# EcheScan User Guide

March 2020

This is partly excerpted and re-edited from Kurihara et al.(2020).

## 1 Overview

We describe the design and implementation of software for the echelon analysis named *EcheScan*. The EcheScan constructs the echelons from input data files which consists of univariate values and neighbors for each lattice and visualizes the result of a dendrogram. It can also detect hotspot clusters based on Poisson model using the echelon scan technique by reading a file describing the observed and expected values of each lattice. In addition, users can obtain the analysis results output as a file. Developed by R Shiny environment and visualized in this web(`https://fishi.ems.okayama-u.ac.jp/echescan`), users can access the software through the Internet. Table 1 summarizes the files required for input and the files that can be output.

Table 1: Input and output files of EcheScan

| I O | File | Contents | Notes |
|---|---|---|---|
| I | Neighborhood information | Neighbor information of each lattice | File format: `txt, csv` |
| I | Univariate | Value ($h$) of each lattice | File format: `txt` |
| I | Case & expectation | Observed($c$) and expected($\lambda$) values of each lattice | For hotspot detection based on Poisson model<br>File format: `txt` |
| O | Echelon table | Details of echelons | File format: `csv` |
| O | Lattices forming echelon | Lattice information within each echelon | File format: `csv` |
| O | Echelon dendrogram | Graphical representation of echelons | File format: `png, pdf, eps` |
| O | Hotspot table | Details of detected hotspots | File format: `csv` |
| O | Echelon dendrogram with scanning | Graphical representation of echelon scan technique | File format: `png, pdf, eps` |

# 2 Input files

## 2.1 Neighbor information file

The neighborhood information file provides the labels and neighbors of each lattice. The first column of each line is the lattice label, and the next columns specify the line numbers that are neighbors of the lattice described in the first column. Therefore, the total number of lines in this file is $NL$. For example, when “lattice1” is neighbor of “lattice2” and “lattice4”, the file becomes

| | | |
|---|---|---|
| lattice1 | 2 | 4 |
| lattice2 | 1 | ... |
| lattice3 | ... | |
| lattice4 | 1 | ... |
| ... | | |

To make the neighbors correspond to the opposite viewpoint, it must be written “1” (meaning “lattice1”) in the lines of “lattice2” and “lattice4”.
Tabs and commas are allowed to separate each item.

## 2.2 Univariate file

The univariate file provides the values ($h$) of each lattice in one column. The number of lines is $NL$. Here, the order of each value must be identical to that of lattices in the neighborhood information file.

## 2.3 Case & expectation file

The case & expectation file provides the observed ($c$) and expected ($\lambda$) values for each lattice to detect the hotspot, which forms $NL$ lines of two columns. Similar to the univariate file, the order must be identical to that of lattices provided by the neighborhood information file.

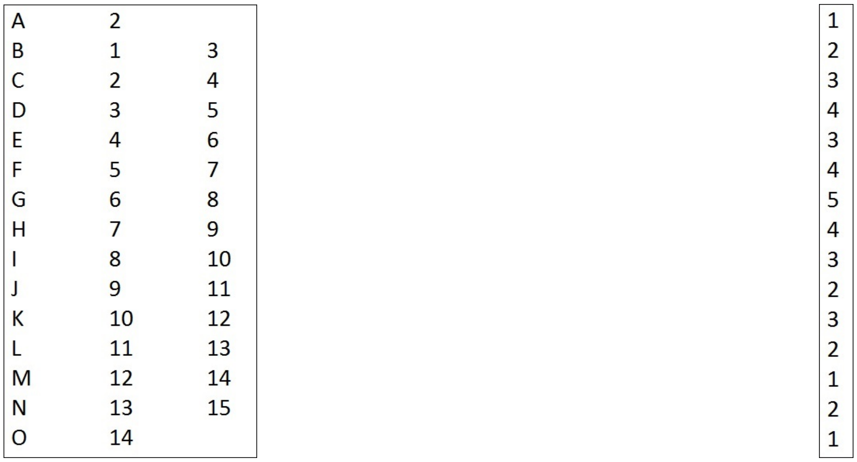
# 3 Execution examples

## 3.1 One dimensional lattice

We introduce how to use the EcheScan for the one dimensional lattice data shown in Table 2. With the software application, the neighborhood information file (`dim1nb.txt`) and the univariate file (`dim1h.txt`) are prepared as shown in Figure 1. In Figure 1(left), for example, lattice "C" (the third line) indicates that it is neighboring to the second line (B) and the fourth line (D).

Table 2: One dimensional spatial lattice data.

| $i$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $ID$ | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O |
| $h(i)$ | 1 | 2 | 3 | 4 | 3 | 4 | 5 | 4 | 3 | 2 | 3 | 2 | 1 | 2 | 1 |



```
A	2
B	1	3
C	2	4
D	3	5
E	4	6
F	5	7
G	6	8
H	7	9
I	8	10
J	9	11
K	10	12
L	11	13
M	12	14
N	13	15
O	14
```

```
1
2
3
4
3
4
5
4
3
2
3
2
1
2
1
```

Figure 1: Contents of the neighborhood information file (`dim1nb.txt`; left) and the univariate file (`dim1h.txt`; right) for the one dimensional lattice data.

