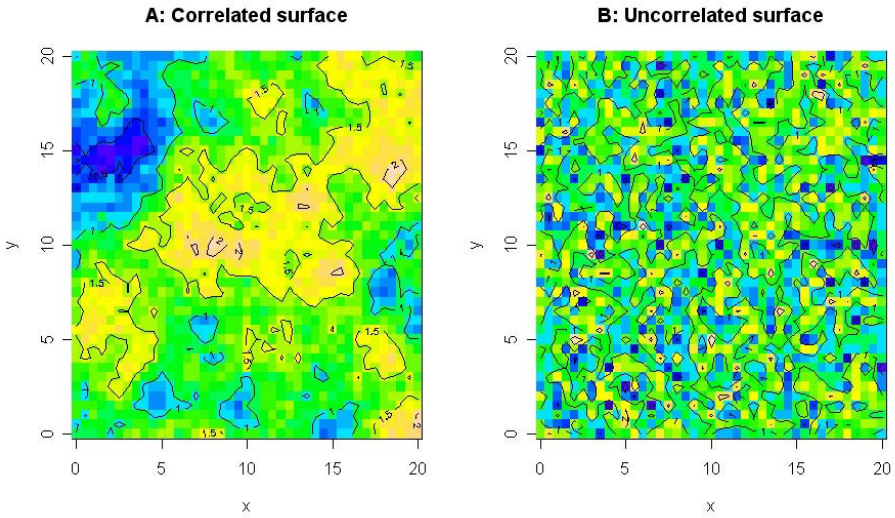
Disease Clustering



What is disease clustering?

(non-focussed clustering (Besag and Newell, 1991))

Definitions:

‘**a geographically bounded group of occurrences of sufficient size and concentration to be unlikely to have occurred by chance**’. (**Knox**,**1989**)

This is a very general definition and it corresponds with ’*hot spot’* clustering. (Note that this definition could include a focused cluster. The main difference is that with focussed clustering the focus is *known.*)

An alternative definition:

‘**The study area has a predefined cluster structure**’

These two definitions correspond to two extremes of clustering: the first assumes little knowledge of the cluster form, whereas the second assumes a form *apriori.*

Reasons for studying Clustering

- Finding the Etiology of Disease

Often there may be unknown factors which affect an disease and clusters can lead to finding these factors

- Evaluating Disease Cluster Alarms

Cluster alarms are now quite common and it is the focus of analysis to ’test’ whether clusters really exist and what is the underlying factors leading to the cluster occurrence.

- Public Health Surveillance

Routine surveillance of disease may require that *unusual* clusters of disease be detected monitored.

Examples of clustering studies where the aetiology was known

Disease	Location	Aetiology
Cholera	London	water borne
Nasal sinus	England	furniture trade
lung cancer	Scotland	air pollution
Leukaemia	Hiroshima	Radiation
mesothelioma	Karain, Turkey	Erionite fibre
lung cancer	Georgia USA	ship asbestos
pneumonia	Los Angeles	HIV
oral cancer	southern USA	snuff dipping
asthma	Barcelona	soya dust
Down’s Synd	Hungary	Trichlorfon-fish

Reference
Snow(1854)
Macbeth(1965)
Lawson and Williams (1994)
Ishimaru et al (1971)
Barris et al (1978)
Blot et al (1978)
CDC (1981)
Winn et al (1981)
Anto et al (1989)
Czeizel et al (1993)

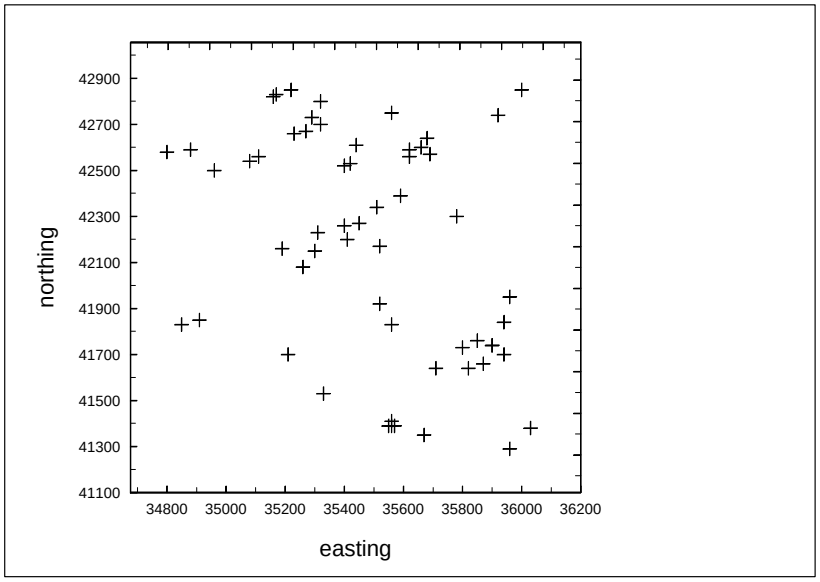
Modelling Issues

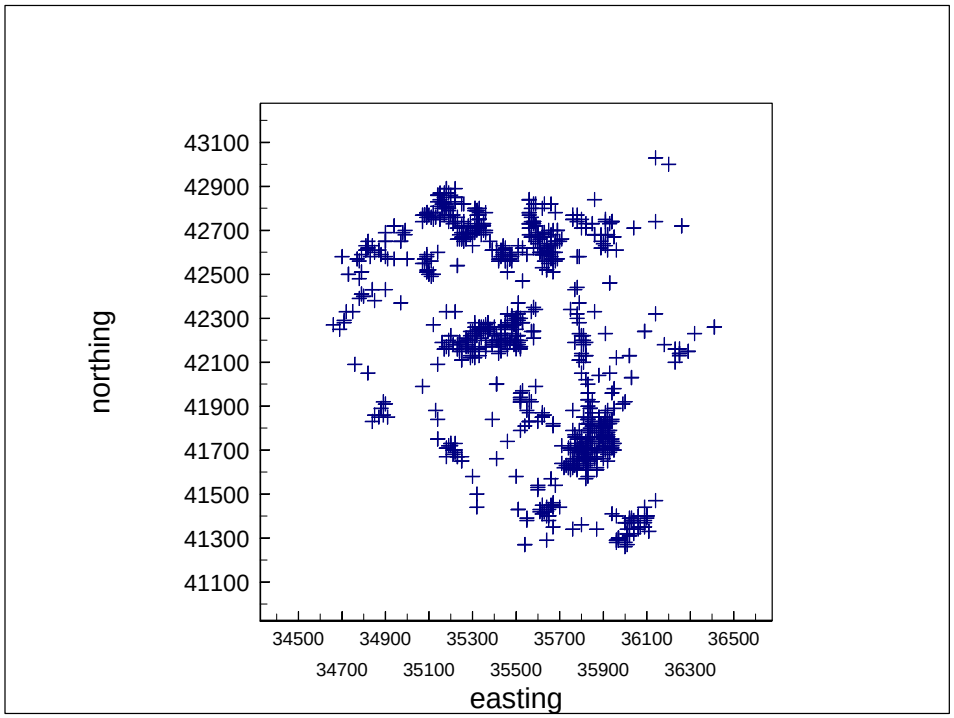
- The development of models for clusters and clustering has seen greater development in some areas than in others. For example, it is straightforward to formulate a non-specific Bayesian model for case events or tract counts that includes heterogeneity.
- Specific models are less often reported.
- It is possible to formulate specific clustering models for the case-event and tract-count situation.
- If it is assumed that the intensity of case events, at location $\mathbf{x}$, is $\lambda(\mathbf{x})$, then by specifying a dependence in this intensity on the locations of cluster centres, it is possible to proceed.

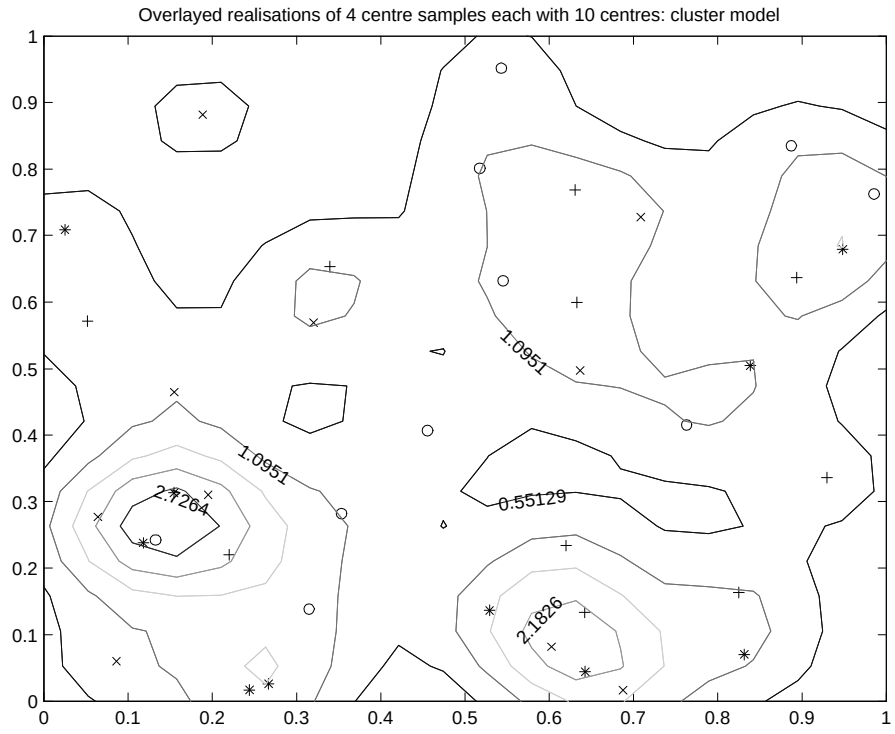
- The distribution of events around a centre is defined by a cluster distribution function $h(.)$. Conditional on the cluster centres, the events can be modelled via a likelihood.
- As the number (k) and the locations of centres are unknown, then with a suitable prior distribution specified for these components, it is possible to formulate this problem as a Bayesian posterior sampling problem, with a mixture of components of unknown number.
- This type of problem is well suited to advanced MCMC sampling methods. The approach can be extended to count data straightforwardly.
- Other approaches are possible (see e.g. Lawson and Denison(2002))

Some pictorial examples:

Larynx cancer cases and lung cancer control locations: Lancashire.




northing
43100
42900
42700
42500
42300
42100
41900
41700
41500
41300
41100
34500
34900
35300
35700
36100
36500
34700
35100
35500
35900
36300
easting


Overlayed realisations of 4 centre samples each with 10 centres: cluster model
1.0951
1.0951
2.7264
0.55129
2.1826
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

Hypothesis Tests for Clustering

- The literature has developed considerably in the area of hypothesis testing for clusters.

