Statistical Methods for Rare-Event Modelling under Missingness and Measurement Drift

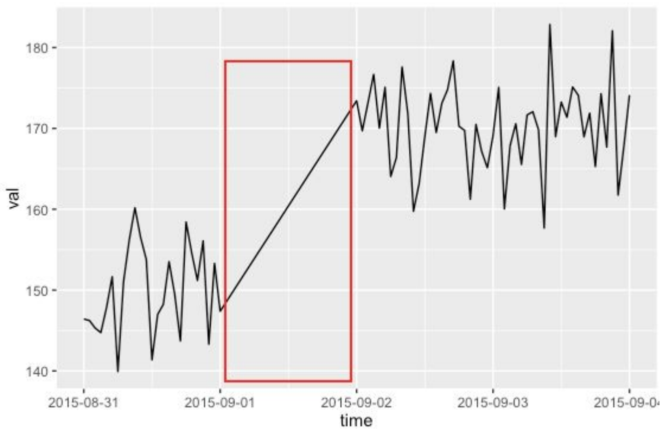
An Integrative Review for MSc Thesis Research

Prepared by: *Kateryna Kaplunova*

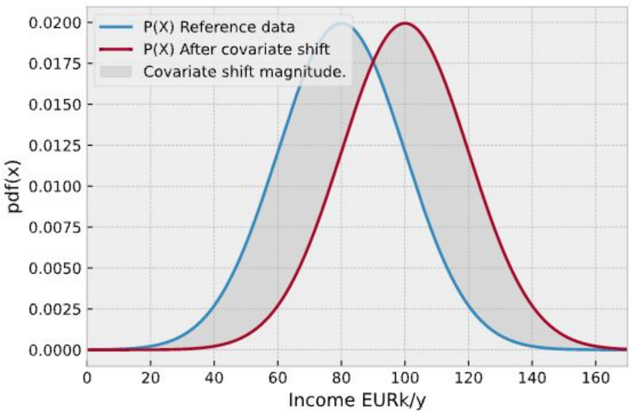
Fragile Tails: Two Data Hazards

Recall: Extreme-value inference is often based on only the top few percent of observations—sometimes less than one percent in large data sets.

