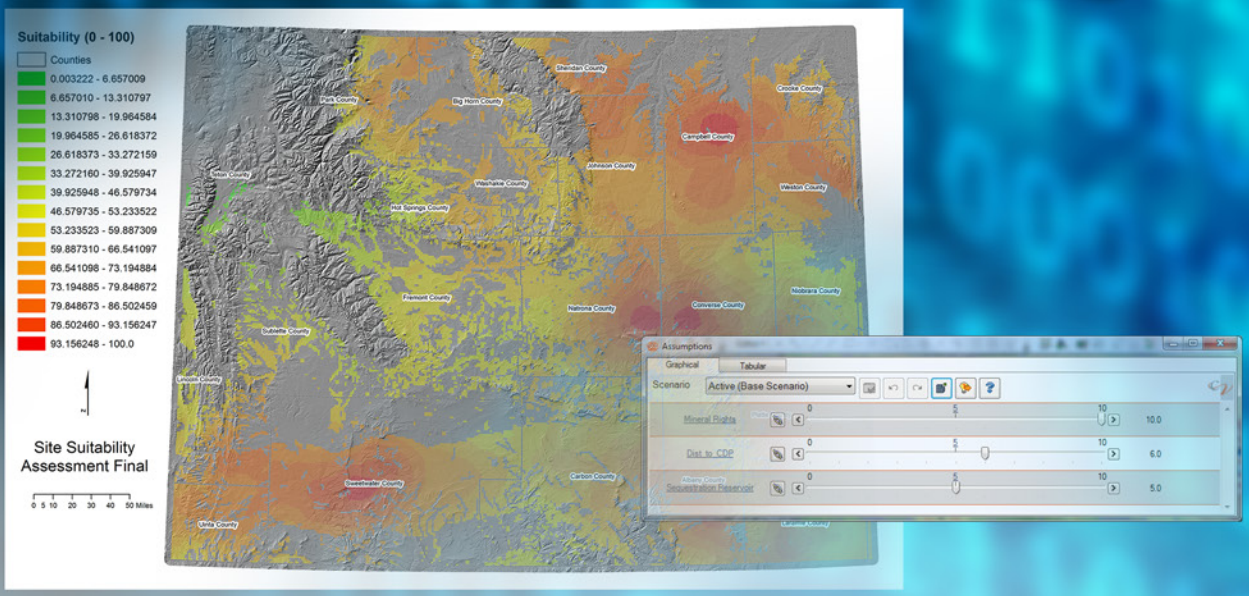
Vol. 26 No. 1

Journal

of the Urban and Regional Information Systems Association

URISA®



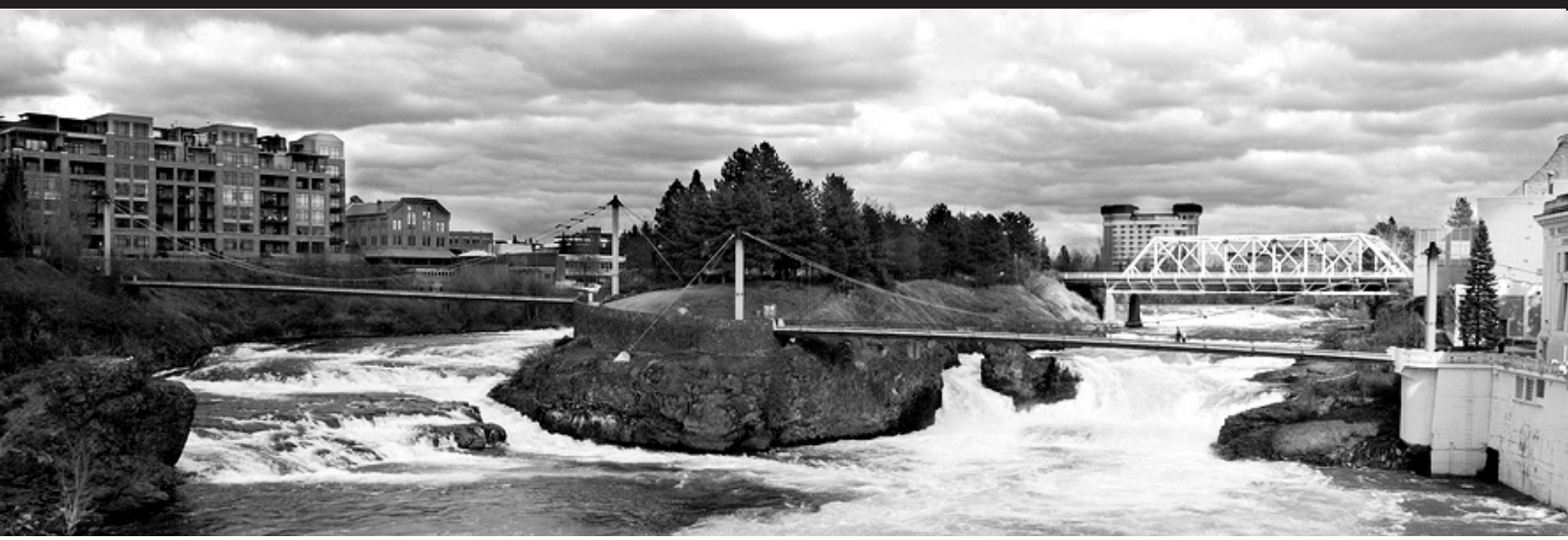
Indicator Analysis for Unpacking Poverty in New York City

Integrating Planning Support Systems and Multicriteria Evaluation for Energy Facility Site Suitability Evaluation

Implementing a Utility Geographic Information System for Water, Sewer, and Electric: Case Study of City of Calhoun, Georgia

Desirable Characteristics of an Online Data Commons for Spatially Referenced, Locally Generated Data from Disparate Contributors

www.urisa.org



Mark Your Calendar!

GIS-Pro 2015: URISA's 53rd Annual Conference

October 18-22, 2015 • Spokane, Washington

We are pleased to announce that next year's URISA annual conference will be presented in partnership with the Northern Rockies URISA Chapter and the Northwest GIS Users' Group, along with the Washington URISA Chapter!

Expect a phenomenal conference program in a beautiful setting! Natural, walkable, friendly… Spokane is a vibrant city, a roaring river, a gateway to the American west. Book a whitewater adventure, shred fresh powder, bike the beautiful Centennial Trail or hike through countless nature preserves. But it's not just about the outdoors! Taste Washington's renowned grapes and hops at any number of local wineries and craft breweries.

Schedule:

Call for Presentations:
Abstract submissions due by March 2, 2015
Exhibitor & Sponsor Opportunities: March 2015
Conference Program & Registration: April 2015

Volume 26 • No. 1

URISA®

Journal of the Urban and Regional Information Systems Association

Contents

5 Indicator Analysis for Unpacking Poverty in New York City
Jochen Albrecht and Mimi Abramovitz

13 Integrating Planning Support Systems and Multicriteria Evaluation for Energy Facility Site Suitability Evaluation
Scott N. Lieske and Jeffrey D. Hamerlinck

25 Implementing a Utility Geographic Information System for Water, Sewer, and Electric: Case Study of City of Calhoun, Georgia
Davie Crawford and Ming-Chih Hung

35 Desirable Characteristics of an Online Data Commons for Spatially Referenced, Locally Generated Data from Disparate Contributors
James Campbell and Harlan Onsrud

URISA Journal

Publisher: **Urban and Regional Information Systems Association**

Editor-in-Chief: **Dr. Piyushimita (Vonu) Thakuriah**

Journal Coordinator: **Wendy Nelson**

Electronic Journal: **http://www.urisa.org/resources/urisa-journal/**

EDITORIAL OFFICE: Urban and Regional Information Systems Association, 701 Lee Street, Suite 680, Des Plaines, Illinois 60016; Voice (847) 824-6300; Fax (847) 824-6363; E-mail info@urisa.org.

SUBMISSIONS: This publication accepts from authors an exclusive right of first publication to their article plus an accompanying grant of non-exclusive full rights. The publisher requires that full credit for first publication in the *URISA Journal* is provided in any subsequent electronic or print publications. For more information, the "Manuscript Submission Guidelines for Refereed Articles" is available on our website, www.urisa.org, or by calling (847) 824-6300.

SUBSCRIPTION AND ADVERTISING: All correspondence about advertising, subscriptions, and URISA memberships should be directed to: Urban and Regional Information Systems Association, 701 Lee Street, Suite 680, Des Plaines, Illinois 60016; Voice (847) 824-6300; Fax (847) 824-6363; E-mail info@urisa.org.

URISA Journal is published two times a year by the Urban and Regional Information Systems Association.

© 2015 by the Urban and Regional Information Systems Association. Authorization to photocopy items for internal or personal use, or the internal or personal use of specific clients, is granted by permission of the Urban and Regional Information Systems Association.

Educational programs planned and presented by URISA provide attendees with relevant and rewarding continuing education experience. However, neither the content (whether written or oral) of any course, seminar, or other presentation, nor the use of a specific product in conjunction therewith, nor the exhibition of any materials by any party coincident with the educational event, should be construed as indicating endorsement or approval of the views presented, the products used, or the materials exhibited by URISA, or by its committees, Special Interest Groups, Chapters, or other commissions.

SUBSCRIPTION RATE: One year: $295 business, libraries, government agencies, and public institutions. Individuals interested in subscriptions should contact URISA for membership information.

US ISSN 1045-8077

Editorial Board

URISA Journal Editor

Editor-in-Chief
Dr. Piyushimita (Vonu) Thakuriah, *Department of Urban Studies and School of Engineering, University of Glasgow, United Kingdom*

Thematic Editors:

Sustainability Analysis and Decision Support Systems:
Timothy Nyerges, *University of Washington*

Participatory GIS and Related Applications:
Laxmi Ramasubramanian, *Hunter College*

Social, Economic, Governance and Political Sciences:
Francis Harvey, *University of Minnesota*

GIScience:
Harvey Miller, *University of Utah*

Urban and Regional Systems and Modeling:
Itzhak Benenson, *Tel Aviv University*

Editorial Board

Jochen Albrecht, *Hunter College*

Peggy Agouris, *Center for Earth Observing and Space Research, George Mason University, Virginia*

David Arctur, *Open Geospatial Consortium*

Michael Batty, *Centre for Advanced Spatial Analysis, University College London (United Kingdom)*

Kate Beard, *Department of Spatial Information Science and Engineering, University of Maine*

Yvan Bédard, *Centre for Research in Geomatics, Laval University (Canada)*

Itzhak Benenson, *Department of Geography, Tel Aviv University (Israel)*

Al Butler, *GISP, Milepost Zero*

Barbara P. Buttenfield, *Department of Geography, University of Colorado*

Keith C. Clarke, *Department of Geography, University of California-Santa Barbara*

David Coleman, *Department of Geodesy and Geomatics Engineering, University of New Brunswick (Canada)*

Paul Cote, *Graduate School of Design, Harvard University*

David J. Cowen, *Department of Geography, University of South Carolina*

William J. Craig, *GISP, Center for Urban and Regional Affairs, University of Minnesota*

Robert G. Cromley, *Department of Geography, University of Connecticut*

Michael Gould, *Environmental Systems Research Institute*

Klaus Greve, *Department of Geography, University of Bonn (Germany)*

Daniel A. Griffith, *Geographic Information Sciences, University of Texas at Dallas*

Francis J. Harvey, *Department of Geography, University of Minnesota*

Bin Jiang, *University of Gävle, Sweden*

Richard Klosterman, *Department of Geography and Planning, University of Akron*

Jeremy Mennis, *Department of Geography and Urban Studies, Temple University*

Nancy von Meyer, *GISP, Fairview Industries*

Harvey J. Miller, *Department of Geography, University of Utah*

Zorica Nedovic-Budic, *School of Geography, Planning and Environmental Policy, University College, Dublin (Ireland)*

Timothy Nyerges, *University of Washington, Department of Geography, Seattle*

Harlan Onsrud, *Spatial Information Science and Engineering, University of Maine*

Zhong-Ren Peng, *Department of Urban and Regional Planning, University of Florida*

Laxmi Ramasubramanian, *Hunter College, Department of Urban Affairs and Planning, New York City*

Carl Reed, *Open Geospatial Consortium*

Claus Rinner, *Department of Geography, Ryerson University (Canada)*

Monika Sester, *Institute of Cartography and Geoinformatics, Leibniz Universität Hannover, Germany*

David Tulloch, *Department of Landscape Architecture, Rutgers University*

Stephen J. Ventura, *Department of Environmental Studies and Soil Science, University of Wisconsin-Madison*

Barry Wellar, *Department of Geography, University of Ottawa (Canada)*

Lyna Wiggins, *Department of Planning, Rutgers University*

F. Benjamin Zhan, *Department of Geography, Texas State University-San Marcos*

Indicator Analysis for Unpacking Poverty in New York City

Jochen Albrecht and Mimi Abramovitz

Abstract: *This article presents work that is part of a larger and ongoing research agenda exploring the persistence of health and social problems in some parts of New York City. To this end, the authors have developed a GIS framework that translates a highly diverse set of variables into neighborhood indicators that can help local residents as well as decision makers to understand the relationship between "place" and individual behavior. Using the example of two new indices, Community Loss and Neighborhood Risks, the readers will learn how data can be transformed to emphasize the communal nature of phenomena that is typically understood only in relations to individuals.*

INTRODUCTION

New York City may not be the most segregated city in the country. But it is notorious for hosting some of the wealthiest and poorest neighborhoods in the country – sometimes in close proximity. Many of these neighborhoods have not changed their economic status for many decades. Some neighborhoods such as the South Bronx or East New York together with East Los Angeles and Chicago's South Side have become synonymous with enduring social problems that persist despite considerable investment in a wide range of interventions. Looking a little closer at, for example, the South Bronx we find that the demographics have changed but the problems have not. These observations led the authors to develop a research framework for investigating the relationship between the local infrastructure or conditions under which people live and the concentration of health and social problems in some but not other New York City neighborhoods. That is, there something about the *place* rather than the *people* that makes the difference?

Part of a larger project designed to unpack poverty (Abramovitz & Albrecht, 2013) this article presents two new social indicators: Community Loss and Neighborhood Risk. Similar to the other place-based indicators in the larger project, they capture phenomena that others have previously studied only in relation to individuals. To ensure a focus on community conditions, the authors structured the project to avoid the "tautology trap" that arises when researchers describe neighborhoods in terms of the behavior of local residents (i.e., teen mothers, criminal behavior, school drop outs, etc.) and then conclude that those behaviors are concentrated in these neighborhoods. To that end, the project's independent variables consist of neighborhood conditions categorized as economic, housing, education, food, health, or environmental insecurities (without reference of the behavior of residents) and its dependent variable consists of problematic behaviors such as lack of self-care, self-medication, school drops outs, mental health problems, risky sexual behavior, criminal activities, interpersonal violence, etc. (see Figure 1). The project includes ameliorating factors such as self-advocacy, civic participation, and neighborhood resources like libraries, community centers, etc. This framework protects against "blaming the victim", which often happens when researchers attribute the concentration of health and social problems in poor neighborhoods to the behavior of local residents. In the final analysis, the Neighborhood Stress Projects asks "what happened to the neighborhood?" rather than "what did the residents do wrong"?

This article begins with a discussion of the conceptual underpinnings of the overall project, which seeks to understand what accounts for the concentration of health and social problems in some New York City neighborhoods. As the authors have described this framework elsewhere (Abramovitz & Albrecht 2013), this article focuses on the development of two new indicators for phenomena that scholars have previously ascribed to individuals but not to neighborhoods and on an outline of the GIS methods used to reframe traditional household-based variables into measures that recognize the role of place as an active actor. The ensuing analysis of the resulting maps confirms the hypothesis that exposure to accumulated disadvantage, i.e., living amidst

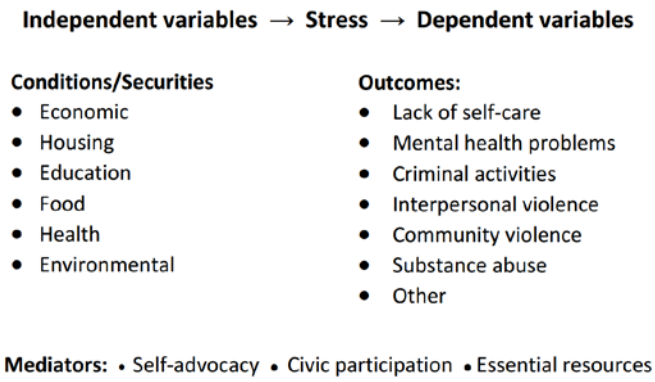


Figure 1. General overview of generic indicator categories (insecurities) for Hunter Neighborhood Stress project.

multiple and persistent adverse conditions at the same time, characterizes New York City neighborhoods known to have the highest concentration of health and social problems. The article ends with a discussion of policy implications and suggested further work on neighborhood indicators.

BACKGROUND

If we assume that nobody purposefully engages in behaviors to harm themselves or others (also known as social problems), individual, communities and policy makers can benefit from a better understanding of drives this kind of behavior that negatively affects communities as well as individuals. The model of Drawing on what is known about the ways in which stress affects behavior, the authors posit that stress operates as a pathway between adverse neighborhood conditions ("Place") and the concentration of health and social problems in some New York City neighborhoods (Abramovitz & Albrecht 2013).

Most discussions of stress focus on the individual and how to reduce the negative consequence of exposure to high levels of singular and acute stress resulting from mass disasters, or from multiple and chronic stress associated with daily life in impoverished neighborhoods. Few scholars examine community level stressors. Even when analyzing such eminently spatial phenomenon as Hurricane Katrina, geographic (GIS) researchers have tended to focus on the individual. Few geoscientists examine how exposure to either mass disaster or the less dramatic adverse local conditions affects the social fabric of neighborhoods. To correct for this singular focus on individuals, the authors ask how exposure to accumulated disadvantage affects community functioning.

The research is informed by Hobfoll's (1989) Conservation of Resource (COR) theory, which is relevant to the experiences of low-income people and communities that are already financially strained. COR theory suggests that the struggle to secure and sustain basic resources can lead to a downward spiral of resource loss that, in turn, may effectively drain an individual's ability to cope effectively. The authors apply COR to neighborhoods and argue that when large numbers of people are exposed to multiple neighborhood-based stressors at the same time, the experience can drain the community's capacity to function. Fullilove (2004) uses the term "root shock" to describe the stress reaction to the loss of one's emotional ecosystem as a consequence of urban renewal projects. The dependent variable list above includes many of the ways that people cope with stress.

The social indicators introduced here move from the study of individuals to the study of communities. The experiences of loss and risk have previously been ascribed to individuals. The Community Loss and Community Risk indicators assume that since local communities are places of interaction and interdependence, something happens to communities when a large number of people living in close proximity regularly suffer multiple losses and risks at the same time. This differs from the ways that most researchers use neighborhood indicators. Where individual data is not available, they describe populations and then infer about individuals. GIS allows us to aggregate individual experiences. With this, the Community Loss and Neighborhood Risk indicators shed light on the ways in which adverse local conditions affect community-wide functioning. The new indicators paint a picture of New York as a "tale of two cities", in which New Yorkers live in different, and some would argue incomparable neighborhoods. By identifying variations in smaller geographic units, the research also unpacks poverty and disrupts the view of poverty as a uniform experience.

COMMUNITY LOSSES AND RISKS

The concept of Community Loss was not part of the original set of insecurities depicted in Figure 1; rather it emerged from the data itself and reflects the notion that there are tangible community-level resources that are an integral component of

Table 1. Community loss variables, their spatial foot print and their sources.

Loss variable	Unit of Measurement	Spatial resolution	Data source
Long-term hospitalization	Hospitalizations lasting longer than 180 days divided by number of households	ZCTA	NY Statewide Planning and Research Cooperative System (SPARCS)
Unemployment	Number of people receiving unemployment insurance divided by the number of households	Census track	US Census ACS
Incarceration	Incarcerations per ZIP code area divided by number of households	Home address	NYS Prison Administration
Foster placement	Placements per ZIP code area divided by number of households	ZCTA	NYC Administration for Children and Families
Untimely death	Given as a rate 1/1,000	Community district*	NYC Department of Health
Foreclosure	Relative need value compared to the neediest in New York State as per HUD calculation	ZCTA	Local Initiatives Support Cooperation Center of Housing Policy, Urban Institute
* see Methods section for redistribution of community district-level data to ZCTAs			

the community beyond the well-studied losses experienced at the individual level. They are grouped here into the removal of people and the removal of material assets (see Table 1). Missing people include individuals removed from the home and community due incarceration, foster care placement, premature death, and long-term hospitalization. The missing assets include loss of job and home due to unemployment and foreclosure. Other measures of missing people were excluded such as college students living out of state, or deployed members of the armed services. Students were excluded as they leave voluntarily, which suggests minimal stress. Armed services personnel were also excluded. To the surprise of the investigators, they turned out to be recruited in almost equal proportions from all parts of New York City. As such deployment was not a spatially distinguishable phenomenon. In the category of missing assets, the project excluded library and hospital closings because they were too rare to have a statistically significant impact; and school closings were more than made up by the creation of new schools[1]. The authors failed to find city-wide high resolution data on business closings and were quite surprised about the lack of job loss data at spatial resolutions smaller than Public Use Microdata Areas (PUMAs).

The authors gained access to census tract-level unemployment figures from both early and late 2008, i.e., just before and after the last recession through a special FOIA request of a colleague working on another project. This was especially helpful given that the US Census American Community Survey (ACS) data compromises on either the temporal or spatial resolution and would therefore not have been useful for this study. The only physical (loss of) assets measure used in this study are foreclosures, which in a city with as high a percentage of rental units as New York City introduces some caveats. Data sources and preparation of each component of the loss indicator will be described in the methods section of this article.

