6.2

DATABASE FEATURES OF THE NATIONAL URBAN DATABASE AND ACCESS PORTAL TOOLS (NUDAPT)

Steven J. Burian[1*], Michael J. Brown[2], David Sailor[3], Ronald Cionco[4], Richard Ellefsen[5], Mark Estes[6], and Torrin Hultgren[7]

[1]Department of Civil & Environmental Engineering, University of Utah, Salt Lake City, Utah
[2]Systems Engineering and Integration Group, LANL, Los Alamos, NM
[3]Department of Mechanical Engineering, Portland State University, Portland, Oregon
[4]Retired Micrometeorologist, Las Cruces, NM
[5]Professor Emeritus, San Jose State University, San Jose, CA
[6]Division of Air Quality, Texas Commission on Environmental Quality
[7]Computer Science Corporation, EPA Contractor, RTP, NC

1. INTRODUCTION

This paper describes the features of the databases included in the National Urban Database and Access Portal Tools (NUDAPT). The Houston prototype of NUDAPT contains three-dimensional building, airborne LiDAR digital elevation model (DEM), derived gridded urban canopy parameters (UCPs), derived micrometerological physical characteristics, land use/cover, population, and anthropogenic heating databases. These form the core databases recommended for future NUDAPT cities with additional public and proprietary datasets to be added as NUDAPT expands. The database characteristics are described and specific problems encountered while integrating these disparate databases into NUDAPT are mentioned.

2. BACKGROUND

Modeling and simulation of meteorology, dispersion, and air quality is limited in the ability and capability to perform accurate air quality and consequence assessments in our urban areas. Improved access to high-resolution urban morphological features and improved ability to incorporate the advanced databases into meteorological, air quality, and human exposure modeling systems will be needed for future urban applications (OFCM 2005). In response, NUDAPT has been developed to address urban database needs (Ching et al. 2007). The initial NUDAPT design concepts and prototype application have been sponsored by the United States Environmental Protection Agency (USEPA) and involves collaborations and contributions from many groups from federal and state agencies, and from private and academic institutions. It is designed to produce gridded fields of UCPs for various new advances in model physics to improve urban meteorological simulations given the availability of new high-resolution data of urban buildings and land usage. An important core-design feature is the utilization of web-based technology to enable NUDAPT to be a “Community” based system. Houston, Texas is serving as NUDAPT’s initial prototype with eventual extensibility to all other cities.

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

3. NUDAPT DATABASES

The NUDAPT Houston prototype contains numerous urban databases including three-dimensional building, airborne LiDAR full-feature digital elevation model (DEM), micrometeorological database, gridded UCPs, population, anthropogenic heating, and land use/land cover. The basic features of these NUDAPT datasets and what data is included in the portal are described and pitfalls associated with implementing them and other datasets are mentioned. In addition, numerous datasets not described below have also been ingested into the portal including freeways, major roads, major thoroughfares, waterbodies, channels, creeks, streams, parks, elevation contour lines, and other publicly-available datasets.

3.1 *Building Database*

Building databases can be obtained from several sources. First, commercial entities can provide three-dimensional building data consisting of footprints with a height attribute in geographic information system (GIS) format. Second, government entities at the local and state level may produce building footprint databases for a variety of purposes (e.g. planning, tax assessment). Third, federal defense agencies (e.g., National Geospatial-Intelligence Agency (NGA)) produce three-dimensional building databases for smaller sections of cities. Researchers also have produced three-dimensional building datasets.

Of the options available to the NUDAPT developers, the use of a building database for Houston developed by researchers at the University of Utah was selected because of the ability to freely distribute the data without concern for licensing. The Houston building database contains more than 650,000 building footprints in shapefile format with height attributes (Fig. 1). This database covers more than 1600 km^2 of the Houston metropolitan area.