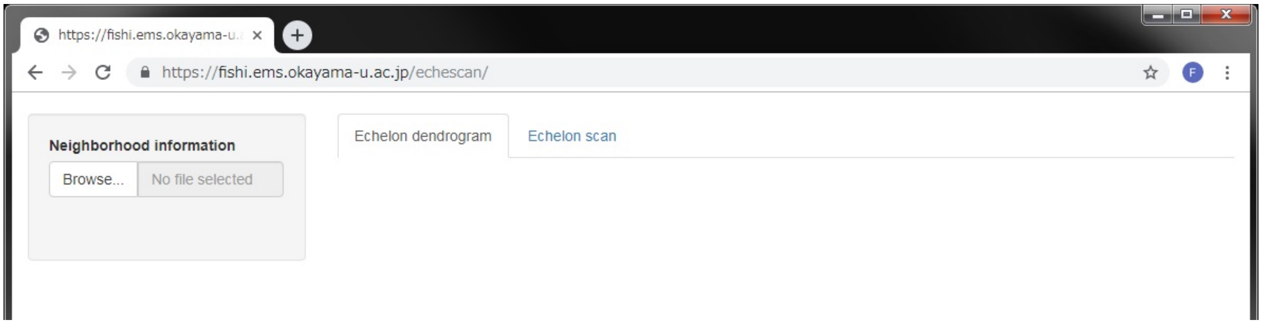
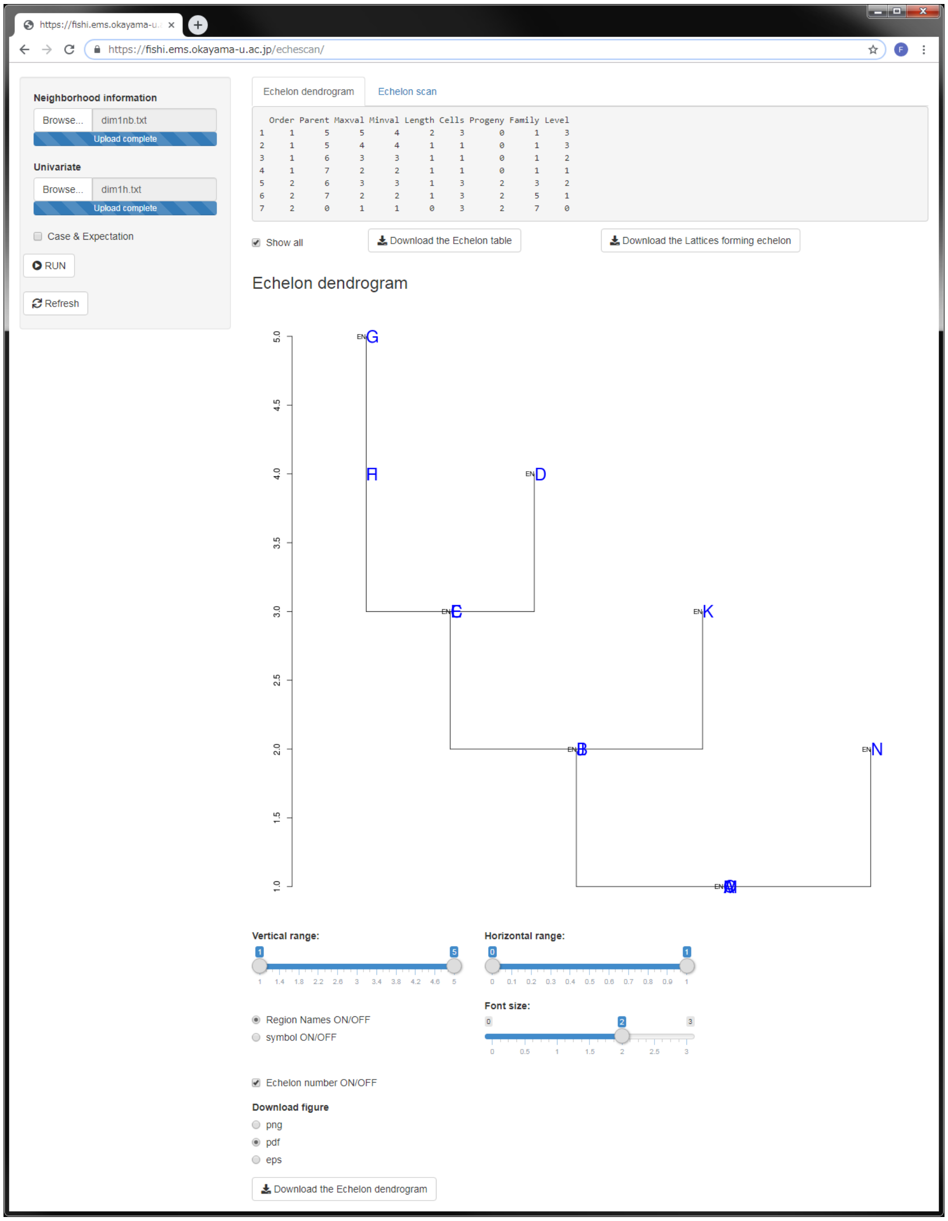


Figure 2: Start screen of the software.

Figure 2 shows the start screen of the software. First, we select the `dim1nb.txt` from [Browse ...] of "Neighborhood information" on the left side of the screen. If there is no error in the contents of the loaded file, "Univariate" appears, and then we select the `dim1h.txt`. Next, when we click [Run], an echelon analysis is executed, and the result is displayed on the [Echelon dendrogram] tab (Figure 3). The table at the top of Figure 3 displays tabular information for each echelon. The first fields is echelon number ($EN$). The second field is $Order$, which gives an integer value greater than or equal to 1; "1" means a peak, "2" means a foundation of order 1s, "3" means a foundation of order 2s, and so on. The third field is $Parent$, which gives the echelon number of the parent. The forth field is $Maxval$, which gives the maximum value. The fifth field is $Minval$, which gives minimum value. The sixth field is $Length$, which is the length of $Maxval - parent's\, Maxval$. The seventh field is $Cells$, which gives the number of lattices. The eighth field is $Progeny$, which gives the number of ascendants (children) for the echelon. The nineth field is $Family$, which gives the number of echelons in the family. The final field is $Level$, which gives the number of echelons in the ancestor. These information and dendrogram of echelons can also be output as a file. The details of $Variable$ are shown in the papers of Myers et al. (1997) and Kurihara et al. (2000).



https://fishi.ems.okayama-u.ac.jp/echescan/

Neighborhood information

Browse... dim1nb.txt

Upload complete

Univariate

Browse... dim1h.txt

Upload complete

☐ Case & Expectation

RUN

Refresh

Echelon dendrogram | Echelon scan

```
  Order Parent Maxval Minval Length Cells Progeny Family Level
1     1      5      5      4      2     3       0      1     3
2     1      5      4      4      1     1       0      1     3
3     1      6      3      3      1     1       0      1     2
4     1      7      2      2      1     1       0      1     1
5     2      6      3      3      1     3       2      3     2
6     2      7      2      2      1     3       2      5     1
7     2      0      1      1      0     3       2      7     0
```

☑ Show all

Download the Echelon table

Download the Lattices forming echelon

Echelon dendrogram

Vertical range: 1 – 5

Horizontal range: 0 – 1

● Region Names ON/OFF

○ symbol ON/OFF

Font size: 2

☑ Echelon number ON/OFF

Download figure

○ png

● pdf

○ eps

Download the Echelon dendrogram

Figure 3: Execution result of echelon analysis for the one dimensional lattice data.

