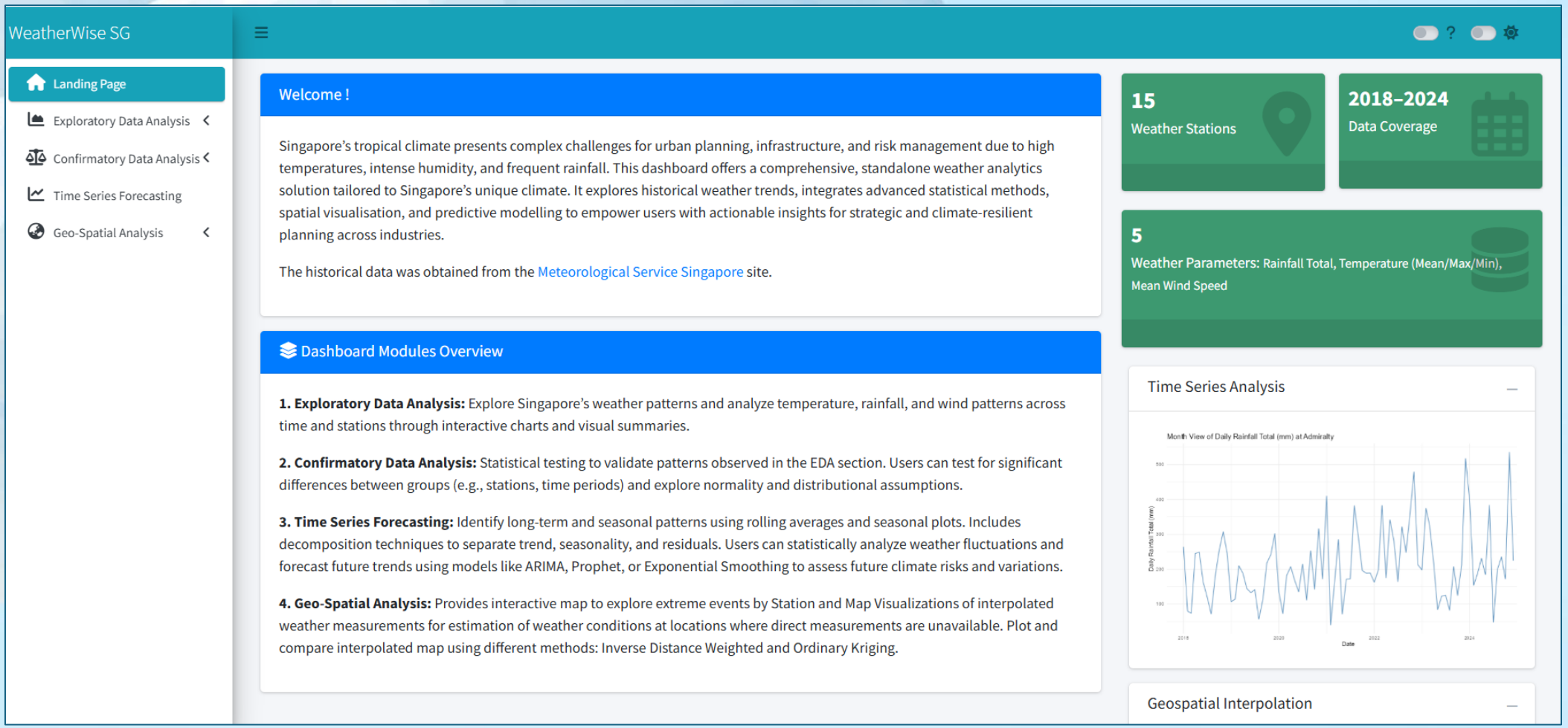
WeatherWise Singapore

Shiny App User Guide

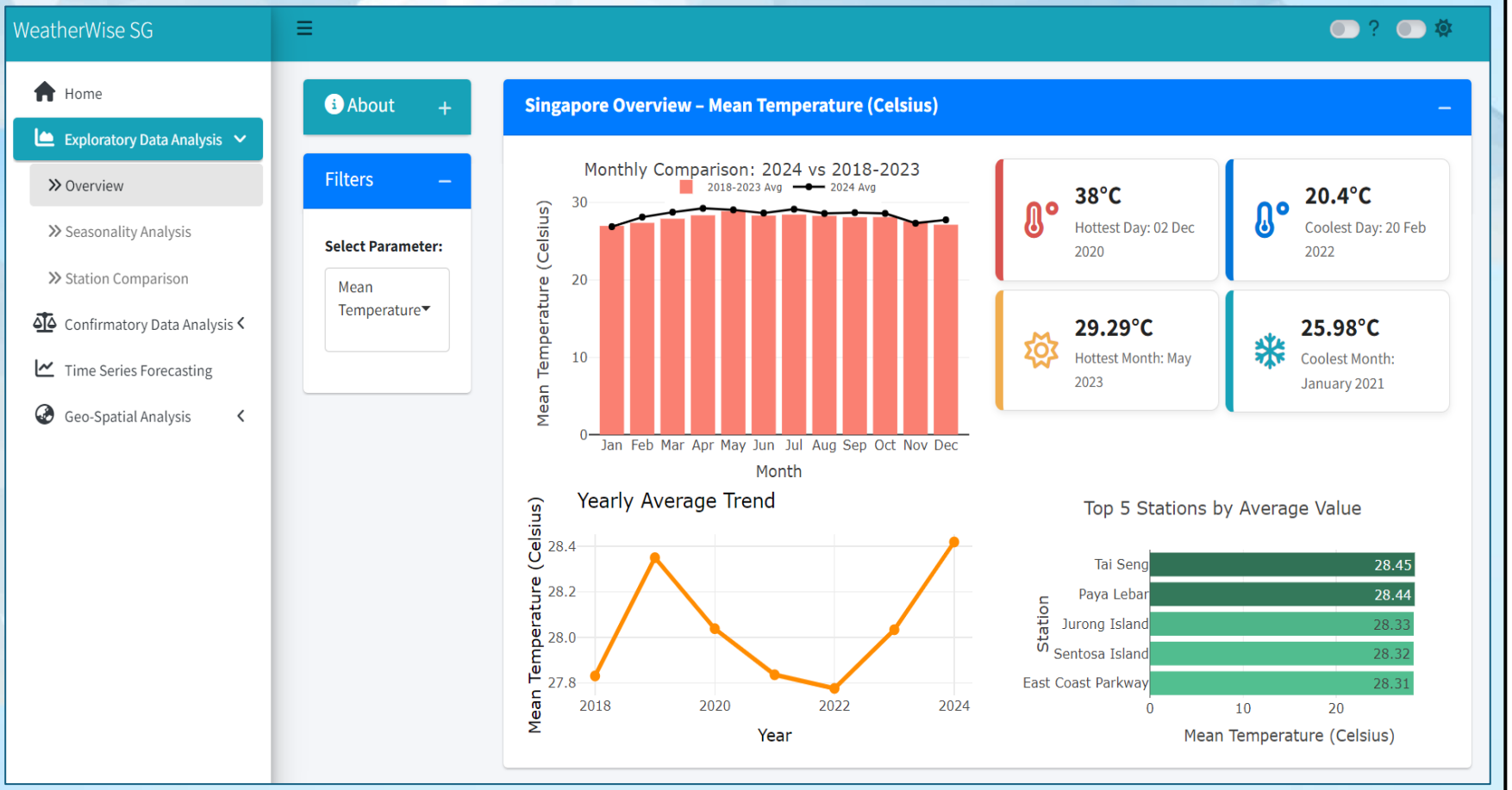
1. Home Page

This page provides a brief overview of the WeatherWise SG dashboard, its objectives, and the four main analytical modules. Key dataset statistics are displayed on the top right, alongside sample visual outputs from the analyses.



