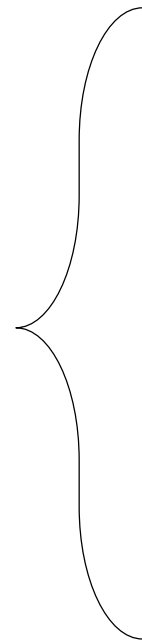HEDA Package {
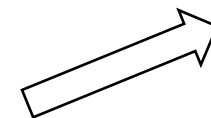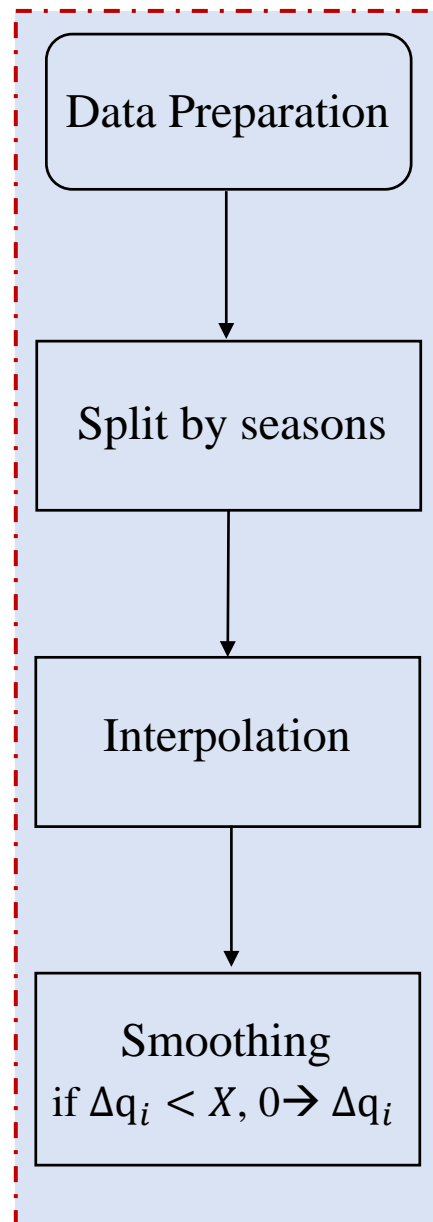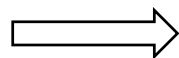
1.HEDA_Tidy

2.ReversalCount

3.clean_position

4. clean_Spt

5. clean_conectD

6.HPK_metrics

7.HPK_plot

Function 1:HEDA_Tidy

Input 1.csv

Input 2.csv

Input n.csv

Data Preparation

Split by seasons

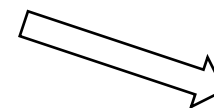Interpolation

Smoothing
if $\Delta q_i < X, 0 \rightarrow \Delta q_i$

Input 1_sm.csv

Input 2_sm.csv

Input n_sm.csv

Input 1_wt.csv

Input 2_wt.csv

Input n_wt.csv

Inputs processed in loop. Apply function doesn't work well, still slow in this case.

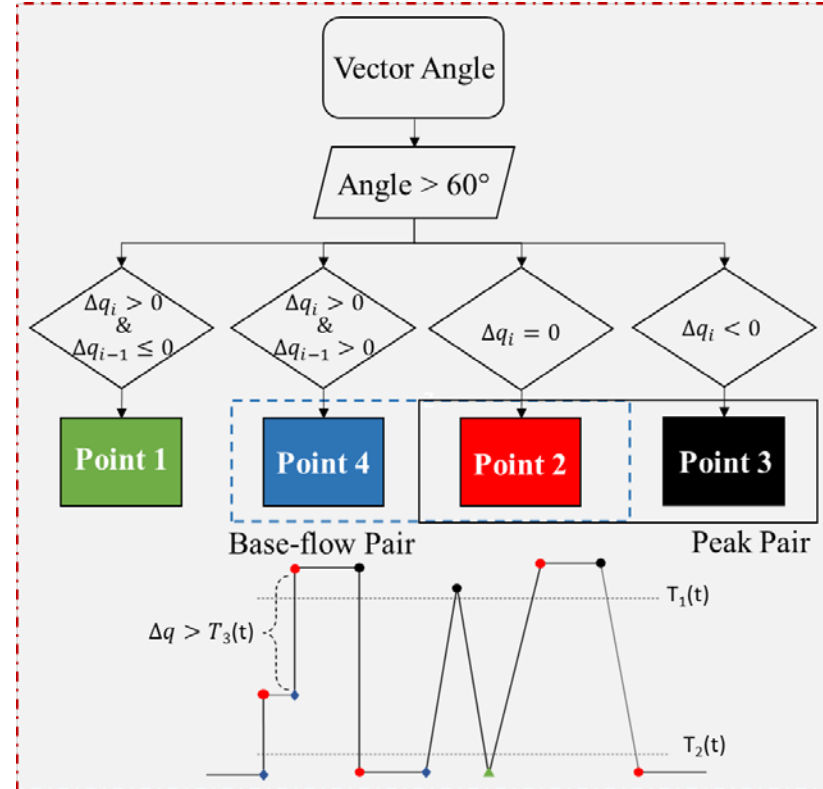| location_id	datetime	parameter_value
text	2020-12-02 1:00:00	(number)

