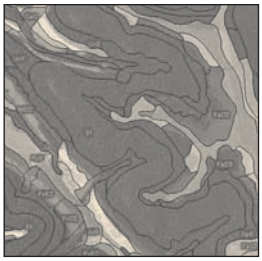
RESEARCH/MANAGEMENT
FINDINGS

"Today, planning and other applied problems do not require soil maps of the traditional type, they generally require specific problem-oriented soil information."

— *European Soil Bureau*



USDA COOPERATIVE SOIL SURVEY

Comparing the Use of STATSGO and SSURGO Soils Data in Water Quality Modeling: A Literature Review

By Adam C. Mednick, Jerry Sullivan, and Dreux J. Watermolen
Bureau of Science Services

INTRODUCTION

The U.S. Department of Agriculture's Natural Resources Conservation Service (NRCS) leads the National Cooperative Soil Survey and is responsible for collecting, storing, maintaining, and distributing soil survey information. As part of this work, the NRCS maintains three soil geographic databases representing the kinds of soil maps produced from different intensities and scales of mapping: the Soil Survey Geographic database (SSURGO), State Soil Geographic database (STATSGO), and National Soil Geographic database (NATSGO).

Soil data comprise a basic input for most hydrologic models. A number of authors have noted that models for predicting runoff, sediment, agricultural chemical yields, and stream flow using STATSGO as opposed to SSURGO can produce different simulation results [1-18]. Additionally, several studies have evaluated the effect of using these different databases on non water-related indicators, such as carbon sequestration [19-22]. To date, however, no comprehensive review has been published on the relative effects of using STATSGO versus SSURGO. Here, we summarize findings from all known studies that have assessed the effects of using STATSGO versus SSURGO as inputs for different water quality-related models. Because the NATSGO database is used primarily for national and multi-state regional analyses, we do not consider its use in this report.

BACKGROUND: THE DATABASES

STATSGO[1] – The NRCS designed the STATSGO database primarily for use in broad-scale resource planning, management, and monitoring efforts (e.g., multistate, statewide, multicounty, and river basin scales) [23]. Soil maps for STATSGO are compiled by generalizing the more detailed SSURGO soil survey maps, using the U.S. Geological Survey's 1:250,000 scale quadrangle map series as base maps. Where more detailed SSURGO maps are unavailable, soil scientists combine data on geology, topography, vegetation, and climate with Land Remote Sensing Satellite (LANDSAT) images and study soils of like areas to determine the probable classification and extent of the soils. They determine map unit composition by transecting or sampling areas on the more detailed maps and expanding the data statistically to characterize the whole map unit [23]. As a result, STATSGO data are not detailed enough to make interpretations at a local or county scale.

SSURGO – The NRCS designed the SSURGO database primarily for smaller scale applications (e.g., parcel, township, or county scale) and it includes the most detailed level of available information [24]. SSURGO maps are made at scales ranging from 1:12,000 to 1:63,360 (those for Wisconsin are typically at the 1:12,000 scale). To generate the SSURGO maps, soil scientists use aerial photographs as base maps. Surveyors observe soils along delineation boundaries and use field traverses and transects to determine map unit composition. The resulting database can be used for "determining erodible areas and developing erosion control practices, reviewing site development proposals and land use potential, making land use assessments and chemical fate assessments, and identifying potential wetlands and sand and gravel aquifer areas" [24].

[1] *The NRCS recently revised and updated the STATSGO spatial and tabular data. STATSGO has been renamed as the U.S. General Soil Map (STATSGO2). These data are available for download from the NRCS Soil Data Mart at* **http://soildatamart.nrcs.usda.gov/USDGSM.aspx**.

Comparison of Databases – For our purposes, the main difference between STATSGO and SSURGO data is the number of different kinds of soils, or "components," that a map unit can contain. In Figure 1 each polygon represents a soil map unit, named for the dominant soil or soils present in the delineated area. Components representing separate soils with distinct, but relatively similar, properties are grouped together for simplicity's sake when characterizing these soil map units. The detailed SSURGO map units (brown polygons) can contain no more than three components, and may be comprised of only a single component [24]. Aggregated STATSGO map units (black polygons), on the other hand, can contain up to twenty-one different component soils [23].

