Univariate Data Presentation: *The Contouring Conundrum and Philosophical Arguments Regarding the Contouring of Geochemical Data*

Clifford R. Stanley

Dept. of Geology, Acadia University, Wolfville, Nova Scotia, B4P 2R6, Canada, cliff.stanley@acadiau.ca

Outline

A) ***To Contour or Not to Contour***

1) regionalization
2) theory justification
3) empirical justification
    a) semivariograms
    b) bubbleplots
4) logical criteria and decision tree
5) contouring pitfalls

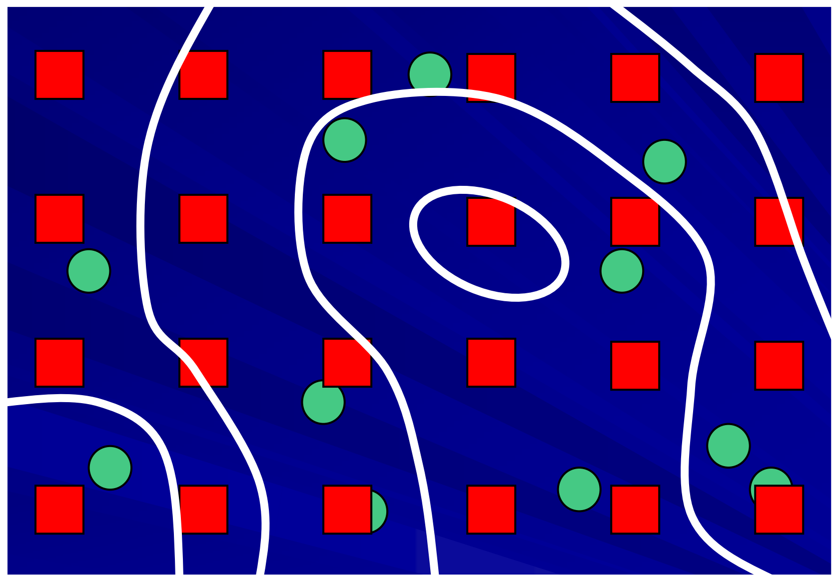
B) ***Bubbleplots***

1) accurate representation
2) aesthetic representation

The Problem With Contours:

Modern computer programs provide us with a myriad of ways (algorithms) to contour geochemical data:

1) *nearest neighbor*
2) *local mean*
3) *inverse distance*
4) *kriging*

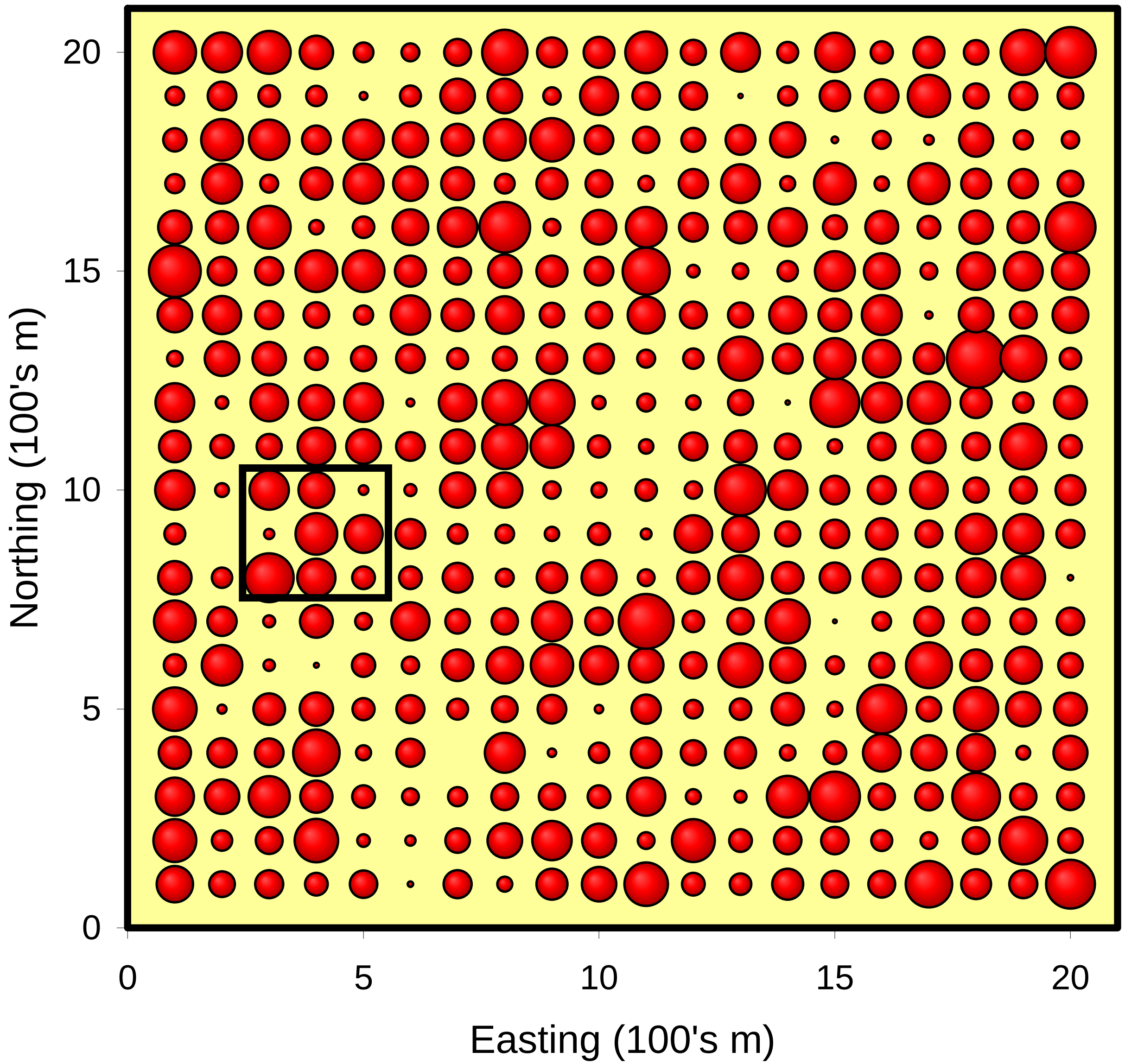


But …

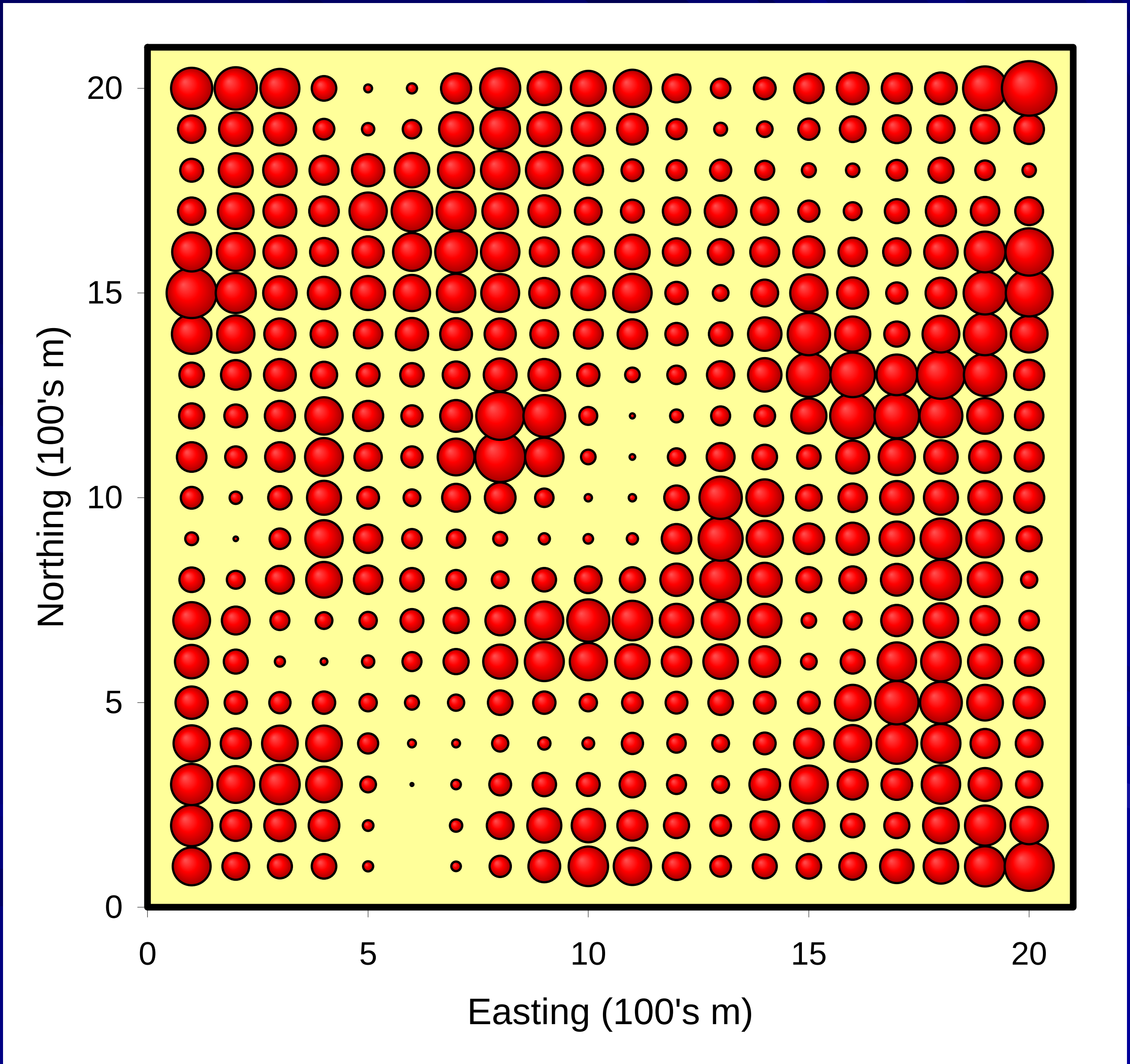
just because we can draw contours,

doesn’t mean we should!

