Rainfall Cluster Trend and Frequency Analysis

Alaa Ali, Stephanya Salazar, Carolina Maran, Nicole Cortez, Walter Wilcox

Presented by

Jagath Vithanage, PhD
Senior Scientist

Water Resources & Systems Modeling Bureau
South Florida Water Management District

aali@sfwmd.gov
jvithana@sfwmd.gov

sfwmd.gov

- SFWMD fourteen rainfall areas are delineated to facilitate operations and associated hydrologic components within them.
- Frequency and trend analyses conducted for rainfall within these areas may not characterize local rainfall behavior under evolving conditions across the region.

- Of interest is to explore frequency and trend behaviors of rainfall grouped by natural areas derived by similar rainfall behaviors.
- Study conducted across the SFWMD domain using rainfall from 1935–2024.
- Extreme rainfall thresholds: 4.00", 5.00", 6.00".
- Spatial clustering using DBSCAN identifies contiguous rainfall hotspots.

DBSCAN Clustering Methodology

- DBSCAN (Density-Based Spatial Clustering of Applications with Noise) identifies clusters based on spatial proximity and density.
- Key Parameters:
  - MinExceedance, minimum exceedance points along time axis
  - MinPts, minimum number of grid cells meeting MinExceedance condition to forma a cluster.
  - ε (epsilon): Maximum radius of the neighborhood around a point.
  - Core Point is a point that has minimum MinPts and MinExceedance points within span *epsilon*
  - Border Point: Is not a Core point but is within epsilon distance to a Core Point
  - All other points are noise

DBSCAN Clustering Methodology

Procedure:

1. All rainfall grid cells exceeding a threshold (e.g., 4.00") are extracted.
2. Grid cells with Minimum exceedance ((MinExceed) points are retained. (aka Exceedance Cells)
3. DBSCAN groups spatially adjacent exceedance cells into contiguous rainfall clusters.
4. Each cluster is assigned a unique ID and analyzed separately.
5. -1 ID is assigned to all points that did not fit in any of the clusters (noise)

DBSCAN Clustering Methodology

Strengths:

- Captures irregular cluster shapes (not constrained to fixed windows).
- Identifies both persistent and emerging hotspots.
- Robust to noise and non-clustered outliers.

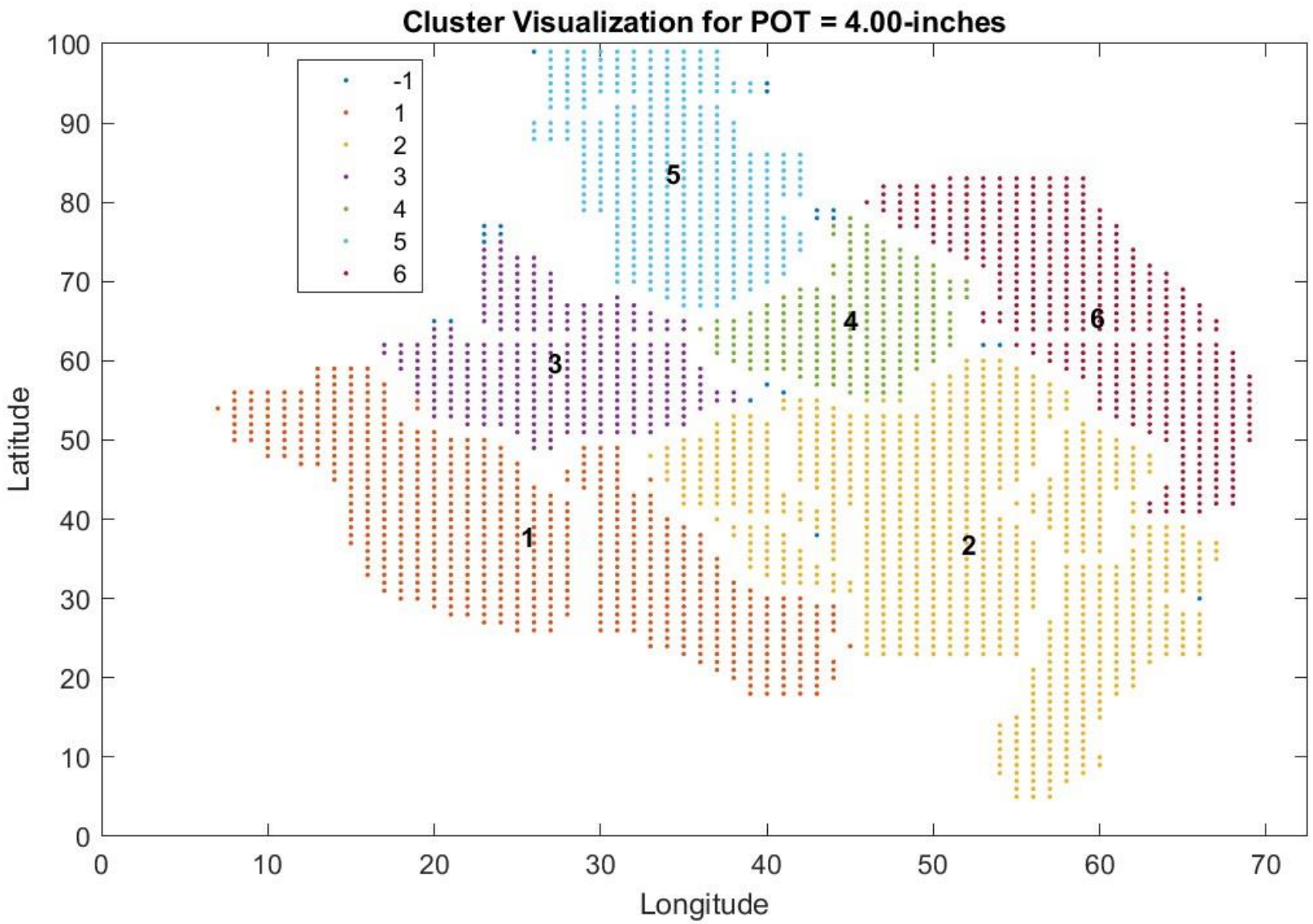
Limitations:

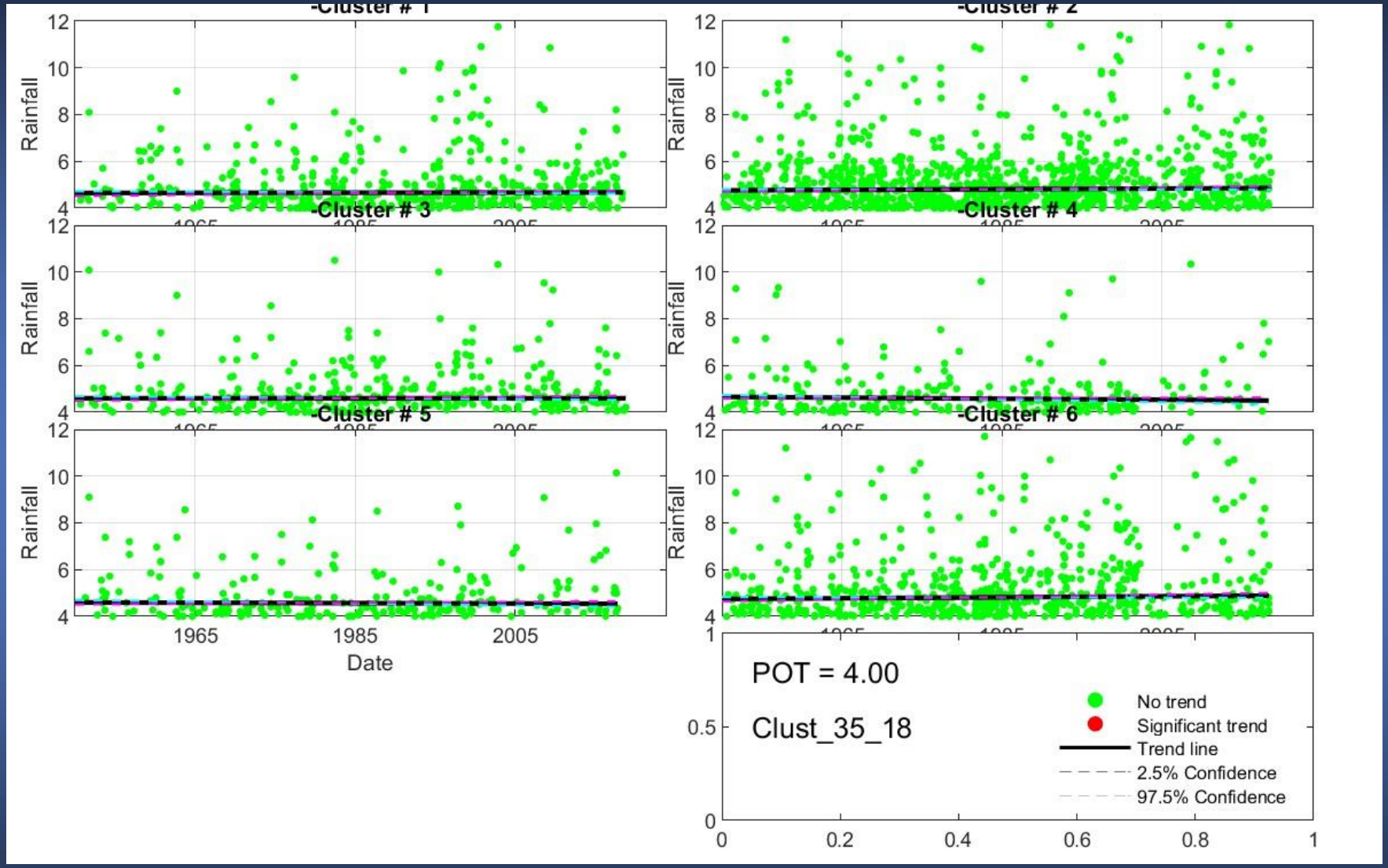
- Cluster definitions vary with parameter choice (ε, minPts, MinExceedance).
- Sensitive to spatial resolution exceedance points scarcity

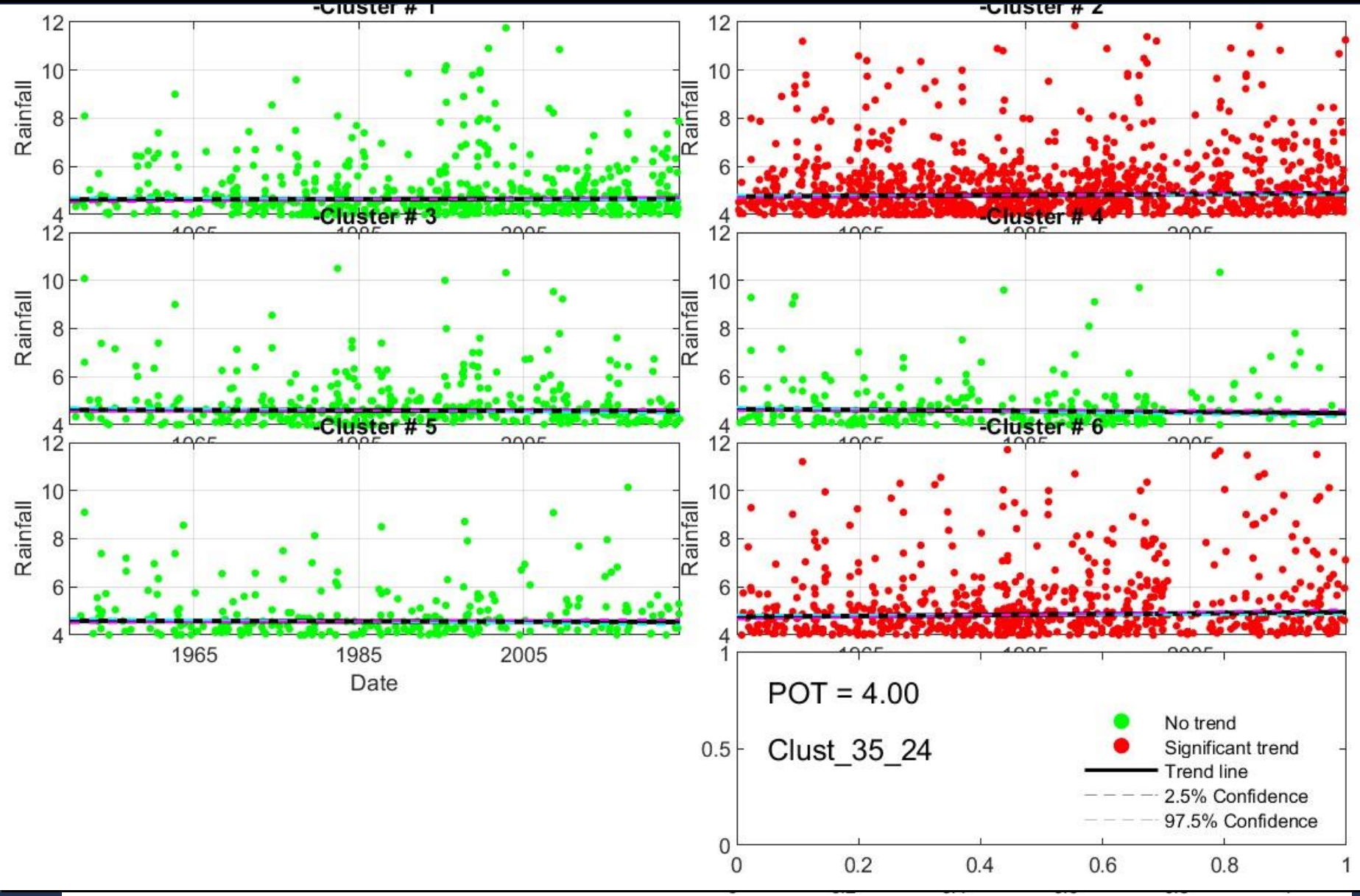
Result:

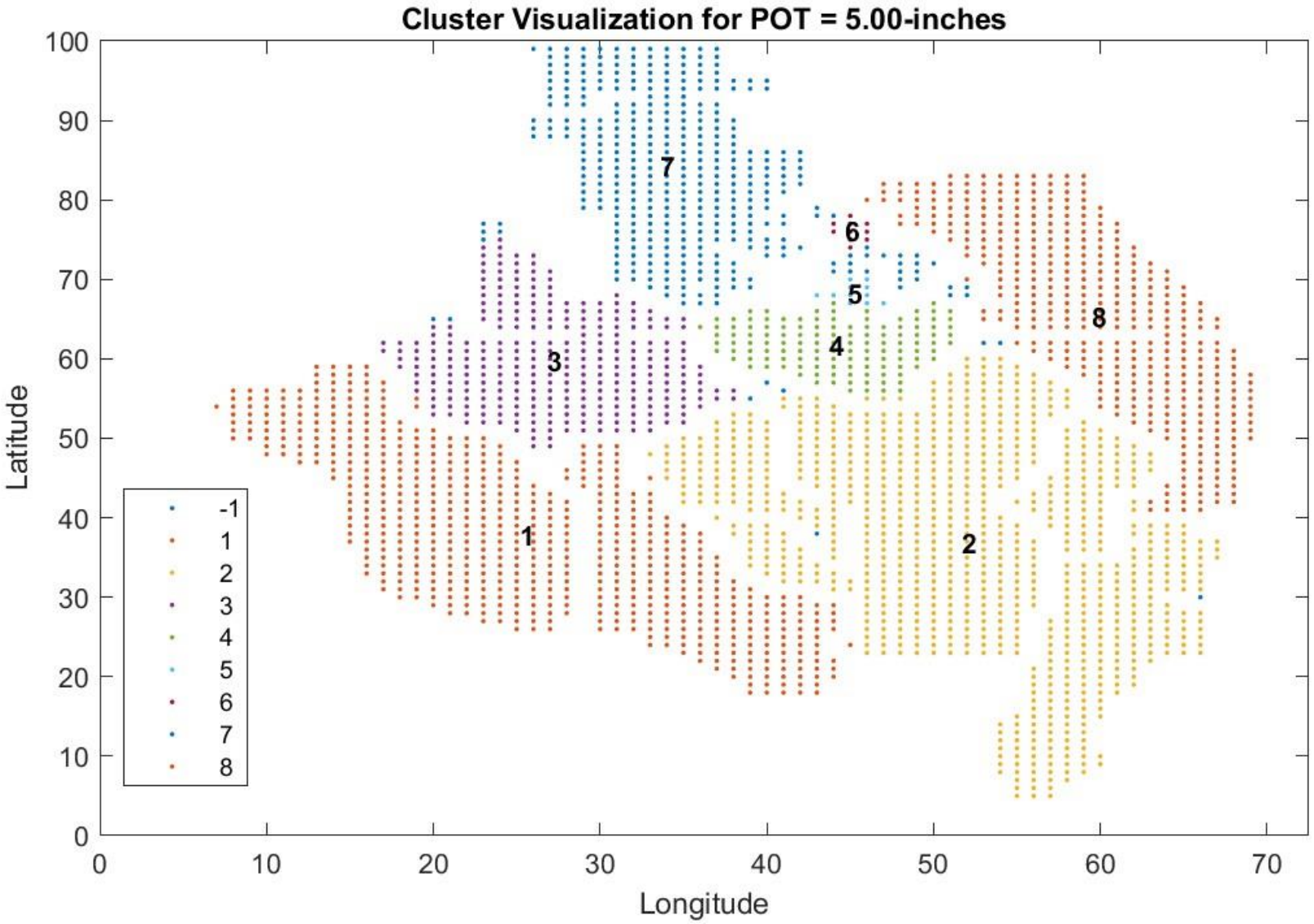
Each DBSCAN cluster generates a time series of exceedance counts, maxima, and mean intensity used in further trend/frequency analysis.

