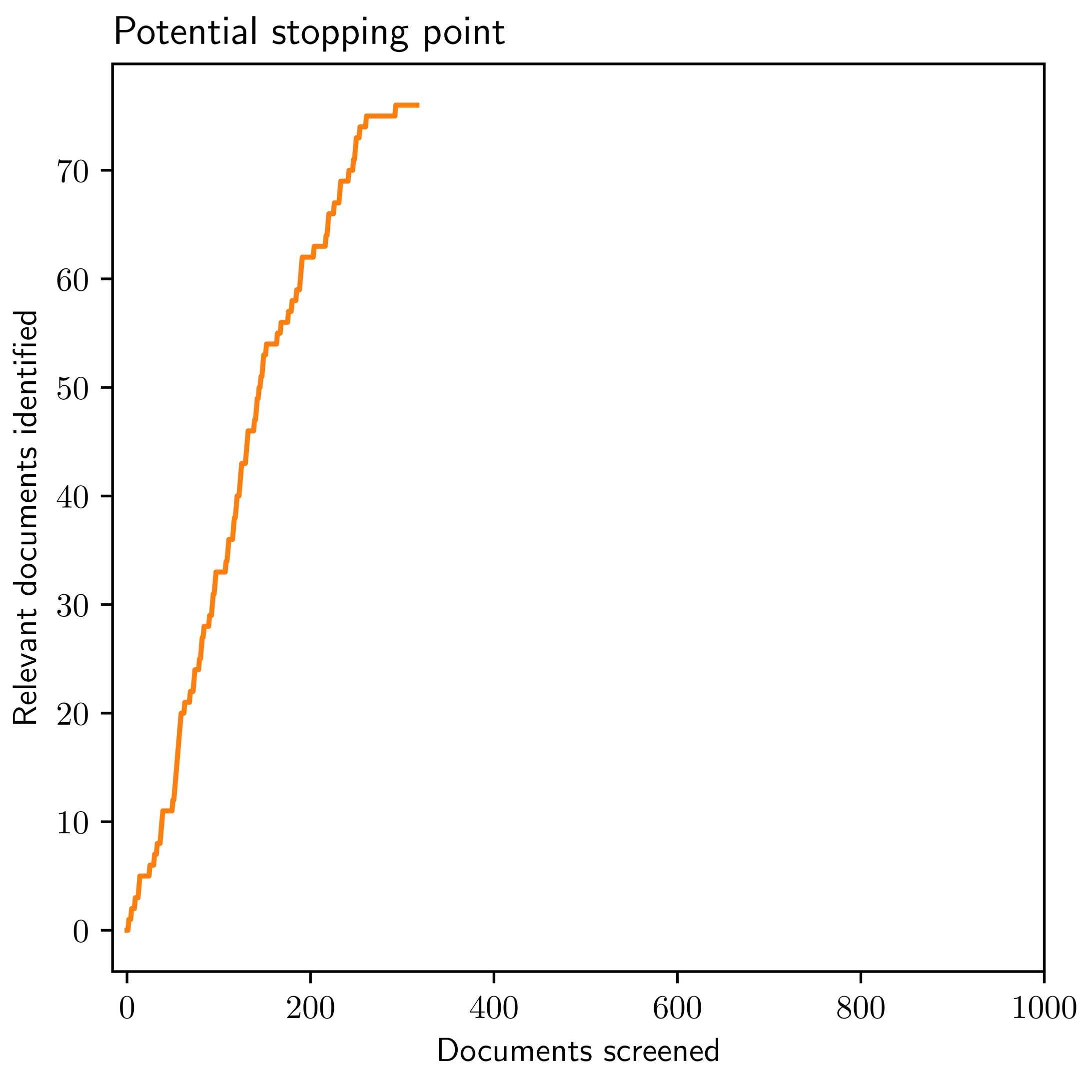
Biased Urns for efficient Stopping Criteria in technology-Assisted Reviews (BUSCAR)

MCC

Problem

Machine-learning priorised screening can save work, but only if we know when it is **safe** to stop



Urns

The hypergeometric distribution models the likelihood of retrieving k red marbles from an urn containing N marbles, of which K are red, if we pick out n marbles without replacing them.

Thinking of relevant documents as red marbles, this provides a conservative stopping criterion (Callaghan and Müller-Hansen, 2020), when we assume (conservatively) that the likelihood of drawing a red or white marble is equal.