To date, relatively little has been published in the peer-reviewed literature on the effect of spatial scale (e.g., 1:250,000 vs. 1:12,000 or 1:24,000) on water quality model predictions. Understanding the effects of differing soil databases on model outputs, however, has implications for the reliability of the models in land planning and storm water regulation. Further, the amount of time and computing resources needed to acquire, process, and use the more complex SSURGO data may be unnecessary. On the other hand, if the models are sensitive to the scale of soils data, it may be worth the modelers' extra effort to use the more detailed data.

REVIEW OF THE LITERATURE

As summarized in Table 1, we focused on studies that include an evaluation of the effect of soil database selection on model-predicted runoff, stream flow, and/or various water quality parameters. Several studies not included in this table set the context by evaluating differences in soil properties as directly reported in STATSGO versus SSURGO. Lin et al [25] measured the "purity" of STATSGO map units relative to underlying SSURGO soils (i.e., the areal percentage of the STATSGO map units matched by SSURGO values). This conformity measure ranged from 90% for hydrologic group down to 60% for soil texture. Murray and McCray [26, 27] conducted similar analyses and found that STATSGO values are generally higher than corresponding spatially-weighted SSURGO values for available water capacity, bulk density, percent organic matter, and hydraulic conductivity. Juracek and Wolock [28] conducted a statewide comparison in Kansas and found that mean values for soil permeability, clay content, and hydrologic group were increasingly dissimilar between the two databases as evaluation areas become smaller and closer to stream networks. Finally, Rosenblatt et al. [29] estimated the water nitrate removal capacity of riparian zones using both databases and found that STATSGO failed to display critical soil features along streams and therefore misrepresented removal capacity, particularly around lower-order streams.

Seven of the 18 studies summarized in Table 1 indicate that using SSURGO in place of STATSGO improves the accuracy of model predictions [1, 2, 10, 14-16, 18]. Of these, the most conclusive are Anderson et al. [1], which evaluated the lumped parameter Sacramento–Soil Moisture Accounting (SAC-SMA) model, and Zhang et al. [18], which evaluated the distributed parameter Hydrology Laboratory Research Modeling System (HLRMS). Peschel et al. [14] and Wang et al. [15] ran the Soil and Water Assessment Tool (SWAT) to predict stream flow in two similarly sized watersheds in different parts of the country and found that the model performed slightly better overall using SSURGO. Lathrop et al. [10] concluded that the use of STATSGO data led to the significant under-prediction of evapotranspiration in a northwestern New Jersey County, although this is only an intermediate variable to runoff and stream flow.

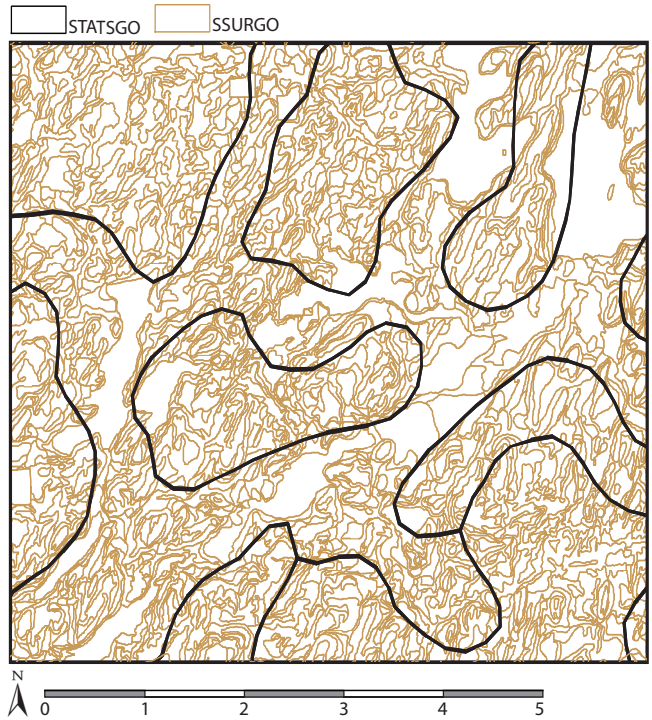


Figure 1. *STATSGO versus SSURGO Map Units: Cottage Grove Township, Dane County, Wisconsin*

One study, Heathman and Larose [9], found that the use of STATSGO data improved model performance. The authors ran the SWAT model for a large watershed in Indiana and found that stream flow predictions were more accurate using STATSGO. The remaining ten studies [3-8, 11-13, 17] are either inconclusive or show mixed results; e.g., one database is more accurate for certain response variables, but less accurate for others.