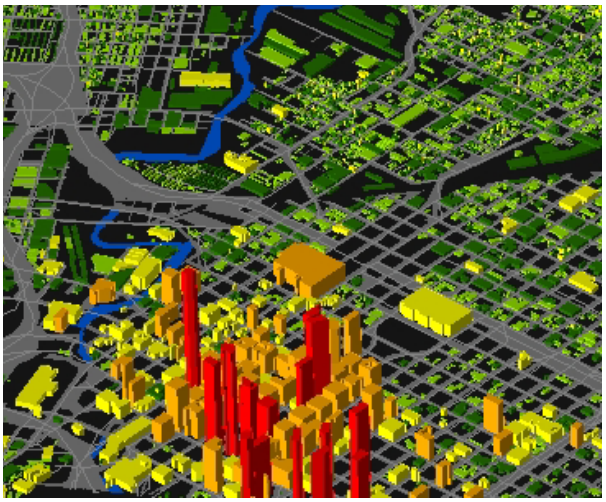

Fig. 1. Illustration of Houston building database included in NUDAPT.

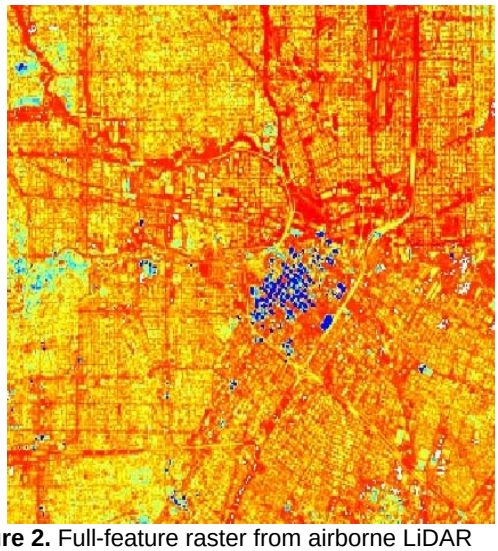

Figure 2. Full-feature raster from airborne LiDAR dataset.

3.2 *LiDAR DEM*

Airborne LiDAR was collected November of 2001 by Terrapoint, LLC as part of an USEPA-sponsored effort to urbanize MM5 and to use a high-resolution gridded set of UCPs. The LiDAR was packaged into several data products including ground, nonground, digital terrain model (bare-earth), and digital elevation model (full feature). Selected LiDAR data products have been incorporated into the portal (Figs. 2, 3 and 4):

- 15-m DTM available for download from Harris County
- 5-m DTM from Terrapoint - missing tile
- 5-m DEM (building heights) from Terrapoint - also missing tile
- 1-m DTM and DEM from Terrapoint - missing tile, not mosaicked or included in portal yet
- 1-m height raster for 220-km^2 area – available but not in portal yet

The height raster was computed by subtracting the ground from the nonground and a patch was created from additional datasets to fill in the missing tile. The difference between ground and nonground results in a full-feature DEM of heights above ground for all roughness features (trees, buildings, etc.).

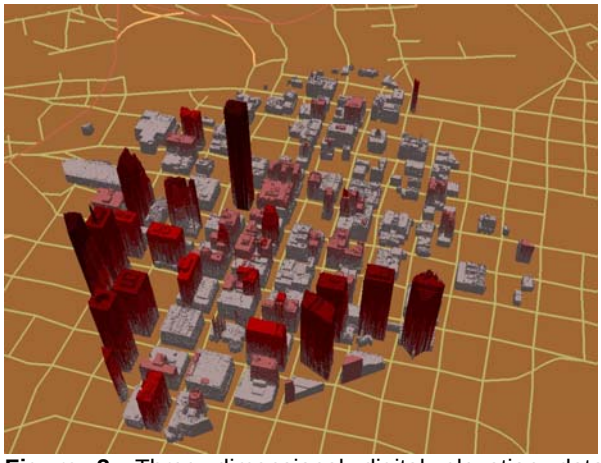

Figure 3. Three dimensional digital elevation data derived from airborne lidar platform for 1X1 km section of downtown Houston.

