



Spatial Data Management

Viacheslav I. Adamchuk

*Biological Systems Engineering
University of Nebraska-Lincoln*

September 23, 2008



Main Tasks of Statistical Analysis

- On-Farm Research
- Year-to-Year (Temporal) Variability
- Spatial Variability (Geostatistics)
- Uncertainty Analysis
- Interpolation and Modeling
- Errors Assessment
- Probability Analysis and Risk Management



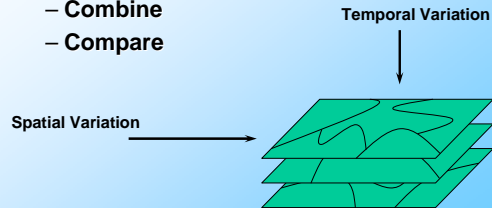
Statistical Parameters

- Mean (AVERAGE) $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$
- Variance (VAR) $s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}$
- Standard Deviation (SDEV) $s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}$
- Coefficient of Variation $CV = \frac{s}{\bar{x}}$



Yield Maps Analysis

- What do I do with five+ years of data?
 - Normalize
 - Combine
 - Compare

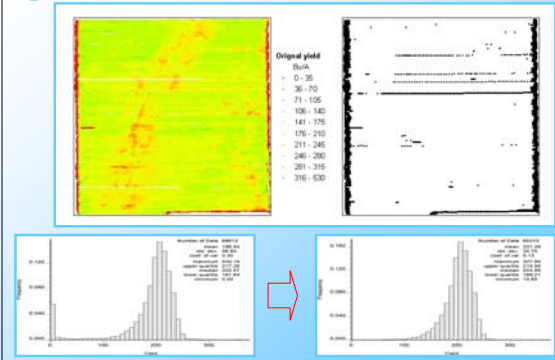


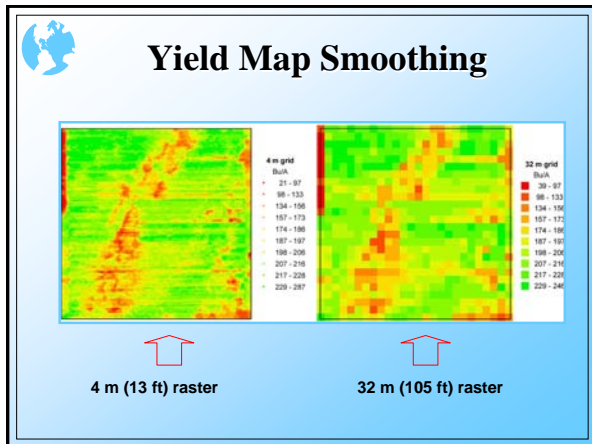
Data Filtering

- Header up points as well as start and end pass delays
- Points with yield values or individual sensor measurements exceeding the possible range
- Outliers based on descriptive statistics.
- Points with detectable misplacement (e.g. co-aligned points)
- Points that do not agree with a predefined statistical estimate based on local neighborhood statistics



Data Filtering

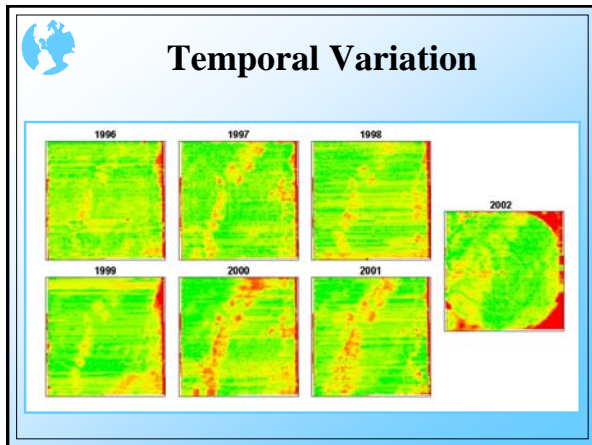




Yield Normalization

$$Yield_{relative} = \frac{Yield_{actual}}{Average_{year}}$$

$Yield_{relative} = 1 \rightarrow$ Average yield
 $Yield_{relative} < 1 \rightarrow$ Below average yield
 $Yield_{relative} > 1 \rightarrow$ Above average yield



