Project Title

Assessing Local Uncertainty in Non-Stationary Scale-Variant Geospatial Data

Principal Investigator

Susan Colarullo, Hydrologist
New Jersey Water Science Center
U.S. Geological Survey
810 Bear Tavern Rd., Suite 206
W. Trenton, New Jersey 08628
(609) 771-3922
colarull@usgs.gov

Primary Contact

Susan Colarullo

Project Description

Background

Watershed managers rely on geospatial datasets depicting land use, hydrography, soil and geologic properties, vegetative cover, elevation, contaminant sources, and other distributed processes to make informed decisions about the allocation of scarce water resources between competing demands of human growth and aquatic ecosystem integrity. These spatial datasets, however, are imperfect. They differ from real-world distributions that they represent because of uncertainties inherent to data collection. Since watershed management decisions inherit the accuracy of geospatial data used to inform them, the reliability of watershed management decisions is only as good as the accuracy of these datasets.

Producers of GIS spatial datasets in the United States are required by Federal Geographic Data Committee (FGDC) metadata standards to report estimates of error (SDTS Task Force, 1996). Elements of the Data Quality section of the FGDC Content Standard for Digital Metadata (CSDGM) include positional and attribute error. Both sources of error contribute to data uncertainty, often in ways that are difficult to distinguish from one another. Error is typically reported using summary statistics derived by comparing ground measurements with dataset values for a subset of the data. This approach treats uncertainty as if it were the same at every location, or as a global phenomenon. In reality, uncertainty in spatially-distributed data is local in nature, varying from one location to another depending on the unique characteristics of the data and on the methods used to measure it or simulate it using deterministic rainfall-runoff models (Heuvelink and others, 1989; Openshaw, 1989; Fisher, 1991; Lee and others, 1992; Englund, 1993; Goodchild, 1995; Unwin, 1995; Burrough and others, 1996; Ehlschlaeger and others, 1997).

Local variations in uncertainty of distributed data occur for a number of reasons. Uncertainty close to measurement locations will tend to be smaller, increasing with greater distance from them, at a rate dictated by the correlation scale of the data. It will be larger in regions where the data is most spatially variable and smaller in areas where it is not. Further compounding the issue of local uncertainty is its tendency to vary with spatial scale, with higher levels of uncertainty typically associated with measurements collected over larger scales of spatial resolution. This dependence of dataset uncertainty on both location and scale should be incorporated into any framework used to assess the uncertainty of data critical to making watershed management decisions.

If spatial variations in the distributed data demonstrate approximate periodicity over space, the data can be treated as a gaussian random process with uniform mean, variance, and correlation scale over space. This property, referred to as second-order stationarity, is the foundation of traditional geostatistical methods used for assessing data uncertainty. Figure 1 shows an example of a stationary gaussian process in the time domain, but the concept is the same when analyzing processes distributed in space. When a process is stationary, the statistics of the underlying joint distribution associated with the process are assumed global, and equally valid at every point in space.

Many traditional geostatistical techniques rely on techniques of Fourier analysis to bring the dataset into the frequency domain, treating the stationary, scale-invariant dataset as the sum of an infinite number of shifted sine basis functions that extend throughout the spatial domain. Hydrologists and other physical scientists have made extensive use of such Fourier techniques to assess the global uncertainty of geospatial datasets (Bakr and others., 1978; Gutjahr and Gelhar, 1981; Dagan, 1982; Mizell and others, 1982; Gelhar and Axness, 1983; Naff and Vecchia, 1986; McLaughlin and Wood, 1988; Vomvoris and Gelhar, 1990; Robin and others, 1993). These estimates of uncertainty represent average error, in some mean-squared sense, and are assumed to apply locally as well as globally. They are well-suited to analyzing distributed processes that were created as the result of large-scale spatially-invariant geophysical influences.

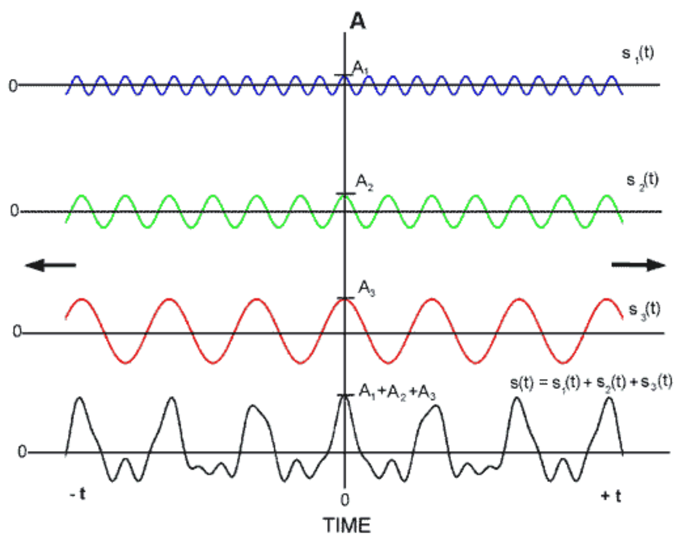

Figure 1. Stationary gaussian process in time as the sum of cosine functions with differing frequencies (from http://130.191.21.201/multimedia/jiracek/dga/spectralanalysis/examples.html).

