Exploring Lightning and Fire Ignition Data Using Data Sharpening Techniques

Workshop on Forest Fires and Point Processes
Fields Institute, Toronto, ON.
May 24 – 24, 2005.

Douglas G. Woolford, W. John Braun & Reg J. Kulperger

Department of Statistical and Actuarial Sciences
The University of Western Ontario

Outline

- Introduction
- Lightning & Fire Data
- Initial Explorations
- Identifying Lightning Cluster ‘Centres’
- Tracking Storms Through Time
- Conclusions

Introduction

Lightning Causes 35-40% of forest fires in Canada; 80% of the burned area.

Ultimate Goal Probability mapping of lightning caused forest fire incidence.

Current Goal Exploratory analysis of lightning and fire ignition data.

What statistical tools might be useful?

What issues arise in analyzing these data?

Lightning – Detection Systems and Data

- 16 direction finders in Ontario
- Cloud to ground discharges are recorded
- Variables: time, location, polarity, signal strength, and multiplicity (for some years)
- Triangulation of multiply detected discharges
- Detection efficiency = 70%
- Location errors could be as large as 10 km

  (Accuracy depends on proximity to direction finders)
- Strokes with LCCs cannot be identified

Lightning and Fire Ignition Data

Fire data variables

Locations, Start times, Out times, Area burned

Weather data variables

- Daily rainfall, Noon temperature, Wind speed, Wind direction, Relative humidity

Fuel Moisture Codes

- derived from weather data
- e.g. Duff Moisture Code (DMC)

  (moisture in the dead organic layer of the forest floor)

Lightning and Fire Ignition

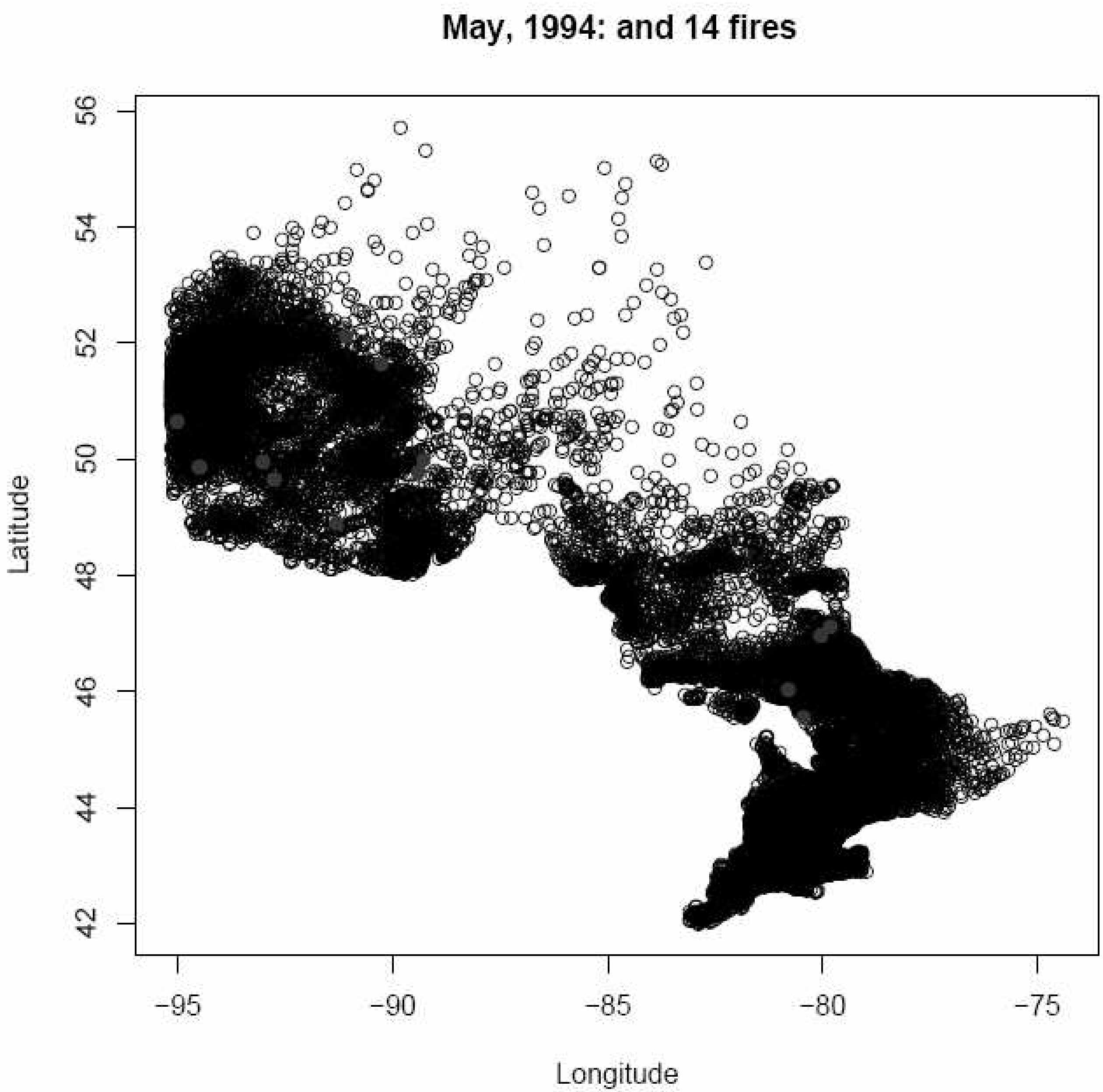
- Lightning-caused fires are caused by LCC strokes
  (LCC = “Long Continuous Current”)
- Anderson (2002, IJWF):
  - Conditional probability of a lightning-caused fire at t given a lightning flash:

$$P_{fire(t)} = P_{LCC}\, P_{ignition}\, P_{survival(t)}\, P_{arrival}$$

  - t could represent several days of smouldering within the Duff layer.

Initial Explorations

- Ontario lightning and fire data available from 1980.
- "Confronted with [such data], the statistician struggles." (R. Peng, 2001)
- e.g. 1994, a test bed for some ideas.
- First rule of data analysis: plot the data.


May, 1994: and 14 fires
Latitude
56
54
52
50
48
46
44
42
Longitude
−95
−90
−85
−80
−75

Initial Explorations (cont'd)

First observations

- Lightning data set is large:
  (157214 observed strikes in July, 1994)
- Fires appear in clusters; none in Southern Ontario
- Lightning clusters are not readily visible from plot (not visible at all in summer months)

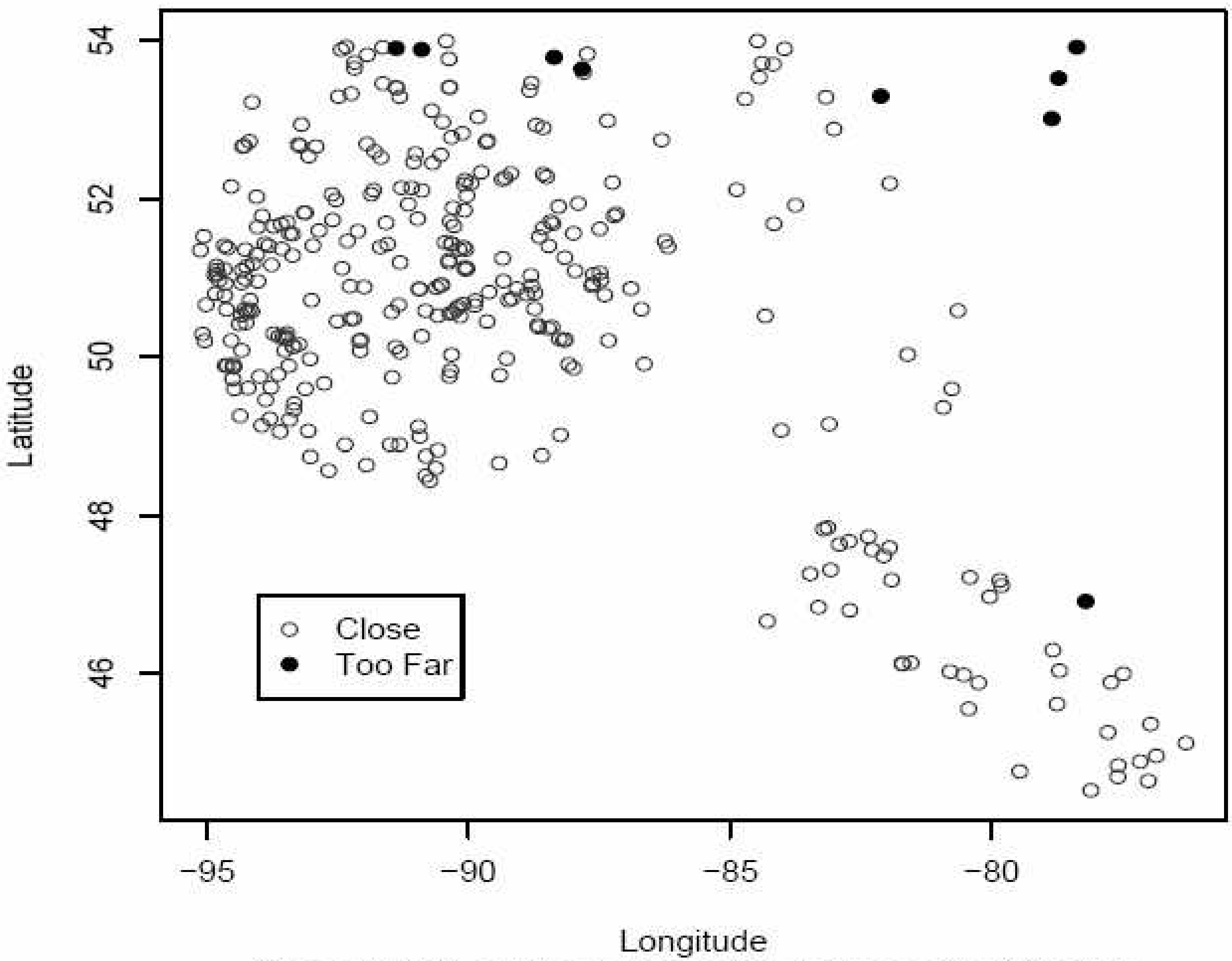
à Difficult to establish a linear association between lightning and fires.

Initial Explorations (cont'd)

- Are there missing data?
- Fire ignitions classified according to whether
  - a lightning strike occurred within 10km (red)
  - within the preceding 12 days
  - or outside a 10 km radius (black)

- 12 days: Smouldering can last for more than 10 days (Flannigan and Wotton, 1990, CJFR).
- 10 km: Approximate max detection system error.

