WeatherWise Singapore

Ee Fook Ming   Nguyen Bao Thu Phuong   Shreya Agarwal

SMU SINGAPORE MANAGEMENT UNIVERSITY

ISSUES & PROBLEMS

Singapore's tropical climate creates frequent and intense weather variations, which significantly impact sectors like urban planning, insurance, agriculture, and infrastructure. However, current weather platforms fall short in supporting data-driven decision-making due to:

- **Basic Functionality**: Existing dashboards offer only real-time observations and short-term forecasts.
- **No Advanced Analytics**: Lack of tools like trend forecasting, anomaly detection, and predictive modeling.
- **Limited Historical Insights**: Inability to analyze long-term climate trends or generate custom comparisons.

MOTIVATION

Existing dashboards lack deep analytical capabilities, trend forecasting, and business-context integration. Given Singapore's tropical climate and its impact on different sectors, there is a need for a comprehensive platform offering statistical insights, spatial visualizations, and predictive modeling.

Our project tackles this challenge by offering a platform with rich analytics, interactive visualizations, and predictive tools tailored to Singapore's context. This empowers businesses, researchers, and policymakers to make timely, strategic, and informed decisions, contributing to national climate resilience and smarter urban development.

APPROACH

DATASET

- Daily historical weather records scraped from weather.gov.sg, chose 15 stations with minimal missing data. Missing values imputed using a 5-day backward and forward moving average
- Additionally, Master Plan Subzone Boundary and weather station location data were incorporated for spatial analysis.

TECHNIQUES

1. Time Series Data Analysis: EDA, CDA

- Exploratory and comparative data analysis to uncover trends in temperature, rainfall, and wind across regions and seasons.
- Visualization of seasonal/weather trends and inter-region comparisons.
- Statistical tests for station-wise comparison across different time intervals and parameters.

2. Time Series Forecasting

- STL: Split data into trend, seasonal and residuals.
- ACF/PACF: Identify lags to set ARIMA/SARIMA p, q parameters.
- ADF Test: Check stationary and fine tune ARIMA/SARIMA (p,d,q)(P,D,Q).
- Benchmarking: Compare with other models – Prophet and ETS; evaluate with MAE, RMSE, MAPE.

3. Geospatial Analysis

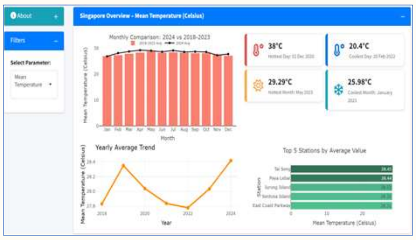
- Plot extreme weather events by station on Singapore map
- Interpolate weather conditions at unmeasured locations based on nearby data points using Inverse Distance Weighted (IDW) and Kriging methods

RESULTS

1. Exploratory Data Analysis

1 Overview

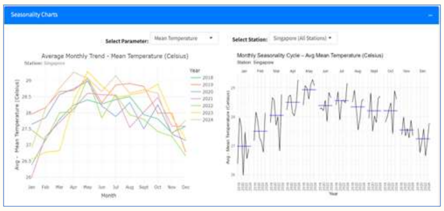
1.1. Monthly Comparison View - (2018–2023 averages) bars overlaid with a 2024-line chart compares current weather to past *(e.g: Apr, July 2024 mean temp. higher than past average value)*



1.2. Extreme Statistics - Locate extreme weather events - (Hottest/ Coolest/ Rainiest day/month etc.)

1.3. Annual Trend Chart - Yearly changes across Singapore to identify long-term trends, such as rising temperatures/ rainfall/wind patterns *(e.g,: 2024 highest and 2022 lowest average mean temperature)*

1.4. Top 5 Stations Bar Chart - Stations with the highest or lowest values *(e.g., top 5 for total rainfall or bottom 5 for average min. temperature)*, for rapid spatial comparison.



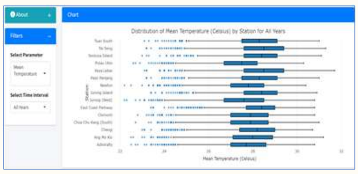
2 Seasonality Analysis

2.1. Monthly Trend - Displays monthly averages/total(rainfall) for each year to detect deviations in specific years and observe potential shifts in seasonal cycles *(e.g., Jan 2021 has lowest mean temperature)*.

2.2. Seasonality Cycle - Month-specific variations across years. Subplots show value changes within a specific month over time. *(e.g., Q2, Q3 have high mean temperature)*.

