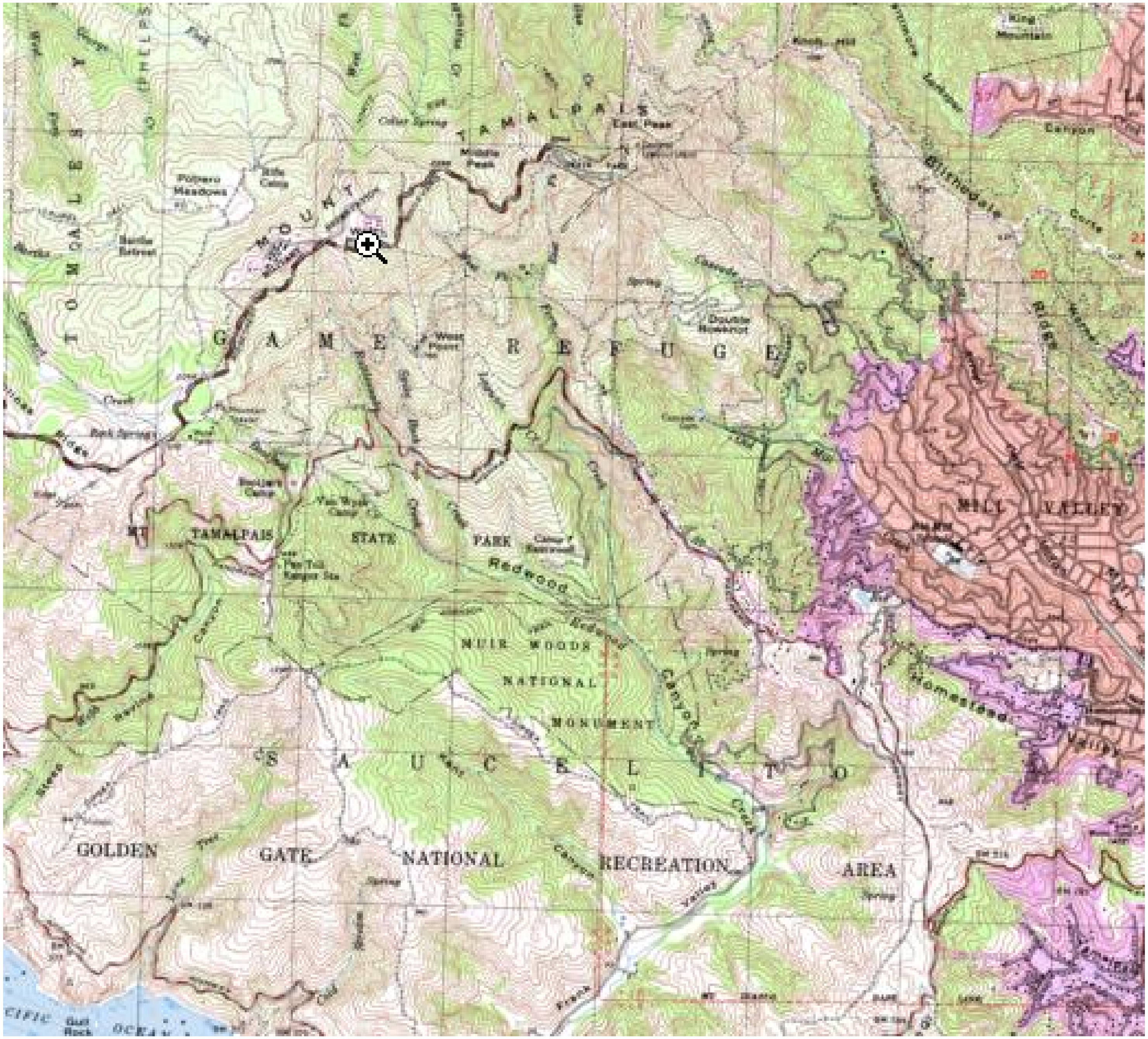
Best Practices for Visualizing Uncertain Information in Geography

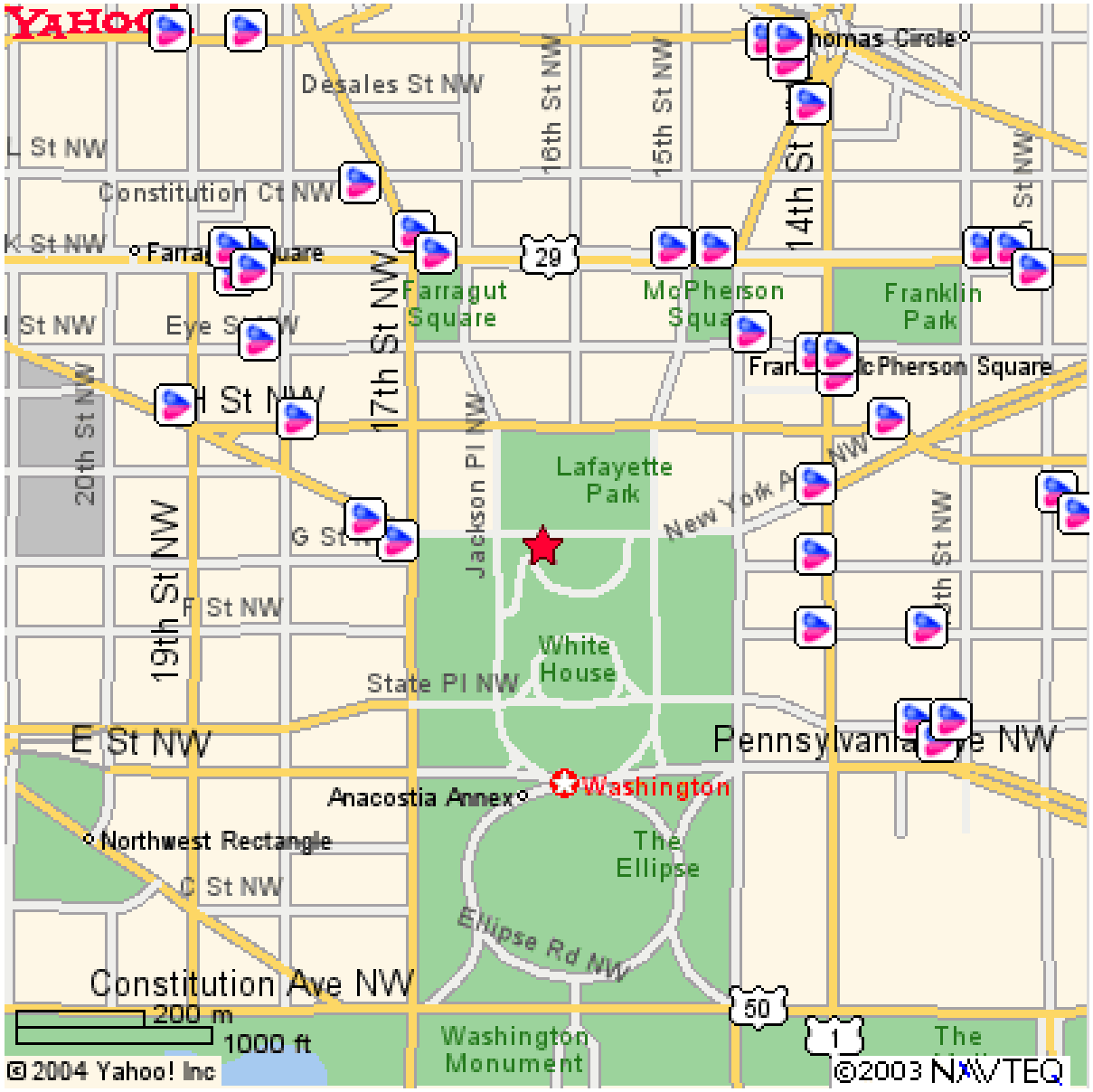
Michael F. Goodchild
University of California
Santa Barbara

Geographic information

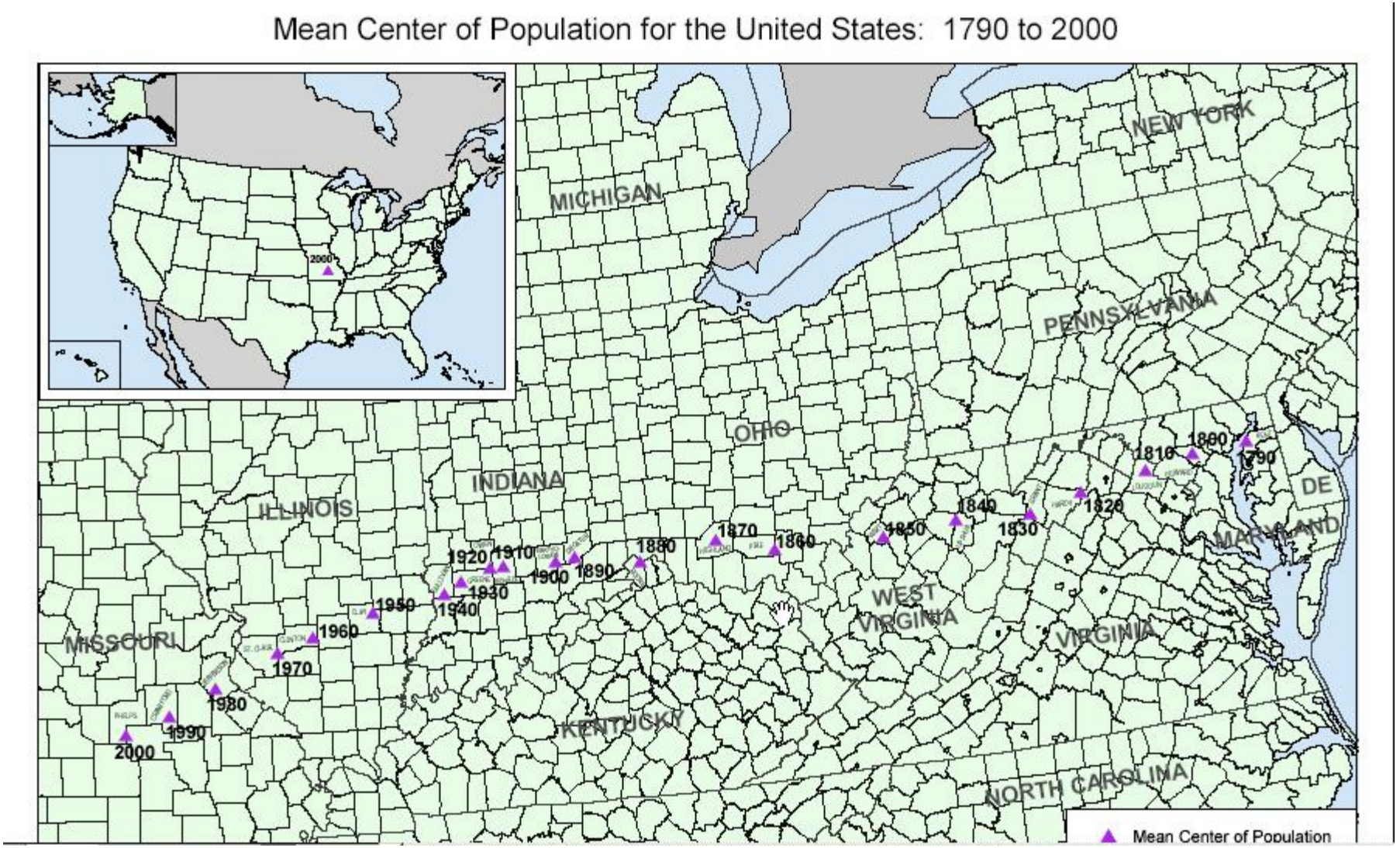
- Maps and images
  - geographically referenced data
  - events and transactions
- The atomic form
  - **<x,z>**
  - observed for point data
  - otherwise aggregated to statements about lines, areas, volumes



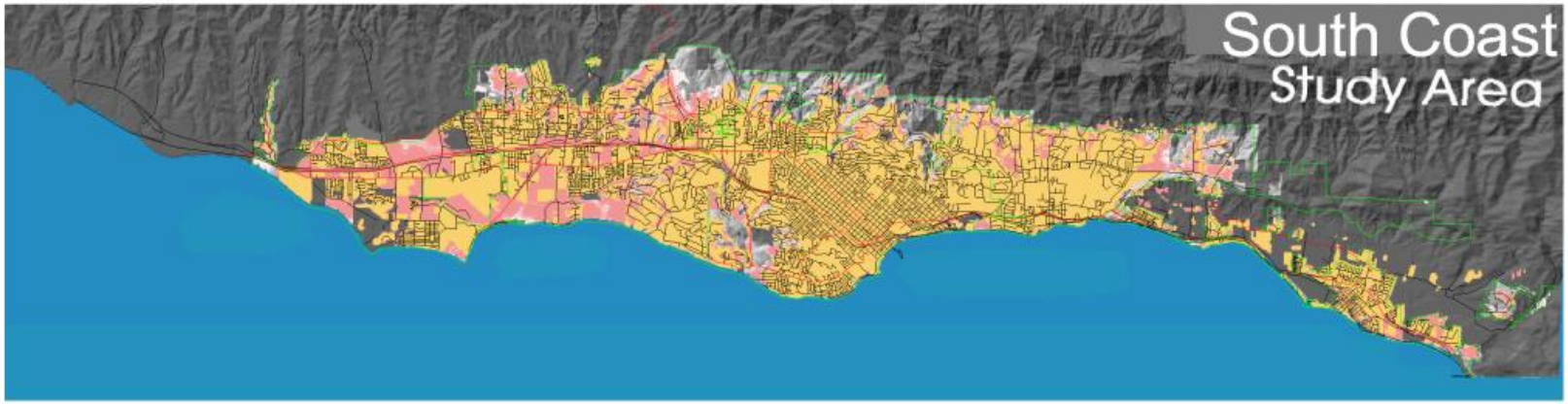
MOUNT TAMALPAIS
GAME REFUGE
East Peak
Middle Peak
King Mountain
Potrero Meadows
West Point
Double Bowknot
Blithedale Ridge
Rock Spring
MT TAMALPAIS STATE PARK
Redwood Creek
MUIR WOODS NATIONAL MONUMENT
MILL VALLEY
Homestead Valley
SAUSALITO
GOLDEN GATE NATIONAL RECREATION AREA


