J 3.2 URBAN NET: URBAN ENVIRONMENT MONITORING AND MODELING WITH A WIRELESS SENSOR NETWORK

Paul J. Croft*, Feng Qi, Patricia Morreale, and Alicja Trzopek
Kean University, Union, New Jersey

1. INTRODUCTION

Use of sensor networks for environmental monitoring of the physical, chemical, or environmental conditions such as temperature, pressure, motion, or pollutant concentrations and transport; has been accomplished in a variety of settings. Whether for agricultural purposes or to understand local micrometeorological and microclimatological phenomena and/or distributions (e.g., see the seminal works of Geiger, Sutton, and Landsberg) in rural or urban or mixed environments; sensor networks for monitoring are critical to both assessments of current conditions as well as to the prediction of future conditions.

In more recent and modern applications (e.g., see the Oklahoma Mesonet and Micronet projects) data collection is assisted by a variety of sensors, some wireless in design, to capture short term variation and fluctuation of multiple parameters. These have included orchard and vineyard systems as well as air quality management platforms. Collection typically focuses on atmospheric variables, their states, and changes over time with regard to characteristic measurements of the phenomena or "issues" of interest (e.g., water stress, leaf temperature, and gas concentrations).

Fundamentally, while these applications are similar in terms of system models of data collection and analysis; they are not so in terms of instrumentation used, sensor exposure, sampling rates/times, nor are they of comparable dimensions as each is directed or focused according to an "issue" of key interest as determined by the situation (or system under consideration). Therefore any monitoring is necessarily driven by the appropriate scenarios which dictate a response by the system (in this case, human intervention). Based on this shift in emphasis, data collection becomes more data synthesis to describe the state and flux of the system rather than one that is concerned with the individual parameters or variables being measured. While this does not negate the importance of quality control and quality assurance, the significance of these are limited in such an operational setting.

This paradigm is particularly relevant to an urban or mixed-zone region in which there are a wide diversity of scenarios possible within the context of a wide variation in landscapes and the affected people or populations (i.e. more generally: plants, animals, et cetera). While select scenarios may be viewed as more regional threats (e.g., local street flooding) others may be considered more localized (e.g., air pollutant levels near a source). In many cases, scenarios that may apply to an entire region will vary in their impacts due to the distribution, character, and frequency of the observed conditions and the reception of these by the affected location (i.e. as per a location's characteristic properties and behaviors).

Within a region the susceptibility in the reception of any range of conditions must then be considered with regard to monitoring, inquiry, analysis, interpretation, and operational decision-making. These are weighted according to the size and nature of populations affected (i.e. whether plants or animals or people or local economic interests) and must be built into an observational system for it to be both comprehensive and responsive to the user needs. Indeed, the repetitive nature of conditions may in themselves create a recurring hazard locally that can be viewed as a public safety issue. Once again, the location's characteristic properties and behaviors will be relevant to the expected outcomes and impacts.

In an urban environment these hazards may range from an underpass that consistently floods during heavy rains to poor air quality that degrades over the course of a day. In a more rural environment it may be high winds causing tree or structural damage. In either situation there is a need to consider any hazard with regard to all baseline information to gauge its seriousness and the level of intervention response that may be needed. If this can be pre-determined, then a hazard or threat may be considered in an operational planning mode to better hone the response – offering a more targeted and

* *Corresponding Author Address*: Dr. Paul J. Croft – Professor of Meteorology, Kean University, Department of Geology and Meteorology, 1000 Morris Avenue, Union, NJ 07083; email <pcroft@kean.edu>.

effective response as well as one that is likely more cost-effective – in order to protect local interests.

To aid in such a response, and to improve the planning process for a response to a predicted hazard, a variety of tools are necessary to tie-in disparate data sets that portray the system in question relative to hazards. The application of a GIS framework is indispensable as it allows for the layering of key information as well as metadata and other information with analytic tools. For the situations described previously the acquisition, integration, and display of traffic and road patterns (e.g., through street-level mapping), identification of any critical receptors (people, economies, plants, and animals) based on specific thresholds, determination of available emergency resources, and cataloging of additional critical factors relevant to the location provide for local avoidance, mitigation, and preventive measures in advance of a hazard.

