Knorr-Held, Rasser:

Bayesian Detection of Clusters and Discontinuities in Disease Maps: Simulations. (REVISED, June 1999)

Sonderforschungsbereich 386, Paper 142 (1999)

Online unter: http://epub.ub.uni-muenchen.de/

Bayesian Detection of Clusters and Discontinuities in Disease Maps

Simulations

Leonhard Knorr–Held and Günter Raßer

Institut für Statistik

Universität München

Ludwigstr. 33, 80539 München

Germany

Email: leo@stat.uni-muenchen.de rasser@stat.uni-muenchen.de

Abstract

This paper is a supplement paper to Knorr-Held and Raßer (1999), "Bayesian Detection of Clusters and Discontinuities in Disease Maps", which describes a novel approach to disease mapping with particular emphasis on the detection of clusters and discontinuities in disease maps. First we investigate several features of the prior by simulations from the clustering model both for the 544 regions of Germany and for the 88 counties of Ohio state. Furthermore the method is applied to various artificial datasets for Ohio with and without spatial patterns. Sensitivity with respect to prior parameters is studied.

1 Simulations from the prior

Here are various simulations from the prior distribution for clustering, both for Germany and for Ohio. The number of neighbors of the regions varies between 1 and 11 in Germany and between 3 and 8 in Ohio. All results were calculated using 100,000 samples, in each one drawing the number of clusters k from the prior distribution (uniform or geometric), the cluster centers uniformly from the the set of regions, and finally assigning each of the remaining regions to a cluster center as described in Knorr-Held and Raßer (1999). From all samples we have calculated

- the empirical distribution of the number of clusters k, which of course is very similar to the theoretical distribution,
- the probability for each region to be selected as a cluster center, which should be (and actually is) uniformly distributed on the total number of regions,
- the probability for each region to be alone in a cluster of size one,

- the average probability for each region to be together with one of its first, second or third order neighbors and
- for each region the average size of the cluster it is assigned to.

To handle the large number of regions (544 in Germany) for the visual presentation of the results we have grouped the regions depending on the number of neighbors.

1.1 Results for Germany

Here are the results of the simulation using a geometric distribution with parameter $c = 0.02$, which is the prior distribution for k we used in the application presented in Knorr-Held and Raßer (1999). The expected number of clusters in this case is nearly 50. With 544 regions the probability for one region to be selected as a cluster center should therefore be approximately 0.1 for all regions. Figure 1 shows a plot of the probabilities for all regions. Not surprisingly we really get an uniform distribution. A plot of the empirical distribution of the number of clusters is also included. For the next results we have grouped the regions depending on the number of neighbors, so the distribution of the number of neighbors might be helpful. From Figure 2 it is obvious that for the extreme cases of 10 or 11 neighbors there are only 9 and 2 observations, respectively, so the results for these cases may be less reliable than for the other regions with a higher number of neighbors. The probability for being alone in a cluster is near to zero for all regions since the average number of clusters is quite small. Yet Figure 3 shows that with growing number of neighbors the probability is getting smaller. This seems intuitive since for a selected region to form a cluster of its own it is necessary that one or more of the regions with small distance to the selected one (mostly neighbors) are cluster centers themselves and therefore regions with a small number of neighbors will more often be alone than regions with a higher number of neighbors.

We also have calculated the probability that a region is in the same cluster as one of its first order neighbors (= neighbors), averaging over all neighbors and all samples. The same has been done for the second order neighbors (= neighbors of neighbors) and the third order neighbors. The results are also plotted in Figure 3. Since the probability of being alone is nearly zero, the probability of being together with one of the neighbors is extremely high. And of course for the extreme case of only one neighbor both probabilities add to 1. All probabilities are getting lower with growing number of neighbors. Only for 10 and 11 neighbors the probabilities are rising again. This may be pure incidence since the number of regions in these groups is very small as mentioned before. A more detailed view on the probabilities for being alone in a cluster is given in Figure 4. The width of the boxplots is proportional to the number of regions in each group. The boxplot on the right shows that, considering all regions, the probabilities are not varying too much. Only the regions with just one neighbor have an unusual high probability. There are also included boxplots for the average size of the cluster the region is assigned to. Obviously, the influence of the number of neighbors is small here.

We have done another simulation using an uniform distribution instead of the geometric prior for k. Altogether the results were quite similar. The probability for one region to be selected as a cluster center should now be exactly 0.5, which can be observed in Figure 5. The empirical distribution of k seems to be noisier, so please note the changed range of the y-axis. The expected number of clusters is now 272.5, so the probabilities of being alone in a cluster will be higher, while the probabilities for being together with the first, second or third order neighbors will be smaller, which becomes obvious from Figure 6. Finally Figure 7 shows that the average size of the clusters are also much smaller than before.

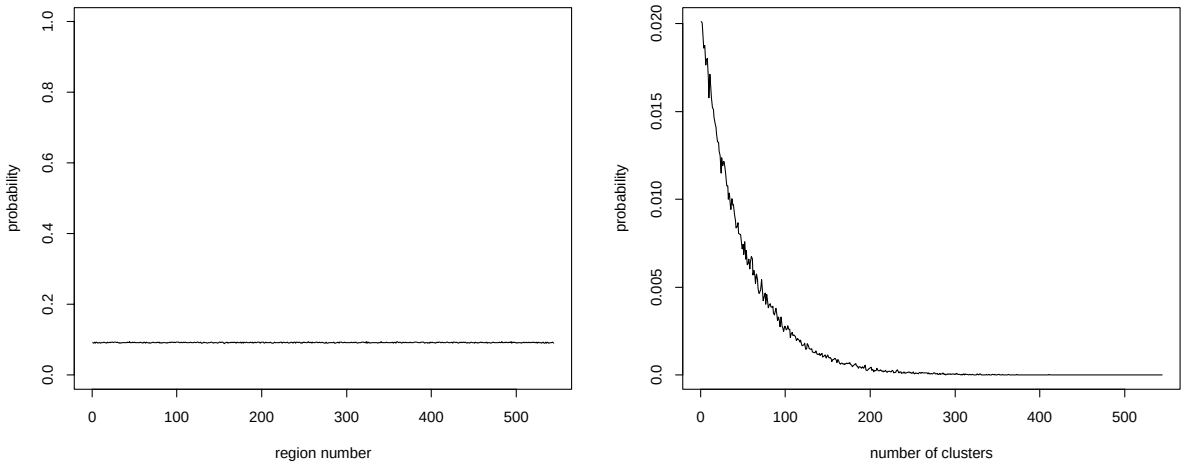


Figure 1: Probabilities for each region to be selected as a cluster center (left) and the distribution of the number of clusters k (right).

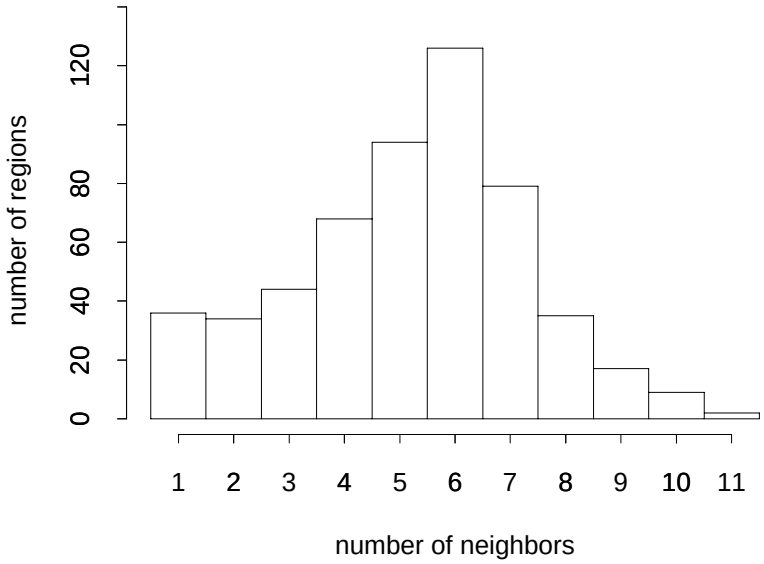


Figure 2: Distribution of the number of neighbors for the 544 regions in Germany.

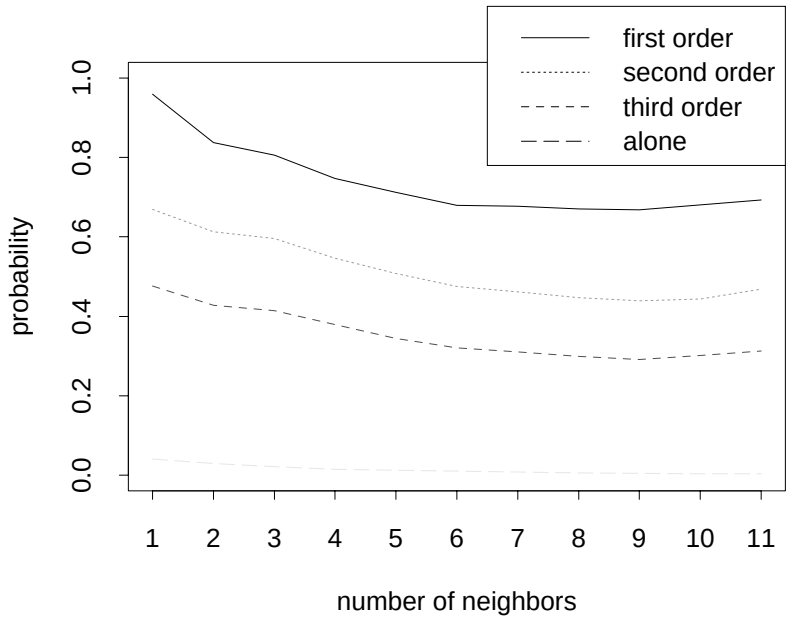


Figure 3: Probabilities for being alone in one cluster, as well as the average probabilities for being together with the first, second or third order neighbors, grouped after the number of neighbors.

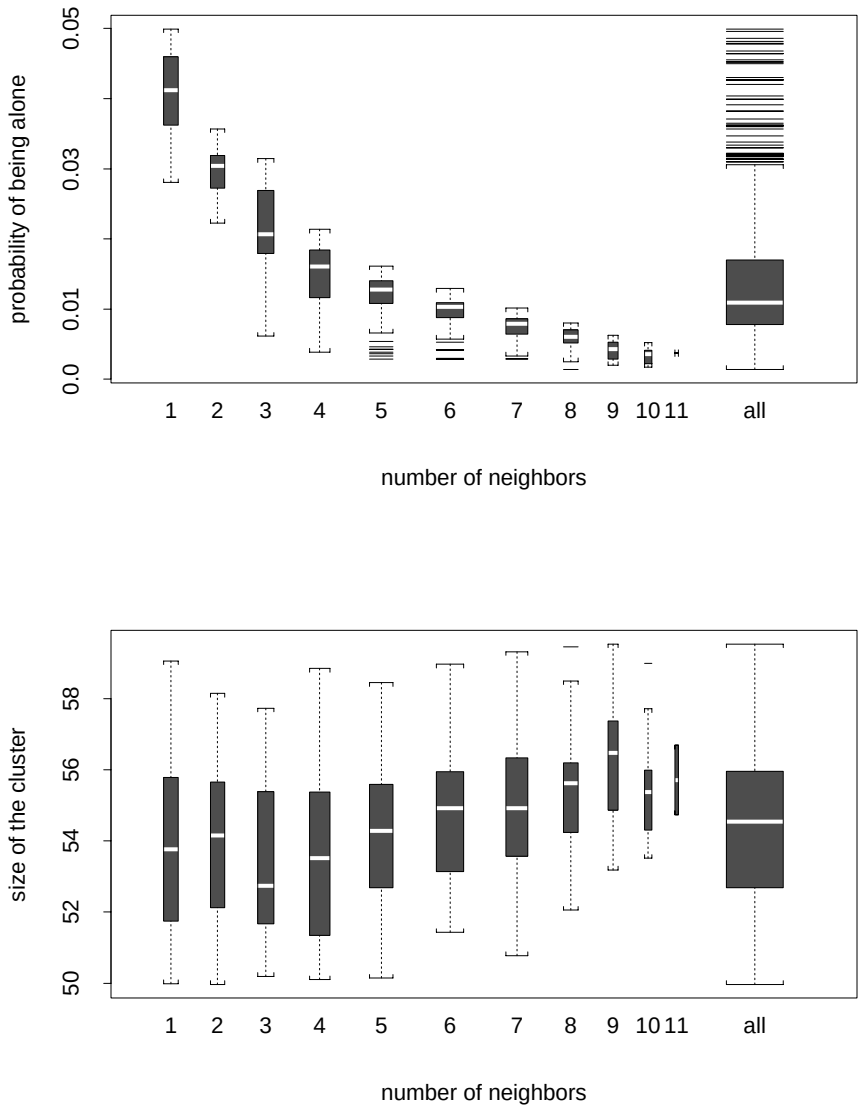


Figure 4: Boxplots of the probability for being alone in one cluster (above) and boxplots for the average size of the cluster the region is in (below).

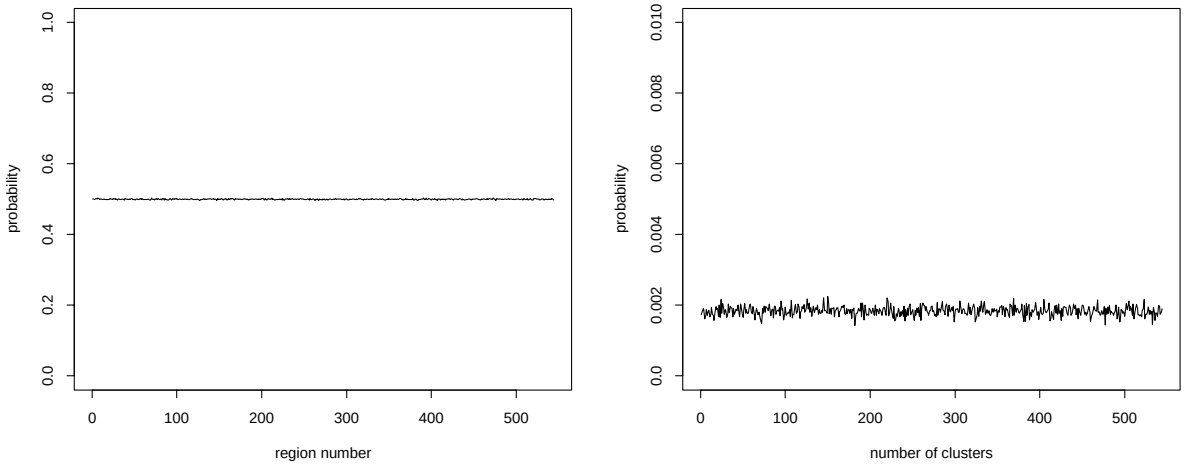


Figure 5: Probabilities for each region to be selected as a cluster center (left) and the distribution of the number of clusters k (right).

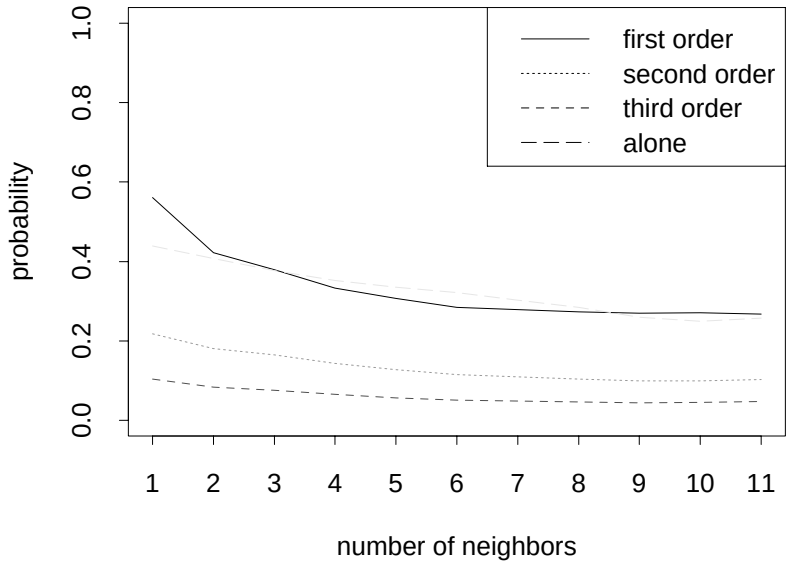


Figure 6: Probabilities for being alone in one cluster, as well as the average probabilities for being together with the first, second or third order neighbors, grouped after the number of neighbors.

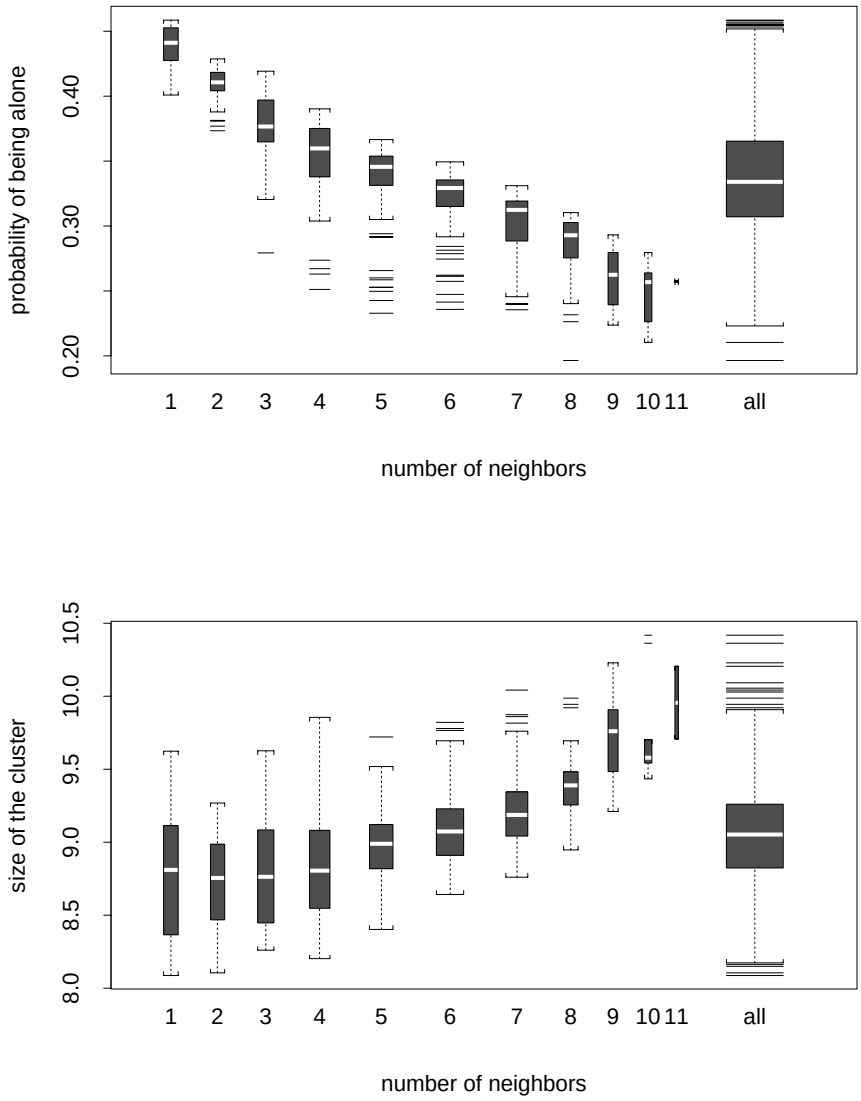


Figure 7: Boxplots of the probability for being alone in one cluster (above) and boxplots for the average size of the cluster the region is in (below).

1.2 Results for Ohio

In Ohio there are only 88 regions and the structure is more regularly than in Germany. The minimum number of neighbors here is 3, while the maximum number is 8, but this time there is only one region with the maximum number of neighbors. Figure 8 shows this distribution, where just like for Germany, we have a modus of 6 neighbors. Again we have applied an uniform and a geometric prior for k. The parameter for the latter one were chosen as $c = 0.1$ and $c = 0.5$, so the expected number of clusters are around 10 and 2, respectively, while it is now 44.5 for the uniform distribution. We have included here the same Figures as for Germany, again using 100,000 samples for each simulation, starting with the results from the geometric prior with parameter $c = 0.1$, which seems to be the best choice according to the simulated datasets presented later. The second simulation indicates that a geometric distribution with parameter $c = 0.5$ is very strong, and should therefore only be used if one has knowledge about the existence of a spatial structure in the risk factors, although this is not reflected by the data. Surprisingly the results from the third simulation with an uniform prior are much the same as for the first simulation.

Generally the results are quite the same as for Germany, though this time the typical features are not so obvious. But since the total number of regions is now much lower the number of regions in the single groups may be too small to draw general conclusions.

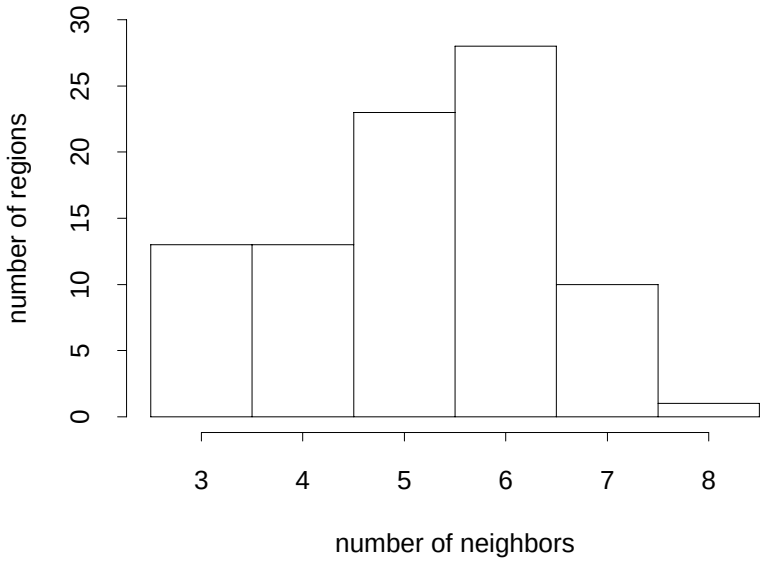


Figure 8: Distribution of the number of neighbors for the 88 regions in Ohio.