2.1 Exploratory Data Analysis - Overview

This page provides a high-level understanding of Singapore's weather patterns between 2018 and 2024 for the selected weather parameter. Users can interactively explore monthly patterns and annual trends using hover and intuitive visualizations.



1. Select Parameter: Use the dropdown on the left to choose a weather parameter to explore: Daily Rainfall, Mean/Max/Min Temperature or Mean Wind Speed

2. Monthly Comparison Chart: Compares the monthly values in the selected year (e.g., 2024) against the historical average (2018–2023).

3. Key Stats Boxes (Top Right): Displays key insights: highest/lowest value and date, month with highest/lowest average value.

- These adapt based on selected parameter where applicable.

4. Yearly Trend: Shows year-over-year average change from 2018 to 2024.

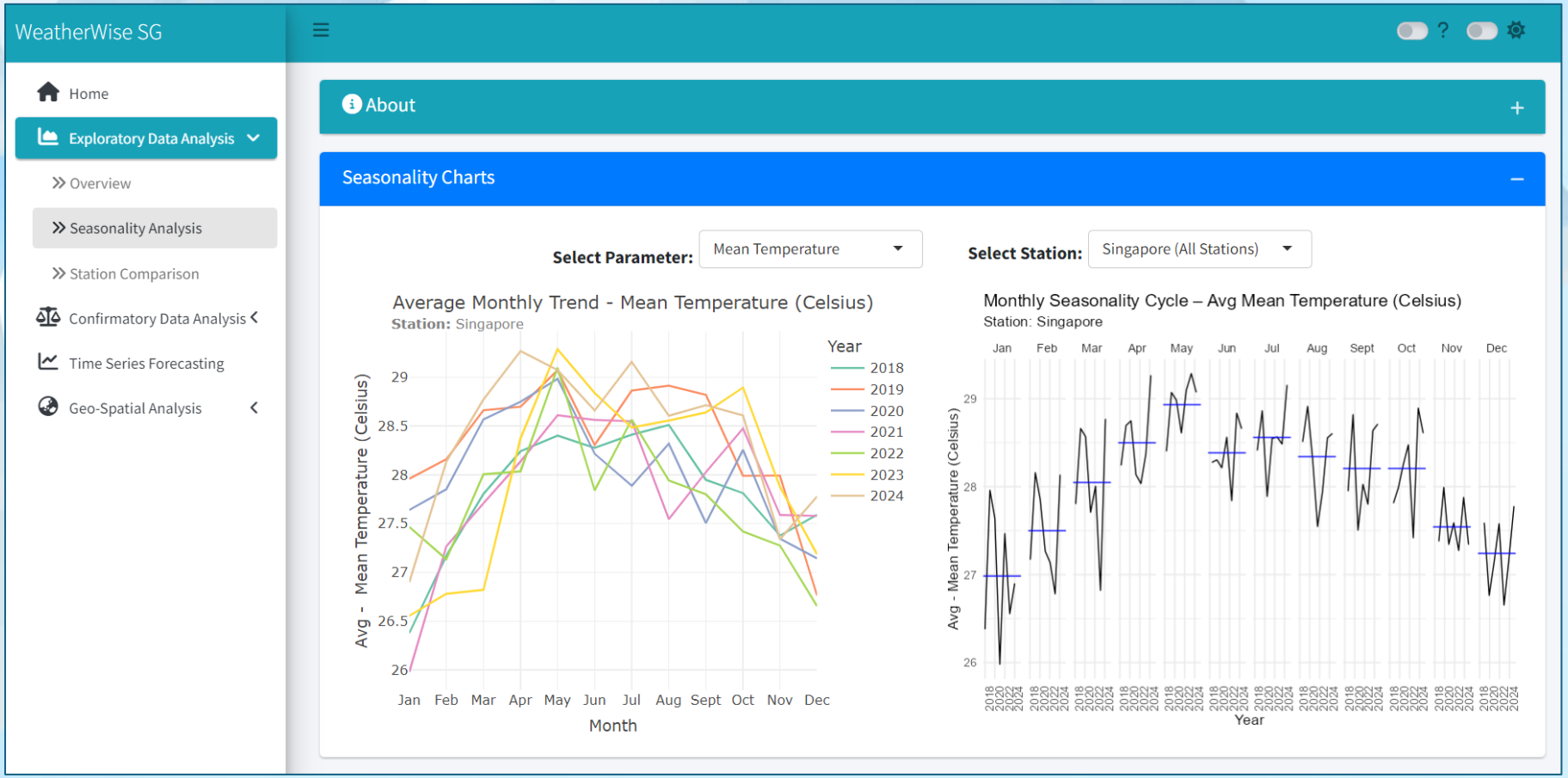
5. Station Rankings: Bar chart showing the top 5 stations by average value.

- For Minimum Temperature, shows the lowest 5 stations instead.

Note: For **Rainfall**, totals (not averages) are shown.

2.2 Exploratory Data Analysis – Seasonality Analysis

This section helps uncover monthly and seasonal patterns across years and stations for various weather parameters.