Desales St NW
16th St NW
15th St NW
14th St
L St NW
Constitution Ct NW
K St NW
Farragut Square
29
McPherson Square
Franklin Park
I St NW
Eye St NW
17th St NW
20th St NW
H St NW
Jackson Pl NW
Lafayette Park
New York Ave NW
G St NW
19th St NW
F St NW
White House
State Pl NW
E St NW
Pennsylvania Ave NW
Anacostia Annex
Washington
Northwest Rectangle
The Ellipse
C St NW
Ellipse Rd NW
Constitution Ave NW
200 m
1000 ft
50
1
Washington Monument
The
© 2004 Yahoo! Inc
©2003 NAVTEQ

Mean Center of Population for the United States: 1790 to 2000



Urban Growth Boundary Exclusion



Only Parks are Excluded

Uncertainty is endemic

- All geographic information leaves some degree of uncertainty about conditions in the real world
- **X**
  - the Greenwich Meridian
  - the Equator
  - standard time
- **Z**
  - definitions of terms
  - errors of measurement

Dear Waldo

View of the Outer Banks of North Carolina from Apollo 9

This photograph was taken on March 1[illegible] 1969 at [illegible]
10:00 a.m. EST, from an altitude of about 120 miles.

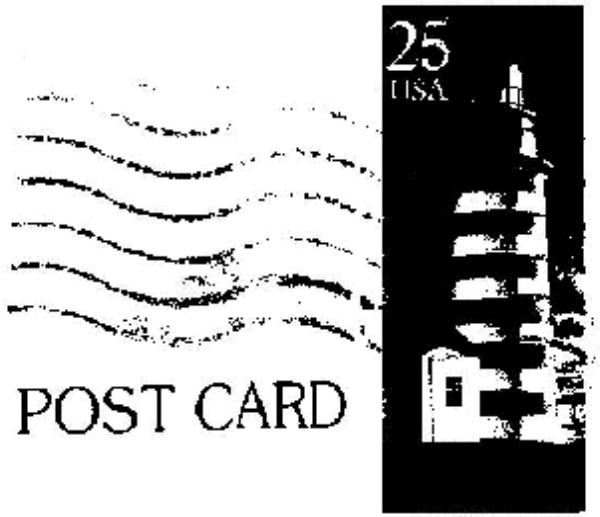
Posted at the old seafaring village of Hatteras, I know this card, with its complete and accurate address will get to you.

A pinhole shows you where we are—

Yours Cartographically

B[illegible]

