Spatial CUSUM for Signal Region Detection

Xin Zhang and Zhengyuan Zhu

Iowa State University

09/25/2018

1 Introduction

2 Background and Problem Formulation

3 Proposed Method: Spatial CUSUM

4 Simulation Study

5 Real Data Application

6 Conclusion

Introduction

Detecting weak clustered signal in spatial data is important but challenging in applications , including

- astrophysics (Abazajian and Kaplinghat (2012); Gladders and Yee (2000));
- brain imaging analysis (Craddock et al. (2012); Zhang et al. (2011); Blumensath et al. (2013); Shen et al. (2013));
- epidemiology (Kulldorff and Nagarwalla (1995); Tango (2000); Wheeler (2007));
- meteorology (Sun et al. (2015)).

A more efficient detection algorithm can provide more precise early warning, and effectively reduce the decision risk and cost.

Introduction: Existing method

To date, many methods have been developed to detect signals with spatial structures:

Spatial Scan Statistics

- known as window statistics, was first proposed in Naus (1965).
- designed to find unusual clusters of randomly positioned points.
- to perform LRT on the scan windows of different sizes and locations and identify the significant windows as clusters.
- computationally too intensive; regular shape.

Introduction: Existing method

Multiple Testing and False Discovery Control

- testing several statistical hypotheses simultaneously
- instead of *p*-value,

$$\mathrm{FDR} = \mathbb{E}[\frac{\#\text{incorrect rejections}}{\#\text{rejected null hypotheses}}]. \tag{1}$$

- in spatial signal detection, the statistical hypotheses are about whether locations belongs to signal region or not.
- the prior knowledge about the null distibution; too conservative for weak signals.

Introduction

In this work,

- we propose a detection method named Spatial CUSUM (SCUSUM).
- SCUSUM employs the idea of the CUSUM procedure and false discovery rate controlling.
- we analyze theoretical properties of SCUSUM which gaurantees the high classification accuracy.
- both the simulation study and real data applilcation illustrate the detection effectiveness of SCUSUM.

1. Introduction
2. Background and Problem Formulation
3. Proposed Method: Spatial CUSUM
4. Simulation Study
5. Real Data Application
6. Conclusion

Background: CUSUM and Changpoint Detection

- The CUSUM procedure is well-known to locate changepoints in time series (Horváth and Hušková (2012); Cho et al. (2016),etc.);
- The CUSUM statistics has asymptotical distribution as Brownian Bridge. (Donsker's Theorem)

$$S_{z,N} = \frac{1}{(\hat{\sigma} N)^2} \sum_{j=1}^{N} (\sum_{i=1}^{j} x_i - j\bar{x})^2 \tag{2}$$

- Changepoint location is

$$t = \arg\max_k (\sum_{i=1}^{k} x_i - k\bar{x})^2 \tag{3}$$

Background: FDR control

- False discovery rate is a criterion designed to control the expected proportion of rejected null hypotheses that are incorrect rejections

$$\text{FDR} = \mathbb{E}[\frac{\#\text{incorrect rejections}}{\#\text{rejected null hypotheses}}]. \tag{4}$$

- Given p-values $\{p_i\}$ and significant level α, a threshold needs to be estimated to have FDR$\leq \alpha$.

Problem Formulation

- Spatial dataset $\{x(s), s \in \mathscr{D}\}$, $\mathscr{D}$ is square and observations are sampled on regular grids.
- Assume that under H_0, there is no signal region (i.e. $\mathscr{D}_{\mathscr{A}} = \varnothing$) and $x(s)$ has the same mean process μ;
- while under H_1, $x(s)$ has mean μ_1 if $s \in \mathscr{D}_{\mathscr{A}}$ and μ_0 if $s \in \mathscr{D}^c_{\mathscr{A}}$. (w.l.o.g, $\mu_0 = 0$)
- consider the following model:

$$x(s) = \mu_0 \mathbb{I}(s \in \mathscr{D}^c_{\mathscr{A}}) + \mu_1 \mathbb{I}(s \in \mathscr{D}_{\mathscr{A}}) + \epsilon(s), \tag{5}$$

- Our goal is to identify the signal region $\mathscr{D}_{\mathscr{A}}$.

