9.1

URBAN MORPHOLOGICAL ANALYSIS FOR MESOSCALE METEOROLOGICAL AND DISPERSION MODELING APPLICATIONS: CURRENT ISSUES

Steven J. Burian[1*], Michael J. Brown[2], Jason Ching[3], Mang Lung Cheuk[4], May Yuan[4], WooSuk Han[1], and Andrew T. McKinnon[1]

[1]Department of Civil & Environmental Engineering, University of Utah, Salt Lake City, Utah
[2]Energy and Environmental Analysis Group, LANL, Los Alamos, NM
[3]ASMD, ARL, NOAA, on assignment to U.S. Environmental Protection Agency
[4]Department of Geography, The University of Oklahoma, Norman, OK

1. INTRODUCTION

Accurate predictions of air quality and atmospheric dispersion at high spatial resolution rely on high fidelity predictions of mesoscale meteorological fields that govern transport and turbulence in urban areas. However, mesoscale meteorological models do not have the spatial resolution to directly simulate the fluid dynamics and thermodynamics in and around buildings and other urban structures that have been shown to modify micro- and mesoscale flow fields (e.g., see review by Bornstein 1987). Mesoscale models therefore have been adapted using numerous approaches to incorporate urban effects into the simulations (e.g., see reviews by Brown 2000 and Bornstein and Craig 2002). One approach is to introduce urban canopy parameterizations to approximate the drag, turbulence production, heating, and radiation attenuation induced by sub-grid scale buildings and urban surface covers (Brown 2000). Preliminary results of mesoscale meteorological and air quality simulations for Houston (Dupont et al. 2004) demonstrated the importance of introducing urban canopy parameterizations to produce results with high spatial resolution that accentuates variability, highlights important differences, and identifies critical areas. Although urban canopy parameterizations may not be applicable to all meteorological and dispersion models, they have been successfully introduced and demonstrated in many of the current operational and research mode mesoscale models, e.g., COAMPS (Holt et al. 2002), HOTMAC (Brown and Williams 1998), MM5 (e.g., Otte and Lacser 2001; Lacser and Otte 2002; Dupont et al. 2004), and RAMS (Rozoff et al. 2003).

The primary consequence of implementing an urban parameterization in a mesoscale meteorological model is the need to characterize the urban terrain in greater detail. In general, urban terrain characterization for mesoscale modeling may be described as the process of collecting datasets of urban surface cover physical properties (e.g., albedo, emissivity) and morphology (i.e., ground elevation, building and tree height and geometry characteristics) and then processing the data to compute physical cover and morphological parameters. Many of the surface cover and morphological parameters required for mesoscale meteorological models are also needed by atmospheric dispersion models. Thus, most of the discussion below is relevant to both types of modeling.

In this paper, the term urban morphological analysis will be used to define the component of urban terrain characterization concerned with the morphological parameters. Furthermore, the focus will be building morphological parameters; therefore, the term urban morphological analysis will refer exclusively to the task of inventorying, computing or estimating ***building*** morphological parameters. Several approaches to perform urban morphological analysis exist; however, all have in common three types of practice issues related to the uncertainty of (1) data, (2) parameter definitions and calculation methods, and (3) extrapolation techniques. The objective of this paper is to describe the state-of-the-practice of urban morphological analysis by reviewing the primary approaches presented in the literature and outlining and commenting on key aspects of the three types of practice issues listed above.

2. BACKGROUND

As described above, the urbanization of numerical models has introduced the problem of defining the urban canopy with a set of representative geometric, radiation, thermodynamic, and surface cover parameters. These urban canopy parameters (UCPs) defined broadly include aerodynamic roughness properties (e.g., roughness length), building height characteristics (e.g., mean height, standard deviation, histograms), building geometry characteristics (e.g., height-to-width ratio, wall-to-plan area ratio, complete aspect ratio), building volume characteristics (e.g. building plan and frontal area densities), radiation trapping parameters (e.g., sky view factor), surface cover properties (e.g., impervious surfaces, albedo), surface material properties (e.g., heat storage capacity, emissivity), vegetation type, height and geometry, and more. The types and attributes of datasets required for urban terrain characterization depends on the processes being simulated and the spatial and temporal scales of interest (Grimmond and Souch 1994). But in general characteristics at the micro-scale (10^{-2} to 10^{3} m) must be determined and aggregated or averaged to the grid cell size of the model. Representing this level of detail in mesoscale simulations is needed to accentuate the spatial heterogeneities of simulated surface temperatures, increase the value of turbulent kinetic

**Corresponding author address*: Steve Burian, 122 S. Central Campus Dr., Suite 104, Salt Lake City, UT 84112; E-mail: burian@eng.utah.edu

energy production, and increase the planetary boundary layer height to better represent urban areas (e.g., Dupont et al. 2004; Holt et al. 2002).

A handful of researchers over the years have pioneered the work on obtaining surface cover and morphological parameters for cities at the micro to neighborhood scale (10^2 to 10^4 m) (e.g., Ellefsen 1990/1991; Theurer 1993,1999). Grimmond and Souch (1994) were among the first researchers to present a geographic information system (GIS)-based technique for representing surface cover and morphological characteristics of the urban terrain for urban climate studies. Petersen and Parce (1994) and Petersen and Cochran (1998) presented software (ROUGH) for estimating the geometric parameters of buildings and structures in urban and industrial sites. Cionco and Ellefsen (1998) and Ellefsen and Cionco (2002) updated Ellefsen's (1990/1991) morphological inventorying procedure using a 100 m X 100 m grid cell size (and then a 50 m X 50 m cell size) for use in a high resolution wind flow model and included more characteristics of urban canopy elements in the database.