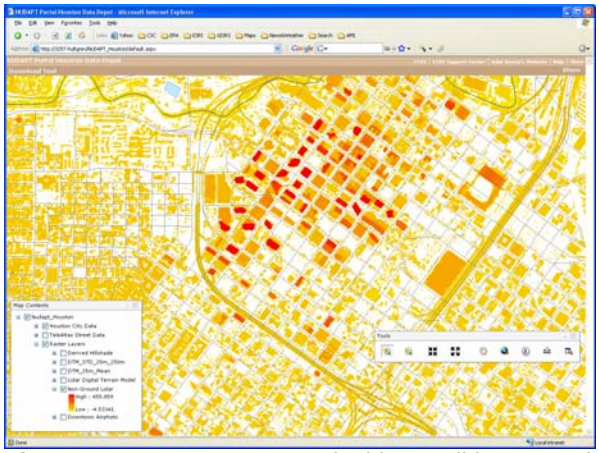

Figure 4. Screen capture of airborne lidar DEM in NUDAPT.

3.3 *Urban Canopy Parameter Database*

The UCP database provides 250-m and 1-km resolution coverage of a selected set of UCPs. The selected set were converted to raster and incorporated into the portal for download:

- Mean Building Height (Fig. 5)
- Standard Deviation of Building Height
- Plan-Area-Weighted Mean Building Height
- Building Height to Width Ratio
- Canopy Top Area Density
- 0 Degrees Mean Street Orientation
- 45 Degrees Mean Street Orientation
- 90 Degrees Mean Street Orientation
- 135 Degrees Mean Street Orientation
- Sky View Factor

The 250-m resolution database covers an area of 1600 km^2 and the 1-km resolution covers Harris County. Beyond these UCPs others are available for eventual upload including building height histograms, plan area fraction and density of buildings, trees, and full canopy, frontal area index and density of buildings, trees, and full canopy, building-to-plan area ratio, roughness length, displacement height, and more.

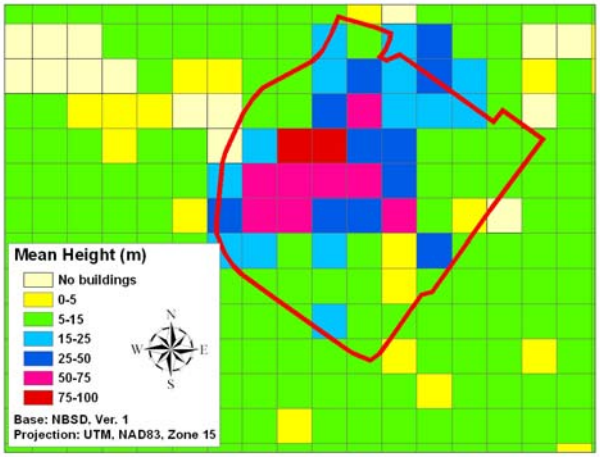

Figure 5. UCP database of mean building height for Houston.

3.4 *Micrometeorological Database*

NUDAPT contains data not only relevant for those modeling in the mesoscale, but also those working at the microscale. Micrometeorological databases include building morphology and non-building morphology (trees, shrubs, crops, grass, bare soil, water, and various types of impervious surfaces). The data are derived from the building database and aerial images to represent all morphology types with high resolution (50-m and 100-m grid cell sizes are commonly used) and high-fidelity data. A sample dataset covering areas in the downtown tall building district of Houston is included in NUDAPT. More details of the Houston micrometeorological database are contained in Cionco and Ellefsen's paper in the proceedings.

3.5 *Population*

The population database created at the Los Alamos National Laboratory (LANL) includes day-night-worker categories at 250-m resolution for the entire United States, although only the Houston area is included in NUDAPT at this time.

3.6 *Anthropogenic Heating*

Anthropogenic heating consists of heat released from three key sectors, buildings, manufacturing/industry, and transportation. Based on annual energy consumption data for the U.S. it is estimated that total anthropogenic heating is divided somewhat evenly among these three sectors. The anthropogenic heating data supplied to NUDAPT linked detailed energy consumption data from multiple sources with fine-scale GIS data (land use, road surface area, building type, and building size).

Building sector data were created using detailed building energy simulation of prototypical buildings. The necessary data to define and characterize the 22 commercial and 8 residential building prototypes were obtained from the Commercial Building and Residential Energy Consumption Surveys (CBECS and RECS), respectively.

