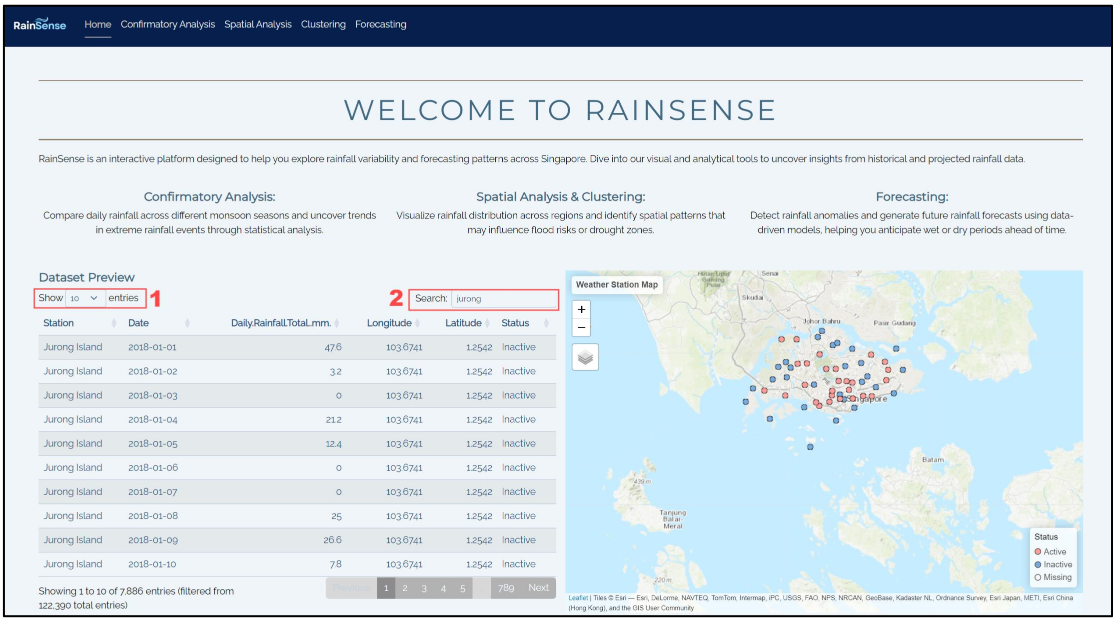
User Guide: RainSense - Rainfall Variability and Forecasting in Singapore

1.Home Page

This page briefly describes the functionality that will be implemented in this app.



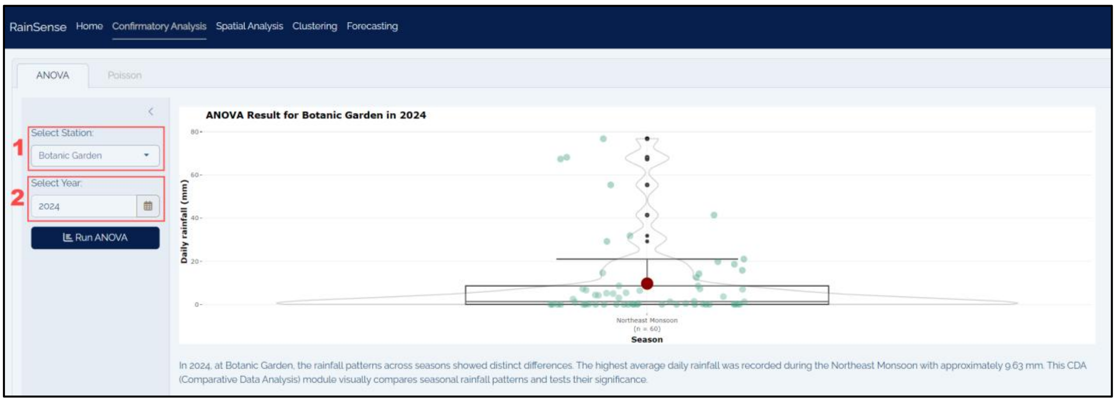
The left part is a preview of the data and allows the user to change the number of rows of data that can be displayed by modifying the parameter in [1]. It is also possible to search for a specific rain station by entering the name of the rain station you want to observe in [2].