1 Introduction

2 Background and Problem Formulation

3 Proposed Method: Spatial CUSUM

4 Simulation Study

5 Real Data Application

6 Conclusion

Framework for SCUSUM

Two steps for SCUSUM:

- Estimate the signal weight for each location, which should be large in $\mathscr{D}_{\mathscr{A}}$, while small in $\mathscr{D}_{\mathscr{A}}^{c}$
    - moving window technique
    - CUSUM cut-off
- Given a significant level α, determining a threshold
    - density estimation with boundary correction
    - false discovery rate control

The first step: signal weight estimation

- Inspired by the CUSUM procedure for changepoint detection in time series.
- Lack of natural order for spatial observations, the conventional CUSUM is impractical.
- Consider the spatial structures of signal regions, local block information is used to order observations.
- Divide $\mathscr{D}$ into b non-overlapping blocks $\{B_i\}_{i=1}^{b}$ of size $k \times k$. ($n = b \times k^2$)

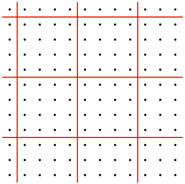


Figure: Block idea

Block information

Block information

- if $B_i \in \mathscr{D}_{\mathscr{A}}$, then $x(s) = \mu_1 + \epsilon(s), \ \forall x \in B_i$;
- if $B_i \in \mathscr{D}_{\mathscr{A}}^c$, then $x(s) = \mu_0 + \epsilon(s), \ \forall x \in B_i$;
- if B_i is at the boundary of $\mathscr{D}_{\mathscr{A}}$ and $\mathscr{D}_{\mathscr{A}}^c$, then

$$x(s) = \begin{cases} \mu_1 + \epsilon(s), & \text{with probability } p_i \\ \mu_0 + \epsilon(s), & \text{with probability } 1 - p_i \end{cases} \tag{6}$$

where p_i is the ratio of signal points inside B_i.

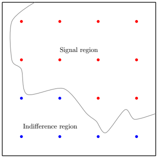


Figure: The block at the boundary. In this case, $p_i = 0.625$.

Two necessary sequences

Based on the division, construct two sequences:

- The first sequence is sample sequence: a 'representative' γ_i is randomly sample from block B_i.
- The second sequence by computing the block mean without the 'representative', $\tilde{\mu}_i = \frac{\sum_{x \in B_i} x - \gamma_i}{\sum_{x \in \mathscr{D}} \mathbb{I}(x \in B_i) - 1}$.

Properities of the sequences

With the above analysis, we can derive the following results:

$$\gamma_i = \begin{cases} \mu_1 + \epsilon, & \text{if } B_i \in \mathscr{D}_{\mathscr{A}} \\ \mu_0 + \epsilon, & \text{if } B_i \in \mathscr{D}_{\mathscr{A}}^c \\ \mu_1 z + \mu_0(1-z) + \epsilon, & \text{if } B_i \text{ at boundary} \end{cases} \tag{7}$$

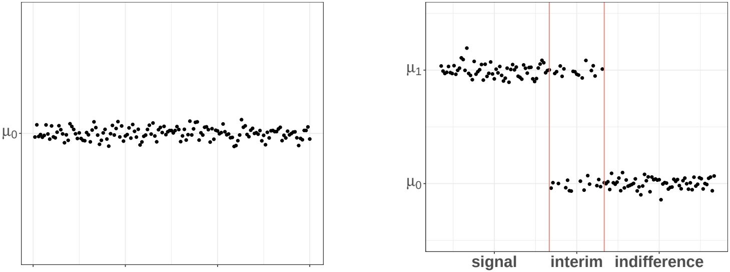
$$\mathbb{E}[\tilde{\mu}_i] \begin{cases} = \mu_1, & \text{if } B_i \in \mathscr{D}_{\mathscr{A}} \\ = \mu_0, & \text{if } B_i \in \mathscr{D}_{\mathscr{A}}^c \\ \approx p_i\mu_1 + (1-p_i)\mu_0, & \text{if } B_i \text{ at boundary} \end{cases} \tag{8}$$

where $z \sim Ber(p_i)$.