Fire Ignitions Close and Too Far from Strikes − Ontario 1994



(Close: Within a 0.05 degree radius in preceeding 12 days)

Identifying Lightning 'Cluster Centres'

- Data Sharpening (Choi and Hall, 1999, Biometrika):
- original motivation - reduce bias in density estimation

Raw Data : $x_1, x_2, \ldots x_n \sim f(x)$

Estimated Density : $\hat{f}(x) = \dfrac{1}{n}\sum_i K_h(x_i - x)$

- $K_h(x)$ is a symmetric pdf with scale parameter h (the 'bandwidth').

- Sharpened Data: $x^*_1, x^*_2, \ldots, x^*_n$ obtained by local constant regression of x on x :

$$x^*_j = \frac{\sum_i K_h(x_i - x_j)\, x_i}{\sum_i K_h(x_i - x_j)}$$

$$\hat{f}_s(x) = \frac{1}{n} \sum_i K_h(x^*_i - x_j)$$

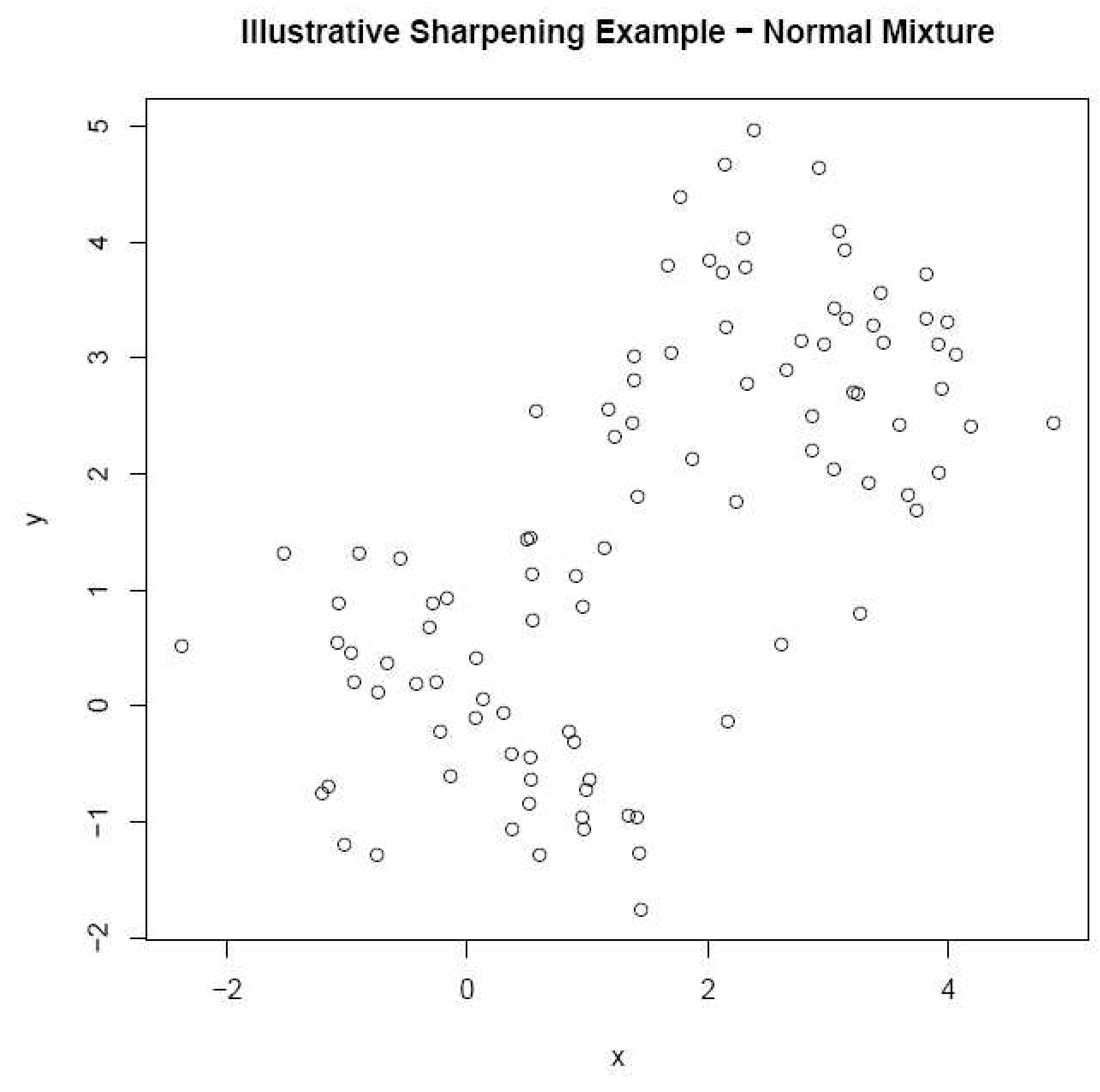
- sharpened density estimate reduces bias in density estimation

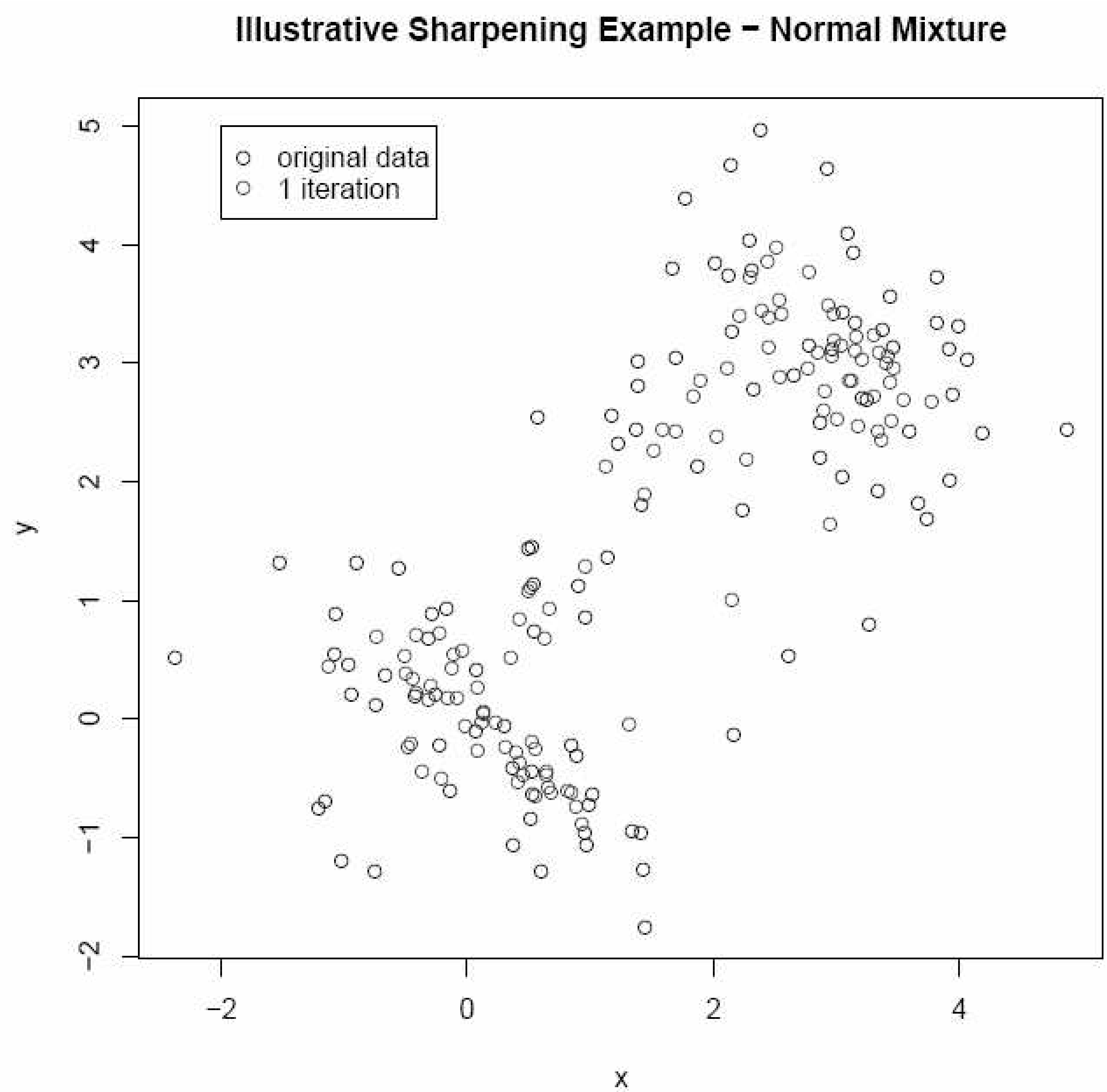
Data-Sharpening (cont'd)