The map on the right shows the current status of all rain stations in Singapore. Red is active status, blue is inactive status, and white is missing data status

2.Confirmatory Analysis

2.1 ANOVA

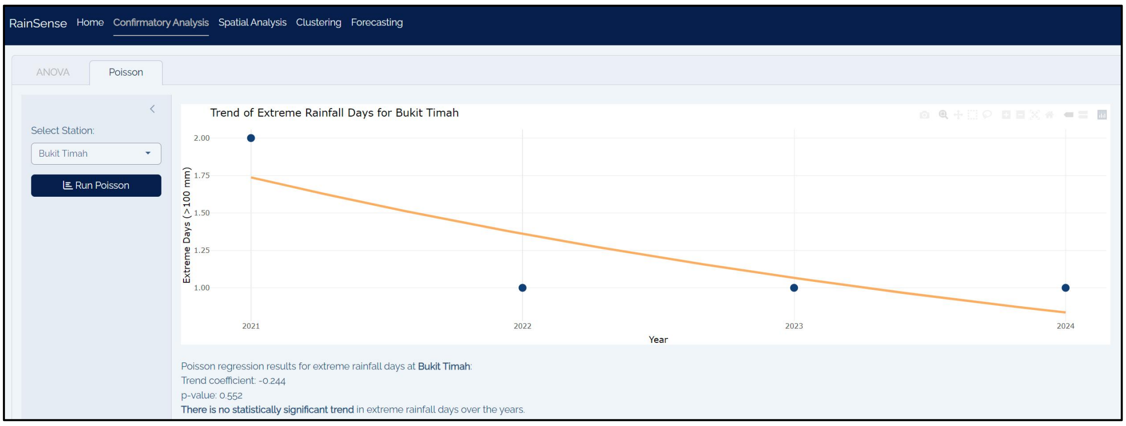
This page is the ANOVA part of the confirmatory analysis, which is mainly used to show the seasonal distribution of rainfall at a particular rain station in a given year.



[1] Users can select the rain station they want to know about from the preset rain stations.
[2] Users can select the years they want to know about, in this feature, the optional years are 2018-2024.

After clicking on the "Run ANOVA" button, the results of the analysis will appear on the right side, along with an explanation of the results of the analysis.

2.2 Poisson



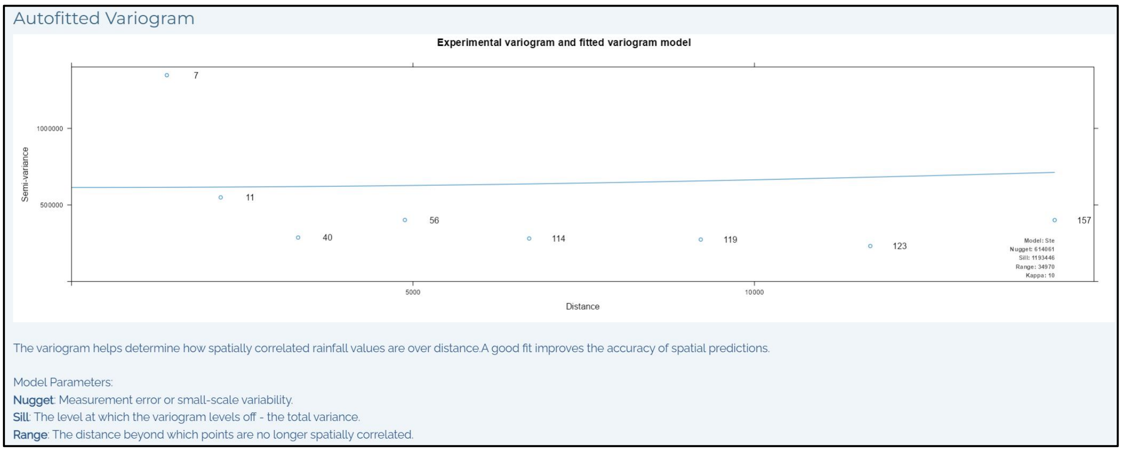
This page shows the Poisson regression analysis of extreme rainfall days (rainfall greater than 100 mm) for all stations. After selecting the rain station users wish to view, click on the "Run Poisson" button and the graph shows the trend of the number of extreme rainfall days in different years. The Poisson regression results are given in the text below.

