

The background is a marbled pattern. The left side features dark, swirling colors of blue, green, and brown. A vertical line divides the image, and the right side features bright, swirling colors of yellow, orange, and blue.

Quality Control for Rainfall Extremes

Kate
Jarryd
Fin

Quality Control

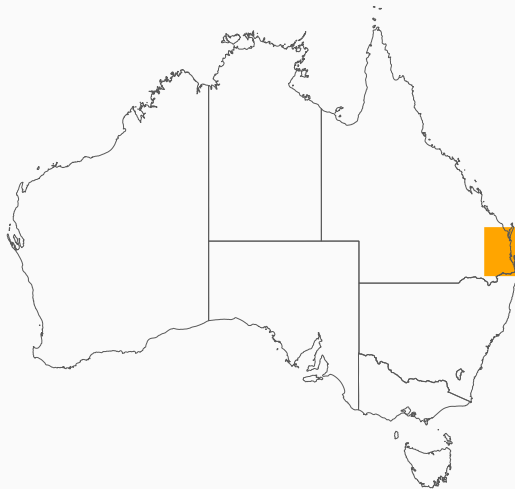
Automated testing

Important for large scale data handling of observations and analysis of extreme events

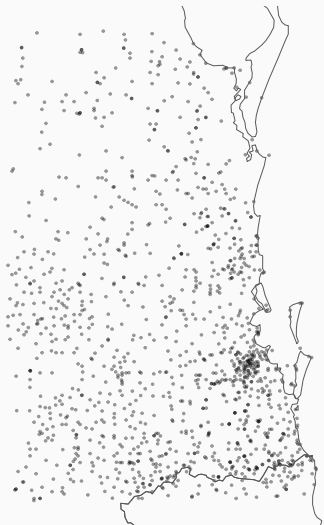
Sequential checks

- Identifies duplicated observations
 - Checks for abnormally large observations
 - Failed spatial consistency check
 - Failed temporal consistency check
 - Temperature too warm for snow
- and so on.

Southeast Queensland



Stations



- 1297 Stations
- Available data ranges from 1750 to 2024
- On average, 102 years of data are available
- 26% of stations have more than 270 years of data

- 1 True extremes values dropped as outliers
- 2 Extreme missings: Data are systematically missing because it is an extreme event
- 3 Tagged accumulations: rainfalls are not recorded and accumulated to the next (next) day with a tag
- 4 Untagged accumulations: rainfalls are not recorded and accumulated to the next without a tag
- 5 Extreme Zeros: Data entries are 0 when it is an extreme event

True extremes values dropped as outliers

- First step in an analysis is commonly to filter the data to remove observations flagged as outlier, but
- **Current outlier test** assumes normality – True extreme observations are being incorrectly flagged as outliers
- People are removing true extreme values from their analysis.

True extremes values are not outliers

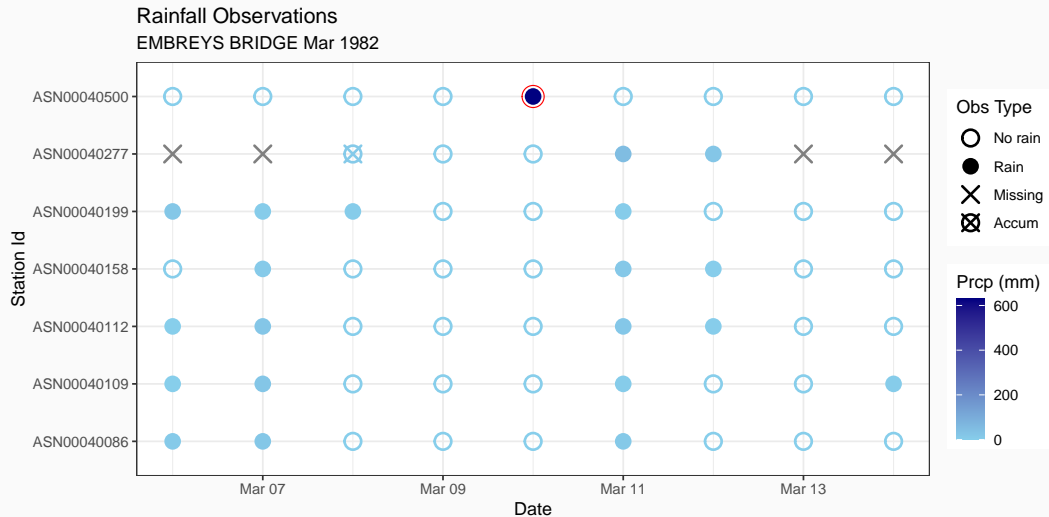
Spatial consistency check

Check if the observation of an “outlier” station is consistent with its neighbouring stations’ observed values.