As with the community loss indicator, the concept of neighborhood risks were not part of the author's original list of insecurities (Figure 1). They too emerged from the subsequent compilation of data, (see Table 2). Cutter (1995), Evans & Marcynyszyn (2001), and Schlosberg (2007) have identified the constant presence of hazards as an environmental justice issue. They are here expanded to include structural fires and traffic injuries, in addition to the environmental nuisances associated with bus/truck depots and garages. Other hazards were explored but rejected because of they had no significant spatial variation or were too similar to those already included in the index. They included school crime rates, bullying in schools, building vacating orders, complaints about rats, traffic deaths, and noise complaints (highly correlated with bar/restaurant activities).

Neighborhood fears include weapons confiscated in one's neighborhood, prosecutions by Immigration and Customs Enforcement (ICE), the presence of registered sex offenders,

[1] School closing may nevertheless be considered as disruptive but without exception, the argument of the authorities was that the schools were failing to provide their students with an adequate education – which arguably would have a larger long-term negative impact.

Table 2. Hazard and fear variables, their spatial foot print and their sources.

Risk variable	Unit of Measurement	Spatial resolution	Data source
Depots and garages	Total number of MTA, NYPD, Sanitation and school bus depots per ZCTA	Address	NYC Department of Information Technology (DOITT)
Ladder runs (fires and building collapses)	Total number of ladder runs divided by number of households	Address	NYC Fire Department
Traffic injuries	Traffic injuries resulting in bodily harm divided by daytime population in census tract	Address	NYC Department of Transportation
Perception of unsafe schools	Percent of parents surveyed that perceived their children's school as unsafe*	School address	NYC Department of Education
Weapons confiscated in a Stop & Frisk (S&F)	S&F incidents where weapons were confiscated divided by number of households	Address of Stop & Frisk	NYC Police Department
Prosecutions by Immigration and Customs Enforcement	ICE Apprehensions divided by number of households	ZCTA	Families for Freedom and the Immigrant Defense Project (who FOILed the data from ICE)
Sex offenders Total number of registered sex offenders per ZCTA Residential address NY State Division of Criminal Justice Services			
* see Methods section for redistribution of address-level data to ZCTAs			

and parental perception of lack of safety in schools. Like loss, fear is a well-known individual stressor (Nasar & Jones 1997, Dohrenwend 1998). As with loss, the impact of omni-present risks (hazards and fears) have on communities had not been studied. A look at social science literature but also real estate reports such as Neighborhood Scout (2014) or Better Homes and Gardens (2014) clearly demonstrate that fear can have a debilitating effect on the neighborhood as a whole.

METHOD

Other than US Census Data, it was not easy to collect the other data needed for this project. Given that the US Census Bureau does not include relevant information at the fine spatial resolution necessary to describe phenomena at the neighborhood level, it was necessary to obtain administrative data from New York City agencies. However, New York City lacks open data for many dimensions of neighborhood life, so for administrative data collection the authors relied on personal networks of professional and academic colleagues as well alumni who now occupy important administrative positions in New York City. For these and other reason research in other part of the country might use a different set of variables.

A classic challenge when working with spatial data from a variety of different source agencies is the change in support of what is also known as the modifiable area unit problem (Openshaw 1983, Wong 2009, Kwan 2012). Similarly, there is no clear definition as to what constitutes a neighborhood. In the case of New York City, the term is used for the political outline of community districts (NYC DCP 2014a), the neighborhood planning areas of the Mayor's Office (NYC DCP 2014b), the marketing terms of real estate agencies (Zillow 2014) and crowd-sourced attempts such as NYCWiki (2014) neighborhood descriptions. With the exception of health data, all other data used in this study is originally available at either the ZIP code tabulation area (ZCTA), US Census tract, or individual address level. ZCTAs were finally chosen as the spatial support for this study based on the following two arguments:

1. For a number of data sets, this spatial footprint is available natively, or can be aggregated to from finer resolution data;
2. The scale of analysis should represent the behavioral space of an average citizen, in New York City that is approximately two square miles[2].

Two data sets required significant spatial adjustments with subsequent uncertainties about the true spatial footprint. Health data, although internally available at the ZCTA scale, is publicly released only at aggregations of on average five ZCTAs. The authors rasterized the data and then used *pycnophylactic* interpolation (Tobler 1979) to redistribute death rates in the Community Loss Index to where people actually live (which is available at a very high spatial resolution). The school-based variables in the Neighborhood Risk Index, available at address level, were redistributed in a three-step process. First, Thiessen polygons where created using enrolment figures as weights. These were then overlaid with Census data on the number of school-aged children to assign each census tract to one school or another, which would then inherit the school-based attributes. The Census tracts were then finally aggregated to ZCTAs resulting in ZIP code-level school data. Both of these methods (especially for the health data) may not pass academic muster. But given the lack of alternatives, they are the best available approximation.[3]

The measurement scales available for each variable vary widely (for example, people per 100,000 households, per capita income, days of hospitalization, etc.). To make them comparable, the data were standardized into deciles (using Jenks natural breaks), where the lowest decile represents the neighborhoods with fewest adverse conditions and the highest decile represents areas with extremely high losses or risks. Jenks is regularly used in spatial data analysis because it divides the data into classes based on natural breaks and thus provides a scale based on actual distribution of the data's characteristics (Jenks 1967, Congalton 1991). This procedure was applied to each of the variables presented here. Thus, every neighborhood can be described and compared to the city as a whole on thirteen attribute dimensions. At the level of ranks, the constituent variables for both indicators (loss and risks) were then aggregated to depict the accumulated loss/risk for each neighborhood. The indices identify ZIP code areas where residents are regularly exposed to multiple losses and risks at the same time, denoting a stressed community. The data are presented in visual form on choropleth maps that use different colors or shades to depict the average values in each area. The maps of Figures 2 and 3 depict the distribution of each ranked variable as well as their accumulation in New York City.

DISCUSSION OF RESULTS

Figure 2 includes nine inset maps (a-i) that visualize community loss in New York City neighborhoods. (1) Six individual maps (b-g) depict the citywide distribution of each of the following losses: foster care placement, incarcerations, unemployment, long-term hospitalizations, pre-mature deaths, and foreclosures. (2) The aggregated loss map (inset a) is a composite of all 6 losses that effectively depicts high loss areas suffering multiple losses at the same time, creating a condition of accumulated disadvantage. (3) Detailed maps of an exemplary high (inset h) and low (inset i) neighborhood with bar charts that depict the variation of losses across different ZIP code areas.

[2] In urban planning, and here in particular in transit-oriented planning, US literature says that Americans are willing to walk ¼ mile to a transit stop. New Yorkers are willing to walk a lot more (on average 20 minutes) and faster (3 miles an hour), which amounts to covering a distance of one mile. Compromising to arrive at a conservative estimate and to include children and the elderly, we used a figure of 0.8 miles, which using the formula for the area of a circle results in approximately two square miles (Thompson 2007, America Walks 2013).

[3] The authors are taking pain to explain the caveats whenever they present the results to decision makers.

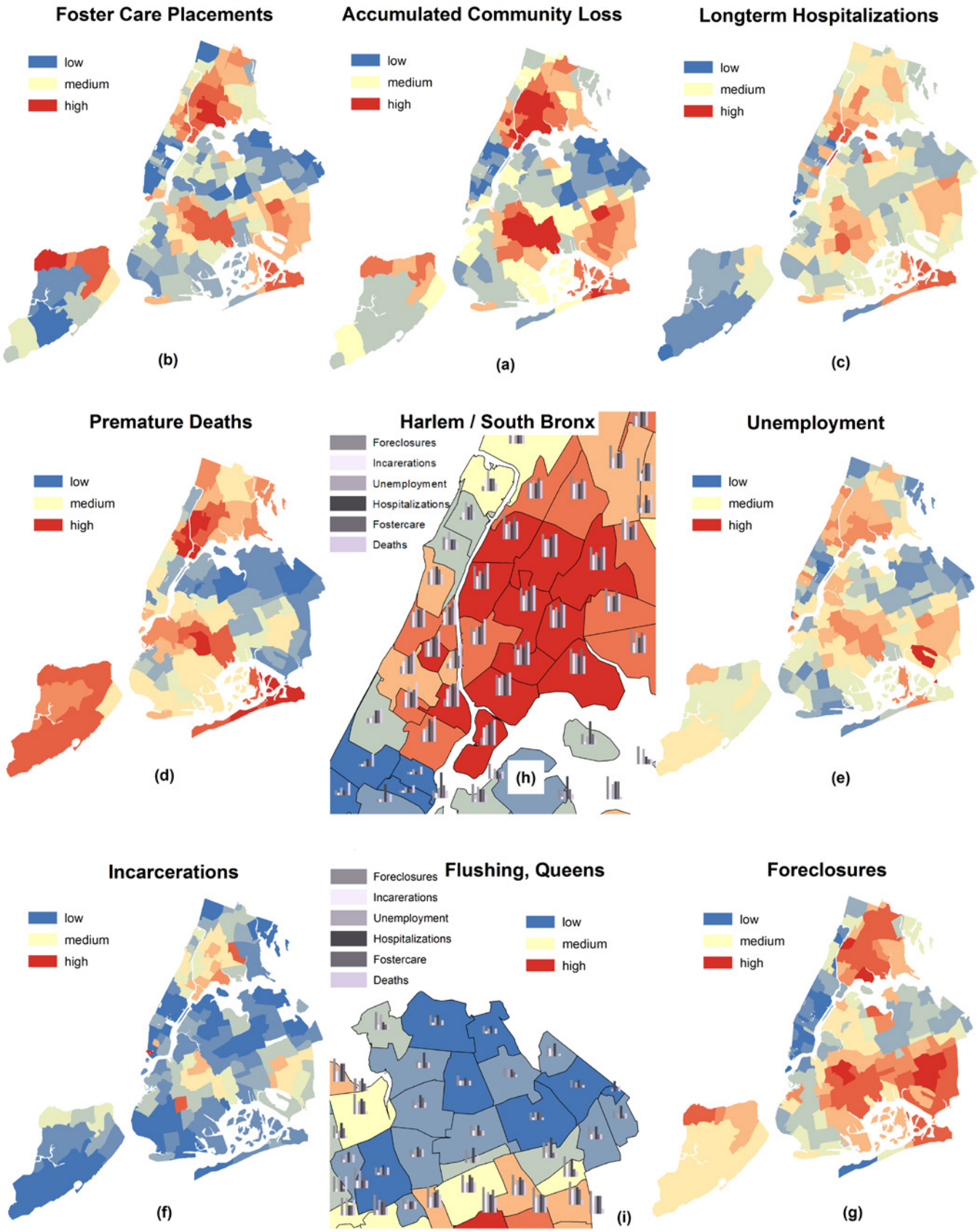


Figure 2. Geography of Community Loss in New York City.

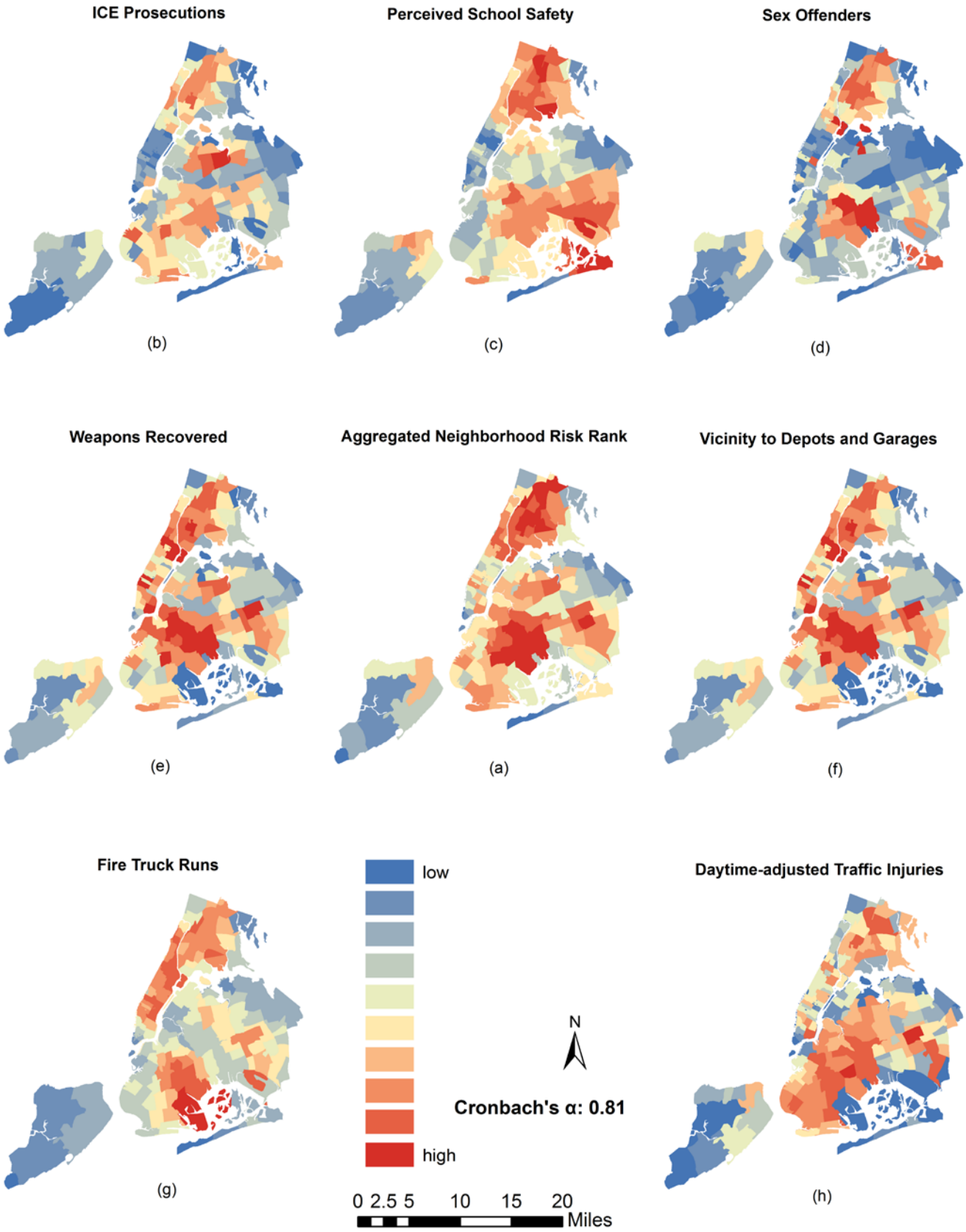


Figure 3. Geography of Community Risks in New York City.

In the high-loss areas, the rank of each of the six losses rises far above the citywide average of five with the exception of incarceration, whose rank of five matches the citywide average. With an average rank of eight, foster care placement consistently accounts for the most severe experience of community loss in the high-loss area. In the low-loss areas, the rank of each of the six losses falls far below the citywide average. Four of the losses (unemployment, foreclosure, untimely deaths, and long term hospitalization) all ranked just above or below three; foster care placement averaged two; and incarceration averaged one.

Taken together, all the Community Loss maps show New York City to be sharply divided by the experience of loss. Digging deeper into smaller spatial units reveals that the high and low loss areas are not all the same. This important variation effectively disrupts poverty as a uniform or singular experience. That is, the new social indicators make it possible to unpack poverty as well as document accumulated disadvantage.

Neighborhood risks are portrayed in Figure 3. Here, a ten-class visualization was chosen to illustrate the detail contained in the data (and somewhat washed over on the maps in Figure 2). The accumulated risks are represented in inset (a) – they show a large agreement with accumulated losses of Figure 2. Inset maps (b-h) render each individual community hazard and fear variable. They show a much higher degree of variation than the loss variables. This variability could be interpreted as the constituting variables to represent different phenomena. However, an analysis of internal consistency (Cronbach's) reveals that there is a very high likelihood for the eight variables to describe the same phenomenon, in this case: neighborhood risks.

The comparison of the two indicator maps with each other raises another question: What is the degree of congruence between the loss and risk areas? A non-spatial correlation analysis results in an r^2 of 0.64; that goes up to an impressive r^2 of 0.9 after accounting for distortions due to spatial autocorrelation. The rank difference between the two indicators rarely reaches 2.0 ranks and can usually be explained by the old housing stock (resulting in more fires) and higher traffic density in Manhattan. The only neighborhood that defies initial explanation for why the losses do not match risks is Corona, Queens. Corona is bordering the highlighted low-loss area in Figure 2 (i), and is the latest candidate for gentrification in New York City.

CONCLUSIONS

The research presented here disrupts the notion of poverty as a uniform event. In spite of significant differences among the contributing factors, there is overwhelming evidence that negative conditions accumulate in exactly those neighborhoods that are known to be the hearth of persistent social (and as we increasingly recognize also health) problems.

The methods are mostly part of the toolset of basic GIS analysis. The challenges (beyond the fact that New York City consists of over 200 neighborhoods resulting in pretty big datasets by the standards of indicator analysis) are mostly on the side of finding appropriate data and developing conceptual models that avoid tautological traps. The ability to drill down and compare areas in both a local (neighborhood) as well as a regional (all of New York City) context opens new doors for policy makers. This has become evident is the uptake of place-based rather than case-based initiatives by health and human services departments in the City as well as non-for-profit organizations.

This is a new chapter in the dialog between service providers, who in the past tended to work with very broad geographic brushes and community-based organizations who were limited by their myopic local knowledge and lacked the means to compare their neighborhoods with others.

Finally, the indicator building method, while not new to an academic audience, has now been demonstrated to and subsequently applied by local residents to allow them to set their own priorities in classic PPGIS fashion.

About the authors:

Dr. Jochen Albrecht is an Associate Professor for Computational and Theoretical Geography at Hunter College, City University of New York. His areas of interest include bridging the quantitative/qualitative divide, spatio-temporal modeling, and geospatial program management.

Dr. Mimi Abramovitz is the Bertha Capen Reynolds Professor of Social Policy at the Silberman School of Social Work at Hunter College and the CUNY Graduate Center. Her interests include social welfare policy, women and the welfare state, the impact of privatization on human service agencies and workers and the impact of social policy on "place". She conceptualized the accumulated disadvantage and stress framework for this study.

References:

Abramovitz M and J Albrecht 2013. The Community Loss Index: a new social indicator. *Social Service Review*, 87(4):677-724. DOI: 10.1086/674112.

America WALKS 2013. *Walking Facts.* http://americawalks.org/?page_id=4, online resource, last accessed 11/30/2014

Better Homes and Gardens 2014. http://www.bhgrealestate.com/Views/Look/, online resource, last accessed 11/30/2014.

Congalton RG 1991. A review of assessing the accuracy of classifications of remotely sensed data, *Remote Sensing of Environment,* 37 (1): 35-46.

Cutter S 1995. Race, class and environmental justice. *Progress in Human Geography,* 19(10: 111-122. DOI: 10.1177/030913259501900111.

Dohrenwend B 1998. Theoretical Investigation. In Dohrenwend B (Ed), *Adversity, Stress and Psychopathology,* pp 539-555. New York: Oxford University Press.

Evans GW and LA Marcynyszyn 2001. Environmental justice, cumulative environmental risk, and health among low-and

middle-income children in upstate New York, *American Journal of Public Health,* 94(11): 1942-1944, DOI: 10.2105/AJPH.94.11.1942.

Goodchild MF 2011. Formalizing place in geographic information systems. In Burton L, Kemp S, Leung MC, Matthews S and D Takeuchi (Eds), *Communities, Neighborhoods, and Health,* pages 21-33. New York: Springer.

Hobfoll SE 1989. Conservation of resources. A new attempt at conceptualizing stress. *American Psychologist,* 44 (3): 513–24. DOI:10.1037/0003-066X.44.3.513.

Jenks GF 1967. The Data Model Concept in Statistical Mapping, *International Yearbook of Cartography,* 7: 186–190.

Kwan MP 2012. The uncertain geographic context problem, *Annals of the Association of American Geographers,* 102(5): 958–968. DOI:10.1080/00045608.2012.687349.

Nasar J and KM Jones 1997. Landscapes of Fear and *Stress. Environment and Behavior,* 29:291-323. DOI:10.1177/001391659702900301.

Neighborhood Scout 2014. http://www.neighborhoodscoutcom, online resource, last accessed 11/30/2014.

NYC DCP 2014a. *DCP Community Portal.* http://www.nyc.gov/html/dcp/html/neigh_info/nhmap.shtml, online resource, last accessed 11/30/2014.

NYC DCP 2014b. *New York City, a city of neighborhoods.* http://www.nyc.gov/html/dcp/html/neighbor/, online resource, last accessed 11/30/2014.

NYCWiki 2014. *NYCWiki community portal.* http://nycwiki.org/wiki/Main_Page, online resource, last accessed 11/30/2014.

Openshaw S 1983. *The modifiable areal unit problem.* Norwich (UK): Geo Books.

Schlosberg D 2007. *Defining environmental Justice: theories, movements, and nature.* Oxford (UK), New York: Oxford University Press.

Thompson C 2007. Why New Yorkers Last Longer. *New York Magazine, News & Features.* Aug 20, 2007. New York. http://nymag.com/news/features/35815/index4.html, online resource, last accessed 11/30/2014.

Tobler W 1979. Smooth Pycnophylactic Interpolation for Geographical Regions. *Journal of the American Statistical Association,* 74: 519–36.

Wong D 2009. The modifiable areal unit problem (MAUP). In Fotheringham A and P Rogerson, *Handbook of Spatial Analysis,* pp. 105–124. London (UK): Sage Publications.

Zillow 2014. *Zillow neighborhood boundaries.* http://www.zillow.com/howto/api/neighborhood-boundaries.htm, online resource, last accessed 11/30/2014.