The implementation of measures to address any hazards, and re-assessment of the local measures taken during the evolution of a hazard, allow for real-time adjustment in the local response. In order to demonstrate the desirability of such an observational system a prototype was developed and implemented in and around Kean University. The campus lies within an urbanized region, typical of the area, but consists of a variety of landscape characteristics to be representative of multiple scenarios. Using both small-scale wireless and conventional sensor network instruments, suitable for urban and more rural environmental monitoring and modeling, a proof-of-concept study was made in a region that is highly diversified (i.e. in terms of populations, land, economies). Data was collected in near real-time for visual display, determination of baseline conditions, and for illustration of how they could be used in predictive models (both statistical and numerical) following their use in data mining and correlation applications.

2. DATA COLLECTION & METHODOLOGY

Kean University (www.kean.edu), located in the city of Union, New Jersey, is found within the New York Metropolitan area and within the megalopolis from Boston, Massachusetts to Washington, D.C. (Fig. 1a). The campus is located in Union County, the most densely populated county in the state with 5,000 people per square mile and is near Newark Liberty International Airport (Fig. 1b). The study required both environmental data collection (using sensor networks and hand-held devices) and compilation of GIS data layers of detailed attribute information. The GIS framework was critical to later visualization of data and its interpretation and crucial to demonstrate the viability of an operational application for use by a variety of decision-makers.

Data collection focused on key and standard environmental variables for the initial demonstration of the monitoring and decision-support system. These included primarily atmospheric data (temperature, light, pressure, and humidity) through hand-held PASCO (www.pasco.com/) and DATAHARVEST (www.dataharvest.co.uk) sensors. These sensors were also left in situ for longer term data collection as needed. Measurements of sound and carbon dioxide were also made, although much less frequently, due to the complexity of the observational methods. These two systems provided in-situ observations as well as extended data collection (unattended) at select locations in and around campus.

Several Crossbow wireless sensors were used (www.xbow.com/) by the Computer Science Department to develop a prototype operational system. The intent was to create a multi-node wireless sensor network that could be calibrated to the environmental feature of interest and deployed for instantaneous reporting of conditions in any setting. Measurements by all systems were made at several locations (including indoors for testing purposes) over multiple days and times to gain a sample data base for analysis and visualization. Sampling sites, although varied for demonstration purposes, are shown in Fig. 2 for the outdoor observations.

GIS database compilation involved the purchase of a high resolution NAIP real color orthoimagery of Union County (captured in spring/summer 2006). This served as a base for the spatial-visual framework in which layers of campus buildings, trees, lawns, stadiums, sidewalks, and other features were digitized based on the orthoimagery. Layers such as flow lines, water bodies and roads were obtained from New Jersey Department of Environmental Protection while landmarks, schools, and other human features were extracted from the ESRI GIS Data archives online (www.esri.com/index.html). Demographic data was extracted from 2002 census. The steps of clipping, digitizing, georeferencing, rectification, overlay, classification, map design, online map publishing, spatial interpolation, and data mining were used to create the root database for the project. This provided information about the study region’s detailed spatial mosaic of a variety of physical and human landscape. Similar procedure and protocol would be followed to scale up the system by covering broader regions (Fig. 3).

An online mapping site was established to provide public access to the base layers as well as environmental monitoring results, as would be expected in real-time applications for practical use, and in order to provide an operational template (http://131.125.4.75/ContourMaps/Default.aspx) for the manner in which data and analyses would be displayed by user communities in the decision-making process. The site allows user-guided inquiry and query. Animated visualization of monitored environmental variables was also created to display dynamics of changes for interpretation of relationships and temporal patterns. The intent is for a user to manipulate the current data with regard to background or baseline information so as to determine an appropriate course of action, if any, is or will be needed in the local area.