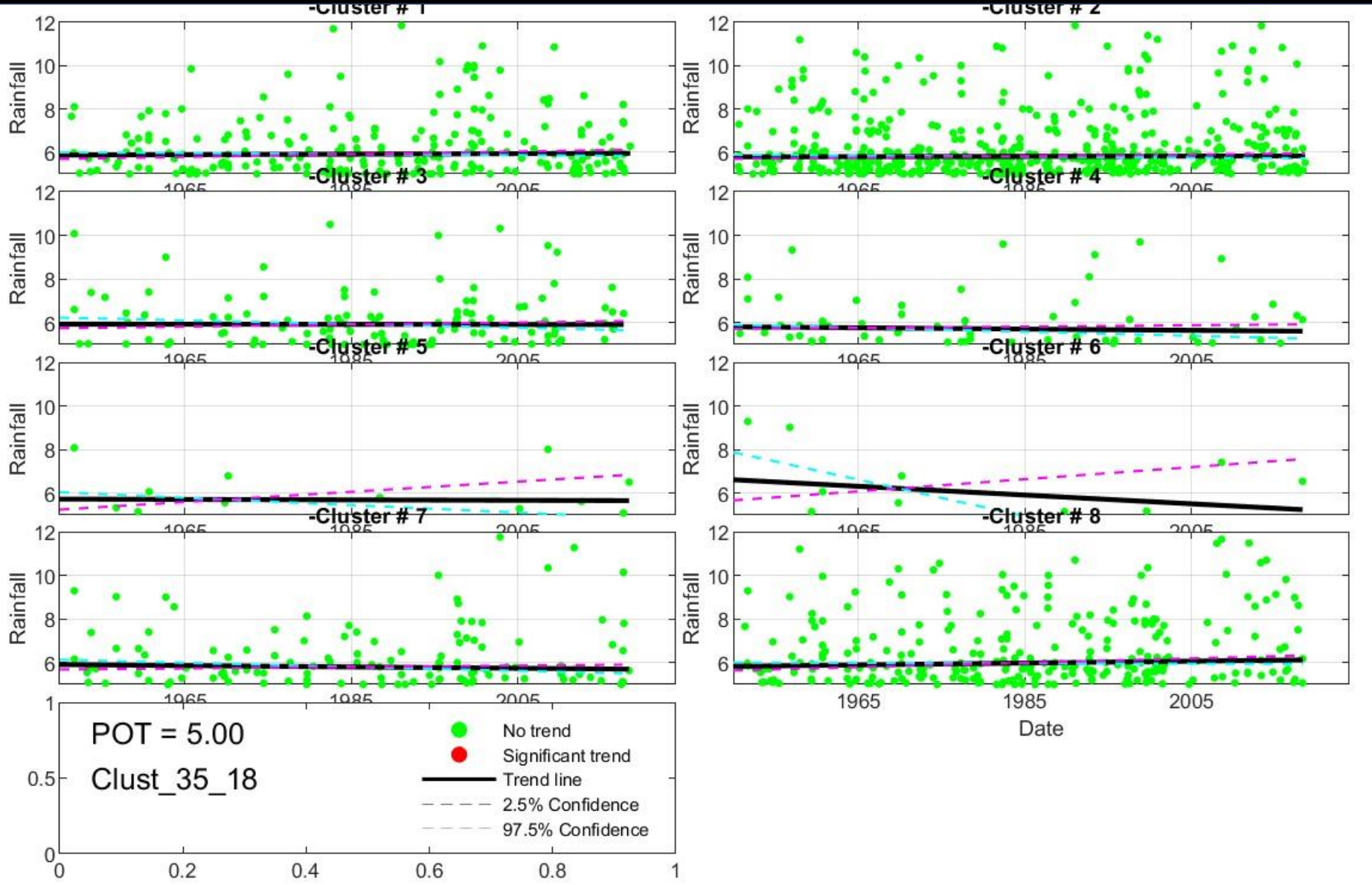
- Daily rainfall data aggregated into threshold exceedances.
- DBSCAN parameters: epsilon (spatial radius), minExceedance (Exceedance cells) and minPts (minimum cells).
- For each cluster: daily POT rainfall time series was constructed.
- minExceedance, minPts, epsilon were determined through sensitivity analysis as 5 points, 10 points and 4 miles respectively