3.Spatial Analysis

3.1 Ordinary Kriging

This page visualizes the spatial distribution of rainfall for 2024 using Ordinary Kriging.

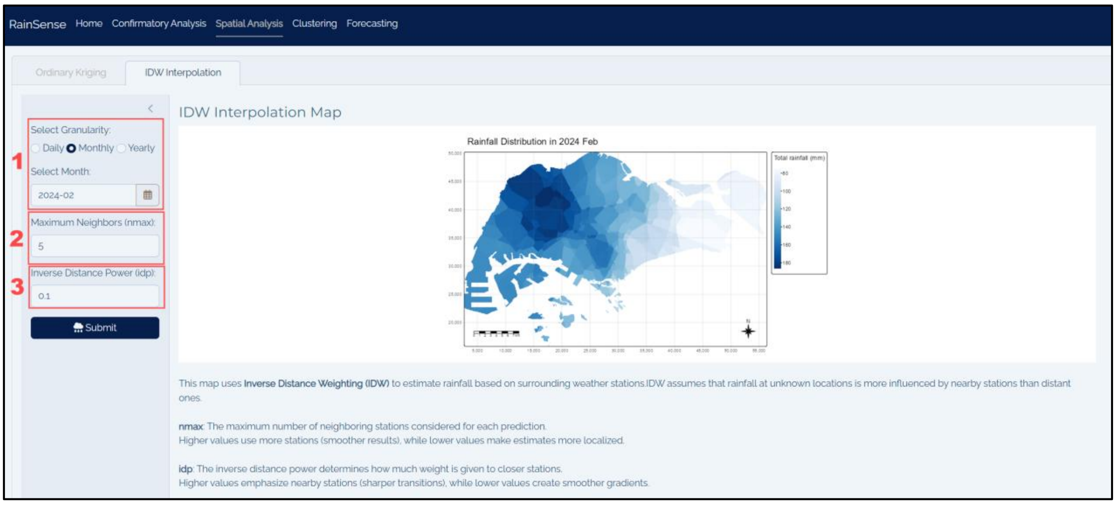
[1]Users can freely choose to view daily, monthly or annual data.
[2]Corresponding to the user's Granularity selection above, the specific time selection changes accordingly, and the user is free to choose a specific time node ranging from 2018-2024.
After clicking the “Submit” button, the results of the analysis will appear on the right side. Based on the known station data and its spatial relationship, the rainfall is estimated for the unmeasured area. Darker areas represent high estimated rainfall and lighter areas represent low rainfall.



Also, the automatically fitted variogram function is displayed at the bottom. By analyzing the variance function, the variation of rainfall data over different distances can be understood, and a good fit helps to improve the accuracy of spatial prediction.

3.2 IDW Interpolation

In this section, the rainfall distribution for February 2024 is interpolated using the IDW (Inverse Distance Weighting) method to show the rainfall in different regions.

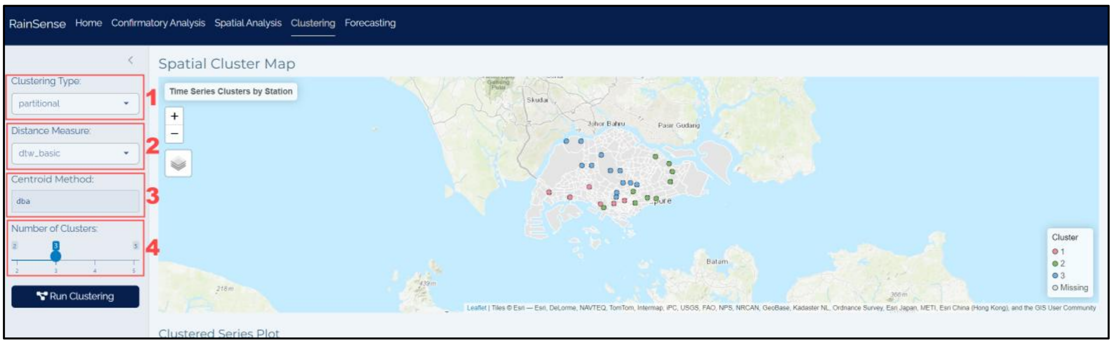


[1]Users are free to view data on a daily, monthly or annual basis, and are free to choose a specific time period, ranging from 2018-2024.
[2]Maximum number of neighbors: the larger the value, the more stations are involved in the estimation and the smoother the results; smaller values make the estimation more localized.
[3]Inverse Distance Power (idp): Used to determine the weights given to neighboring sites. The higher the value, the more weight is given to the neighboring sites and the more pronounced the transition is; a lower value produces a smoother gradient.