DISCUSSION

On balance, the 18 studies summarized in Table 1 seem to confirm the *a priori* assumption that higher resolution soils data are more often than not preferable for modeling hydrologic and water quality parameters, from an accuracy standpoint. That said, the available findings are far from unanimous and reveal no clear pattern with respect to response variable, model type, units of analysis, or scale of analysis, as these factors relate to the variation in model predictions using STATSGO versus SSURGO. A potentially confounding issue is the fact that some authors calibrate their STATSGO and SSURGO-based models separately before comparison [6, 7, 15]. Others avoid doing so for the express purpose of maintaining a clear signal of soil data resolution effects in their comparisons [5, 11-14]. Yet, pairs of studies with the same models—calibrated and un-calibrated—report generally similar results [e.g., 5, 6, 14, 15].

A more likely cause for this lack of an explanatory pattern is the small sample size within and across the different studies. Fourteen of the 18 studies relied on a single watershed or study area. The others ranged from 3 to 11 watersheds. Sensitivity analyses with larger sample sizes are needed, but many of the models are both data and time-intensive and do not lend themselves to widespread replication. A forthcoming study from our team will evaluate the sensitivity to soil data resolution of a simple but spatially distributed rainfall/runoff model run for 300 of the 312 watersheds that fall completely within Wisconsin, and the nearly 1,600 subwatersheds nested therein.

Table 1. *Studies comparing the use of STATSGO versus SSURGO in water quality modeling.*

Reference	Response Variable(s)	Model(s)	Study Area(s)	State(s)	Analysis Extent	Level of Aggregation[1]	Key Findings
Anderson et al. [1]	Daily and Monthly Surface Runoff Volume	SAC-SMA[2]	6 Watersheds	KY, OH, WV	145 to 590 km^2	Lumped	SSURGO-based parameters improve the accuracy of model predictions.
Bolstad et al. [2]	Sediment Yield	WCS-SED[3]	5 Watersheds	NC	1 to 7 km^2	Distributed	Both over predict, but SSURGO less so; consistently 10-30% lower than STATSGO.
Breiby [3]	Soil Loss	GIS-RUSLE[4]	Zumbro River Subwatershed	MN	29 km^2	Distributed	STATSGO- and SSURGO-based results similar ($R^2 > 0.91$). Less so at very high levels of estimated soil loss.
Di Luzio et al. [4]	Monthly Stream Flow, Sediment and Pollutant Loadings	SWAT[5]	Scarborough Creek Watershed	TX	<10 km^2	Partially-Distributed	Little difference in flow or sediment. SSURGO-based nitrogen considerably higher; phosphorous higher; nitrates lower.
Geza & McCray [5]	Mean Daily Stream Flow and Pollutant Loadings	SWAT	Turkey Creek Watershed	CO	126 km^2	Partially-Distributed	STATSGO-based flow more accurate before calibration; SSURGO after. SSURGO-based sediment and sediment-attached nutrients lower; dissolved nutrients higher.
Ghidey et al. [6]	Daily, Monthly, and Annual Stream Flow; and Sediment Yield	SWAT	Long Branch Watershed	MO	462 km^2	Partially-Distributed	STATSGO- and SSURGO-based flow comparable/good. Both under predicted sediment yield over the calibration period and over predicted over the validation period.
Gowda & Mulla [7]	Monthly and Annual Stream Flow; Sediment, Nitrogen, and Phosphorus Loss	ADAPT[6]	High Island Creek Watershed	MN	39 km^2	Distributed	STATSGO- and SSURGO-based predictions were comparable/good. Under alternative management scenarios, STATSGO-based predictions of nitrate loss were consistently higher than SSURGO; vice-versa for phosphorus.
Grove et al. [8]	Direct Runoff Depth	L-THIA[7]	Little Eagle Creek Watershed	IN	71 km^2	Distributed	STATSGO predictions are lower, but difference is minimal.
Heathman & Larose [9]	Monthly and Annual Stream Flow	SWAT	Cedar Creek Experimental Watershed	IN	707 km^2	Partially-Distributed	STATSGO-based models had lowest errors.
Lathrop et al. [10]	Evapotranspiration and Net Primary Productivity	PnET[8]	Hunterdon County	NJ	350 mi^2	Lumped	STATSGO appears to under predict. SSURGO models predicted +19% evapotranspiration and +20% net primary productivity.