Raw Data

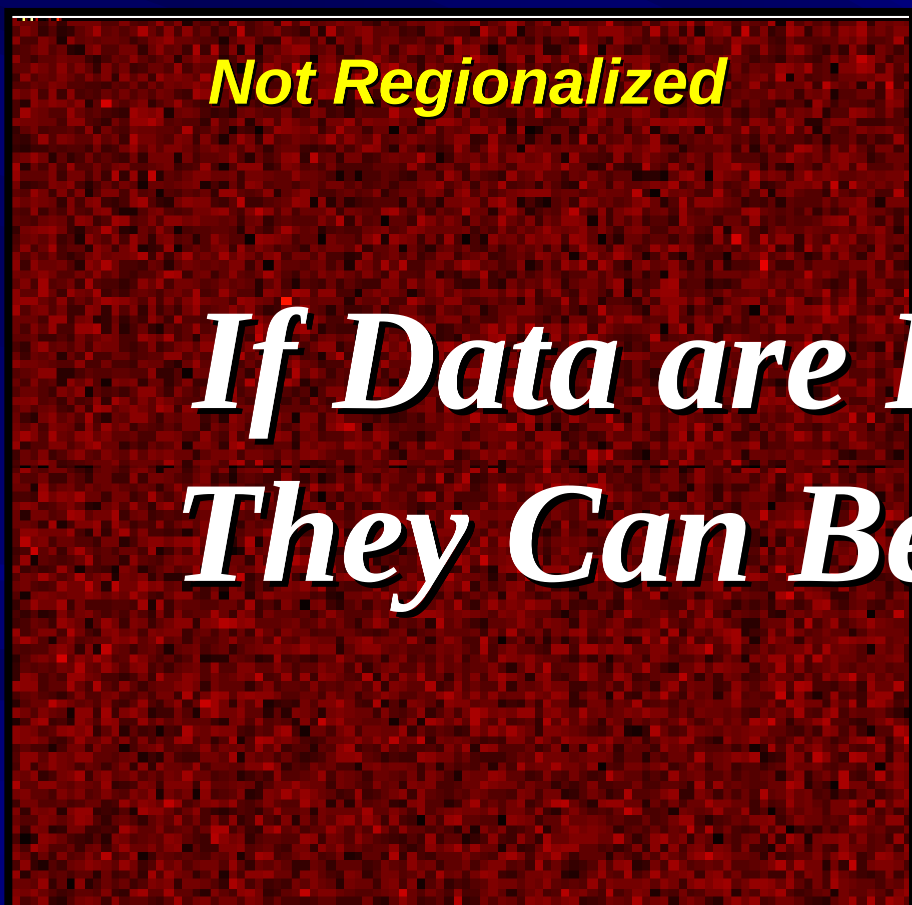


Smoothed Data



Foundational ***Assumption*** of Contouring

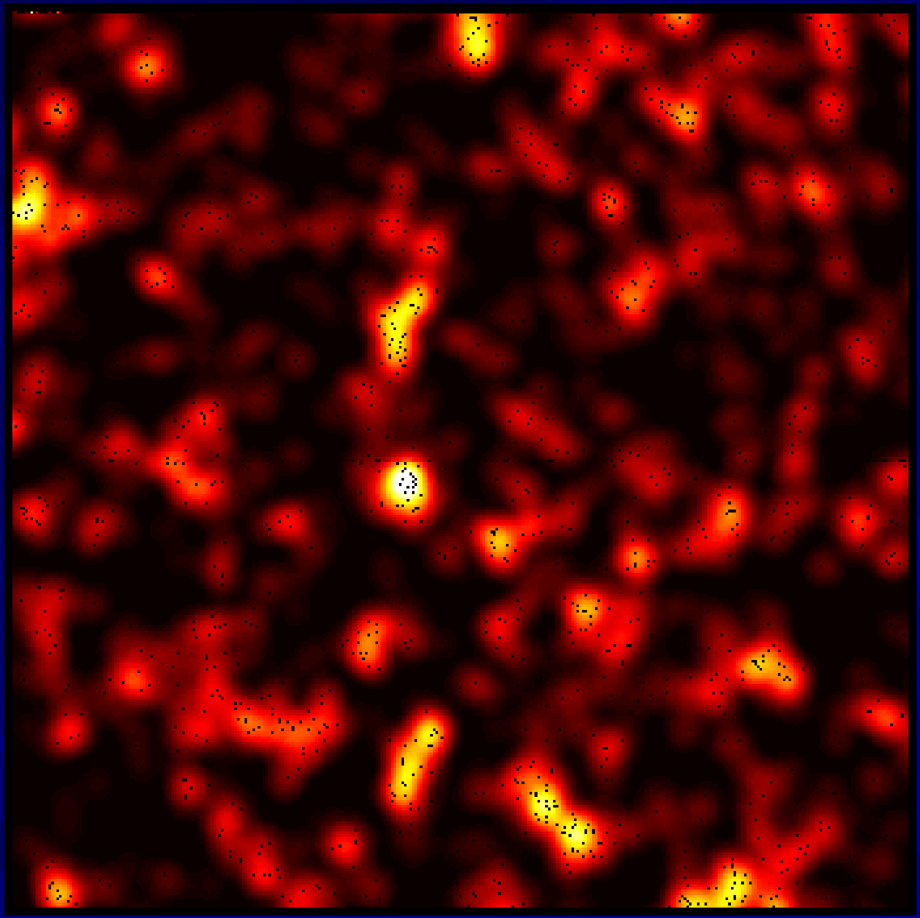
the data are ***regionalized*** (*if the variable is plotted in space, it describes a relatively smooth surface*)



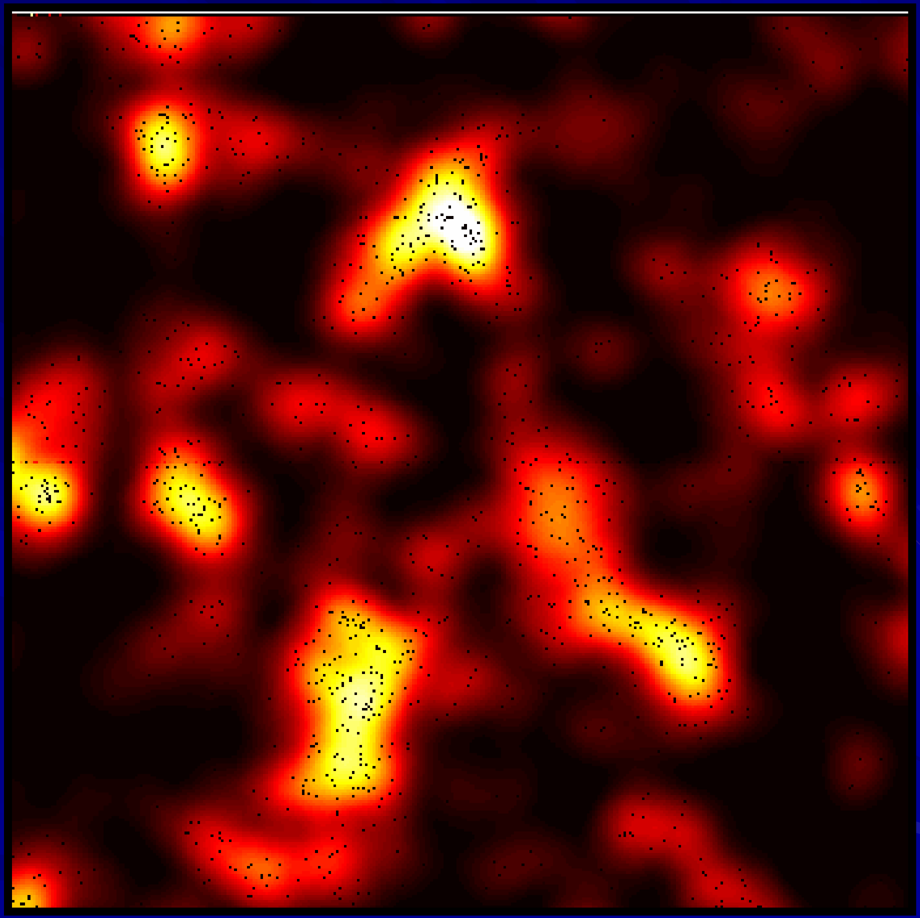
If Data are Regionalized, They Can Be Contoured !