After clicking the "Submit" button, the results of the analysis will appear on the right side. The darker the color represents the amount of rainfall, the darker the color the higher the amount of rainfall, the lighter the color the lower the amount of rainfall. The color bar shows the range of rainfall. The text below explains the IDW method and describes the role of the nmax and idp parameters to help understand the map results and the effect of parameter adjustments on the estimates.

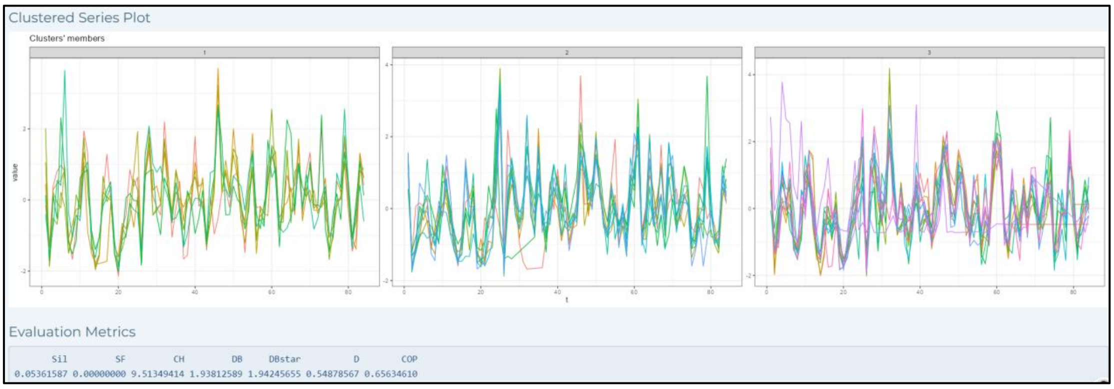
4.Clustering

This section is a spatial clustering map for clustering time series data from meteorological sites.



[1]Clustering Type: Users can choose from partitional, hierarchical and fuzzy options.
[2]Distance Measure: Users can choose dtw (Dynamic Time Warping), softdtw, sbd, euclidean, gak method.
[3]Centroid Method: Adaptive selection of the Centroid according to the distance measurement method selected by the user in the previous step.
[4]Number of Clusters: Adjust the number of clusters by adjusting the slider in the range of 2-5.

After clicking the Submit button, the map on the right will show the clustering results. Different color points represent different clusters, such as red for Cluster 1, green for Cluster 2, blue for Cluster 3, and gray for missing data. Through this visualization, the spatial distribution of the clusters at different sites can be visualized, which helps to analyze which sites have similar rainfall time series characteristics.



The Clustered Series Plot below contains three subplots corresponding to each of the clustered clusters set up in the previous figure. Each subplot shows the variation of each member (weather station time series data) in the cluster to which it belongs, and the fluctuation of the time series values of different stations with a certain variable (horizontal axis “t”, which may be the time step, etc.) is indicated by the lines, which facilitates the comparison of the similarity and difference of the time series of the stations in the same cluster intuitively.

