On the visual inspection of streamflow time series

Laurent Strohmenger, Guillaume Thirel

Preprint in Hydrology and Earth System Sciences
https://doi.org/10.5194/hess-2023-58

Laurent Strohmenger[1]
Guillaume Thirel[1]
and the Explore2 consortium
[1]Université Paris Saclay, INRAE, UR HYCAR, Antony, France
Contact: Laurent.Strohmenger@inrae.fr

On the visual inspection of streamflow time series
Distributions and impacts of non-natural flow records

Objectives

Large datasets of streamflow measurements are widely used to infer hydrological processes (frequency of floods and droughts, river flows regimes, ...).

Streamflow measurements may suffer from what we can consider as anomalies due to instrument failure, missing data interpolation, and human influences.

Visual inspection of streamflow time series by humans is highly recommended. Yet, no study has examined:
(1) how evaluators differ in detecting anomalies
(2) the temporal distribution of anomalies
(3) the influence of anomalies on hydrologic indicators

Methods

Large visual inspection of streamflow time series

- 611 stations in France labeled as "good quality"
- 43 evaluators (academic and operational hydrologists)
- each time series was analyzed by two evaluators

Hydrological indicators

QMNA5: annual min. monthly flow with a 5-year return period
VCN30$_5$: annual min. of a 30-day moving average of flow with a 5-year return period
Q1-99: 1st, 5, 50, 95 and 99th daily flow quantiles
QJXA10: annual max. daily flow with a 10-year return period

Results

(1) Subjectivity of the evaluators

- evaluators reported from 0.04 to 2.92 % (on median) of time series as anomaly, with an overall median of 0.7 % of time steps reported.
- evaluators agreed on only 12 % (overall median) of time steps reported as anomaly

(2) Temporality of anomalies

- annual frequency of anomalies decreases from 1976 to 2019
- anomalies more frequently reported during summer
- more anomalies reported for years with climatic events such as heatwaves and droughts (1976, 1978, 1985, 1989, and 2003)

(3) Impact of anomalies on hydrological indicators

- higher change rates on low-flow indicators (Q1, Q5, QMNA5)
- change rates < 1 % on high-flow indicators (Q95, Q99, QJXA10)

