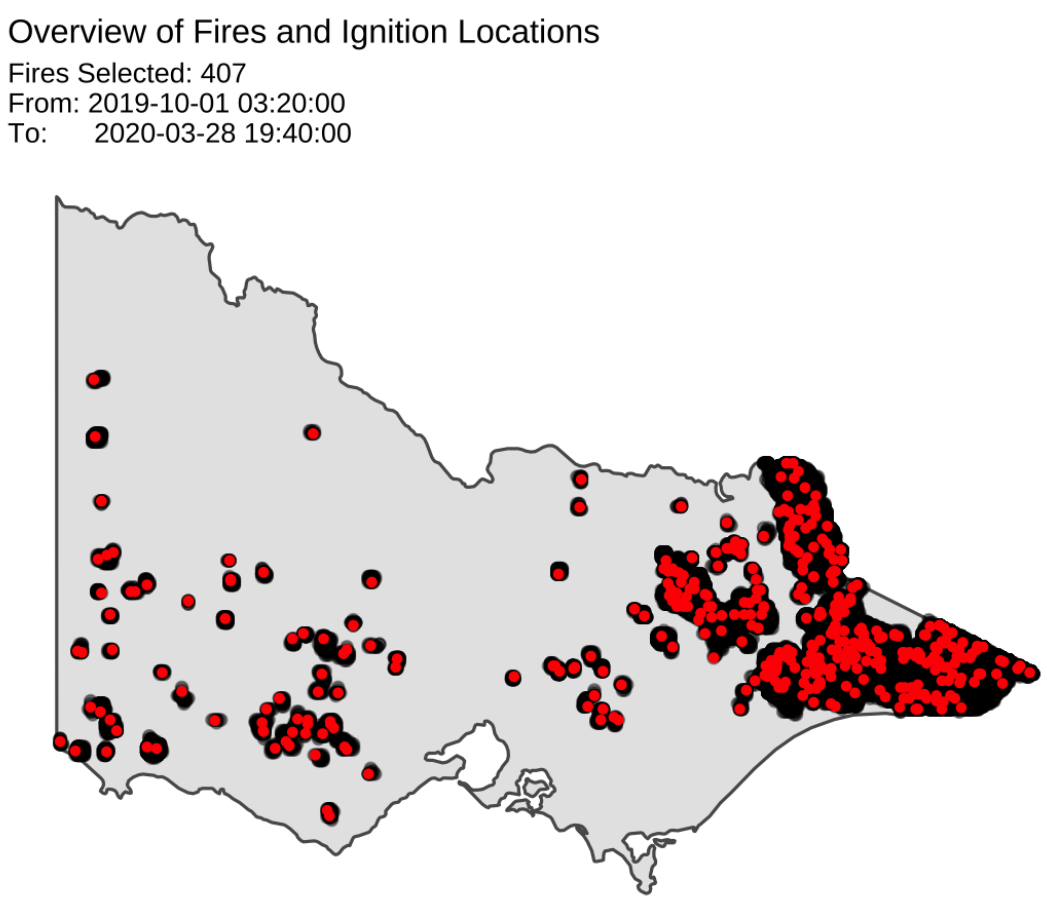
spotoroo: spatiotemporal clustering in R of hotspot data

Di Cook award presentation



Weihao (Patrick) Li

Department of Econometrics and Business Statistics, Monash University

weihao.li@monash.edu

May 17, 2022

2019-2020 Australian bushfire season

A catastrophic and unprecedented disaster.

By the end of 2020,

- almost 19 million hectares of land burned
- over 3000 homes destroyed
- AUD $ 1.7 billion in insurance losses
- an estimated 1 billion animals killed