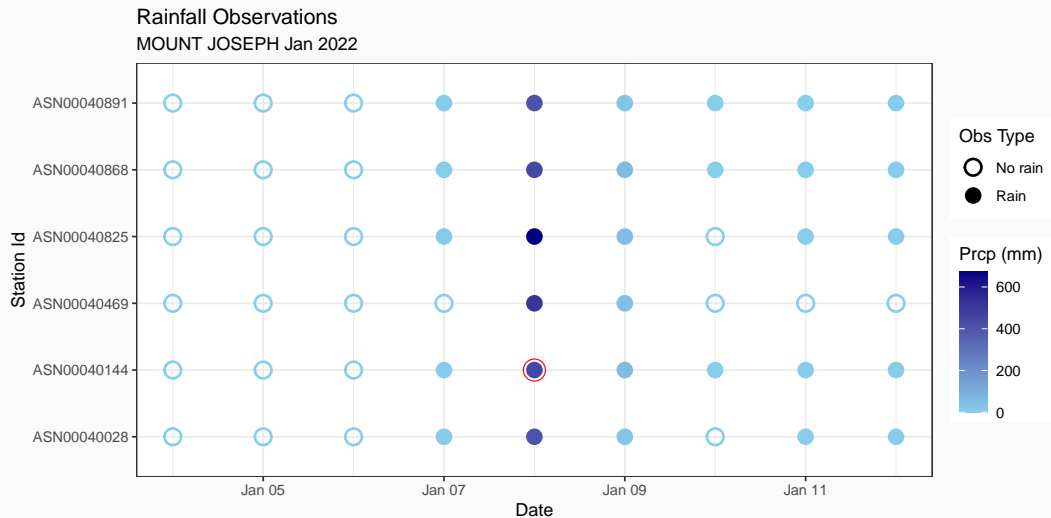
Southeast Queensland

- 60 observations were flagged as outliers and have sufficient neighbours for spatial consistency check
- 37 observations pass the spatial consistency check and should be extreme values and not outliers
- Outlier test has a false-positive rate of 62%

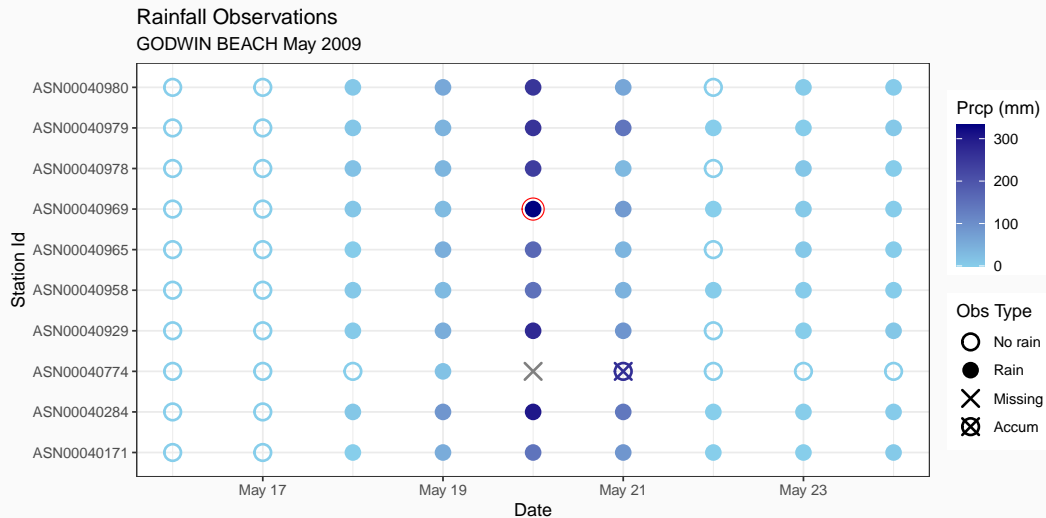
Example 1: Correct outlier flag



Example 2: False outlier (spatially consistent)



Example 3: False outlier (spatially inconsistent)



Solution

Where possible a spatial consistency check should be used,
or

just don't remove outliers.