Properities of the sequences

- the closer B_i is to $\mathscr{D}_{\mathscr{A}}$, the more likely it has large γ_i and $\tilde{\mu}_i$, and vice versa.
- rearrange $\{\gamma_i\}_{i=1}^b$ according to the decreasing order of $\{\tilde{\mu}_i\}_{i=1}^b$, denoted as $\{\gamma_i^*\}_{i=1}^b$.
- if there is no signal region, then $\{\gamma_i^*\}$ should be around μ_0;
- With signal region, $\{\gamma_i^*\}_{i=1}^b$ should have three parts: signal blocks is around μ_1, the interim from μ_1 to μ_0 and indifference blocks around μ_0

Diagram for sequences



(a) under H_0 (b) under H_1

Figure: The possible patterns of $\{\gamma_i^*\}_{i=1}^b$: (a) presents the scenario without signal region and $\{\gamma_{(i)}^*\}_{i=1}^b$ are around μ_0; (b) shows the pattern with signal region and there are three parts: signal, interim and indefference.

Theoretical results

Lemma

Based on assumed model, $\{\gamma_i\}$ and $\{\tilde{\mu}_i\}$ are indepedent. As the number of observations in each block n_i goes to infinity,

- *under H_0 : there is no signal region, then $\{\gamma_i^*\}$ is an i.i.d sequence;*
- *under H_1 : signal region exists, then*

$$\mathbb{E}[\gamma_i^*] = \begin{cases} \mu_1, & \text{if } 0 \le i < l_1 \\ \in (\mu_0, \mu_1) & \text{if } l_1 \le i < l_2 \\ \mu_0, & \text{if } l_2 \le i \le n, \end{cases} \tag{9}$$

and l_1 is the number of blocks inside $\mathscr{D}_{\mathscr{A}}$, and $(b - l_2)$ is the number of blocks inside $\mathscr{D}_{\mathscr{A}^c}$.

Theoretical results

- Lemma shows that under H_1, the projected sequence $\{\gamma_i^*\}_{i=1}^b$ has a changepoint in $[l_1, l_2]$.
- The conventional CUSUM could help locate a cut-off index near or inside the interval.
- Compute the CUSUM statistics for $\{\gamma_i^*\}_{i=1}^b$ at each location:

$$\tilde{\gamma}_r = |\sum_{i=1}^{r} \gamma_i^* - \frac{r}{b} \sum_{i=1}^{b} \gamma_i^*|. \tag{10}$$

- The cut-off index is $t = \arg\max_i \tilde{\gamma}_i$.

Theoretical results

Theorem

Under the alternative hypothesis: if the signal region exists, as the number of block b and the number of observations in each block n_i go to infinity, then the cut-off index t based on CUSUM procedure will fall into the interval $[l_1, l_2]$ with probability 1, i.e. $\mathbb{P}(l_1 \leq t \leq l_2) = 1$ as $b \to \infty$ and $\min n_i \to \infty$.

- The theorem ensures that the cut-off procedure could asymptotically separate the signal region and indifference.
- With the given spatial domain, as the block size becomes finer (i.e. $b \to \infty$), the ratio of the blocks at the boundary is decreasing, $(l_2 - l_1)/b \to 0$.
- Hence, the ratio of misclassified locations goes to zero.

Practical Problems

Above theoretical results require $b \to \infty$ and $\min n_i \to \infty$.
In practice, with limited observations, the detected result might be affected by

- the block division, especially when the signal region is not regular;
- representatives selection.

Practical Solution

- Consider the moving window technique.
- Select one point (x, y) from $\{(p + \frac{1}{2}, q + \frac{1}{2}),\ p = 1 : k,\ q = 1 : k\}$ as inital point.
- Eliminate the effect of initial point by going through all the possible inital points.
- Extract the local information and eliminate the effect of randomly sampled representatives repeated the previous procedure more than once.

The signal weights $\{w(s)\}$ (or $\{\tilde{w}(s)\}$) can be computed as the detected frequency of each location.

Diagram for Moving Window

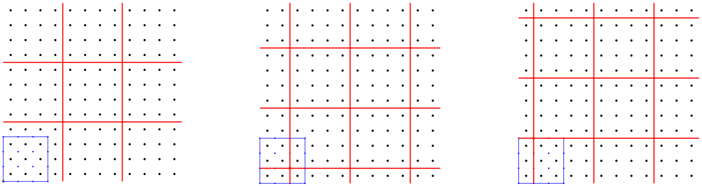


Figure: Moving window to divide spatial domain: The blue points inside blue square is $\{(p + \frac{1}{2}, q + \frac{1}{2}),\ p = 1 : k,\ q = 1 : k\}$, we select initial grid point from the set; The black points present the observed locations; red grids are boundary lines for blocks.