Most naturally occurring distributed datasets, however, are not gaussian (see Figure 2). Their mean, variance, and correlation scale change over space in response to the local geophysical processes that formed them. They may exhibit non-stationary behavior at a variety of scales, including hierarchies of continuously-evolving scales of spatial variation (Neuman, 1990). Such scale-dependent structure often occurs when a variety of geophysical processes, each dictated by its own natural scale of variation, collectively produce the observed geospatial dataset. Geologic processes responsible for soil and aquifer formation, for example, range from small-scale fluvial sediment transport mechanisms to large-scale tectonic influences. In the presence of such non-stationary, scale-variant behavior, global estimates of error derived using standard gaussian geostatistical techniques will not accurately reflect local data uncertainty, and water resource management decisions made at township and watershed scales may be seriously compromised.

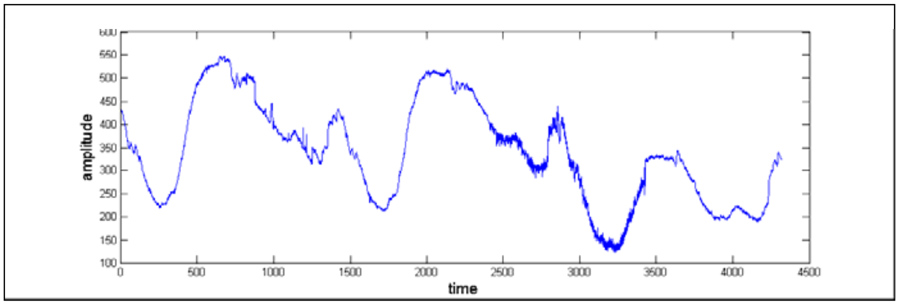

Figure 2. Non-stationary process in time
(from http://www.iro.umontreal.ca/~lisa/seminaires/22-07-2005.pdf).

As described previously, traditional Fourier methods of estimating non-local uncertainty are based on the assumption that the joint statistical behavior of the distributed process does not change over space or scale, and are appropriate only for application to stationary, scale-invariant geospatial datasets that exhibit strong global periodicities. Wavelets, on the other hand, are small waves characterized by locally-periodic structure that dampens out over space, as shown in Figure 3. The compact support of wavelet basis functions over space is what makes them useful for analyzing local structure (Walker, 1999; Walnut, 2004). Wavelets range from short, high-frequency functions capable of resolving small-scale features to long, low-frequency functions able to preserve large-scale structure. Since wavelets characterize data at a variety of different scales of resolution, dataset variation at many frequencies can be simultaneously accounted for at each location.

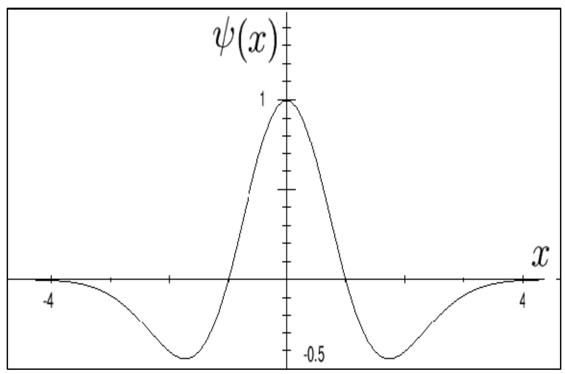

Figure 3. Mexican Hat wavelet
(from http://www.lac.inpe.br/~margarete/JASRWavelet.pdf).

How do wavelets handle local spatial variations at multiple scales? Figure 4 illustrates a set of wavelets obtained by successively re-scaling or 'dilating' a continuous Mexican Hat mother wavelet by powers of two. Again noting that we can substitute space for time, each wavelet describes data variability at a particular scale of spatial resolution, independent of variability at all other spatial scales. Some wavelets, such as the continuous Haar wavelet transform presented in Figure 5, are particularly adept at characterizing spatial variability in the presence of sharp discontinuities frequently encountered in natural distributed processes influenced by large-scale formative processes.

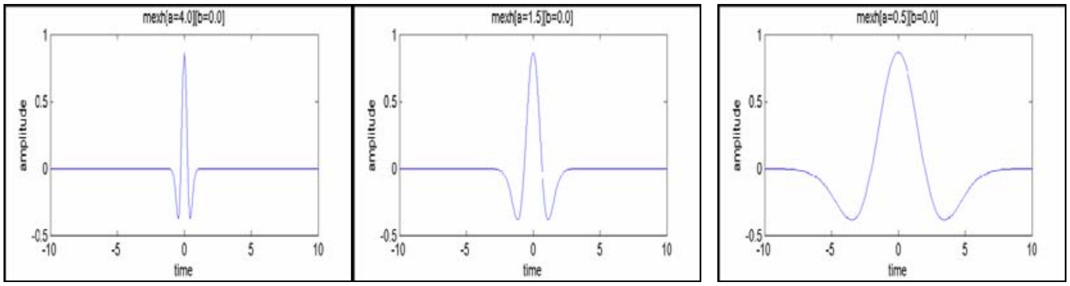

Figure 4. Wavelets obtained by scaling a Mexican Hat wavelet
(from http://www.iro.umontreal.ca/~lisa/seminaires/22-07-2005.pdf).