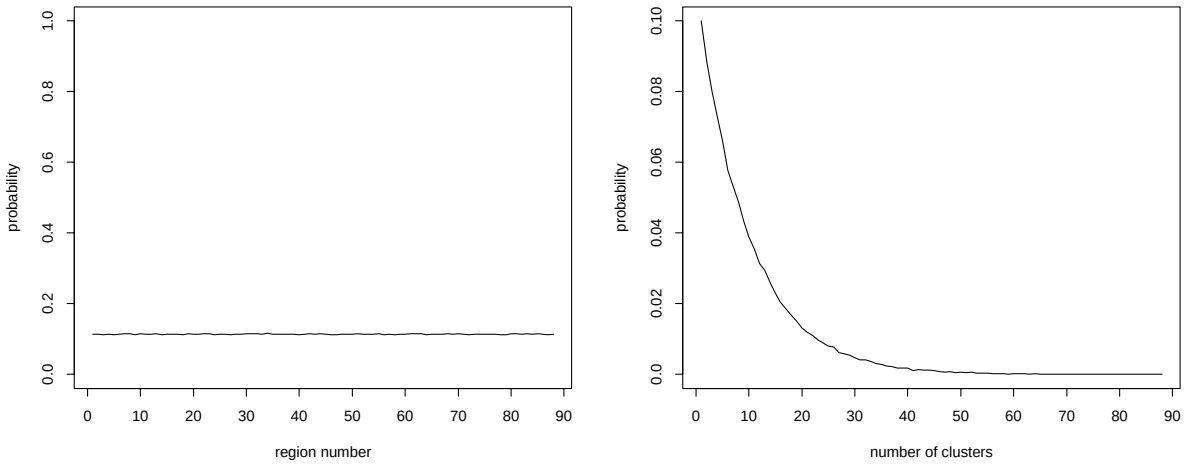


Figure 9: Probabilities for each region to be selected as a cluster center (left) and the distribution of the number of clusters k (right), using a geometric prior for k with parameter $c = 0.1$.

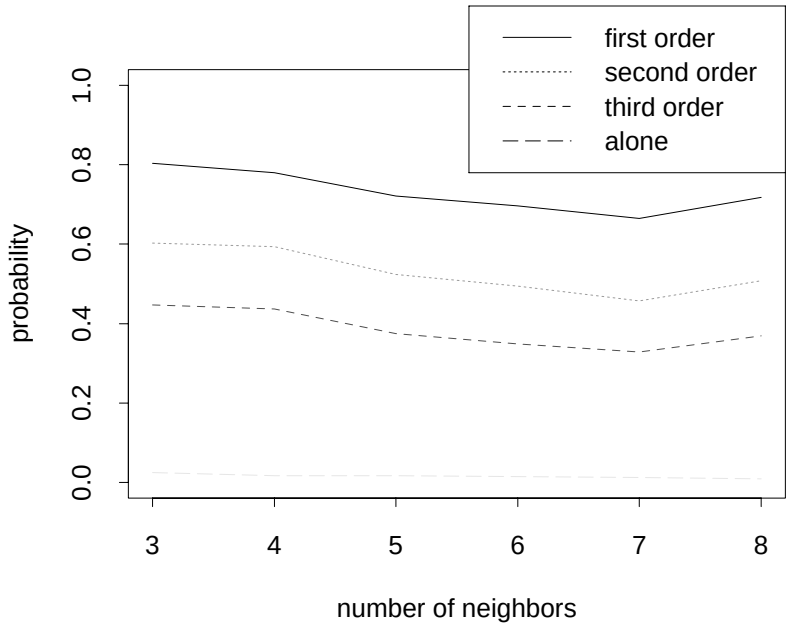


Figure 10: Probabilities for being alone in one cluster, as well as the average probabilities for being together with the first, second or third order neighbors, grouped after the number of neighbors. Same prior as in Figure 9.

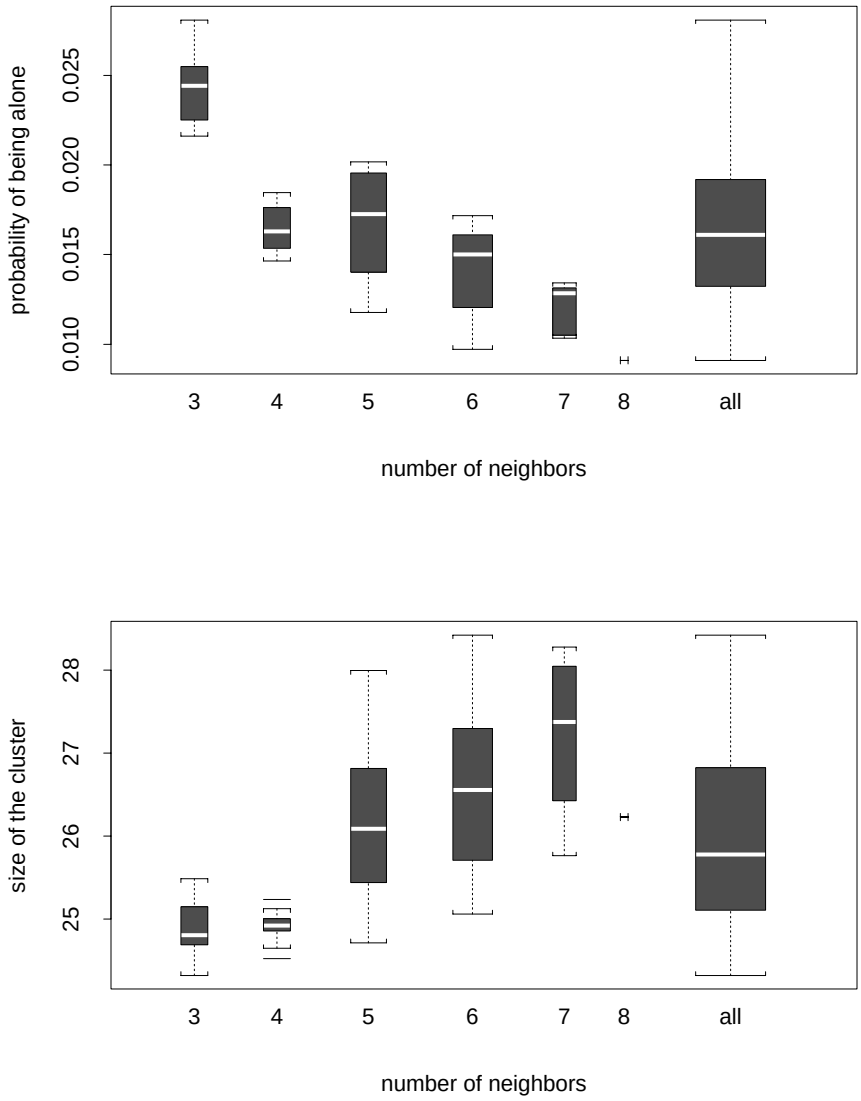


Figure 11: Boxplots of the probability for being alone in one cluster (above) and boxplots for the average size of the cluster the region is in (below). Same prior as in Figure 9.

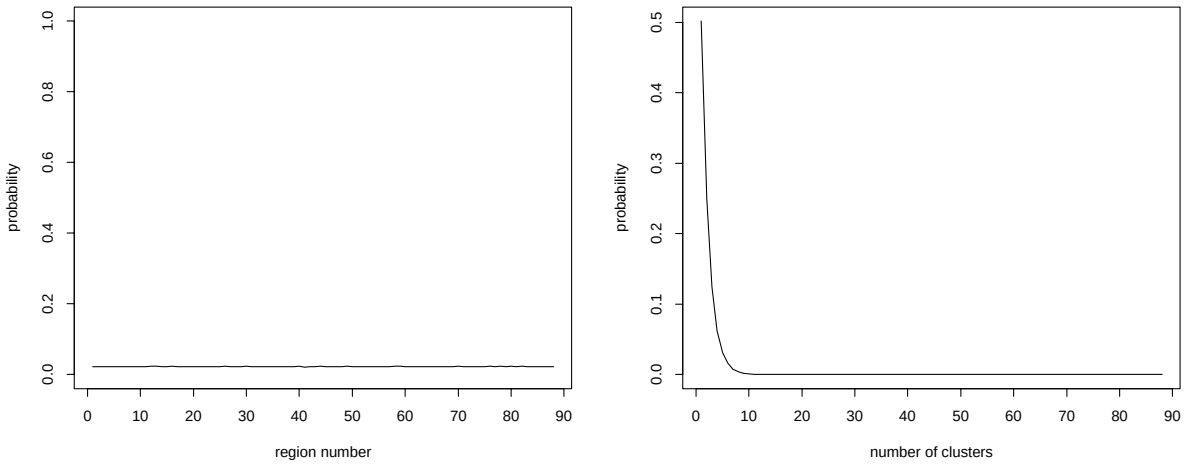


Figure 12: Probabilities for each region to be selected as a cluster center (left) and the distribution of the number of clusters k (right), using a geometric prior for k with parameter $c = 0.5$.

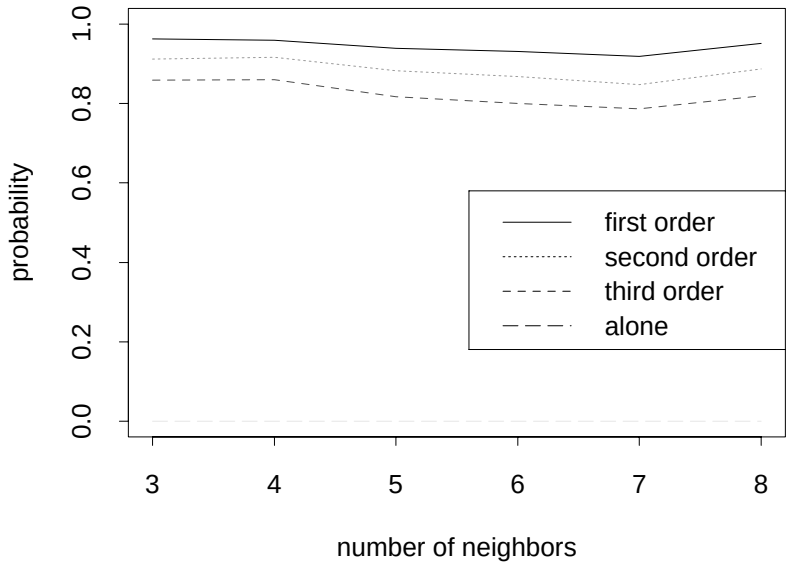


Figure 13: Probabilities for being alone in one cluster, as well as the average probabilities for being together with the first, second or third order neighbors, grouped after the number of neighbors. Same prior as in Figure 12.

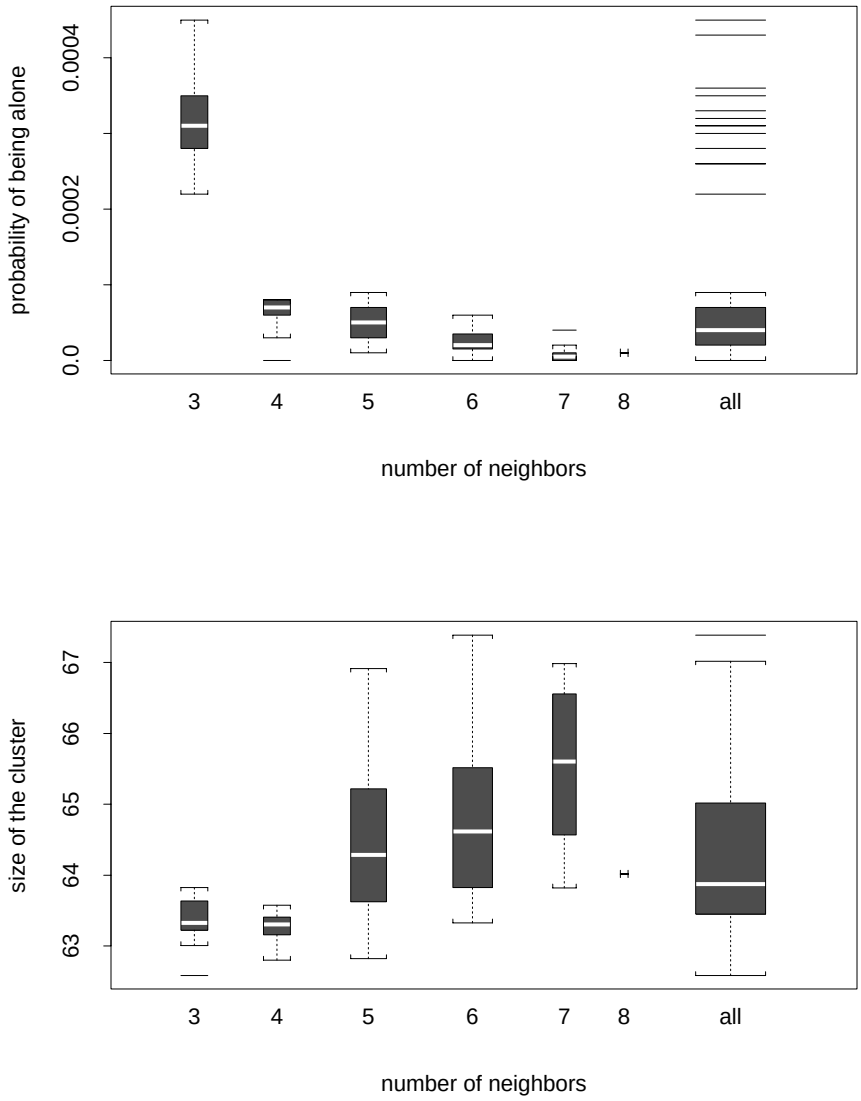


Figure 14: Boxplots of the probability for being alone in one cluster (above) and boxplots for the average size of the cluster the region is in (below). Same prior as in Figure 12.

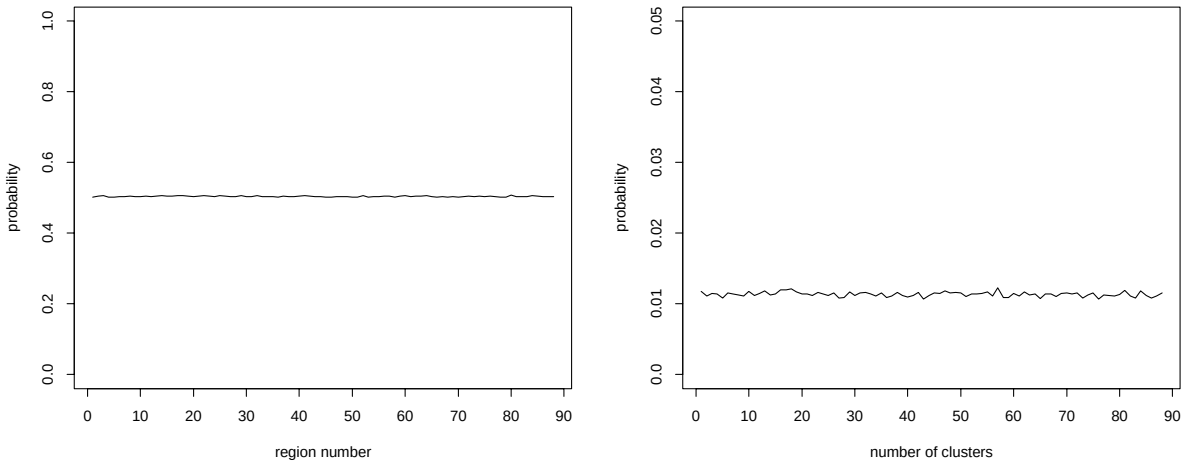


Figure 15: Probabilities for each region to be selected as a cluster center (left) and the distribution of the number of clusters k (right), using an uniform prior for k.

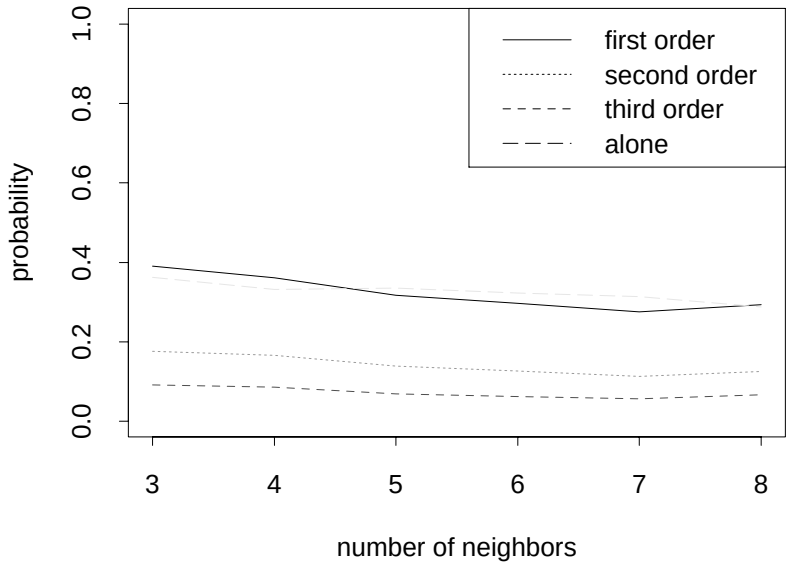


Figure 16: Probabilities for being alone in one cluster, as well as the average probabilities for being together with the first, second or third order neighbors, grouped after the number of neighbors. Same prior as in Figure 15.

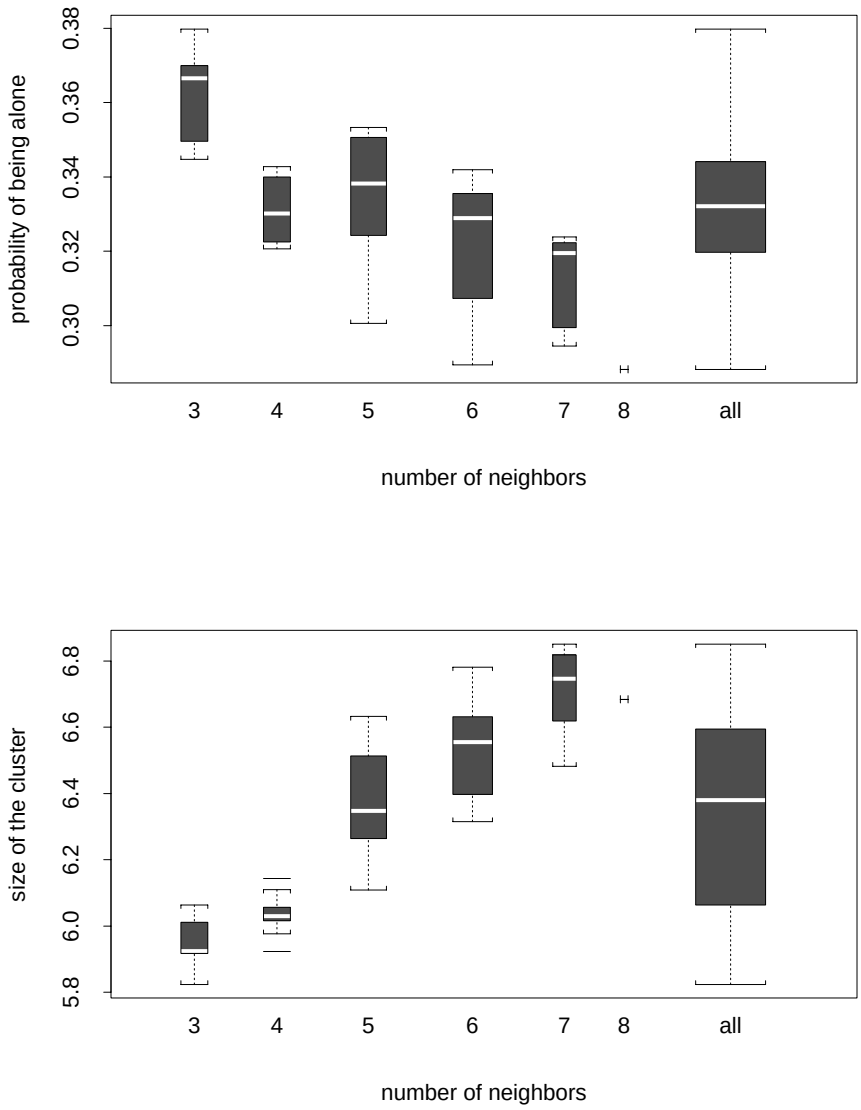


Figure 17: Boxplots of the probability for being alone in one cluster (above) and boxplots for the average size of the cluster the region is in (below). Same prior as in Figure 15.

2 Simulated datasets

To investigate the performance of the algorithm in reconstructing the true relative risks, we defined four different risk patterns and simulated datasets based on these patterns. For this purpose we used a dataset from Ohio instead of Germany in order to simplify the presentation of the results.

We constructed four different risk patterns,

(P1) constant risk over the whole state,

(P2) two different risk levels in areas of different shapes and sizes,

(P3) slowly rising risk from the west to the east,

(P4) heterogeneity in all regions.

A detailed description of each is given together with the results later. Based on these given risk patterns the observed number of cases for each region was drawn from a Poisson distribution with parameter $e_i h_i$, where e_i denotes the expected number of cases (based on lung cancer rates for white males in 1968) and h_i the relative risk in region i according to the given pattern. For each pattern these simulations have been done three times, so we get a total of 12 different datasets. We also used three different prior distribution for the number of clusters k,

(C1) a uniform distribution on $\{1, \ldots, 88\}$,

(C2) a geometric distribution with parameter $c = 0.1$,

(C3) a geometric distribution with parameter $c = 0.5$.

For the inverse gamma distribution (the prior distribution for the dispersion parameter σ^2 of the log-normal prior for the heights) we also needed to specify the parameters, using

(S1) IG(0.25,0.00025),

(S2) IG(1,0.01),

(S3) IG(5,0.125).

This means we get 9 possible combinations of prior distributions and therefore 108 runs of the algorithm, applying each prior combination to each dataset. In the following sections results will be given, showing plots of the

(a) true underlying risk pattern,

(b) simulated dataset,

(c) posterior median estimates of the relative risks,

(d) prior and posterior distribution for k,

(e) scatter plot of SMR's versus true risks

(f) scatter plot of median estimates versus true risks

(g) scatter plot of SMR's versus log expected values

(h) scatter plot of median estimates versus log expected values

(i) values of k for all iterations,

(k) values of σ^2 for all iterations,

(l) deviance of the model for all iterations,

(m) values of the (unnormalized) posterior for all iterations,

(n) autocorrelation-function of k,

(o) autocorrelation-function of σ^2,

(p) autocorrelation-function of the deviance,

(q) autocorrelation-function of the posterior,

(r) variance σ^2 versus k,

(s) deviance versus k,

(t) posterior versus k,

(u) acceptance-rates for the six moves.

Note that in Figure (u) the six moves are represented by numbers in the order (1) hyper, (2) height, (3) shift, (4) birth, (5) death and (6) switch. The acceptance probability for the hyper move of course is always 1. Since the output of one single run is very rich, we present only results from some selected runs. For each pattern (P1) - (P4) we present one run in full length and the estimates from the other two priors from (C1) - (C3) with the same prior (S1), (S2) or (S3).

All results were gained by runs of the algorithm using a burn-in length of 1,000,000 iterations. Form the total of 10,000,000 iterations after burn-in we used just each 10,000th step, so we get a sample size of 1000 iterations for each run.