Algorithm of Moving Window detecting method

Algorithm 1 Moving Window detecting method for signal weight

Require: observed data $\{x(s)\}$ on grid $\{(p,q),\ p=1:n,\ q=1:n\}$, neighbor size k, repeat times m;

Ensure: corresponding signal weights $\{w(s)\}$ or $\{\tilde{w}(s)\}$;

for (x,y) in $\{(p+\frac{1}{2}, q+\frac{1}{2}),\ p=1:k,\ q=1:k\}$ **do**

  Divide $\mathscr{D}$ into blocks $\{B_i\}_{i=1}^{b}$ of size $k \times k$ based on (x,y);

  Sample one observation from each block γ_i;

  Estimate the block mean $\tilde{\mu}_i$;

  Reorder $\{\gamma_i\}_{i=1}^{b}$ according to $\{\tilde{\mu}_i\}_{i=1}^{b}$ decreasingly as $\{\gamma_i^*\}_{i=1}^{b}$;

  Conduct CUSUM transformation on $\{\gamma_{(i)}^*\}_{i=1}^{b}$ as $\{\tilde{\gamma}_{(i)}^*\}_{i=1}^{b}$;

  Find the location t where $\{\tilde{\gamma}_i^*\}_{i=1}^{b}$ reaches maximum;

  Define the blocks corresponding to the first t elements in $\{\gamma_i\}_{i=1}^{b}$ as signal block, and so do the observations in these blocks;

end for

Compute corresponding signal weight $w(s) = \frac{\text{detected times for } x(s)}{k^2}$;

(Option)

Repeat above produce m times and obtain $\{w^i(s)\}_{i=1}^{m}$;

Compute the average signal weights at each location $\{\tilde{w}^i(s)\}$: $\tilde{w}^i(s) = \sum_{i=1}^{m} w^i(s)/m$;

The second step: Threshold estimation with FDR

With the estimated signal weight and a given significant level α, we could identify the signal region with false discovery control.

- Weights $\in [0, 1]$, considered as the possibilities that the locations have signals.
- Density estimation with the boundary correction to estimate the distribution of the signal weights, $f(x)$.
    - the local polynomial density estimation method from Cattaneo et al. (2017)

The second step: Threshold estimation with FDR

Under the null hypothesis, $\{\gamma^*_{(i)}\}_{i=1}^{b}$ are i.i.d. and the correponding blocks are random indexed.
Hence, the distribution for signal weight $f(x)$ is symmetric and has lower value with $x = 0$ and $x = 1$.

Lemma

Under the null hypothesis H_0, as the number of block b and the number of observations in each block n_i go to infinity, then the density for signal weights $f(x)$ is symmetric.

The second step: Threshold estimation with FDR

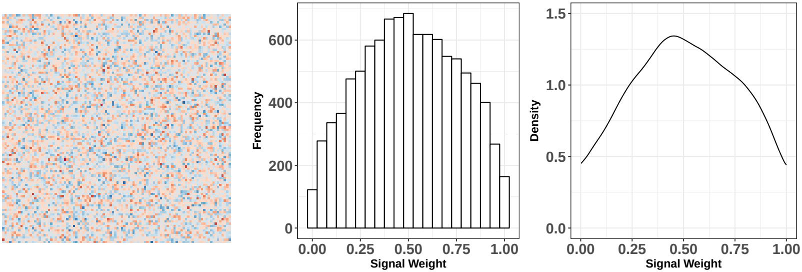


Figure: The idea for the second step: Under H_0, there is no signal (first column.) The histogram of signal weights are shown in the second column and it's symmetric with 'peak' around 0.5. The estimated density is shown in the third column.

The second step: Threshold estimation with FDR

Under the alternative hypothesis, $f(x) = f_{H_0}(x) + f_{H_1}(x)$.

- The observations inside the signal region are more likely to be detected, i.e. the corresponding signal weight gets close to 1 and vice versa for the observations inside the indifference region.
- With finer block division, the fraction of the observations in the blocks at the boundary goes to 0.
- $f_{H_1}(x)$ has the 'peak' near 1 and $f_{H_0}(x)$ has the 'peak' around 0.