Levick et al. [11]	Surface Runoff from individual storm-events	AGWA -KINEROS2[9]	3 Watersheds	AZ	8 to 114 km^2	Partially-Distributed	For two small watersheds, both predictions similar, accurate. For large watershed, both significantly under predict; SSURGO more so.
Mukundan [12]	Daily Stream Flow and Sediment Load	SWAT	North Fork Broad River Watershed	GA	180 km^2	Partially-Distributed	SSURGO data improved model predictions only slightly and differentials were not statistically significant.
Peschel et al. [13]	Daily Water Yield	SWAT	Leon Creek Watershed	TX	536 km^2	Partially-Distributed	SSURGO-based water yield considerably higher in 3 of 5 sub-watersheds, slightly higher in one, and nearly identical to STATSGO in the fifth. Size of the differential appears related to land cover and impervious surface.
Peschel et al. [14]	Daily and Monthly Stream Flow	SWAT	Upper Sabinal Watershed	TX	541 km^2	Partially-Distributed	SSURGO-based flow more closely correlated with gauge. SSURGO runoff lower, but water yield higher. Caused by higher percolation and groundwater return.
Wang & Melesse [15]	Monthly, Seasonal, and Annual Stream Flow	SWAT	Upper Portion of Elm River Watershed	ND	515 km^2	Partially-Distributed	SSURGO slightly better predictions overall. Both did a comparable/good job predicting high stream flows. STATSGO more accurate for low flows and slightly better over the validation period.
Williamson & Odom [16]	Stream Flow	TOPMODEL[10]	South Fork Kentucky River Basin	KY	1,938 km^2	Distributed	Combination of SSURGO soils and high resolution (10-m) elevation data produced most accurate results.
Wilson et al. [17]	Depth of Picloram Movement	CMLS[11]	Teton County	MT	320 km^2	Distributed	STATSGO- and SSURGO-based predictions significantly different.
Zhang et al. [18]	Hourly Stream Flow	HLRMS[12]	11 Watersheds	AR, OK	37 to 2,484 km^2	Distributed	SSURGO-based parameters improve the accuracy of model predictions.

1. Level of Aggregation refers to the spatial units of analysis used in the models. "Distributed" models estimate parameters for individual pixels or hydrologic response units (e.g., unique combinations of soil and land cover classes). "Lumped" models estimate parameters for the entire analysis extent (e.g., a watershed). "Partially-Distributed" models estimate parameters for intermediate spatial units (e.g., sub-watersheds or catchments).
2. Sacramento–Soil Moisture Accounting model
3. Watershed Characterization System–Sediment Tool
4. Geographic Information System application of the Revised Universal Soil Loss Equation
5. Soil and Water Assessment Tool
6. Agricultural Drainage and Pesticide Transport model
7. Long-term Hydrologic Impact Assessment tool
8. Photosynthetic EvapoTranspiration model
9. Automated Geospatial Watershed Assessment tool–Kinematic Runoff and Erosion model v.2
10. Topography based hydrologic model
11. Chemical Movement through Layered Soils model
12. Hydrology Laboratory Research Modeling System

LITERATURE CITED

1. Anderson, R.M., V.I. Koren, and S.M. Reed. 2006. Using SSURGO data to improve Sacramento Model *a priori* parameter estimates. *Journal of Hydrology* 320(1-2):103-116.

2. Bolstad, P., A. Jenks, M. Reidel, and J.M. Vosc. 2006. Estimating sediment yield in the southern Appalachians using WCS-SED. Pp. 136-143 In Proceedings of the Eighth Federal Interagency Sedimentation Conference. Reno, NV. Available online at **http://pubs.usgs.gov/misc_reports/FISC_1947-2006/pdf/1st-7thFISCs-CD/8thFISC/Session%2011A-1_Bolstad.pdf**.

3. Breiby, T. 2006. Assessment of soil erosion risk within a subwatershed using GIS and RUSLE with a comparative analysis of the use of STATSGO and SSURGO soil databases. Unpubl. Paper. Dept. of Resource Analysis, St. Mary's University, Winona, MN. 22 pp. Available online at **http://www2.smumn.edu/ra/gis/GradProjects/BreibyT.pdf**.

4. Di Luzio, M., J.G. Arnold, and R. Srinivasan. 2004. Integration of SSURGO maps and soil parameters within a geographic information system and nonpoint source pollution model system. *Journal of Soil and Water Conservation* 59(4):123-133.

5. Geza, M. and J.E. McCray. 2007. Effects of soil data resolution on SWAT model stream flow and water quality predictions. *Journal of Environmental Management* 88(3):393-406.