Grimmond and Oke (1999) reviewed several methods to define aerodynamic characteristics of urban areas using morphometric approaches. The work compared several methods to determine the roughness length, displacement height, depth of roughness sublayer and aerodynamic conductance based on measures of building and tree morphology. GIS was developed for 11 sites in seven North American cities and were used to characterize the morphological characteristics of the terrain and, using the morphometric equations, the aerodynamic parameters.

With recent advancements in data collection and management, digital 3D building and tree datasets have been developed for many locations in the U.S. providing an available data pool for automated and semi-automated analyses to compute UCPs. Computer software products including GIS and image processing tools have also been enhanced and now large areas covered by 3D digital building and tree datasets can be analyzed automatically to extract morphological information (e.g., building height and geometry characteristics, roughness length). Several researchers have developed automated and semi-automated computational procedures to process the 3D building and tree data to obtain UCPs. Ratti and Richens (1999), for example, built upon the initial effort of Richens (1997) to implement efficient urban terrain analysis algorithms in an image processing framework built within the MATLAB software package. Ratti et al. (2002) used the image processing approach to compute building plan and frontal area densities, distribution of heights, standard deviation, aerodynamic roughness length, and sky view factor for three European cities (London, Toulouse, and Berlin) and two U.S. cities (Salt Lake City and Los Angeles). The results illustrated the roughness length differences between European and U.S. cities.

Burian et al. (2002) presented an approach using GIS to process 3D building datasets to compute building height characteristics (mean, standard deviation, plan-area-weighted mean, histograms), plan area density, frontal area density, wall-to-plan area ratio, complete aspect ratio, height-to-width ratio, roughness length, and displacement height. The UCPs were calculated for each grid cell in predefined grid meshes and the average values for each land use type were also determined. This automated GIS approach was used to compute UCPs for Los Angeles, Phoenix, Salt Lake City, Portland, Albuquerque, Oklahoma City, Seattle, and Houston. Comprehensive reports are available for each city (e.g., Burian et al. 2002; visit www.civil.utah.edu/~burian for copies of the reports). The GIS approach has recently been expanded to include analysis of 3D vegetation, other 2D GIS datasets (e.g., roads) and multi-spectral imagery to compute an expanded set of parameters including surface cover fractions, impervious surfaces, sky view factor, predominant street orientation, and more (Burian et al. 2003). The processing capability continues to be enhanced and is currently available as a graphical user interface tool using a VBA macro for the ESRI ArcGIS software package.

Long et al. (2002) developed and tested the DFMap software to process vector building and vegetation data (BDTopo) available from the French National Geographic Institute (IGN). With DFMap, a user can select a cell size and wind direction to compute a series of morphometric and aerodynamic roughness parameters. Long (2003) used the DFMap software to compute morphological statistics and define urban land use/cover types using an unsupervised k-means analysis. The analysis tools and approach were tested using data for the city of Marseille. Long et al. (2003) extended the DFMap application by incorporating the analysis of multi-spectral and panchromatic imagery in an attempt to improve the definition of urban surface cover.

Urban morphological approaches have evolved in less than two decades from detailed inventorying using aerial photographs and extensive field surveys to computationally intensive processing of integrated 2D and 3D GIS and multi-spectral imagery datasets. The users of such data have also expanded to include federal agencies (e.g., LANL, LLNL (e.g., Chin et al. 2000), DTRA (Pace 2002), U.S. Environmental Protection Agency (e.g., Ching et al. 2002; Dupont et al. 2004), U.S. Army and Naval Research Laboratories (e.g., Cionco and Luces 2002; Holt et al. 2002)), university research centers (e.g., University of Houston Institute of Multidimensional Air Quality Studies (IMAQS) (Daewon Byun, personal communication)), university researchers (e.g., Rozoff et al. 2003), and private consultants (Haider Taha and Robert Bornstein, personal communication). During this time, the basic concept of defining UCPs for given homogeneous areas (e.g., land use, land cover, terrain zone) or model grid cells has remained the same, but new developments have improved the methods used to calcualte the UCPs. Even with the advancements, the need for datasets covering large areas and extensive data management and processing requirements limit the ability to derive gridded parameter datasets for entire mesoscale model domains. Further developments in data collection and processing are needed. In addition, there is a need to further refine urban morphological analysis methods by standardizing processes and reducing the uncertainty of methods. The following section outlines several issues that are currently of interest to those conducting urban morphological analyses for mesoscale meteorological or atmospheric dispersion modeling applications. Resolving these issues will contribute to improvements of analysis methods and likely increase the level of accuracy of calculated UCPs.

3. MORPHOLOGICAL ANALYSIS ISSUES

Urban morphological analysis methods have matured during the past two decades, but there still remain several operational issues to be resolved. Three primary types of issues facing those deriving UCPs are:

(1) **Data Issues.** What data are available, what are the available levels of resolution, accuracy and detail, and what levels of resolution, accuracy and detail are necessary for UCP calculations?

(2) **Definition and Calculation Issues.** What are the current UCP calculation methods and what are the effects of ambiguity and uncertainties in parameter definitions and calculation methods on UCP values and mesoscale meteorological and dispersion simulation results?

(3) **Extrapolation Issues.** What are the most accurate means to extrapolate UCPs to mesoscale meteorological and dispersion model domains?

The salient points associated with the three types of issues listed above are described in the following three subsections.

