Stat645

Linked brushing

Hadley Wickham

Basic brushing

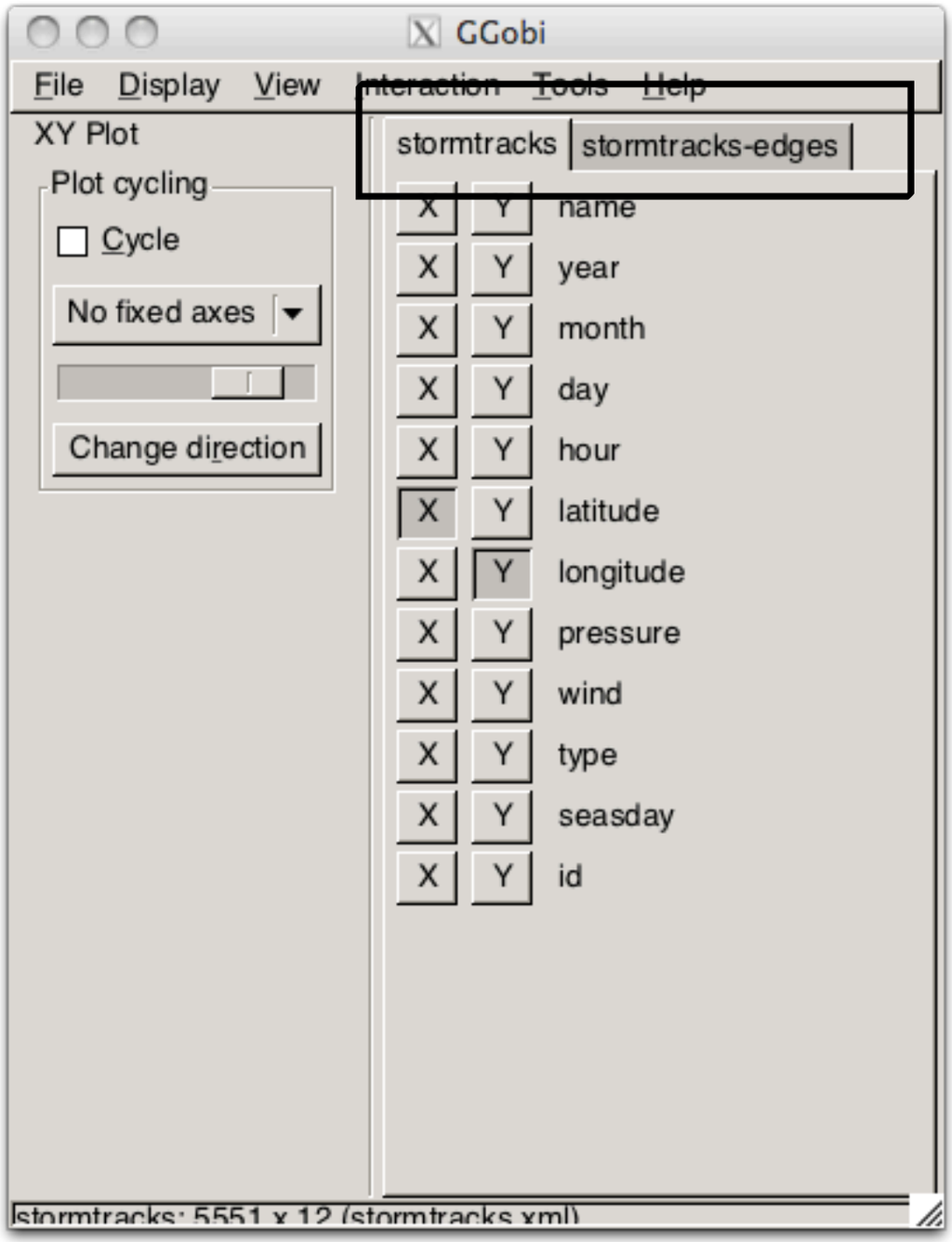
- Open places.csv in ggobi
- How is climate related to location?
- Are there clusters in the data (excluding location?)
- Are nearby cities similar?