1. Select Parameter: Use the dropdown to choose a weather parameter to explore.

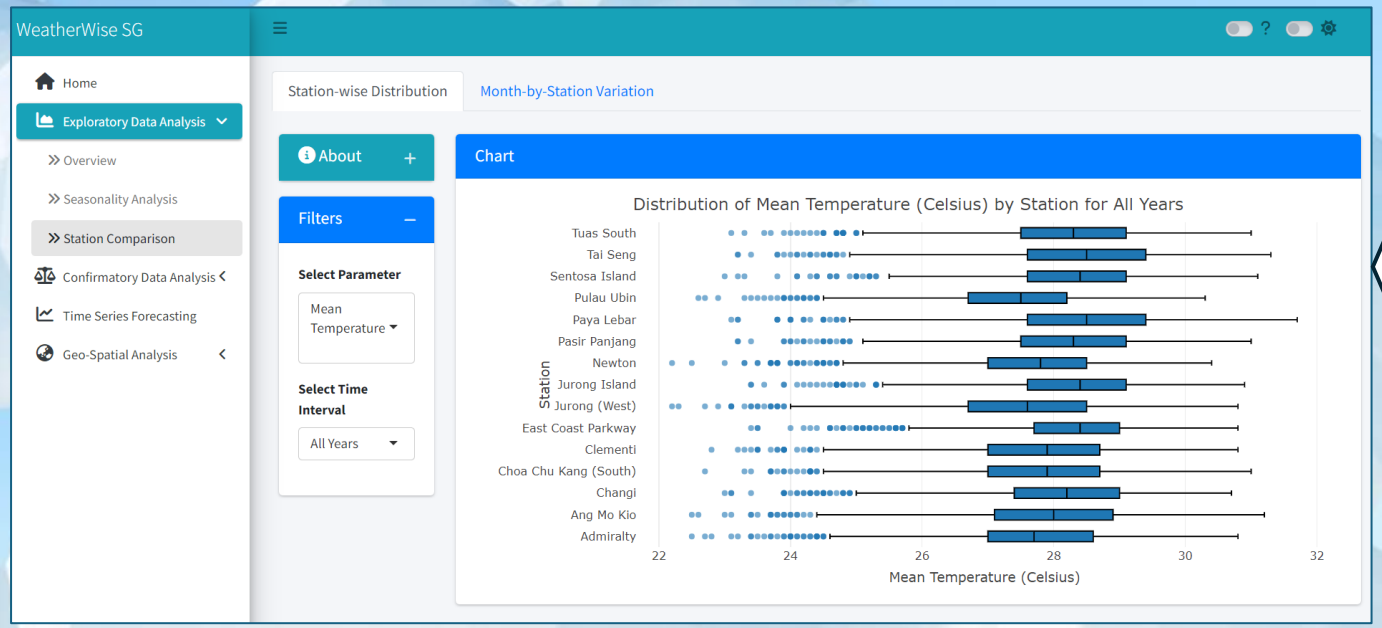
2. Select Station: View trends for Singapore (All Stations) - aggregated national trends or select an individual weather station for localized patterns

3. Average Monthly Trend (Left Plot): Multi-line plot showing average value per month for each year (2018–2024). Hover to see the values.

4. Monthly Seasonality Cycle (Right Plot): Highlights monthly seasonality cycles across all years. Shows year-over-year variability for each month (black line) and monthly averages (blue horizontal bars)

Note: Rainfall shows monthly total values instead of averages.

2.3 Exploratory Data Analysis – Station Comparison



Distribution across stations (Boxplot): identify which stations consistently record higher/lower values and where extreme conditions are more frequent.

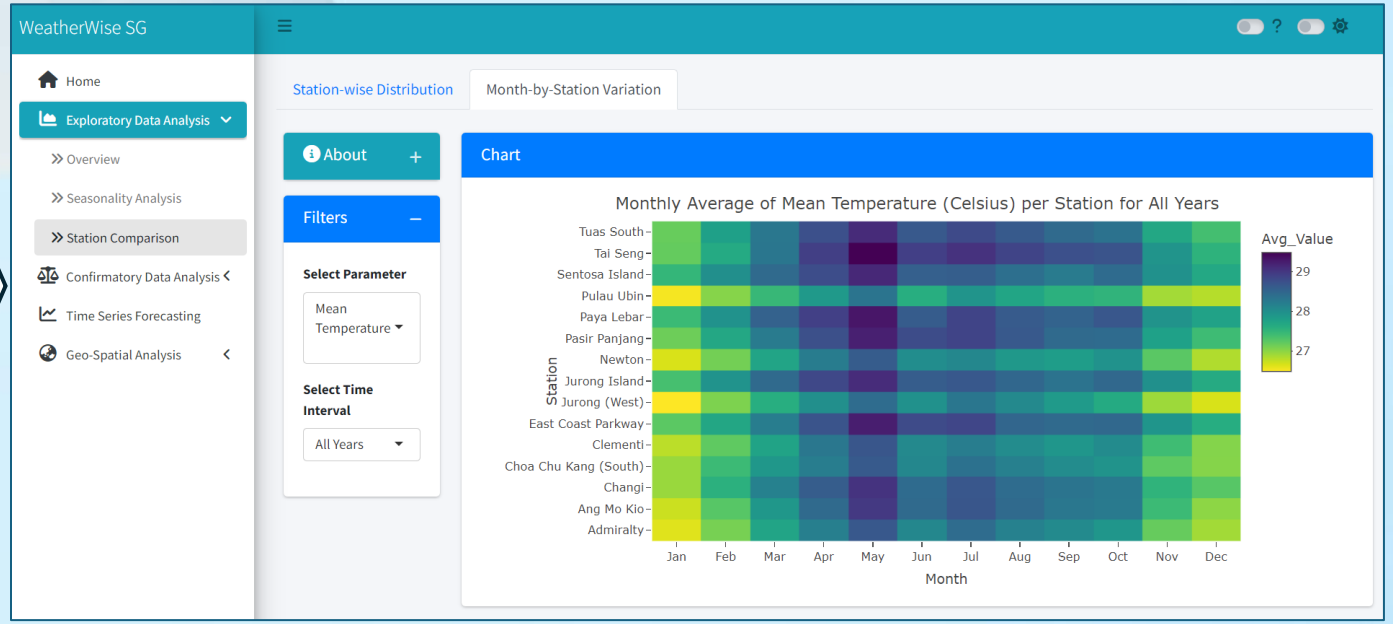
1. Select Parameter: Use the dropdown to choose a weather parameter to explore.
2. Select Time Interval: View distribution across *All Years* (aggregated) or filter by a specific year.
3. Interpret distributions: Each station has a boxplot showing Median temperature (center line), Interquartile range (IQR) (box), Whiskers showing spread of typical values, Dots indicating outliers (extreme values).

- *Hover on boxplot to see these values.*

Monthly Average by Station (Heatmap): Identify patterns of monthly variation across stations and compare seasonal behaviors.

1. Select Parameter: Use the dropdown to choose a weather parameter to explore.
2. Select Time Interval: View distribution across *All Years* (aggregated) or filter by a specific year.
3. Interpret heatmap: Rows represent individual stations. Columns represent months. Color intensity indicates average value: Lighter/warmer colors show lower values while Darker/cooler colors show higher values.

