

CLOUD COMPUTING, WEB-BASED GIS, TERRAIN ANALYSIS, DATA FUSION, AND MULTIVARIATE STATISTICS FOR PRECISION CONSERVATION IN THE 21ST CENTURY

Tom Mueller, Surendran Neelakantan, Eduardo Rienzi, and Blazan Mijatovic

*Plant and Soil Science Department
University of Kentucky
Lexington, Kentucky*

Annamaria Castrignano

*Agricultural Research Council - Research
Unit for Cropping Systems in Dry Environments
Bari, Italy*

Adam Pike

*LiDAR group
PhotoScience
Lexington, Kentucky*

Cody Bumgardner

*Enterprise Architecture
University of Kentucky
Lexington, Kentucky*

INTRODUCTION

Precision conservation involves utilizing a set of tools in agriculture that can help meet the great challenges of this century including a growing world population, increasing demand for food and biofuels, global climate change, and the global water crisis. Precision conservation seeks to efficiently target conservation practices at specific locations within landscapes where they can help maintain or improve soil quality, increase crop productivity, and protect water and air resources. The objective of this paper is to discuss how cutting edge techniques (e.g., cloud computing, geospatial analysis, environmental modeling, and precision technologies) could be used together to help facilitate more effective precision conservation management.

TECHNOLOGY ENABLED PRECISION CONSERVATION

Wireless Technologies and Cloud Computing

Equipment manufacturers are designing tractors and combines with wireless communication, allowing transmission of yield monitor and sensor data to “the cloud” where farmers, consultants, and researchers can access information for real-time decision making. The concept of “the cloud” developed when companies like Amazon had extra server capacity and they realized they could “time sharing” their server resources when they were not in use (e.g., at night). One advantageous characteristic of cloud applications is that they are allocated hardware resources from the cloud as software requests increase. Resources are released when they are no longer needed. This effectively maximizes performance while minimizing costs. Load balancers constantly distribute and redistribute Internet traffic across multiple servers thus optimizing performance. Cloud computing techniques could substantially reduce the infrastructure costs required to perform management and conservation planning in agriculture today because contracting national budgets for government conservation programs will require more efficient use of resources. We have developed a sample cloud computing application for soil survey data for an MLRA-121 (located in parts of Kentucky, Indiana, and Ohio) that can be found at go.uky.edu/LandUse. The architecture for this website is shown in Figure 1.

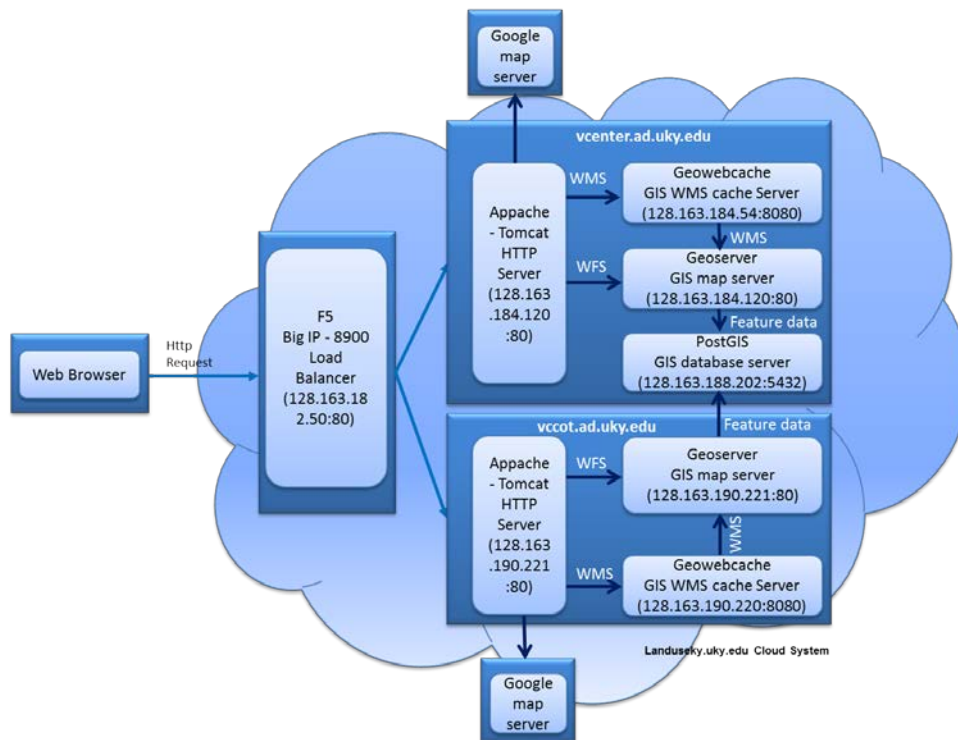


Figure 1. Cloud computing architecture for a Land Use and Management website.