## 3.2 5-by-5 lattice

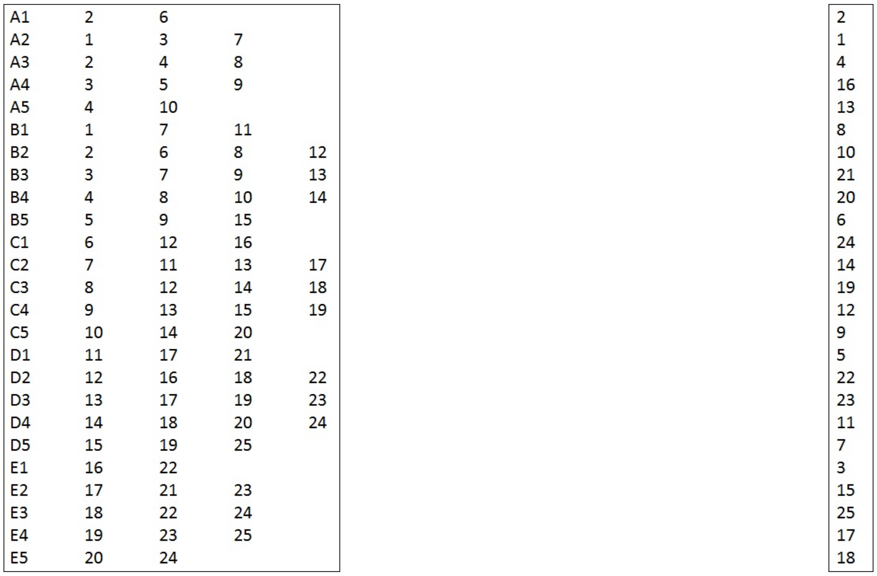
Two-dimensional lattice data such as remote sensing data are pixels of digital values over the $m$-by-$n$ regularly spaced lattice area.

Table 3: Digital values and lattice ID (right side) for a 5-by-5 lattice.

| | A | B | C | D | E |
|---|---|---|---|---|---|
| 1 | 2 | 8 | 24 | 5 | 3 |
| 2 | 1 | 10 | 14 | 22 | 15 |
| 3 | 4 | 21 | 19 | 23 | 25 |
| 4 | 16 | 20 | 12 | 11 | 17 |
| 5 | 13 | 6 | 9 | 7 | 18 |

| | A | B | C | D | E |
|---|---|---|---|---|---|
| 1 | A1 | B1 | C1 | D1 | E1 |
| 2 | A2 | B2 | C2 | D2 | E2 |
| 3 | A3 | B3 | C3 | D3 | E3 |
| 4 | A4 | B4 | C4 | D4 | E4 |
| 5 | A5 | B5 | C5 | D5 | E5 |

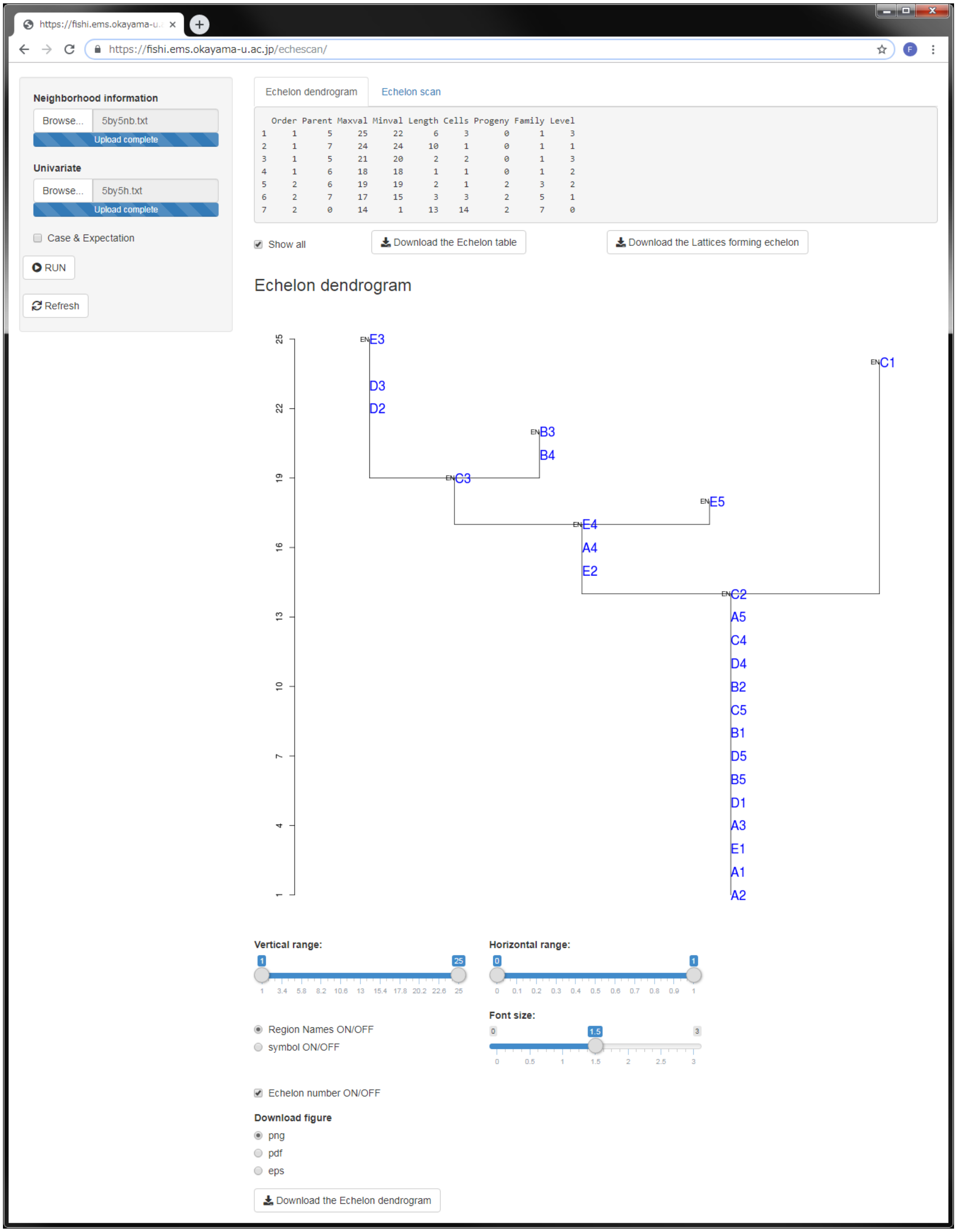
We apply the EcheScan to the 5-by-5 lattice data ($m = 5, n = 5$) in Table 3. Here, each cell is given a rook-type neighbor. The neighborhood information file (`5by5nb.txt`) and the univariate file (`5by5h.txt`) we prepare are shown in Figure 4, respectively. As in the case of one dimensional lattice data, we move the `5by5nb.txt` to "Neighborhood information" and the `5by5h.txt` to "Univariate". By clicking on [Run], the result of echelon table and dendrogram are displayed on the [Echelon dendrogram] tab (Figure 5).



```
A1   2    6
A2   1    3    7
A3   2    4    8
A4   3    5    9
A5   4    10
B1   1    7    11
B2   2    6    8    12
B3   3    7    9    13
B4   4    8    10   14
B5   5    9    15
C1   6    12   16
C2   7    11   13   17
C3   8    12   14   18
C4   9    13   15   19
C5   10   14   20
D1   11   17   21
D2   12   16   18   22
D3   13   17   19   23
D4   14   18   20   24
D5   15   19   25
E1   16   22
E2   17   21   23
E3   18   22   24
E4   19   23   25
E5   20   24
```

```
2
1
4
16
13
8
10
21
20
6
24
14
19
12
9
5
22
23
11
7
3
15
25
17
18
```