Pattern 1

Here we assume the same relative risk 1.0 in all regions. We give three results for one dataset using the three different prior distributions for the number of clusters k along with a gamma prior distribution for σ^2 with parameters $a = 5$ and $b = 0.125$, i.e. (S3). In Figure 18 and on the following pages, full results are given for the analysis with a geometric prior distribution with parameter $c = 0.1$ for k. This prior distribution for k has an expected number of clusters of nearly 10, but for the posterior distribution the values are much lower, mostly 1. This means the algorithm was able to detect the pure randomness of the data variation and therefore we get a perfect reconstruction of the true risks. Since in this run there is mostly just one cluster, a death move is rarely proposed, so we get a high acceptance rate for this move, while a birth move has a very low acceptance rate, indicating that it is improbable to increase the number of clusters. For all other datasets with higher values of k, the acceptance rates for the birth and the death move were nearly the same.

Yet in this case the algorithm appeared quite insensitive to the choice of the prior for k. Figure 19 shows the data and the estimates from the analysis with an uniform prior for k. The posterior distribution for k here is bimodal with local modes at $k = 1$ and $k = n$. Nonetheless, the reconstruction of the true relative risks is very good. Finally, Figure 20 shows the results using a geometric prior distribution with parameter $c = 0.5$.

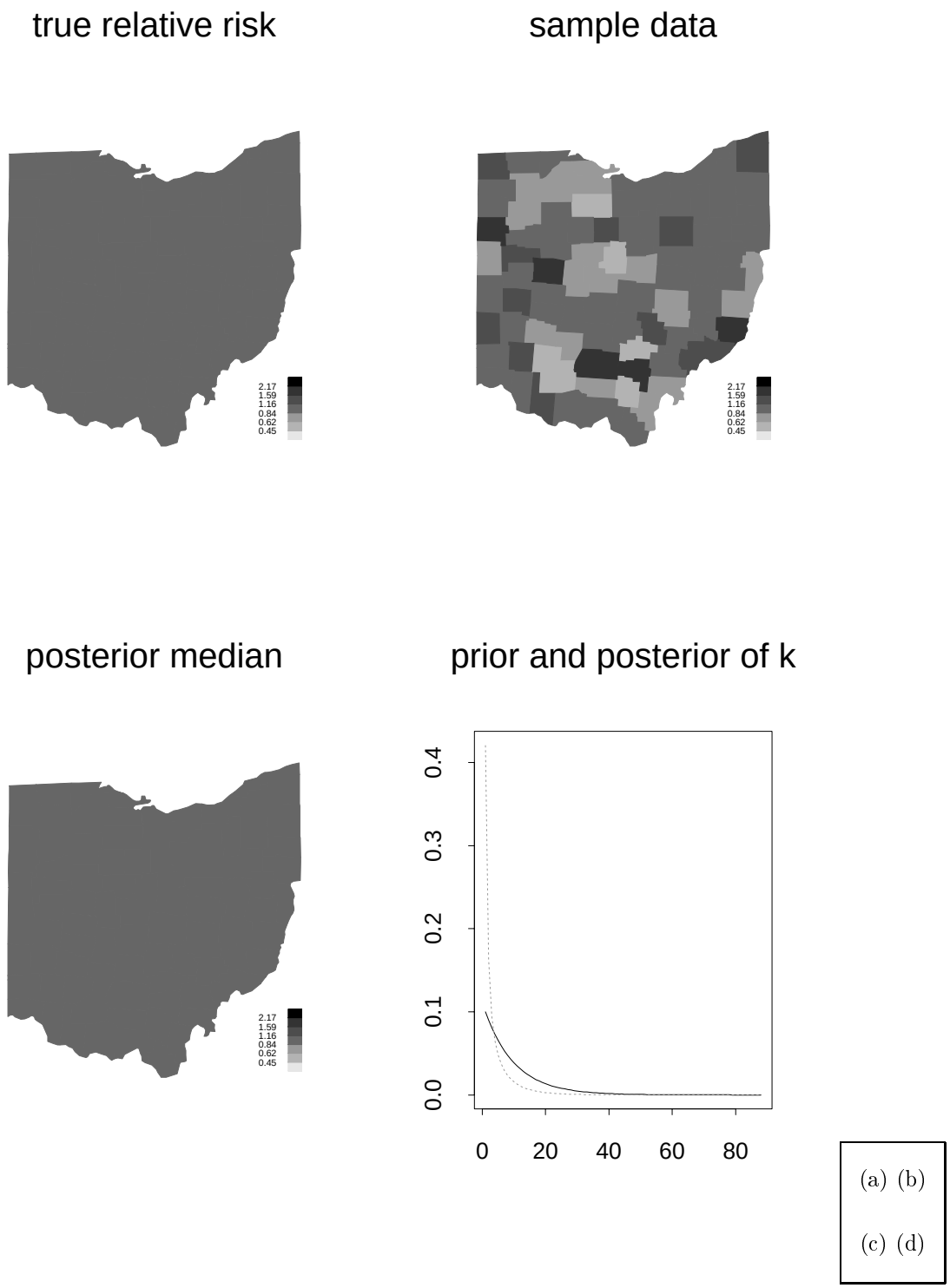
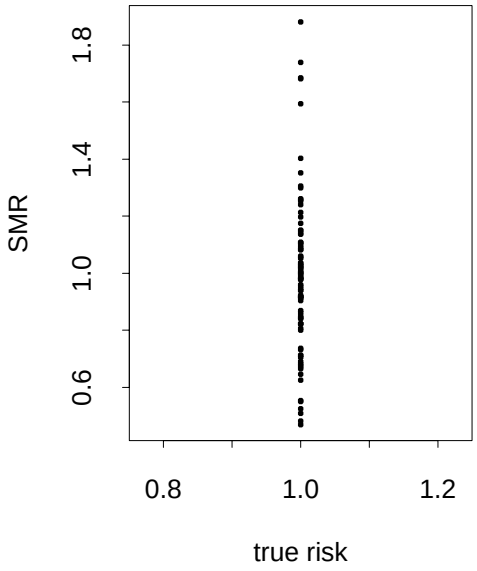
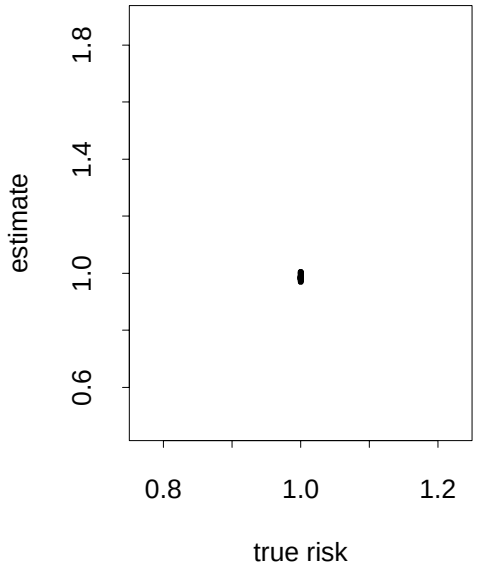
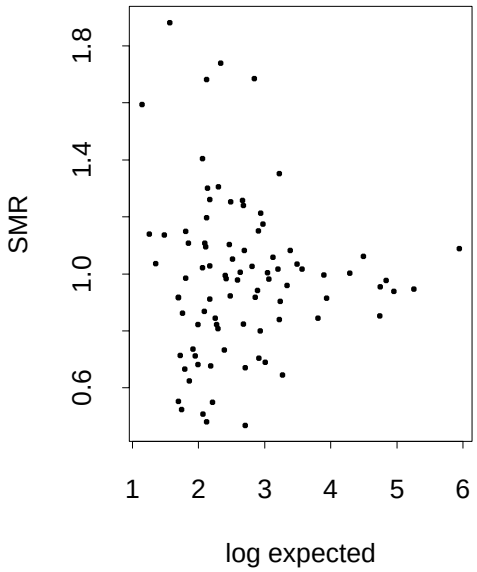
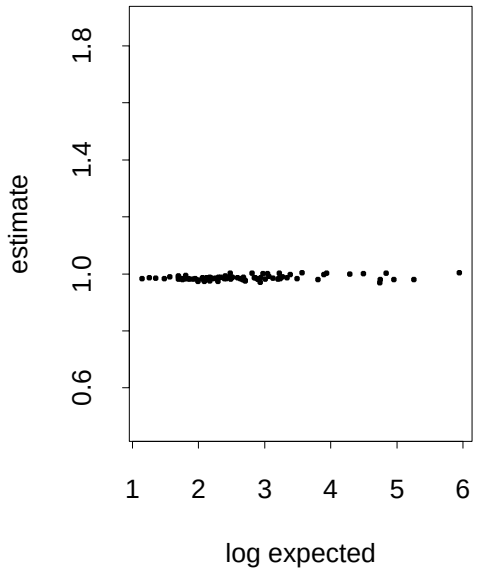


Figure 18: Analysis with (C2) and (S3).


SMR
0.6
1.0
1.4
1.8
0.8
1.0
1.2
true risk

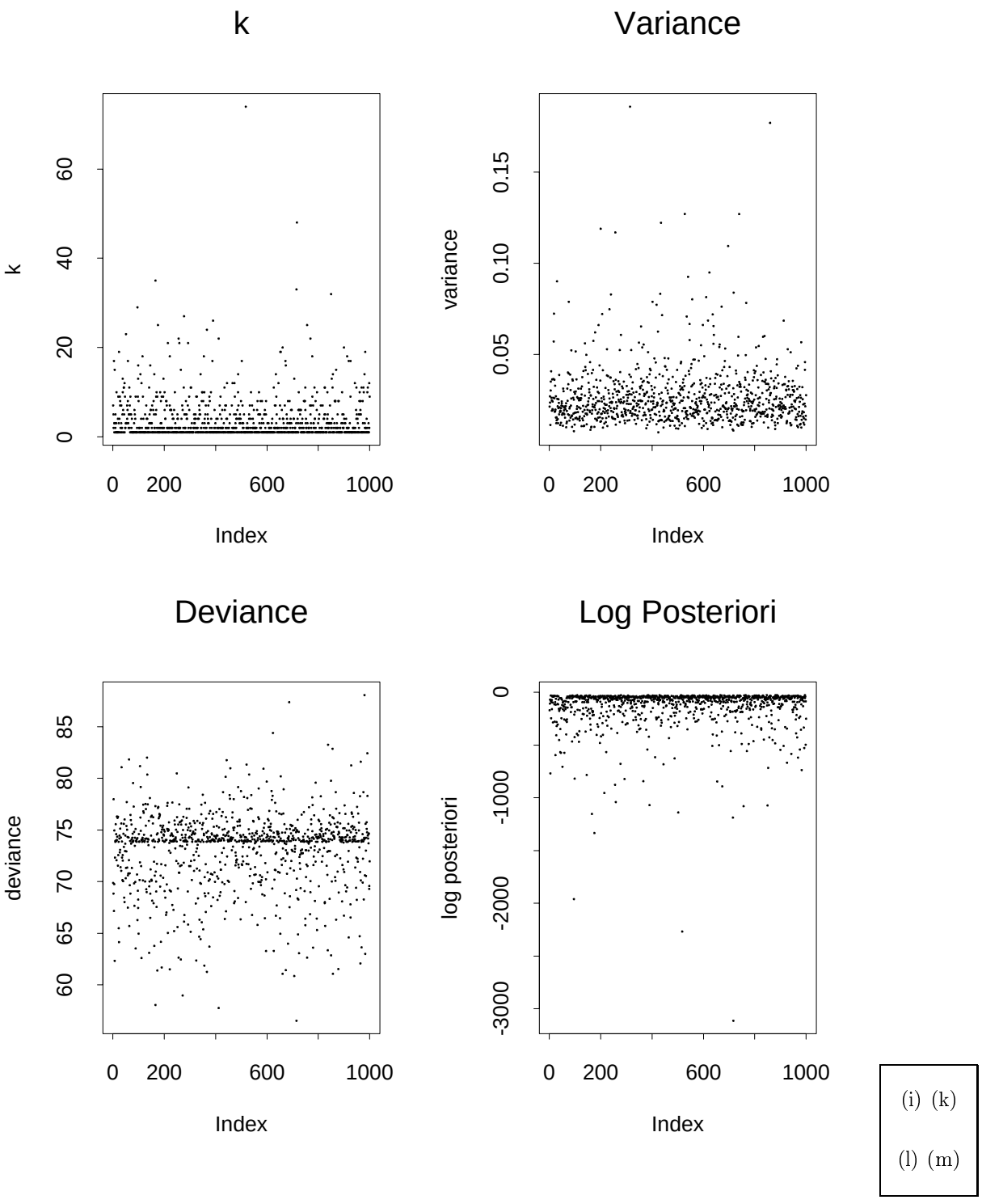

estimate
0.6
1.0
1.4
1.8
0.8
1.0
1.2
true risk

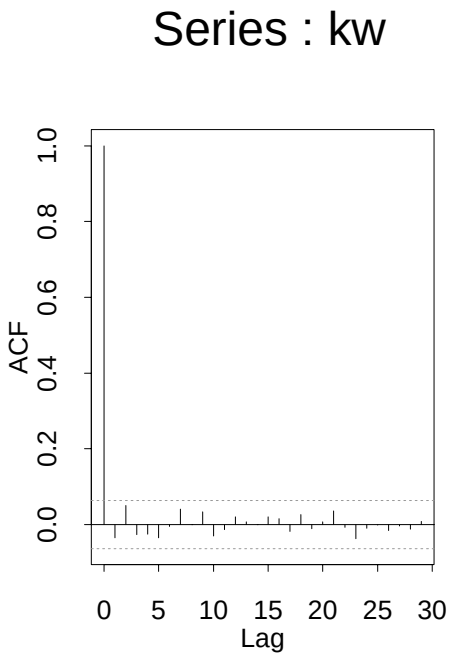

SMR
0.6
1.0
1.4
1.8
1
2
3
4
5
6
log expected

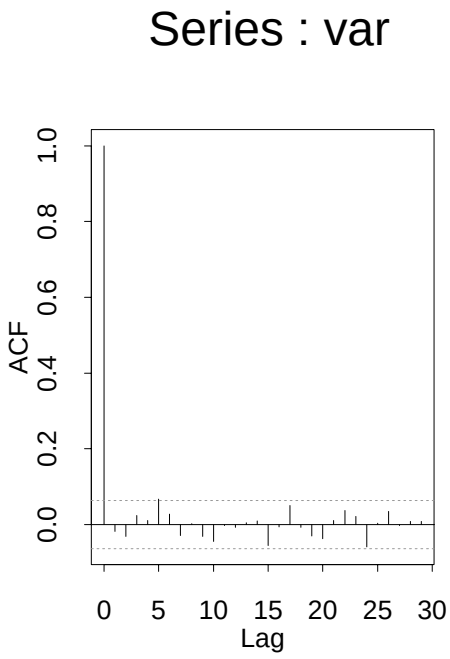

estimate
0.6
1.0
1.4
1.8
1
2
3
4
5
6
log expected

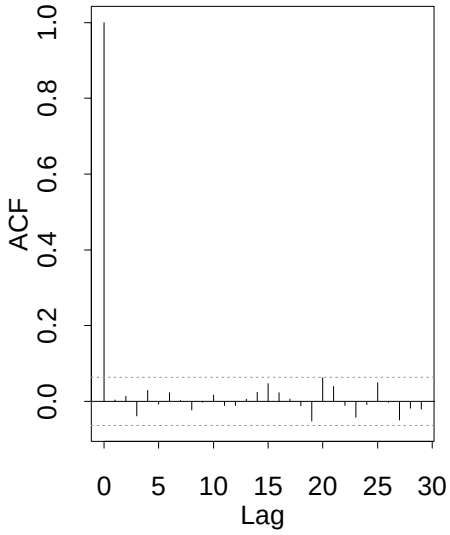
(e) (f)

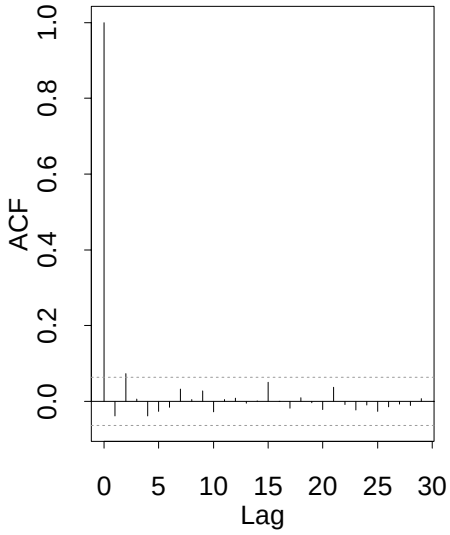
(g) (h)


k
k
0
20
40
60
0
200
600
1000
Index
Variance
variance
0.05
0.10
0.15
0
200
600
1000
Index
Deviance
deviance
60
65
70
75
80
85
0
200
600
1000
Index
Log Posteriori
log posteriori
-3000
-2000
-1000
0
0
200
600
1000
Index
(i) (k)
(l) (m)


Series : kw
ACF
0.0
0.2
0.4
0.6
0.8
1.0
0
5
10
15
20
25
30
Lag

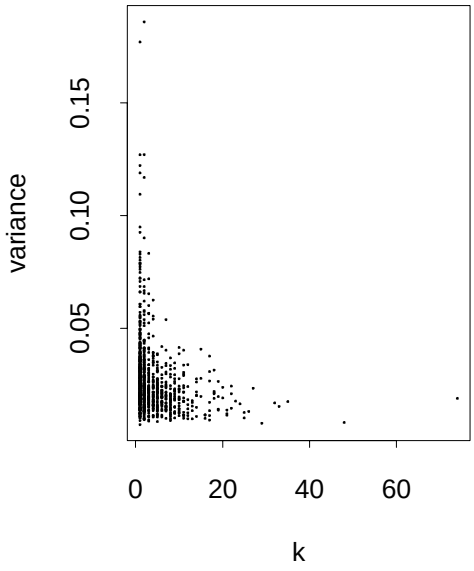

Series : var
ACF
0.0
0.2
0.4
0.6
0.8
1.0
0
5
10
15
20
25
30
Lag

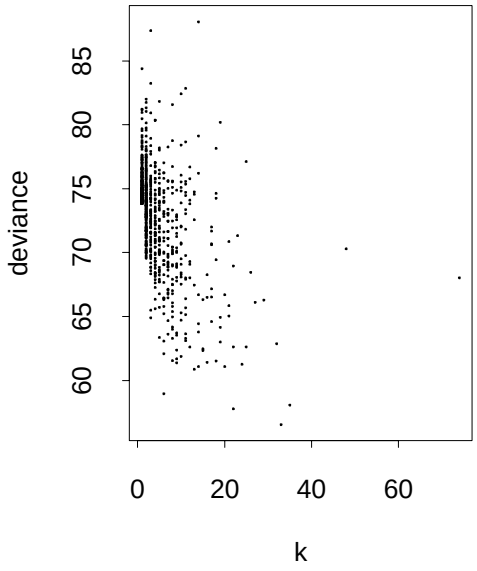

Series : dev
ACF
0.0
0.2
0.4
0.6
0.8
1.0
0
5
10
15
20
25
30
Lag


Series : post
ACF
0.0
0.2
0.4
0.6
0.8
1.0
0
5
10
15
20
25
30
Lag

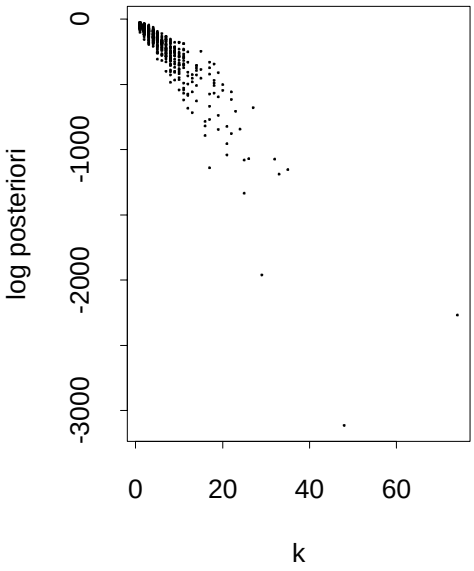
(n) (o)

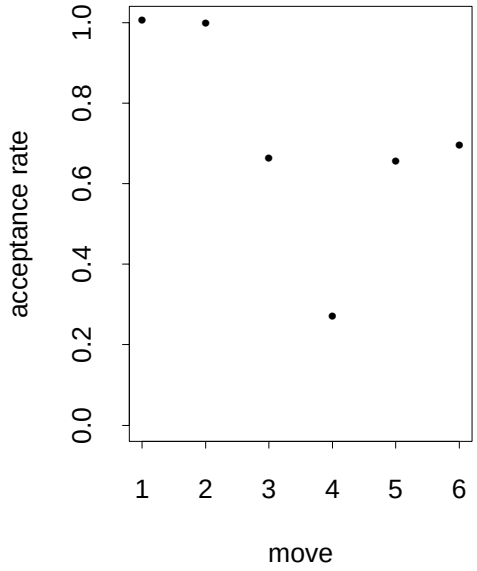
(p) (q)


variance
0.05
0.10
0.15
0
20
40
60
k


deviance
60
65
70
75
80
85
0
20
40
60
k

Rates


log posteriori
-3000
-2000
-1000
0
0
20
40
60
k


acceptance rate
0.0
0.2
0.4
0.6
0.8
1.0
1
2
3
4
5
6
move

(r) (s)

(t) (u)