- *Hover to observe the exact values.*

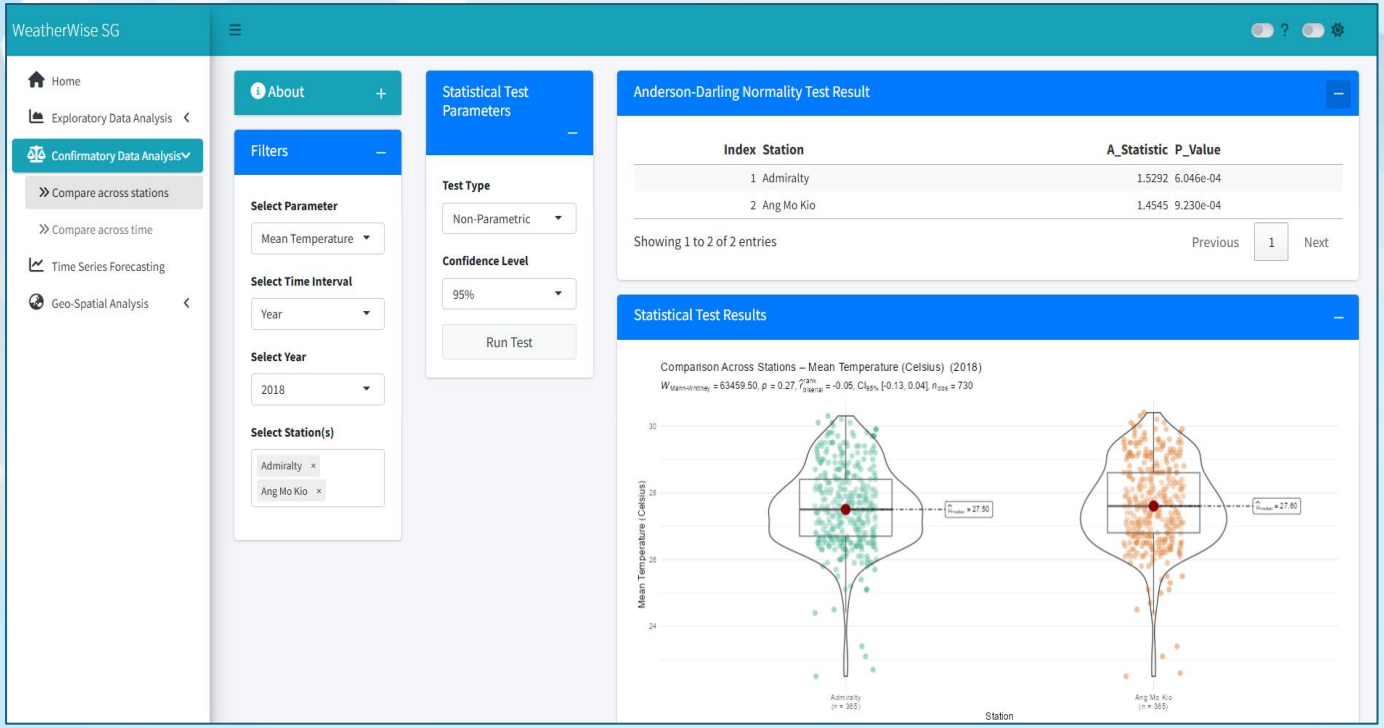


3.1 Confirmatory Data Analysis – Compare across stations

This section enables statistical comparison of weather metrics across different weather stations to detect significant differences or patterns.

1. Filters Panel – Data Selection: Use the filters on the left to select data inputs.

- **Select Parameter:** Choose a weather parameter *(e.g. Mean Temperature)* to explore
- **Select Time Interval:** Year (*2018,...)* or Month *(Jan 2018,...)*
- **Select Year/Month:** Displays based on the user selected time interval.
- **Select Station(s):** Choose 1 to 8 stations (maximum). If more than 8 are selected, a warning is triggered when you run the test.



2. Statistical Test Parameters:

- **Test Type:** Select the type of test to apply *(Parametric, Non-Parametric, Robust, or Bayes-Factor)*
- **Confidence Level:** Choose your preferred confidence level *(90%, 95%, or 99%)* for statistical testing.
- **Run Test Button:** Click this after all the selections. A loading overlay appears during processing.

3. Normality Test Section (Anderson-Darling Test):

- **Collapsed by default:** Click the "+" icon to expand and view the normality results for each selected station. *(uncollapsed in the image for display purposes)*
- **Columns Explained:** A_Statistic – Test statistic value and P_Value of normality test
- If any p_value is < 0.05, it indicates that the sample failed to confirm normality and should perform non-parametric test.

4. Statistical Test Results: Visual summary

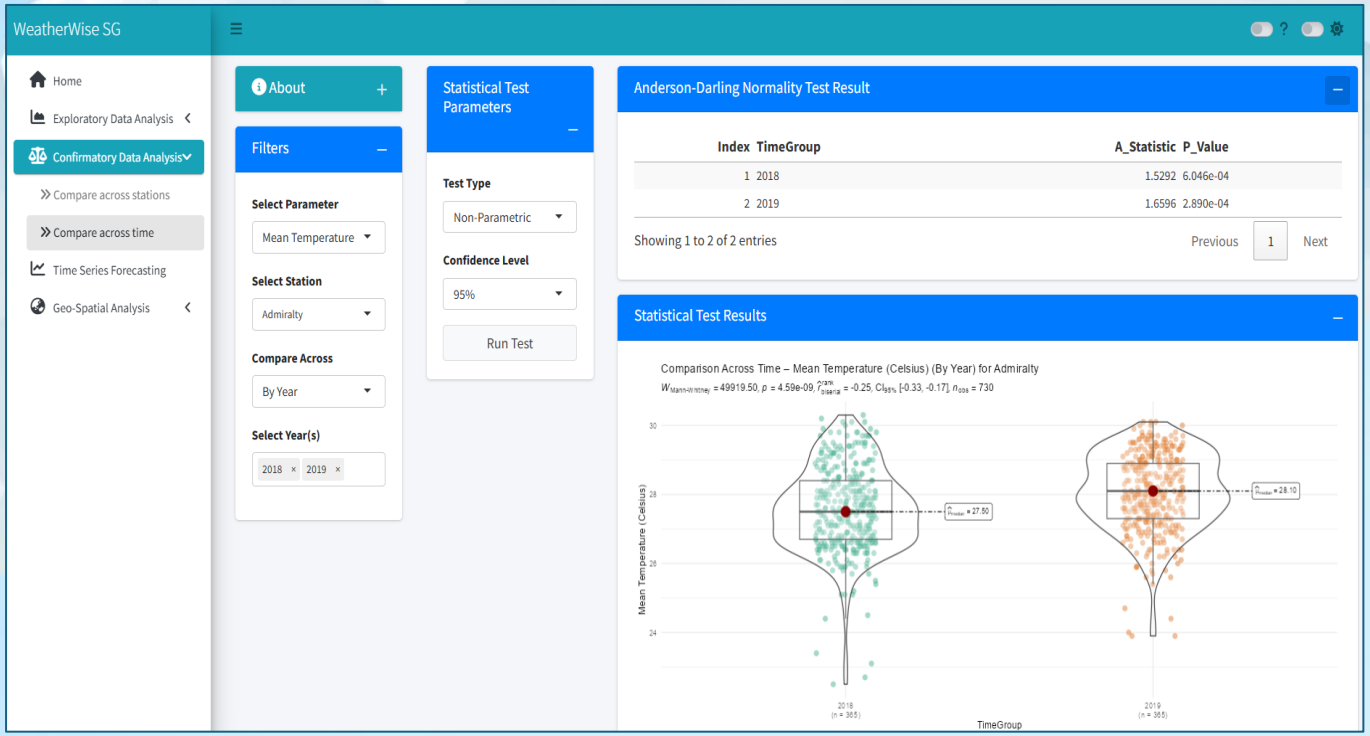
- **If 1 station selected:** A histogram plot is displayed of one-sample test comparing observed values to the mean.
- **If 2+ stations:** A boxplot-violin chart is shown for each station.
  - Displays distribution, median, test statistics.
  - **Pairwise statistical comparisons** are performed to check for significant differences across stations.
  - **Interpretation tips:** Look for differences in distribution and significance indicators on the plot.