Figure 4: Contents of the neighborhood information file (`5by5nb.txt`; left) and the univariate file (`5by5h.txt`; right) for the 5-by-5 lattice data.



```
  Order Parent Maxval Minval Length Cells Progeny Family Level
1     1      5     25     22      6     3       0      1     3
2     1      7     24     24     10     1       0      1     1
3     1      5     21     20      2     2       0      1     3
4     1      6     18     18      1     1       0      1     2
5     2      6     19     19      2     1       2      3     2
6     2      7     17     15      3     3       2      5     1
7     2      0     14      1     13    14       2      7     0
```

Figure 5: Execution result of echelon analysis for the 5-by-5 lattice data.

### 3.3 Lung cancer data in New Mexico

As a real data analysis, we apply echelon analysis and the scanning technique to lung cancer data in New Mexico available on the SaTScan website(`https://www.satscan.org/datasets/nmlung/`). The data consist of the number of malignant lung cancer cases and the number of populations of 32 counties from 1973 to 1991. A total of 9,254 cancer cases and 25,604,291 populations were recorded during this period, and they are provided with separate multiple categorical covariates such as 18 age groups (1 = ages <5, 2 = ages 5–9, 3 = ages 10–14, …, 17 = ages 80–84, 18 = ages 85+) and sex (1 = male, 2 = female). To simplify the interpretation of output analysis, we aggregate the years into six time periods: 1st = 1973–1975, 2nd = 1976–1978, 3rd = 1979–1981, 4th = 1982–1984, 5th = 1985–1987 and 6th = 1988–1991. The cancer cases and the populations for each data point are denoted by $c_{itjk}$ and $n_{itjk}$, respectively (county: $i = 1, 2, \ldots, 32$; time period: $t = 1, 2, \ldots, 6$; age group: $j = 1, 2, \ldots, 18$; sex: $k = 1, 2$). With covariate adjustment, the expected number of cases in a county $(i, t)$ is calculated using age group $j$ and sex $k$.

$$\lambda_{it} = \sum_{j=1}^{18} \sum_{k=1}^{2} n_{itjk} P_{jk} \tag{1}$$

where $P_{jk} = \sum_{i=1}^{32} \sum_{t=1}^{6} c_{itjk} / \sum_{i=1}^{32} \sum_{t=1}^{6} n_{itjk}$ is the incidence of cancer. Similarly, the standardized mortality ratio (SMR), which is the most common health index of a county $(i, t)$ is given by

$$\mathrm{SMR}_{it} = \frac{c_{it}}{\lambda_{it}} = \frac{\sum_{j=1}^{18} \sum_{k=1}^{2} c_{itjk}}{\lambda_{it}} \tag{2}$$

Figure 6 shows the spatial distribution of SMR for each time period.

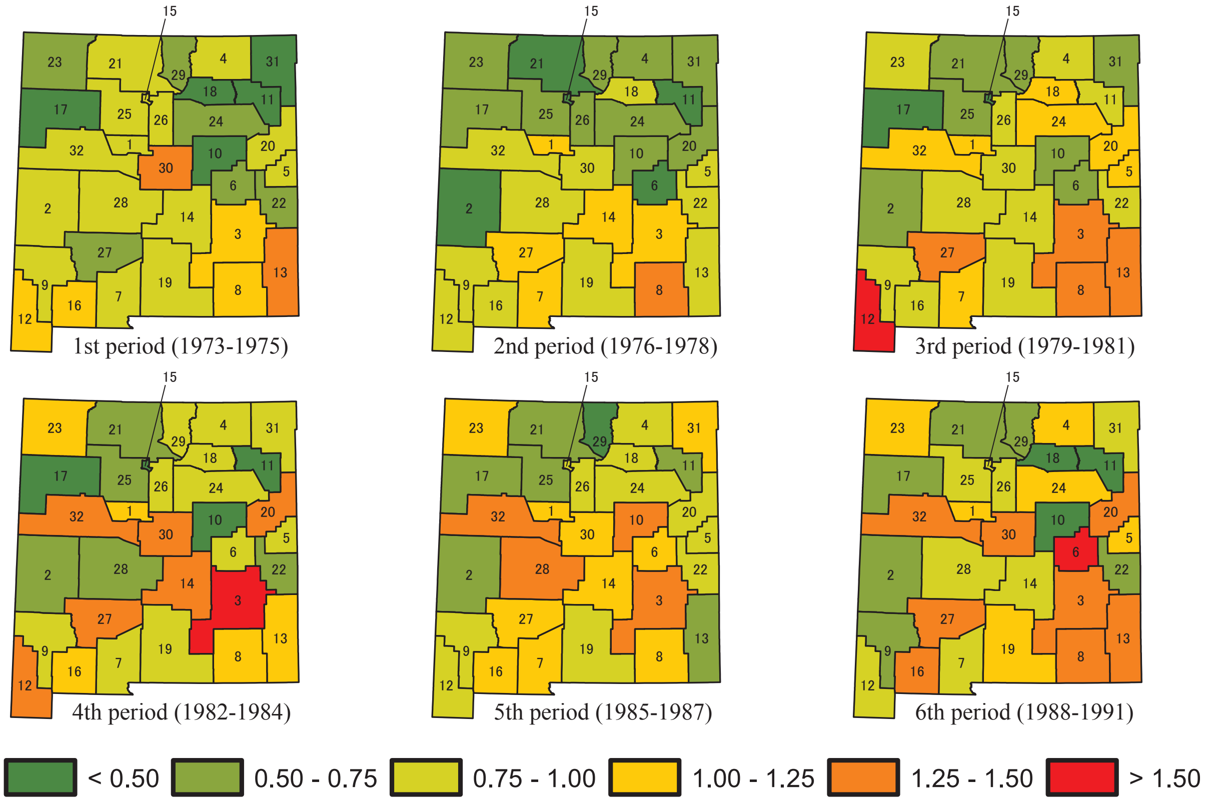

1st period (1973-1975)
2nd period (1976-1978)
3rd period (1979-1981)
4th period (1982-1984)
5th period (1985-1987)
6th period (1988-1991)

Figure 6: SMR geographical distribution for each time period and county numbers for each of the 32 counties of New Mexico.

