GeoComputation 99

Program & Abstracts

4th International Conference
On
GeoComputation

25-28 July 1999

Mary Washington College
Fredericksburg, Virginia
USA

J. Ehlen
J. Nicholas
D. Caldwell
J. Diaz
P. Krause
J. Eichholz

Contents

Welcome

Welcome to the 4th International Conference on GeoComputation, Mary Washington College, and Fredericksburg. We have an excellent technical program, led by six distinguished keynote lectures, that we hope you will find worthwhile. In addition, we have a number of social activities planned to let you meet and discuss your research.

Judy Ehlen
Joe Nicholas

Registration

Most of you will have preregistered for the conference. For those who haven't or who need information, the registration desk will be located at the Main Entrance of the Woodward Campus Center (see sketch on page 4) from 1000 to 1700 on Sunday, July 25 and from 0830 to 1200 on Monday morning, July 26. Please make sure you sign up for one of the Walking Tours when you register.

Conference Program

Each half-day of the conference will begin with a keynote lecture. These lectures will be followed by parallel technical sessions interspersed withcoffee breaks to allow time for discussion. Keynote lectures will be held in the Great Hall of the Woodward Campus Center. A poster session will be held from 1515 to 1700 Tuesday afternoon in the Exhibit Area of the Great Hall.

The time table for keynote lectures and technical sessions can be found on page 5 in this volume. The technical program is on pages 6-15, an alphabetical listing of abstracts by author is on pages 16-22, the abstracts for each presentation are on pages 23-107, and the author index is on pages 108-110.

Exhibits

The exhibit space will be in the Great Hall in the Woodward Campus Center in the same area where coffee breaks and the poster session will be held. Exhibitors will be presentthroughout the conference.

Social Activities

Welcoming Party. The Welcoming Party will be held Sunday evening July 25 between 1800 and 2200 at Belmont Plantation in Falmouth, across the Rappahannock River from Fredericksburg. Buses will leave at 1730 from Mary Washington College to carry you to and from Belmont.

Belmont is an 18th century plantation house that was purchased by internationally known artist Gari Melchers in 1916. It is currently owned by the Commonwealth of Virginia and is administered by Mary Washington College. Mr. Melchers' studio and the house will be open for your pleasure and food and drink will be provided.

Representatives from the Fredericksburg Wine Festival Consortium (Hartwood Winery, Ingleside Plantation Vineyards, and Lake Anna Winery) will be present so that you can sample and learn about Virginia wines. Wine tasting

is sponsored by the Silver Cos. of Fredericksburg.

Dress will be casual.

Walking Tours. Walking tours on Monday afternoon are sponsored by the City of Fredericksburg. Two historical tours will be on offer. One will introduce conference participants to the Revolutionary War and will be led by Mr. George Washington. Ms. Jane Beale, a resident of the City during the Civil War (1860-1865), will lead the second tour.

Sign up sheets will be available on Sunday at the registration desk for these tours. Buses will transport participants to and from the downtown area. The buses will leave at 1515. Participants may wish to stay downtown after the walking tours to shop and have their evening meal. A map of down town restaurants is provided in the registration package.

Dress will be causal.

Conference Dinner. The Conference Dinner will be held in Seacobeck Dining Hall on Tuesday evening between 1900 and 2200.

Meals

The following meals are included in your registration fee: Breakfast on Monday, Tuesday and Wednesday; Lunch on Monday and Tuesday, the Welcoming Party Sunday evening and the Conference Dinner Tuesday evening.

Breakfast, lunch, and the Conference Dinner will be served in the Faculty/Staff Dining Room in Seacobeck Dining Hall (see Campus Map on Page 3).

Morning and Afternoon coffee breaks will be held in the Exhibit Area of the Great Hall in the Woodward Campus Center except for Monday (because of the Walking Tours).

Weather

Summer in Fredericksburg is typically hot, hazy, and humid. Daytime temperatures in July are usually in the high-80's to low-90's (F) and nighttime temperatures, in the upper 60's to low 70's. Precipitation is often in the formof showers and in the late afternoon and early evening. Lightweight clothing is appropriate.

Conference Sponsors

U.S. Army Topographic Engineering Center, Alexandria, Virginia, USA

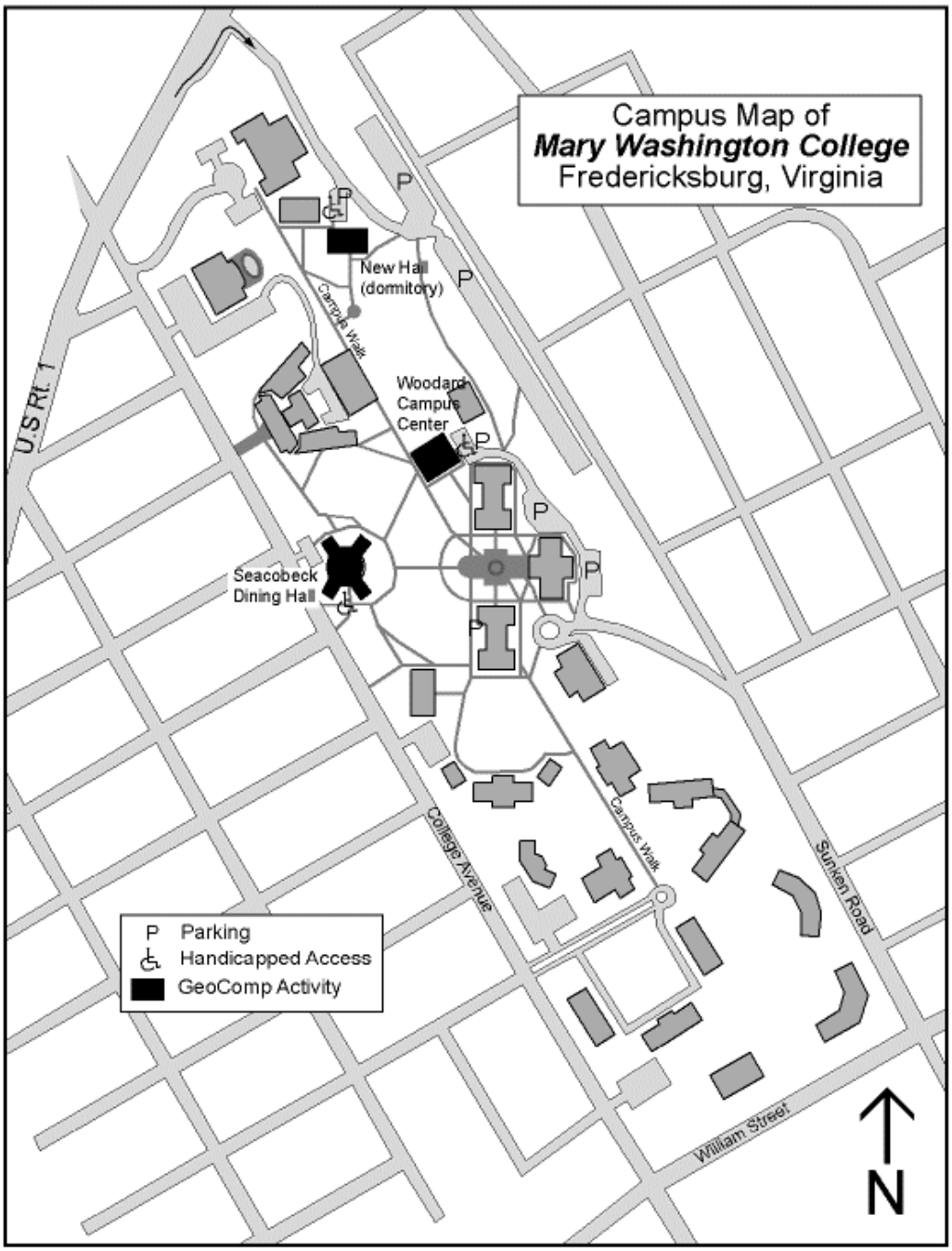
U.S. Army Research Office. Research Triangle Park, North Carolina, USA

U.S. Army European Research Office, London, England

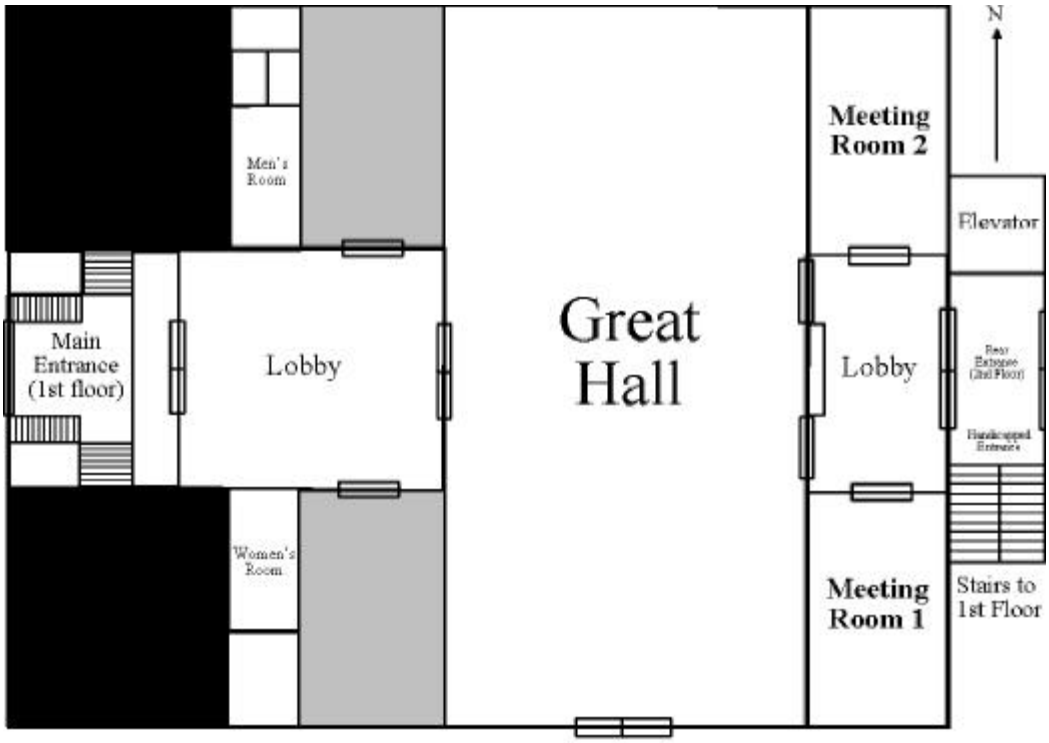
City of Fredericksburg, Fredericksburg, Virginia, USA

Silver Cos, Fredericksburg, Virginia, USA

Map of Mary Washington College Campus



Map of Woodward Campus Center



Conference Time Table

Sunday July 25		**Monday July 26**		**Tuesday July 27**		**Wednesday July 28**	
		Keynote Speaker (0830-0915)		Keynote Speaker (0830-0915)		Keynote Speaker (0830-0915)	
		Technical Session 3A (0920-1020)	Technical Session 3B (0920-1020)	Technical Session 6A (0920-1020)	Technical Session 6B (0920-1020)	Technical Session 10A (0920-1020)	Technical Session 10B (0920-1020)
		Coffee Break (1020-1040)		Coffee Break (1020-1040)		Coffee Break (1020-1040)	
		Technical Session 4A (1040-1200)	Technical Session 4B (1040-1140)	Technical Session 7A (1040-1200)	Technical Session 7B (1040-1200)	Technical Session 11A (1040-1140)	Technical Session 11B (1040-1200)
		Lunch (1200-1300)		Lunch (1200-1300)		Closing Remarks (1200-1215)	
Welcome (1300-1315)		Keynote Speaker (1300-1345)		Keynote Speaker (1300-1345)			
Keynote Speaker (1315-1400)		Technical Session 5A (1350-1450)	Technical Session 5B (1350-1450)	Technical Session 8A (1350-1510)	Technical Session 8B (1350-1510)		
Technical Session 1A (1400-1520)	Technical Session 1B (1400-1520)						
Coffee Break (1520-1540)		Walking Tours (1530-1700)	Buses Depart 1515	Technical Session 9	Posters (1515-1730)		
Technical Session 2A (1540-1640)	Technical Session 2B (1540-1640)						
Welcome Reception (1800-2200)	Buses depart 1730			Conference Dinner (1900-2200)			

"A" Sessions in Meeting Room 1

"B" Sessions in Meeting Room 2

Keynote Lectures in Great Hall

GeoComputation 99 Technical Program

SUNDAY July 25, 1999

Registration (Woodward Campus Center) 1000-1700

Welcome (Great Hall) 1300-1315 Judy Ehlen/Joe Nicholas

Keynote Presentation (Great Hall)

1315-1400 Geographical Data Mining
Author Stan Openshaw
Presenter Ian Turton*

Exploratory Spatial Data Analysis (Technical Session 1A, Meeting Room 1)

1400-1420 Web-Based Multi-Engine Spatial Analysis Tools
Authors James Macgill, Stan Openshaw, and Ian Turton

1420-1440 Multiresolution Spatial Analysis
Authors Mitchell J. Morehart, Fionn Murtagh, and Jean-Luc Starck

1440-1500 Augmented Seriation: Usability of a Visual and Auditory Tool for Geographic Pattern Discovery with Risk Perception Data
Authors Myke Gluck, Lixin Yu, Boryung Ju, Woo-Soeb Jeong, and Crystal Chang

1500-1520 Exploratory Spatial Optimization and Site Search: A Neighborhood Operator Approach
Authors Thomas J. Cova and Richard L. Church

Modeling the Physical World (Technical Session 1B, Meeting Room 2)

1400-1420 Toward an Object-Oriented Toolkit for Distributed Environmental Modeling
Authors Greg Tucker, Nicole Gasparini, Rafael Bras, Scott Rybarczyk, and Stephen Lancaster

1420-1440 Object-Oriented Modeling of Geodata as a Basic Tool for the Integration of Heterogeneous Paleoecological Information
Authors Holger Gaertner and Andreas Bergmann

* The presentor is indicated by underlining.

1440-1500	Investigations in Slope Development Through Landslide Activity - Concepts, Methods and Implications for Interdisciplinary and Interoperable Data Management
Authors	Jochen Schmidt, Holger Gaertner, and Richard Dikau
1500-1520	Interoperable Geospatial Objects
Author	Jonathan Doughty
Break	1520-1540

Spatial Data Processing (Technical Session 2A, Meeting Room 1)

1540-1600	Spatial Context Awareness in Feature Simplification
Authors	David McKeown, Jeff McMahill, and Douglas Caldwell
1600-1620	An Improved Algorithm for Calculating the Perimeter and Area of Raster Polygons
Author	Steven Prashker
1620-1640	Towards a Framework for High-Performance Geocomputation: Handling Vector-Topology within a Distributed Service Environment.
Authors	Steve Dowers, Bruce Gittings, and Mike Mineter

Terrain Analysis (Technical Session 2B, Meeting Room 2)

1540-1600	Principles of Semantic Modeling of Landform Structures
Authors	Martin Dehn, Holger Gaertner, and Richard Dikau
1600-1620	Spatial Pattern Analysis of Compositional Landforms in the Mojave Desert, California
Authors	Roy K. Dokka, Joseph M. Watts, and Denise R. LaDue
1620-1640	Long-Range Persistence of Elevation and Relief Values along the Continental Divide in the Conterminous U.S.
Authors	Scott Rice-Snow and Joshua Russell

Welcome Reception at Belmont Plantation (1800-2200)

MONDAY July 26, 1999

Keynote Presentation (Great Hall)

0830-0915 High Resolution Integrated Modelling of the Spatial Dynamics of Urban and Regional Systems
Author Roger White

Cellular Automata (Technical Session 3A, Meeting Room 1)

0920-0940 Cellular Strategies for the Simulation of Human Spatial Systems
Author W.D. Macmillan

0940-1000 Integrating Cellular Automata and Spatial Optimization for Evaluating Rapidly Urbanizing Regions
Authors Douglas P. Ward, Alan T. Murray, and Stuart R. Phinn

1000-1020 Empirical Cellular Automata (CA) Simulation from a High Resolution Population Surface
Authors David Martin and Fulong Wu

Neural Networks and Fuzzy Modeling (Technical Session 3B, Meeting Room 2)

0920-0940 An Investigation Into the Use of Data Spaces as an Organisational Concept for the Classification of Geographic Datasets
Authors Mark Gahegan and Masahiro Takatsuka

0940-1000 Neural Network Classifiers for GIS Data: Improved Search Strategies
Authors Gordon German

1000-1020 A Fuzzy Modelling Approach to Wild Land Mapping in Scotland
Authors Steffen Fritz, Linda See, and Steve Carver

Coffee Break 1020-1040

Cellular Automata and Modeling (Technical Session 4A, Meeting Room 1)

1040-1100 Exploring the Structure of Space: Towards Geocomputational Theory
Author David O'Sullivan

1100-1120 An Agent-Based Model of Residential Mobility in the Tel-Aviv Metropolitan Area
Authors Itzhak Benenson, Itzhak Omer, and Juval Portugali

1120-1140 Cyber-Spatial Analysis: Modelling Web Server Information Flows
Author Shane Murnion

1140-1200 Database Fusion for the Comparative Study of Migration Data.
Authors Oliver Duke-Williams and M. Blake

Remote Sensing 1 (Technical Session 4B, Meeting Room 2)

1040-1100 Developing Quality Training Data for a Statistical Decision Tree Classifier in a Spatial Environment
Authors Paul Crowther, Jacky Hartnett, Ray Williams, and Steve Pendelbury

1100-1120 A Comparison of Supervised Imagery Classification Using Analyst-Chosen and Geostatistically Chosen Training Sets
Authors James A. Shine and Gery I. Wakefield

1120-1140 Error-Constrained Change Detection
Authors Mark J. Ware and Christopher B. Jones

Lunch 1200-1300

Keynote Presentation (Great Hall)

1300-1345 Pluralism in Spatial Information Systems
Author Gail Langran Kucera

Distributed Computing (Technical Session 5A, Meeting Room 1)

1350-1410 Linking Process and Content in a Distributed Spatial Production System
Authors Henry Kucera and Pierre Lafond

1410-1430 Second Generation Spatial Information Warehousing Architectures
Authors Edric Keighan and Henry Kucera

1430-1450 Data Access, Querying, Analysis and Data Mining in a Distributed Framework for Earth System Science Support
Author Menas Kafatos

Spatial Modeling (Technical Session 5B, Meeting Room 2)

1350-1410 A Predictive GIS Model for Potential Mapping of Gold and Base Metal Mineralization in Takab Area, Iran
Authors Hooshang Asadi Haroni and Martin Hale

1410-1430 Geologically-Constrained Probabilistic Mapping of Gold Potential, Baguio District, Philippines
Authors Emmanuel John M. Carranza and Martin Hale

1430-1450 The Application of Cellular Automata Modeling for Enhanced Land Cover Classification in the Ecuadorian Amazon
Authors Joseph P. Messina, Stephen J. Walsh, Greg Taff, and Gabriela Valdivia

Walking Tours (1515-1700)

TUESDAY July 27, 1999

Keynote Presentation (Great Hall)

0830-0915 On Information Extraction Principles for Hyperspectral Data
Author David Landgrebe

Accessibility and Interaction (Technical Session 6A, Meeting Room 1)

0920-0940 Measuring Accessibility Using GIS
Author Rui Pedro Juliao

0940-1000 Towards a Percolation Model of Accessibility
Authors Kingsley Haynes, Roger Stough, and Rajendra Kulkarni

1000-1020 Parameter Estimation in Neural Spatial Interaction Modelling by a Derivative Free Global Optimization Method
Authors Manfred M. Fischer and Martin Reismann

Remote Sensing 2 (Technical Session 6B, Meeting Room 2)

0920-0940 A Practical Method of Calibrating Airborne Hyperspectral Imagery
Author Yu Jiang

0940-1000 Road Extraction from Digital Images using Linear Programming and Cost Functions
Authors Venkat Chalasani and Peter Beling

1000-1020 Texture Information and Supervised Classification of Hyperspectral Imagery by Means of Neural Networks
Author Edward H. Bosch

Break 1020-1040

Crime Analysis (Technical Session 7A, Meeting Room 1)

1040-1100 Smart Crime Pattern Analysis Using the Geographical Analysis Machine
Authors Ian Turton, Stan Openshaw, and James Macgill

1100-1120 First- and Second-Order Properties of Spatial Point Patterns: The Application of Crime Data from Baton Rouge, LA
Author Michael Leitner

1120-1140 Spatial Crime Displacement Resulting fromLarge-Scale Urban Renewal Programs in the City of Baltimore, MD: A GIS Modeling Approach
Authors Brian J. Schumacher and Michael Leitner

1140-1200 CrimeStat: A Spatial Statistics Program for the Analysis of Crime Incident Locations
Author Ned Levine

Hydrological Applications 1 (Technical Session 7B, Meeting Room 2)

1040-1100 Applying Saliency Analysis to Neural Network Rainfall - Runoff Modelling
Authors Robert J. Abrahart, Linda See, and Pauline E. Kneale

1100-1120 Multi-Model Data Fusion for Hydrological Forecasting
Authors Linda See and Robert J. Abrahart

1120-1140 A Spatial Solution to Calculate Optimum Surface Water Target Levels Using a Water Management Decision Support System
Authors J.W.J. van der Gaast and J.G. Kroes

1140-1200 Compositional Data Analysis of the Influence of Drainage Area and Stream Order on Hydraulic Geometry
Author Gregory S. Ridenour

Lunch 1200-1300

Keynote Presentation (Great Hall)

1300-1345 Investigating the Spatial Variation in Soil Radon Geostatistically.
Authors M. A. Oliver and A. L. Kharyat

Geostatistics (Technical Session 8A, Meeting Room 1)

1350-1410 Performance Comparison of Geostatistical Algorithms for Incorporating Elevation into the Mapping of Precipitation
Author Pierre Goovaerts

1410-1430 Designing Optimal Sampling Configurations with Ordinary and Indicator Kriging
Authors Christopher D. Lloyd and Peter M. Atkinson

1430-1450 Modern Enviroinformatics: BME Mapping in the Light of Uncertain Physical Knowledge Bases - The Equus Beds Case Study
Authors George Christakos and Mark Serre
Presenter Alexander Kolovos

1450-1510 Estimating a Non-Stationary Spatial Structure by Simulated Annealing
Authors Olivier Perrin and Serge Iovleff

Hydrological Applications 2 (Technical Session 8B, Meeting Room 2)

1350-1410 Self-Organization in Fluvial Landscapes: Sediment Dynamics as an Emergent Property
Author Dirk DeBoer

1410-1430 The Validity of Using a Simplified Distributed Hydrological Model for Estimation of Landslide Probability Under a Climate Change Scenario.
Authors James A. Griffiths, A.J.C. Collison, and S.W. Wade

1430-1450 Dynamic Modelling of the Spatio-Temporal Distribution of Phytoplankton in a Small Productive English Lake
Authors R.D. Hedger, T.J. Malthus, N.R.B. Olsen, D.G. George, and P.M. Atkinson

1450-1510 Flat Feature Processes from Triangulated Irregular Networks for Hydrological Modeling
Authors Honglei Zhu and Kristin Schneider

Poster Session (Technical Session 9, Great Hall)

1515-1730	Assessing Uncertainty in Fuzzy Land Cover Maps Obtained from Remotely Sensed Imagery
Author	Peter M. Atkinson
1515-1730	3D City: Prototyping Techniques for Urban Design Modeling
Author	Weiso Chen
1515-1730	A Stochastic-Dynamical System for River Basin Runoff in Complex Terrain: A Case for Low-Dimensional, Physically-Based Models
Author	Christopher J. Duffy, Kris Sedmera, and Tongying Shun
1515-1730	Image Understanding for Battlefield Awareness Video Surveillance and Monitoring
Authors	Joseph H. Findley III and Steven C. Haes
1515-1730	Quantifying and Visualizing Terrain Fabric from Digital Elevation Models
Author	Peter L. Guth
1515-1730	Environmental Modeling and Monitoring using a High-Resolution Hydrometeorological Data System (HRHDS)
Authors	David R. Legates, Kenneth R. Nixon, Geoffrey E. Quelch, and Thomas D. Stockdale
1515-1730	Reliability-Based Vegetation Mapping: The Integration of Predictive Modeling and Digital Image Processing
Authors	Joseph Watts, Joni Jarrett, Gery Wakefield, Kevin Slocum, Francis Precht, and John Fels
1515-1730	Web-Based GIS Used to Enhance Public Democratic Involvement
Authors	Andrew J. Evans, Richard Kingston, Steve Carver, and IanTurton
1515-1730	CrimeStat: A Spatial Statistics Program for the Analysis of Crime Incident Locations (Demonstration)
Author	Ned Levine

Conference Dinner 1900-2200

WEDNESDAY July 28, 1999

Keynote Presentation (Great Hall)

0830-0915 Exploring Geo-Data Spaces - The Search for Meaning
Author Alan M. MacEachren

Visualization and Multidimensional Data (Technical Session 10A, Meeting Room 1)

0920-0940 Mapping the Environment through 3-D Space and Time
Authors Kevin Morris, David Hill, and Tony Moore

0940-1000 Developing a Visualisation and Analysis System for Simulating the Growth of Coastal Spits
Authors David Livingstone and Jonathan Raper

1000-1020 A 3-D Topology Model for Vector Data
Authors Kevin C. Trott and Ian Greasley

DEM Generation and Applications (Technical Session 10B, Meeting Room 2)

0920-0940 What's the Point? Interpolation and Extrapolation with a Regular Grid DEM
Authors David Kidner, Mark Dorey, and Derek Smith

0940-1000 Failure Prediction in Automatically Generated Digital Elevation Models
Authors Michael Gooch and Jim Chandler

1000-1020 A Distributed Approach For Planning Radio Communications
Authors David Kidner, Ian Fitzell, Miqdad Al Nuami, Phil Rallings, and Andrew Ware

Break 1020-1040

Remote Sensing 3 (Technical Session 11A, Meeting Room 1)

1040-1100 Operational Interferometric SAR Data Processing for RADARSAT Using a Distributed Computing Environment
Authors Stephen C. Taylor, Bernard Armour, William H. Hughes, Andrew Kult, and Chris Nizman

1100-1120 Remote Sensing and Process-BasedModelling
Authors Paul J. Curran, Terry P. Dawson, and Toby E. Wicks

1120-1140 Estimating Subpixel Geospatial Features
Authors Henry Berger and James Shine

Uncertainty and Error (Technical Session 11B, Meeting Room 2)

1040-1100 Stochastic Modeling of Alternate Methods for the Generation of Event Patterns with Complete Spatial Randomness
Author Frank Hardisty

1100-1120 Representing Uncertainty of Qualitative Thematic Maps with an Inter-map Cell Swapping Heuristic
Authors Charles R. Ehlschlaeger

1120-1140 Spatially Assessing Model Error Using Geographically Weighted Regression
Author Shawn W. Laffan

1140-1200 Exploring Microsimulation Methodologies for the Estimation of Household Attributes
Authors Dimitris Ballas, Graham Clarke, and Ian Turton

Closing Remarks (Great Hall) 1200-1215 Judy Ehlen/Joe Nicholas

List of Abstracts (Alphabetical by Principal Author)

Paul Crowther, Jacky Hartnett, Ray Williams, and Steve Pendelbury
Developing Quality Training Data for a Statistical Decision Tree Classifier in a Spatial Environment

Paul J. Curran, Terry P. Dawson, and Toby E. Wicks
Remote Sensing and Process-Based Modelling

Dirk DeBoer
Self-Organization in Fluvial Landscapes: Sediment Dynamics as an Emergent Property

Martin Dehn, Holger Gaertner, and Richard Dikau
Principles of Semantic Modeling of Landform Structures

Roy K. Dokka, Joseph M. Watts, and Denise R. LaDue
Spatial Pattern Analysis of Compositional Landforms in the Mojave Desert, California

Jonathan Doughty
Interoperable Geospatial Objects

Steve Dowers, Bruce Gittings, and Mike Mineter
Towards a Framework for High-Performance Geocomputation: Handling Vector-Topology within a Distributed Service Environment

Christopher J. Duffy, Kris Sedmera, and Tongying Shun
A Stochastic-Dynamical System for River Basin Runoff in Complex Terrain: A Case for Low-Dimensional, Physically-Based Models

Oliver Duke-Williams and M. Blake
Database Fusion for the Comparative Study of Migration Data

Charles R. Ehlschlaeger
Representing Uncertainty of Qualitative Thematic Maps with an Inter-map Cell Swapping Heuristic

Andrew J. Evans, Richard Kingston, Steve Carver, and Ian Turton
Web-Based GIS Used to Enhance Public Democratic Involvement

Joseph H. Findley III and Steven C. Haes
Image Understanding for Battlefield Awareness Video Surveillance and Monitoring

Manfred M. Fischer and Martin Reismann
Parameter Estimation in Neural Spatial Interaction Modelling by a Derivative Free Global Optimization Method

Steffen Fritz, Linda See, and Steve Carver
A Fuzzy Modelling Approach to Wild Land Mapping in Scotland

Holger Gaertner and Andreas Bergmann
Object-Oriented Modeling of Geodata as a Basic Tool for the Integration of Heterogeneous Paleoecological Information

Mark Gahegan and Masahiro Takatsuka
An Investigation Into the Use of Data Spaces as an Organisational Concept for the Classification of Geographic Datasets

Gordon German
Neural Network Classifiers for GIS Data: Improved Search Strategies

Myke Gluck, Lixin Yu, Boryung Ju, Woo-Soeb Jeong, and Crystal Chang
Augmented Seriation: Usability of a Visual and Auditory Tool for Geographic Pattern Discovery with Risk Perception Data

Michael Gooch and Jim Chandler
Failure Prediction in Automatically Generated Digital Elevation Models

Pierre Goovaerts
Performance Comparison of Geostatistical Algorithms for Incorporating Elevation into the Mapping of Precipitation

James A. Griffiths, A.J.C. Collison, and S.W. Wade
The Validity of Using a Simplified Distributed Hydrological Model for Estimation of Landslide Probability Under a Climate Change Scenario.