3.1 Confirmatory Data Analysis – Compare across time

This section enables statistical comparison of weather metrics across different time intervals for a particular station to detect significant differences or patterns.

1. Filters Panel – Data Selection: Use the filters on the left to select data inputs.

- **Select Parameter:** Choose a weather parameter *(e.g. Mean Temperature)*
- **Select Station:** Any station *(eg. Admiralty)*
- **Compare Across:** Select the time category to compare across *(By Year – eg 2018, By Month (All Years) – eg Jan, or By Month (Selected Year) – eg Jan 2018)*
- **Select Year(s)/Month(s) or Select Year and Month(s):** Displays either of the 3 options based on the time category selected.



2. Statistical Test Parameters:

- **Test Type:** Select the type of test to apply *(Parametric, Non-Parametric, Robust, or Bayes-Factor)*
- **Confidence Level:** Choose your preferred confidence level *(90%, 95%, or 99%)* for statistical testing.
- **Run Test Button:** Click this after all the selections. A loading overlay appears during processing.

3. Normality Test Section (Anderson-Darling Test):

- **Collapsed by default:** Click the "+" icon to expand and view the normality results for each selected time group. *(uncollapsed in the image for display purposes)*
- **Columns Explained:** A_Statistic – Test statistic value and P_Value of normality test
- If any p_value is < 0.05, it indicates that the sample failed to confirm normality and should perform non-parametric test.

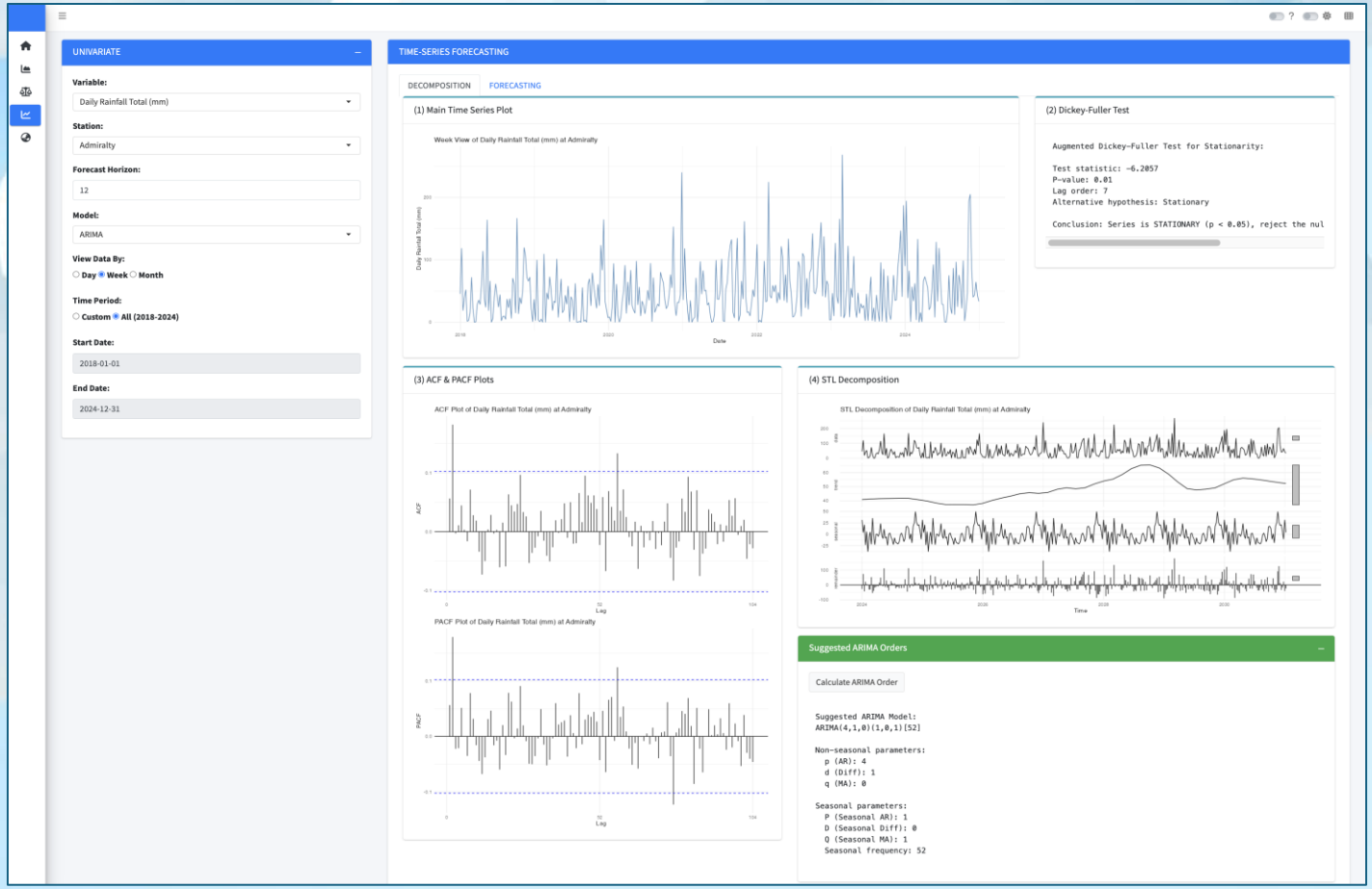
4. Statistical Test Results: Visual summary

- **If 1 time group selected:** A histogram plot is displayed of one-sample test comparing observed values to the mean.
- **If 2+ time groups:** A boxplot-violin chart for each time group.
  - Displays distribution, median, test statistics.
  - **Pairwise statistical comparisons** are performed to check for significant differences across time groups.
  - **Interpretation tips:** Look for differences in distribution and significance indicators on the plot.