Regionalization at Various Scales

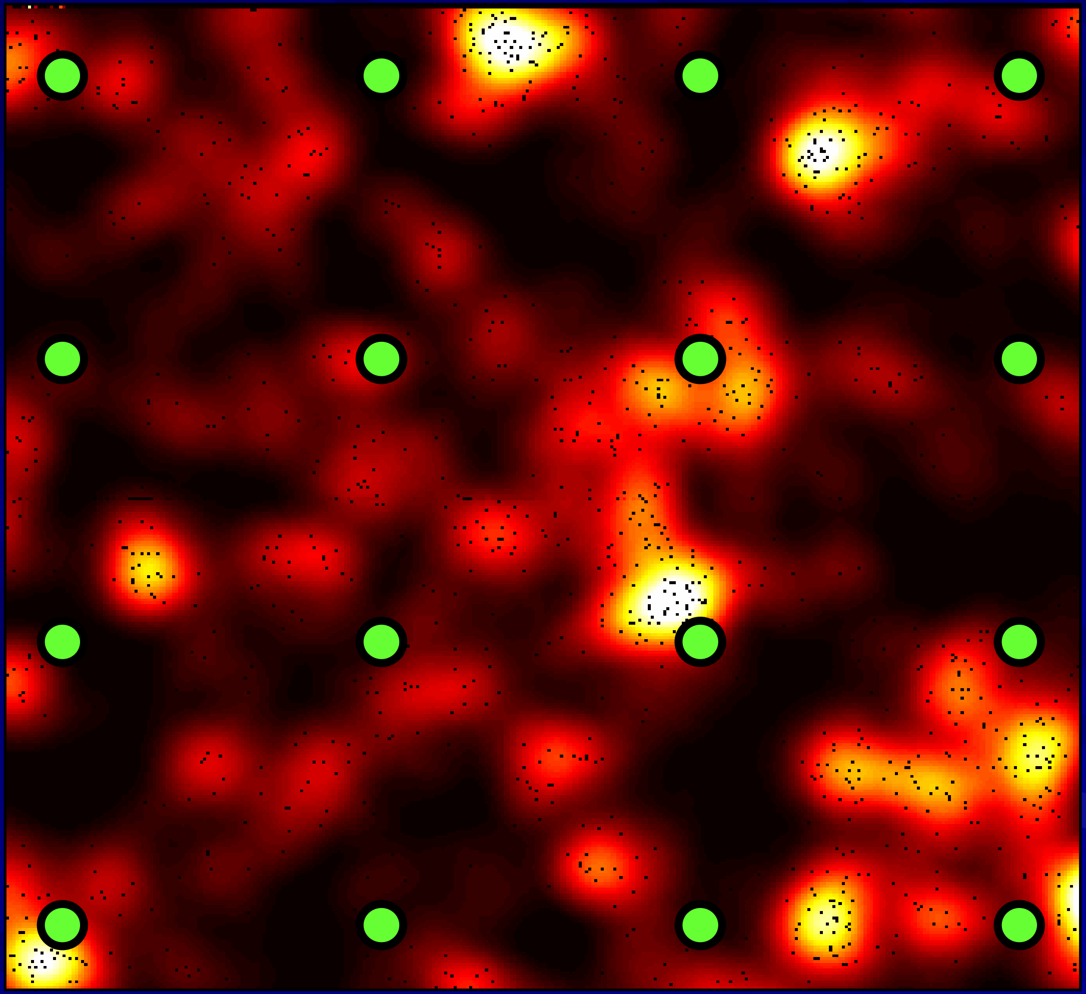
Small Scale Regionalization



Large Scale Regionalization

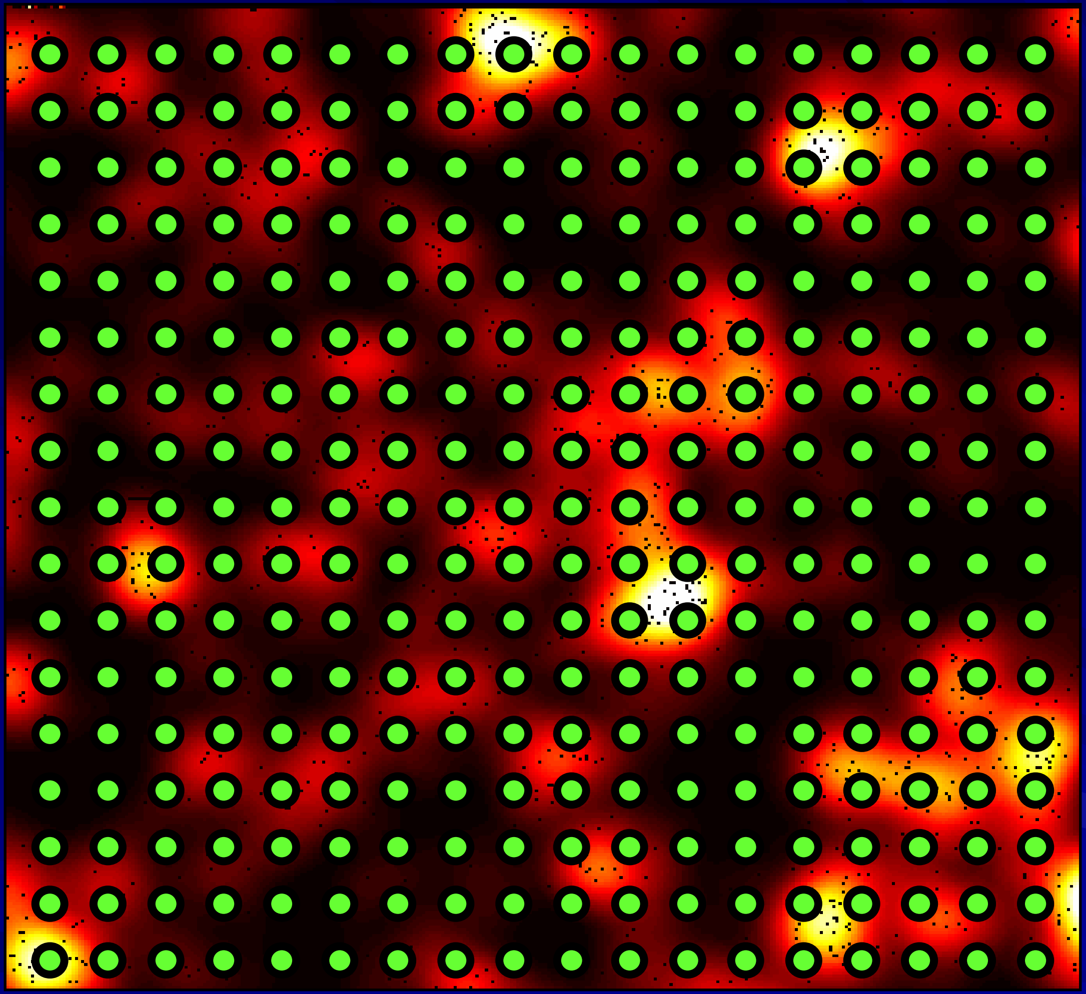


Regionalization Scale vs. Survey Scale



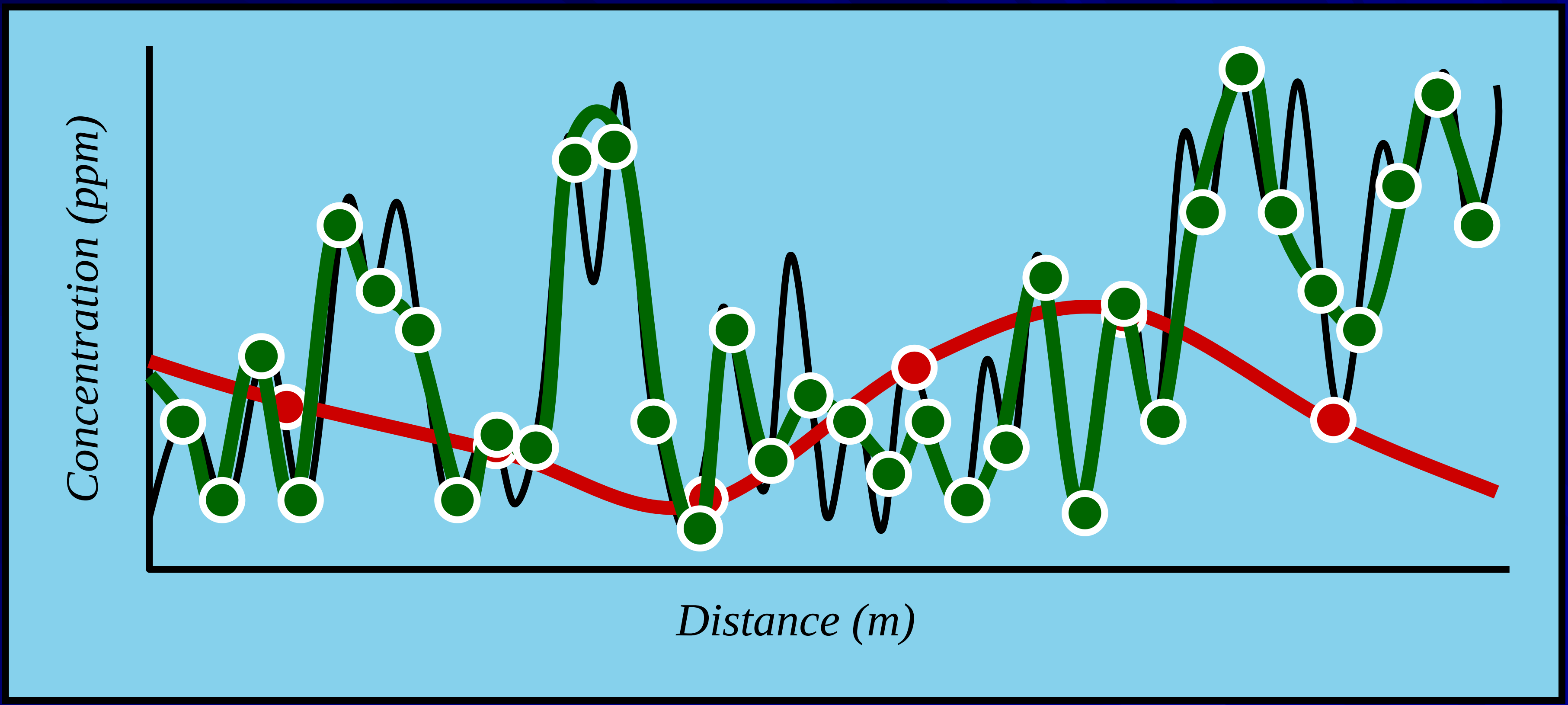
Scale of the Survey > Scale of Regionalization

Regionalization Scale vs. Survey Scale



Scale of the Survey < Scale of Regionalization

Aliasing

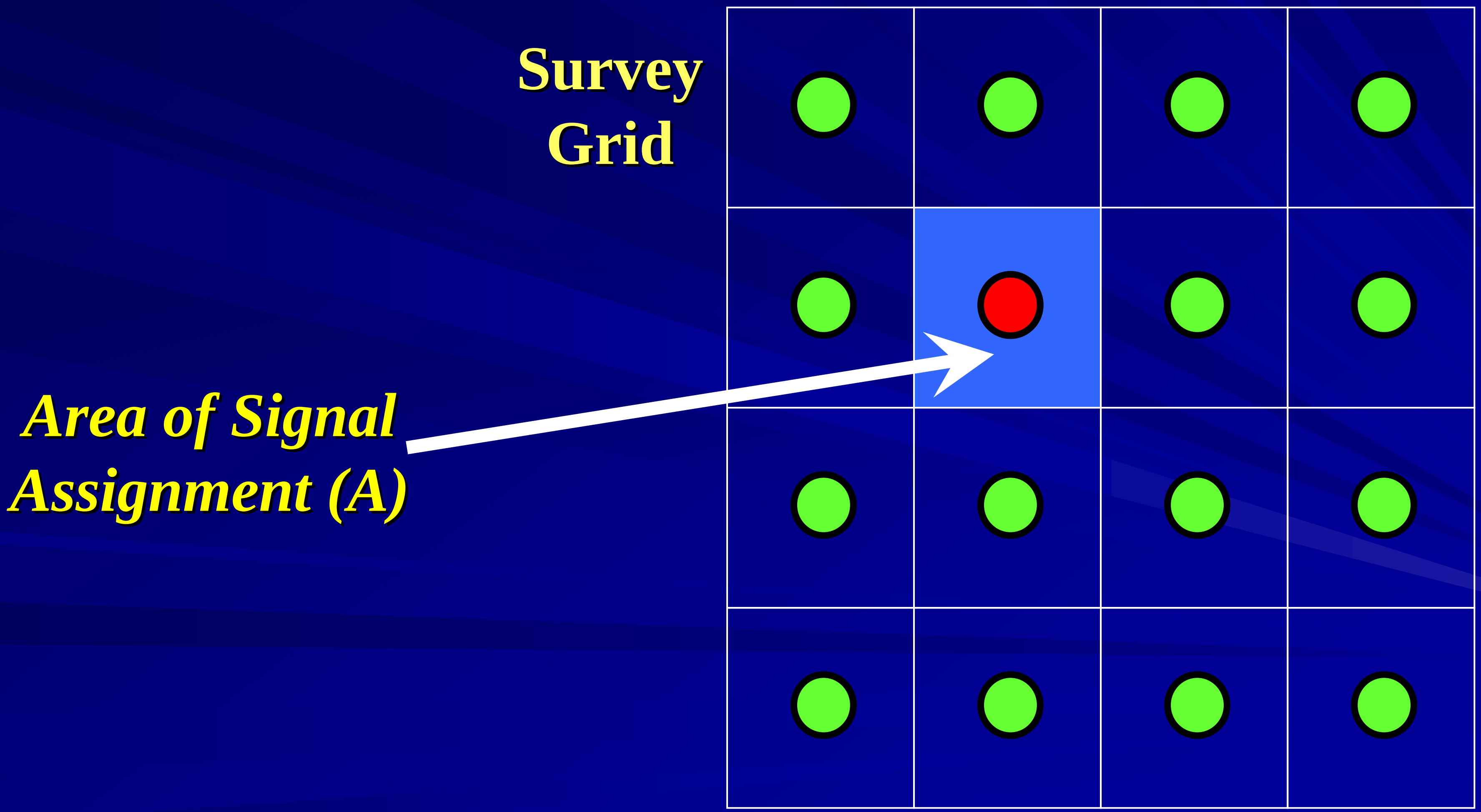


high density samples ⇒ more accurate representation of small scale regionalization