Input 1.csv

- One sample of the data frame of input data.
- 3 columns.
- location id stores the gauge id information, 1 gauge 1 ID
- datetime is time format. Year-month-day Hour:00:00
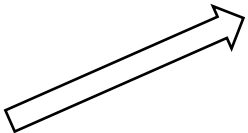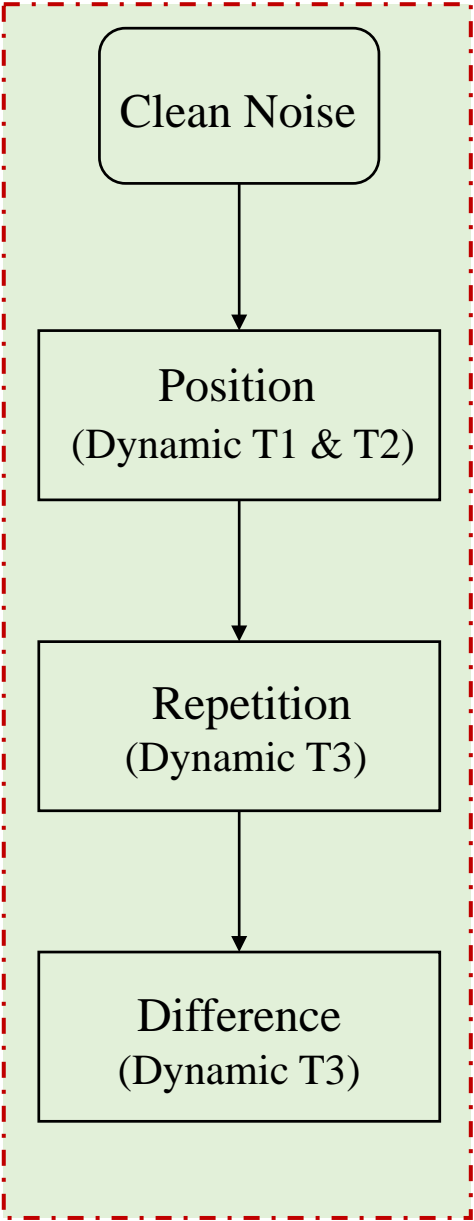- Parameter_value is discharge. Float format

2.ReversalCount

Input 1_sm.csv

Input 2_sm.csv

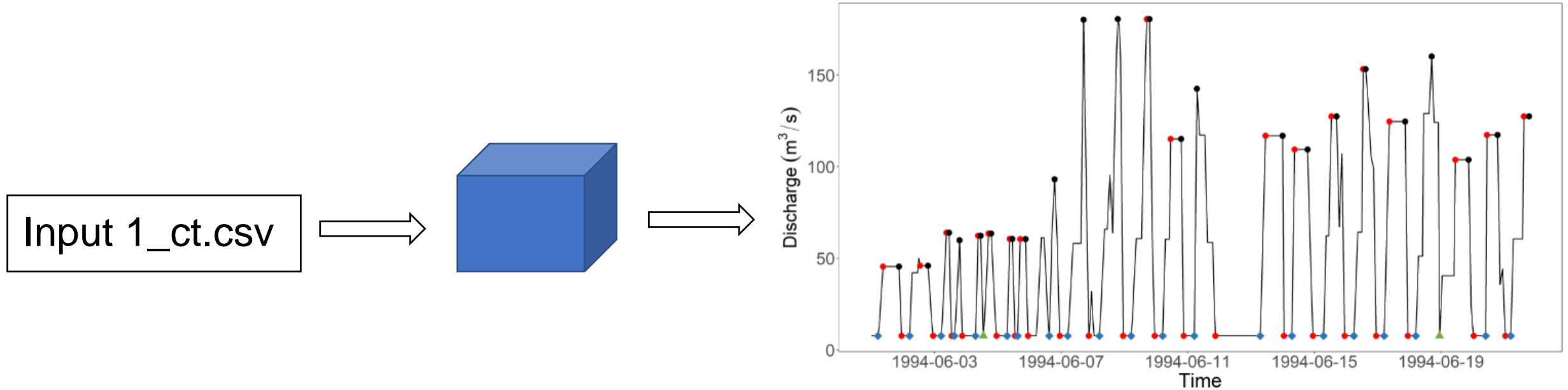Input n_sm.csv

3.clean_position
4. clean_Spt
5. clean_conectD

Clean Noise

Position
(Dynamic T1 & T2)