Elizabeth City News Co. Elizabeth City, North Carolina



POST CARD

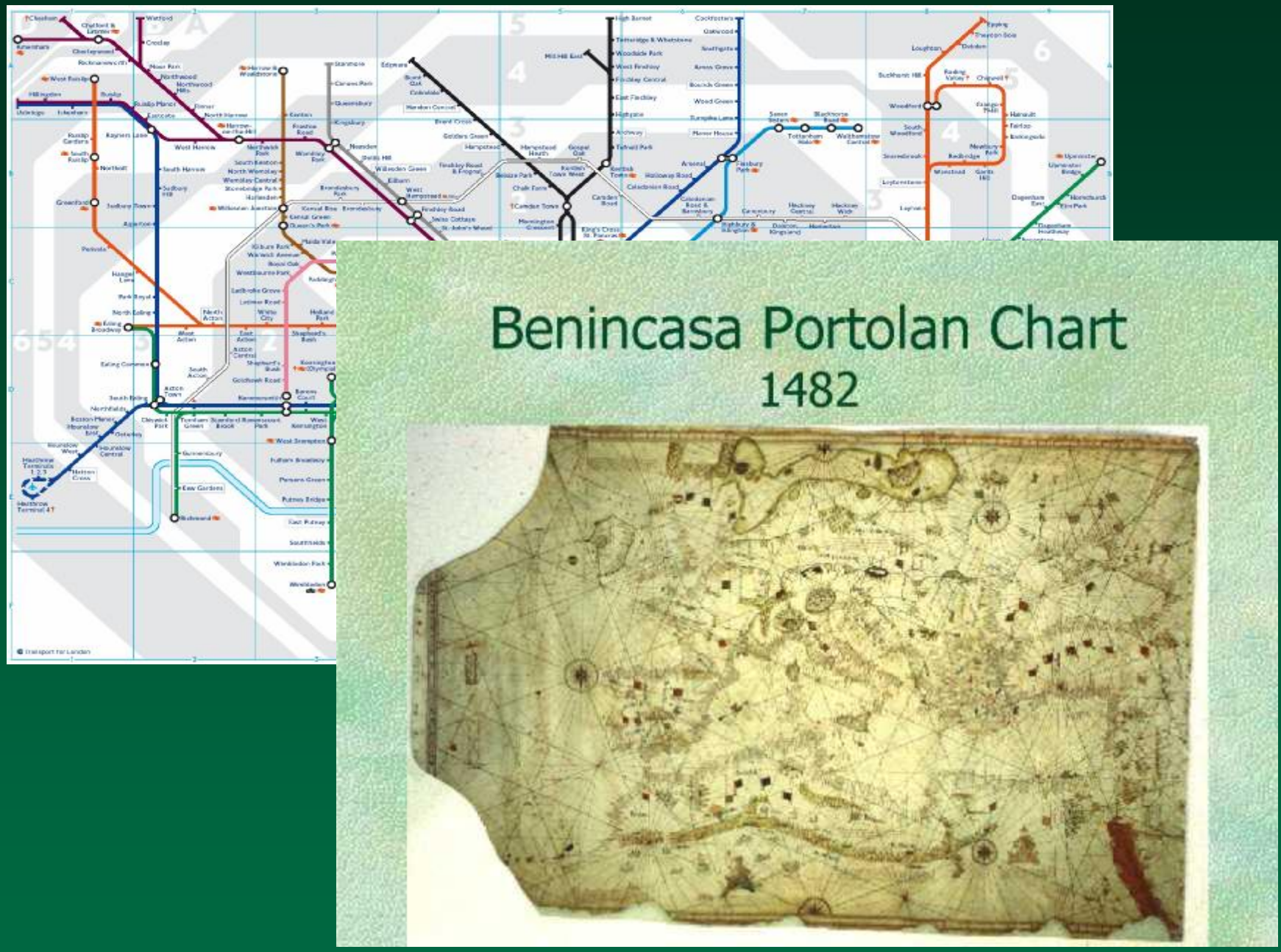
Professor Waldo Tobler

34° 26' 41" N

119° 48' 26" W

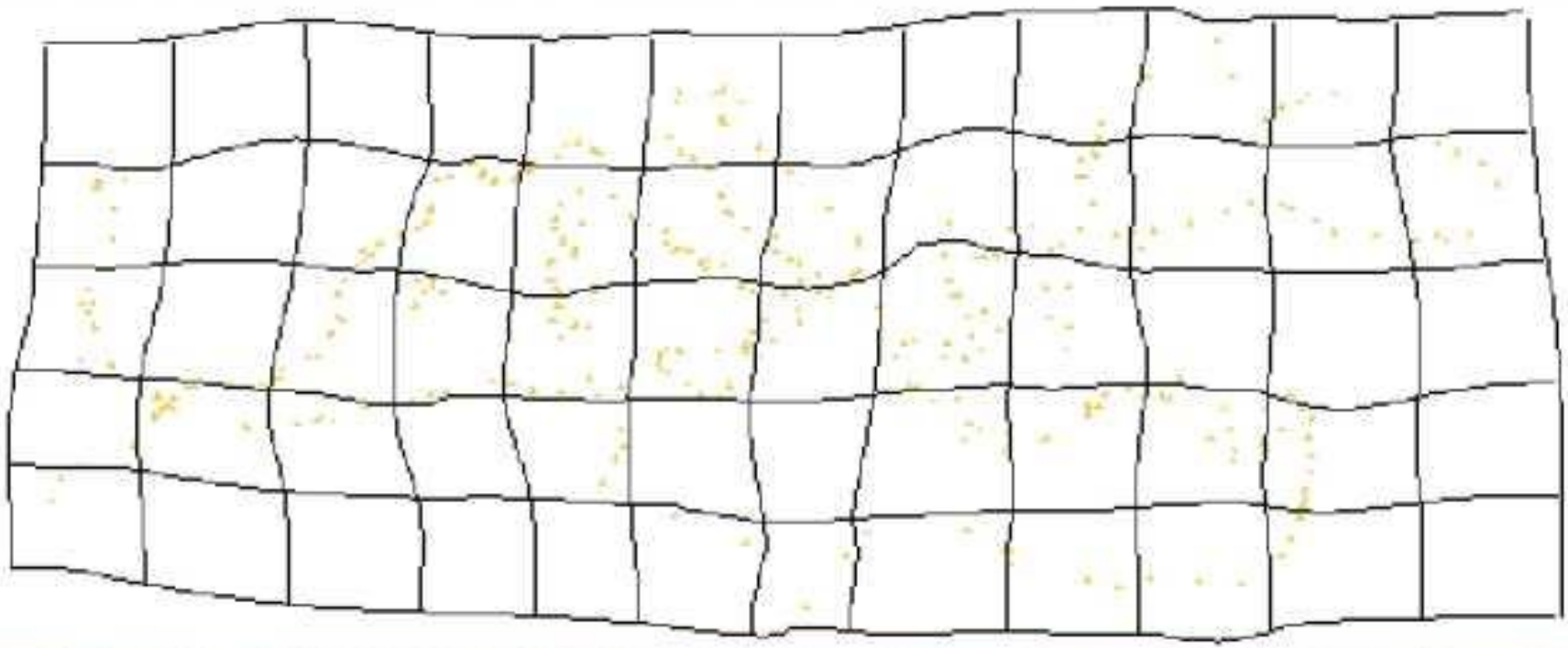
Plastichrome®

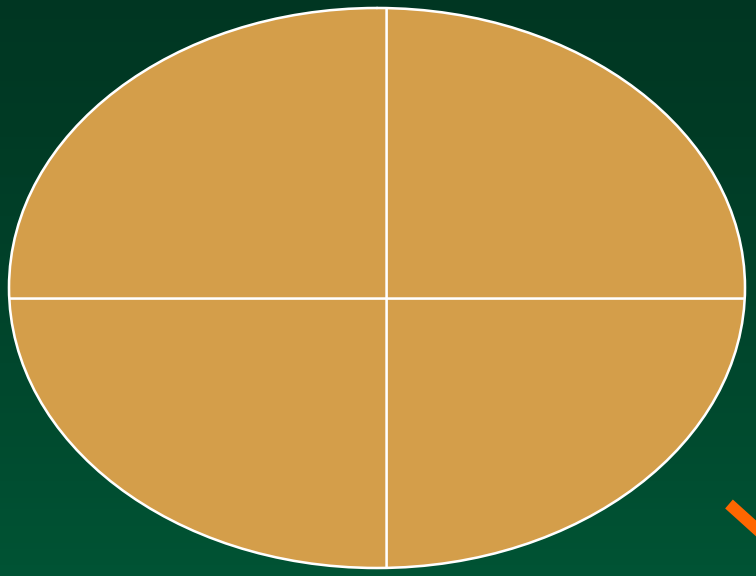
PRINTED IN IRELAND


Benincasa Portolan Chart
1482

Warped Grid of Portolan Chart

As 'pushed' by the interpolated vector field





Clarke Ellipsoid of 1866

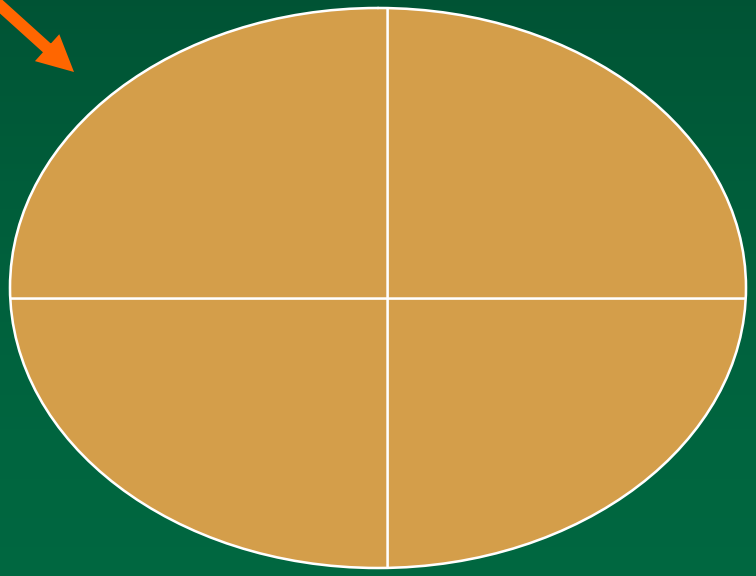
a = 6378206 m

1/*f* = 294.98

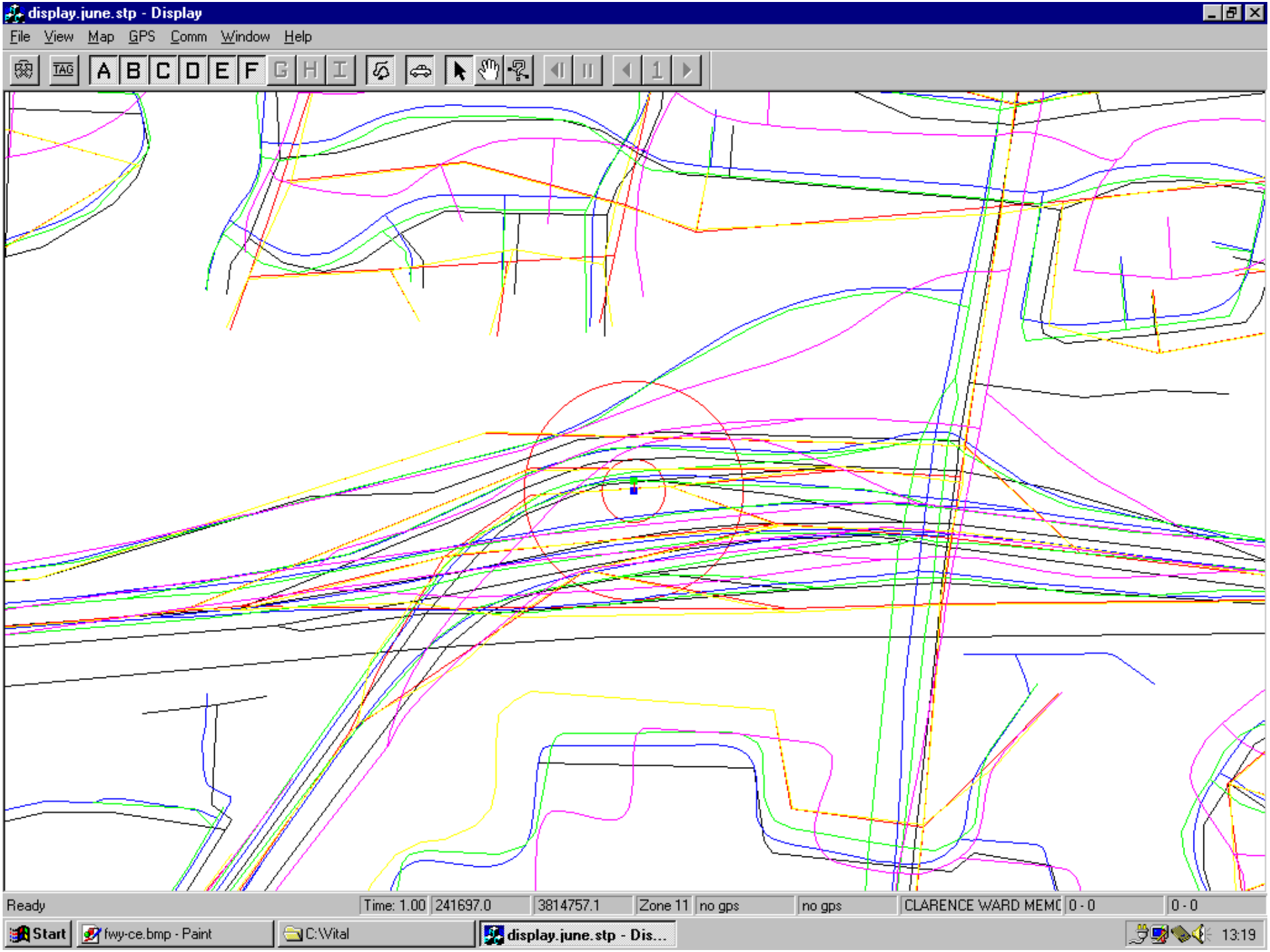
World Geodetic System of 1984

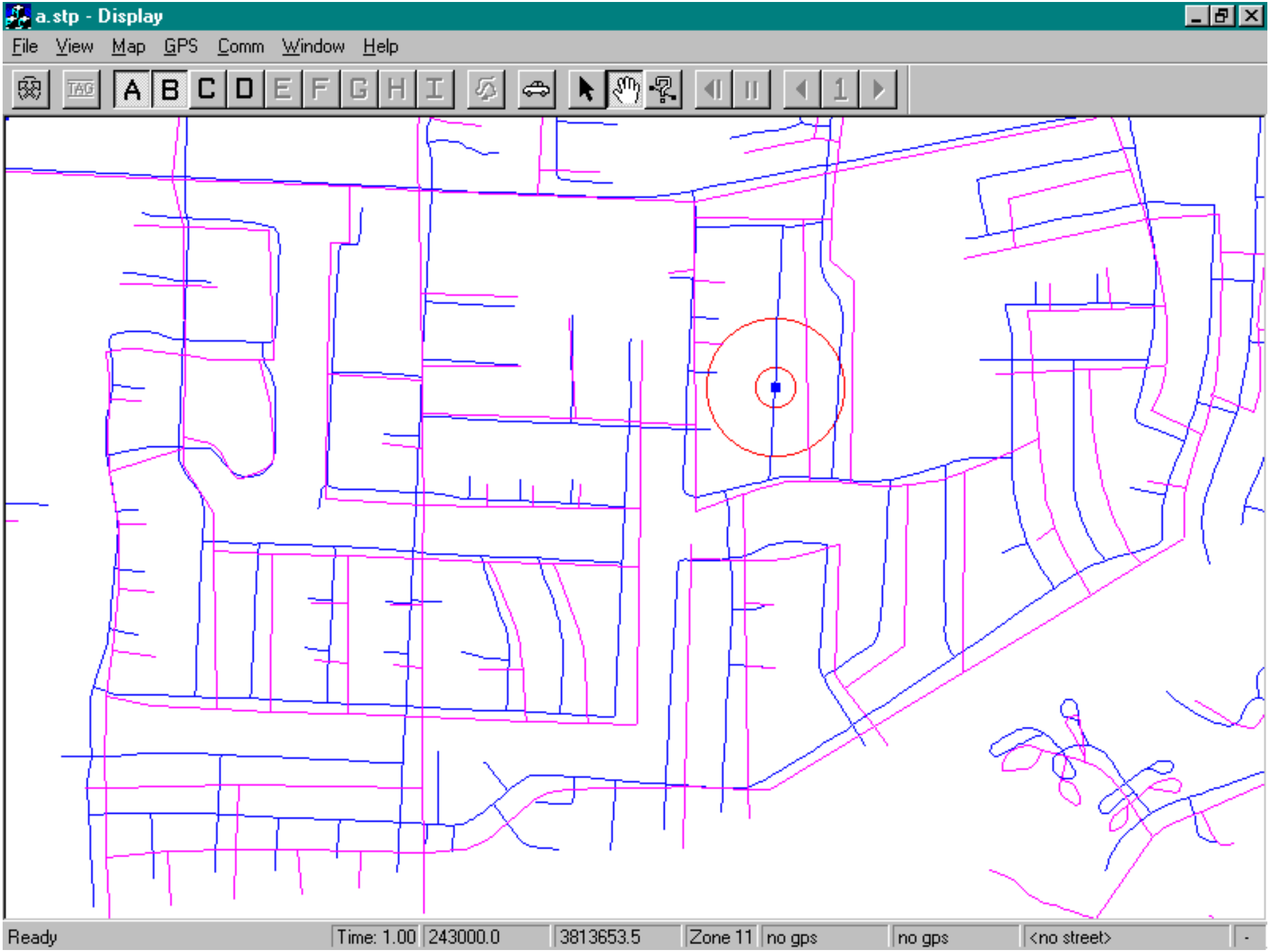
a = 6378137 m

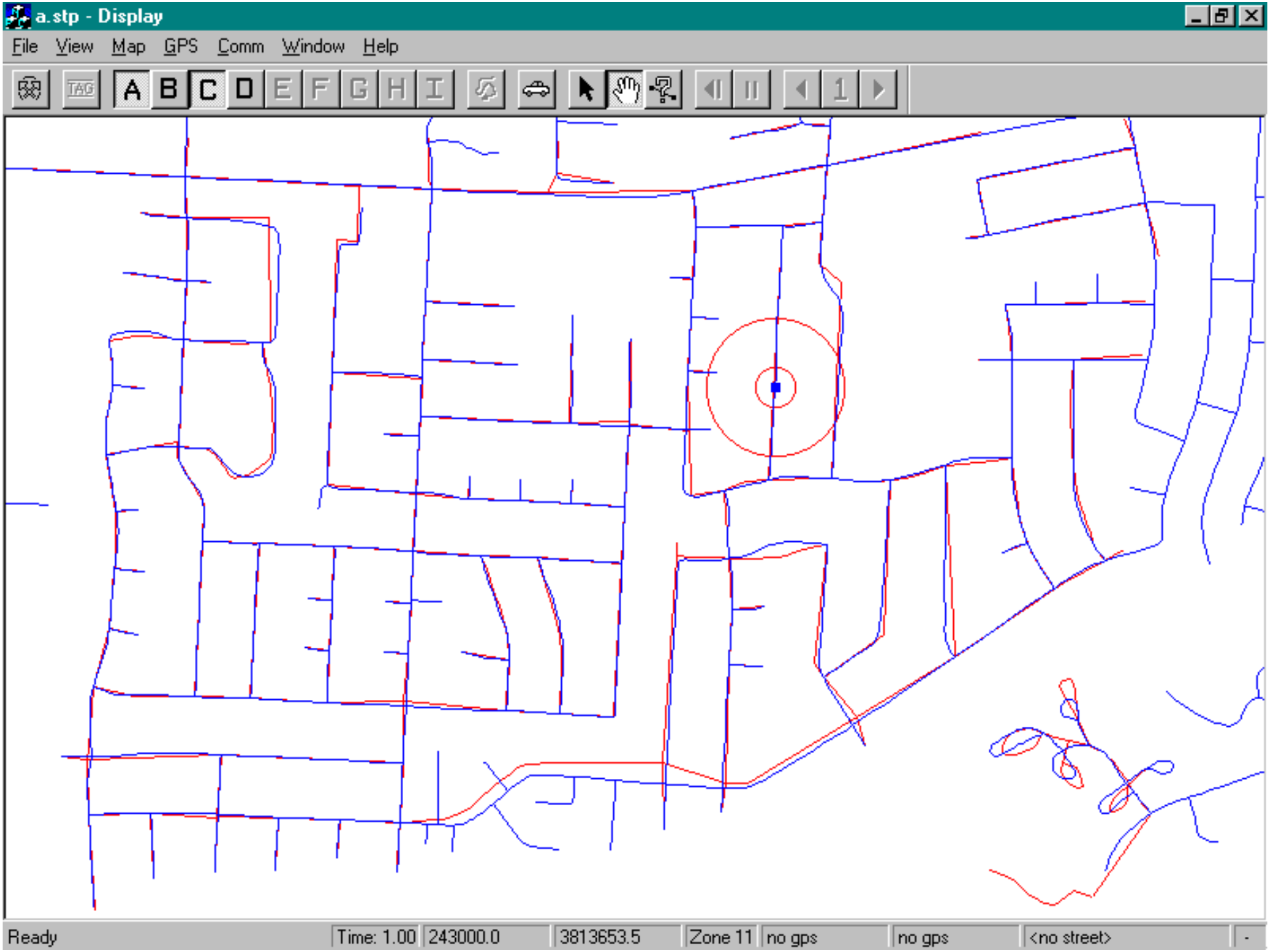
1/*f* = 298.26

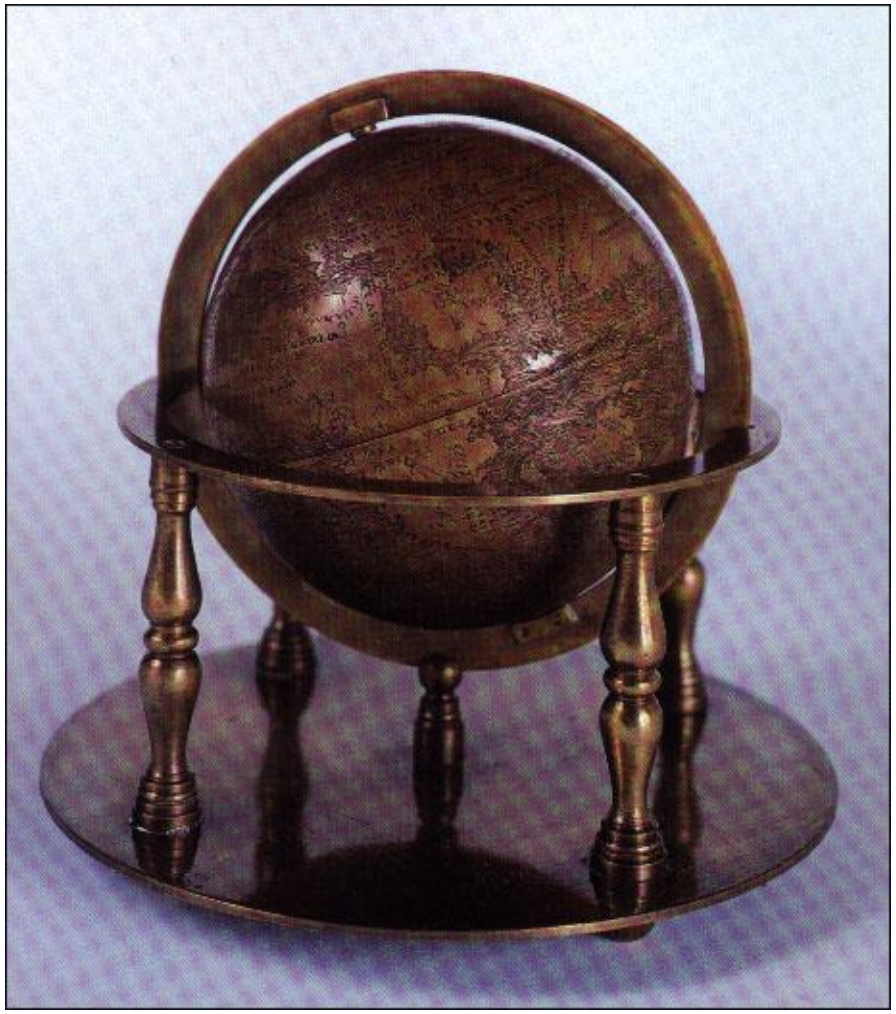


http://www.thesalmons.org/lynn/wh-greenwich.html


display.june.stp - Display
File View Map GPS Comm Window Help
A B C D E F G H I
TAG
1
Ready
Time: 1.00
241697.0
3814757.1
Zone 11
no gps
no gps
CLARENCE WARD MEMC
0 - 0
0 - 0
Start
fwy-ce.bmp - Paint
C:\Vital
display.june.stp - Dis...
13:19