- Very early developments in this area arose from the application of statistical tests to spatio-temporal clustering, a particularly strong indicator of the importance of a *spatial* clustering phenomenon.

- Distinction should be made between tests for general (*non specific)* clustering, which assess the overall clustering pattern of the disease, and the *specific* clustering tests where cluster locations are estimated.

- For case events, a few tests have been

developed for non-specific clustering.

- Cuzick and Edwards (1990) developed a test which is based on a realisation of cases and a *sample* of a control realisation. Functions of the distance between case locations and *k* ‘nearest’ cases were proposed as test statistics(as opposed to controls).

- Anderson and Titterington (1997) have proposed the use of a simple integrated squared distance statistic (ISD) for cluster assessment. This is closely related to the analysis of density ratios in exploratory analysis, and could be regarded as a type of nonparametric assessment of clustering.

- The advantage of this approach is that the assessment is not tied to a specific cluster model but detects departures from

background. The major disadvantage, shared with all such statistics, is its low power against specific forms of clustering.

- Other simple forms of a global test can be proposed, where density estimates of cases are compared to intensity estimates of case events simulated from the control background. These could provide pointwise confidence intervals as well as global tests.

- There appears to have been little development of tests that detect uncorrelated heterogeneity in the intensity of the case-event process as a form of spatial clustering. It is unclear what aetiological difference would be inferred when uncorrelated rather than correlated forms of heterogeneity were found.

- The general tests for overall clustering so

far proposed, suffer from the following problem. Often underlying unobserved heterogeneities are common in the data and the general tests do not provide mechanisms for the incorporation of such effects.

- For example, if non-stationarity were present in the case events, then this effect could be confounded with cluster effects.

- One solution to this is to adopt a full clustering model such as, which can be expanded easily to include such effects as non-stationarity and heterogeneity, and to test for inclusion of effects within MCMC algorithms.

- General clustering tests based on tract counts, can be classified into tests for correlated heterogeneity and tests for uncorrelated heterogeneity.

- Tests of the latter are not *spatial* in origin but are included here for completeness.

- Tango (1995) developed a modified general class of tests for general and focussed clustering. All these tests make approximating assumptions (e.g., that counts are independently Poisson distributed with constant expectation within each tract), and are unlikely to perform well against specific clustering alternatives.

- Also they assume that clusters yield a total increase in divergence between count and expectation over the whole study region, while equivalent degrees of divergence could be due to non-cluster processes also, and hence this could lead to misinterpretation.

- Some use has been made of tests for uncorrelated heterogeneity to assess clustering of tract counts. For example, the Euroclus project (Alexander et al (1996)), has invested considerable effort in testing for such heterogeneity across European states using the Potthoff-Whittinghill test and score tests for Poisson versus negative binomial distributions for the marginal count distribution.

- These tests are approximate in that they assume constant-within-region expected rate, and they may suffer from considerable interpretational problems when *a priori* there is likely to be some non-specific heterogeneity in small-area data.

- In addition, the evidence of Euroclus suggests that for certain important forms of non-Poisson alternatives within the negative binomial family, these tests perform poorly.

- Finally, *at least for rare diseases, it is easily possible that the marginal count distribution would not follow a negative binomial distribution and could even display multimodality.*

Specific cluster tests

- address the issue of the location of putative clusters. These tests produce results in the form of locational probabilities or significances associated with specific groups of tract counts or cases.

- Openshaw et. al.(1987) first developed a general method that allowed the assessment of the location of clusters of cases within large disease maps.

- The method was based on repeated testing of counts of disease within circular regions of different sizes. Whenever, a circle contains a significant excess of cases, it is drawn on the map.

- After a large number of iterations, the resulting map can contain areas where a concentration of overlapping circles suggests localised excesses of a disease.

- The statistical foundation of this method has been criticised and an improvement to the method was proposed by Besag and Newell (1991).

- Their method involves accumulating events (either cases or counts) around individual event locations. These could be cases or tracts. Accumulation proceeds up to a fixed number of events or tracts (k). The number k is fixed in advance. The method can be carried out for a range of k values.

- While the local alternative for this test is increased intensity, there appears to be no specific clustering process under the alternative and in that sense the test procedure is nonparametric, except that a monotone cluster distance distribution is implicit.

- One advantage of the test is that it can be applied to focussed clusters also, while a disadvantage is that an arbitrary choice of k must be made and the results of the test must depend on this choice.

Scan Tests

- An alternative statistic has been proposed by Kulldorff and Nagarwalla(1995), who employ a likelihood ratio test for the comparison of an overall binomial model for the number of cases found in the study region population(the null hypothesis), to a model that has different binomial components depending on being inside or outside a circular zone of defined size. *The scan statistic*

- The test can be applied to both case events and tract counts.

- The advantage of the test is that it examines a potentially infinite range of zone sizes and does rely on a formal model of null and alternative hypotheses.

- However, some limitations of the method relate to the use of circular regions, which tends to emphasise *circular* clusters (as does the Openshaw test), and to the choice of crude population as the expression of the background ‘at risk’ structure.

- The software which implements the spatial scan statistic is *SatScan* and is available at *http://www.satscan.org/*

Session: Disease Clustering case study

In this section we will examine some real clustering data and attempt to develop some methods for the analysis of clustering

- Some example data sets:

Figure 1 to Figure 8 represent different examples of clustering and methods of representing clusters

- Figure 1 and 2:

Describe the patterns seen: are they clustered?

- Figure 3

This is a simulated data set consisting of 200 cases of larynx cancer; the simulation is based on taking respiratory cancer control disease locations (978)

and smoothing these and then generating cases from the smoothed surface. This is a typical map of case locations generated from a population and it does not contain any unusual artifacts (in the sense that the cases purely arose independently from the population).

- Figure 4 and 5

These are simply contour plots of the case distribution using different degrees of smoothing (h=0.005 and 0.05 respectively)

- Figure 6

Clearly it is important to relate the case distribution to the control distribution. In Figure 6 the ratio of the two surfaces are plotted and it can be seen that there is a very smooth result with little expression of excess risk.

- Figure 7 and 8

Another simulation: Is there any evidence here of clustering?