Peter L. Guth
Quantifying and Visualizing Terrain Fabric from Digital Elevation Models

Frank Hardisty
Stochastic Modeling of Alternate Methods for the Generation of Event Patterns with Complete Spatial Randomness

Kingsley Haynes, Roger Stough, and Rajendra Kulkarni
Towards a Percolation Model of Accessibility

R.D. Hedger, N.R.B. Olsen, D.G. George, P.M. Atkinson, and T.J. Malthus
Dynamic Modelling of the Spatio-Temporal Distribution of Phytoplankton in a Small Productive English Lake

Yu Jiang
A Practical Method of Calibrating Airborne Hyperspectral Imagery

Rui Pedro Juliao
Measuring Accessibility Using GIS

Menas Kafatos
Data Access, Querying, Analysis and Data Mining in a Distributed Framework for Earth System Science Support

Edric Keighan and Henry Kucera
Second Generation Spatial Information Warehousing Architectures

David Kidner, Mark Dorey, and Derek Smith
What's the Point? Interpolation and Extrapolation with a Regular Grid DEM

David Kidner, Ian Fitzell, Phil Rallings, Miqdad Al Nuami, and Andrew Ware
A Distributed Approach For Planning Radio Communications

Gail Langran Kucera
Pluralism in Spatial Information Systems

Henry Kucera and Pierre Lafond
Linking Process and Content in a Distributed Spatial Production System

Shawn W. Laffan
Spatially Assessing Model Error Using Geographically Weighted Regression

David Landgrebe
On Information Extraction Principles for Hyperspectral Data

David R. Legates, Kenneth R. Nixon, Geoffrey E. Quelch, and Thomas D. Stockdale
Environmental Modeling and Monitoring using a High-Resolution Hydrometeorological Data System (HRHDS)

Michael Leitner
First- and Second-Order Properties of Spatial Point Patterns: The Application of Crime Data from Baton Rouge, LA

Ned Levine
CrimeStat: A Spatial Statistics Program for the Analysis of Crime Incident Locations

David Livingstone and JonathanRaper
Developing a Visualisation and Analysis System for Simulating the Growth of Coastal Spits

Christopher D. Lloyd and Peter M. Atkinson
Designing Optimal Sampling Configurations with Ordinary and Indicator Kriging

Alan M. MacEachren
Exploring Geo-Data Spaces - The Search for Meaning

James Macgill, Stan Openshaw, and Ian Turton
Web-Based Multi-Engine Spatial Analysis Tools

W.D. Macmillan
Cellular Strategies for the Simulation of Human Spatial Systems

David Martin and Fulong Wu
Empirical Cellular Automata (CA) Simulation from a High Resolution Population Surface

David McKeown, Jeff McMahill, and Douglas Caldwell
Spatial Context Awareness in Feature Simplification

Joseph P. Messina, Stephen J. Walsh, Greg Taff, and Gabriela Valdivia
The Application of Cellular Automata Modeling for Enhanced Land Cover Classification in the Ecuadorian Amazon

Mitchell J. Morehart, Fionn Murtagh, and Jean-Luc Starck
Multiresolution Spatial Analysis

Kevin Morris, David Hill, and Tony Moore
Mapping the Environment through 3-D Space and Time

Shane Murnion
Cyber-Spatial Analysis: Modelling Web Server Information Flows

David O'Sullivan
Exploring the Structure of Space: Towards Geocomputational Theory

M. A. Oliver and A. L. Kharyat
Investigating the Spatial Variation in Soil Radon Geostatistically.

Stan Openshaw
Geographical Data Mining

David O'Sullivan
Exploring the Structure of Space: Towards Geocomputational Theory

Olivier Perrin and Serge Iovleff
Estimating a Non-Stationary Spatial Structure by Simulated Annealing

Steven Prashker
An Improved Algorithm for Calculating the Perimeter and Area of Raster Polygons

Scott Rice-Snow and Joshua Russell
Long-Range Persistence of Elevation and Relief Values along the Continental Divide in the Conterminous U.S.

Gregory S. Ridenour
Compositional Data Analysis of the Influence of Drainage Area and Stream Order on Hydraulic Geometry

Jochen Schmidt, Holger Gaertner, and Richard Dikau
Investigations in Slope Development Through Landslide Activity - Concepts, Methods and Implications for Interdisciplinary and Interoperable Data Management

Brian J. Schumacher and MichaelLeitner
Spatial Crime Displacement Resulting from Large-Scale Urban Renewal Programs in the City of Baltimore, MD: A GIS Modeling Approach

Linda See and Robert J. Abrahart
Multi-Model Data Fusion for Hydrological Forecasting

James A. Shine and Gery I. Wakefield
A Comparison of Supervised Imagery Classification Using Analyst-Chosen and Geostatistically Chosen Training Sets

Stephen C. Taylor, Bernard Armour, William H. Hughes, Andrew Kult, and Chris Nizman
Operational Interferometric SAR Data Processing for RADARSAT Using a Distributed Computing Environment

Kevin C. Trott and Ian Greasley
A 3-D Topology Model for Vector Data

Greg Tucker, Nicole Gasparini, Rafael Bras, Scott Rybarczyk, and Stephen Lancaster
Toward an Object-Oriented Toolkit for Distributed Environmental Modeling

Ian Turton, Stan Openshaw, and James Macgill
Smart Crime Pattern Analysis Using the Geographical Analysis Machine

J.W.J. van der Gaast and J.G. Kroes
A Spatial Solution to Calculate Optimum Surface Water Target Levels Using a Water Management Decision Support System

Douglas P. Ward, Alan T. Murray, and Stuart R. Phinn
Integrating Cellular Automata and Spatial Optimization for Evaluating Rapidly Urbanizing Regions

Mark J. Ware and Christopher B. Jones
Error-Constrained Change Detection

Joseph Watts, Joni Jarrett, Gery Wakefield, Kevin Slocum, Francis Precht, and John Fels
Reliability-Based Vegetation Mapping: The Integration of Predictive Modeling and Digital Image Processing

Roger White
High Resolution Integrated Modelling of the Spatial Dynamics of Urban and Regional Systems

Honglei Zhu and Kristin Schneider
Flat Feature Processes from Triangulated Irregular Networks for Hydrological Modeling

ABRAHART, Robert J. (bob@ashville.demon.co.uk), University of Greenwich, School of Earth and Environmental Sciences, U.K.; SEE, Linda, and KNEALE, Pauline E., University of Leeds, School of Geography, Leeds LS2 9JT, U.K.

Applying Saliency Analysis to Neural Network Rainfall - Runoff Modelling

Key Words: rainfall-runoff modelling, neural networks, saliency analysis

Explanation and understanding are two important objectives of model building. Most current hydrological model buildingutilises knowledge of the physical system to build a set of cause and effect relationships based on mathematical equations. But the fact that a model works is not a sufficient justification for the existence of causal relationships. This must be inferred from logical reasoning. Such arguments also will be true for neural networks that are often criticised on the grounds that these tools are nothing more than "black box" models. This paper considers the nature of explanation from a model building standpoint and provides an illustrated example of how a neural network can be disaggregated in terms of its forecasting inputs. This disaggregation, termed "saliency analysis," is based on a neurocomputing method for assessing the relative importance of different neural network components and, in this work, is used to examine the significance of various internal relationships. For example, it could be assumed that connections with low weights are less important than those with large weights, and then use the magnitude of a weight value as a measure of its importance. But this approach has limited theoretical justification.

A more useful alternative would be to define relative importance in terms of the effect that each item has on the network error function. In particular, the importance of a weight or node could be defined in terms of the change in the error function that results from small changes to, or deletions of, individual weights or nodes. Deletion in a practical sense could involve setting the weights to zero, setting all node output values to zero, or passing zero values to the nodes. Most standard neural network packages do not contain the tools that are needed to perform these types of operations on an iterative basis; however, converting a neural network solution into 3GL code allows the manipulation of input values using program loops and embedded functions to achieve this aim.

It must be stressed that "saliency analysis" is not a form of "sensitivity analysis" since the purpose of the exercise is not to examine "the rate of change in one factor with respect to change in another." Saliency analysis is concerned with excluding one or more variables, and then exploring inherent variation associated with the remainder. Neural networks are ideal tools for this type of analysis because of the distributed nature of their information processing structure, which enables inputs to be omitted while the model is running.

Neural network solutions for generating one-step-ahead river flow predictions based on TOPMODEL hydrological inputs were used in this analysis. Embedded functions were used to perform input "saliency analysis" assessment. Time series plots

showed that certain inputs were influential controls in different parts of the hydrograph. This technique is useful for building parsimonious neural network modelling solutions and for rapid prototyping of more complex mechanisms.

ASADI HARONI, Hooshang (Harouni@itc.nl), Ministry of Culture and Higher Education, Dr. Beheshti Avenue, Shahid Sabounchi Cross, Tehran, Iran; and HALE, Martin, International Institute for Aerospace Survey and Earth Sciences (ITC), Kanaalweg 3, 2628 EB, Delft, The Netherlands

A Predictive GIS Model for Potential Mapping of Gold and Base Metal Mineralization in Takab Area, Iran

Key Words: sediment-hosted gold deposits, epithermal mineralization, hydrothermal alteration mapping, structural mapping, magnetic signatures, weights of evidence analysis, predictive GIS model

Interpretation of aeromagnetic, Landsat Thematic Mapper (TM), geological and mineral occurrence data are used to recognize the combination of mapped geological features, hydrothermal alteration, and magnetic signatures that could be associated with the known epithermal gold, arsenic, antimony, and base metal deposits in a 1,600 km^2 area near Takab, northwest Iran. This area, part of the suture zone located between the Afro-Arabian and Iranian plates, which extends more than 2,000 km to northeast of, and parallel to, the Zagros thrust fault, comprises Precambrian metamorphics, Tertiary volcano-sedimentary rocks, and Quaternary sediments.

A principal component (PC) transformation applied to selected reflective Landsat-TM bands identifies iron oxides and hydroxyl-bearing minerals, some of which are hydrothermal alterations associated with the known mineralization in the area. Gradient images of aeromagnetic data depict major structures in the area, which appear to control the hydrothermal fluid flow and mineral deposition. Reduction to the pole of total magnetic intensity aids recognition of weak magnetic responses that could be associated with zones of iron oxide hydrothermal alteration; two such zones characterize major mineral deposits in the area. Analytical signal analysis of aeromagnetic data reveals that high amplitude anomalies characterize many of the known mineral deposits and their associated high magnetic-susceptibility igneous rocks both at surface and at depth.

Geological data such as permeable and reactive host rocks, surface igneous rocks, and structures are integrated with the hydrothermal alteration and the subsurface igneous heat sources, and four important parameters that could act as indicators of the occurrence of further mineralization are determined. Four binary maps representing diagnostic deposit recognition criteria have been combined using a weight of evidence model. The model uses the spatial distribution of 19 known mineral occurrences to calculate a final map of gold and base metal potential, based on the binary predictor patterns, for undiscovered mineral deposits in the Takab area.

ATKINSON, P.M. (pma@soton.ac.uk), University of Southampton, Department of Geography, Highfield, Southampton SO17 1BJ, U.K.

Assessing Uncertainty in Fuzzy Land Cover Maps Obtained From Remotely Sensed Imagery

Key Words: error, fuzzy classification, remote sensing,uncertainty

A common objective of remote sensing is the mapping of land cover. Traditionally, this has been achieved using hard classifiers such as the maximum likelihood classifier. In these circumstances, accuracy is readily assessed using the confusion matrix or contingency table. Based on this matrix the overalluser's and producer's accuracies can be estimated as can the same accuracies per class. Further, Kappa statistics can be estimated that allow for randomness in the matrix. During the last 10 years or so, researchers have increasingly applied fuzzy classifiers, such as the fuzzy c-means classifier and the mixture model that estimates the membership of a pixel, to each class. Clearly, where a pixel represents a mixture of land cover classes (for example, 20 percent cereals, 50 percent heathland, and 30 percent forest) fuzzy classifiers provide the opportunity for greater accuracy, and for this reason, fuzzy classification has become a fundamentally important approach in remote sensing. There is no accepted standard method for assessing the accuracy of a fuzzy classification. A scatterplot between the observed values and the estimated values provides a useful graphical representation of the accuracy of the estimates; however, quantitative summary of the information in the scatterplot has proved elusive and researchers have often used ad-hoc combinations of statistics, such as the mean error, root mean square (RMS) error, and correlation coefficient.

The RMS error is insufficient because it is insensitive to the variance per class and the overall number of classes. For example, a class that has small membership x in all pixels (say $0\%<x<10\%$) may be estimated with small RMS error by setting all pixels to the mean membership for that class (say 5 percent) even though the correlation between observed and estimated memberships is zero. The likelihood of having small memberships per class increases with the number of classes c.The correlation coefficient is insufficient because it is insensitive to bias such that r may be large when the scatterplot lies away from the 1:1 line. A particular problem arises for classes whose memberships are bimodal (i.e., little mixing as is often the case for "water" classes where pixels are either close to 0 percent or 100 percent water but rarely lie between). Then r can be large when the correlation within each "cluster" is zero. Several statistics were evaluated that address these issues including schemes forstandardising the RMS error and weighting schemes for estimating overall precision and bias. First, the above issues were explored using fuzzy maps of land cover obtained for four different sites representing different biomes (New Forest, U.K.; Cukurova Deltas, Turkey; BOREAS site, Canada; HAPEX-Sahel site, Niger). These fuzzy maps were obtained from the

National Oceanic and Atmospheric Administration's (NOAA) Advanced Very-High Resolution Radiometer (AVHRR) imagery (trained on classified fine spatial resolution imagery) using a standard feed-forward back-propagation artificial neural network. Second, the proposed statistics were evaluated using simulated data involving classes with different distributions of memberships and different numbers of these classes.

BALLAS, Dimitris (d.ballas@geography.leeds.ac.uk), CLARKE, Graham, and TURTON, Ian, University of Leeds, School of Geography, Leeds LS2 9JT, U.K.

Exploring Microsimulation Methodologies for the Estimation of Household Attributes

Key Words: microsimulation, micro-data, urban modelling, micro-scale, list-based representation

Microsimulation is a rapidly expanding area of spatial modelling that seems to offer great potential for applied policy analysis; however, currently there is considerable debate on the most appropriate methodology for estimating micro-data. Household or individual attribute data can be represented both as lists and/or tabulations. It has long been argued (Birkin and Clarke, 1995; Clarke, 1996; Williamson et al., 1998) that the representation of information on households and individuals in the form of lists offers greater efficiency of storage and spatial flexibility as well as an ability to update and forecast. This paper reviews the possibilities and methodologies of building list-based population micro-data for small areas. First, it evaluates the methods that have been developed and employed thus far for the estimation of population micro-data, outlining the advantages and drawbacks of each. Then the paper investigates the comparison of methods for generating conditional probabilities by statistical matching techniques, or by using probabilities directly from household data sets such as the Samples of Anonymised Records (SARs) and the Small Area Statistics (SAS) from the U.K. Census of Population. In addition, it explores the combination of these methods in a microsimulation framework and presents micro-data outputs from a local labour market microsimulation model for Leeds. Finally, it highlights the difficulties of calibrating this kind of model and validating the model outputs, given the absence of suitable observed statistics

---

BENENSON, Itzhak (bennya@post.tau.ac.il), OMER, Itzhak and PORTUGALI, Juval, Tel-Aviv University, Department of Geography and the Human Environment, 69978, Ramat-Aviv, Tel-Aviv, Israel

An Agent-Based Model of Residential Mobility in the Tel-Aviv Metropolitan Area

Key Words: agent-based modeling, GIS-based model, residential mobility, residential segregation

An attempt to construct an agent-based model of a residential mobility in Tel-Aviv is described. It is based on a hybrid of a PC-based GIS and High-Performance (HP) parallel computer. The GIS serves as a tool to represent and model immobile infrastructures, and, as a user's interface while the HP is a computation engine for simulating residential mobility. At the current stage of the research we have accounted for urban physical structure and residential distribution only (ignoring, for instance, the distribution of jobs).

An urban system in the model is decomposed into two compartments, i.e., physical properties and population, according to different time scales in which their dynamics take place. For example, in Tel-Aviv, during last 30 years, new dwellings appear at a yearly rate of 1 percent and below, while the rates of internal residential mobility, out-migration and in-migration, each remain at an annual level of 5 percent. To represent infrastructure, land parcels, buildings, and the links of the street networks are considered as basic units. They are organized as a set of layers within the high-resolution (vector) GIS of the Tel-Aviv metropolitan area. GIS serves as the ground upon which to place the model of residential mobility. Descriptions of urban population mobility are based on the description of the residential behavior of an individual decision-maker (householder).

Data on the city's inhabitants are based on the Israeli Central Bureau of Statistic (CBS) databases, which are partially available for supervised research. Model agents have the ability to estimate the state of the city and to behave in accordance with their own properties, the local information regarding the state of the neighborhood, and of the neighbors and the global information based on the state of the city as a whole. Based on this information, agents immigrate into the city, occupy and change residential locations there, and leave the city when conditions become unsatisfactory. When residing in the city, agents instantaneously re-estimate the available information on the neighborhood, and on the whole city, and react to changes by adjusting their properties and residential behavior. As a result, the dynamics of the entire urban system are described as the collective consequence of complex interactions of numerous immobile infrastructure units and mobile human agents. Several algorithms aimed to determine the neighboring buildings, based on visibility, adjacency, and the other geographic (geometric) relations among houses, streets, and open spaces and are implemented within infrastructure GIS. We assume that the influence of a city's global structure on an agent's properties increases with the level of residential segregation according to this property. Residential segregation is recognized and described in the model by means of local measures of spatial autocorrelation.

Simulation experiments are held for neighborhoods in the southern part of Tel-Aviv-Jaffo- with its heterogeneous population distribution. They demonstrate that if we introduce sufficiently strong tendencies of individual agents to depart from alienating neighbors, and to resettle in more fitting neighborhoods, then the likelihood clustering of agents of similar properties are establishing. In this way we can easily obtain spatial clusters of individuals of similar levels of income, origin, etc. The extension and form of the clusters is essentially influenced by the local topology of a city space. A distance-

independent search entails non-diffusion dynamics of urban residential distribution. A relatively homogeneous group of agents can be established in the model far from the current residence of the majority of the group members.

BERGER, Henry and SHINE, James (jshine@tec.army.mil), U.S. Army Topographic Engineering Center, 7701 Telegraph Road, Alexandria, VA 22315-3864

Estimating Subpixel Geospatial Features

Key Words: small-scale geospatial non-uniformities, derivative-as-limit algorithms, radiance, irradiance

Geocomputational literature describes spatial data acquisition from motion video with pixel-level detail that aims at building 3-D models of ground surface features that may be used in applications to develop topographic maps for geographic information systems (GIS). This paper describes a method for estimatingsubpixel geospatial features from spatial imagery when the details of these features are too small, or have spatial nonuniformities that vary so rapidly, that normal pixel-level spatial imagery has insufficient resolution to determine these special features.

The paper "Spatial Data Acquisition From Motion Video," by M. Williams, in Volume II of the proceedings of GeoComputation 1996, lists and differentiates between a number of methods used to obtain spatial data for GIS from imagery. While conventional photogrammetry requires a widely spaced pair of images, the method employing motion video proposed in that paper, which is a form of photogrammetry that avoids the correspondence problem, uses a dense image set; thus, the availability ofsubpixel information may lead to more accurate, detailed results.

The current paper describes a method that estimates a mini-image within each pixel. Normally an image consists of an orderly array of pixels. Each pixel in an image is seen as a single dot on a display unit. This corresponds to an average value of all that is seen within the nonzero-sized instantaneous-field-of-view (IFOV) of a single detector, or smallest element of film represented in a digitization process. Such detectors or film elements represent the most basic sensing element within an array of elements that comprise the total optical sensor.

There may be many reasons for concern about representing all that is seen within the IFOV by a single averaged numerical value for each wavelength band. Some examples are when there are fine surface features, relatively rapid changes in ground composition causing corresponding changes in ground reflectance, relatively small structures and objects, or relatively abrupt changes in ground surface altitude (such as may occur with hills, buildings, and mountains).

The November 1998 Optical Engineering article, "Data Analysis Systems for Films," by D. Zhang, et al., on subpixel estimation for sequential spatial imagery of moving objects recorded on film, then converted to digital imagery, uses a correlation technique to better locate a small number of key spatial points in the image. The current

paper describes a method that estimates, for the region within each pixel/IFOV, spatial coordinates of effective centers'-of-gravity of the optical illumination based on illumination data from neighboring pixels/IFOVs.

These, in turn, form the basis for setting up local coordinate systems in which interpolation based on data from neighboring sensing elements are used to estimate the illumination intensity at each spatial point within each IFOV for all of the detectors or film elements in the array. This all can be done recursively. Center-of-gravity techniques have been used very successfully in sky-viewing optical trackers tracking a single source of illumination with some instruments yielding accuracies better than 1/100 of a pixel; however, this seems to be the first method for ground-viewing sensors that require determination of centers-of-gravity within each pixel.

BOSCH, Edward H. (edward.h.bosch@usace.army.mil), U.S. Army Topographic Engineering Center, 7701 Telegraph Road, Alexandria, VA 22315-3864

Texture Information and Supervised Classification of Hyperspectral Imagery by Means of Neural Networks

Key Words: neural networks, hyperspectral imagery, supervised classification, texture measures, spatial information

In numerous studies, including those of the U.S. Army Topographic Engineering Center (TEC), researchers have used the spectral dimension of hyperspectral imagery for the purpose of classifying such data sets. In some cases, hyperspectral data sets are made up of hundreds of narrow spectral bands. The totality of the useful spectral bands are treated as a high-dimensional vector, which may be modeled by means of different mathematical and statistical methods. In the case of a supervised classifier, a very simple and often used method consists in obtaining the spectral mean for each of the training classes, and then computing the angle between the spectra in question and each of the corresponding spectral means. Since thespectra is being treated as a vector, the above classifier literally corresponds to the inverse cosine of the inner product of two vectors. The sample being classified is assigned to the class associated with the smallest angle. Each spectral sample (vector) also is normalized.

At TEC, we have developed and extensively employed a version of the back propagation algorithm to obtain neuro-spectral models, which are used to classify hyperspectral imagery. In this study, we investigate the potential of neural networks to discriminate texture information using hyperspectral data. The data associated with texture will be obtained from one of the bands of the hyperspectral data set and will be used in both the training and classification phases. Since texture is related to spatial information, one of the questions thatneeds to be addressed is the pixel size of the search window (kernel).

Clearly, if the size of the window is just one pixel in length and width, the amount of data in this window is not sufficient to be considered texture information.

Nevertheless, this case is not unreasonable if we use the remaining spectral bands to perform the analysis; however, this corresponds to the case of classifying hyperspectral imagery with spectral data as opposed to spatial data. Conversely, if the size of the window is too large, the network may not be able to correctly classify the texture data when there is more than one type of texture within the window. Although this situation is inevitable, a way to circumvent this problem may be provided by generating a texture class, which is a combination of at least two textures.

Since contiguous hyperspectral bands are so highly correlated, we will determine how well a neural network, obtained from one band, classifies such contiguous bands. This will help us further analyze surfaces whose textures are similar but whose mean amplitudes are different.

CARRANZA, Emmanuel John M. (carranza@itc.nl), HALE, Martin, Mines and Geosciences Bureau, Region 5, Philippines, PRESENT ADDRESS: International Institute for Aerospace Survey and Earth Sciences (ITC), Kanaalweg 3, 2628 EB Delft, The Netherlands

Geologically-Constrained Probabilistic Mapping of Gold Potential, Baguio District, Philippines

Key Words: Bayesian probability theory, spatial correlation, mineral potential mapping, Baguio Gold District

An approach using the Bayesian probability theory for mapping mineral potential is demonstrated in the Baguio mineral district of the Philippines. The approach involves extraction of binary predictor geological patterns based on the quantified spatial correlation between a number of geological features and the locations of mineral occurrences.

Whereas in its simplest form the binary code of a unit cell or pixel represents presence or absence of a mineral occurrence, the binary pattern of a predictor curvilinear geological feature represents the presence or absence of mineral occurrence points in unspecified cell(s) or pixel(s) within the area of the pattern. The areal extent of such patterns needs to be set, for example by using appropriate cutoff distance away from the predictor curvilinear geological features.

Using Bayes' rule, two probabilities can be computed in which a binary predictor pattern contains a mineral occurrence. The loge of each of these probabilities are the weight for the binary predictor pattern present, W+, and the weight for the binary predictor pattern absent, W-, respectively. If a binary predictor pattern is positively correlated with mineral occurrence points, W+ is positive and the contrast, C=3DW+-W- is a measure of the spatial correlation between a geological feature and a set of mineral occurrence points. The Studentized C (i.e., the ratio of C to its standard deviation) provides the basis for determining cutoff distances from the linear and curvilinear geological features when these features are converted into binary predictor patterns. The

loge of the posterior odds of a mineral occurrence given the presence/absence of a binary predictor pattern is then obtained by adding the weights of the binary predictor patterns to the loge of the prior odds. Combining the binary predictor patterns results in a map of posterior odds, which when converted to posterior probabilities, represents favourability for mineral potential. Combining the binary predictor maps assumes that these maps are conditionally independent from one another with respect to the mineral occurrence points. Conditional independence is tested through pairwise calculation of chi-square values and, from the map of posterior probabilities, by comparing the predicted and observed number of mineral occurrences.

The Baguio datasets consist of (1) geological maps, (2) lineaments representing mapped faults/fractures, (3) large-scale gold occurrences data, and (4) small-scale gold occurrences data. The results show that the input binary maps of importance for predicting known large-scale and small-scale gold occurrences are (1) proximity to Late Miocene - Pleistocene intrusive complexes, (2) proximity to the Late Oligocene - Early Miocene Agno Batholith, (3) proximity to NE-trending lineaments, (4) presence of the Zigzag and Pugo Formations, and (5) proximity to NW-trending lineaments. Pairwise chi-square tests show that the binary predictor patterns are conditionally independent with respect to the large-scale gold occurrences. On the other hand, the map of proximity to intrusive complexes and the map of proximity to the Agno Batholith show conditional dependence with respect to the small-scale gold occurrences. However, the resulting map of posterior probabilities indicates that conditional independence is not violated when these maps are used. The resulting map of posterior probabilities based on the large-scale gold occurrences shows that 79 percent of the known large-scale gold occurrences are associated with zones having posterior probabilities greater than, or equal to, the prior probability. In this map, 58 percent of the known small-scale occurrences are associated with the zones of gold potential. The resulting map of posterior probabilities based on the small-scale gold occurrences shows that 70 percent of the known small-scale gold occurrences are associated with zones having posterior probabilities equal to, or greater than, the prior probability. In this map, 56 percent of the known large-scale gold occurrences are associated with the zones of gold potential. Both predictive maps indicate that the probabilistic approach to mineral potential mapping is effective for the Baguio district datasets in that they predict several zones of gold potential that are close to the known gold occurrences.