a.stp - Display
File View Map GPS Comm Window Help
A B C D E F G H I
TAG
1
Ready
Time: 1.00
243000.0
3813653.5
Zone 11
no gps
no gps
<no street>


a.stp - Display
File View Map GPS Comm Window Help
A B C D E F G H I
TAG
1
Ready
Time: 1.00
243000.0
3813653.5
Zone 11
no gps
no gps
<no street>

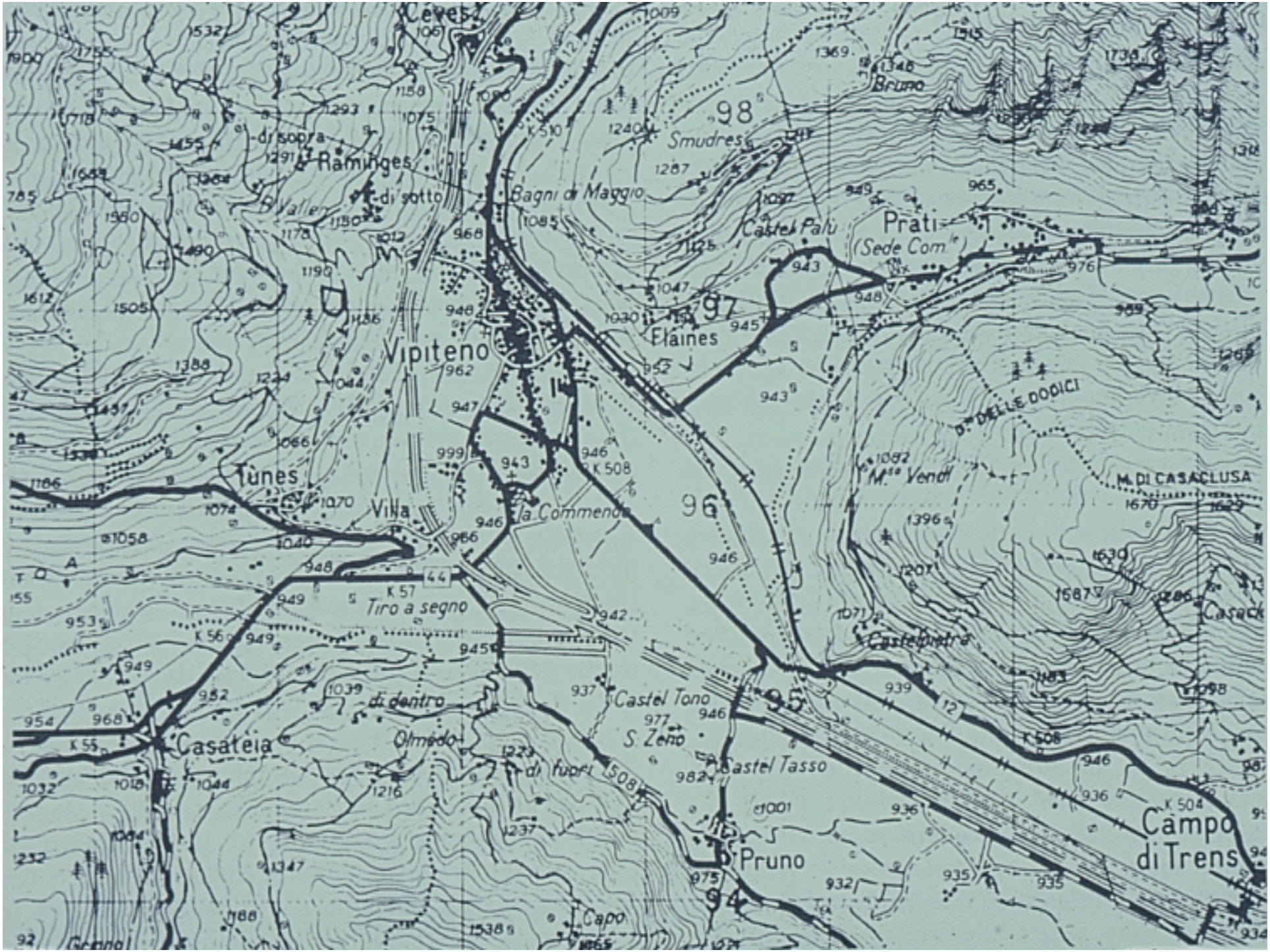


"hc sunt dracones"

The Lenox Globe of 1503-1507

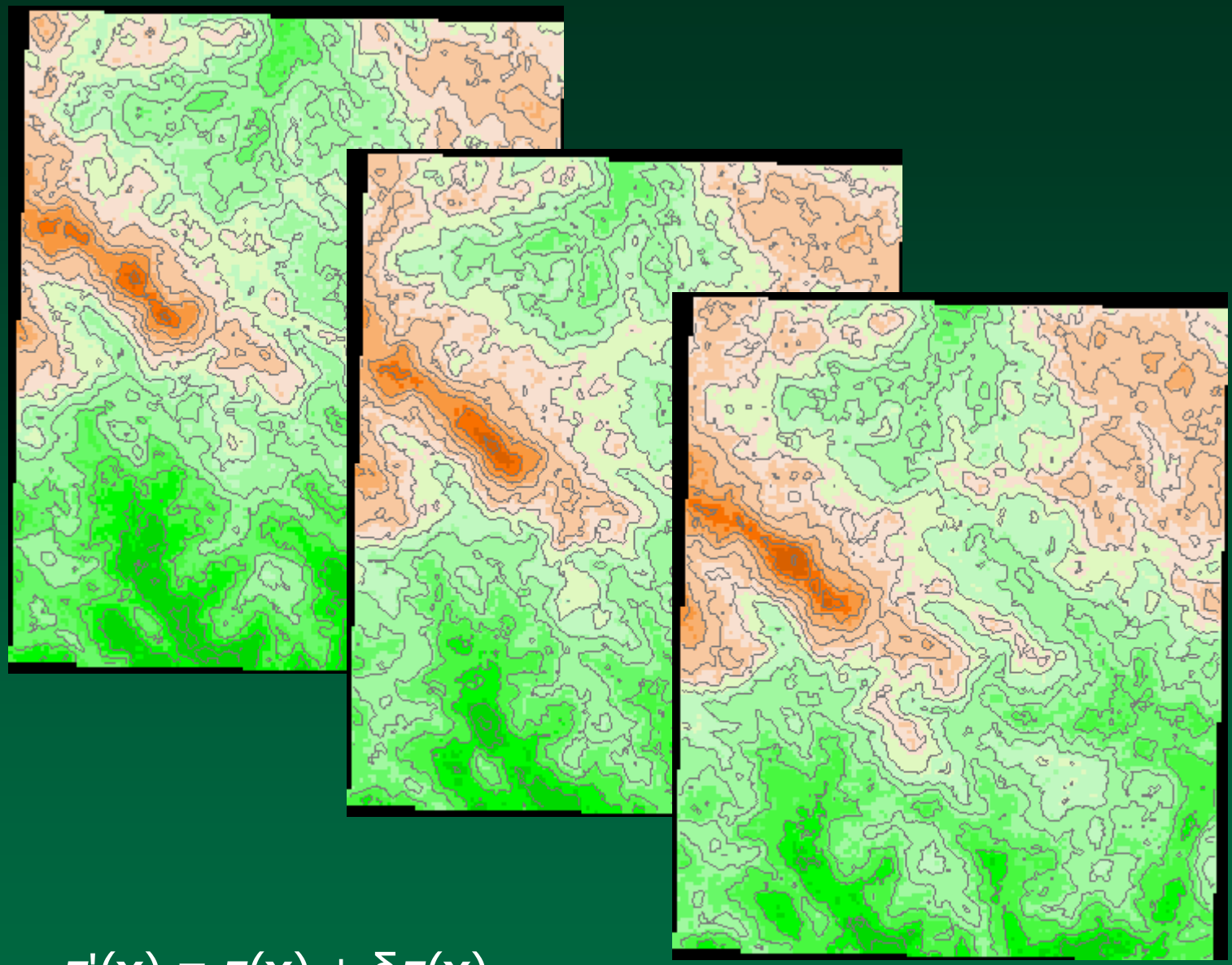
http://www.henry-davis.com/MAPS/Ren/Ren1/314.html


Tschöfs
Schnuders
Braunhof
Ramings
Flains
Moos
Wiesen
STERZING
E.St.N.
Pfoltnock
Thuins
Deutschhaus
Werfelhof
Sprechenstein
Elisabeth Q.
Reifenstein
Elsack
Elzenbaum
Rudolfs Qu.
Freienfeld


Ceves
Raminges
di sopra
di sotto
Bagni di Maggio
Smudres
Bruno
Castel Palù
Prati
(Sede Com.)
Elaines
Vipiteno
Tunes
Villa
la Commenda
M. DELLE DODICI
M.te Venodi
M. DI CASACLUSA
Tiro a segno
Castelpietra
Casateia
di dentro
Olmedo
di fuori
Castel Tono
S. Zeno
Castel Tasso
Pruno
Campo di Trens
Capo
12
44
95
96
97
98

The "five-fold way"

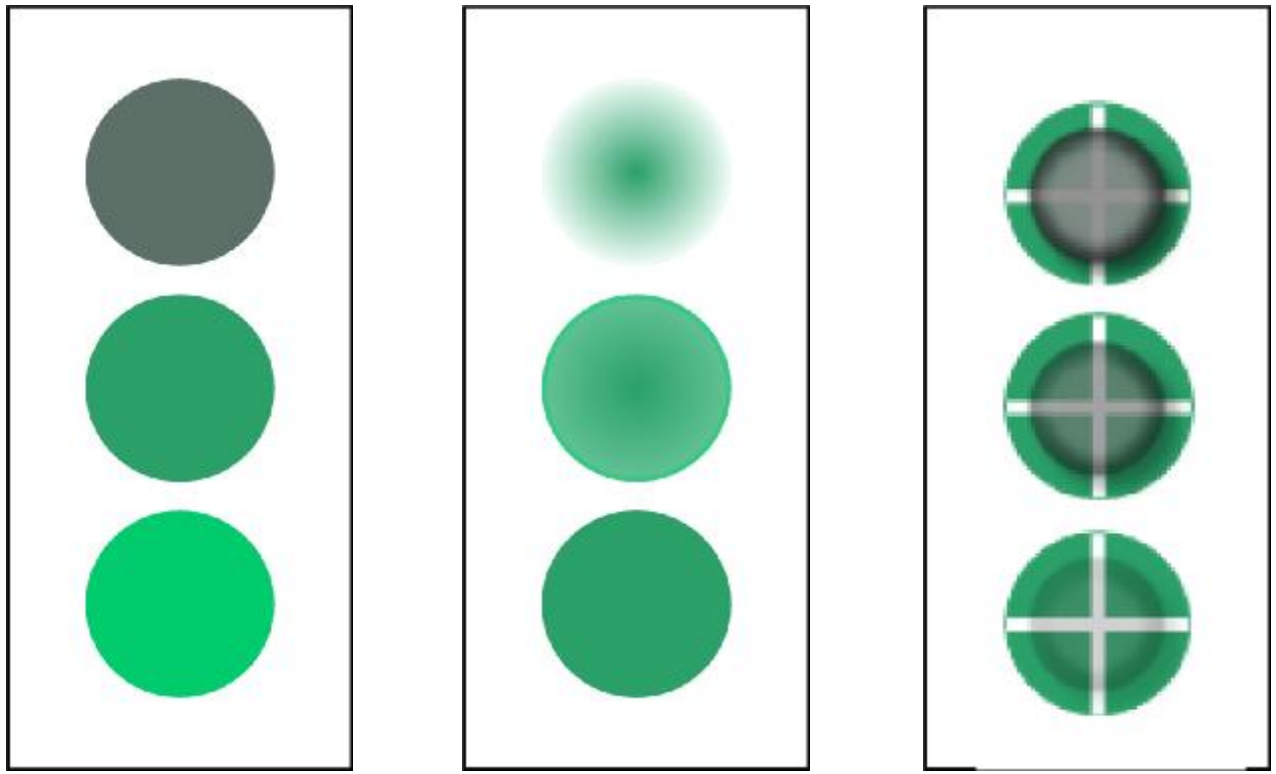
- Positional accuracy
- Attribute accuracy
- Logical consistency
- Completeness
- Lineage
- Federal Geographic Data Committee
  - Spatial Data Transfer Standard
  - Content Standard for Digital Geospatial Metadata
  - www.fgdc.gov