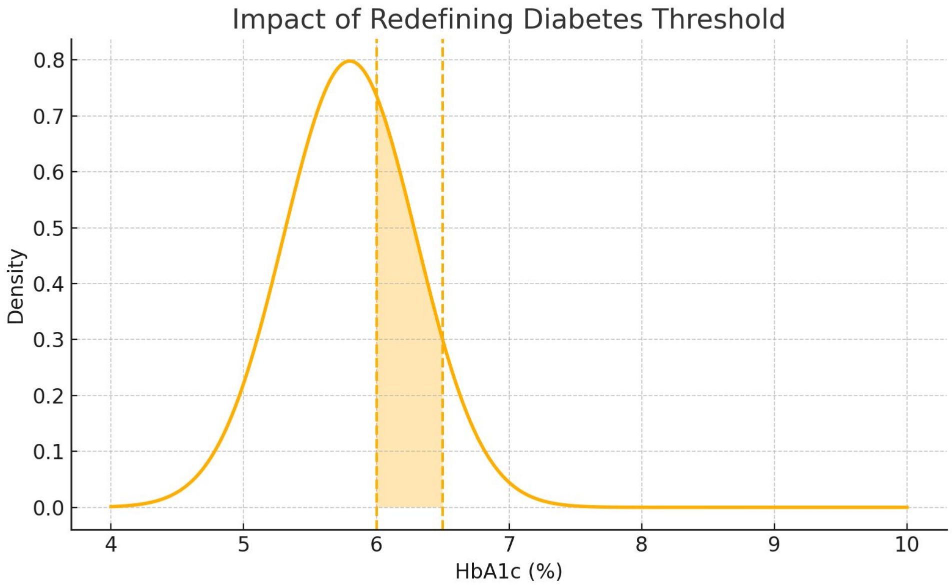
Hazard 1: Missing observations



Hazard 2: Measurement drift



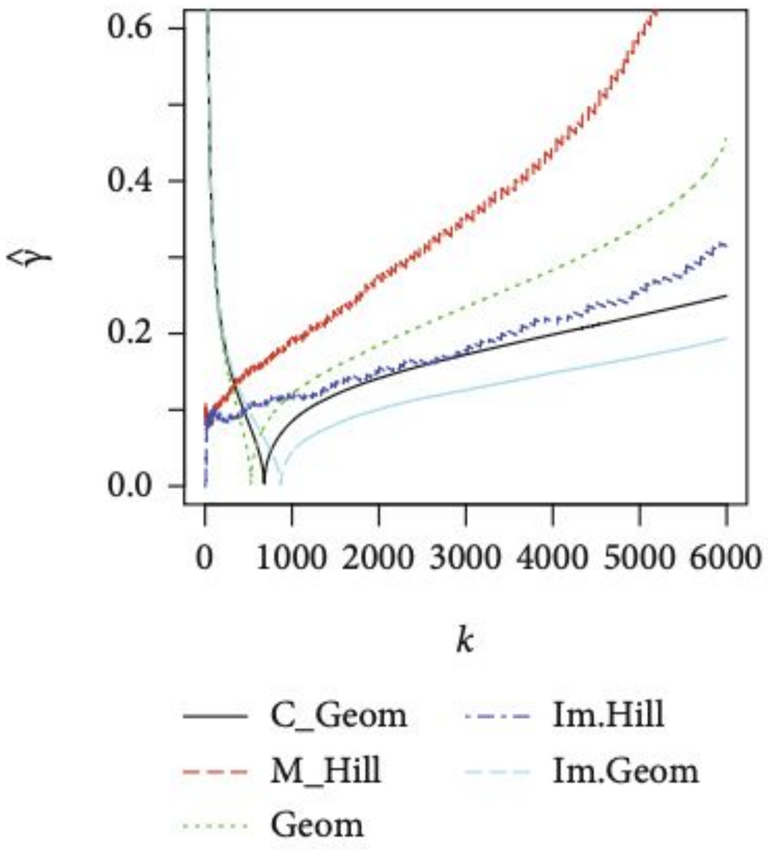
Real-World Example: Redefined Thresholds



Missing Data. Impute-then-Estimate Workflow.

Case study: Ayiah-Mensah et al. 2021 – 30-year Brest wind-speed series with 45 % values deleted

Estimator	Definition
C_Geom	*Complete-data geometric-type estimator – our benchmark.*
M_Hill	*Modified Hill on 45 % gappy data.*
Geom	*Modified Geometric-type on 45 % gappy data.*
Im.Hill	*Hill on MICE imputation dataset.*
Im.Geom	*Geometric-type on MICE imputation dataset.*



Estimator	Definition
C_Geom	*Complete-data geometric-type estimator – our benchmark. (black)*
M_Hill	*Modified Hill on 45 % gappy data. (dashed red)*
Geom	*Modified Geometric-type on 45 % gappy data. (dashed green)*
Im.Hill	*Hill on MICE imputation dataset. (dashed dark blue)*
Im.Geom	*Geometric-type on MICE imputation dataset. (dashed light blue)*

Hazard 1: Missing observations

Other Fixes for Missingness. 1

Tail-aware multiple imputation via quantile regression

Method

Active research; increasingly used in applied biomed/finance papers.

Maturity in the literature

Kleinke K., Fritsch M., Stemmler M., Reinecke J. & Lösel F. (2021) “Quantile-Regression-Based Multiple Imputation of Missing Values”

Reference

*reach for quantile-imputation when your covariates really predict extremes

Other Fixes for Missingness. 2

Censored Peaks-Over-Threshold (POT) likelihood

Method

Well-established for flood/insurance data since Beirlant et al. (2007 introduces the approach); extensions still appearing.

Maturity in the literature

Beirlant J., Guillou A. & Toulemonde G. (2010) “Peaks-Over-Threshold Modelling under Random Censoring”

Reference

*reach for censored-likelihood when the threshold is objective and recorded

Other Fixes for Missingness. 3

Inverse-Probability Weighting (IPW) for quantiles /extremes

Method

Classic missing-data tool; fewer extreme-value applications, but conceptually simple and fast.