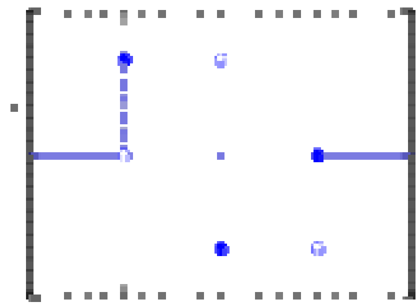

Figure 5. The Haar wavelet
(from http://en.wikipedia.org/wiki/Haar_wavelet).

Hypothesis

The local support of wavelets and their ability to handle more than one scale of spatial variation make them well-suited to analyzing local, scale-dependent behaviors of natural distributed processes (Katul and others, 1994; 1997; 2001). Many of the natural processes used to make water resource decisions exhibit non-stationary, scale-variant structure because the very processes that formed them varied over both space and scale. Wavelets provide a useful framework for explicitly quantifying local uncertainty in natural processes and for assessing the fitness of datasets for making sound water-resource decisions. An understanding of the distribution of local uncertainty is critical to making reliable local management decisions, particularly if the dataset is used as input to a highly nonlinear rainfall-runoff model. Under such conditions, local errors may become amplified and can propagate to other scales as they cascade through the model.

Figure 6 illustrates how wavelet analysis provides information about local, scale-variant structure. To perform a continuous wavelet transform, Morlet wavelet functions of varying scale, with s=1 for the mother wavelet and larger values of s for increasingly dilated wavelets, are translated at regular intervals beginning at the start of the process. The wavelet function characterized by scale s=1 is multiplied by the signal, integrated over all space, and multiplied by 1/sqrt(s). The resulting transformed value, a measure of the degree of correlation between the wavelet and the process, is plotted as amplitude in the space-translation/scale wavelet domain shown in Figure 8. The same wavelet is then shifted towards the right by a small amount, and a new transform value obtained and plotted. This procedure is repeated until the wavelet reaches the end of the signal.

When the transforms of the translated wavelet characterized by scale s=1 have been determined, the wavelet is dilated to scale s=2, and the procedure outlined above repeated, as illustrated in Figure 7. The wavelet is successively dilated to larger scales until the scale of the wavelet becomes so large that it no longer contains any useful information about the underlying geospatial process. All space- and scale-dependent behavior of the process is contained in the wavelet domain shown in Figure 8, which can be used to reconstruct the process or to generate many equally-probable alternate realizations by randomly selecting subsets of wavelet amplitudes. In practice, discrete wavelet transforms, such as Daubechies wavelets, are used rather than continuous wavelets to avoid the large computational burdens associated with oversampling in highly-redundant low frequency ranges (Daubechies, 1992). Extension of wavelet transforms to two-dimensional geospatial processes, which involves using higher-dimensional versions of the wavelets shown in Figure 6 and Figure 7, is straightforward.

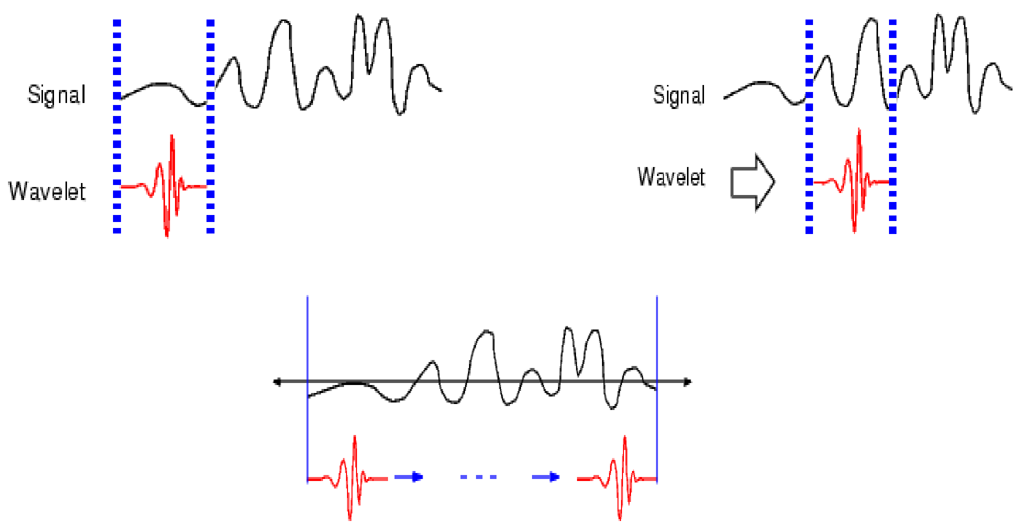

Figure 6. Translating a Morlet wavelet with scale s=1 through a process
(from http://www.mathworks.com/access/helpdesk/help/toolbox/wavelet/wavelet.html?/access/helpdesk/help/toolbox/wavelet/ch01_i15.html).

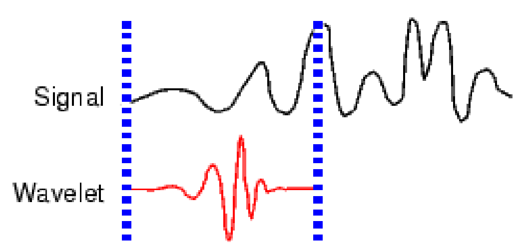

Figure 7. A Morlet wavelet with scale s=2
(from http://www.mathworks.com/access/helpdesk/help/toolbox/wavelet/wavelet.html?/access/helpdesk/help/toolbox/wavelet/ch01_i15.html).