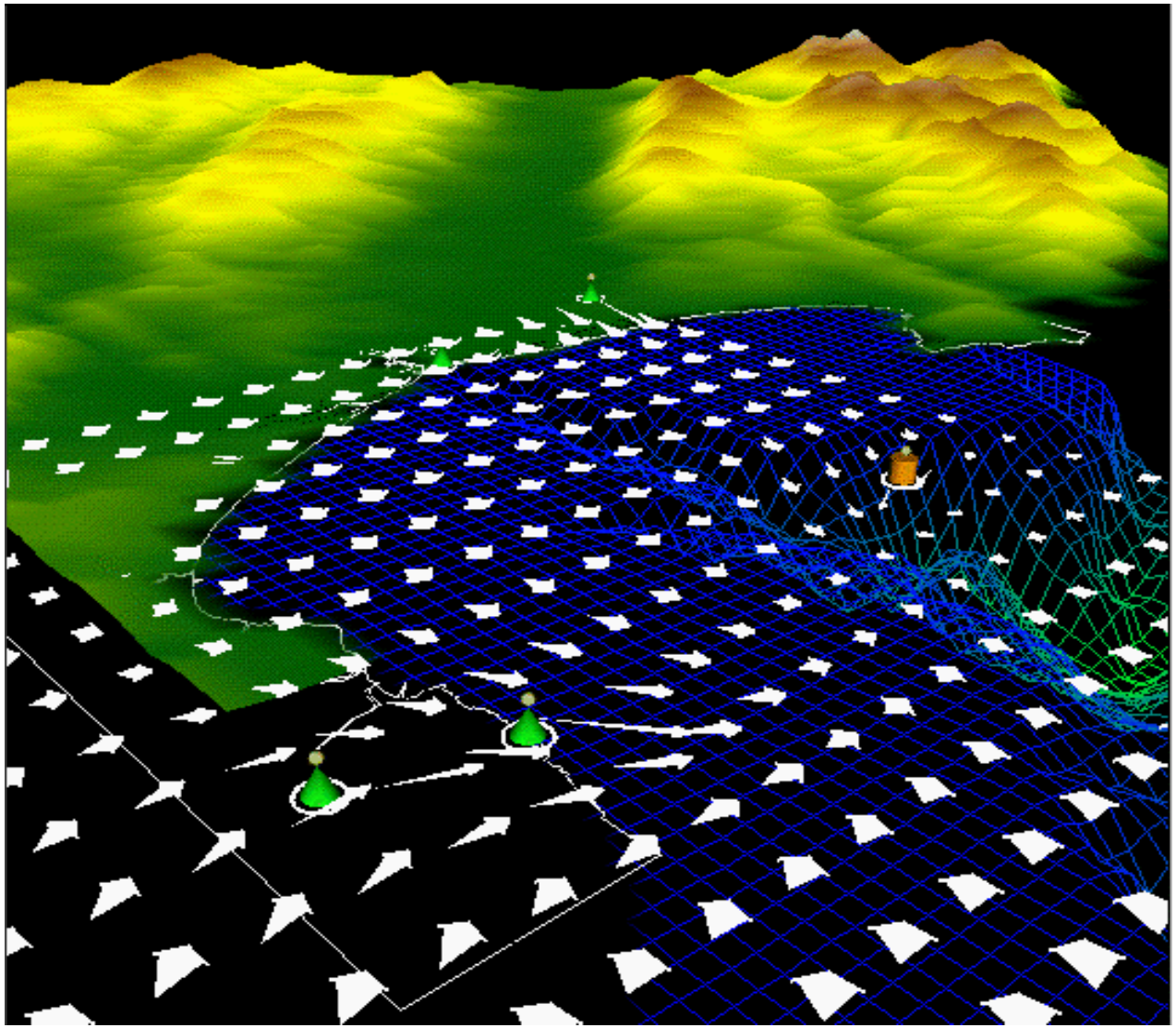
$$z'(\mathbf{x}) = z(\mathbf{x}) + \delta z(\mathbf{x})$$

Data Type / Data Quality	Positional Accuracy	Attribute Accuracy	Logical Consistency	Completeness	Lineage
Discrete Points and Lines	Size Shape (Error ellipses) (Epsilon bands)	Value Color Saturation (Feature code checks)	Color mixing Redundancy by overprinting Slivers by solid fills (Topological cleaning)	Mapping Technique Density traces Marginalia Generalization algorithm Mapping tolerance Buffer size	Mapping Technique Minimum Bounding Rectangles
Categorical Aggregation & Overlay (Tesselation, tiling, Areal coverages)	Texture Value (Certainty of boundary location)	Color mixing (Attribute code checks) (Topographic classifier)	lack error models	Mapping Technique Missing values Logical adjacency surface Marginalia Discrete model weights	
Partitioning & Enumeration (Metric class breaks)	not meaningful	Size = height (Blanket of error)	Size = height (Maximum likelihood prism maps)	Mapping Technique Missing values Missclassification matrix Classing scheme OAL/TAI	Marginalia Source of data Scale/Resolution Date Geometry
Continuous Interpolation (Surfaces and volumes)	no clear distinction b/w the two Value Color Saturation (Continuous tone vignettes) (Continuous tone isopleths)		Size = line wt Color Shape = compactness (TIN links)	not possible by definition Mapping Technique Surface of search attenuation Marginalia Interpolation algorithm	

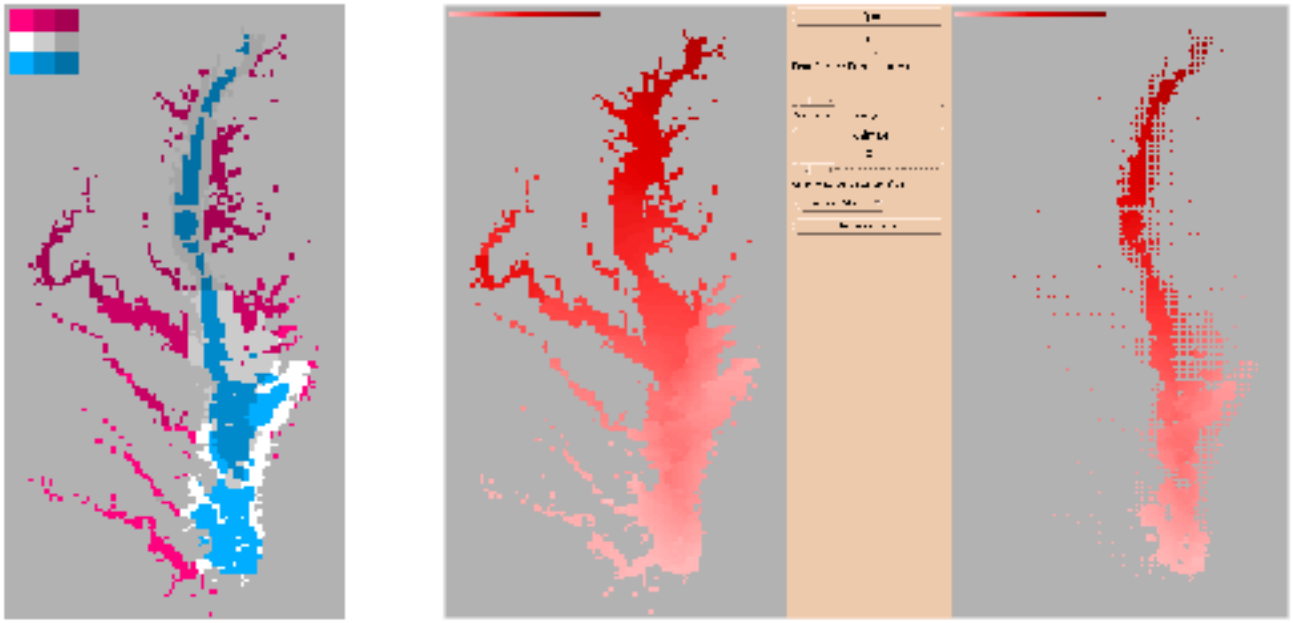
Graphical Syntax ← → **Graphical/Lexical Syntax** ← →



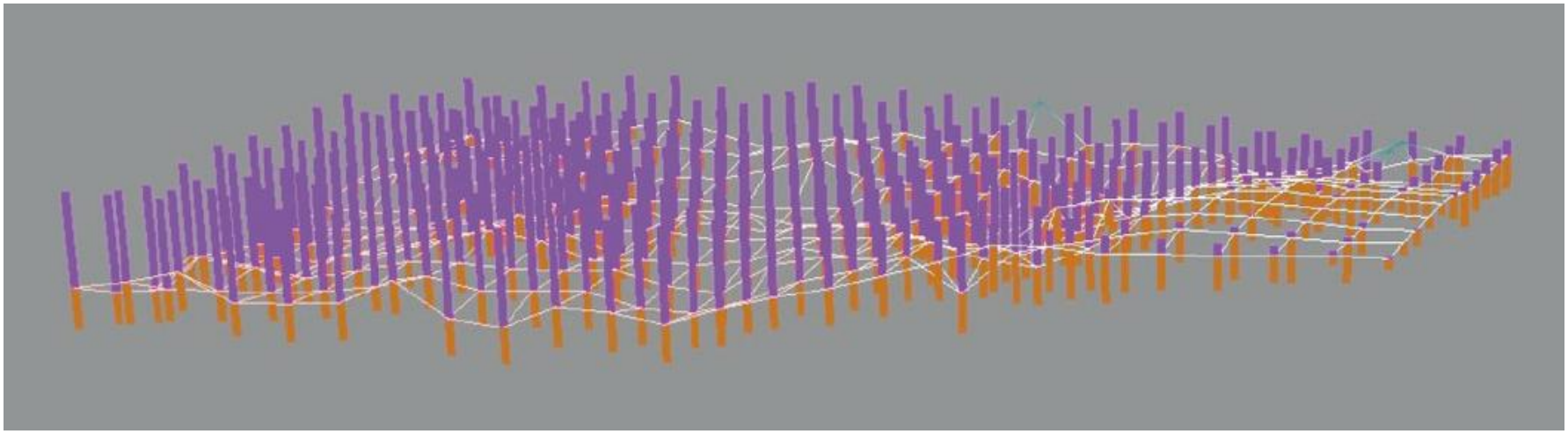
Point symbol sets depicting uncertainty with variation in (a) saturation (colors vary from saturated green, bottom, to unsaturated, top); (b) crispness of symbol edge – middle; and (c) transparency of symbol – right.



Glyphs indicating wind direction, magnitude, and uncertainty (Pang, 2001)



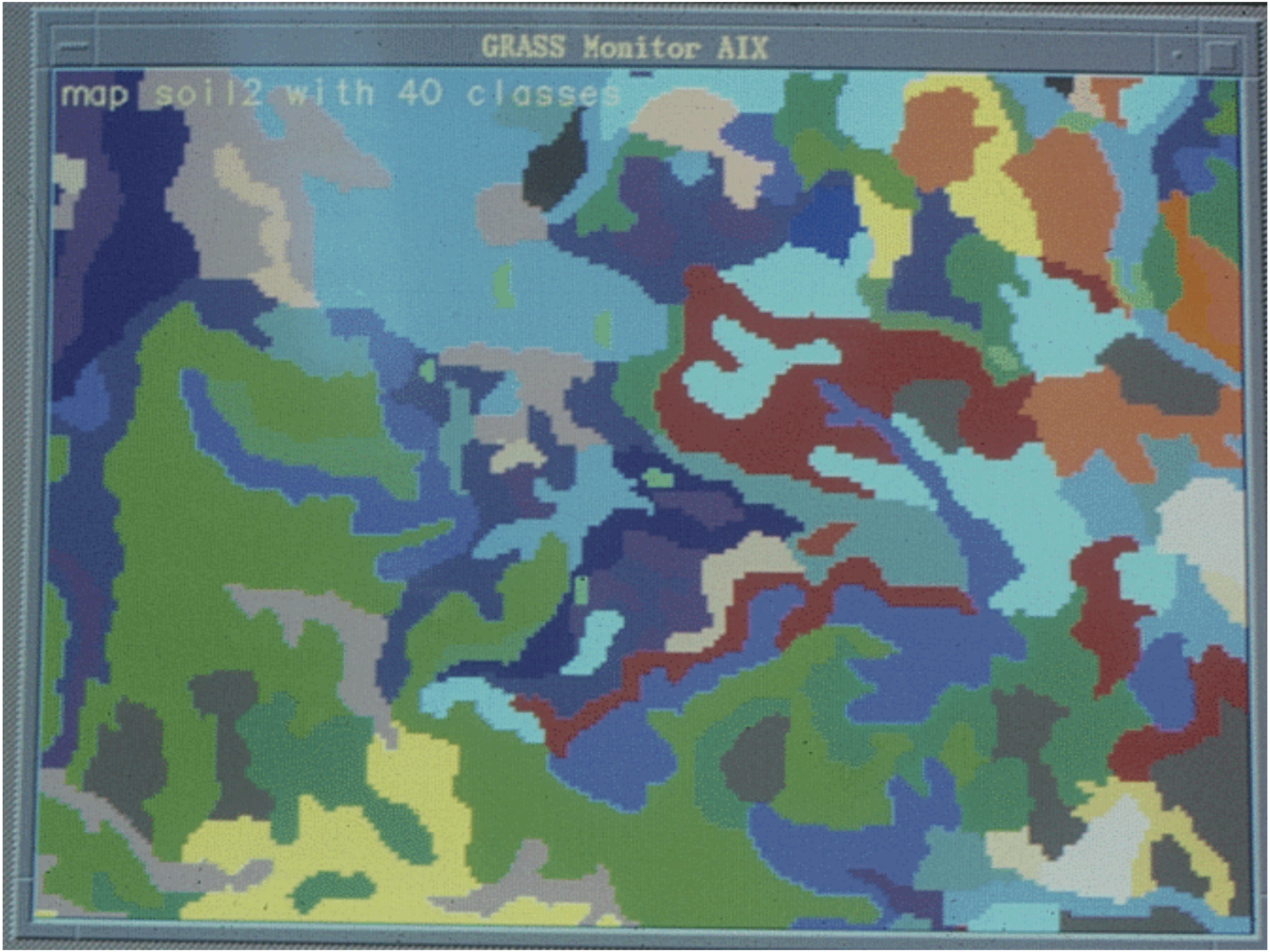
Alternative depictions of data (inorganic nitrogen in Chesapeake Bay) and uncertainty of data interpolated from sparse point samples. Left view shows bivariate depiction in which dark=more nitrogen and certainty is depicted with a diverging color scheme (blue = most certain and red = most uncertain). Right view depicts data in both panels (dark = more nitrogen), with the right panel showing the results of interactive focusing on the most certain data.