Fig. 1: *An out of control bushfire in Werombi, NSW (Wikimedia Commons, 2019).*

Remote sensing data

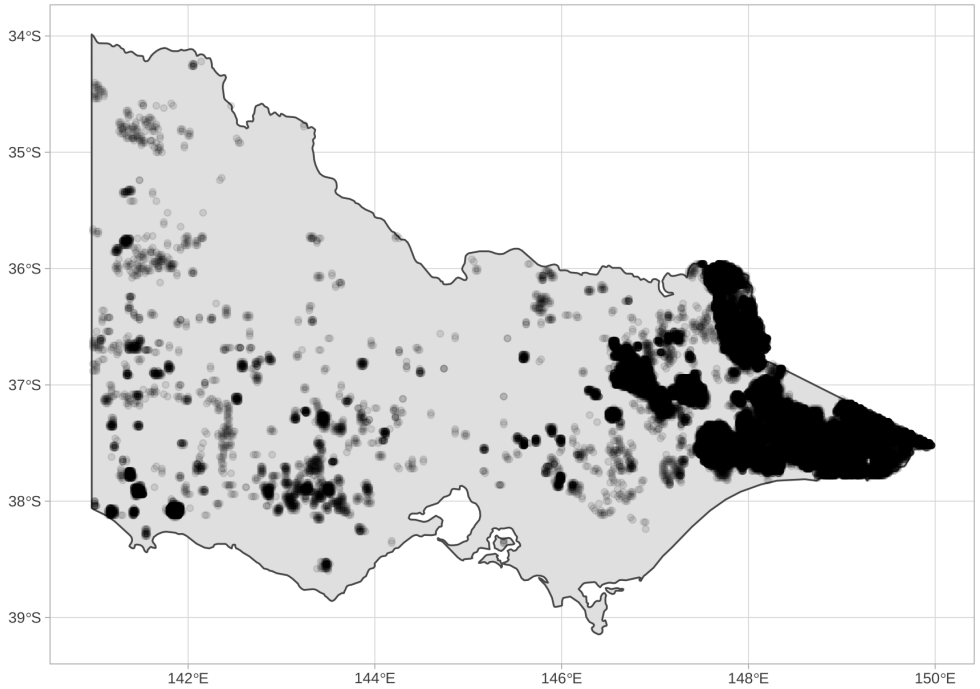


Fig. 2: *Himawari-8 hotspot data in Victoria from October 2019 to March 2020.*

Remote satellite data provides a potential solution to the challenge of active fire detection and monitoring.

Japan Aerospace Exploration Agency (JAXA) Himawari-8 satellite wildfire product:

- 1989572 hotspots in Australia from October 2019 to March 2020
- $0.02°$ (~ 2 kms) spatial grid resolution
- 10 minutes temporal resolution

👨‍💻 Authors of spotoroo 👩‍💻

The clustering algorithm was developed in 2019, and made available in the `spotoroo` package in March, 2021.

```
install.packages("spotoroo")
library(spotoroo)
```

- Author, maintainer: Weihao (Patrick) Li
- Contributor: Prof. Di Cook
    - Professor of Business Analytics at Monash University.
- Contributor: Emily Dodwell
    - Principal Inventive Scientist at AT&T.

1️⃣ Divide hotspots into intervals

`activeTime`: the maximum amount of time a fire may stay smoldering but undetectable by satellite before flaring up again

$$
\begin{aligned}
S_1 &= \text{hotspots observed from hour 0 to hour 1} \\
S_2 &= \text{hotspots observed from hour 0 to hour 2} \\
&\vdots \\
S_{activeTime} &= \text{hotspots observed from hour 0 to hour } activeTime \\
S_{activeTime+1} &= \text{hotspots observed from hour 1 to hour } activeTime + 1 \\
&\vdots \\
S_T &= \text{hotspots observed from hour } T - activeTime \text{ to hour } T
\end{aligned}
$$

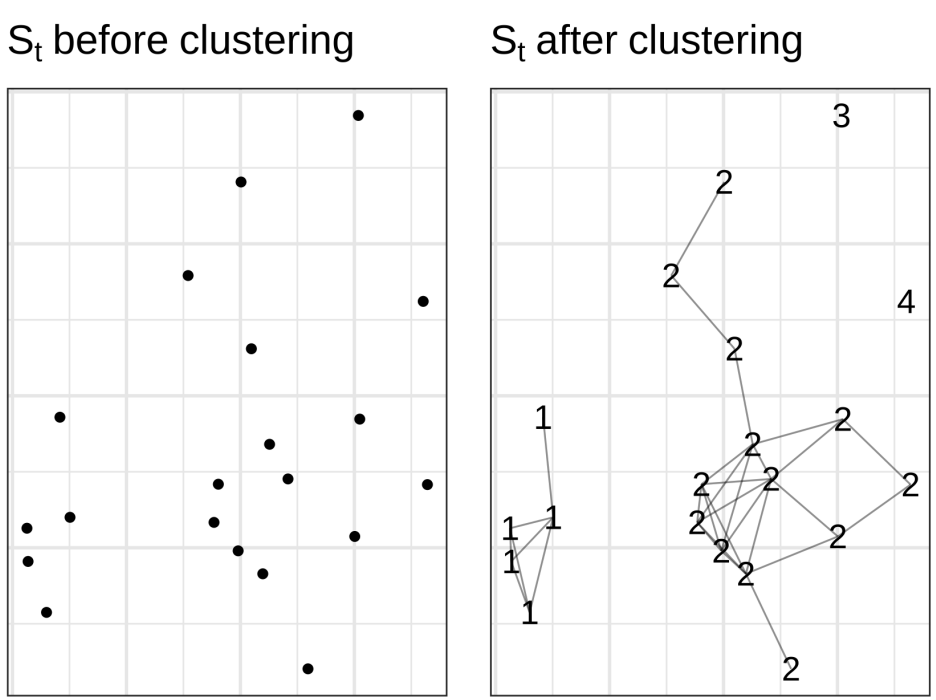
In summary, data will be divided into T intervals

