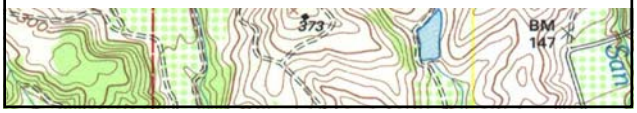
Geography 12: Maps and Spatial Reasoning

# Lecture 19: Applications of Feature Measurements

Professor Keith Clarke




## Measurement

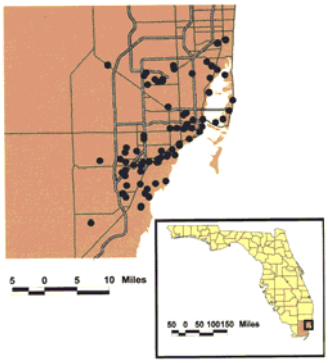
- Shape
  - Miller
  - Bunge
  - Boyce-Clark
  - Fourier measures
- Distribution
  - Quadrat analysis
  - Nearest neighbor analysis

**Geographic Information Systems and Ciguatera Fish Poisoning in the Tropical Western Atlantic Region**

John F Stinn, Donald P de Sylva, Lora E Fleming, Eileen Hack

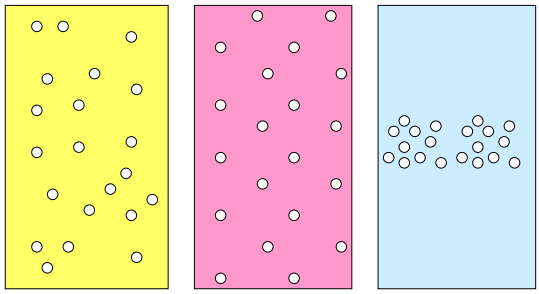
To evaluate the geographic distribution by residence at the time of illness, cases from 1978 to 1981 within Miami-Dade County, a ciguatera endemic region, were analyzed (Figure 1). Of the 304 index cases, 169 occurred in Miami-Dade County, with 102 (60.4% of Miami-Dade County cases) of these cases occurring during the specified time period. A nearest-neighbor analysis was performed in an attempt to show a random distribution of cases in the county. However, despite various attempts to adjust for population density and lack of habitability (e.g., airports, Everglades, and ocean areas), **the R-value was 0.10, indicating a strong clustering pattern**. Nevertheless, the clustering pattern closely followed densely populated roadways that

pass through highly varied neighborhoods.



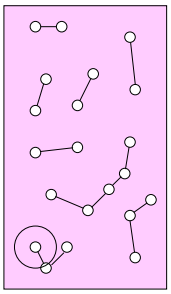

## Nearest-Neighbor Analysis

- Unlike quadrat analysis uses distances between points as its basis
- The mean of the distance observed between each point and its nearest neighbor is compared with the expected mean distance that would occur if the distribution were random
- Also needs a reference area



RANDOM UNIFORM CLUSTERED

## Distance to the nearest neighbor



| Point | Nearest Neighbour | Distance ( r ) |
|---|---|---|
| 1 | 2 | 1 |
| 2 | 3 | 0.1 |
| 3 | 2 | 0.1 |
| 4 | 5 | 1 |
| 5 | 4 | 1 |
| 6 | 5 | 2 |
| 7 | 6 | 2.7 |
| 8 | 10 | 1 |
| 9 | 10 | 1 |
| 10 | 9 | 1 |
| | | **10.9** |
| r | 1.09 | |
| Area of Region | 50 | |
| Density | 0.2 | |
| Expected Mean | 1.118034 | |
| R | 0.9749256 | |