Maturity in the literature

Vansteelandt S., Carpenter J. & Kenward M.G. (2010), "Analysis of Incomplete Data Using Inverse Probability Weighting and Doubly Robust Estimators"

Reference

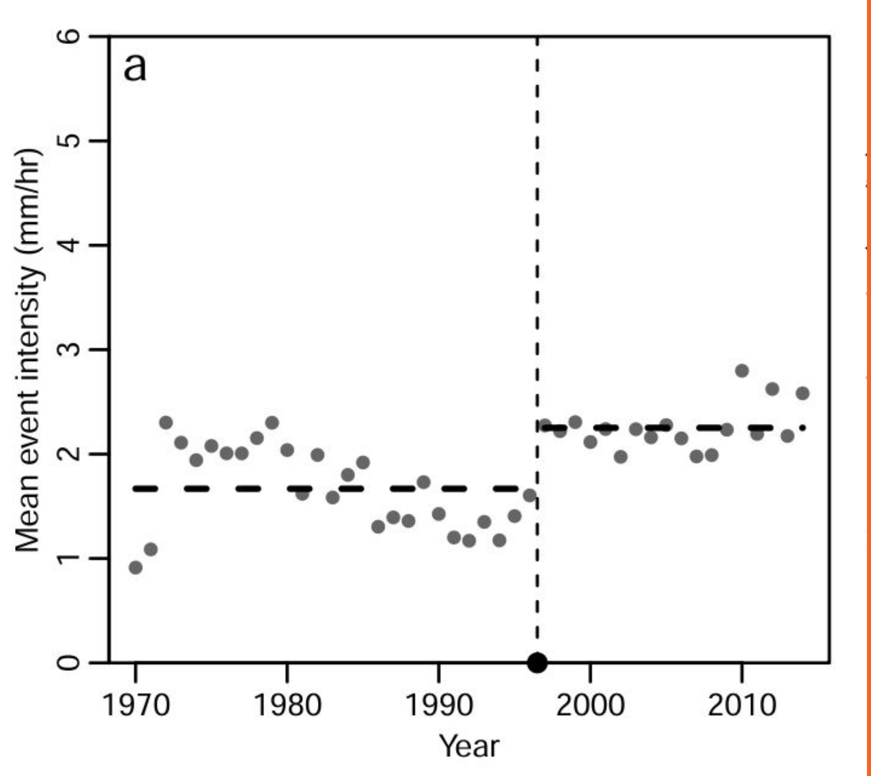
*reach for IPW when you can model the missingness process more confidently than the outcome itself

Measurement drift. Impute-then-Estimate Workflow.

Study set-up:
(a) 6-min rainfall from 71 stations, 1970-2014;
(b) the switch from manual Dines pluviographs (0.01 mm resolution) to digital tipping-bucket gauges (0.2 mm per tip);
(c) the swap peaked in late 1996, so that year is treated as the known breakpoint.

Case study: Wasko et al. (2022) "Automating Rainfall Recording: Ensuring Homogeneity when Instruments Change"

(a) Pre-aggregated, raw data



Hazard 2: Measurement drift

Measurement drift. Impute-then-Estimate Workflow.

Five corrections tested:

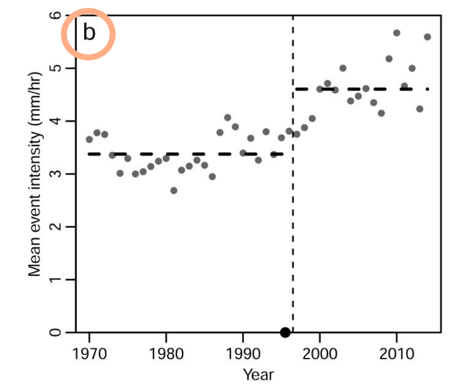
(b) **Tip filter**: Keep only tips bigger than 0.2 mm

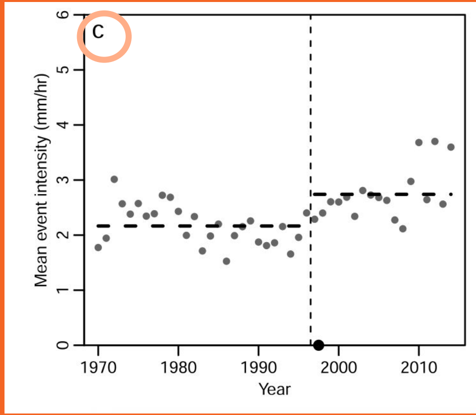
(c) **Event filter**: Restrict whole events ≤ 1 mm