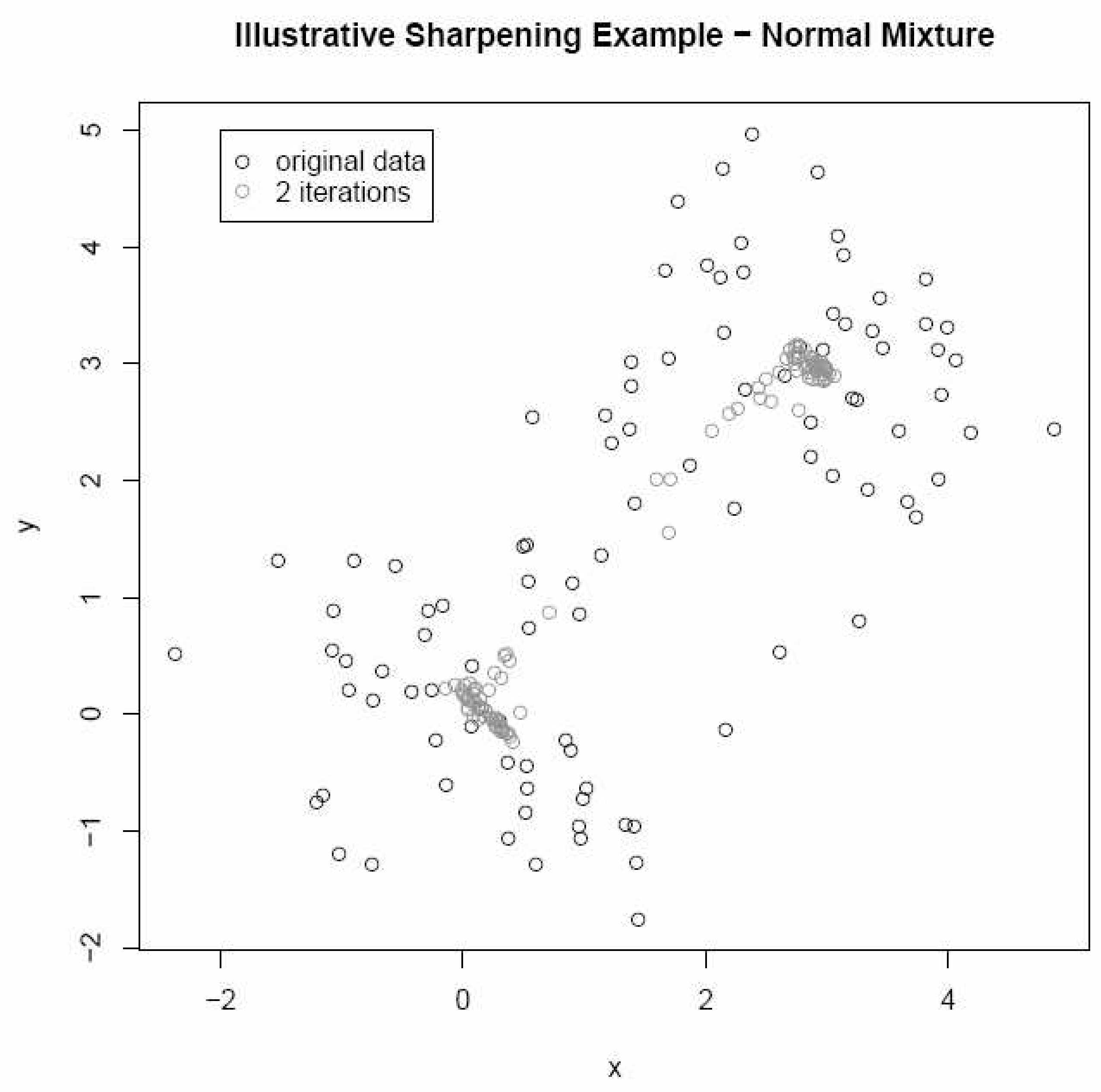
- observations are moved towards local modes
- sharpening procedure can be iterated
  - identical procedure identified earlier as a clustering algorithm (Fukunaga & Hostetler, IEEE 1975)

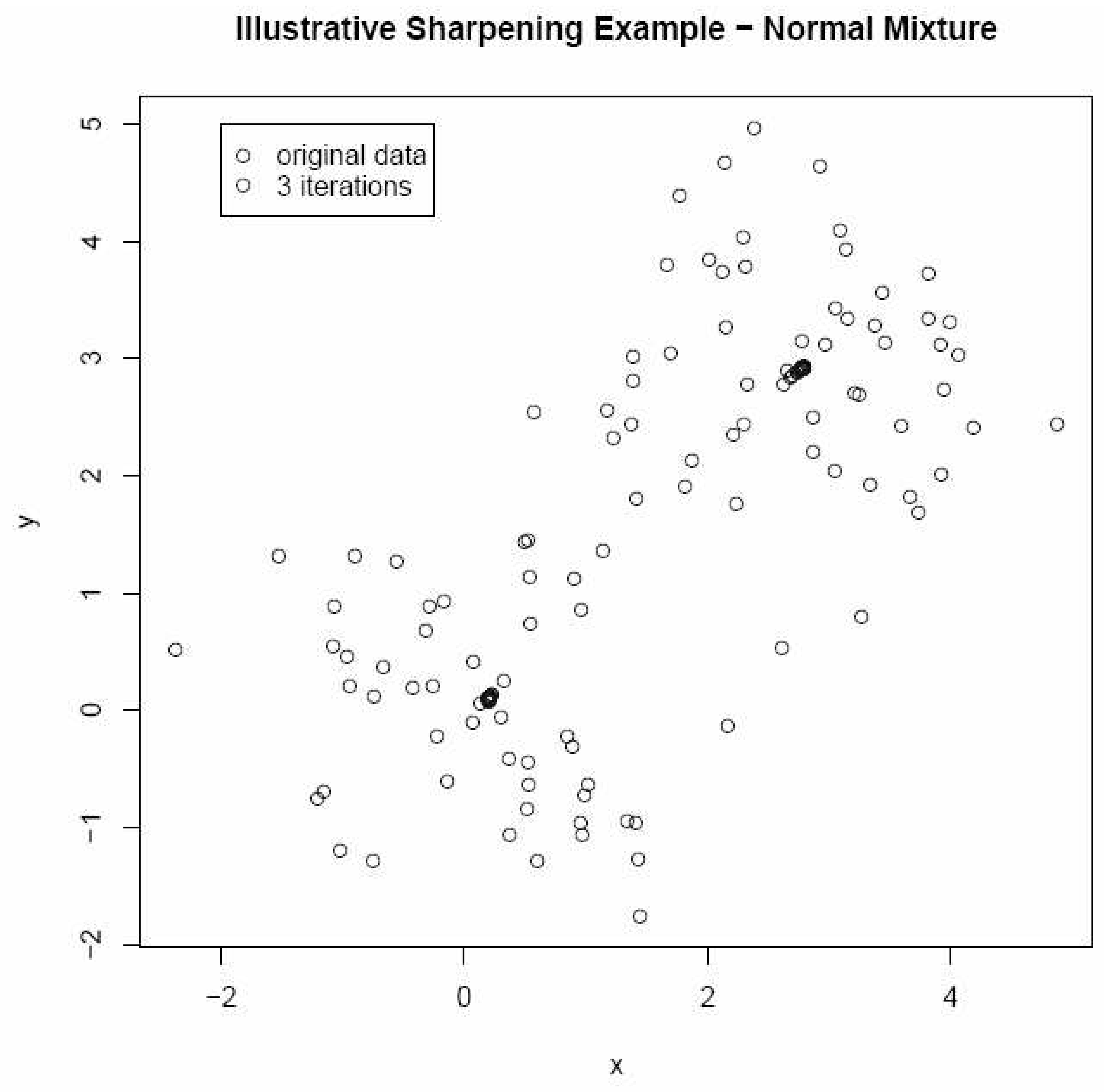
- 2-Dimensional Example
  - mixture of 2 Normal Distributions

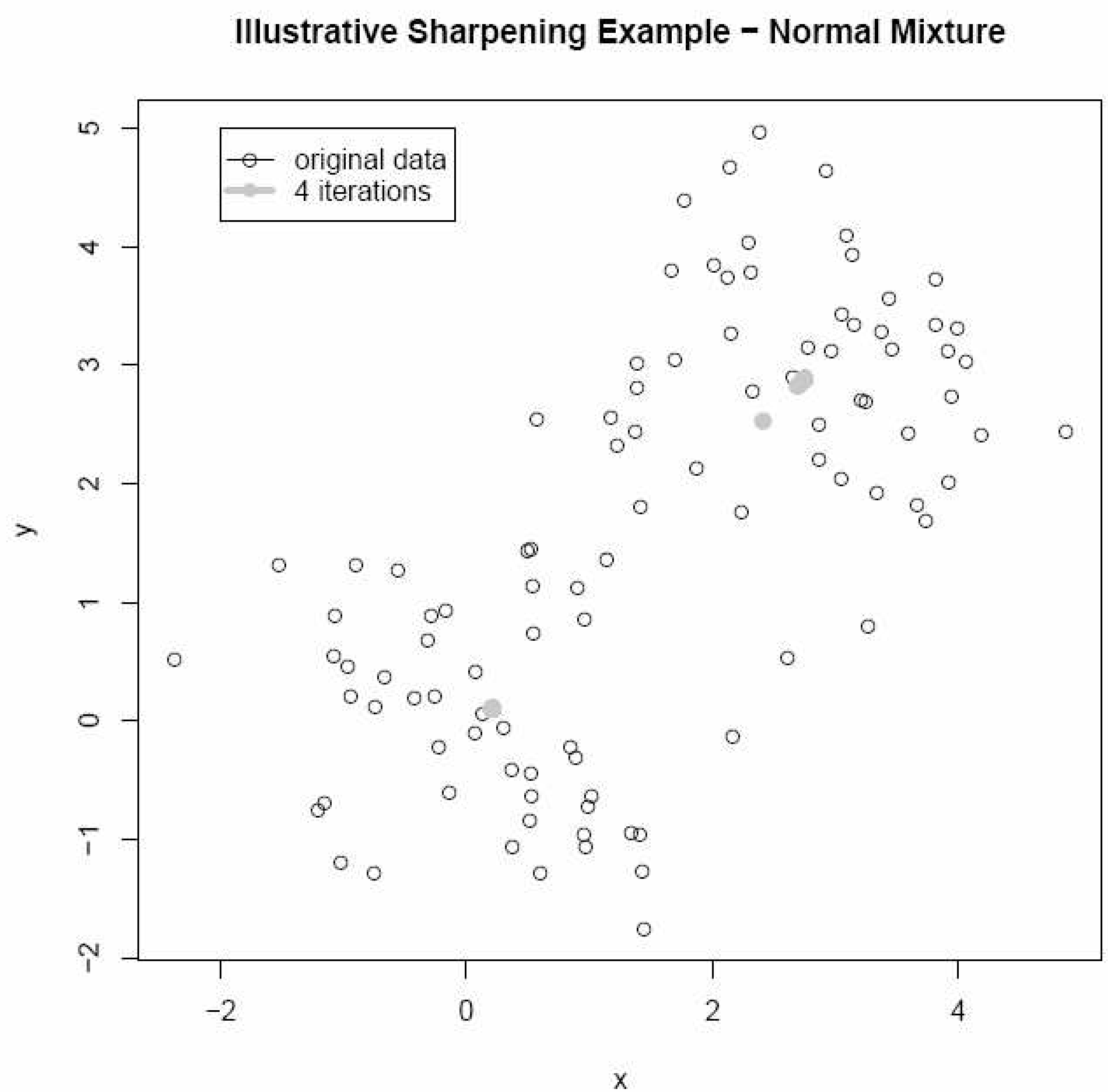
    (means of 0 and 3)
  - iterate to identify local modes

