9.6 A FEDERATED PARTNERSHIP FOR URBAN METEOROLOGICAL AND AIR QUALITY MODELING

*David J. Williams[1], Jason Ching[2]

[1]Environmental Sciences Division, EPA NERL, RTP, NC
[2]Atmospheric Sciences Modeling Division, NOAA ARL (on assignment to NERL, US EPA, RTP, NC)

1. INTRODUCTION

Recently, applications of urban meteorological and air quality models have been performed at resolutions on the order of 1 km grid sizes. This necessitated development and incorporation of high resolution landcover data and additional boundary layer parameters that serve to describe the influence of urban morphological structures on the flow, transport and energetic details in urban areas. The task of determining the structure and composition of urban areas has been termed urban morphological analysis. The data requirements for urban morphological parameters for advanced mesoscale grid models has lead to the generation of very large datasets for urban canopy parameters (UCPs) used to drive models such as CMAQ/MM5, e.g., Dupont et al., 2004, and Otte et al., 2004. Recently, means to obtain high resolution information (of order 1 m) for building and other urban structures are now operationally feasible using a variety of methods including photogrammetry and LiDAR on airborne platforms (Gamba et al., 2003; Thomas et al., 2003). Databases of varying sizes, but typically very large, have been collected for most of the major cities in the continental US and Europe and are being continually upgraded (airbornlasermapping.com, 2004; Renslow, 2002; Vexcel, 2004). Derivative information from such databases such as gridded parameterizations for mesoscale models can quickly increase the size of databases. In some current and future applications, standardization of databases and model inputs may be a recommended operational procedure for nationwide use.

Given these circumstances and considerations, data storage and dissemination issues become important and can be problematic for some applications requiring these large datasets. These high resolution datasets are often expensive to procure, process, efficiently archive, and maintain, especially with the evolving size and complexity of urban centers. For example, data sharing amongst collaborators can become problematic when transferring 10-100 GB files. What is desired is an efficient means to obtain relevant high resolution landcover data, LiDAR based digital elevation models (DEM), and satellite imagery to generate the required UCPs, or previously created UCPs, or both. This database would take care of the data procurement, storage, archiving, dissemination, and curation issues (quality control), allowing the researchers to concentrate more time on modeling and less time on data management. The number of urban areas that have current high resolution (5-m or less) DEM and landcover data is increasing steadily (Terner et al., 2004). It is conceivable that in the near future the database preparations and simulation modeling accomplished as a prototype for Houston, TX using high resolution datasets (Burian, et al., 2004) can be done on a national basis. Any data sharing mechanism should scale accordingly, allowing for nationwide urban meteorological and air quality modeling network.

Given the anticipated need for such detailed urban morphological databases, we take this opportunity to engage in a public discussion the examination and exploration of the feasiblity for creating an Urban Meteorological and Air Quality Information Partnership for the establishment of national database of urban morphological parameters. This database is only in the planning phase, but the opportunity exists now for interested participants to guide the formation of the Partnership and database. The creation of this database would allow for the dissemination of data in various stages of processing, from minimally processed LiDAR DEM and orthorectified airborne and satellite imagery to the highly refined parametric sets into the public domain. In our opinion, we see a need and an opportunity to establish such a national database to serve a wide set of user communities that potentially would be interested in this, including but not be limited to urban air quality modelers, developers of modeling tools for Homeland Security, urban design and energy planners, flood control agencies, and metropolitan emergency coordination agencies. We now examine specific features that we envision for such a database.

2. DATA AND DATASET REQUIREMENTS

At least two communities of users are envisioned as the primary clients for this database. Each community has its own data requirements and the database must satisfy both if it is to be useful. The first

*corresponding author and address: David J. Williams, EPA, Research Triangle Park, NC, 27711; email: Williams.davidj@epa.gov

community is the computational fluid dynamics (CFD) modelers who are interested mainly in minimally processed source data for dispersion modeling at the building level. These users require a standard line of products at native (non-aggregated) spatial resolution. The CFD modelers are expected to extract the information they require from the data. The second community is the mesoscale flow pattern modelers who define urban canopy parameters from morphologic information but do not necessarily process the source data to determine the needed parameters. Other communities of users may be involved in policy and planning or management and require further refined data for visualization rather than modeling. The primary goal of the database will be to satisfy each user's data requirements through a common interface and data transfer mechanism.

High resolution topographic and landcover data useful for determining urban morphological parameters vary in quality, procurement cost, license policy, and file size (which in turn affects data processing, storage, retrieval, and dissemination). The quality of a DEM or landcover dataset is related to geopositional accuracy, image fidelity, and the technical details of the data acquisition such as view geometries. The generation of urban morphological parameters varies with the requirements of the end user. This diverse set of source data, analytical methods, and end data products can create barriers to scientific collaboration and information dissemination. A primary requirement of a national database is to record the attributes of all data types in the system so data products are not used inappropriately. To achieve this, metadata documentation will accompany each dataset distributed by the database. Any submissions to the database will be expected to adhere to the metadata standards and accurately describe the processing that occurred to the submitted dataset.

