# Read Me

In this example, the [emerging hotspot analysis](https://pro.arcgis.com/en/pro-app/latest/tool-reference/space-time-pattern-mining/emerginghotspots.htm) from ArcGIS Pro was applied to the weekly methane concentrations data (i.e., weekly interpolated images) from 2019-2021. The data for the monsoon season (May-October) were excluded due to extensive missing data in those months. The emerging hotspot analysis evaluates spatiotemporal patterns in methane concentrations using a combination of two statistical measures 1) the [Getis-Ord Gi\* statistic](https://pro.arcgis.com/en/pro-app/2.8/tool-reference/spatial-statistics/h-how-hot-spot-analysis-getis-ord-gi-spatial-stati.htm) to identify hot spots and cold spots of methane concentrations for each week, and 2) the [Mann Kendall trend test](https://pro.arcgis.com/en/pro-app/latest/tool-reference/space-time-pattern-mining/learnmorecreatecube.htm#ESRI_SECTION2_4518F0A12E194690AA986118D508E9F7) to examine how hot spots and cold spots have evolved over time.

The term ‘hot spot’ has been used generically across disciplines to describe a location has a value that is higher relative to its surroundings. However, a location with a high value may not be a statistically significant hot spot. In this case, a hot spot was defined as an area that have a high value of methane concentrations and is surrounded by other areas with high values as well.

This folder contains the following sub-folders:

1. Result

* Result.gdb contains a series of output feature classes from the emerging hot spot analysis using different neighbour distances (10 km, 15 km (default), 20 km, 30km, 50 km). They all share and the same parameters on the neighbourhood time interval (1 step: 6 days) and defined global window (individual time step)
* [netCDF](https://pro.arcgis.com/en/pro-app/3.0/tool-reference/multidimension/what-is-netcdf-data.htm) space time cube (.nc) which was used as an input for the emerging hot spot analysis and includes several resultant variables from the emerging hot spot analysis (using the default neighbourhood distance of 15 km)

1. Map: Several sample maps based on the analysis result:

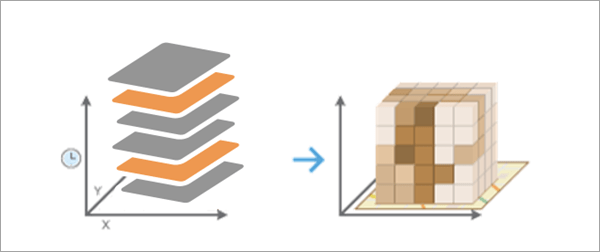
* Categories of hot and cold spots of methane concentrations between 2019 and 2021 (Figure 1)
* Trends of hot/cold spot z-scores for methane concentrations between 2019 and 2021 (Figure 2)
* Percent significant hot spots between 2019-2021 (Figure 3)
* Average weekly methane concentrations between 2019 and 2021 (Figure 4)
* Trends of methane concentrations from 2019-2021 (Figure 5)

3) Literature: Contains several research articles which use the emerging hotspot analysis tool

1. Readme document

The following figure illustrates the general workflow Using [Emerging Hotspot Analysis](https://pro.arcgis.com/en/pro-app/latest/tool-reference/space-time-pattern-mining/emerginghotspots.htm) Tool in ArcGIS Pro

Create Space-Time Cube



# Procedures

The steps for running the [emerging hotspot analysis tool](https://pro.arcgis.com/en/pro-app/latest/tool-reference/space-time-pattern-mining/emerginghotspots.htm) in ArcGIS Pro are relatively straightforward. First, you need to crate a space time cube from the time-series of raster imagery data (e.g., interpolated weekly methane concentrations images). Please refer to the ESRI blog post for detailed instructions about each step for creating the space-time cube: <https://www.esri.com/arcgis-blog/products/arcgis-pro/analytics/explore-your-raster-data-with-space-time-pattern-mining/>

There are two things need to be noticed:

1. Creating and populate a Timestamp field in the mosaic Footprints table in this case is a bit complicated than as described in the above blog link. Please refer to “Calculating Timestamp Field.doc” for details.
2. When running the **Create Space Time Cube From Multidimensional Raster Layer** tool, you need to specify the **Fill Empty Bins Method:**

The space-time cube will have one space-time bin per raster cell. Any raster cell that has NoData values for every time interval will be excluded from the output, and no space-time bin will be assigned. There are several options to fill in values for raster cells with NoData values.