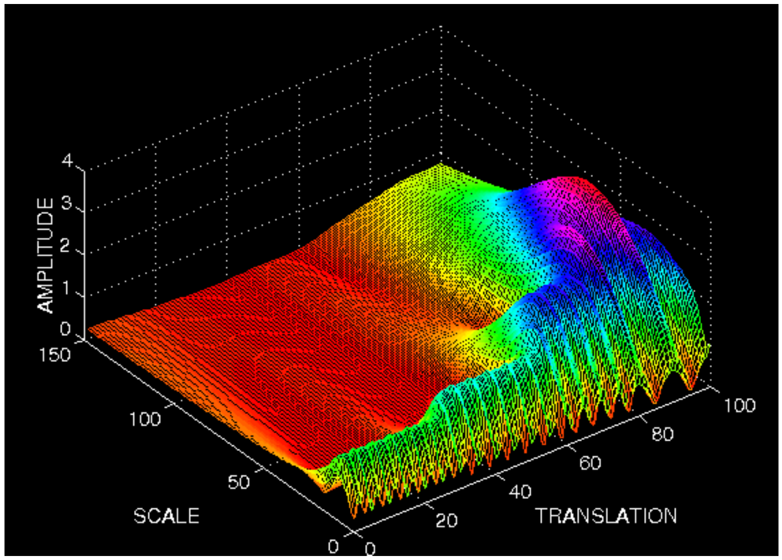

Figure 8. Plot of wavelet coefficients over scale and space
(from http://users.rowan.edu/~polikar/WAVELETS/WTpart3.html).

Objectives and Approach

The objectives of the proposed research are to:

- Develop a methodology for assessing local uncertainty in non-stationary, scale-variant natural processes; and
- Build a user-friendly GIS tool within the ArcGIS framework to implement the local uncertainty analysis.

As discussed previously, wavelets provide a useful tool for assessing local spatial variations in non-stationary, scale-variant natural processes. They differ from standard geostatistical techniques that assume the distributed process is statistically stationary and governed by a single characteristic scale of spatial variation. How do we translate understanding of non-stationary, scale-variant spatial structure obtained from wavelet analysis into estimates of local uncertainty?

Any distributed dataset represents a single possible set of geospatial data from among all possible datasets. To see that this is true, imagine the addition of a piece of data to the dataset. The resulting dataset will differ from the one that contained all the original data. Likewise, removal of a piece of data will produce a different dataset. This dependency of the dataset on the particular pieces of data collected occurs whenever we lack complete knowledge about the distributed process. It remains the principal source of uncertainty in most geospatial datasets.

We can mimic this lack of knowledge by randomly re-sampling the observed dataset, artificially producing an alterative dataset that is close to the original dataset but containing local error. If we do this repeatedly and in the same manner each time using a Monte Carlo approach, an ensemble of equally-probable datasets, or realizations, can be generated. Figure 9 illustrates how N alternative datasets might be generated by repeatedly re-sampling 23 of the 30 original data points, with replacement of all data points prior to each episode of re-sampling. In effect, this process of re-sampling with replacement, called bootstrapping, locally perturbs the dataset (Davison and Hinkley, 1997; Efron and Tibshirani, 1993). Collectively, the ensemble of N perturbed datasets provides us with information needed to assess local dataset uncertainty, by allowing us to estimate the variance at each location, as shown in Figure 10.

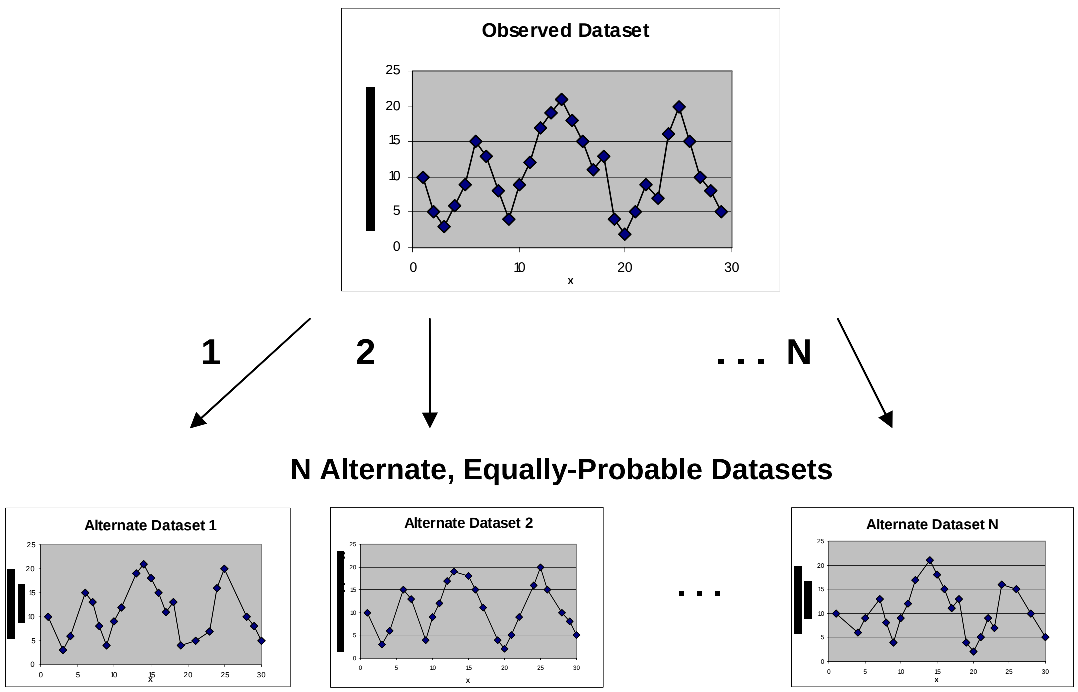

Figure 9. Bootstrappping generates N equally-probable alternative datasets.