Biased Urns

Biased Urn Theory (Fog, 2023), models what happens when the likelihood of drawing a random item of interest is higher than the likelihood of drawing a random irrelevant item. The ratio between these two is the **bias**

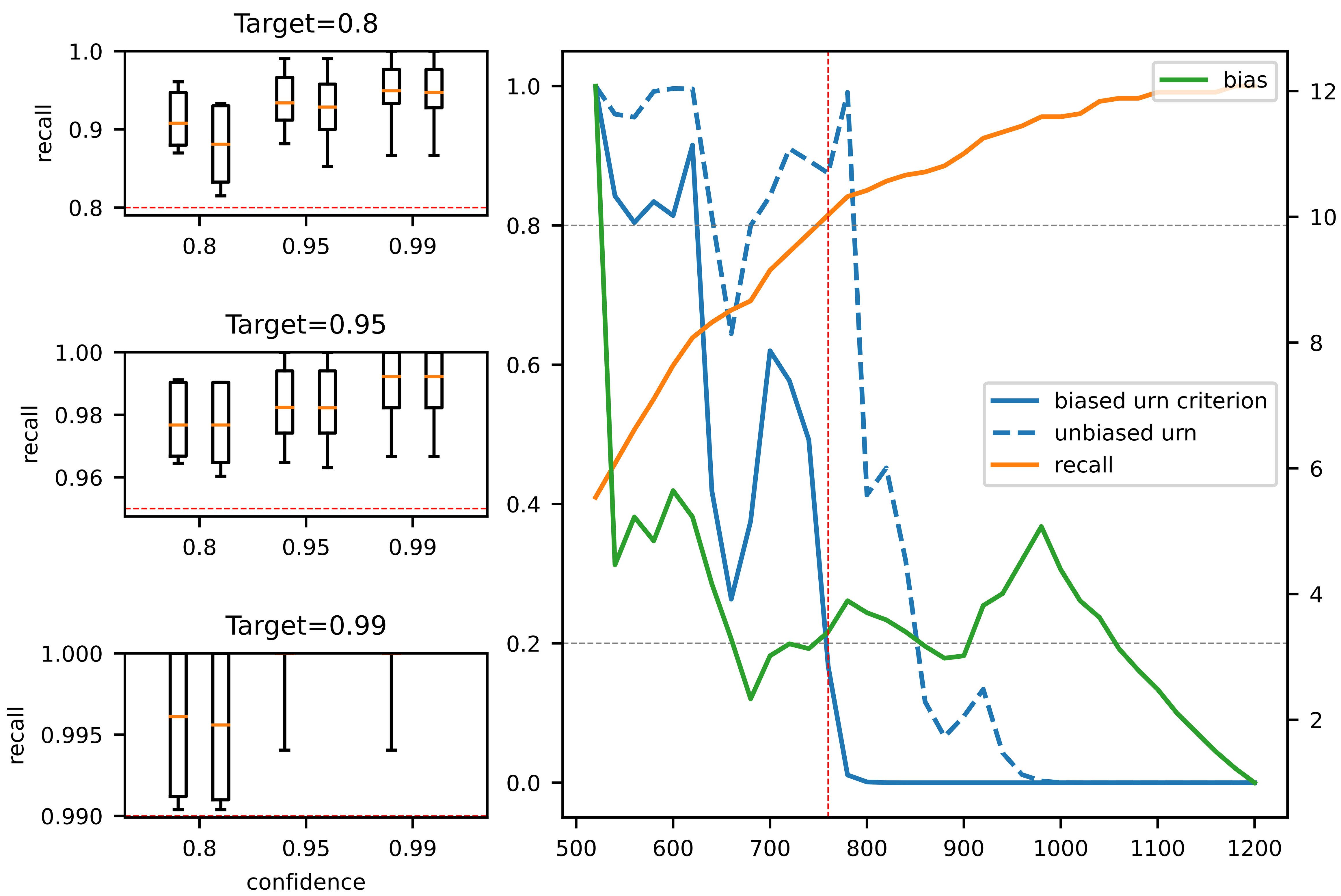
Method

1. Estimate maximum number of relevant documents in a given confidence interval from random sample
2. Look back at the last 1/2 of documents since the random sample, and use Wallenius' non-central hypergeometric distribution and Maximum Likelihood Estimation to find the level of bias below which 95% of the probability distribution occurs
3. Look back at previous batches (last 1, last 2, last 3...) to see the likelihood of the observed number or relevant documents occuring given the null hypothesis that we have missed our target.

Experiments

We test our stopping criteria with 100 runs on complete systematic review datasets (from Cohen, 2006), with different recall targets (0.8, 0.95, 0.99) and different confidence levels (0.8, 0.95, 0.99). We show the distribution of additional work savings compared to an omniscient criterion, as well as the distribution of recall, both on average, and for each dataset.

Results



Using a biased urn results in higher work savings (see example pathway, while still maintaining recall levels above a target level at a given confidence level).

Email me for how to use the criterion in your project. R and python packages forthcoming.

Max Callaghan & Finn Müller-Hansen
callaghan@mcc-berlin.net, @MaxCallaghan5