Currently, several agencies are involved in producing and developing highly detailed building and urban structure information for most of the US (Vernon, 2004). The method of data collection and the handling of such data are not at this time standardized. Efforts to evaluate the degree of relative differences as to their resolution, accuracy, and precision would be a basic requirements.

3. DATA FEDERATION PARTNERSHIPS

We propose the creation of a national database of urban morphological parameters based on the model of federated partnerships as implemented by the Earth Science Information Partnership or ESIP. Information science defines a data federation as a system having the ability to access multiple databases using one query (Kerschberg, 2001). The policy analogy is described as a confederation of research organizations affiliated through their use and dissemination of scientific data and information (ESIP, 2003). This policy sets the stage for a framework where interested parties can collaborate and share data in an organized way. Each party in the federation has an equal voice in discussions. Determination of standards and interoperability issues can be worked out in an efficient and equitable fashion. In the end, the implementation of this national database of urban morphological parameters may be a centralized data repository or a distributed system where each organization in the federation hosts and maintains its own data and links to a common data portal for search and retrieval. The federation is responsible for deciding data standards, formats, and implementation. An example of one possible implementation, which will be described next for illustrative purposes, is based on the Seasonal to Interannual ESIP (Kafatos et al., 1999).

4. URBANAIR ESIP: A POSSIBLE IMPLEMENTATION

The premise of this national database of morphological parameters is that there are several levels of data requirements and informational needs. Some members of the federation will be interested in performing urban morphological analysis and generating parameters such as UCPs for various applications. Others will be interested in the parameters only, while still others will primarily be interested in the model simulation results for analysis or informational purposes. The framework of the UrbanAir ESIP will allow for data needs assessments or cataloging, standardization of UCP calculation methods if that is deemed useful, and format standards. The database will include metadata that records how parameters were generated and what datasets were used for the analysis.

The architecture of a centralized data repository with distributed nodes is an idea that works well for systems that gather varied types and amounts of data but disseminates them through a common portal or search and retrieval mechanism. The central repository idea has the benefits for database administration, leveraging in data procurement, and data curation. Data used in urban morphological analysis can be costly and have restrictive licensing that prohibits wide dissemination, as is the case with high resolution satellite imagery. These use restrictions can be overcome by a variety of means such as payment of an extra fee for unlimited distribution, or by only disseminating data products derived from the source imagery. A central data repository can institute dissemination controls that limit access to certain users in order according to the data license agreement. There are other holders of landcover and topographic data that prefer to interact with a single organization rather than field requests from multiple entities. This is the case for data held by the Homeland Security Infrastructure Program (HSIP) (R. Tugwell 2004, personnel communication). Distributed nodes for the federation, such as NASA's Distributed Active Archive Center (DAACs), NOMADS (NOAA Operational Model Archive and Distribution System) and collaborating university databases, will handle requests for data using an agreed upon interoperability standard like XML. A common data portal will be the user interface for all queries. In this case the proposed portal is the Global Change Master Directory (GCMD), which can be accessed through NASA, the ESIP Federation, and Mercury from Oak Ridge National Labs. These portals are metadata search engines that synchronize with the federation's metadata servers.

The proposed implementation would be comprised of multiple layers for servicing several types of queries (figure 1). As was previously described, the premise is that a significant proportion of the users will be interested in the processed urban morphological parameters, and not in the source data used to generate them. These users are easily serviced via the web

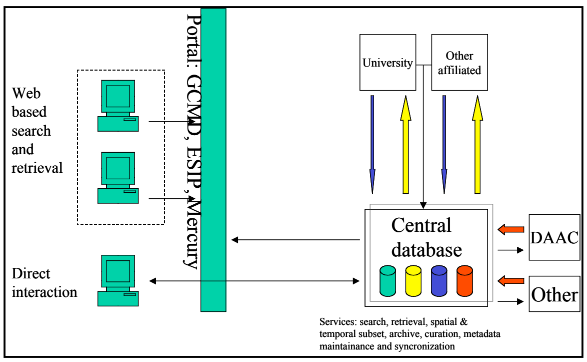

Figure 1. Schematic of proposed database. Green objects are metadata searches, portals and servers. Yellow objects are high bandwidth data pulls and storage, blue objects are uploads of products such as UCPs and storage. Red objects are data pulls from NASA DAAC and storage.

based portal. Other users may be more sophisticated and will want to directly interact with the database via database to database XML calls and bypass the portal. The users interested in performing urban morphological analysis require source data such as LiDAR based DEMs, airborne or satellite imagery, and in some cases large GIS files such as municipal building planimetric data and cadastral information. The data requirements for these users entails using a provisioning approach to the location and dissemination of data (Hardy and Groom, 2004). The provisioning approach is a data staging and user retrieval method that uses more bandwidth and disk space than the other users. These users will also be allowed to upload data to the central database such as the generated UCPs, once completed.