CHALASANI, Venkat, (vsc4d@virginia.edu) and BELING, Peter (beling@virginia.edu), University of Virginia, ATTN: Systems Engineering, Olsson Hall, Charlottesville VA 22903

Road Extraction from Digital Images Using Linear Programming and Cost Functions

Key Words: road extraction, linear programming, classification, cost functions

Remote sensing data of the Earth's surface is readily available in digital format. These data can be used for identifying certain features of interest in the image with the assistance of computers. To identify a feature of interest we not only have to classify individual pixels as belonging to a specific class, but also identify a set of such pixels as a part of the feature. We refer to this process as feature extraction. Roads have certain interesting properties that can help us in extracting the roads from the image. This can be practical in terms of creating and updating digital maps.

In this paper we propose a two-step process for extraction of roads and demonstrate the results from an Airborne Visible Infra Red Imaging Spectrometer (AVIRIS) image of an area near Williamsburg, Virginia. The first step in our methodology is a classification step in which the bands of the image are treated as attributes. We create a training set by manually identifying certain coordinates as belonging to a specific class, such as road or soil, and then list the attributes for the specific coordinates obtained from the bands of the image. We use linear programming to create a discriminating surface between different classes. The linear program is formulated to find a surface in which all the points of one class in the training set are on one side of the discriminating surface, and all points of the other classes in the training set are on the other side of the surface. If it cannot find such a surface, it finds a surface that minimizes the average sum of errors, where an error is the distance to the separating surface if a point falls on the wrong side of the surface. By separating out the classes one after another, the separating surfaces form a decision tree. The decision tree is then applied to all the pixels of the image to identify pixels belonging to the class road. We use the distance to discriminant surfaces for each pixel to assign it a probability measure of being a road, which is then converted to a gray scale value and is used to rebuild the image with only roads. We also show results of using LP for feature selection in a nearest neighbor-based approach.

As a second step, certain properties of the roads, such as their elongation and contrast to the adjoining pixels, and the grayness measure created as a result of step one, are used to create cost functions. The function has a lower value if the configuration of pixels is closer to what we would expect from a road, and is higher otherwise. We then use dynamic programming to trace a minimum cost path through the pixels of the image produced by classification. We show that this approach is effective in tracing roads in a road network.

CHEN, Weiso (w.chen@mluri.sari.ac.uk), Macaulay Land Use Research Institute, Craigiebuckler, Aberdeen, AB15 8QH, U.K.

3-D City: Prototyping Techniques for Urban Design Modeling

Key Words: visualization, urban design, 3-D modeling, choice experiment

Architects and planners have recognized the advantages of computer-aided design (CAD) packages in their daily practices. Many CAD packages are now necessities in

most design firms and planning authorities, replacing drawing boards. Use of CAD systems do improve practitioners' work in many aspects, however, the essence of the design and planning work does not change. Theuse of CAD packages transform a designer's desk into an electronic drawing board, and more.For example, visual aid techniques and visual impact analysis have long been used to assist professionals such as urban designers, town planners, and landscape architects in their design and planning practices; however, traditional visual aid techniques such as physical modeling, perspective drawings, photo montage, and visibility determination have proven either inaccurate or too costly to manage. What makes it even worse is that once a model is built, any changes to the original design cannot be easily applied to the already constructed model. These shortcomings can be overcome because of the development of computer hardware and software along with advances in spatial database design and structure which, we believe, will fundamentally change the way designers work.

The new technology, dynamic real-time 3-D modeling, as an emerging modeling technique, has been used in scientific and engineering fields for more than two decades to assist scientists in getting insight into object relationships that otherwise cannot be revealed. The 3-D trend is influencing the urban design community as well, as a promising 3-D, real-time interactive urban design modeling technique. This paper presents an outline of an urban design project aiming at using the state-of-the-art 3-D modeling techniques with an option of choice experiment methodology to assist local enterprise organizations to develop strategic plans to revitalize local economy. We conclude this presentation with future work and directions.

CHRISTAKOS, George (george_christakos@unc.edu), SERRE, Mark (marc_serre@unc.edu), University of North Carolina, 111 Rosenau Hall, CB# 7400, DESE, Chapel Hill, NC 27599-7400

Modern Enviroinformatics: BME Mapping in Light of Uncertain Physical Knowledge Bases - The Equus Beds Case Study

Key Words: spatiotemporal mapping, stochastic, uncertainty, enviroinformatics, terrain systems

This work is concerned with the Bayesian maximum entropy (BME) method, which can rigorously and efficiently handle space/time mapping applications of considerable practical importance. BME, which belongs to the field of Enviroinformatics, can integrate and process physical knowledge that belongs to two major bases: general knowledge (i.e., obtained from general principles and laws, summary statistics, and background information) and specificatory knowledge (i.e., obtained through experience with the specific situation). BME allows considerable flexibility regarding the choice of an appropriate spatiotemporal map, offers a complete assessment of the mapping uncertainty, and contributes to the scientific understanding of the underlying natural

phenomenon. Valuable insight is gained by studying a spatiotemporal data set representing water-level elevations at theEquus Beds aquifer (Kansas). Numerical results show that, as was expected in theory, classical geostatistics results are obtained as special cases of the BME approach. In addition, a more accurate and informative analysis is possible by incorporating various sources of physical knowledge that cannot be processed by classical geostatistics methods.

COVA, Thomas J. (cova@geog.utah.edu), University of Utah, Department of Geography, 260 S. Central Campus Dr., Room 270, Salt Lake City, UT 84112-9155; and CHURCH, Richard L., (church@geog.ucsb.edu), University of California Santa Barbara, Department of Geography, Santa Barbara, CA 93016

Exploratory Spatial Optimization and Site Search: A Neighborhood Operator Approach

Key Words: site search, spatial optimization,exploratory spatial data analysis

This paper describes a framework for searching a discrete representation of geographic space for an optimal site of finite area in which to locate a particular land use or activity. The approach relies on a point-to-area neighborhood operator that is capable of identifying the best contiguous site that includes a given point in a study area, where the area and shape of the site may vary. The framework represents a synthesis of concepts drawn from cartographic modeling (Tomlin, 1990), optimal land allocation (Wrightet al., 1983), and land suitability mapping (Lyle and Stutz, 1983). The result of applying an optimal site search neighborhood operator to every point in a study area is a mathematical field of best site boundaries, one for eachpoint, that we refer to as a site field. A site field does not lend itself to static visualization in two dimensions, and we describe a means for interactively exploring one. A significant benefit of adopting a neighborhood operator approach in a discrete geographic data modeling context is that it serves to spatially decompose the global site search problem into a set of smaller local problems; thus, the computational effort required to solve an optimal site search problem at a point is not tied to the number of spatial units in the global data set. Using this approach, we can efficiently search large spatial data sets for an optimal site in which to locate a land use or activity.

References

Diamond, J.T., and J.R. Wright, 1989, "Efficient Land Allocation," Journal of Urban Planning and Development, v. 115, pp. 81-96.

Lyle, J., and F.P. Stutz, 1983, "Computerised Land Suitability Mapping," The Cartographic Journal, v. 20, pp. 39-49.

Tomlin, C.D., 1990, "Geographic Information Systems and Cartographic Modeling," Englewood Cliffs, Prentice Hall.

CROWTHER, Paul (P.Crowther@utas.edu.au), HARTNETT, Jacky, WILLIAMS, Ray, PENDELBURY, Steve, University of Tasmania, School of Computing, P.O. Box 1214, Launceston, Tasmania, 7250, Australia

Developing Quality Training Data for a Statistical Decision Tree Classifier in a Spatial Environment

Key Words: training data, remote sensing, statistical classifier

One of the strengths of a remotely sensed data set is that it represents a complete spatial population; however, in order to make sense of this population, the most successful classifiers first require that areas of the image be selected as training data. This paper describes a method that solves the problem of selecting suitable training sites when it is not practical to select these using ground truth. The method was developed to provide a training data set for a statistical decision tree classifier toanalyse the National Oceanic and Atmospheric Administration's (NOAA) Advanced Very-High Resolution Radiometer (AVHRR) multispectral satellite images of Antarctica. This was in response to a user request for a tool to develop a data set for a statistical decision tree package, "S-Plus." It was found that the same tool could be used to create training data for use with other systems.

Statistical classifiers are dependent for their accuracy on the quality of the training data rather than on the algorithm used for classification. The training data set needs to be representative of the whole area to be classified. The populations of pixels used for training must be statistically significant. This means that there is a need to know the minimum number of observations required to characterise a particular site to an acceptable level of error.

Ground-truthed training data in Antarctica is difficult to obtain. In other domains, for example, vegetation classification training data are developed using ground truth combined with aerial and satellite interpretation to identify representative vegetation types. Vegetation is relatively stable over time. In Antarctica there can be rapid changes in features such as sea ice that cannot easily be verified on the ground; therefore, training data must be developed directly from satellite images.

The tool allows an expert image interpreter to choose sample points on an image and apply a label to each one. The label and pixel values for that point on all bands are then stored. The tool's operation gives the user the choice of automatically sampling an image or allowing a user to sample the image in a directed way. In the former case the user can specify the coarseness of the sampling grid and then is asked to supply names at each sampling point. The user can switch between image bands to help decide on the label to be attached to the sampling point. It is possible to switch between automatic

sampling and user-directed sampling. This feature was added because if a coarse grid was chosen for automatic sampling, some important features could be missed.

This paper will describe the operation of the spatial sampling tool and the results of using the sample to develop a decision tree by submitting this sample to the S-Plus decision tree package. The results of applying this decision tree to classify Antarctic sea-ice will be presented.

CURRAN, Paul J. (P.Curran@soton.ac.uk), Department of Geography, University of Southampton, Highfield, Southampton SO17 1BJ, U.K.; DAWSON, Terry P. (terry.dawson@ecu.ox.ac.uk), Environmental Change Unit, University of Oxford, Mansfield Road, Oxford OX1 3TB, U.K.; WICKS, Toby E. (tew@soton.ac.uk), Department of Geography, University of Southampton, Highfield, Southampton SO17 1BJ, U.K.

Remote Sensing and Process-Based Modelling

Key Words: remote sensing, process-based models, LIBERTY, FLIGHT, FOREST-BGC, BOREAS

The use of remote sensing in ecology is predicated both on a need to understand the interaction of radiation with vegetation, and to couple this with aspects of environmental understanding. A key to the ecological use of remote sensing has become the application of process models that enable us to understand the interaction of radiation with a leaf and a canopy and to use remotely sensed data to drive an a-spatial ecosystem simulation model over large areas of terrain. By way of illustration, this paper will focus on the application of three process models to contemporary ecology. The first is the Leaf Incorporating Biochemistry Exhibiting Reflectance and Transmittance Yields (LIBERTY) model that can be used to determine the effect of key variables (e.g., water or chlorophyll content) on leaf spectra. The second is the Forest LIGHT interaction (FLIGHT) model that can be coupled with LIBERTY to determine the effect of key variables (e.g., water or chlorophyll content) on canopy spectra. The model is the FOREST-Bio-Geo-Chemical (FOREST-BGC) model that enables, via the coupling of remotely sensed data with ecosystem understanding, the spatial estimation of vegetation productivity.

The utility of these three models will be illustrated, by reference to research undertaken, as part of an international environmental research programme called BOReal Ecosystem Atmosphere Study (BOREAS). Specifically the LIBERTY model will be used to investigate the spectral sensitivity of conifer needles to changes in foliar biochemical content. The combined FLIGHT and LIBERTY models will be used to determine the degree to which foliar biochemical information is retained in the spectrum of a conifer canopy. The FOREST-BGC model will be driven, using leaf area index estimates from the National Oceanic and Atmospheric Administration's (NOAA) Advanced Very High

Resolution Radiometer (AVHRR) data and used to estimate the net primary productivity of boreal forest.

The paper ends with discussions of themodelling approaches mentioned above, to the understanding and use of remotely sensed data from European Space Agency's (ESA) Envisat satellite.

DEBOER, Dirk H., (deboer@duke.usask.ca), University of Saskatchewan, Department of Geography, 9 Campus Drive, Saskatoon, Saskatchewan, S7N 0J1, Canada

Self-Organization in Fluvial Landscapes: Sediment Dynamics as an Emergent Property

Key Words: cellular automata, emergent properties, complexity, fluvial erosion, sediment load

Landscapes are the end product of the interaction of form and process at a variety of temporal and spatial scales. As a consequence, landscapes are inherently complex. Nevertheless, at least part of the complexity of a landscape arises from processes that follow very simple rules such as: the pull of gravity results in a down slope transfer of water and sediment; a steeper slope angle results in a faster flow of water and erosion; and sediment deposition occurs when the slope angle decreases. One of the effects of these rules is that when they are applied at small scales, the resulting landscape has large-scale properties such as drainage density and drainage network configuration. These emergent properties are not part of the basic, small-scale rules but, instead, result from the application of these rules and the ensuing self-organization of the landscape.

This paper discusses a cellular model of the long-term evolution of a fluvial landscape. The model is set in motion by selecting a cell at random and applying rainfall to this cell. Erosion takes place as the water moves as runoff to the lowest neighboring cell. During the next time step, the runoff is routed from the lowestneighbour to its lowest neighbour, and in addition, a new cell to receive rainfall is selected at random. In the model, sediment is routed downslope according to a transport equation with the transport rate proportional to the slope to the power of an exponent; thus, the model allows both erosion and deposition of sediment depending on the difference between the sediment input and output of a cell. The model allows runoff from cells to converge resulting in increased sediment transport rates downstream. Starting with a block-faulted landscape, over time a drainage network evolves. Sediment export records of the drainage basins display a complex behaviour, even though there are no external factors that would explain the variations in sediment export. The complexity of the sediment dynamics in the model arises from self-organization within the modeled system itself. Studies like these are a first step towards separating the impact of this aspect of complexity on the sediment export and depositional record from the impact of external factors associated with global change.

DEHN, Martin (martin@giub.uni-bonn.de), GAERTNER, Holger (holger@slide.giub.uni-bonn.de), and DIKAU, Richard (rdikau@slide.giub.uni-bonn.de), University of Bonn, Department of Geography, Meckenheimer Allee 166, D-53115 Bonn, Germany

Principles of Semantic Modeling of Landform Structures

Key Words: landform structure, semantic modeling,geomorphometry, scale

Landform as a result of various processes acting on the surface is also seen as a static boundary condition for processes in geomorphology, hydrology, meteorology, and others. The description, parameterization, and modeling of landform structures, as well as the definitions of nomenclatures, are fitted to the requirements of the mentioned disciplines, and therefore are often strongly divergent. As a consequence, representations of landform structures of different disciplines often are not compatible and require frequent revisions and adaptations. Being aware of these difficulties, an interdisciplinary Postgraduate School (Graduiertenkolleg) on landform structures was established at the University of Bonn. Within this school,geoscientific disciplines are working together with mathematics, photogrammetry, and computer science in order to derive a comprehensive insight into some of the problems. The presented study is carried out in the frame of this Postgraduate School.

Principles of the semantic approach will be presented in this paper. The main objective is a semantically correct description of landforms that is useful to all disciplines related to surface structure.

The presented approach is not based on traditional terms, such as peak, valley, or hillslope, in the initial phase of definitions; in contrast, it is focusing on geometric forms as a basic property in a much more objective way. Form, position, and the spatial arrangement of forms (topology) are considered. The hierarchical approach is starting with form facettes that can be clustered to form elements based on specific rules. A larger set of form elements can be grouped to landform associations depending on topological conditions that have to be defined based on the theoretical background ofgeoscientific disciplines. While geometric forms and parts of positions are already defined to a certain degree, topology and the structure of landform componentsare still a field of research. The outlined basic approach enables the development of a modular modeling framework in which the focus is on semantic representation in the first place and, only thereafter, special features of DEMs, tools, and algorithms have to be considered.

Scale is not explicitly addressed in the basic semantic representation of landform structure. It is discussed if this provides a possibility to apply the concept from small to large scales.

DOKKA, Roy K. (rkdokka@biotite.geol.lsu.edu), Louisiana State University, Department of Geology and Geophysics, Baton Rouge, LA 70803; WATTS, Joseph M., (watts@tec.army.mil), U.S. Army Topographic Engineering Center, Topographic Research Division; LADUE, Denise R. (dladue@tec.army.mil), Operations Division, Alexandria, VA 22315

Spatial Pattern Analysis of Compositional Landforms in the Mojave Desert, California

Key Words: spatial pattern analysis, surface composition, geomorphic landforms, Mojave Desert

The Louisiana State University and the U.S. Army Topographic Engineering Center (TEC) have recently mapped geomorphic landforms, surface compositions, and geologic ages of the California portion of the Mojave Desert in support of the Mojave Desert Ecosystem Project. Using an integrated combination of spaceborne spectral scanners, airphoto interpretation, and geological field techniques, we have delineated ~30,000 polygons at a minimum mapping unit of 10 hectares for an area ~150,000 km^2. These GIS coverages are intended to impact a wide range of public and private sector functions including decision support for regional land managers, mineral exploration, education, and land use planning.

A detailed, regionally uniformdataset, such as this product, also contains a wealth of spatially explicit information that can be used for hypothesis construction and testing by research scientists. The objective of this paper is to present an exploratory spatial pattern analysis of Mojave Desert landforms to facilitate future scientific inquiries regarding causal relationships between tectonic process and geomorphic pattern. Our focus is on the identification and quantification of similarities and differences in physiographic composition (number of categories, proportions, diversity) and configuration (size, shape, adjacency, sinuosity) at sample locations. Hydrologically defined landscape units will be selected in the tectonically dissimilar western, central, and eastern sections of the Mojave Desert. For each sample, standard metrics of physiographic heterogeneity will be computed and compared for statistical similarity. In addition to tabular analysis, cartographic visualizations of the compositional landform database will be presented.

DOUGHTY, Jonathan (jwd@mitre.org), The MITRE Corporation,1820 Dolley Madison Boulevard, McLean, VA 22102

Interoperable Geospatial Objects

Key Words: integration, metadata, interoperability, expert system, distributed computing

Geospatial interoperability, the ability for two or more heterogeneous data sets to interact with one another, is increasingly an issue for users of spatial data. Commercial and Government sources are making increased amounts of geospatial data and applications available, in many cases outside the distribution channels that heretofore provided control on data misuse in the past. Emerging products are designed for use by non-GIS professionals who are generally unaware of the complexities of map projections, datums, scale, topology, and accuracy.

The research project that this paper describes is developing a geospatial information interoperability envelope to surmount these problems: an object that contains the methods that can act upon geospatial information in addition to the information itself. This envelope contains knowledge about the geospatial information that it accompanies and can ensure the data's integrity and pedigree. The envelope and the information it accompanies becomes "Interoperable Geospatial Objects" that can migrate from system to system.

This research seeks to demonstrate the integration of independent geospatial and imagery components by making use of emerging distributed objects and web technologies. The first stage of the research has focused on developing an interoperability model for geospatial data use: a mechanism to validate integration operations in the context of intended use and metadata about geospatial data. Geospatial objects in the current prototype are made up of metadata about the underlying data sets combined with the interoperability characteristics that are validated by the prototype as being an appropriate use of the data given the user's context.

Underlying the research prototype are a small but growing set of software geospatial "experts" that understand interoperability problems and attempt to warn the user when integration problems, e.g., mismatched datums, projections, or inappropriate scale changes, are identified.

DOWERS, Steve (sd@geo.ed.ac.uk), GITTINGS, Bruce and MINETER, Mike, University of Edinburgh, Department of Geography,Drummond Street, Edinburgh, EH8 9XP, U.K.

Towards a Framework for High-Performance Geocomputation: Handling Vector-Topology within a Distributed Service Environment

Key Words: high-performance computing, vector topology,OpenGIS, software libraries, service architectures, algorithms

High performance computing has been shown to be a critical technology to realise current and future spatial data handling requirements. While commercial parallel processing approaches within DBMS are providing the performance for underlying integrated data models, the acceleration of spatialanalyses remains to be addressed. This paper investigates the scope for a framework that integrates parallel implementation within a service architecture such as that proposed by the Open GIS Consortium.

The application of parallel processing to spatial analyses requires flexible strategies for decomposition of data if a scalable solution is to be achieved. Such strategies are relatively well-known for grid, raster, and digital terrain data, and for the spatial analysis of point data. These techniques are not established for vector-topological data. Furthermore, the application of parallel processing to geocomputational algorithms has been limited to specialists and has not made the transition to the operational area of the end-user as has happened with database technology. Previous work by the authors has explored the issues involved in designing a generic approach to the handling of vector topological data in parallel and to the development of parallel libraries that encapsulate the complexity of the parallel implementation. The goal has been to allow the algorithm implementor to concentrate on the specifics of themodelling process, and also take advantage of the performance gains from parallel architectures in a portable manner.

Some success in this approach has already been reported in previous publications, while current projects are testing implementations of potential components of a parallel library concerned with input, distribution, collation, and output. Central to the details of these implementations is the choice of data models and the representations of features, geometry, and the topological relations between them. Existing work by the authors has been based on the use of Neutral Transfer Format (NTF) Level 4 model which is part of a British Standard. Current work is exploring the issues of the interface between parallel applications and other systems and services. These systems may include thin clients, spatial analysis engines, and databases with spatial capabilities. TheOpenGIS Abstract Specification Model addresses many of the issues associated with these interfaces, including data models and data exchange. This paper will explore the implications of this and related models for our existing design work, and will identify the role of parallel libraries and high-performance applications in the general framework of a service architecture. Debates are ongoing as to the benefit of stored versus generated topology with integrated spatial data stores; we shall explore the generic nature of algorithm designs and the extent to which they can apply in these differing contexts.

References

British Standards Institution (1992) Electronic Transfer of Geographic Information (NTF, BS 7567, British Standards Institution, London, U.K.

Healey, R., S. Dowers, B. Gittings, and M. Mineter, (1998) Parallel Processing Algorithms for GIS, Taylor & Francis, London, U.K.

Mineter, M.J. and S. Dowers (in press) Parallel Processing for Geographical Applications: A Layered Approach, Journal of Geographic Systems.

Open GIS Consortium (1998) The OpenGIS Abstract Specification Model, http://www.opengis.org/techno/specs.htm

DUFFY, Christopher J. (cxd11@psu.edu), SEDMERA, Kris, Pennsylvania State University, Civil and Environmental Engineering, and SHUN,Tongying, Geoscience Department, 212 Sackett Building, University Park, PA 16802

A Stochastic-Dynamical System for River Basin Runoff in Complex Terrain: A Case for Low-Dimensional, Physically-Based Models

Key Words: stochastic forcing, runoff, physiography,dynamical systems

This paper develops a strategy for the formulation of physically-based, low-dimensional models of precipitation-storage-runoff in regions of complex topography or drainage density. The goal of the research is to improve river flow and stage forecasting by explicitly including the physical processes of soil moisture and groundwater in addition to surface and channel processes. The basic concept is that soil-moisture and groundwater storage form an extended dynamic state-space forced by stochastic weather and climate conditions, and that runoff is a nonlinear function of these states (storage). The general model is in the form of an n-dimensional, nonlinear, discrete dynamical system. For a given river basin, the particular state-space dimension (n) is estimated from the eigenspectrum and dominanteigenmodes of historical river discharge, soil moisture, groundwater level and/or remotely-sensed time series. The regional physiography and conceptual geohydrology are used as a basis to qualitatively interpretspatio-temporal modes found in the data. Once formed, the parameters of the dynamical system are estimated using the genetic algorithm and the stability of the model is examined. Data analysis, calibration, and validation of runoff forecasts from river basins within diverse physiographic regions and space-time scales are presented including: Appalachian Plateau (PA), Valley and Ridge (PA), and Basin and Range (Utah). Results demonstrate that runoff records are generally well-described by a low-dimensional model provided that the dominant state variables (e.g. time and space scales) are incorporated.

DUKE-WILLIAMS, Oliver (oliver@geog.leeds.ac.uk), University of Leeds, Centre for Computational Geography, Leeds LS2 9JT, U.K.; and BLAKE, M. (mblake@gisca.adelaide.edu.au), University of Adelaide, Key Centre for Social Applications in GIS, 10 Pulteney Street, Adelaide SA 5005, Australia

Database Fusion for the Comparative Study of Migration Data

Key Words: migration, time-series, WWW, visualization, distributed databases, demographic modelling

Internal migration is one of the key processes influencing the population of small areas, together with the natural agents of birth and death. In most post-industrial countries it is both the most significant and the least predictable component of change. The study of migration is often made difficult by the particular characteristics of data collected- alternative datasets may not be directly comparable, data may be of limited extent, or where data does exist, it may do so in sufficiently large volumes so as to make analysis daunting. In addition, both the data themselves, and the results of analysis, can be complex, and require large amounts of metadata to support their explanation.

This paper describes a software system designed to facilitate the comparative study of two time series of migration data, respectively, detailing patterns of internal migration in Australia and the U.K. Both datasets stretch over a concurrent 20-year period (1976-1996), but they are different in conceptual structure: the Australian data describes migration transitions over four inter-Censal 5-year periods, whereas the U.K. data consists of 20 years of migration event data compiled from administrative registers. A database has been designed that contains both source data and data transformed through demographic models, which is intended to generate migration statistics, and measures which can be directly compared, in order to permit the study of features and patterns of change in migration in the two countries, with the primary aim being the elucidation of age-period-cohort relationships in the two sets of data.

The system is constructed from web servers coupled via server-side scripts to RDBMS servers in a manner that is intended to be flexible in terms of hardware and software used, and also to allow servers in the two countries to offer complementary services. The use of a hypertext system allows familiar interfaces to be offered to the user, and as much supporting explanatory text as is needed.The system is designed to support a wide variety of requirements ranging from the extraction of subsets of data to be imported into external analysis packages, and the generation of statistics describing the intensities and effects of migration for selected data to visualization of the data for speculative and exploratory analysis of the data. Migration data is multidimensional, thus, visualization through means such as VRML models allows a rich summary of the data to be presented to the user. This multidimensional nature of the data includes both demographic detail and derived information such as measures of change over time, and also geographic detail. The effects of geographic scale are of considerable importance when studying migration data, and it is presumed that causal factors have differing priorities for local and long distance moves. The system is designed to allow re-aggregation of the migration interaction matrices to suit the users requirements.In addition, many measures of migration intensity and impact rely on distance as a component, and studies have shown that alternative calculations of distance between origins and estimations can have very significant affects on the measures subsequently calculated; the system has been designed to accommodate a variety of methods for calculating distances between locations.

The system described in the paper uses data fusion and integration techniques, together with a variety of presentation approaches, to offer new analytic tools for researchers, allowing large volumes of data to be studied and compared in a manner that was not previously available.

GeoComputation 99

4th International Conference on GeoComputation
Mary Washington College
Fredericksburg, Virginia
July 25-28, 1999

EHLSCHLAEGER, Charles R. (chuck@everest.hunter.cuny.edu), Hunter College, Department of Geography, New York, NY 10021

Representing Uncertainty of Qualitative Thematic Maps with an Inter-map Cell Swapping Heuristic

Key Words: uncertainty, spatial data, thematic maps, area class maps, categorical coverage maps, data fusion, conflation, conditional simulation, unconditional simulation, geostatistics, statistical modelling, Monte Carlo simulation

Monte Carlo simulation, a technique that generates many versions of possible application results, is a popular method for representing application uncertainty. Monte Carlo simulation can be used to generate many sets of application input maps with each input map consisting of a possible version of reality. Representing the total distribution of potential application results during the Monte Carlo process requires two critical components: a thematic class probability model and a model of spatial autocorrelation for these same classes. This paper introduces a combination of techniques including an inter-map cell-swapping algorithm, a class probability model, and a new spatial statistic. Together, these techniques allow for the generation of spatially autocorrelated random qualitative thematic maps for the purpose of representing spatial application uncertainty.