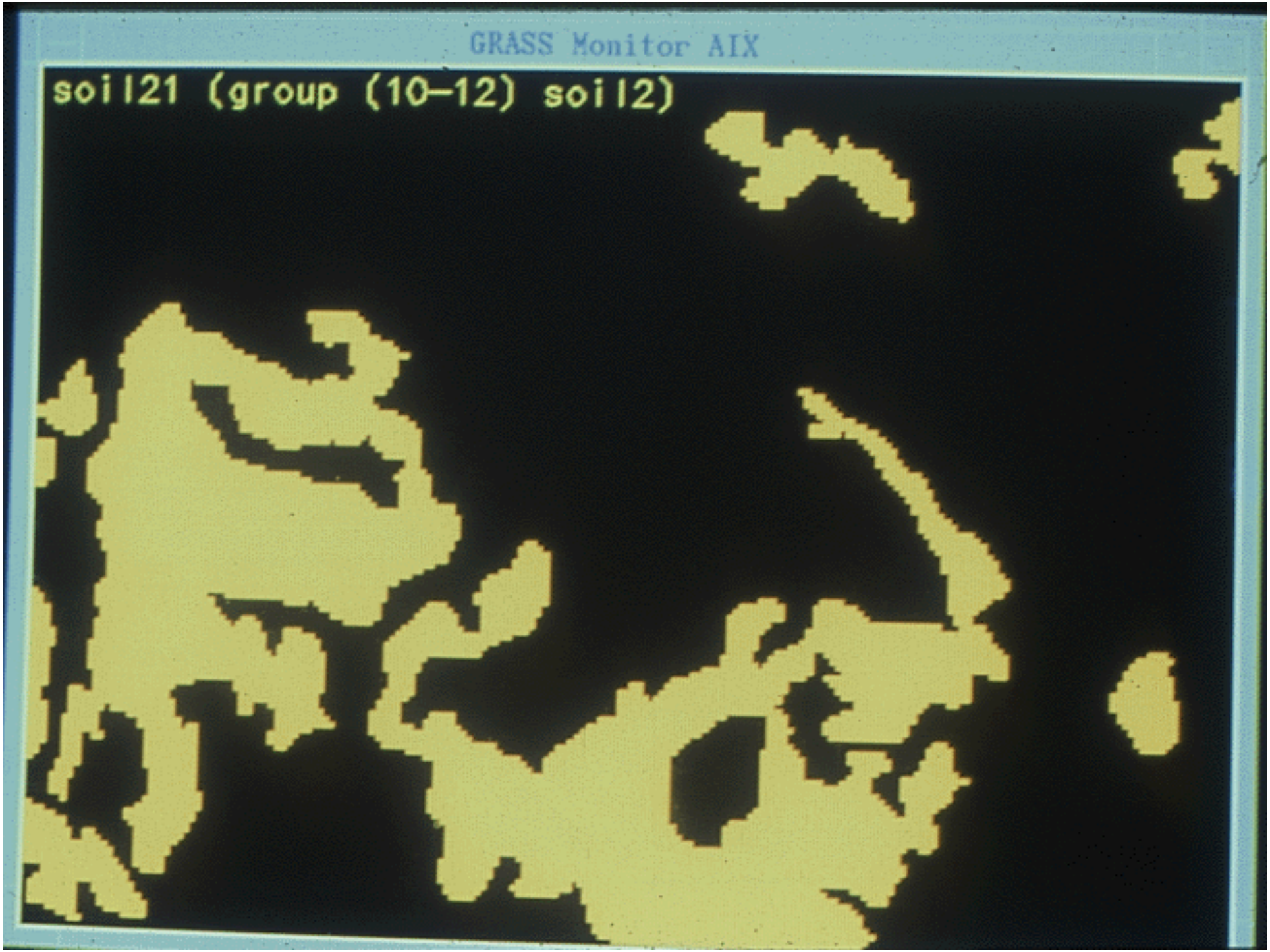


Representation of estimated water balance surplus/deficit (using a mesh surface) and uncertainty in the estimates (using bars above and below the surface). The bars depict the range of a set of model predictions with those predictions above the mean in purple and those below the mean in orange.

Fauerbach, Edsall, Barnes, MacEachren animation

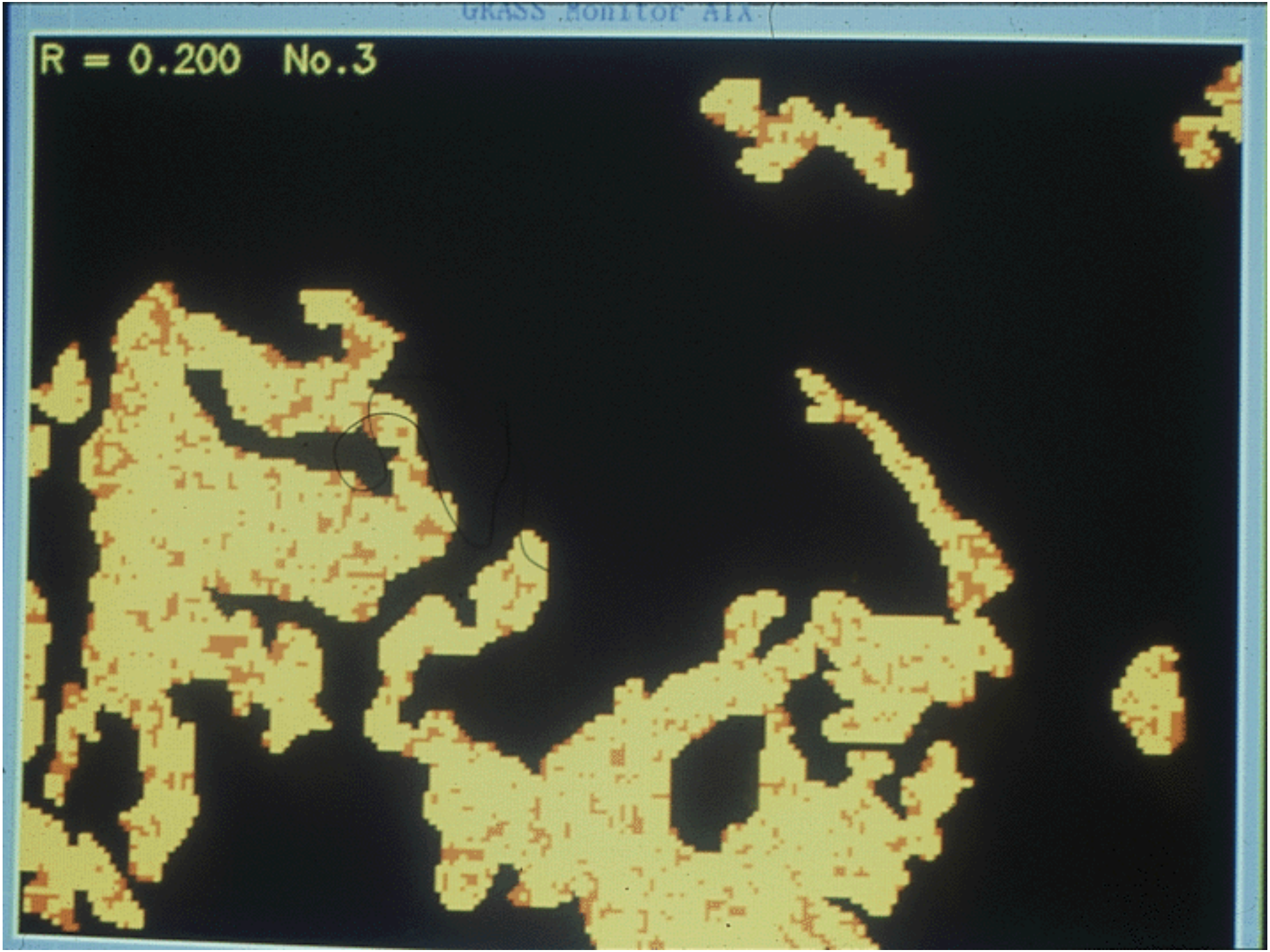
Ehlschlaeger simulation of Boston harbor