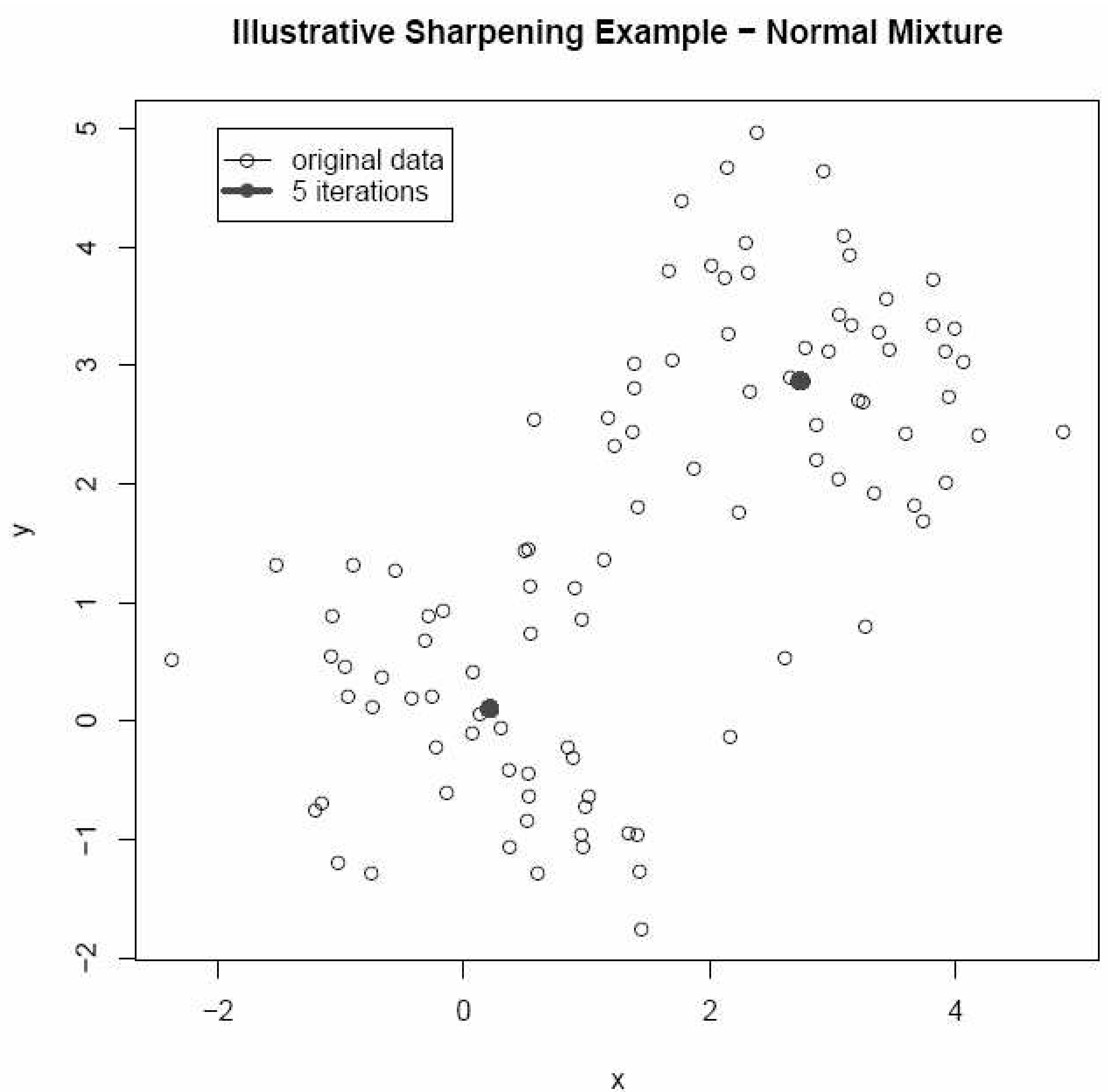

Illustrative Sharpening Example − Normal Mixture
y
5
4
3
2
1
0
−1
−2
−2
0
2
4
x


Illustrative Sharpening Example − Normal Mixture
original data
1 iteration
y
x
−2
0
2
4
−2
−1
0
1
2
3
4
5


Illustrative Sharpening Example − Normal Mixture
original data
2 iterations
y
x
−2
0
2
4
5
3
1
−1


Illustrative Sharpening Example − Normal Mixture
original data
3 iterations
y
x
−2
0
2
4
−2
−1
0
1
2
3
4
5


Illustrative Sharpening Example − Normal Mixture
original data
4 iterations
y
x
−2
0
2
4
−2
−1
0
1
2
3
4
5

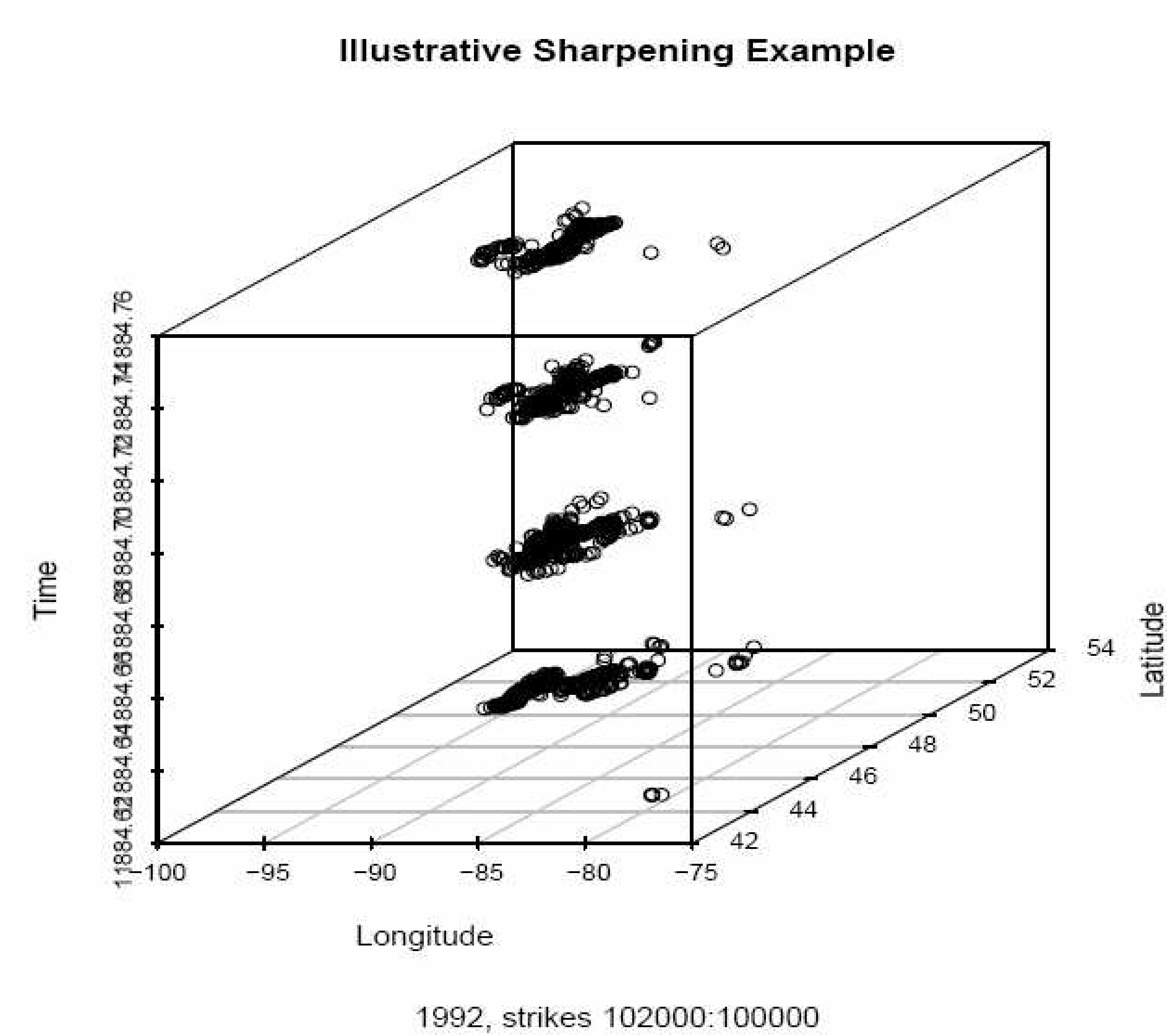


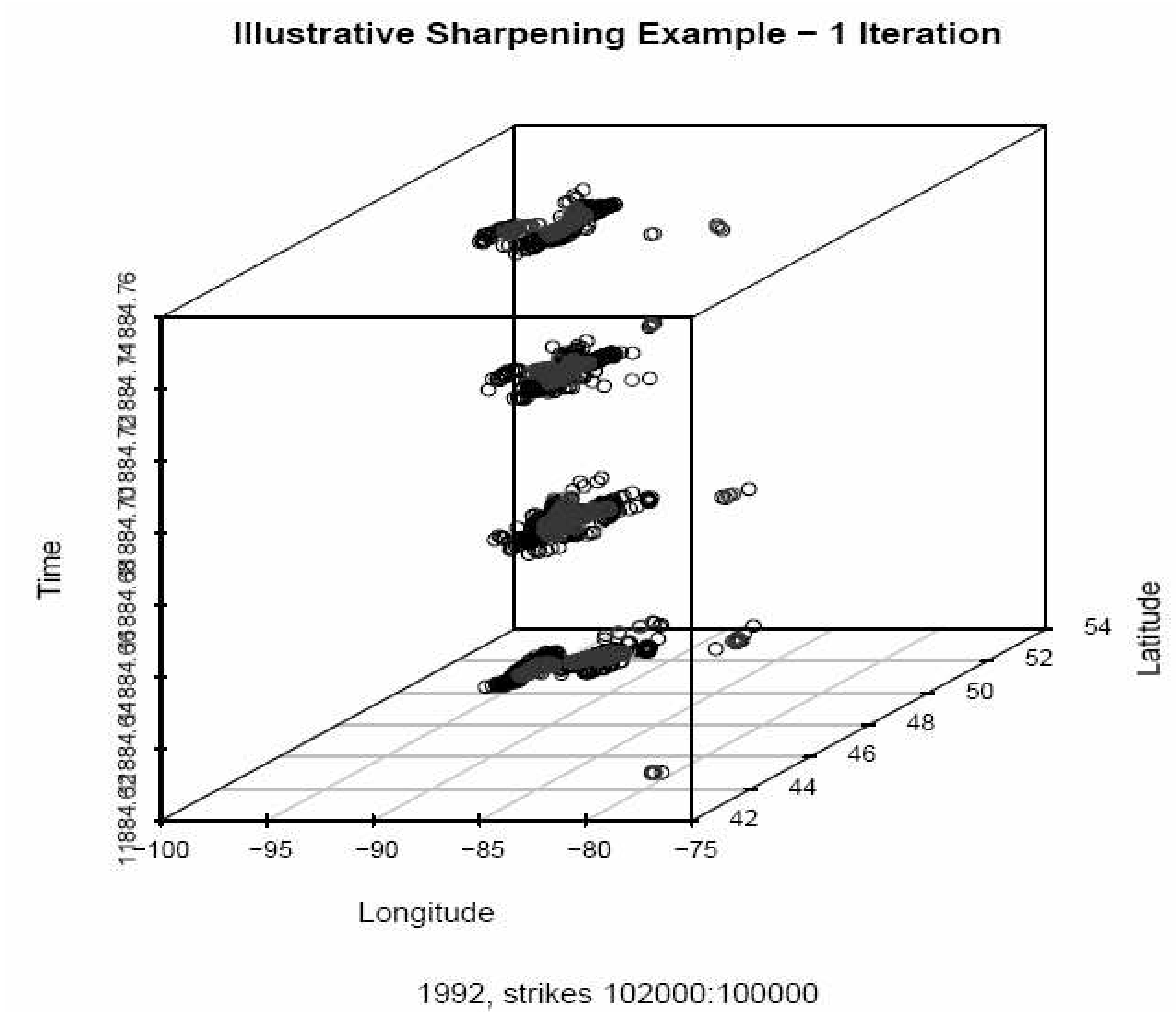
Illustrative Sharpening Example – Normal Mixture
original data
5 iterations
y
x
−2
−1
0
1
2
3
4
5
−2
0
2
4