6. Ghidey, F., E.J. Sadler, R.N. Lerch, and C. Baffaut. 2007. Scaling up the SWAT model from Goodwater Creek Experimental Watershed to the Long Branch Watershed. In 2007 ASABE Annual International Meeting. Minneapolis, MN. Available online at **http://agspace.nal.usda.gov:8080/bitstream/10113/8911/1/IND43994364.pdf**.

7. Gowda, P.H. and D.J. Mulla. 2005. Scale effects of STATSGO VS. SSURGO soil databases on water quality predictions. Pp. 579-587 In P.W. Gassman (ed.). Watershed Management to Meet Water Quality Standards and Emerging TMDL (Total Maximum Daily Load): Proceedings of the Third Conference. Atlanta, GA.

8. Grove, M., J. Harbor, B.A. Engel, and S. Muthukrishnan. 2001. Impacts of urbanization on surface hydrology, Little Eagle Creek, Indiana, and analysis of L-THIA model sensitivity to data resolution. *Physical Geography* 22(2):135-153.

9. Heathman, G.C. and M. Larose. 2006. Modeling streamflow using SWAT with different soil and land cover geospatial datasets. In H.Q. Tian (ed.). Proceedings of the Second IASTED International Conference on Environmental Modeling and Simulation. St. Thomas, U.S. Virgin Islands. Abstract available online at **http://www.ars.usda.gov/research/publications/publications.htm?seq_no_115=204907**.

10. Lathrop, R.G., Jr., J.D. Aber, and J.A. Bognar. 1995. Spatial variability of digital soil maps and its impact on regional ecosystem modeling. *Ecological Modeling* 82(1):1-10.

11. Levick, L., D.P. Guertin, S. Scott, D.J. Semmens, D.C. Goodrich. 2006. Automated geospatial watershed assessment tool (AGWA): Uncertainty analysis of common input data. In Third Federal Interagency Hydrologic Modeling Conference. 2006. Reno, NV. Available online at **http://acwi.gov/hydrology/mtsconfwkshops/conf_proceedings/3rdFIHMC/9D_Levick.pdf**.

12. Mukundan, R. 2008. Effect of soil data resolution on hydrologic modeling in a Southern Piedmont watershed. In GSA/SSSA-ASA-CSSA/GCAGS Joint Annual Meeting. Houston, TX. Abstract available online at **http://a-c-s.confex.com/crops/2008am/webprogram/Paper45902.html**.

13. Peschel, J.M., P.K. Haan, and R.E. Lacey. 2003. A SSURGO pre-processing extension for the ArcView Soil and Water Assessment Tool. Paper No. 032123. ASAE Annual International Meeting. Las Vegas, NV. Available online at **http://lcluc.tamu.edu/ssurgo/2003.ASAE.032123.paper.pdf**.

14. Peschel, J.M., P.K. Haan, and R.E. Lacey. 2006. Influences of soil dataset resolution on hydrologic modeling. *Journal of the American Water Resources Association* 42(5):1371-1389.

15. Wang, X. and A.M. Melesse. 2006. Effects of STATSGO and SSURGO as inputs on SWAT model's snowmelt simulation. *Journal of the American Water Resources Association* 42(5):1217-1236.

16. Williamson, T. and K.R. Odom. 2007. Implications of SSURGO vs. STATSGO data for modeling daily streamflow in Kentucky. In ASA-CSSA-SSSA 2007 International Annual Meetings. New Orleans, LA. Abstract available online at **http://a-c-s.confex.com/crops/2007am/techprogram/P32289.htm**.

17. Wilson, J.P., W.P. Inskeep, J.M. Wraith, and R.D. Snyder. 1996. GIS-based solute transport modeling applications: Scale effects of soil and climate data input. *Journal of Environmental Quality* 25(3):445-453.

18. Zhang, Z., V. Koren, S. Reed, M. Smith, F. Moreda. 2006. Comparison of simulation results using SSURGO-based and STATSGO-based parameters in a distributed hydrologic model. In Third Federal Interagency Hydrologic Modeling Conference. Reno, NV. Available online at **http://acwi.gov/hydrology/mtsconfwkshops/conf_proceedings/3rdFIHMC/8D_ZhangZi.pdf**.

19. Davidson, E.A. and P.A. Lefebvre. 1993. Estimating regional carbon stocks and spatially covarying edaphic factors using soil maps at three scales. *Biogeochemistry* 22(2):107-131.