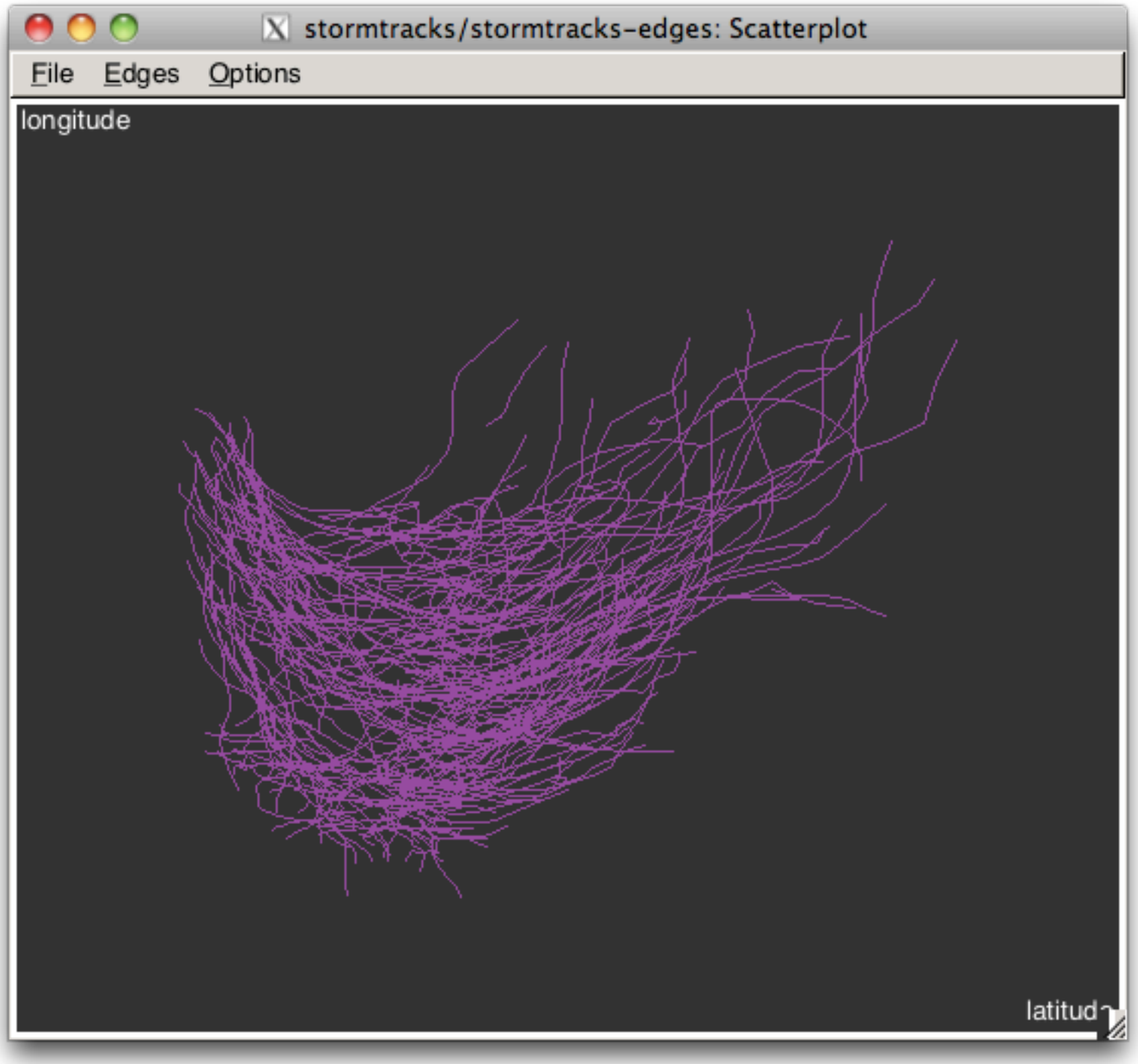
http://www.jstor.org/pss/2685098

Stormtracks

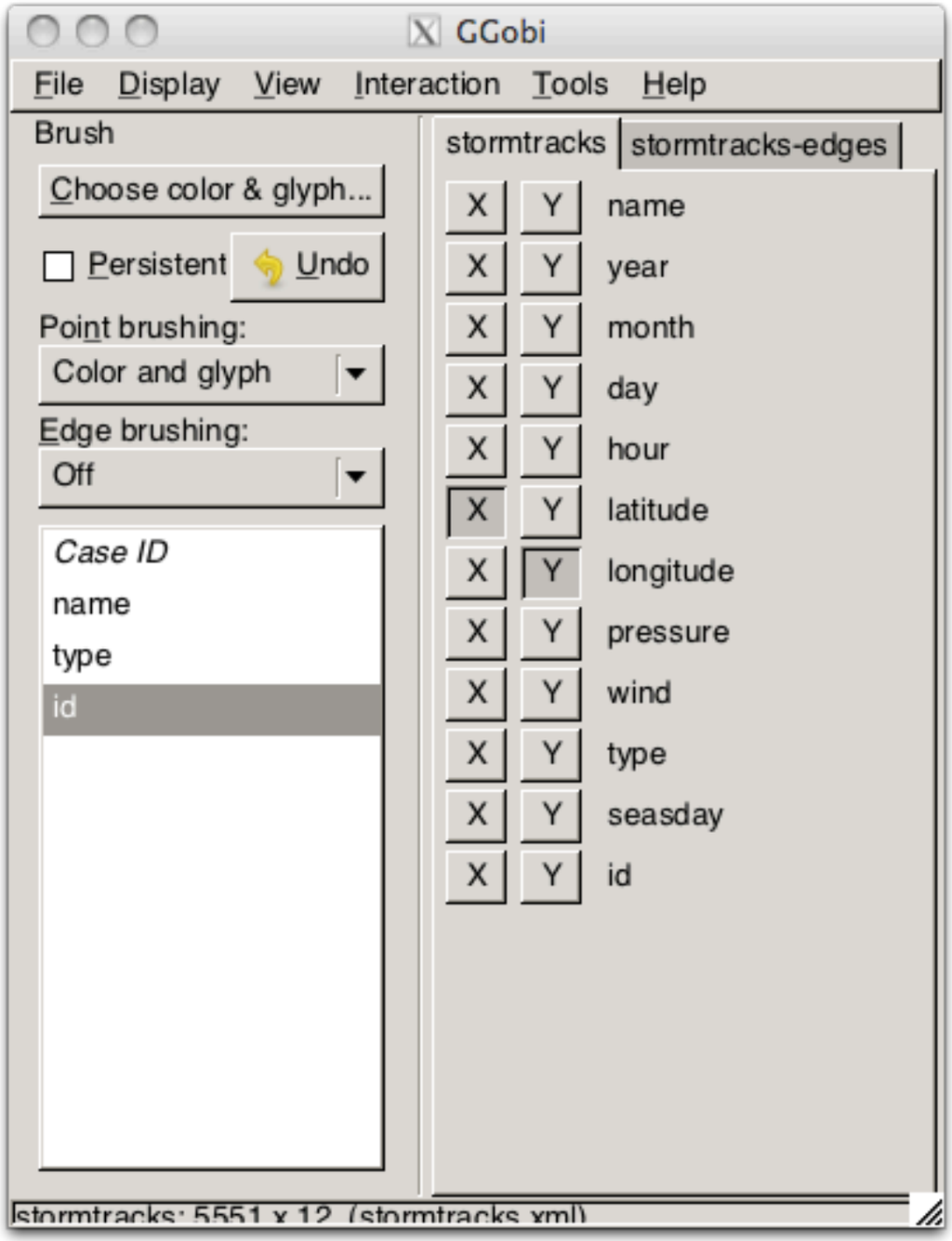
Each row in stormtrack-edges represents one storm