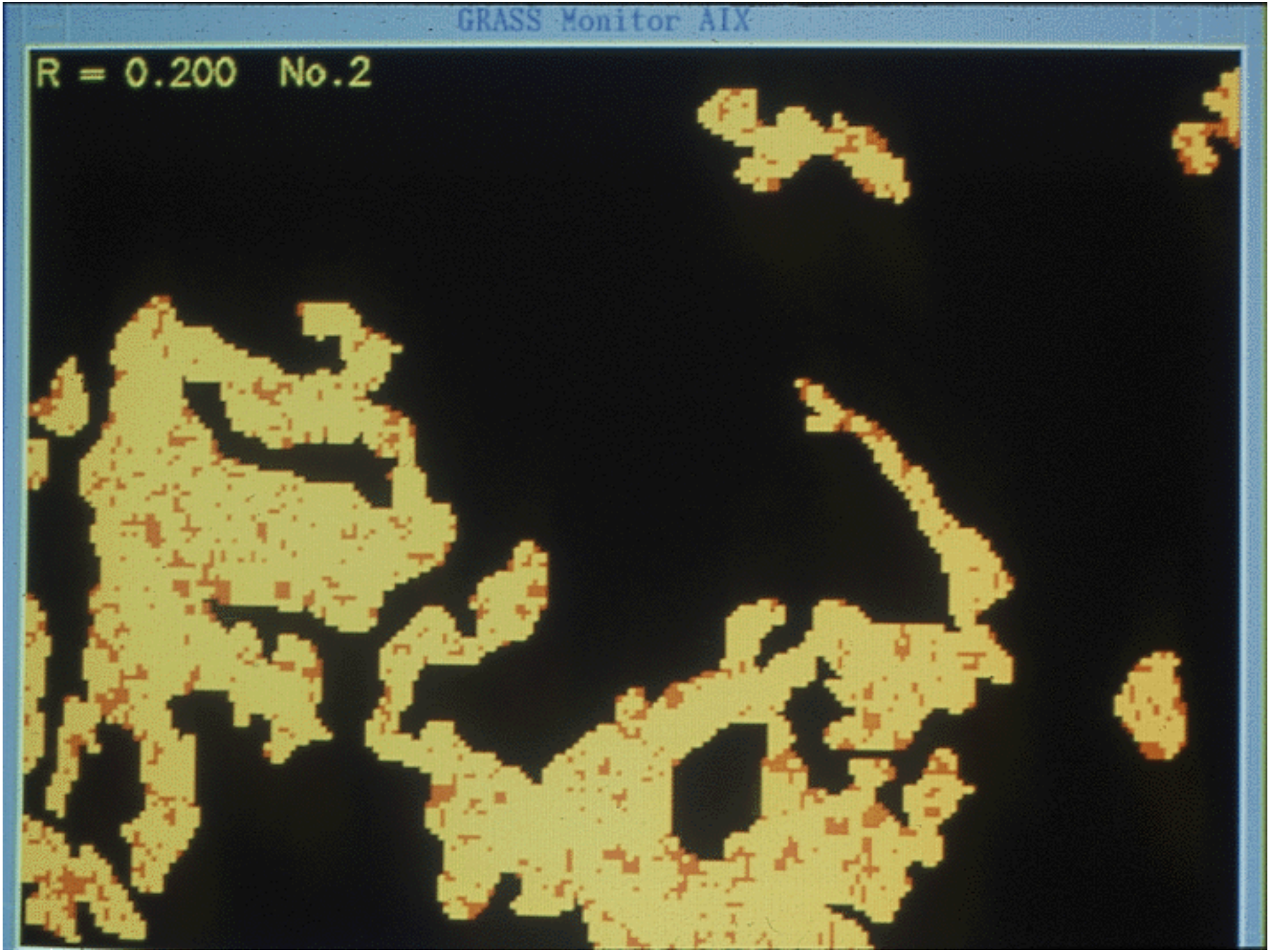

GRASS Monitor AIX
map soil2 with 40 classes

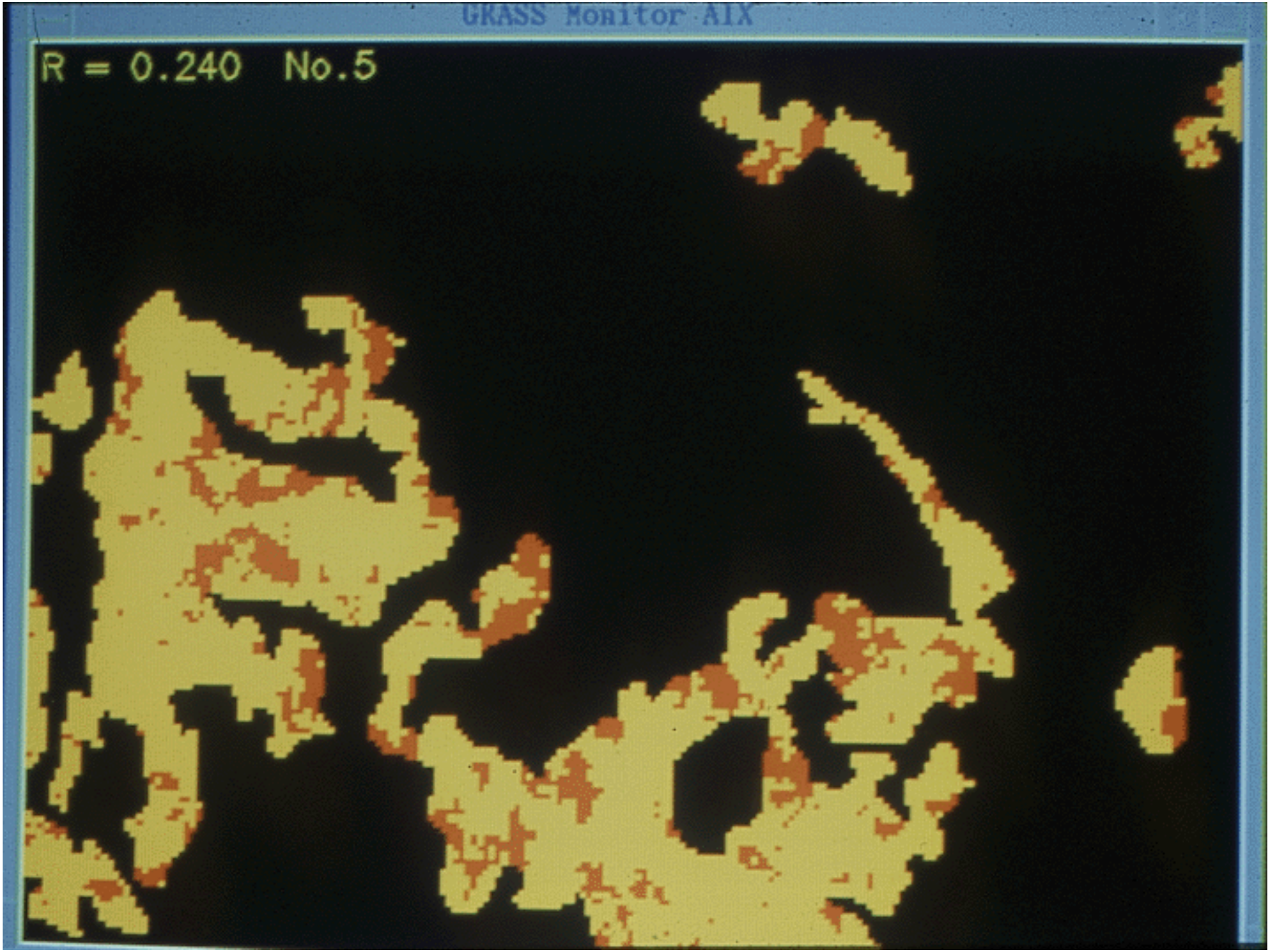

GRASS Monitor AIX
soil21 (group (10–12) soil2)

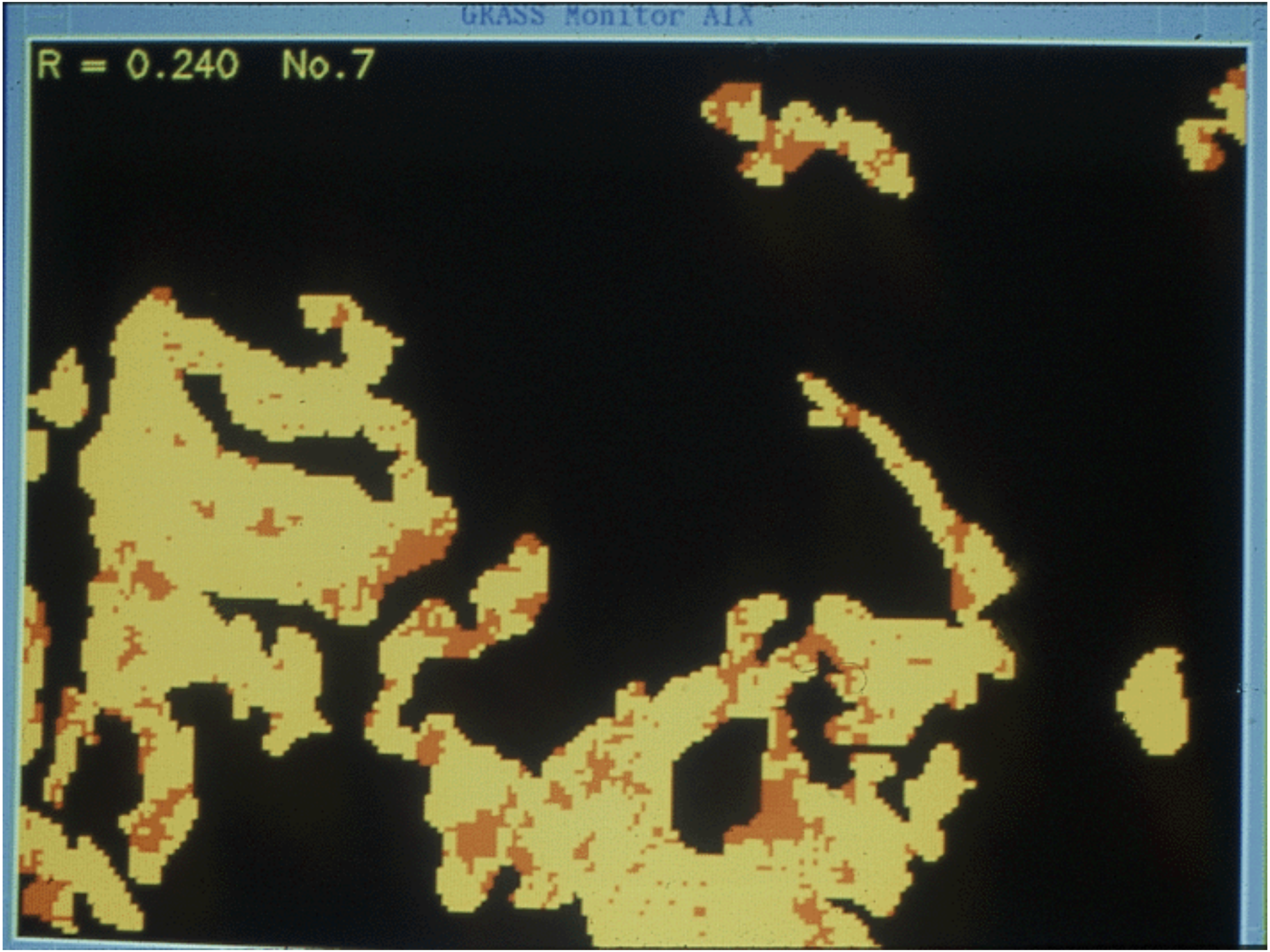
Model

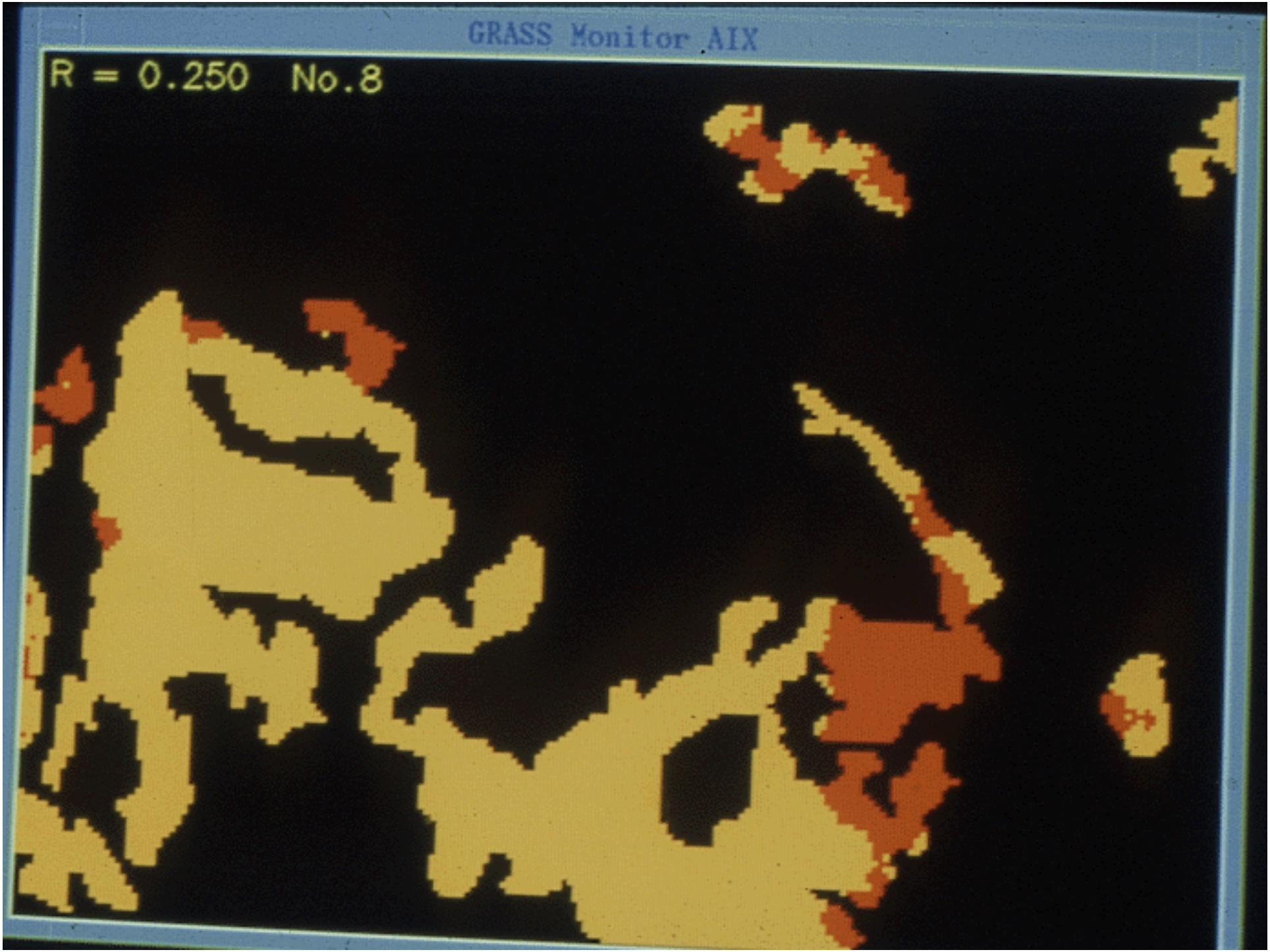
- $\{p_1, p_2, \ldots, p_n\}$
- correlation in neighboring cell outcomes
- posterior probabilities equal to priors
- 80% sand, 20% inclusions of clay
- no knowledge of correlations