20. Hunt, E.R., P.C. Doraiswamy, C.S. Daughtry, G.W. McCarty, J.L. Hatfield, and R.C. Izaurralde. 2005. Simulation of erosion and soil carbon sequestration over an agricultural landscape in Iowa. In Third USDA Symposium on Greenhouse Gases and Carbon. Baltimore, Maryland. Abstract available online at **http://www.ars.usda.gov/research/publications/Publications.htm?seq_no_115=177575**.

21. Rasmussen, C. 2006. Distribution of soil organic and inorganic carbon pools by biome and soil taxa in Arizona. *Soil Science Society of America Journal* 70(1):256-265.

(continued on next page)

Bureau of Science Services
Wisconsin Department of Natural Resources
P.O. Box 7921
Madison, WI 53707

Presorted Standard
U.S. Postage
Paid
Madison, WI
Permit No. 906

Printed on recycled paper.

PUB-SS-760 2008

Literature Cited *(continued)*

22. Wang, F., Y. Xu, and B. Zhong. 2005. Scales and uncertainties in estimating soil carbon with STATSGO database for Louisiana. In American Geophysical Union, Spring Meeting. Boston, MA. Abstract available online at **http://adsabs.harvard.edu/abs/2005AGUSM.H43B..03W**.

23. USDA. 1994. *State Soil Geographic (STATSGO) Database: Data Use and Information*. Miscellaneous Publ. No. 1492. National Soil Survey Center, Natural Resources Conservation Service. Washington, DC. Available online at **http://www.nrcs.usda.gov/technical/techtools/statsgo_db.pdf**.

24. USDA. 1995. *Soil Survey Geographic (SSURGO) Data Base: Data Use and Information*. Miscellaneous Publ. No. 1527. National Soil Survey Center, Natural Resources Conservation Service. Washington, DC. Available online at **http://tahoe.usgs.gov/files/ssurgo_database.pdf**.

25. Lin, H., D. Wheeler, J. Bell, and L. Wilding. 2005. Assessment of soil variability at multiple scales. *Ecological Modelling* 182(3-4):271-290.

26. Murray, K.E. and J.E. McCray. 2002. Statistical evaluation of uncertainty in small-scale GIS-based soil property databases. In Rocky Mountain 54th Annual Meeting, Geological Society of America. Cedar City, UT. Abstract available online at **http://gsa.confex.com/gsa/2002RM/finalprogram/abstract_34081.htm**.

27. Murray, K.E. and J.E. McCray. (no date). Comparison of Hydrologic Properties Represented by Two Scales of Digital Soil Databases. Unpubl. Manuscript. Dept. of Earth and Environmental Science, Univ. of Texas at San Antonio, San Antonio. 27 pp.

28. Juracek, K.E. and D.M. Wolock. 2002. Spatial and statistical differences between 1:250,000- and 1:24,000-scale digital soil databases. *Journal of Soil and Water Conservation* 57(2):89-94.

29. Rosenblatt, A.E., A.J. Gold, M.H. Stolt, P.M. Groffman, and D.Q. Kellogg. 2001. Identifying riparian sinks for watershed nitrate using soil surveys. *Journal of Environmental Quality* 30(5):1596-1604.

ACKNOWLEDGMENTS

The U.S. Environmental Protection Agency's Office of Wetlands, Oceans, and Watersheds (Washington) and Region 5 Headquarters (Chicago) provided funding for our efforts through Grant No. C6–96539401-0. Points of view expressed in this report do not necessarily reflect the views or policies of the U.S. Environmental Protection Agency. Mention of trade names and commercial products does not constitute endorsement of their use. Our manuscript benefited from a prepublication review by Sally J. Kefer, Bureau of Science Services. We appreciate her insights and contributions.

Graphic Design: Michelle Voss

ABOUT THE AUTHORS

Adam C. Mednick, AICP, is a Natural Resources Educator in the Bureau of Science Services and a PhD candidate in Urban and Regional Planning at the University of Wisconsin-Madison. **Jerry Sullivan** is a GIS project manager for the Bureau of Science Services. **Dreux J. Watermolen** is an ecologist and the chief of the Science Information Services section.

AUTHOR CONTACT INFORMATION

Address:
Wisconsin DNR, SS/7
Bureau of Science Services
PO Box 7921, Madison, WI 53707-7921

E-mails:
Adam.Mednick@Wisconsin.gov
Jerry.Sullivan@Wisconsin.gov
Dreux.Watermolen@Wisconsin.gov