Estimating Urban Canopy Parameters Using Synthetic Aperture Radar Data

INDUMATHI JEYACHANDRAN AND STEVEN J. BURIAN
Department of Civil and Environmental Engineering, University of Utah, Salt Lake City, Utah

STEPHEN W. STETSON
Global Environmental Management, Inc., Whistler, British Columbia, Canada

(Manuscript received 10 July 2008, in final form 19 October 2009)

ABSTRACT

This paper introduces a remote sensing–based approach to rapidly derive urban morphological characteristics using radar satellite data. The approach is based on the expectation that the magnitude of the synthetic aperture radar (SAR) backscatter can be related to urban canopy parameters (UCPs) describing the height, density, and roughness of buildings, trees, and other objects in cities. This hypothesis was tested with full-feature terrain elevation and SAR datasets for the Houston, Texas, metropolitan area. The backscatter magnitude was found to vary as expected across the city with higher backscatter values in the downtown tall building district relative to adjacent residential and commercial areas. To demonstrate the concept of using radar backscatter to estimate UCPs, relationships were derived between SAR backscatter and mean height, plan area fraction, and frontal area index of roughness elements (e.g., buildings and trees). In addition, SAR backscatter relationships were derived with roughness lengths computed using morphometric approaches. In all cases, the derived relationships were found to provide estimates of UCPs acceptable for use in meteorological models. Further testing using data from the Salt Lake City, Utah, metropolitan area validated the relationships and identified key areas for improvement for future research, including SAR instrument view angle differences and buildings split between SAR pixels.

1. Introduction

Mesoscale meteorological codes and transport and dispersion models are increasingly being applied in urban areas. Representing urban terrain characteristics in the models is critical for accurate predictions of wind flow, energy exchange, heating and cooling, and airborne contaminant concentrations. Urbanization of the models has introduced urban canopy parameterizations into the codes to simulate drag, momentum, and radiation trapping effects (Brown 2000). Many research and operational numerical weather, air quality, and atmospheric dispersion models have been equipped with urban canopy parameterizations [e.g., Coupled Ocean–Atmosphere Mesoscale Prediction System (COAMPS; Holt et al. 2002), Higher Order Turbulence Model for Atmospheric Circulation (HOTMAC; Brown and Williams 1998), fifth-generation Pennsylvania State University–National Center for Atmospheric Research Mesoscale Model (MM5; Otte et al. 2004; Lacser and Otte 2002; Dupont et al. 2004), Regional Atmospheric Modeling System (RAMS; Rozoff et al. 2003), and Weather Research and Forecasting model (WRF; Chen et al. 2004)]. Furthermore, studies have demonstrated the importance of using urban canopy parameterizations to produce results with high spatial resolution that accentuate variability, highlight important differences, and precisely identify critical areas (Dupont et al. 2004; Holt and Shi 2004; Holt and Pullen 2007).

Implementation of urban canopy parameterizations in mesoscale meteorological models requires defining surface cover and geometric parameters describing the geometric, radiation trapping, thermodynamic, and surface cover properties of urban areas. The spatial coverage of these urban canopy parameter (UCP) datasets extends thousands of square kilometers, and ranges from the ground surface to above the tallest structure in the city. The wide area coverage of UCP datasets combined with their high spatial resolution creates substantial database
requirements. For example, implementing an urban canopy parameterization in MM5 (Dupont et al. 2004) for a study of the Houston, Texas, metropolitan area required hundreds of gigabytes of airborne light detection and ranging (lidar) data to derive the UCPs, and months to compute more than 80 million parameter values (Burian et al. 2004b).

Recent advances in geospatial data collection and analysis have widely expanded the use of morphometric methods to calculate grid-based coverage of UCPs (Burian et al. 2004a). Automated methods have been developed to further advance the use of morphometric methods in modeling (Burian et al. 2006). One approach uses the MATLAB software package to compute building plan and frontal area densities, distribution of heights, roughness length, sky view factor, and more (Richens 1997; Ratti and Richens 1999; Ratti et al. 2002). Another approach uses geographical information system (GIS) software to process three-dimensional building and tree datasets to compute building height characteristics (mean, standard deviation, plan area–weighted mean, histograms), plan area density, frontal area density, height-to-width ratio, roughness length, displacement height, sky view factor, and several other UCPs (Burian et al. 2002, 2004a, 2006, 2008). In similar GIS-based work, Long et al. (2003) developed and tested the DFMap software to process building and vegetation data to compute selected UCPs.

Although automated morphometric approaches have been effectively used to derive UCPs for cities throughout the world, they are limited by their data and computer processing requirements. Recent UCP derivation advances have relied on the use of airborne lidar data to produce the full-feature and bare-earth digital elevation models used in the morphometric methods to compute the UCPs (Burian et al. 2003). However, airborne lidar and other datasets used to compute UCPs (e.g., stereographic paired images) in general are expensive to collect, difficult to manage and process because of their immense size, and quickly become outdated as urban areas change and expand (Burian et al. 2006). Consequently, one limitation to the continued advancement of model applications in urban environments is access to updated or new UCP databases. Therefore, UCP derivation methods are needed that can 1) efficiently provide UCP estimates covering a mesoscale modeling domain, 2) be applied in data-poor regions of the world, and 3) be updated periodically without extensive data collection efforts. The use of remote sensing data is proposed as a potential approach to meet these needs.