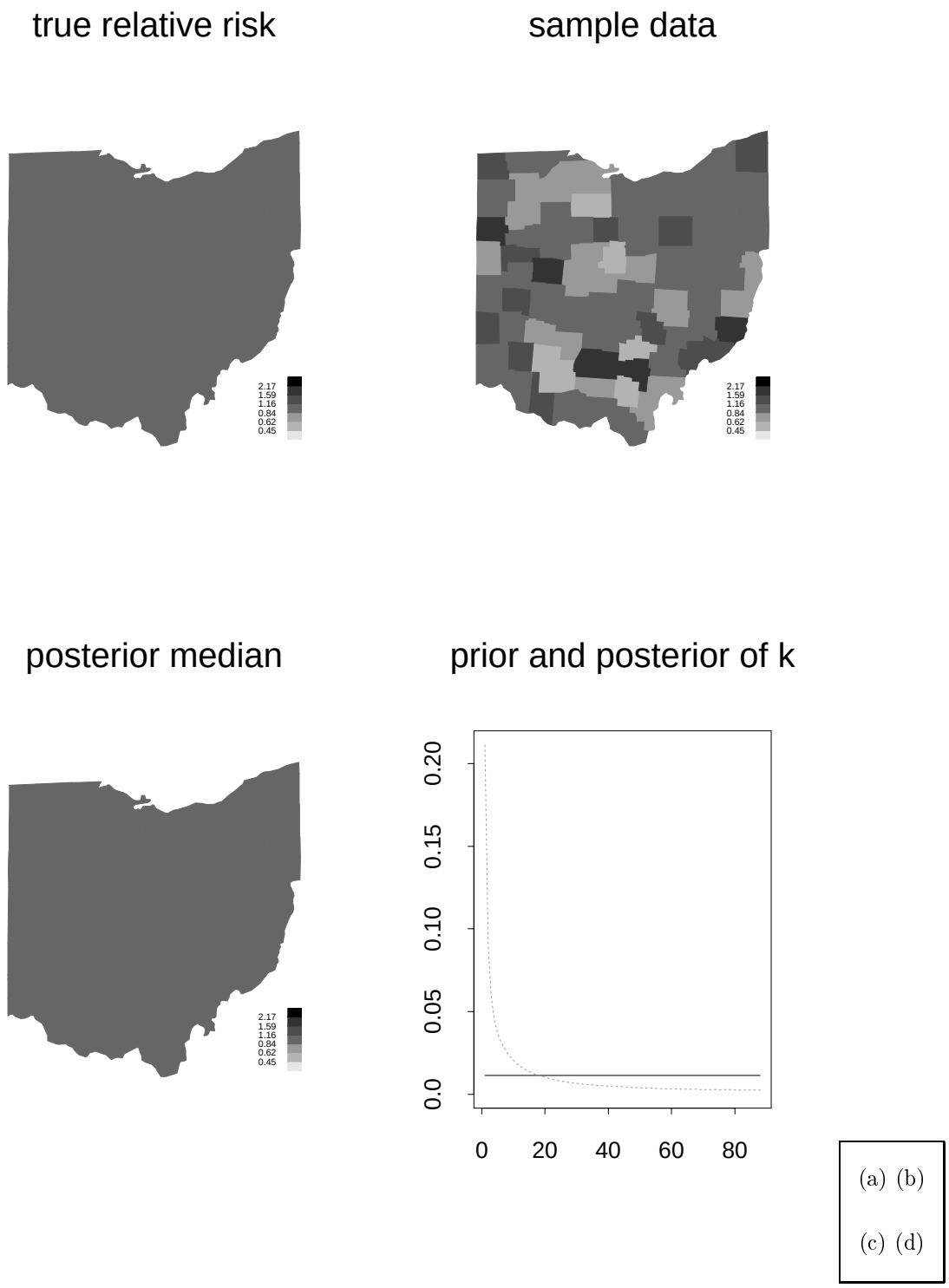
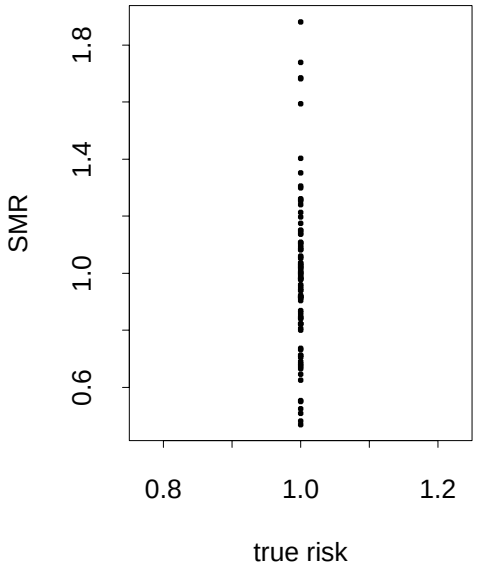
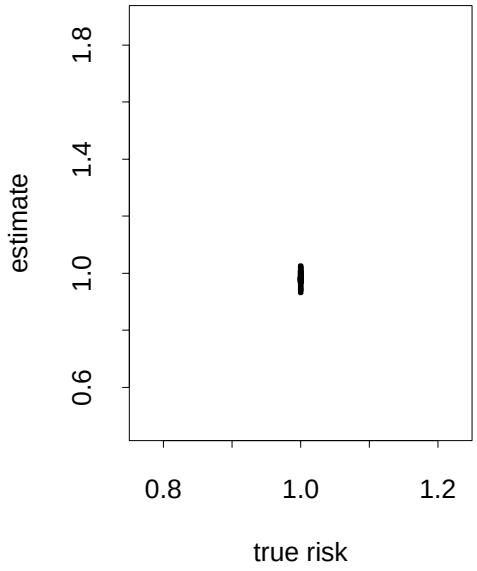
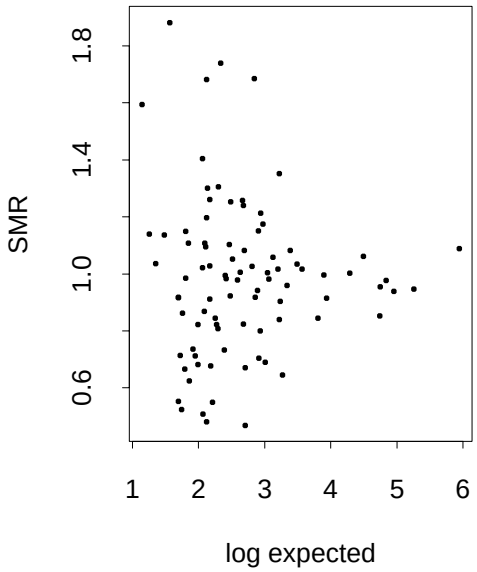
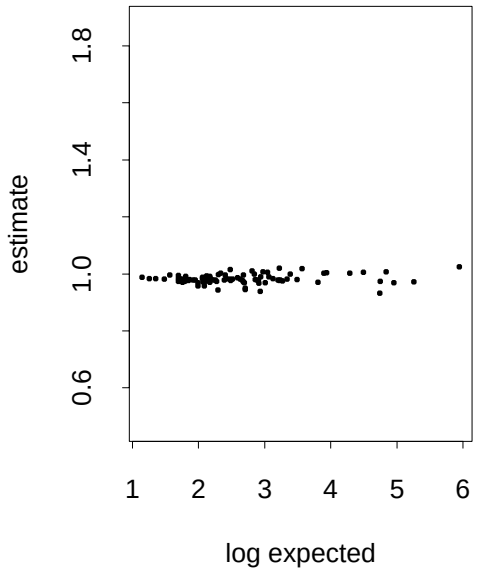


Figure 19: Analysis with (C1) and (S3).


SMR
0.6
1.0
1.4
1.8
0.8
1.0
1.2
true risk


estimate
0.6
1.0
1.4
1.8
0.8
1.0
1.2
true risk


SMR
0.6
1.0
1.4
1.8
1
2
3
4
5
6
log expected


estimate
0.6
1.0
1.4
1.8
1
2
3
4
5
6
log expected

(e) (f)

(g) (h)

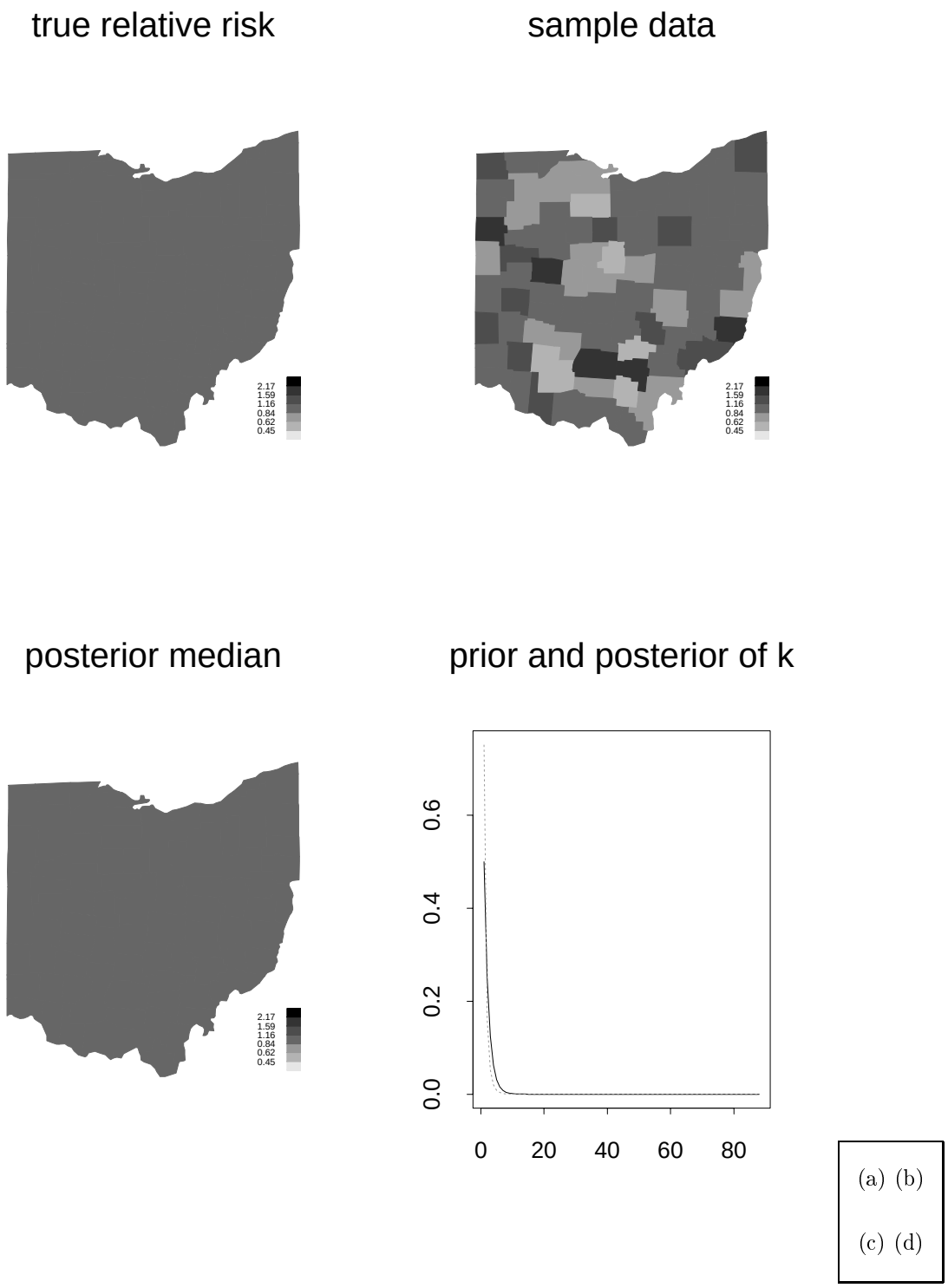
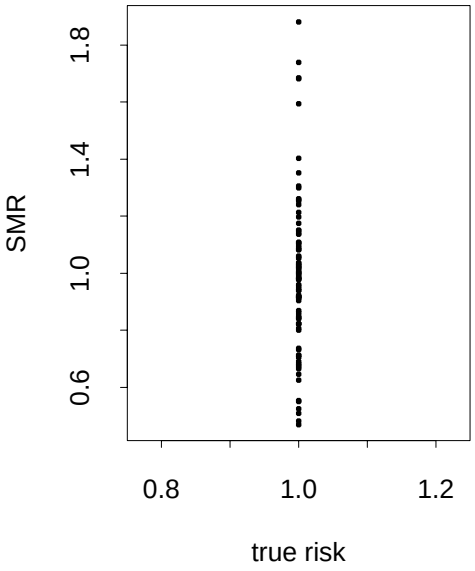
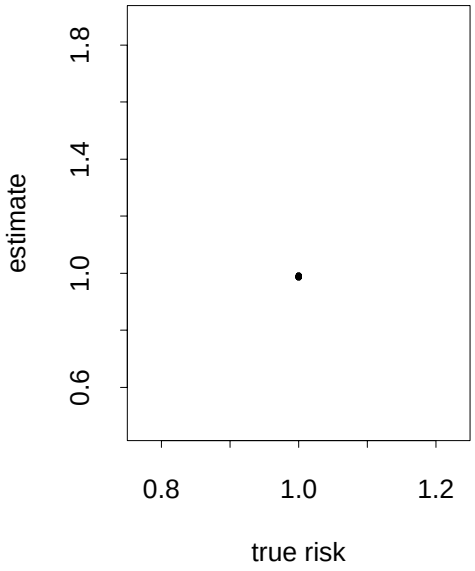
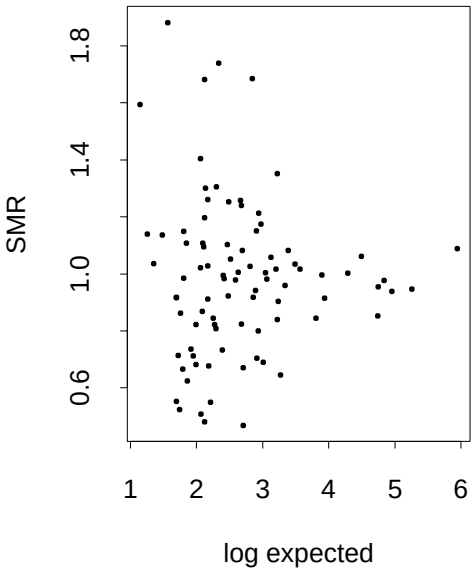
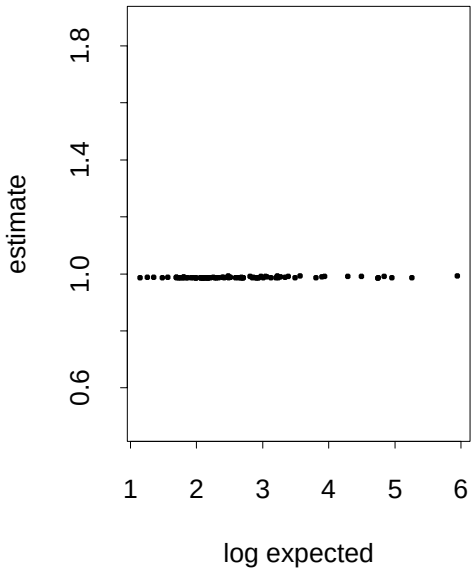


Figure 20: Analysis with (C3) and (S3).


SMR
0.6
1.0
1.4
1.8
0.8
1.0
1.2
true risk


estimate
0.6
1.0
1.4
1.8
0.8
1.0
1.2
true risk


SMR
0.6
1.0
1.4
1.8
1
2
3
4
5
6
log expected


estimate
0.6
1.0
1.4
1.8
1
2
3
4
5
6
log expected

(e) (f)

(g) (h)

Pattern 2

To examine whether the algorithm really can find discontinuities in the spatial pattern or not, we have given low relative risk of 0.5 in most parts of Ohio and a high relative risk of 2.0 in some clusters of different sizes. So there appear sharp edges in the true relative risk pattern. Included here is an analysis of a dataset based on these relative risks using an uniform prior distribution for k and setting the values of the parameters for the gamma distribution to $a = 1$ and $b = 0.01$. In this case other prior distributions showed nearly the same results. Neither a geometric prior for k, nor different values for a and b had any influence on the posterior distribution for k or on the estimated relative risks.

This time the reconstruction is also very good. The posterior distribution for k has a very small variance, which results in quite small acceptance rates for the birth and death move.

Figure 22 and Figure 23 show the estimates using the two different geometric distributions.

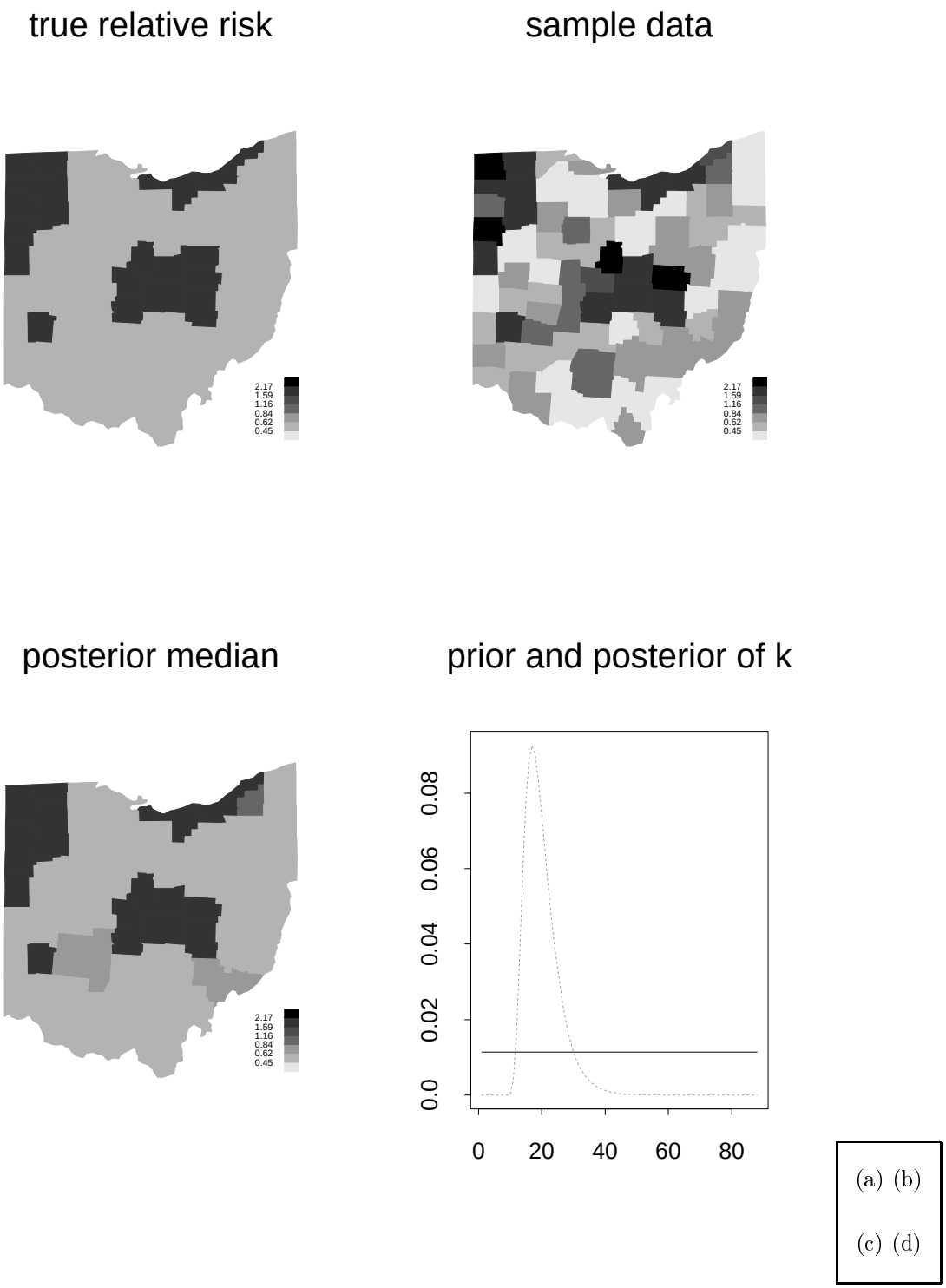
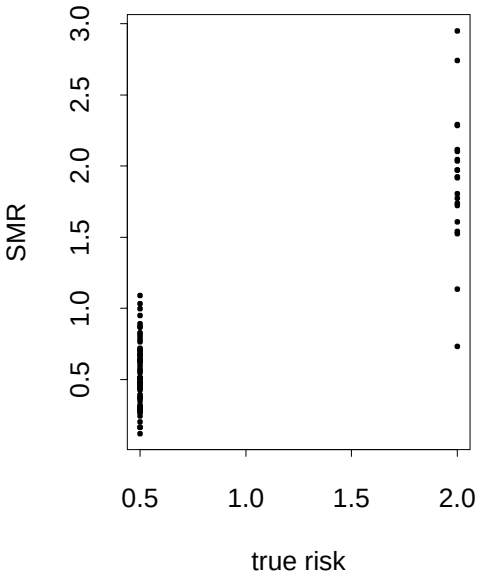
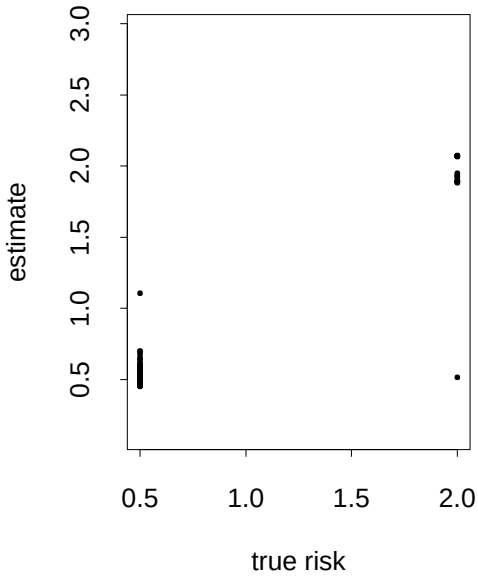
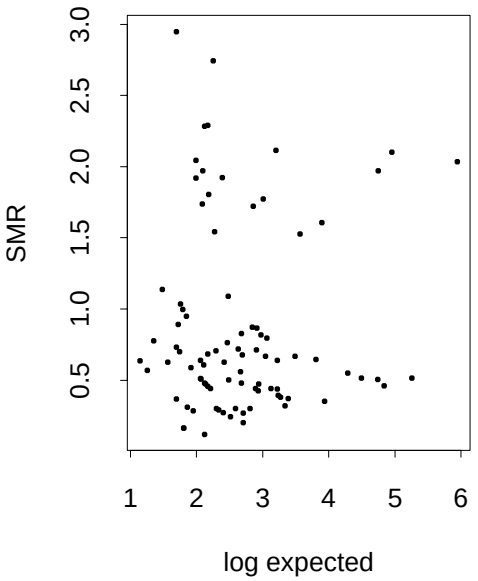
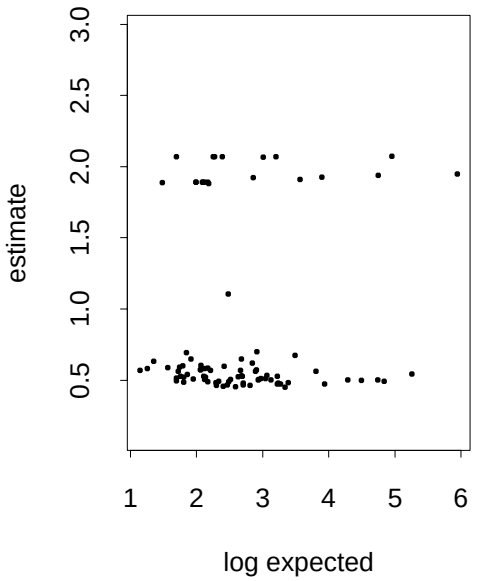


Figure 21: Analysis with (C1) and (S2).