The second step: Threshold estimation with FDR

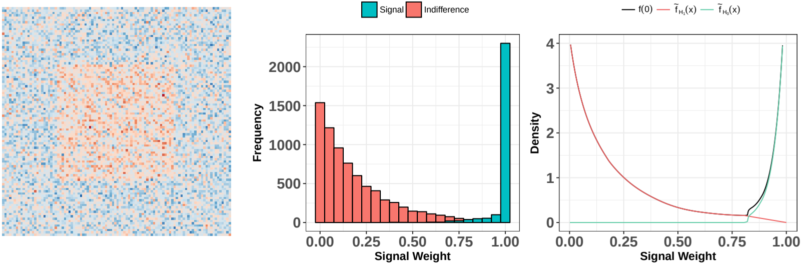


Figure: The idea for the second step: Under H_1, the histogram (the second column) is composited with two parts and has higher values around boundaries 0 and 1, lower values at the middle part. The estimated densities are shownd in third column: the black curve is $f(x)$, green one is estimated null density $\tilde{f}_{H_0}(x)$ and red one is estimated alternative density $\tilde{f}_{H_1}(x) = f(x) - \tilde{f}_{H_0}(x)$.

The second step: Threshold estimation with FDR

- $f(x)$ has two 'peaks' near the boundaries seperately and a 'valley' in the middle of $[0,1]$.
- Using the line search to locate the 'valley', say $(t^*, f(t^*))$, and conducting linear interpolation between the two points $(t^*, f(t^*))$ and $(1,0)$.
- the null density $f_{H_0}(x)$ is controlled by

$$\tilde{f}_{H_0}(x) = \begin{cases} f(x), \text{ if } 0 \leq x \leq t^* \\ f(t^*)(1 - \frac{x - t^*}{1 - t^*}), \text{ if } t^* < x \leq 1 \end{cases} \tag{11}$$

The second step: Threshold estimation with FDR

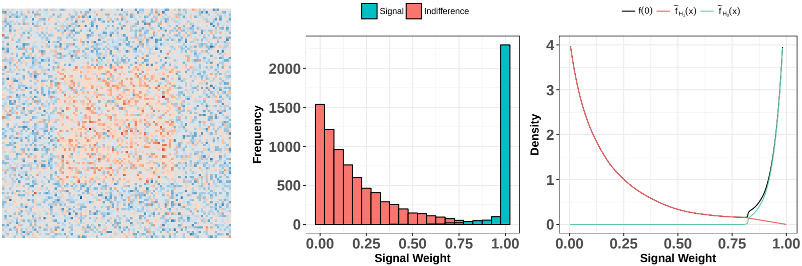


Figure: The idea for the second step: Under H_1, the histogram (the second column) is composited with two parts and has higher values around boundaries 0 and 1, lower values at the middle part. The estimated densities are shownd in third column: the black curve is $f(x)$, green one is estimated null density $\tilde{f}_{H_0}(x)$ and red one is estimated alternative density $\tilde{f}_{H_1}(x) = f(x) - \tilde{f}_{H_0}(x)$.

The second step: Threshold estimation with FDR

- Recall the definition of the marginal false discovery rate (mFDR) (Genovese and Wasserman (2002); Sun et al. (2015)):

$$\text{mFDR} = \frac{\mathbb{E}[\#\text{false positive}]}{\mathbb{E}[\#\text{rejected}]}. \tag{12}$$

- Control mFDR with given significant level α by finding a threshold c so that

$$c = \arg\min_{x}\left(\frac{\tilde{f}_{H_0}(x)}{f(x)} \leq \alpha\right). \tag{13}$$

- The observations with signal weight larger than c are the detected signals.

The signal region detection method

The second step is summarized in the following algorithm.

