

Assemble Data

Station observation

- **Network:** GHCN-D stations
- **Variables:** Daily precipitation, temperature, latitude, longitude, elevation
- **Record:** 1991-2012

Reanalysis datasets

- **Regional:** EMDNA, PRISM, RDRS
- **Global:** ERA5, CHIRPS, MERRA-2

Spatial attributes

- **Elevation:** Sampled from MERIT DEM (≈ 90 m)
- **Domain mask:** GLB boundary + 100 km buffer (extract reanalysis grids)

Grid-Station Mapping

Inputs for mapping

- Station lat, lon, elev (from MERIT DEM).
- All reanalysis grid centroids within domain mask.

Candidate mapping methods

- **Neighbor search (feeds the regression):** 8-12 grids within 25 km to 100 km radius

Grid Value at Station

- **Locally Weighted Regression:** Fit
" $\beta_0 + \beta_1 \text{Lat} + \beta_2 \text{Lon} + \beta_3 \text{Elev} + \epsilon$ " on chosen grids (distance-weighted)
- Plug station's attributes to predict reanalysis value at the station location
- Repeat for each dataset & neighbor option

Performance metrics

Statistical metrics*

- MBE
- RMSE
- Std
- CC
- Index of agreement (d)

* Station vs mapped grid series

Climatic extremes indices*

- **Precipitation:** RX1Day, RX5Day, R95p, R99p, CWD, Drydays, Wetdays
Temperature: TN10p, TX90p, TNn, TXx, CSDI, WSDI.

* ETCCDI subset

Temporal & spatial summaries

- **Annual & seasonal aggregates:** DJF, MAM, JJA, SON
- **Visual Diagnostics:** Maps, CDFs, Taylor diagrams, Boxplots.

Evaluation & Guidance

Overall Skill Scores

- Aggregate the statistical metrics into a composite score for each dataset.
- Rank datasets separately for precipitation & temperature.

Recommended Strategy

- **Primary choice:** dataset with highest statistical metric score.
- **Extremes focus:** dataset with top-ranked extremes indices.
- **Spatial gaps:** Highlight parts of the basin where another dataset outperforms the primary choice.

Actionable Outputs

- Final pick for each variable/season
- Map layers ready for modelling & planning.