Data Fusion and Multivariate Statistical Analyses

Data fusion techniques will be helpful for fast, efficient, accurate, and automatic data analysis for conservation planning because data are collected at a wide variety of spatial and temporal scales. For example, LiDAR generally does not require repeated measurements and are usually provided at a resolution of less than 1 m. Yield is measured once a growing season and for a 12-row combine travelling at 5 mph harvesting corn planted on 2.5-ft centers, yield data will have a resolution of approximately 7 ft in the direction of travel and 30 ft between passes. Soil EC_a is generally only needs to be measured once and this occurs in the spring or fall. The spatial resolution of the EC survey data depends on ground speed and the distance between passes (generally 40 and 60 feet). Soil samples are often collected in the fall or spring, every 2 to 4 years with a grid sampling (e.g. 330 foot spacing) or with a regular or irregular zone sampling approach (e.g., 5 to 20 acre zones). Landsat 7 imagery has a resolution of about 100-ft in the visible and near infrared ranges and 200-ft in the thermal infrared (TIR) range. Soil survey data is available for many areas within the US at scales between 1:12,000 and 1:26,000, but the intensity of ground soil profile observations varied substantially and depended on many factors.

These scale mismatches can be addressed with data fusion techniques such as block interpolation (e.g., kriging) and multi-located cokriging. Multi-located factorial co-kriging is a data fusion technique that also allows efficient synthesis of many variables in ways that would allow the creation of management zones for soil and water conservation (Castrignano et al., 2009). This procedure was conducted for a field in Central Kentucky (Figure 2). Input data included LiDAR derived terrain attributes, soil EC data, and grid soil sample data. The first factor (F1) at short range (140 m) and the first factor (F1) at long range (500 m) were the only factors that were statistically significant. The map of F1 (140-m) mostly summarized the distribution of EC and Mg and Ca, with the highest values on the lower parts of the field, whereas the map of F1 (500-m) is characterized by a wide middle area with low values corresponding to ridge tops with higher contents of phosphorus. Isofrequency maps could be created for the short and long range F1 maps and the delineation in homogeneous zones could be used for site-specific management and could serve as the basis for sampling for conservation soil fertility management. Potentially, multi-located factorial co-kriging, using as finer-scale auxiliary variable a terrain attribute or remote/proximal sensing image, could be conducted in the cloud with GSLib, the source code for which is freely available.

Terrain Analysis for Soil and Water Conservation

Terrain analysis can be used to identify environmentally sensitive areas within landscapes. Several studies have used terrain attributes to determine the locations where concentrated flow erosion (e.g., ephemeral gully erosion) is likely to occur (Thorne et al., 1986; Moore et al., 1988; Srivastava and Moore, 1989; Berry et al., 2005). These studies overlaid maps depicting the locations where eroded channels were observed in the field on the top of terrain attribute maps. The authors of

these studies visually determined terrain attribute thresholds above which concentrated flow erosion was likely to occur. The practical problems that conservation planners would have employing this type of approach over large areas are that it 1) requires site-specific thresholds, 2) would be very time consuming, 3) requires a high level of GIS knowledge, and 4) does not easily lend itself to automation.