Stetson (2004) introduced a technique to estimate roughness lengths using synthetic aperture radar (SAR) data for the Houston metropolitan area in support of an air quality modeling project. For eight counties in the vicinity of Houston, the SAR backscatter value for each 150-m pixel in the dataset was converted into a roughness coefficient by computing the quotient of the individual pixel’s backscatter value and the mean backscatter of its respective land use–land cover (LULC) class. Within each LULC class the roughness coefficient was then used to derive the roughness length $z_o$:

$$z_o = RC \times z_o^*, \quad (1)$$

where RC is the roughness coefficient for each LULC (SAR backscatter of pixel/mean SAR backscatter of land use class) and $z_o^*$ is the index roughness length for the given LULC class, which is obtained from the literature. While this method resulted in improvements to the meteorological modeling output, particularly wind speeds, it was found that the relatively low pixel resolution of the SAR data and the generic base values from the literature allowed room for improvement of the process in the high spatial variability, both horizontal and vertical, of the urban LULC class. Differences were especially noted for commercial and residential land uses and the tall building districts (Burian et al. 2003, 2004b). Moreover, the technique was tied to LULC data, which are inherently a poor estimator of UCPs because of the wide variability of height and density of buildings and trees within urban land uses from city to city and within a city (Velugubantla et al. 2004). In this paper, an advancement of the preliminary SAR approach is introduced to improve the estimates of roughness length and to expand the approach to include other UCPs. The objective is to provide a proof-of-concept of the use of SAR to derive UCPs and to identify key areas of research needed to continue to advance the use of satellite data to derive UCPs.

2. Methods

a. SAR and urban morphology relationship

SAR operates by sending out a radar signal that hits the target and is scattered and reflected with a portion of the backscatter ($\sigma_v$) being received by the radar (Oliver and Quegan 2004). Radar backscatter is influenced by the radar system parameters and target characteristics. The major system parameters that influence radar signal return are wavelength or frequency of the signal, incident angle, and look direction. The target parameters that influence the backscatter are surface roughness, complex dielectric constant, slope angle, and orientation of the target.

The signal wavelength influences radar backscatter following from the Rayleigh criterion (Sabins 1987).
Targets appear rougher when exposed to radar signals with shorter wavelengths. L-band radar has been found to be more accurate for urban applications compared to C-band radar because vegetation effects are less pronounced (Xia and Henderson 1997). The incident angle ($\theta$) is the angle between the radar beam and a line perpendicular to the surface at the point of incidence (Henderson and Lewis 1998). It is considered to be a major factor influencing the radar backscatter and the appearance of objects in the imagery. Look direction refers to the orientation of the feature with respect to the flight line (Henderson and Lewis 1998) and it has been found to influence radar signal return (Xia and Henderson 1997). In urban areas, the orientation of streets and buildings relative to the look direction influences the radar backscatter (Sabins 1987). Also, metal features (e.g., fences and power lines) have bright signatures when they are oriented perpendicular to the look direction and are not significantly recognizable when they are oriented parallel to the look direction (Sabins 1987).

Among the target characteristics influencing backscatter, surface roughness is one of the dominant features. The roughness level of the target determines the backscatter value and dictates the tone of the image. In general, rough surfaces reflect more of the radar signal back to the sensor, and hence they appear bright in the imagery. Smooth surfaces reflect most of the energy away from the sensor resulting in a very low backscatter and they appear dark. The Rayleigh roughness criterion is an often-quoted guideline for determining whether a surface is smooth or rough (Sabins 1987). Rayleigh criterion states that surfaces can be considered “rough” and act as diffuse reflectors, if the root-mean-square (RMS) height of the surface variation exceeds one-eighth of the wavelength of sensing ($\lambda/8$) divided by the cosine of the local incident angle. Such surfaces scatter incident energy in all directions and return a significant portion of the incident energy to the radar antenna. Surfaces are considered “smooth” by the Rayleigh criterion and act as specular reflectors when their RMS height variation is less than approximately $\lambda/8$ divided by the cosine of the local incident angle. Such surfaces reflect most of the energy away from the sensor, resulting in a very low return signal. Typical values of radar backscatter in decibel (dB) range from $+5$ dB for rough surfaces to $-40$ dB for smooth surfaces.

Because of the ability of radar images to capture roughness, SAR data have been used by numerous researchers for estimating surface roughness (Evans et al. 1992; Dubois et al. 1995; Mo et al. 1988; Weeks et al. 1996, 1997; Deroïn et al. 1997; Shi et al. 1997). Most of the applications have been made to nonurban areas, while several have investigated backscatter–surface roughness relationships in urban areas (Xia and Henderson 1997; Franceschetti et al. 2002; Ainsworth et al. 2008). SAR images have been successfully used to detect and extract building features (Bryan 1975; Gamba and Housham 2000; Gamba et al. 2000; Quartulli and Dateu 2004; Ainsworth et al. 2008). In addition, building heights have also been retrieved using SAR (Luckman and Grey 2003; Simonetto et al. 2005). Previous studies identified that major challenges of SAR in estimating urban morphology are accounting for corners of buildings and at street intersections, which confuse the signal by incorrectly increasing roughness (Lillesand and Kiefer 2002; Sabins 1987). The approach presented herein follows from previous research estimating absolute surface roughness, but extends the application to the derivation of UCPs for meteorological models. In addition, the relationships are developed for grid cells ranging in size from 25 to 250 m, which avoids the problem of corners because the backscatter is averaged over an area containing numerous buildings, trees, and other roughness elements.