The data are considered as spatio-temporal lattice data with 192 irregular lattices (32 counties × 6 time periods). Each lattice is denoted by $l_3(i, t)$ $(i = 1, 2, \ldots, 32; t = 1, 2, \ldots, 6)$. Here the simplest example of defining the neighbors of $l_3(i,t)$ is given as

$$NB(l_3(i,t)) = \begin{cases} \{l_3(k,t) | \text{county } i \text{ and } k \text{ are neighbor}\} \cup l_3(i,t+1), & t = 1 \\ \{l_3(k,t) | \text{county } i \text{ and } k \text{ are neighbor}\} \cup l_3(i,t+1) \cup l_3(i,t-1), & 1 < t < 6 \\ \{l_3(k,t) | \text{county } i \text{ and } k \text{ are neighbor}\} \cup l_3(i,t-1), & t = 6 \end{cases} \tag{3}$$

### Spatial scan statistics

Several models have identified areas with statistically significantly high values (hotspots) for a spatially distributed response variable. The spatial scan statistics (Kulldorff 1997) detect the disease cluster used in epidemiologic and disease surveillance studies. The spatial scan statistic detects the hotspots or clusters of lattices by use of scan window $Z$ on each centroid of aggregated lattices. Let the lattice in $Z_i$ be the connected subset of index set $D$ and satisfy $D = \cup_{i=1}^{M} Z_i$, $i = 1, 2, \ldots, M$. We define a family of a subset

$$\mathcal{Z} = \{Z_1, Z_2, \ldots, Z_M\}, \quad Z_i \subset D \tag{4}$$

as candidates of a hotspot. The hypothesis is threre is no hotspot under the null hypothesis $H_0$, versus alternative hypothesis $H_1$ where there is at least one window $Z$ of the hotspot. We find the window $\hat{Z}$ that maximize the likelihood function $L(Z)$. The test statistic is based on the likelihood ratio (LR),

$$LR(Z) = \frac{\text{the likelihood under } H_1}{\text{the likelihood under } H_0} = \frac{L(\hat{Z})}{L_0} \tag{5}$$

The logarithm of $LR(Z)$, (i.e, $LLR$) is for computational simplicity. We define a hotspot as a window $Z$ with a maximum value of $LLR$. Some probability models of spatial scan statistic have been proposed depending on the feature of data so far. For example, $LR$ based on Poisson model can be defined in

$$LR(Z) = \left(\frac{c(Z)}{\lambda(Z)}\right)^{c(Z)} \left(\frac{c - c(Z)}{c - \lambda(Z)}\right)^{c - c(Z)} I\left(\frac{c(Z)}{\lambda(Z)} > \frac{c - c(Z)}{c - \lambda(Z)}\right) \tag{6}$$

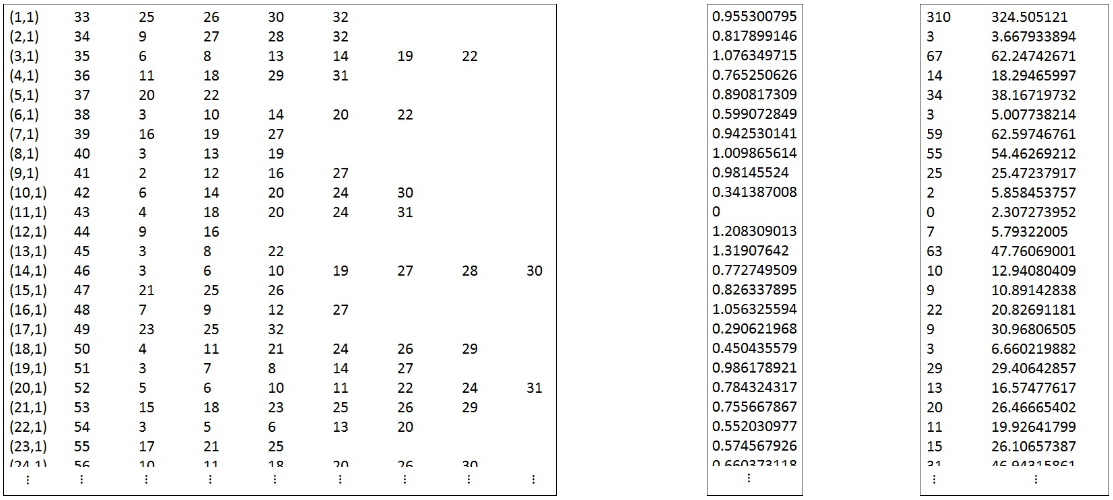
where $c(Z)$ and $\lambda(Z)$ denote a observed number of cases and an expected number of cases, respectively, within the specified window $Z$. $c$ is a total number of the observed number of cases. $I()$ is the indicator function. To evaluate the statistical significance of detected hotspots, Monte Carlo test under the null hypothesis is typically used to estimate p-values since it is difficult to obtain the exact distribution of the spatial scan statistic.

### Echelon scan technique

The scan window is an important research topic that satisfies several properties to detect a candidate of a hotspot. First, the window should comprise a geographically connected subset of the index set. Second, the window should be medium. This restriction is achieved by limiting the hotspot search to lattice that does not comprise more than 50% of the lattices. The difficult part of hotspot estimation lies in maximizing $LLR$ as $Z$ varies over the collection of all lattices in $\mathcal{Z}$. The family of subset $\mathcal{Z}$ is a finite set that has high computational costs to maximize $LLR$ through an exhaustive search. The traditional spatial scan statistic uses expanding circles to determine the candidate of window $Z$. The candidate window $Z$ may do a poor job of approximating actual hotspots. The echelon scan technique (Kurihara 2003, Ishioka and Kurihara 2012) has been proposed to reduce the size of the scanned window $Z$ and computational costs. This technique is performed to find the candidate of the hotspot by scanning from the peak .