Since bootstrapping does not involve making any assumptions about the global probabilistic behavior of dataset error, it would appear to be well-suited to the problem of assessing local uncertainties caused by non-stationary scale-variant influences. However, the random sampling inherent to bootstrapping requires that each piece of data be independent of all other data in the dataset. When the data are correlated over space, as they typically are for natural processes, bootstrapping can produce biased estimates of statistical properties at a point. To overcome this problem, we re-sample in the wavelet domain where the data are uncorrelated, rather than in the space domain (Percival and others, 2001; Breakspear and others, 2003). When the wavelet coefficients are brought back into the space domain, they provide an alternative dataset that preserves overall spatial structure contained in the original dataset, minus local structure associated with wavelet coefficients not re-sampled by the bootstrapping algorithm.

To implement the non-stationary, scale-variant uncertainty analysis, either an ArcMap UIButtonControl tool or a stand-alone VB.NET application will be developed within the ArcGIS framework. In either case, a button control will be attached to visual basic code and triggered by its click event. Initially, the code will determine whether the user-specified data is a feature class that must be converted into the raster format required to perform the uncertainty analysis. After raster conversion, the tool or application will then perform a single wavelet transform that uniquely defines the space- and scale-dependent characteristics of the input dataset. Localization of wavelets in space and frequency tends to produce sparseness in the distribution of non-zero wavelet coefficients over location and scale, a phenomenon that is well documented in the literature (Mallat, 1998). To account for such sparseness, the bootstrapping algorithm will automatically reject zero-valued coefficients during re-sampling. Depending on user-input specifications, the code will then repeatedly apply the bootstrapping algorithm in the wavelet domain to produce an alternate realization, followed by an inverse wavelet transform to bring each realization back into the space domain. Finally, ensemble statistics at each point will be estimated, and when the user moves the screen cursor over a particular location in the map view, the variance at that point will be displayed via the interface of the identify object class.

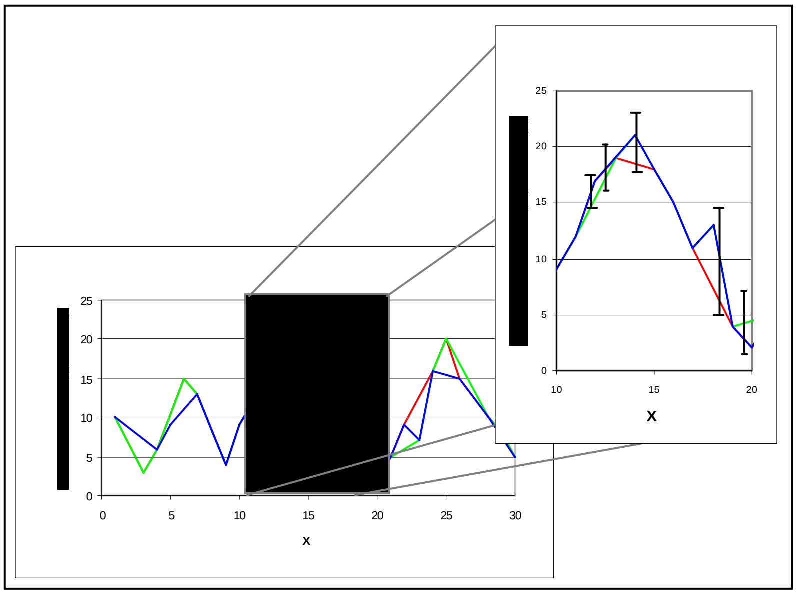

Figure 10. Estimating local variance from the ensemble of alternative equally-probable datasets.

Expected Results/Products

The PI is an experienced ArcObjects programmer, and has written ArcObjects code within both VBA and VB.NET programming environments. A sample of code recently written by the PI to test for a vector dataset layer and convert it to raster format is included in the appendix. The code references methods and properties on interfaces of the map document, map, layer, raster layer, raster, raster workspace, raster dataset, feature class, feature class descriptor, raster conversion operation, and raster analysis environment object classes.

Testing for and conversion of a feature class to raster format is the first step in bringing the spatial dataset into the wavelet domain, where it will be repeatedly re-sampled via bootstrappping, taken back into the space domain, and incorporated into the uncertainty analysis. Code to perform forward and inverse wavelet transforms is widely available on internet software developer user forums, including public domain fortran code distributed by Press and others (1992) for implementing Daubechies discrete wavelet transforms. This fortran code can easily be converted to VBA or Visual Basic .NET and included in the code executed by the uncertainty tool. Additional code to perform bootstrapping and other operations essential to the analysis will be added during the course of the project. The final ArcGIS uncertainty tool or application will be distributed in the form of an ArcMap template (.mxt) document or as a dll binary executable, along with a user's manual that describes output and provides guidance on how the user should interpret results of the uncertainty analysis.