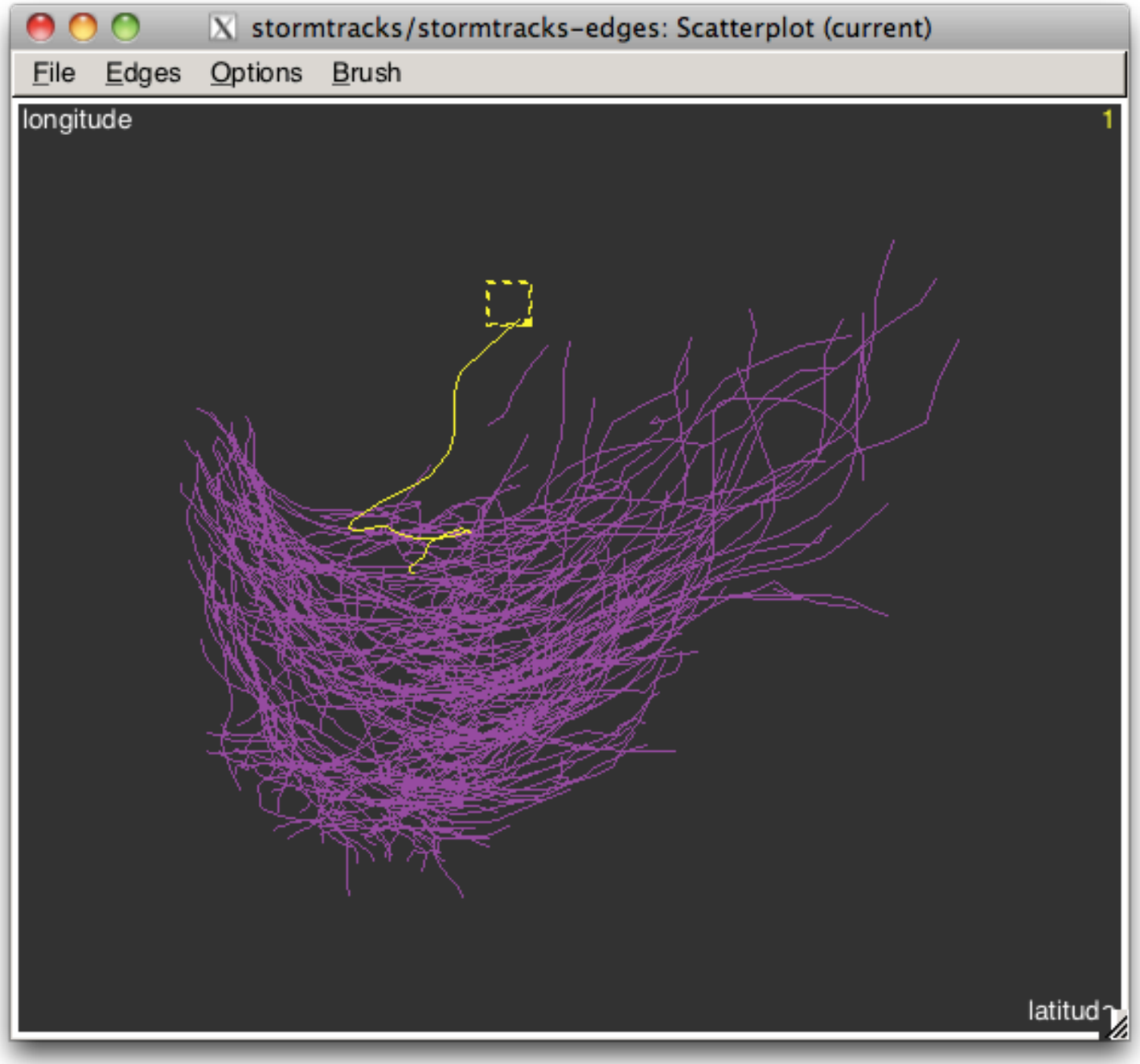
Each row in stormtrack represents a measurement at one time point



