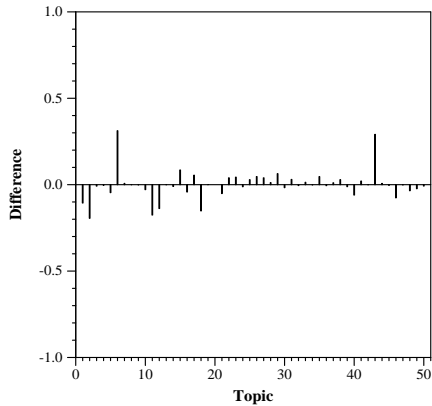


Relevance Feedback Track results

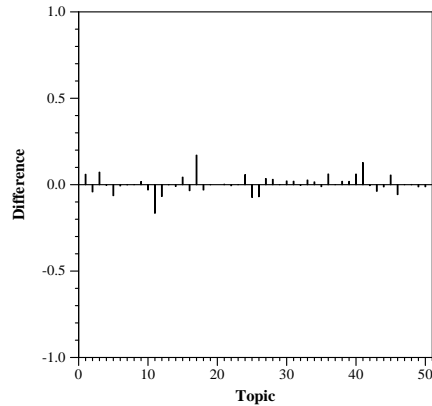
Phase 2: Each group ran with 7 to 8 different relevance feedback input documents, and ran a base case with no relevance feedback. Evaluated with two measures. If the group ran on the full collection, the measures were MAP and P(10). If the group ran on the B subset, the measures were statAP and eMAP (Million Query style evaluation).

In the Per topic Median Difference graphs, the median used is the global median measure (over all Phase 1 sets and base case) for each topic. Thus it remains constant between graphs.

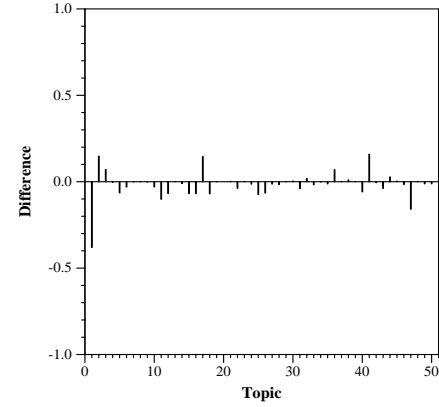
| Phase 2 Run Summary Statistics | | |
|--|--------|--------|
| Document Collection : A (Full English) | | |
| Run ID | MAP | P(10) |
| WAT2.base | 0.0526 | 0.2755 |
| WAT2.CMIC.2 | 0.0566 | 0.3857 |
| WAT2.MSRC.2 | 0.0364 | 0.2388 |
| WAT2.UCSC.1 | 0.0887 | 0.4367 |
| WAT2.UPD.1 | 0.0446 | 0.2714 |
| WAT2.WatS.2 | 0.1047 | 0.4816 |
| WAT2.YUIR.1 | 0.0184 | 0.1755 |
| WAT2.hit2.2 | 0.0410 | 0.2449 |
| WAT2.udel.1 | 0.0924 | 0.4388 |



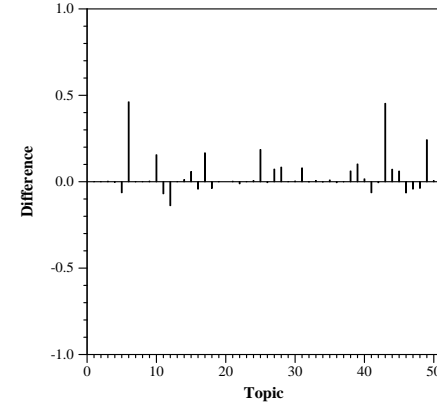
MAP Per Topic Median Diff, WAT2.base



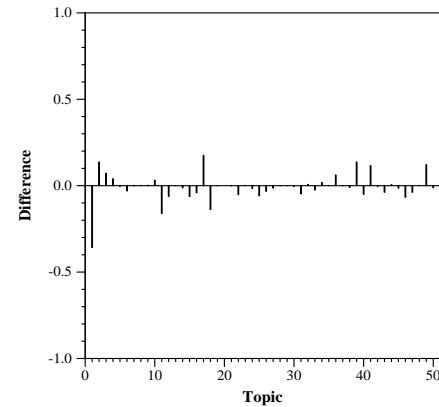
MAP Per Topic Median Diff, WAT2.CMIC.2



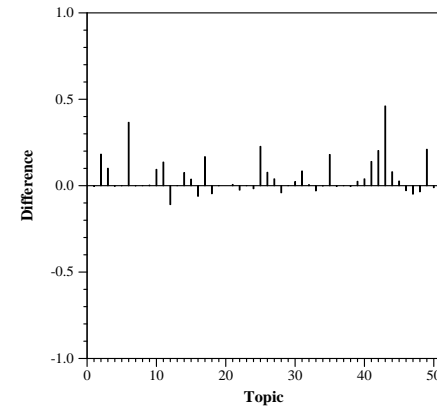
MAP Per Topic Median Diff, WAT2.MSRC.2



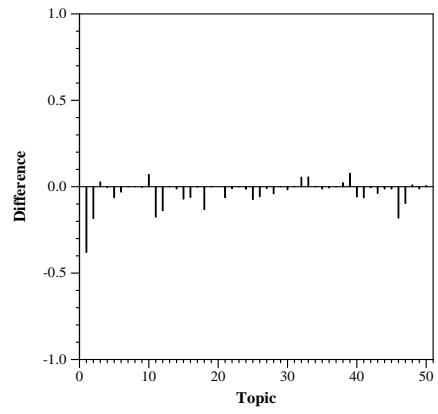
MAP Per Topic Median Diff, WAT2.UCSC.1



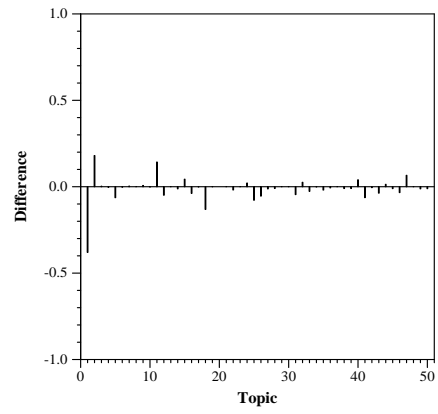
MAP Per Topic Median Diff, WAT2.UPD.1



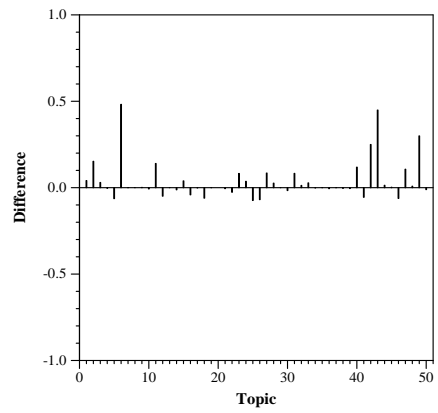
MAP Per Topic Median Diff, WAT2.WatS.2



MAP Per Topic Median Diff, WAT2.YUIR.1



MAP Per Topic Median Diff, WAT2.hit2.2



MAP Per Topic Median Diff, WAT2.udel.1