3.1 *Data Issues*

3.1.1 *What Data are Available?*

The building elevation data used to define the urban morphology is usually obtained by:

- Analyzing aerial photographs to estimate building data (time consuming, feasible only for very small areas)
- Performing ground surveys (time consuming, feasible for very small areas; elevation can be more accurately obtained than by analysis of aerial photographs)
- Analyzing stereographic images (can be time consuming to perform manual digitization; elevation can be accurately estimated)
- Analyzing airborne LIDAR data (requires managing and processing massive datasets; newness of LIDAR technology presents several problems; elevation measurements can be obtained with vertical and horizontal accuracies of 15 cm RMSE)
- Obtaining/purchasing datasets from government or municipal agencies (may be outdated or lack information necessary to compute all UCPs)
- Purchasing datasets from commercial vendors (potentially high cost; enhanced data products can be delivered)

The data are obtained in two basic forms: raster and vector. Raster data are usually square-grid cells with one elevation attribute per cell. Vector data use one or more polygons to represent the building footprint and rooftop. Vector data can represent the geometry and details with higher precision than raster data. In addition, vector data can represent multiple building layers and can contain multiple attributes per polygon (e.g., height, roof pitch, color, material). Building polygon data in vector form can be obtained from municipal governments and commercial vendors. Data products range from the raw stereographic pairs or airborne LIDAR data to the finished end products of building polygon vector or full-feature raster digital elevation model (DEM) datasets. Other options for obtaining building data are federal government agencies. For example, a potential major source of building morphology data in the U.S. is the so-called "120 cities" project. This project involves a consortium of federal agencies to collect and prepare airborne LIDAR databases for cities for domestic preparedness applications. Unfortunately, the current availability of the data and the ultimate distribution policy is uncertain.

3.1.2 *How Accurate are the Data?*

A critical question to be asked of all available building datasets is the accuracy. The primary errors possible include:

1. Buildings may currently exist, but are not included in the dataset
2. Buildings may not currently exist, but are included in the dataset
3. Several individual buildings may be represented as a single building
4. A single building may be represented as several individual buildings
5. Building outline may not accurately represent the actual building shape
6. Location of building polygon or parts of the polygon may be inaccurate
7. Building heights may be inaccurate

The causes of these data errors could be operator error during extraction, poor quality assurance/quality control (QA/QC) procedures, limitations of building feature extraction methods, changes of morphology during data collection, and the age of the building dataset (which may cause the data to not represent the morphology for the time period of interest). To illustrate several of the common building morphological dataset errors, Figure 1 was created showing red outlines representing building footprints extracted using an automated LIDAR DEM processing technique overlaid onto an aerial photo. Of the potential errors listed above, numbers 3, 4, 5, and 6 can be directly observed in the figure. Although not observable the other three errors were also contained in this sample dataset.

Figure 1. Sample residential area in Salt Lake City. Red outlines represent a vector dataset derived through an automated analysis of a full-feature DEM.

Airborne LIDAR has emerged as one of the primary sources for the development of 3D urban models (Priestnall et al. 2000). In most cases, both horizontal and vertical resolutions of LIDAR data can be down to one meter. Consequently, the high resolution may be misunderstood as high accuracy. It should be understood that LIDAR data processing is a nontrivial task. Three general approaches have been developed for urban 3D data analysis using LIDAR data, categorized as using an edge operator, mathematical morphology, and height bins, respectively (Zhou et al. 2004). In practice, LIDAR data processing may involve all three approaches. To assess possible LIDAR data inaccuracies a comparison was made between building heights derived from LIDAR data for Oklahoma City and field observations. The LIDAR data was acquired in late October 2001 by an Optech Airborne Laser Terrain Mapper (ALTM) 2033 sensor with a differential Global Positioning System. The Joint Precision Strike Demonstration (JPSD) Program Office of the U.S. Army processed the raw LIDAR data and provided data of LIDAR estimated building heights in downtown Oklahoma City. Preliminary findings identified four categories of significant differences between building heights estimated from LIDAR data and those from field observations. Significant differences were defined as those greater than 5 meters, which represents at least one floor of the structure. The four types of differences are briefly described below; more details and examples will be included in the presentation.

The first LIDAR inaccuracy found from the Oklahoma City analysis is common to all data collection and building extraction approaches. It involves the sate of the building morphology changing from the time of data collection. Buildings that were recorded in the dataset may have been demolished or new buildings may have been built in areas where buildings previously did not exist. The second type of LIDAR inaccuracy was observed at locations where buildings were constructed of highly reflective materials. Figure 2 shows two buildings that had heights under-estimated by the LIDAR data by more than 5 m. The building shown on the left side of the figure is a greenhouse with a curved roof-top. A significant portion of the south end of the building has an under-estimated height. The second example is a building nearly completely glass-covered on the south side. The end of the building covered by glass is noted to have a height difference of greater than 5 m. The third type of difference noted was associated with buildings having complex or narrow structures, e.g., passages, gaps, platforms, extensions. The LIDAR-derived building heights contained both under- and over-estimations in the vicinity of such complex building elements. And the fourth difference between the LIDAR-derived heights and field measured heights was noted in the vicinity of vegetation adjacent to buildings. Trees adjacent to buildings that overhang the building rooftop will cause the LIDAR-derived building height to be over-estimated because the LIDAR does not differentiate between objects.

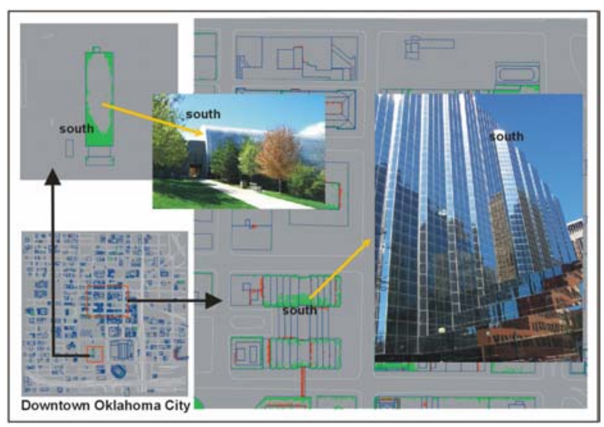

Figure 2. Examples of association between building materials and under-estimated building heights based on LIDAR data.