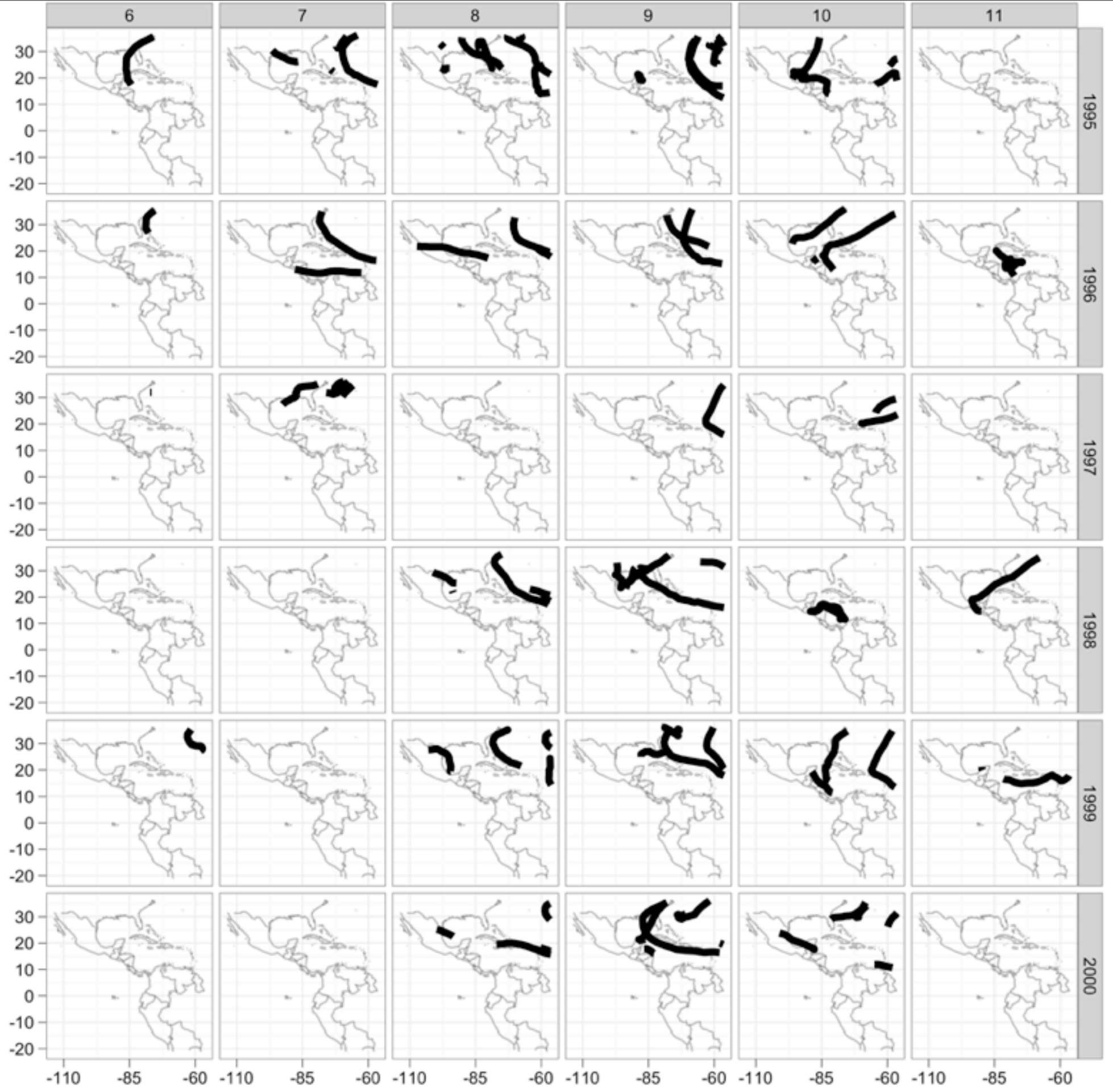


Turn points off and edges on





Brush by storm id to highlight entire stormtrack


6
7
8
9
10
11
1995
1996
1997
1998
1999
2000
30
20
10
0
-10
-20
-110
-85
-60

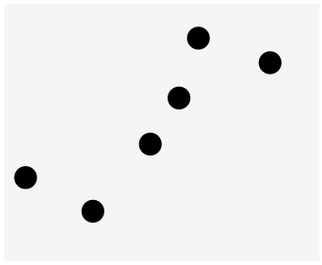
Questions

Do storms in different seasons take different paths?

Do storms that start in the same place finish in the same place?

Is speed or pressure related to location?

Meifly



Observation

Obs ID

Original data

Model-observation summaries

1

many

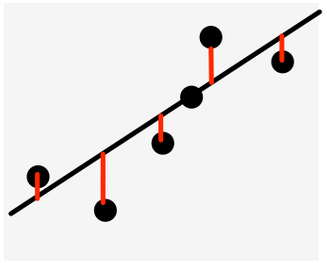
Model-Observation

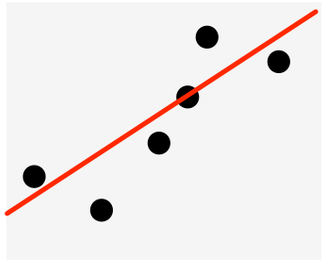
Obs ID

Model ID

Diagnostics

Fit quality





Model

Model ID

Model fit statistics

1

many

Model-Estimate

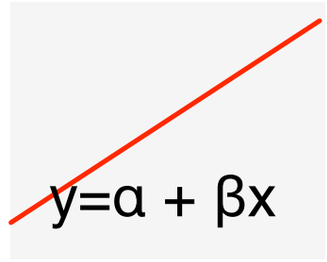
Model ID

Estimate ID

Raw

Standardised

Uncertainty



f(**α**, **β**)