The Two Critical Areas

Area (Volume, Length) of Signal Assignment

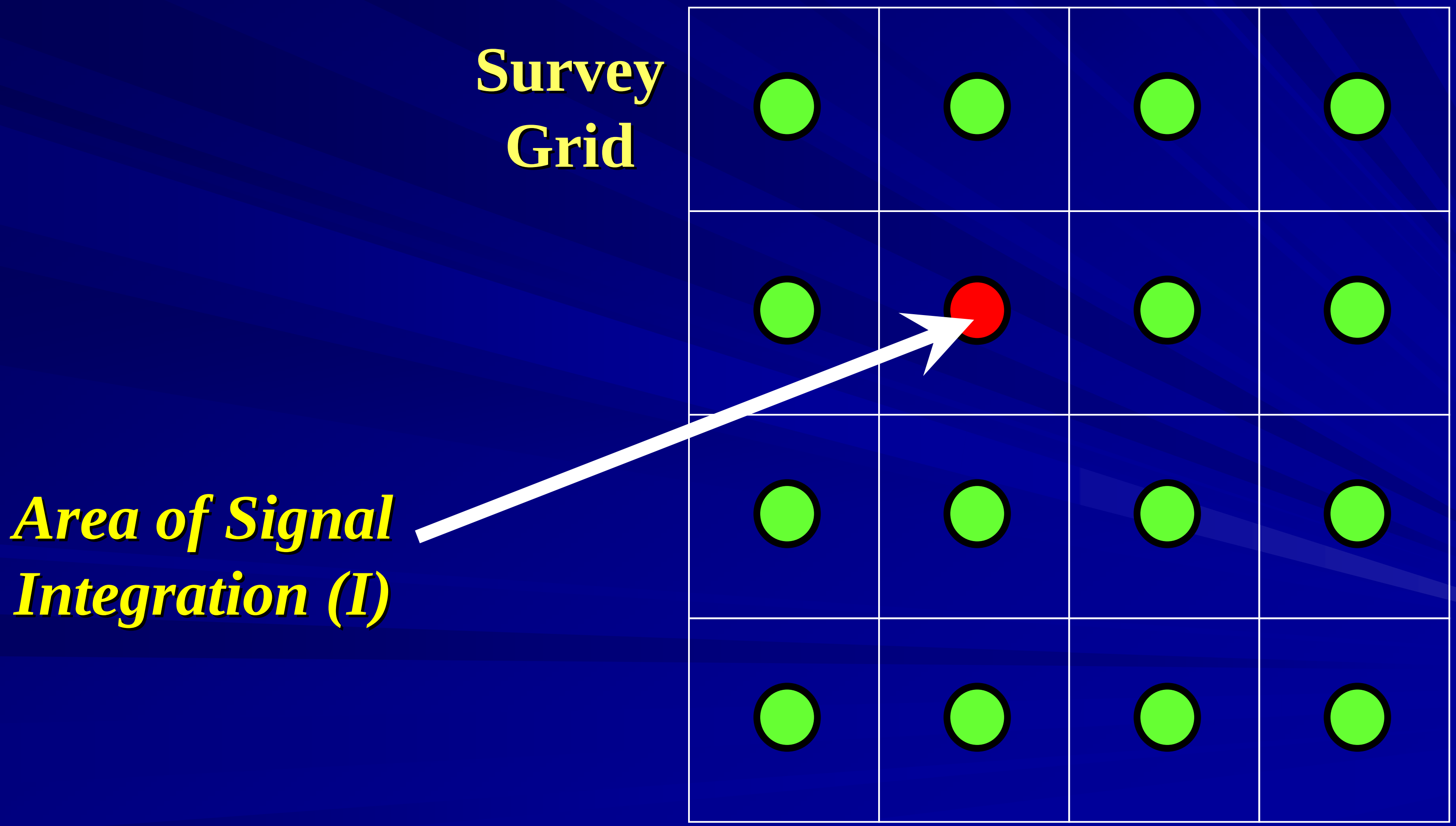
dependent on survey scale (*sample density, sample spacing*)



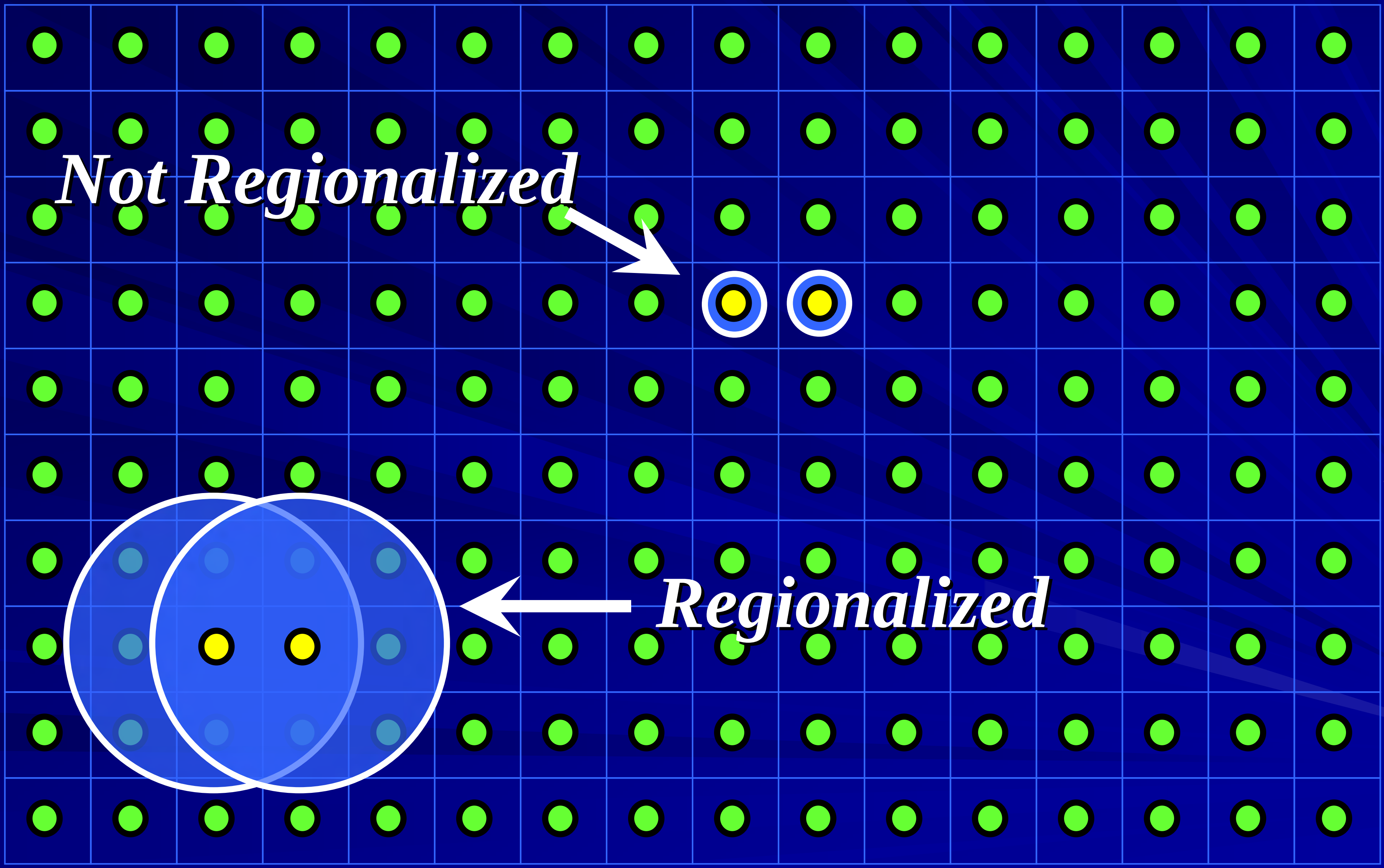
The Two Critical Areas

Area (Volume, Length) of Signal Integration

dependent on measurement scale (*how big is the sample*?)