Figure 8 is the ratio of the case to control surfaces

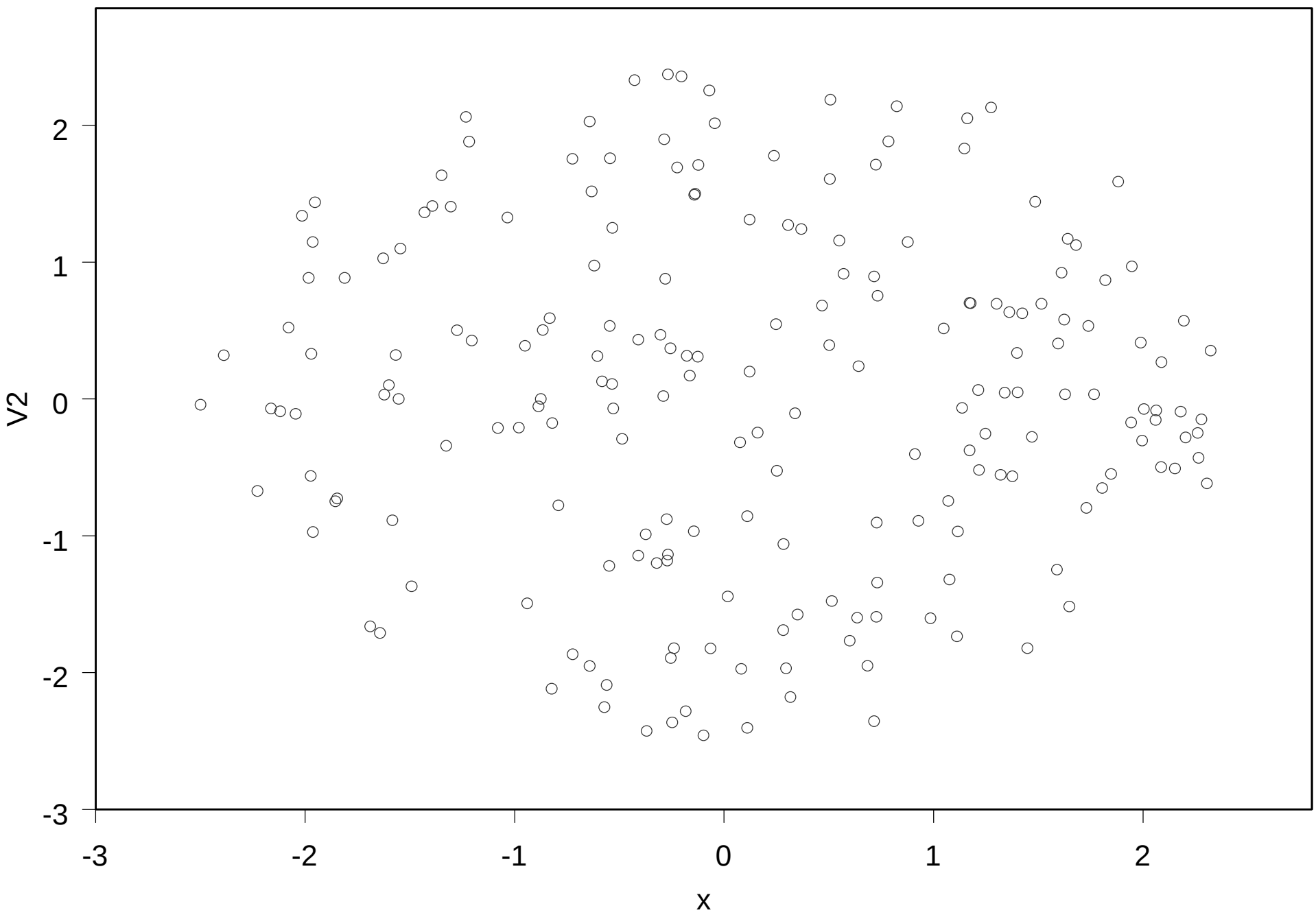
Technical data for figures 7 and 8

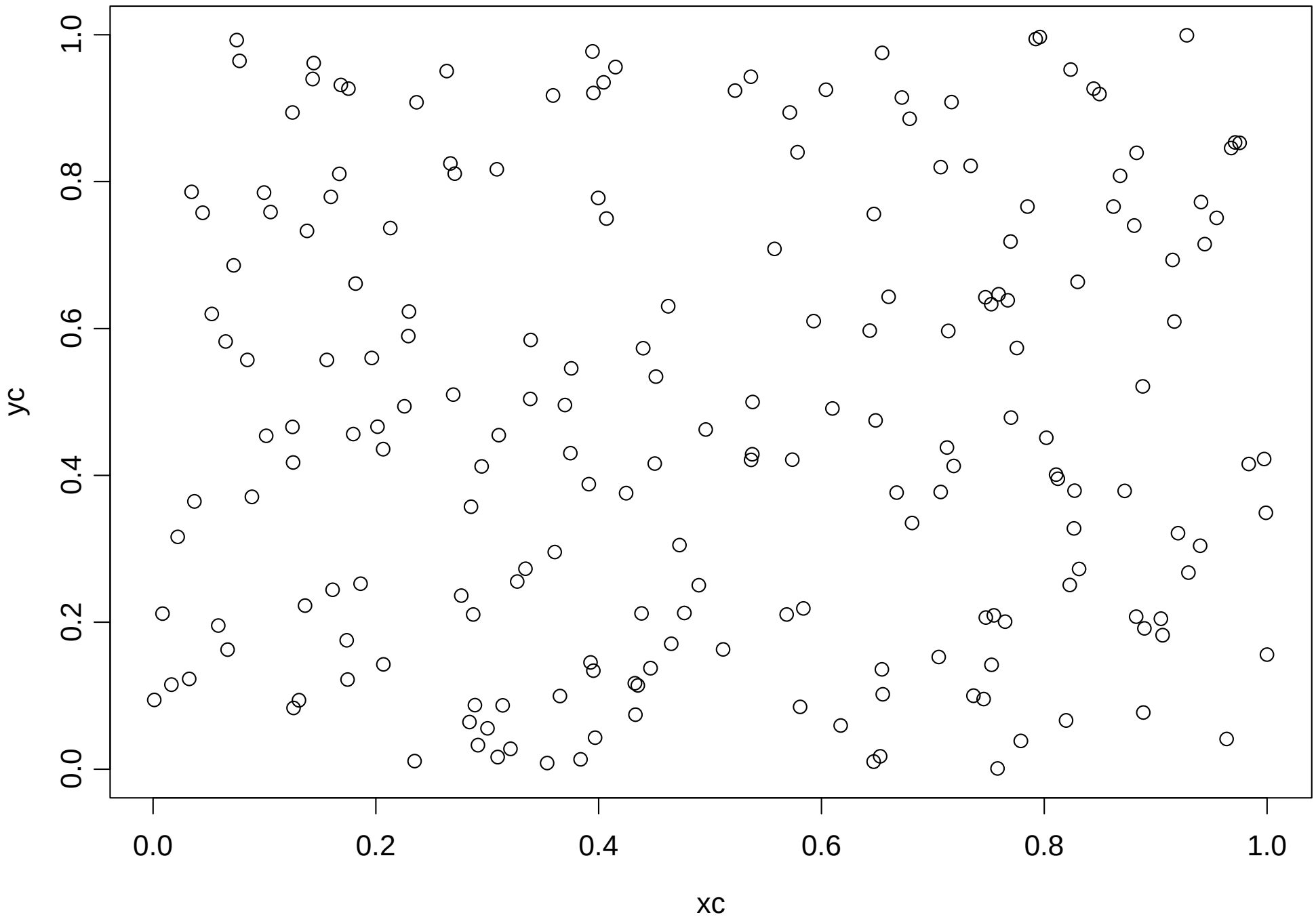
- In addition to simulation of 200 cases from the lung cancer control:
  clusters were induced at 5 locations with small variance (0.03), and 225 points were simulated:

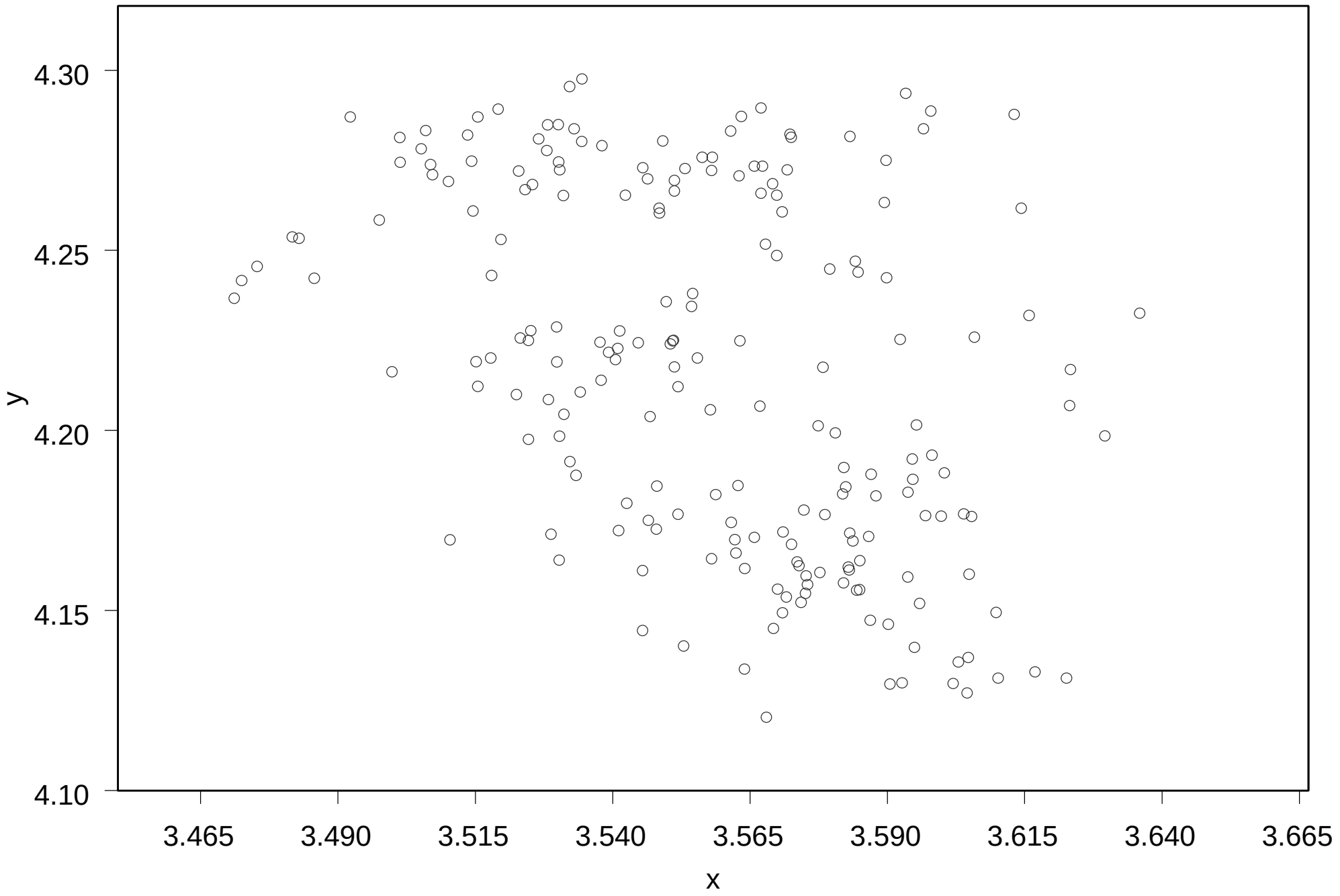
- the cluster locations were;

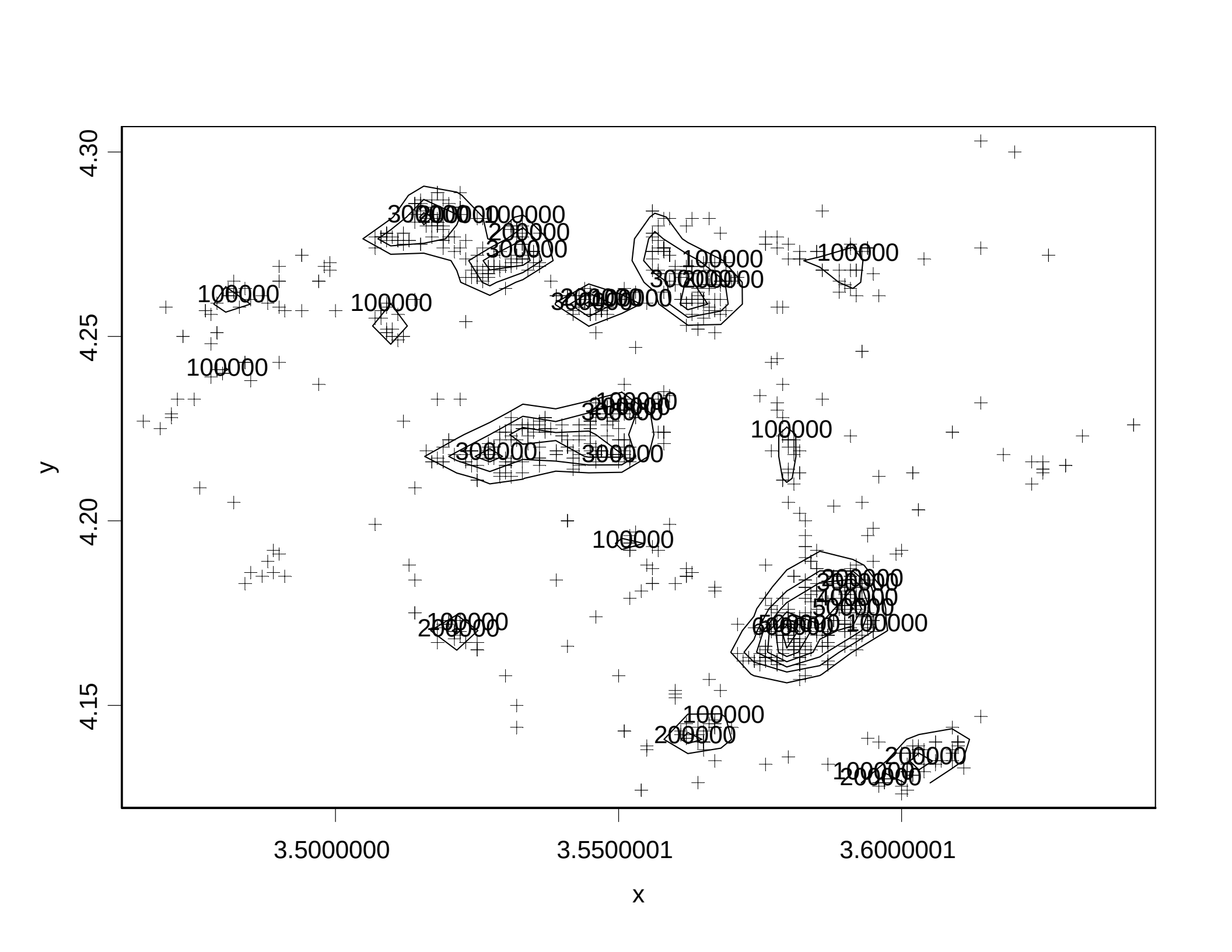
x	y
3.515	4.150
3.565	4.150
3.565	4.280
3.614	4.200
3.470	4.250