Estimate

Estimate ID

n

model-estimate summaries

1

many

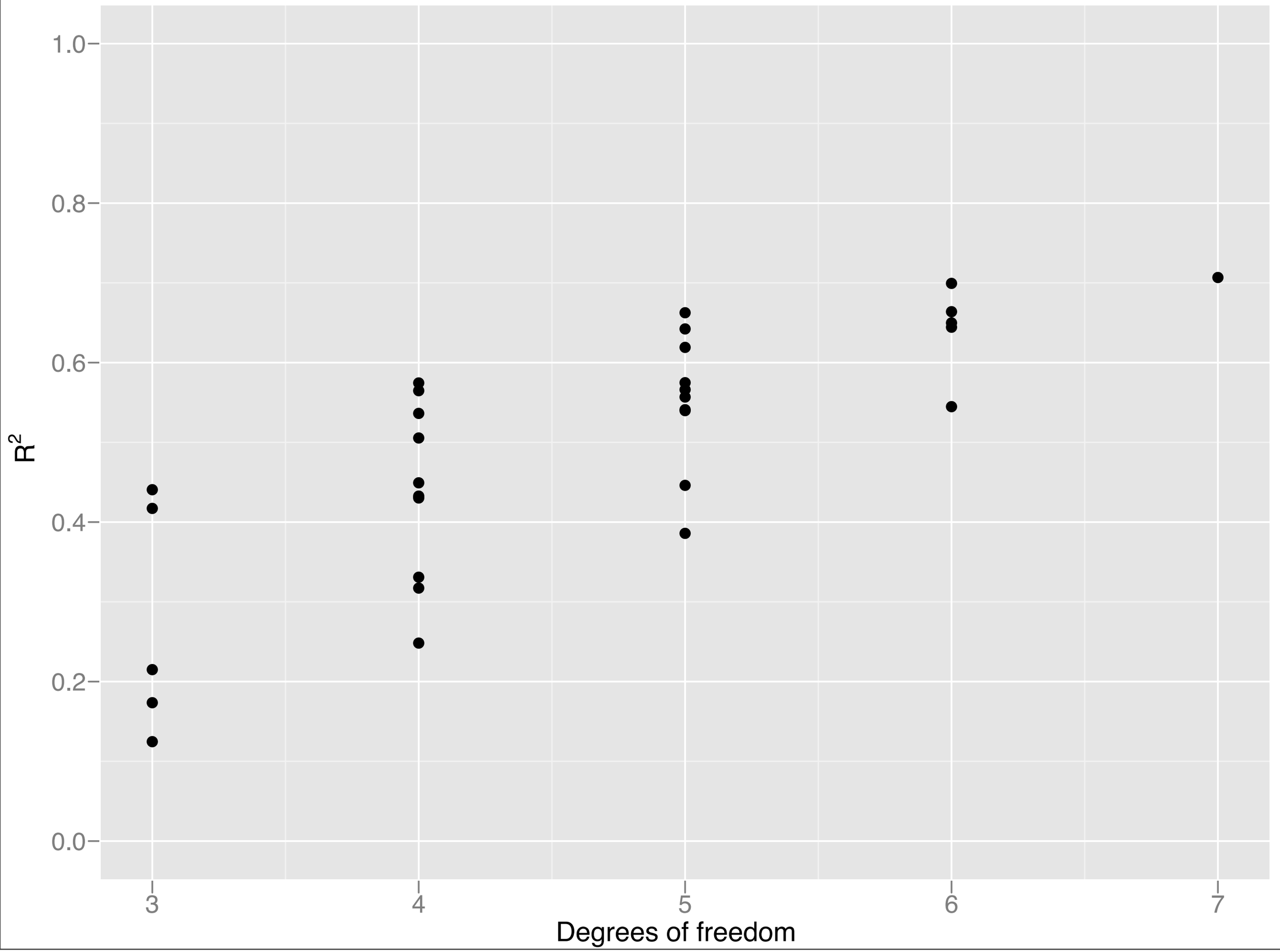
Example

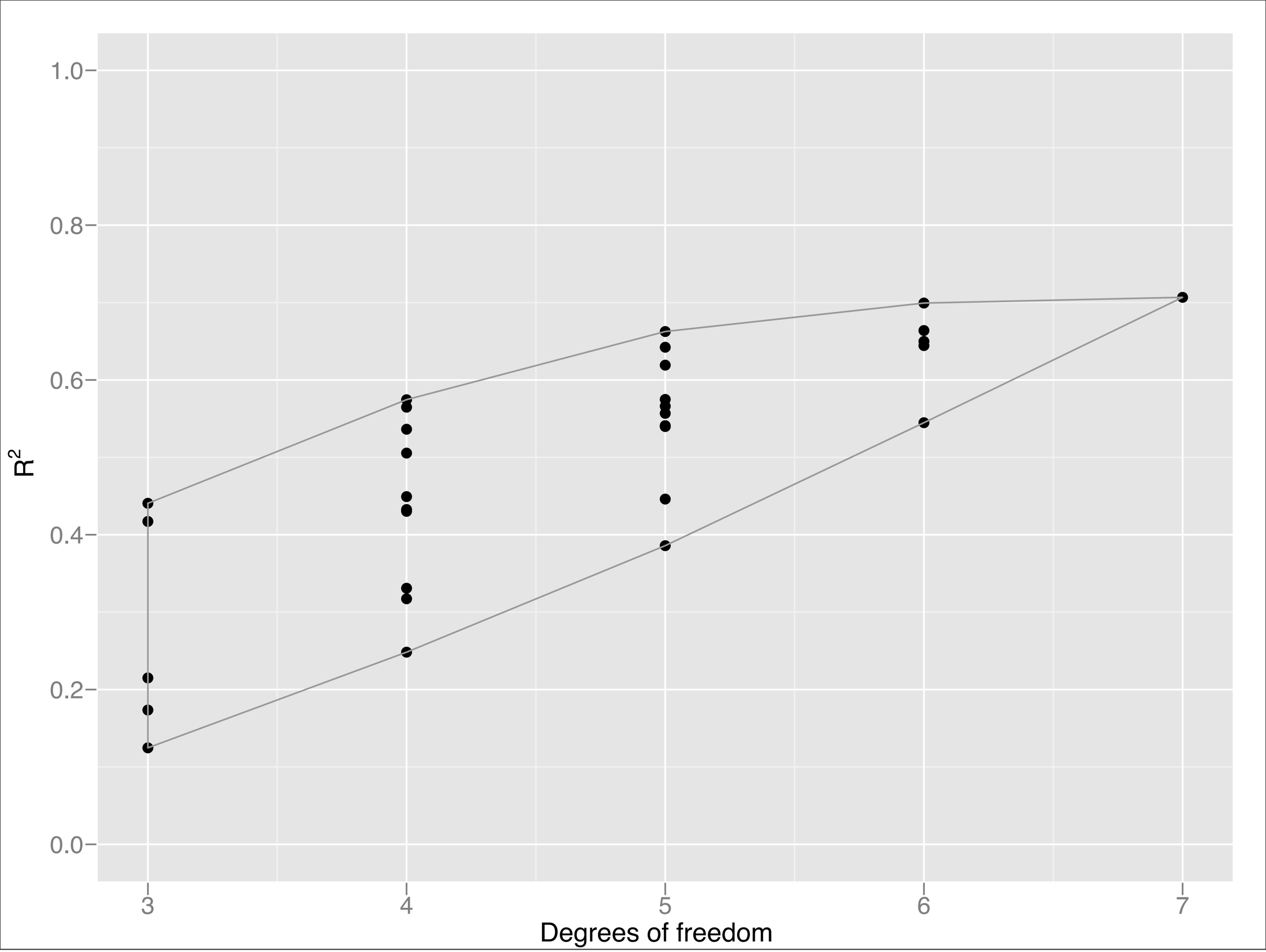
- Data from French-speaking Swiss provinces in the late 1800's
- Want to understand the relationship between fertility and:
  - proportion of agricultural workers
  - performance on army examination
  - higher education
  - proportion of Catholics
  - infant mortality