### b. Deriving SAR backscatter–UCP relationships

The development of the preliminary SAR backscatter–UCP relationships requires having a UCP dataset for a selected study area to intersect with the SAR backscatter dataset for the same study area. Four UCPs are used in this study—mean height, plan area fraction, frontal area index, and roughness length—to assess the potential of SAR data to estimate height, density, drag, and roughness parameters. The mean height is calculated by averaging the height of the roughness elements (e.g., buildings, trees) in each analysis area (i.e., model grid cell):

$$\bar{h} = \frac{1}{N} \sum_{i=1}^{N} h_i,$$  \hspace{1cm} (2)

where $\bar{h}$ is the mean roughness element height, $h_i$ is the height of the roughness element $i$, and $N$ is the total number of roughness elements. The plan area fraction ($\lambda_p$) is computed as the ratio of the plan area of the roughness elements to the plan area of the analysis area:

$$\lambda_p = \frac{A_p}{A_T},$$  \hspace{1cm} (3)

where $A_p$ is the plan area of roughness element at ground level (i.e., the footprint area) and $A_T$ is the total plan area of the region of interest. The frontal area index ($\lambda_f$) is computed as the total area of roughness elements projected into the plane normal to the approaching wind.
direction \((A_{\text{proj}})\) divided by the plan area of the study site \((A_T)\):

\[
\lambda_f(\theta) = \frac{A_{\text{proj}}}{A_T}.
\]

where \(\theta\) is the wind direction. The frontal area index computation using the Urban Morphological Analysis Processor (UMAP; Burian et al. 2005) is explained in more detail in the appendix.

Roughness length is a measure of the aerodynamic roughness of a surface, defined as the height at which the neutral wind speed extrapolates to a zero wind speed. Three widely used methods to estimate roughness length are listed below.

1) SIMPLE COEFFICIENT

The roughness length is often estimated by multiplying a coefficient by the mean height of roughness elements (buildings):

\[
z_o = f_o \bar{\zeta}_H,
\]

where \(\bar{\zeta}_H\) is the average building height and \(f_o\) is an empirical coefficient. The coefficient \(f_o\) in urban areas is typically 0.1, which is the value used for this study. Beyond the limitations of applying these equations to horizontally inhomogeneous urban areas, these equations also only hold for medium building density situations, as it is known that \(z_o\) varies with building spacing. Therefore, the equations in section 2b(2) are expected to provide better function fits to the SAR backscatter for a wider range of building and tree heights and density.

2) RAUPACH EQUATIONS

Raupach (1994) presented the following set of equations to determine displacement height and roughness length:

\[
\frac{z_d}{z_H} = 1 - \left( \frac{1 - \exp\left[-(c_{d1}2\lambda_f)^{0.5}\right]}{c_{d1}2\lambda_f^{0.5}} \right), \tag{6}
\]

\[
\frac{z_o}{z_H} = \left( 1 - \frac{z_d}{z_H} \right) \exp\left(-k \frac{U}{u_b} + \psi_k\right), \tag{7}
\]

where \(\psi_k\) is the roughness sublayer influence function, \(U\) and \(u_b\) are the large-scale wind speed and the friction velocity, respectively, \(c_S\) and \(c_R\) are drag coefficients for the substrate surface at height \(z_H\) in the absence of roughness elements and of an isolated roughness element mounted on the surface, respectively, and \(c_{d1}\) is a free parameter. Raupach (1994) suggested \(\psi_k = 0.193, (u_b/U)_{\text{max}} = 0.3, c_S = 0.003, c_R = 0.3,\) and \(c_{d1} = 7.5\). These parameter values were also used in this study.

3) MACDONALD EQUATIONS

The third set of equations introduced by Macdonald et al. (1998) incorporate the drag coefficient and displacement height into the expression for roughness length \((z_o)\):

\[
\frac{z_d}{z_H} = 1 + \alpha^{-1/2}(\lambda_p - 1) \quad \text{and} \quad (8)
\]

\[
\frac{z_o}{z_H} = \left( 1 - \frac{z_d}{z_H} \right) \exp\left\{ - \left[ 0.5\beta \frac{C_D}{k^2} \left( 1 - \frac{z_d}{z_H} \right) \lambda_f \right]^{0.5} \right\}, \tag{9}
\]

where \(\alpha\) is an empirical coefficient, \(C_D\) is a drag coefficient, \(k\) is the von Kármán constant, and \(\beta\) is a correction factor for the drag coefficient (the net correction for several variables, including velocity profile shape, incident turbulence intensity, turbulence length scale, and incident wind angle, and for rounded corners). Macdonald et al. (1998) recommended for staggered arrays of cubes that \(\alpha \approx 4.43\) and \(\beta \approx 1.0\), and these are the values used in this study. The Raupach and Macdonald methods may be applied at different wind directions. Although the roughness length values were computed using all three described methods and for different wind directions, the results were fairly similar. To eliminate redundancy, the results presented below only include the Raupach roughness lengths computed for an incident wind angle from the north (0°).