Integrating Planning Support Systems and Multicriteria Evaluation for Energy Facility Site Suitability Evaluation

Scott N. Lieske and Jeffrey D. Hamerlinck

Abstract: *Suitability modeling seeks to identify the continuum of best or worst sites for a facility, development, activity, or particular use based on site characteristics and preference assessments. The purpose of this paper is to assess the quality of a planning support system-based suitability model through presentation of an energy infrastructure development case study. The presentation of the case includes the suitability model and a second, complementary approach, the Kepner-Tregoe problem-solving and decision-analysis framework. The suitability model is presented within the context of the methods, assumptions, and best practices of multicriteria evaluation and weighted linear combination modeling. The primary finding of this research is the CommunityViz planning support system suitability model is a valid tool for spatial multicriteria evaluation and demonstrably effective at producing trusted outcomes. The contributions of this paper are an evaluation of the scientific rigor of the CommunityViz suitability model as well as discussion of planning support system–based interactive and iterative model development within a multicriteria evaluation process.*

INTRODUCTION

The core of suitability modeling is the analysis and interpretation of data to produce information useful to decision makers and stakeholders in a decision process (Malczewski 2004). Suitability modeling may consider a number of geographic conditions, including location, development actions, and environmental elements (Collins et al. 2001), as well as legal requirements and social factors reflecting the values and interests of decision makers, individuals, or other stakeholders. While the use of the word suitability often refers specifically to the idea of site selection and development, the analytical concepts are more general (Hopkins 1977) and applications more wide ranging.

Discussing spatial expert systems, Malczewski (1999) notes a number of decision-making obstacles relevant to suitability modeling: spatial decision problems are not well understood; knowledge of spatial processes and decisions includes causal, common sense, and meta-knowledge but differs from person to person; people will approach and solve spatial problems differently; and communication barriers may exist between experts and people who operationalize decision support. Some of these obstacles can be overcome using an information-structuring process such as multicriteria evaluation (MCE). Geographic information systems (GIS)–based spatial decisions support systems (SDSS) (Densham 1991) also are useful to apply to siting problems to bridge the gap between decision makers and complex quantitative analytic models (Maniezzo et al. 1998). With long-standing motivation for research on SDSS stemming from the recognition that some spatial decision problems are characterized by many of the previously mentioned problems, MCE has come to be recognized as an inherent part of SDSS (Jankowski et al. 2008).

Developed as a subset of SDSS, planning support systems (PSS) are a special type of planning information technology consisting of geospatial application software and information frameworks designed to support planning processes (Klosterman 1997, Geertman and Stillwell 2003). PSS extend GIS capabilities in analysis and problem solving, and add design, decision-making and communication capabilities (Nedovic-Budic 2000). Unlike complex land-use or resource modeling software, PSS often take the form of a toolbox from which decision makers can draw for assistance in decision management, modeling, analysis and design, communication, visualization, and information dissemination (Klosterman 1997, Batty 2003).

The purpose of this paper is to assess the quality of a PSS-based suitability model. While the utility of PSS is broadly supported in the literature, implementation of PSS technologies has been slow and often unsuccessful (Geertman 2013, Te Brömmelstroet 2012). Vonk et al. (2006) mentioned a number of bottlenecks to PSS usage, including lack of experience, lack of awareness, and problems or uncertainties with instrument quality. Following Vonk and Geertman (2008), we assess the quality of the CommunityViz® suitability model with: (1) a literature-based overview of MCE, weighted linear combination modeling, the Kepner-Tregoe (K-T) decision-analysis framework, CommunityViz and the CommunityViz suitability model, and uncertainty evaluation in MCE; (2) a stepwise presentation of PSS and K-T methods; and (3) a comparison of outputs between the PSS and K-T decision-making frameworks. Methods and outputs are presented using a case example of an energy facility siting decision situation in the U.S. West.

BACKGROUND

Multicriteria Evaluation

MCE is defined by Voogd (1983) as a flexible framework for appraisal of a set of decision options using a number of criteria. MCE techniques are able to accommodate the political, social,

and values dimensions of a decision process or problem-solving situation. In discussing the theory underpinning MCE, Voogd (1983) argues that classification theory, not decision theory, provides the basis for MCE work. MCE assists with inventory, classification, and arrangement of the information needed to make choices. This added structure can produce a deeper knowledge of the decision situation, which would not have been obvious, given its complex nature. A key caveat, however, is that while MCE provides a structure for solving a problem, it does not provide the solution per se (Voogd 1983).

There are a number of benefits to using MCE. MCE is seen as a transparent and systematic approach that increases objectivity and yields reproducible results (Janssen 2001). As detailed in Kiker et al. (2005), the use of MCE to structure a problem improves on heuristic approaches to reducing complexity in problem solving. MCE processes are a means of getting greater insight into value judgments, incorporating differing views in an analytical framework, providing a tangible means of demonstrating openness in decision making, and reducing information incorporated in decision situations. Incorporating social and political concerns in an evaluation structure can generate circumstances that lead to acceptance, adoption, and implementation of resulting decisions. Integrating preferences with geographic data yields results that are feasible and accurate as well as acceptable to decision makers (Jankowski and Richard 1994) and the public (Lieske et al. 2009). MCE is a means to both justify and account for policy decisions (Voogd 1983). MCE facilitates the documentation of decision processes and enables decision-maker learning (Hajkowicz 2007). MCE may, through evaluation of alternatives, facilitate compromise (Malczewski 1996). Another benefit of MCE is bringing scientific information to situations or people who might not otherwise have it. Most important, MCE processes are a way to arrive ". . . at substantially better considered decisions" (Voogd 1983, p. 33).

There also are potential disadvantages of MCE. MCE may lead to premature or over disclosure of information or intentions; MCE may be seen as too complex and/or technocratic. MCE may be seen as providing a false sense of accuracy, be subject to manipulation (Janssen 2001), and, like any research, MCE may be used as ". . . a 'scientific sauce' over a decision already made" (Voogd 1983, p. 34).

GIS-based MCEs are distinctive because results depend on the patterns of spatial data-based evaluation criteria and how spatial data and preferences are combined (Malczewski 2011). Voogd (1983) defines an evaluation criterion as "a measurable aspect of judgment by which a dimension of the decision options under consideration can be characterized" (p. 55). Evaluation criteria used in GIS-based MCE are based on spatial relationship tests, including simple location factors such as proximity, conditional location factors, overlap, conditional overlap, Boolean tests, complex factors, or numerical or lexical data attributes. With complex factors, evaluation criteria are determined using a separate model (Walker and Daniels 2011). Baban and Flannagan (1998) also mention consideration of criteria that are not site-specific such as impacts on human health and the environment.

Evaluation criteria may be differentiated between benefit criteria and cost criteria and further differentiated between requirements and preferences. With benefit criteria, higher data values are correlated with better performance. With cost criteria, lower data values are correlated with better performance (Nyerges and Jankowski 2010). Requirements are evaluation criteria in a decision situation that are absolute and not subject to preferences or tradeoffs. While MCE most often is focused on preferences, identification of requirements, especially in spatial modeling, is extremely useful for it can speed up processing by a priori elimination of unsuitable decision options.

The MCE literature provides a number of recommendations for establishing a set of evaluation criteria. The set of criteria should cover all aspects of the decision problem. Criteria should be able to be included in an analysis in a meaningful way. Criteria should be comprehensive, measurable, and nonredundant (Malczewski 2000). The definition and measure of criteria should be in accord with their intended use (Voogd 1983) and the overall set of evaluation criteria should be minimized (Malczewski 1999). Per Voogd (1983), limiting the number of criteria minimizes uncertainty.

Voogd (1983) offers several recommendations for addressing uncertainty in MCE: comparison of initial with final evaluation criteria, sensitivity analysis, and comparison of multiple MCE methods. Comparison of the initial list of evaluation criteria with the final list of evaluation criteria allows assessment of whether all pertinent criteria have been considered (Voogd 1983). Comparison of multiple methods helps minimize what Jiang and Eastman (2000) call decision risk, the probability that a decision will be made incorrectly. Sensitivity analysis is an exploratory process that allows one to gain a deeper understanding of a problem structure through evaluation of how changes in inputs (evaluation criteria and weights) affect changes in outputs. The purpose of MCE sensitivity analysis is to facilitate uncertainty evaluation and assess the spatial impact of differing weights (Jankowski et al. 2008). If small changes in evaluation criteria or weights result in no changes in the preferred decision option, one may have more confidence in output rankings (Nyerges and Jankowski 2010). If small changes in inputs result in changes in outputs, it may be necessary to reevaluate the structure of the model. Sensitivity analysis also helps to indicate which criteria have more and which criteria have less influence on model outcomes. Sensitivity analysis can reduce complexity by enabling the identification of criteria that do and do not influence decision-option ranking. Criteria with minimal influence on outcomes may be removed (Nyerges and Jankowski 2010). In general, there are two types of sensitivity analysis, one-at-a-time (OAT) factor analysis and global sensitivity analysis, with OAT being more common and easy to implement (Ligmann-Zielinska and Jankowski 2014). Sensitivity analysis couches MCE outputs by making clear outputs depend on the technique employed, the criteria chosen, criteria scores, and data quality, as well as weights (Voogd 1983). Outputs, therefore, are conditional.

WEIGHTED LINEAR COMBINATION

One of the more widely used MCE methods is weighted linear combination (WLC) modeling. With WLC, evaluation criteria are standardized to a common numeric range, weighted, and combined to create a composite score for each decision option. Weights indicating relative importance are assigned to each evaluation criteria. The larger the weight, the more important a criterion is. For each decision option, a score for each criterion is calculated by multiplying the weight by the standardized value of that criterion. Scores are summed for all criteria to generate an overall suitability score for each decision option. The result is a continuous measure of suitability. Results generally are not compared with a separate benchmark or empirical standard (Hopkins 1977). WLC is one of the most straightforward and often-used GIS-based MCE methods (Malczewski 2011); WLC is easy to implement within GIS, is easy to understand, and is intuitively appealing to decision makers (Nyerges and Jankowski 2010, Malczewski 2004, Voogd 1983). It also has been described as methodologically sound and transparent (Janssen 2001). WLC-based results derived from GIS often are presented visually, using maps where scores are displayed with a graduated color ramp. Importantly, WLC and similar techniques provide reasonable problem solutions (Janssen 2001).

Primary assumptions of WLC modeling are the linearity and independence of evaluation criteria. The linearity assumption means a change in desirability of an attribute is constant for any change in the level of an attribute. For example, the change from zero to one acres of buildable land has the same impact on the model as the change from 999 to 1,000 acres of buildable land. The independence assumption means there are limited to no interaction effects among evaluation criteria. Results may be incorrect if interaction among attributes has not been taken into account (Malczewski 2000) through multiple counting of like or near-identical criteria. The independence assumption is conceptually similar to the assumption of no perfect correlation among independent variables in ordinary least squares regression analysis. With MCE, if there is a high measure of correlation between two criteria, one may be excluded from the set of evaluation criteria (Malczewski 1999). However, correlated criteria may be both incorporated in an analysis if they are likely to receive different weightings (Voogd 1983).

Table 1. Steps in weighted linear combination modeling (modified from Malczewski 1999, p. 199)

1.	State the decision.
2.	Define the set of evaluation criteria and the set of decision options.
3.	Standardize each criterion map layer.
4.	Define the criterion weights.
5.	Construct the weighted standardized map layers.
6.	Generate the overall score for each alternative.

Table 1 lists key steps in WLC modeling. In the first step, it is necessary for decision makers and stakeholders, in the language of Drobne and Lisec (2009), to recognize and agree on the problem to be addressed. WLC step two is actually three tasks, establishing the evaluation criteria, establishing the set of decision options, and calculating raw suitability scores. In a GIS-based suitability model evaluation, criteria typically are spatial layers and decision options are areal units. WLC step three is criterion standardization where raw suitability values are transformed to comparable units (Malczewski 1999). Many criteria, for example distance to infrastructure and slope, use different measurement scales. Raw suitability scores more often than not require transformation to a common scale suitable for direct comparison. There are two scale transformation techniques, linear and nonlinear standardization. Nonlinear standardization is the common approach used in suitability modeling (Walker and Daniels 2011). With nonlinear standardization, criteria are standardized to a consistent range, often zero to one or zero to 100. Nonlinear standardization makes weights more easily understandable and removes potential problems with differences stemming from a lack of knowledge or confusion over units of measure (Hopkins 1977). When raw data values include both negative and positive numbers, nonlinear standardization should be used (Nyerges and Jankowski 2010). Disadvantages of nonlinear standardization include the loss of clear meaning of well-understood measurement scales (Malczewski 1996) and that model outputs do not relate to the raw scores in a linear fashion (Nyerges and Jankowski 2010). While there is some obfuscation associated with the loss of well-understood measurement scales, the issue of model outputs not relating to raw scores in a linear fashion does not ordinarily appear to be a problem. The latter issue especially is more than compensated for by the easier interpretation of evaluation criteria including relaxed requirements for knowledge of the units of the evaluation criteria. It also is noted that scores standardized with a nonlinear transformation will not necessarily be normally distributed. Negatively skewed standardized criteria will impact an analysis as though they are given a high weight, while positively skewed standardized criteria will impact an analysis as though they are given a low weight.

WLC step four is assigning weights. Preferences may be captured in MCE numerically, using ordinal expressions (e.g., low, medium, high), or as Boolean values. In MCE, quantitative values are referred to as weights while ordinal and other expressions of value are referred to as priorities (Voogd 1983). Weights and priorities improve an analysis by enabling a better understanding of tradeoffs among evaluation criteria as well as the consequences of different preferences (Hajkowicz 2007). A common option for incorporating weights in a WLC model is a numeric point scale where respondents indicate a number for each evaluation criterion on a one to X or zero to X scale. Osgood et al. (1957) found a seven-point number scale augmented with semantically differentiated (opposite) labels allowed respondents to adequately express their preferences. Voogd (1983) presents the results of an empirical comparison of several methods of

measuring preferences that indicates a seven-point scale is one of two methods that perform better, take less time, and are less difficult than other methods. WLC steps five and six, constructing weighted standardized map layers and generating scores for each decision option, may be automated with GIS-based weighted overlay technologies, including purpose-built PSS.

KEPNER-TREGOE DECISION ANALYSIS

The Kepner-Tregoe decision model is part of a broader organizational management framework first conceptualized at the RAND Corporation in the 1950s by Drs. Charles Kepner and Benjamin Tregoe. Grounded in the rational theory of organizational behavior (Dawson 1996), the K-T framework was formalized in the 1960s and made widely available through a popular business literature monograph (Kepner and Tregoe 1965). Decision analysis is one of four analytic processes that make up the K-T framework, the others being Problem Analysis, Potential Problem (or Opportunity) Analysis, and Situation Appraisal (Kepner and Tregoe 1997). The framework has been extensively applied in a diversity of business-management applications when issues are complex and when a number of solution options exist (Kepner and Tregoe 1997, Finlow-Bates et al. 2000), in environmental management and remediation (Linkov et al. 2004, Kiker et al. 2005), and physical infrastructure development (Thorpe and Kumar 2002). Watson (1987) points out that much of the success of the framework is because of its approach in structuring individual and organizational thought processes in a highly systematic manner.

Table 2. Kepner-Tregoe decision analysis steps (Source: Kepner and Tregoe 1997, pp. 85-86)

1. State the decision.
2. Develop objectives.
3. Classify objectives into MUSTs and WANTs.
4. Weigh the WANTs.
5. Generate alternatives.
6. Screen alternatives through the MUSTs.
7. Compare alternatives against the WANTs.
8. Identify adverse consequences.
9. Make the best-balanced choice.

Table 2 outlines the nine steps of a traditional K-T decision-analysis process. The first step in the K-T process is identical to the first step in WLC modeling: to recognize and agree on the problem to be addressed. K-T step two involves developing objectives that are identical to evaluation criteria in MCE. In K-T step three, objectives are categorized as requirements ("musts") and operational objectives ("wants"). In step four, wants are ranked and assigned relative weights. In step five, alternatives are generated that in step six are screened against the musts. In step seven, alternatives are compared against the wants by assigning relative scores for each alternative on an objective-by-objective basis and calculating weighted scores for each of the alternatives to identify the top-scoring choices. Step eight involves identifying adverse consequences for each top alternative and evaluating risk probability (high, medium, low) and severity (high, medium, low), before making a final, single choice between top alternatives (step nine).

The K-T framework shares many characteristics with WLC modeling. K-T modeling has predominately been operationalized in business applications using common spreadsheet technology. While K-T does not enable the direct incorporation of spatial data, the framework may represent spatial concerns in the abstract, for example, by considering travel time between locations.

COMMUNITYVIZ®

The CommunityViz suitability model is a spatial MCE framework built on a WLC model. Developed by the Orton Family Foundation (Rutland, Vermont), CommunityViz is a modular system built on the ArcGIS platform (ESRI Inc., Redlands, California). It consists of two integrated extensions to ArcGIS: Scenario 360 and Scenario 3D. The Scenario 360 module of CommunityViz extends the quantitative capabilities of ArcGIS by allowing formula-based spreadsheet-like calculations to be performed on geographic data. Formula-based GIS data attributes allow on-the-fly adjustment of geographic and numeric inputs as well as automated recalculation of maps and quantitative output in a process referred to as "dynamic analysis" (Walker and Daniels 2011, pp. 32-35). Scenario 3D allows for three-dimensional display of the built environment and landscape with real-time movement and object manipulation in a semi photo-realistic setting. CommunityViz is a promising tool for suitability modeling and spatial MCE generally because of the ability of the software to link weights with geographic data and automatically update the model when there are changes in either weights or geographic data inputs.

Sitting within the Scenario 360 module, the CommunityViz suitability model generates two kinds of evaluation criteria scores, raw and standardized. A raw evaluation criterion score is a direct query based on spatial relationships or attribute values. CommunityViz uses a formula-based dynamic attribute to calculate raw scores. Raw scores may be specified as a benefit or cost by indicating whether lower or higher suitability scores result from the calculation of a suitability criterion value.

As shown in Figure 1, evaluation criterion weighting is incorporated in CommunityViz with easily changeable "assumptions" (Walker and Daniels 2011, p. 34) linked to dynamic attributes via a slider-bar interface. Weight sliders provide a graphical display of values as well as an easy means of adjusting weights. On changing a weight or attribute value, the CommunityViz suitability model will recalculate the suitability analysis based on the new input(s) then graphically display the new results in maps and charts. Weighted assumptions in the CommunityViz model often are set up using a numeric point scale. Given the ability to rapidly recalculate a model, a numeric point scale that includes a zero

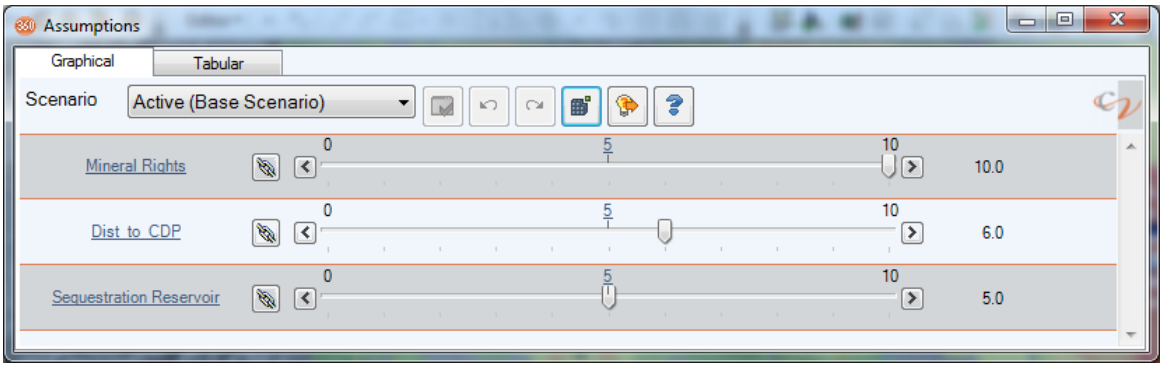


Figure 1. Representative CommunityViz weight sliders on an 11-point scale

value allows one to easily temporarily or permanently remove a criterion from the analysis.

This technology invites interactive experimentation, supports discussion of the relative importance of each criterion, provides an approach for working through the difficulty of conflicting preferences, supports sensitivity analysis, and enables PSS-based suitability analysis to be used as a thinking tool in site selection.

SUITABILITY MODEL CASE EXAMPLE

The High Plains Gasification-Advanced Technology Center (HPG-ATC) was envisioned as a $120 million synthesis gas research and development facility in the state of Wyoming. Goals of the facility were to advance both the technical understanding of the conversion of feedstocks (e.g., coal) by gasification into synthetic gas (or syngas) for use in power generation, subsequent downstream conversion of syngas into liquid fuels and chemicals, and to increase in-state utilization of Wyoming minerals. As a research and development facility, the HPG-ATC was planned to be approximately 1/100th the size of a comparable commercial facility. Major components identified as part of the facility were feedstock storage, rainwater retention, feedstock processing, industrial gas processing, a gasifier, gas flare, byproducts handling, a control center, and electrical, maintenance, and educational facilities.

In February of 2008, the University of Wyoming (UW) entered into a partnership with a U.S.-based energy company to design, construct, and operate the HPG-ATC. The project utilized a Front End Engineering Design (FEED) approach for determining the technical requirements and estimated costs of the facility (Plummer 2007). The FEED process addresses all aspects of facility construction, from process design, equipment and material selection, to plant layout, health, safety and environment (HSE) planning, and civil, mechanical and electrical engineering (Baron 2010). The purpose of the FEED process is to develop the necessary strategic information for developers to address risk and commit resources to maximize the potential for a successful project. A completed FEED process serves as the basis for the start of facility construction (CII 2012). For the HPG-ATC, the development of a project FEED plan involved completing a number of preliminary or pre-FEED steps. These included analysis of facility requirements in tradeoff studies, determination of facility capabilities and configurations, total construction costs estimations, permitting process initiation, and site selection. The site-selection process is the focus here.