5.Forecasting

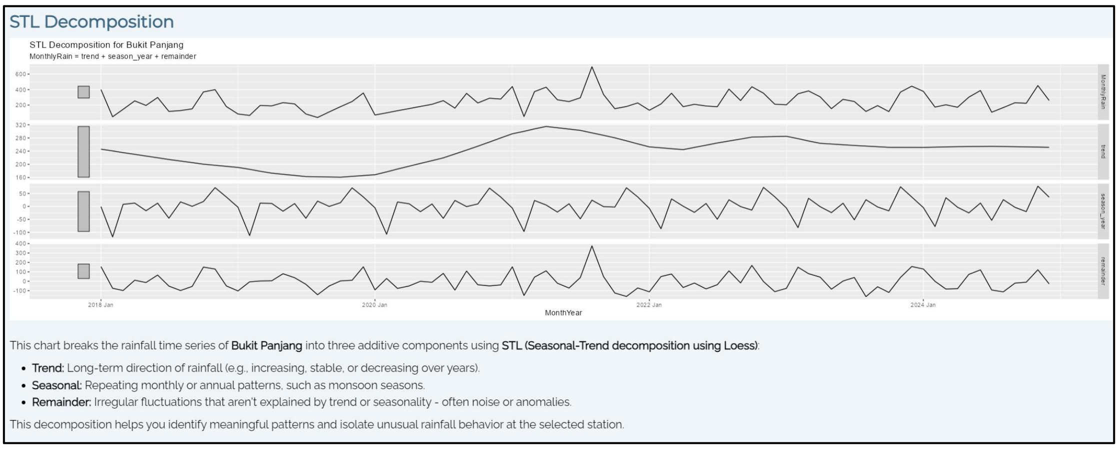
5.1 Decomposition

This section mainly displays the Time-series rainfall data of the rainfall station selected by the user.

Upper line graph: shows the raw data of monthly rainfall for the selected area for the period 2018 - 2024, which visualizes the fluctuation of rainfall over time.

Middle Autocorrelation Function (ACF) plot: analyzes the correlation of rainfall data with past months, which can help to find out if there are repeating cycles or persistent patterns in the rainfall pattern.

Bottom right line graph: Compare the seasonal rainfall trend of each month in different years, with different color lines representing different years, which makes it easy to find out the months with consistently higher or lower rainfall.



The graph below uses the STL (Seasonal - Trend decomposition using Loess) decomposition method to split the rainfall time series into three additive components:

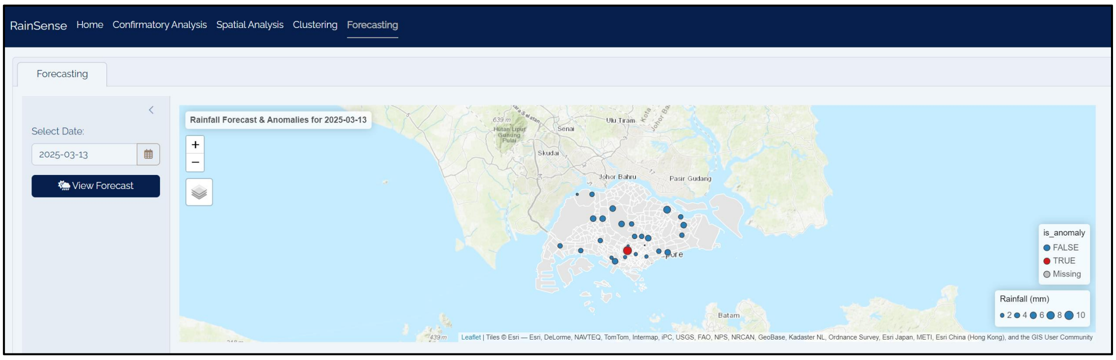
Trend: Reflects the long-term changes in rainfall over a number of years. Through the trend graph, it can grasp the macro changes in rainfall and provide reference for long-term water resources planning.

Seasonal: Reflects the recurring pattern of rainfall on a monthly or yearly basis. It helps to understand the regular fluctuation of rainfall at different times of the year, which is a guide for agricultural production arrangements, urban flood preparedness, etc. .

Residuals (Remainder): represent irregular fluctuations that cannot be explained by trends and seasonality, mostly noise or outliers. It can help to identify anomalies in rainfall data and troubleshoot the effects of data collection errors or special events such as extreme weather.

5.2 Forecasting

This section is mainly used to demonstrate rainfall prediction and outlier detection.



Users are free to choose the dates, where 2018-2024 are actual data and 2025 are predicted data based on historical data using the ARIMA model. After clicking the “View Forecast” button, the map on the right side will show the rainfall on that day. The larger the rainfall data recorded at the rain station, the larger the blue circle will be. A red colored rain station means that it has an outlier.