After the tool has been developed and fully documented, it will be applied to the two Highlands projects discussed in the Project Support section of this proposal. Specifically, the tool will be used to flag locations and scales at which uncertainty in datasets used to inform water resource plans in the Highlands is most dominant. At these locations and scales, collection of additional data will be most critical to reducing uncertainties in Highlands planning processes at the state, county, watershed, and township scales.

Based on results of the proposed work, a journal article will be submitted to a peer-reviewed journal specializing in GIS content, such as the *International Journal of Geographical Information Science.* The paper will describe details of the wavelet bootstrapping algorithm and its application to various Highlands datasets critical to members of the Highlands Council and to water managers at the New Jersey Department of Environmental Protection (NJDEP) responsible for making decisions about the allocation of water resources in the Highlands area.

Significance to the USGS Mission

In support of the USGS mission to enhance our understanding of the Earth, mitigate losses from natural disasters, manage the allocation of scarce natural resources, and enhance environmental quality, the proposed study will provide a user-friendly tool to assess local uncertainty in non-stationary, scale-variant geospatial datasets typically used to make critical water resource decisions. More specifically, the proposed research is expected to offer watershed planners and managers a better understanding of the accuracy of distributed data used to make water resource decisions, including those made to minimize flood- and drought-related loss of life and property, manage surface-water resources for human and ecological uses, protect and enhance water resources needed to sustain human health, aquatic health, and environmental quality, and promote well-informed development of the Nation’s resources for the benefit of present and future generations. A high degree of dataset uncertainty at a particular location and scale will indicate the need to collect additional data at or near that location at a scale of measurement close to the scale for which uncertainty is high, allowing greater reliability in plans developed at that site and scale. Conversely, small uncertainty in data observed at a given point and scale of resolution will suggest a much lower priority be placed on data collection at or near the point at the measurement scale used to collect the data, and that resource allocation plans based on the data can be made with greater confidence.

The proposed research will also contribute to the CEGIS mission by providing a tool that can help water resource managers more fully understand linkages between anthropogenic factors and environmental quality, particularly in the context of how distributed but poorly measured landscape disturbances may affect the integrity of biotic ecosystems, enhancing GIS expertise currently lacking in water resource and other USGS disciplines.

References

Bakr, A., L. Gelhar, A. Gutjahr, and J. MacMillan (1978), Stochastic analysis of spatial variability in subsurface flows: 1. A comparison of one and three-dimensional flows, *Water Resour. Res*., 14, 263– 271.

Breakspear, M., Brammer, M., and P.A. Robinson (2003), Construction of multivariate surrogate sets from nonlinear data using the wavelet transform, *Physica*, D 182, 1, 1-22.

Burrough, P.A., van Rijn, R, M. Rikken (1996), Spatial data quality and error analysis issues: GIS functions and environmental modeling. In Goodchild M. and others (eds.), *GIS and Environmental Modeling: Progress and Research Issues,* GIS World Books, Fort Collins, Colorado, 29-34.

Dagan, G. (1982), Stochastic modeling of groundwater by unconditional and conditional probabilities, 1, Conditional simulation and the direct problem. *Water Resour. Res.*, 18, 813-833.

Daubechies, I. (1992), *Ten Lectures on Wavelets,* Society of Applied and Industrial Mathematics, Philadelphia, 357 p.

Davison, A.C., and Hinkley, D.V. (1997), Bootstrap methods and their application, *Cambridge Series on Statistical and Probabilistic Mathematics*, 1. Cambridge: Cambridge University Press.

Efron, B., and R.J. Tibshirani (1993), *An introduction to the bootstrap.* Monographs on Statistics and Applied Probability, 57, New York, NY: Chapman & Hall.

Ehlschlaeger, C.R., Shortridge, A.M., and M.F. Goodchild (1997), Visualizing spatial data uncertainty using animation, *Computers & Geosciences*. 23(4), 387-395.

Englund, E.J. (1993), Spatial simulation: environmental applications, in Goodchild, M.F., Parks, B.O., L.T. Steyaert (eds.), *Environmental modeling with GIS,* Oxford Press, New York, 432-437.

Fisher, P.F. (1991). First experiments in viewshed uncertainty: the accuracy of the viewshed area, *Photogrammetric Engineering & Remote Sensing*. 57(10), 1321-1327.

Gelhar, L., and C. Axness (1983), Three-dimensional stochastic analysis of macrodispersion in a stratified aquifer, *Water Resour. Res*., 19, 161–180.

Goodchild, M.F. (1995), Attribute accuracy, in Guptil, S.C. and J.L. Morrison.(eds.), *Elements of Spatial Data Quality*. Elsevier, London, 59-79.

Gutjahr, A. L., and L. W. Gelhar (1981), Stochastic models of subsurface flow: Infinite versus finite domains and stationarity, *Water Resour. Res*., 17, 337– 350.

Heuvelink, G.B., Burrough, P.A., and A. Stein (1989), Propagation of errors in spatial modeling with GIS,. *International Journal of Geographical Information Systems*. 3(4), 303-322.

Katul, G.G., Lai, C.T., Schafer, K., Vidakovic, B., Albertson, J.D., Ellsworth, D., and R. Oren (2001), Multiscale Analysis of Vegetation Surface Fluxes: from Seconds to Years, *Adv. Water Resour.* , 24, 1119-1132.

Katul, G.G., Hsieh, C.I., and J. Sigmon (1997), Energy-inertial scale interaction for temperature and velocity in the unstable surface layer, Boundary-Layer, *Meteorol*. 82, 49-80.