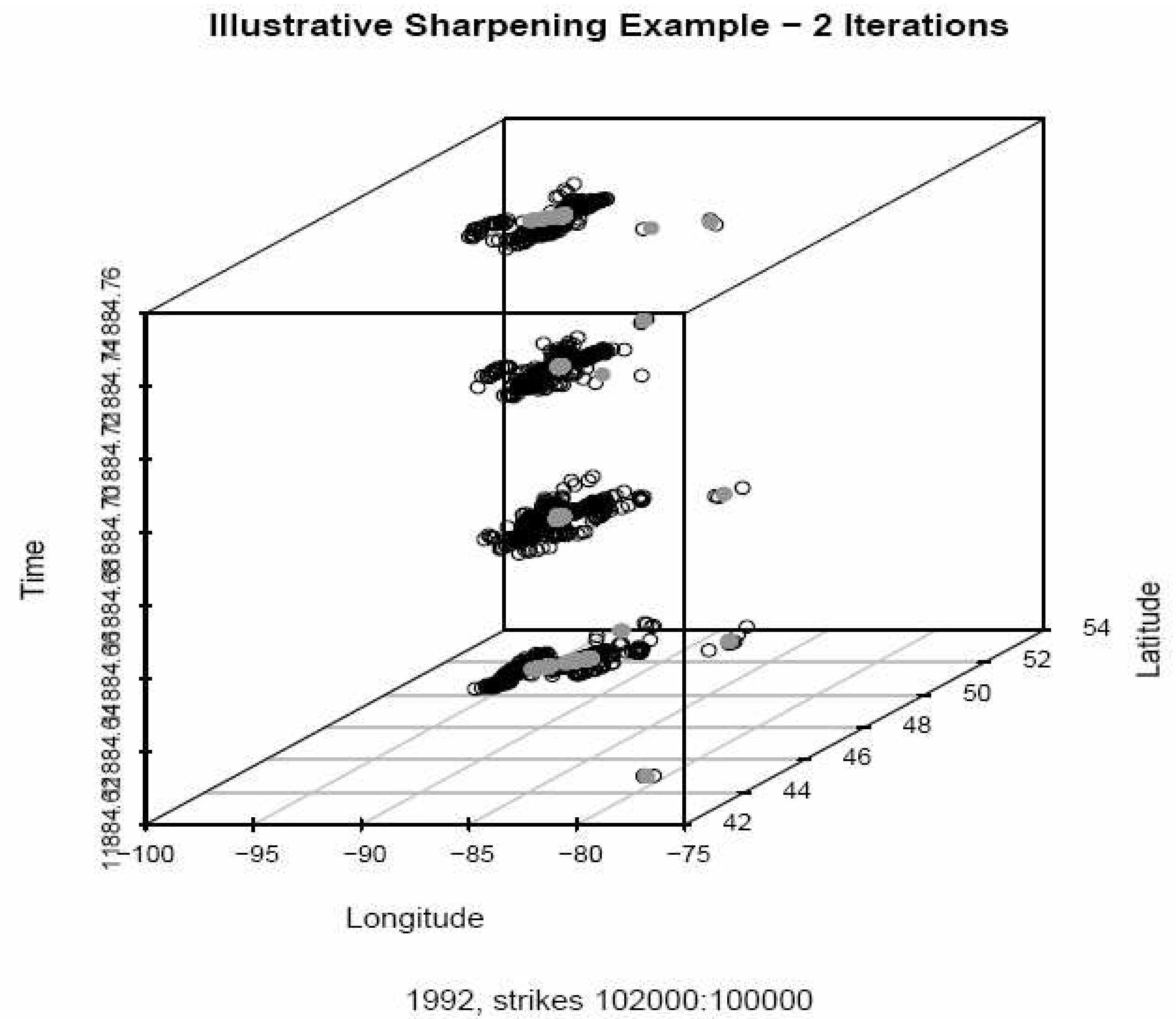
Data Sharpening (cont'd)

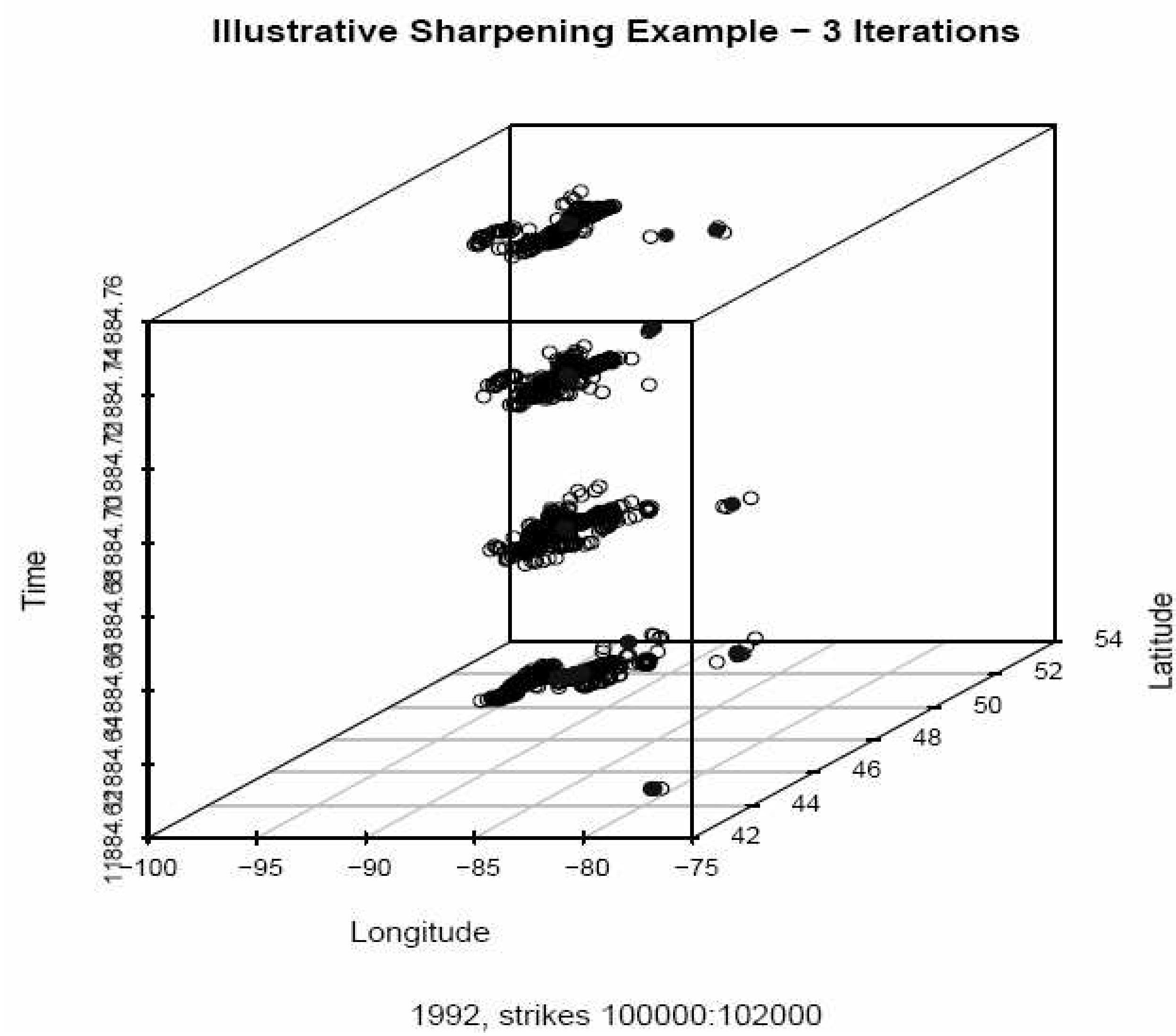
- iteration converges (proof by induction)
- applicable in any number of dimensions
  - bandwidth parameter can vary across dimensions
- data reduction strategy
  - sharpen lightning data in 3-dimensions (space-time)
    - space, bandwidth = 1 degree (about 100 km)
    - time, bandwidth = 1 day