Algorithm 2 The signal region detection method

Require: Signal weights $\{w(s)\}$ or $\{\tilde{w}^i(s)\}$, signficant level α;

Ensure: corresponding detected result;

1: Estimate the density curve $f(x)$ based on signal weights, $x \in [0, 1]$;

2: Estimate null density $f_{H_0}(x)$ and alternative density $f_{H_1}(x)$;

3: Compute mFDR and find the threshold c;

4: Obtain the detected result with the threshold c;

Neighbor size selection

- Basic idea: the larger k means the larger block and tends to over-smooth; while the smaller k implying the small block might lose spatial information.
- To make the "right" cut-off:
    - variance of $\{\tilde{\mu}_i\}$ should be small, so that we could reasonably rearrange the 'representatives' $\{\gamma_i\}$;
    - the length of $\{\gamma_i^*\}$ should be long, which could ensure the cut-off location t fall into $[l_1, l_2]$ with probability 1.
- For the first point, the number of observations in each block n_i go to infinity;
- For the second point, the number of blocks b go to infinity;

Neighbor size selection

The relationship between b, k and n_i could be approximated as:

$$\begin{cases} \min n_i \approx k^2 \\ b \approx n/k^2 \end{cases} \tag{14}$$

The trade-off:

$$k_{opt} = \arg\min_k k^2 + C_1 \frac{n}{k^2} = \sqrt[4]{C_1 n}, \tag{15}$$

where C_1 is a given weight to reflect which part we want to emphasize and n is the total number of observations.

1 Introduction

2 Background and Problem Formulation

3 Proposed Method: Spatial CUSUM

4 Simulation Study

5 Real Data Application

6 Conclusion

Simulation Study: Independent noise

Model Setting

- All the examples are simulated in the image with 100×100 pixels.
- Data generating model:

$$x(i,j) = \mu(i,j) + \epsilon(i,j), \quad i,j = 1, ..., 100. \tag{16}$$

- Independent standard normal distribution $N(0,1)$ for noise term ϵ.
- $\mu(i,j) = 0$ for $(i,j) \in \mathscr{D}^c_{\mathscr{A}}$, and $\mu(i,j) \neq 0$ for $(i,j) \in \mathscr{D}_{\mathscr{A}}$.

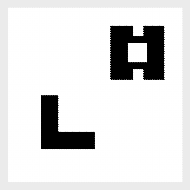


Figure: Ground truth

Simulation Study

Simulation Setting

- The repeated time m in Algorithm 1 as 50.
- μ from 0.8 to 2 and k from $\{3, 5, 10\}$.
- The significant level α as 0.05.
- Repeating simulation 100 times under each setting.
- Comparing with FDR_L Zhang et al. (2011), in which the spatial structures are considered. (a standard normal test on each pixels first).

Simulation Study: Detected accuracy comparision

Table: Detected accuracy comparision between SCUSUM and FDR_L

neighbor size	Signal μ	SCUSUM FN	SCUSUM FP	SCUSUM FDR	FDR_L FN	FDR_L FP	FDR_L FDR
k=3	0.8	0.6338	0.0010	0.0186	0.9831	0.0003	0.0917
	1	0.4286	0.0010	0.0115	0.9380	0.0006	0.0484
	1.5	0.1953	0.0005	0.0044	0.5158	0.0039	0.0508
	2	0.1271	0.0006	0.0042	0.1478	0.0072	0.0535
k=5	0.8	0.3750	0.0009	0.0094	0.8034	0.0020	0.0586
	1	0.2750	0.0009	0.0082	0.5351	0.0047	0.0612
	1.5	0.1599	0.0019	0.0147	0.0971	0.0112	0.0770
	2	0.1014	0.0027	0.0194	0.0159	0.0150	0.0928
k=10	0.8	0.3240	0.0066	0.0588	0.1433	0.0286	0.1793
	1	0.2526	0.0084	0.0692	0.0569	0.0394	0.2163
	1.5	0.1642	0.0133	0.0968	0.0076	0.0651	0.3051
	2	0.1288	0.0149	0.1034	0.0027	0.0751	0.3358

(FN: false negative; FP: false positive; $\alpha = 0.05$)

Simulation Study: Detected accuracy comparision

For SCUSUM,

- FDR controlled under 0.05 with small neighbor size $k = 3, 5$;
- outperforms both in false negative and false positive with small k;
- FN$= 0.6338$ and FP$= 0.001$ when $k = 3$ and $\mu = 0.8$;
- FN< 0.50 while FP< 0.003 when $k = 5$;

While for FDR$_L$

- have FDR a little bit higher than 0.05;
- have large false negative with small k and μ;
- FN$= 0.9831$ and FP$= 0.003$ when $k = 3$ and $\mu = 0.8$;
- FN> 0.50 and FP≈ 0.003 with $k = 5$ and low signal strength 0.8 and 1.0;