3. ANALYSIS OF EVENTS

While the demonstration system was meant for multiple scenario applications, one basic selection was examined with regard to its attendant conditions (i.e. temperature, humidity, and carbon dioxide) to provide a simple system diagnostic that would show the variations of temperature and air quality across and near campus as a function of the atmosphere, traffic flow, land cover, and related variables. The scenario offers flexibility in that it may be used in other regions or specific locations and is centered on the user community. This implies that the system would be helpful for high impact, short duration, and highly localized phenomena and/or hazards of interest.

Data for this basic scenario was collected on several days (see sample in Fig. 4) and then provided for data mining. Data mining is a knowledge discovery process that may include various methods including cluster analysis, predictive modeling, anomaly detection, classification, and association analysis. In the prototype here both predictive modeling and anomaly detection were used to ascertain any discrepancies or identify any unusual evolutions of the raw data in terms of its values, rates of change, interactions, and spatial distributions. These were also considered with regard to "normal" or expected evolutions as defined by micrometeorological processes.

This sort of application allows for the objective identification of critical threshold values upon which decisions can be made – and upon which advance planning may operate. By doing so, crucial thresholds can be monitored and assessed in real-time so that any action or alert may be automatic and highly responsive. The detection of anomalies also helps to isolate events or measured characteristics which are different from the rest of the data or the expected measurement. These anomalies are often the source of the understanding of rare or infrequent events that merit specific courses of action. However, since not all anomalies are critical events, some require further investigation. A good anomaly detection mechanism must be able to detect non-normal events or measurements, and then validate such events as being outside of expectations – a high detection rate and low false alarm rate is desired, as these define the critical success rate of the application.

Initial analyses indicated that discernment of anomalies and any unusual evolutions was difficult given their dependence on each location's specific characteristics and behaviors (i.e. how each of these interacted with the prevailing environmental regime). This verified that the significance of data collection lay in the distribution of the observed parameter and its change with time and in space. Therefore, it was more important to capture the field of each variable rather than rely on specific values at any one location. To do so, spatial interpolation was conducted to create contour maps that display in a toggle fashion so that a user might assess the importance of any one variable with regard to the location under consideration (Fig. 5). The contour maps were also published online for public access. This approach provides timely and direct insight to the environmental "issue" of significance and thus would help to direct the appropriate response to a given hazard (e.g., avoid, mitigate, prevent).

ACKNOWLEDGEMENTS

The Department of Geology and Meteorology and the Department of Computer Science faculty and staff at Kean University were critical in their assistance and support infrastructure. The assistance of Mr. William Heyniger, Professional Services Specialist was greatly appreciated as was help provided by several other students during the course of this project.