4.1 Time-Series Forecasting - Decomposition

The Time-Series Forecasting Dashboard is designed to help you predict future values based on historical data. It is intuitive and user-friendly, even if you have limited experience with data analysis or forecasting.
Access this module by selecting Time Series Forecasting from the navigation menu on the left. The module contains two primary sections:

1. **Univariate Settings Panel** (left side)
2. **Time-Series Forecasting Tab Panel** (right side): Decomposition & Forecasting



Using the Univariate Settings Panel

Variable Selection:

- Click on the dropdown under Variable and Choose the type of data you wish to forecast (e.g., "Daily Rainfall Total (mm)").

Station Selection:

- Click the Station dropdown and Select your preferred station (e.g., "Changi").

Forecast Horizon:

- Specify how far into the future you want predictions by entering a numerical value (e.g., 12 weeks).

Model Selection:

- Click the Model dropdown and Choose the forecasting model to apply:
    - ARIMA: Advanced statistical model suitable for data with patterns.
    - Prophet: Good for time series data that exhibit strong seasonal effects and anomalies.
    - ETS: Exponential smoothing suitable for data with clear trends and seasonality.

View Data By:

- Select how you want to aggregate data: by Day/Week/Month

Time Period:

- Custom: Set a specific date range for historical data.
- All: Uses all available historical data.

Setting Dates (Custom Period only):

- Enter your preferred Start Date and End Date (fixed).

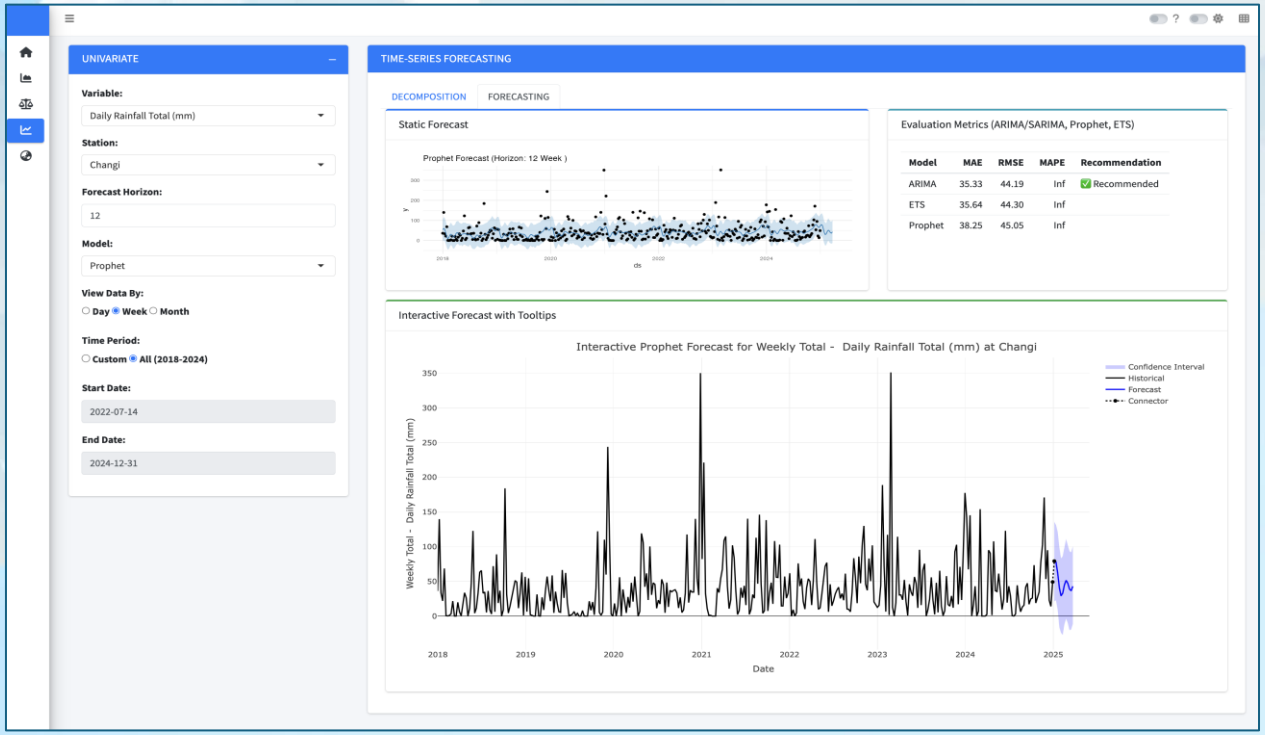
On the DECOMPOSITION tab-panel, the selected time-series is plotted. ACF/PACF and STL of the time-series is displayed for visual analysis. It is automatically checked for stationary or non-stationary. Compute Arima order seasonal and non-seasonal for Arima/Sarima modelling

4.2 Time-Series Forecasting - Forecasting

Performance comparison of models using MAE, RMSE, and MAPE (lower values indicate better accuracy). Recommended model marked by ✔. Displays historical and predicted data through static and interactive visuals, including confidence intervals.

Tips for Best Results

- Choose your forecast horizon carefully based on your planning needs.
- Compare different models (ARIMA, Prophet, ETS) using provided evaluation metrics.
- Always check data stationarity (using Dickey-Fuller Test) before deciding on a model.



Evaluation Metrics

- Compares performance of forecasting models based on error metrics:
  - MAE (Mean Absolute Error): Lower is better.
  - RMSE (Root Mean Squared Error): Lower indicates more accurate forecasts.
  - MAPE (Mean Absolute Percentage Error): Measures prediction accuracy as a percentage; lower values are ideal.
- Recommendation:
  - A green tick ✔ indicates the model recommended based on accuracy metrics.