Simulation Study: Probability map

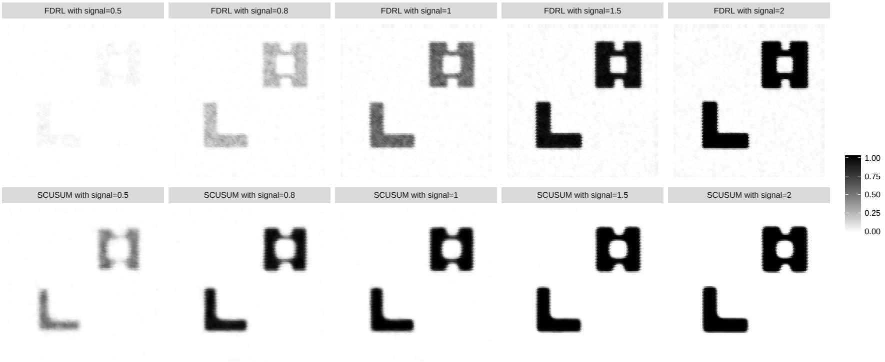


Figure: Comparision of the detection probability for SCUSUM and FDR_L: the darker the color means the higher detected probability.($k = 5, \alpha = 0.05$)

Simulation Study: Probability map

- With hige signal strength $\mu \geq 1.5$, the two methods have almost the same performance; while the signal is much too low (e.g.$\mu = 0.5$,) SCUSUM has a better detected result than FDR_L.
- In the results of FDR_L there are some shadows outside of 'L' and 'H' signal region, which are false positive; while for SCUSUM, the detections for indifference region are more 'white' (no shadows.)

Simulation Study: Weak dependent noise

- Using the Exponential Covariance Model (Gelfand et al. (2010)) to generate dependence noise.
- Generating data from multivariate normal distribution with mean, μ_0 for $\mathscr{D}_{\mathscr{A}^c}$ and μ_1 for $\mathscr{D}_{\mathscr{A}}$, and dependent covariance matrix.
- Covariance matrix

$$C(s_i, s_j) = \exp(-\frac{\|s_i - s_j\|}{r}), \tag{17}$$

  $\|s_i - s_j\|$ is the distance between the two location, the larger r means the stronger dependence.

- r from $\{0.1, 0.3, 0.5\}$, and the corresponding covariances for unit distance $\{0.00004, 0.03567, 0.13533\}$.

Simulation Study: Probability map

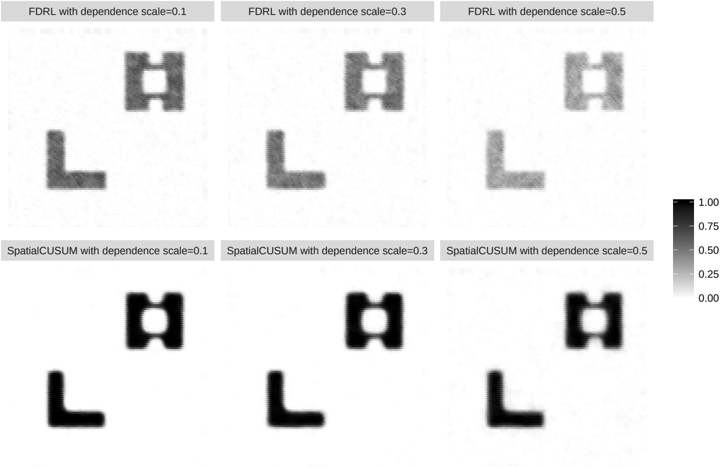


Figure: Comparision of the detection probability for SCUSUM and FDR_L with different dependence scales

Simulation Study: Detected accuracy comparision

Table: Summary for detection probabilities on dependence noise

Dependence Scale		$r = 0.1$	$r = 0.3$	$r = 0.5$
false negative	SCUSUM	0.2721	0.2820	0.3108
	FDR_L	0.5213	0.5879	0.7317
false positive	SCUSUM	0.00095	0.00154	0.00418
	FDR_L	0.00499	0.00411	0.00280
FDR	SCUSUM	0.0086	0.0141	0.0386
	FDR_L	0.0622	0.0604	0.0590

- with the weak dependence, SCUSUM could still detect the signal region efficiently while control the false posive.
- larger dependence scale leads to larger false positive, false negative and FDR for SCUSUM. (for larger scale dependence noise, choose relatively larger blocks.)
- SCUSUM could recognize the signal region with higher probability than FDR_L, when the noise dependence is weak.