* Zeros —Empty bins with be filled with zeros. This is the default.
* Spatial neighbors —Empty bins will be filled with the average value of spatial neighbors (the tool estimates based on the closest 8 nearest neighbors. A minimum of 4 of those spatial neighbors must have values to fill the empty bin using this option).
* Space-time neighbors —Empty bins will be filled with the average value of space-time neighbors (the tool estimates based on the closest 8 nearest neighbors. Additionally, temporal neighbors are used for each of those bins found to be spatial neighbors by going backward and forward 1 time step. A minimum of 13 space time neighbors are required to fill the empty bin using this option).
* Temporal trend —Empty bins will be filled using an interpolated univariate spline algorithm (the first two time periods and last two time periods at a given location must have values in their bins to interpolate values at other time periods for that location <https://pro.arcgis.com/en/pro-app/2.8/tool-reference/3d-analyst/how-spline-works.htm>)

*Note: We recommend using either Space-time neighbors or Temporal trend method. The results and sample maps included in this folder were based on the Space-time neighbors.*

1. After running the **Create Space Time Cube From Multidimensional Raster Layer** tool, you will get a warning message saying that “This tool requires projected data to accurately measure distances. The Input Multidimensional Raster Layer will be projected to the WGS 1984 World Equidistant Cylindrical projection (WKID 4087).”

Reprojecting the multidimensional raster layer will also change the cell size (spatial resolution) of the data. The cell size of the resultant space time cube is 1 km.

After you created the space time cube, you can use it as an input for running the [emerging hot spot analysis tool](https://pro.arcgis.com/en/pro-app/latest/tool-reference/space-time-pattern-mining/emerginghotspots.htm). There are several important parameters you need to specify that will affect the results.

1. To determine which bins will be included in each analysis neighborhood, the tool first finds neighboring bins that fall within the specified **Conceptualization of Spatial Relationships**. Next, for each of those bins, it includes bins at those same locations from N previous time steps, where N is the **Neighborhood Time Step** value you specify.

* Your choice for the Conceptualization of Spatial Relationships parameter should reflect inherent relationships among the features you are analyzing. The more realistically you can model how features interact with each other in space, the more accurate your results will be.

The default **Conceptualization of Spatial Relationships** is **Fixed distance**. A bin is considered a neighbor if its centroid falls within the Neighborhood Distance and its time interval is within the Neighborhood Time Step you specify. When you do not provide a Neighborhood Distance value, one is calculated for you based on the spatial distribution of your point data. The **Number of Neighbors** parameter may override the **Neighborhood Distance** for the **Fixed distance** option or extend the neighbor search for the **Contiguity edges only** and **Contiguity edges corners** options.

*Note: In this case, if you don’t specify anything, the default is the fixed distance of 15247.763236 meters. We also tried using different distance of 10 km, 20 km, 30km, 50 km. As you may find out from the results, as neighborhood distance increases, hot spots will become larger and fewer while smaller neighborhood distances capture more localized trends. You can determine the most suitable neighbourhood distance based on your research goal, study context etc. All the sample maps were based on the default distance of 15 km.*

* Neighbourhood time step: The number of time-step intervals to include in the analysis neighborhood. This value determines which features are analyzed together in order to assess local space-time clustering.

*Note: If you don’t specify anything, the default value is 1 step (6 days). As a result, all bins within the Neighborhood Distance, and all of their associated bins for the previous one time-step interval will be included in the analysis neighborhood.*

1. Define global window: the statistic works by comparing a local statistic calculated from the neighbors for each bin to a global value. This parameter can be used to control which bins are used to calculate the global value.

• Entire cube—Each neighborhood is analyzed in comparison to the entire cube. This is the default.

• Neighborhood Time Step—Each neighborhood is analyzed in comparison to the bins contained within the Neighborhood Time Step specified.

• Individual time step—Each neighborhood is analyzed in comparison to the bins in the same time step.