Comparing The Sizes of The Two Areas



Comparing The Sizes of The Two Areas

if (I > A)

adjacent samples will be at least partially correlated because their areas of signal integration overlap

data exhibit regionalization => ***justified to contour (on theoretical grounds)***

if (A > I)

there is no guarantee that adjacent samples will be correlated because their areas of signal integration do not overlap

data may not exhibit regionalization => ***no a priori justification to contour***

In Applied Geochemistry

- ***I*** is small (its the size of the sample)
- ***A*** is large (it's a function of the sample spacing)
- => in geochemistry, regionalization is not guaranteed
- no theoretical justification for contouring geochemical data
- This doesn't mean that we can't contour geochemical data
- It just means that we cannot demonstrate geochemical data to be regionalized on theoretical grounds
- If we can demonstrate data to be regionalized on empirical grounds, then it can be contoured

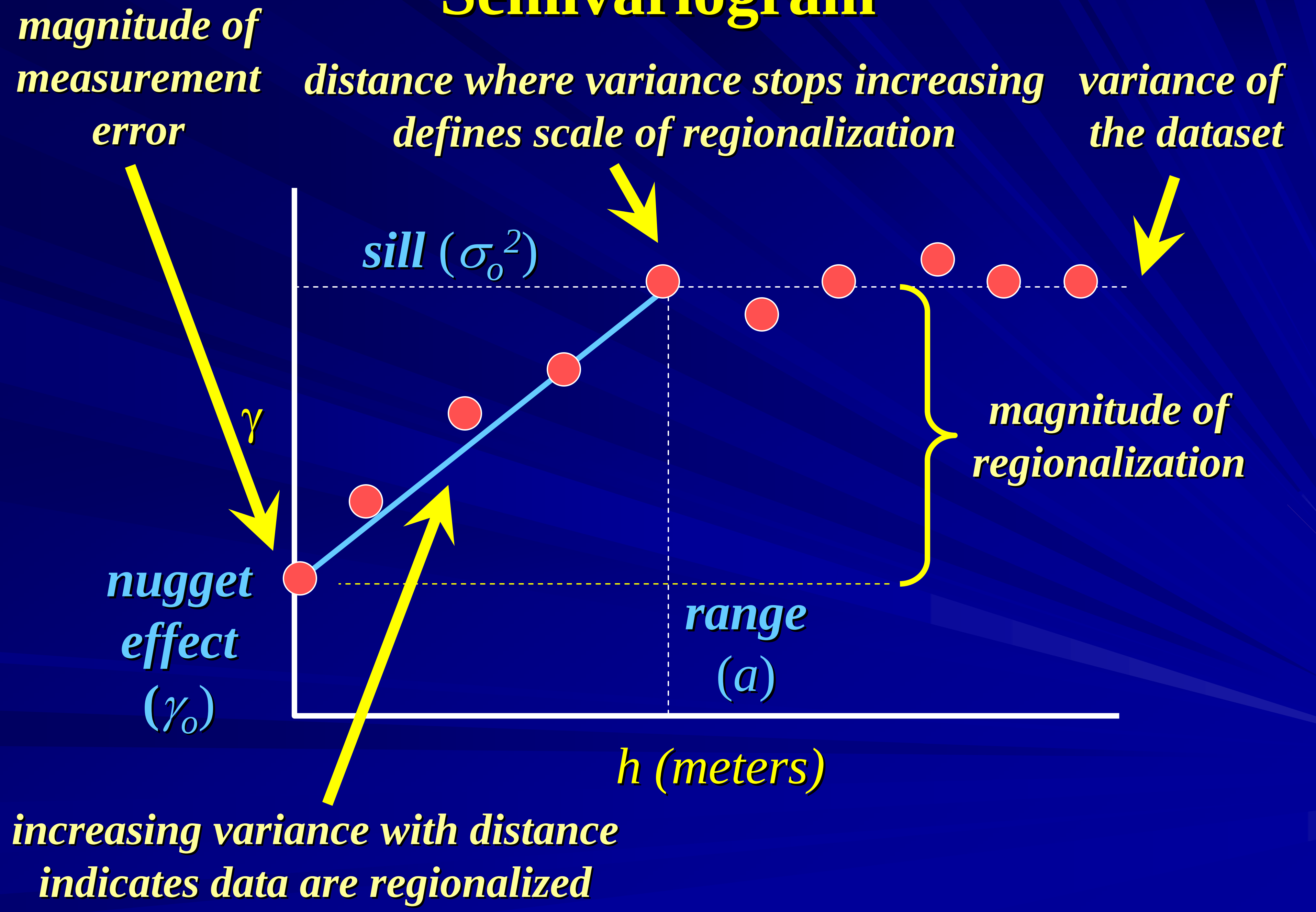
Empirical Assessment of Regionalization

Several ways to demonstrate geochemical data are regionalized:

1.) ***semivariograms*** (**geostatistics** => *average variance of pairs increases with distance between pairs*)

2.) ***bubbleplots*** (*plot circles at sample locations with the size of the circle proportional to the value of the variable*)

Semivariogram



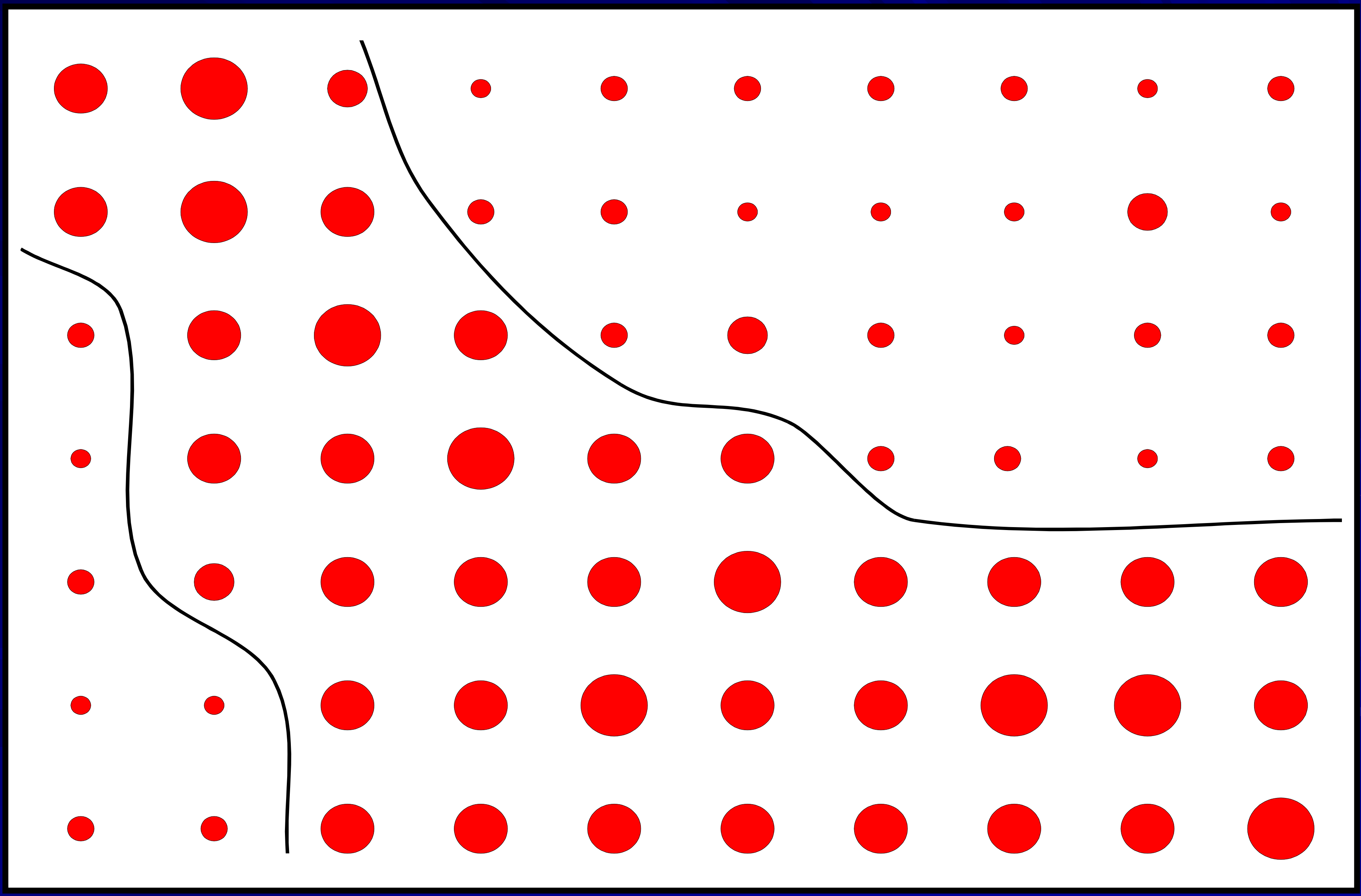
Empirical Assessment of Regionalization

Several ways to demonstrate geochemical data are regionalized:

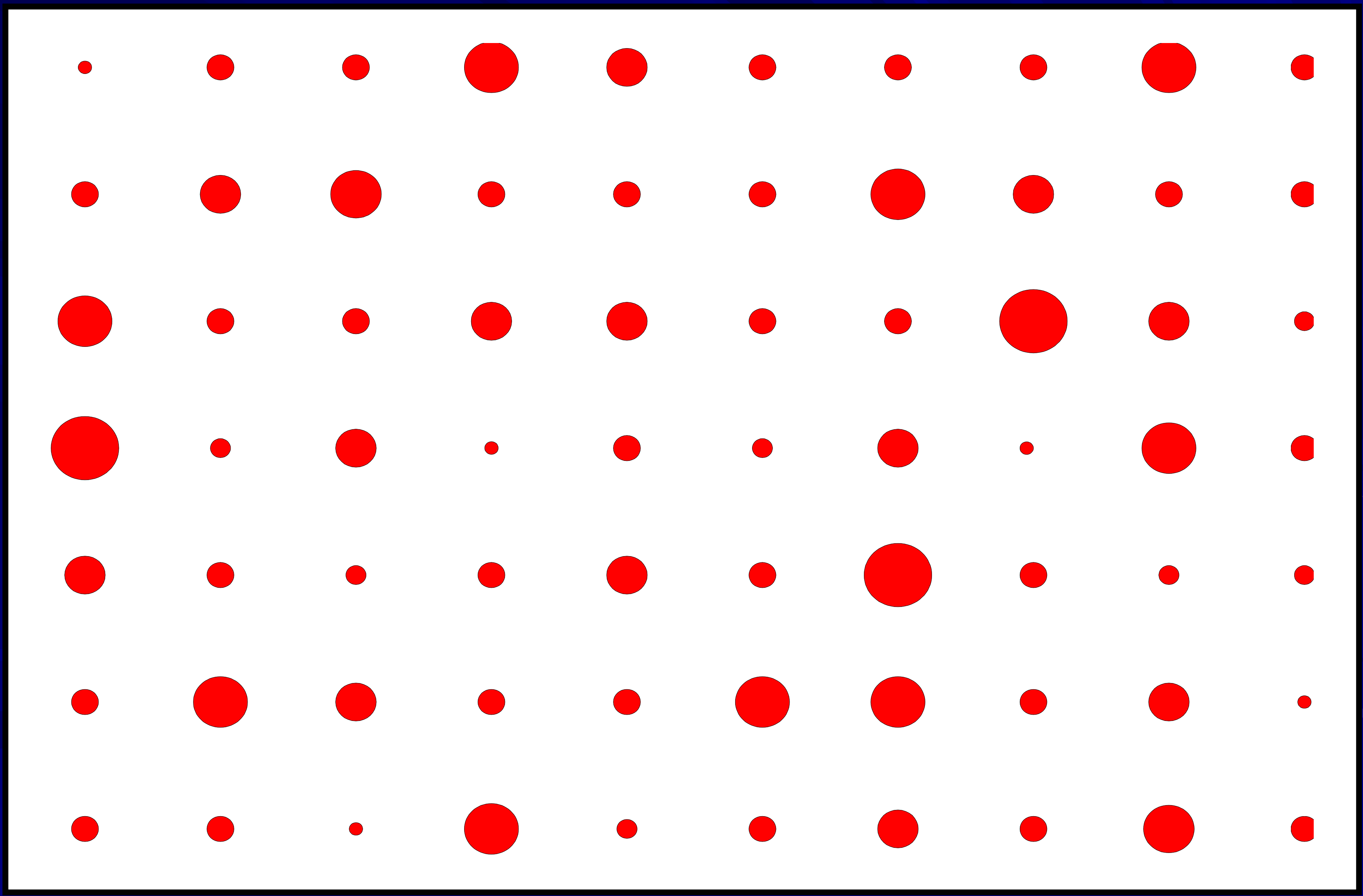
1.) ***semivariograms*** (**geostatistics** => *average variance of pairs increases with distance between pairs*)

2.) ***bubbleplots*** (*plot circles at sample locations with the size of the circle proportional to the value of the variable*)

Bubbleplot - Regionalized



Bubbleplot – Not Regionalized



Logic Stream

Are the data theoretically regionalized? ($I > A$)

N

Determine whether data are empirically regionalized (*via semivariograms or bubbleplots*)

Y

Y

N

Contour

DO NOT CONTOUR!
Theoretical and empirical justification is lacking; results may misrepresent data! (*use bubbleplots*).

