar measurement of wind velocity over the ocean—an overview of the NSCAT scatterometer system. Proc. IEEE 79 (6), 850–866.

Nghiem, S.V., 2015. Global mega urbanization and impacts in the 2000s. IEEE International Geoscience and Remote Sensing Symposium 83–85. https://doi.org/10.1109/IGARSS.2015.7325703.

Nghiem, S.V., Balk, D., Rodriguez, E., Neumann, G., Sorichetta, A., Small, C., Elvidge, C.D., 2009. Observations of urban and suburban environments with global satellite scatterometer data. Isprs J. Photogramm. Remote. Sens. 64, 367–380. https://doi.org/10.1016/j.isprsjprs.2009.01.004.

Nghiem, S.V., Sorichetta, A., Elvidge, C.D., Small, C., Balk, D., Deichmann, U., Neumann, G., Njoku, E., 2014. Remote sensing of urban environments - the Beijing case study. In: Njoku, E. (Ed.), Encyclopedia of Remote Sensing. Springer, New York, NY, pp. 869–878.

Nghiem, S.V., Jacobson, M.Z., Mathews, A.J., Small, C., Balk, D., Sorichetta, A., Masetti, M., Stevenazzi, S., Oda, T., 2018. Four-dimensional satellite observations of Urban change for global-through-Urban nested modeling of environmental impacts. 38th EARSeL Symposium and 3rd Joint EARSeL LULC & NASA LCLUC Workshop.

Nguyen, L.H., Nghiem, S.V., Henebry, G.M., 2018. Expansion of major urban areas in the Great Plains from 2000 to 2009 using satellite scatterometer data. Remote Sens. Environ. 204, 524–533. https://doi.org/10.1016/j.rse.2017.10.004.

Pham, A.T., 2017. Vietnam roadmap to master satellite technology. Vietnam. J. Sci. Technol. Eng. 60 (1), 6–9. https://doi.org/10.31276/VJSTE.59(1).06.

Qiu, F., Sridharan, H., Chun, Y., 2010. Spatial autoregressive model for population estimation at the census block level using lidar-derived building volume information. Cartogr. Geogr. Inf. Sci. 37 (3), 239–257. https://doi.org/10.1559/152304010792194949.

Şen, Z., 2017. Spatial trend analysis. Innovative Trend Methodologies in Science and Engineering. Springer, Chamhttps://doi.org/10.1007/978-3-319-52338-5_6.

Seto, K.C., Dhakal, S., Bigio, A., Blanco, H., Delgado, G.C., Dewar, D., Huang, L., Inaba, A., Kansal, A., Lwasa, S., McMahon, J., Muller, D.B., Murakami, J., Nagendra, H., Ramaswami, A., 2014. Human settlements, infrastructure and spatial planning. Climate Change 2014: Mitigation of Climate Change. IPCC Working Group III Contribution to AR5. Cambridge University Press, Cambridge, UK.

Stevenazzi, S., Masetti, M., Nghiem, S.V., Sorichetta, A., 2015. Groundwater vulnerability maps derived from time dependent method using satellite scatterometer data. Hydrogeol. J. 23 (4), 631–647. https://doi.org/10.1007/s10040-015-1236-3.

Tsai, W.-Y., Nghiem, S.V., Huddleston, J.N., Spencer, M.W., Stiles, B.W., West, R.D., 2000. Polarimetric scatterometry: a promising technique for improving ocean surface wind measurements. Ieee Trans. Geosci. Remote. Sens. 38 (4), 1903–1921.

Zhao, Y., Ovando-Montejo, G., Frazier, A.E., Mathews, A.J., Flynn, K.C., Ellis, E.A., 2017. Estimating home and work population using lidar-derived building volumes. Int. J. Remote Sens. 38 (4), 1180–1196. https://doi.org/10.1080/01431161.2017.1280634.