The UCPs were computed using the UMAP (Burian et al. 2005) by processing the lidar-based digital terrain model of the Houston study area at 25-m resolution (each grid cell is 25 m × 25 m). For each grid cell, the mean height, plan area fraction, frontal area index, and roughness length are computed. Also, for each grid cell, the SAR backscatter magnitude was determined, providing a backscatter value and corresponding UCP values for each grid cell. To assess the potential use of SAR backscatter to estimate UCPs, the variability of backscatter across the study area and as a function of land use was determined. Five urban land use types were included in the study. Three were taken from the National Land Cover Dataset (NLCD) classifications: low-intensity residential, high-intensity residential, and commercial. In addition, major roads and tall building districts were also included. The second part of the analysis focused on the relationships between backscatter magnitude and the UCP values across Houston as a function of the same five land use categories.
3. Case study

a. Study area

The study area for this project was the metropolitan area of Houston, located in the southwestern United States. Houston was selected for this study because it provides a wide variation of land use types to test the SAR–UCP relationships. Furthermore, the size of the Houston metropolitan area (>20,000 km², >4 million people) and the tall building district make it suitable as a case study location. A subset of the Houston metropolitan area covering approximately 210 km² was carefully chosen for model development and testing to be sure it included sufficient amounts of commercial and low- and high-intensity residential land uses and the tall building district.

b. Lidar data for UCP computation

The effect of trees on the UCPs was expected to be important for the study area. Therefore, a 1-m-resolution airborne-lidar-based digital terrain model (acquired November 2001), incorporating buildings, trees, and other roughness elements, was used to compute the UCPs. The dataset covers approximately 5200 km², from which a smaller area encompassing the tall building district and adjacent commercial and residential land uses (210 km²) was chosen for this study.

c. SAR data

Spaceborne Imaging Radar-C (SIR-C) data were obtained from the U.S. Geological Survey (USGS) EROS Data Center in CEOS format for a 5936-km² area centered on the Houston tall building district. The SIR-C data were collected on April 1994 with an incident angle of 58.1° using an L band resulting in a Multi Look Complex (MLC) dataset at 25-m pixel size. The L-band image was chosen as it is the preferred band for urban applications. Also, MLC data were used to reduce the noise in the image. The incidence angle of the image used was 58.1°, which is quite high and in the range preferred for urban applications. The discrepancy of SIR-C image acquisition (1994) and UCP data acquisition (2001) required the selection of locations to derive and test the regression relationships that were not expected to have changed significantly from 1994 to 2001. Aerial photos dating back to 1992 and 2000 were used to verify that the locations included in the study were unchanged.

After synthesis of the SIR-C image in ENVI 4.2 (Research Systems Inc. 2005), the image consisted of radar backscatter values ranging from 0 to −26 dB. The image was rectified using a high-resolution aerial image to overlay with the lidar-derived UCP raster. Georeferencing was performed in ArcGIS using 13 ground control

<table>
<thead>
<tr>
<th>Land use</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low-intensity residential</td>
<td>−18.7</td>
<td>−11.9</td>
<td>−15.6</td>
</tr>
<tr>
<td>High-intensity residential</td>
<td>−18.8</td>
<td>−8.6</td>
<td>−14.8</td>
</tr>
<tr>
<td>Commercial</td>
<td>−18.6</td>
<td>−1.3</td>
<td>−10.0</td>
</tr>
<tr>
<td>Tall building district</td>
<td>−14.0</td>
<td>−1.2</td>
<td>−9.1</td>
</tr>
<tr>
<td>Roads</td>
<td>−26.0</td>
<td>−19.0</td>
<td>−21.0</td>
</tr>
</tbody>
</table>
points selected from the road network line feature dataset verified with a high-resolution aerial image. The first-order polynomial transformation resulted in a root-mean-square error at the ground control points of 0.02 (a low value). A preliminary examination of the image indicates that the SAR image is able to represent rough areas containing tall buildings distinctly from the smooth surfaces (e.g., roads). The SAR image for Houston was manually verified with a road dataset and aerial photos. In the SAR image shown (Fig. 1), the backscatter values of the smooth surfaces range from $-19$ to $-26$ dB, whereas rough surfaces range from $0$ to $-18$ dB. Higher backscatter values indicate rougher building surfaces and lower backscatter values indicate smoother road surfaces.

d. Variability of SAR backscatter and UCPs with land use type

Specific portions of the SAR image representing low- and high-intensity residential and commercial land uses were chosen by examining the aerial photos and the NLCD digital image. In addition, the tall building district (i.e., central business district) was also included in this analysis. The radar backscatter values were analyzed to observe backscatter variation across the land use categories. The variability of the UCPs (mean height, plan area fraction, frontal area index, and roughness length) was also analyzed as a function of land use.

e. Relationships between SAR backscatter and UCPs

To highlight the relationships, redundant cells values were eliminated by randomly selecting a smaller number of cells to include in this analysis. Trial-and-error tests found that including approximately 30 grid cells in the relationships provided similar insight and fit to the relationships, but eliminated the overlap of data points. Therefore, 34 grid cells were randomly selected from the set of cells that had not had modifications to its built characteristics between 1994 and 2001. The backscatter and UCP relationships for this subset of 34 cells were analyzed to determine the fit of several functional relationships. In addition, an independent set of 26 cells was also randomly selected to test the ability of the derived relationships to estimate UCPs given only SAR backscatter magnitudes. In addition, the models were applied to an areal extent of 5 km² consisting of nearly 3500 commercial, high-intensity residential, and low-intensity residential land use units (25-m pixel size) and difference images of the UCPs were created.