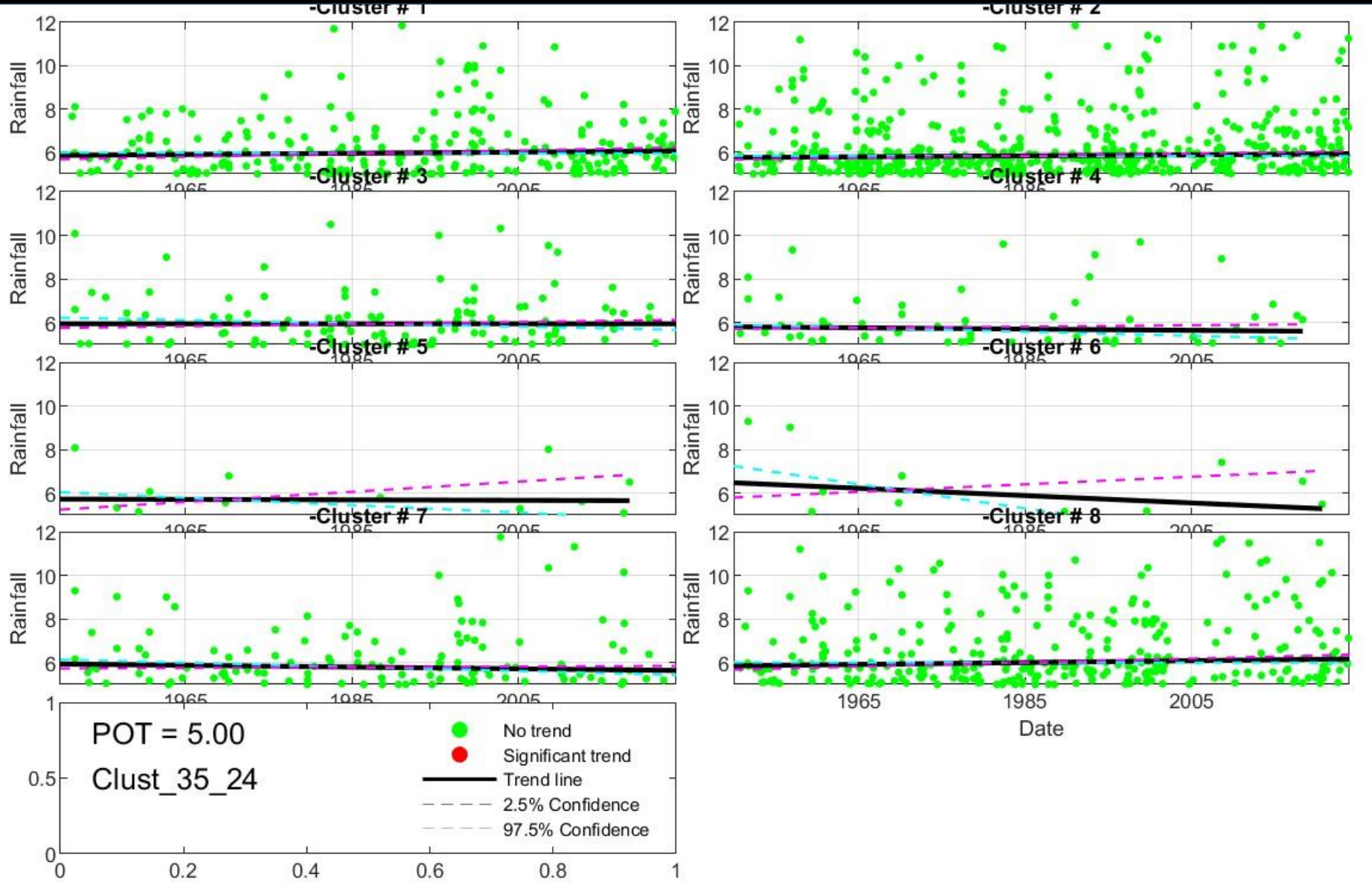

Cluster Visualization for POT = 4.00-inches
-1
1
2
3
4
5
6
Latitude
Longitude
0
10
20
30
40
50
60
70
80
90
100
1
2
3
4
5
6

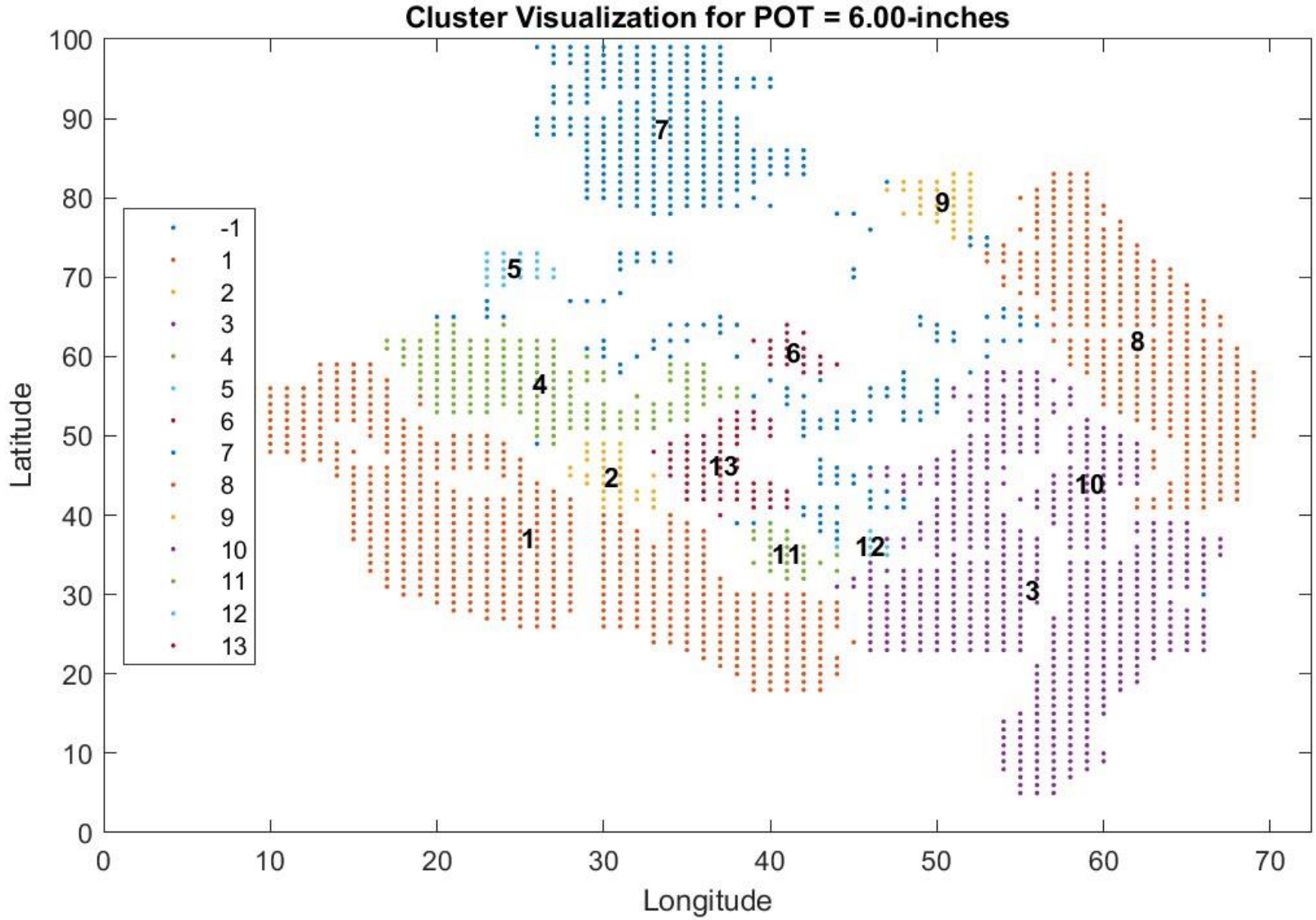

-Cluster # 1
-Cluster # 2
-Cluster # 3
-Cluster # 4
-Cluster # 5
-Cluster # 6
Rainfall
Date
1965
1985
2005
POT = 4.00
Clust_35_18
No trend
Significant trend
Trend line
2.5% Confidence
97.5% Confidence