SMR
0.5
1.0
1.5
2.0
2.5
3.0
0.5
1.0
1.5
2.0
true risk

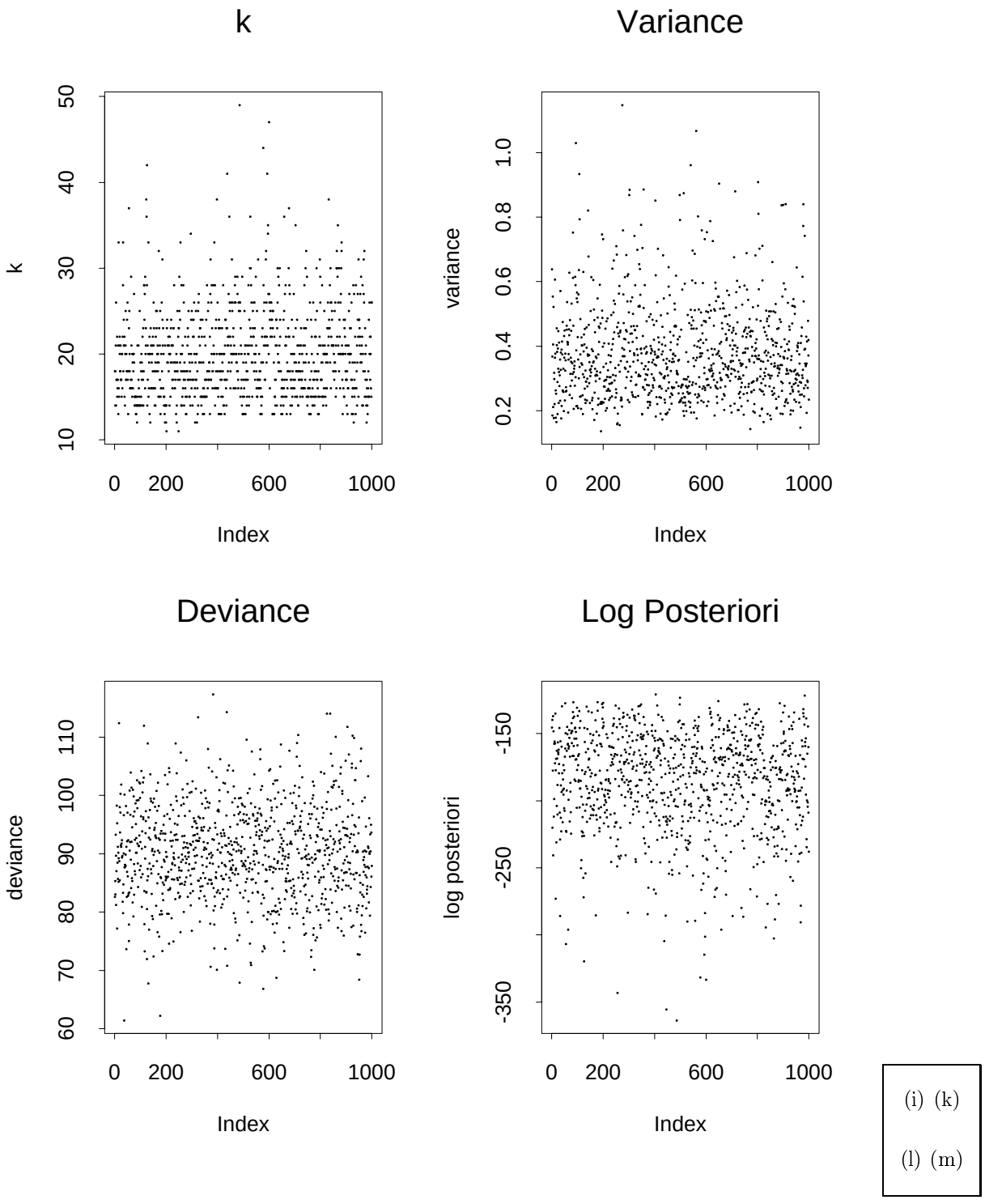

estimate
0.5
1.0
1.5
2.0
2.5
3.0
0.5
1.0
1.5
2.0
true risk

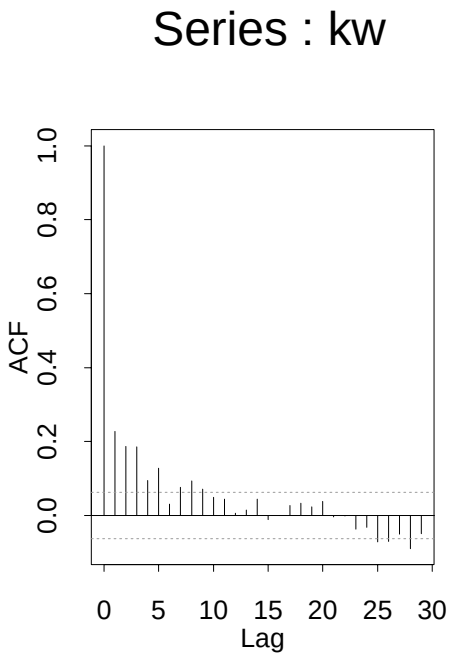

SMR
0.5
1.0
1.5
2.0
2.5
3.0
1
2
3
4
5
6
log expected

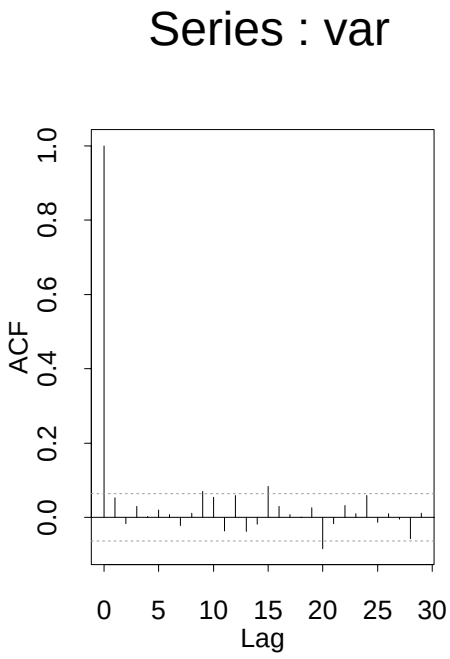

estimate
0.5
1.0
1.5
2.0
2.5
3.0
1
2
3
4
5
6
log expected

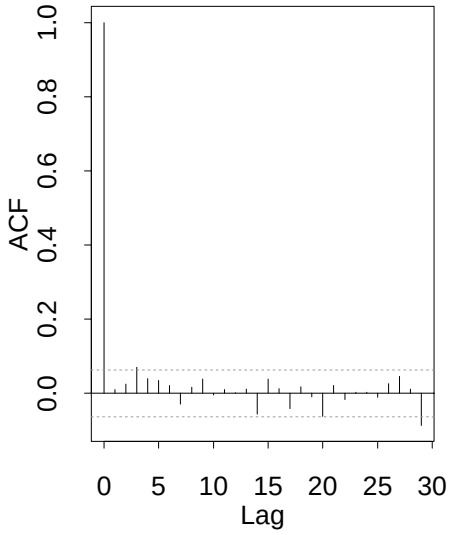
(e) (f)

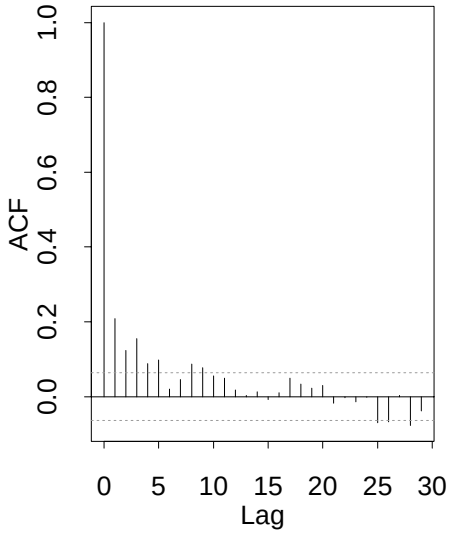
(g) (h)


k
k
10
20
30
40
50
0
200
600
1000
Index
Variance
variance
0.2
0.4
0.6
0.8
1.0
0
200
600
1000
Index
Deviance
deviance
60
70
80
90
100
110
0
200
600
1000
Index
Log Posteriori
log posteriori
-350
-250
-150
0
200
600
1000
Index
(i) (k)
(l) (m)


Series : kw
ACF
0.0
0.2
0.4
0.6
0.8
1.0
0
5
10
15
20
25
30
Lag


Series : var
ACF
0.0
0.2
0.4
0.6
0.8
1.0
0
5
10
15
20
25
30
Lag


Series : dev
ACF
0.0
0.2
0.4
0.6
0.8
1.0
0
5
10
15
20
25
30
Lag


Series : post
ACF
0.0
0.2
0.4
0.6
0.8
1.0
0
5
10
15
20
25
30
Lag

(n) (o)

(p) (q)

variance
0.2
0.4
0.6
0.8
1.0
10
20
30
40
50
k

deviance
60
70
80
90
100
110
10
20
30
40
50
k

log posteriori
-150
-250
-350
10
20
30
40
50
k

Rates
acceptance rate
0.0
0.2
0.4
0.6
0.8
1.0
1
2
3
4
5
6
move

(r) (s)

(t) (u)

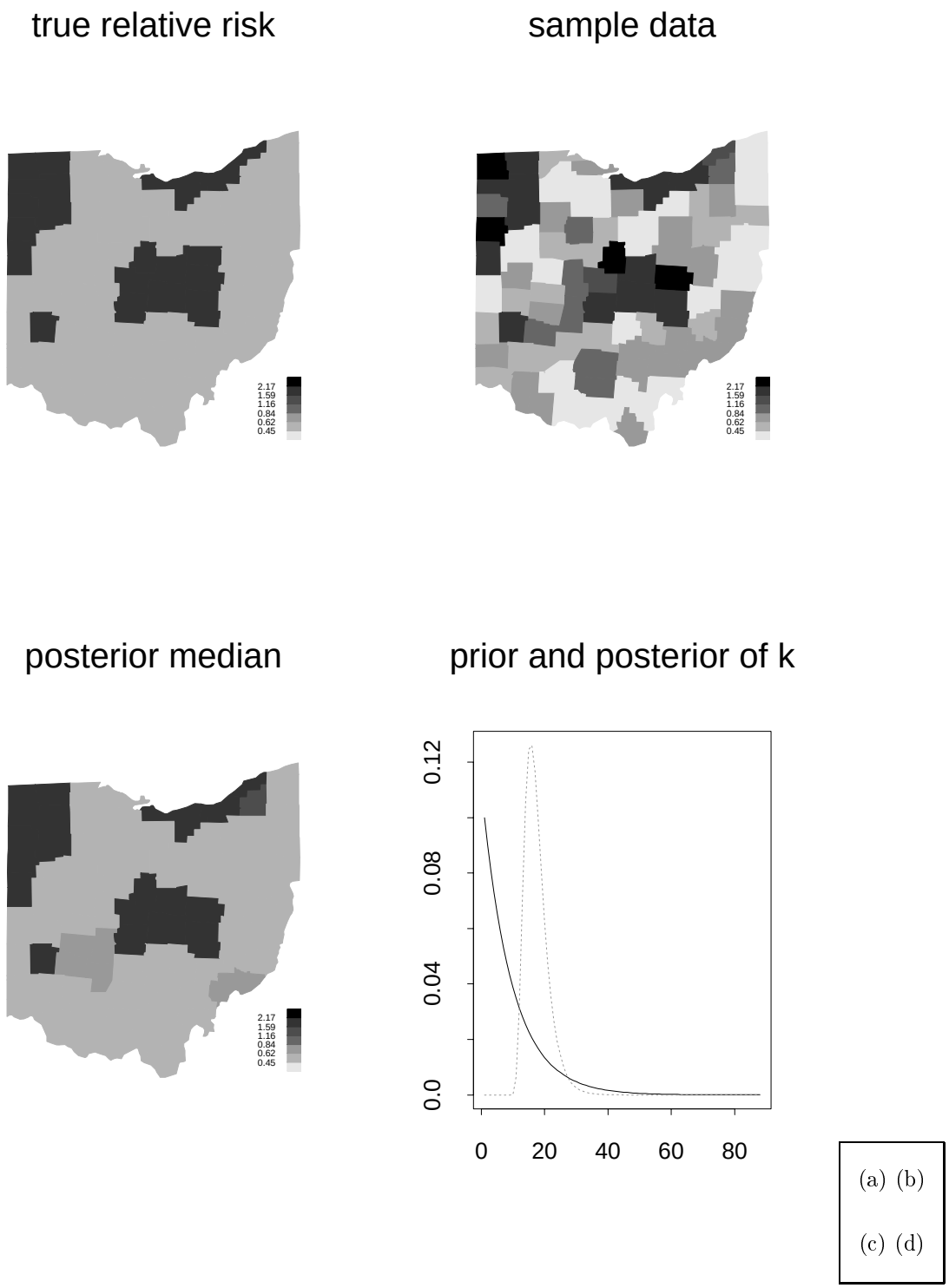
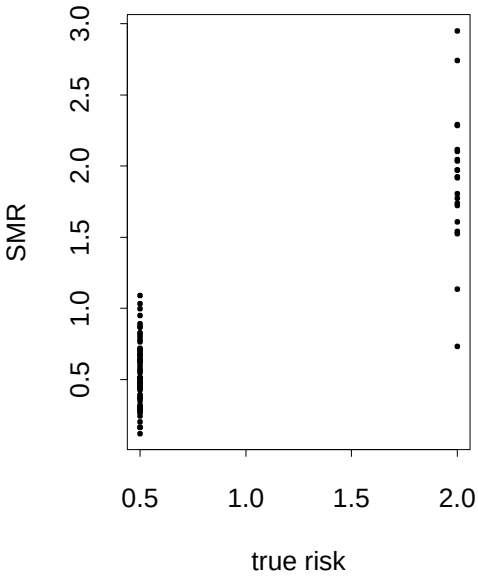
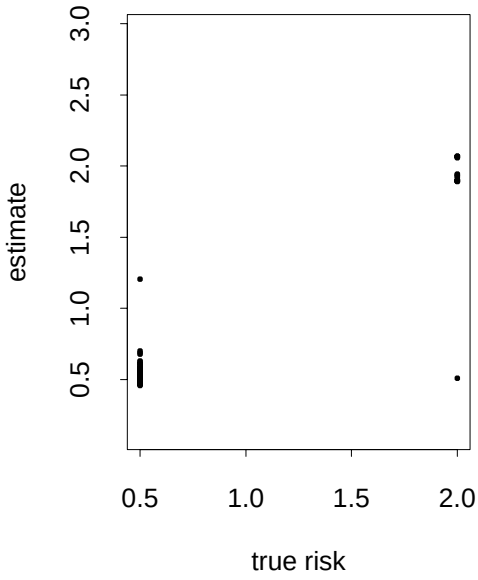
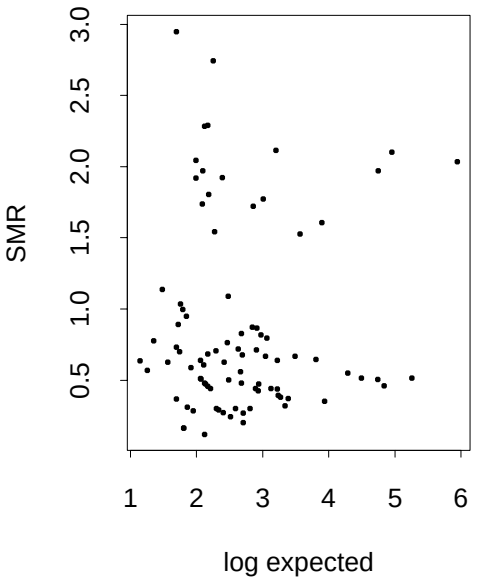
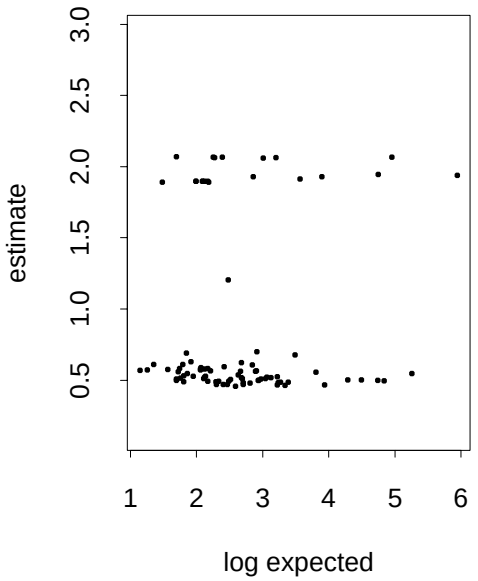


Figure 22: Analysis with (C2) and (S2).


SMR
0.5
1.0
1.5
2.0
2.5
3.0
0.5
1.0
1.5
2.0
true risk


estimate
0.5
1.0
1.5
2.0
2.5
3.0
0.5
1.0
1.5
2.0
true risk


SMR
0.5
1.0
1.5
2.0
2.5
3.0
1
2
3
4
5
6
log expected


estimate
0.5
1.0
1.5
2.0
2.5
3.0
1
2
3
4
5
6
log expected

(e) (f)

(g) (h)

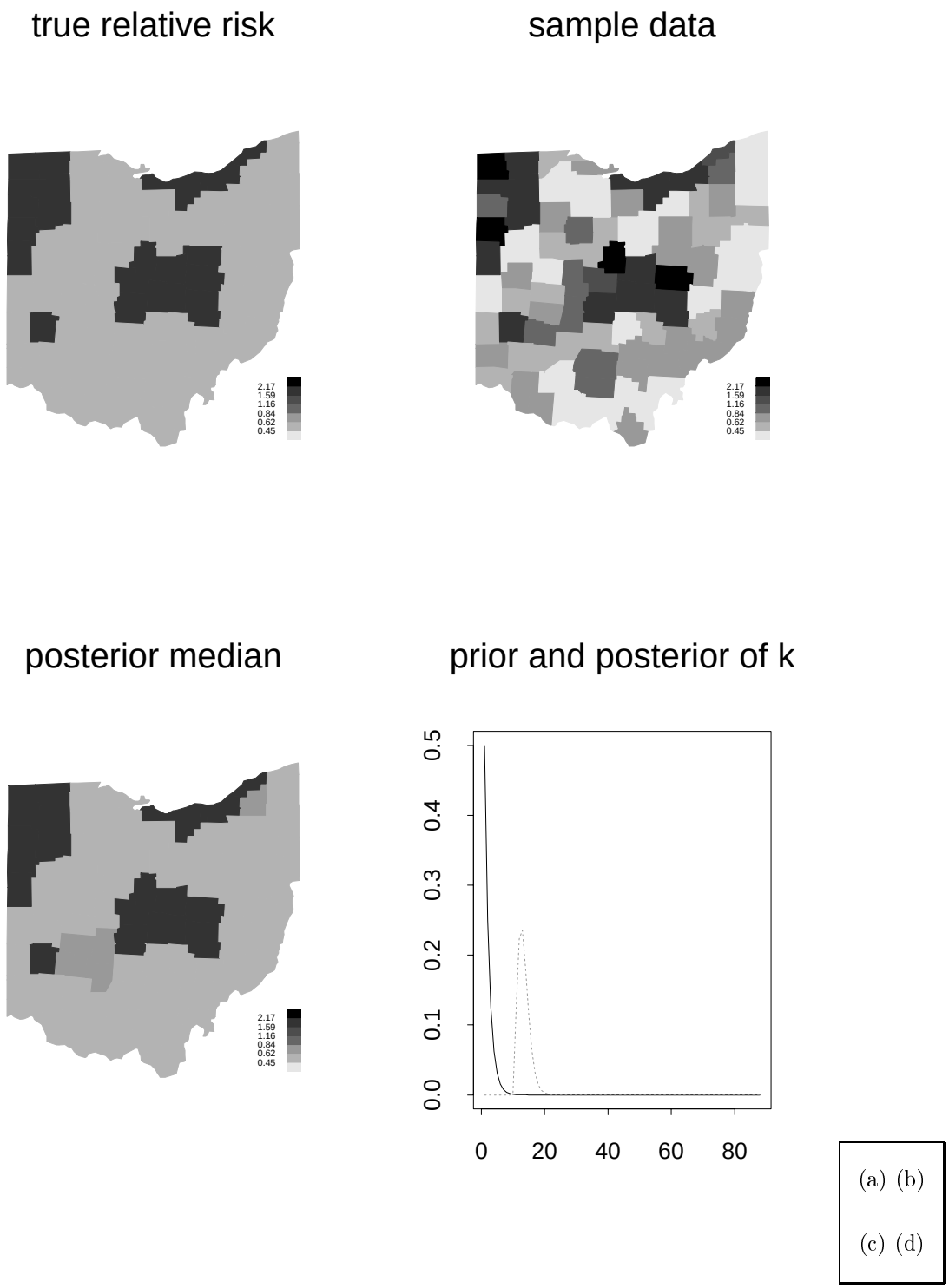
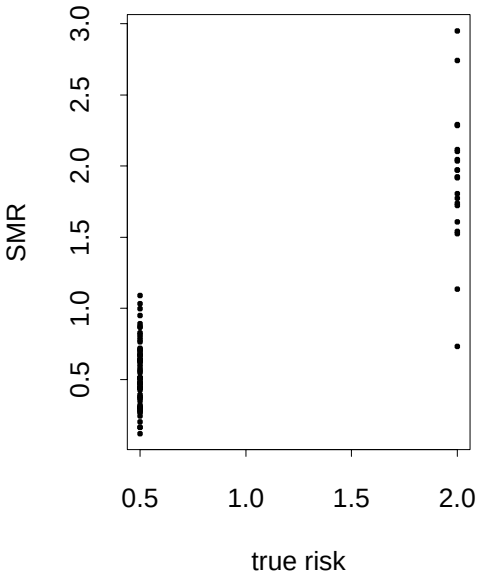
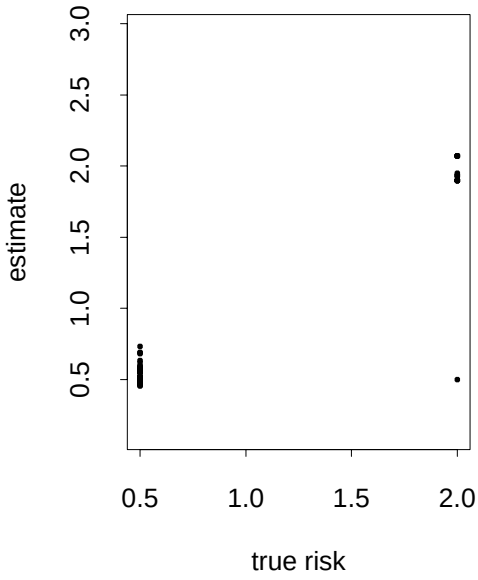
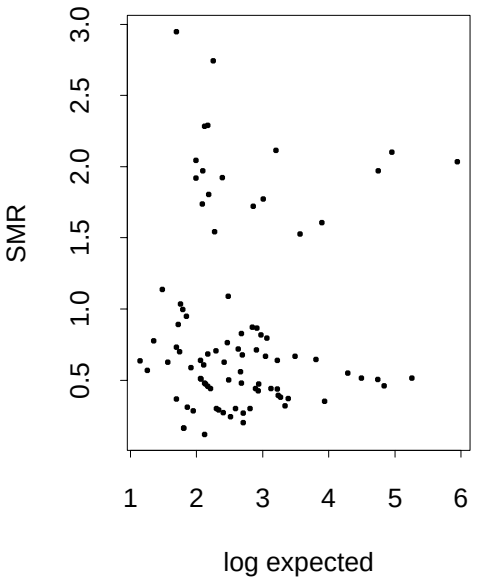
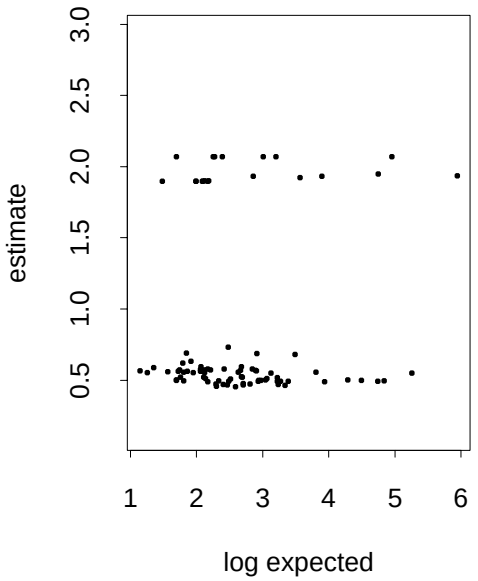


Figure 23: Analysis with (C3) and (S2).