While four categories of differences between LIDAR estimates and field observations have been presented here, these differences represent preliminary findings and are by no means exhaustive. Current work focuses on trying to understand the physical causes of the observed uncertainties associated with the LIDAR-derived building heights. It is also expected that other uncertainties tied to building morphological characteristics will be noted. Inaccurate data is prevalent but the effect of data inaccuracies on UCP values and numerical model results is being studied. Results will be included in the presentation.

3.1.3 *What are the Available Levels of Data Detail?*

The detail of the data used to compute the UCPs may also influence the resulting UCP value. In terms of detail, the necessity to represent precise building geometry and ancillary attached structures is the primary issue. For example, a building dataset might represent the change in building geometry with height or include rooftop structures. The alternative to this level of detail is to simply represent each building with a single polygon that has a uniform shape with height. An example of the difference in detail for a small section of downtown Salt Lake City is shown in Figure 3. The top part of the figure shows a CAD dataset with a very high level of detail included, while the bottom part of the figure shows a GIS shapefile with a relatively low level of detail. Visual comparison of the top and bottom parts of the figure suggests that the low-level of detail captures the general form of the location. Therefore, average UCPs for a 1-km model grid cell size are likely to not be affected by approximating each building as a uniform polygon with a single height. This needs to be quantitatively confirmed.

One question to consider if a uniform polygon with single height attribute is to be used to calculate the UCP is the choice of average or maximum height as the attribute used in the calculation. For some UCPs (e.g., sky view factor), rooftop structures are probably not important and in areas with tall buildings, the rooftop structures are also probably negligible. In these cases, the choice of average or maximum building height will probably not impact the UCP value significantly. In residential land uses however the use of average or maximum height may alter the UCP values. For example, consider Figure 1, which contained an aerial photograph of a residential block in downtown Salt Lake City consisting of predominantly single-story, single-family homes with basements and pitched roofs. The buildings were extracted from LIDAR data using an automated approach and the heights were determined in two ways: (1) finding the average height of raster cells inside each building polygon boundary and (2) selecting the maximum raster height inside each building polygon boundary. Using the mean height for each building, the mean and standard deviation of building height for the entire block were calculated to be 3.8 m and 0.7 m, respectively. However, using the maximum height of the rooftop, the mean and standard deviation were found to be 5.5 m and 1.5 m. These differences are significant for the UCP value (~45% and 100% differences), but the significance of the magnitude of differences is uncertain for mesoscale meteorological or dispersion model results considering that the UCP value will be aggregated with other data for a grid cell size on the order of ~1 km^2. Further analysis is needed.

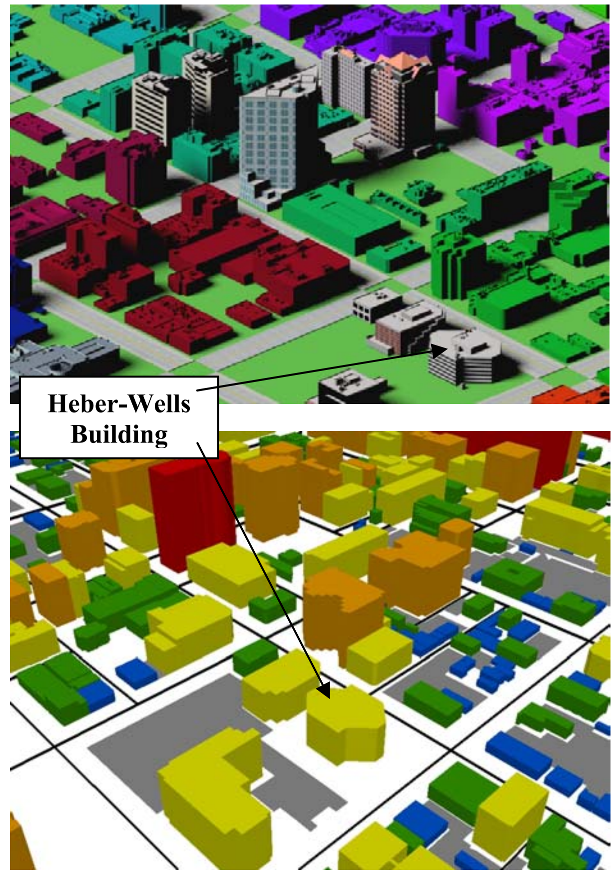

Figure 3. Comparison between a high-detail building dataset (top) and a coarse dataset (bottom) for small section of downtown Salt Lake City. The Heber-Wells Building is identified in both for frame of reference.

3.2 *UCP Definition and Calculation Issues*

A second concern facing those deriving UCPs involves the decisions of what are the precise UCP definitions, what calculation methods and tools to choose, what formats of input data to expect, and what outputs need to be produced. The following questions are pertinent to these concerns.

3.2.1 *What is the Effect of Ambiguous UCP Definitions on Calculated Values?*

The definition of the UCP may be ambiguous and thus not account for all potential real-world building arrangements. For these circumstances calculation of the UCP will be subjective because the individual will need to decide whether to ignore the unforeseen situation, implement a simplified calculation approach, or develop a modified calculation approach accounting for the encountered circumstance. The potential impact of UCP definition ambiguity on UCP values will be considered herein using mean building height. The presentation will contain the assessment of other UCPs including height-to-width ratio and frontal area index.

Mean building height is seemingly a well-defined

UCP. However, there are several ways in which the mean can be computed. The most common are: simple average, weighting by plan area (top or bottom or average), or weighting by frontal area. Based on analysis of real building data the values can be significantly different depending on the definition chosen. For example, consider Table 1 which lists the computed mean building heights using a simple average and weighted by plan area for the downtown core areas of eight cities in the U.S. The plan area weighted mean building height ranges from being approximately equal to the simple mean building height to being more than 40% greater. What these differences mean in terms of mesoscale simulation results must be determined.

Table 1. Comparison of mean building heights for downtown core areas of U.S. cities.