RANDOM

| Point | Nearest Neighbour | Distance |
|---|---|---|
| 1 | 3 | 2.2 |
| 2 | 4 | 2.2 |
| 3 | 4 | 2.2 |
| 4 | 5 | 2.2 |
| 5 | 7 | 2.2 |
| 6 | 7 | 2.2 |
| 7 | 8 | 2.2 |
| 8 | 9 | 2.2 |
| 9 | 10 | 2.2 |
| 10 | 9 | 2.2 |
| | | **22** |
| r | 2.2 | |
| Area of Region | 50 | |
| Density | 0.2 | |
| Expected Mean | 1.118034 | |
| R | 1.9677398 | |

UNIFORM

| Point | Nearest Neighbour | Distance |
|---|---|---|
| 1 | 2 | 0.1 |
| 2 | 3 | 0.1 |
| 3 | 2 | 0.1 |
| 4 | 5 | 0.1 |
| 5 | 4 | 0.1 |
| 6 | 5 | 0.1 |
| 7 | 6 | 0.1 |
| 8 | 9 | 0.1 |
| 9 | 10 | 0.1 |
| 10 | 9 | 0.1 |
| | | **1** |
| r | 0.1 | |
| Area of Region | 50 | |
| Density | 0.2 | |
| Expected Mean | 1.118034 | |
| R | 0.0894427 | |

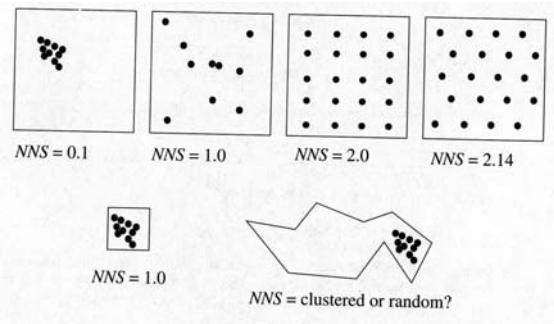
CLUSTERED

$$NNS = 2 \times \frac{\sum_{i=1}^{npts} d_i}{npts \times \sqrt{\frac{A}{npts}}}$$

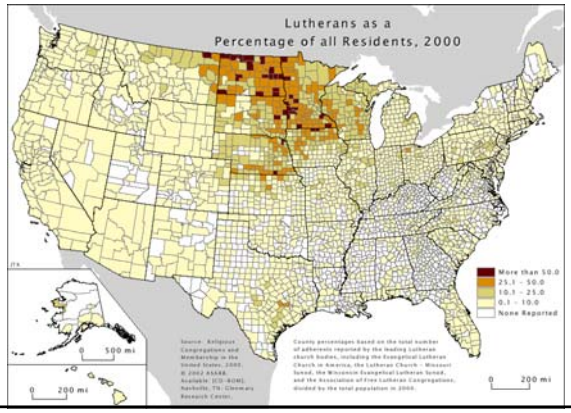
## Advantages of Nearest Neighbor over Quadrat Analysis

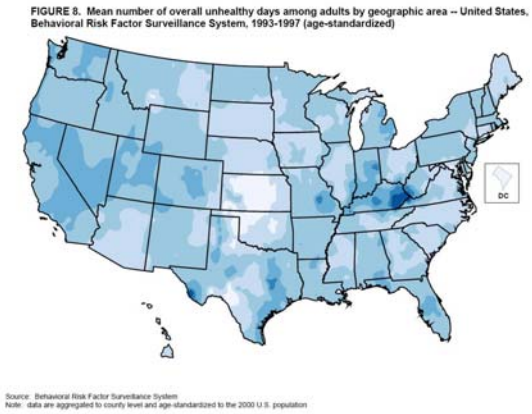
- No quadrat size problem to be concerned with
- Takes distance into account
- Problems
  - Related to the entire boundary size
  - Must consider how to measure the boundary
    - Arbitrary or some natural boundary
  - May not consider a possible adjacent boundary