### Space-time hotspot detection using EcheScan

We use the EcheScan for the detection of space-time hotspot of the lung cancer data in New Mexico. First, we prepare the neighborhood information file (`NMnb.txt`) and the univariate file (`NMsmr.txt`) for the echelon analysis, and also use the case & expectation file (`NMCasExp.txt`) to calculate the spatial scan statistics as shown in Figure 7.



```
(1,1)   33   25   26   30   32
(2,1)   34   9    27   28   32
(3,1)   35   6    8    13   14   19   22
(4,1)   36   11   18   29   31
(5,1)   37   20   22
(6,1)   38   3    10   14   20   22
(7,1)   39   16   19   27
(8,1)   40   3    13   19
(9,1)   41   2    12   16   27
(10,1)  42   6    14   20   24   30
(11,1)  43   4    18   20   24   31
(12,1)  44   9    16
(13,1)  45   3    8    22
(14,1)  46   3    6    10   19   27   28   30
(15,1)  47   21   25   26
(16,1)  48   7    9    12   27
(17,1)  49   23   25   32
(18,1)  50   4    11   21   24   26   29
(19,1)  51   3    7    8    14   27
(20,1)  52   5    6    10   11   22   24   31
(21,1)  53   15   18   23   25   26   29
(22,1)  54   3    5    6    13   20
(23,1)  55   17   21   25
(24,1)  56   10   11   18   20   26   30
  ⋮      ⋮    ⋮    ⋮    ⋮    ⋮    ⋮    ⋮    ⋮
```

```
0.955300795
0.817899146
1.076349715
0.765250626
0.890817309
0.599072849
0.942530141
1.009865614
0.98145524
0.341387008
0
1.208309013
1.31907642
0.772749509
0.826337895
1.056325594
0.290621968
0.450435579
0.986178921
0.784324317
0.755667867
0.552030977
0.574567926
0.660373118
     ⋮
```

```
310     324.505121
3       3.667933894
67      62.24742671
14      18.29465997
34      38.16719732
3       5.007738214
59      62.59746761
55      54.46269212
25      25.47237917
2       5.858453757
0       2.307273952
7       5.79322005
63      47.76069001
10      12.94080409
9       10.89142838
22      20.82691181
9       30.96806505
3       6.660219882
29      29.40642857
13      16.57477617
20      26.46665402
11      19.92641799
15      26.10657387
31      46.94315861
⋮           ⋮
```

Figure 7: Part of the contents of the neighborhood information file (`NMnb.txt`; left), univariate file (`NMsmr.txt`; center) and case & expectation file (`NMCasExp.txt`; right). These files consist of 192 lines. Each lattice label is given by "(county number, time period)". For example, "(1,1)" corresponds to the county 1 at the 1st period. The univariate file consists of SMR. In the case & expectation file, the two values of $c_{it}$ and $\lambda_{it}$ corresponding to each lattice (the each line) are described.

If the neighbor information and univariate data succeed in reading without error, "Case & Expectation" appears. By clicking on "Case & Expectation", we select the `NMCasExp.txt` for the case & expectation file. Next, we click on the [Echelon scan] tab and click on [Run] to execute the echelon scan technique. Figure 8 shows the execution result when setting the significance level = 0.05, the maximum hotspot size = 30, and the Monte Carlo replications = 999. By changing the settings of "Vertical range:" and "Horizontal range:" at the bottom of the screen, we can display part of the echelon dendrogram interactively. Figure 9 is an enlarged view of the echelons and the lattices recognized as the hotspot.

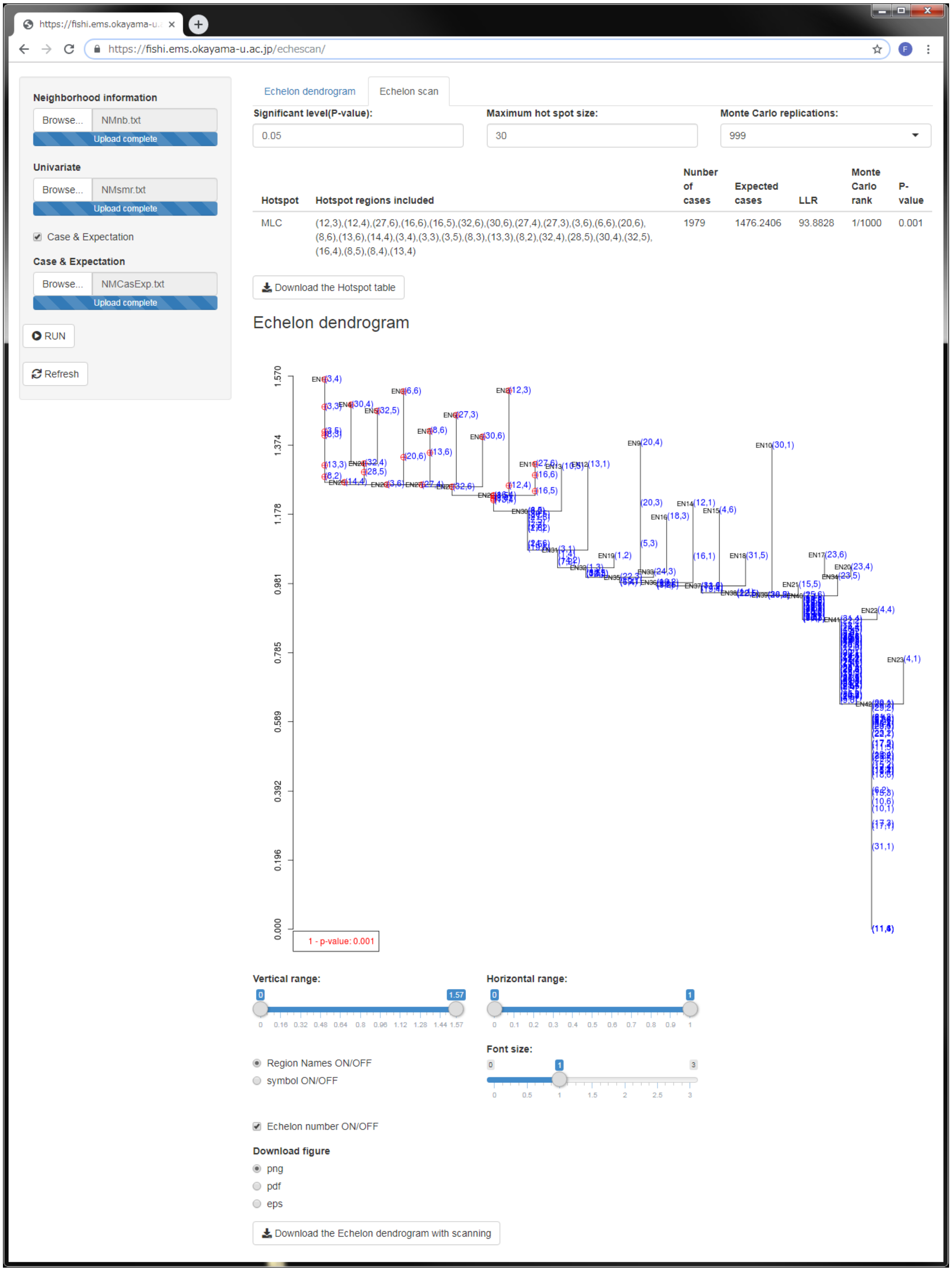
The detected hotspot cluster has $LLR = 93.883$ and $p = 0.001$, and we visualized it on the map using ArcMap software by Esri. As shown in Figure 10, the echelon scan technique revealed that the counties recognized as the hotspot cluster have a complex variation over time. There were no hotspot counties in the 1st period, but as time goes on, we can see that it spreads in the southeast beginning at the county 8. The hotspot counties in the southwest centered on the county 16 was detected based on the 4th period, and at this period the hotspot was covered a wide range from east to west. Furthermore, it can be seen that the hotspot was split east and west after the 5th period.