City	Mean Building Height (m) – Simple Average	Mean Building Height (m) – Plan Area Weighted
Albuquerque, NM	15.2	24.6
Houston, TX	22.8	36.0
Los Angeles, CA	45.0	44.1
Oklahoma City, OK	19.4	25.4
Phoenix, AZ	17.2	21.2
Portland, OR	18.1	18.5
Salt Lake City, UT	23.6	41.5
Seattle, WA	21.1	31.9

3.2.2 *What is the Effect of Ambiguous UCP Calculation Methods?*

In addition to the ambiguities associated with UCP definitions, there are ambiguities associated with UCP calculation methods. Most notable is the many equations available to compute the aerodynamic roughness parameters (roughness length and displacement height), with confusing guidance on relative applicability and accuracy. Grimmond and Oke (1999) compared many of the methods to compute the aerodynamic roughness parameters and could not conclusively identify an order of preference, although several of the more intuitively accurate methods were highlighted as recommended. Additional studies must be conducted to investigate the differences between the results produced by the various approaches for real cities and how the differences affect mesoscale meteorological and dispersion model results.

3.2.3 *What are the Advantages and Disadvantages of UCP Derivation Approaches?*

Several approaches have been presented in the literature to derive UCPs (see Background section above):

1. Conduct a field survey to measure building geometry and height information and record construction materials.
2. Analyze high-resolution aerial photographs to estimate building geometry and height information and record construction materials.
3. Perform automated or semi-automated image analysis of remotely sensed data to compute the building information.
4. Perform automated or semi-automated analysis of full-feature digital elevation models (DEM) or vector representations of buildings using GIS or image processing software.

Conducting a field survey can provide a high level of detail, yet it is time consuming/labor intensive and requires being in the location of the buildings. It is feasible for analysis of very small areas of cities. Analysis of high-resolution aerial photographs is also time consuming/labor intensive, but it does not require the analyst to be present at the building location. It is however similar to field surveys in that it is only applicable for small areas of cities. Automated or semi-automated analysis approaches are the most time efficient and once the codes are written to perform the automated calculations, labor requirements are limited (computer time is the only requirement). However, removing the trained analyst from the interpretation and measurement of building characteristics eliminates subjectivity at the cost of reducing quality assurance and quality control and oversight. In addition, simplified approximations to the calculations may be necessary to execute an automated approach and this may further reduce the accuracy of the results.

Despite the reduction in accuracy potentially caused by using an automated analysis approach, it is the only way to derive UCP coverage for large areas in a reasonable amount of time. Besides being able to process large areas, the image and GIS processing approaches have other advantages including:

- Spatial data in GIS compatible formats (e.g., roads) is readily available for incorporation into the analysis and for map making.
- The quality and availability of building morphological data is continuously improving and the automated approaches are specifically designed to work with these datasets.
- GIS and image processing software continues to improve and computer processing speed continues to be increased, all of which enhances the ability of automated approaches to provide more accurate UCP values more efficiently.

Some questions may arise regarding the use of image processing versus GIS software. The image processing software may provide more efficient computation than the GIS, but with the speed of today's computers the difference in time is negligible compared to the time to be invested in the modeling and analysis activities.

3.2.4 *How Important Are Trees for Accurate UCP Values?*

Because of their relative importance in downtown areas, building morphological data have been collected in cities more often than tree morphological data. Yet tree morphology can play an important role in the overall morphological characteristics of an urban area, especially residential zones and other areas with significant vegetation elements. Including trees is recommended by Grimmond and Oke (1999) in order to better represent the overall roughness of the urban surface. However, due to the difficulty to obtain sufficient tree morphological information for large areas, the importance of sparse coverage of urban trees remains relatively uncertain. Further study is needed.

3.2.5 *How Sensitive are UCP Values to Data Resolution?*

Building datasets typically are vector products (perhaps derived from raster data, e.g., LIDAR), but the computation of many UCPs is most easily performed on raster data. For such circumstances, the vector data must be converted into a raster dataset with a selected horizontal resolution. Higher resolution data should produce a more accurate UCP value, but at the cost of higher computational requirements. A study is needed to determine the methods of rasterization and how using coarser cell sizes will affect the building data and then how raster cell size affects processing time, accuracy of the UCP value, and the ultimate mesoscale meteorological or dispersion model results.

3.2.6 *Is a Consistent Data Model/Format Necessary?*

Currently, a standard data model for urban morphological input data or output format does not exist. Although some agencies have identified the need for standardization of the collection of building and morphological data (e.g., DTRA, John Pace, personal communication), consensus has not been established by those performing urban morphological analysis in support of mesoscale modeling. In terms of the input data used to compute UCPs, having a standard data model would promote data sharing and the reuse of developed software products and methodologies for urban morphological analysis. Further, reaching consensus on base data requirements would encourage data collectors to obtain information needed by all users thus making the collected data more valuable. In terms of output data, a consistent format may be unnecessary because the advancements in software products has enabled numerous forms of building morphological data to be accessed and processed with a variety of computational tools. For example, data in CAD form can be accessed in GIS and processed and vice-versa. The use of proprietary data forms (e.g., ESRI shapefile) may limit the applicability of some software products, but conversion tools are generally available. More consideration is needed to develop data standards that meet the needs of the widest section of the mesoscale meteorological and atmospheric dispersion model user community.

3.3 *Extrapolation Issues*

3.3.1 *What is the Appropriate Size Area for UCP Analysis to Obtain Meaningful Building Statistics?*

Time or budget constraints often limit the size of the area that can be morphometrically analyzed. Use of a small analysis area to compute mean parameters as a function of land use/cover types (or other homogeneous units) can lead to errors when extrapolating to larger areas of the modeling domain outside of building data coverage (Burian et al. 2003). The heterogeneity of urban terrain may cause the resulting land use-specific mean UCPs to vary depending on the size of the area included in the calculation.

