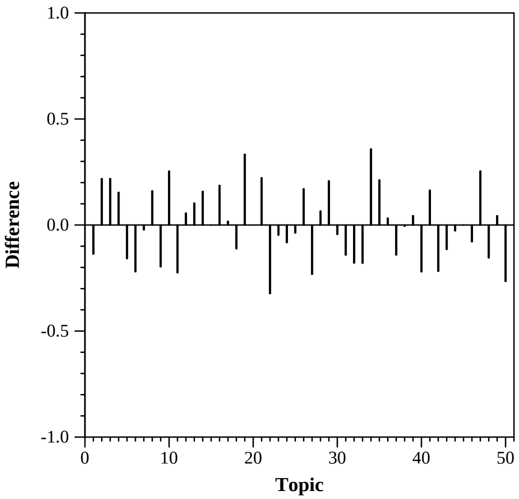
Phase 1: Each group submitted a set of 5 documents per topic to be used as relevance feedback input in Phase 2 by 3 to 5 groups. One or two sets submitted. Evaluation output includes number of relevant documents in set, and how well other groups did on this set compared to the other sets that that group ran (each group ran 7 to 8 Phase 1 sets). Comparison numbers totaled among the collection and evaluation measures used in Phase 2. Total score = B / (B + W) where B is the total number of runs/measures this set did better than, and W is the number this set did worse on.

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).

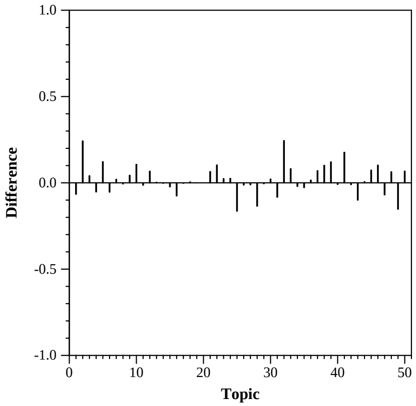
In the Per topic Median Difference graphs, the median used is the global median measure (over all Phase 1 sets and base case) for each topic. THus it remains constant between graphs.

Phase 1 Summary Statistics			
RF Input Set		UPD.1	
Total Num Rel in Set		115	
Measure	Coll	Num Worse Than	Num Better Than
MAP(all)	Full	4	3
P(10)(all)	Full	4	3
statMAP (NEU) (all)	B	12	8
eMAP (UMass) (all)	B	13	7
Measure			Score
Score (all)			0.3889
Score (average over q)			0.4847

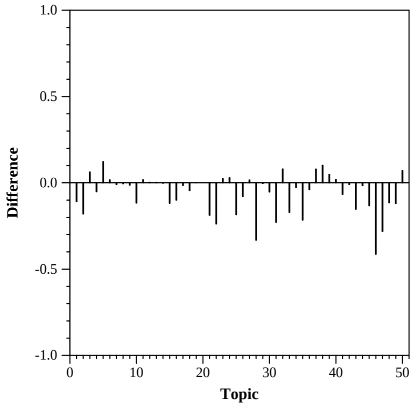


Score Per Topic Diff from Median, UPD.1

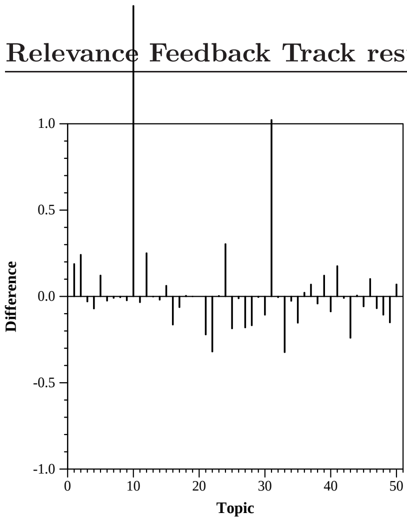
Phase 2 Run Summary Statistics		
Document Collection : B (English1 Subset)		
Run ID	statAP	eMAP
UPD9RF.base	0.1655	0.0409
UPD9RF.CMU.1	0.0827	0.0279
UPD9RF.PRIS.1	0.1778	0.0270
UPD9RF.QUT.1	0.1349	0.0298
UPD9RF.UMas.1	0.0970	0.0218
UPD9RF.UPD.1	0.1015	0.0273
UPD9RF.hit2.1	0.0938	0.0259
UPD9RF.ilps.2	0.0984	0.0242