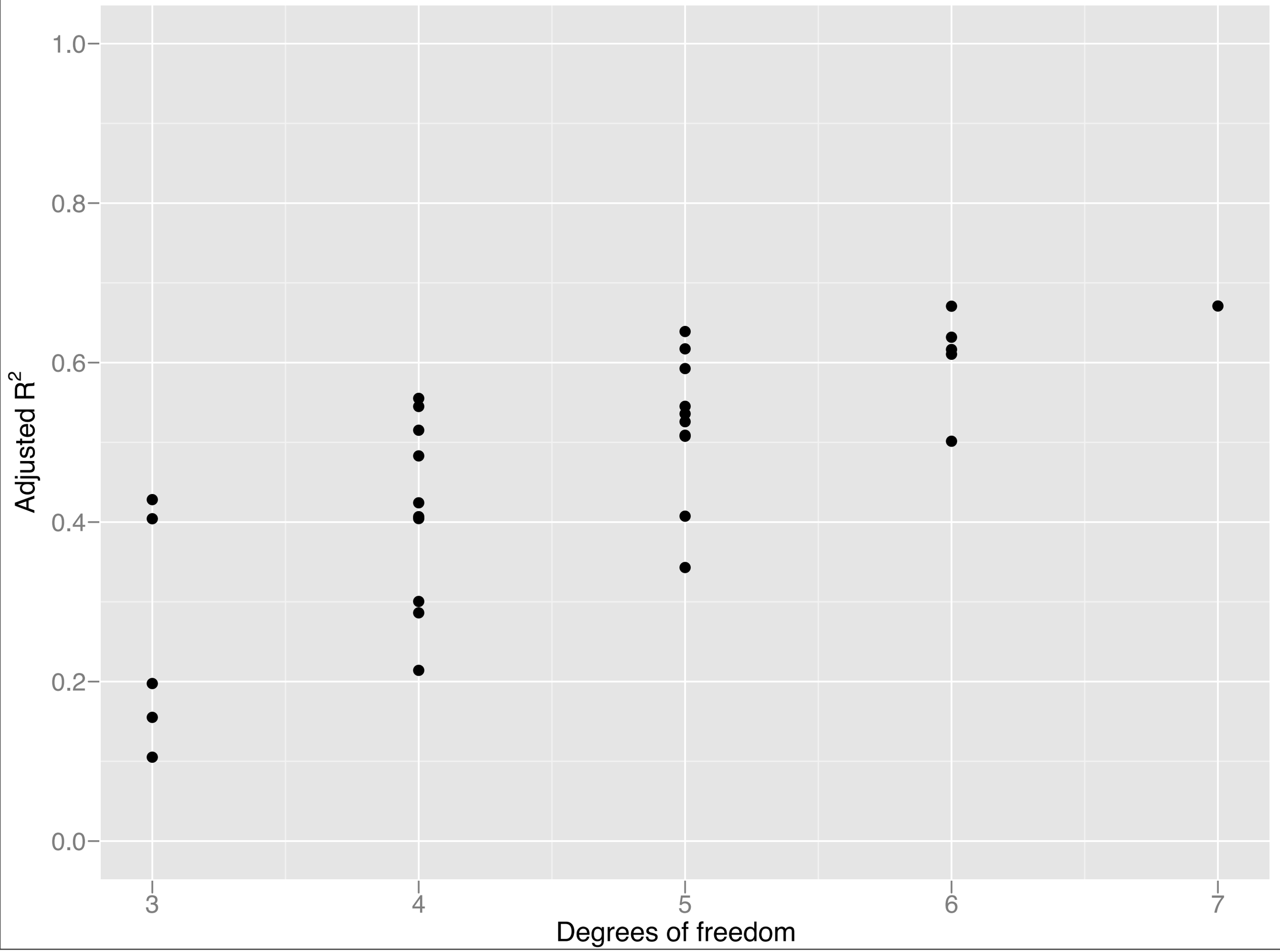
Focus on understanding, not prediction

Model level

- Fit all $2^5 - 1 = 31$ possible linear models
- Summarise with:
  - degrees of freedom
  - R^2, adjusted R^2
  - Log-likelihood, BIC, AIC