## NNS problems




## Clustering

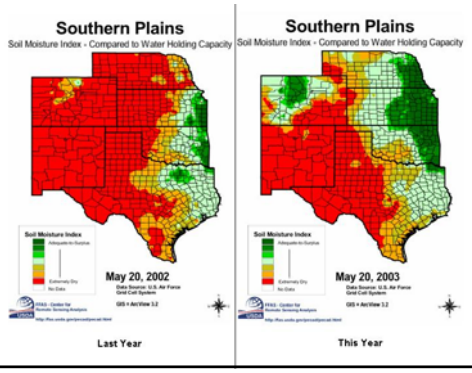


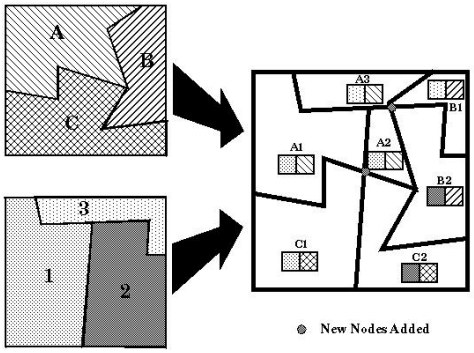

## Measurement

- Does A relate geographically to B
- Spatial Correspondence
  - Coefficient of areal correspondence (Set theory, intersection / union
  - Chi-square
  - Yule's Q

Compare over time, scale, theme, etc.




## Overlay




## Text Q example

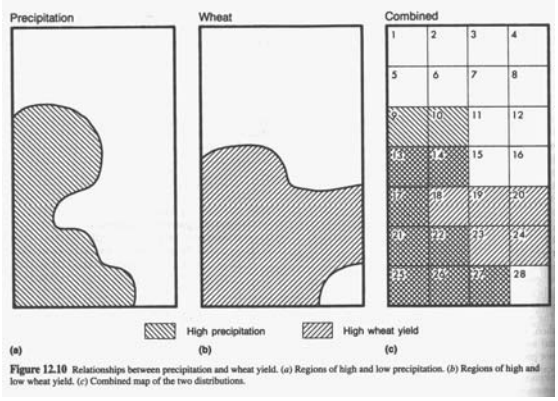



Figure 12.10 Relationships between precipitation and wheat yield. (*a*) Regions of high and low precipitation. (*b*) Regions of high and low wheat yield. (*c*) Combined map of the two distributions.

## Example Test of Spatial Pattern

- Is there a relationship between the distribution of rainfall and the wheat yield in the area shown?
- NULL HYPOTHESIS: There is no relationship
- ALTERNATIVE HYPOTHESIS: There is a relationship

## Chi-square

- Make assumption that there is no relation between maps A and B
- Compute statistics that allow the assumption to be rejected
- Chi-square is the sum of the (Observed value – Expected value)^2/Expected value
- Can check value against table for actual likelihoods

## Calculating Chi-Squared : text p193
## Observed frequencies

| Right: Wheat Yield<br>Below: Rainfall | High | Low | Total |
|---|---|---|---|
| High | 8 | 2 | 10 |
| Low | 5 | 13 | 18 |
| Total | 13 | 15 | 28 |

## Calculating Chi-Squared: text p193
## Expected frequencies

| Right: Wheat Yield<br>Below: Rainfall | High | Low | Total |
|---|---|---|---|
| High | 5 | 5 | 10<br>(35.7%) |
| Low | 8 | 10 | 18<br>(64.3%) |
| Total | 13 | 15 | 28 |

e.g. High yield is 10/28 cells = 35.7% times the total high yield of 13 = 5

## Calculating Chi-Squared : text p193
## Observed and Expected frequencies

| Right: Wheat Yield<br>Below: Rainfall | High | Low | Total |
|---|---|---|---|
| High | 8 (5) | 2 (5) | 10 |
| Low | 5 (8) | 13 (10) | 18 |
| Total | 13 | 15 | 28 |

## Calculating Chi-Squared : text p193
## Observed and Expected differences

| Right: Wheat Yield<br>Below: Rainfall | High | Low | Total |
|---|---|---|---|
| High | 8-5=3<br>[9/5] | 2-5=-3<br>[9/5] | 10 |
| Low | 5-8=-3<br>[9/8] | 13-10=3<br>[9/10] | 18 |
| Total | 13 | 15 | 28 |