3 Station Comparison

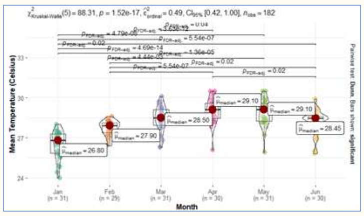
3.1. Station-wise Distribution - Distribution of the selected parameter across stations and time interval - median, spread, and outliers for each station. *(e.g – Pulau Ubin has lowest median for mean temperature)*



3.2. Month-by-Station Heatmap - Weather parameter variation across stations and months for different years. Color intensity indicates average values *(e.g., May has highest mean temperature of all stations)*.

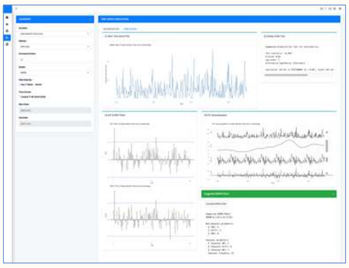
2. Confirmatory Data Analysis

Statistical Comparison (Stations & Time) - statistically compare weather parameters across stations or time intervals *(e.g. compare mean temperature between 6 months)*.



- **Normality Testing -** Checks if data follows a normal distribution using tests like Anderson-Darling.
- **Group Comparison -** Apply tests (e.g., Kruskal-Wallis, ANOVA) to detect significant differences across stations or time periods.
- **Outputs -** Box-violin plots and statistical summaries

3. Time Series Forecasting



Smart Forecasting for Climate Variables

- Forecasts rainfall, temperature and wind at daily, weekly, or monthly levels
- Supports flexible date range selection for exploratory analysis

Multiple Forecasting Models Compared

- Models used: **ARIMA**, **ETS**, and **Prophet**
- Automatically tuned models with visual diagnostics
- Dynamic comparison across models using MAE, RMSE, and MAPE

Model Evaluation Panel

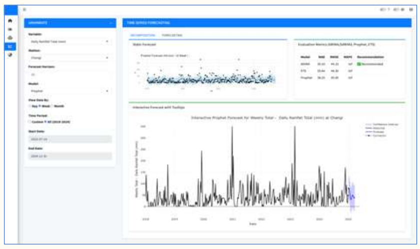
- Transparent comparison of model accuracy
- Automatically recommends the best model
- Clearly displays where each model excels (ARIMA for stability, Prophet for seasonality, etc.)

Interactive Forecasting Visualization

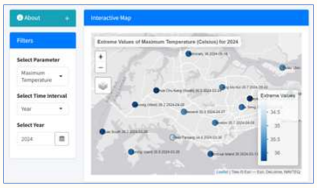
- Zoomable and tooltip-enabled plots with confidence intervals
- Dotted connector shows transition from historical to forecast
- Forecasts displayed alongside confidence bands (80% & 95%)

Integrated Stationarity Check

- Augmented Dickey-Fuller test to assess if series is stationary
- STL decomposition reveals trend and seasonal patterns
- ACF & PACF used to identify autoregressive behaviour



4. Geo Spatial Analysis

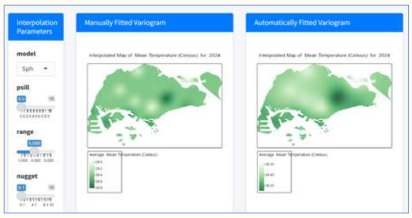


1. Interactive Extreme Weather Map - An interactive map with zoom option that visualizes extreme weather events by station, allowing users to explore the hottest, wettest, or windiest days based on selected weather variables and time periods.

2. Spatial Interpolation –

- Visualize interpolated map based on user selection of weather measurements, time horizon and interpolation parameters
- Provide map visualizations of two different interpolation methods for comparison

2.1. Inverse distance weighted method: a simple and efficient interpolation method for a quick overview



- **2.2. Kriging method:** Accounts for spatial autocorrelation, improve prediction accuracy by considering both distance and statistical relationships between data points.
- Allows users to visualize and compare interpolation of weather measurements using both manual and automatically fitted variogram.

FUTURE WORK

Future Work: Toward a Next-Gen Weather Platform

- **Multimodal Data Fusion –** Combine ground-based station data with satellite imagery, radar, and IoT sensors for enhanced spatial coverage and accuracy.
- **Multivariate & AI-Enhanced Forecasting -** Incorporate temperature, wind, humidity, and external climate signals into ML and deep learning models for improved accuracy. Leverage geographically weighted predictive modelling
- **Integrated Decision Support Tools -** Deliver forecasts through APIs and interactive dashboards for government agencies, urban planners, and climate-sensitive sectors.
- **Real-Time Learning & Auto-Retraining -** Enable continuous model updates from live station feeds to improve accuracy and adapt to changing patterns.