For the vehicle sector, Department of Transportation data for per capita vehicle miles traveled were mapped onto roadways based on length and type of roadways (freeways vs. major arterials) within each grid cell. Diurnal traffic pattern shapes were applied based on national traffic count data.

Industrial energy consumption was assumed to be constant in time and was divided equally among all industrial land-use parcels.

In all cases, total anthropogenic heating at the parcel scale was aggregated up to the 500 m scale for presentation in NUDAPT. The anthropogenic heating data currently available in the portal includes:

- August Daily Average
- August Hourly (24 individual rasters)
- January Daily Average
- January Hourly (not yet included in portal)

More may be added in the future as NUDAPT evolves.

3.7 *Land Use/Land Cover*

Land use/cover databases from the National Land Cover Dataset (NLCD), the U.S. Geological Survey (USGS), and a specialized database from the Texas Commission on Environmental Quality are included in NUDAPT:

- NLCD 2001 Land Cover
- NLCD 2001 Canopy

- NLCD 2001 Impervious
- NLCD 1992 Land Cover
- TCEQ 30-m East Texas Land Cover

The LULC datasets have proven to be very straightforward to work with in the portal because their size permits wide areas being manipulated with few processing challenges.

4. IMPLEMENTATION PITFALLS

One of the first pitfalls encountered was the different spatial references of the databases. With different coordinate systems or coordinate systems specifications the datasets in their raw form presented a spatial mapping problem. Changing spatial references created was problematic because of the wide spatial scales covered by the databases, which in some cases required geographic coordinates and in others a local coordinate system (e.g., State Plane).

Similar to the spatial reference discrepancy, the spatial resolution of the raw data presented a pitfall for eventual model use. Retaining the LiDAR data in 1-m resolution for a large area (thousands of square kilometers) presented a data management and transfer problem. Further, the need for spatial resampling presented a challenge. Resampling to modify the spatial resolution is being addressed by a tool called the Spatial Allocator. However, the effect of smoothing of data upon resampling to coarser resolutions and methods to prevent impacts has not been adequately studied.

The value of the very high resolution data is also unknown for mesoscale model applications, yet essential for microscale and computational fluid dynamics applications. Thus, a dilemma evolved related to which model scale to address. NUDAPT meetings addressed this issue, but a standard approach has not been determined. Currently, the database defaults to the needs of the mesoscale modeling community with most of the databases, although some do address the finer scale (see Section 3.4).

One common problem with the data is missing data or data errors. With any massive database project proper quality assurance and quality control of the database is an extreme challenge. NUDAPT is no different. One problem in particular is archived data that was collected years ago and proven to be difficult to obtain. As NUDAPT moves towards operation this challenge will grow as users identify more data problems and the potential to permit users to upload data to the portal to supplement core datasets is explored.

The technical issues with any web-based application require a full-time staff member to manage. The NUDAPT interface was originally developed in a simple form, but has been advanced as the web application was developed. But, new bugs in the software are found each week and must be dealt with on a case-by-case basis.

Finally, metadata is a major issue currently being resolved. The current lack of metadata is a problem even for internal testers of NUDAPT because often little information is available about a particular dataset. Transferring critical spatial reference information such as origin and projection must be incorporated into the portal and the ability to define specific attributes must also be addressed.

5. SUMMARY

This paper described the core databases currently included in NUDAPT portal. Several problems associated with data acquisition, management, and dissemination were highlighted and the current status of NUDAPT to deal with these problems was described. The NUDAPT portal will continue to be advanced and the databases being incorporated will expand, especially as the problems identified in this paper become more routine with experience.

6. REFERENCES

Office of the Federal Coordinator for Meteorology (2005): Proceedings from the Forum on Urban Meteorology. www.ofcm.gov/homepage/text/pub.htm.

Ching, J., Chen, F., Kiley, C., and Burian, S. (2007). "Advancing urban dispersion and air quality models using a community-based high resolution building and urban data base system." *11th International Conference on Harmonization within Atmospheric Dispersion for Regulatory Purposes*, 2-5 July 2007, Cambridge, UK.