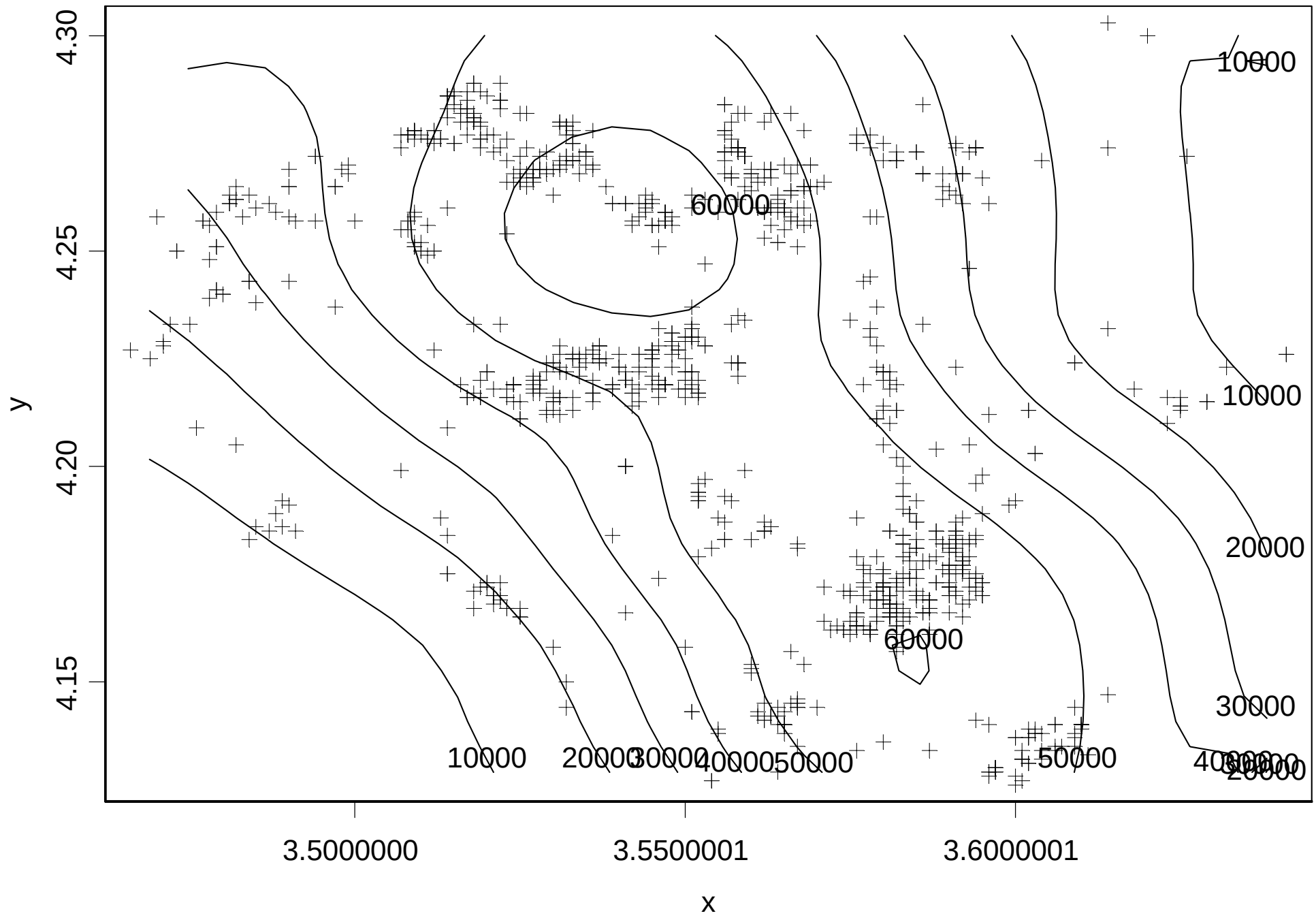
- kernel ratio estimation using SPlancs function : *kernrat* with h1=0.05,h2=0.05