## Chi-squared

Chi-square = $\Sigma[(O-E)^2/E]$

For the example = 9/5 + 9/5 + 9/8 + 9/10
1.8 + 1.8 + 1.125 + 0.9 = 5.625

This value is then compared to a table of chi-squared to See if the value allows us to reject the null hypothesis that the observed values are not those expected based on proportions

## Chi-squared tables

| $\nu$ | 0.995 | 0.990 | 0.975 | 0.950 | 0.900 | 0.100 | 0.050 | 0.025 | 0.010 | 0.005 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.000 | 0.000 | 0.001 | 0.004 | 0.016 | 2.705 | 3.841 | 5.024 | 6.635 | 7.879 |
| 2 | 0.010 | 0.020 | 0.051 | 0.103 | 0.211 | 4.605 | 5.991 | 7.378 | 9.210 | 10.597 |
| 3 | 0.072 | 0.115 | 0.216 | 0.352 | 0.584 | 6.251 | 7.815 | 9.348 | 11.345 | 12.838 |
| 4 | 0.207 | 0.297 | 0.484 | 0.711 | 1.064 | 7.779 | 9.488 | 11.143 | 13.277 | 14.860 |
| 5 | 0.412 | 0.554 | 0.831 | 1.145 | 1.610 | 9.236 | 11.070 | 12.832 | 15.086 | 16.750 |
| 6 | 0.676 | 0.872 | 1.237 | 1.635 | 2.204 | 10.645 | 12.592 | 14.449 | 16.812 | 18.548 |
| 7 | 0.989 | 1.239 | 1.690 | 2.167 | 2.833 | 12.017 | 14.067 | 16.013 | 18.475 | 20.278 |
| 8 | 1.344 | 1.646 | 2.180 | 2.733 | 3.490 | 13.362 | 15.507 | 17.535 | 20.090 | 21.955 |
| | | | | | 4.168 | 14.684 | 16.919 | 19.023 | 21.666 | 23.589 |

Two by two table has four values so three degrees of freedom

Chi-squared of zero is no relationship.
Higher the value the stronger the relationship.

## Conclusion

- Using Chi-squared it is not possible to reject the NULL hypothesis that there is no relationship between wheat yield and high precipitation
- Test statistic fails, but only just
- Use another method?

## Yule's Q

- Divide world into high/low (2 classes)
- Overlay two maps gives four classes
- Count quadrats in the four classes in a 2 x 2 table (with cells a,b,c,d) (i.e. Observed only)
- Q = (ad – bc) / (ad + bc)
- Value lies between -1 and +1
- -1 is perfect inverse relationship, +1 is perfect positive

## Calculating Yule's Q : text p193
Observed frequencies

| Right: Wheat Yield<br>Below: Rainfall | High | Low | Total |
|---|---|---|---|
| High | 8 (a) | 2 (b) | 10 |
| Low | 5 (c) | 13 (d) | 18 |
| Total | 13 | 15 | 28 |

## Calculating Q

- Q = (ad – bc) / (ad + bc)

(8 x 13) – (2 x 5)

---------- = 94/114 = 0.82

(8 x 13) + (2 x 5)

Close to +1, so can conclude that there is a positive relationship

## Testing spatial relationships

- Is there a relationship between geographical location and the price of gas?
- Are apartment rents less as distance from the campus increase?
- Are grocery store prices higher in poorer areas?
- Are the increased cancer death rates in a district caused by water contamination?
- Is there a relationship between hydrocarbon emissions and decreased upper atmosphere ozone in the polar regions?

## Summary

- Distributions can be quantified, using NNS or other means
- Maps can be compared using Chi-squared, Yule's Q etc.
- Allows cartometry of higher order structures on maps: shape, distribution, arrangement and pattern