$$S_t = [max(1, t - activeTime), t] \quad t = 1, 2, \ldots, T.$$

2️⃣ Cluster hotspots spatially within each time interval

For each time interval S_t, connect all hotspots within `adjDist` to form a graph

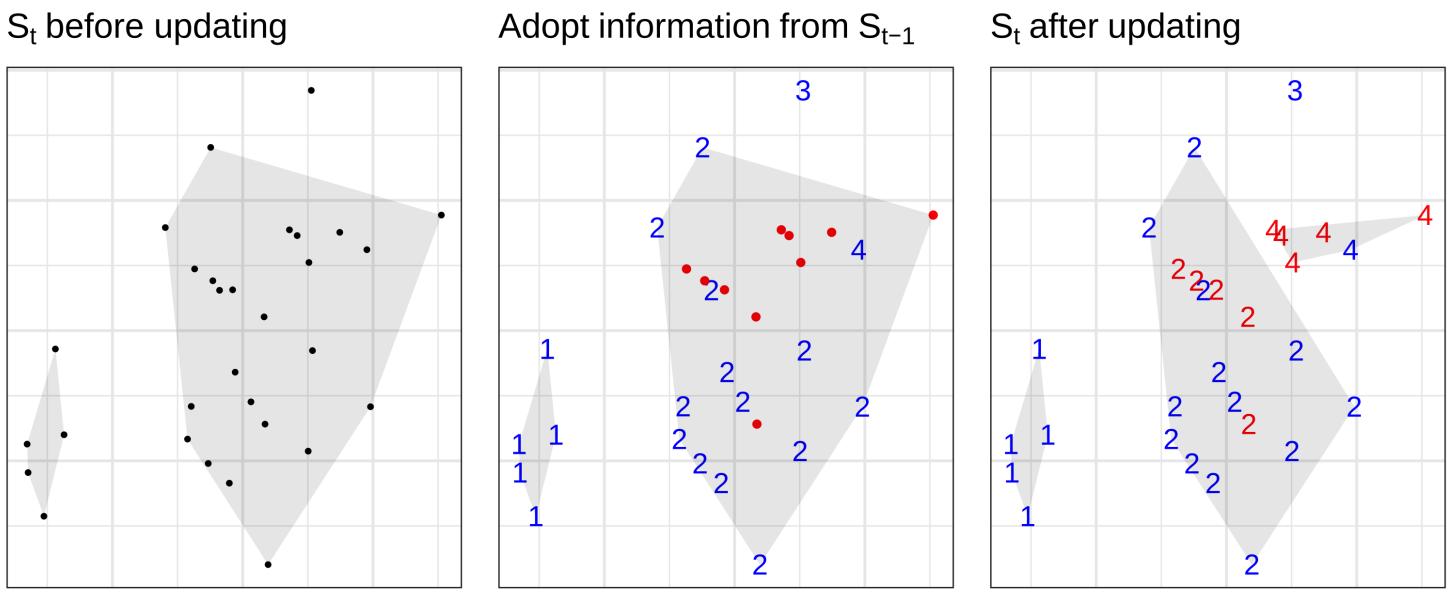
Every connected component of the graph is an individual cluster



3 Update memberships for hotspots

Some hotspots in S_t have been clustered in S_{t-1}

Update the membership by finding the nearest label



4️⃣ Handle noise in the clustering result

- `minPts`: the minimum number of hotspots in a cluster
- `minTime`: the minimum time a cluster can exist and still be considered a bushfire

Any cluster that does not satisfy these two conditions will be reassigned membership label -1 to indicate they represent noise.