4. Results

a. Variability of SAR backscatter and UCPs with land use type

Table 1 displays the SAR backscatter magnitudes for the five land use types. As expected the roads had the lowest backscatter, ranging from $-19$ to $-26$ dB. This was expected because the highways and major roads are the smoothest of the five land uses. The SAR backscatter magnitudes for the tall building district are the highest, ranging from $0$ to $-13$ dB. This also was expected. The commercial land use overlaps the range of the tall building district. This is also consistent with the morphological characteristics of Houston commercial areas. The Galleria shopping mall part of the study area was classified as a commercial area. This location contains numerous tall buildings, many of which are densely packed. The SAR backscatter in that location is relatively high. The two residential land use classifications do not display much difference in SAR backscatter, except for a small difference in the higher end of the range. The high-intensity residential contains a few more densely spaced and taller building areas (i.e., apartment complexes) that caused the range to display higher SAR values consistent with rougher surfaces.

<table>
<thead>
<tr>
<th>Land use</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low-intensity residential</td>
<td>0.01</td>
<td>0.96</td>
<td>0.49</td>
</tr>
<tr>
<td>High-intensity residential</td>
<td>0.14</td>
<td>0.90</td>
<td>0.57</td>
</tr>
<tr>
<td>Commercial</td>
<td>0.13</td>
<td>0.76</td>
<td>0.32</td>
</tr>
<tr>
<td>Tall building district</td>
<td>0.15</td>
<td>0.97</td>
<td>0.57</td>
</tr>
</tbody>
</table>

Table 2. Mean height (m) variation across four land use classes.

<table>
<thead>
<tr>
<th>Land use</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low-intensity residential</td>
<td>0.01</td>
<td>0.17</td>
<td>0.08</td>
</tr>
<tr>
<td>High-intensity residential</td>
<td>0.03</td>
<td>0.15</td>
<td>0.09</td>
</tr>
<tr>
<td>Commercial</td>
<td>0.02</td>
<td>0.16</td>
<td>0.09</td>
</tr>
<tr>
<td>Tall building district</td>
<td>0.04</td>
<td>0.89</td>
<td>0.30</td>
</tr>
</tbody>
</table>

Table 4. Plan area fraction variation across land use classes.

<table>
<thead>
<tr>
<th>Land use</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low-intensity residential</td>
<td>0.01</td>
<td>1.98</td>
<td>1.20</td>
</tr>
<tr>
<td>High-intensity residential</td>
<td>0.51</td>
<td>1.61</td>
<td>1.22</td>
</tr>
<tr>
<td>Commercial</td>
<td>0.49</td>
<td>1.98</td>
<td>1.20</td>
</tr>
<tr>
<td>Tall building district</td>
<td>0.30</td>
<td>5.55</td>
<td>2.29</td>
</tr>
</tbody>
</table>

Table 3. Mean height (m) variation across four land use classes.

<table>
<thead>
<tr>
<th>Land use</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low-intensity residential</td>
<td>0.20</td>
<td>1.98</td>
<td>1.20</td>
</tr>
<tr>
<td>High-intensity residential</td>
<td>0.51</td>
<td>1.61</td>
<td>1.22</td>
</tr>
<tr>
<td>Commercial</td>
<td>0.49</td>
<td>1.98</td>
<td>1.20</td>
</tr>
<tr>
<td>Tall building district</td>
<td>0.30</td>
<td>5.55</td>
<td>2.29</td>
</tr>
</tbody>
</table>

Table 5. Plan area fraction variation across land use classes.
Tables 2–5 display the variation of the four UCPs as a function of land use type. The roads land use was not included because there were few buildings and trees present (meaning the UCPs were essentially zero or very near to zero). The results confirmed earlier studies of UCP variation in other cities that found significant variation of UCPs within each land use category. Here we noted similar ranges for the two residential and the commercial classes, as well as the tall building district. The residential land classes and the commercial land classes had nearly the same mean height. These results confirmed our previous conclusion that land use data do not provide an adequate proxy for morphological parameters in urban environments. What was also confirmed was the significant difference in mean height noted in the tall building district.

The results for the three other UCPs—plan area fraction, frontal area index, and roughness length—further confirm our previous studies of urban morphology with the ranges and means of UCPs being similar between NLCD land classes. The plan area fraction variation was consistent with expectations of high-intensity residential having on average greater density than the low-intensity residential. The commercial land use also registered on average the lowest density. This occurred in Houston because of the presence of extensive surface parking lots, which tend to reduce the magnitude of the roughness parameters. The within-land class variation far exceeded the variation between land classes, suggesting the need to avoid using LULC datasets to set morphological parameters in meteorological models.

A selected set of grid cells was examined closer to determine the reasons for the high ranges. A few grid cells classified as low-intensity residential with high plan area fractions of 0.80–0.96 also had relatively high backscatter magnitudes from −15 to −13 dB. Examination of aerial photos indicated the cells had substantial vegetation and a few buildings with similar heights (approximately 7–10 m). These cells also had higher than average roughness lengths. The relatively high ranges of values of SAR backscatter and UCPs within the land classes suggest there is significant variability. Examination of aerial photos also suggested that the SAR and UCP values were related, with consistently higher values in rougher areas and smaller values in smoother areas. The examination of selected cells suggested that the SAR–UCP relationships may be stronger than the relationships of SAR and UCP to LULC.