To investigate this question an analysis of a large 650,000 building dataset for Houston was performed. The downtown core area of the city was delineated and a set of UCPs were calculated for the delineated area and the mean value was determined for each land use type. Then the boundary of analysis was increased incrementally and the UCPs re-calculated (see Table 2 for a summary of the size characteristics of the incremental analysis zones). The trends of the UCP values as the boundary of the analysis zone increased were quantified. The choice of the initial analysis zone to be the downtown core area may influence the results of this analysis. But, the selection of the downtown core area is appropriate for this preliminary analysis because in most cases the downtown will be of interest from a modeling perspective (and will have unique morphological characteristics). To eliminate this subjectivity, the analysis will be repeated using a series of randomly selected initial analysis zones and the results will be reported in the presentation.

Based on results using the downtown core area as the center of the analysis zones, Figure 4 shows the trend of the mean building height for several urban land use types as the size of analysis area increases. The mean building height in Residential and Industrial land uses was found to not be significantly sensitive to the size of the analysis zone, but the Commercial & Services values do change significantly as the size of the analysis area increases. In the initial analysis area extent encompassing the downtown core area, the buildings are predominantly high-rise and the land use is predominantly Commercial & Services. As the analysis extent increases the character of the Commercial & Services land use changes from high-rise to shopping malls, strip malls, and other forms with much shorter building heights (and less variable heights) than the downtown core area. Thus, the observed change to mean building height of the Commercial & Services land use is understandable. However, Residential and Industrial land uses are usually not prevalent in the downtown core area and as the analysis extent increases the building heights typically do not become significantly smaller. This observation is also noted in Figure 4.

Table 2. Characteristics of analysis zones.

Zone	Area (km^2)	Number of Buildings	Building Density (# $/km^2$)
1	3.4	489	144
2	7.5	1539	205
3	13.1	5126	391
4	27.0	15888	588
5	53.3	46058	864
6	105.9	99868	943
7	207.0	180746	873
8	371.4	278266	749
9	722.1	434285	601

Figure 5 suggests that building plan area fraction is sensitive to analysis area for all three land use types. The results indicate that the plan area fraction decreases as the analysis zone increases, reflecting a decreasing building density with distance from the downtown core area. This is most likely due to including areas of lower density housing and industrial parks normally found on the outskirts of urban areas. Analysis of the behavior of other UCPs as the analysis zone size increases will be reported in the presentation.

3.3.2 *Is a Standard Urban LULC Classification Needed?*

Land use and land cover have served for many years as surrogate data layers to define surface parameters in mesoscale meteorological and dispersion models. The approach involves defining model parameters for LULC classes based on analyzed samples of the use/cover or literature values representative of the use/cover (e.g., a roughness length for urban land use). The LULC dataset then serves as the extrapolation medium to parameterize the entire model domain using the established parameter values. This approach has been necessary because datasets describing these model parameters or datasets that they are derived from (e.g., 3D building databases) were not available or could not be efficiently analyzed for areas covering the extent of mesoscale model domains.

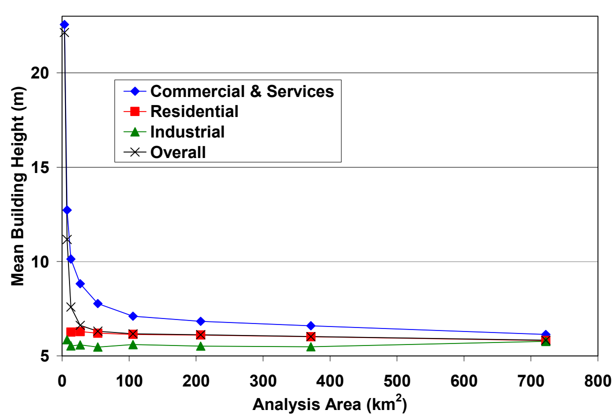

Figure 4. Mean building height as a function of analysis area size.

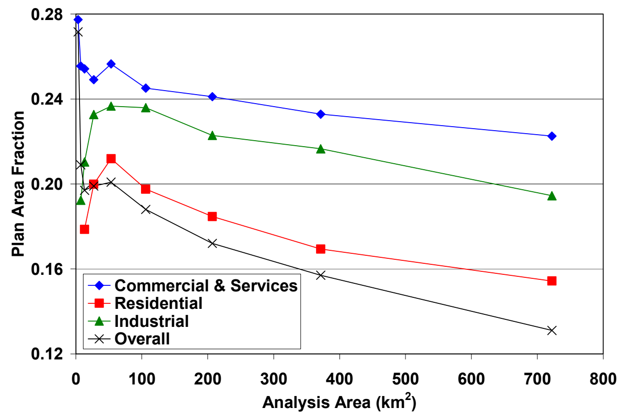

Figure 5. Building plan area fraction as a function of analysis area size.

The available LULC datasets themselves, however, may potentially have limitations. For example, the U.S. Geological Survey (USGS) LULC dataset has been commonly used and is freely available in nationally consistent form for the U.S. But, the primary source of data for the USGS LULC dataset were NASA high-altitude aerial photographs and National High-Altitude Photography (NHAP) program photographs collected mostly during the 1970s. These data are outdated for representing current conditions in areas at the urbanizing fringe of cities, especially those that have experienced massive growth during the past 30 years. Moreover, the urban land use categories in the Anderson classification scheme used by the USGS (Anderson et al. 1976) are not based on morphological characteristics. A potential source of more updated LULC data is the National Land Cover Dataset (NLCD) (Vogelmann et al. 1998). However, the NLCD is based on semi-automated classification of remote sensing data with some incorporation of ancillary data; consequently, the definition of urban land use types is limited to a small number because of the heterogeneous surface properties (Vogelmann et al. 1998). The four urban land use types represented in the NLCD (Low Intensity Residential, High Intensity Residential, Commercial/ Industrial/Transportation, and Urban Vegetation) in most cases will include a wide range of urban surface fractions (e.g., paved areas, rooftops, landscaped areas, bare soil, etc.) with different reflective properties within each class, which will confuse classification algorithms and potentially cause misclassification. Moreover, the aggregation of Commercial, Industrial, and Transportation land uses into a single land use category will mask the heterogeneity of the urban surface because the building morphological characteristics of these three individual land uses are not similar (e.g., see building statistics in Burian et al. 2002 or Grimmond and Oke 1999).