Words of Caution

1) Standard contouring algorithms can create trends in the data
2) We should evaluate whether data exhibits regionalization at an appropriate scale before contouring
3) Otherwise, trends might be created during contouring
4) => possibly misinterpreted as being caused by real geological phenomena (when none exist)

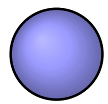
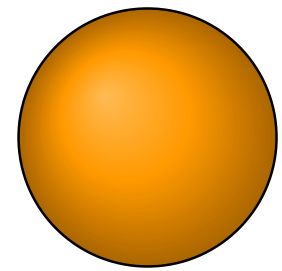
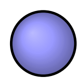
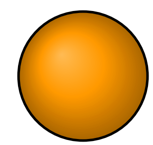
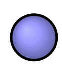
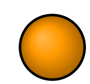
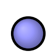
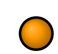
So if one shouldn't contour a dataset, how should we represent the data?

Use Bubbleplots!

1) They don’t smooth the data, and so don’t create trends
2) Bubble size is proportional to the geochemical variable, so any trends or patterns are accurately represented
3) Data transformations can be used to improve geochemical contrast

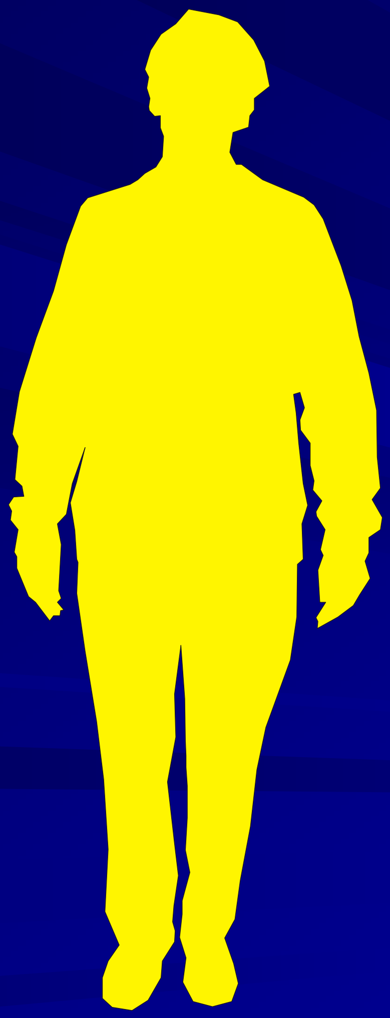
But how do we define/represent bubble size?

Example Bubble Sizes

Value	*Bubble Area Doubles*	*Bubble Diameter Doubles*
8		
4		
2		
1		

Perception of Size

Dr. Charles Butt
190 cm tall
7481 cm^2



Dr. David Cohen
167 cm tall
6026 cm^2

The Golden Ratio = ϕ = 1.618034

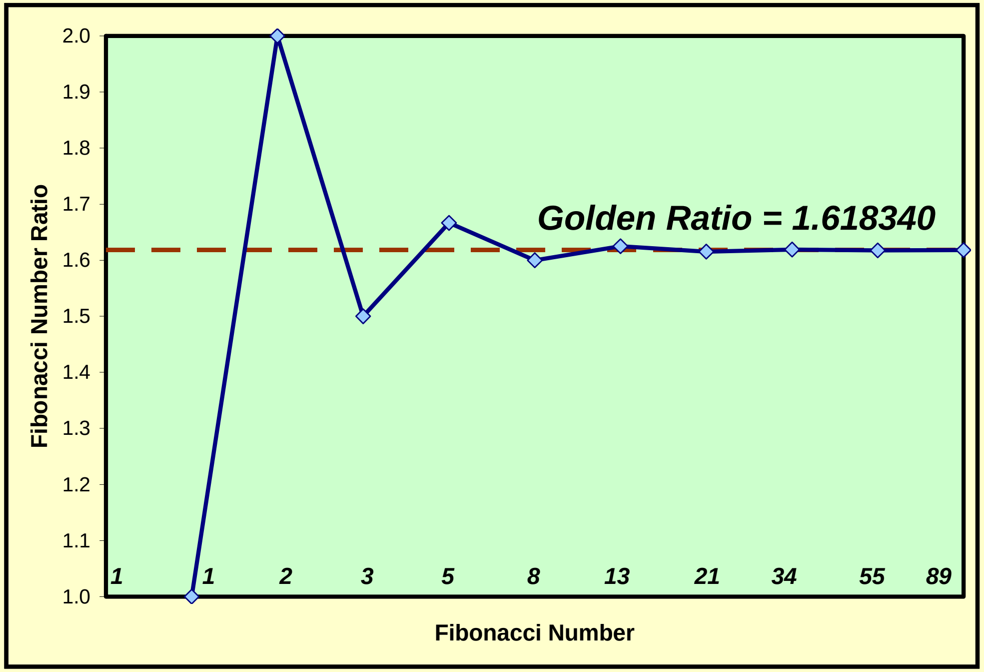
Fibonacci Numbers:

1, 1, 2, 3, 5, 8, 13, 21, 34, 55,

89, 144, 233, 377, 610, 987, 1597…

$$\frac{F_{i+1}}{F_i} \Rightarrow \phi$$

e.g., 1597/987 = 1.618034



The Golden Ratio = ϕ

$$\phi = 1.618034 = \frac{1}{2}\left(1+\sqrt{5}\right) = 2\cos\left(\frac{\pi}{5}\right)$$

$$\frac{1}{\phi} = \frac{1}{1.618034}$$

$$= 0.618034$$

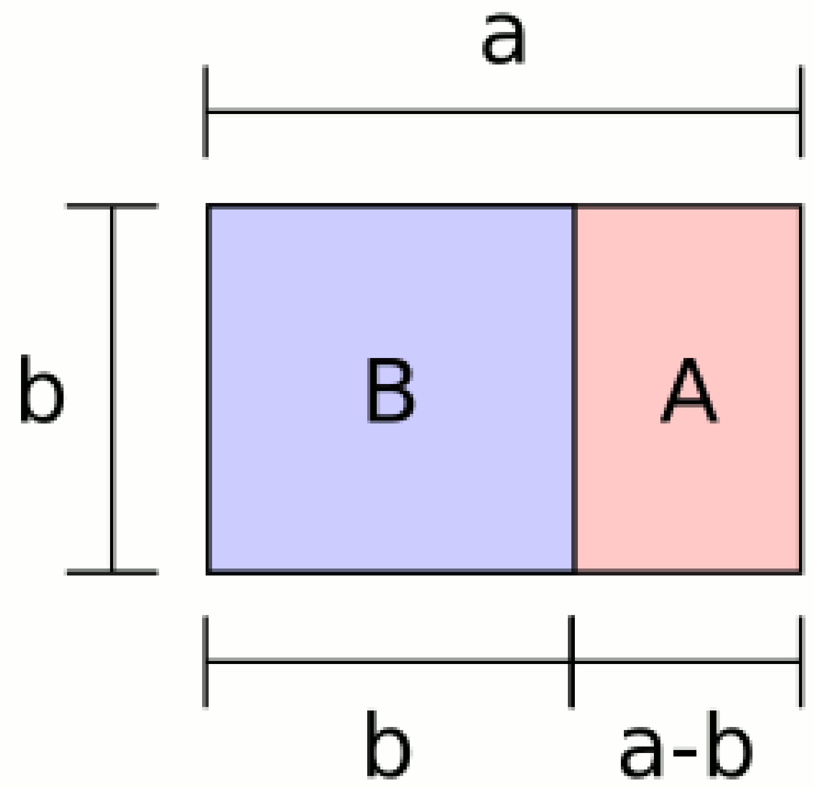
$$= \phi - 1$$

$$\phi = \frac{13}{8} + \sum_{n=0}^{\infty} \frac{(-1)^{n+1}(2n+1)!}{(n+2)!\,n!\,4^{2n+3}}$$