b. Relationships between SAR backscatter and UCPs

SAR backscatter–UCP relationships were developed for the Houston study area, and then by land use class. The relationships for the Houston area and the residential and commercial land use classes did not provide an adequate fit with any of the functions tested. The problem identified was the relatively small range of SAR magnitudes in the residential and commercial categories (which occupied a large fraction of the Houston study area). Numerous data points plotted would result in a dense cloud. The result of applying these relationships would thus be a constant value of a UCP regardless of the SAR magnitude. This is essentially the approach used when the base LULC is used to assign the UCPs in mesoscale meteorological and dispersion models, except in this case some variability is captured. Relationships were established in the tall building district and adjacent

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Bias</th>
<th>CRE</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean height (m)</td>
<td>−1.30</td>
<td>0.34</td>
<td>9.48</td>
</tr>
<tr>
<td>Plan area fraction</td>
<td>0.22</td>
<td>0.87</td>
<td>0.52</td>
</tr>
<tr>
<td>Frontal area index</td>
<td>−0.01</td>
<td>0.34</td>
<td>0.12</td>
</tr>
<tr>
<td>Roughness length (m)</td>
<td>−0.24</td>
<td>0.40</td>
<td>1.01</td>
</tr>
</tbody>
</table>

Fig. 2. (a) Regression relationship between SAR backscatter and mean height and (b) scatterplot showing the results of the test of the relationship.
land area and those are the focus of this paper. It is most critical to represent the tall building district accurately for modeling because it has the highest roughness parameters that must be represented in the models to accurately simulate the wind flow and pollutant dispersion.

Future work will address improvement of the satellite–UCP relationships for areas outside the tall building districts.

The SAR–UCP relationships for the tall building district are analyzed qualitatively and quantitatively by UCP.

![Difference image of mean height for Houston study area.](image1)

![Aerial photo of Houston study area corresponding to the difference image area.](image2)
The relationships are also tested by estimating UCPs for an area not included in the derivation of the relationships. To quantify the performance of the relationships, three comparative statistics are computed: bias, cumulative relative error (CRE), and root-mean-square error (RMSE). The bias is computed as

$$\text{bias} = \frac{1}{n} \sum_{i=1}^{n} (\text{UCP}_i - \text{UCP}'_i),$$

where $\text{UCP}_i$ is the model-predicted UCP value for the $i$th of $n$ grid cells, $\text{UCP}'_i$ is the corresponding UCP value from lidar (i.e., the "true" value), and $n$ is the number of grid cells included in the analysis.

Bias is determined by averaging the differences between the SAR- and lidar-based UCP values for the set of grid cells included in the analysis. A positive value indicates that the model on average is overestimating the parameter being predicted. A negative value indicates that the model on average is underestimating the parameter. A zero value indicates the model is neither overestimating nor underestimating.

The CRE is computed as

$$\text{CRE} = \frac{\sum_{i=1}^{n} \text{absolute}(\text{UCP}_i - \text{UCP}'_i)}{\sum_{i=1}^{n} \text{UCP}'_i}.$$

CRE is the sum of the absolute differences between the SAR- and lidar-based UCP values for the set of grid cells included in the analysis. Values closer to zero indicate a better match between the SAR- and lidar-based UCP values. This is a relative measure dependent on the magnitude of the lidar-based UCP values. RMSE is computed as

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} (\text{UCP}_i - \text{UCP}'_i)^2}{n}}.$$

RMSE is essentially a standard deviation from the mean-square error (MSE). Therefore, the number is the amount of error that occurs from the mean value. A zero value indicates a perfect match between the SAR- and lidar-based UCP values.

### 1) SAR Backscatter–Mean Height Relationship

Figure 2a presents the SAR backscatter–mean height relationship (for the tall building district). An exponential function is fit to the data (several functional fits were tested including linear and power, but the exponential function had the best fit). The SAR backscatter magnitude, as expected, increases (becomes less negative) as the mean height increases. Thus, the taller the buildings in a grid cell the greater the scatter of the radar signal. The exponential relationship has an $R^2$ value of 0.7, indicating a relatively strong correlation. Testing of the relationship for the independent set of grid cells indicated acceptable accuracy (Fig. 2b). The accuracy was especially good for cells with higher mean heights, which were expected to be less accurate because fewer cells with high heights were included in the derivation of the relationship. Overall, the average percent difference between the SAR-estimated mean heights and the lidar computed mean heights is 7%. Analyzing the statistics of the radar backscatter–mean height relationship shown in Table 6 indicates the tendency to underestimate mean height (relatively high negative bias) and the RMSE is quite high.

The difference image of mean height for an areal extent of 5 km² consisting of nearly 3500 commercial, high- and low-intensity residential land use units (25-m pixel size) and the corresponding aerial photo image of the city are shown in Figs. 3 and 4, respectively. The blank areas in the difference image are areas where the UCP value would essentially be zero (area lacks trees and buildings) and the SAR backscatter indicates a smooth
surface. Most of the grid cells exhibit a difference between −3 and 3 m, especially on the right side of the image (an area of predominantly single-story buildings). Grid cells with absolute differences greater than 3 m (underestimate or overestimate) correspond mostly to the commercial areas with taller buildings. Cross-referencing the difference image with the aerial photo indicates that the underestimation is primarily due to the presence of tall buildings being split between analysis grid cells (i.e., SAR pixels). Taller buildings split between grid cells would cause the mean height estimated by the lidar to be elevated in those grid cells, while the SAR backscatter will be reduced if the remainder of the pixel is a flat smooth surface. In one case, a grid cell with an underestimated mean height had a tall building that was not split between cells. In that case the explanation for the underestimation was possibly a pool on the roof of the building (water absorbs the radar signal making the pixel appear smoother). A similar cross-referencing of cells having significant overestimation errors indicated that a significant fraction of the cells was covered by trees and bushes, which likely contributed to the higher backscatter not captured in the lidar-based UCP.