SITE-SELECTION PROCESS

The purpose of the site-selection process was to identify the most preferred land parcel or set of contiguous parcels for HPG-ATC construction and operation based on criteria mutually agreed on by UW and the industry partner. This multi-scale internal evaluation process involved three distinct yet overlapping analyses: (1) a PSS-based statewide suitability assessment, (2) an evaluation of site proposals offered by local government and economic development entities through a public request for proposals (RFPs) process, and (3) parallel evaluation of the final six decision options using both PSS at the parcel level and K-T methods. The major activities, workflow, and approximate timeline of the suitability analysis are presented in Figure 2. The overall site-selection process was structured around the steps of the K-T decision analysis process, presented in Table 2.

A generalization of the evaluation criteria used in the HPG-ATC site assessment is presented in the RFP (UW 2008). The site was to be at least 35 acres in size, level ground with minimal vegetation, at or above 4,000-feet elevation. The elevation requirement came from the U.S. Energy Policy Act of 2005. The act specified a national research and development focus on high-elevation integrated gasification combined cycle plants that would be carbon-capture and sequestration-capable, driven in part to tackle technology shortcomings in gasification of high-moisture coals such as those abundant in the state of Wyoming (CRN 2009). Other influences on criteria development were HSE, greenfield status, suitable power, transportation infrastructure, distance to commercial air service, availability of natural gas fuel,

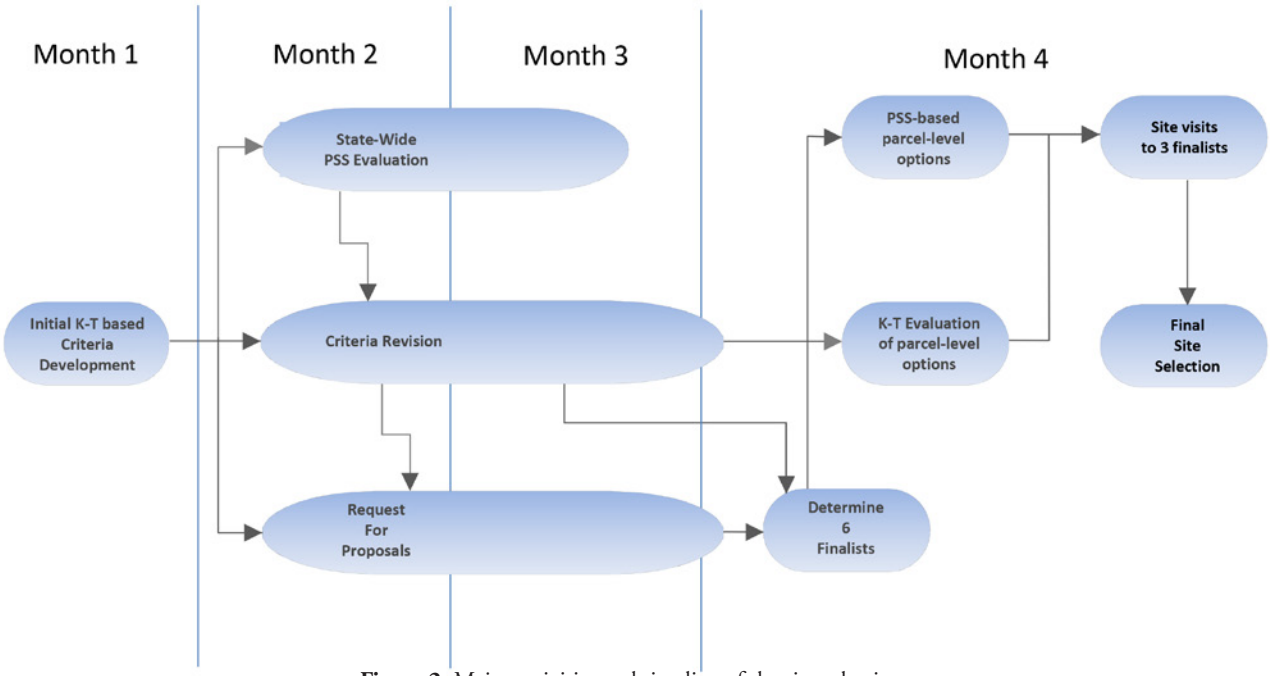


Figure 2. Major activities and timeline of the site-selection process

public utility water and sewer, the quality and locations of wells and aquifers, landfill requirements, and distance to laboratory facilities. Anthropological, archaeological, historical, and cultural resources, as well as compatibility with natural areas, parks, and monuments, were also of concern. Proximity criteria included distance to wetlands, threatened and endangered species, species of critical concern, and wildlife migration corridors. Criteria were generated based on the amenities of nearby communities, including the availability of emergency medical services, groceries, health care, housing, and restaurants. Criteria also were developed based on legal encumbrances, including zoning, air quality, and noise restrictions. Other infrastructure criteria included roads, flood management, and telecommunications availability (UW 2008).

K-T CRITERIA DEVELOPMENT

The evolution of thought surrounding evaluation criteria and weights occurred in a series of meetings of the site-selection team between November of 2008 and February of 2009. The process was similar to that described by Erdoğan (2009) where the knowledge of an interdisciplinary group of experts is modeled and refined over the course of the modeling process.. Originally (month one), 75 evaluation criteria were identified. During month two, the number of evaluation criteria had expanded to 97. At the same time, it was becoming clear that discussions of the statewide suitability model were causing experts to begin to think more spatially. For example, the month-one criteria specified proximity to CO_2 sink. The month-two criteria refined proximity to CO_2 sink as a cost criterion. Wetlands changed from a proximity-based criterion to a Boolean criterion for the team decided distance to wetland was not of concern as long as the facility was outside of the wetland. The month three criteria were annotated with a Boolean value indicating the availability of GIS data. This version of the criteria also indicated thresholds for a number of criteria, for example, distance no greater than 20 miles. Weighting the wants (K-T step four) proceeded from the evolution of K-T criteria over the course of the decision process. During month three, weights were specified as one (low), two (medium), or three (high) importance. By month four, there was a substantial paring down of the number of criteria driven by data availability and the articulated need to consider independence given obvious redundancy in the original 75 criteria.

Assignment of criteria attribute values initially were categorical, based on specific conditions, and were transformed to numerical values. For example, site conditions where the site is level were given a value of nine where they meet specifications, three where they require work, or one where they require substantial construction or improvement. Final scores were calculated by multiplying the criterion attribute values by the weight. As part of the process, different components of the team determined weights separately. UW and the industry partner scorecards differed slightly in what they considered to be low-impact, medium-impact, and high-impact criteria. Weights were reconciled during the middle of month four. K-T criteria were contracted to 47 early in month four then further reduced to 31 by the middle of month four. The final list of criteria included weights and explanations of the attribute value designation for each criterion. Following K-T step seven, alternatives were compared with the wants by calculating weighted scores for each of the alternatives.

STATEWIDE SUITABILITY

As the RFP was being circulated for responses and the set of evaluation criteria were evolving, the CommunityViz suitability model was used to develop a suitability map to guide the selection team on suitable locations for the HPG-ATC across the state of Wyoming. The steps used in this statewide model follow Malczewski's (1999) steps for WLC modeling summarized in Table 1.

The base layer used in the statewide model was a dataset of public land-survey system (PLSS) sections. Standard sections are one square mile in size. The raw data contain nearly 99,000 records. To speed up processing, this layer was made smaller by removing unsuitable data records where (a) elevations are < 4,000 feet, (b) most public lands, and (c) big-game migration corridors. The resulting data layer contained 55,892 records. Removing clearly unsuitable decision options at the beginning of a GIS-based suitability analysis minimized the processing time required for subsequent calculations.

WLC steps five and six, which result in a suitability score for each decision option, are operationalized in CommunityViz with two dynamic attributes, raw suitability score and suitability. Both scores are calculated for each decision option (in this case, areal unit). The raw suitability score is determined by first calculating proportional weights (criterion weight divided by the sum of suitability weights) then multiplying the proportional weight by the standardized score for each evaluation criterion. Using CommunityViz, evaluation criteria were weighted using an 11-point scale where values range from 0 to 10 in increments of 0.1. Raw suitability scores are standardized with the suitability dynamic attribute using nonlinear standardization formulas for benefit criteria (Equation 1) and cost criteria (Equation 2):

$$x'_{ij} = \frac{x_{ij} - x_j^{\min}}{x_j^{\max} - x_j^{\min}} \tag{1}$$

$$x'_{ij} = \frac{x_j^{\max} - x_{ij}}{x_j^{\max} - x_j^{\min}} \tag{2}$$

(Nyerges and Jankowski 2010, Malczewski 1999)

The result is a suitability score assigned to each decision option. There are advantages and disadvantages to the use of a final suitability score to implement nonlinear standardization of raw suitability score results. The primary advantage of standardized suitability scores is being able to directly compare alternative combinations of evaluation criteria and weights on a standardized suitability output scale. A drawback to this standardization is that while key ordinal results do not change, the standardized scores suggest a larger range of variation between the sites than do the raw scores. While the consequences of this transformation are beneficial for the direct comparison of differing evaluation criteria and weights, the consequences of the transformation are more ambiguous for the presentation of suitability results. For maximum clarity, one may present both the raw and standardized scores when evaluating specific choice possibilities.

To summarize, CommunityViz creates a raw score (direct measurement), a standardized score (nonlinear standardization of the raw scores), a raw suitability score, and a suitability score (the raw suitability scores transformed via nonlinear standardization). The majority of the evaluation criteria used in the spatial models were a subset of the criteria employed in the K-T analysis. Criteria were included in the spatial model where spatial data of sufficient quality were available or could be developed within the scope of the project.

PARCEL-LEVEL SUITABILITY

The statewide PSS analysis was used in conjunction with the K-T analysis to develop and assess evaluation criteria, including, as shown in Figure 2, the site-selection criteria put forward in the RFP. The RFP process was the means of generating site-specific decision options. The RFP resulted in 15 responses, each indicating specific parcels for potential construction of the HPG-ATC.

Following K-T step six, decision options that did not meet a must were dropped from consideration. Of the 15 responses to the RFP, six met the musts and subsequently were evaluated as choice possibilities, using both the PSS and K-T frameworks. The PSS-based parcel-level evaluation was a modification of the statewide model that assessed suitability for the final six decision options. The parcel-level analysis offered direct comparison between PSS and K-T outputs. In addition to testing multiple methods, the rationale for developing a parcel-level PSS-based model stemmed in part from the observation that MCE results are not necessarily consistent across spatial scales (Malczewski 2000). Weights incorporated in the parcel-level PSS analysis were based on the designations of criteria as low impact, medium impact, or high impact from K-T step four.

The final evaluation criteria incorporated in the K-T analysis, the PSS-based statewide HPG-ATC suitability model, and parcel-level site selection are presented in Table 3. There were a total of 36 evaluation criteria incorporated in the three analyses. Thirty-one evaluation criteria were used in the K-T analysis. Fifteen evaluation criteria were incorporated in the statewide suitability model and the parcel-level model, 12 of which were present in both models. The site-selection team worked to make criteria as consistent as possible across the three analyses. Differences between the criteria incorporated in the K-T analysis and the two PSS models were because of limitations in the availability and quality of spatial data. Differences in criteria between the statewide suitability model and the parcel-level site selection model were the result of the inappropriateness of some evaluation criterion being included in a meaningful way at multiple spatial scales. For example, spatial data on soils were incorporated within the parcel-level model but not in the statewide model. The high degree of spatial heterogeneity in the soils data within the areal

Table 3. Evaluation criteria used in the K-T, PSS statewide site suitability model, and PSS parcel-level site selection mode

Criterion Name	Criterion Type	K-T	Statewide Site Suitability	Parcel-level Site Selection
Highway	Proximity Benefit	✓	✓	✓
Power	Proximity Benefit	✓	✓	✓
Fuel	Proximity Benefit	✓	✓	✓
Transportation	Proximity Benefit	✓	✓	✓
Landfill	Proximity Benefit	✓	✓	✓
Seismic	Complex	✓	✓	✓
CO2	Proximity Benefit	✓	✓	✓
Air Quality	Proximity Cost	✓	✓	✓
Public Receptor	Proximity Cost	✓	✓	✓
Municipality	Proximity Benefit		✓	✓
Hospital	Proximity Benefit		✓	✓
Fire Department	Proximity Benefit		✓	✓
Rail	Proximity Benefit		✓	
Analytical and Lab Statewide	Proximity Benefit		✓	
Analytical and Lab Local	Proximity Benefit	✓	✓	
Air Quality	Boolean	✓		✓
Soil	Complex	✓		✓
Surface Water	Proximity Cost	✓		✓
Land Use	Attribute	✓		
Greenfield	Attribute	✓		
Water quality	Attribute	✓		
Grey Water	Attribute	✓		
Sanitary	Attribute	✓		
Site Condition	Attribute	✓		
Public Support	Attribute	✓		
Fire Hydrants	Attribute	✓		
Neighboring Site Information	Attribute	✓		
Chemical Supply	Attribute	✓		
Slag	Attribute	✓		
Local Climate	Attribute	✓		
Local Support	Attribute	✓		
Labor	Attribute	✓		
Syngas	Attribute	✓		
Laydown	Boolean	✓		
Food and Shelter	Proximity Benefit	✓		
Emergency Services	Proximity Benefit	✓		

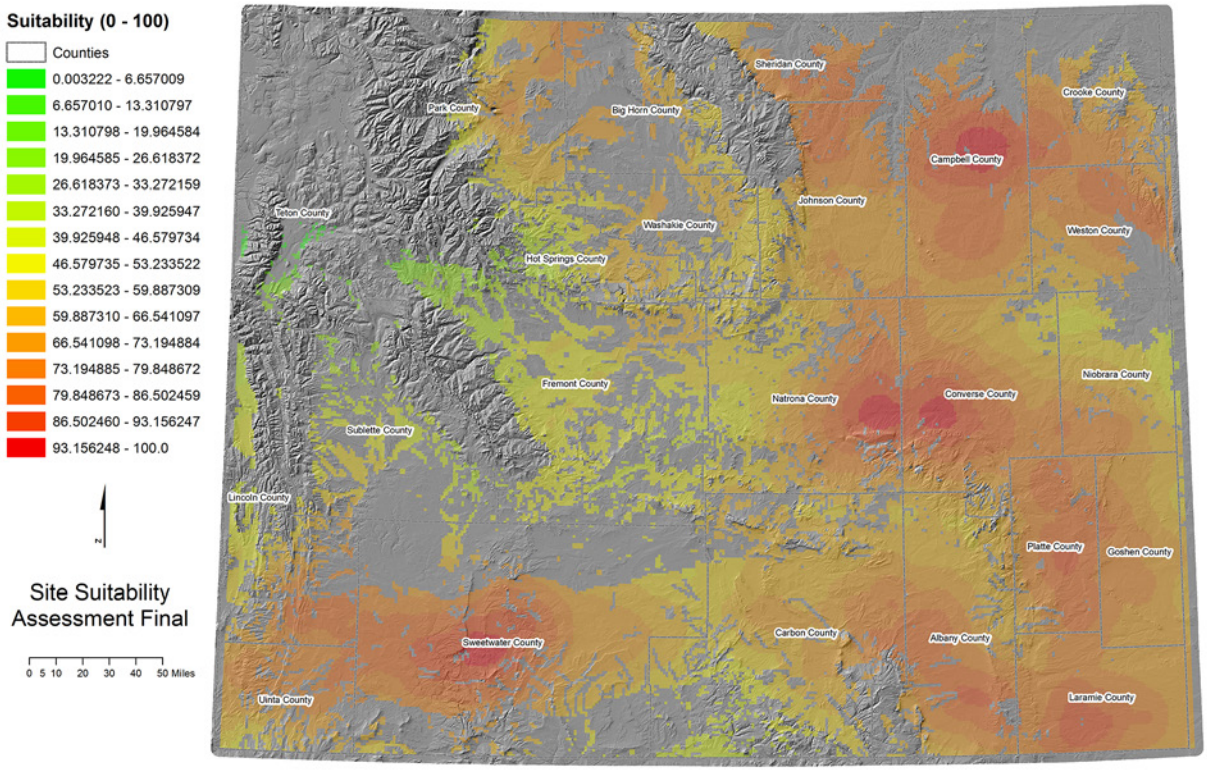


Figure 3. Statewide suitability standardized scores

units of the statewide model (approximately one square mile in size) made consideration of soil characteristics problematic at the statewide scale. Specific weights for the evaluation criteria are not shown because of the confidentiality constraints associated with a nondisclosure agreement governing the facility development partnership.

RESULTS

Results of the statewide suitability model are presented in Figure 3. Areas presented as gray hillshade are outside the study area based on elevation, public lands, and/or the presence of migration corridors as described previously. Suitability results are presented with a green to red color ramp where dark green areas identify the least suitable lands and dark red areas identify the most suitable lands. Figure 3 allowed the site-selection team members to see a clear visual representation of the implications of their collective preferences.

Figure 4 presents the results of the PSS-based site-selection analysis (raw suitability score, upper panel) and the K-T–based results (lower panel). Although the K-T analysis incorporated 31 evaluation criteria and the PSS site-selection analysis only 15, the processes led to similar outcomes. Site E located in Laramie County and site C located in Campbell County were the top two sites in both the PSS parcel-level model and the K-T analysis. Sites A, B, and F (all located in Albany County) and site D (located in Goshen County) were ranked differently by the K-T and PSS analyses. Top-rated appraisal scores that are reasonably close to one another (e.g., less than 15 percent difference) should invite additional scrutiny such as verifying evaluation criteria have been assessed properly and that no relevant evaluation criteria have been excluded.

In this case, multiple methods were demonstrated to yield similar site-selection outcomes. The number of final evaluation criteria in all the analyses (presented in Table 3) was considerably less than the number of initial set of 75 criteria considered in the K-T analysis. The difference between initial and final criteria was because of the elimination of redundant criteria and criteria eliminated because of poor quality or unavailable data. The final set of evaluation criteria were viewed as both comprehensive and nonredundant by the site-selection team.

One-at-a-time (OAT) sensitivity analysis was performed during site-selection team interactive discussions by reducing the weight of individual criterion to zero and observing the effect on suitability outputs. The OAT approach to sensitivity analysis is easily implemented using CommunityViz because of on-the-fly input adjustment and automated recalculation of maps and quantitative output. By applying sensitivity analysis to the statewide model, each of the 15 evaluation criteria incorporated in the PSS model may be mapped, analyzed, and evaluated separately. One is able to see the contribution of the individual components to

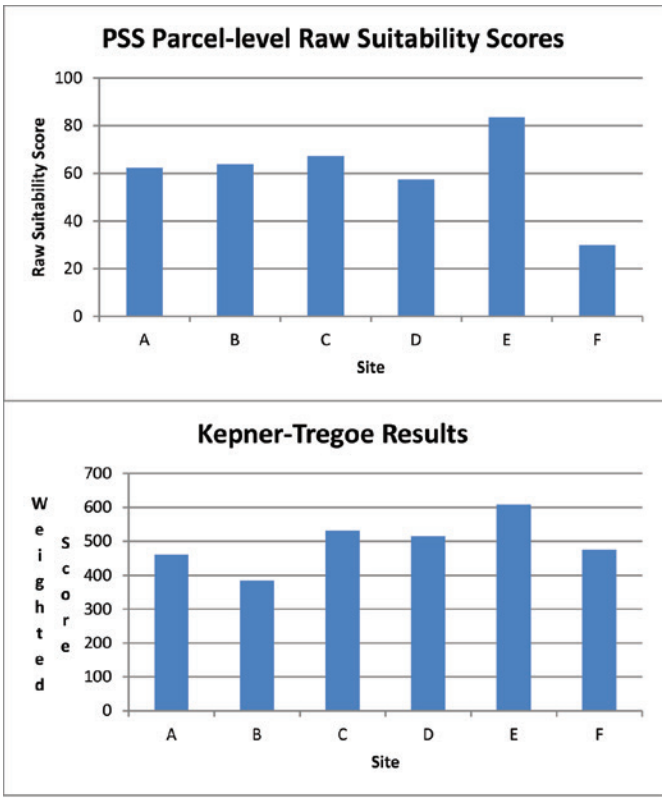


Figure 4. PSS-generated raw suitability and Kepner-Tregoe scores

the overall analysis, inspect the components and formulas created by CommunityViz, and make changes if needed to improve accuracy or performance. Sensitivity analysis allowed the team members to investigate the drivers of their suitability assessment, primarily negatively skewed attributes of spatial data that tended to overwhelm weights in determining suitability model outputs. Finding similar results from different methods, well-developed understanding of the evaluation criteria, consideration of multiple input alternatives, and rapid assessment of the resulting impacts on the outputs of these alternatives helped the team become very confident in the process and the modeling efforts helped inform the best possible choices.

DISCUSSION AND CONCLUSION

This assessment of the CommunityViz suitability model covered methodological foundations, a stepwise walk-through of methods, and a comparative analysis augmented by consideration of uncertainty and assessment of best practices. The primary findings of this research are that the CommunityViz suitability model closely follows the methods of multicriteria evaluation and weighted linear combination modeling, is a beneficial thinking and spatial decision support tool for facility site selection, and, therefore, is more broadly a valid tool for spatial multicriteria evaluation. Congruence with MCE and WLC methods serves to validate the CommunityViz suitability modeling framework. Comparison of PSS results with K-T results, especially as the models were built with differing criteria (see Table 3) both served to validate PSS outputs and assisted in reducing decision risk.