As briefly noted, the USGS and NLCD datasets have potential drawbacks for UCP definition and extrapolation to mesoscale meteorological and dispersion model domains. From a meteorological or dispersion modeling perspective, numerous researchers have devised more detailed LULC classification schemes to overcome the potential deficiencies of the aforementioned national datasets and provide more accuracy for their analyses. Ellefsen (1990/1991) for example, derived an urban LULC dataset for urban climate and meteorological modeling applications. His classification system included 17 urban terrain zones (UTZ) that were homogeneous units from a building morphological perspective. Each zone was defined to have a distinctive mix of function, development age, street pattern, lot configuration, and type of construction and density and height of buildings. A primary difference between UTZ categories is whether the buildings are attached or detached. UTZs with detached structures were further subdivided by the closeness of the structures.

Using a morphological-based classification approach similar to Ellefsen, Theurer (1993, 1999) identified a set of nine typical building arrangements for cities in Germany. Another classification approach incorporating morphological characteristics was introduced by Grimmond and Souch (1994). They used a 36-category urban land use classification scheme to represent urban terrain in a GIS-based surface characterization methodology. The first level of land use categories included traditional single and apartment residential, commercial/industrial, institutional, transportation, vacant, vegetated, impervious, and water. Sub-categorization was based on building height and density and cover fractions.

Burian et al. (2002) also used a two-tier classification approach, but their scheme was based on the Andersen Level 2 urban land use categories as the first tier. Morphology was then incorporated to subdivide traditional USGS urban land use classes into morphologically-based sub-categories. For example, the Residential land use category was subdivided into low-density or high-density based on a chosen threshold building density level. Long et al. (2003) used two automated approaches to classify land use, one based on morphological parameters and the other based on analysis of multi-spectral imagery (20-m resolution SPOT). Originally, a 7-class scheme based on morphological parameters was tested, but was expanded to 10-classes after high parameter variability was noted in each class. The morphological classification identified urban land use types well, but in general did not define land cover well. The imagery approach did give accurate information about the cover type, but was unable to accurately identify urban elements and their combinations to form the urban terrain classes.

This sampling of urban LULC classification approaches indicates that morphologically-based approaches are available and can be used in urban morphological analysis and may have application in mesoscale modeling. The question then becomes the need for standardization of classification and the creation of repeatable, objective, automated approaches to derive a spatially-consistent LULC dataset. Standardization would benefit morphological analysis and may benefit the mesoscale meteorological modeling community, but perhaps would decrease the value of previously collected data and derived relationships unless the new land use/cover categories could be chosen to be consistent with categories from other classification approaches (or categories from other classification approaches can be generalized into the categories of the new classification approach). The popularity of the USGS LULC dataset with the mesoscale modeling community might also influence the selection of classification categories. One possible solution is to use morphological characteristics (e.g., building density, mean building height, vegetative cover, impervious surface fraction) to subdivide a generalized first level of land use/cover, but this would require a nationally consistent morphological database.

4. SUMMARY

This paper reviewed the historical advancements to urban morphological analysis approaches and discussed the details of several practice issues. The issues discussed are not considered exhaustive, but do provide a sampling of current practice inconsistencies and uncertainties that could potentially reduce the accuracy of urban canopy parameters and mesoscale model results. Further analysis of these issues and others will be included in the presentation.

Acknowledgements

This work was partially supported by the Department of Homeland Security Chemical and Biological Countermeasures Program. The United States Environmental Protection Agency through its Office of Research and Development also partially funded and collaborated in the work described here. It has been subjected to Agency review and approved for publication. The authors wish to thank Tanya Otte for insightful comments that helped to improve the manuscript.

5. REFERENCES

Anderson, J. R., E. E. Hardy, J. T. Roach, R. E. Witner, 1976: *A land use and land cover classification system for use with remote sensor data*. USGS Prof. Paper 964.

Bornstein, R. D., 1987: Mean diurnal circulation and thermodynamic evolution of urban boundary layers. *Modeling the Urban Boundary Layer*, AMS, Boston, MA.

Bornstein, R. D. and K. J. Craig, 2002: Survey of the history of the urbanization of numerical mesoscale models. Preprints, *Fourth Symposium on the Urban Environment*, AMS, Norfolk, VA.

Brown, M. J., 2000: Urban parameterizations for mesoscale meteorological models. *Mesoscale Atmospheric Dispersion*. Ed., Z. Boybeyi., pp. 193-255.

Brown, M. J. and M. D. Williams, 1998: An urban canopy parameterization for mesoscale meteorological models, Preprints, *Second Urban Environment Conference*, AMS, Albuquerque, NM.

Burian, S. J., M. J. Brown, and S. P. Linger, 2002: *Morphological analysis using 3D building databases: Los Angeles*, California. LA-UR-02-0781. Los Alamos National Laboratory, 66 pp.

Burian, S. J., W. Han, S. P. Velegubantla, S. R. K. Maddula, 2003: *Development of gridded fields of urban canopy parameters for Models-3/CMAQ/MM5*. U.S. EPA Internal Report, RTP, North Carolina.

