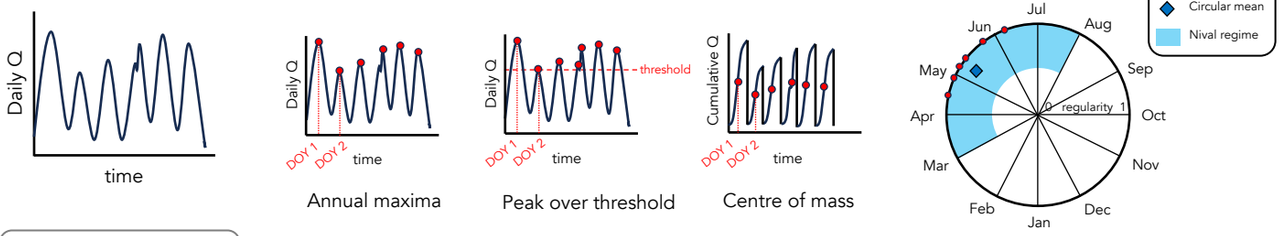
# FROSTBYTE

## Forecasting River Outlooks from Snow Timeseries: Building Yearly Targeted Ensembles

## 1. Regime classification

*FROSTBYTE/notebooks/1_RegimeClassification.ipynb*

Circular statistics are computed on daily streamflow station observations to identify nival basins. We then sub-select nival basins with at least 20 years of overlapping SWE and Q data.

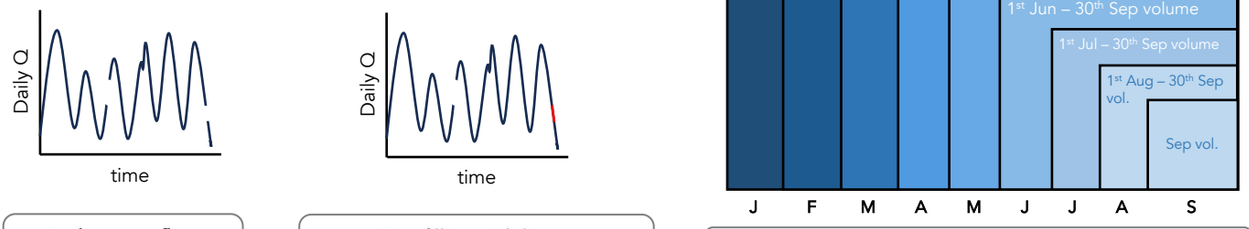



Daily streamflow station observations → 3 streamflow peak metrics → Circular statistics

## 2. Streamflow pre-processing

*FROSTBYTE/notebooks/2_StreamflowPreprocessing.ipynb*

Daily streamflow timeseries are aggregated to seasonal volumes.




Daily streamflow station observations → Gap filling with linear interpolation for gaps ≤ 15 days → Aggregations for target periods (example for 1 year)

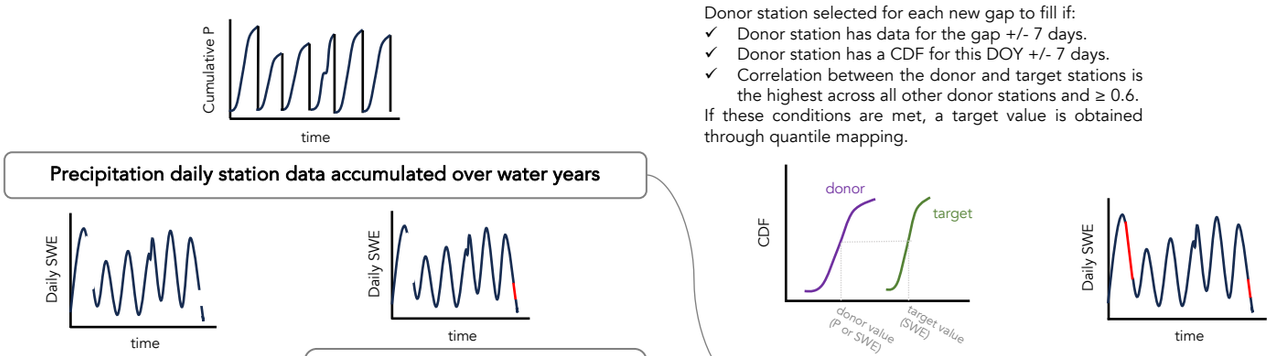
## 3. SWE pre-processing

*FROSTBYTE/notebooks/3_SWEPreprocessing.ipynb*

SWE station timeseries are gap filled using a combination of linear interpolation and quantile mapping from neighbour SWE and P stations.