**Note:** *The default option (entire cube) might not be useful unless you expect a very consistent pattern over the entire study time period. In this case, we recommend using the Individual time step. For more information, please refer to this post:* [*https://community.esri.com/t5/spatial-statistics-questions/emerging-hot-spot-analysis/td-p/1171928*](https://community.esri.com/t5/spatial-statistics-questions/emerging-hot-spot-analysis/td-p/1171928)

# Results

The Emerging hot Spot Analysis produces a feature class which contains the following variables as shown in the screenshot below.

Table

Description automatically generated

You can create maps based on some of the above variables to show important insights on the spatio-temporal patterns in the data.

1. PATTERN: The result category used to classify each location as one of [17 categories](https://pro.arcgis.com/en/pro-app/3.0/tool-reference/space-time-pattern-mining/learnmoreemerging.htm#GUID-09587AFC-F5EC-4AEB-BE8F-0E0A26AB9230) of hot spot and cold spot trends.

As shown in Figure 1, the persistent hot spot and intensifying hot spot of methane concentrations were prominent within the Dhaka Division. Specifically, the persistent hot spot was found around the Narayanganj district, which was a hot spot for at least 90% of the study period, while the intensity of clustering was not trending significantly upward or downward over time. An intensifying hot spot of methane was found within the Dhaka district, which indicates that it was a hot spot for at least 90% of the study period and there was a statistically significant increase in the intensity of clustering of high concentrations over the three years. In contrast, there was a large area of cold spots in the southeastern part of the country (e.g., Chittagong Division), which are mainly persistent and intensifying cold spots.



Figure 1- Categories of hot and cold spots of methane concentrations between 2019 and 2021

1. TREND\_BIN: The result category used to classify each location as having a statistically significant upward or downward trend for hot/cold spot z-scores.

-3: down trend, 99 percent confidence

-2: down trend, 95 percent confidence

-1: down trend, 90 percent confidence

0: no significant trend

1: up trend, 90 percent confident

2: up trend, 95 percent confident

3: up trend, 99 percent confident

***Note:*** *The hot/cold spot z-scores are from the Getis-Ord Gi\* statistic. For statistically significant positive Gi\* statistic z-scores, the larger the z-score is, the more intense the clustering of high values (hot spot). For statistically significant negative z-scores, the smaller the z-score is, the more intense the clustering of low values (cold spot).*

Map

Description automatically generated

Figure 2 –Trends of hot/cold spot z-scores for methane concentrations between 2019 and 2021

1. PERC\_HOT: Percentage of time intervals a location being a significant hot spot. For example, as shown in the map below, areas in dark red have been significant hot spots for more than 80 percent of the weeks within the study period.



Figure 3-Percent significant hot spots between 2019-2021

1. MEAN\_VALUE: Mean weekly methane concentrations over the study period.

For example, this map below shows the average weekly methane concentrations by stand division from 2019-2021.

Map

Description automatically generated

Figure 4 – Average weekly methane concentrations between 2019 and 2021

Running Emerging Hot Spot Analysis also adds some analysis results back to the [netCDF](https://pro.arcgis.com/en/pro-app/3.0/tool-reference/multidimension/what-is-netcdf-data.htm) input space time cube. A summary of the variables added to the **Input Space Time Cube** can be found here: <https://pro.arcgis.com/en/pro-app/latest/tool-reference/space-time-pattern-mining/emerginghotspots.htm>

You can visualize these variables by the space-time cube data in either 2D or 3D using the tools in the [Utilities toolset](https://pro.arcgis.com/en/pro-app/3.0/tool-reference/space-time-pattern-mining/an-overview-of-the-utilities-toolset.htm) or by downloading the [Space Time Cube Explorer](https://spatialstats.github.io/addin/).

For example, the map below shows the trend for methane concentrations over the study period. Over 99% of the cells have statistically significant increase in the methane concentrations within this time.



Figure 5-Trends of methane concentrations between 2019 and 2021

***Note:***

1. *For the above maps, you may want to also add the boundaries of divisions/subdivisions as a reference layer.*
2. *You can also run the emerging hot spot analysis for each year separately and compare the results among them.*