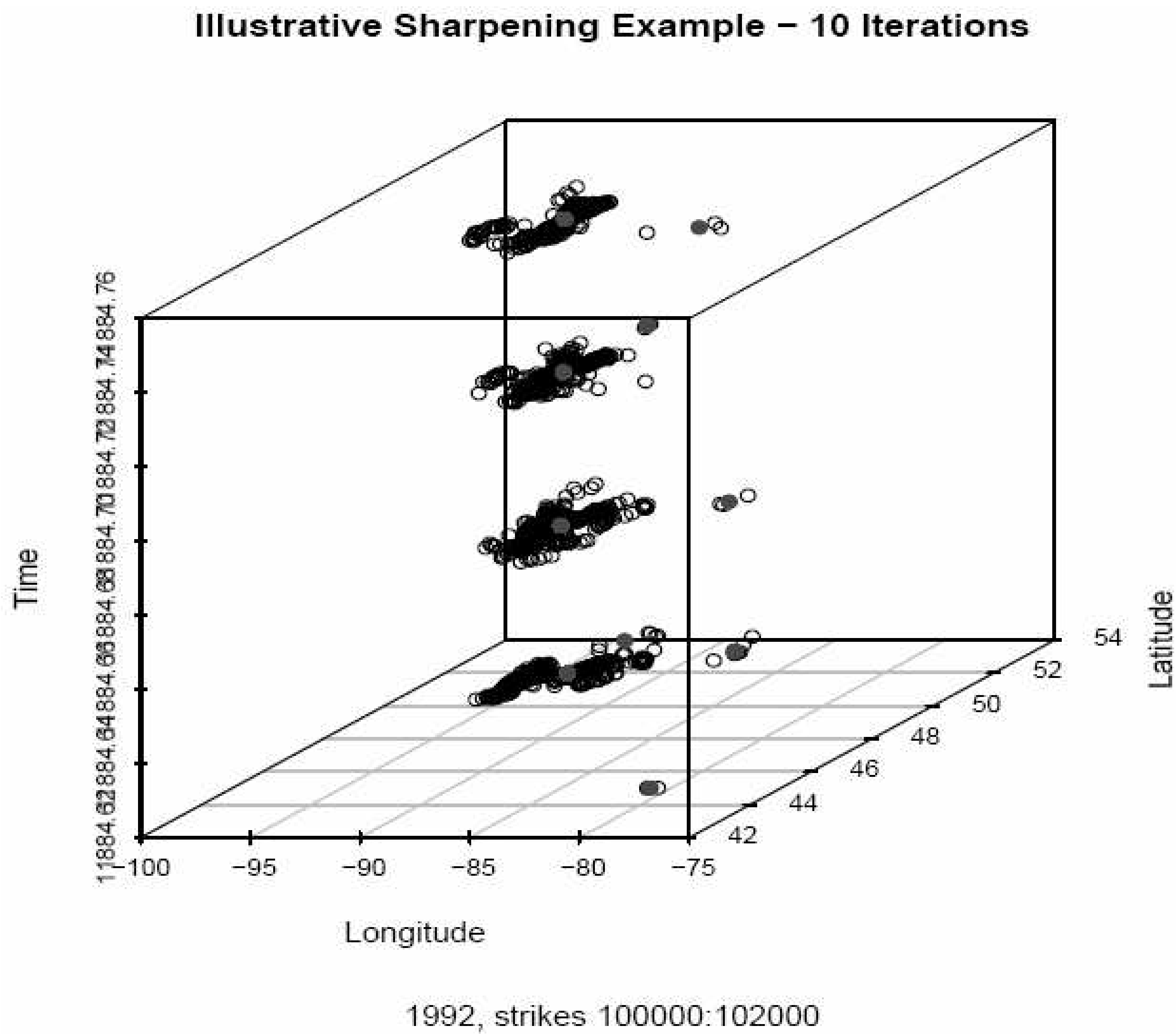
Purpose: Identify 'storm centres' of lightning activity.


Illustrative Sharpening Example
Time
Longitude
Latitude
−100
−95
−90
−85
−80
−75
42
44
46
48
50
52
54
1992, strikes 102000:100000

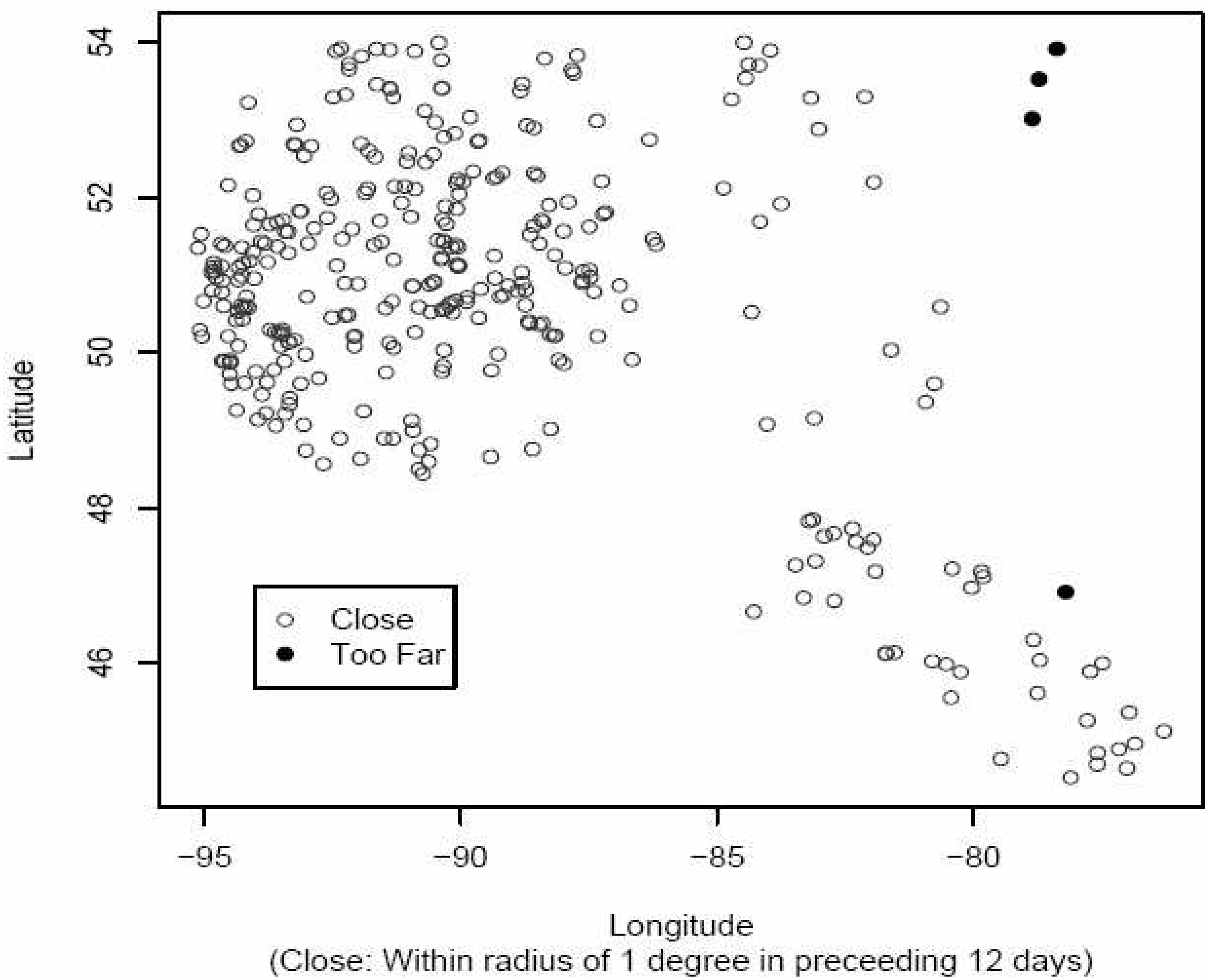

Illustrative Sharpening Example − 1 Iteration
Time
Longitude
Latitude
−100
−95
−90
−85
−80
−75
42
44
46
48
50
52
54
1992, strikes 102000:100000

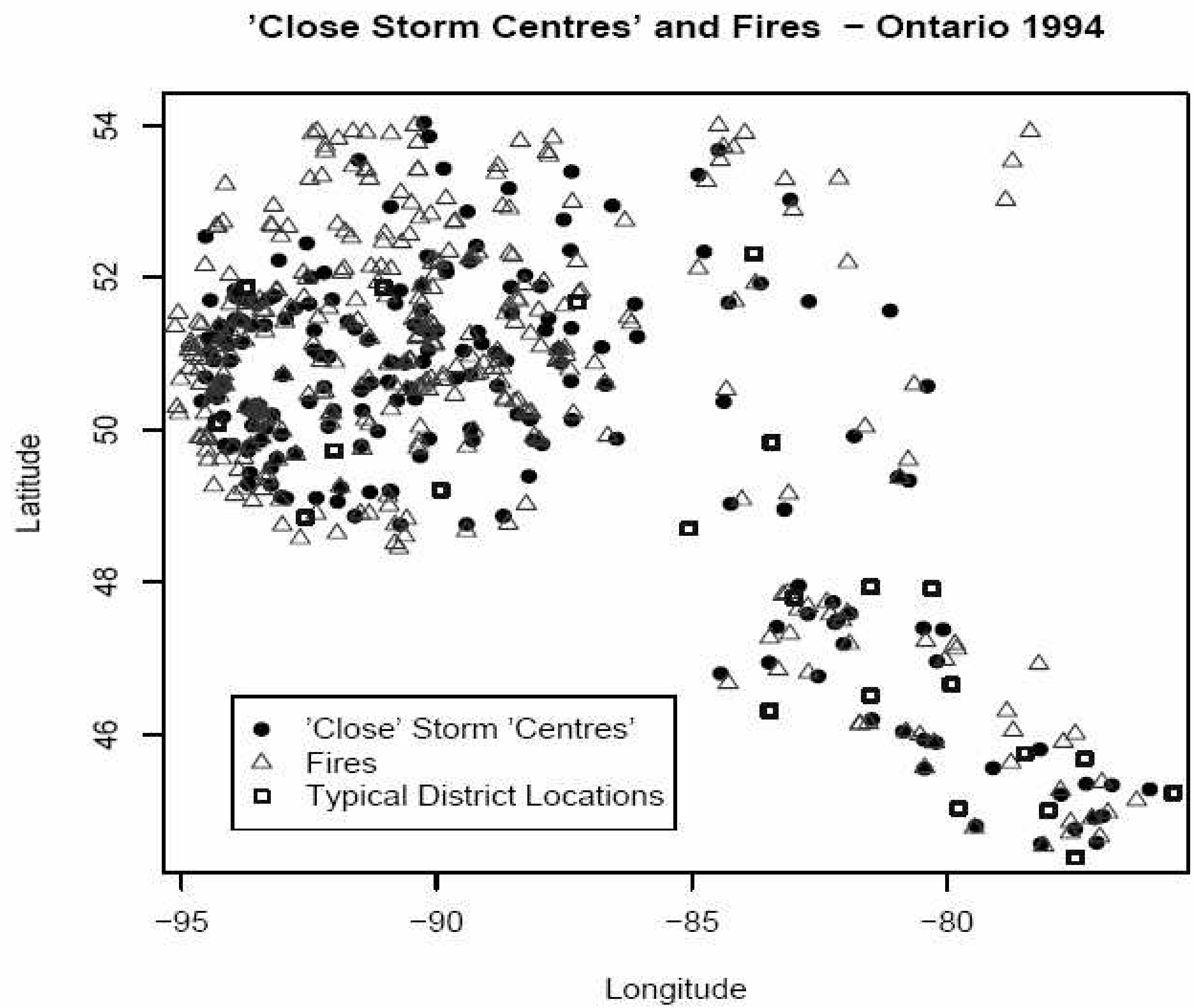

Illustrative Sharpening Example − 2 Iterations
Time
Latitude
54
52
50
48
46
44
42
−100
−95
−90
−85
−80
−75
Longitude
1992, strikes 102000:100000


Illustrative Sharpening Example − 3 Iterations
Time
Longitude
Latitude
−100
−95
−90
−85
−80
−75
42
44
46
48
50
52
54
1992, strikes 100000:102000