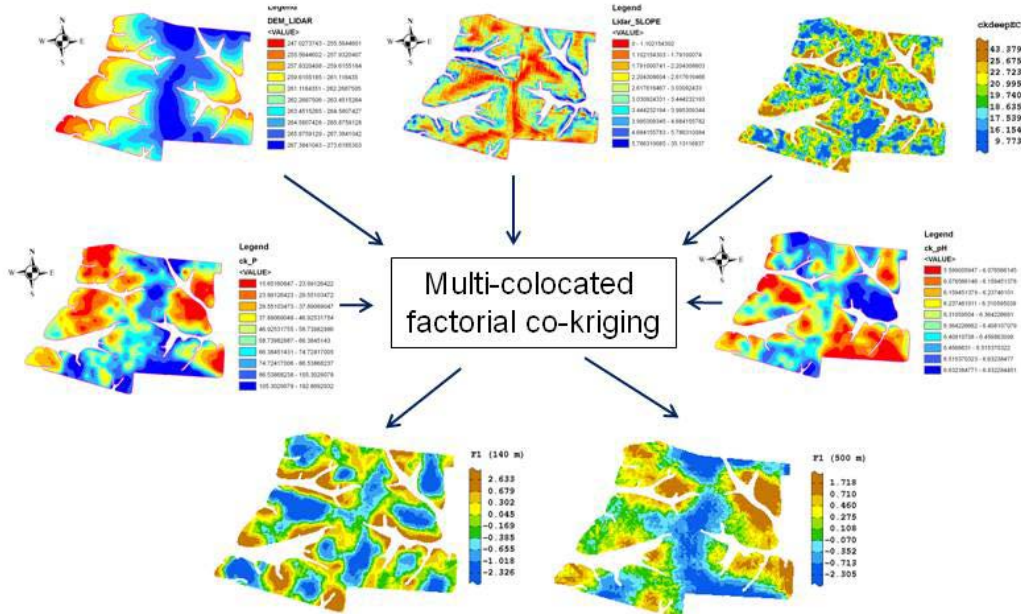


Fig. 2. Multivariate co-located factorial co-kriging analysis to summarize nutrient, soil EC_a, and LiDAR derived terrain attributes maps created at different scales.

To overcome these problems, we developed a multivariate statistical approach to determine where concentrated flow erosion was likely to occur in agricultural fields, based on terrain attributes (Pike et al., 2009). We compared logistic regression and artificial neural network analysis as prediction methods and found that both procedures performed very well as determined with a leave-one-field-out validation procedure. Next, we evaluated how these procedures performed with 10-m US Geological Survey (USGS) digital elevation models (DEMs; freely available on the Internet for most of the United States) with those made from elevation measurements created using Real-Time Kinematic (RTK) survey-grade Global Positioning System (GPS) (Pike et al., 2010). Although we found that RTK predictions were superior to those made with the USGS data. USGS predictions would be adequate for many agricultural applications. For example, most of the waterways that are visually apparent in Figure 3a could be predicted by the 10-m USGS map (Figure 3b), but were more clearly delineated by the RTK data (Figure 3c). In another study, we found that the model developed for the farm in Shelby County, Kentucky (Pike et al., 2010) also worked well 80 miles away in Hopkins County, where soils differed substantially in properties including texture (Luck et al., 2010). We have observed that the waterway prediction models also work well in other Kentucky counties.

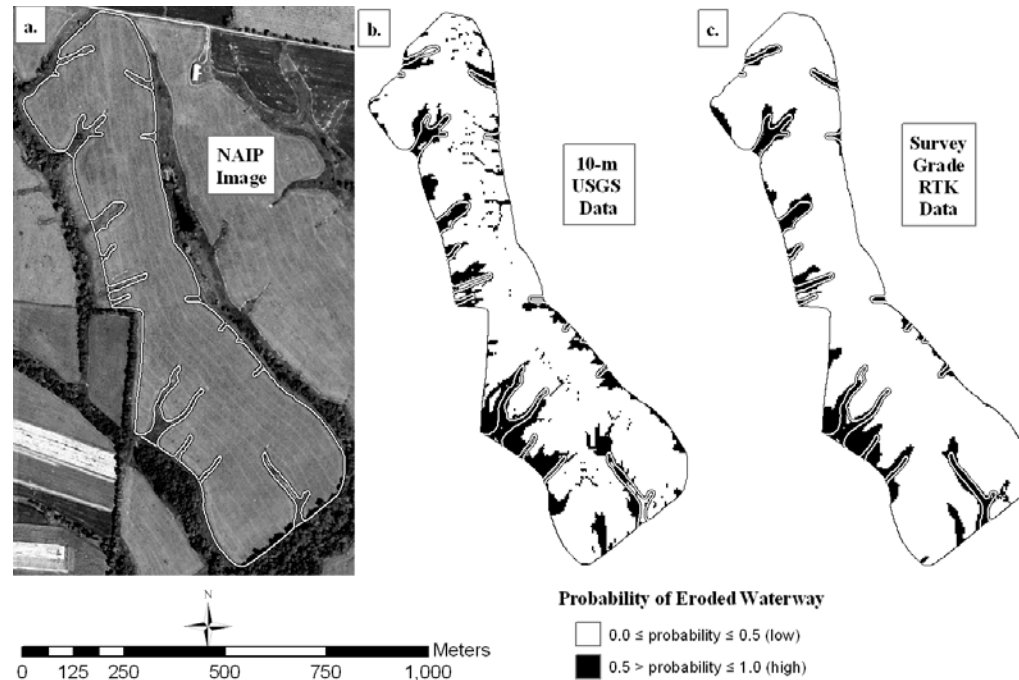


Figure 3. The outlines of the locations of observed waterways overlaid on a) an aerial photograph, b) USGS and c) Survey grade GPS derived maps of the probability of the occurrence of eroded waterways (from Pike et al., 2009, 2010). The probability analysis is from a leave-one-field out validation analysis. The logistic regression equations to calculate the likelihood of erosion were and

$$\text{Probability of Erosion}_{\text{USGS}} = \frac{1}{1 + e^{-(7.55 + 0.94 \times \text{LS} + 0.84 \times \text{WET})}}$$

$$\text{Probability of Erosion}_{\text{RTK}} = \frac{1}{1 + e^{-(12.53 + 2.25 \times \text{LS} + 1.12 \times \text{WET})}}$$

where LS and WET are both terrain attributes and LS is the estimated length slope factor from the USLE equation and WET is the topographic wetness index.