SMR
0.5
1.0
1.5
2.0
2.5
3.0
0.5
1.0
1.5
2.0
true risk


estimate
0.5
1.0
1.5
2.0
2.5
3.0
0.5
1.0
1.5
2.0
true risk


SMR
0.5
1.0
1.5
2.0
2.5
3.0
1
2
3
4
5
6
log expected


estimate
0.5
1.0
1.5
2.0
2.5
3.0
1
2
3
4
5
6
log expected

(e) (f)

(g) (h)

Pattern 3

Here we defined a true relative risk pattern showing slowly rising values from the west to the east. Starting with 0.6 in the west and rising to 0.7, 0.8, 0.9, 1.0, 1.2, 1.4, 1.6 to a maximum of 2.0 in the east, so we have 9 different risk levels over the whole state. This time we used a geometric prior distribution with parameter $c = 0.1$ for k. But also in this case the other prior distributions led to very similar results for the posterior distribution for k as well as for the estimated relative risks. As a prior for σ^2 we applied a gamma distribution with parameters $a = 5$ and $b = 0.125$.

Since the changes between the areas of different risk levels are quite small, these areas are often connected in one cluster, So the values of the posterior distribution for k are also small. Yet the rising risk level from the West to the East is still clearly recognizable in Figure 24 in the estimated risk pattern.

Again the results using an uniform prior given in Figure 25 are nearly the same. The strong geometric prior with parameter $c = 0.5$ in Figure 26 leads to slightly lower values for k, yet the estimates of the relative risks are not differing to much form the other results.

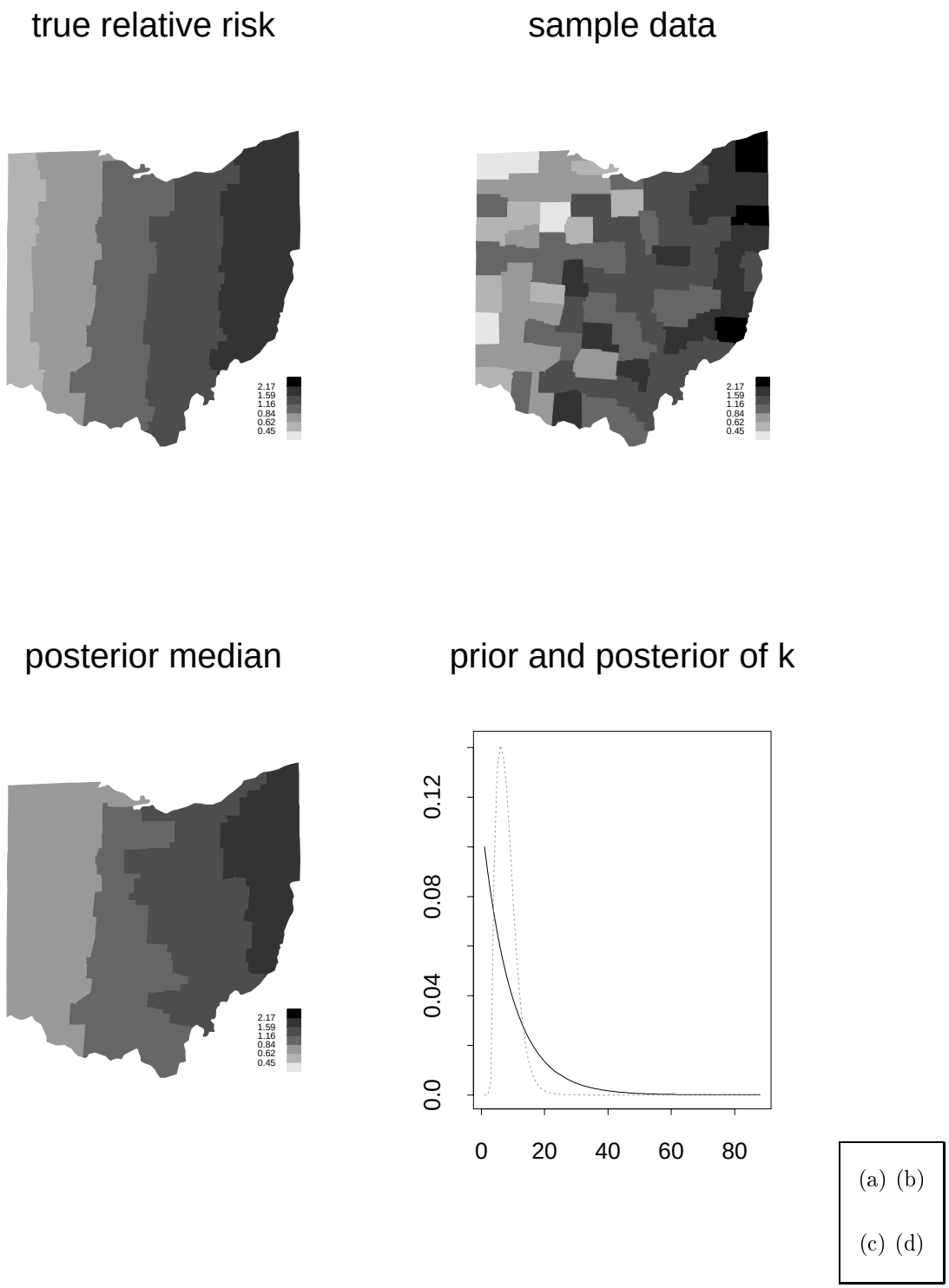
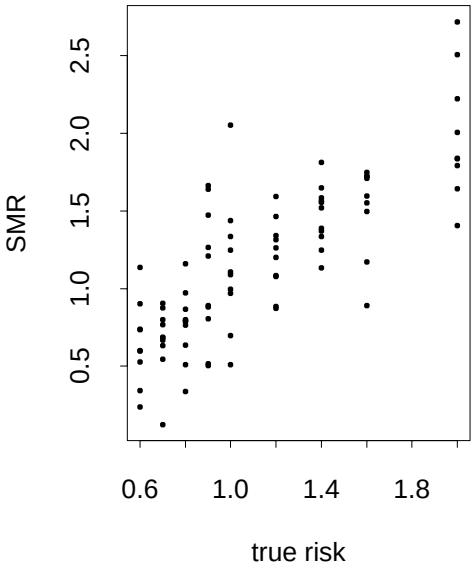
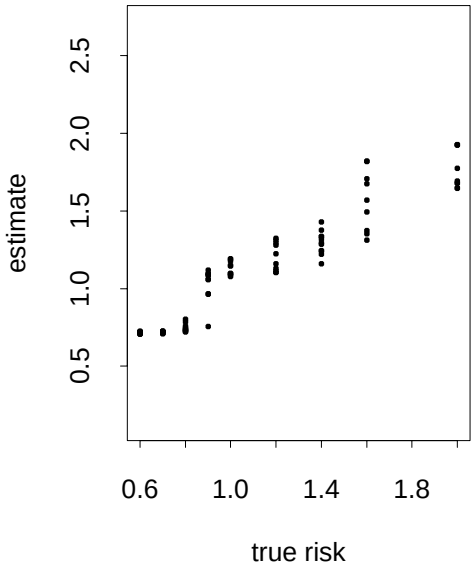
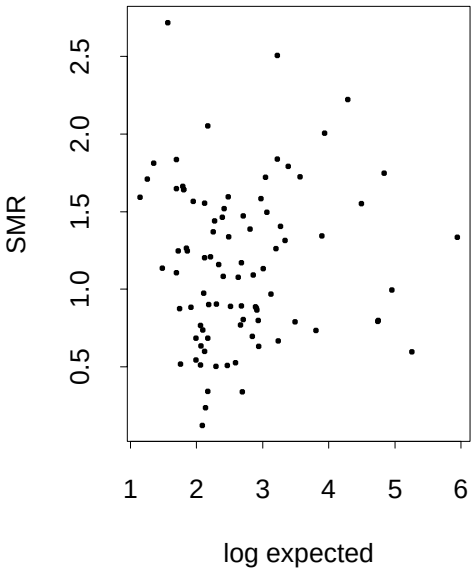
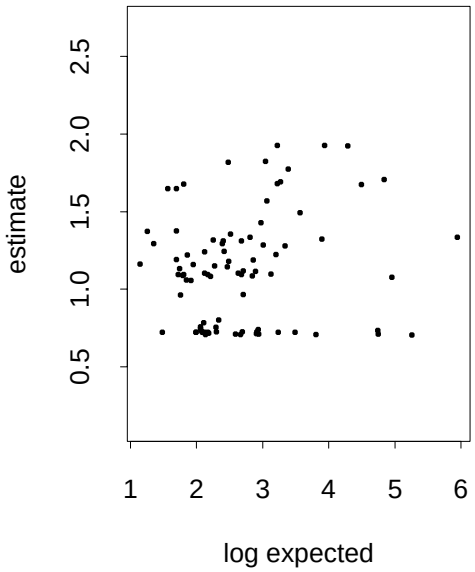


Figure 24: Analysis with (C2) and (S3).


SMR
0.5
1.0
1.5
2.0
2.5
0.6
1.0
1.4
1.8
true risk

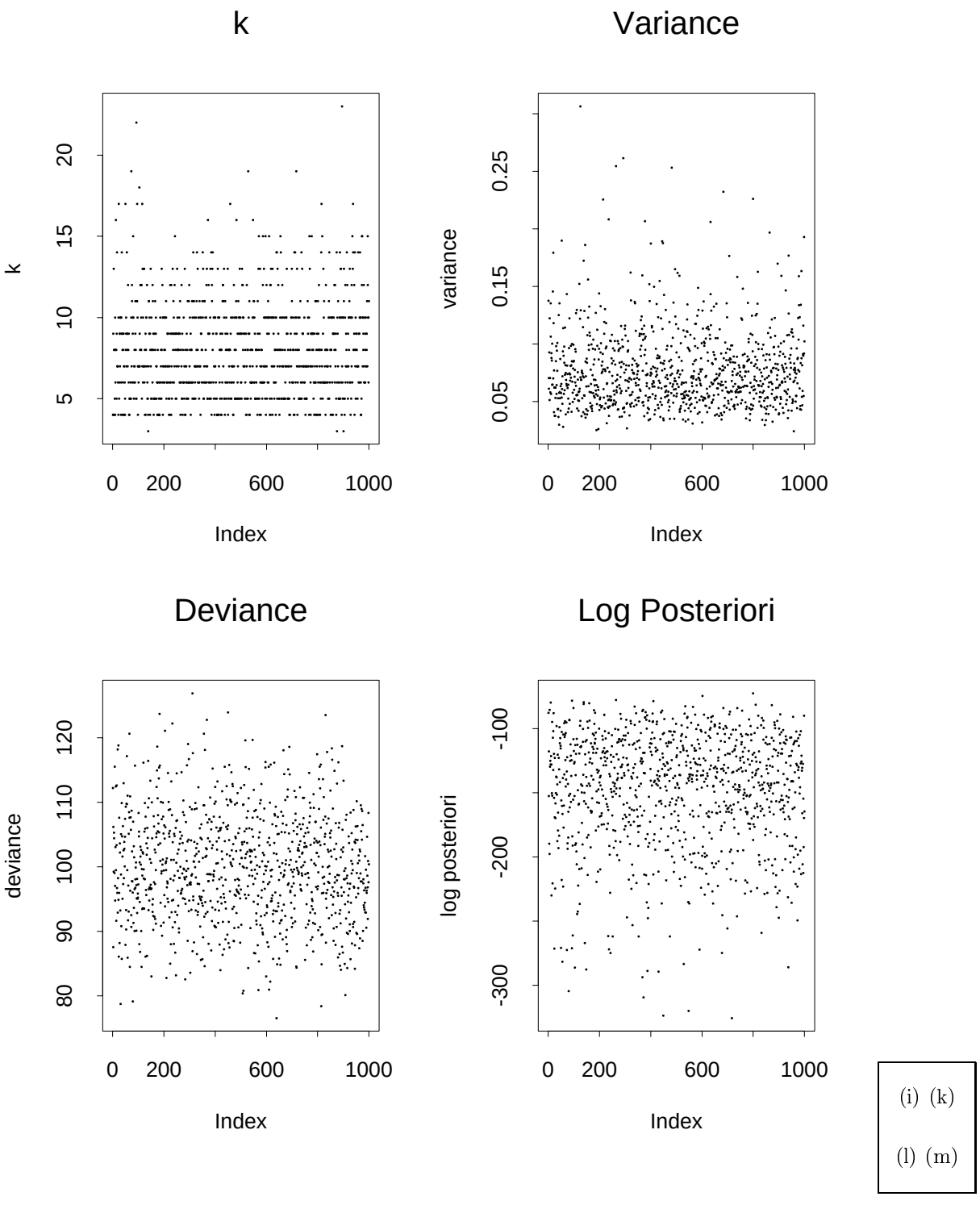

estimate
0.5
1.0
1.5
2.0
2.5
0.6
1.0
1.4
1.8
true risk

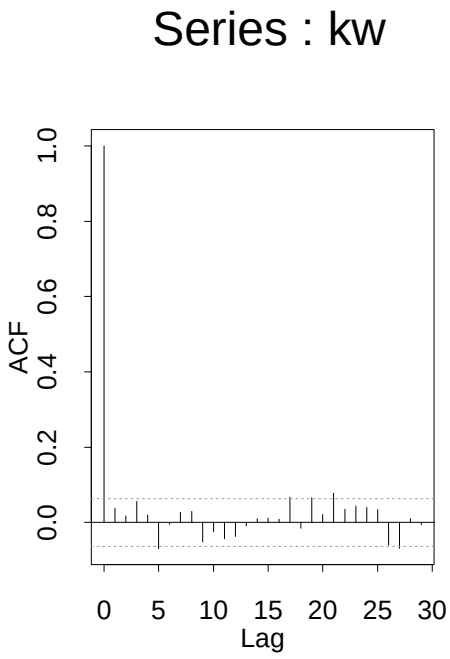

SMR
0.5
1.0
1.5
2.0
2.5
1
2
3
4
5
6
log expected

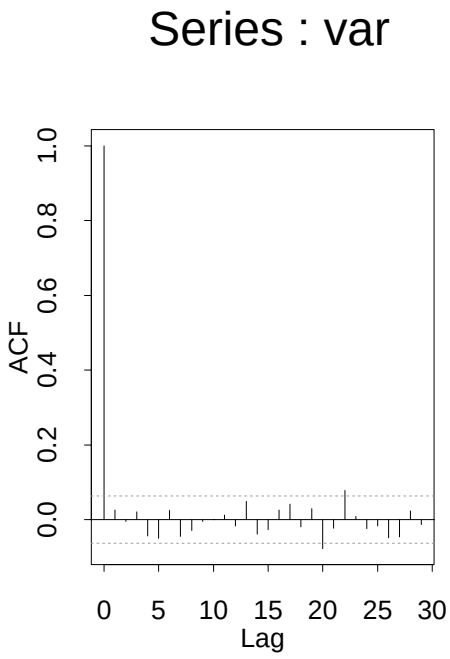

estimate
0.5
1.0
1.5
2.0
2.5
1
2
3
4
5
6
log expected

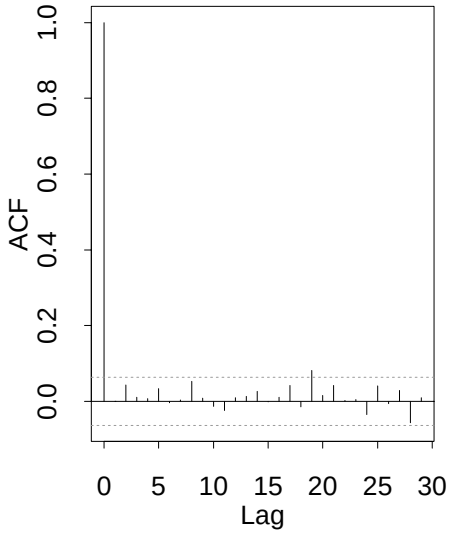
(e) (f)

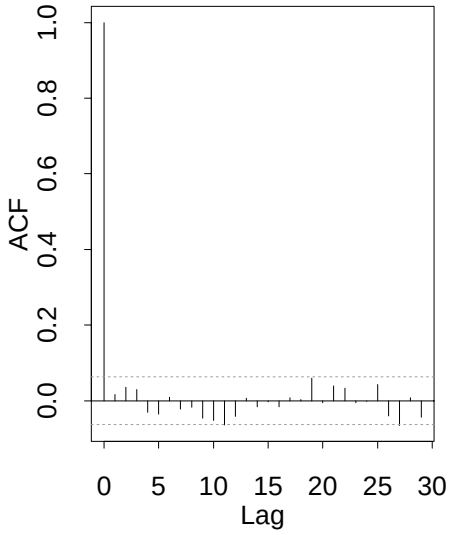
(g) (h)


k
Variance
k
variance
Index
Index
Deviance
Log Posteriori
deviance
log posteriori
Index
Index
(i) (k)
(l) (m)


Series : kw
ACF
0.0
0.2
0.4
0.6
0.8
1.0
0
5
10
15
20
25
30
Lag


Series : var
ACF
0.0
0.2
0.4
0.6
0.8
1.0
0
5
10
15
20
25
30
Lag


Series : dev
ACF
0.0
0.2
0.4
0.6
0.8
1.0
0
5
10
15
20
25
30
Lag


Series : post
ACF
0.0
0.2
0.4
0.6
0.8
1.0
0
5
10
15
20
25
30
Lag

(n) (o)

(p) (q)

variance
0.05
0.15
0.25
5
10
15
20
k

deviance
80
90
100
110
120
5
10
15
20
k

log posteriori
-300
-200
-100
5
10
15
20
k

Rates
acceptance rate
0.0
0.2
0.4
0.6
0.8
1.0
1
2
3
4
5
6
move

(r) (s)

(t) (u)