Temporal Statistics

Average

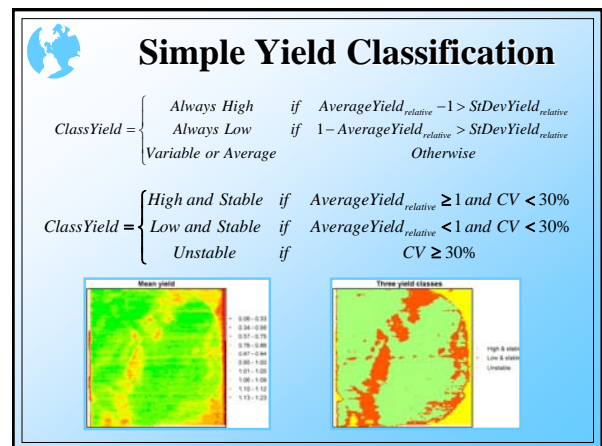
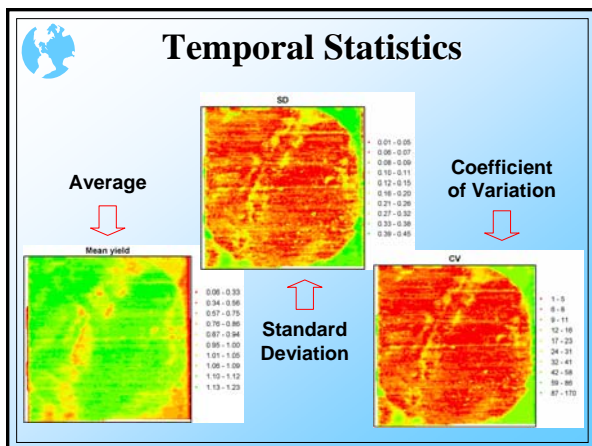
$$AverageYield_{relative} = \frac{Yield_{relative}^{year1} + Yield_{relative}^{year2} + \dots + Yield_{relative}^{yearN}}{N}$$

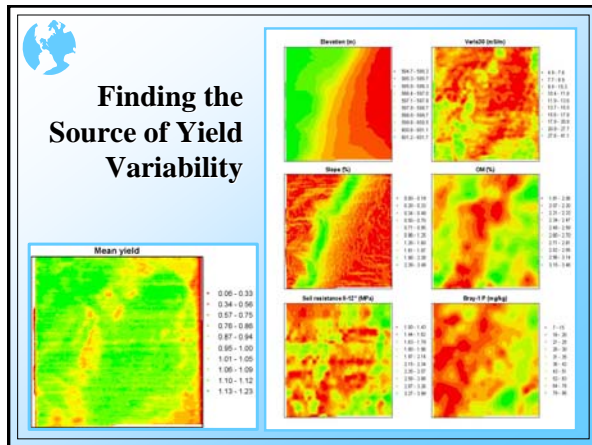
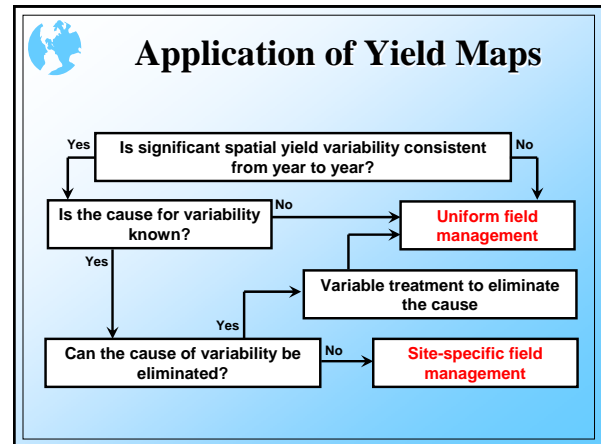
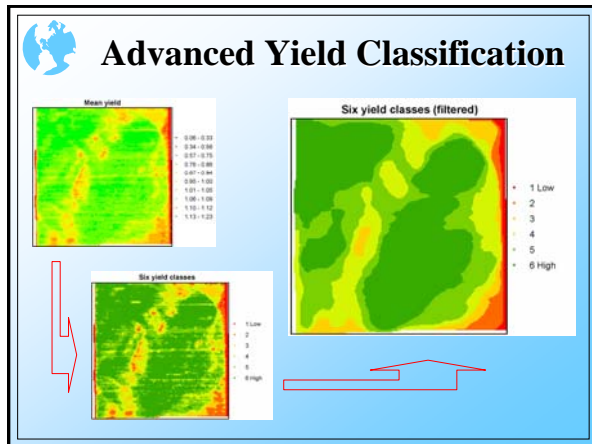
Standard Deviation

$$StDevYield_{relative} = \sqrt{\frac{(Yield_{relative}^{year1} - AverageYield_{relative})^2 + \dots + (Yield_{relative}^{yearN} - AverageYield_{relative})^2}{N-1}}$$

Coefficient of Variation

$$CV[\%] = \frac{StDevYield_{relative}}{AverageYield_{relative}} \cdot 100$$





- ## Map Interpolation
- Exact Interpolators
 - Nearest Neighbor
 - Inverse Distance Weighted with no smoothing factor specified
 - Splines (Radial Basis Functions) when you do not specify an R^2 value
 - Kriging when you do not specify a nugget effect
 - Smoothing Interpolators
 - Moving Average
 - Global or Local Polynomial Regression
 - Inverse Distance Weighted when you specify a smoothing factor
 - Splines when you specify an R^2 value
 - Kriging when you specify an error nugget effect

- ## Data Processing Recommendations
- Always filter data first
 - Check point data for validity
 - Do not use simple interpolators for multi-layer data analysis
 - Avoid too much smoothing
 - Use fixed size of interpolation grid
 - While selecting interpolator parameters keep in mind sampling density
 - Use common sense