These conservation analyses should be available to conservation planners on hand held computing devices in the field. Cloud computing technologies could be used to distribute data efficiently by adding ArcGIS servers to cloud systems similar to the one depicted in Figure 1. Elevation data could be pulled directly from a national USGS server or a local database.

Processed Based Environmental Models

Environmental models (e.g., SWAT, WEPP, RUSLE2, and ANNAGNPS) could provide valuable information for assessing conservation practices in the field. Unfortunately, these models may be too complex and cumbersome to be operated in real time in the cloud in the field by conservation planners on mobile computing devices. An alternative approach would be to adapt environmental

models for particular watersheds or regions and then to develop simple relationships to predict the environmental impacts of conservation practices which would be more suited to for rapid decision making in the field via the internet. For example Dosskey et al. (2011) developed a design aid for sizing filter strips based on the process-based Vegetative Filter Strip Model.

CONCLUSIONS

There is a new wave of the information technology revolution in agriculture today including web-based mapping and GIS, cloud computing, and wireless technologies. Conservation professionals could utilize these technologies as environmental and agricultural modeling tools to improve the efficacy and economics of conservation planning in a way that will help meet the challenges of the 21st century.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge support from the NRCS, Kentucky SB-271 Water Quality program, and the University of Kentucky Division of Information Technology. We would personally like to thank Steve Workman, Steve Crabtree, Steve Blanford, Paul Finnell, Jon Hempel, Doyle Friskney, Victoria Banks, Nick McClure and Carey Johnson.

REFERENCES

- Berry, J.D., F.J. Pierce, and R. Khosla. 2005. Applying spatial analysis for precision conservation across the landscape. *Journal of Soil and Water Conservation* 60(6):363-370.
- Castrignano, A., Costantini, E., Barbetti, R., Sollitto, D. (2009). Accounting for extensive topographic and pedologic secondary information to improve soil mapping. *CATENA*, vol. 77; p. 28-38.
- Dosskey, M.G., M.J. Helmers, and D.E. Eisenhauer. 2011. A design aid for sizing filter strips using buffer area ratio. *Journal of Soil and Water Conservation*. 66:1-29.
- Luck, J.D., T.G. Mueller, S.A. Shearer, and A.C. Pike. 2010. Grassed waterway planning model evaluated for agricultural fields in the western coal field physiographic region of Kentucky. *Journal of Soil and Water Conservation* 65:280-288.
- Moore, I.D., G.J. Burch, and D.H. Mackenzie. 1988. Topographic effects on the distribution of surface soil water and the location of ephemeral gullies. *Transactions of the American Society of Agricultural Engineers* 31(4):1098-1119.
- Neelakantan, S., T.G. Mueller, B. Lee, B. Lee, P. Finnell, V. Bumgardner, and D. Carey. 2011. Web 2.0 spatial data browser for visualizing land-use assessment information from soil surveys. *JSWCS*. 66:37A-39A.
- Srivastava, K.P., and I.D. Moore. 1989. Application of terrain analysis to land resource investigations of small catchments in the Caribbean. In *Proceedings of the 20th International Conference of the Erosion Control*

Association, Vancouver, BC, Canada: 229-249, International Erosion Control Association, Steamboat Springs, CO.

- Pike, A.C., T.G. Mueller, A. Schörgendorfer, J.D. Luck, S.A. Shearer, and A.D. Karathanasis. 2009. Erosion index derived from terrain attributes using logistic regression and neural networks. *Agronomy Journal* 101:1068-1079.
- Pike, A.C., T.G. Mueller, A. Schörgendorfer, J.D. Luck, S.A. Shearer, A.D. Karathanasis. 2010. Eroded Waterways with USGS Elevation Data. *Agronomy Journal*. 102:1269-1273.
- Thorne, C.R., L.W. Zezenbergen, E.H. Grissinger, and J.B. Murphey. 1986. Ephemeral gullies as sources of sediment. In *Proceedings of the 4th Federal Interagency Sedimentation Conference, Las Vegas, Nevada, Vol. 1:301-309*. Washington, DC, Government Printing Office.