Usage

```
result <- hotspot_cluster(hotspots,
                          activeTime = 24,
                          adjDist = 3000,
                          minPts = 4,
                          minTime = 3,
                          timeUnit = "h",
                          timeStep = 1)
```

```
──────────────────────────────────── SPOTOROO 0.1.2 ────────────────────────────────────

── Calling Core Function : `hotspot_cluster()` ──

── "1" time index = 1 hour
✓ Transform observed time → time indexes
ℹ 4313 time indexes found

── activeTime = 24 time indexes | adjDist = 3000 meters
✓ Cluster
ℹ 1055 clusters found (including noise)

── minPts = 4 hot spots | minTime = 3 time indexes
✓ Handle noise
```

📖 Summary of the clustering result

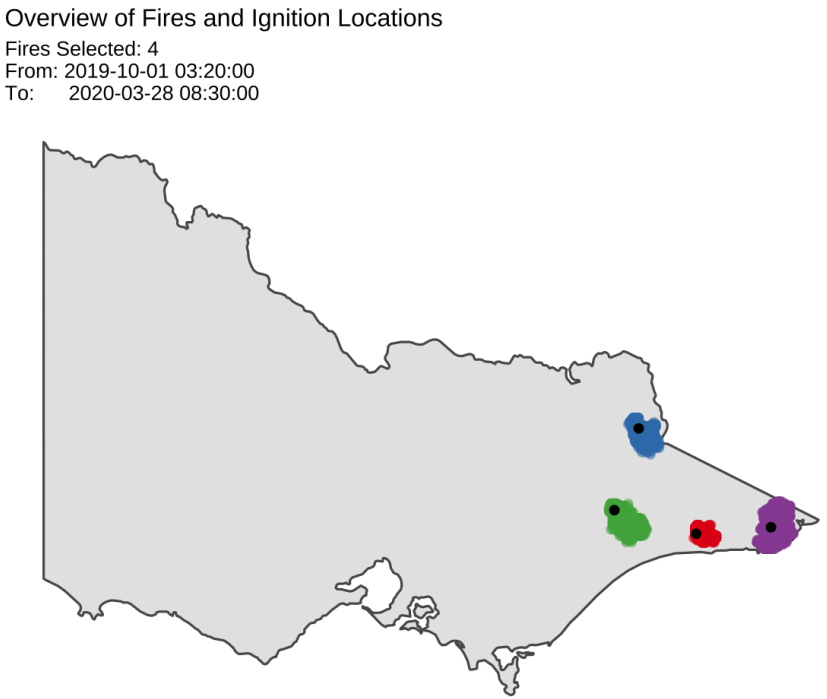
```
summary(result)
```

```
──────────────────────────────────────── SPOTOROO 0.1.2 ────────────────────────────────────────

── Calling Core Function : `summary_spotoroo()` ──

CLUSTERS: ALL
OBSERVATIONS: 75936
FROM: 2019-10-01 03:20:00
TO:   2020-03-28 19:40:00


── Clusters
ℹ Number of clusters: 407

Observations in cluster
        Min.      1st Qu.         Mean      3rd Qu.         Max.
         4.0         17.0        178.5        159.0      3863.0
Duration of cluster (hours)
        Min.      1st Qu.         Mean      3rd Qu.         Max.
         2.3          9.7         43.4         63.4        285.5
```

Visualizing the clustering result

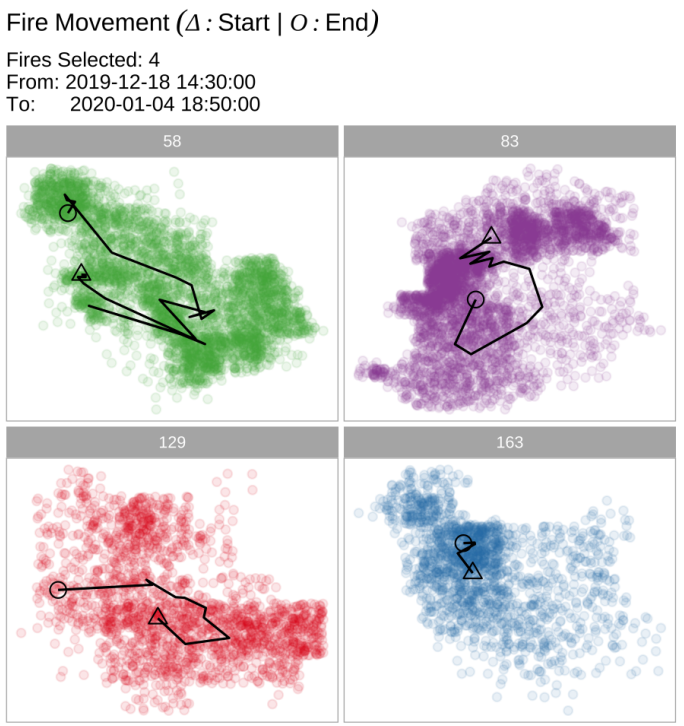
Spatial distribution of clusters

```
plot(result, bg = plot_vic_map(),
     cluster = c(58, 83, 129, 163))
```