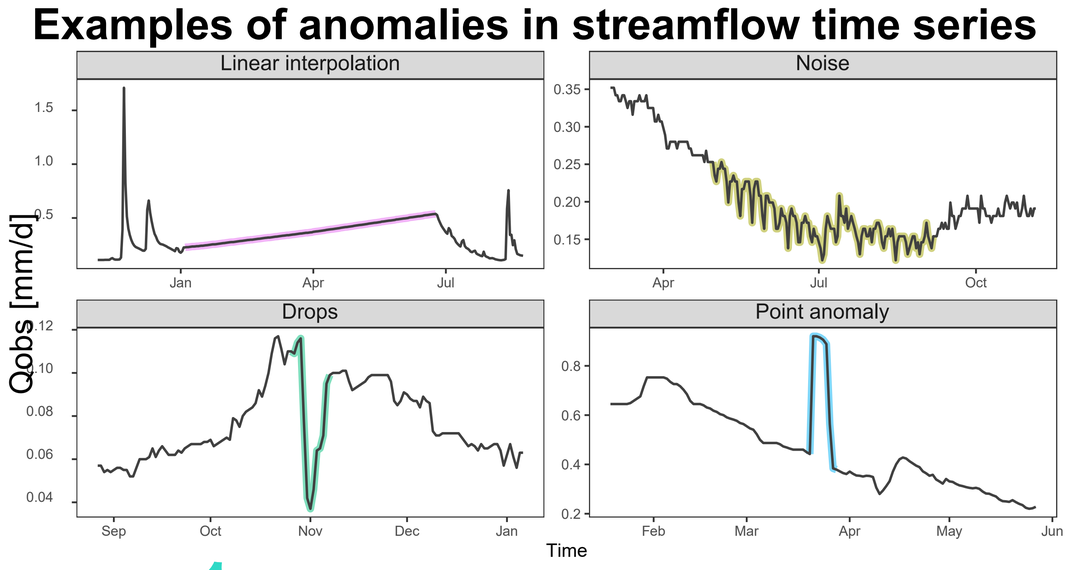
Conclusions

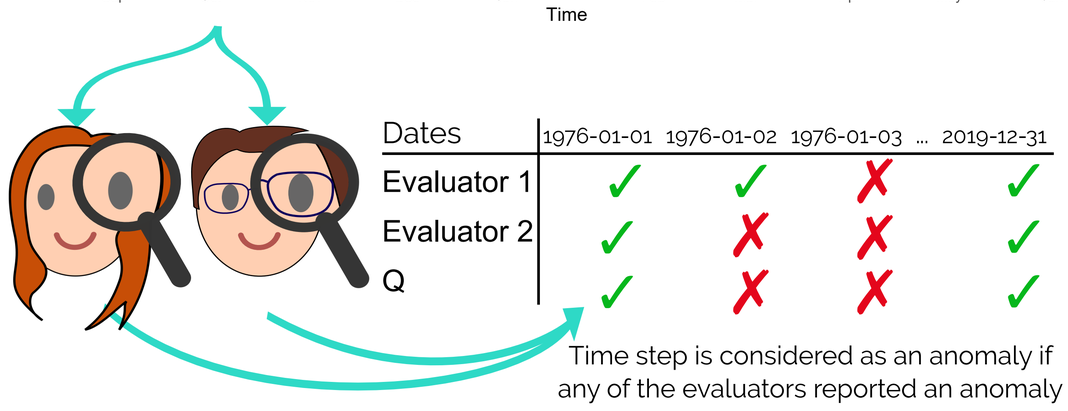
Detection of anomalies is highly dependent on the evaluator, raising questions on best practices for data cleaning.

The anomalies were more frequently reported in summer: challenges to monitor low flows, higher human influences.

An automatic detection of anomalies would be very valuable for cleaning streamflow time series, although its training raises questions.

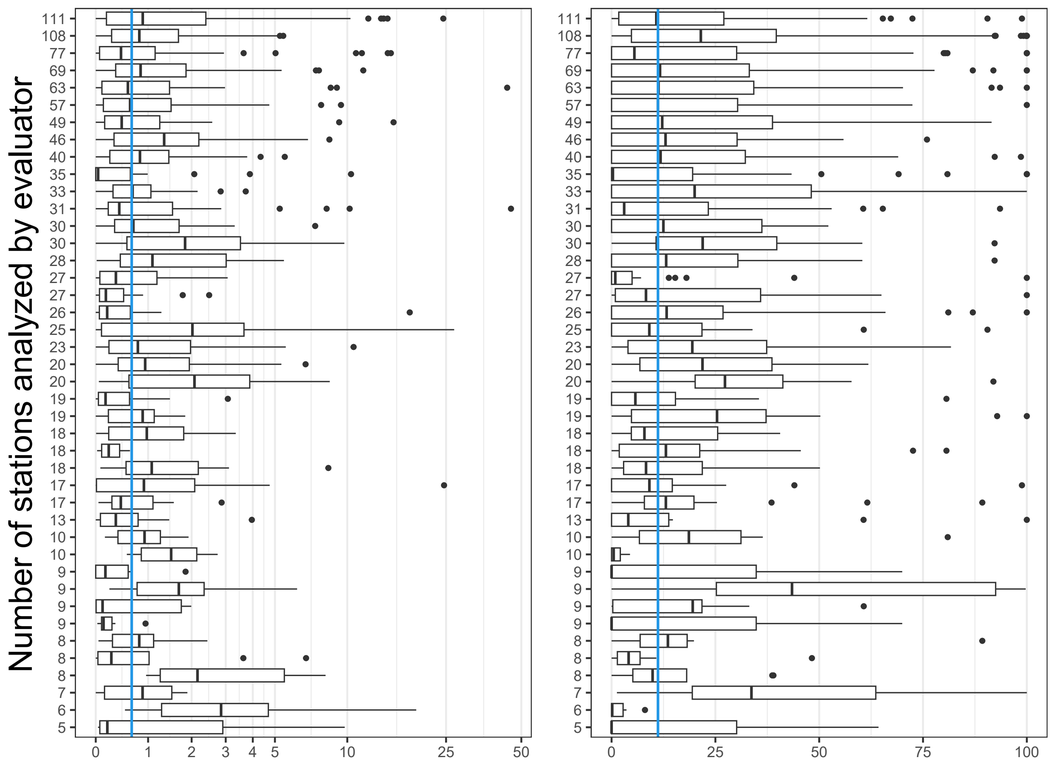
Examples of anomalies in streamflow time series