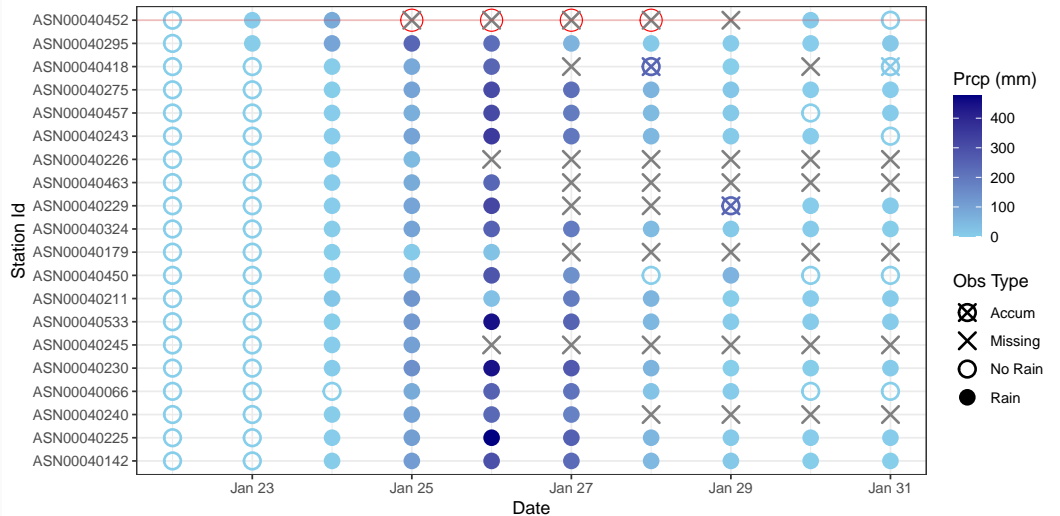
Determine extreme values

For a specific station

- 1 Find neighbouring stations within a 20 km radius
- 2 Take the 99-th percentile of the daily rainfalls on rainy days of these neighbouring stations as the threshold
- 3 On a particular day, if any station among those neighbouring stations has a rainfall record larger than the threshold, we determine that the centre station experiences extreme rainfall on this day

The procedure is repeated for every station.

Extreme missings



Extreme missings

Southeast Queensland

- 6396 missing observations are flagged as extreme values
- Ranging from 1750-02-01 to 2023-12-24
- From 753 stations

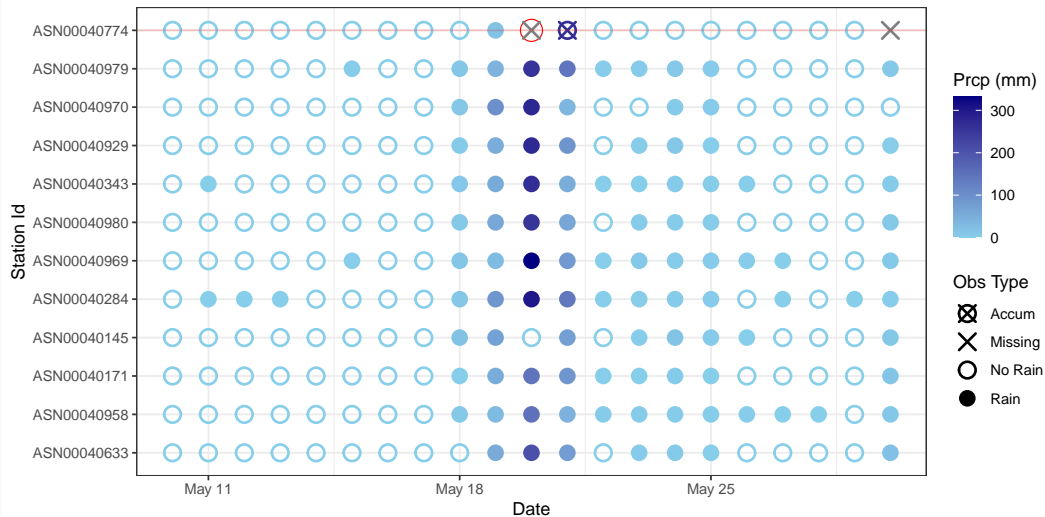
Tagged accumulation

Essentially the same as extreme missings with an additional accumulation tag.