This paper assumes a generalized thematic map of the study area is available. A generalized map, for the purpose of this paper, is a map containing qualitative thematic information at a resolution too coarse and/or attribute information too inaccurate to be useful for a particular application. Qualitative thematic maps also are known as "area class maps" or "categorical coverage" maps. Samples of application quality data must also be available. For the purpose of this paper, application quality data are qualitative thematic information accurate enough and with a resolution fine enough to achieve useful application results. Class information from the application quality data is used to determine the spatial statistics of all generated maps. Samples of application quality data can either be in the study area, or outside the study area if the application quality data intersects generalized thematic data. If the application quality data are inside the study area, the Monte Carlo process will be a conditional simulation, otherwise, the Monte Carlo process will be an unconditional simulation. Intersecting application quality data and the generalized thematic map allows a conflation technique defining an algorithm that will generate the probability vector for any generalized map cell. A probability vector is a set of values representing the likelihood each class will be located in that particular cell. This probability model analyzes each application quality data cell intersecting the generalized map by looking at the distance within a generalized map class polygon, as well as the distance away from the closest cell of other classes in the generalized map. During these analyses, the generalized map will be resampled to the resolution of the application quality data.

Spatial autocorrelation is measured with a new spatial statistic named densogram. A densogram measures the proportion of each application quality cell's class (MG1) at various distances from cells of the same class. A densogram and a variogram of a binary map provides the same measure while the densogram is computationally quicker when calculating statistic changes during the cell-swapping heuristic. The quality of a randomly generated Monte Carlo input map will be based on how closely its densogram matches the densogram of application quality data using a weighed least squares analysis.

The inter-map cell swapping heuristic begins with a set of random thematic maps. Each cell within each random thematic map is given a class value based on the probability vector defined by the location of class polygons in the surrounding generalized map. Cells are swapped at the same location from different maps if the resulting maps' densograms provide a better overall fit to the application quality data's densogram. This heuristic is continued until no more possible cell swaps are possible. The inter-map swapping technique described in this paper works for both conditional and unconditional simulation of qualitative thematic data because the proportion of class values within a cell will never change, preserving the probability model. The inter-map swapping technique is flexible, allowing for any spatial statistic to be used, whether global or regional, alone or in combination, with other spatial statistics.

Source code and sample data demonstrating the inter-map cell-swapping heuristic is included in this paper's appendix. The software is written in Java, and works for versions 1.1.7a and later. The software is part of the Research Geographic Information System (RGIS). RGIS is a public domain GISystem designed by the author as a research and educational tool. RGIS is used in Hunter College's geographic programming class to demonstrate object-oriented programming and will run on all commonly used workstations. RGIS allows GISystem raster data from Arc/Info and ArcView as inputs.

EVANS, Andrew J. (aevans@geog.leeds.ac.uk), KINGSTON, Richard, CARVER, Steve, TURTON, Ian, University of Leeds, School of Geography, Leeds LS2 9JT, U.K.

Web-Based GIS Used to Enhance Public Democratic Involvement

Key Words: GIS, cyberdemocracy, public participation

Increasingly, the World Wide Web is being used to disseminate information of a spatial nature; however, there are relatively few systems that allow for the public to both query and manipulate information, let alone submit their ideas on spatial problems to those with the power to enact solutions. Ideally, systems should be available for the public to participate in finding solutions to a wide variety of real-world spatial problems. Such systems should allow users to:

- examine objective bipartisan information, and meta-information, on a problem in a way that encourages comfortable exploration

- add their own information in multimedia formats, discuss the datasets, and form interested communities
- reach an informed decision, submit this to those responsible for implementing it, see the results, and receive feedback on the reasons for the final choice.

This paper outlines research aimed at distributing such systems to the public to aid in the solution of environmental problems with a spatial component. Initial results and future developments also are presented.

FINDLEY, Joseph H. III (jfindley@tec.army.mil) and HAES, Steven C.(shaes@tec.army.mil), TRW, Inc., 4301 North Fairfax Drive, Suite 640, Arlington, VA 22203

Image Understanding for Battlefield Awareness Video Surveillance and Monitoring

Key Words: image understanding, synthetic environments, video surveillance, synthetic theater of war, ModSAF, terrain database, VSAM, IU, SE, STOW

In conjunction with the Defense Advanced Research Projects Agency's (DARPA) Image Understanding for Battlefield Awareness-Video Surveillance and Monitoring (VSAM) project, the Software Experimentation, Evaluation and Demonstration Site (SEEDS) at the U.S. Army Topographic Engineering Center (TEC), Alexandria, Virginia, is extending current Advanced Distributed Simulation (ADS) capabilities so that VSAM researchers may benefit from state-of-the-art simulation technology. The objective is to develop a cooperative, multisensor video surveillance system that provides continuous coverage over urban and battlefield environments within the virtual world.

SEEDS is: (1) implementing real-world sensors and relevant behaviors into existing ADS technology (i.e., Modular Semi-Automated Forces (ModSAF) a behavioral tool and OpenScene ModStealth-a 3-D Visualization tool); (2) creating a robust communication conduit enabling a VSAM Operator Control Unit (OCU) to efficiently relay information on sensor(s) and target(s) geoposition, orientation, heading, and identification; and (3) constructing high resolution compact terrain databases (CTDBs) or synthetic environments (CMU Campus, Schenley Park, Pittsburgh, PA; Fort Benning, GA; etc.). Ultimately, this effort will enable VSAM researchers to conduct pre-mission experimentation, run-time visualization, and after-action review (AAR). Additionally, this effort provides VSAM researchers with valuable data on optimal sensor geolocation, performance, occlusion, and intervisibility predictions.

FISCHER, Manfred M. (Manfred.M.Fischer@wu-wien.ac.at), Department of Economic and Social Geography, Wirtschaftsuniversitat Wien, Augasse 2-6, A-1090 Vienna, Austria; and REISMANN, Martin, (reismann@wigeo1.wu-wien.ac.at), Institute for Urban and Regional Research, Austrian Academy of Sciences,Postgasse 7/4/2, A-1010 Vienna, Austria

Parameter Estimation in Neural Spatial Interaction Modelling by a Derivative Free Global Optimization Method

Key Words: neural spatial interaction modelling, parameter optimization, evolutionary compu-tation, backpropagation of conjugate gradient descent errors, real world application performance test

Parameter estimation is one of the central issues in neural spatial interaction modelling. Current practice is dominated by gradient-based local minimization techniques. They find local minima efficiently and work best in unimodal minimization problems, but can get trapped in multimodal problems. Global search procedures provide an alternative optimization scheme that allows escape from local minima. Differential evolution recently has been introduced as an efficient direct search method for optimizing real-valued multi-modal objective functions (Storn and Price 1997). The method is conceptually simple and attractive, but little is known about its behaviour in real-world applications. This paper explores this method as an alternative to current practice for solving the parameter estimation task, and attempts to assess its robustness, measured in terms of in-sample and out-of-sample performance. A benchmark comparison against back propagation of conjugate gradients is based on Austrian interregional telecom-munication traffic data.

FRITZ, Steffen (pgsf@geog.leeds.ac.uk), SEE, Linda, and CARVER, Steve, University of Leeds, School of Geography, Leeds, LS2 9JT, U.K.

A Fuzzy Modelling Approach to Wild Land Mapping in Scotland

Key Words: wilderness mapping, fuzzy logic,multicriteria evaluation

The use of GIS for wilderness mapping is a recent development, but several attempts have already been made that cover a range of different areas across the globe (Lesslie, 1988, 1995; Henry and Husby, 1994; Kliskey and Kearsley 1993, 1994; Carver 1996). Methodologies range from the mechanistic and rigorous approach taken by the Australian Heritage Commission in 1988, where the mapping was based on a number of deterministic, yet arbitrary, parameters, to the approach of Kliskey and Kearsley (1993), who have taken into account the subjective nature of wilderness by using Stankey's wilderness purism scale (Stankey, 1973). There are, however, drawbacks to the way in which Kliskey and Kearsley (1993) have translated the perceived levels of wilderness to

the spatial domain. Moreover, none of these methodologies are directly applicable to Scotland, where the term "wild land" is proposed as a better representation of the landscape (Aitken,1977) due to its long history of settlement and rural land use.

This paper describes an alternative approach to capturing qualitative perceptions of Scottish wild land using a fuzzy logic framework. Fuzzy logic is particularly well suited to wild land mapping because it is very difficult to determine a crisp boundary between wild and non-wild land with existing methods. Using fuzzy logic, wild land quality corresponds to the degree to which a certain point in the landscape belongs to a set of wild land quality membership functions with corresponding labels such as low, medium, and high. To generate these membership functions, an internet questionnaire has been designed for acquiring public perceptions of human impact within an otherwise wild landscape. These are then used in a fuzzy multicriteria evaluation (FMCE) for determining wild land quality in Scotland. Respondents are divided into different "purism" groups, and multiple perceptions of wild lands are mapped. It is shown that it is possible to quantify the wild land resource according to multiple perceptions of wilderness in addition to the more easily quantifiable natural resources such as ecology, land capability, and geology. For the purpose of establishing national parks in Scotland and for their potential zoning, this type of wild land area mapping will be a valuable and necessary technique. Moreover, the methodology is widely applicable and will be extended to other European countries in the future.

References

Aitken, R. (1977) Wilderness Areas in Scotland, unpublished Ph.D. Thesis.University of Aberdeen.

Carver, S. (1996) Mapping the Wilderness Continuum Using Raster GIS.Eds: S.Morain and S. Lopez Baros, Raster imagery in Geographic Information Systems, Onword Press, New Mexico.

Henry, D. and E. Husby (1994) Wilderness Quality Mapping in the Euro-Arctic Barents Region: A Potential Management Tool, http://www.esri.com/recourses/userconf/proc95/to150/p113.html

Kliskey, A.D., and G.W. Kearsley (1993) Mapping Multiple Perceptions of Wilderness in Southern New Zealand, Applied Geography, 13, 203- 223.

Kliskey, A. D. (1994) A Comparative Analysis of Approaches to Wilderness Perception Mapping, Journal of Environmental Management, 41, 199 - 236.

Lesslie, R. G., B.G. Mackey, and K.M. Preece (1988) A Computer-Based Method of Wilderness Evaluation, Environmental Conservation, 15, 3, 225 - 232.

Lesslie, R., and M. Maslen (1995) National Wilderness Inventory Handbook of Procedures, Content, and Usage (Second Edition), Commonwealth Government Printer, Canberra, Australia.

Stankey, G.H. (1973) Visitor Perception of Wilderness Recreation Carrying Capacity, Research Paper, INT-142.Ogden, Utah: USDA Forest Service.

GAERTNER, Holger (holger@slide.giub.uni-bonn.de), University of Bonn, Department of Geography, Meckenheimer Allee 166, D-53115 Bonn, Germany; BERGMANN, Andreas, University of Bonn, Institute of Computer Science III, Römerstrasse 164, D-53117 Bonn, Germany

Object-Oriented Modeling of Geodata as a Basic Tool for the Integration of Heterogeneous Paleoecological Information

Key Words: object-oriented modeling, UML, data integration, distributed databases, geomorphology

The increasing use of new techniques within Geo Sciences for capturing field data, and the ongoing development of geographic information systems (GIS), is coupled with a growing diversity of methods such as geomorphological mapping, analyses of sediment covers, hydrological modeling, or dating methods. Furthermore, capturing and analyzing data followed by knowledge-based interpretations of these analyses means high investment of equipment, manpower, and specific knowledge to gain new information about natural processes. The diversity of techniques and methods leads to a vast and further growing amount of heterogeneous data that are stored in diverse formats and in different databases, often even in a single research group. As a consequence, high investments have been made to create central databases to avoid the loss of data as well as to get the possibility to use these data for further research in the same or different research areas. Further specific query components were coupled to these databases for getting easy data access. The existence of several worldwide databases and query systems leads to a number of problems. First of all, the structure of different central databases is often diverse. Furthermore, this argument also applies to the query components. This means, to connect databases or even data sets, special transfer algorithms have to be developed, and the possibility of losing data in the sense of not getting information about their existence is high. An important precondition in order to facilitate integrated access to heterogeneous information is to adequately model and represent the semantics of the different source data sets. The paper focuses on our experiences gained using an object-oriented modeling technique to describe diverse data used in paleoecologic research. The work is carried out in the context of a joint research project, named OPALIS, by Geographers and Computer Scientists at the University of Bonn. The overall goal is to develop an open GIS architecture supporting uniform exploration of distributed

heterogeneous data pools. We present problems in sampling, documenting, and modeling both existing data and methods for data storage; therefore, several German research groups within the IGBP-PAGES program were selected as cooperative partners. The advantage of creating basic semantic models (including all metainformation) and representative object-oriented models for each group using the Unified Modeling Language (UML) will be discussed. The UML has been used as a "common language" between computer scientists and geo scientists in the project. The use of this modeling technique allows us to represent different data pools in basically unique models without changing the data sets themselves in any way, as must be done in common databases.

GAHEGAN, Mark (mark@geog.psu.edu) and TAKATSUKA, Masahiro, Pennsylvania State University, Department of Geography, 302 Walker Building, University Park, PA 16802

An Investigation Into the Use of Data Spaces as an Organisational Concept for the Classification of Geographic Datasets

Key Words: classification, neural networks, self organising maps, dataspaces

Traditional geography and environmental analyses use the idea of a data space in order to add structure to the understanding of complex natural systems. Adataspace is some subset of the overall attribute or feature space where some geographical theme is represented. Examples are spectral space, geographic space, and environmental space (Aspinall and Lees, 1994). It has been suggested that the relationships within these spaces are not the same as those between them, due to the inherently different structure of each separate domain.

Many types of statistical classifiers (such as maximum likelihood) and inductive learning techniques (such as decision trees and neural networks) treat the entire attribute space as comprising a single descriptive vector. In geographic terms this is equivalent to the construction of a single and all-inclusive dataspace. While this approach has obviously met with a good deal of success, it seems likely that the mixing of concepts has two distinct disadvantages, namely:

1. The classification tool has to effectively learn to disentangle data used in training. This increases the complexity of the learning task, which, in turn, may reduce accuracy.

2. The functioning of the classifier is difficult to explain in any language that is geographically meaningful.

This paper will investigate the use of a neural classification technique known as Self Organising Maps (Kohonen, 1995) to provide a structure to the classification problem based around the dataspace. The Self-Organising Map (SOM) is a set of artificial neurons, each of which is an ordered classifier in feature space. A 2-D SOM is commonly used to find (construct) which is an ordered classifier in feature space. A 2-D

SOM is commonly used to find (construct) classifiers, hence, provide a continuous topological mapping between the dataspace and the 2-D space. It is known that the SOM may develop a poor mapping if elements of an input vector have different scales (Kohonen, 1995). This situation arises when an input vector consists of signals from different domains (Takatsuka, 1996). In this study, separate SOMs are used for classifying and discovering underlying structures from different domains. These are then combined to produce an overall solution.

The following hypothesis is tested: that classification based on dataspaces can outperform traditional approaches in terms of accuracy, computing efficiency, and understandability, by providing much-needed structure to the problem. Results from a landcover classification exercise will be presented and compared to established neural and statistical classifier benchmarks (Gahegan and West, 1998).

References

Aspinall, R.J. and B.G. Lees (1994) Sampling and Analysis of Spatial Environmental Data. Proc. 6th International Symposium on Spatial Data Handling, Ed: Waugh, T.C. and R.G. Healey, Edinburgh, Scotland, pp. 1086-1098.

Gahegan, M. and G. West (1998) The Classification of Complex Geographic Datasets: An Operational Comparison of Artificial Neural Network and Decision Tree Classifiers, Proc. GeoComputation98, 3rd International Conference on GeoComputation. http://www.cs.curtin.edu.au/~mark/geocomp98/geocom61.html

Kohonen, T. (1995) Self-Organizing Maps. Springer-Verlag, Berlin, Germany.

Takatsuka, M. (1996) Free-Form Three-Dimensional Object Recognition Using Artificial Neural Networks. Ph.D thesis, Dept. of Electrical and Computer Systems Eng., Monash University, Clayton 3168 VIC, Australia.

GERMAN, Gordon (gordon@cs.curtin.edu.au), Curtin University of Technology, School of Computing, Kent Street, Perth, Western Australia

Neural Network Classifiers for GIS Data: Improved Search Strategies

Key Words: GIS classification, neural networks, remote sensing, function

Artificial neural networks have several advantages when used as classifiers of complex geographic and remotely-sensed datasets. They normally require no assumptions on the data distribution and can be trained with relatively small sample sets. Further, they are robust classifiers that require little data preparation prior to use; however, the selection of a suitable architecture and the subsequent lengthy training time of the

network have often been perceived as a disadvantage to the acceptability of such classifiers. The author has previously shown a methodology for selecting an architecture and reducing the time required for training by specific manipulation of the network parameters prior to the training phase.

The training of these classifiers is a process that typically involves the use of a minimisation routine to search for an acceptable minimum-error position, in some highly multidimensional weight space. The variation of these weights affects the positioning of separating hyperplanes in attribute-space, which are generated by the network's hidden layer. Although successful in many applications where a mechanistic, "goal-seeking" search is applicable (the algorithm generally seeks for a best global outcome regardless of the path required to attain it), the use of specific knowledge of the form of the solution can be used to improve the performance of the search. Specifically, the solution requires the hyperplanes to model class surface boundaries in attribute-space. Despite this, the minimisation routine makes no use of any prior information on the class separability, which is available from the sample sets and could be used to provide a gross positioning of the hyperplanes prior to training. The previously developed methodology partially addresses this. Initialisation of the network is done with regard to the class spread in attribute-space and the calculated redundancy available in the network. The method first positions the separating hyperplanes generated by the network in a sub-optimal, linear-separable configuration prior to training. It then assigns additional hyperplanes for classes displaying more complex separation boundaries, enabling a more efficient buildup of piecewise linear separating surfaces, reducing the training time and increasing the chance of convergence on a suitable minimum.

Even so it is often the case, especially when dealing with complex geographical datasets, that during training, hyperplanes that are already in an optimal position (and thereby contributing minimal local error) will be perturbed in an effort to reduce the *overall* global error; hence, there is not a monotonic increase in classification performance as training progresses. This reduces the efficiency of the training process and can limit the classification performance attainable. In an extension of the prior methodology, hyperplanes that already produce adequate separation in attribute-space are "frozen" in position prior to, or during, training. This allows the network to focus on areas of poor separability and spawn additional hyperplanes if required. Results obtained on several complex datasets show a significant improvement of classification performance, as well as a reduction in training time.

GeoComputation 99

4th International Conference on GeoComputation
Mary Washington College
Fredericksburg, Virginia
July 25-28, 1999

GLUCK, Myke (mgluck@lis.fsu.edu), YU, Lixin, JU, Boryung, JEONG, Woo-Soeb, Jeong, and CHANG, Crystal, Florida State University, 244 Shores Building, Tallahassee, FL 32306-2100

Augmented Seriation: Usability of a Visual and Auditory Tool for Geographic Pattern Discovery with Risk Perception Data

Key Words: seriation, multimedia, visualization, auditory tools, geographical data analysis, spatial analysis, usability

Geographic data are often presented in maps and tables using geographic information system (GIS) viewers, spreadsheets, and other software. Are there other ways of presenting spatial data to users to improve perception and interpretation of the information? This study developed and tested a data presentation tool that integrated a map viewer, seriation matrix, and sound generator to study how users receive and perceive spatially related information.

Seriation is the process of looking for univariate sequences in data, such as a chronological sequence of pottery shards at an archeological site.Seriation, also called reorderable or permuted matrices, may be used as a tool for discovering patterns in numerical data. Seriation data matrix cells contain icons proportional in size to data values rather than actual data values. Users of a seriation tool permute rows and columns of the matrix to visually discover graphical patterns of the data with meaningful interpretations.

Augmented seriation uses a computer graphical user interface to add interactive manipulation, color, sound, and map views to the basicseriation techniques. Augmented seriation permits simultaneous display of maps andiconized data matrices. Augmented seriation's maps and matrices are graphically linked to permit highlighting (brushing) of related cells in the matrix and map elements (points, lines, or polygons), expanding the use of seriation for more visual analyses of spatial and temporal phenomena; thus, augmented seriation permits users to perform a visual principal component-like analysis of numerical and spatial data concurrently. Augmentedseriation also employs sound supplementing the visual pattern discovery techniques with aural analyses.

We conducted usability tests of our software implementation with naïve geographic users and have found favorable results. Results indicate that after receiving an explanation of seriation, augmented seriation, and a basic introduction to the tool's menus, users can effectively and efficiently employ the concepts ofseriation and this tool to discern numerical as well as spatial patterns. In our tests, users discerned patterns in hazards and risk data especially well with the visual tools and, to a lesser degree, the auditory tools. The presentation includes a demonstration of the current augmented seriation software implementation

GeoComputation 99
4th International Conference on GeoComputation
Mary Washington College
Fredericksburg, Virginia
July 25-28, 1999

GOOCH, Michael (M.J.Gooch@lboro.ac.uk) and CHANDLER, Jim, Loughborough University, Department of Civil and Building Engineering, Loughborough, LE11 3TU, U.K.

Failure Prediction in Automatically Generated Digital Elevation Models

Key Words: digital photogrammetry, DEM, accuracy, classification, strategy parameters

Developments in digital photogrammetry have provided the ability to automatically generate Digital Elevation Models (DEMs) that are increasingly used by geoscientists (Chandler, in press). Using overlapping imagery, dense grids of coordinates can be collected at high speeds (150 points per second) with a high level of accuracy. The trend towards using PC-based hardware, the widespread use of Geographical Information Systems, and the forthcoming availability of high resolution satellite imagery over the internet at ever lower costs, means that the use of automated digital photogrammetry in elevation modeling is likely to become more widespread. Automation can reduce the need for an in-depth knowledge of the subject, thus rendering the technology an option for more users.

One criticism of this trend towards the automated black-box approach is the common lack of quality control procedures within the software (Cooper 1998), particularly to identify areas of the DEM with a low accuracy. The traditional method for accuracy assessment is through the use of check point data (data collected by an independent means of a higher level of accuracy against which the DEM can be compared); however, in most cases, check point data are not available and the user is recommended to manually check and edit the data using stereo viewing methods, a potentially lengthy process that can negate the obvious speed advantages brought about by automation.

Work at Loughborough on accuracy optimization (Gooch et al., in press) has primarily been carried out using the ERDAS Imagine OrthoMAX digital photogrammetric system. This software uses an area correlation-based algorithm over which the user has a certain amount of control through the use of a set of strategy parameters (Gooch and Chandler, 1998). These control the acceptance and quality control requirements of the points in the DEM, and early research work assessed the effect of altering these parameters on the resulting accuracy of the DEMs. It was found that altering the parameters has the most effect on points where the algorithm has difficulty in estimating the elevation (areas with a low image content, a sudden elevation change, or where movement has occurred between the capture of the two images). More significantly, in areas where accurate height estimates were generated, changing the strategy parameters was found to have little impact on the elevation of the point.

A software data processing model has been developed that uses this phenomena to identify areas where elevations are unreliable and to which the user should pay attention when editing and checking the data. The user simply inputs two DEMs of the area (each generated with a different strategy parameter specification) and the software identifies where height estimates are likely to be unreliable due to (a) the software having

interpolated a point in areas with a sudden elevation change, and (b) low image content or object movement between the two images. It overlays this information on an orthophoto of the area thus enabling a hardcopy visual output to be printed for easy identification by the user. The user can then focus on these areas when manually editing the DEM, reducing the need for check data and to check the whole DEM.

The software model developed from this research will be explained and described in detail in the presentation. Results from tests on different scales of imagery, different types of imagery, and other software packages also will be presented to demonstrate the efficacy and generality of the technique.

References

Chandler, J.H. (in press). "Effective Application of Automated Digital Photogrammetry for Geomorphological Research," Earth Surface Processes and Landforms.

Cooper, M.A.R. 1998. Datums, Coordinates and Differences in Landform Monitoring, Modelling and Analysis, S.N. Lane,K.S. Richards, and J.H. Chandler, J. Wiley and Sons, Chichester, Cambridge, 1998, pp.1-17.

Gooch, M. J. and J.H. Chandler, 1998. "Optimization of Strategy Parameters Used in Automated Digital Elevation Model Generation," ISPRS International Archives of Photogrammetry and Remote Sensing, 32(2): 88-95.

Gooch, M. J., M.J. Stojic, and J.H. Chandler (in press). "Accuracy Assessment of Digital Elevation Models Generated Using the ERDAS Imagine OrthoMAX Digital Photogrammetric System," Photogrammetric Record

---

GOOVAERTS, Pierre (goovaert@engin.umich.edu), The University of Michigan, Department of Civil and Environmental Engineering, Ann Arbor, MI 48109-2125

Performance Comparison of Geostatistical Algorithms for Incorporating Elevation into the Mapping of Precipitation

Key Words: geostatistics, co-kriging,precipitation, digital elevation model (DEM)

Until the late 1980s, techniques such as the inverse distance, Thiessen polygons, or isohyetals, have been used to interpolate rainfall data. Geostatistics, which is based on the theory of regionalized variables, is increasingly used because it allows one to capitalize on the spatial correlation between neighboring observations to predict attribute values at unsampled locations. Several authors have shown that geostatistics provide better estimates of precipitation than conventional methods. A major advantage of geostatistical prediction (kriging) is that sparsely sampled observations of the primary attribute can be complemented by secondary attributes that are more densely sampledA

valuable and cheap source of information for many climatic attributes is provided by digital elevation models (DEM).

In this paper, three geostatistical algorithms are introduced to incorporate an exhaustive secondary information (DEM) into the spatial prediction of precipitation: simple kriging with varying local means (SKlm), kriging with an external drift (KED) and colocated cokriging. The techniques are illustrated using annual and monthly rainfall observations measured at 36 climatic stations located in the Algarve region, Portugal. Cross validation is used to compare the prediction performances of the three geostatistical interpolation algorithms with the straightforward linear regression of rainfall against elevation and three univariate techniques: Thiessen polygons, inverse square distance, and ordinary kriging. Larger prediction errors are obtained for the two algorithms (inverse square distance, Thiessen polygons) that ignore both elevation and rainfall records at surrounding stations. The three multivariate geostatistical algorithms outperform other interpolators, in particular linear regression, which stresses the importance of accounting for spatially dependent rainfall observations in addition to the colocated elevation. Last, ordinary kriging yields more accurate predictions than linear regression when the correlation between rainfall and elevation is moderate (less than 0.75 in the case study).

GRIFFITHS, James A.(Jim.Griffiths@kcl.ac.uk), COLLISON, A.J.C. and WADE, S.W., King's College London, Geography Department, Strand, WC2R2LS, U.K.

The Validity of Using a Simplified Distributed Hydrological Model for Estimation of Landslide Probability Under a Climate Change Scenario

Key Words: hydrological modelling, landslides, climate change, validation

In order to assess the implications of envisaged future climate change on the stability of shallow-translational landslides, a simple distributed hydrological model was designed for use within humid-temperate and sub-humid environments. In order to simulate the hydraulic response of a landslide over relatively long timescales, however, it was necessary to simplify both its structure and related hydrological processes. The vertical soil profile of the feature was represented by just three layers (root zone, colluvium, and underlying impermeable layer), while the landslide surface was divided into 10 by 10m cells using a digital elevation model (DEM).

Though not strictly physically-based, processes of infiltration, unsaturated and saturated flow, and throughflow, all were represented in some way within the model. Vertical water movement between soil layers was simulated using a tank-model approach; i.e., downward moisture movement is gravity driven at a rate determined by soil conductivity and the capacity of the underlying layer to accept moisture. Lateral movement occurs in the direction of adjacent cells with the lowest moisture content at a rate determined by a derivative of Darcy's Law. For each daily timestep then, moisture

movement was modelled vertically between layers but not between cells; and horizontally between cells, butnot between layers.

In the first instance the suitability of the model under present day conditions was addressed. Using the dynamic modelling-based GIS package "PCRaster," timeseries output of predicted local water table heights could be obtained for any cell within the DEM, and then compared with observations made within the field. Model output was optimised by using different interpolation techniques applied to sampled soil and vegetation properties.