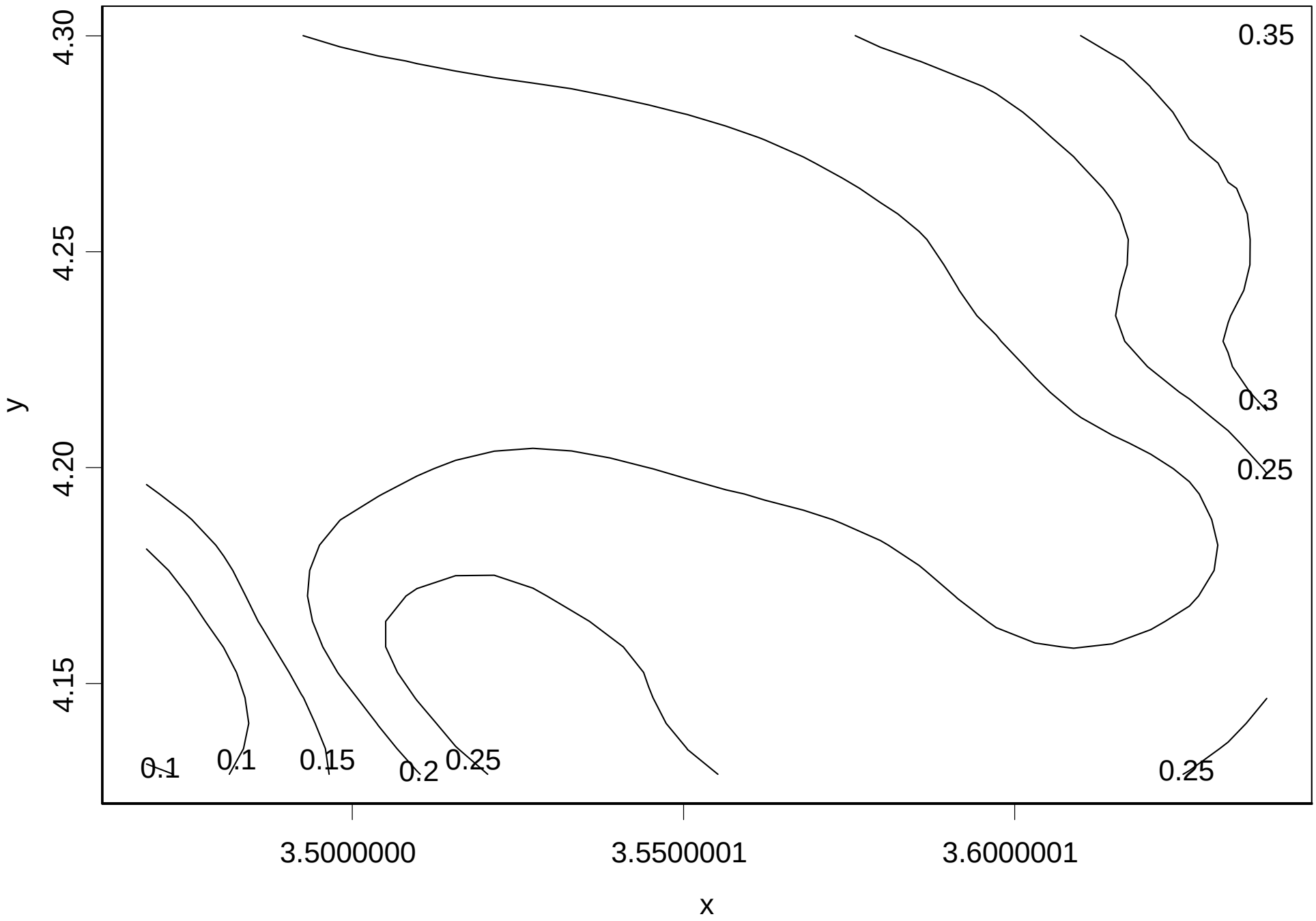

2
1
0
-1
-2
-3
V2
-3
-2
-1
0
1
2
x

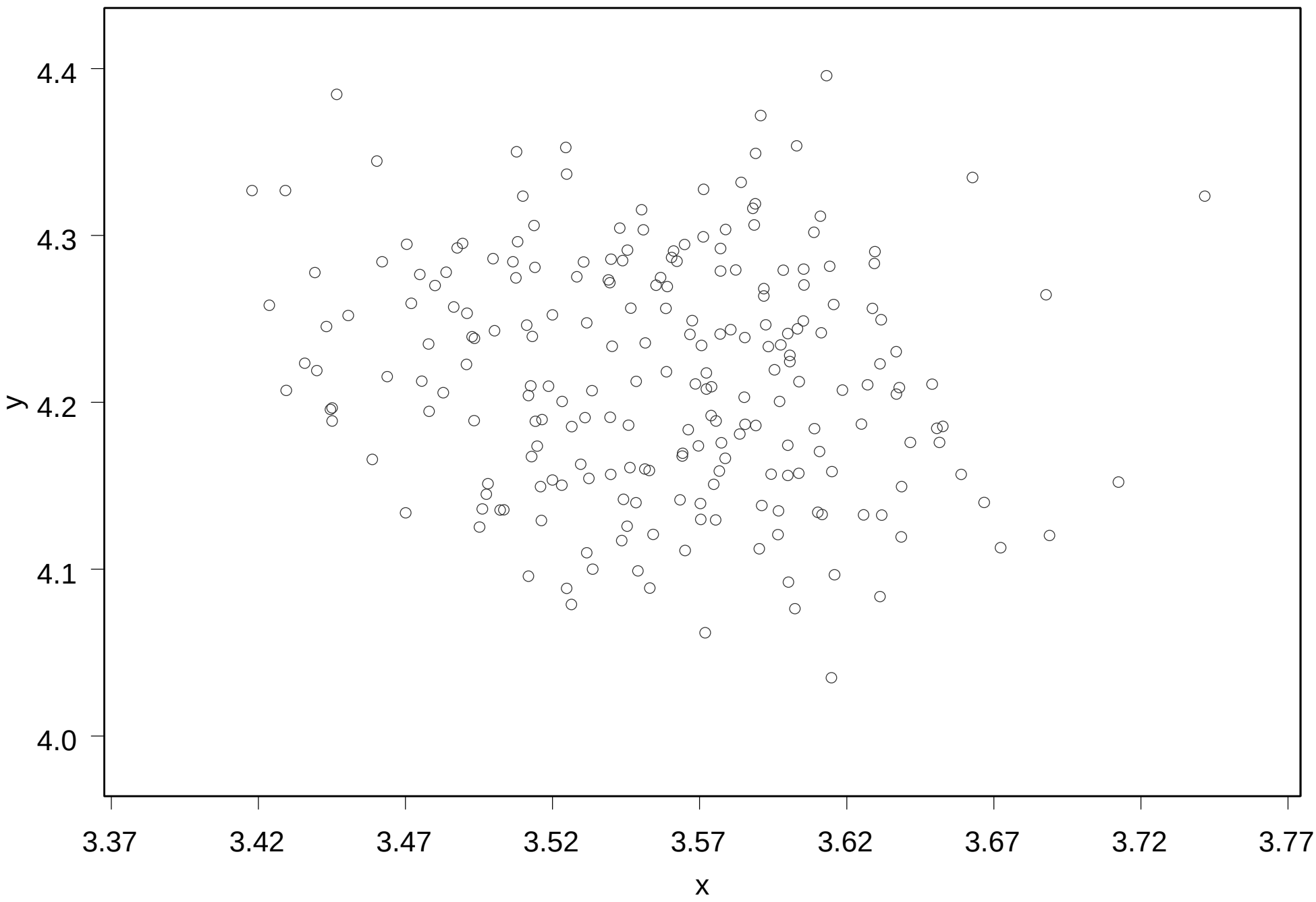

yc
xc
0.0
0.2
0.4
0.6
0.8
1.0

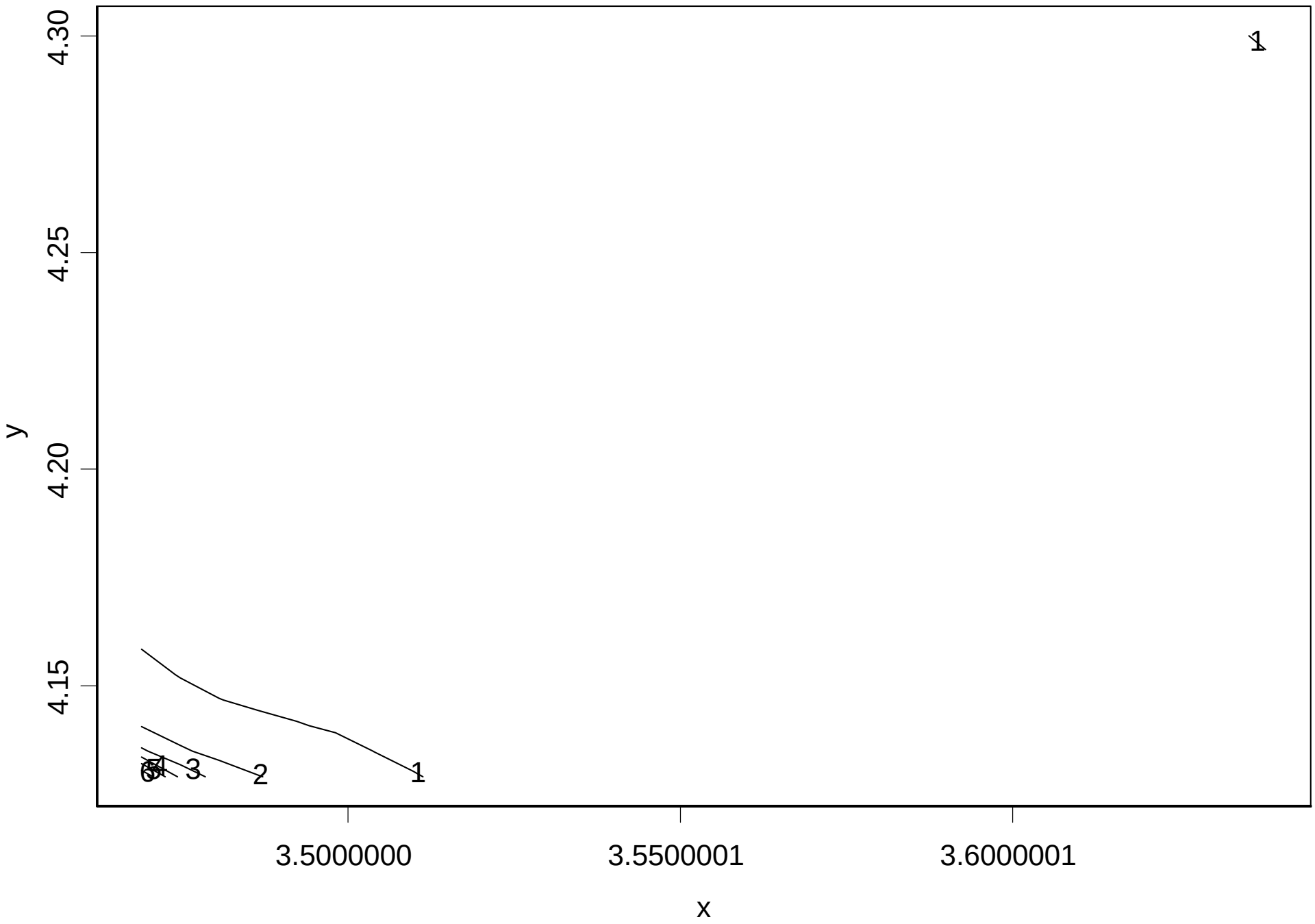

4.30
4.25
4.20
4.15
4.10
y
3.465
3.490
3.515
3.540
3.565
3.590
3.615
3.640
3.665
x