Chin, H. N., M. J. Leach, and M. J. Brown, 2000: A sensitivity study of the urban effect on a regional-scale model: An idealized case. Preprints, *Third Symposium on the Urban Environment*, Davis, CA.

Ching, J., T. L. Otte, S. Dupont, S. Burian, and A. Lacser, 2002: Urban morphology for Houston to drive Models-3/CMAQ at neighborhood scales. Preprints, *Fourth Symposium on the Urban Environment*, AMS, Norfolk, Virginia.

Cionco, R. M. and R. Ellefsen, 1998: High resolution urban morphology data for urban wind flow modeling. *Atm. Env.*, **32**(1), 7-17.

Cionco, R. M. and S. Luces, 2002: Coupled mesoscale-microscale model to compute neighborhood scale wind fields. Preprints, *Fourth Symposium on the Urban Environment*, AMS, Norfolk, VA.

Dupont, S., J. Ching, and S. Burian, 2004: Introduction of urban canopy parameterization into MM5 to simulate urban meteorology at neighborhood scale. Preprints, *Symposium on Planning, Nowcasting, and Forecasting in the Urban Zone*, 84th AMS Annual Meeting, Seattle, WA.

Ellefsen, R., 1990/1991: Mapping and measuring buildings in the canopy boundary layer in ten U.S. cities. *Energy and Buildings*, **15-16**, 1025-1049.

Ellefsen, R. A. and R. M. Cionco, 2002: A method for inventorying urban morphology. Preprints, *Fourth Symposium on the Urban Environment*, AMS, Norfolk, VA.

Grimmond, C. S. B. and T. Oke, 1999: Aerodynamic properties of urban areas derived from analysis of surface form. *J. Appl. Met.*, **38**, 1262-1292.

Grimmond, C. S. B. and C. Souch, 1994: Surface description for urban climate studies: A GIS based methodology. *Geocarto Int.*, **9**, 47-59.

Hanna, S. R. and R. E. Britter, 2002: *Wind flow and vapor cloud dispersion at industrial and urban sites*. American Institute of Chemical Engineers, New York, NY.

Holt, T. R., S. Chin, M. Leach, and G. Sugiyama, 2002: Sensitivity of mesoscale real-data simulations to an urban canopy parameterization. Preprints, *Fourth Symposium on the Urban Environment*, AMS, Norfolk, VA.

Lacser, A. and T. L. Otte, 2002: Implementation of an urban canopy parameterization in MM5. Preprints, *Fourth Symposium on the Urban Environment*, AMS, Norfolk, VA.

Long, N., 2003: *Analyses morphologiques et aérodynamiques du tissue urbain – Application á la climatologie urbaine de Marseille pendant la campagne ESCOMPTE*. Thesis, École Centrale de Nantes, Nantes, France.

Long, N., S. Kermadi, C. Kergomard, P. G. Mestayer, and A. Trébouet, 2003: Urban cover modes and thermodynamic parameters from urban database and satellite data: A comparison for Marseille during ESCOMPTE, Proceedings, *Fifth Int. Conf. Urban Climate*, Łódź, Poland, September 2003.

Long, N., Mestayer, P. G., and C. Kergomard, 2002: Development of a software to describe the city morphology and to compute aerodynamic parameters from an urban database. Preprints, *Fourth Symposium on the Urban Environment*, AMS, Norfolk, Virginia.

Otte, T. L. and A. Lacser, 2001: Implementation of an urban canopy parameterization in MM5 for meso-gamma-scale air quality modeling application. Preprints, *Ninth Conference on Mesoscale Processes*, AMS, Ft. Lauderdale, FL, 78-81.

Pace, J. C. and R. J. Babarsky, 2002: DTRA urban terrain database project. Preprints, *Fourth Symposium on the Urban Environment*, AMS, Norfolk, VA.

Petersen, R. L. and B. C. Cochran, 1998: *Development and testing of methods for estimating surface roughness length at refineries and other industrial facilities*. CCP Project 95-1214, CPP, Ft. Collins, CO.

Petersen, R. L. and D. K. Parce, 1994: *Development and testing of methods for estimating surface roughness length at refineries*. CCP Project 92-0890, CPP, Ft. Collins, CO.

Priestnall, G., J. Jaafar and A. Duncan, 2000: Extracting urban features from lidar digital surface models. *Computers, Environment and Urban Systems*, **24**(2), 65-78.

Ratti, C. and P. Richens, 1999: Urban texture analysis with image processing techniques, *Proceedings of the CAADFutures99 Conference*, Atlanta, GA.

Ratti, C., S. Di Sabatino, R. Britter, M. Brown, F. Caton, and S. Burian, 2002: Analysis of 3-D urban databases with respect to pollution dispersion for a number of European and American cities. *Water, Air, and Soil Pollution: Focus*, **2**, 459-469.

Richens, P., 1997: Image processing for urban scale environmental modeling. *Proceedings of the Int. Conf. on Building Simulation '97*, Prague, 163-171.

Rozoff, C. M., W. R. Cotton, and J. O. Adegoke, 2003: Simulation of St. Louis, Missouri, land use impacts on thunderstorms. *J. Appl. Met.*, **42**, 716-738.

Theurer, W., 1993: *Ausbreitung bodennaher Emissionen in komplexen Bebauungen*. Dissertation, IHW, Universität Karlsruhe.

Theurer, W., 1999: Typical building arrangements for urban air pollution modeling. *Atm. Env.*, **33**, 4057-4066.

Vogelmann, J. E., T. L. Sohl, P. V. Campbell, and D. M. Shaw, 1998: Regional land cover characterization using Landsat Thematic Mapper data and ancillary data sources. *Environmental Monitoring and Assessment*, **51**, 415-428.

Zhou, G., C. Song, J. Simmers, P. Cheng, 2004: Urban 3d GIS from lidar and digital aerial images. *Computers & Geosciences*, **30**(4), 345-353.