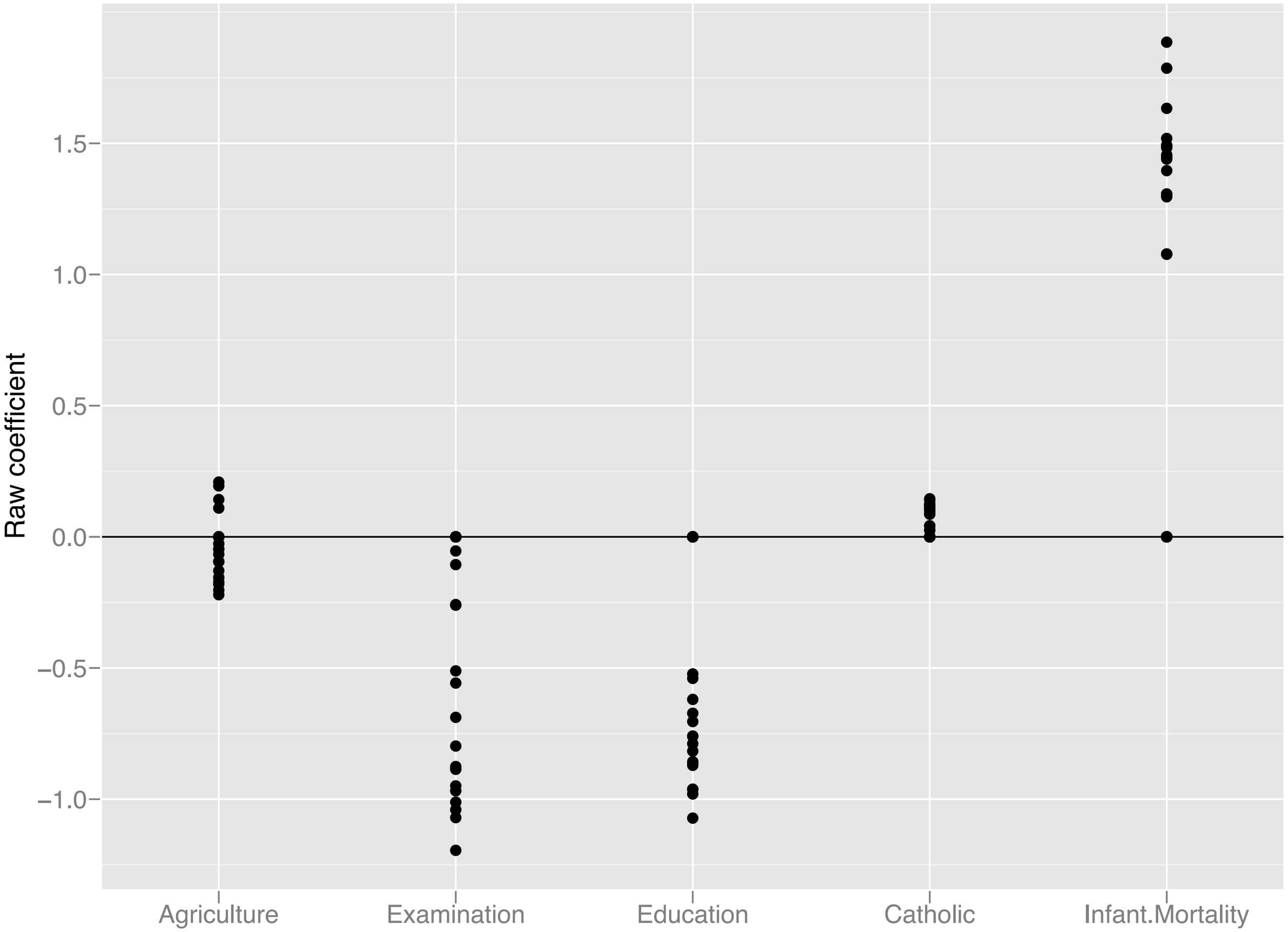

1.0
0.8
0.6
0.4
0.2
0.0
R2
3
4
5
6
7
Degrees of freedom

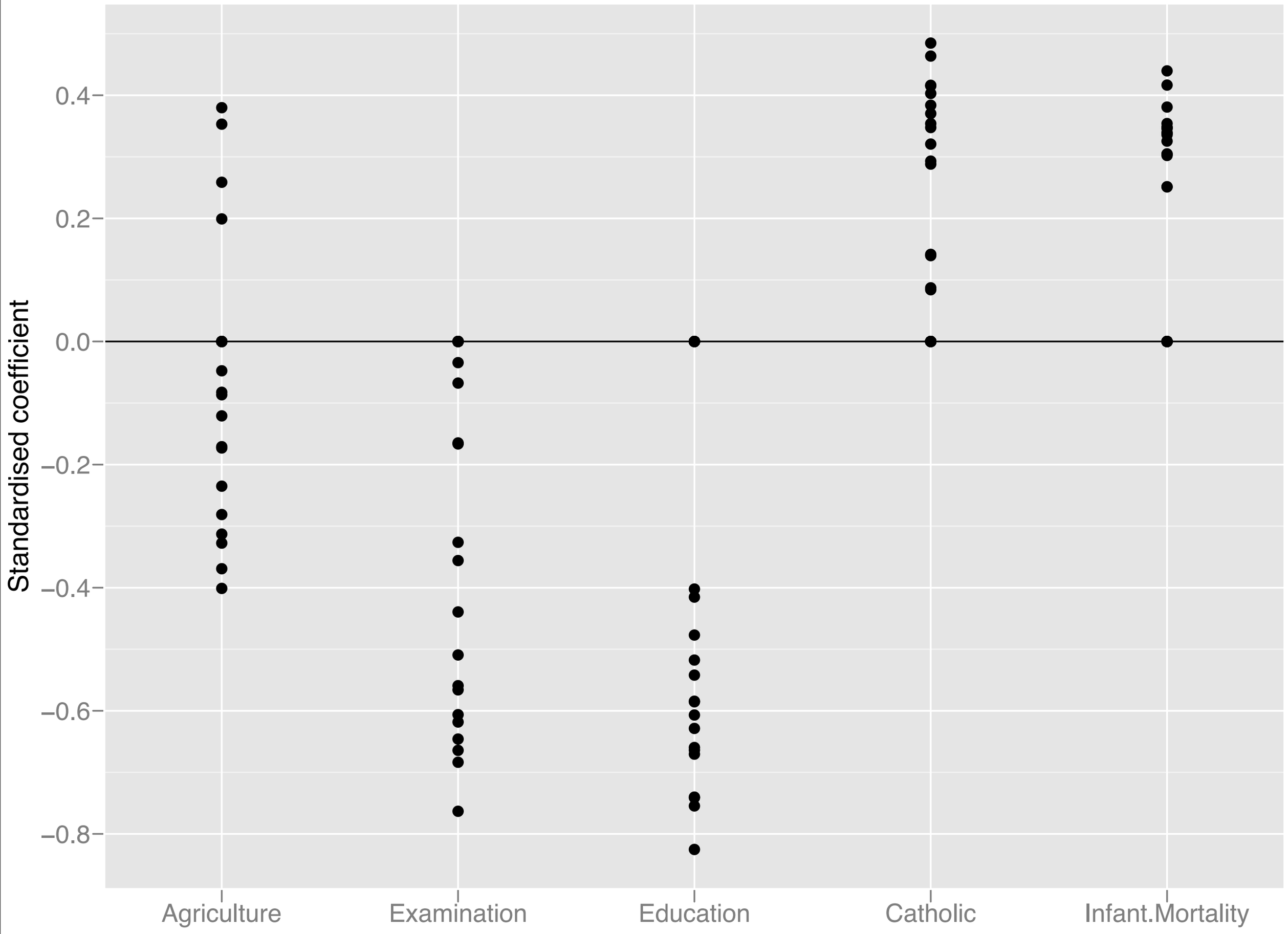

1.0
0.8
0.6
0.4
0.2
0.0
R^2
3
4
5
6
7
Degrees of freedom

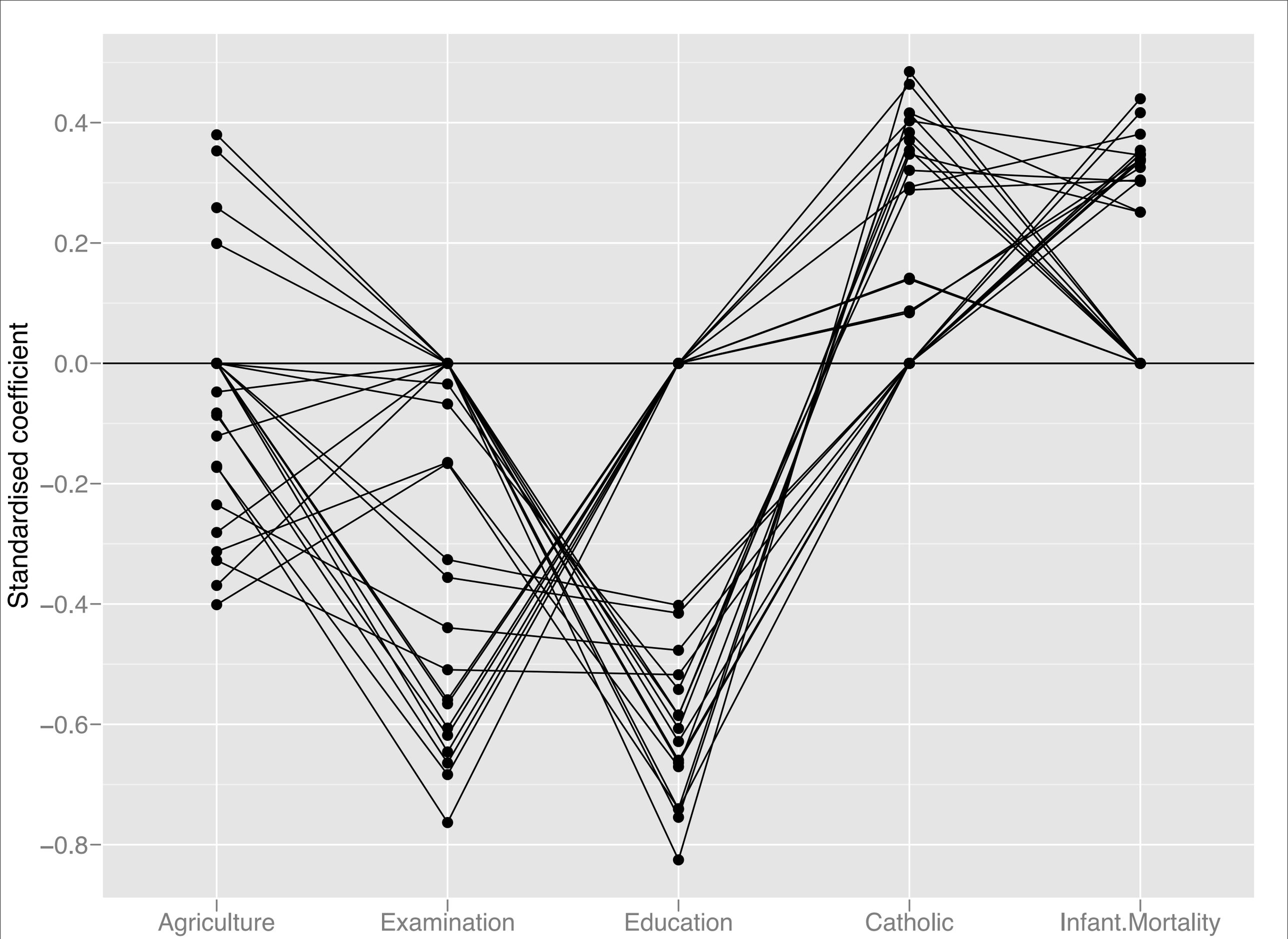

1.0
0.8
0.6
0.4
0.2
0.0
Adjusted R^2
3
4
5
6
7
Degrees of freedom

Model-estimate level

- Raw and standardised estimates
- Standard error
- t-value, absolute t-value
- Explore variance-covariance matrix of predictors


Raw coefficient
1.5
1.0
0.5
0.0
−0.5
−1.0
Agriculture
Examination
Education
Catholic
Infant.Mortality


Standardised coefficient
0.4
0.2
0.0
−0.2
−0.4
−0.6
−0.8
Agriculture
Examination
Education
Catholic
Infant.Mortality


Standardised coefficient
0.4
0.2
0.0
−0.2
−0.4
−0.6
−0.8
Agriculture
Examination
Education
Catholic
Infant.Mortality

GGobi

- Open swiss-meifly.xml
- Recreate previous plots in GGobi
- Pay attention to the linking variables