With the HPG-ATC facility siting, the CommunityViz suitability model was demonstrably effective at producing trusted outcomes. As the CommunityViz suitability model and the K-T decision analysis framework both lack a built-in quantitative assessment of uncertainty, the site-selection team followed Voogd's (1983) recommendations for addressing uncertainty in MCE: comparison of initial with final evaluation criteria, sensitivity analysis, and comparison of multiple MCE methods. The addition of the K-T framework to the CommunityViz analysis addressed method uncertainty through separate verification of outputs. The collaborative internal decision nature of the HPG-ATC site-selection process assisted with mitigating problems associated with the interdependence of evaluation criteria and developing weights that accurately reflected requirements and preferences. The evolution of evaluation criteria as part of an interactive and iterative model development process over several months, coupled with the sensitivity analysis of the final model enabled by the dynamic analysis capabilities of CommunityViz, resulted in a transparent process and built confidence among the team members. These observations are congruent with Kleinmuntz (2007), who notes that considering the effects of uncertainty helps build confidence in a model. This occurs, in part, because outputs may be viewed more broadly than a single modeling process resulting in a specific result, but as a framework where varied inputs may consistently produce similar results. Sensitivity analysis and the exploration of alternative inputs deemphasizes the outputs of any specific combination of inputs but bolsters the decision process when there is consistency of outputs.

Determining the set of evaluation criteria serves as a basis for MCE best practices, including quality documentation, easy repetition, objectivity, and transparency (Janssen 2001). Additional best practices suggested by this research include removing clearly unsuitable decision options at the beginning of an analysis, which minimizes the processing time required for subsequent calculations. Enabling faster processing can be important when challenging the processing capability of a computer in an analysis with a larger number of decision options and when waiting on the results of a dynamic update in a meeting setting.

This assessment shows the CommunityViz suitability model meets the requirements for planning methods proposed by Voogd (1983), including increasing insight to a decision situation, the ability to quickly handle changing inputs, transparency, and making values incorporated in a decision process explicit. A potential drawback to the use of PSS for MCE is that it is a tool rather than a problem-driven approach. Voogd (1983) recommends drivers of an MCE process should be the characteristics of the problem and not the characteristics of the problem-solving technique. This is an inherent challenge for planning support systems. Consideration of the flexibility of PSS for adapting MCE modeling to specific decision situations may be part of the answer but this remains an area for future research. It also is noted that while the PSS modeling process was transparent to the site-selection team, nonexpert and/

or third-party audiences would likely require considerable effort to ensure the workings of the model were clearly understood.

Final selection of the HPG-ATC site, site E from Table 4, was based on criteria guided by the PSS and K-T frameworks but ultimately went beyond these methods. After using the PSS and K-T process to identify the top three ranked proposals, visits were made by the team to each potential site. Final selection occurred after these visits based in part on information acquired during the visits and not exclusively based on criteria included in the models. The final decision on site selection was outside the bounds of MCE. At the same time, the final decision incorporated options and choices from the common and understood MCE framework. The process corroborates Voogd's (1983) argument that MCE is a tool for classifying the information needed for choice and providing a structure for solving a problem rather than a decision-making tool that provides a "correct" solution.

Acknowledgments

The authors thank Robert "Bob" Ballard, HPG-ATC project manager at the University of Wyoming, for introducing them to the Kepner-Tregoe decision-analysis framework. We also thank the two anonymous reviewers whose comments and suggestions helped improve and clarify this manuscript.

About the Authors

Dr. Scott N. Lieske is a research fellow at the Sustainability Research Centre within the Faculty of Arts and Business at the University of the Sunshine Coast, Australia. His research interests center around the use of geographic information science and spatial tools to support place-based decision making.

Corresponding Address:
Sustainability Research Centre - ML28
University of the Sunshine Coast
Locked Bag 4, Maroochydore DC
Queensland, Australia 4558
Telephone: 61 7 5456 5812
Fax: 61 7 5430 2887
E-mail: slieske@usc.edu.au

Dr. Jeffrey D. Hamerlinck is the Director of the Wyoming Geographic Information Science Center and Senior Research Scientist in the Department of Geography at the University of Wyoming, Laramie. His research interests are in the area of geographic information science with an emphasis on spatial data infrastructures, decision support, and geographic visualization, as well as planning and resource management, including planning support systems, land use, landscape ecological planning, and regional planning.

Wyoming Geographic Information Science Center
Department 4008
1000 E. University Avenue
University of Wyoming
Laramie, Wyoming 82071
Telephone: (307) 766-2736
Fax: (307) 766-2744
E-mail: jeff.hamerlinck@uwyo.edu

References

Baban, S. M. J., and J. Flannagan. 1998. Developing and implementing GIS-assisted constraints criteria for planning landfill sites in the UK. Planning Practice and Research 13: 139–151.

Baron, H. 2010. The oil and gas engineering guide. Paris: Editions Technip.

Batty, M. 2003. Planning support systems: Technologies that are driving planning. In Geertman, S., and J. C. H. Stillwell, Eds. Planning support systems in practice, advances in spatial science. Berlin; New York: Springer, pp. v-viii.

Collins, M. G., F. R. Steiner, and M. J. Rushman. 2001. Land-use suitability analysis in the United States: Historical development and promising technological achievements. Environmental Management 28: 611-21.

Construction Industry Institute (CII) Implementation Strategy Committee. 2012. IR166-3—CII Best practices guide: Improving project performance, Version 4.0.

Cooperative Research Network (CRN). 2009. Strategic issues for integrated gasification combined cycle. Arlington, VA: National Rural Electric Cooperative Association.

Dawson, S. 1996. Analysing Organisations. Basingstoke: Macmillan Business.

Densham, P. J. 1991. Spatial decision support systems. In Maguire, D. J., M. F. Goodchild, and D. Rhind, Eds. Geographical information systems: Principles and applications. Wiley, Harlow, Essex, England ; New York; Longman Scientific and Technical, pp. 403-12.

Drobne, S., and A. Lisec. 2009. Multi-attribute decision analysis in GIS: Weighted linear combination and ordered weighted averaging. Informatica 33: 459-74.

Erdoğan, A. 2009. Modelling of expert knowledge in geographic information systems–based planning of the Tuz Lake special environmental protection area, Turkey. Planning Practice and Research 24: 435-54.

Finlow-Bates, T., B. Visser, and C. Finlow-Bates. 2000. An integrated approach to problem solving: Linking K-T, TQM and RCA to TPM. The TQM Magazine 12: 284-89.

Geertman, S. 2013. Planning support: From systems to science. Proceedings of the ICE. Urban Design and Planning 166: 50-59.

Geertman, S., and J. C. H. Stillwell. 2003. Planning support systems: An introduction. In Geertman, S., and J. C. H. Stillwell, Eds. Planning support systems in practice, advances in spatial science. Berlin ; New York: Springer, pp. 3-21.

Hajkowicz, S. 2007. A comparison of multiple criteria analysis and unaided approaches to environmental decision making. Environmental Science and Policy 10: 177-84.

Hopkins, L. D. 1977. Methods for generating land suitability maps: A comparative evaluation. Journal of the American Institute of Planners 43: 386-400.

Jankowski, P., and L. Richard. 1994. Integration of GIS-based suitability analysis and multicriteria evaluation in a spatial decision support system for route selection. Environment and Planning B: Planning and Design 21: 323-40.

Jankowski, P., A.-L. Zielinska, and M. Swobodzinski. 2008. Choice modeler: A Web-based spatial multiple criteria evaluation tool. Transactions in GIS 12: 541-61.

Janssen. R. 2001. On the use of multi-criteria analysis in environmental impact assessment in The Netherlands. Journal of Multi-Criteria Decision Analysis 10: 101-9.

Jiang, H., and R. Eastman. 2000. Application of fuzzy measures in multi-criteria evaluation in GIS. International Journal of Geographical Information Science 14: 173-84.

Kepner, C. H., and B. B. Tregoe. 1965. The rational manager: A systematic approach to problem solving and decision making. McGraw-Hill.

Kepner, C. H., and B. B. Tregoe. 1997. The new rational manager: An updated edition for a new world. Princeton, NJ: Princeton Research Press.

Kiker, G. A., T. S. Bridges, A. Varghese, T. P. Seager, and I. Linkov. 2005. Application of multicriteria decision analysis in environmental decision making. Integrated Environmental Assessment and Management 1: 95-108.

Kleinmuntz, D. N. 2007. Resource allocation decisions. In Edwards, W., R. F. J. Miles, and D. von Winterfeldt, Eds. Advances in decision analysis. Cambridge: Cambridge University Press, pp. 400-18.

Klosterman, R. E. 1997. Planning support systems: A new perspective on computer-aided planning. Journal of Planning Education and Research 17: 45-54.

Lieske, S. N., S. Mullen, and J. D. Hamerlinck. 2009. Enhancing comprehensive planning with public engagement and planning support integration. In Geertman, S. and J. Stillwell, Eds. Planning support systems best practice and new methods. Dordrecht: Springer Netherlands, pp. 295-315.

Ligmann-Zielinska, A., and P. Jankowski. 2014. Spatially explicit integrated uncertainty and sensitivity analysis of criteria weights in multicriteria land suitability evaluation. Environmental Modelling and Software 57: 235-47.

Linkov, I., A. Varghese, S. Jamil, T. P. Seager, G. A. Kiker, and T. Bridges. 2004. Multi-criteria decision analysis: A framework for structuring remedial decisions at contaminated sites. In Linkov, I., and A. B. Ramadan, Eds. Comparative risk assessment and environmental decision making. The Netherlands: Springer, pp. 15-54.

Malczewski, J. 1996. A GIS-based approach to multiple criteria group decision-making. International Journal of Geographical Information Systems 10: 955-71.

Malczewski, J. 1999. GIS and multicriteria decision analysis. Wiley.

Malczewski, J. 2000. On the use of weighted linear combination method in GIS: Common and best practice approaches. Transactions in GIS 4: 5-22.

Malczewski, J. 2004. GIS-based land-use suitability analysis: A critical overview. Progress in Planning 62: 3-65.

Malczewski, J. 2011. Local weighted linear combination. Transactions in GIS 15: 439-55.

Maniezzo, V., I. Mendes, and M. Paruccini. 1998. Decision support for siting problems. Decision Support Systems 23: 273-84.

Nedovic-Budic, Z. 2000. Geographic information science implications for urban and regional planning. Journal of the Urban and Regional Information Systems Association 12: 81-93.

Nyerges, T. L., and P. Jankowski. 2010. Regional and urban GIS: A decision support approach. New York: Guilford Press..

Osgood, C. E., G. J. Suci, and P. H. Tannenbaum. 1957. The measurement of meaning. Urbana: University of Illinois Press.

Plummer, F. B. 2007. Project engineering: The essential toolbox for young engineers. Amsterdam; Boston: Butterworth-Heinemann/Elsevier.

Te Brömmelstroet, M. 2012. Performance of planning support systems: What is it, and how do we report on it? Computers, Environment and Urban Systems.

Thorpe, D. S., and A. Kumar. 2002. A life cycle model for asset investment decision making. In Wang, K. C. P., S. Madanat, and G. Spring, Eds. Proceedings of the Seventh International Conference on Applications of Advanced Technology in Transportation. Presented at the Seventh International Conference on Applications of Advanced Technology in Transportation, American Society of Civil Engineers, Boston, MA, pp. 576-83.

University of Wyoming. 2008. Request for proposals (RFPs) for site selection of high plains gasification—Advanced Technology Center.

Vonk, G., and S. Geertman. 2008. Improving the adoption and use of planning support systems in practice. Applied Spatial Analysis and Policy 1: 153-73.

Vonk, G., S. Geertman, and P. Schot. 2006. Usage of planning support systems: Combining three approaches. In Leeuwen, J. van, and H. J. P. Timmermans, Eds. Innovations in design and decision support systems in architecture and urban planning. Dordrecht: Springer, pp. 263-74.

Voogd, H. 1983. Multicriteria evaluation for urban and regional planning. London: Pion.

Walker, D., and T. L. Daniels. 2011. The planners guide to CommunityViz: The essential tool for a new generation of planning. Chicago: Planners Press, American Planning Association.

Watson, S. R. 1987. Decision synthesis: The principles and practice of decision analysis. Cambridge (Cambridgeshire) ; New York: Cambridge University Press.

Implementing a Utility Geographic Information System for Water, Sewer, and Electric: Case Study of City of Calhoun, Georgia

Davie Crawford and Ming-Chih Hung

Abstract: *This paper describes the design and implementation of a geographic information system (GIS) for the Water, Sewer, and Electric Departments for the city of Calhoun, Georgia. The objective of this paper is to explain how the design and implementation of a GIS for the city of Calhoun was established to efficiently manage its utility distribution systems and replace the existing computer-aided design (CAD) system. It also provides other small municipalities with an understanding of what it takes to design and implement a utility GIS. The design and implementation were divided into a set of phases that were carried out to ensure a successful completed system. The methodology used in the development of the GIS has been acquired through reviewing and evaluating other similar systems that involve utility data. The utility departments have relied on inaccurate CAD data for years. The departments all agreed that a more accurate and up-to-date system would help manage their assets. The conclusion of this paper demonstrates the improved efficiency after implementing the GIS compared with the previous CAD system.*

INTRODUCTION

Most utilities throughout the United States and abroad are planning or implementing an automated mapping-facilities management geographic information system (GIS) according to Cannistra (1999). Over the years, many organizations have come to realize that GIS not only helps manage the existing utility infrastructure, but also can help in the design for future expansion (Shamsi 2002, Croswell 1991). The utility industry is a major consumer of GIS because almost all utilities can be spatially referenced. For example, more than 80 percent of all the information that is within water and wastewater utilities is geographically referenced (Shamsi 2005). Utility organizations not only use GIS for the spatially referenced data but also for any information that could be used to carry out further analysis if needed. A GIS allows utility operators and managers not only to determine where their assets are located but to analyze attributes about those assets (Hughes 2006). The majority of the utility organizations reside in municipal governments. Traditionally, utility organizations managed their systems by paper maps.

Many municipal governments provide their citizens with public utilities whether electric, water, sewer, telecommunications, or gas. The size of the utility system depends on the size of the area and the population it serves. Typically, the local governments over the years have managed these utility systems using hard-copy paper maps. The hard-copy paper maps usually were produced by using a computer-aided design system (CAD). The CAD system has helped the utility organizations throughout the years with managing their assets, but it lacks the ability to provide the organizations with database technology. The database technology that is incorporated into GIS has greatly extended the ability to effectively manage the utility assets. Many nonprofit organizations and municipalities are drawn to employing a GIS because it has the ability to combine large amounts of data from different sources and on different media, order them into layers or themes, and analyze or display various relationships (Sieber 2000).

GIS is able to provide the utility organizations with endless amounts of information about their assets, whether spatial or nonspatial (Environmental Systems Research Institute, ESRI, 2003). Utility organizations spend a large amount of money and time on maintaining their infrastructures. By using GIS, these organizations are able to greatly reduce the amount of time and money involved on maintenance. Many of the organizations incorporate their work-order and billing systems into the GIS, which saves even more time and resources. The organizations are able to use one system to effectively manage all their utilities. Whether as a means of data dissemination or acquiring new data, data sharing has become an essential element of local government GIS processes (Tulloch and Harvey 2007).

The city of Calhoun, Georgia, has always utilized a CAD system to manage its utilities. The utility departments realized that the data in the CAD system was not accurate. The city started researching ways to improve the data and efficiency within the departments and wanted a centralized system that could be accessed across all departments in the city. While researching, the city found GIS and decided that it was the type of system it wanted to implement.

OBJECTIVE

The GIS implementation process for a municipal government's utility system can be very complex, expensive, and time consum-

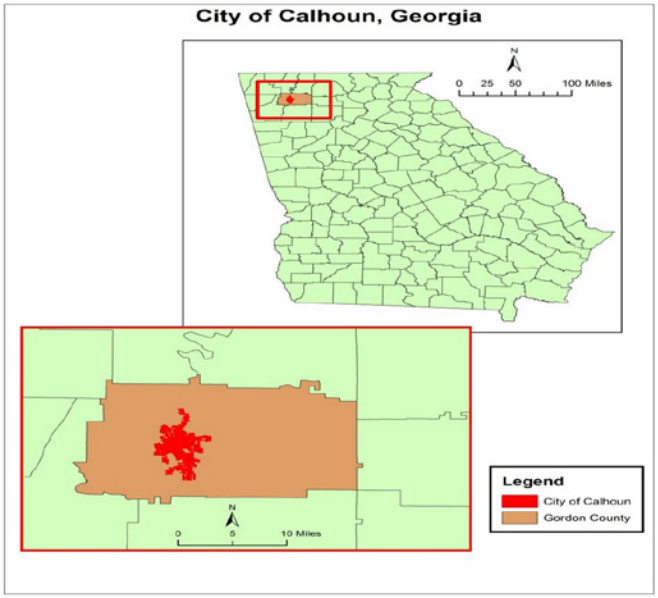


Figure 1. Location of the city of Calhoun, Georgia

ing, depending on what the organization is prepared to manage with the system (Uhrick and Feinberg 1997). The research objective is to review, explain, and provide an example of the implementation process for the water, sewer, and electric utilities within the city of Calhoun, Georgia. These utility areas are very common among a large percentage of the local governments in the United States. The implementation process involves determining the needs of each department and constructing an implementation plan to help track and determine the outcome of the overall system (Tomlinson 2003). The research also reviews the database development process for each of the utility departments.

STUDY AREA

The city of Calhoun, Georgia—located about 60 miles northwest of Atlanta—is in Gordon County, which covers approximately 356 square miles, with 2.5 square miles consisting of water. Approximately 53,000 people live in Gordon County, with roughly 14,000 residing within the city limits of Calhoun (Gordon County 2008). Figure 1 shows the location of Gordon County and the city of Calhoun. Major carpet and flooring industries account for the majority of the workforce in the city.

Established in 1852, the city of Calhoun is governed by a mayor and four council members. The general administration of Calhoun includes the functions of the mayor and the council, city administrator, finance advisor, tax administrator, human-resources personnel, and risk-management personnel. The public works consist of Highway and Streets, Recycling Center, Animal Control, and Cemetery Departments. The public safety and development area includes the Police, Fire, and Community Development Departments. The city of Calhoun utilities consist of the Water, Sewer, Electric, Telecommunications, Engineering, and GIS Departments.

The Water, Sewer, and Electric Departments were the areas of concentration for this research. The Water Department serves more than 20,000 customers in Calhoun and Gordon County. There are a total of about 800 miles of water mains throughout the county. The purpose of the Water Department is to provide clean, pure drinking water to customers and to protect health; to maintain the water distribution system, adding new lines and connections; to add new customers; and to provide proper pressure and clean water at all times.

The Sewer Department serves approximately 6,500 customers. The sewer infrastructure is made up of approximately 3,000 manholes, 9 lift stations, and 150 miles of sewer line. Some of the maintenance involved includes the department responding to approximately 300 utility locate requests each month. The closed-circuit television or CCTV crew performs inspections on approximately 2,500 linear feet of main-line sewer each month. The jet/vacuum crew cleans approximately 6,000 linear feet of main-line sewer each month.

The Electric Department serves approximately 5,000 customers throughout the city. The Electric Department handles approximately 1,000 service calls annually. The department estimates that there are approximately 5,700 poles in the electric system. The electrical system consists of 88 miles of primary overhead wire, 30 miles of primary underground wire, 90 miles of secondary overhead wire, and 25 miles of secondary underground wire. The department also maintains a number of the streetlights and security lights. There are approximately 1,000 streetlights and 200 security lights. The goal of the Electric Department is to employ properly trained personnel and to secure a safe environment for those employees and the community. This will ensure that the distribution system service is maintained at the highest level of quality and reliability. The Electric Department is committed to customer satisfaction and a state-of-the-art approach to power supply.

The Water, Sewer, and Electric GIS was developed and implemented by following three phases, adapted from suggestions by Tomlinson (2003) and Harmon and Anderson (2003). The first phase of the project was to conduct a needs assessment for each of the departments. The second phase was geodatabase design. The third phase was data development and conversion. The phases used for the project were performed once the hardware and software were installed and configured. The city of Calhoun already had a server available for the GIS implementation. A Microsoft SQL Server Enterprise license was purchased and installed on the server. The ESRI software that was installed for the implementation included ArcGIS Server Advance and a number of ArcINFO seats in each department. ArcSDE, which is bundled with ArcGIS Server Advance, also was installed and configured. The advance version of ArcGIS Server provided the ability to develop a Web-mapping application that each department can easily access and use.

NEEDS ASSESSMENT OVERVIEW

The list that follows identifies the steps that were taken to complete the needs assessment for each of the departments:

Interviewed the staff within each department to identify the current document workflows and conditions.

- Established an overall set of goals and objectives for each of the departments.
- Gained an understanding of their business processes and identified redundancy within these processes.
- Evaluated the existing data and data formats currently being used.
- Identified applications that the departments can benefit from.
- Determined the number of users who will be accessing the system from the departments.

WATER NEEDS ASSESSMENT

The needs assessment meeting with the Water Department identified a number of functions and processes that will be achieved with the implementation of the GIS. The following list outlines the major functions and processes identified.

- Preparing maps for work orders
- Integration of billing data
- Future planning for system expansion
- Asset inventory
- Hydrant locations
- Water main isolation
- New meter locations
- Meter-replacement scheduling
- Linking meters and valves to the parcels they serve
- Water distribution and usage analyses
- Pressure zone mapping
- Water main break reporting application
- Crew routing
- Maintenance tracking and inventory
- Web-mapping application for easy access to data

The statistics in Table 1 show an average of the Water Department's operations during any given month. Table 2 lists the estimated number of features within the water system during the time of the assessment. The statistics in Table 2 provide an overall evaluation of the amount of data that was involved in the implementation.

Table 1. Average monthly statistics for water department

Item	Number
Customers Served	22,000
Number of Service Calls	40
Number of Leaks	50
New Meters Added	75
Meter Repairs	35
Number of Utility Locate Requests Requese= RRequests	500
Average Number of New Pipe in Feet Added	5000

Table 2. Feature statistics for water department

Item	Number
Water Lines in Miles	800
Number of Meters	22,000
Number of Hydrants	1600
Number of Valves	3700
Number of Pump Stations	9
Number of Tanks	17

The assessment also identified the number and type of users of the GIS. The department stated that one person would be responsible for creating and updating the data while at least five or more people would be using the system at any given time. The format and use of the existing data also was discussed. The data is in CAD format and is not based on a coordinate system and is not to any kind of scale. The attribute information for the CAD data is in the form of labels in the drawings. The CAD drawings are old and outdated. The locations of the water mains were sketched in from employees' memory over the years. The valve and hydrant drawings were developed using field sketches. The department stated that the drawings are inaccurate and have been pieced together over the years and they would like access to updated data for the electric and sewer systems throughout the city and county. Having access to landowner and parcel information also would be helpful when planning new construction. The water data needs to be updated on a daily basis. The overall goal of the Water Department is to have easy access to accurate and updated information on its entire water system.