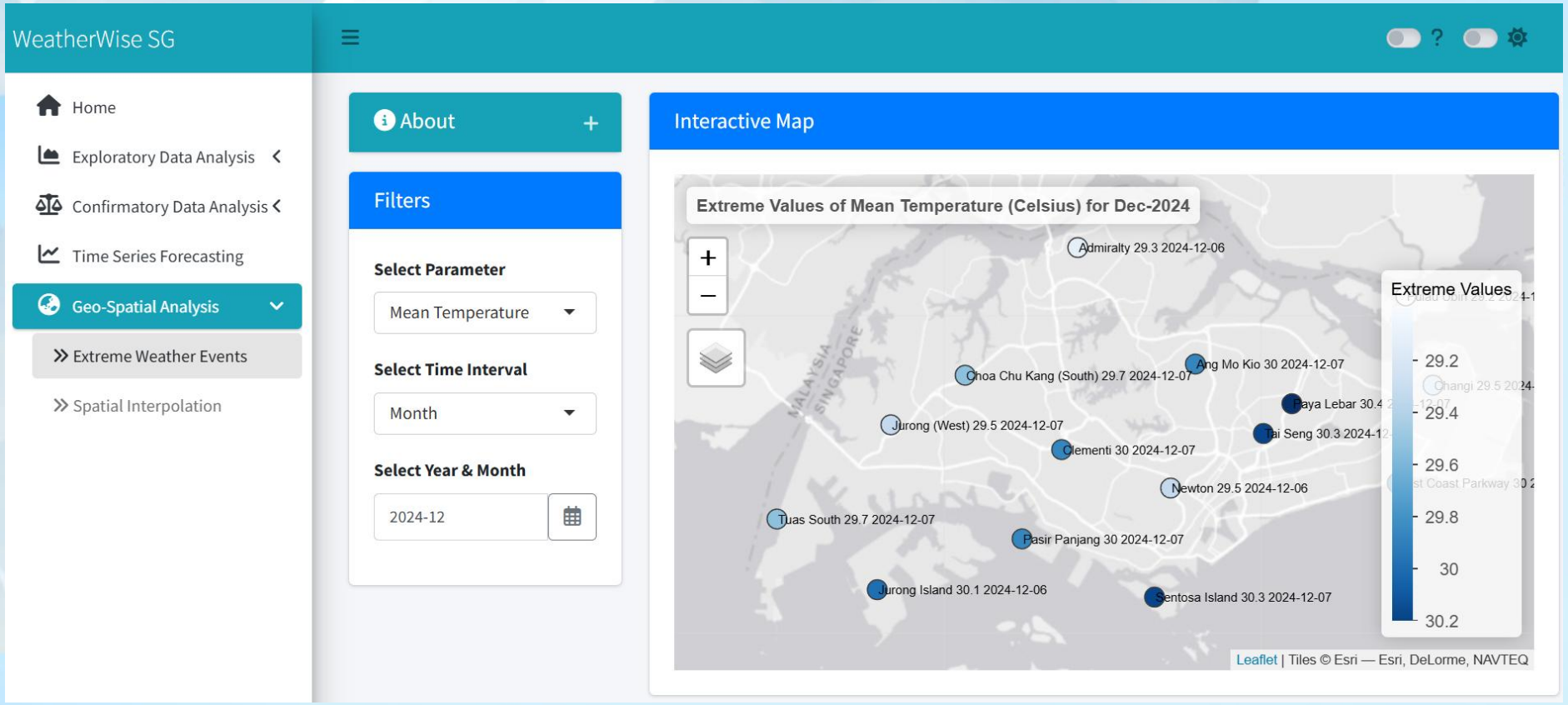
Forecasting Results

- Main Time Series Plot:
  - Visualizes the actual historical data alongside the forecasted future values.
- Static Forecast:
  - Shows a snapshot of non-interactive prediction for your selected forecast horizon.
- Interactive Forecast:
  - Hover over points to see precise forecast values.
  - Visualizes confidence intervals, historical data, and future predictions.

This is a quick prototyping App serves as a Minimum Viable Product (MVP) for demo. It is not designed for performance and scalability. Please note that Arima model performs poorly when historical daily/weekly data over several years are chosen, as it requires a lot of compute power & memory to forecast.

5.1 Geo-Spatial Analysis - Extreme Weather Events

This section provides and interactive map with that visualizes extreme weather events by station, allowing users to explore the hottest, wettest, or windiest days based on selected weather variables and time periods.



1. Select Parameter: Use the dropdown on the left to choose a weather parameter to explore: Daily Rainfall, Mean/Max/Min Temperature or Mean Wind Speed

2. Select Time Interval: Use the dropdown box to select the time interval of interest: Year or Month

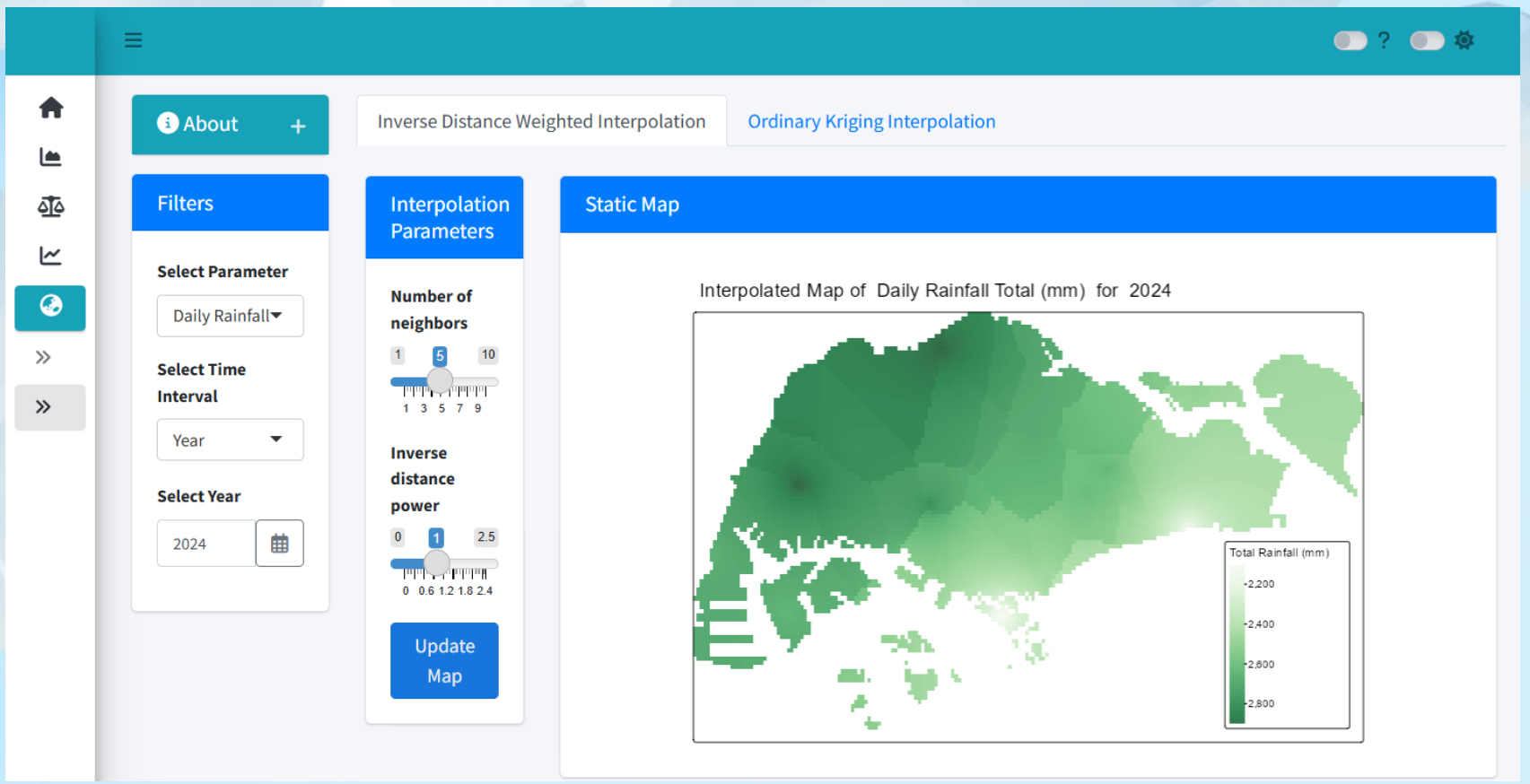
3. Select Year/Select Year & Month: Based on the Time Interval selected in 2., use the Date picker to select the Year or Month to view extreme events by station