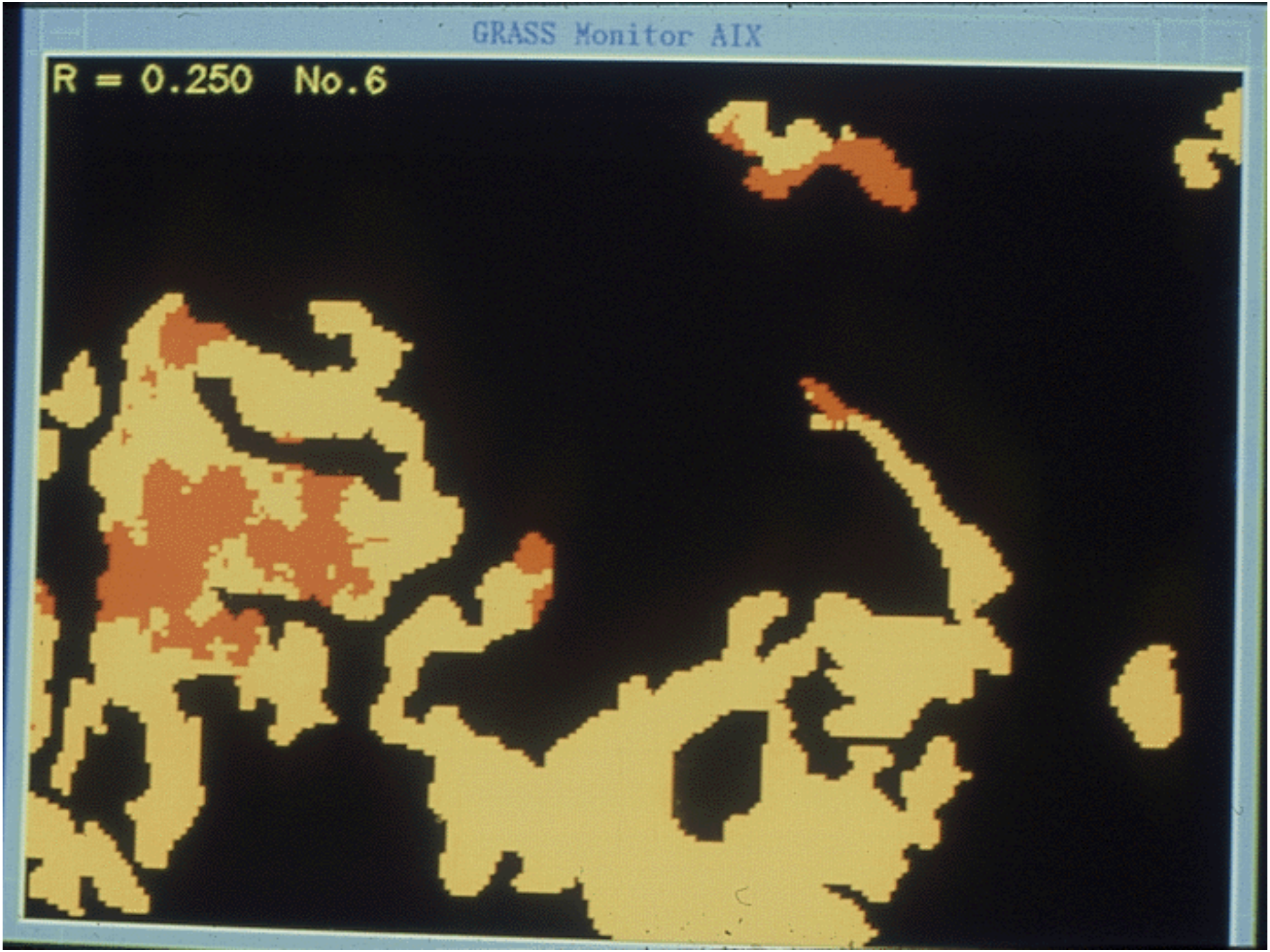

R = 0.200 No.3

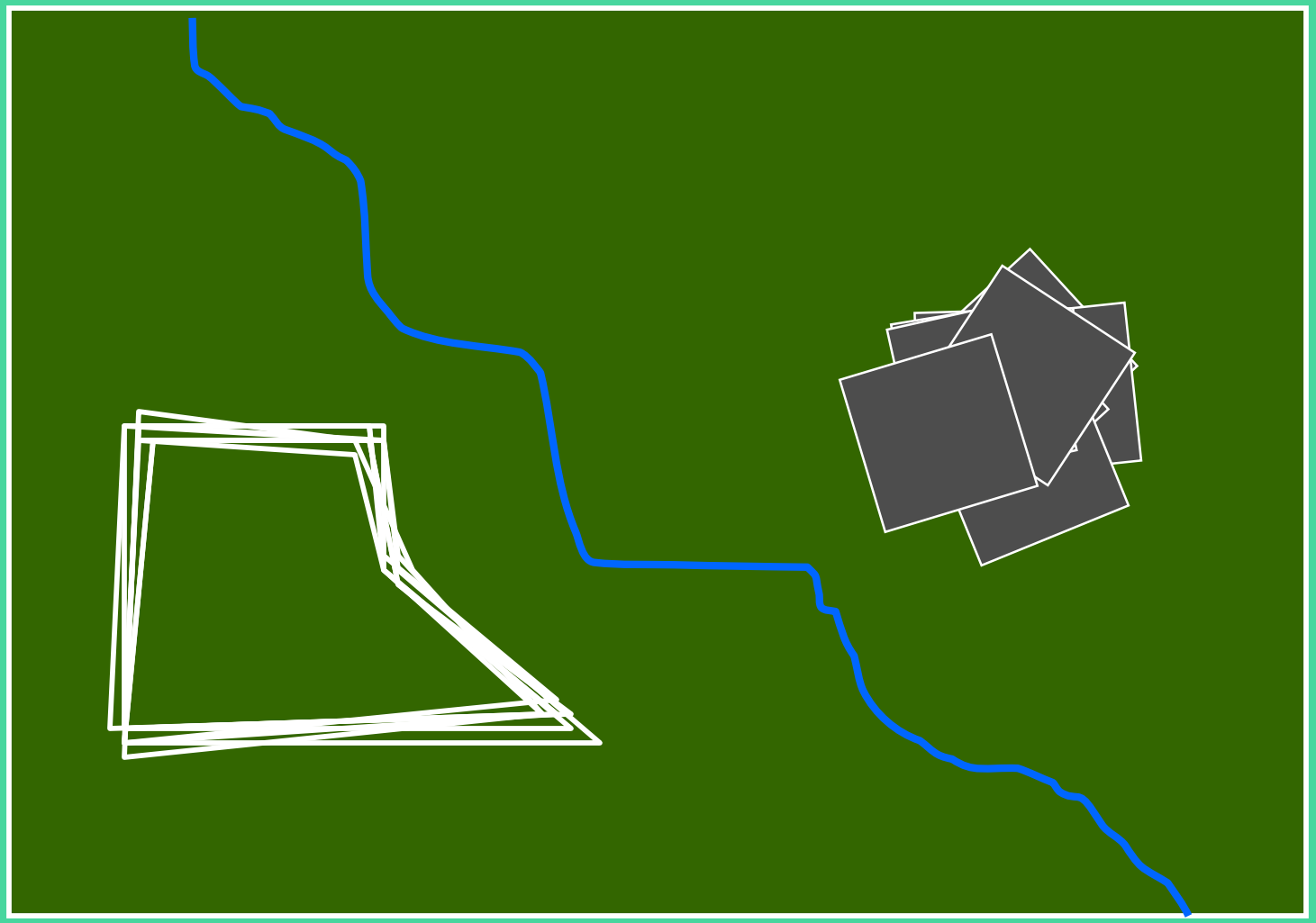

GRASS Monitor AIX
R = 0.200 No.2


GRASS Monitor AIX
R = 0.240 No.5


GRASS Monitor AIX
R = 0.240 No.7


GRASS Monitor AIX
R = 0.250 No.8


GRASS Monitor AIX
R = 0.250 No.6

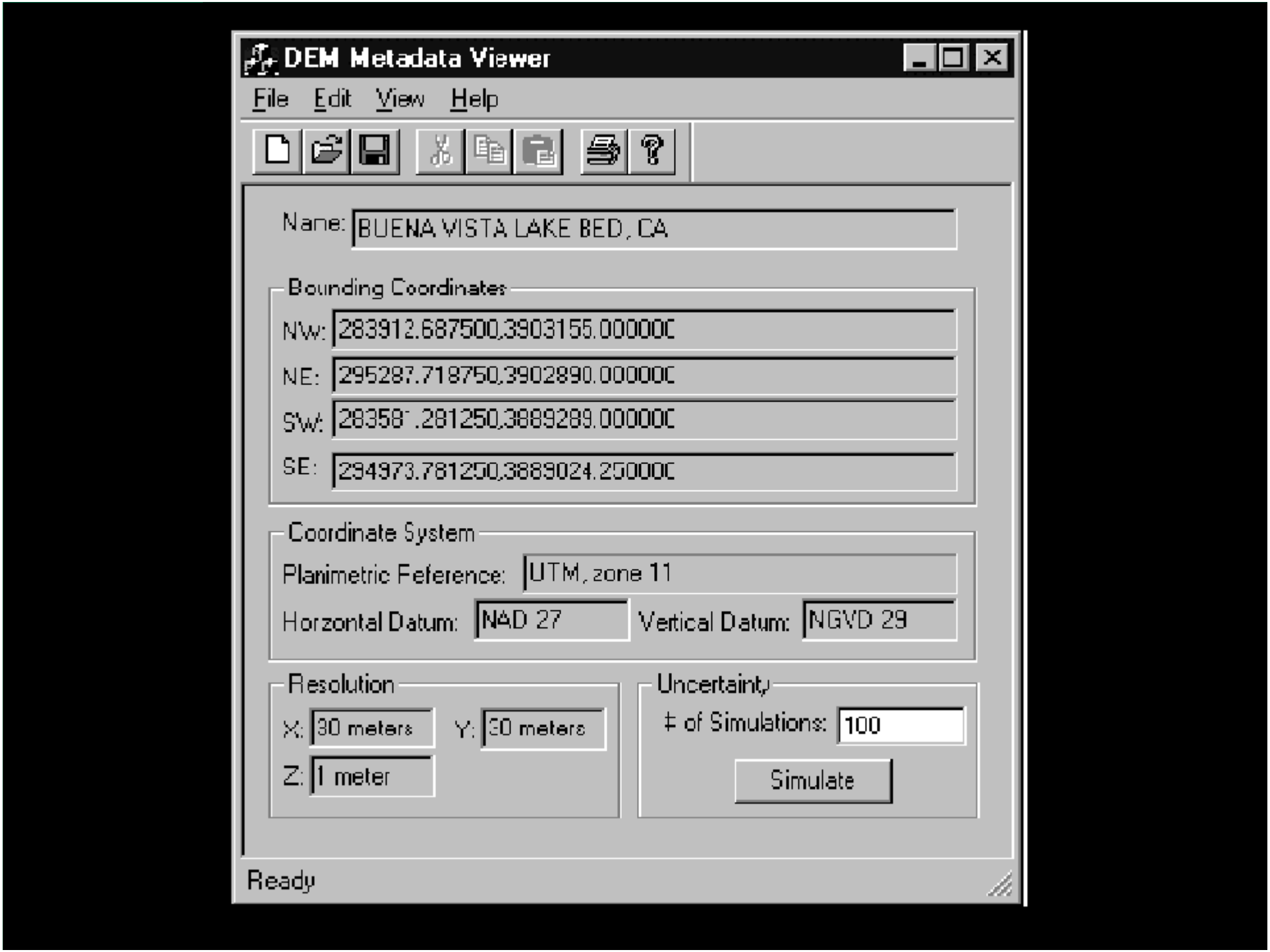


Options for uncertainty

- Ignore
- Present parameters
- Present simulation
  - confidence intervals

Communication of uncertainty

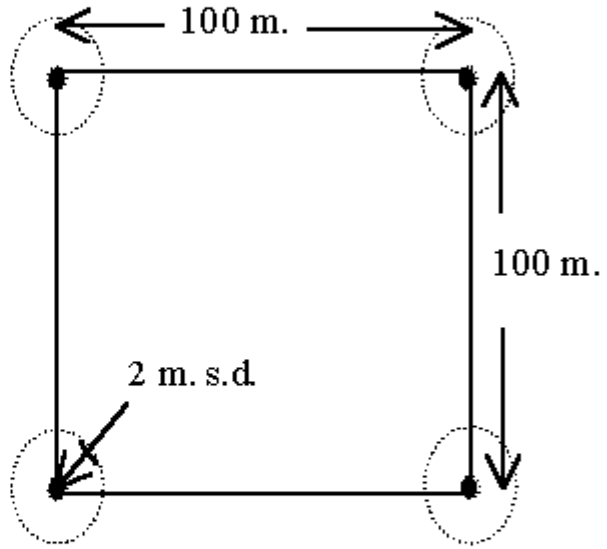
- Producer to user
  - abilities
- Metadata standards
  - parameters of complex models
- Assertion:
  - all knowledge of uncertainty can be expressed in a suitable simulation model
  - equally likely realizations


DEM Metadata Viewer
File Edit View Help
Name: BUENA VISTA LAKE BED, CA
Bounding Coordinates
NW: 283912.687500,3903155.000000
NE: 295287.718750,3902890.000000
SW: 28358 .281250,3889289.000000
SE: 294973.781250,3889024.250000
Coordinate System
Planimetric Reference: UTM, zone 11
Horzontal Datum: NAD 27
Vertical Datum: NGVD 29
Resolution
X: 30 meters
Y: 30 meters
Z: 1 meter
Uncertainty
# of Simulations: 100
Simulate
Ready

Online Java demonstration

http://www.ncgia.ucsb.edu/~ashton/demos/propagate.html

Quadrilateral parcel defined by four surveyed points. The data consist of 4 (x,y) locations, and indicate they form a square 100m on a side, with an area of 10,000 square meters.



Uncertainty in point location is characterized by a Gaussian distribution, mean of zero and s.d. of 2 meters.

Question: What is the uncertainty associated with the area of the land parcel, given the positional uncertainty information?

Summary points

- Uncertainty endemic
  - counter to cartographic tradition since 1900
- Numerous forms
  - attributes versus position
  - discrete objects versus continuous fields
  - absolute versus relative
- Strong positive spatial autocorrelation
  - marginal versus joint distributions
- Rich variety of potential methods
  - value of animation and simulation