Figure 1a. Kean University campus in Union, New Jersey

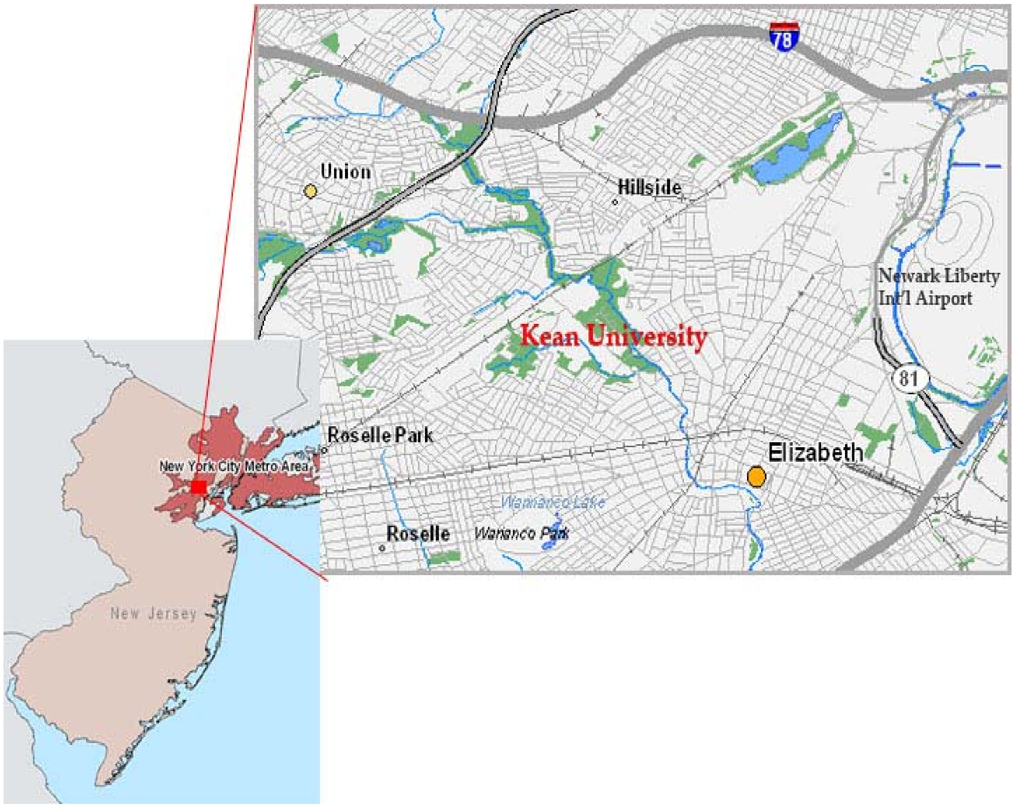

Figure 1b. Kean University campus with regard to Newark Liberty International Airport and nearby monitoring sites (indicated by stars).