**Note**

• Each input file is allowed up to 5MB.

• In a 10 × 10 lattice data with randomly generated observed and expected values, the average execution time and its SD for hotspot detection of 50 trials were 12.72 seconds and 0.34 respectively when we set to the maximum hotspot size = 50% of the total lattices and 999 Monte Carlo replications (using a google chrome browser on a PC windows7 Intel(R), Core(TM)i7 CPU X990 (3.47GHz) and 24GB memory).

• Figure 6 and Figure 10 are drawn using Esri's arcmap software.



Neighborhood information

Browse... NMnb.txt

Upload complete

Univariate

Browse... NMsmr.txt

Upload complete

☑ Case & Expectation

Case & Expectation

Browse... NMCasExp.txt

Upload complete

RUN

Refresh

Echelon dendrogram | Echelon scan

Significant level(P-value): 0.05

Maximum hot spot size: 30

Monte Carlo replications: 999

| Hotspot | Hotspot regions included | Nunber of cases | Expected cases | LLR | Monte Carlo rank | P-value |
|---|---|---|---|---|---|---|
| MLC | (12,3),(12,4),(27,6),(16,6),(16,5),(32,6),(30,6),(27,4),(27,3),(3,6),(6,6),(20,6),(8,6),(13,6),(14,4),(3,4),(3,3),(3,5),(8,3),(13,3),(8,2),(32,4),(28,5),(30,4),(32,5),(16,4),(8,5),(8,4),(13,4) | 1979 | 1476.2406 | 93.8828 | 1/1000 | 0.001 |

Download the Hotspot table

Echelon dendrogram

Vertical range: 0 – 1.57

Horizontal range: 0 – 1

● Region Names ON/OFF

○ symbol ON/OFF

Font size: 1

☑ Echelon number ON/OFF

Download figure

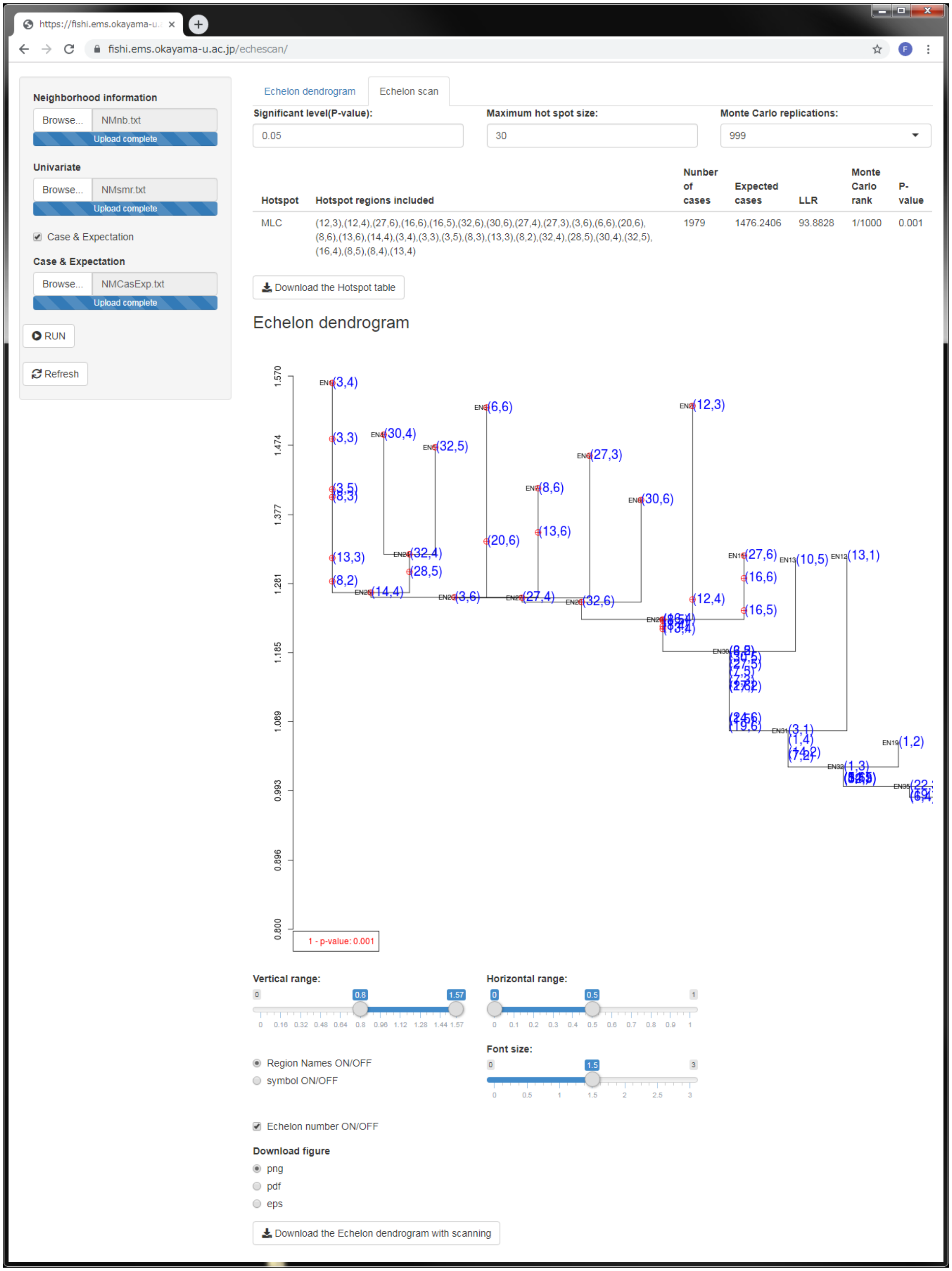
● png

○ pdf

○ eps

Download the Echelon dendrogram with scanning

Figure 8: Execution result of hotspot detection using the echelon scan technique for lung cancer data in New Mexico.