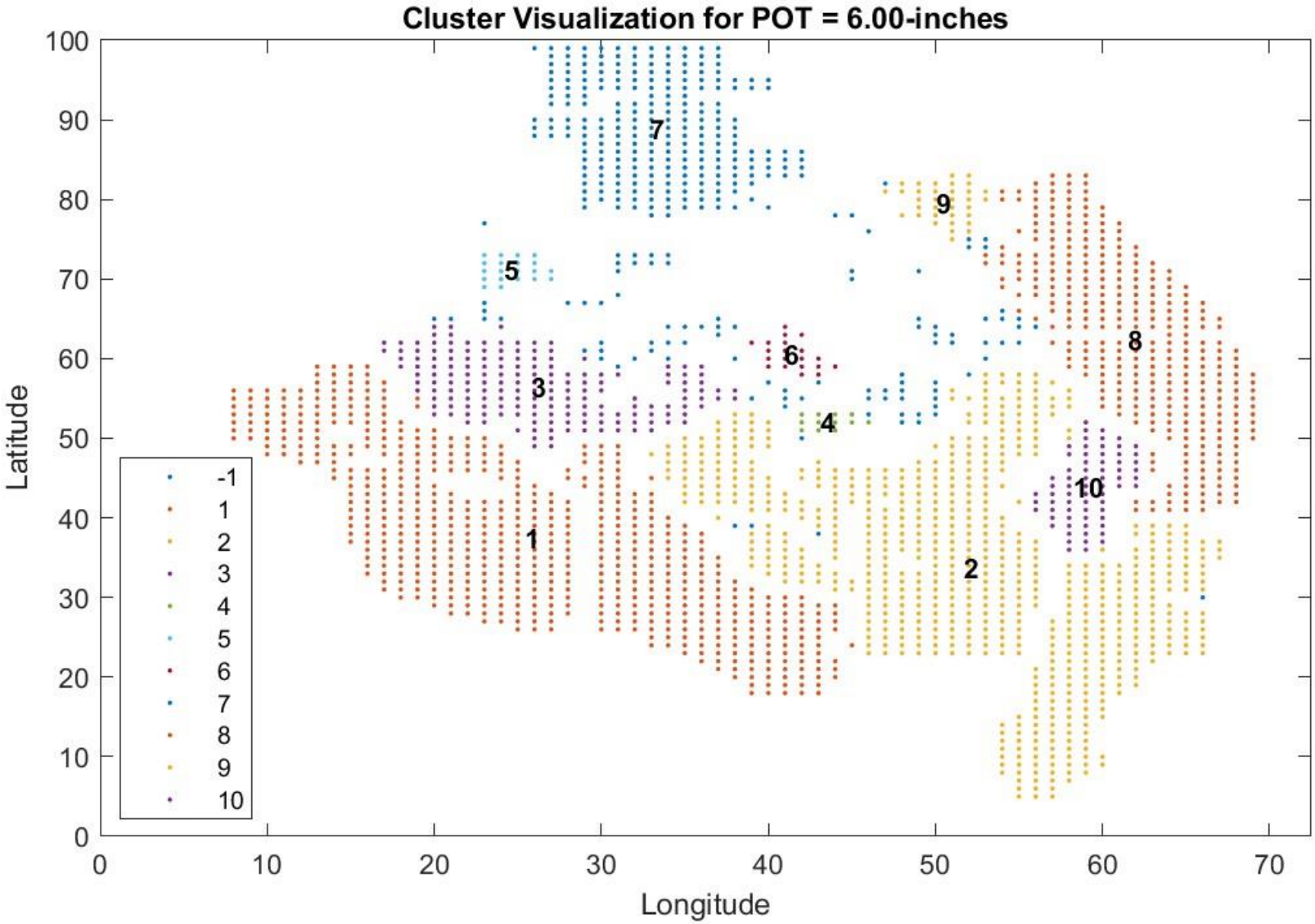

-Cluster # 1
-Cluster # 2
-Cluster # 3
-Cluster # 4
-Cluster # 5
-Cluster # 6
Rainfall
Date
1965
1985
2005
POT = 4.00
Clust_35_24
No trend
Significant trend
Trend line
2.5% Confidence
97.5% Confidence