Katul, G.G., Parlange, M.B., and C.R. Chu (1994), Intermittency, local isotropy, and non-Gaussian statistics in atmospheric surface layer turbulence, *Physics of Fluids,* 6, 2480-2492.

Lee, J., Snyder, P.K., P.F. Fisher (1992), Modeling the effect of data errors on feature extraction from digital elevation models, *Photogrammetric Engineering & Remote Sensing*, 58(10), 1461-1467.
Mallat, S. (1998), *A Wavelet Tour of Signal Processing*, Academic Press, 637 p.
McLaughlin, D., and E. F. Wood (1988), A distributed parameter approach for evaluating the accuracy of groundwater model predictions: 1. Theory, *Water Resour. Res*., 24, 1037–1047.
Mizell, S., A. Gutjahr, and L. Gelhar (1982), Stochastic analysis of spatial variability in two-dimensional steady groundwater flow assuming stationary and nonstationary heads, *Water Resour. Res*., 18, 1053– 1067.
Naff, R., and A. Vecchia (1986), Stochastic analysis of three-dimensional flow in a bounded domain, *Water Resour. Res*., 22, 695– 704.
Neuman, S.P. (1990), Universal scaling of hydraulic conductivities and dispersivities in geologic media, *Water Resour. Res.,* 26(8), 1749-1758.
Olden, J.D. and N.L. Poff (2003), Redundancy and the choice of hydrologic indices for characterizing streamflow regimes: *River Research and Applications*, 19, 101-121.
Openshaw, S. (1989), Learning to live with errors in spatial databases, in Goodchild, M.F. and S. Gopal (eds.), *Accuracy in Spatial Database,*. Taylor & Francis, 263-276.
Percival, D.B., Sardy, S., and A.C. Davison (2001). Wavestrapping Time Series: Adaptive Wavelet-Based Bootstrapping, in *Nonlinear and Nonstationary Signal Processing*, W.J. Fitzgerald, R. L. Smith, A. T. Walden and P. C. Young (eds.), Cambridge University Press.
Poff, N. L, Allan, J. D. Bain, M. B. Karr, J. R. Prestegaard, K.L. Richter, B.D. Sparks, R.E. and J.C. Stromberg (1997), The natural flow regime: A paradigm for river conservation and restoration, *BioScience*, 47, 769-784.
Poff, N.L. and J.V. Ward (1989), Implication of streamflow variability and predictability for lotic community structure: A regional analysis of streamflow patterns, *Canadian Journal of Fisheries and Aquatic Sciences*, 46, 1805-1818.
Press, W. H., Flannery, B. P., Teukolsky, S. A., W.T. Vetterling (1992), *Numerical Recipes in FORTRAN*, Second Edition, Cambridge University Press, 584-599.
Robin, M. J., Gutjahr, A. L., Sudicky, E. A. and J.L. Wilson (1993), Cross-correlated random field generation with the direct Fourier transform method, *Water Resour. Res.*, 29, 2385-2397.
SDTS Task Force. (1996). The spatial data transfer standard: guide for technical managers, U.S. Dept. of the Interior. <ftp://sdts.er.usgs.gov/pub/sdts/articles/pdf/mgrs.pdf>
Unwin, D.J. (1995), Geographical information systems and the problem of ‘error and uncertainty’, *Progress in Human Geography*, 19(4), 549-558.
Vomvoris, E. G. and L.W. Gelhar (1990), Stochastic analysis of the concentration variability in a three-dimensional heterogeneous aquifer, *Water Resour. Res*., 26, 2591-2602.
Walker, J.S. (1999), *A Primer on Wavelets and Their Scientific Applications*, CRC Press LLC, 168 p.
Walnut, D. (2004), *An Introduction to Wavelet Analysis*, Birkhauser, 465 p.

Project Support

This research will support several ongoing USGS studies being conducted in the New Jersey Highlands, a 1,250 square mile area that provides water for a population of more than 5 million people in New Jersey (http://www.highlands.state.nj.us/njhighlands/actmaps). Undisturbed forests in the Highlands have historically protected both surface-water and groundwater supply and quality, by acting as a natural buffer to extreme flow events and environmental degradation. However, these forested lands are rapidly being lost to subdivisions and urban development. The U.S. Forest Service estimated that roughly 5,000 acres of forest and farmland was lost annually between 1994 and 2004 to urban and suburban development in the Highlands, threatening New Jersey's water supply and the health of its aquatic ecosystems (http://www.state.nj.us/dep/newsrel/2004/crowntowers200405.htm). In response to this growing challenge, the USGS is cooperating with the New Jersey Highlands Council and the NJDEP to assess the consequences of lost habitat and natural buffering capacity on water resource availability and quality in the Highlands.

Under the scope of one of these cooperative efforts, the New Jersey Water Science Center (NJWSC) of the USGS provides critical geospatial data to Highlands Council decision makers responsible for defining optimal rates of growth based on water availability, sustainability, and susceptibility. In accordance with the goals of the New Jersey Highlands Water Protection and Planning Act of 2004, the Highlands Council is developing a Regional Master Plan (RMP) based on scientifically robust principles. The RMP will ensure continued economic opportunity while protecting water supply and sustaining aquatic ecosystem viability. In the absence of meaningful estimates of geospatial data uncertainty, Council members must assume that their plans to limit or accommodate development are equally reliable across the full extent of the Highlands, without regard for the impact of local data uncertainties on statewide, county, township, and watershed zoning plans. The proposed CEGIS study will provide a tool to Highlands Council decision makers that will better inform critical growth decisions made at all locations and scales relevant to the planning process. It will also provide quantitative information regarding locations where additional data collection would most improve reliability of the RMP.