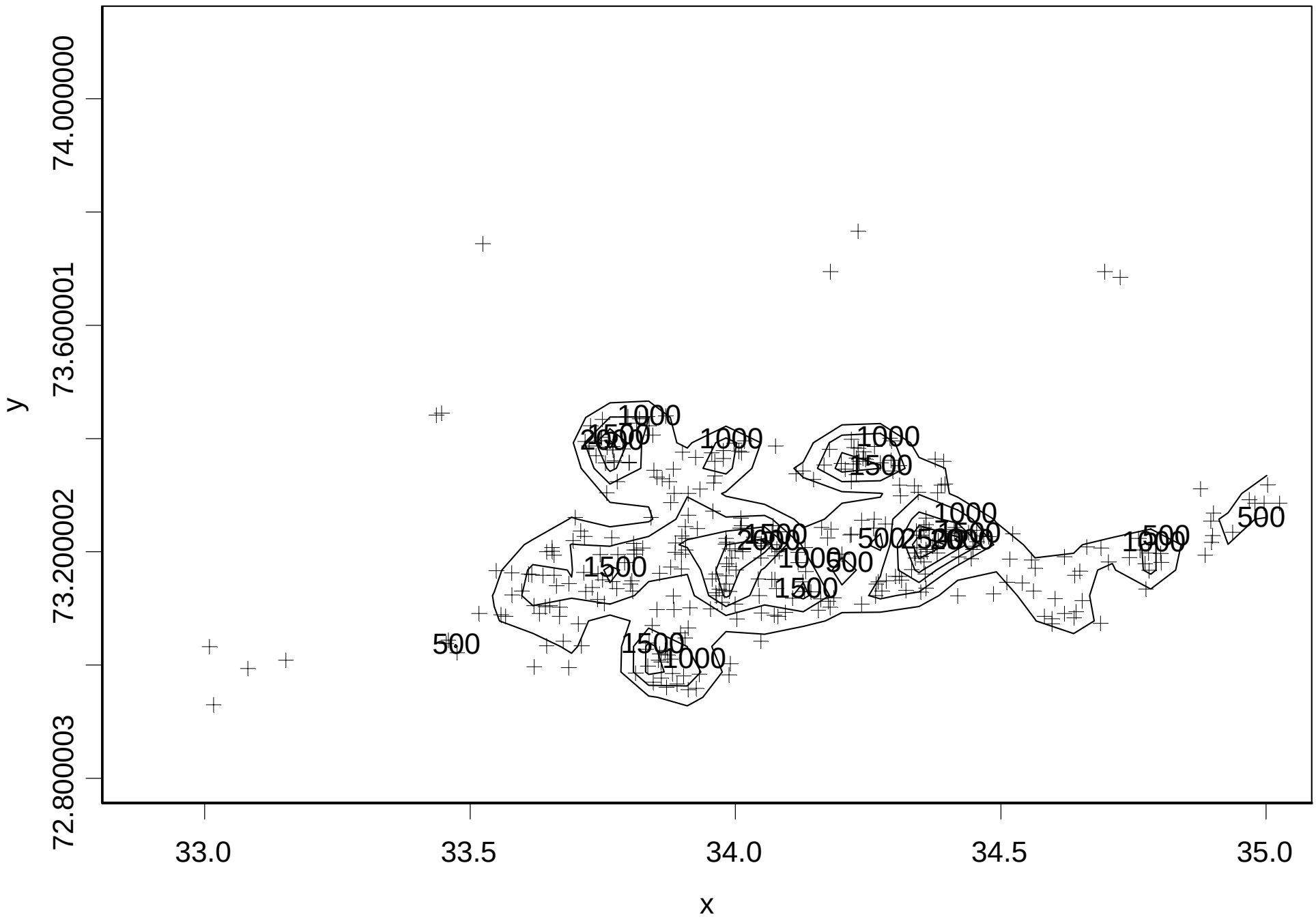

4.30
4.25
4.20
4.15
y
3.5000000
3.5500001
3.6000001
x

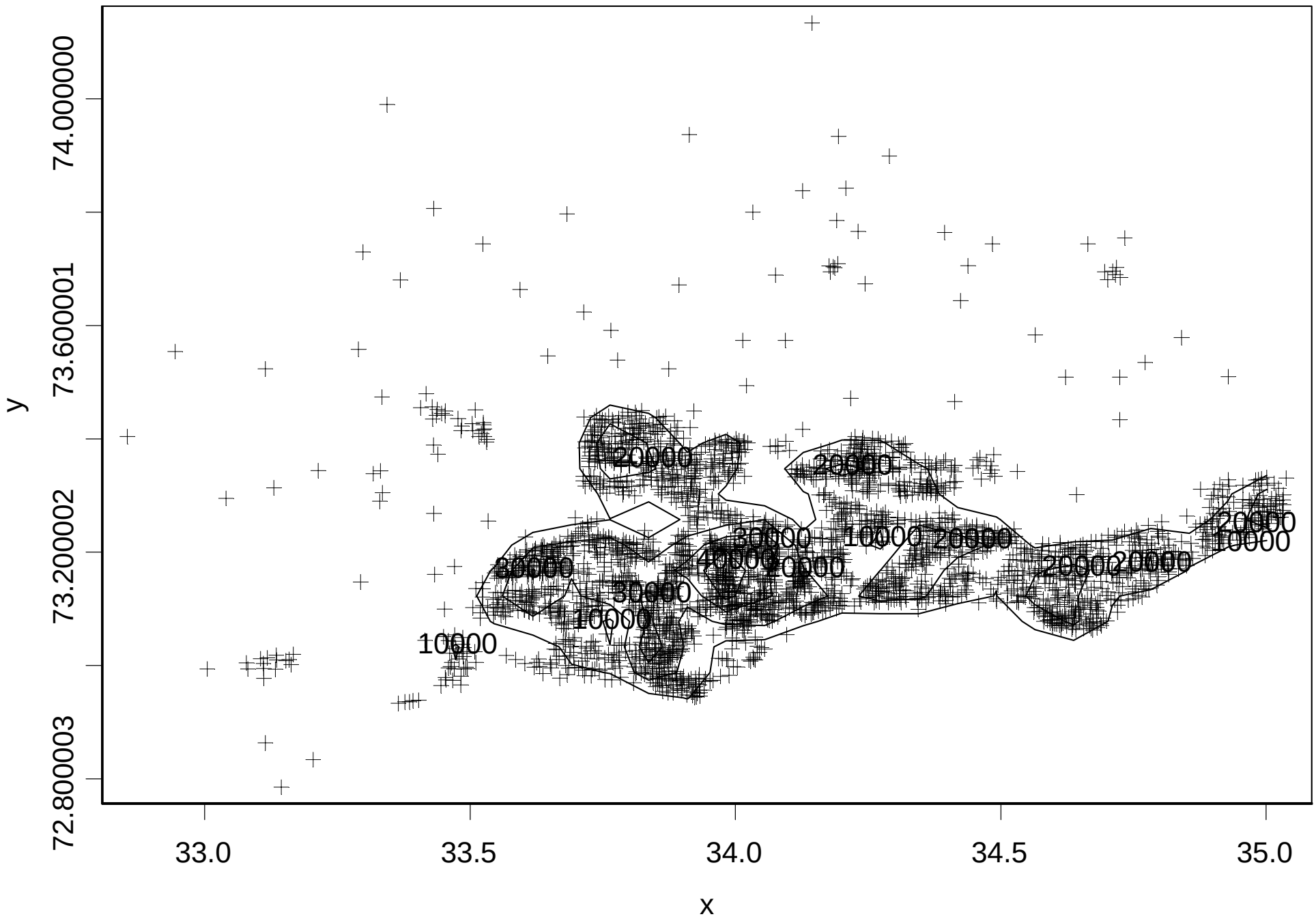

4.30
4.25
4.20
4.15
y
10000
60000
10000
20000
60000
30000
10000
20000
30000
40000
50000
50000
40000
60000
3.5000000
3.5500001
3.6000001
x

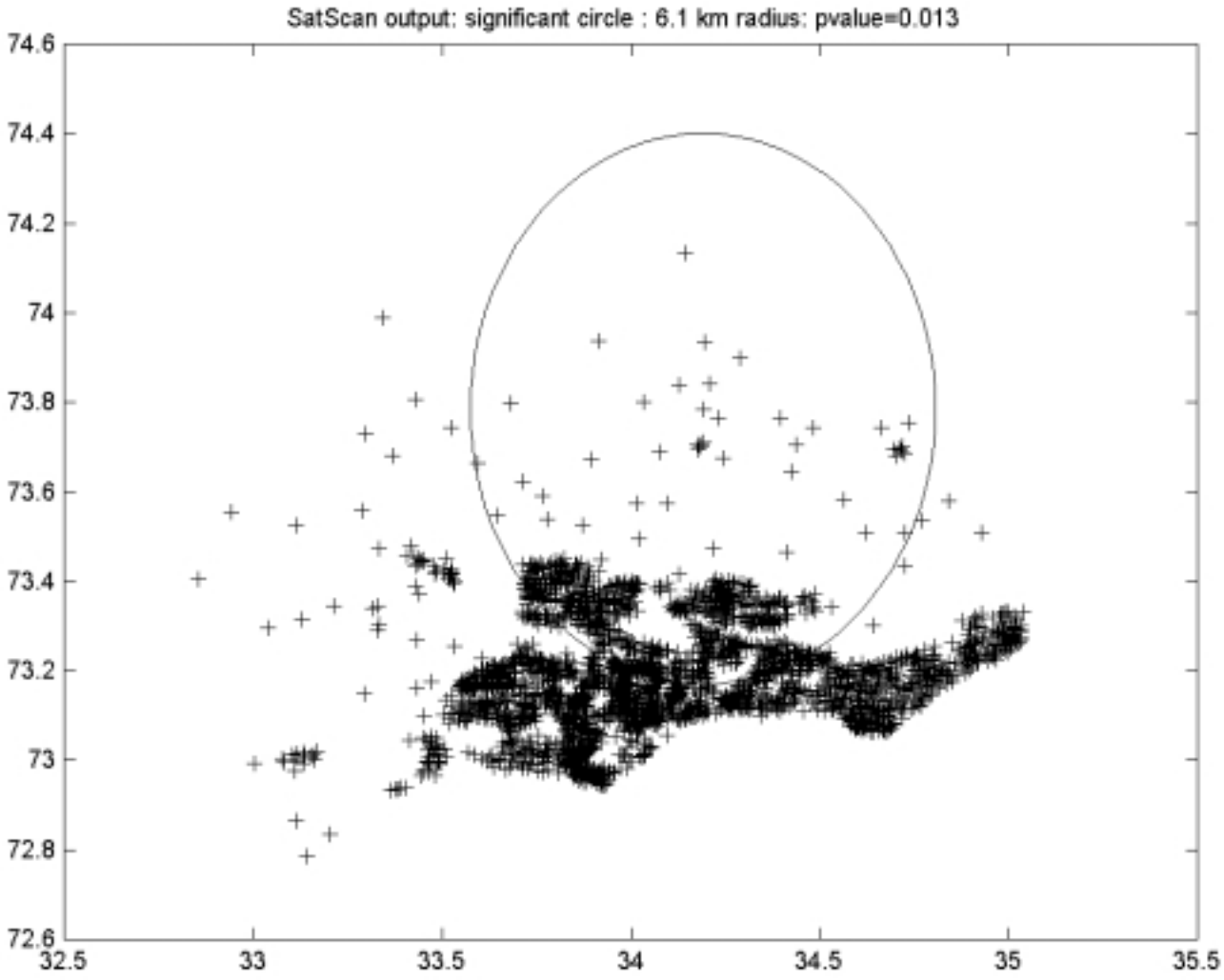

4.30
4.25
4.20
4.15
y
0.35
0.3
0.25
0.1
0.1
0.15
0.2
0.25
0.25
3.5000000
3.5500001
3.6000001
x

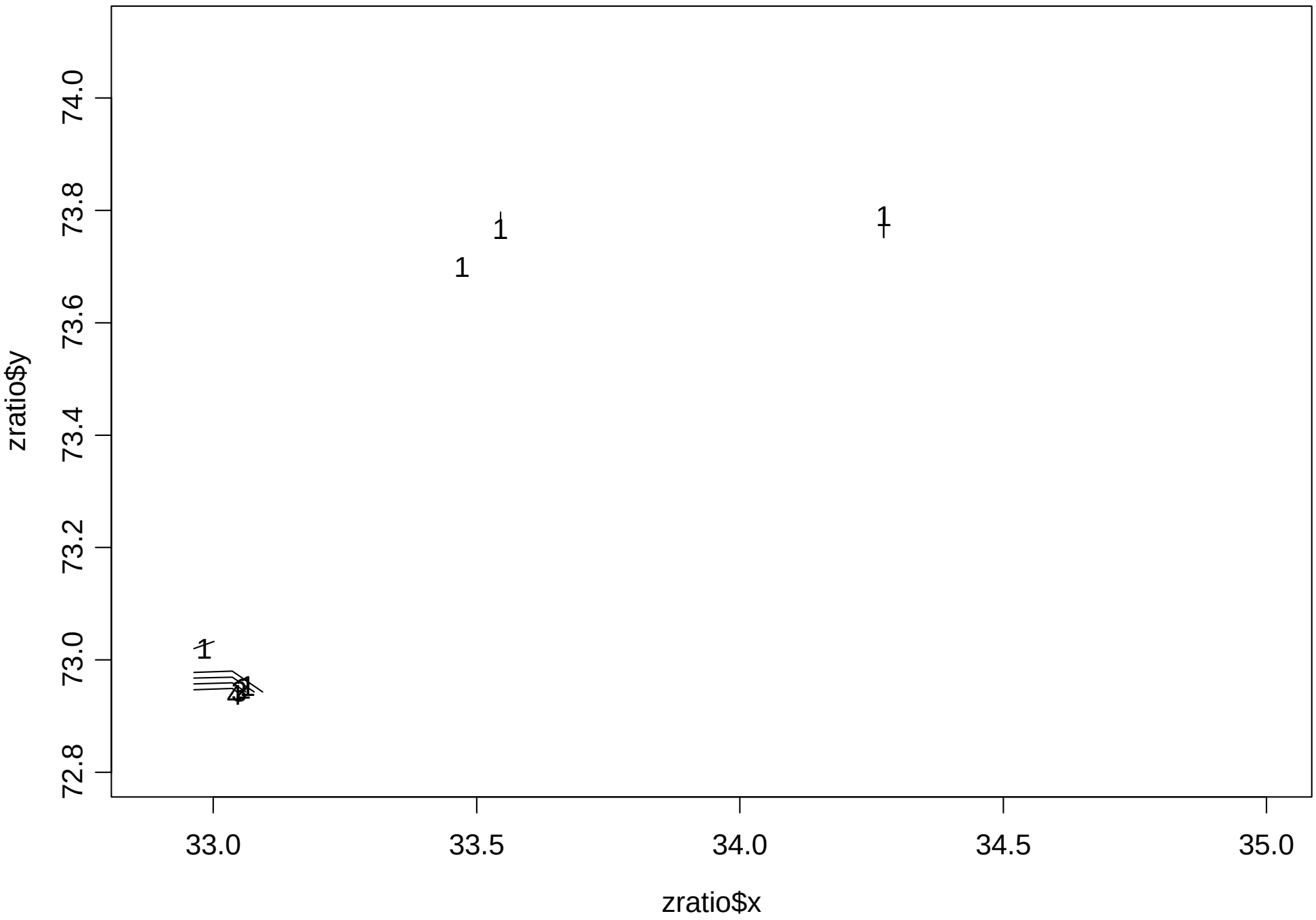

4.4
4.3
4.2
4.1
4.0
y
3.37
3.42
3.47
3.52
3.57
3.62
3.67
3.72
3.77
x


4.30
4.25
4.20
4.15
y
1
1
2
3
3.5000000
3.5500001
3.6000001
x


74.000000
73.600001
73.200002
72.800003
y
33.0
33.5
34.0
34.5
35.0
x
1000
1500
2000
500


y
74.000000
73.600001
73.200002
72.800003
x
33.0
33.5
34.0
34.5
35.0
10000
20000
30000
40000
10000
20000
30000
10000
20000
10000
20000
20000
20000
10000


SatScan output: significant circle : 6.1 km radius: pvalue=0.013
74.6
74.4
74.2
74
73.8
73.6
73.4
73.2
73
72.8
72.6
32.5
33
33.5
34
34.5
35
35.5


zratio$y
74.0
73.8
73.6
73.4
73.2
73.0
72.8
33.0
33.5
34.0
34.5
35.0
zratio$x

A Real Data set

- Figure 9,10,11,12

  - This is a real anonymised data set: Figure 9 is a map of the cases of particular cancer (lets call it the *case* cancer), while Figure 10 is a map of the a composite control based on lower body cancers (the *control* cancers).
  - Figure 11 is a display of the results of a SatScan run (the out put from SatScan is also included) while Figure 12 is the results of using extraction mapping of the cases compared to the controls (using *kernrat* from SPlancs)

Comments:

SatScan output for the real data

SaTScan V.2.1.3

Program run on: Sat Nov 03 15:58:16 2001
Purely Spatial analysis scanning for clusters with
high rates using the Bernoulli model.