The suitability of the model for longer term prediction was then assessed with respect to stochastically generated daily rainfall and temperature data, derived from downscaled mean monthly GCM predictions. Initial results assessed through water-table ascribed thresholds of instability indicate that though climate change increases the probability of slope instability through greater precipitation, this effect may be offset by greater levels of evapotranspiration. Modelled using a temperature dependent empirical relationship, initial sensitivity analysis of evapotranspiration concurs with previous research that variation in parameters of vegetation height and canopy conductivity are of greatest significance to ET prediction.

To conclude, it is recognised that though relevant hydrological processes are numerically modelled to an acceptable standard within the model, the results obtained describe changes to slope stability in relation only to rainfall and temperature. In order to account for the full effects of climate change, a better knowledge of the potential variation in relative humidity; net radiation; cloud cover; and regional patterns of rainfall intensity, duration and seasonality, are needed. As much of this information has yet to be quantitatively determined, the described model may be regarded as a best first estimate given the data currently available.

GUTH, Peter L. (pguth@nadn.navy.mil), U.S. Naval Academy, Department of Oceanography, 572 Holloway Road, Annapolis, MD 21402-5026

Quantifying and Visualizing Terrain Fabric from Digital Elevation Models

Key Words: digital elevation models (DEMs), fabric, eigenvector analysis, terrain analysis

Digital elevation models (DEMs) yield a terrain classification based on three variables: elevation, ruggedness, and topographic fabric or grain (tendency to form linear ridges). To quantify fabric, the analysis extracts eigenvectors and eigenvalues from a 3 by 3 matrix of the sums of the cross products of the directional cosines of the surface normals at each point in the DEM. For topography eigenvalue S1 is much greater than S2 and S3, and S2 and S3 are approximately equal. The ratios ln(S1/S2) and ln(S1/S3) correlate highly with relief (difference between highest and lowest elevations), standard deviation of elevation, average slope, and standard deviation of slope; any of these could

categorize ruggedness. Uncorrelated with ruggedness, ratios (ln(S1/ S2)/ln(S2/ S3)) and ln(S2/ S3) measure the fabric or grain. Orientations of S2 and S3 define the dominant grain of the topography, and the ratio of S2 to S3 determines the strength of the grain. Average elevation correlates poorly with all other measures, defining the third element of the classification.

This fabric measures a point property of the DEM and the underlying topographic surface, but the property depends on the size of the region considered. This property varies in a systematic way, both spatially over a region and at a single point as the region size varies. With 30-m DEMs, regions as small as 100 elevations (300-m squares) produce meaningful results. The analysis appears relatively insensitive to DEM quality (U.S. Geological Survey Level 1 and Level 2 DEMs produce very similar results) or DEM spacing (10-m and 30-m DEMs also produce similar results).

Visualizing this organization of topography requires a variety of graphical techniques, including contoured stereographic projections of surface normals, graphs showing the variation in ruggedness and fabric strength as a function of analysis region size, colored maps showing the strength and orientation of the fabric, and contour or reflectance maps with the direction and strength of the fabric overlaid by vectors. Animations showing how these parameters vary with region size greatly assist in the analysis process.

HARDISTY, Frank (Hardisty.1@osu.edu), The Ohio State University, Department of Geography, Columbus, OH 43210

Stochastic Modeling of Alternate Methods for the Generation of Event Patterns with Complete Spatial Randomness

Key Words: CSR, stochastic, event patterns

Many spatial statistical tests call for comparison of a given event pattern to a pattern or patterns with the characteristic of complete spatial randomness (CSR). Other spatial statistical tests rest on claims about characteristics of patterns under CSR that we may wish to test; therefore, the generation of CSR patterns is therefore of some interest. Unfortunately, not all writers on spatial statistics agree on accurate methods for generating such patterns. Three algorithms for the generation of CSR patterns are described, and then simulated under 1,200,000 trials of the three-event generation processes, generating nine test statistics for each of the processes in each trial. A new test statistic, Frank's K, for characterizing event patterns, is described.

HAYNES, Kingsley (khaynes@gmu.edu), Director, The Institute of Public Policy, STOUGH, Roger, (rstough@gmu.edu), and KULKARNI, Rajendra, (rkulkarn@gmu.edu), George Mason University, ATTN: MS 3C6, Fairfax, VA 22030-4444

Towards a Percolation Model of Accessibility

Key Words: accessibility, site, bond percolation

Transportation networks are characterized by two complementary functions: accessibility to economic activities and mobility to carry out these activities. One of the ways to measure accessibility is the number of opportunities available to a user of a transportation network. But for users to carry out economic activities, i.e., exercise mobility, they have to at least have conceptual or perceived links to the activities. Given this, we ask what is the process that leads to the development of these perceived links? We present an analytical approach based on percolation theory to explain the evolution of perceived links to economic activities. A subjective accessibility measure, i.e., a network of perceived links, is developed based on economic opportunities and individuals' income levels in a transport region. For this research a perceived link to an opportunity exists if the user's income is above some threshold. Accessibility to opportunities may be viewed as analogous to the sites and bonds percolation model where sites are opportunities and bonds are conceptual links. In the context of percolation theory, the complex network of sites and bonds is viewed as subject to the critical phenomenon of phase transitions. This paper describes and applies, through numerical analysis, i.e., computer simulation, an application of the sites and bonds percolation model to accessibility. The paper provides an original approach to the investigation of the evolution of perceived links in a transport network and the related mobility of network users.

HEDGER R.D. (rdh@geo.ed.ac.uk) and MALTHUS, T.J., Department of Geography, University of Edinburgh, Edinburgh, EH8 9XP, U.K.; OLSEN, N.R.B., Norwegian University of Science and Technology, Norway; GEORGE, D.G., Institute of Freshwater Ecology, Far Sawrey, Ambleside, Cumbria, LA22 OLP, U.K.; ATKINSON, P.M., Department of Geography, University of Southampton, Southampton, SO17 1BJ, U.K.

Dynamic Modelling of the Spatio-Temporal Distribution of Phytoplankton in a Small Productive English Lake

Key Words: computational fluid dynamics, phytoplankton spatio-temporal distributions, velocity field, irradiance

The relationships between the spatio-temporal distribution of phytoplankton concentration and the environmental conditions of Esthwaite Water (a small eutrophic lake in the English Lake District, U.K.) were examined using a 3-D computational fluid dynamics (CFD) model. The water velocity field was obtained through solving the 3-D Navier Stokes equation for turbulent flow on a finite-volume, unstructured non-orthogonal grid. The spatio-temporal distributions of two types of phytoplankton were modelled: the cyanobacterium Microcystis, and the dinoflagellate Ceratium. Cyanobacterial buoyancy was estimated according to the Kromkamp and Walsby model, and dinoflagellate motility was estimated according to a model that we devised using empirical data from Esthwaite Water and other similar lakes. Circulation patterns of water and phytoplankton, as simulated by the CFD model, were similar to those obtained through field observations.

Downwind surface drift currents were initiated by wind stress, with sub-surface return gradient currents initiated near the thermocline. Near-surface accumulations of cyanobacteria were pushed downwind by the surface currents and accumulated at downwelling areas, and near-thermocline accumulations of dinoflagellates were pushed upwind by the sub-surface return currents, and accumulated at upwelling areas. In all cases, the Coriolis force greatly influenced patterns, causing a clockwise deflection of water flow and phytoplankton accumulation. Through the use of the CFD model, it was possible to conclude that the horizontal and vertical phytoplankton distributions resulted from the interaction between the vertical motility of the phytoplankton (dependent on the light environment) and the velocity vectors at the depths at which the phytoplankton accumulated (dependent upon wind stress and basin morphometry).

JIANG, Yu (yjiang@iftd.org), Institute for Technology Development, Building 1103, Suite 118, Stennis Space Center, MS 39529

A Practical Method of Calibrating Airborne Hyperspectral Imagery

Key Words: hyperspectral, distortion, calibration

This paper introduces a simple but practical method of calibrating airborne hyperspectral imagery acquired by a push-broom airborne hyperspectral imaging sensor (RDACS/H3). The sensor generates imagery via a line-by-line scanning method. Usually, an image generated by the RDACS/H3 has a wavy look due to the constant aircraft attitude changes during the scanning process, therefore, a calibration is essential before a geo-referencing process can be undertaken to match an image to a map or reference imagery. However, most airborne hyperspectral data lack the attitude changing information of the aircraft, which makes it impossible to conduct the calibration by conventional means. This paper addresses this problem by fitting a linear feature on a raw image to a corresponding linear feature on the reference image. This method eliminates the major wavy distortions but reserves the necessary features for the later geo-

referencing process. The results proved that a wavy, distorted hyperspectral image can be efficiently calibrated and accurately geo-referenced to match a reference image.

JULIAO, Rui Pedro (rpj@fcsh.unl.pt) Departamento de Geografia e Planeamento Regional, Universidade Nova de Lisboa/FCSH, Av. Berna, 26 C, 1050 Lisboa, Portugal

Measuring Accessibility Using GIS

Keywords: accessibility, planning, regional development

Today it is widely understood that one cannot promote development regardless of different territorial specifications that make the nation mosaic. We cannot promote the development of the whole without the knowledge of each small territory and the relationship between them. This is why recent territorial planning is facing these principles. Planning policies are concerned with equity and a better distribution of people and activities in the territory, which is why accessibility, regardless if it is measured in time, cost, distance, or population, are the most important variables that one must consider in the early stages of planning. In what concerns population studies there are several different methodologies for its analysis and understanding, considering spatial interaction. Unfortunately, accessibility is widely evaluated by methodologies that do not consider a real spatial model. The traditional methods for accessibility evaluation do not consider the whole territory; they are mainly based in matrix methods and in node/arc logic. By using this type of method, one cannot get information to the whole territory. It was usual to get an accessibility index for a specific node of the network, but not for any point. Today accessibility measurement, regardless if its unit is time, cost, or distance, must be evaluated for the whole territory and not only on the network.

The usual methods for accessibility evaluation, based in graph theory, are quite easy to essay in a vector format GIS analysis, but if one wants to create a continuous model, we must work in a raster environment. This, of course, will reduce the geometrical accuracy of the information; however, it opens a wide range of new analysis capabilities. This is why although the original information was in vector format, the analysis was made mainly using raster data.

The general structure of the accessibility evaluation methodology has 3 phases:

- Data Acquisition and Integration
- Cost Surface Modelling
- Accessibility Analysis

This paper describes the development and application of a GIS-based methodology for accessibility evaluation, and its different potential applications in planning studies using the Lisbon and Tagus Valley Region as a case study area.

KAFATOS, Menas (mkafatos@gmu.edu), George Mason University, Center for Earth Observing and Space Research, 4400 University Drive, MSN 5C3, Fairfax, VA 22030

Data Access, Querying, Analysis, and Data Mining in a Distributed Framework for Earth System Science Support

Key Words: data access, data mining, remote sensing, distributed databases

Over the next decade, remote sensing missions of the National Aeronautical Space Administration (NASA), European Space Agency (ESA), National Space Development Agency (NASDA), and other international missions, will be observing the Earth's oceans, lands, and atmosphere and producing very large volumes of geoscience data sets. These data will be used by many communities beyond the traditional science communities, including interdisciplinary and process studies and applications. To facilitate data access, users need to query search engines and obtain information on the content of data before they proceed to order data sets that may or may not serve their needs. At George Mason University, we have developed the concept of distributed data access; querying and on-line analysis through a working prototype termed the Virtual Domain Application Data Center (VDADC)

The VDADC maintains global geophysical, gridded L3 data sets supporting interdisciplinary Earth science, and provides on-line analysis capabilities of these data sets.

KEIGHAN, Edric (Ekeighan@cubewerx.com), CubeWerx, Inc., R-13, 200 Montcalm Hull, Quebec, Canada; and KUCERA, Henry (hkucera@ibm.net), Mercator Systems, Ltd., 2936 Phyllis Street, Victoria, V8V 4L8, B.C., Canada

Second Generation Spatial Information Warehousing Architectures

Key Words: multiresolution, seamless, spatial, warehousing, distributed

The challenge for spatial data users is to access accurate data-trusted information in support of their specific needs. This paper describes the architecture for spatial information access based on a data warehouse driven by an open data access interface. Requested data sets, in vector, raster, matrix, and textual formats are accessible through on-line network gateways, compliant with international, national, industry, and government standards for exchange and interconnection.

Data Warehousing differs from traditional reporting and transactional systems in three significant ways. A warehouse provides a separate decision support database developed by integrating data from one or more operational systems. This supports synthesis and integration of the data. A warehouse also allowsdecision makers efficient

on-line navigation into time-variant corporate data without impacting the operational systems. The warehouse also can be optimized for speed to allow distributed access to information that would be unthinkable in today's GIS or statistical applications. These three principles fundamentally alter the way business and government users interact with corporate data and how data can be leveraged. With data warehousing, users access data directly, when and how they want; instead of submitting requests to their Information Systems (IS) that might take weeks to fulfill. Decision makers can execute queries and build reports on their own workstations or thin clients, freeing the IS department to focus on other tasks such as maintenance and administration. If the spatial data warehouse is properly configured, users can issue one query after another in an ad-hoc manner to explore trends, identify problems, evaluate market opportunities, and/or order data to respond to their application specific requirements. Within a distributed data warehousing architecture, a spatial database populated with a basic set of feature data provides the framework while intensification or vertical integration of subjects are performed through access to federated databases in a seamless design.

The authors believe this type of system is dependent on a database architecture that is extensible (multiple data types), scalable (terabytes plus) and multiresolution (drilldown and rollup on the fly). The system also requires open interfaces to support thin clients and must be supported by a metacontent repository to support user interaction with the information. The first generation of this type of system was delivered through a project in Canada called Mercator I. The next generation architectures are presented and discussed within the context of ongoing research in Canada and the U.S.

KIDNER, David (dbkidner@glam.ac.uk), DOREY, Mark, University of Glamorgan, School of Computing, SMITH, Derek, Division of Mathematics & Computing, Pontypridd, Rhondda Cynon Taff, Wales, CF37 1DL, U.K.

What's the Point? Interpolation and Extrapolation with a Regular Grid DEM

Key Words: digital terrain modelling, DEMs, interpolation, surface polynomials and splines, extrapolation, linear and non-linear prediction, Lagrange multipliers, data compression

This paper advocates the use of more sophisticated approachesto mathematical modelling of elevation samples for applications that rely on interpolation and extrapolation. The computational efficiency of simple, linear algorithms is no longer an excuse to hide behind today's advances in processor performance. The limitations of current algorithms are illustrated for a number of applications ranging from contouring and visibility analysis to data compression.

A regular grid digital elevation model (DEM) represents the heights at discrete samples of a continuous surface. As such, there is not a direct topological relationship between points; however, for a variety of reasons, users consider these elevations to lie at the vertices of a regular grid, thus imposing an implicit representation of surface form.

For most GIS, a linear relationship between vertices is assumed, while a bilinear representation is assumed within each DEM cell. The consequences of imposing such assumptions can be critical for those applications that interpolate unsampled points from the DEM. Small interpolation or elevation errors can propagate through to large application errors. This is true for intervisibility analysis and even more traditional GIS algorithms, such as contouring. While these applications are well understood and have been with us for many years, bilinear interpolation is still the most commonly used method for retrieving surface estimates from DEMs. The first part of the paper demonstrates the problems of interpolation within a DEM and evaluates a variety of alternative approaches such as bi-quadratic, cubic, and quintic polynomials and splines that attempt to derive the shape of the surface at interpolated points.

Extrapolation is an extension of interpolation to locations outside the current spatial domain. One can think of extrapolation as standing in the terrain and given my field of view, what is my elevation at a location one step backwards? This approach to elevation prediction is at the heart of many new techniques of data compression applied to DEMs. The demand for better data compression algorithms is a consequence of finer resolution data, e.g. LiDAR, and the wider dissemination of DEMs by intranet and internet. In a similar manner to the interpolation algorithms, the basis of elevation prediction is to determine the local surface form by correlating values within the field of view. The extent of this field of view can be the nearest three DEM vertices that are used to bilinearly determine the next vertex. The second part of the paper evaluates this approach for more extensive fields of view, using both linear and non-linear techniques.

KIDNER, David (dbkidner@glam.ac.uk), FITZELL, Ian and AL NUAMI, Miqdad, University of Glamorgan, School of Computing; RALLINGS, Phil and WARE, Andrew, Division of Mathematics and Computing, Pontypridd, Rhondda Cynon Taff, Wales, CF37 1DL, U.K.

A Distributed Approach For Planning Radio Communications

Key Words: digital terrain modelling, 3-D modelling, parallel and distributed computing, radio propagation

This paper focuses on integrating detailed 3-D topographic features with high-resolution digital terrain models, radio path propagation algorithms with respect to terrain and topographic features, and a parallel implementation of these algorithms across a distributed network of Pentium PCs to predict broadcast coverages over a 'cluttered' field-of-view.

Radio communications planning has been a prominent application area of digital terrain modeling for the last 30 years. Path loss models describe the signal attenuation between the transmitter and receiver as a function of the propagation distance, and other parameters related to the terrain profile and its surface features. Many of the problems of radio propagation are akin to the problems ofintervisibility analysis, i.e., greater accuracy

can be obtained with higher resolution digital terrain models (DTMs), such as LiDAR-derived DTMs, and detailed 3-D surface features or "clutter." Clutter can include buildings and other relevant properties including roof structure, building materials and elevations; vegetation and shrubbery, including individual trees; and other 2-D surface features that have distinctive propagation properties, such as roads and water bodies; however, the complexity of these 3-D data models and propagation algorithms means increased computational workloads. This is particularly significant for determining broadcast coverages in the field-of-view. One solution is to parallelise the application. The need for specialised hardware is no longer an obstacle to producing parallel solutions, as recent advances in modern operating systems allied with PC networks are available in most organisations. The real problem is identifying the best strategy for assigning the workloads to machines, which minimises both processor redundancy and communication overheads. The paper discusses the issues for developing a distributed GIS for radio propagation modelling and presents a domain decomposition strategy that optimises performance.

KUCERA, Gail Langran (kucera@islandnet.com), Mercator Systems Ltd., 2936 Phyllis Street, Victoria, BC V8N 1Z1, Canada

Pluralism in Spatial Information Systems

Key Words: pluralism, versioning, multiresolution, multitemporal

An increasingly data-rich environment requires more than a simple monolithic representation of information. The need to extend an information system to model a more pluralistic world can occur because the application domain itself are pluralistic or because the available data are pluralistic. The following are some examples.

1. The application needs to track change over time, meaning that multitemporal versions of features or attributes are required.

2. Multiple data sources are available, but no single source can be considered to be definitive. One approach is to maintain data from all sources concurrently within the database, and create on-demand reconciled views using application software.

3. Data at multiple resolutions are available, but a lower-resolution data source might contain features not available on a higher-resolution data source.

4. Users need the ability to drill down and up, changing not just the resolution, but also the content of the data being analyzed. Rapid performance could require roll-ups of generalized data, or data from a lower-resolution source.

The most common and simple way to manage pluralistic information is to manage a monolithic database and reconcile any pluralism as part of the update process. New information that differs from existing information is evaluated, and if "better," it replaces the "worse" information. The monolithic database holds a single "best map" compiled by integrating available information.

A number of techniques are available to model a more pluralistic world within the database. Examples include:

- Time-bracketed non-persistent features
- Multitemporal linked lists of persistent feature versions- "Same-as" networks that link different versions of different portions of the same feature
- Source goodness hierarchies to describe the relative reliability of data sources
- Feature goodness hierarchies that facilitate drill-down and drill-up within a data warehouse
- Various roll-ups of pluralistic data to present an authorized monolithic view.

This paper will draw examples from three projects that, combined, use all the above techniques. The first project resulted in an operational system to manage 100-plus years of information on the crown lands that encompass more than 90 percent of British Columbia. The second project involves ongoing Research and Development funded by the U.S. Army Topographic Engineering Center, and will result in a pluralistic spatial data warehouse to manage multisource, multiresolution information with linked "same-as" feature networks to support spatial drill-down/up, and generation of a "best map" at a requested scale. The third is a spatial data warehouse for Canadian Forest Service inventory information and forest management criteria and indicators.

KUCERA, Henry (hkucera@ibm.net), Mercator Systems, Ltd., 2936 Phyllis Street, Victoria, V8V 4L8, B.C, Canada; and LAFOND, Pierre (plafond@holonics.ca), Holonics Data Management Group, Ltd., 200 Montcalm, Suite, 105, Hull, J8Y 3B5, Q.C., Canada

Linking Process and Content in a Distributed Spatial Production System

Key Words: metacontent, distributed computing, multi-resolution, repository

Advances in technology have so improved the efficiency of data collection that organizations find it difficult to keep up with the growing flood of data. Many organizations continue to prepare the data for single narrowly focused applications in which only the local creators of particular databases may be very knowledgeable about their characteristics and legacy. While local analysts may use a particular data set with confidence, consideration is seldom given to recording the history of the database in a form that can be conveyed to remote users. As a potential user of this data, how do I know if these data are good? Can I easily find out where it came from? How was it

transformed and cleansed before being loaded in the database? What other activities are currently going on in the geographic area that I am working in? These are questions that many people may be interested in, and should be easily answered by simply querying an existing database, using standard query tools, over a local-area network, or the Internet.

Remote users who have had no part in collecting or storing the data originally, and are separated from the data collection activity by time and space, may not be experts in a given subject area that the data represents. Remote information analysts who wish to query and/or process the data are dependent on information about the assumptions made prior to collecting the data, the sampling procedure, potential sources of bias, inconsistencies and/or error. Lack of knowledge of these factors can limit the manner in which particular data can be manipulated and interpreted. Because of this, a wealth of information collected at enormous expense remains locked in databases and cannot be mined to extract additional knowledge.

This paper presents a process management approach that transforms information to knowledge through the concept of "Metacontent Management." The approach combines standard methods for planning and tracking of spatial data acquisition, validation and production activities, with cutting edge tools for managing complex information. The paper discusses the steps we have taken in ongoing projects to move beyond the primitive use of "metadata," and demonstrates how process management tools can be used to facilitate collaboration between people involved in all phases of the spatial data management life cycle. This approach ensures quality and repeatability of the compilation process and streamlines analysis and synthesis of the resulting information.

LAFFAN, S.W. (Shawn.Laffan@anu.edu.au), Australian National University, Department of Geography, Canberra, 0200, Australia

Spatially Assessing Model Error Using Geographically Weighted Regression

Key Words: geographically weighted regression, error assessment, neural network, expert system

This research develops a method to identify local areas where a model has performed well when predicting some spatial distribution.

There is increasing interest in the use of non-spatial tools such as expert systems and artificial neural networks for mapping continuous or fuzzy spatial properties. This is because they can deal with more ancillary variables than spatial interpolation techniques such as co-kriging. Non-spatial tools do not provide truly spatial error measures so it is important to assess their spatial performance in order to identify local areas of acceptable prediction. These areas may then be used with confidence. Areas where there is not acceptable prediction may indicate the need for other variables in the model, or a different approach.

The method uses a variant of Geographically Weighted Regression (GWR) (Brunsdon et al., 1996). Measured values are compared with predicted values within

circles of increasing size to allow thevisualisation of error at increasing scales. No spatial weighting scheme is used. GWR is used because it allows the calculation of error values without transforming the original values to error residuals, and over locations where there is no measured data. The error is the square root of the area between the optimal 1:1 line and a fitted line, bounded by the maximum and minimum predicted values. The R-squared value indicates confidence in the assessment and is used to determine the best spatial scale.

The method is demonstrated using results from an artificial neural network trained to infer aluminum oxide percentages across a 1,100-square-km area inWeipa, far north Queensland, Australia.

Reference

Brunsdon, C., A.S. Fotheringham, and M.E. Charlton, 1996; Geographically Weighted Regression - A Method for Exploring Spatial Nonstationarity, Geographical Analysis, 28 (4), 281-298.

LANDGREBE, David (landgreb@ecn.purdue.edu), Purdue University, School of Electrical & Computer Engineering, West Lafayette IN, 47907-1285

On Information Extraction Principles for Hyperspectral Data

Key Words: hyperspectral data, signal spaces, hyperspectral analysis

Means for optimally analyzing hyperspectral data have been a topic of study for some years. Our work has specifically focused on this topic since 1986. The point of departure for our study has been signal theory and the signal processing principles that have primarily grown from the communicationsciences area over the last half century. The basic approach has been to seek a more fundamental understanding of high dimensional signal spaces in the context of multispectral remote sensing, and then to use this knowledge to extend the methods of conventional multispectral analysis to the hyperspectral domain in an optimal or near optimal fashion. The purpose of this paper is to outline what has been learned thus far.

The introduction of hyperspectral sensors that produce much more detailed spectral data than those previously, provides enhanced abilities to extract useful information from the data stream they produce. In theory, it is possible to discriminate successfully between any specified set of classes of data by increasing the dimensionality of the data far enough. In fact, current hyperspectral data, which may have from a few to several hundreds of bands, essentially makes this possible; however, it also is the case that this more detailed data requires more sophisticated data analysis procedures if their full potential is to be achieved. Much of what has been learned about the necessary procedures is not particularly intuitive, and indeed, in many cases is counter-intuitive. In

this paper, we attempt not only to illuminate some of these counter-intuitive aspects but to illustrate the practical methods that will make optimal analysis procedures possible.

LEGATES, David R. (legates@bayamo.srcc.lsu.edu), Louisiana State University, Southern Regional Climate Center, Baton Rouge, Louisiana 70803; NIXON, Kenneth R., QUELCH, Geoffrey E., and STOCKDALE, Thomas D., Computational Geosciences, Inc., 330 West Gray, Suite 500, Norman, Oklahoma 73069

Environmental Modeling and Monitoring Using a High-Resolution Hydrometeorological Data System (HRHDS)

Key Words: water resource management, WSR-88D weather radars, spatial interpolation

Surface observations have usually been the sole source of meteorological data for environmental modeling and monitoring applications. Development of a national network of Doppler weather radars, coupled with advances in GIS methodology and sophisticated spatial interpolation techniques, has provided improved, high-resolution weather data for use in a new generation of distributed (grid-based) environmental models. The High-Resolution Hydrometeorological Data System (HRHDS) is a real-time, site-specific, operational system that couples this new weather information with surface observations, hydrological modeling, and an interface to facilitate more informed decision-making tasks. Spatially and temporally distributed meteorological and hydrological fields produced by the HRHDS include precipitation (radar-based, gage-based, and a composite radar/gage product), wind velocity (speed and direction), air temperature, atmospheric humidity, soil moisture content, and runoff potential. Derived fields include crop stress factors (both temperature and moisture stress), soil water deficit, and streamflow modeling. Both digital and graphical products are produced that can be used for monitoring and analyzing meteorological and hydrological conditions for a particular location or region.

This system is designed to improve site-specific water resource management for a variety of purposes including river management for optimal hydroelectric power generation, soil moisture monitoring for optimal irrigation scheduling or for wildfire prediction, wind (speed and direction) estimation for pesticide drift applications, and rainfall/flood monitoring for enhanced emergency management. Presently, the HRHDS and its components have been used in a wide variety of applications. For example, it operates as the front-end of a river management system that models the real-time water flow for the Catawba River Basin in North Carolina (by Duke Energy Corporation) and has been successfully applied to assess the spatial and temporal distribution of rainfall for several flooding events in Texas. Applications of irrigation scheduling and pesticide drift in southwestern Oklahoma also are being developed.

The HRHDS was software engineered using the latest software development methodologies. As a spatial tool, it incorporates geographic information systems (GIS) tools and its products can be easily input to commercial GIS packages for further analysis

and presentation. The HRHDS also incorporates a modular, extendible architecture so that it can easily accommodate the development of new products and incorporate new algorithms. Its research extends from a successful technology transfer project that first began in the Center for Computational Geosciences at the University of Oklahoma.