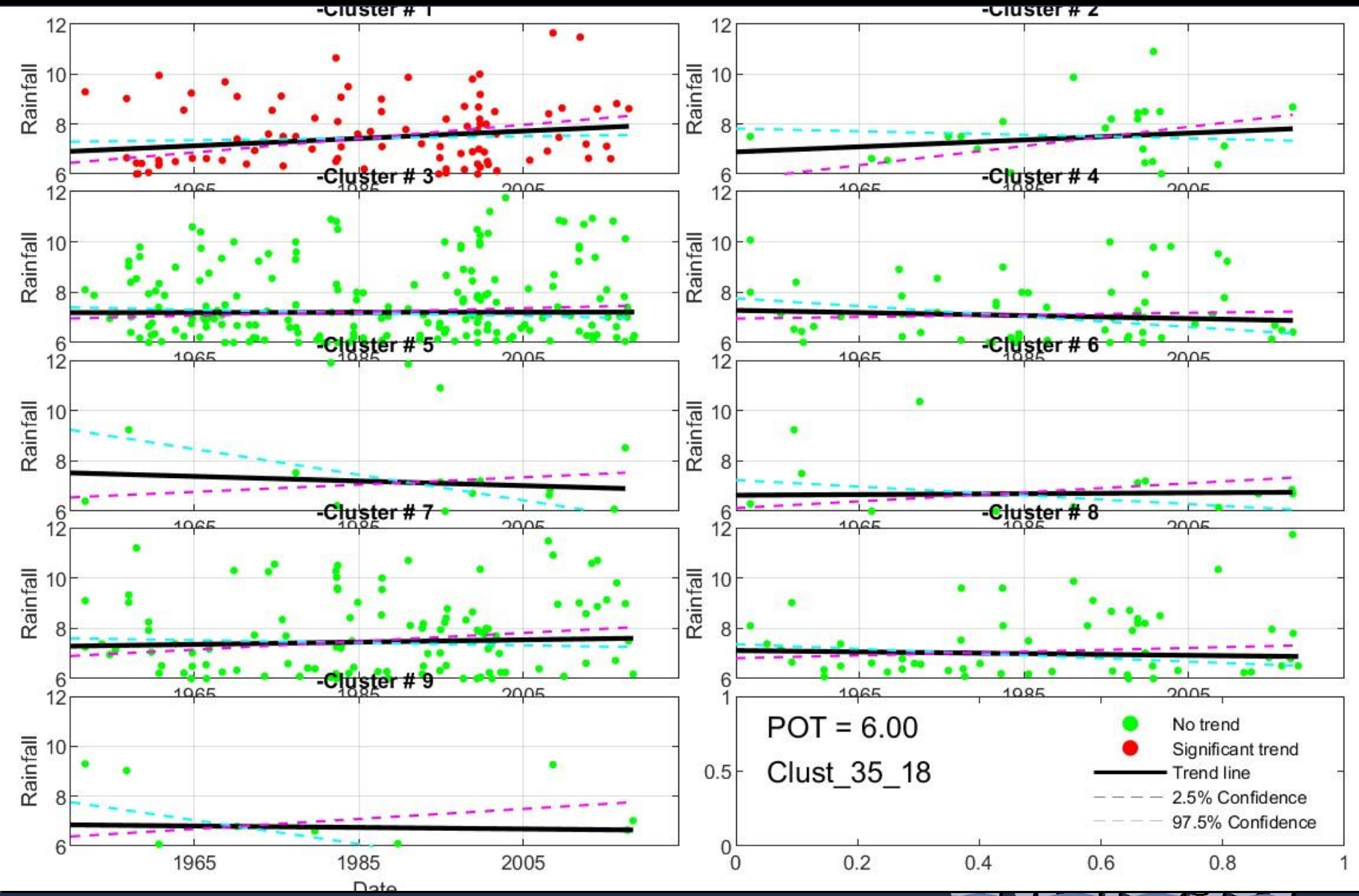

Cluster Visualization for POT = 5.00-inches
Latitude
Longitude
0
10
20
30
40
50
60
70
80
90
100
-1
1
2
3
4
5
6
7
8

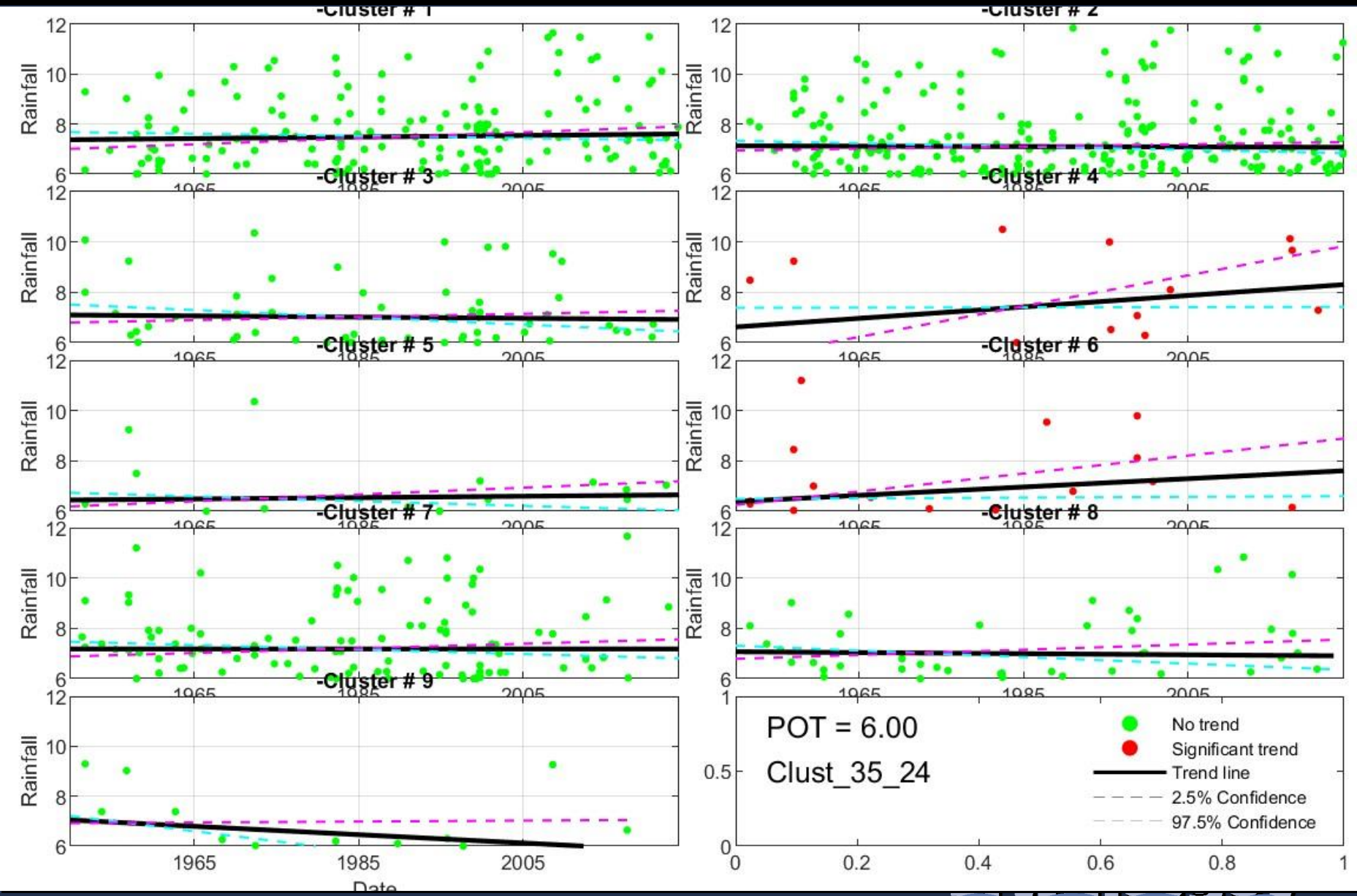

-Cluster # 1
-Cluster # 2
-Cluster # 3
-Cluster # 4
-Cluster # 5
-Cluster # 6
-Cluster # 7
-Cluster # 8
Rainfall
Date
1965
1985
2005
POT = 5.00
Clust_35_18
No trend
Significant trend
Trend line
2.5% Confidence
97.5% Confidence