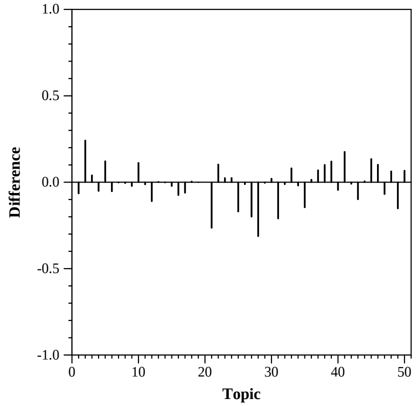
statAP Per Topic Median Diff, UPD9RF.base



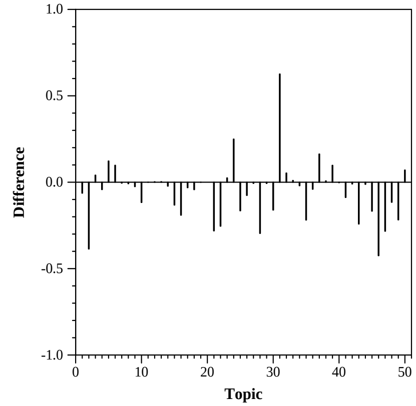
statAP Per Topic Median Diff, UPD9RF.CMU.1



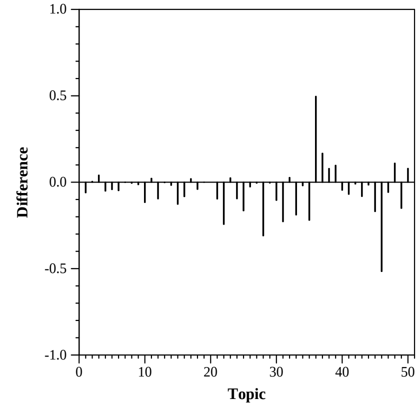
statAP Per Topic Median Diff, UPD9RF.PRIS.1



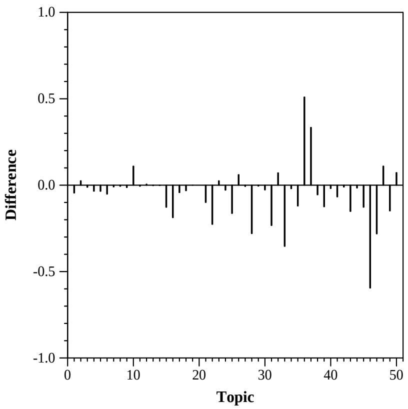
statAP Per Topic Median Diff, UPD9RF.QUT.1



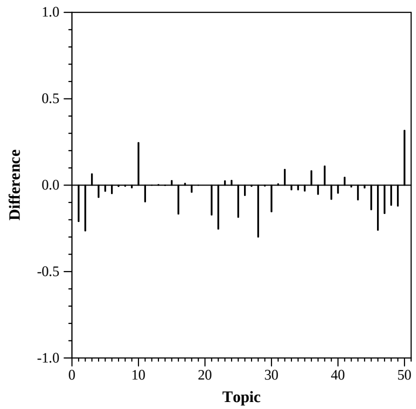
statAP Per Topic Median Diff, UPD9RF.hit2.1



statAP Per Topic Median Diff, UPD9RF.ilps.2



statAP Per Topic Median Diff, UPD9RF.UMas.1



statAP Per Topic Median Diff, UPD9RF.UPD.1