4. Extreme Weather Events Map: Allow zooming in and out, enabling users to focus on specific areas or have an overview of extreme weather events across Singapore for the selected weather parameter and time period

5.2 Geo-Spatial Analysis: Spatial Interpolation – Inverse Distance Weighted Method

The section provides users with a map visualization of interpolated weather measurements at locations where no station/direct measurements are available. This would enable a view of continuous weather surfaces that reveal localized variations and pattern across Singapore.

The first tab plot the map using **Inverse Distance Weighted** Method (IDW), which is a simple and fast way to interpolate weather parameters at locations without direct measurements.



1. Select Parameter: Use the dropdown on the left to choose a weather parameter to explore: Daily Rainfall, Mean/Max/Min Temperature or Mean Wind Speed

2. Select Time Interval: Use the dropdown box to select the time interval of interest: Year or Month

3. Select Year/Select Year & Month: Based on the Time Interval selected in 2., use the Date picker to select the Year or Month

4. Select Number of Neighbors: select the number of nearest observations that should be used for interpolation

5. Select Inverse distance power (idp) as below:

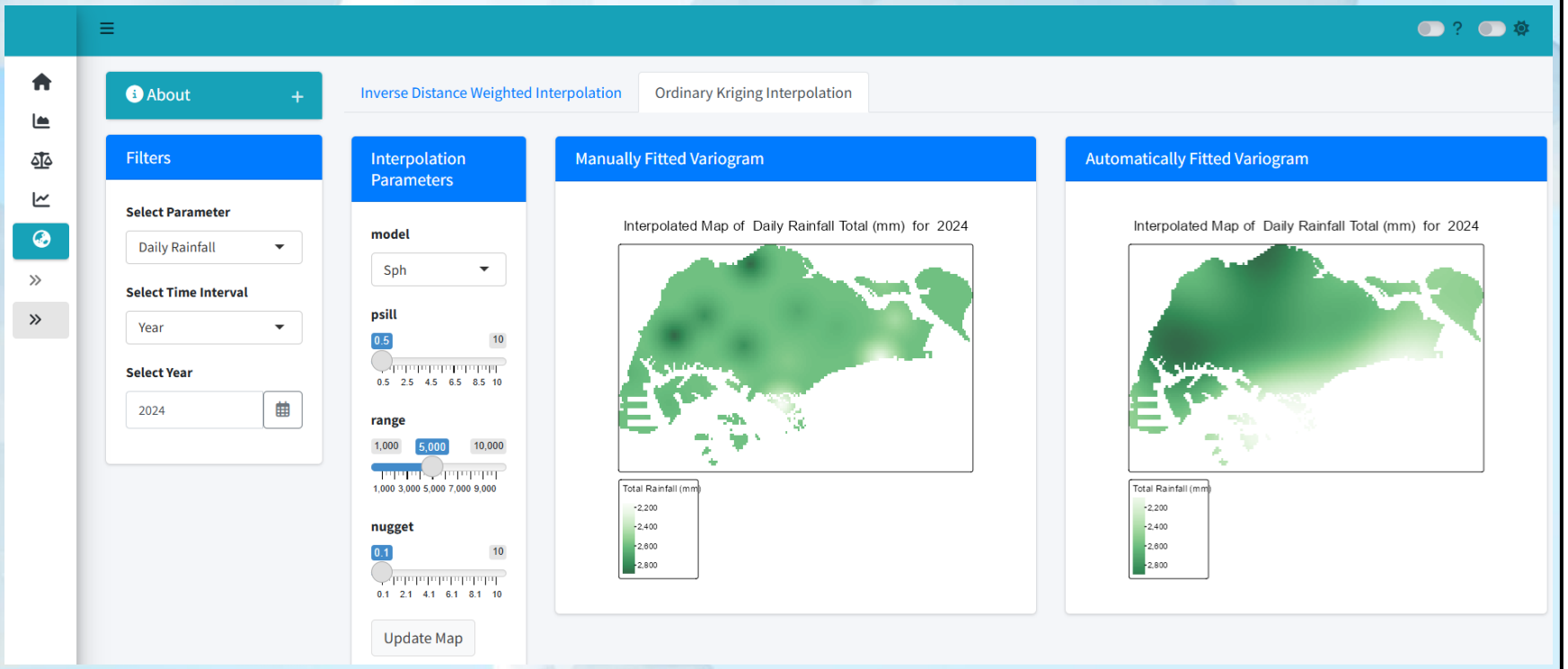
- idp = 0: All points are weighted equally, resulting in a simple average of values.
- idp = 1: Nearby points have more influence, but far points still contribute.
- idp > 1: Nearby points dominate, making the interpolation sharper.

6. Click **"Update Map"** to display the map

7. Interpolated Map: Visualize interpolated value of selected weather parameter across Singapore using the above selections.

5.2 Geo-Spatial Analysis: Spatial Interpolation – Ordinary Kriging Method

The second tab plots the map using **Ordinary Kriging** method , which accounts for spatial autocorrelation, improving prediction accuracy by considering both distance and statistical relationships between data points. This results in more reliable and precise interpolations, especially in areas with sparse weather stations.



1. Select Parameter: Use the dropdown on the left to choose a weather parameter to explore: Daily Rainfall, Mean/Max/Min Temperature or Mean Wind Speed.

2. Select Time Interval: Use the dropdown box to select the time interval of interest: Year or Month.

3. Select Year/Select Year & Month: Based on the Time Interval selected in 2., use the Date picker to select the Year or Month.

4. Select Variogram parameters: model, psill, range, nugget.

5. Click **"Update Map"** to display the map.

6. Manually Fitted Variogram Map: Visualize interpolated value of selected weather parameter using manually fitted variogram as selected in 4.

7. Manually Fitted Variogram Map: Visualize interpolated value of selected weather parameter using automatically fitted variogram. The parameters in 4 are not considered.