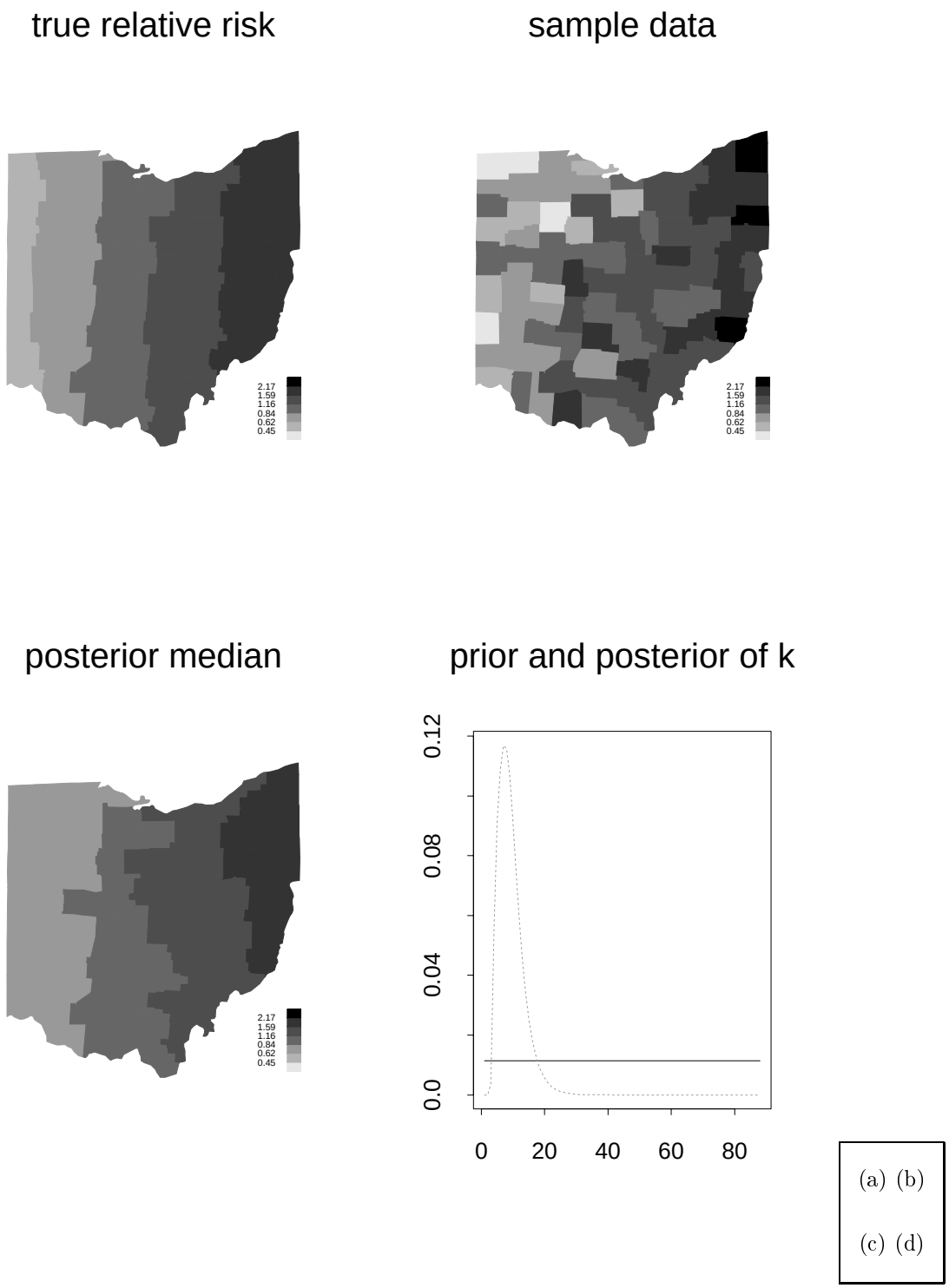
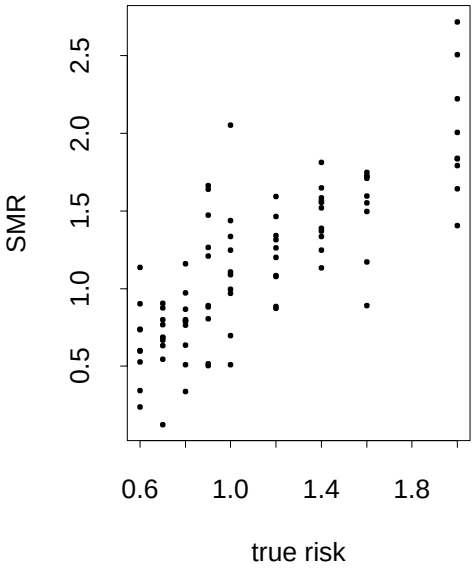
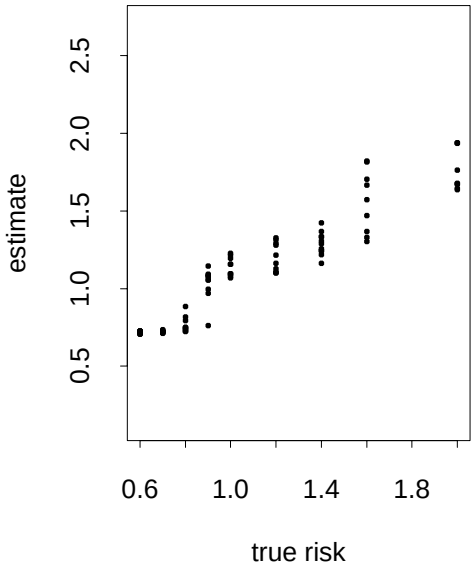
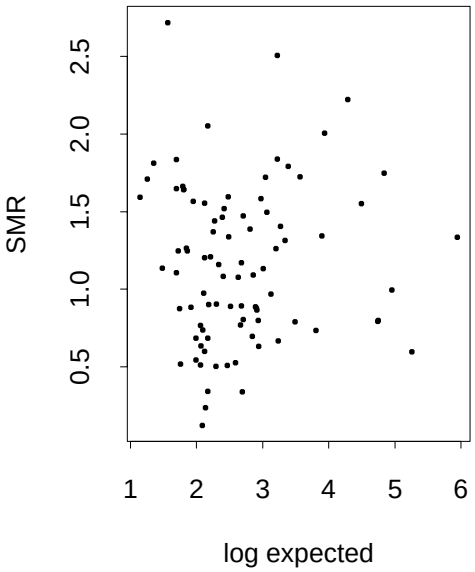
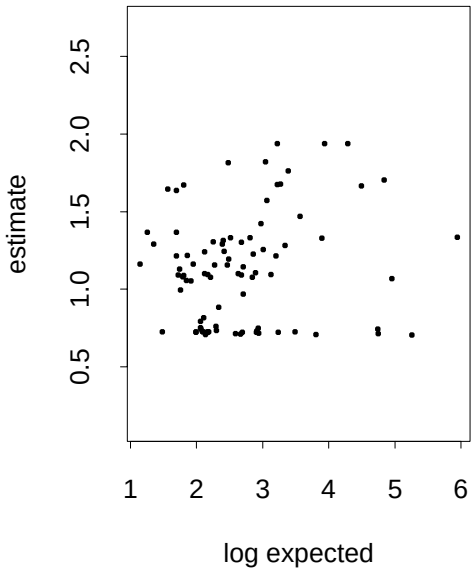


Figure 25: Analysis with (C1) and (S3).


SMR
0.5
1.0
1.5
2.0
2.5
0.6
1.0
1.4
1.8
true risk


estimate
0.5
1.0
1.5
2.0
2.5
0.6
1.0
1.4
1.8
true risk


SMR
0.5
1.0
1.5
2.0
2.5
1
2
3
4
5
6
log expected


estimate
0.5
1.0
1.5
2.0
2.5
1
2
3
4
5
6
log expected

(e) (f)

(g) (h)

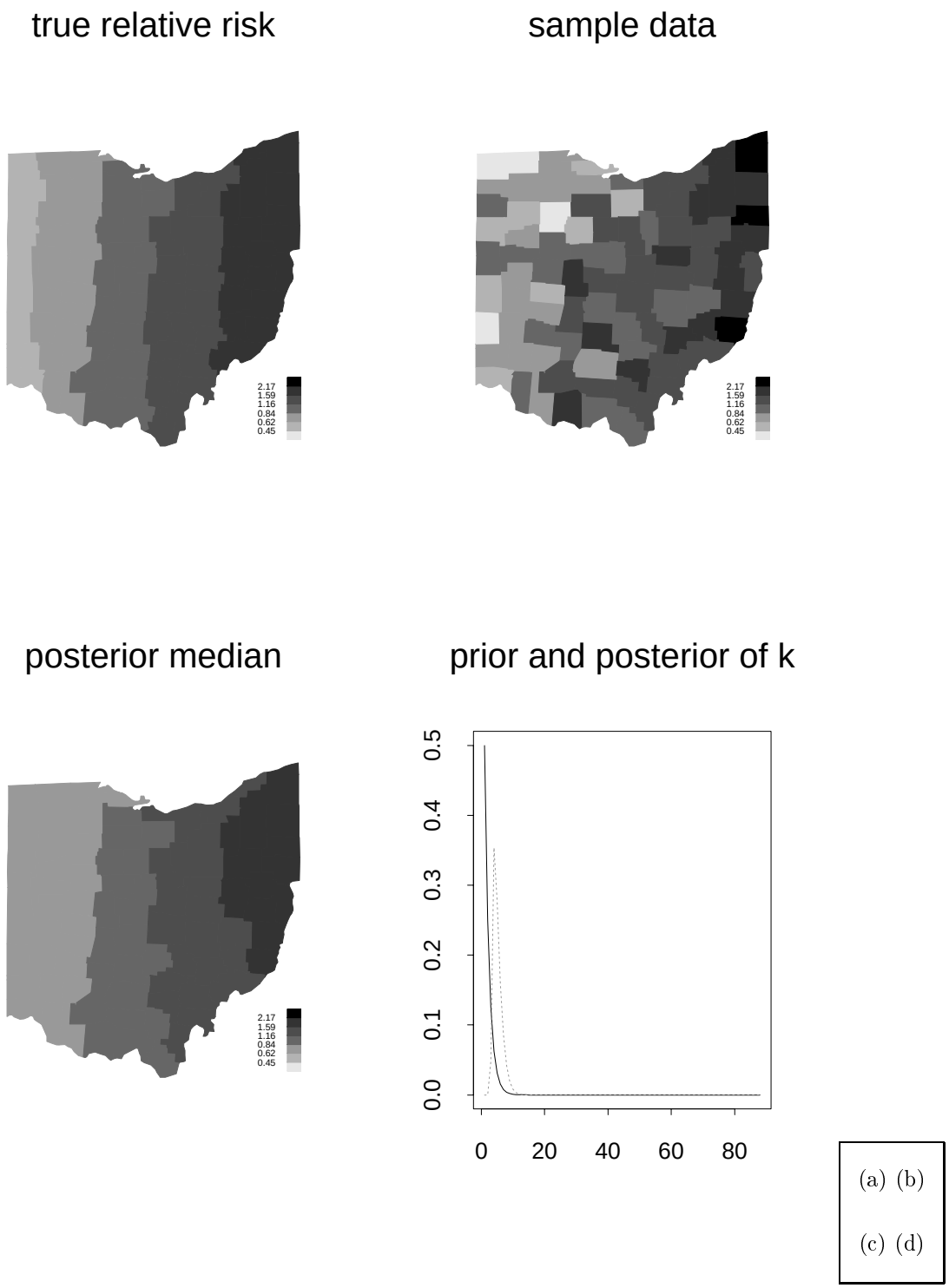
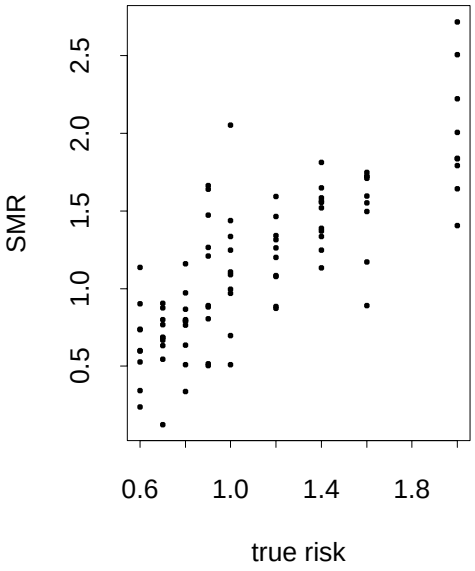
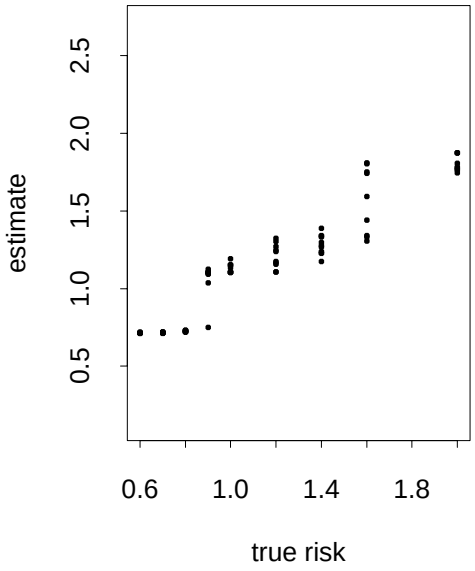
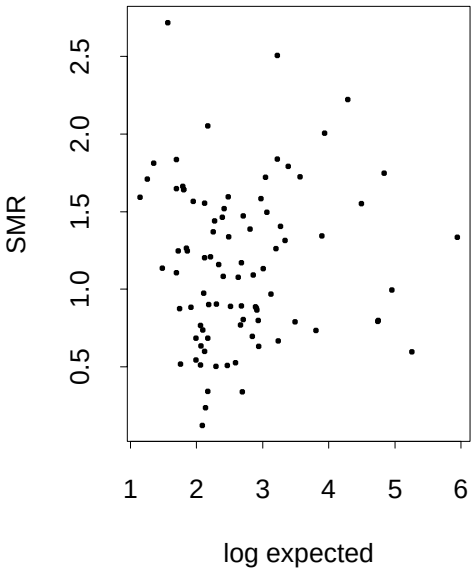
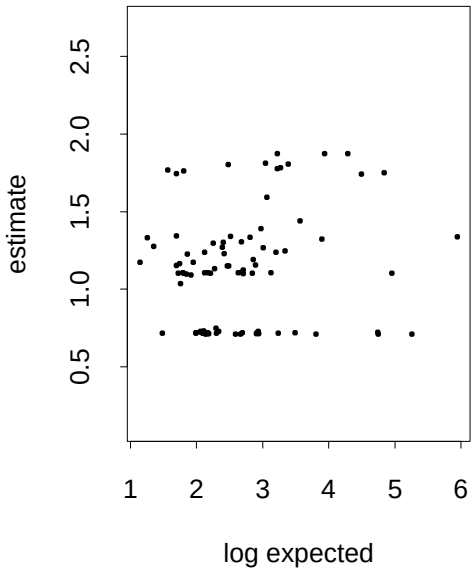


Figure 26: Analysis with (C3) and (S3).


SMR
0.5
1.0
1.5
2.0
2.5
0.6
1.0
1.4
1.8
true risk


estimate
0.5
1.0
1.5
2.0
2.5
0.6
1.0
1.4
1.8
true risk


SMR
0.5
1.0
1.5
2.0
2.5
1
2
3
4
5
6
log expected


estimate
0.5
1.0
1.5
2.0
2.5
1
2
3
4
5
6
log expected

(e) (f)

(g) (h)

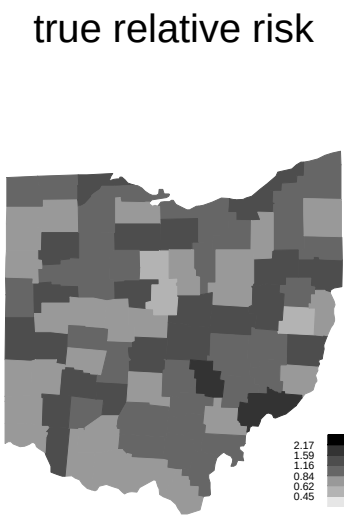
Pattern 4

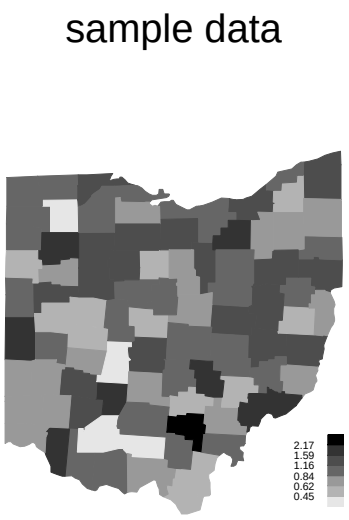
Finally, we want to investigate whether the algorithm is able to detect the nonexistence of a spatial pattern. So in this case we assume total heterogeneity of the relative risks. Therefore, the log relative risks have been independently generated from a normal distribution with mean zero and standard deviation 0.25. Of course the algorithm will smooth the risks and therefore one can not expect to restore the original pattern completely.

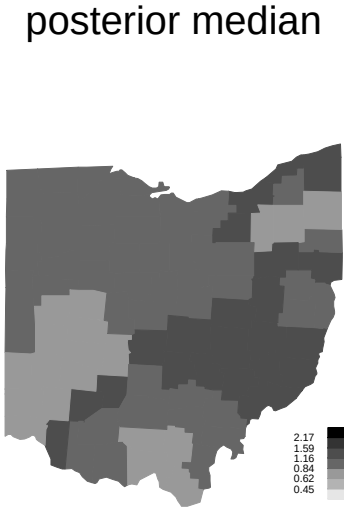
Here the output of the algorithm really depends on the choice of the prior distribution for k. And so the mode of the posterior distribution for k is getting lower when using a stronger prior, decreasing from around 80 with an uniform prior to around 45 with a geometric prior with parameter $c = 0.1$ to under 20 using a geometric distribution with parameter $c = 0.5$. And with decreasing k, the estimated relative risks are getting smoother, of course. But since this strong influence of the prior distribution only appears if there is no spatial pattern recognizable, the conclusion seems appropriate that if the choice of the prior distribution has only minimal effect on the estimates there really seems to be a spatial structure, which is reconstructed quite closely by the algorithm.

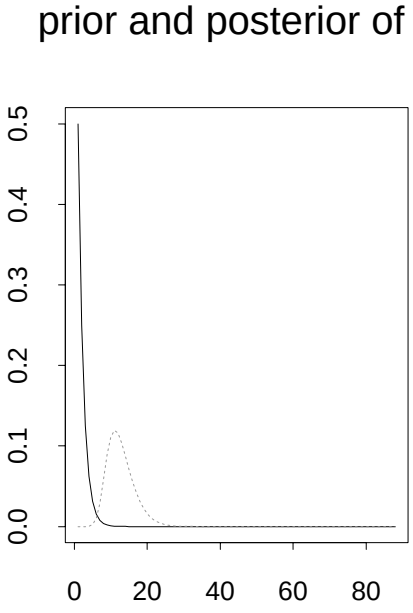
Given in Figure 27 is the analysis with a geometric prior with parameter $c = 0.5$, a very strong prior for k. The parameters for the gamma distribution were set to $a = 0.25$ and $b = 0.00025$. So these are the most extreme results we reached for this pattern.

The influence of the choice of the prior distribution becomes obvious from Figure 28 and Figure 29, showing the results from the uniform and the geometric prior with parameter $c = 0.1$, respectively. The patterns of the estimates are noisier in both cases.





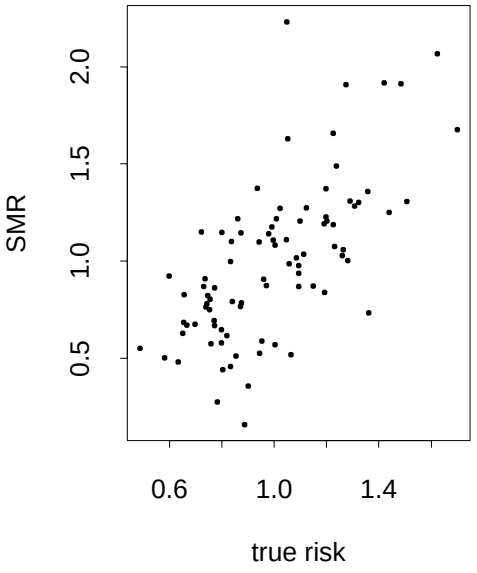


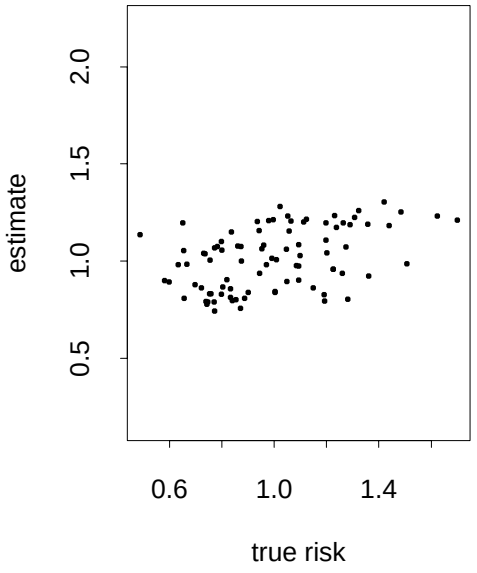


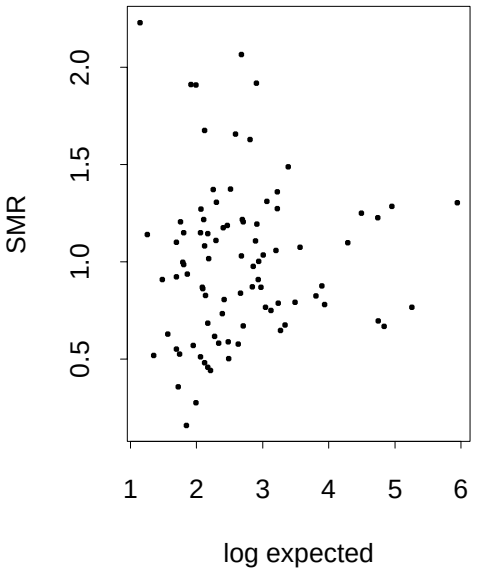
(a) (b)

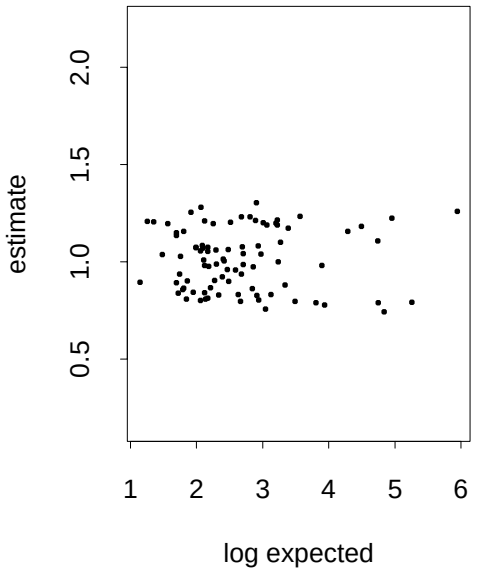
(c) (d)

Figure 27: Analysis with (C3) and (S1).