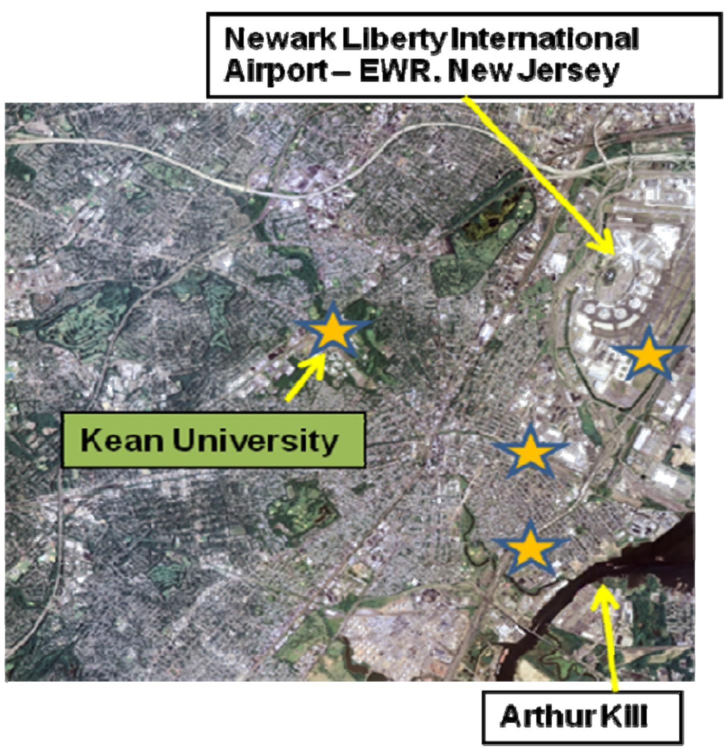

Figure 2. Sampling sites on Kean University Campus

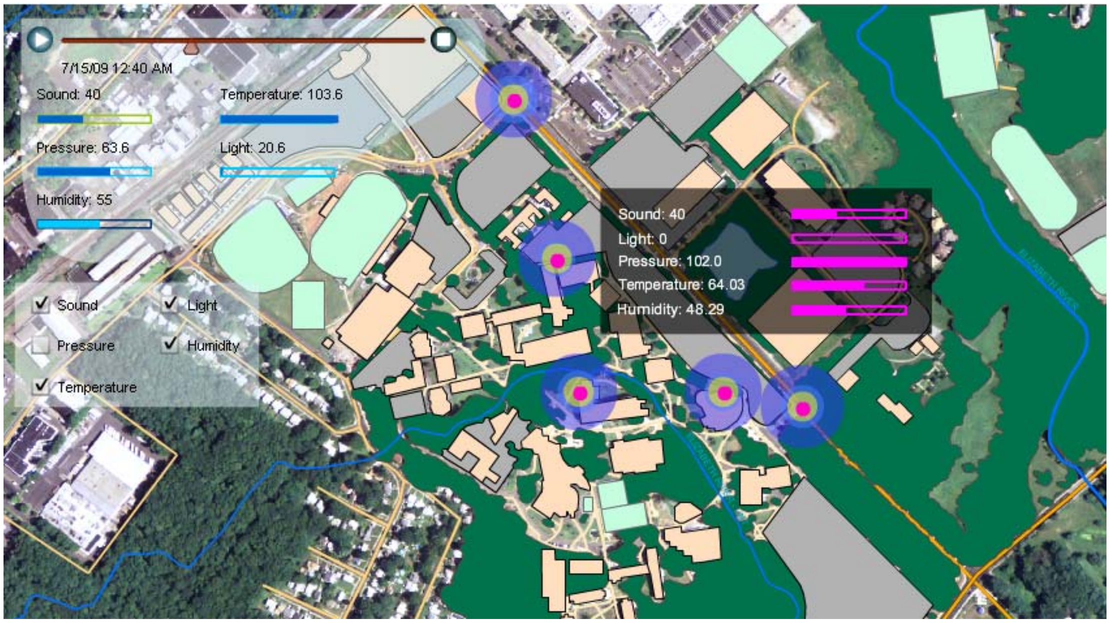

Figure 3. GIS data layers and development process.