2) SAR BACKSCATTER–PLAN AREA FRACTION RELATIONSHIP

Figure 5a displays the SAR backscatter–plan area fraction relationship. Similar to the SAR backscatter–mean height relationship, an increase in plan area fraction (increase in roughness element density) leads to an increase in backscatter magnitude. A linear model provided the best fit for this relationship, with an $R^2$ value of 0.75. Testing of the relationship showed poorer accuracy than the SAR backscatter–mean height model (Fig. 5b). The average percent difference in plan area fraction for the test grid cells was 20%. The poorer fit for plan area fraction is suspected to be caused by pockets of density that may appear relatively homogenous and different levels of vegetation plan area that may impact backscatter values differently but register similar plan area fraction. The statistical measures from the testing of the SAR backscatter–plan area fraction relationship (see Table 6) indicate overestimation of plan area fraction on average (suggested by bias value), while the RMSE is relatively low, indicating a smaller cumulative error compared to the mean height. The difference image is shown as Fig. 6. Spatially the differences are more significant and distributed throughout the study area more evenly when compared to the differences in mean height (Fig. 3). Similar to the spatial distribution of differences in the mean height, the east side of the study area (right half of image) displays greater accuracy than the west side; most of the grid cells possessed a plan area fraction difference of less than 0.2. Cross-referencing the aerial photograph suggests that the reasons for the underestimation are splitting of buildings between grid cells and possibly building orientation effects. One reason causing overestimation appears to be the presence of
rather tall buildings in combination with isolated trees and bushes, which appears to contribute to a higher backscatter value that does not correspond to the plan area fraction of the roughness elements as calculated from the lidar data.

3) SAR BACKSCATTER–FRONTAL AREA INDEX RELATIONSHIP

Figure 7a displays the SAR backscatter–frontal area index relationship, which appears similar to the SAR backscatter–plan area fraction relationship. There is a direct relationship with increasing frontal area index being related to increasing backscatter magnitude. An exponential function, similar to mean height, provided the best-fit model of those tested \( R^2 = 0.7 \). Another similar trend to the mean height relationship is the increase in scatter about the fitted function as the magnitude of the frontal area index increases. This suggests that the use of SAR backscatter to estimate UCPs will produce greater errors in the extremely dense areas of cities, although this may be corrected through calibration in the future. Testing of the frontal area index model indicates that the performance is similar to the mean height (Fig. 7b). The testing suggests the SAR–frontal area index relationship underestimates the values (see bias and RMSE values in Table 6). The difference image shown as Fig. 8 is similar to the mean height, displaying a fairly accurate estimate of frontal area index on the east half of the study area (right side of image); most of the grid cells possessed a frontal area index difference of less than 0.05. Cross-referencing with the aerial photo indicated similar reasons for the mean height differences for underestimation (splitting of buildings) and overestimation (presence of vegetation).

4) SAR BACKSCATTER–ROUGHNESS LENGTH RELATIONSHIP

Figure 9a shows the relationship between SAR backscatter and roughness length (computed using the Rau-pach equations). The aerodynamically rougher surfaces produce greater SAR backscatter values. An exponential model again provided the best fit, but the fit was not as good as the other UCPs \( R^2 = 0.62 \). Testing of the relationship (Fig. 9b) suggests that the accuracy of the SAR-derived roughness lengths may be prone to underestimation compared to the morphometrically computed values (see bias in Table 6). It is important to note that the morphometric relationships themselves are prone to error and the equations used to compute roughness length provide a range of answers. Figure 10 presents the difference image for roughness length, with the same reasons for underestimation (split buildings) and overestimation (trees and bushes) being noted from the aerial photographs.

c. Validation for Salt Lake City metropolitan area

The radar backscatter–UCP relationships derived for Houston were tested for their generality by applying them to estimate UCPs in the Salt Lake City (SLC), Utah, metropolitan area. Again, SIR-C data were obtained from USGS for an area encompassing the SLC metropolitan area. The pixel size of the image is 25 m and the incidence angle is 35.3°; the processed scene range is 38 km and the processed scene azimuth is 100 km. The incidence angle is significantly smaller than Houston, which is expected to present a problem for the relationships, but this analysis is necessary to confirm this expectation and begin to formulate solutions to the problem.

The tall building district SAR–UCP relationships from Houston were applied directly to the tall building district and surrounding areas of SLC. From the SLC study area, 40 grid cells were randomly selected and the radar backscatter–UCP models were used to compute the four UCPs: mean height, plan area fraction, frontal area index, and roughness length. The estimated UCPs were compared with the morphometrically computed UCPs.
(the same process conducted for Houston). Performance assessment of the models was again based on the bias, CRE, and RMSE. The results shown in Table 7 indicate that the Houston relationships underestimate the mean height (indicated by a highly negative bias value), which is similar to the model performance in Houston. Grid cells with higher differences in mean height when examined in conjunction with aerial photos indicated that the disparity can again be attributed to the splitting of the buildings between the grid cells (the same problem observed in Houston).