-Cluster # 1
-Cluster # 2
-Cluster # 3
-Cluster # 4
-Cluster # 5
-Cluster # 6
-Cluster # 7
-Cluster # 8
Rainfall
Date
1965
1985
2005
POT = 5.00
Clust_35_24
No trend
Significant trend
Trend line
2.5% Confidence
97.5% Confidence


Cluster Visualization for POT = 6.00-inches
Latitude
Longitude
-1
1
2
3
4
5
6
7
8
9
10
11
12
13


Cluster Visualization for POT = 6.00-inches
Latitude
Longitude
-1
1
2
3
4
5
6
7
8
9
10
0
10
20
30
40
50
60
70
80
90
100


-Cluster # 1
-Cluster # 2
-Cluster # 3
-Cluster # 4
-Cluster # 5
-Cluster # 6
-Cluster # 7
-Cluster # 8
-Cluster # 9
Rainfall
Date
1965
1985
2005
POT = 6.00
Clust_35_18
No trend
Significant trend
Trend line
2.5% Confidence
97.5% Confidence


-Cluster # 1
-Cluster # 2
-Cluster # 3
-Cluster # 4
-Cluster # 5
-Cluster # 6
-Cluster # 7
-Cluster # 8
-Cluster # 9
Rainfall
Date
1965
1985
2005
POT = 6.00
Clust_35_24
No trend
Significant trend
Trend line
2.5% Confidence
97.5% Confidence

Key Takeaways

- Use multiple thresholds to identify both chronic and acute rainfall zones.
- Recent data alters cluster structure, POR length matters.
- Support adaptive infrastructure and resilience planning using these insights.
- Cluster Analysis needs to expand in a similar way the OMD basin based analysis.

Discussion

- Persistent clusters appear across both 1935–2018 and 1935–2024 datasets.
- These clusters indicate geospatial zones with long-term exposure to heavy rainfall.
- Emerging clusters (only in 1935–2024) likely reflect recent rainfall intensification post-2018.
- This pattern supports recent changes in rainfall behavior due to climate variability or urbanization.

Discussion

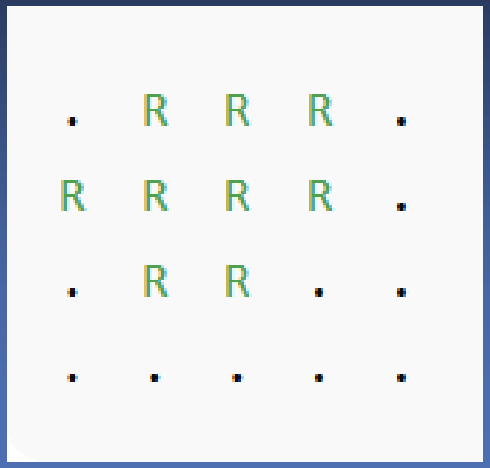
- At 4.00", many clusters are spatially diffused, covering larger areas.
- At 6.00", clusters are compact, located over known heavy-rainfall zones.
- This threshold filtering helps distinguish chronic vs. acute flood risk zones.
- Planning should treat low-threshold clusters as broader vulnerability zones, and high-threshold clusters as acute hotspots.
- More in depth Cluster Analysis is needed to address the above issues.

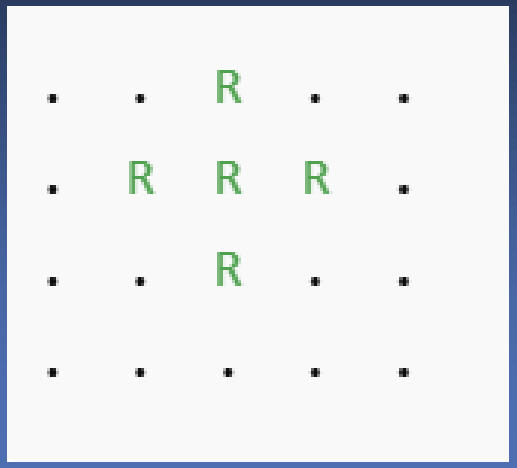
Discussion

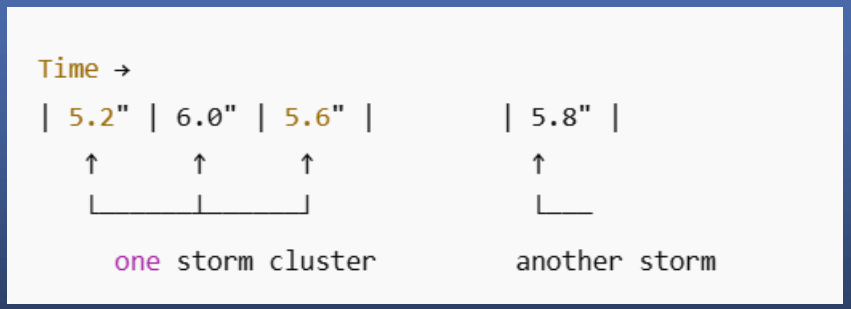
- Higher mean exceedance in newer POR suggests increased hazard frequency.
- Increased spatial density of clusters may overwhelm infrastructure in new locations.
- Updated cluster maps provide a data-backed communication tool for stormwater and resiliency planning.
- Clear visuals (CDFs and histograms) enable dialogue with stakeholders unfamiliar with raw rainfall data.

Thank you

- Questions






Time →
| 5.2" | 6.0" | 5.6" |
| 5.8" |
one storm cluster
another storm