DISCOVERING

DESCRIBING

UNDERSTANDING

SPATIAL-

TEMPORAL

PATTERNS

OF

DISEASE

USING

DYNAMIC

GRAPHICS

WILLIAM F. EDDY

AUDRIS MOCKUS

Carnegie Mellon University

DATA

NNDSS

CSTE

MMWR

DETAILS

57 DISEASES

WEEKLY BY STATE

1980 — 1994

783 (REPORT) WEEKS

MIASMA

MULTIVARIATE

INTERACTIVE

ANIMATION

SYSTEM

FOR

MAP

ANALYSIS

SPECIAL FEATURES

SPATIAL-TEMPORAL MODELS

ANIMATED MAPS

SIMULTANEOUS VIEWS

Overview

- Structure of observations
  - Dynamic Map data (values in space and time)
  - Averaged data (averages over regions)
- Tools
  - Multiple Views
  - Model fitting
  - Transformations
  - Handling of missing values
  - Superpositions of several quantities

Views

Show a $quantity(latitude, longtude, time)$ where attribute may be color, pattern, ...

- Dynamic map

$$\begin{aligned} quantity &\rightarrow attribute \\ lattitude &\rightarrow x \\ longitude &\rightarrow y \\ time &\rightarrow framenumber \end{aligned}$$

- Static map

$$\begin{aligned} quantity &\rightarrow attribute \\ lattitude &\rightarrow x \\ longitude &\rightarrow y \end{aligned}$$

- Time Series (X-Y plot)

$$\begin{aligned} quantity &\rightarrow y \\ time &\rightarrow x \end{aligned}$$

- Other (e.g., scatterplots, histograms)

Modelling

- How:
  - Smoothing (prediction)
  - ANOVA
  - Other
- Why:
  - To see the model
  - To inspect the fit (residuals)
  - To hide region boundary artifacts
  - To get a qualitative picture

Why to smooth or to interpolate

- Why not?
- To hide region boundary artifacts
- To fit a model
- To give a qualitative picture (without too much or unnecessary detail)

Inspection of the residuals

Dynamic maps of:

- Region (state) effects
- Time (month) effects
- Residuals

Transformations

- Why:
  - To fit the available range of attribute values.
  - To emphasize different features
- What:
  - Power (Linear, Logarithmic)
  - Rank

Aggregation

- Why:
  - Several quantities
  - Several time moments
  - Several spatial locations
- Over:
  - Different quantities
  - Time of observation
  - Location of observation
  - Combination of the above

Aggregation Functions

- Arithmetic (sum, variance)
- Order (minimum, median)
- Selection (section, several sections)
- Composition of all of the above

Display of Missing Values

- Leave out
  - Use neutral attribute
  - Use background attribute
- Fill in
  - Impute the value from available data
  - Indicate that the value was imputed

Display of Vector Quantities

- Side by side
- Alternate in time
- Use different attributes (color, pattern, height, transparency)
- Combine

INPUT

DATABASE

E.G. COUNTS

INCIDENCE RATES

FITTED VALUES

RESIDUALS

ETC.

Mumps Example

- Available Data.
  - Numbers per month-state for 1968-1988.
  - State population in 1970 and 1980.
- Objective
  - Detect spatial and temporal patterns
- Important Features.
  - High in spring and low in autumn.
  - Dramatic decrease due to vaccination.
  - Large outbreak in 1986-1987.

Lets look at the data:

slide of a map

Raw Data Animation

(Can not wait 21 years, so

one month takes 2/3 of a second,

human life would take 10 min).

Spatial temporal trends obscured by noise

Why state boundaries?

How to model the data?

Show Raw data:

Explain Display:

USA map,

color scale,

background missing,

transformation (transformed by ranks to use color scale more uniformly),

main features:

seasonal (spring to autumn)

High in the beginning to low in the end

Outbreak in the 86-87

Smooth Data Animation

A model Easier to perceive