$$\phi = 1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \cdots}}}}$$

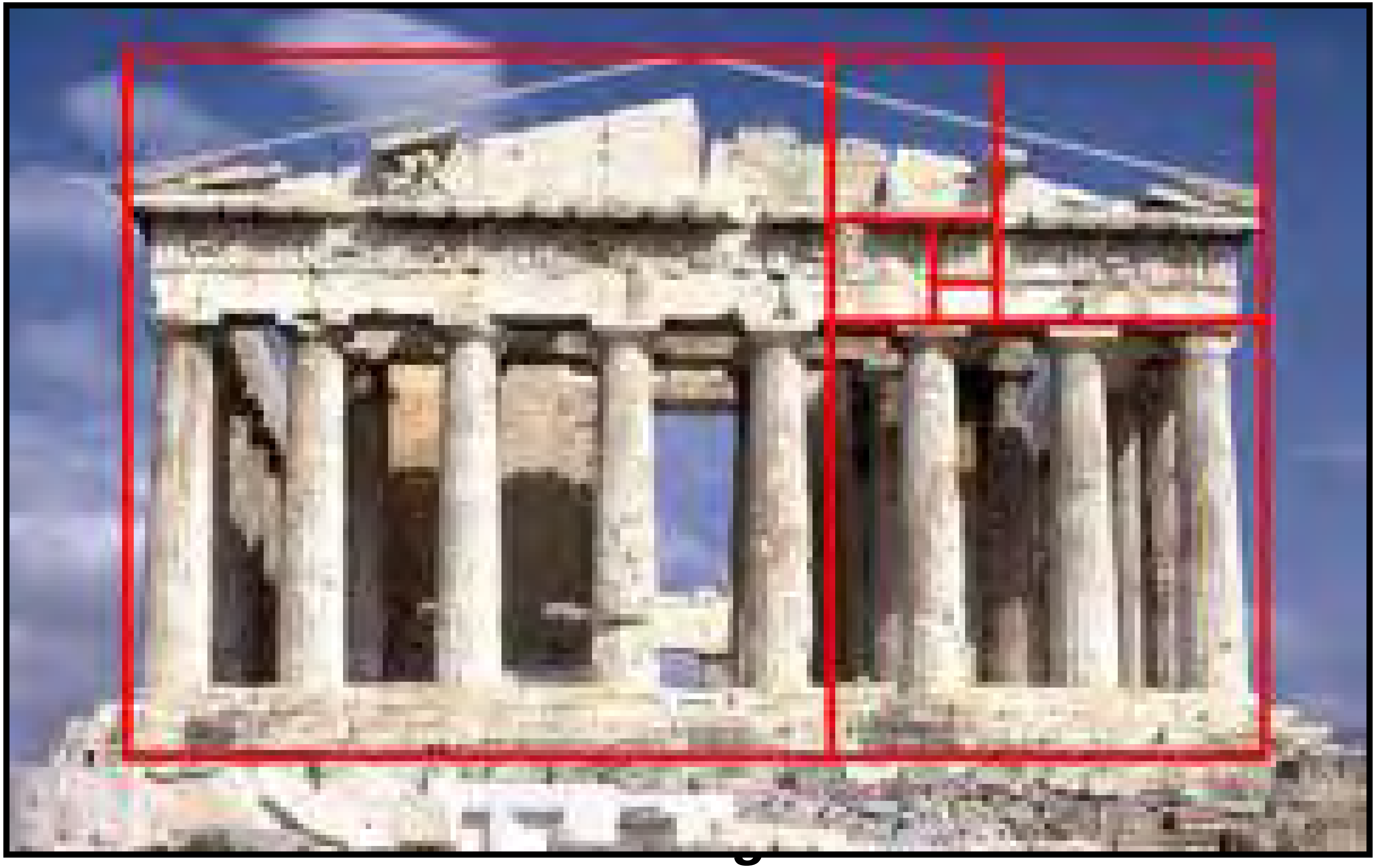
$$\phi = \sqrt{1+\sqrt{1+\sqrt{1+\sqrt{1+\cdots}}}}$$