SEWER NEEDS ASSESSMENT

The assessment for the Sewer Department identified a number of functions that could benefit from the implementation of a GIS. The list that follows shows some of the examples discussed during the needs assessment. Table 3 shows some of the more important monthly statistics on the functions of the Sewer Department. The numbers in Table 4 give an idea of the number of features that make up the sewer system.

- Access to accurate system mapping
- Integration of billing data
- Work order mapping
- System modeling
- Crew scheduling
- Inflow and infiltration planning
- Planning future expansion of system
- Maintenance tracking and inventory
- Sewer distribution analysis
- Inspection reporting tools
- Web-mapping application for easy access to data

Table 3. Average monthly statistics for sewer department

Item	Number
Customers Served	6500
Avg. # of Service Calls	27
Number of Utility Locate Requests	300
Amount of Pipe Camera Inspected in Feet	2500
Amount of Pipe Cleaned in Feet	6000

Table 4. Feature statistics for sewer department

Item	Number
Manholes	3000
Sewer Lines in Miles	150
Lift Stations	9
Monitor Locations	17
Wet Wells	9
Cleanouts	500

The sewer data is similar to the water data except that roughly 35 percent of the manholes in the system have been located using a real-time kinematic global positional system (RTK GPS) by the Engineering Department. The attributes of the manholes and pipes are in the form of labels in the CAD system. A number of subdivision drawings have been submitted by developers in CAD format and placed in the overall CAD system drawing. All lift stations and wet wells have been located using GPS. Information on the remaining 65 percent of the manholes and pipes is not accurate. The department uses AutoCad for all its mapping and information needs. The Sewer Department would like access to updated data for the Electric and Water Departments for planning new construction. They also would like access to parcel information for easement purposes. The department would like to have the data updated daily as projects are completed in the field. The data will be maintained and edited by one person and accessed by a number of people within the city. The overall goal of the Sewer Department is to have accurate and up-to-date data on its system and easy and quick access to the data.

ELECTRIC NEEDS ASSESSMENT

The needs assessment for the Electric Department began with identifying processes within the department that could benefit from a GIS. The following list shows a number of the processes identified. The statistics in Table 5 show the average monthly number of items that occur on a normal basis for the department. Table 6 lists some of the more important features in the electric system data to give an idea of the size of the system.

- Work order mapping Integration of billing data Inventory
- New electric utility expansionImproved location of utilities in the field Locate trouble calls and outagesProvide better customer service Crew routing
- Electrical distribution analysisLinking meters to the parcels they serve Maintenance tracking
- Web-mapping application for easy access to data

The electric data is in CAD format with no coordinate system or scale. The attributes for the electric features are in the form of labels on the CAD drawings. The CAD data is updated on a daily basis by making the changes to the system drawings. Work crews are given a work order with a printed-out CAD map explaining what work needs to be done. Once the work is completed, it is noted on the CAD drawing and work order in the field. The work order and drawing then are given back to the CAD operator to make changes in the CAD drawings. The Electric Department would like to have access to parcel, water, and sewer data to help support the planning of new construction. The electric GIS data will need to be updated on a daily basis. The department will use one data maintainer and editor to perform the daily updates. The main goal for the Electric Department is to have accurate and updated data on the electric system with easy access to it.

Table 5. Average monthly statistics for electric department

Item	Number
Customers Served	5000
Avg. # of Service Calls	80
Number of Outages	35
Number of Locates	100
Street Light Replacements	20

Table 6. Feature statistics for electric department

Item	Number
Poles	5700
Transformers	1700
Meters	5030
Miles of Primary Wire	120
Miles of Secondary Wire	120

GEODATABASE DESIGN

The second phase of this project, geodatabase design, began with reviewing each department's existing CAD data structures. Databases will be implemented in relational database as it is used in most GIS (Zeiler 1999, Harmon and Anderson 2003, Arctur and Zeiler 2004). The feature classes were determined by meeting with the departments on a number of occasions to decide which features should be included in the system. The feature dataset figures and attribute field tables listed below are for each department's dataset. A number of the attribute fields noted in the design were actually labels in the CAD data for the features. Some of the attribute fields were added items that the departments would like to start keeping track of in the future.

The planimetric data (e.g., roads, buildings, parcels, and

bridges) that each department uses was available from the Gordon County GIS Department. Aerial photography from 2005 also was already in place for Gordon County and the city of Calhoun. The spatial reference of each dataset is based on the existing planimetric and aerial photography data. The coordinate system used for the geodatabase is the Georgia West State Plane and the projection used is the Transverse Mercator. All the datasets will be based on the State Plane Coordinate System.

The water geodatabase design was developed based on the information gathered during the needs assessment and the existing CAD data. The attributes for each of the feature classes in the water geodatabase were determined by examining the existing CAD layer labels. Attributes were added to the features in the areas where the department wanted more information than what was available in the CAD drawings. Figure 2 shows the entity relationship (ER) diagram. According to this ER diagram, tables are created for the following entities (Longley et al. 2005, Bolstad 2008): tanks, pump stations, water mains, valves, abandoned water lines, laterals, leaks, fittings, backflow points, backflow fire taps, hydrants, and water meters.

The development of the sewer geodatabase also was developed based on the existing CAD data and information that was obtained during the needs assessment. Figure 3 shows the entity relationship diagram for the sewer dataset. According to this ER diagram, tables are created for the following entities: service laterals, clean outs, gravity mains, lift stations, manholes, abandoned lines, wet wells, abandoned points, monitor locations, and service mains.

The electric geodatabase design also was created by reviewing the existing CAD data and determining what feature classes and attributes needed to be added. Figure 4 shows the entity relationship diagram for the electric dataset. According to this ER diagram, tables are created for the following entities: support structures, dynamic protective devices, PF correcting equipments, switches, primary OH electric lines, anchor guys, primary UDG electric lines, open points, bus bars, span guys, fuses, surface structures, transformers, secondary OH electric lines, secondary UDG electric lines, streetlights, and electric meters.

DATA DEVELOPMENT AND CONVERSION

The third phase of this project is data development and conversion. The city of Calhoun and Gordon County have existing aerial photography that is tied to the Georgia State Plane Coordinate System. The aerial photography overlaid with Gordon County's planimetric data serves as the base-map foundation for the GIS. The Gordon County GIS Department already has completed the majority of the base-map dataset work with the construction of the parcel data and attributes associated with the data.

The city of Calhoun's Engineering Department already has completed a number of feature locations throughout the city. The department has a GPS base station installed for locating utilities with RTK GPS. By using the RTK GPS, the features located are

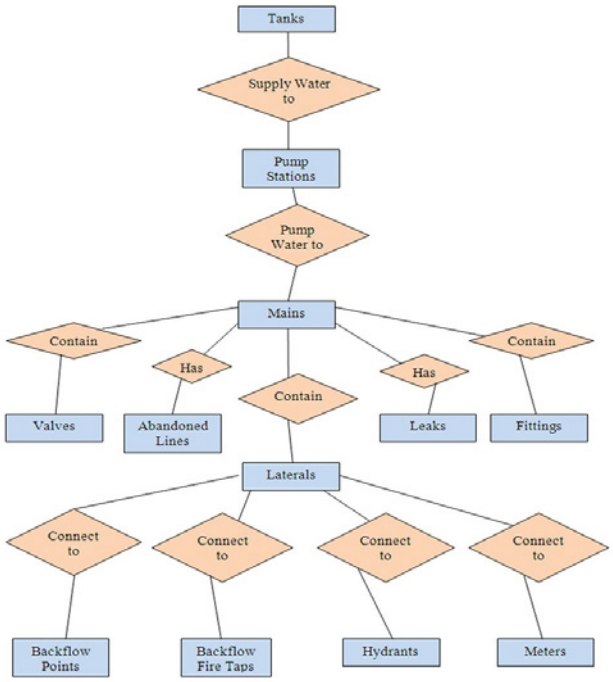


Figure 2. Water Department ER diagram

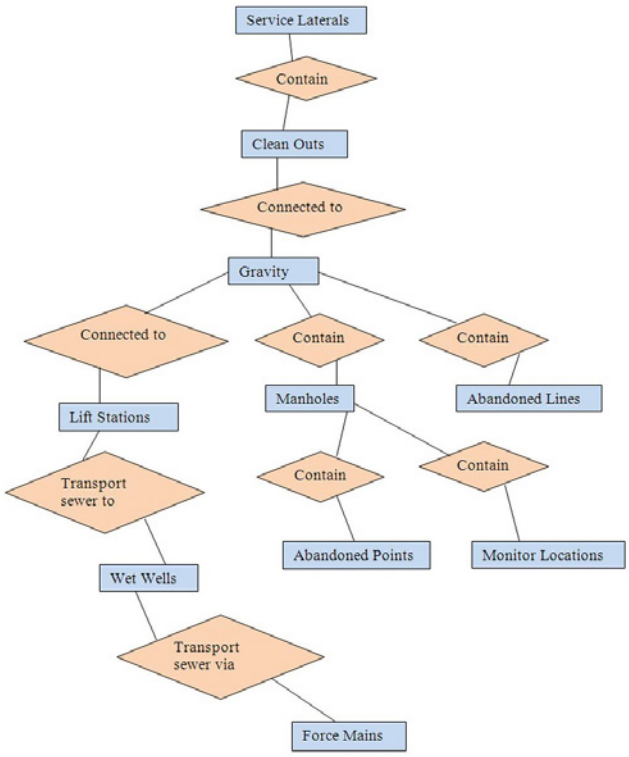


Figure 3. Sewer Department ER diagram

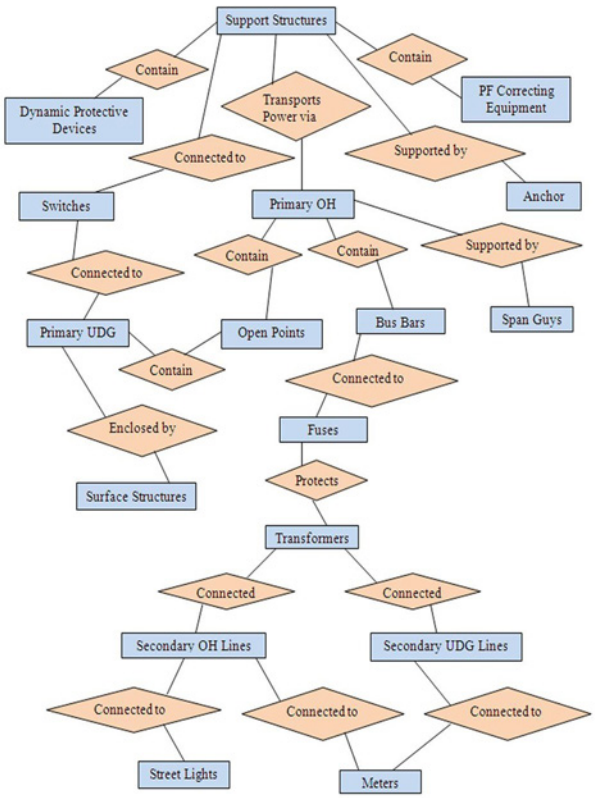


Figure 4. Electric Department ER diagram

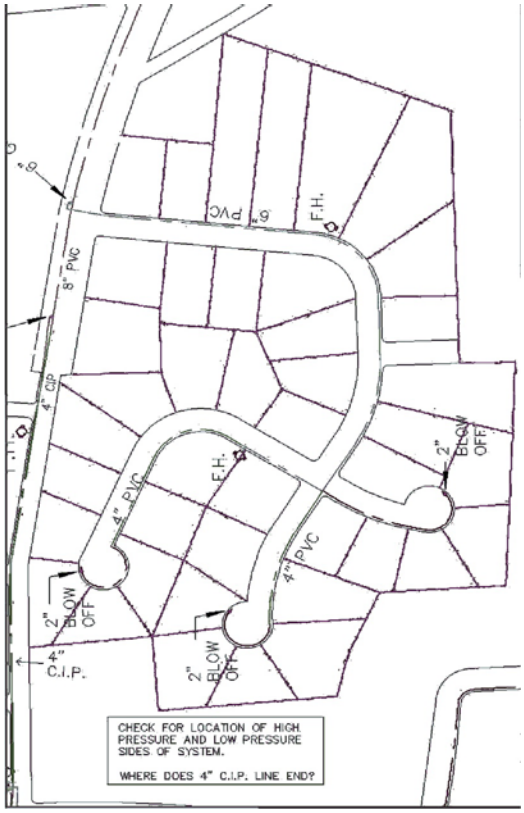


Figure 5. Example water map in CAD not to scale

highly accurate to within a few inches. Some of the features located by the Engineering Department include wastewater manholes, utility poles, and street centerlines. The features were located with RTK GPS on the Georgia State Plane Coordinate System. The Sewer and Water Departments have already begun using the location data and aerial photography. This data will be very helpful in referencing other data throughout the city for the GIS.

WATER SYSTEM DATA CONVERSION AND DEVELOPMENT

The existing CAD data for the water system is based on an assumed coordinate system, which is not on a defined coordinate system. A field inventory of the valves and hydrants was completed to accurately place the lines. The water-line attributes were populated by using the existing CAD data. The valve diameters were field-verified while determining the RTK GPS locations. By using the aerial photography, planimetrics, and the RTK GPS valve and hydrant data, the water dataset was populated in GIS. A large amount of the line locations came from meetings involving the water crews who actually installed the lines or who have performed repairs on the lines.

The water meters were located using a handheld GPS in which the data had to be differentially corrected using postprocessing software. The handheld GPS unit produced accuracy within a foot compared to inches of the RTK GPS. The handheld unit was used because of the obstruction of trees and bushes usually around meters. The RTK GPS needs at least the visibility of four satellites to get a fixed location. The locations of the meters were incorporated into the procedure of meter replacements for the Automated Meter Reading (AMR) system. During the AMR project, each meter had to be changed out with a new meter, and a GPS location was taken and attributes noted.

Figure 5 shows an example of the water CAD drawings the department has been using for many years. The map displays water main locations with hydrants and valves. The drawing was created without using any scale. The GIS map in Figure 6 shows the newly acquired data in the same area on the State Plane Coordinate System. The GIS map not only shows the water main, hydrant, and valve locations, but also the meter locations. The backflow fire taps and backflow points also were added to the water dataset. The backflow features allow the Water Department to keep track of testing dates and the history on each backflow location.

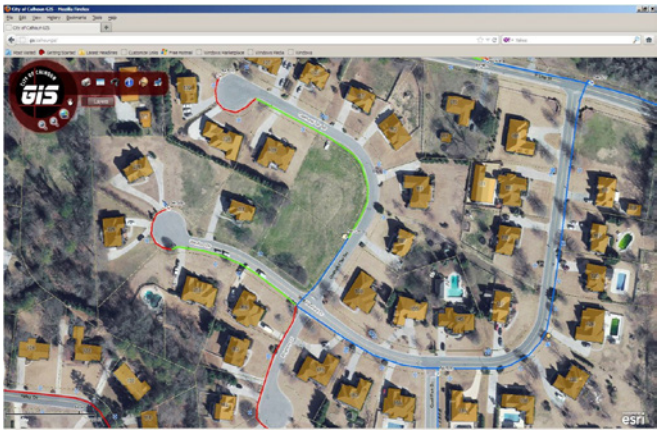


Figure 6. Example water map in GIS

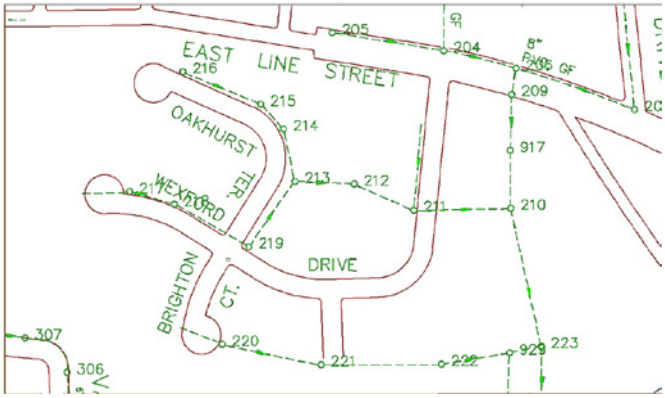


Figure 7. Example sewer map in CAD not to scale

SEWER SYSTEM DATA CONVERSION AND DEVELOPMENT

The sewer data was much easier to convert and develop than the water data because of the size of the system. The sewer system only serves the residences within the city limits. Approximately 35 percent of all sewer manholes already were located using RTK GPS and attributed by the Engineering Department. Also a number of developers over the past few years have submitted drawing plans on subdivisions in the State Plane Coordinate System showing the locations of the sewers installed.

A field inventory located and attributed the remaining 65 percent of the manholes. During the field inventory, each manhole was located with RTK GPS, which provided accurate and precise location and elevation. The elevation data for each manhole had to be accurate for modeling purposes in the future. The manhole covers were removed and a measurement of the depth was taken. The material of the pipe and condition of the manhole also were noted during this time.

The CAD map in Figure 7 shows an example of a sewer drawing. The CAD drawing displays the manholes and gravity mains. The CAD drawing was created without any scale. Figure 8 shows an example of the same area in GIS. The GIS map not only shows the manholes and gravity mains but also the lateral lines that are displayed in yellow.

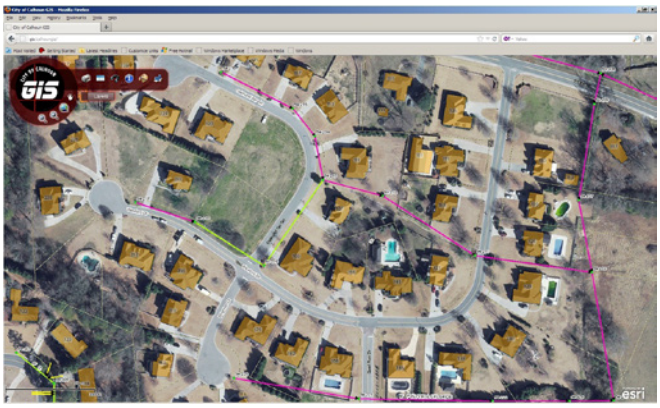


Figure 8. Example sewer map in GIS

ELECTRIC SYSTEM DATA CONVERSION AND DEVELOPMENT

The electric data like the water and sewer data existed in CAD format only with assumed coordinates. The attribute information for all the electric data was in the form of labels in the CAD system. The data conversion began with georeferencing all the poles and then applying the attributes based on the CAD data. The aerial photography provided enough detail to pick out poles and lines in many of the areas. Poles had to be located by GPS in the areas where they could not be identified by the aerial photography.

The development of the GIS began by adding each layer of data one at a time by circuit. The electric system is made up of three substations and 18 different circuits. After the poles were referenced and located, the other features were added. The feature classes that make up the electric datasets are mostly attached to poles except for surface structures and underground transformers. All the features associated with the underground lines were located in the field with GPS. The phasing of the primary lines was field-verified by visiting each of the three substations and tracing out the circuits. A number of phasing errors was identified during this process and changed in the GIS.

Figure 9 shows an example of a CAD electric drawing. The CAD drawing was drawn similar to the water and sewer drawings with no scale. The map in Figure 10 is an example of the new GIS data developed for the same area. The symbology remained as close to the old CAD drawings as possible for the field crews.