**Neighborhood information**

Browse... NMnb.txt

Upload complete

**Univariate**

Browse... NMsmr.txt

Upload complete

☑ Case & Expectation

**Case & Expectation**

Browse... NMCasExp.txt

Upload complete

RUN

Refresh

Echelon dendrogram | Echelon scan

**Significant level(P-value):** 0.05

**Maximum hot spot size:** 30

**Monte Carlo replications:** 999

| Hotspot | Hotspot regions included | Nunber of cases | Expected cases | LLR | Monte Carlo rank | P-value |
|---|---|---|---|---|---|---|
| MLC | (12,3),(12,4),(27,6),(16,6),(16,5),(32,6),(30,6),(27,4),(27,3),(3,6),(6,6),(20,6),(8,6),(13,6),(14,4),(3,4),(3,3),(3,5),(8,3),(13,3),(8,2),(32,4),(28,5),(30,4),(32,5),(16,4),(8,5),(8,4),(13,4) | 1979 | 1476.2406 | 93.8828 | 1/1000 | 0.001 |

Download the Hotspot table

## Echelon dendrogram

**Vertical range:** 0.8 – 1.57

**Horizontal range:** 0 – 0.5

**Font size:** 1.5

● Region Names ON/OFF

○ symbol ON/OFF

☑ Echelon number ON/OFF

**Download figure**

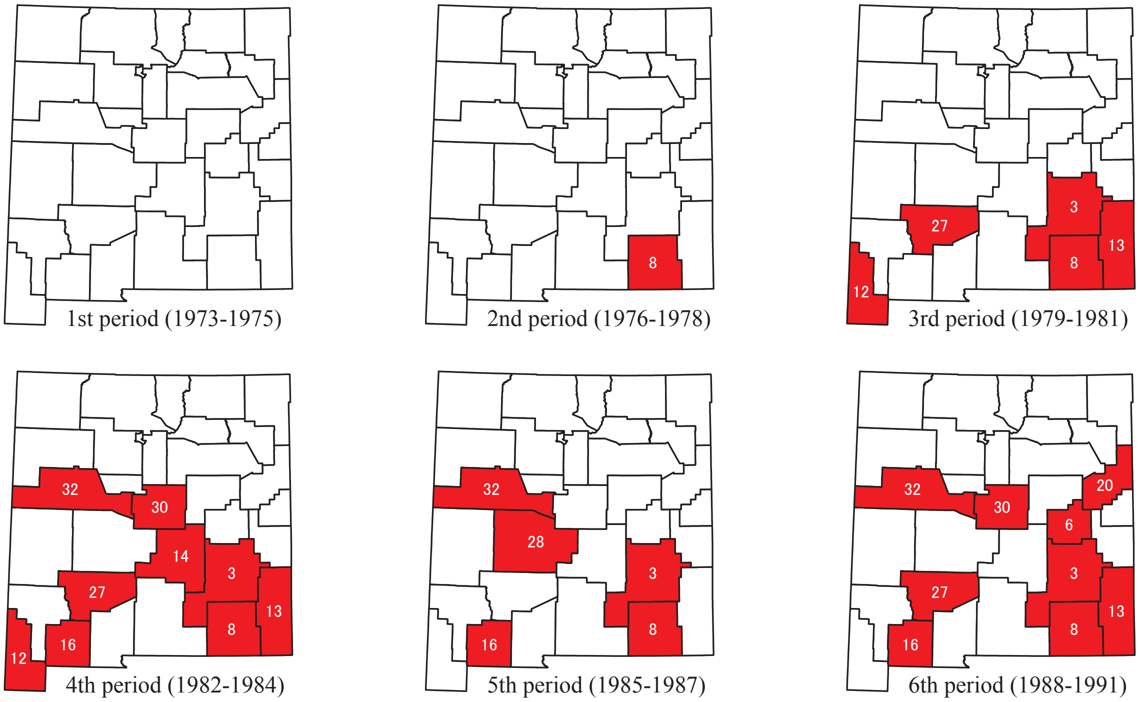
● png

○ pdf

○ eps

Download the Echelon dendrogram with scanning

Figure 9: Enlarged view of the echelons in which the hotspot was recognized.



1st period (1973-1975)

2nd period (1976-1978)

3rd period (1979-1981)

4th period (1982-1984)

5th period (1985-1987)

6th period (1988-1991)

Figure 10: Detected hotspot cluster of malignant lung cancer data in New Mexico. County numbers of the hotspot are indicated. 1st period: no counties are detected; 2nd period: 1 county; 3rd period: 5 counties; 4th period: 9 counties; 5th period: 5 counties; 6th period: 9 counties.

# References

Ishioka, F., Kawahara, J., Mizuta, M., Minato, S., & Kurihara, K. (2019). Evaluation of hotspot cluster detection using spatial scan statistic based on exact counting *Japanese Journal of Statistics and Data Science, 2(1)*, 241–262. DOI: 10.1007/s42081-018-0030-6

Ishioka, F., & Kurihara, K. (2012). Detection of spatial clustering using echelon scan. *Proceedings of the 20th International Conference on Computational Statistics (COMPSTAT2012) (Edited by Colubi, A. et al.), Heidelberg : Physica-Verlag*, 341–352.

Kulldorff, M. (1997). A spatial scan statistic. *Communications in Statistics: Theory and Methods, 26(6)*, 1481–1496. DOI: 10.1080/03610929708831995

Kurihara, K., Myers, W. L., & Patil, G. P. (2000). Echelon analysis of the relationship between population and land cover patterns based on remote sensing data. *Community Ecology, 1*, 103–122. DOI: 10.1556/ComEc.1.2000.1.14

Kurihara, K. (2003) The detection of hotspots based on the hierarchical spatial structure. *Bulletin of the Computational Statistics of Japan, 15(2)*, 171–183. DOI: 10.20551/jscswabun.15.2_171

Myers, W. M., Patil, G. P., & Joly, K. (1997). Echelon approach to areas of concern in synoptic regional monitoring. *Environmental and Ecological Statistics, 4*, 131–152. DOI: 10.1023/A:1018518327329