(d) **Aggregate** the 0.01 mm pluviograph traces to mimic a 0.2 mm tipping-bucket

(e) **Aggregate and apply** the 0.2 mm **tip filter**

(f) **Aggregate and apply** the 1 mm **event filter**





Hazard 2: Measurement drift

Measurement drift. Impute-then-Estimate Workflow.

Five corrections tested:

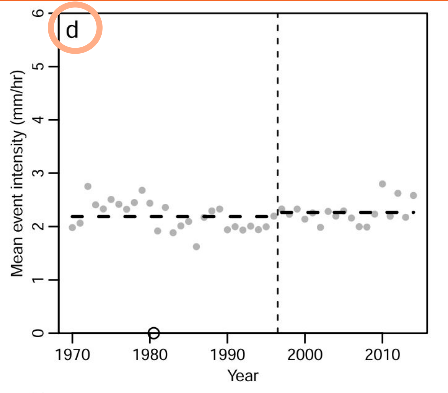
(b) **Tip filter**: Keep only tips bigger than 0.2 mm

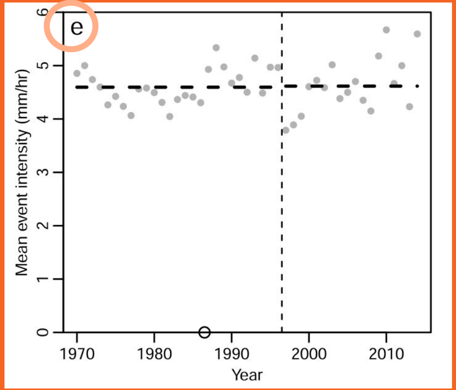
(c) **Event filter**: Restrict whole events ≤ 1 mm

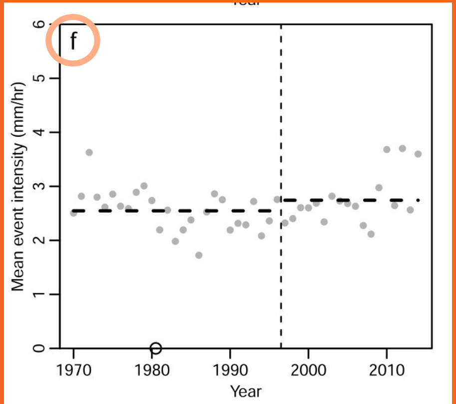
(d) **Aggregate** the 0.01 mm pluviograph traces to mimic a 0.2 mm tipping-bucket