SUMMARY OF DATA
Study period .....: 1979/1/1 - 1996/12/31
Number of census areas: 3542
Total population .....: 8243
Total cases .....: 387

MOST LIKELY CLUSTER
1.Census areas included.: XX40QY, XX40QX, XX40QU, XX40PZ, XX40QT,
XX40QS, XX40RH, XX40QZ, XX40QP, XX40QR,
XX40QN, XX40QA, XX40QB, XX40QL, XX40RD,
XX40PW, XX40PT, XX40RE, XX30QE, XX40PP,
XX40PN, XX30QA, XX40PR, XX40QG, XX40RG,
XX30PY, XX40PQ, XX30QB, XX40QE, XX53PP,
XX49RA, XX40PJ, XX49QP, XX49QN, XX49QW,
XX49QT, XX49QS, XX40PF, XX49QU, XX49QJ,
XX49QY, XX49QQ, XX49QH, XX40EZ, XX49QZ,
XX49PW, XX49PN, XX40XA, XX49PR, XX49NN,
XX49QD, XX49QG, XX49PS, XX49DB, XX40EY,
XX49PX, XX40EN, XX40EW, XX49PT, XX49PQ,
XX49PP, XX49QE, XX49NX, XX40HA, XX49PH,
XX49DG, XX49PG, XX49RD, XX49NT, XX49NL,
XX30PQ, XX49PU, XX40EG, XX49NU, XX40HB,
XX49PZ, XX49PE, XX49PJ, XX49DE, XX49NH,
XX40EH, XX49NY, XX49DF, XX40XN, XX49PF,
XX49DJ, XX49NS, XX49NJ, XX49NP, XX40EB,
XX49PD, XX49HJ, XX49DQ, XX40HL, XX40ED,
XX49NZ, XX49HS, XX49NR, XX49QA, XX49HP,
XX49EA, XX40EJ, XX49XX, XX49HT, XX49EP,
XX40RR, XX49HH, XX49EB, XX49NQ, XX40HE,
XX49EW, XX49PA, XX53PD, XX40EA, XX49NF,
XX49HL, XX49DP, XX49HU, XX49DR, XX40HF,
XX49HW, XX49LL, XX49DW, XX40ER, XX49HY,
XX49NE, XX40EL, XX49HX, XX49EN, XX30SN,
XX30SW, XX40HJ, XX30PD, XX40UG, XX49HD,
XX49HQ, XX49HZ, XX49BZ, XX49DY, XX49HN,
XX49LU, XX40HG, XX30SG, XX40UH, XX49LJ,

XX49ND, XX49SS, XX30SD, XX49JB, XX49TB,
XX49JA, XX30NA, XX30SB, XX49TA, XX40HS,
XX30LT, XX49SA, XX30SQ, XX49DX, XX49HG,
XX30NR, XX49HE, XX30LS, XX49ED, XX49NB,
XX49HF, XX49DN, XX30LY, XX49LH, XX40HR,
XX49LS, XX49LT, XX30SF, XX49LQ, XX49DU,
XX49DZ, XX49HB, XX30LX, XX40HQ, XX49EL,
XX40EX, XX30ND, XX49SR, XX49SP, XX49LG,
XX49SZ, XX49SU, XX30LU, XX49SX, XX30SH,
XX30SE, XX49EJ, XX49SL, XX49EQ, XX30LP,
XX49LP, XX30NB, XX49JD, XX30LW, XX40EE,
XX49LZ, XX49NA, XX30LR, XX49SY, XX40RJ,
XX40DT, XX49JE, XX49LR, XX30LZ, XX40EF,
XX30NX, XX30LL, XX49EY, XX40ET, XX49EH,
XX40HP, XX30LN, XX40XR, XX40DU, XX49SW,
XX40DY, XX49DT, XX40ES, XX49LW, XX49SH,
XX40LL, XX49SN, XX30NG, XX40DS, XX40DZ,
XX49SJ, XX49LY, XX49SF, XX49EE, XX49EX,
XX49EF, XX30HF, XX40DR, XX30NE, XX49DS,
XX49SG, XX30LH, XX49JF, XX40DW, XX49SB,
XX30HG, XX49EU, XX49SD, XX49LX, XX49SE,
XX40LS, XX49SQ, XX30HH, XX40LN, XX30NQ,
XX49ET, XX30HB, XX40DX, XX30NF, XX49JP,
XX49JG, XX40LZ, XX49JT, XX30EX, XX30HQ,
XX40NB, XX30NH, XX49ES, XX49JR, XX39EA,
XX30PJ, XX49BY, XX40LU, XX30LQ, XX30EY,
XX30DH, XX30HE, XX30NJ, XX40LR, XX40LB,
XX49JS, XX49JJ, XX40LT, XX39EB, XX40NN,
XX40DP, XX30EZ, XX30EU, XX40NA, XX30DW,
XX40LW, XX40LP, XX30LF, XX39EE, XX30ER,
XX30HN, XX30HD, XX49JL, XX49JW, XX39ED,
XX40AQ, XX49JU, XX30EE, XX53PN, XX30HW,
XX49ER, XX49JN, XX30DJ, XX40AN, XX30EB,
XX49JZ, XX40LE, XX30NN, XX40LY, XX40LX,
XX30NW, XX30LE, XX40AH, XX39DZ, XX30HA,
XX30LB, XX49LB, XX30EP, XX48BJ, XX30EW,
XX39TA, XX30JA, XX30DQ, XX30PH, XX39EG,
XX39EF, XX30JD, XX49LF, XX40AJ, XX40BH,
XX30EQ, XX40BN, XX40BJ, XX30EF, XX30PE,
XX40BL, XX30ET, XX30DZ, XX49JX, XX48BG,
XX49JY, XX30EA, XX30EN, XX30DY, XX30DL,
XX40AW, XX39TE, XX48BH, XX30EG, XX49LD,
XX48BQ, XX39EJ, XX30DP, XX40LH, XX40LF,
XX49LE, XX40AS, XX40LD, XX30HZ, XX30EJ,
XX39EQ, XX30DR, XX48BL, XX48BD, XX40AT,
XX39EL, XX39DX, XX30JE, XX40LG, XX40BW,
XX40BG, XX30DX, XX30LD, XX40DN, XX30BB,
XX30EL, XX30DF, XX30DG, XX48BE, XX39TD,
XX40LJ, XX40AU, XX40AR, XX30JB, XX40DH,

XX30DT, XX30HY, XX30AG, XX30DU, XX40SE,
XX40DJ, XX40AP, XX30JF, XX30DB, XX40BR,
XX39SB, XX40DL, XX48BB, XX30JG, XX40JG,
XX40BS, XX39RD, XX48EZ, XX40TF, XX39EN,
XX40TE, XX30DE, XX40AY, XX39QW, XX40DQ,
XX40BE, XX40XX, XX48BN, XX39QN, XX48EE,
XX40JL, XX30JQ, XX48ET, XX39DU, XX30LA,
XX40JH, XX48BW, XX40NG, XX40JE, XX40AZ,
XX40BT, XX39TG, XX30JZ, XX39TH, XX40DG,
XX40DF, XX30JX, XX40DE, XX30DS, XX40BA,
XX30EH, XX39EP, XX30PU, XX39QL, XX48EY,
XX40JW, XX40BD, XX49BX, XX48BA, XX40DB,
XX39ER, XX40BB, XX48EX, XX30BG, XX30BQ,
XX39QP, XX40JF, XX40BU, XX30AE, XX39QJ,
XX39ET, XX30AW, XX30JH, XX40TH, XX40JT,
XX40BX, XX48BP, XX30JU, XX40DA, XX39QH,
XX40SG, XX40JN, XX40JZ, XX30BE, XX39RL,
XX48EG, XX30RA, XX48EQ, XX40ND, XX40BY,
XX48EL, XX30HT, XX40TP, XX40JA, XX48AZ,
XX39QZ, XX40NJ, XX39RB, XX48EN, XX40NE,
XX30AD, XX48AX, XX48EH, XX30AZ, XX53PB,
XX40NQ, XX30BH, XX30AH, XX40BZ, XX53PE,
XX30JY, XX40JB, XX39EU, XX40SH, XX40JR,
XX39QF, XX39QG, XX40JS, XX48BS, XX39ES,
XX48NQ, XX39HA, XX30JN, XX30BJ, XX30JR,
XX39DH, XX48BT, XX39EZ, XX30BS, XX48EW,
XX30AQ, XX48EJ, XX39QE, XX40JD, XX48ND,
XX40JP, XX30AB, XX39DQ, XX48AU, XX39DJ,
XX30HR, XX39RJ, XX39DG, XX30BT, XX40NX,
XX39QR, XX48ES, XX30BP, XX53PT, XX40NH,
XX40JY, XX48AW, XX39PY, XX39EX, XX48LY,
XX39HD, XX30JW, XX48NB, XX39BS, XX48JH,
XX30BR, XX39QS, XX49BS, XX39RH, XX39DE,
XX48LZ, XX48NF, XX48JQ, XX39HP, XX39HW,
XX48AN, XX48AT, XX40JX, XX48NA, XX48BU,
XX30BW, XX53PG, XX39EY, XX39DN, XX39BN,
XX40TA, XX39PZ, XX30BX, XX39BW, XX39BP,
XX39QT, XX39PX, XX48AP, XX30BN, XX39RX,
XX48AL, XX40PB, XX48DG, XX39DB, XX39DF,
XX40TD, XX48JG, XX48JF, XX48LU, XX48AJ,
XX39HF, XX39BT, XX53PJ, XX53PH, XX48DR,
XX30DA, XX48JN, XX39HG, XX39DA, XX39XX,
XX38LR, XX39HN, XX39PR, XX48DL, XX39BU,
XX30BU, XX48AS, XX48LR, XX48JE, XX39AX,
XX30AA, XX48AG, XX48DN, XX48HF, XX48HQ,
XX39PS, XX39RY, XX48JD, XX48DW, XX48JW,
XX38LX, XX39HL, XX48DH, XX48AF, XX39BZ,
XX48AR, XX40TB, XX39BJ, XX48HG, XX48LL,
XX39BG, XX38NE, XX39QA, XX48LP, XX38NB,