Technologies currently exist that allow for a certain level of data processing on the server side. Users will be able to interactively subset larger datasets using their own geospatial coverages or shapefiles (ESRI, 2004), manually enter coordinates or select an area on a map. Users may then retrieve data the requested in a variety of file formats such as binary, ASCII, geo-TIFF, geo-JPEG and others along with the necessary metadata. Minimal data processing will be available that allow users to perform simple analysis operations such as aerodynamic roughness calculations (mean building height and standard deviation) at multiple scales and surface materials (vegetation indices).

5. ISSUES

Implementation of a national database of urban morphologic parameters and accompanying landcover and topographic data requires institutional commitment and financial support to become an operation reality. Potential affiliates in the UrbanAir federation must commit to working within the federation's framework and participate in the necessary discussions of standards and interoperability. A long-term funding mechanism must be developed to procure data, operate the database systems, and maintain and archive the data. Coordination with agencies with related mandates may assist in funding and ensure the long-term use and viability of the system. The intergovernmental Group on Earth Observations (GEO) may have keen interest in the federation as this ties directly into the GEO's concept of a system of systems to provide timely, quality, long-term global information for sound decision making (GEO, 2003). Issues of data security, privileged versus open access, and legal matters dealing with licensed imagery need to be addressed.

The Partnership may want to appoint a Board of Scientific Advisors to facilitate strategic planning and coordinate efforts with GEO and activities at the National Academy of Sciences. This board may also serve as a peer-review of the Partnership over its lifetime as personnel, political administrations, and agencies mandates change.

6. CONCLUSION

The urban meteorological and air quality modeling community will greatly benefit from the creation of a national database of urban morphological parameters. The organizational framework that accompanies the implementation of this database, centered on the federated model of database interoperability, will allow for a robust and useful system that can benefit both the modeling community as well as homeland security, urban planning and policy, emergency management communities. The UrbanAir federation may also increase the use of this type of data in unforeseen ways, which often is the mark of a truly useful resource. Commitment by interested partners and a long-term funding strategy are the keys to the creation and success of the system. This paper has proposed the goals and objectives of such as system along with an example of a possible implementation. Further discussion and debate is encouraged by the authors.

7. REFERENCES

Airbornelasermapping.com. 2004. 2004 LiDAR industry directory. http://www.airbornelasermapping.com/ALMID.html

Burian, S. J., 2004: Urban canopy parameter assessment for Houston, Texas. Research report prepared for Daewon Byun, Institute for Multidimensional Air Quality Studies (IMAQS), University of Houston, Houston, Texas.

Dupont, S., Otte. T., and J.K.S. Ching, 2004. Simulation of meteorological fields within and above urban and rural canopies with a mesoscale model (MM5) In press, 113:1 (October 2004) Boundary layer Meteorology.

ESIP, 2003. Constitution of the Federation of Earth Science Information Partners. http://www.esipfed.org/business/governance/constitution_01_14_03.html

ESRI, 2004. Environmental Systems Research Institute, Redlands, CA, USA

Gamba, P., O. Hellwich, and P. Lombardo. 2003. Theme issue on algorithms and techniques for multi-source data fusion in urban areas. ISPRS Journal of Photogrammetry and Remote Sensing, 58(1-2).

Hardy, M and T. Groom. 2004. On demand spatial data provisioning. Sanz Corp. whitepaper.

Kafatos. M., M., L. Chiu, X. Deng, B. Doty, Tarek El-Ghazawi, A. Jearanai, O. Kelley, S. Kempler, J. Kinter, J. Kwiatkowski, Z. Liu, C. Lynnes, K. Matsuura, J. McManus, Prachya, P. Schopf, G. Serafino, C.Wang, X.S. Wang, H. Weir, C. Willmott, K-S Yang, R. Yang. 1999. System Description for the Seasonal to Interannual Earth Science Information Partner (SIESIP), Supporting the Science, Data, and Information Needs of the S-I Community. Fall Meeting of the American Geophysical Union (AGU), San Francisco.

Kerschberg, L. 2001. Knowledge management in heterogeneous data warehouse environments. in International Conference on Data Warehousing and Knowledge Discovery, Lecture Notes in Computer Science 2114, (Munich, Germany, 2001), Springer-Verlag, 1-10.

Otte, T.A., Lacser, S. Dupont, and J.K.S. Ching, 2004. Implementation of an urban canopy parameterization in a mesoscale meteorological model. Accepted by JCAM Mar-04.

Renslow, M. 2002. Mapping the LiDAR market. Geospatial Solutions

Terner M., R. Sutton, B. Hebert, J. Bailey, H. Gilbert and C. Jacqz. 2004. Protecting critical infrastructure. GeoIntelligence

Thomas, N., C. Hendrix and R. G. Congalton. 2003. A comparison of urban mapping methods using high-resolution digital imagery. Photogrammetric Engineering & Remote Sensing, 69(9).

Vexcel, 2004. Vexcel Corporation, Boulder, Colorado, USA.

Vernon, D. E. 2004. Geospatial technologies in homeland security. Earth Observation Magazine, 13(1).