Donor station selected for each new gap to fill if:
- ✓ Donor station has data for the gap +/- 7 days.
- ✓ Donor station has a CDF for this DOY +/- 7 days.
- ✓ Correlation between the donor and target stations is the highest across all other donor stations and ≥ 0.6.

If these conditions are met, a target value is obtained through quantile mapping.

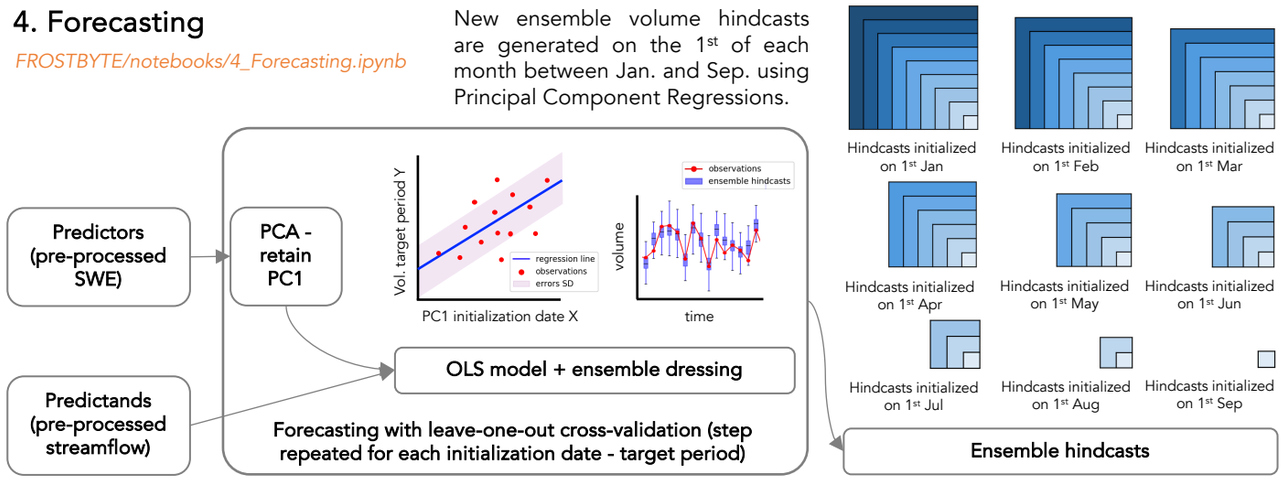



Daily SWE station data → Linear interpolation for gaps ≤ 15 days → Quantile mapping → Gap-filled SWE

## 4. Forecasting

*FROSTBYTE/notebooks/4_Forecasting.ipynb*

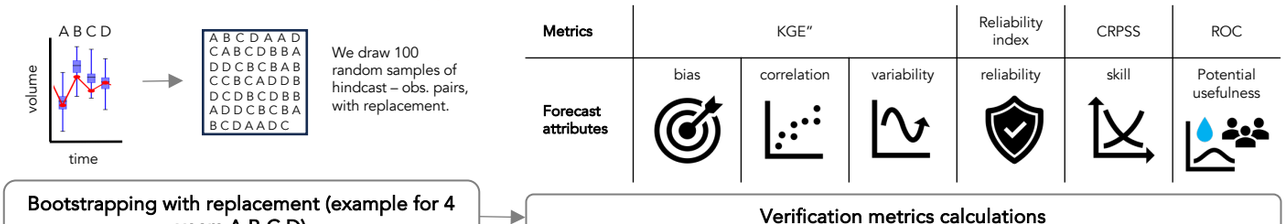
New ensemble volume hindcasts are generated on the 1st of each month between Jan. and Sep. using Principal Component Regressions.




## 5. Hindcast verification

*FROSTBYTE/notebooks/5_HindcastVerification.ipynb*

Deterministic and probabilistic verification metrics are computed to measure various aspects of the forecast quality, with bootstrapping.




Bootstrapping with replacement (example for 4 years A B C D) → Verification metrics calculations