Time step is considered as an anomaly if any of the evaluators reported an anomaly

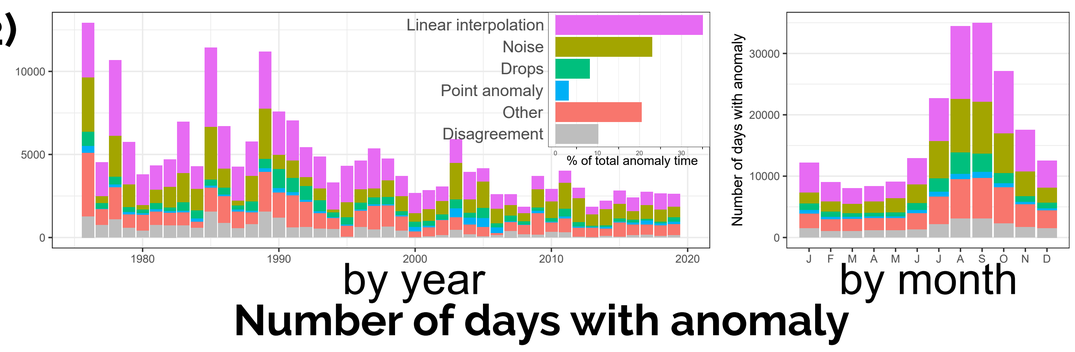
(1)



time reported as anomaly [% of time series]

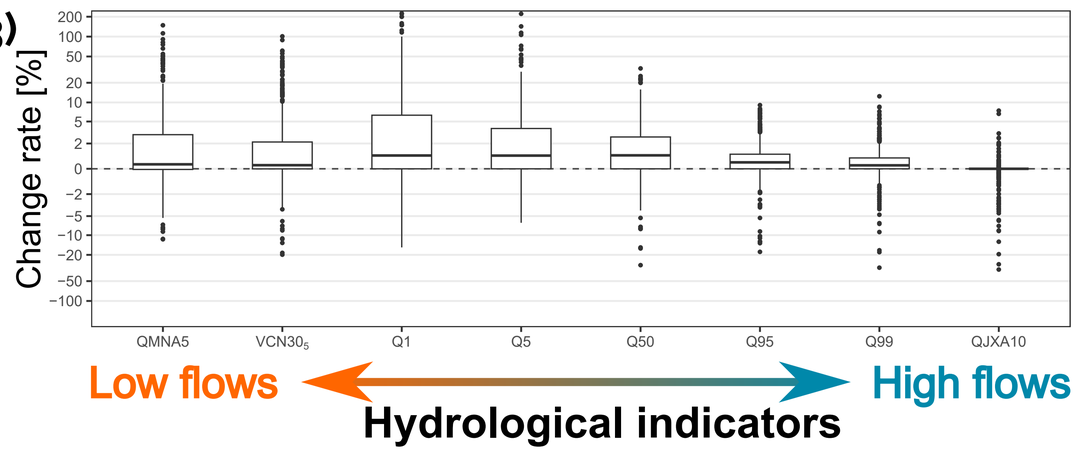
inter−evaluator agreement [%]

(2)



Number of days with anomaly

(3)



Île-de-France – Jouy-en-Josas – Antony

Laurent.Strohmenger@inrae.fr
1 Rue Pierre Gilles de Gennes
92160 Antony, France