XX39PP, XX48UA, XX39BX, XX48AQ, XX53PL,
XX38NF, XX49AN, XX49BR, XX53QF, XX39HJ,
XX49BT, XX48DS, XX48JB, XX49AW, XX38LJ,
XX48DJ, XX39BY, XX30QY, XX39QY, XX39HR,
XX49BP, XX39BA, XX48UB, XX49AJ, XX48DT,
XX39AZ, XX38LH, XX38LS, XX38LU, XX48AH,
XX38JR, XX48AE, XX48LN, XX48LW, XX39HH,
XX38NA, XX38LP, XX48DU, XX38LL, XX48HN,
XX48HW, XX48HJ, XX48LJ, XX39AT, XX47DE,
XX39PU, XX49BN, XX48HS, XX48AD, XX49BL,
XX49BH, XX39RQ, XX38NG, XX38RS, XX39BE,
XX39PT, XX38ND, XX38LZ, XX38QJ, XX48PY,
XX39BH, XX48DX, XX39RF, XX38LQ, XX39AA,
XX39RU, XX48PU, XX39BQ, XX48QD, XX38JP,
XX48DE, XX49BJ, XX38LW, XX38JS, XX38JL,
XX48JA, XX39HS, XX38JJ, XX49BW, XX49BQ,
XX39NA, XX47DB, XX38RR, XX38LN, XX39BB,
XX49BG, XX38NQ, XX48HH, XX39AB, XX38LY,
XX48EB, XX48PT, XX38QL, XX48EA, XX39NB,
XX48PZ, XX38JT, XX39AR, XX39SA, XX48XG,
XX38JW, XX48DF, XX39BD, XX39NT, XX48PX,
XX48BX, XX38JY, XX47BG, XX49BA, XX38NS,
XX48DY, XX49BB, XX39AP, XX53NX, XX48HL,
XX38RP, XX47EJ, XX38JN, XX49BE, XX48QQ,
XX39AL, XX38QN, XX39AD, XX39JQ, XX39RG,
XX48DZ, XX48BZ, XX49AY, XX47ER, XX39AN,
XX47EH, XX47BE, XX39HZ, XX38LG, XX49AZ,
XX38JX, XX47EW, XX39AW, XX47EN, XX39LZ,
XX38RW, XX47TX, XX38NX, XX48QA, XX38QW,
XX38NU, XX48PS, XX53NY, XX48QE, XX47EG,
XX38NT, XX39JG, XX49BD, XX39AE, XX39HT,
XX38JU, XX47BX, XX38NL, XX39RN, XX49AX,
XX48PD, XX39AQ, XX49AH, XX48NZ, XX39HX,
XX38RN, XX39AJ, XX39NU, XX38JZ, XX38RJ,
XX39AH, XX49AE, XX47EF, XX48HY, XX49AT,
XX47JD, XX39JH, XX39NZ, XX48PB, XX39NX,
XX38LD, XX48HX, XX39RW, XX48HU, XX38JH,
XX47EQ, XX38QP, XX48DB, XX47EL, XX47BR,
XX38PJ, XX48HZ, XX49AG, XX48XB, XX48PP,
XX39BL, XX49AS, XX39RS, XX49AF, XX39PN,
XX48NY, XX48PR, XX39ND, XX38NZ, XX38NR,
XX38QE, XX47BW, XX49AU, XX39RT, XX39JB,
XX38NY, XX38LA, XX38LB, XX39HY, XX48QF,
XX38NJ, XX38JQ, XX38QT, XX48QB, XX38PA,
XX47DX, XX48JR, XX39AF, XX38NP, XX38RL,
XX49AR, XX48QG, XX47BL, XX39JA, XX48PE,
XX49AQ, XX48JP, XX39JL, XX38NW, XX39NE,
XX47BQ, XX48PA, XX39JJ, XX48JS, XX38QF,
XX48RR, XX38RG, XX38QQ, XX48PN, XX38RH,

XX39RR, XX39JE, XX48NX, XX39LY, XX38QY,
XX48DA, XX39AG, XX23QX, XX47BJ, XX39DS,
XX48PL, XX38NH, XX23QY, XX48PF, XX23RA,
XX38RQ, XX39PD, XX39PA, XX38QG, XX39PE,
XX48RP, XX23QZ, XX38PZ, XX38PD, XX38PE,
XX47DT, XX47AN, XX48RS, XX49AD, XX39PL,
XX38QA, XX39JD, XX48QL, XX37NU, XX49AA,
XX39JP, XX47EE, XX38QX, XX39JS, XX23RD,
XX49RE, XX38TJ, XX38PG, XX48PG, XX48RW,
XX23RB, XX47AP, XX47DZ, XX38JB, XX49XA,
XX38PW, XX39JT, XX39JF, XX48QJ, XX39NF,
XX37PA, XX38RF, XX39NG, XX47AJ, XX47BA,
XX48RN, XX47DY, XX48SN, XX38PQ, XX39PG,
XX47JA, XX23RE, XX37NZ, XX37NX, XX23RF,
XX48PH, XX48QH, XX37SN, XX47BD, XX39JR,
XX47EY, XX38PF, XX23RH, XX39PH, XX47DU,
XX47DA, XX38PX, XX48NS, XX39NH, XX39JX,
XX39RP, XX48SW, XX47DS, XX37NY, XX39JW,
XX48NL, XX38RE, XX23RG, XX38PP, XX48RU,
XX38HL, XX39JU, XX38PY, XX37PD, XX38RB,
XX47AH, XX48PQ, XX39NJ, XX37NT, XX23RQ,
XX39LU, XX48SJ, XX39PQ, XX37LL, XX37SW,
XX39LX, XX48QW, XX23RL, XX37SZ, XX38PN,
XX37TA, XX37PL, XX47BU, XX23RJ, XX23RP,
XX47EA, XX48QN, XX39NQ, XX39LE, XX37LJ,
XX37SX, XX47UP, XX30PW, XX47AQ, XX39JY,
XX47ED, XX39JZ, XX37PE, XX48NW, XX48PJ,
XX47EZ, XX39LT, XX38RA, XX48NR, XX48QX,
XX23RR, XX47HB, XX38HP, XX48SH, XX48QS,
XX47AG, XX48RY, XX39LS, XX48NH, XX39NW,
XX47SW, XX47EB, XX39LB, XX38PR, XX37PF,
XX48RX, XX37LP, XX47DR, XX37UE, XX37SY,
XX37PJ, XX39LA, XX47HA, XX38HR, XX23RS,
XX23RU, XX23RN, XX47AA, XX38PU, XX48QR,
XX39LR, XX37LZ, XX38HW, XX47AZ, XX23RT,
XX47AF, XX39LW, XX47TU, XX53PA, XX38PS,
XX39NN, XX48NP, XX37LY, XX47AY, XX47AE,
XX37UD, XX37PH, XX48RZ, XX48QU, XX37LQ,
XX47DN, XX47DQ, XX48NN, XX48SG, XX47TW,
XX39NL, XX48RG, XX47DG, XX37NR, XX47TS,
XX39LD, XX38BE, XX37LN, XX47DP, XX37NS,
XX37NA, XX37NQ, XX23RW, XX53PF, XX38PT,
XX48SQ, XX23TJ, XX47DL, XX38PH, XX47DJ,
XX38BD, XX47TP, XX47TT, XX48QT, XX23TF,
XX37LR, XX39LP, XX37LW, XX47DH, XX38AH,
XX39LF, XX38AG, XX23RY, XX39LG, XX47SN,
XX47AT, XX48SA, XX23RX, XX37LG, XX37NB,
XX37LX, XX23TH, XX23TL, XX37QU, XX48QY,
XX47HY, XX38BG, XX47AD, XX30PS, XX47SZ,

XX38BB, XX38AZ, XX38BA, XX48UH, XX47TR,
XX37LS, XX23SA, XX37NF, XX39NP, XX38AT,
XX48SB, XX23SD, XX37LH, XX37NP, XX23RZ,
XX39LH, XX37ND, XX38EJ, XX48SD, XX48QZ,
XX47HD, XX38BH, XX47TB, XX37LE, XX38BP,
XX37LU, XX47TN, XX39LL, XX37LD, XX38AY,
XX37NN, XX39NR, XX48RE, XX47TQ, XX38BQ,
XX38AJ, XX47TA, XX48UY, XX23TE, XX37NE,
XX37LT, XX38AS, XX47ST, XX48RQ, XX48RF,
XX38BJ, XX23SB
Coordinates / radius..: (341908,737855) / 6152.84
Population.....: 2435
Number of cases.....: 159 (114.32 expected)
Overall relative risk.: 1.391
Log likelihood ratio..: 12.265537
Monte Carlo rank.....: 13/1000
P-value.....: 0.013
SECONDARY CLUSTERS
2.Census areas included.: XX23LF, XX23LG
Coordinates / radius..: (338090,732030) / 50.00
Population.....: 4
Number of cases.....: 3 (0.19 expected)
Overall relative risk.: 15.975
Log likelihood ratio..: 6.985582
Monte Carlo rank.....: 744/1000
P-value.....: 0.744
3.Census areas included.: XX22BU
Coordinates / radius..: (338430,730700) / 0.00
Population.....: 2
Number of cases.....: 2 (0.09 expected)
Overall relative risk.: 21.300
Log likelihood ratio..: 6.122324
Monte Carlo rank.....: 989/1000
P-value.....: 0.989
The log likelihood ratio value required for an observed
cluster to be significant at level
... 0.01: 12.554408
... 0.05: 10.652960

For further study using a GIS or database program, an ASCII
format GIS file has been created, describing the detected clusters.
The name of this file is E:\work\nonhodgpost79.gis.

PARAMETER SETTINGS
Input Files

Case File : E:\work\nonhodgpost79.cas
Control File : E:\work\cont.post79.ctl
Coordinates File : E:\work\work.coord.geo

Precision of Times : Years
Coordinates : Cartesian
Analysis
———
Type of Analysis : Purely Spatial
Probability Model : Bernoulli
Scan for Areas with : High Rates
Start Date : 1979/1/1
End Date : 1996/12/31
Number of Replications : 999
Scanning Window
——————
Maximum Spatial Cluster Size : 50.00
Output
———
Results File : E:\work\nonhodgpost79.res
GIS File : E:\work\nonhodgpost79.gis
———————————————————————————————————————
Program completed : Sat Nov 03 16:12:12 2001
Total Running Time : 13 minutes 56 seconds