(e) **Aggregate and apply** the 0.2 mm **tip filter**

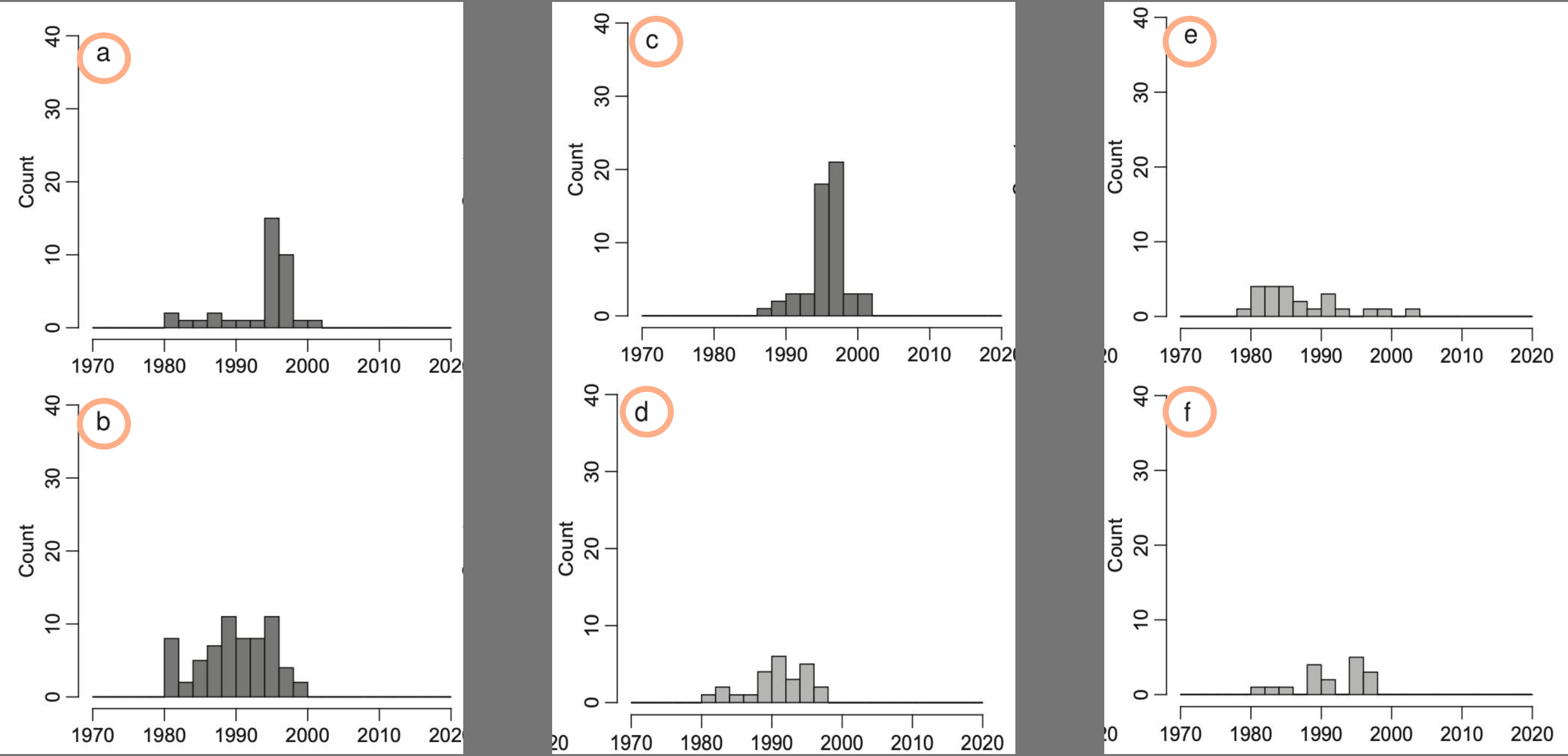
(f) **Aggregate and apply** the 1 mm **event filter**







Histogram of statistically significant change points for each correction



How well do the fixes work?

discussing the findings

➔ **Raw Data**

About 60 % of the 71 gauges show a statistically significant Pettitt change point in mean event intensity right on the 1996 instrument swap.

➔ **Aggregation only (d)**

Cuts the problem roughly in half (count the statistically significant change points from the previous slide)), but change points around 1996 are still common.

➔ **Aggregation + either filter (e)/(f)**

Drops the count to "very few stations" (can be think of as a background noise) and) flattens the Australia-wide map of change points

➔ **Best fit (f)**

Pettitt p-value moves from 0.00 (significant) in the raw series to 0.24 (non-significant) after the aggregation + event filter, confirming the shift is gone (Table 2 in the article).

Other Fixes for Instrument-Induced Inhomogeneity. 1

Standard-Normal Homogeneity Test + step adjust

Method

Univariate change point with step correction

Category

Pandžić, K., et all (2019). “Standard normal homogeneity test as a tool to detect change points in climate-related river discharge variation”

Reference

Other Fixes for Instrument-Induced Inhomogeneity. 2

RHtestsV4 (penalised t/F)

Method

Recursive multi-break adjustment

Category

Wang, X. L. and Y. Feng, published online July 2013: RHtestsV4 User Manual. (software package)

Reference

Other Fixes for Instrument-Induced Inhomogeneity. 3

Bayesian hierarchical change-point modelling

Method

Model-based / probabilistic

Category

Zhao et al. 2019, Remote Sensing of Environment - A Bayesian ensemble algorithm.

Reference

Thank you!