The performance assessment of the SAR backscatter–plan area fraction model shows slight underestimates of the plan area fraction (low negative bias value shown in Table 7). The frontal area index and roughness length parameters had relatively poor predictions, both with high negative bias values (Table 7). The significant underestimation of the mean height and roughness length values in SLC is attributed to the lower incident angle (35.3°) of the SIR-C image used as compared to that of Houston (58.1°). Grid cells with higher mean height and roughness length were exhibiting lower radar backscatter values. This is because images with lower incident angles exhibit poorer range resolution, making the detection of structures difficult. Although this application to SLC provides a limited validation of the model performance, it confirms a key concern—that incidence angle differences will affect the applicability of SAR–UCP relationships to other cities.

5. Conclusions

This paper presented a new approach to estimate UCPs using satellite data. The ability of SAR backscatter to represent the variability of height, density, drag, and roughness elements in cities was tested for the Houston metropolitan area. The magnitude of the SAR backscatter was found to vary significantly across the city and between selected land use classes. The SAR variability was consistent with the variability of the UCPs, suggesting the potential for SAR backscatter to represent UCPs.

The SAR backscatter magnitudes in Houston were related to the height, density, drag, and roughness characteristics of the city by comparing lidar-derived mean height, plan area fraction, frontal area index, and aerodynamic roughness lengths to the SAR backscatter magnitude. This analysis confirmed the potential of radar backscatter to represent urban morphological characteristics in the tall building district, but indicated that the variability of SAR magnitude in residential and commercial areas was too small to establish accurate relationships in Houston and possibly other cities.

Functions were fit to the relationships of SAR backscatter and UCPs in the tall building district of Houston. Exponential functions provided the best fit for mean height, frontal area index, and roughness length, and a linear fit was the best for plan area fraction. The derived relationships were tested and found to provide
acceptable estimates of UCPs for use in meteorological models (based on how closely they match UCPs estimated using accepted approaches). A key problem identified in the assessment was the splitting of buildings between SAR pixels that acted to reduce the backscatter magnitude in the pixel. The relationships derived in Houston were tested for generality in Salt Lake City. Also tested was the influence of incidence angle on the performance of the model. The application to Salt Lake City confirmed the ability to rapidly apply the technique and produce reasonable values, but with significant errors for some UCPs. The significant errors corresponded to the UCPs expected to be most influenced by the incidence angle used to collect the image.

This paper introduced the concept of using SAR backscatter to estimate UCPs in urban areas. Although problems were identified and the concept needs significant improvement to be acceptable for widespread application, the preliminary results indicate improvement over the use of a constant literature value for UCPs in the models (Hasager et al. 2003). To advance the use of SAR and satellite data to estimate UCPs, future work must address the splitting of buildings between SAR pixels and the inability of SAR backscatter by itself to represent UCPs in residential and commercial areas. Standard reflectivity images and textural analysis are currently being tested in concert with SAR backscatter to provide improved multivariable regression relationships.

Acknowledgments. This research has been supported by the Defense Threat Reduction Agency. The suggestions of three anonymous reviewers significantly improved the quality of the manuscript.

Fig. 9. As in Fig. 2, but for roughness length.

Fig. 10. As in Fig. 3, but for roughness length.
To compute frontal area index, UMAP processes a rasterized building height dataset, which contains the building height in a two-dimensional array (see Fig. A1). To find the building faces exposed to the wind direction, UMAP steps through each raster pixel searching for a building height (a nonzero value). When a building height is located in a pixel and the immediately adjacent neighbor pixel (toward the approaching wind) does not contain a building, the pixel is used to compute the projected frontal area. For example, in Fig. A1 assume pixel \((i, j)\) has a nonzero height value indicating that it contains a building. For a westerly wind direction (from the left) the code would check pixel \((i - 1, j)\) for a nonzero height. If pixel \((i - 1, j)\) does not contain a building (it has a value of zero), then pixel \((i, j)\) is included in the calculation of frontal area. If the pixel does contain a building height and it is less than the height in the current pixel, the difference in height (i.e., exposed frontal area) is used in the calculation. Multiple neighboring pixels may need to be checked depending on the wind direction approach.

To use this method to find an exposed building surface area in the projected plane perpendicular to the approach wind direction requires defining cases for different wind directions and systematically checking them based on input from the user. The UMAP code considers four possible sectors of approach wind direction: \(0^\circ–90^\circ\), \(90^\circ–180^\circ\), \(180^\circ–270^\circ\), and \(270^\circ–360^\circ\) where \(0^\circ\) is north. Each pixel side has two possible wind directions and each “pixel side–wind direction sector” combination has a different calculation. For example, for a southwesterly wind direction in Fig. A1 the angle of the wind striking pixel \((i, j)\) is contained in the \(0^\circ–90^\circ\) sector. Based on the user-defined approach wind direction the algorithms corresponding to the relevant sectors and pixel sides are selected. Figure A2 illustrates the calculation of the length of the pixel side projected into the plane perpendicular to the approach wind direction. The projected length is multiplied by the building height in the pixel (or the exposed building height if partially blocked by an upstream building) to determine the projected frontal area. The frontal area index for the pixel is then found by dividing the frontal area by the plan area of the pixel. The user can select an analysis grid that is different from the pixel size, in which case the projected

![Fig. A1. The 3 x 3 section of a 2D building height array (north is at the top).](image)

![Fig. A2. Illustration of approach to compute projected length of pixel side exposed to approaching wind.](image)
frontal area within the analysis grid cell (which may contain several raster pixels) is summed and divided by the plan area of the analysis grid cell.

REFERENCES