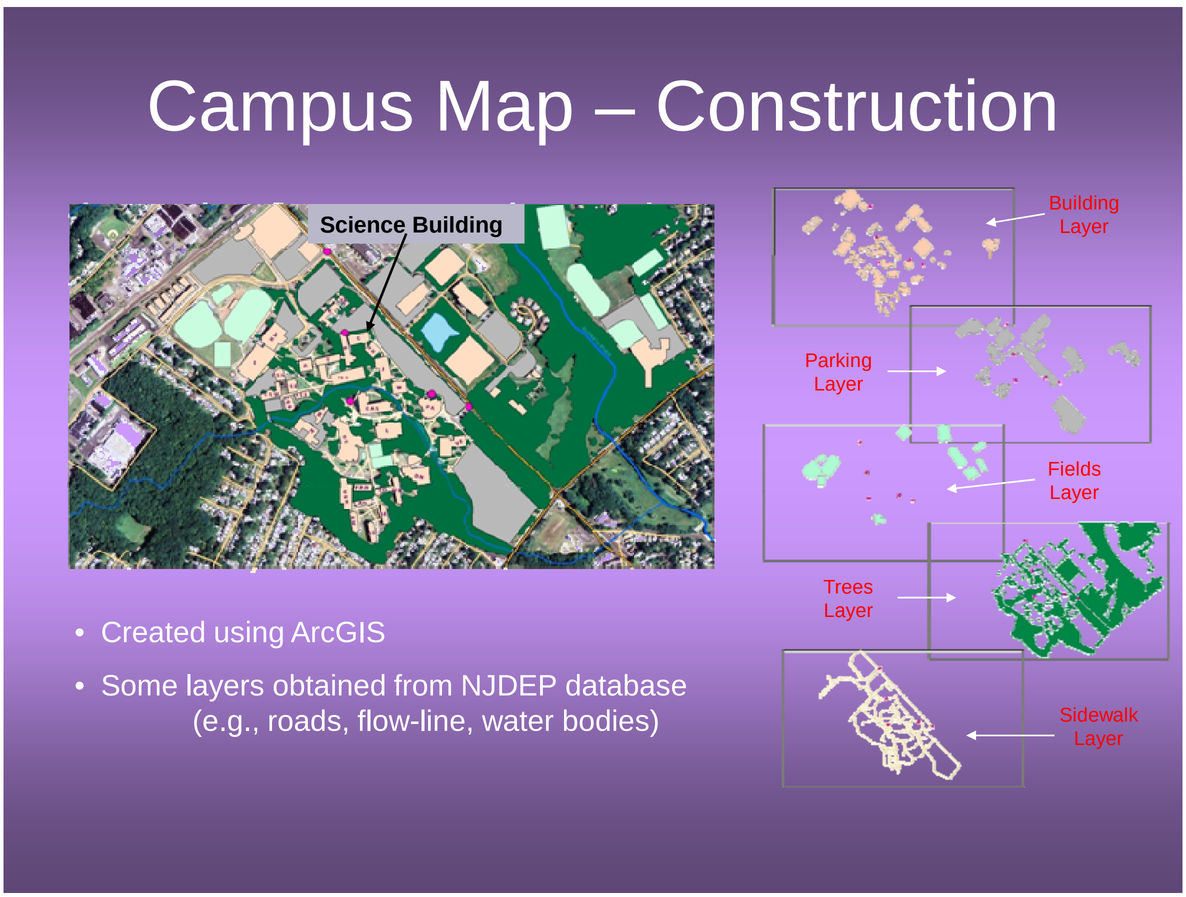

Figure 4. Sample of the data plots (as indicated) used in Data Mining applications for analysis.

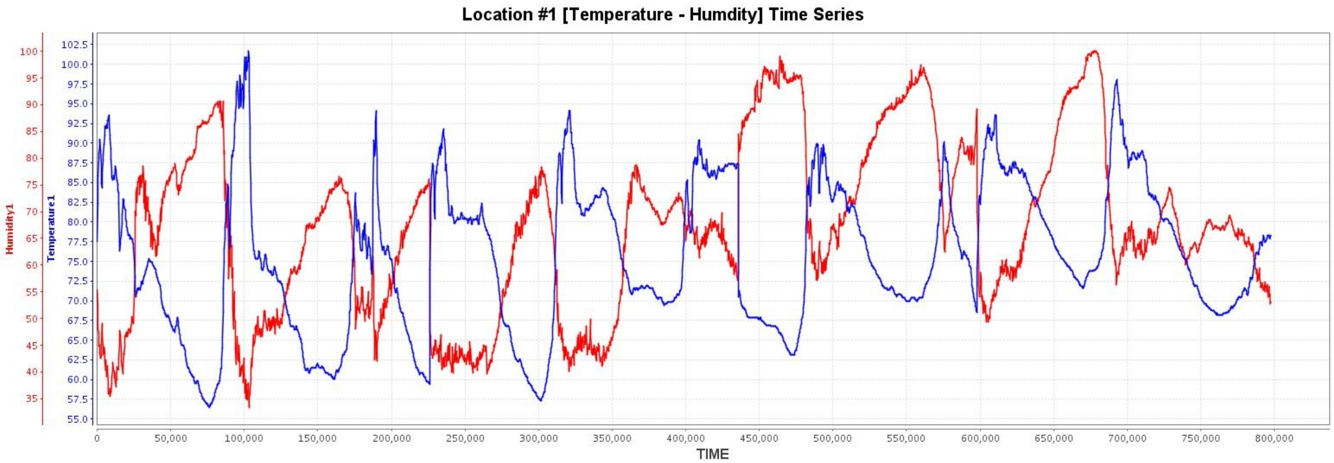

Figure 5. Sample contour analysis of field variables of interest (carbon dioxide) as associated with varying conditions in time (days)