The USGS New Jersey Water Science Center is also working in collaboration with NJDEP to develop innovative methods for establishing passing streamflow requirements. Such requirements will help preserve natural streamflow variability critical to maintaining the integrity of aquatic ecosystems in the state. The NJDEP study relies on USGS continuous streamflow data to calculate ecologically relevant hydrologic indices describing natural flow variability (Poff and Ward, 1989; Olden and Poff, 2003). Predicted alterations in these hydrologic indices caused by projected changes in water demand, land use, and other anthropogenic influences represent measures of growth-related streamflow disturbance that will be used as surrogates for federally-mandated TMDLs. To estimate these 'hydro-TMDLs' in ungaged basins where streamflow records are not available, indices estimated in gaged basins have been regressed on various spatial datasets, including geospatial data characterizing stream channel and basin morphology, land use, soil and geologic properties, vegetative cover, recharge, and precipitation. Streamflow indices and hydro-TMDLs estimated for ungaged basins using these regression datasets will tend to inherit local dataset uncertainties. The proposed CEGIS study will help quantify the reliability of hydrologic disturbance indices inferred for ungaged basins, and also provide information about where additional collection of data would most improve the reliability of hydro-TMDLs.

Budget - Year 1

Budget Request

Fiscal Year 2007 Budget	Budget Amounts by Participating Cost Center Code		Total Year 1
	2454		
Personnel Salary	93,160		93,160
Other Expenses: travel, equipment, and supplies	0		0
TOTAL DIRECT	**93,160**		**93,160**
Gross Assessment Rate for Each Participating Cost Center	25.58%		
INDIRECT COSTS ESTIMATE (Gross Assessment Rate Times Total Direct)	23,830		23,830
TOTAL	**116,990**		**116,990**

Appendix: Sample ArcObjects Code

```
Private Sub RasterizeDataset_Click()

  Dim pDoc As IMxDocument
  Set pDoc = ThisDocument
  Dim pMap As IMap
  Set pMap = pDoc.FocusMap
  Dim pLayer As ILayer
  Set pLayer = pMap.Layer(0)

  'Allow user to input name of workspace, then open it.
  Dim DirName As String
  DirName = InputBox("Input the name of the directory that you'd like
        to use as the workspace:", "Input Name of WorkSpace")
  Dim pWSF As IWorkspaceFactory
  Dim pWS As IRasterWorkspace
  Set pWSF = New RasterWorkspaceFactory
  Set pWS = pWSF.OpenFromFile(DirName, 0)

  'If feature layer, ask user to specify name of raster output,
        cellsize, and attribute name, then convert to raster layer
  If TypeOf pLayer Is IGeoFeatureLayer Then

  'User inputs the name of the output grid, cellsize, and attribute
        name.
  Dim GridName As String
  GridName = InputBox("Your dataset is in vector format, and must be
        converted to a grid. Input the name of the grid where you'd like
        to store the rasterized coverage:", "Input Name of Output Grid
        for Vector Dataset")
  Dim CellSize As Integer
  CellSize = InputBox("Input the integer cell size.  Note that the
        smaller the cell size, the more accurate local uncertainty
        estimates will be:", "Input Cell Size")
  Dim AttributeName As String
  AttributeName = InputBox("Input the name of the attribute that you'd
        like to use to rasterize your dataset:", "Input Attribute Name")

  'Get the featureclass for the featureclassdescriptor.
  Dim pFeatLayer As IFeatureLayer
  Set pFeatLayer = pLayer
  Dim pFeatureClass As IFeatureClass
  Set pFeatureClass = pFeatLayer.FeatureClass

  'Create a FeatureClassDescriptor
  Dim pFCDesc As IFeatureClassDescriptor
  Set pFCDesc = New FeatureClassDescriptor
  pFCDesc.Create pFeatureClass, Nothing, AttributeName

  'Test for old raster file left from previous runs, and delete if it
        exists.
  On Error Resume Next
  Dim pRasterDS As IRasterDataset
  Set pRasterDS = pWS.OpenRasterDataset(GridName)
  If Not pRasterDS Is Nothing Then
  Set pDataset = pRasterDS
  If pDataset.CanDelete Then pDataset.Delete
  End If
```

```
'Set up the conversion environment.
Dim convert As IConversionOp
Set convert = New RasterConversionOp
Dim pEnv As IRasterAnalysisEnvironment
Set pEnv = convert
pEnv.SetCellSize esriRasterEnvValue, CellSize

'Convert to raster dataset.
Set pRasterDS = convert.ToRasterDataset(pFeatureClass, "GRID", pWS,
        GridName)

'Convert the output raster dataset to layer and add to map.
Dim pRasterLayer As IRasterLayer
Set pRasterLayer = New RasterLayer
Dim pRaster As IRaster
Set pRaster = pRasterDS.CreateDefaultRaster
pRasterLayer.CreateFromRaster pRaster
pMap.DeleteLayer pLayer
Set pLayer = pRasterLayer
pMap.AddLayer pLayer

End If

End Sub
```