SMR
0.5
1.0
1.5
2.0
0.6
1.0
1.4
true risk

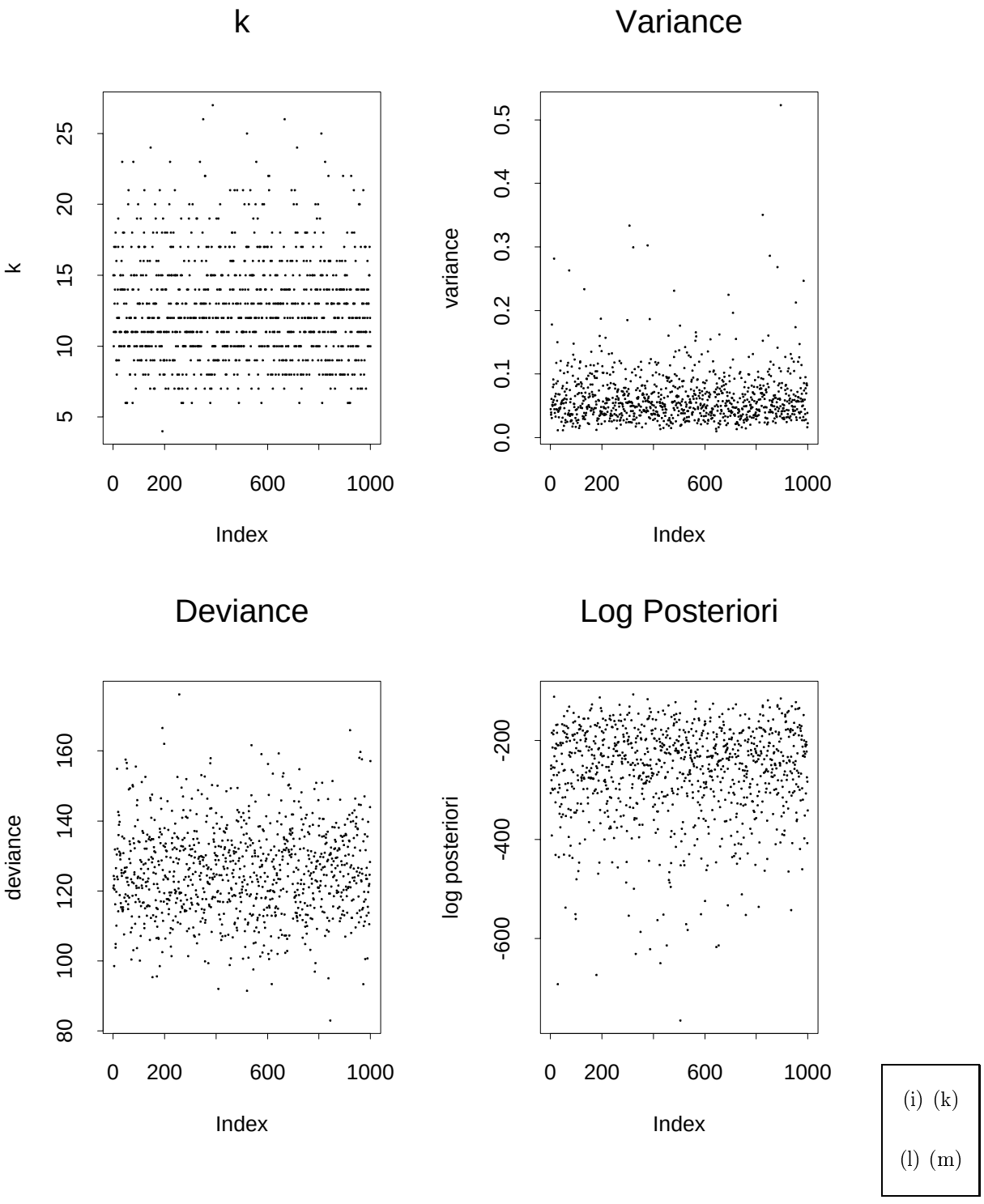

estimate
0.5
1.0
1.5
2.0
0.6
1.0
1.4
true risk

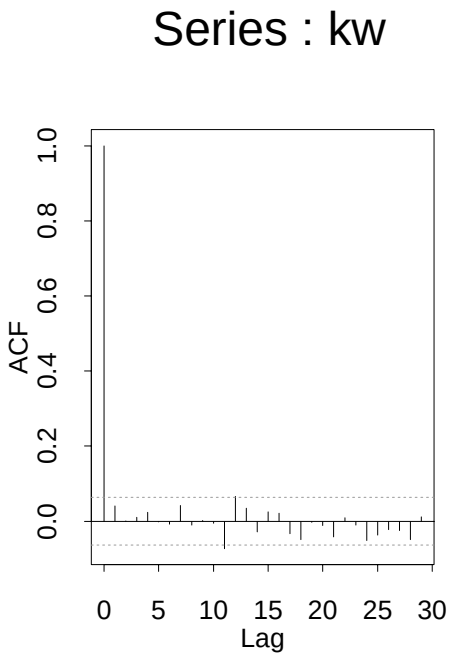

SMR
0.5
1.0
1.5
2.0
1
2
3
4
5
6
log expected

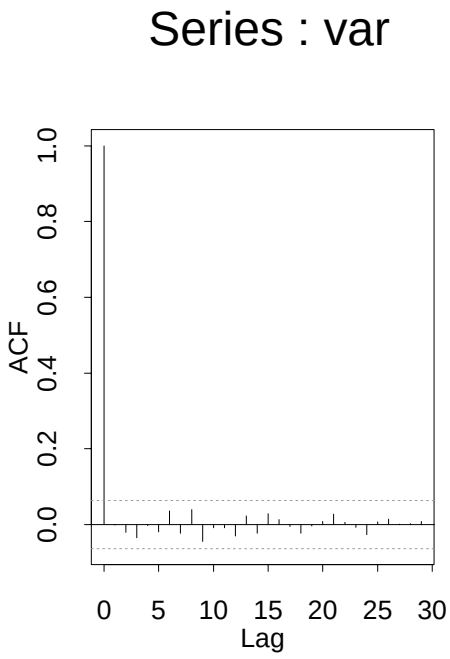

estimate
0.5
1.0
1.5
2.0
1
2
3
4
5
6
log expected

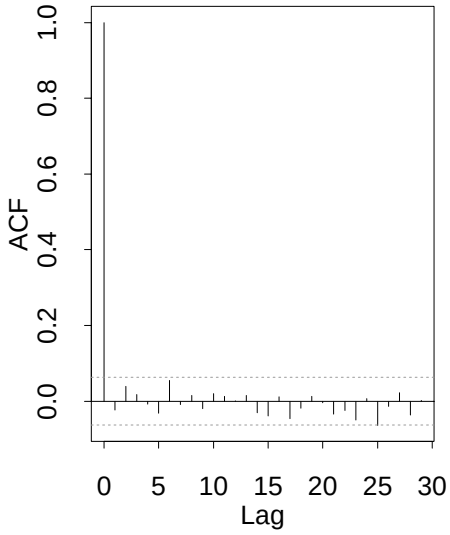
(e) (f)

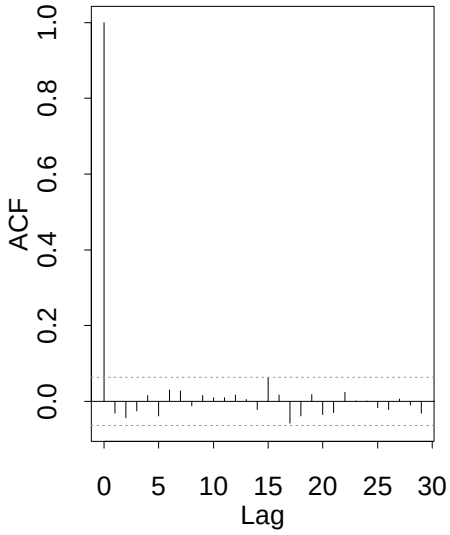
(g) (h)


k
k
5
10
15
20
25
0
200
600
1000
Index
Variance
variance
0.0
0.1
0.2
0.3
0.4
0.5
0
200
600
1000
Index
Deviance
deviance
80
100
120
140
160
0
200
600
1000
Index
Log Posteriori
log posteriori
-600
-400
-200
0
200
600
1000
Index
(i) (k)
(l) (m)


Series : kw
ACF
0.0
0.2
0.4
0.6
0.8
1.0
0
5
10
15
20
25
30
Lag


Series : var
ACF
0.0
0.2
0.4
0.6
0.8
1.0
0
5
10
15
20
25
30
Lag


Series : dev
ACF
0.0
0.2
0.4
0.6
0.8
1.0
0
5
10
15
20
25
30
Lag


Series : post
ACF
0.0
0.2
0.4
0.6
0.8
1.0
0
5
10
15
20
25
30
Lag

(n) (o)

(p) (q)

variance
0.0
0.1
0.2
0.3
0.4
0.5
5
10
15
20
25
k

deviance
80
100
120
140
160
5
10
15
20
25
k

Rates

log posteriori
-600
-400
-200
5
10
15
20
25
k

acceptance rate
0.0
0.2
0.4
0.6
0.8
1.0
1
2
3
4
5
6
move

(r) (s)

(t) (u)

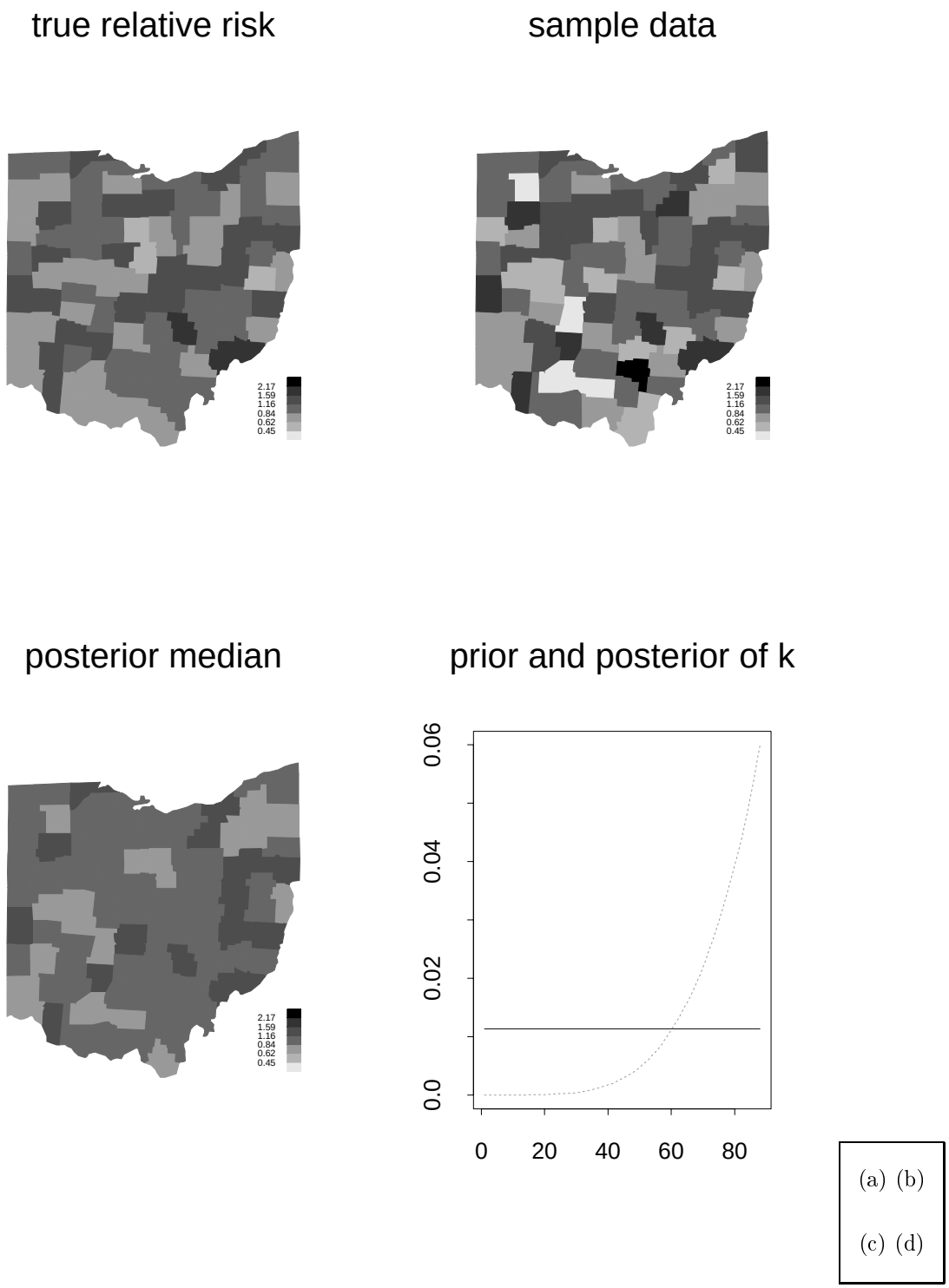
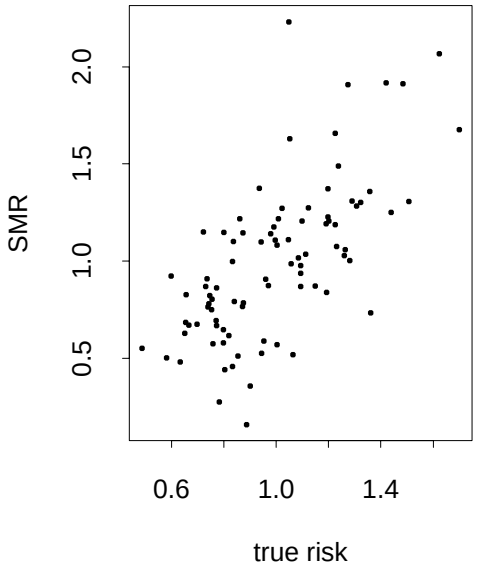
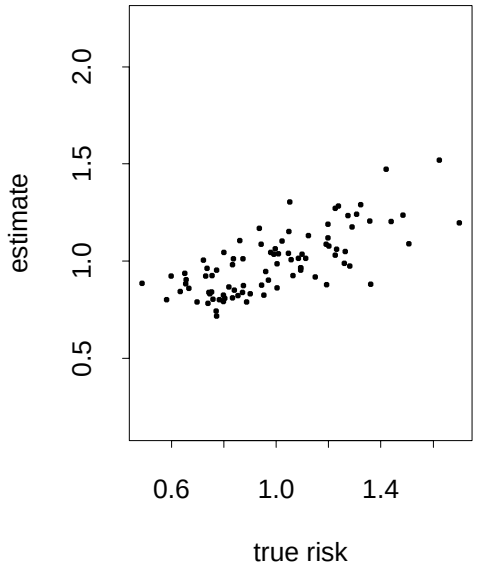
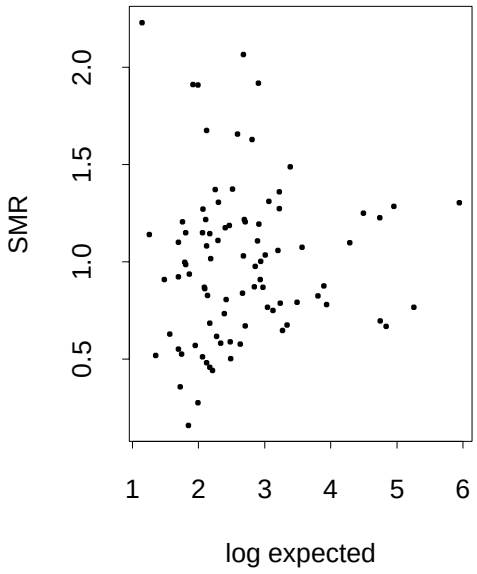
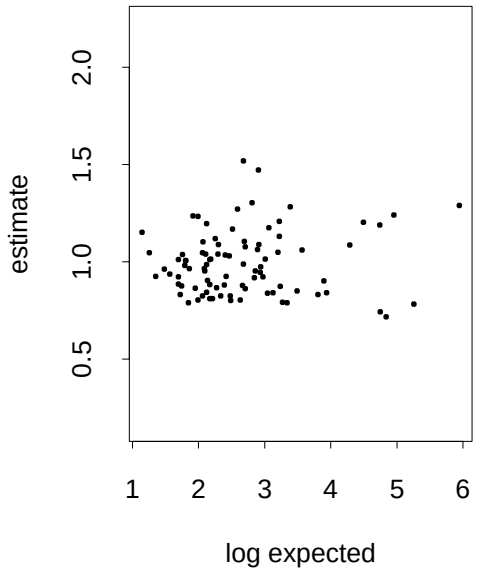


Figure 28: Analysis with (C1) and (S1).


SMR
0.5
1.0
1.5
2.0
0.6
1.0
1.4
true risk


estimate
0.5
1.0
1.5
2.0
0.6
1.0
1.4
true risk


SMR
0.5
1.0
1.5
2.0
1
2
3
4
5
6
log expected


estimate
0.5
1.0
1.5
2.0
1
2
3
4
5
6
log expected

(e) (f)

(g) (h)

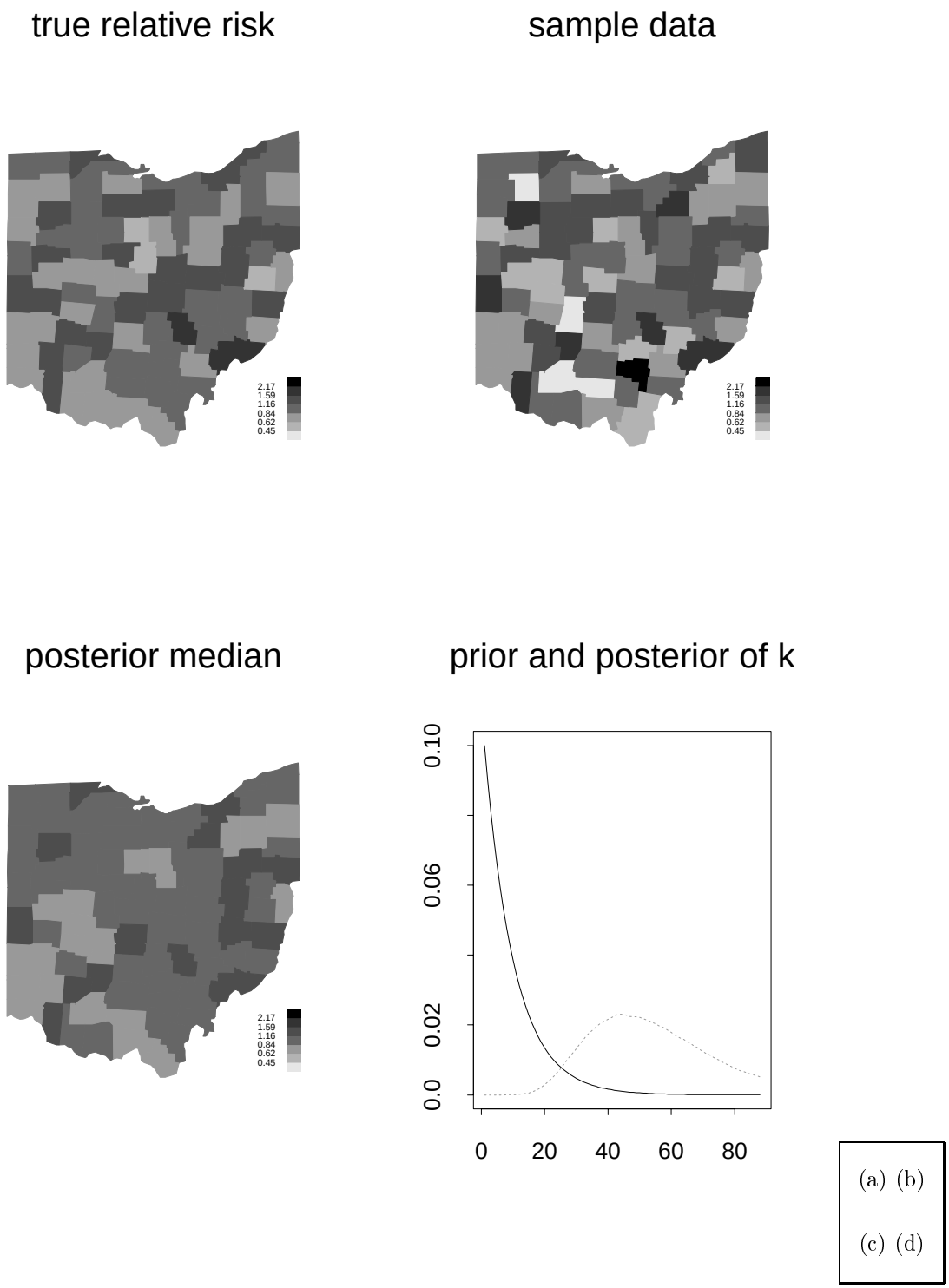
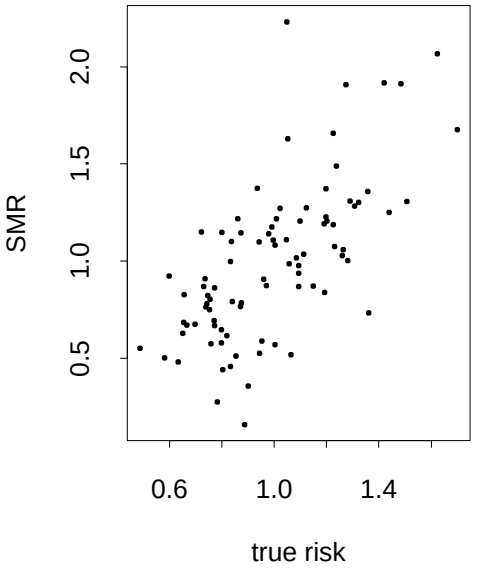
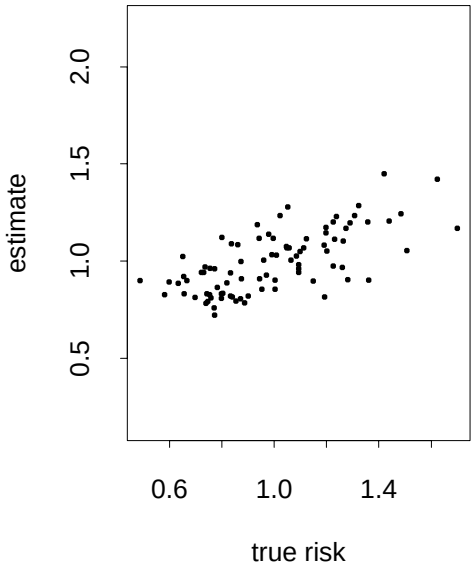
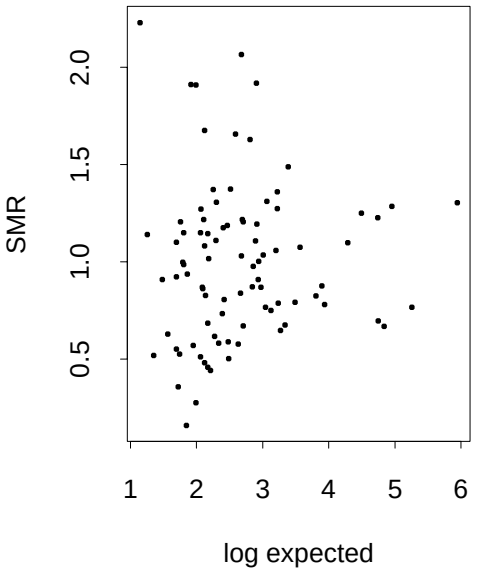
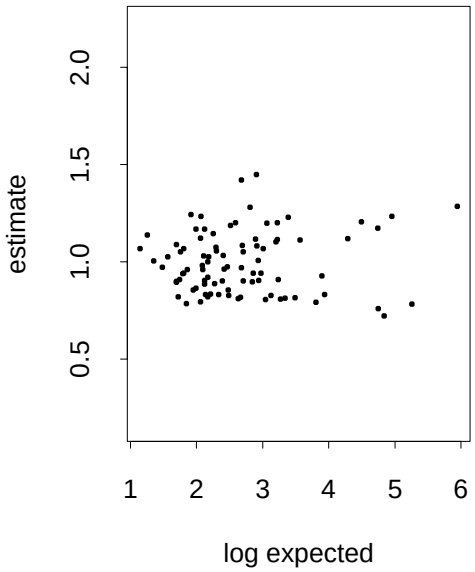


Figure 29: Analysis with (C2) and (S1).


SMR
0.5
1.0
1.5
2.0
0.6
1.0
1.4
true risk


estimate
0.5
1.0
1.5
2.0
0.6
1.0
1.4
true risk


SMR
0.5
1.0
1.5
2.0
1
2
3
4
5
6
log expected


estimate
0.5
1.0
1.5
2.0
1
2
3
4
5
6
log expected

(e) (f)

(g) (h)

3 Conclusions

The prior simulations seem to indicate that the degree of smoothing is approximately the same for all regions with only slight differences depending on the number of neighbors. The prior probabilities for forming a region on its own and the probability that two adjacent regions are within the same cluster depend more strongly on the number of neighbors and we therefore recommend to report both posterior and prior probabilities in any given application.

The analyses of the various simulated datasets suggest that our method seems to reconstruct a given risk surface quite well. Of course, the prior model assumes some sort of spatial structure, and therefore the performance of the method is not as good for pattern 4, where each latent relative risk is generated independently from the location of the region. However, it is a more philosophical question if the corresponding sample data could have been generated by a different risk surface *with* spatial structure.

Regarding sensitivity, it seems that the prior for k has only weak influence on the estimates of the relative risks as long as c is not too large. We therefore recommend to use rather small but positive values for c, in order to avoid problems with mulitmodality of the posterior for k (see the results for pattern 1). We did not find much evidence for sensitivity with respect to $p(\sigma^2)$. However, this might be different for other datasets.