LEITNER, Michael (mleitne@unix1.sncc.lsu.edu), Louisiana State University, E108 Howe-Russell Geoscience Complex, Baton Rouge, LA 70803

First- and Second-Order Properties of Spatial Point Patterns: The Application of Crime Data from Baton Rouge, LA

Key Words: crime analysis systems, spatial point pattern, Baton Rouge, LA

In recent years the United States (U.S.) has seen rapid development in the areas of crime analysis and mapping using geographic information systems (GIS) technology. Such development was in response to increasing crime rates especially in urban areas of the U.S. Instrumental in this development was the establishment of the Crime Mapping Research Center (CMRC) by the U.S. National Institute of Justice (NIJ) in 1996. The mandate of the CMRC is to promote research, evaluation, development, and dissemination of GIS for use in the field of criminal justice. The goal of this center and other law enforcement organizations has been to develop fully functional crime analysis systems (CAS) with standardized data collection and reporting mechanisms, tools for spatial and temporal analysis, visualization of data, and much more. The majorproblem of current CAS's are their lack of tools for spatial analysis, spatial modeling, and forecasting capabilities.

The purpose of this paper is to derive the first- and second-order properties of three different crime data sets and to investigate if these properties can be used to differentiate the spatial distribution between the data sets. This paper will further address if first- and second-order properties provide new insights into the spatial distribution of criminal activities previously not known. The first-order properties describe the way in which the expected value (mean or average) of the spatial point pattern varies across space (i.e., the intensity of the spatial point pattern). Such properties are usually measured with the so-called kernel estimation. Second-order properties describe the covariance (or correlation) between values of the spatial point pattern at different regions in space and are usually measured with the K function.Applied to crime data, both properties could be used to explore the spatial variation in the risk of being victimized by a crime, spatial and space-time clustering of criminal activities, and the raised incidence of criminal activities around point sources, such as robberies around ATM machines, subway entrances and exits, etc.

The three crime data sets used in this study were collected from police reports made available by the Homicide/Armed Robbery Division, Baton Rouge Police Department (BRPD). These data sets include the location of the homicide, the location of the victim's residence, and the location of the offender's residence. Altogether 497

homicide cases from the City of Baton Rouge, Louisiana, spanning 1991-1997, are included in the analyses.

This paper will address the following conference topics: exploratory spatial data analysis and data mining; advances in geographical information systems, particularly in the area of spatial analysis; and geostatistics.

LEVINE, Ned (ned@nedlevine.com), NedLevine & Associates, Annandale, VA

CrimeStat: A Spatial Statistics Program for the Analysis of Crime Incident Locations

Key Words: spatial statistics, crime analysis, GIS modeling

CrimeStat is a spatial statistics package for the analysis of crime incident locations. It is a public domain program developed for the Crime Mapping Research Center of the National Institute of Justice, Washington, D.C. The program is a stand-alone Windows NT/9x program that can interface with most desktop GIS packages. It can read ASCII and 'dbf' files and can write graphical objects to ArcView, MapInfo, Atlas*GIS, Surfer for Windows, and Spatial Analyst. The program is multi-threading and can support multiple processors.

The program takes as input the X and Y coordinates of incident locations (i.e., points). For example, the points could be the locations of crime incidents, motor vehicle accidents, clients, patients, facility locations, or any other characteristic that can be assigned an X and Y coordinate. The points also can be weighted. For example, the points can be the location of police stations with the weights being the number of crime incidents handled by stations in a month. Spherical or projected geometry can be used. A secondary and a reference file also can be allocated. The key to the program is a 'distance engine' and CrimeStat calculates direct and indirect distances. The calculations have been optimized for speed and the program can handle large data sets.

The statistics are grouped into four categories. First, there are a number of spatial distribution statistics, such as the center of minimum distance, standard deviational ellipse, angular mean, and Moran's "I" spatial autocorrelation index. These measure 'first-order' properties of a spatial distribution. Second, there are a number of statistics for distance analysis, such as K-order nearest neighbors, linear nearest neighbors, and Ripley's "K" statistics. These can be used to examine whether points are clustered closer than an expected random distribution or another distribution (e.g., population, employment). Third, there are three 'Hot Spot' routines for identifying clusters of points that group together, and include a nearest neighbor hierarchical clustering routine, a K-means clustering routine, and a local Moran index. Fourth, there are single- and dual-variable kernel density interpolation routines. Both allow quartic and normal kernels, fixed and adaptive bandwidths, and differential outputs. The single-variable kernel routine can be used to model the density of a distribution (e.g., density of burglaries)

while the dual-variable kernel routine can relate two different distributions (e.g., the density of burglaries relative to the density of households). The latter can be used for risk analysis.

The author will demonstrate the program, which should be released by the National Institute of Justice during the late summer or early Fall.

LIVINGSTONE, David (david.livingstone@port.ac.uk), Portsmouth University, Department of Geography, Buckingham House, Lion Terrace, Portsmouth, PO1 3HF, U.K.; RAPER, Jonathan, City University, Department of Information Science, School of Informatics, Northampton Square, London EC1V 0HB, U.K.

Developing a Visualisation and Analysis System for Simulating the Growth of Coastal Spits

Key Words: coastal simulation, process modelling, visualisation, GIS

The paper describes the development of a system to enable a coastal simulation model to be optimised for investigating the growth of structures with a morphometry similar to that of coastal spits. The system has evolved from initial investigations using Sedsim/Wave and Visual Sedsim (Martinez and Harbaugh, 1993, Pflug, Klein, Ramshorn et al., 1992) and was originally employed for simulating the growth of deltas in order to investigate the formation of geological structures (Komar and Inman, 1970). The core modelling procedures have a long history and have been developed and modified in FORTRAN during the past 25 years. The visualisation procedures developed for Sedsim3.03 were written specifically for a previous generation of Silicon Graphics so, in the new system, IDL is used for this task. This approach has a number of benefits. IDL is a multiplatform system that also can be used for more detailed exploratory data analysis, is widely used for a range of scientific applications, and has a flourishing community of developers. The main drawback is that a runtime version of IDL is required in order for the developed procedures to be usable.

Martinez developed and validated the principles behind the Sedsim/Wave models, in part, by experiments using topographies based upon contemporary situations (Martinez and Harbaugh, 1993) using data collected for Silver Strand and El Moreno beaches (Komar and Inman, 1970). The system described here is required to combine environmental models, and digital elevation data measured for a spit on the North Norfolk coast of U.K., in order to investigate the processes that control spit development. The major issues that the system addresses are:

- Managing and analysing model variables such as wave angles, topography, and time steps

- Visualising and measuring model-generated processes such as transport rates and deposition
- Combining and comparing model results with measured field data
- Defining, tracking and comparing sub-features in both the simulated and measured data

The system has four major components:

- The process simulation models: SEDSIM/WAVE
- A model management database: SEDLINKS
- A geographical information system to create model landform situations: ARCVIEW
- A data visualisation and analysis system: IDL

A major consideration was to produce a system that was flexible enough to expand beyond its initial requirements for a generic range of applications. While the components of the system need to work together specifically to perform the task, independently they need to be able to perform functions with potential for a wider range of modelling scenarios. This objective is much more easily stated than achieved and compromises have had to be adopted for pragmatic considerations; however, the philosophy behind this approach is important, as it tries to address some of the issues that have restricted the widespread use of GIS with respect to particular kinds of statistical and mathematical environmental modelling techniques.

References

Martinez, P.A. and J.W. Harbaugh (1993), Simulating Nearshore Environments, In: Computer Methods in the Geosciences, vol.12, pp. 265, Pergamon Press, 1993.

Pflug, R., H. Klein, C.H. Ramshorn, M. Genter, and A. Stark (1992), 3-D Visualisation of Geologic Structures, In: Computer Graphics in Geology, Eds: R. Pflug and J.W. Harbaugh, Lecture Notes in Earth Sciences 41, p. 29-39, Springer Verlag 1992.

Komar, P.D. and D.L. Inman (1970), Longshore Sediment Transport on Beaches, Journal of Geophysical Research, 75, p. 5914-5927.

LLOYD, C.D. (cdl195@soton.ac.uk) and ATKINSON, P. M., University of Southampton, Department of Geography, Southampton, SO17 1BJ, U.K.

Designing Optimal Sampling Configurations with Ordinary and Indicator Kriging

Key Words: ordinary kriging, indicator kriging, stationarity, sampling

The objective of this paper is to examine the applicability of two geostatistical approaches, ordinary kriging (OK) and ordinary indicator kriging (IK), to the design of optimal sampling strategies. A by-product of OK is the OK variance. The OK variance is a measure of confidence in estimates. It is a function of (i) the form of spatial variability of the data (modelled, for example, by the variogram), and (ii) the spatial configuration of the samples. The disadvantage of the OK variance is that it is independent of the magnitude of data values locally. For a given sampling configuration, the OK variance will be the same irrespective of the data values; thus, if data are measured on a regular grid, the maximum OK variance will be identical; however, much of the data vary locally. An approach that is conditional on the data values would be more suitable in such cases. This paper uses the conditional variance of the conditional cumulative distribution function (ccdf) derived through IK to assess local uncertainty in estimates. Since the conditional variance of the ccdf is conditional on the data values, the problem of OK variance data independence is overcome. Previously, to determine an acceptable sample grid spacing, investigators have plotted the maximum OK variance for a range of sample spacings and used the plot to select a sample spacing that achieved a given precision of estimation; however, where the spatial variability is not stationary across the region of concern, the OK variance will be biased as it is independent of the data values locally. The maximum conditional variance was used in the same way to account for the magnitude of data values as well as the form of spatial variability and the spatial configuration of the data. A photogrammetically derived digital terrain model (DTM) was sampled on a regular grid, and the success of the OK and IK approaches in ascertaining optimal sampling intervals was examined and compared with reference to the DTM. Once the variogram and indicator variograms were computed for the sample data, mathematical models were fitted and the model coefficients were used for kriging. The performance of the two approaches was assessed in three separate ways: (i) The model coefficients were used to ascertain the maximum OK and conditional variance for several sampling intervals; (ii) The DTM was then sampled at several (progressively smaller) spacings, and estimates were made from the samples. The differences between the estimates and the population (that is, the complete DTM) were then computed and the errors using OK and IK were related to the maximum error that was predicted by the OK variance or the conditional variance. The proportion of estimates that fell outside the estimation variances were quantified and the different results compared; (iii) The estimation errors for each grid cell were plotted against the OK variance and conditional variance, and the form of the relationships was assessed. Finally, the implications of using the two approaches were discussed.

MACEACHREN, Alan M. (alan@essc.psu.edu), Pennsylvania State University, 302 Walker, Department of Geography, University Park, PA 16802

Exploring Geo-Data Spaces-The Search for Meaning

Key Words: geographic visualization, data mining, spatiotemporal representation, exploratory data analysis

Technology for collection, transmission, and display of georeferenced data, at scales from architectural to global, has advanced rapidly in the past decade-resulting in an order of magnitude increase in availability of georeferenced data. While computational methods allow us to extract *information* from these masses of data, domain expertise coupled with the power of human vision provides the necessary complement to these computational methods that allow us to extract *meaning* from the information. Recent and anticipated developments in geographic visualization (GVis) provide the mechanism for linking human experts to computational tools, facilitating the critical step from information extraction to knowledge construction. These developments and their potential are presented in this paper.

GVis synthesizes perspectives on visual representation and analysis, thus far primarily from visualization in scientific computing (ViSC), cartography, and exploratory data analysis (EDA). For GVis to be an effective method in knowledge construction, a closer coupling is needed between the visual methods at the core of GVis and the analytical methods of geocomputation. Beyond this integration, for geo-knowledge construction environments to reach their potential, a more complete understanding must be achieved concerning how people (particularly domain specialists) conceptualize problems and interact with computer systems.

Within the International Cartographic association, a four-component research agenda for GVis is under development (focusing on issues in representation, interface design, GVis-database integration, and cognitive aspects of visualization method development and use). Elements within each component of this research agenda have direct implications for design of knowledge construction environments that involve integration of visualization with computational methods. In this paper, the author provides a brief overview of the ICA research agenda, then focuses specifically on selected issues that underlie design of knowledge construction environments linking analytical with visual methods. Particular attention will be directed to three topics: design of exploratory spatiotemporal data analysis methods (particularly the adaptation of EDA to space-time data); integration of GVis with knowledge discovery in database methods; and the cognitive issues that must be addressed if we are to achieve the next generation of geo-knowledge construction environments that create a more effective conjunction of human and machine capabilities.

MACGILL, James (J.Macgill@geog.leeds.ac.uk), OPENSHAW, Stan (stan@geog.leeds.ac.uk), TURTON, Ian (ian@geog.leeds.ac.uk), University of Leeds, Center for Computational Geography, Leeds LS2 9JT, U.K.

Web-Based Multi-Engine Spatial Analysis Tools

Key Words: smart spatial analysis, web, GAM, GEM, MAPEX, Flocks

The last few years have seen the development and enhancement of a number of tools for the analysis of spatial data. This has been a particular focus of the Center for Computational Geography of which GAM is the most well known. Other exploratory geographical analysis tools have been developedthat use smart search methods while trying to track down patterns in GIS databases. Methods such as MAP Explorer (MAPEX) (which uses a genetic search procedure) and Flock (which uses swarm optimization) power fast and intelligent searches. This paper considers a different approach that is both novel and promises robust performance wheremultiple different pattern The idea is simple. If there is a selection of tools to choose from for any given analysis application, then there is a much better chance of having one that will best suit the problem under analysis; however, the difficulty is knowing in advance which method will perform best in any given situation, as each different technology has its own advantages and disadvantages. In reality, it is likely that different techniques would not just perform differently on different data sets, but also on different parts of individual data sets and at different stages during the exploratory analysis. These problems are likely to become far more severe as the analysis task moves from purely spatial analysis to space-time analysis and higher dimensional data spaces. An ideal system would be able to take the best from each engine while overcoming the shortcomings of each method. This paper proposes a Multi-Engine Spatial Analysis Tool (MESAT) where the different search techniques can be made to work together in a single unified system in which the different search engines cooperate to produce an optimal hybrid spatial analysis technology. The conjecture is that the power of the whole will be greater than that of any of the individual components. The power comes from using multiple search procedures that cooperate by sharing their results.

The paper describes a three-tier architecture for MESAT, whereby the data to be analyzed is stored on a central server (acting as a data warehouse or a GIS), and is attached closely to this system as clients will be the different spatial search engines. Finally, there is a tier, which consists of a visualization system that allows users to observe and interact with both the results and the search mechanisms. This final tier is a web-based visualization system allowing all three tiers to be initiated, run, modified, and observed from almost anywhere in the world. This three-tier architecture allows multiple clients to tackle the same problem simultaneously, and then use multiple visualizations to look at the progress being made. The key to successful inter-search-client communication is a central shared result surface, which is built from the combined input of each separate search engine when they find interesting results. In addition, this surface acts as a form of message board allowing search engines to call in the help of other systems. For example, the fast but less rigorous systems could use the slow but exhaustive search engines when they hunting exploratory search engines work together on a common problemencounter regions of the data set that look potentially interesting to ensure nothing is missed.

The paper also reports the results of applying this MESAT system to analyze synthetic highly complex spatio-temporal-multivariate data sets. The analysis of these sets would almost certainly have been out of reach if any one of the tools had been used individually.

MACMILLAN, W.D. (bmacmill@nta.geog.ox.ac.uk), University of Oxford, School of Geography, Mansfield Road, Oxford OX1 3TB, U.K.

Cellular Strategies for the Simulation of Human Spatial Systems

Key Words: multi-agent simulation, cellular modelling, irregular networks, scale, economic agency

Numerical systems in which the state of each cell in a 2-D or 3-D grid is a function of its previous state(s) and that of its neighbours are well suited to the description of hydrological, geomorphological, and atmospheric processes (see Burrough, 1998); however, they are of limited value in representing human spatial systems. Multiagent simulation has the advantage of separating the treatment of populations of agents from that of the territory that they occupy (see Benenson, 1998). Helpful though this is, it is not sufficient, on its own, to allow a thorough exploration of many economic and social processes. Part of the difficulty lies in the fact that economic and social actions are not generally confined to neighbourhoods. This applies both in a spatial sense, where action at a distance is important and, in time, where actions may occur over different time horizons. The spaces of human geographical activity tend to be complex (see Cliff and Haggett, 1998, on the use of multidimensional scaling to produce metrics based on diffusion rates). The complexity has a number of sources, one of the most important of which can be thought of as the vector structure of human landscapes. Some work has been done using vector structures in the form of regular networks, which is broadly analogous to cellular modelling in terms of strategy (see Peeters et al., 1998) but this approach is still in its infancy.

The purpose of this paper is to explore the scope and limitations of cellular modelling (and the related approaches referred to above) with respect to spatial economic processes. It looks at questions of spatial and temporal scale, the treatment of economic agency, the replacement of conventional approaches to pricing and allocation by iterative, agent-based simulation, and the range of problem types that might be tackled by a cellular strategy. It also looks at the problem of devising a visual vocabulary for human spatial simulation in a variety of contexts.

References

Benenson, I. (1998), Multiagent Simulations of Residential Dynamics in the City, Computers, Environment and Urban Systems, 22(1), pp. 25-42.

Burrough, P. (1998), Dynamic Models and Geocomputation, in Geocomputation: A Primer (P. Longley, S. Brooks, R. McDonnell, and W. MacMillan, eds.), Wiley.

Cliff, A. D. and P. Haggett. (1998) On Complex Geographical Space: Computing Frameworks for Spatial Diffusion Processes, in Geocomputation: A Primer (P. Longley, S. Brooks, R. McDonnell, and W. MacMillan, eds.), Wiley.

Peeters, D., J.F. Thisse, and I. Thomas (1998), Transportation Networks and the Location of Human Activities, Geographical Analysis, 30(4), pp. 355-371.

MARTIN, David (D.J.Martin@soton.ac.uk) and WU, Fulong, University of Southampton, Department of Geography, Southampton, SO171BJ, U.K.

Empirical Cellular Automata (CA) Simulation from a High Resolution Population Surface

Key Words: population density surface, cellular automata, SE England, urban growth, simulation

There is growing literature on the application of cellular automata (CA) to simulate the growth of urban settlement forms (e.g. Batty, 1998). CA allows researchers to view the city as a self-organizing system in which the basic land parcels are developed into various land-use types. A model of the urban system is constructed by the aggregation of uncoordinated local decision-making processes. One of the most important potential uses for such simulations is their ability to model the impact of alternative planning regimes on the development process. CA applications, based on hypothetical urban forms, can provide valuable insights, but the interpretation of such modelling is hampered by difficulties in relating themodelled form to empirical combinations of settlement and constraints. The use of CA methods to model the future development of real urban systems is made particularly complex by the tension between self-organization and the application of empirical constraints.

This paper describes the application of CA to the simulation of urban growth in the southeast region of the U.K., an area currently subject to considerable development pressure. The actual settlement pattern is initiallymodelled as a fine resolution grid using a population surface modelling technique originally developed for use with census area centroid data (Bracken and Martin, 1995), but here applied to unitpostcode information that offers greater spatial and temporal resolution than that available from the population census or conventional land-use mapping. This application is a further development of SimLand (Wu, 1998), which makes use of Arc/Info for spatial data management, with AML programs to permit the evaluation of a range of alternative local and regional constraints on the development process. We classify the wide variety of factors affecting development into static and dynamic ones. The success or failure of a seed becoming a developed land use depends on their combined effect on the self-organised process of local growth. This is further dependent upon the threshold that allows such a process to

proceed. From the observation of land use states in two time periods, the distribution of land use changes is identified. In general, the threshold and its transformation are used to reflect three types of inputs: the growth rate that is related to economic activity, regional variation, and policy control. With different thresholds applied, simulation can generate a series of scenarios of urban development. Development scenarios are treated not as place-specific predictions, but as possible realizations of the development process from which a number of structural indicators can be derived.

This research has a number of original features: it uses detailed empirical spatial data on a fine resolution grid; integrates global effects with the local self-organisation mechanism of urban growth in a more explicit and parameterised way; uses GIS functionality, thus, provides a closer integration with other decision-making tasks; searches the parameter space through a computationally intensive approach; and uses structural indicators to compare the simulation results with the reality.

References

Batty, M. (1998) Urban Evolution on the Desktop: Simulation with the Use of Extended Cellular Automata Environment and Planning A 30, 1943-1967.

Bracken, I. and D. Martin (1995) Linkage of the 1981 and 1991 U.K. Censuses Using Surface Modelling Concepts, Environment and Planning A 27, 379-390.

Wu, F. (1998) SimLand: A Prototype to Simulate Land Conversion through the Integrated GIS and CA with AHP-Derived Transition Rules, International Journal of Geographical Information Science, 12, 63-82.

McKEOWN, David (dmm@maps.cs.cmu.edu), McMAHILL, Jeff, Carnegie Mellon University, Digital Mapping Laboratory, Pittsburgh, PA 15213; CALDWELL, Douglas (caldwell@tec.army.mil) U.S. Army Topographic Engineering Center, 7701 Telegraph Road, Alexandria, VA 22315

Spatial Context Awareness in Feature Simplification

Key Words: simplification, generalization,topology

The automated generalization of spatial data has been an active topic of research in the geographic information system (GIS) community since the early 1970s, when Douglas and Peucker published their initial results on algorithms to simplify lines. Today, generalization is seen as a tool for decluttering cartographic presentations at smaller scales, reducing the information content of data representations regardless of scale, and compressing volumes of spatial data. The bulk of generalization research has focused on the problem of line simplification, primarily because it is more easily understandable and tractable than other generalization transformations. More complicated generalization

transformations, such as aggregation, amalgamation, exaggeration, enhancement, and displacement, have received less attention, primarily because of the difficulties in defining and implementing computational solutions. The focus in linear feature simplification has been on the development of algorithms to manipulate the coordinate information in a single feature. This focus does not fully address the simplification issue and has resulted in anomalies where the topological structure of an individual feature is altered and/or the topological relationships among features are changed. The presence of altered topological structure necessitates the development of additional algorithms to identify the errors and use of editing tools to correct the problems.

This paper addresses the issue of using spatial context in line simplification. It is the result of a research effort designed to produce high quality 3-D simulation databases, not a research effort on generalization. The simulation databases required low information density to support real-time, distributed rendering of the data. They were constructed from highly detailed source data, which contained linear features that needed to be carefully selected and simplified. The use of standard GIS tools for simplification created data with topological errors and unrealistic representations of features with regard to the underlying terrain. The errors resulted in longer simulation database production timelines and required operator intervention to locate and correct the problems. The solution to the data editing problem was the development of smarter algorithms that considered spatial context when performing simplification. Carnegie Mellon University approached the problem in two ways, with the development of topologically-aware and terrain-aware simplification algorithms. The topologically-aware simplification algorithm preserves the relationships of the source data in the generalized data. It supports a per-edge tolerance specification, prevents the insertion of new intersections, and performs proximity checks to keep features' user-specified distances apart in the generalized output. The terrain-aware simplification algorithms preserve relationships with features and the underlying terrain. They handle situations where a 2-D approach to generalization would produce undesirable results. For example, when a road runs around a spur, a 2-D approach to simplification could produce a road that runs through the spur. This creates significant problems in a 3-D simulation database where deep and unnatural cuts are created when the roads are automatically integrated with the terrain.

The Carnegie Mellon University research has not only reduced the time required for the production of simulation databases, it has contributed significantly to the advancement of generalization research by incorporating spatial context in the simplification process.

GeoComputation 99

4th International Conference on GeoComputation
Mary Washington College
Fredericksburg, Virginia
July 25-28, 1999

MESSINA, Joseph P. (messina@email.unc.edu), WALSH, Stephen J. (swalsh@email.unc.edu), VALDIVIA, Gabriela, and TAFF, Gregory, University of North Carolina-Chapel Hill, Department of Geography, CB#3220, Chapel Hill, N.C. 27599-3220

The Application of Cellular Automata Modeling for Enhanced Land Cover Classification in the Ecuadorian Amazon

Key Words: cellular automata, Ecuador, landscape, remote sensing

The Ecuadorian Amazon region is experiencing rapid development as the result of both resource exploitation and official settlement programs. Frontier settlement in the study site lends itself to a clearer understanding of landuse and landcover change. The landscape of the region is heterogeneous and differs structurally in the distribution of species, energy, and materials among the patches, corridors, and matrix present. This particular region is well suited for evaluation as the variables defining settlement expansion and resource development are well controlled and tightly defined.

The discrimination of landcover and landuse has historically been accomplished using single or possibly multidate remotely sensed imagery, systematic ground truthing, or a combination of both. The complexity of the landscape in the Ecuadorian Amazon is such that ground truthing in the historic sense is not easily accomplished, if at all possible, in any systematic manner. With this research, we propose to demonstrate the utility of cellular automata modeling by enhancing the landcover characterization of the region.

Landsat Thematic Mapper data spanning the 10-year period from 1986-96 is used as base data for the development of an initial landcover classification scheme. Traditional classification procedures and validation routines are applied on the collected remotely sensed data. These traditional methods are used as a comparative baseline to evaluate the statistical validity of the cellular automaton enhancement method.

The cellular automata model rules used for this research are more complex than those of a typical cellular automaton, and involve the use of multiple data sources, including topography, road networks, and existing settlement distributions, and their modification over time. The existing settlement patterns are defined by an official government frontier settlement program and exhibit consistency not found in other Amazonian regions. Using traditional urban geography parameters, the model is redefined for the region and allowed to run providing a prior probability component to the classification scheme. Furthermore, the model's control parameters are allowed to self-modify in order to adapt itself to the circumstances it generates, in particular, over the productive lifespan of given settlement and agricultural components. The combination of traditional classification techniques with enhanced cellular automata modeling will significantly improve the local landcover classification accuracy and may prove extendible to widely varying regional conditions.

MOREHART, Mitchell J. (morehart@econ.ag.gov), U.S. Department of Agriculture, Economic Research Service, 1800 M Street N.W., Room 4121N, Washington, D.C. 20036-5831; MURTAGH, Fionn (fmurtagh@acm.org), The Queen's University, Department of Computer Science, BT7 1NN, Belfast, Ireland; STARCK, Jean-Luc (jstarck@cea.fr), CEA/DSM/DAPNIA, 91191, Gif-Sur-Yvette Cedex, France

Multiresolution Spatial Analysis

Key Words: wavelets, spatial analysis, multiresolution, nonparametric regression, geographic information systems

Geographic information systems (GIS) are increasingly used as tools for topographical applications and research. A comprehensive GIS is characterized by its capabilities in the areas of data processing, analysis, and post processing. This paper explores the use of the wavelet transform as a spatial analysis tool for modeling complex multivariate geographic relationships. The use of wavelets in spatial statistics is a relatively recent phenomenon that is rapidly developing. The appeal of wavelet methods stems from their ability to process noisy data with local structures and represent discontinuities such as jumps or peaks in a function. Several examples from agricultural data are used to illustrate the exploratory data analysis inherent in the wavelet transform. The resulting maps provide a convenient means of visually conveying tremendous amounts of information. The redundant `atrous discrete wavelet transform is shown to aid enormously in feature detection and exploration in the succession of resolution views of the data. Analysis is carried out through use of the MR/1 multiresolution image and data analysis package.

MORRIS, Kevin (K.Morris@ccms.ac.uk), Centre for Coastal and Marine Sciences, Plymouth Marine Laboratory, Plymouth, Devon, PL1 3DH, U.K.; and HILL, David and MOORE, Tony, Institute of Hydrology, Crowmarsh Gifford, Oxfordshire, OX10 8BB, U.K.

Mapping the Environment through 3-D Space and Time

Key Words: 4-D, temporal, depth, mapping, information systems

Traditional geographical information systems (GIS) employ a 2-D, or, at best, 2.5-D framework, which is fine for many applications. However, mapping the environment introduces a number of problems that are not easily managed within existing systems. The natural environment is constantly changing and requires a more dynamic way of handling such data (one of the main problems facing temporal GIS (Kemp & Kowalczyk, 1994)). For instance, land-use may change from season to season or year to year. Similarly, the water-flow in a river is a constantly changing phenomenon.

Environmental media, such as the oceans and the atmosphere, complicate matters further as processes that occur within them vary through 3-D space and time. In the past, time and depth have been handled as attributes to a feature. This can be verylimiting as there is no ready dimensional structure against which features can be displayed or manipulated relative to time and depth. In nearly all conventional GIS, the x and y dimensions solely are used to display spatial data in map form at any one time. This paper describes a GIS system that handles time or depth visualization of a feature at the same time as mapping the feature horizontally. This treats time or depth as a dimension rather than an attribute, which is a prerequisite to effective multidimensional visualization and analysis (Raper & Livingstone, 1995).