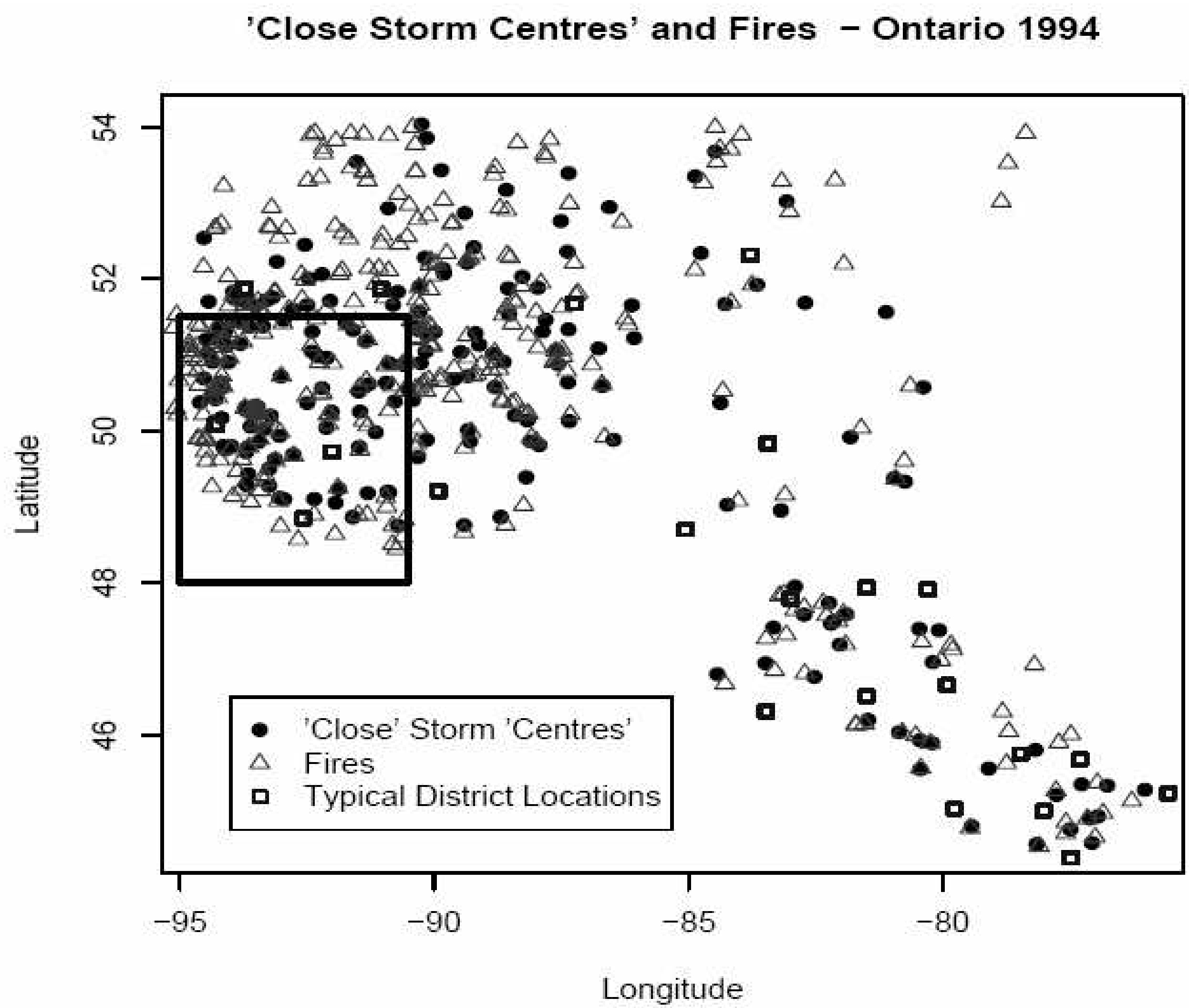

Illustrative Sharpening Example − 10 Iterations
Time
Longitude
−100
−95
−90
−85
−80
−75
Latitude
42
44
46
48
50
52
54
1992, strikes 100000:102000

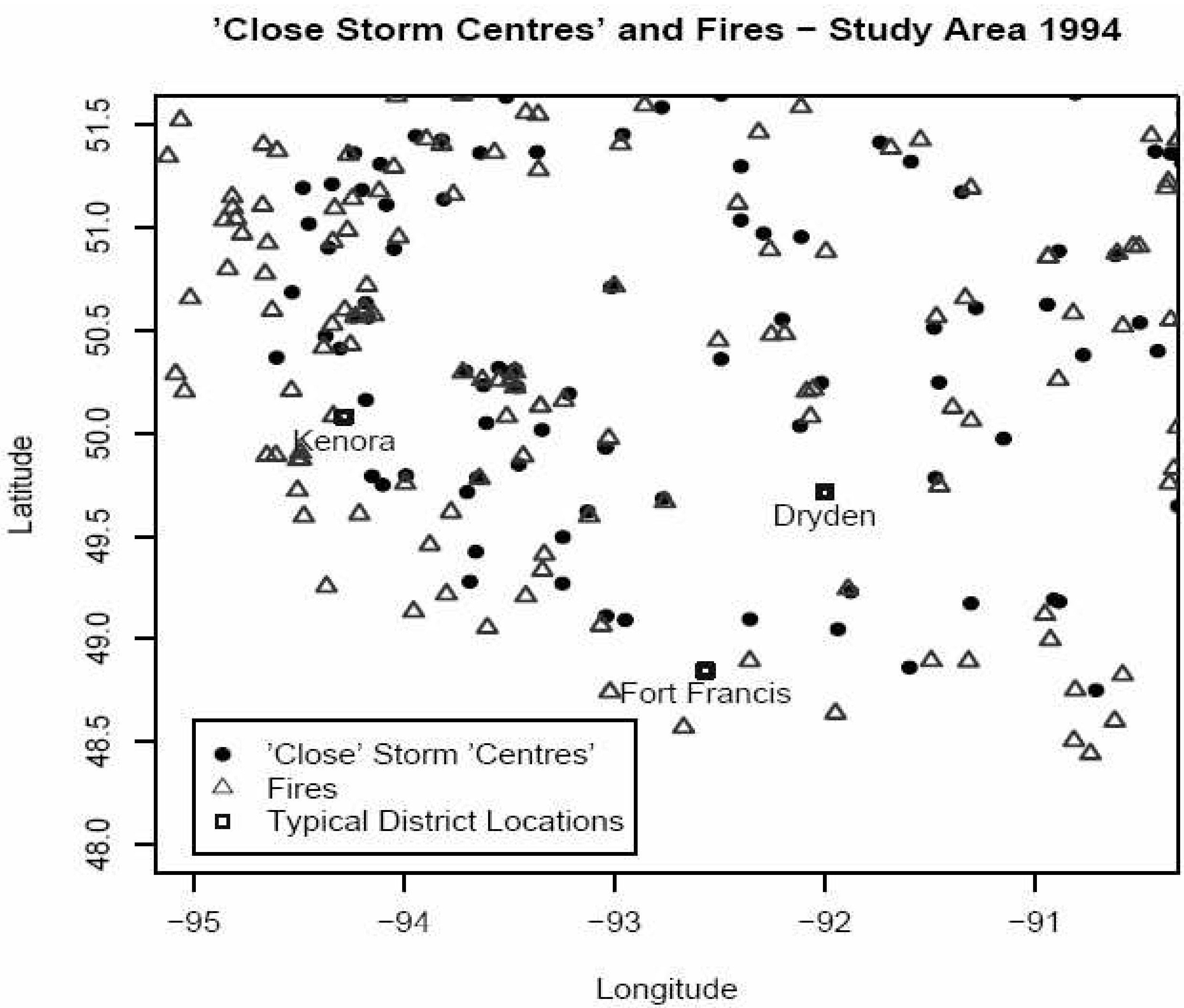
Fire Ignitions Close and Too Far from Storm Centres − 1994





'Close Storm Centres' and Fires - Ontario 1994
Latitude
54
52
50
48
46
−95
−90
−85
−80
Longitude
'Close' Storm 'Centres'
Fires
Typical District Locations


'Close Storm Centres' and Fires - Ontario 1994
Latitude
54
52
50
48
46
-95
-90
-85
-80
Longitude
'Close' Storm 'Centres'
Fires
Typical District Locations


'Close Storm Centres' and Fires – Study Area 1994
Latitude
Longitude
51.5
51.0
50.5
50.0
49.5
49.0
48.5
48.0
−95
−94
−93
−92
−91
Kenora
Dryden
Fort Francis
'Close' Storm 'Centres'
Fires
Typical District Locations

