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

Easting (100's m)

Northing (100's m)
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 $\Rightarrow$ low frequency representation of small scale regionalization

low density samples $\Rightarrow$ low frequency representation of small scale regionalization
The Two Critical Areas

Area (Volume, Length) of Signal Assignment
dependent on survey scale (sample density, sample spacing)

Survey Grid

Area of Signal Assignment (A)
The Two Critical Areas

**Area (Volume, Length) of Signal Integration**
dependent on measurement scale (*how big is the sample?*)

<table>
<thead>
<tr>
<th>Survey Grid</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="Survey Grid Diagram" /></td>
</tr>
</tbody>
</table>

**Area of Signal Integration (I)**
Comparing The Sizes of The Two Areas

Not Regionalized

Regionalized
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)
- \( \Rightarrow \) 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

- **nugget effect** ($\gamma_0$)
- **range** ($a$)
- **sill** ($\sigma_0^2$)

**magnitude of measurement error**

**distance where variance stops increasing**

**defines scale of regionalization**

**variance of the dataset**

**increasing variance with distance**

**indicates data are regionalized**
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 – Not Regionalized
Are the data theoretically regionalized? 

\[ (I > A) \]

Determine whether data are empirically regionalized 

(\textit{via semivariograms or bubbleplots})

\[ \text{Y} \]

\[ \text{N} \]

\[ \text{DO NOT CONTOUR!} \]

Theoretical and empirical justification is lacking; results may misrepresent data!

(\textit{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

<table>
<thead>
<tr>
<th>Value</th>
<th>Bubble Area Doubles</th>
<th>Bubble Diameter Doubles</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td><img src="" alt="8" /></td>
<td><img src="" alt="8" /></td>
</tr>
<tr>
<td>4</td>
<td><img src="" alt="4" /></td>
<td><img src="" alt="4" /></td>
</tr>
<tr>
<td>2</td>
<td><img src="" alt="2" /></td>
<td><img src="" alt="2" /></td>
</tr>
<tr>
<td>1</td>
<td><img src="" alt="1" /></td>
<td><img src="" alt="1" /></td>
</tr>
</tbody>
</table>
Perception of Size

Dr. Charles Butt
190 cm tall
7481 cm²

Dr. David Cohen
167 cm tall
6026 cm²
The Golden Ratio = $\phi = 1.618034$

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 = $\phi$

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

\[
\frac{1}{\phi} = \frac{1}{1.618034} = 0.618034 = \phi - 1
\]

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

\[
\phi = 1 + \frac{1}{1 + \frac{1}{1 + \frac{1}{1 + \frac{1}{1 + \cdots}}}}
\]

\[
\phi = \sqrt{1 + \sqrt{1 + \sqrt{1 + \sqrt{1 + \cdots}}}}
\]
The Golden Ratio = $\phi$

\[
\frac{a}{b} = \frac{b}{a - b}
\]
The Golden Ratio $= \phi$
The Golden Ratio = $\phi$

Pyramids
The Golden Ratio = $\phi$
The Golden Ratio = $\phi$
The Golden Ratio = $\phi$

Spirals
The Golden Ratio = $\phi$

- Ergonomic Chairs
- Human Body Proportions
- Stock Market
- Tesla Coil
- Flags
- Fruit
- Fine Art
- Nature (coneflower)
- DNA Symmetry
## The Golden Ratio $\phi$

<table>
<thead>
<tr>
<th>Value</th>
<th>Improper</th>
<th>Aesthetic</th>
<th>Accurate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bubble Area Doubles</td>
<td>Bubble Diameter Increases by Golden Ratio</td>
<td>Bubble Diameter Doubles</td>
</tr>
<tr>
<td>8</td>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
</tr>
<tr>
<td>4</td>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
</tr>
<tr>
<td>2</td>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
</tr>
<tr>
<td>1</td>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
</tr>
</tbody>
</table>
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!