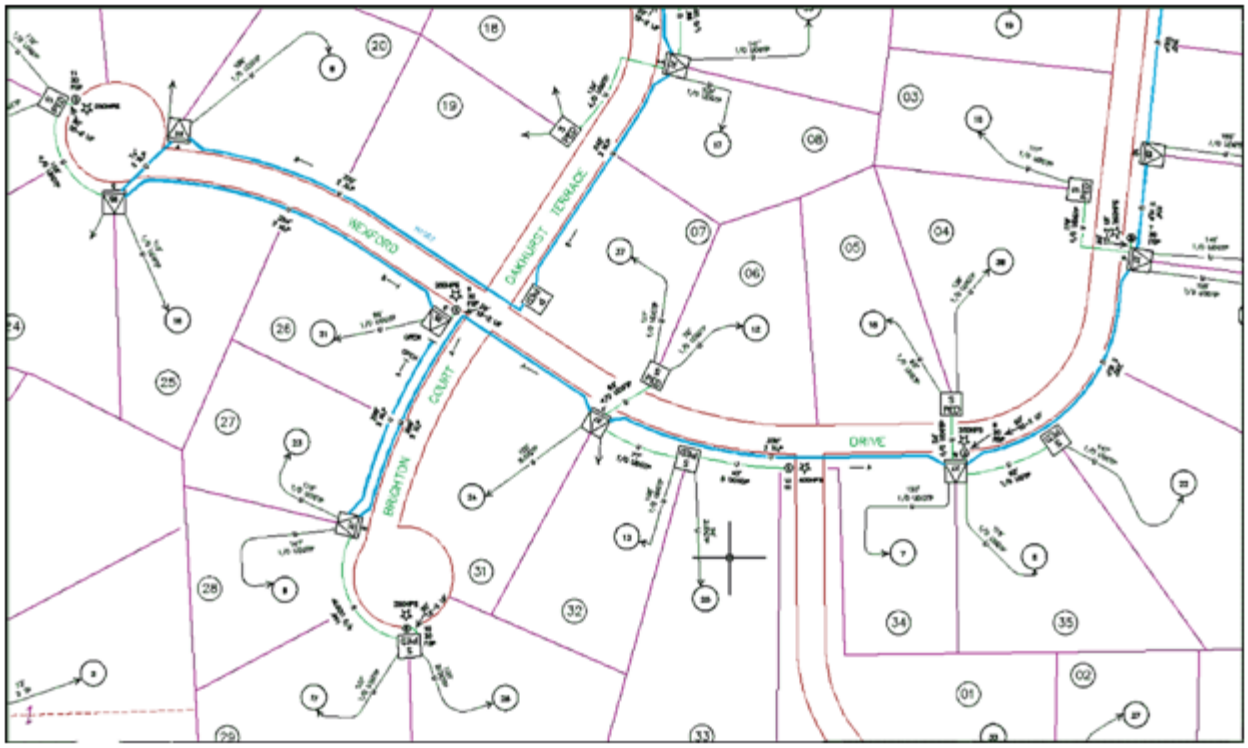


Figure 9. Example electric CAD map not to scale

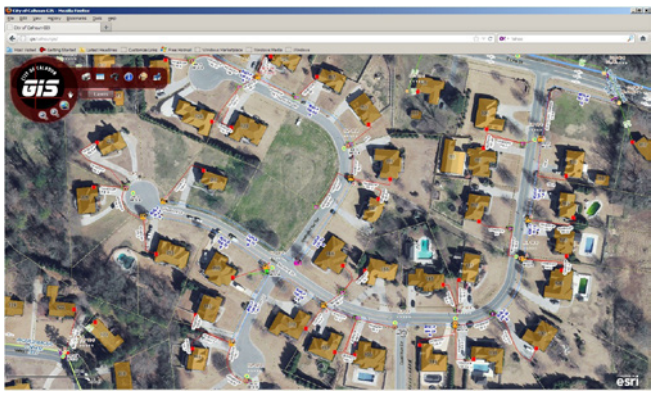


Figure 10. Example electric map in GIS

RESULTS AND CONCLUSION

The implementation and development of the utility GIS was completed in a three-year timeframe. The system is continuously being updated with new layers of data. The development of the GIS is a never-ending, ongoing project because of the applications, tools, and analyses each department requests. The GIS today is being used by all departments on a daily basis. The system has become a tool that the departments depend on to carry out their daily functions and decision making. The GIS provides these departments with the ability to analyze and manage their entire infrastructure. The system provides each department with linked data from the billing, work order, and AMR systems. The locations of the features are very accurate and precise compared to the old CAD mapping techniques used. The data now is stored in a geodatabase that provides an unlimited amount of information to be linked to the features in the infrastructure.

The two main types of benefits gained from the GIS are efficiency and effectiveness. Efficiency benefits occur when a GIS is used for a task that was not previously performed with GIS and the output quality is same, but at a lower cost. The effectiveness benefits result when a GIS is used to improve the quality of an output or to produce the output that was not previously available (Gillespie 2000).

Compared to the old CAD system, the GIS has improved the efficiency of the departments greatly by reducing the amount of time needed to carry out daily processes. A few examples of the time saved by using the GIS are listed below by departments.

Water Department

- The CAD drawings had no meter locations. The GIS now provides the department with accurate location of all meters for replacement purposes.
- Address lookup for work orders saves time for the work crews when responding to a service call. The CAD system provided no physical address data.
- Real-time reading of meters allows the Electric and Water Departments to view the reading by clicking on the meter in the GIS. Prior to the AMR and GIS implementation, meter readers were used in the field to record readings.
- The Water Department now is able to easily perform consumption analysis by using the readings from the meters in GIS.

- Water valve isolation of system when leaks occur was a huge problem before the GIS was implemented. The location of the valves in the CAD drawings was not accurate, which wasted a lot of time in the field trying to locate them.
- Mapping of leaks in the GIS now provide the Water Department with the location of future rehabilitation of certain pipes in the system.
- Management of backflow testing for businesses was managed by a long paper trail in the past. The GIS now provides the Water Department with locations of all the backflows and links to the test results.
- Utility locate requests now are being done mostly in the office using GIS instead of in the field.

Sewer Department

- Inflow and infiltration projects are managed based on data from GIS. The drainage basins were developed by using a combination of elevation and feature data. The department had no inflow and infiltration mapping in the past.
- GIS provides the crews with physical address lookups for work orders. Prior to GIS, the crews had to use paper street maps with no house numbers.
- Inspection videos now are attached to each pipe segment, allowing quick access to viewing.
- The sewer CAD drawings had no locations of the service laterals in the system. The GIS now has a number of service lateral locations that were identified by reviewing the inspection videos.
- The CAD drawings had no monitor manhole locations. The GIS has all monitor manhole locations with data attached.
- Utility locate requests now are being done mostly in the office using GIS instead of in the field.

Electric Department

- Daily mapping of work orders in the past with the CAD data was not accurate, which led to time wasted in the field. The GIS has provided the crews with accurate mapping that can be relied on in the field.
- The department now has the ability to read the meters from the GIS instead of having to read them in the field.
- Measuring distances for proposed lines in the office instead of in the field has saved the department valuable time.
- Physical address data in the GIS has helped the crews quickly identify the location of outages.
- Utility locate requests now are being done mostly in the office using GIS instead of in the field, saving time and resources.

The focus of this research has been to review, explain, and show an example of how a GIS was developed for the city of Calhoun utility departments. This research showed how the GIS helped improve the efficiency within each of the utility departments for the city of Calhoun. The implementation process outlined in this research can provide other cities and municipalities with the knowledge and foundation to develop their own GIS. This research can provide individuals with a reference to implement a utility GIS from the beginning to the end. Using the GIS to manage the city of Calhoun's utility data has allowed more flexibility over the previous CAD system when analyzing data. The management of the utility data has become much more efficient compared to previous management with CAD. The utilities departments of the city of Calhoun have embraced GIS and look forward to advancements to the system.

About the Authors

Davie Crawford graduated from Northwest Missouri State University's online Master of Science in GIS program in 2012. He currently works as the GIS manager for the city of Calhoun, Georgia.

Corresponding Address:
City Of Calhoun GIS
700 West Line Street
Calhoun, GA 30701
E-mail: dcrawford@calnet-ga.net

Ming-Chih Hung, Ph.D., is an associate professor of geography/geographic information science at Northwest Missouri State University, Maryville, Missouri. His major research interest is GIScience (e.g., GIS, remote sensing, GPS, cartography, visualization) and urban environments.

Corresponding Address:
Northwest Missouri State University
800 University Drive
Maryville, MO 64468
E-mail: mhung@nwmissouri.edu

References

Arctur, D., and M. Zeiler. 2004. Designing geodatabases: Case studies in GIS data modeling. Redlands, CA: ESRI Press.

Bolstad, P. 2008. GIS fundamentals: A first text on geographic information systems. Third Ed. White Bear Lake, MN: Eider Press.

Cannistra, J. R. 1999. Converting utility data for a GIS. American Water Works Association Journal 91(2): 55-64.

Croswell, P. L. 1991. Obstacles to GIS implementation and guidelines to increase the opportunities for success. Urban and Regional Information Systems Association Journal 3(1): 43-56.

Environmental Systems Research Institute (ESRI). 2003. Shredding the map: Building an enterprise geographic information system for utilities. Available from http://www.esricanada.com/documents/shredding-the-map.pdf. Accessed October 10, 2009.

Gillespie, Stephen R. 2000. An empirical approach to estimating GIS benefits. Urban and Regional Information Systems Association Journal 12(1): 7-13.

Gordon County. 2008. Gordon County detailed profile. Http://www.city- data.com/county/Gordon_County-GA.html.

Harmon, J. E., and S. J. Anderson. 2003. The design and implementation of geographic information systems. Hoboken, NJ: John Wiley & Sons.

Hughes, J. 2006. GIS combines geography and information for effective utility management. American Water Works Association Journal 32(12): 10-11.

Longley, P. A., M. F. Goodchild, D. J. Maguire, and D. W. Rhind. 2005. Geographic information systems and science. Second Ed. England: John Wiley & Sons Ltd.

Shamsi, U. M. 2002. GIS tools for water, wastewater, and stormwater systems. Reston, VA: American Society of Civil Engineers.

Shamsi, U. M. 2005. GIS applications for water, wastewater, and stormwater systems. Boca Raton, FL: Taylor and Francis.

Sieber, R. E. 2000. GIS implementation in the grassroots. Urban and Regional Information Systems Association Journal 12(1): 15-29.

Tomlinson, R. 2003.Thinking about GIS: Geographic information system planning for managers. Redlands, CA: Environmental Systems Research Institute.

Tulloch, D., and Harvey F. 2007. When data sharing becomes institutionalized: Best practices in local government geographic information relationships. Urban and Regional Information Systems Association Journal 19(2): 51-59.

Uhrick, S., and D. Feinberg. 1997. Integrating GIS with utility information management systems. Available from http://www.gisdevelopment.net/proceedings/gita/1997/bepm/bepm20pf.htm. Accessed September 17, 2009.

Zeiler, M. 1999. Modeling our world: The ESRI guide to geodatabase design. Redlands, CA: Environmental Systems Research Institute.

Desirable Characteristics of an Online Data Commons for Spatially Referenced, Locally Generated Data from Disparate Contributors

James Campbell and Harlan Onsrud

Abstract: *A significant body of spatially referenced, locally produced data in small isolated collections exists on the hard drives and backup systems of individual researchers, nonprofit groups, private associations, small companies, universities, and nongovernmental organizations across the United States. From a practical perspective, that data currently is unavailable to professional scientists and to the general public. If there were an online environment where that data could be deposited or registered and readily found, what infrastructure characteristics might potential users find desirable for them to be willing and interested in finding, consulting, and using such data? While there are major national and international initiatives such as the Global Earth Observation System of Systems (GEOSS) that are providing a gateway for access to millions of spatially referenced datasets, primarily from national government data sources, a similar gateway to access spatially referenced, locally produced datasets from disparate private and nonprofit sources has yet to emerge. If one or more were to emerge, what characteristics should be incorporated into the design to make it useful to users of the portal or gateway?*

Based on data-preservation literature, this study posits three potential characteristics as desirable: make conditions of use of data files clear to potential users; provide a variety of ways to search for data; and enable users to access comments and feedback from prior users, and add comments of their own. These three characteristics were examined because they often are not provided or inadequately provided in general-purpose portals for finding geographic data and services. A combination of qualitative and quantitative methods was used and the results of the analysis using both methods support the hypothesis.

INTRODUCTION

Background

A significant body of spatially referenced, locally produced, small-scale data developed for specific local purposes exists on the hard drives and backup systems of individuals, nonprofit groups, private associations, universities, private companies, and other nongovernmental organizations across the United States. Spatially referenced data, as the term is used here, is data that refers to a particular physical location. Examples might include a university botany class project that locates and catalogs all the trees more than 15 feet tall in a small town; a homeowners' association that monitors the water quality and plant growth of the lake on which members' properties are located; a land trust that records environmental easements; or a historical museum that ties its photographic images to their physical locations, among many others.

In all these cases, the data gathered by these small local originators could be of great value to others if its existence were known. At present, however, very little of this data is available from a practical perspective to other scientific researchers and potential users. It is, for all intents and purposes, completely or partially "invisible."

While much emphasis has shifted in recent years to providing geospatial services, there still is a strong need for service developers to be able to find and exploit existing geographic data that would make those services more effective and efficient. Many efforts at the national and state levels are being made to make government-generated spatially referenced data available to the public. In the United States and in other countries around the world, initiatives are under way to make geographic information more freely available to scientists and to the general public. In English-speaking countries, for example, UK Location (http://location.defra.gov.uk) in the United Kingdom, the Atlas of Canada (http://atlas.gc.ca/site/english/index.html), and Geoscience Australia (www.ga.gov.au) provide open access to some government-generated spatially referenced data. In the United States, initiatives such as the National Map (http://nationalmap.gov), the National Atlas (www.nationalatlas.gov), and the geospatial section of data.gov (http://www.data.gov/geospatial/) serve similar functions. These U.S. sites contain a wider array of data than many other national portals because the U.S. federal government cannot hold copyright on materials it generates, and because some state governments make their state-level data visible through these gateways. Efforts also are under way to make international sharing of large datasets more viable, especially with regard to divergent approaches to data licensing and use rights (Onsrud et al. 2010). GEOSS Data

Collection of Open Resources for Everyone (GEOSS Data-CORE 2014) is an example of an international initiative to support open access to geographic data gathered by governments across nine societal benefit areas (GEOSS 2014).

Similarly, disciplinary and special purpose repositories exist to capture large sets of spatially referenced data. Examples include PANGAEA (http://www.pangaea.de) and OneGeology (http://www.onegeology.org).

Google Maps, Google Earth, Virtual Earth, and Open Street Maps provide structured environments where the user may take advantage of a data-gathering and display infrastructure to contribute data or volunteer effort to a commercial or open-data environment. In these information infrastructure environments, legal and data management issues as well as data format issues are closely controlled by the infrastructure system provider. These are not infrastructure environments for depositing or finding diverse geographic datasets, and this article does not address such environments.

We conclude that no gateway exists analogous to the Global Earth Observation System of Systems (GEOSS) that could provide more visible and efficient access to millions of spatially referenced datasets drawn from disparate locally generated sources. Note that the GEOSS is a *portal* or *gateway* for finding relevant geographic data and services rather than a *repository* of geographic data itself. Furthermore, the metadata on geographic data and services contained within the GEOSS is provided or mined from primarily national and international government members and participating organizations of the Group on Earth Observations (GEO). The GEOSS serves as an exemplar of the kind of infrastructure that can make geospatial data files and services from widely disparate cooperating sources much more readily findable.

VOLUNTEERED GEOGRAPHIC INFORMATION (VGI)

In the past decade, regular people have become producers as well as consumers of geospatial data, a phenomenon variously called neogeography (Turner 2006, Sui 2008), ubiquitous cartography (Gartner et al. 2007), collaboratively contributed geographic information (Bishr and Mantelas 2008), and volunteered geographic information (Goodchild 2007). VGI seems to be the most widely used term at present.

Affordable, portable GPS devices have made it possible for anyone to make a quite accurate observation of the position of an object on the face of the earth. Simple-to-use infrastructures that use Google Maps, Open Street Maps, or similar frameworks make it easy to add those observations to a map, and to attach notes or information to the location. To date, the great bulk of VGI activity has involved this form of adding locations and labels of features within a mapping facilitation framework or to already existing maps. At the observation level, then, VGI contributors can contribute data in many situations as well as trained geographers could in pre-GPS days.

Adding or correcting locations, names, and characteristics of features on a map base such as Google Maps or Open Street Maps is a type of spatially referenced data but there are many other types including complete datasets of various kinds such as the examples mentioned previously. Most of the examples involve "asserted" rather than "authoritative" data (Bishr and Mantelas 2008). In VGI-contributed environments, where disparate datasets are only asserted as potentially useful and not vouched for, context becomes crucial. VGI data, or any data, collected for one specific purpose may not be relevant or useful or even accurate for a different purpose. Potential online environments that may feature collections of data generated locally for disparate purposes need to contextualize that data for the data to be useful.

DESIRABLE CHARACTERISTICS OF AN ONLINE SPATIALLY REFERENCED DATA REPOSITORY

Simply having an online gateway or home for widely disparate, spatially referenced, locally generated datasets could be of significant use for providing access to this type of data. It probably would be of greatest use to geospatial specialists and professionals desiring to find and draw from existing spatially referenced data to provide further products and services. We refer to this perceived online gateway or home as a Commons of Geographic Data (CGD). However, if such a facility or capability, centrally located or distributed, is to be of maximal use over time to both professional scientists and to interested nonprofessionals, a number of studies and reports suggest that it should include functionality that enables users to know usage rights and search for and discover data using standards-based metadata, and provide users with a way to access evaluation commentary from previous users of the datasets and offer comments of their own. See these common elements in, for example, Report of the Workshop on Opportunities for Research on the Creation, Management, Preservation and Use of Digital Content (IMLS 2003), Licensing Geographic Data and Services (NRC 2004), and To Stand the Test of Time: Long Term Stewardship of Digital Data Sets in Science and Engineering (ACRL 2006).

In a commons-type environment for data users, data is made available under a license—if a license is necessary to use the data—that grants permission for use as long as any stipulated conditions are adhered to. This makes it possible for potential users to be sure that they may use any data found in such a commons environment without seeking additional permission from the owner. In such environments, permission already has been granted as long as any conditions specified in the license are respected. Creative Commons licenses are one example of so-called "some rights reserved" license types typically found in a commons environment for materials that are not in the public domain. Creative Commons licenses currently are used in more than half a billion digital works. Creative Commons and its affiliate, Science Commons, have designed several licenses specifically

applicable to datasets (Creative Commons 2014) that could be used in a Commons of Geographic Data.

An online Commons of Geographic Data with the characteristics listed previously does not exist at present. If such an environment were contemplated as a future project, based on the reports previously cited, important questions arise almost immediately. If there were such an online data commons repository for small, privately generated datasets, would people who are interested in spatially referenced data be willing to access and use the data in such a repository? What type of functional characteristics of such a repository or gateway would help to motivate those potential data users to actually examine and possibly use the data located there for their own purposes?

It may seem reasonable to assume that such characteristics would be desirable to potential users, but at this point in time, reasonable or not, this still is an assumption. The goal of this research is to address this question empirically.

HYPOTHESIS

The purpose of this research is quite practical. It is hoped that the results may provide some guidance for future architects of an online Commons of Geographic Data about functionality that potential users would be interested in finding in an online commons environment for spatially referenced small datasets from disparate sources, if and when such a commons environment is constructed. The results could suggest several areas for future research, and might also be of use to those who currently operate data gateways or repositories that they would like to make more responsive to users' interests.

Based on common elements in the reports noted previously as well as in other data-preservation related studies (e.g., Committee on Science, Engineering, and Public Policy (U.S.) 2009, Interagency Working Group on Digital Data 2009), we hypothesized that potential data users would be willing to consider using data accessed through an online gateway or data repository if such a facility included:

(a) a simple, clear licensing mechanism that reveals ownership of, and conditions for use of, the contributed data;
(b) a simple, effective searching/finding mechanism that provides an option to search using either *Thesaurus*-controlled vocabulary, "plain English" keywords, or location; and
(c) a simple postpublication peer-evaluation mechanism that will provide information on quality and suitability for purpose for users.

METHOD

To test this hypothesis, we used a combination of qualitative and quantitative research procedures (Onwuegbuzie and Leech 2004; Ragin, Nagel, and White 2004). Personal interviews were conducted with ten people who were regular users of spatially referenced data. These particular interviewees also were generators of spatially referenced data. The findings from these qualitative interviews were used to construct an online questionnaire, and results from that questionnaire with responses from a much larger group (139 people) were compared with the results from the interviews to see if the qualitative results were supported by quantitative data.

METHODOLOGICAL LIMITATIONS

The respondents in this study are not in any way meant to be considered a statistical or otherwise representative sample of potential data users of an online commons gateway or repository for spatially referenced datasets from disparate sources. The major reason for not attempting to select a representative sample of potential users is that the universe of such users is unknown and probably unknowable. Thus, the combination of qualitative in-depth interviews with quantitative data was chosen to produce findings that would be informative, even though not "proven" in a statistical sense, for future designers of an online commons-type geospatial data environment, and that could suggest directions for future study.

All participants in the study were self-selected. In addition, to generate quantitative responses online, given the reverse traceability of personal user information in today's online environment, potential respondents were guaranteed anonymity by requesting no geographic, employment, or other demographic information. This makes some types of statistical analysis impossible.

INTERVIEWEES AND DATA TYPES

Interviewees were selected based on a "snowball technique" (Maxwell 2005). Interviewees were referred by word of mouth from those interested in spatially referenced data who were located in geographic areas accessible to the authors. Those who agreed to participate were asked if they could recommend others who might be potential interviewees. In the final group of ten interviewees, seven were from Maine, one from Massachusetts, one from Pennsylvania, and one from North Carolina.

One interviewee was a graduate student working on a spatial-data research project; one regularly dealt with spatially referenced data as part of the respondent's employment, although the role the respondent held in this study was as a volunteer citizen on a municipal committee. About half the respondents were familiar with and used GIS software to a greater or lesser degree; about half did not. Four were involved with land trusts of one type or another, one was an author of nature books, one a high school teacher, one a local museum curator, and the others were involved with other types of local civic groups. All the spatially referenced data that these originators were gathering were deemed by the investigators and the gatherers to be of potential interest to others in the future but none of the data was available on the Web.

QUALITATIVE DATA-COLLECTION PROCESS

The purpose of these qualitative interviews was to test whether the hypothesis above would hold, and to discover if other important

desirable characteristics arose spontaneously in the interviews. All interviews were conducted from the same interview instrument by the same interviewer. The interviews were transcribed and coded, and then the transcripts were checked against the voice recordings for accuracy. A summary of key points then was sent to each interviewee for correction, if necessary, and for confirmation. None of the interviewees who responded submitted any corrections other than spelling errors.

Because all interviewees were asked the same set of questions, initial top-level codes were based on those questions, e.g., "conditions" (which owners might put on use of contributed data); "metadata" (short description, keywords, search order, etc.); "evaluation" (valuable or not, amount of time willing to spend commenting, etc.). As additional aspects of responses appeared, subcategories for the major categories were added to make meanings more precise, and a few additional top-level codes added for topics that emerged that were not specific responses to asked questions but that were relevant to overall online data commons use.

QUANTITATIVE DATA-COLLECTION PROCESS

Based on the information generated in the analysis of the qualitative data, an online questionnaire was constructed to see if others who identified themselves as users of spatially referenced data would agree with the responses of the ten interviewees regarding the hypothesis points. Notice of the existence of the questionnaire along with an invitation to participate in the research was sent out to listservs of those concerned with geographic information of different types, specifically to members of the Global Spatial Data Infrastructure Association and to members of the Maine Geolibrary listserv. In addition, printed flyers inviting participation were distributed at a conference of the Maine GIS User Group and the Maine Municipal Association.