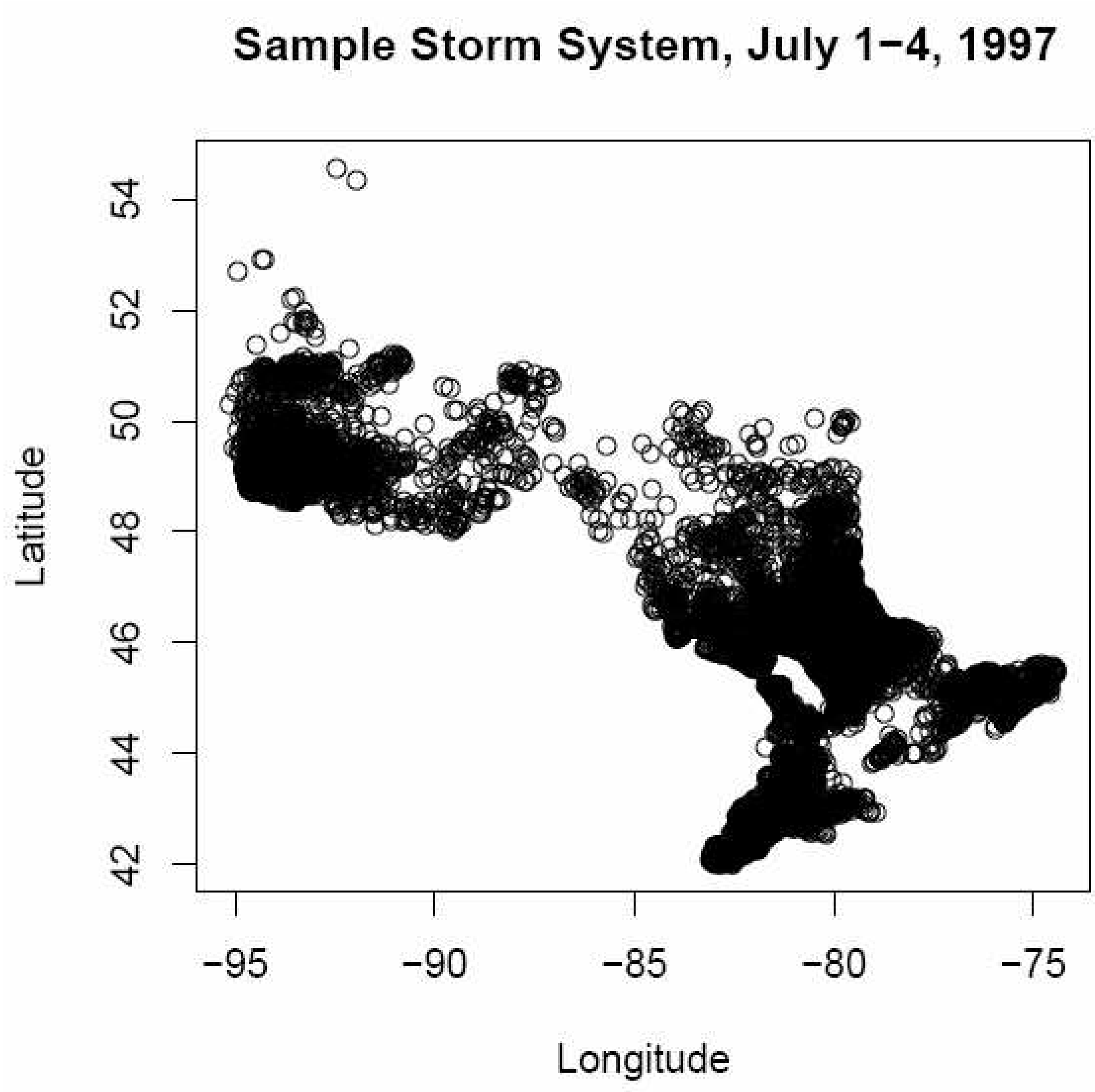
Tracking Storms Through Time

Another Data Sharpening Application

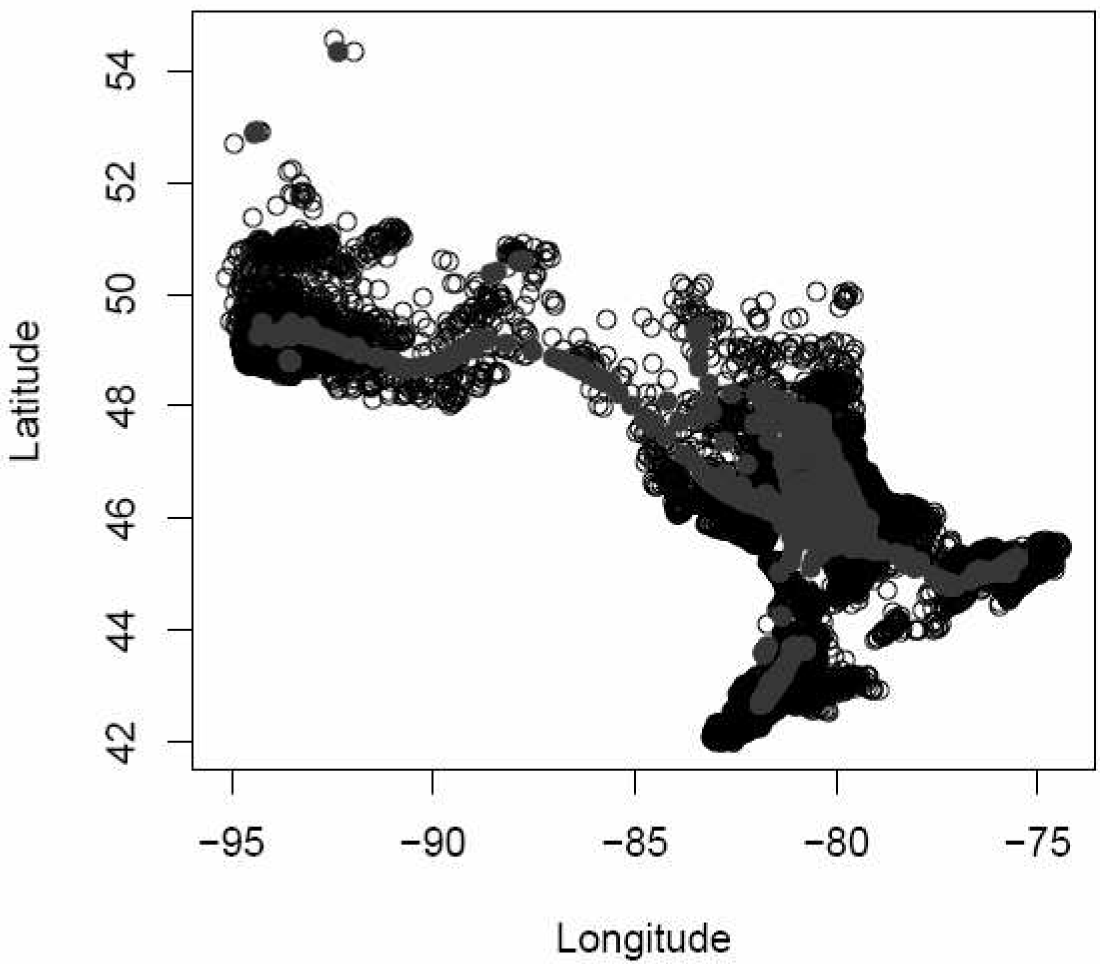
- data filtering (Fukunaga & Hostetler, IEEE 1975)
  - reduce noise in data
  - can reduce noise only in minor dimensions if desired
  - use to reveal “dominant data surface”

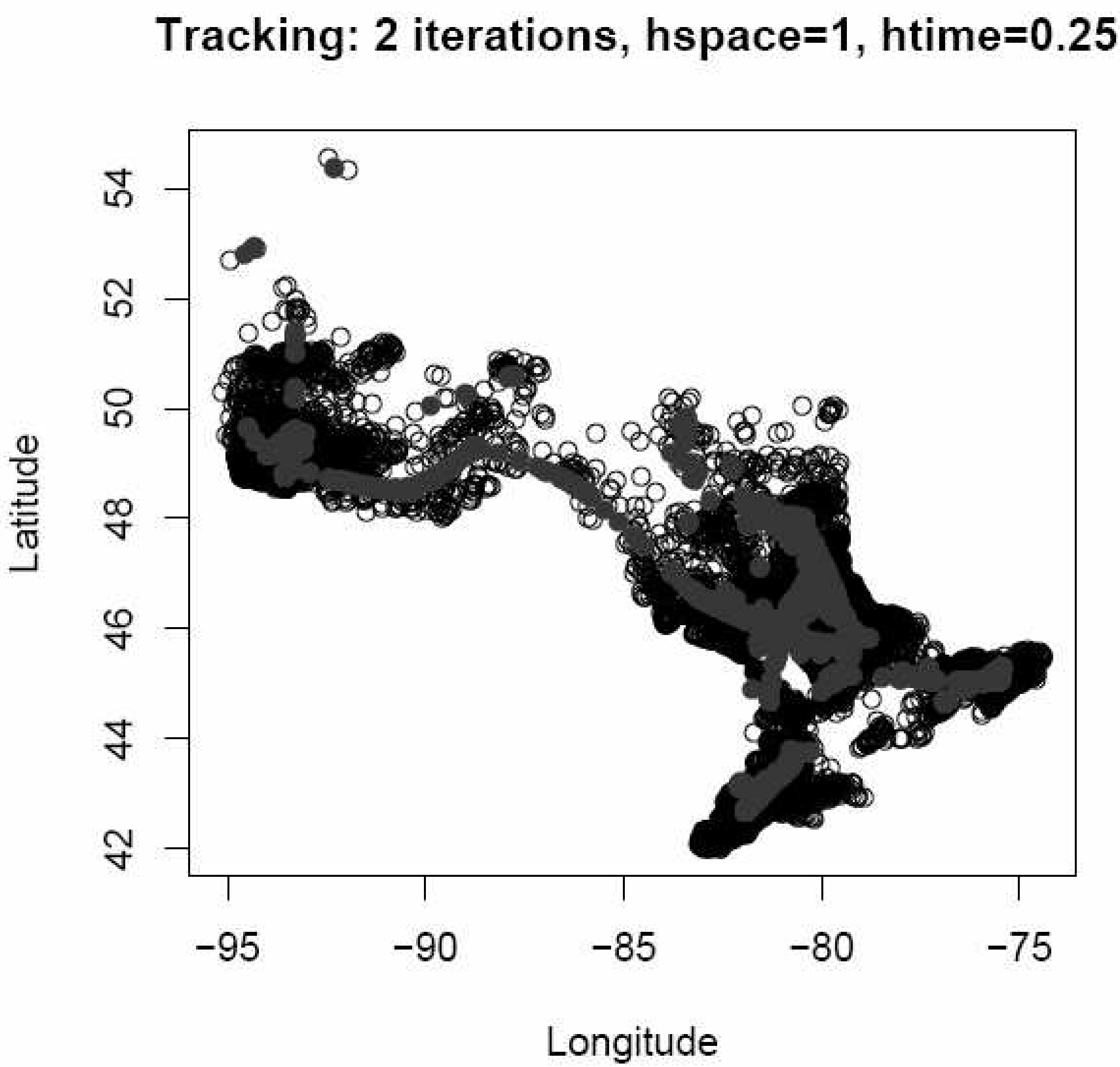
Tracking Storms

- systems move across Ontario
- by reducing time bandwidth & not iterating to convergence can “track” the path of the storm centre through space


Sample Storm System, July 1−4, 1997
Latitude
54
52
50
48
46
44
42
−95
−90
−85
−80
−75
Longitude

Tracking: 2 iterations, hspace=1, htime=0.5




Tracking: 2 iterations, hspace=1, htime=0.25
Latitude
42
44
46
48
50
52
54
Longitude
−95
−90
−85
−80
−75

Conclusions

- Data issues:
  - Summarizing/reducing and declustering large lightning observations.
  - Missing and inaccurate lightning data.
  - Nonstationarity.

- Tools:
  - Converged data sharpening ð Identify cluster centres.
  - Fire ignitions ð Locate ‘close’ cluster centres.

Further Work:

- Can lightning be modelled as a Poisson cluster process?
- Sharpened storm centres allow us to study ‘clusters’.
- Converged sharpening depends on bandwidth (3D).
- Smouldering time distribution. Cross-intensity function?
- Nonstationarity, e.g. is ‘district’ a potentially significant covariate?
- Identifying other covariates of significance

  e.g. weather variables, fuel moisture codes, polarity.
- Other wildfire questions, e.g. growth, spot ignitions, containment, optimal resource allocation.

Acknowledgements

The support of the Natural Sciences and Engineering Research Council of Canada is gratefully acknowledged.