The Space and Time Environmental Mapper (STEM) has been developed for Land-Ocean Interaction Study (LOIS), a U.K. research project investigating features and processes in the coastal zone. STEM is a GIS data viewer fronting a database containing the highlights of LOIS. STEM owes its flexibility to two key design objectives: a simple yet powerful query expression, retrieval and visualisation interface, and secondly, a generic database design that provides the core of the data-driven system. The database represents the real world in terms of objects (features) and properties (attributes). Both features and attributes can vary in space and time. The generic data model is interfaced to the application shell by a database application programming interface (DBAPI) that lies behind the query interface. The results of a query are returned to the interface shell for graphical presentation to the user where innovative display techniques make the exploration of relationships between objects and properties in the temporal and depth dimension achievable. The system incorporates time and depth bars to represent the additional dimensions. These bars act as an index to the data, and changing the time or depth on these bars will result in the display of the relevant x-y feature data in the mapping window. Alternatively, themultitemporal data can be animated and graphed.

References

Kemp, Z. and A. Kowalczyk, "Incorporating the Temporal Dimension in a GIS," in M.F. Worboys (ed.) "Innovations in GIS 1," (London: Taylor and Francis, 1994), 89-103.

Raper, J. & D. Livingstone, "Development of a Geomorphological Spatial Model Using Object-Oriented Design," (1995), 9: 4: 359-383.

MURNION, Shane (Shane.murnion@port.ac.uk) University of Portsmouth, Department of Geography, U.K.

Cyber-Spatial Analysis: Modelling Web Server Information Flows

Keywords: internet, latency, location, e-commerce,cybergeography

Inequalities between the various global sub-networks may well result in basic asymmetries between service and demand for on-line information services. The basic nature of these asymmetries should directly influence the accessibility of information services in emerging information technology regions to global markets. The resulting development and placement of such services will directly influence the level of accessibility to information and information services for users in these Internet-developing regions. Unfortunately, at the present time it is quite difficult, or perhaps impossible, to determine the nature and effects of these inequalities since little quantitative information is currently available on the quality of remote networks. Historically, measurement tools have arisen from a requirement by system administrators to be able to determine and measure the quality of connectivity between a local computer and some remote computer or computers. This has resulted in a wide variety of tools that allow us to measure this one (local) to many (remote) relationships. A wide variety of these services are producing what are often termed latency or "Internet Weather" maps available at the current time. However, if we wish to determine the quality of a many (remote and local) to many (remote and local) relationship then we discover that there are a very limited number of tools and results available for our use. In this work an attempt is made to apply an old geographical technique of triangulation and adapt it to the new medium of the Internet to predict the latency between two remote computer systems. The main difficulty in adapting such a method is that the information does not travel in a direct line from one system to another, but rather travels along a network between the two computers.

Five "Internet triangulation" stations were set up and distributed around the Internet. The distance between two locations in cyberspace and each of these stations was measured simultaneously using simple ping latency measurements. Each of the stations also monitored the latency from their location to the other monitoring stations. Using the simultaneous measurements, an attempt at using a neural network was made to predict the latency between the two locations being observed. Direct measurements of the latency between the two points were also gathered to compare with the predicted results. The results show that the method shows promise, but did not produce results of sufficient quality for immediate practical use. Some of the possible causes for error were examined. One major reason for interest in getting this method to operate successfully is that in order for the method to work, it requires the location of each object being observed in cyberspace to be explicitly defined. By measuring the position of a large number of Internet addresses for which geographical locations are available, it may be possible using neural computing techniques to develop a model that can map cyberspace locations as defined by the monitoring stations to geographical locations and vice versa. This would allow us to monitor the temporal changes in cyberspace and its changing relationship with geographical space. Analysis of short-term variations throughout the day may reveal interesting patterns of WWW usage in different parts of the globe and might provide insights into the daily cycles of Internet use. Long-term trends may reveal how the quality of the Internet is developing worldwide. Of particular interest is whether or not the disparity between developing and developed nations is increasing or decreasing with time. An automatic gazetteer of this type should allow us to provide reasonably

detailed maps of the density of Internet locations and also of the extent of the current and future Internet domains.

The author wishes to acknowledge the support of the following institutions in this work: The NCGIA at UCSB, and the Curtiss University of Technology in Western Australia.

OLIVER, M. A. (oliverma@isdux1.bham.ac.uk), University of Reading, Whiteknights, Department of Soil Science, Reading, RG6 6DW, U.K.; and KHARYAT, A.L., Taiz University, Department of Physics, College of Science, Republic of Yemen

Investigating the Spatial Variation in Soil Radon Geostatistically

Key Words: geostatistics, kriging, nested survey, radon, spatial variation,variogram

Increasing concern about possible links between emissions of radon and certain types of malignant disease has led to local and regional surveys to measure radon concentrations in the soil and in dwellings. We selected an area in the Midlands of England (Derbyshire), designated as a radon affected area, to explore the nature of the spatial variation of the gas using solid state nuclear track detection. The aim of the first of three surveys was to discover the approximate scale of resolution in the variation using an unbalanced nested sampling scheme and a hierarchical analysis of varianceAn area of 7 km by 7 km with two different limestone formations was sampled by a range of sampling intervals: 1 m, 4 m, 15 m, 60 m, 240 m, 950 m and 3,750 m. The components of variance from the analysis of variance were accumulated and plotted against distance to give a reconnaissance variogram. The results suggested that geologyexerts a strong control on the variation, but that there are other factors involved giving rise to spatially correlated variation within the lithological units over distances of 60 m to 240 m.

To gain further insight into the variation of radon, we sampled an area of 2 km by 2 km on the Monsal Dale limestone only. The variogram was computed andmodelled with a power function. It was then used with the data to estimate radon concentrations at unsampled locations by kriging (geostatistical estimation). The map showed a distinctive pattern that suggested an inverse relation between elevation and radon concentration. The large radon values in the western part were associated with low elevation, and the small values occurred where the relief is higher. This suggested that factors other than lithology might account for the variation in radon concentration, for instance soil thickness, bulk density, particle size distribution, and so on.The third survey was a 2 km long transect along which the soil radon concentration was measured together with soil thickness, conductivity, and slope angle every 20 m. The radon concentrations along the transect varied considerably and the variogram was periodic with a wavelength of about 263 m. Ground conductivity and slope angle also were periodic with wavelengths of about 180 m. The periodic variation of radon and ground conductivity might result from variation in soil thickness associated with joint controlled solution in the underlying limestone.

Resistivity measurements suggested that where the radon values were large in the West the soil is thick, and where they were small in the East the soil is shallow. Particle size distribution of the soil also showed some relation with radon: where the soil contains more sand and silt, the radon values are larger than where it contains more clay.

PERRIN, Olivier, (olivier@math.chalmers.se) INRA - Unite de Biometrie, Domaine Saint-Paul, Site Agroparc, 84914 Avignon Cedex 9, France; and IOVLEFF, Serge, SABRES - IUP de Vannes, Rue Yves Mainguy-Tohannic, 56000 VANNES, France

Estimating a Non-Stationary Spatial Structure by Simulated Annealing

Key Words: bijection, estimation,prediction

The underlying correlation structure of spatial environmental processes often exhibits non-stationarity. Sampson and Guttorp (1992) model the correlation r(x,y) of the spatial random field Z={Z(x), x belonging to |R^2}, as a function of Euclidean distance between locations in a bijective deformation of the geographic coordinate system. The model for r(x,y) is r(x,y)=K(||f(x)-f(y)||) where ||.|| represents the classical Euclidean norm in |R^2, f represents a bijective bi-continuous deformation and K is a known stationary and isotropic correlation function.

We propose to estimate f with a simulated annealing according to the rules of the Metropolis algorithm, with non-folding constraints, when the random field Z is observed at a finite number of geographical sites. These non-folding constraints ensure the bijection condition of the space deformation f. They consist of building the Delaunay triangulation associated to the geographical sites and of imposing that our algorithm let the topological structure of the triangulation be the same.

Our results are illustrated through spatio-temporal precipitations from 20 sites in the Languedoc-Roussillon region of France. We propose a cross-validation study to demonstrate the improvements in predictions due to the coordinate deformation.

OPENSHAW, Stan (Stan@geog.leeds.ac.uk), University of Leeds, Center for Computational Geography, School of Geography, Leeds LS2 9JT

Geographical Data Mining

Key Words: geographical data mining, artificial intelligence, GIS, geographical analysis, high performance computing

The immense explosion in geographically referenced data and lack of many suitable analysis tools in current Geographic Information Systems (GIS) software is resulting in many important data sets not being fully or appropriately analyzed. In the

machine age, information is the raw material for creating new knowledge, new discoveries, and for developing new products and services; however, most of the existing data mining tools are not suitable for making the most ofgeodata. Many data mining tools claim to function well with any data, but this overlooks the fact that geographical data are different and special. If the data riches created by GIS are to be fully used, then it seems essential to try and develop a suitable Geographical Data Mining (GDM) technology that will meet at least some of the needs for exploratory spatial analysis. This review paper examines the issues and outlines some possible solutions.

It is well known that GIS data consists of three broad classes of data types: (1) geographical coordinates, (2) temporal coordinate(s), and (3) multivariate attributes relating to the geographical entities. All three data types have unrelated measurement scales. The traditional difficulty is that most geographical analysis only starts after necessary, but data damaging, decisions concerning data selection have strangled the data. Indeed, most users are so pre-conditioned by traditional thinking that they seldom even realize how they have unwittingly harmed the unknown patterns and structures that once existed in their data. The problem is that the spatial patterns, found in geography map-space, depend on decisions made in temporal space and the multivariate data space; and vice versa. For example, if you study a disease data for the wrong time period you might find it does not cluster. Change the time period and it does cluster. Unfortunately, the choice of time period is subjective; however, the problem becomes worse. Change the definition of the disease (i.e., add or subtract one or two possibly related types to it) and you may get totally different patterns. Unfortunately, the choice of disease is subjective; change both disease classification and time period selections and you may even discover different spatial patterns in different parts of the map. It is not easy and as you gain access to more and more data at finer and finer levels of resolution, so the problems become more severe. Once data were so restricted you had no choice other than to analyze whatever you had access to. Now, thanks to developments in Information Technology (IT) and GIS, you have so much choice that conventional tools cannot cope without making the most outrageous data reduction decisions based on ignorance.

Currently, there is only one method known to the author that has been developed to explore all three spaces simultaneously; the so-calledSpace Time Attribute Creature (STAC) dating from the early 1990s. This paper seeks to describe the original STAC idea, how it worked, and then how it can be generalized and improved. Consideration is given to devising practical solutions to the problems ofparameterising the search process, leaving a trail that can be visualized in hyperspace, devising an appropriate objective function, searching for multiple near optima, how to handle data uncertainty, and coping with the results. The hypersearch method uses fuzzy logic to handle some of the problems. The outcome is a system called Geographical Data Miner (GDM/1). It is tested on both real and synthetic data sets.

GeoComputation 99

4th International Conference on GeoComputation
Mary Washington College
Fredericksburg, Virginia
July 25-28, 1999

O'SULLIVAN, David (david.osullivan@ucl.ac.uk), University College London, Centre for Advanced Spatial Analysis (CASA), Gower Street, London, WC1E 6BT, U.K.

Exploring the Structure of Space: Towards Geocomputational Theory

Key Words: spatial models, cellular automata, Geocomputation

This paper describes ongoing research that seeks to answer Couclelis' (1998a, 1998b) call for a specifically geocomputational theory in her Challenges for Geocomputation, through a phenomenological investigation of a new class of models of spatial processes.

The models themselves are not novel, but are a recently proposed adaptation of cellular automata (CA) models familiar from the field of computation. Conventional cellular automata consists of three elements: a cell space, a set of allowed cell states, and a set of rules governing the evolution of cell states over time. In these traditional CA, the cell space is an infinite regular lattice such as a grid, in which each cell has an identical neighbourhood composed of the cells that are immediately adjacent to it in the lattice. Although CA have been successfully applied in geography (see for example: Tobler, 1979, Clarke et al., 1997, White and Engelen, 1993, Xie, 1996, Batty, 1996, Batty and Xie, 1997), Couclelis (1997) points out that the spatial model implied by such models is somewhat simplistic, and that it may be appropriate to relax the requirement for a regular lattice (Couclelis, 1997, Takeyama and Couclelis, 1997). In this paper such models are referred to as graph-based CA, because this draws attention to the underlying relational structure implicit in them.

The remainder of this paper proposes a method by which the properties of such models might be usefully explored, and concludes with some initial results from such an exploration. The approach adopted seems relevant to geocomputation in that it seeks to develop our understanding of the relationships between the structure and dynamics of spatial models, and so may help in the development of more coherent perspectives on geographic space. Model structure may be characterised by using some of the numerous graph structural measures that have been proposed. A brief survey of some graph measures of the sort that may be of interest is presented.

The dynamics of model behaviour may be characterised by reference to Wolfram's (1983) classification of (regular) CA behavioural classes. Wolfram and others (for example Wuensche, 1998) have suggested entropy-like measures that may be used to classify the dynamic behaviour of these discrete systems as variously homogeneous, cyclic, complex, or chaotic. An adaptation of these entropy measures is proposed that takes into account the variations in cell neighbourhood sizes present in graph-based CA.

It is suggested that it may be profitable to bring these two sets of mathematical ideas together in a phenomenological investigation of graph-based CA models. By building a graph-based CA, running a number of starting configurations, and using the proposed entropy measures, its behavioural class can be determined. Such a set of experiments constitutes one observation. Further observations are generated by successively deforming the underlying graph, while holding the behavioural rules

constant, and running the same set of starting configurations on each deformed graph-based CA. Graph structure measures may then be used as the independent variable, as the dependent variable of dynamicbehavioural class is investigated.

Such a program of research is analagous to the search in the life sciences for a relationship between the structure of CA rules and their dynamicbehaviour (Langton, 1990), and might be described as an exploration of spatial structure.

The final part of the paper will present initial results from some experiments of the type described.

A final section will suggest practical additions to this class of models, and possible avenues for further research.

References

Batty, M. (1996), Urban Evolution on the Desktop: Simulation with the Use of Extended Cellular Automata, Environment and Planning A, 30, 1943-1967.

Batty, M. and Y. Xie, (1997), Possible Urban Automata, Environment and Planning B: Planning & Design, 24(2), 175-192.

Clarke, K. C., S. Hoppen and L. Gaydos, (1997), A Self-Modifying Cellular Automation Model of Historical Urbanization in the San Francisco Bay Area, Environment and Planning B: Planning & Design, 24(2), 247-262.

Couclelis, H. (1997), From Cellular Automata to Urban Models: New Principles for Model Development and Implementation, Environment and Planning B: Planning & Design, 24, 165-174.

Couclelis, H. (1998a), Geocomputation and Space, Environment and Planning B: Planning & Design, 25th anniversary issue, 41-47.

Couclelis, H. (1998b), Geocomputation in Context, pp. 17-29 in Geocomputation A Primer, Eds: P.A. Longley, S.M. Brooks, R. McDonnell, and B. MacMillan, (John Wiley & Sons: Chichester, England).

Langton, C. G. (1990), Computation at the Edge of Chaos: Phase Transitions and Emergent Computation, D.Physica, 42, 12-37.

Takeyama, M. and H. Couclelis, (1997), Map Dynamics: Integrating Cellular Automata and GIS through Geo-Algebra, International Journal of Geo-Information Science, 11, 73-91.

Tobler, W. R. (1979), Cellular Geography, pp. 379-386 in Philosophy in Geography,Eds: S. Gale and G. Olsson, (D. Reidel Publishing Company:Dordrecht, Holland).

White, R. and G. Engelen, (1993), Cellular Automata and Fractal Urban Form: A Cellular Modelling Approach to the Evolution of Urban Land-Use Patterns, Environment and Planning A, 25(8), 1175-1199.

Wolfram, S. (1983), Statistical Mechanics of Cellular Automata, Review of Modern Physics, 55, 601-643.

Wuensche, A. (1998), Classifying Cellular Automata Automatically, Santa Fe Institute, Working Paper 98-02-018.

Xie, Y. (1996), A Generalized Model for Cellular Urban Dynamics, Geographical Analysis, 28(4), 350-373.

PRASHKER, Steven (steve_prashker@carleton.ca), Carleton University, Department of Geography, 1125 Colonel By Drive, Ottawa, Ontario, K1S 5B6, Canada

An Improved Algorithm for Calculating the Perimeter and Area of Raster Polygons

Key Words: algorithm, polygon, perimeter,area

The algorithm presented here was designed to improve the results obtained in calculating the perimeter or area of raster polygonal areas represented in a Geographic Information System (GIS). Generally, raster GIS systems use simplistic methods for calculating the perimeter and area of a polygon. For the area calculation, they typically count the number of pixels that make up the polygon, then multiply that total by the area of one pixel to get the area of the polygon. For the perimeter calculation, they typically count the number of edges of individual pixels that are considered on the boundary of the polygon, then multiply that total by the length of the side of the pixel (assuming square pixels) to get the perimeter of the polygon. The problem is that these simplistic methods do not take into account the aliasing error that is introduced for non-vertical or non-horizontal sides of polygons. Any polygon that has diagonal edges will usually have an overestimate for the perimeter, due to counting two edges per pixel (2 times the edge length) contributing to the perimeter, rather than counting the diagonal distances (sqrt(2) times the edge length). In severe cases, this can lead to a 50 percent error for the perimeter statistic. The error is not as severe for area calculations, since only the edge pixels of a polygon contribute to the aliasing error, but it can be as high as 20 percent, if the area of the polygon is small relative to the perimeter.

The method presented here compensates for the aliasing by visiting each pixel in an image, and surrounds it with an eight-pixel grouping, thereby forming a 3 by 3 grid. This results in up to 256 different combinations of pixels of the same class value surrounding the centre pixel. For each one of these unique combinations, a value or factor is associated with that combination that will compensate for the aliasing. Once the sums

of these pixels or pixel edge factors are totaled for each unique class, they are then multiplied by the length or area of the pixel resulting in the corrected perimeter or area, respectively. The paper will examine extreme polygonal shape examples, such as square and diamond shaped polygons, in vector space, then rasterize these shapes, at varying resolutions, and compare the computed perimeters and areas with their vector equivalents. Tests will be performed on a typical raster map. Conclusions will be drawn based on empirical experiments in altering the combinations' values or factors. Preliminary results have shown the average error of raster polygon perimeter calculations approaching 2 percent, compared to the approximately 20 percent error reported from typical GIS software packages.

RICE-SNOW, Scott (ricesnow@wp.bsu.edu), Ball State University, Department of Geology, Muncie, IN 47306; RUSSELL, Joshua, Rocky Mountain Consultants, 825 Delaware Avenue, Suite 500, Longmont, CO 80501

Long-Range Persistence of Elevation and Relief Values Along the Continental Divide in the Conterminous U.S.

Key Words: drainage divides, hausdorff dimension, self-affinity, R/S analysis

We have collected a variety of elevation and relief statistics for 336 points (measurement stations) spaced at 10-km linear intervals along the U.S. Continental Divide extending from the Mexican to the Canadian border. The information was derived from 1:100,000 topographic quadrangles, with vertical resolution on the order of map contour intervals (20-50 m). These statistics can be treated as sequential data sets and tested for self-affine characteristics by rescaled range (R/S) analysis. Results show a good fit to the self-affine model for all measures tested, with sample intervals of 5 to 336 stations. For the smallest intervals available (3-5 stations) there is a somewhat lower plot slope, corresponding to higher fractal dimension (D) value. This drift at small intervals is notably opposite in sense to the "initial transients" in time-series examples given by Mandelbrot and Wallis.

Elevation data series for the Continental Divide crest yield very uniform, low, D values in the approximate range 1.02-1.04. This indicates a very high degree of persistence in values, with consistent self-affine character bridging the scales of mountain range segments, entire ranges, and full physiographic provinces. The minor variations in results for different statistics indicate slightly higher persistence for minimum Divide elevation within a 5-km radius of station, than for station point elevation or maximum Divide elevation within 5 km.

Various local relief statistics also show significant long-range persistence along the Divide, though with D values higher than obtained for elevation series. Among the statistics, relatively low fractal dimension values (approx. 1.05) indicate very high along-divide persistence for overall relief measures, including relief from Divide station to lowest point within a 10-km radius, and maximum relief available within 5 km. The least

persistence (D range 1.12-1.18) is indicated for more subtle relief along the Divide crest within 5 km of station, and relief in the immediate vicinity of the Divide, from station to lowest point within 1-km radius.

Unlike Continental Divide elevation series and associated overall relief measures, the more subtle Divide relief statistics, with their reduced degree of long-range persistence, show little systematic variation from one physiographic province to another. This suggests that divide-crest and very local divide-area relief characteristics are largely independent of regional geologic and climatic controls.

RIDENOUR, Gregory S. (ridenourg@apsu02.apsu.edu), Department of Geology & Geography, P.O. Box 4418, Austin Peay State University, Clarksville, TN 37044

Compositional Data Analysis of the Influence of Drainage Area and Stream Order on Hydraulic Geometry

Key Words: compositional data analysis, stream order, hydraulic geometry, unit-sum constrained, simplex, ternary diagram

The width, mean depth, and mean velocity of water in a stream channel are typically power functions of discharge, producing three equations collectively referred to as hydraulic geometry. Theoretically, the exponents of these three equations are unit-sum constrained. Compositional data sets are 2-D arrays whose rows sum to one, or one hundred percent. Compositional data are restricted to a subset of Cartesian space known as the simplex. Because standard multivariate statistical methods were designed for data in Cartesian space, they are inappropriate for the analysis of compositional data. The appropriate diagram for the display of a three-part composition is the ternary diagram, which was first used for comparative hydraulic geometry in 1977. A methodology for parametric, multivariate statistical inference from unit-sum constrained data was not generally available until the publication of a manuscript on compositional data analysis in 1986, which was first applied to hydraulic geometry (by this author) in 1991.

Several empirical studies have suggested that stream order (an ordinal class variable assigned to a stream on the basis of its position in the hierarchy of the branching of a stream network) is related to hydraulic geometry. For example, the hydraulic geometry between the main channel of the Suia-Missu River in Brazil differs significantly from its smaller tributaries; however, this hypothesis was tested with three separate F-tests, a shortcoming referred to as "partial analysis." Multivariate discriminant function analysis was employed to test the hypothesis that stream order in the Pecatonica drainage basin in southwestern Wisconsin is dependent upon hydraulic geometry parameters, including the logarithms of width, depth, velocity, suspended load, slope, and roughness. Though all of these variables, except load, were found to be fairly effective in discriminating streamorder, the strength of the relationships were judged to be less than expected.

If systematic relationships exist between stream order, drainage area, and hydraulic geometry, then qualitative or semi-quantitative approximations for the outcomes of such models in stream reaches for which input data (width, depth, velocity, and discharge) is not available, becomes possible by modeling other streams of similar order and/or drainage area within the same basin. In this study, compositional data analysis was used to test three intrabasin hypotheses: that hydraulic geometry (1) is correlated with the drainage area above a given point on a stream (higher stream orders should drain larger areas), (2) differs significantly between a main channel and its tributaries, and (3) is functionally related to proportional stream order, which uses rational numbers to account for the variation within a stream segment of a given order. The hypotheses were tested using previously unpublished data presented for the upper Sabine River Basin in northeastern Texas and from data in the literature for the Sangamon River Basin in central Illinois. This study differs from the previously mentioned multivariate analysis in that it is the unit-sum constrained exponents of the hydraulic geometry equations, rather than their dependent variables (width, depth, and velocity) that were analyzed, thus requiring the use of compositional data analysis.

Hydraulic geometry was determined to have no correlation with proportional stream order nor drainage area, and showed no significant difference between the main channel and its tributaries, suggesting that more site-specific factors (e.g., channel shape, bed composition, bank stability) are probably controlling hydraulic geometry in these stream networks.

SCHMIDT, Jochen (jochen.schmidt@uni-bonn.de), GAERTNER, Holger and DIKAU, Richard, University of Bonn, Department of Geography, Meckenheimer Allee 166, D-53115 Bonn, Germany

Investigations in Slope Development through Landslide Activity - Concepts, Methods, and Implications for Interdisciplinary and Interoperable Data Management

Key Words: landslide investigation, slope stability, landform evolution, data integration

It is an open research question, to which extent landslide activity contributes to landform evolution, especially under moderate humid climatic conditions? In a multidisciplinary research project at the University of Bonn, we are trying to get insight into the process of slope development through mass movements. Research methods include local field investigation and stability analysis, mapping and modelling of landslide susceptibility, geophysical subsurface monitoring, and geomorphometric slope profile analysis. The research aim of the coupled use of these different techniques is to model quantitative measures for sediment transport through landslides on hillslopes in the Bonn area. In this paper, the research approach, methods, and a few first results are presented. Special emphasis is given to the data management. The scheme and system

used to store and manage the data and analysis results is discussed. Additionally, an alternative, object-oriented approach is presented.

In quantifying the sediment transport through mass movements over longer time scales, we have to cope with several problems, e.g., (1) unknown boundary conditions in time, (2) discontinuity of the process, (3) different process types, and (4) coupling with other (slope) processes; therefore, a statistical approach, coupled with slope stability analysis is used to estimate the contribution of landslides to slope evolution under variable (i) climatic, (ii) morphometric, and (iii) geologic boundary conditions. First, several indices describing mass movements over longer time scales were identified and selected for the study. Slope profile types are extracted for the study area using several morphometric algorithms. Material properties from laboratory tests are related to geologic units. Several slope stability models based on different approaches are used to model the selected landslide indices under different morphometric and geologic conditions. The models are calibrated using data from several landslides in the Bonn area. Results show the dependency of landslide occurrence in the Bonn area on hillslope morphometry and geology. These findings are used in a sensitivity study on the influence of different climatic conditions, which shall lead to an estimation of spatio-temporal landslide activity. Investigations on sites near Bonn are carried out, producing a large amount of field and laboratory data. Supplementary information is available by climate data, geologic maps, topographic maps, DEMs, etc. Additionally, data resulting from interpretation, analysis, and modelling of the field data must be handled.

In our project, storage, visualization, and analysis of these data is realized using GIS (Arc/Info, GRASS), geotechnical software (GeoDIN), databases (Access, Oracle), slope stability programs, and several other software products. This shows that landslide investigation is a typical example for the heterogeneity of data and methods used in Geosciences, which necessitates a careful and consistent data management. Presently, a relational data model and diverse structures, methods, and tools to handle and analyze the data are used. This (common) practice complicates the exchange of data, methods, and research results. In contrast, an object-oriented approach, developed in cooperation with a project on Open Information Systems, is compared with classical concepts. The results show that object-oriented data modelling can facilitate user access to multiple datasets, support integrated use of different analysis technologies, and could aid in the development of standards for exchanging data in a multidisciplinary environment.

SCHUMACHER, Brian J. (bjschu@juno.com), LEITNER, Michael, Louisiana State University, 113 Howe-Russell Geoscience Complex, Department of Geography and Anthropology, Baton Rouge, LA 70803

Spatial Crime Displacement Resulting from Large-Scale Urban Renewal Programs in the City of Baltimore, MD: A GIS Modeling Approach

Key Words: spatial crime displacement; GIS modeling; Baltimore, MD

The city of Baltimore, Maryland has received widespread acclaim for the aggressive programs of urban renewal it has implemented since the mid-1970s, most notably around the Inner Harbor and the Fells Point sections of its downtown area. However, while the aesthetic improvements and the influx of revenue generated by the programs have been studied at length, the effect that these programs have had on the spatial distribution of criminal activity in the city has remained largely unknown. This is especially important because while crime has decreased (due to the urban renewal programs around the Inner Harbor and Fells Point), the city's overall crime rates have remained unchanged and among the highest in the country. The main objective of this research is to determine how the spatial distribution of crime has changed in the wake of the downtown redevelopment and to propose models of spatial crime displacement. Such models will help city planning bureaus and law enforcement agencies in Baltimore, and other urban areas, with pro-active decision making with the goal to decrease overall criminal activities.