Fire movement plot

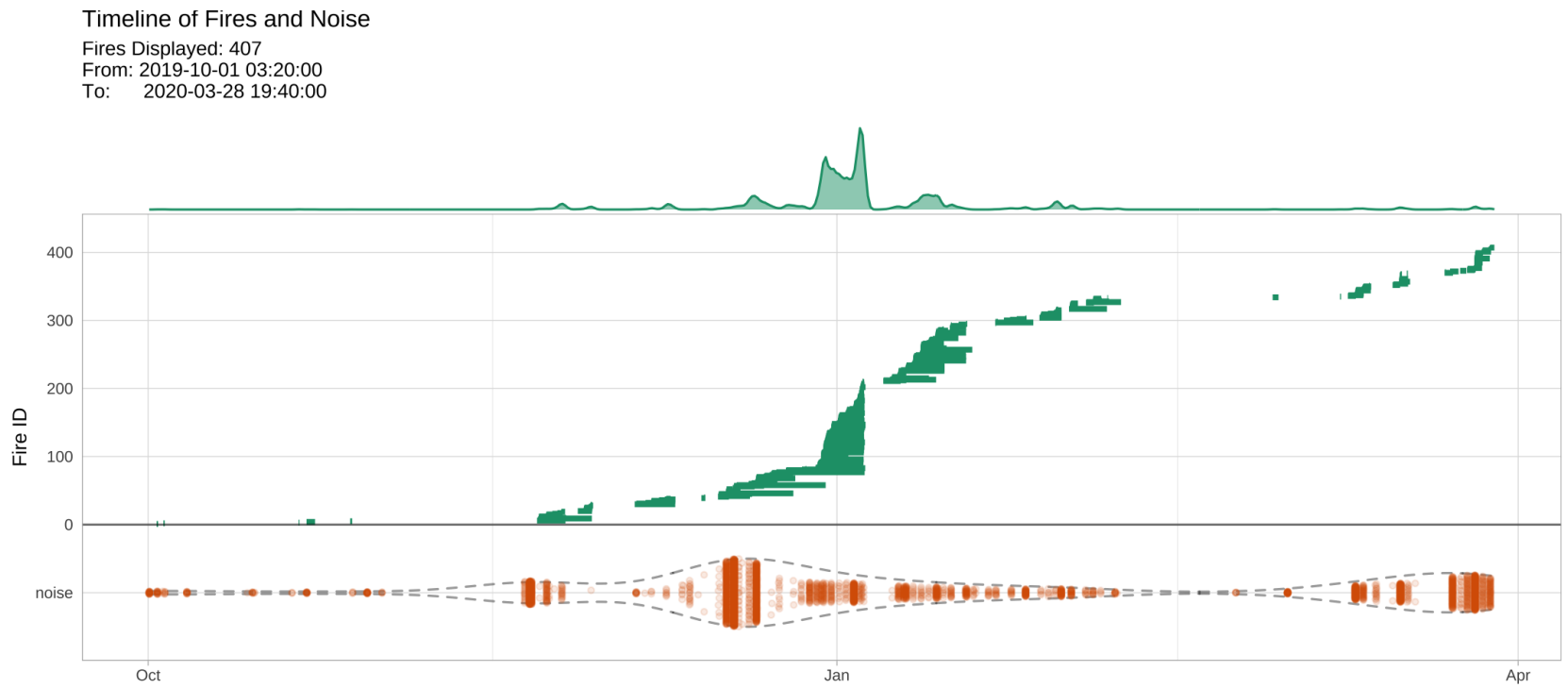
```
plot(result, type = "mov", step = 12,
     cluster = c(58, 83, 129, 163))
```



Visualizing the clustering result

Timeline plot

```
plot(result, type = "timeline")
```



Thanks!

Bibliography

Wikimedia Commons (2019). "Werombi Bushfire". https://commons.wikimedia.org/wiki/File:Werombi_Bushfire.jpg

Li, W., Dodwell, E., & Cook, D. (2021). A Clustering Algorithm to Organize Satellite Hotspot Data for the Purpose of Tracking Bushfires Remotely. https://github.com/TengMCing/Hotspots-Clustering-Algorithm/blob/master/li-dodwell-cook/RJwrapper.pdf