Repetition
(Dynamic T3)

Difference
(Dynamic T3)

Input 1_ct.csv

Input 2_ct.csv

Input n_ct.csv

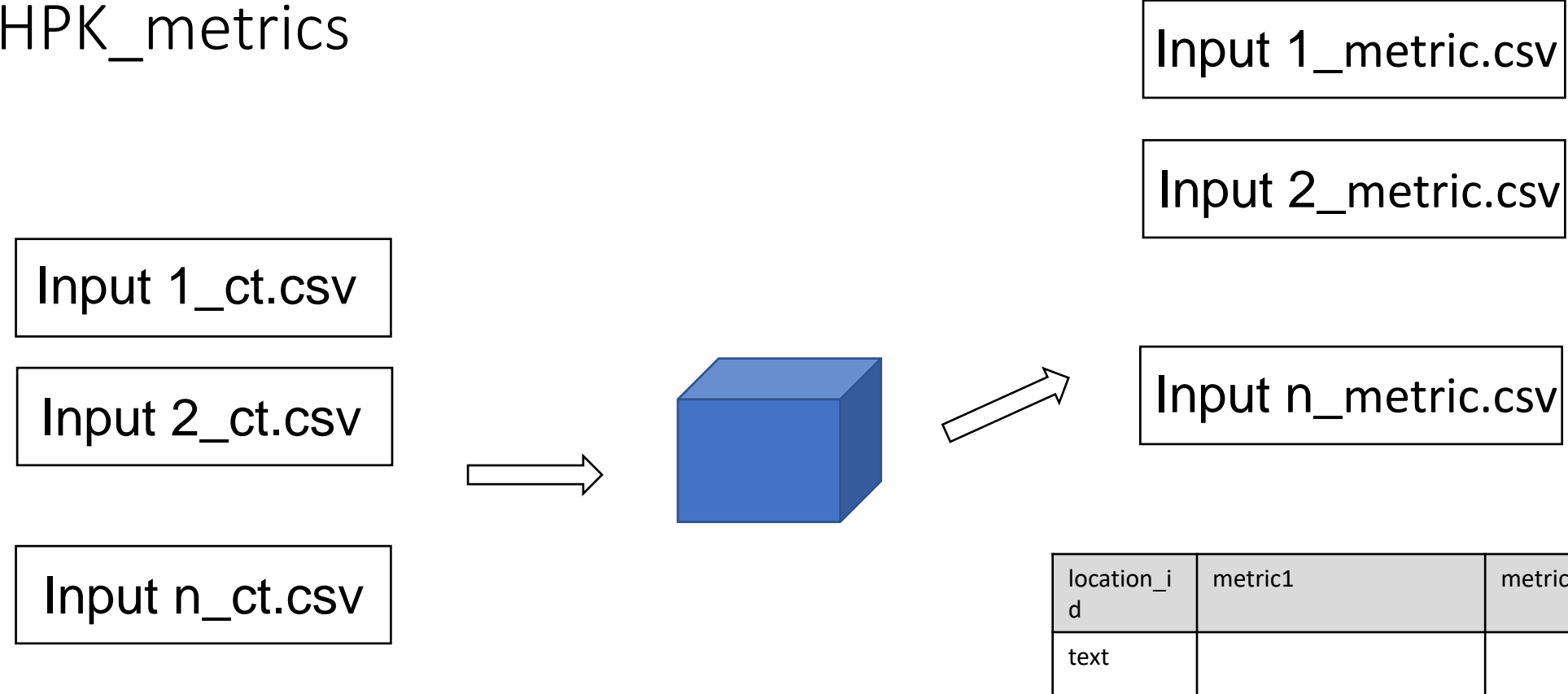| location_id	datetime	parameter_value	
text	2020-12-02 1:00:00	(number)	

It has similar structure as input but with more columns

7.HPK_plot

Input 1_ct.csv

Visualization of reversal count. A sample plot

6.HPK_metrics

Input 1_ct.csv

Input 2_ct.csv

Input n_ct.csv

Input 1_metric.csv

Input 2_metric.csv

Input n_metric.csv

location_id	metric1	metric2	...
text			

For each file, output file has 16 columns and one row.