Southeast Queensland

- 5804 missing observations are flagged as extreme values
- Ranging from 1892-04-02 to 2023-11-21
- From 622 stations

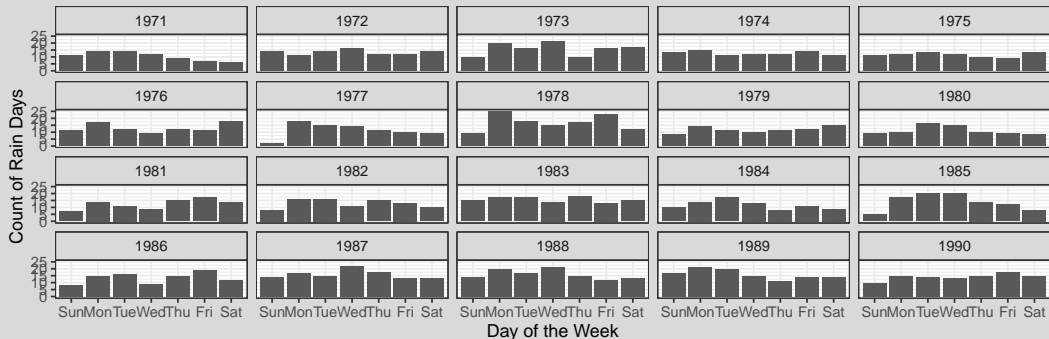
Tagged accumulation



Untagged accumulations

Rainfall observations on Sundays are often recorded on Mondays without being noted by the observer in the records

Sunday-Monday untagged accumulations



Extreme zeros

Southeast Queensland

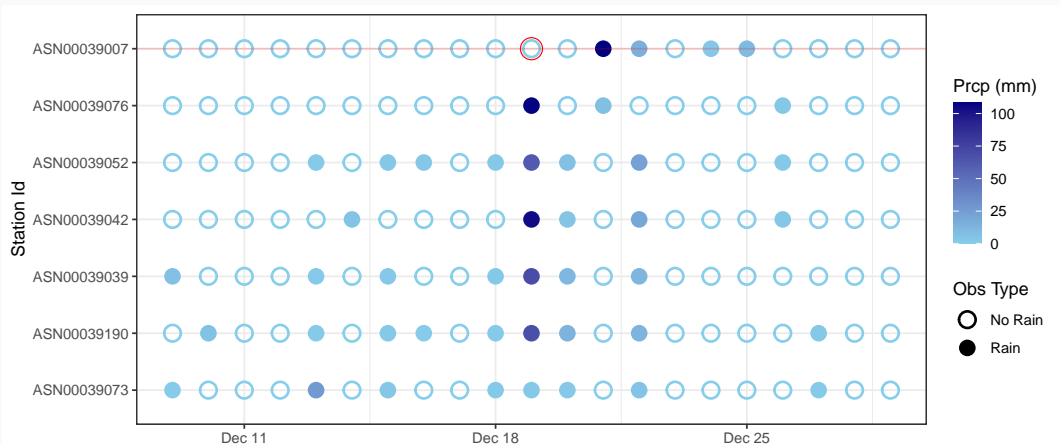
- 14287 missing observations are flagged as extreme values
- Ranging from 1750-03-15 to 2023-12-26
- From 1145 stations

Issues

Usually combined with other problems, especially untagged accumulation, which makes it difficult to further categorise different types of extreme zeroes.

Extreme zeros with untagged accumulation

The zero value that is identified as extreme value is silently accumulated to days later



Extreme zeros with untagged accumulation

The zero value that is falsely identified as extreme value because of untagged accumulation in a neighbouring station