The Golden Ratio = ϕ

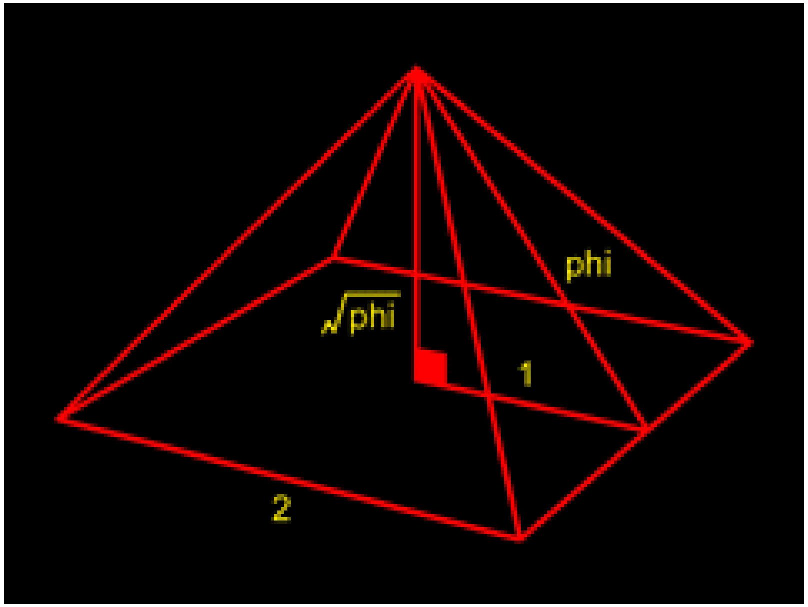


$$\frac{a}{b} = \frac{b}{a-b}$$

The Golden Ratio = ϕ

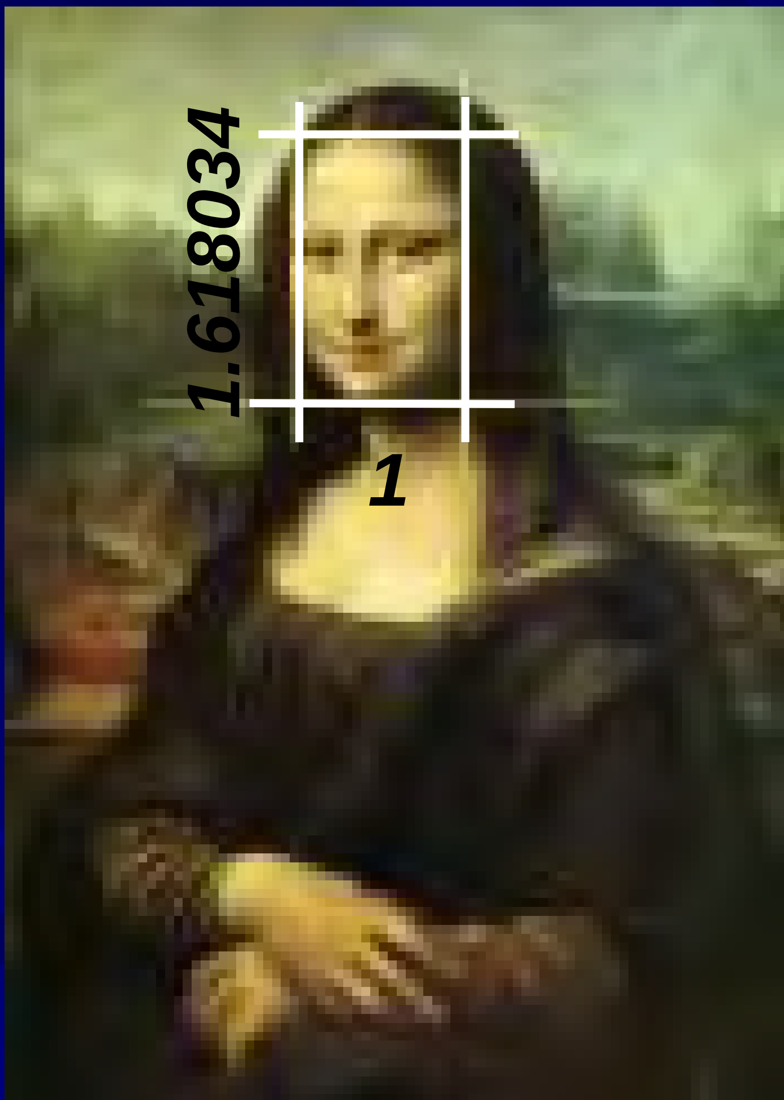


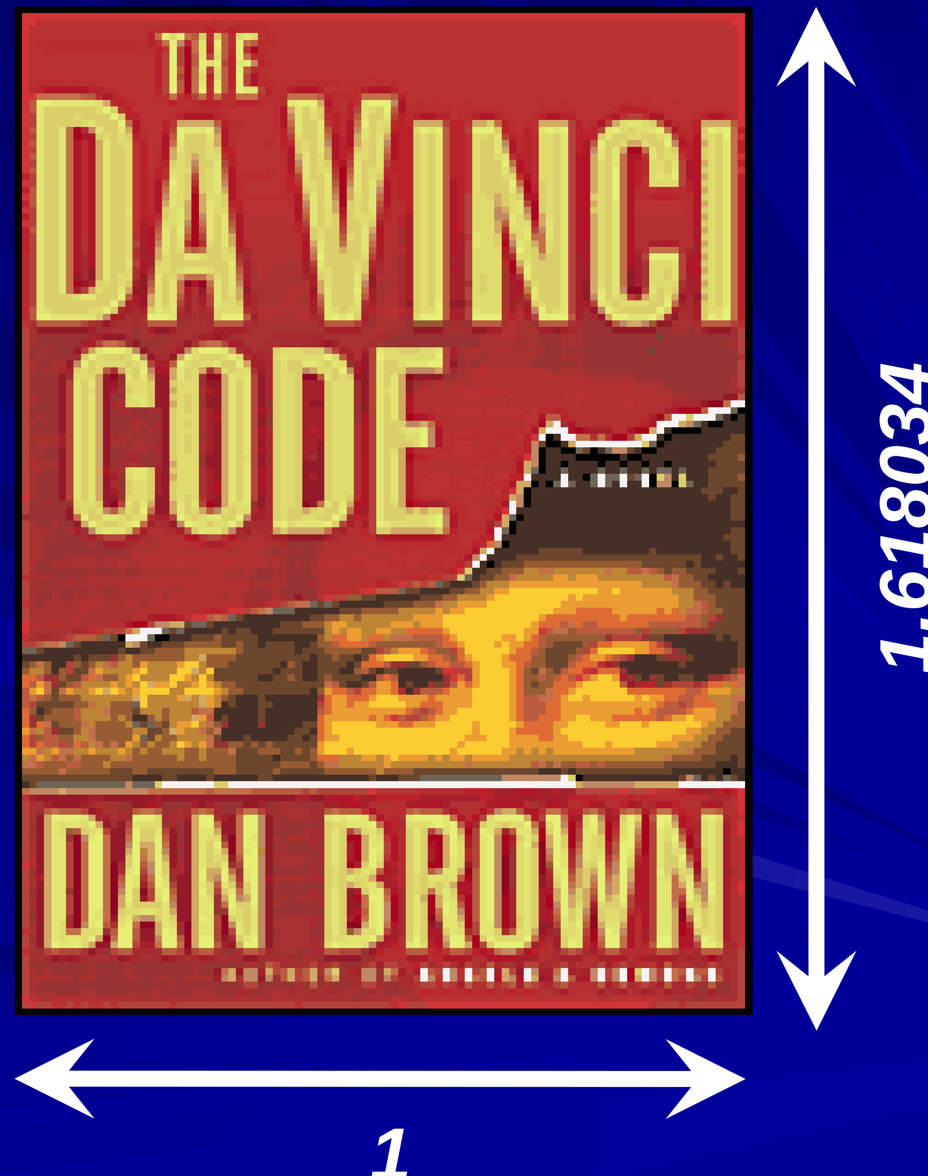
The Golden Ratio = ϕ



Pyramids

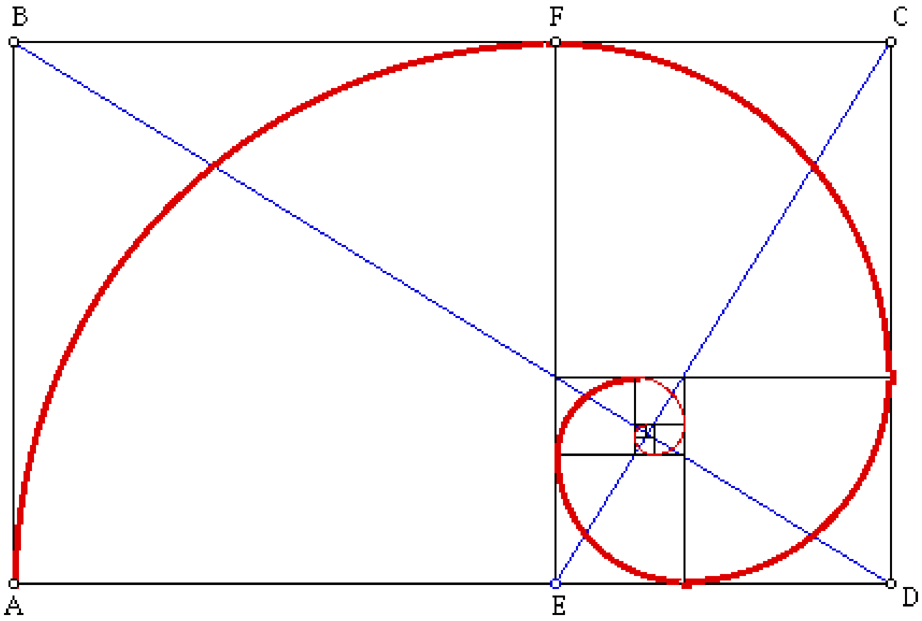
The Golden Ratio = ϕ

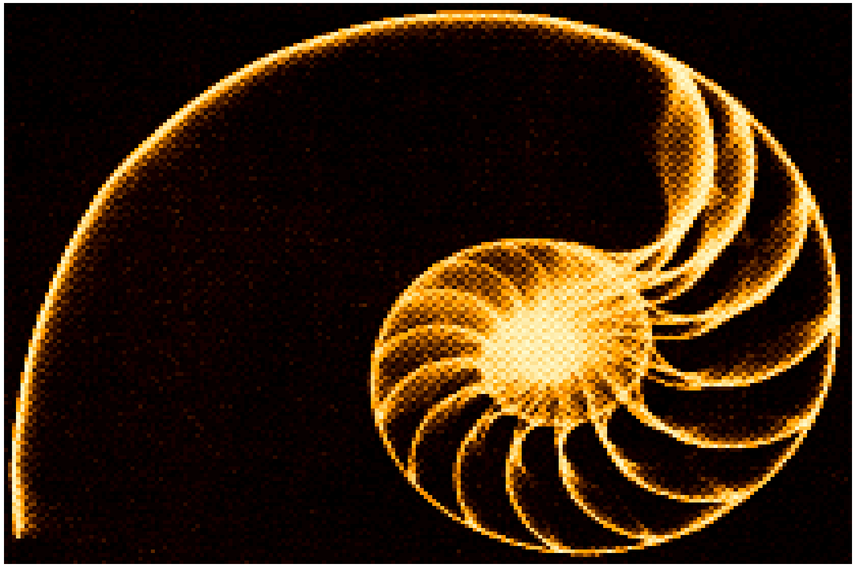




The Golden Ratio = ϕ

The Golden Ratio = ϕ

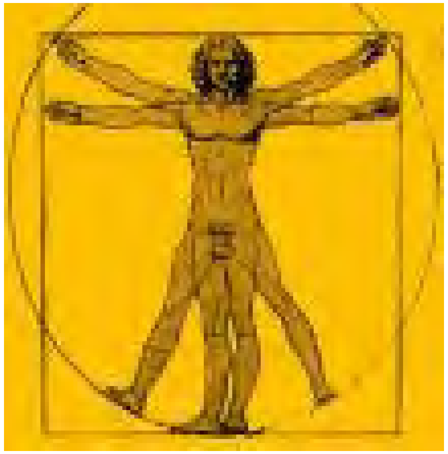




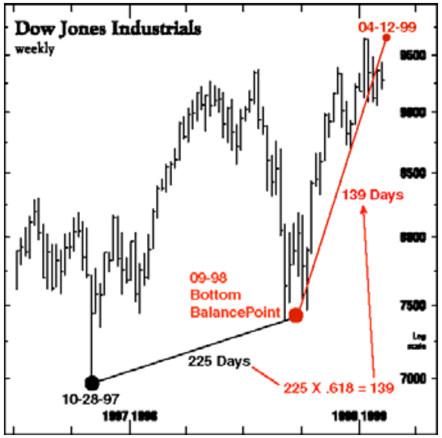
Spirals

The Golden Ratio = ϕ

Ergonomic Chairs



Human Body Proportions

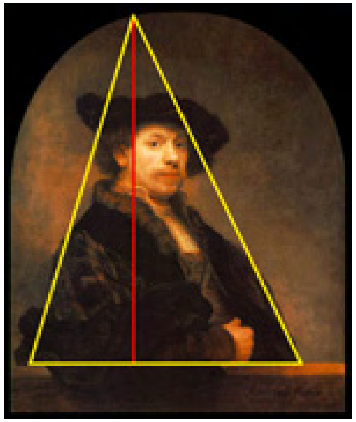


Stock Market

Tesla Coil

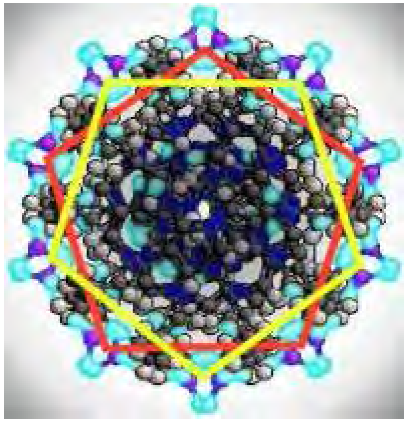
Flags

Fruit



Fine Art

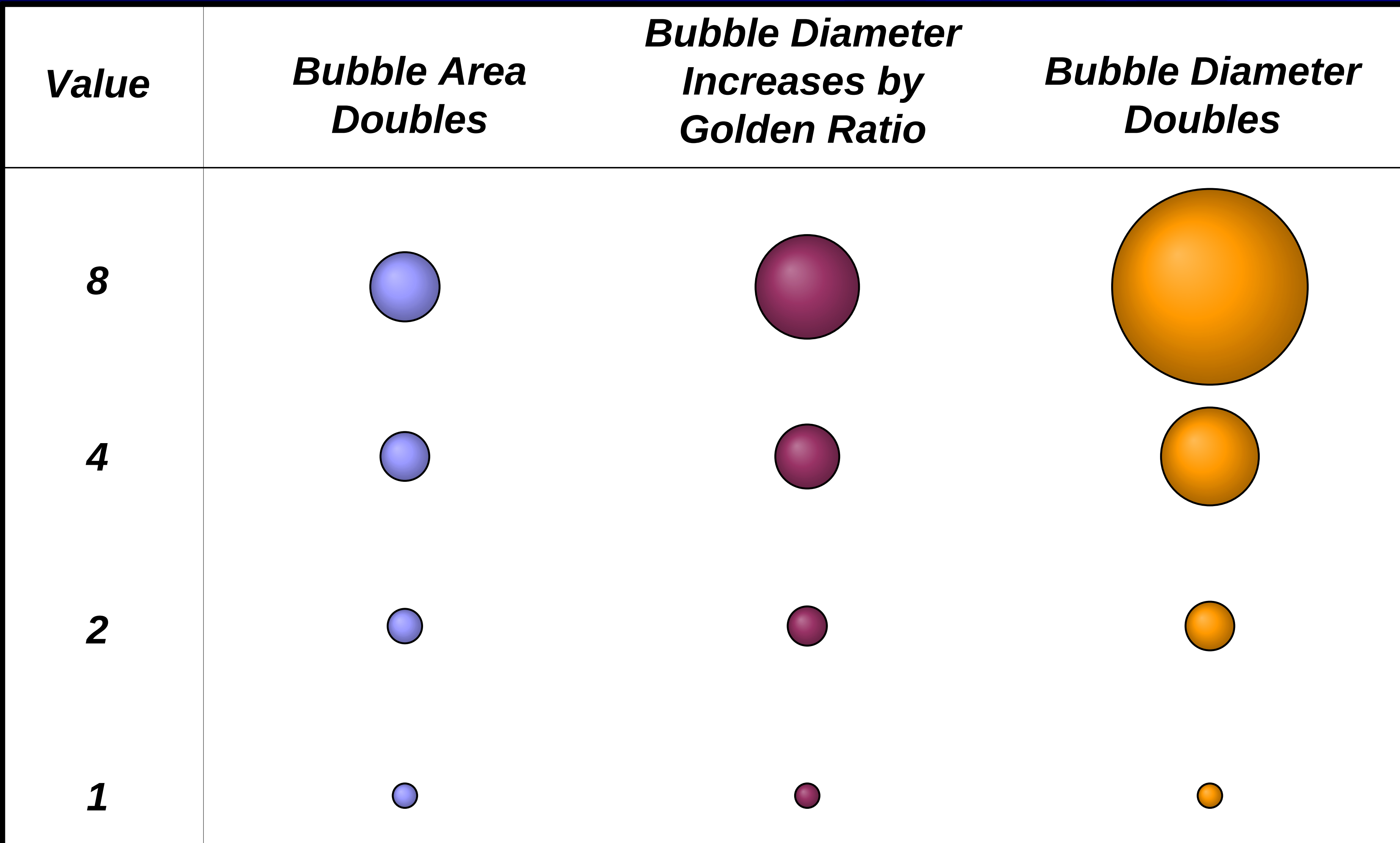
Nature (coneflower)



DNA Symmetry

The Golden Ratio = ϕ

Improper *Aesthetic* *Accurate*



Two Points

- Don’t contour your geochemical data unless the data are regionalized at an appropriate scale
- Use bubbleplots with sizes proportioned by diameter or by the Golden Ratio!

Thank You!