The survey instrument used the first question to separate those who were owners of, or who had significant influence on data sharing in their organizations (potential contributors), from those who considered themselves only potential data users.

All those who identified themselves as potential contributors also considered themselves potential users, and there were additional respondents who considered themselves users only. We report on the results of the questions answered by all users, including those who also identified themselves as owners or controllers of spatially referenced data. There were 11 questions data users were asked to answer in the survey, of which three requested text-based responses.

As in the qualitative portion of the research, no attempt was made to construct a statistically valid sample. Rather, the goal was to gather a reasonable number of responses from self-identified potential users of spatially referenced data to either support or invalidate the qualitative research findings.

There was a total of 197 click-throughs from the survey splash page to the actual survey instrument. Each click-through response was given a specific ID for analysis purposes. Of 197 click-throughs, 139 completed some or all of the questions put to users.

RESULTS

We review the results by each hypothesis subpart. Although the prior discussion refers to both portals and repositories for geographic data, with the human subjects we focused on the simpler concept of data repositories. However, we believe the results are generalizable for also guiding feature developments for portals or gateways such as GEOSS that lead to distributed repositories or portals.

HYPOTHESIS SUBPART (A): SIMPLE CLEAR TERMS OF USE

Hypothesis: Data users would be willing to consider using data in an online data repository if such a repository included a simple, clear licensing mechanism that reveals ownership of, and conditions for use of, the contributed data.

QUALITATIVE RESULTS

All ten of the interviewees indicated that they would want to be able to check license conditions before they decided to download and use data, and that they would respect any conditions that were put on the use of the data in a particular file. Most indicated that they would want a simple-to-understand statement of what they could or could not do with a data file. In the words of one interviewee: "I would want to be able to identify the conditions or at least get a sense of the conditions very quickly . . . I am not going to spend a lot of time reading a three-page license agreement."

Several assumed that any conditions for use would be stipulated when a file was found, and certainly by the time it was opened, although another interviewee said that the interviewee always scans the Web page a file appears on to see if, for example, attribution is required.

Several interviewees referred to ethical considerations when describing whether and why they would check any licensing conditions before using the data in any but a personal way. Two of the interviewees indicated specifically that they would not bother to check for licensing conditions if they were just looking at the data for their own information, but if they contemplated using it in any additional way, they would check and respect any conditions of use.

Interviewees were asked if the presence of conditions of use that were clearly stated before opening a file might impact whether they would choose to look at a data file or not. Responses were evenly divided between those who would look at the data anyway and those who would not bother if they felt the conditions would preclude the use that they might wish to put the data to.

QUANTITATIVE RESULTS

Results from responses to the online questionnaire are consistent on this topic with those gleaned from the personal interviews.

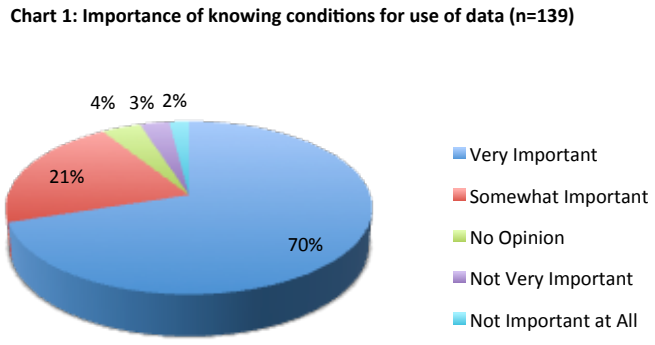


Chart 1. Importance of knowing conditions for use for data (n=139)

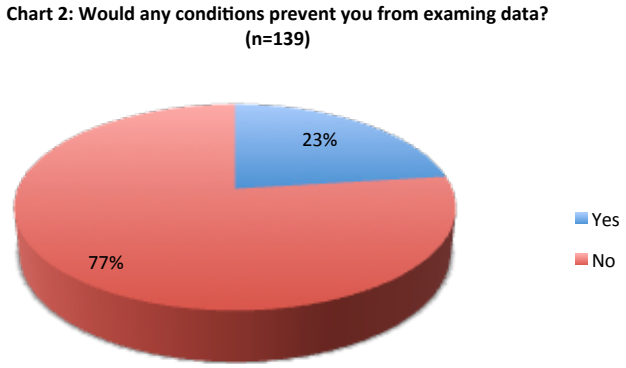


Chart 2. Would any conditions prevent you from examining data? (n=139)

Users were asked in each question "If you were looking for data that others had contributed to an online commons-type environment, please indicate how important each of the following would be in your decision of whether to access and/or use such data . . ."

Users were given five choices:

- Very Important
- Somewhat Important
- No Opinion
- Not Very Important
- Not Important at All

This first question asked how important it would be that "Conditions for the use of the data are clear." See Chart 1. (Note that all the following chart percentages are rounded.)

The importance of knowing the conditions for use expressed by interviewees is mirrored in the larger population of questionnaire respondents, with 91 percent indicating that such knowledge would be "Very Important" or "Somewhat Important" to them.

Addressing the question of whether licensing conditions put on the use of the data would affect potential users from accessing the data, respondents were asked: "If conditions for use of the data were clear, e.g., requiring attribution or noncommercial use only, might there be any conditions that would prevent you from examining the data?" (See Chart 2.)

Of those questionnaire respondents who responded "Yes" to this question, examples of conditions that might prevent users from examining a data file varied. The predominant response concerned limitations on commercial use. Some other reasons included cost, administrative requirements, concern about data quality, limited bandwidth that would preclude downloading large files, and inability to modify the data for their own use.

HYPOTHESIS SUBPART (B): SEARCH MECHANISM

Hypothesis: Data users would be willing to consider using data in an online data repository if such a repository included a simple, effective searching/finding mechanism that provides an option to search using either Thesaurus-controlled vocabulary, "plain English" keywords, or location.

QUALITATIVE RESULTS

None of the interviewees said that they would search for data based on Thesaurus-controlled vocabularies. All would begin searches using either natural language keywords and phrases, or location terms. All interviewees indicated that they might use either strategy first depending on what they were looking for at a particular time. About half indicated that they usually would begin with topic keywords, about half with location. However, each group then would use the other strategy to help narrow their results.

For example, an interviewee who served on a municipal recreation committee interested in resident uses of lakes described a strategy for finding that type of information: "So when we start to look out and search the Internet we throw a broad net at the beginning based on certain things like those lake management plans but when we get down to specifics we start looking at information of lakes that are more in the same latitude or in close proximity to where the municipality that we live is." Another interviewee who worked with a local land trust took a different approach: "In terms of my work and the way I would do it, it would be place based; it would be coming from the place to the information."

In either case, interviewees found being able to begin their searches either by topic or place keywords was important for their search strategies.

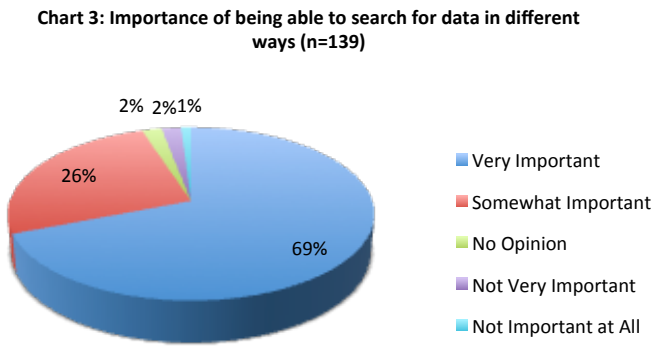


Chart 3. Importance of being able to search for data in different ways (n=139)

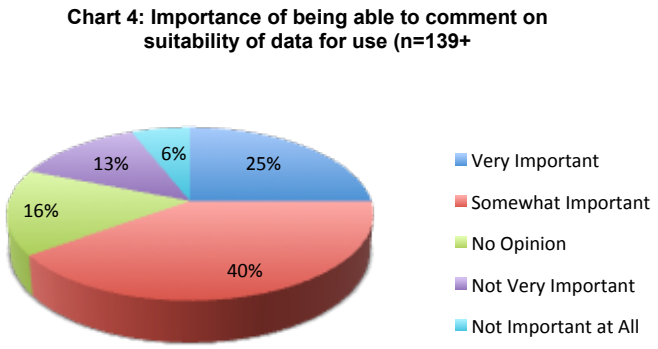


Chart 4. Importance of being able to comment on suitability of data for use (n=139)

QUANTITATIVE RESULTS

Questionnaire respondents were asked how important the "Ability to search for data in different ways, e.g., by location, keyword, etc." would be to them. The results are consistent with those from the interview phase of this research. (See Chart 3.)

Being able to conduct searches using different starting points, including location and natural language keywords, appears to be an important functional capability for an online repository for locally generated, small-scale spatially referenced data.

HYPOTHESIS SUBPART (C): PEER EVALUATION

Hypothesis: Data users would be willing to consider using data in an online data commons environment if such an environment included a simple post-publication peer-evaluation mechanism that would both provide feedback for contributors, and provide information on quality and suitability for use for users.

QUALITATIVE RESULTS

In this age of Amazon and online shopping, it is no surprise that interviewees used online shopping comments as an analog to looking at comments/evaluations in an online commons environment for spatially referenced data. Half of the interviewees made comments similar to this one: "I mean I buy CDs on Amazon.com" that indicated familiarity with commercial online retailer commenting systems that they found useful, and indicating that they would consult peer comments and evaluation of data files if such comments were available.

Half of the respondents, however, said that they would look at the data themselves if it were data that might suit their needs, no matter what the comments said. Two indicated that they would look at the data first and only subsequently consult other user comments to see if those corresponded with their own judgments.

Only one interviewee said that the interviewee would be unlikely to consult comments made by others because the interviewee preferred to form a personal opinion directly from the data.

One interviewee indicated that "junk comments" were always a potential problem in evaluation systems and recommended that any such system have a moderator who would screen comments for civility, relevance, and, if possible, quality before posting them.

Other interviewees who would consult comments made by others indicated that while they would not view it as necessary, they would prefer to know who the commenter was so that they could form an opinion about the relevance or quality of the comment source if the commenter were known to them.

Nine of the interviewees indicated that they would be willing to make comments if they felt that they had something useful to say about a file. Most said that they would be willing to spend a limited amount of time, 5 to 15 minutes, to input a comment if there were a simple way to do so.

Consistent with the desire to know who made a comment, all nine said that they would be willing to use their own names rather than to use a screen name in offering a comment.

In summary, the majority of interviewees would find a commenting/evaluation system valuable in an online commons repository.

QUANTITATIVE RESULTS

Support for the "Ability to comment on the suitability of the data for your uses" was not so strong among survey respondents as among interviewees, although it was substantial, with 65 percent finding that capability "Very Important" or "Somewhat Important." (See Chart 4.)

The amount of time that survey respondents would be willing to spend providing a comment generally mirrored what most interviewees would spend, 5 to 15 minutes. Given 139 responses rather than 10 as in the personal interviews, however, it is not surprising that there were a few outliers who would commit anywhere from "no time" to "as much as would be needed."

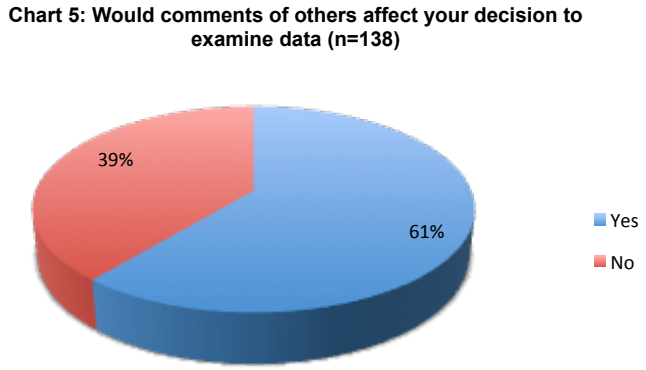


Chart 5. Would comments of others affect your decision to examine data? (n=138)

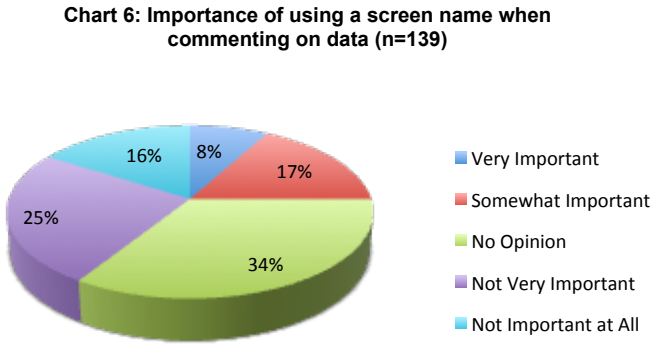


Chart 6. Importance of using a screen name when commenting on data (n=139)

In response to the question "Would the comments of other users affect your decision about whether to examine data that is available in the repository?" of 138 responses, 61 percent replied "Yes" and 39 percent said "No" (see Chart 5).

When asked to "explain how comments of others might affect your decision about whether to examine data further," a large majority of those who answered (78 of 84) cited comments that dealt with data quality and accuracy. Here, again, the analogy of online commerce sites came up: "Same as eBay. If someone says the data are junk, I'll probably be reluctant to use them."

The other major reason expressed by respondents was not the quality of the data itself but rather the lack of suitability for purpose, e.g., "how the data fits with my base maps."

The "Ability to use a screen name rather than your actual name when commenting" was more of an issue to survey respondents than it was with the interviewees. (See Chart 6.)

While nine of ten interviewees would use their own names rather than a screen name when making comments and preferred to know the identity of those making comments when possible, 25 percent of questionnaire respondents felt it would be "Very Important" (8 percent) or "Somewhat Important" (17 percent) to be able use screen names when commenting, and a third did not express any opinion. The reason for this divergence from the attitudes of interviewees is not explainable based on the data this research gathered. The location of the questionnaire respondents might be an issue for commenting using one's real name, or employment status, or some other variable for which this research did not gather any data.

SUMMARY AND CONCLUSIONS

This research, subject to the caveats listed below, empirically suggests that it would be desirable from the perspective of potential users of spatially referenced data in an online commons-type environment to provide infrastructure capability that would:

- make conditions of use of files clear to potential users,
- provide a variety of ways to search for data, and
- enable users to access comments and feedback from prior users, and to add comments of their own.

There are other desirable features of a commons-type online infrastructure, as the reports cited previously outline. This research addressed only these three.

LIMITATIONS

As noted earlier, this research has several limitations that prevent any assertion that the hypothesis is "proven" in the usual meaning of that term. However, we can assert that the hypothesis is supported by the results of this study.

These limitations do not, we feel, limit the usefulness of the research results for their intended purpose: to provide guidance to those who may in the future choose to construct an online commons environment for locally generated, small-scale spatially referenced data that anyone, nonprofessional and professional alike, can use.

DIRECTIONS FOR FUTURE RESEARCH

This research is based on interviews and on online questionnaire results. Results from the interviews generally are confirmed by the survey results. Although percentages differed slightly, opinions about the hypotheses generally were shared both in the interviews and in the survey responses.

However, there was a noticeable disparity in the perception of the importance of being able to use a screen name rather than a real name to make comments, although because a large number of questionnaire respondents expressed "No Opinion," it is difficult to tell if the disparity was important. The absence of demographic, employment, or geographic location information for interviewees and questionnaire respondents makes it impossible to explain that

divergence based on those characteristics. This is an area in which additional research may be fruitful.

This study made no effort to directly ask comparative questions, e.g., is one factor, such as clarity of conditions, more important than another to respondents? Answers to such questions may be inferred from the responses in the importance respondents placed on each factor, but it also could be desirable to ask comparative questions directly.

POSSIBLE WIDER APPLICATIONS

While this research focused on a possible future online commons-type environment for spatially referenced data from widely disparate sources, the results could be of some use to operators of existing online spatial-data services. Understanding what is desirable to users in approaching data with which they are not familiar, especially non-GIS professionals, could be helpful for existing services to, for example, make clear in an obvious way any restrictions on use of their data. Portals that do not presently enable users to search for data in different ways may wish to evaluate whether such functionality would be desirable to their existing user base, and whether it might help to increase usage among current nonusers of their services. Sites that do not offer commenting capability may wish to investigate if that functionality might increase usage.

For designers of potential future online environments for spatially referenced data, which might include, for example, university libraries or state library systems, and possibly for operators of existing portals as well, we hope this research, though not designed to be statistically "proven," offers some empirical insight into what online characteristics users find valuable for spatially referenced data repositories and/or portals.

About the Authors

James Campbell (james.j.campbvell@maine.edu) is a PhD Candidate, and Harlan Onsrud (harlan. onsrud@maine.edu) is a Professor in Spatial Informatics, School of Computing and Information Science, University of Maine, 5711 Boardman Hall, Orono, ME 04469. Campbell's research interests focus on access to information from both a technical and a law and policy perspective. Onsrud's research focuses on the analysis of legal, ethical, and institutional issues affecting the creation and use of digital databases and the assessment of the social impacts of spatial technologies.

References

ARL Workshop on New Collaborative Relationships: The Role of Academic Libraries in the Digital Data Universe. Friedlander, A., and P. Adler, Association of Research Libraries & National Science Foundation (U.S.). 2006. To stand the test of time: Long-term stewardship of digital datasets in science and engineering. Washington, D.C.: Association of Research Libraries.

Bishr, M., and L. Mantelas. 2008. A trust and reputation model for filtering and classification of knowledge about urban growth. GeoJournal 72.

Committee on Science, Engineering, and Public Policy (U.S.). 2009. Ensuring the integrity, accessibility, and stewardship of research data in the digital age. Washington, D.C.: National Academies Press.

Creative Commons. 2014. Http://wiki.creativecommons.org/Data.

Gartner, G., D. Bennett, and T. Morita. 2007. Toward ubiquitous cartography. Cartography and Geographic Information Science 34: 247-57.

GEOSS Data Sharing Action Plan, as accepted at GEO-VII, November of 2010, https://www.earthobservations.org/documents/geo_vii/07_GEOSS%20Data%20Sharing%20Action%20Plan%20Rev2.pdf.

Goodchild, Michael F. 2007. Citizens as sensors: The world of volunteered geography. GeoJournal 69: 211-21.

Interagency Working Group on Digital Data. 2009. Harnessing the power of digital data for science and society. Washington, D.C.: National Science and Technology Council, Executive Office of the President.

Institute of Museum and Library Services (U.S.). 2003. Report of the workshop on opportunities for research on the creation, management, preservation, and use of digital content. Washington, D.C.: Institute of Museum and Library Services.

Maxwell, J. A. 2005. Qualitative research design: An interpretive approach, 2nd Ed., Vol. 41, Applied Social Research Series. Thousand Oaks, CA: Sage Publications.

National Research Council (U.S.) Committee on Licensing Geographic Data and Services. 2004. Licensing geographic data and services, Washington, D.C.: National Academies Press.

Nielsen, J. 2000. Designing Web usability. Indianapolis: New Riders.

Nielsen, J. 1993. Usability engineering. Boston: Academic Press.

Onwuegbuzie, A. J., and N. L. Leech. 2004. Enhancing the interpretation of "significant" findings: The role of mixed methods research. The Qualitative Report 9(4): 770-92.

Onsrud, H., et al. 2004. Public commons of geographic data: Research and development challenges. In Egenhofer, M. J., C. Freska, and H. J. Miller, Eds. Geographic information science, Berlin: Springer-Verlag.

Onsrud, H., and J. Campbell. 2007. Big opportunities in access to "small science" data. Data Science Journal 6, http://www.jstage.jst.go.jp/browse/dsj/6/0/_contents/2.

Onsrud, H., J. Campbell, and B. van Loenen. 2010. Towards voluntary interoperable open access licenses for the Global Earth Observation System of Systems (GEOSS). International Journal of Spatial Data Infrastructure Research 5: 194-215.

Ragin, C., J. Nagel, and P. White. 2004. Workshop on scientific foundations of qualitative research, Washington: National Science Foundation.

Sui, D. 2008. The Wikification of GIS and its consequences; or Angelina Jolie's new tattoo and the future of GIS. Computers, Environment and Urban Systems 32: 1-5.

Turner, A. 2006. An introduction to neogeography. Sebastapol, CA: O'Reilly Media.

GISCORPS® URISA

GIS Professionals Volunteering for a Better World Since 2003

Operating since 2003 as a program of the Urban and Regional Information Systems Association (URISA), GISCorps coordinates short-term, volunteer-based GIS services to underprivileged communities worldwide. GISCorps supports humanitarian relief, emergency response, health and education, local capacity building, and community development.

To find out how you can get involved— as a volunteer, donor or partner agency— visit www.giscorps.org or email info@giscorps.org.

Upcoming URISA® Education Galore!

March 2-5, 2015
GIS/CAMA Technologies Conference
Oklahoma City, Oklahoma

June 1-3, 2015
2015 CalGIS Conference
Sacramento, California
www.calgis.org

June 22-26, 2015
URISA Leadership Academy
Denver, Colorado

September 1-3, 2015
GIS in Transit Conference
Washington, DC

October 18-22, 2015
GIS-Pro 2015 & NWGIS 2015 Conference
Spokane, Washington

www.urisa.org

Esri® CityEngine®

Create High-Quality 3D Content

Design urban layouts in 3D for analysis and review.

Model 3D environments for entertainment and simulation.

Quickly create 3D models using real-world 2D GIS data.

Download your free 30-day trial at

esri.com/cityenginetrial

esri®

Copyright © 2015 Esri. All rights reserved. Esri, the Esri globe logo, and esri.com are trademarks, registered trademarks, or service marks of Esri in the United States, the European Community, or certain other jurisdictions. CityEngine is a registered trademark of Procedural AG and is distributed under license by Esri. Other companies and products mentioned herein may be trademarks or registered trademarks of their respective trademark owners.