1 Introduction

2 Background and Problem Formulation

3 Proposed Method: Spatial CUSUM

4 Simulation Study

5 Real Data Application

6 Conclusion

Real Data: FMRI data

- Functional MRI (fMRI) is a well-known technique to measures brain activity.
- Each individual image has 128×128 pixels.
- The pixels' values are the transformations of p-values from a previous study, following $N(0, 1)$.
- We only care about detection of the regions with positive values ('darker' pixels).

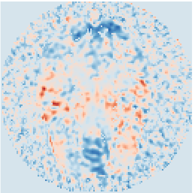


Figure: An example of the raw image

Real Data: Comparision

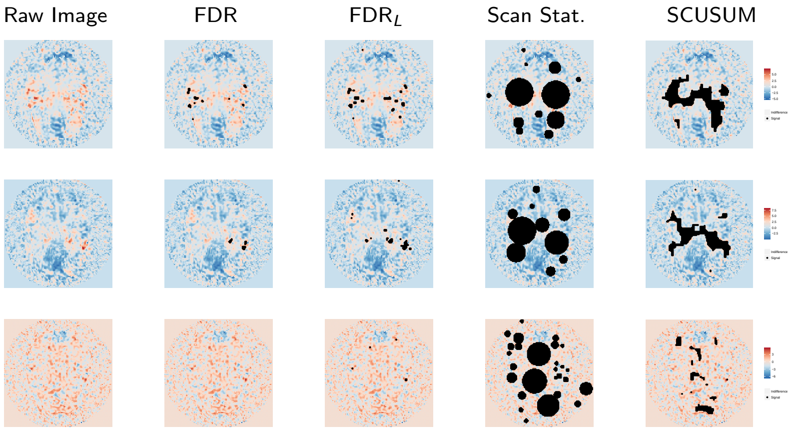


Figure: Detection comparision ($\alpha = 0.0001$)

Real Data: Comparision

- For the conventional FDR:the lack of identification phenomenon happened (eg. the third slice); without considering spatial correlation, some detected 'signals' are scattered around (eg. the first slice).
- For FDR_L, many weak signals are missed, e.g. the bottom part in the third slice.
- For Scan Stat., even though almost all the 'hot' pixels are detected, the 'signal' regions are too large, which is doubtable.
- SCUSUM is more likely to detect weak signals and identifiy several irregularly shaped clusters.

1. Introduction
2. Background and Problem Formulation
3. Proposed Method: Spatial CUSUM
4. Simulation Study
5. Real Data Application
6. Conclusion

Conclusion

- In this work, we proposed a spatial signal detection method, SCUSUM.
- SCUSUM consists of two parts, signal weights estimation and threshold estimation.
- SCUSUM is noise model-free.
- SCUSUM tends to detect spatially gouped and weak signals, which might be missed by some existing methods.
- Our simulation study shows that our method has a better performance compared to FDR_L method.
- SCUSUM is applied to a real fMRI data to illustrate its detection effectiveness and many weak signals are detected.

References

Abazajian, Kevork N and Manoj Kaplinghat (2012), 'Detection of a gamma-ray source in the galactic center consistent with extended emission from dark matter annihilation and concentrated astrophysical emission', *Physical Review D* **86**(8), 083511.

Aue, Alexander, Robertas Gabrys, Lajos Horváth and Piotr Kokoszka (2009), 'Estimation of a change-point in the mean function of functional data', *Journal of Multivariate Analysis* **100**(10), 2254–2269.

Benjamini, Yoav and Daniel Yekutieli (2001), 'The control of the false discovery rate in multiple testing under dependency', *Annals of statistics* pp. 1165–1188.

Benjamini, Yoav and Yosef Hochberg (1995), 'Controlling the false discovery rate: a practical and powerful approach to multiple testing', *Journal of the royal statistical society. Series B (Methodological)* pp. 289–300.

Blumensath, Thomas, Saad Jbabdi, Matthew F Glasser, David C Van Essen, Kamil Ugurbil, Timothy EJ Behrens and Stephen M Smith