The study area for this project includes the Inner Harbor, Fells Point, and adjacent areas, covering three police districts (the first, the second, and the ninth) and fifty census tracts. For this area, address-level crime data (robberies and burglaries) have been obtained from the Baltimore City Police Department for the period 1988 through 1998. Additionally, various economic, educational, social, and physical variables have been collected from the 1980 and 1990 U.S. Population Census. The methods used to develop models of crime displacement include measures of spatial distribution, hot-spot analyses, and multiple regression analyses. Hot-spot analyses are applied to identify the highest concentrations of robberies and burglaries within the study area and how they changed over time due to changes in the census variables. The regression analyses are used to determine the specific characteristics of places within the study area witnessing changes in their crime rates. Geographic Information Systems (GIS) technology is used for the storage and analysis of the data, as well as for the development and visualization of the model results.

This paper will address the following conference topics: advances in Geographical Information Systems, particularly in the area of spatial analysis; statistical modeling (predictive and descriptive); and geostatistics.

SEE, Linda (L.See@geog.leeds.ac.uk), ABRAHART, Robert J. (babrahart@hotmail.com), Center for Computational Geography, School of Geography, University of Leeds, Leeds, LS2 9JT, U.K.

Multi-Model Data Fusion for Hydrological Forecasting

Key Words: data fusion, neural networks, hydrologicalmodelling

Data fusion is an expanding area of research that deals with the integration of information from multiple sensors and/or data sources to obtain a better solution than could otherwise be achieved with the use of single-source data on its own. The different

approaches that could be adopted for fusing multisource data include conventional, statistical, and numerical methods as well as artificial intelligence techniques such as neural networks, genetic algorithms, and fuzzy-logic. The emergence of new sensors, advanced fusion algorithms, together with improved hardware and software, make real-time data fusion a practical option for automated target recognition systems, applied robotics, and remote sensing. The opportunities that are afforded from data fusion also are important in other areas of commerce and science, and are seen to have clear implications for different types of geographical research and in application-related fields such as hydrological modelling. Data fusion can be implemented at different scales of operation and with different strengths and complexities. The simplest type of data fusion operation would involve the integration of input data from multiple sources to produce output data. This is referred to as the "data in - data out" fusion architecture. But data fusion also can operate at more complicated or elaborate feature-based or decision-based levels which could, in turn, have significant implications for the design and construction of enhanced automated flood prediction and flood warning systems.

This paper provides results from some initial explorations that have been undertaken on the use of neural networks for fusing hydrological river flow forecasts, and involves a comparison between two different river catchments in the United Kingdom-the Upper River Wye and the River Ouse. The data sources in this instance were not sensors but different predictive models that have been run in parallel. The operation was based on fusing multisource model output data, obtained from conventional statistical and numerical methods, and from various artificial intelligence techniques. The new data fusion predictions were based on different combinations of these single-model predictions, with some additional experiments being carried out in association with the historical hydrological record.

SHINE, James A. (jshine@tec.army.mil), WAKEFIELD, Gery I., U.S. Army Topographic Engineering Center, ATTN: CETEC-TR-G, 7701 Telegraph Road, Alexandria VA 22315-3864

A Comparison of Supervised Imagery Classification Using Analyst-Chosen and Geostatistically Chosen Training Sets

Key Words: supervised classification, geostatistics, spatial variation,imagery analysis

A continuing challenge in image processing is the classification of spatial imagery into categories. Examples of these categories are: roads, urban areas, evergreen trees, deciduous trees, water, and grasslands. The accurate classification of images has a wide range of applications, including reconnaissance, assessment of environmental damage, land use monitoring, urban planning, and growth regulation.

One classification approach is supervised classification. The imagery is divided into training data and test data. The correct categories are known for the training data, and some classification approach is specified based on this data. This approach is then

used to classify the test data. Some approaches include classification trees, minimum distance statistical approaches, and neural networks. The choice of a good training set can have significant influence on the success of a classification approach.

A common technique in imagery classification is selection of good test data points by an experienced analyst. An image or set of registered images is viewed in image processing software such as ERDAS Imagine, and some pixels that are unambiguous in each of the desired classification categories are selected and used for training data. The entire image is then classified based on the training data metrics. In cases where ground truth is available, classification accuracy can be assessed by use of an error matrix. This method can be time consuming and requires expertise on the part of the analyst, something not available in all classification settings.

A recent approach uses spatial variation scales from geostatistical analysis to choose the training points rather than an analyst's choices. A semivariogram is computed on the pixel values of an image, and a spatial variation scale is determined from this semivariogram. A grid of points chosen from this scale (usually 50 percent of the scale) is then selected for the training data and is then used for the classification. This approach does not require an experienced analyst.

Experiments comparing these two approaches have been conducted using several registered images of Fort A.P. Hill, Virginia, which have accompanying accurate ground truth. The experiments were performed in ERDAS Imagine 8.3 using a maximum distance supervised classifier. The results of the error matrices for the two approaches were not statistically different. Geostatistically chosen training data has the potential to reduce the need for experienced image analysts to perform imagery classification. Further developments may make further automation of the imagery classification process possible.

TAYLOR, Stephen C. (staylor@atlsci.com), ARMOUR, Bernard, HUGHES, William H., KULT, Andrew, NIZMAN, Chris, Atlantis Scientific, Inc., 20 Colonnade Road, Suite 110, Nepean, Ontario, K2E 7M6, Canada

Operational Interferometric SAR Data Processing for RADARSAT Using a Distributed Computing Environment

Key Words: distributed computing, SAR interferometry

There has recently been a surge of interest in interferometric value-added products for SAR data by RADARSAT and other spaceborne SAR sensors. These products primarily include digital elevation models (DEMs) and deformation maps. The high level of demand has put a strain on traditional interferometric SAR (InSAR) processing tools, which have been designed for research purposes. We present a production, or operational, approach to InSAR processing that emphasizes maximal throughput and algorithmic robustness, as well as around-the-clock operation requiring relatively little/unskilled supervision. The basis for this approach is an application

environment known as A Performance Emphasized Production Environment for Remote Sensing, or peppers ™.

Peppers™ uses a distributed parallel computing tile-based pipelined approach to processing remotely sensed image data products, which are characterized by their very large size (often more than 500MB). It makes use of the inherently localized nature of most image processing algorithms to formulate the division labour both spatially across the image and temporally along the pipeline of transformations. This formulation can be used to split the load among all available processing units, either using multiple threads of execution on a multi-CPU computer, or using a message passing technique (MPT) to coordinate multiple processes distributed across a networked cluster of computers.

Peppers™ represents an end-to-end approach to operational processing of remote sensing data. It's components manage every aspect of the production run, including such things as automatic update of intermediate products following a change, auditing of historical results, and the ability to trap exceptional conditions and redirect the approach to overcome these conditions. It is anticipated thatpeppers™ will form the basis for future operational remote sensing applications.

The development of peppers™ is supported in part by the RADARSAT User Development Program of the Canadian Space Agency.

TROTT, Kevin C. (kevin_trott@partech.com) and GREASLEY, IAN (ian_greasley@partech.com), PAR Government Systems Corporation, 314 S. Jay Street, Rome, NY 13440

A 3-D Topology Model for Vector Data

Key Words: topology, 3-D, volumes, vector data, NIMA, VPF

This paper describes an object-oriented 3-D spatial data model, based on the National Imagery and Mapping Agency's (NIMA) Vector Product Format (VPF), which is capable of supporting high-resolution 3-D representations of natural and man-made environments with full 3-D topology. The development of this spatial data model was driven by the requirements of creating timely, accurate, and highly detailed synthetic environment databases that provide realistic representations of real-world locations for military training and mission rehearsal purposes. Simulated, computer-generated military forces must be able to perform topological reasoning on the contents of such databases in order move and fight in a realistic manner. These databases often violate the assumptions that are relied upon by traditional 2-D digital topographic databases. For example, the assumption that elevation has a single value at a specified 2-D location does not necessarily hold true. Structures such as bridges, overpasses, tunnels, and the interiors of buildings cannot be adequately represented using 2-D topology; therefore, a spatial data model that supports 3-D topology is needed.

Previous work in the development of data models that support 3-D topology, from both the GIS and computer graphics communities, is briefly reviewed. The 3-D

topological data model is then described in detail. The spatial entities that make up the data model, including nodes, edges, faces, and volumes, are each defined. Organizational groupings of spatial entities, including rings, which group and order the edges that bound a face, and shells, which group the faces that bound a volume, also are defined.

All of the topological relationships among the 3-D spatial entities are then defined. The 3-D topological relationship between a node and its collection of connected edges is no longer ordered, since the relationship between edges and faces becomes much more complex than in 2-D planar topology. Each edge has an ordered collection of zero, or more, adjacent faces, while each face is bounded by an ordered collection of one or more edges. Two additional levels of topology, beyond planar topology, are defined: Level 4, which provides partial 3-D topology, in which the rules of planar topology no longer fully hold, and Level 5, which provides full 3-D topology, in which space is partitioned into a collection of mutually exclusive and exhaustive volumes. An extended set of access primitives to support 3-D topological operations also is defined.

Volume features, such as buildings and bodies of water, are defined and discussed. Several examples are presented that illustrate the use of the 3-D topology data model to represent situations that cannot be adequately represented using 2-D planar topology.

TUCKER, Greg (gtucker@mit.edu), GASPARINI, Nicole, BRAS, Rafael, RYBARCZYK, Scott, Massachusetts Institute of Technology, Department of Civil and Environmental Engineering, Building 48, Room 429, Cambridge, MA 02139; LANCASTER, Stephen, Oregon State University, Department of Geosciences, Forest Sciences Laboratory, 3200 SW Jefferson Way, Corvallis, OR 97331

Toward an Object-Oriented Toolkit for Distributed Environmental Modeling

Key Words: simulation, model,geomorphology, runoff, erosion, sedimentation, software, Delaunay triangulation, TIN

In recent years, spatially distributed models of land surface processes, such as runoff and erosion, have come into widespread use in the Earth and environmental sciences. As these models grow in sophistication, the software engineering effort required to implement them also expands. Therein lies the need for portable, modular codes that can implement many of the basic requirements of a distributed model in a flexible, efficient, and application-independent manner. Here, we describe a simple prototype of such a system and its use in modeling long-term landscape evolution and short-term flood forecasting.

Distributed models of surface processes such as runoff, vegetation growth, soil erosion, forest fires, landscape evolution, and other processes typically share a number of important features in common: all involve (1) spatial division of terrain into discrete elements, (2) storage of mass and/or energy within landscape elements, (3) routing of flows of mass (e.g., water) and/or energy among landscape elements, (4) dynamic

updating of boundary conditions (e.g., rainfall input), and (5) dynamic updating of state variables (e.g., soil moisture and surface elevation) through time. Often, the programming effort required to implement these features is non-trivial and quite labor intensive. This is particularly true when the underlying spatial representation is irregular; for example, the case of models based on triangulated irregular networks. Although current GIS systems provide sophisticated capabilities for spatial representation of data, performance and other limitations make them unsuitable for computationally intensive dynamic (i.e., time evolving) simulations; thus, to reduce software development times and minimize duplication of effort, it would be advantageous to develop application-independent modeling routines that would provide the underlying space and time structure for distributed models without dictating the processes or state variables.

We present a simple prototype of one such system. The system is an outgrowth of parallel ongoing efforts in (1) modeling rainfall, runoff, andstreamflow for real-time flood forecasting, and (2) modeling long-term drainage basin evolution via uplift, erosion, and sedimentation. The system is based on a triangulated irregular network (TIN) representation of terrain using the Delaunay triangulation. The TIN representation is implemented via a set of C++ classes, which are in effectdecoupled from the process routines. Efficient storage of mesh elements is accomplished by using a "dual edge" data structure, an adaptation of the well known Quad Edge structure for Delaunay triangulations. Capabilities of the system include the ability to construct triangulations from a given set of points, calculation of polygon areas and edge lengths, and the ability to move, add, or subtract points dynamically during a simulation. A class inheritance hierarchy is developed in order to enable new applications (such as models of watershed slope stability) to take advantage of existing code for more general processes (such as routing of overland flow across a topographic surface).

The system is illustrated through several examples, including simulation of long-term river meandering within a floodplain, and simulation of event-driven rainfall and runoff. We discuss some of the advantages and disadvantages of the modular, object-oriented approach in terms of system performance, ease of development, and related issues.

TURTON, Ian, OPENSHAW, Stan, and MACGILL, James (pgjm@geog.leeds.ac.uk), University of Leeds, Centre for Computational Geography, Leeds, LS2 9JT, U.K.

Smart Crime Pattern Analysis Using the Geographical Analysis Machine

Key Words: smart pattern analysis, clusters, crime

The presentation will give an introduction to automated crime pattern hot-spot detection using methods that are easily accessible with a very short learning curve. The focus of the presentation will be a case study of crime pattern analysis using the Geographical Analysis Machine (GAM) and other related methods developed at the Centre for Computational Geography (CCG). The example used is a crime data set for

Baltimore County, with a variety of crime types being analyzed for patterns. This allows the detailed examination of the methods developed by the CCG.

The GAM (Openshaw et al., 1987, 1988) was an early attempt at automated exploratory spatial data analysis that was easy to understand. The GAM sought to answer a simple practical question; namely given some point referenced data of something interesting. Where might there be evidence of localized clustering if you do not know in advance where to look because of the lack of knowledge of possible causal mechanisms, or if prior knowledge of the data precluded testing hypotheses on the database. More simply put, here is a geographically referenced database, now tell me if there are any clusters or crime hot spots and, if so, where are they located? It offers a solution to those researchers and users of GIS who want to perform a fast exploratory geographical analysis of their data with a minimum of effort. It is an automated procedure that is designed to yield safe results that are largely self-evident.

GAM reflects the view that useful spatial analysis tools have to be able to cope with both the special nature of spatial data and end-users who do not have degrees in statistics. The results also have to be easily understood and self-evident so that they can be readily communicated to other non-experts. This need has been clearly expressed as follows: We want a push-button tool of academic respectability where all the "heavy stuff" happens behind the scenes but the results cannot be misinterpreted (Adrian Mckeon, Infoshare: email: 1997). There also is a requirement for results expressed as pretty pictures rather than statistics.

The GAM have been further developed to produce the Geographical Explanations Machine (GEM) that attempts to explain, in a geographical way, associations between hot spots and other GIS data coverages such as socioeconomics and other geodata.

Further developments have been made recently to improve the search methods used in GAM/GEM, since as dataset sizes grow, it becomes computationally unfeasible to exhaustively search the entire data space. Two of these new search techniques will be briefly discussed; the first MAPEX uses a genetic algorithm to control the search that reduces the search times by an order of magnitude compared to GAM. The second is a more experimental method using smart agents to explore the dataset in an intelligent manner.

Van der GAAST, J.W.J. (J.W.J.vanderGaast@sc.dlo.nl) and KROES, J.G., DLO Winand Staring Centre, 6700 AC

Wageningen, The Netherlands

A Spatial Solution to Calculate Optimum Surface Water Target Levels Using a Water Management Decision Support System

Key Words: ArcView, crop yields, decision support, drainage area, fuzzy classification, groundwater, hydrological modelling, lowland areas, surface water, SWAP, topography, water management

The interaction between ground water and surface water in lowland areas plays an important role in planning surface water levels for fixed drainage level areas. One important factor in the interaction between ground water and surface water is topography. High-resolution topographic data in combination with hydrological models and Geographic Information Systems (GIS) makes it possible to calculate spatial effects of surface water target levels. The Water Management Decision Support system integrates a hydrological dynamic model with ArcView GIS to provide decision support for optimum target-level calculations. The hydrological model (SWAP) simulates the interaction between one-dimensional soil moisture, heat, and solute transport in the unsaturated-saturated upper part of the soil profile. The model was used to calculate the ground water regime for different target levels. The calculated ground water regime for a long-term period was used to calculate crop yields. The relative crop yields for the calculated groundwater regime provides a fuzzy classification for differences in topography within an area using specified drainage levels. Combinations of the fuzzy classification and high-resolution topographic data provides spatial information on crop yields. This paper describes a water management decision support system for lowland areas. This will be the first step to a valuable tool for water managers. It will provide information on the spatial effects on crop yields for a given surface water level. It can be used for a regional evaluation and design of drainage systems with optimum surface water target levels.

WARD, Douglas P. (s195448@student.uq.edu.au), MURRAY, Alan T. and PHINN, Stuart R., University of Queensland, Department of Geographical Sciences and Planning, Brisbane, Queensland 4072, Australia

Integrating Cellular Automata and Spatial Optimization for Evaluating Rapidly Urbanizing Regions

Key Words: cellular automata, spatial optimization, urban growth, sustainable development

Rapidly urbanized regions face particular issues associated with sustainable development in which the spatial nature of urban form has economic, environmental, and social implications. Models of landscape transformations that characterize urban growth in terms of socio-economic and bio-physical factors can provide valuable tools for planners to explore urban scenarios that result from different land-use policies. Urban development can be conceived as a self-organizing system in which natural constraints and institutional controls associated with land-use policies temper the way in which local decision making processes produce macroscopic patterns of urban form. In this paper, a cellular automata (CA) model that simulates local decision making processes is integrated with an optimization framework that addresses issues of sustainable urban development. In the model, CA transition rules are modified in accordance with the outcomes of the optimization of economic, social, and environmental target thresholds

associated with sustainable urban development. The model provides a means for simulating different land-use scenarios to reveal the implications of different land-use policies. Application results evaluating possible growth scenarios for a rapidly urbanized region in eastern Australia are given.

WARE, J. Mark (jmware@glam.ac.uk), and JONES, Christopher B., University of Glamorgan, School of Computing, Pontypridd CF37 1DL, Wales, U.K.

Error-Constrained Change Detection

Key Words: equivalence testing, data uncertainty, Bayesian classification, locational error

We consider the problem of detecting real-world changes that occur over a period of time. This is achieved by comparing a pair of vector-defined polygon coverages, A and B; A represents a given theme for a particular region at the start of the time-period (t_1), and B represents the same theme and region at the end of the time-period (t_2). The standard technique for detecting change is to intersect A and B so as to produce a third coverage C. Each intersection-polygon in C is associated with two theme classification codes; one describing theme class at t_1, the other describing theme class at t_2. If these classification codes differ, then the polygon is labeled as corresponding to an area that has undergone change. A problem with adopting this approach is that it assumes the source coverages are free of error. This will not usually be the case; it is likely that A and B are subject to both classification error (i.e., some polygons are assigned to a wrong class) and positional error (i.e., some polygon boundaries are in the wrong location).

In this paper we hope to improve on the standard change detection procedure. To date, we concern ourselves with positional error only. We try to achieve improvement by matching and aligning boundaries (or parts of boundaries) in A with equivalent boundaries (or parts of boundaries) in B. The term equivalent boundaries is used here to describe a pair of boundaries that are intended to be representing the same real-world phenomena. It is most probable that, because of locational error, equivalent boundaries will not match exactly with regards to their geometry. Boundary matching is achieved using Bayesian multivariate classification, in which candidate matching pairs are classed as being actual matches or not. To begin, conditional a priori probability is calculated using geometric signatures obtained from a training set of manually classified boundary pairs. The signatures used include comparisons based on length, sinuosity, bandwidth, anchor length, and angle. Next, candidate matching boundary pairs are found using a crude feature matching procedure based on buffering. In essence, this procedure deems two boundaries (or parts of boundaries), bA and bB, to be a candidate match if all of bA lies within a pre-defined distance of bB and all of bB lies within a pre-defined distance of bA. Each candidate matching pair (bA, bB) is then assigned a posteriori probability of it being an actual match. This probability is calculated using signatures derived from (bA, bB) and the previously derived a priori probability. Having been calculated, the a

posteriori probabilities of all candidate matching boundary pairs are examined and compared against a pre-defined probability threshold value, and a list of matching boundary pairs is produced. Each matching boundary pair is then aligned using a boundary merging procedure. This procedure makes use of weighted interpolation, thus, allowing for a range of results (i.e., bA replaces bB in B, or bB replaces bA in A, or bA and bB are both replaced by a weighted average). Having updated A and B, we can produce a modified intersection coverage CM. The assumption is that when compared to C, the change statistics derived from CM will more accurately represent change that has taken place in reality.

The above method has been implemented as a collection of C functions. These functions are callable from within an ArcView Script and can be applied to ArcView Shapefiles. Testing is being carried out on land-cover data supplied by the Macaulay Land Use Research Institute. Three data sets are available, each representing the same region of the Cairngorms (Scotland) at different times (1946, 1964 and 1988). These data have been manually processed, and all matching boundary pairs were recorded. A subset of the data was set aside for training purposes. The full paper will report on a comparison of the results of automatic change detection, based on the training set, with the independent data produced by expert manual interpretation.

WATTS, Joseph (watts@tec.army.mil), JARRETT, Joni (jarrett@tec.army.mil), WAKEFIELD, Gery (wakefield@tec.army.mil), SLOCUM, Kevin (slocum@tec.army.mil), U.S. Army Topographic Engineering Center, Topographic Research Division, Alexandria, VA 22315; PRECHT, Francis (f_precht@frostburg.edu), Frostburg State University, Department of Geography, Frostburg, MD 21532; FELS, John (fels@unity.ncsu.edu), North Carolina State University, Raleigh, NC 27695

Reliability-Based Vegetation Mapping: The Integration of Predictive Modeling and Digital Image Processing

Key Words: vegetation response modeling, predictive vegetation mapping, classification tree analysis, discriminant analysis, digital image processing, accuracy assessment, Huachuca Mountains

Army tactical operations and military training land management demand geospatial data of known accuracy on vegetation type and structure. To address these requirements, we have integrated predictive vegetation modeling and digital image processing for the Garden Canyon watershed on the Fort Huachuca Military Reservation in southeastern Arizona. Vegetation plots, topographic data layers, ancillary GIS layers on geology, geomorphology, soils, and medium resolution multispectral images have served as input to classification tree, discriminant analysis, and maximum likelihood algorithms. We examined vegetation classifications in terms of accuracies and spatial structures for separate trials of each algorithm and then for integrated algorithmic approaches.

The techniques used to relate heuristic and empirical associations or rules of vegetation types to environmental conditions can be classified as pre-, during, and post-classification procedures. Pre-classification rule application involved masking image data based on elevation and substrate zones and buffering riparian zones. During the classification process, the data layer selection was guided by empirical associations of base imagery, linear image transforms, and topographic data layers to field vegetation data. Post-classification overlay operations combined predictions attributed by classification probabilities. Our overall objective is to define optimal vegetation classification procedures that integrate complementary algorithms at field and landscape scales.

WHITE, Roger (roger@riks.nl), Memorial University of Newfoundland, Department of Geography, St. John's, Newfoundland, Canada

High Resolution Integrated Modelling of the Spatial Dynamics of Urban and Regional Systems

Keywords: cellular automata, land use, integrated models, process models, spatial dynamics

An emerging branch of geocomputing involves the modelling of spatial processes. This work builds on the resources provided by remote sensing and GIS, and, to a lesser extent, spatial statistics, to show how spatial systems evolve in time. The fundamental logic of most spatial process models represents causal relations, and the models are, in a formal sense, predictive; they are also, to a lesser degree, empirically predictive. A variety of techniques are being used, the most important being traditional regionalized system dynamics approaches, multiagent systems (MA), and cellular automata (CA). The techniques are frequently combined to model processes operating at different spatial scales.

Cellular automata are dynamical systems defined on a raster space. Cell states typically represent land use and land cover, and the transition of a cell from one state to another depends on the states of cells within a neighbourhood of the cell. The cell space may either be homogeneous, in the case of theoretical applications designed to investigate basic properties of dynamic spatial systems, or inhomogeneous, in the case of most applications to actual geographical systems, where the inhomogeneities may represent such factors as suitability, accessibility, or legal restrictions on land use; thus, a cellular model may be thought of as running on top of a GIS. In a pure CA, the number of cells in a particular state is determined endogenously by the cellular dynamics; however, this is not realistic for many geographical applications, since the number of cells usually reflects the level of demand for the activity that is carried out on the cells. Most applied cellular models are constrained to generate particular numbers of cells in each state, with the target cell numbers determined exogenously, usually by another model.

Urban and regional models based on cellular automata give good representations of the spatial dynamics of land use and land cover, as judged by a variety of measures, from the Kappa index to various fractal dimensions and visual inspection. In a current application, a cellular model of The Netherlands as 500-m resolution is driven by a macro-scale dynamic spatial interaction model defined on 40 economic regions; this model is in turn driven by national planning projections and policy goals. Given the national totals, the macro scale model generates regional demands for population and economic activity. These demands are translated into demands for cell space, which the CA then attempts to allocate. In turn, information on conditions at the cellular level, such as the quantity and quality of land available to various activities and actual densities at the cellular scale, are returned to the regional model to modify parameter values there. Linking the two models operating at the two scales improves the performance of both; for example, regional population estimates are improved by about 65 percent when the macro scale model is linked to the cellular level. The purpose of this multiscale model, developed for the Netherlands Institute of Health and the Environment, is to permit the evaluation of national and regional policies in terms of their effects on the natural and human environment. This is only possible because the CA model is able to translate the national trends and policies into potential consequences at the micro scale where most effects will be experienced.

The results of high resolution modelling of spatial dynamics raise a number of methodological issues. One of the most pressing concerns evaluation of the results. In calibration and testing, a model-generated map must be compared to an actual land use/land cover map, but current pixel by pixel techniques are unable to capture patterns and often misrepresent the degree of similarity, while other measures, such as fractal dimensions, are too general. Approaches based on fuzzy logic and pattern recognition techniques show some promise. Another issue concerns predictability. To the extent that these models capture the evolving, innovative nature of real cities and regions, they cannot be strictly predictive, nor can they reliably be characterized by statistical measures because the statistical patterns may undergo punctuated evolution as the system's history unfolds. We know that the world is a complex mix of predictability, uncertainty, and novelty. A CA-based modelling approach captures this, but at the same time it implies a different kind of science, for which the methodology and standards have not yet been fully worked out.

ZHU, Honglei (hzhu@clarku.edu) and SCHNEIDER, Kristin, Clark University, Clark Labs, 950 Main Street, Worcester, MA 01610

Flat Feature Processes from Triangulated Irregular Networks for Hydrological Modeling

Key Words: triangulated irregular network, flat triangle,digital elevation model

Regular grid digital elevation models (DEM) and triangulated irregular networks (TIN) are widely used in a geographic information systems (GIS) to represent a digital terrain model. A GIS can be used to extract geomorphic information from a DEM or a TIN for hydrologic modeling. During processing, depressions and flat areas in such models prevent us from obtaining the information that hydrologic modeling requires. This paper proposes algorithms for removing flat features in a TIN model. Flat triangles are removed, then flat ridges and channels are processed.

1. Remove flat triangles in a TIN model. Flow directions cannot be calculated if a triangle is flat. A TIN model must be flat-triangle-free before it is used to extract features such as channel networks or watershed boundaries. Two cases are considered in this study, and algorithms are developed.

A) TIN model generated from a discrete point data set: Flat triangles are processed according to the status of flat triangles' neighboring triangles. Flat triangles without any flat triangle as neighbors are always given higher priority; therefore, stand-alone flat triangles are first processed in flat feature removal procedures. Flat triangles with three flat neighboring triangles are not considered until their neighbors are processed.

i) If a flat triangle is a stand-alone triangle, a centroid point at the center of a flat triangle is added, the point's elevation is interpolated, the flat triangle is split, and the TIN is adjusted.

ii) If a flat triangle has one or two non-flat triangles, a vertex's elevation of the flat triangle is adjusted, or a point at the middle of a flat edge is added, or two points at the middle of two flat edges are added, and then, elevation values are interpolated, and the TIN is adjusted.

B) TIN model generated from a contour data set: Edges known as bridge and tunnel edges form flat triangles in a TIN when a TIN is generated from a contour data set. The contour data set is used to assist the process of flat triangle removal. By inserting a point at the middle of bridge or tunnel edges, flat triangles can be removed from a TIN model. The elevations of added points are interpolated, and the TIN is adjusted.

2. Flat Ridge and Channel Processing. Flat ridges and channels may prevent us from locating depression points and from obtaining the flow direction information based on a TIN model. Algorithms for processing such features are proposed. Neighboring triangles are used to assist the process.

Validation of the methodology is provided via a case study.

Index of Authors