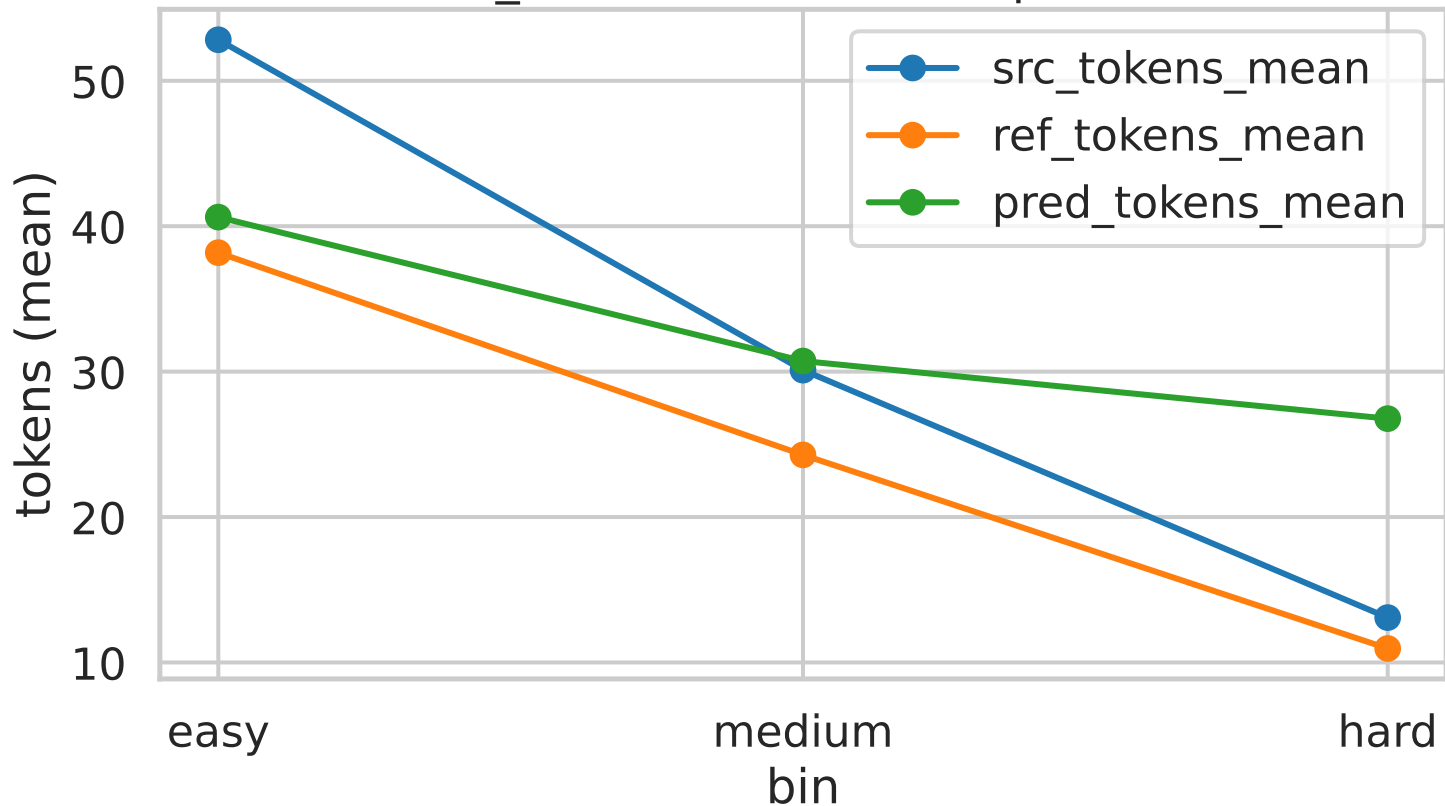
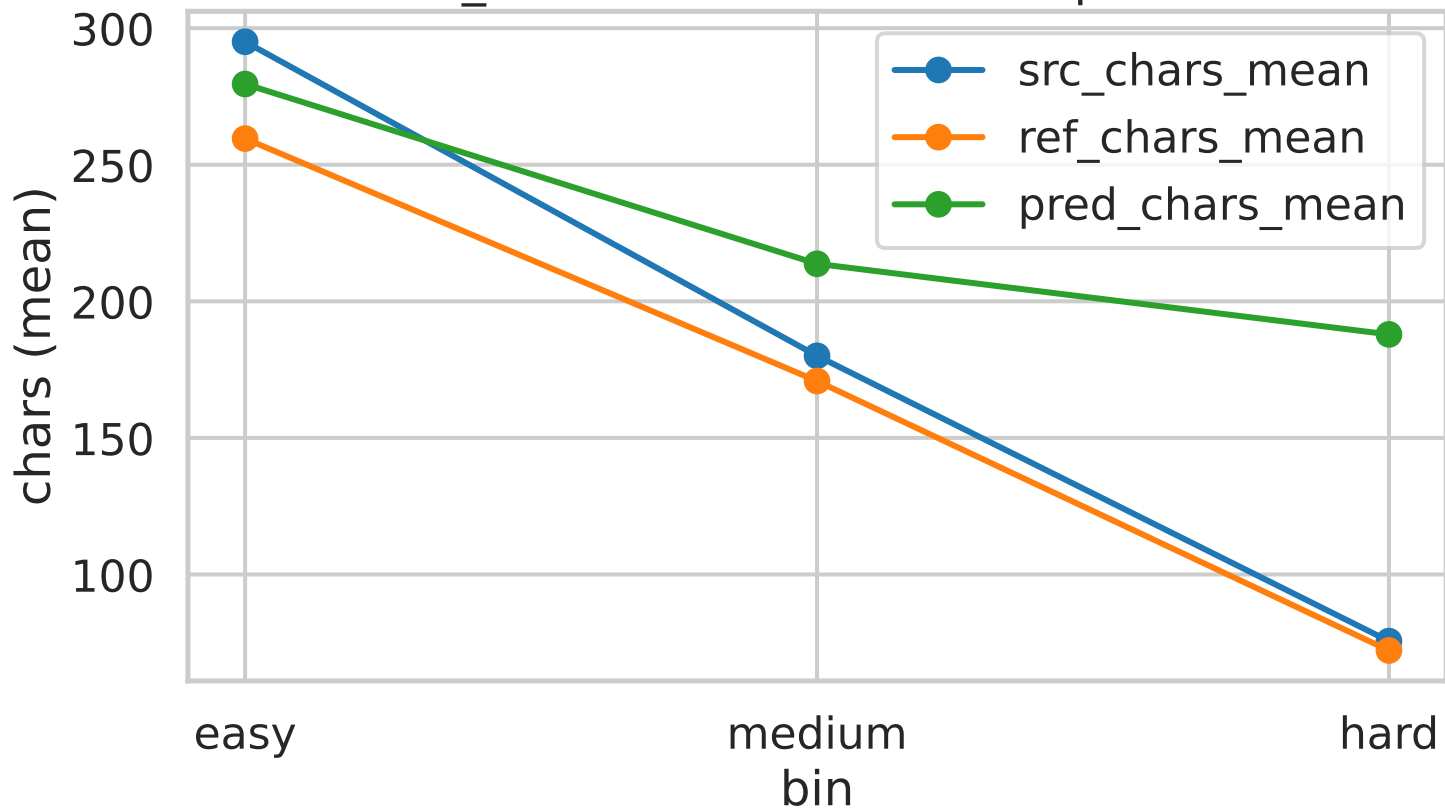


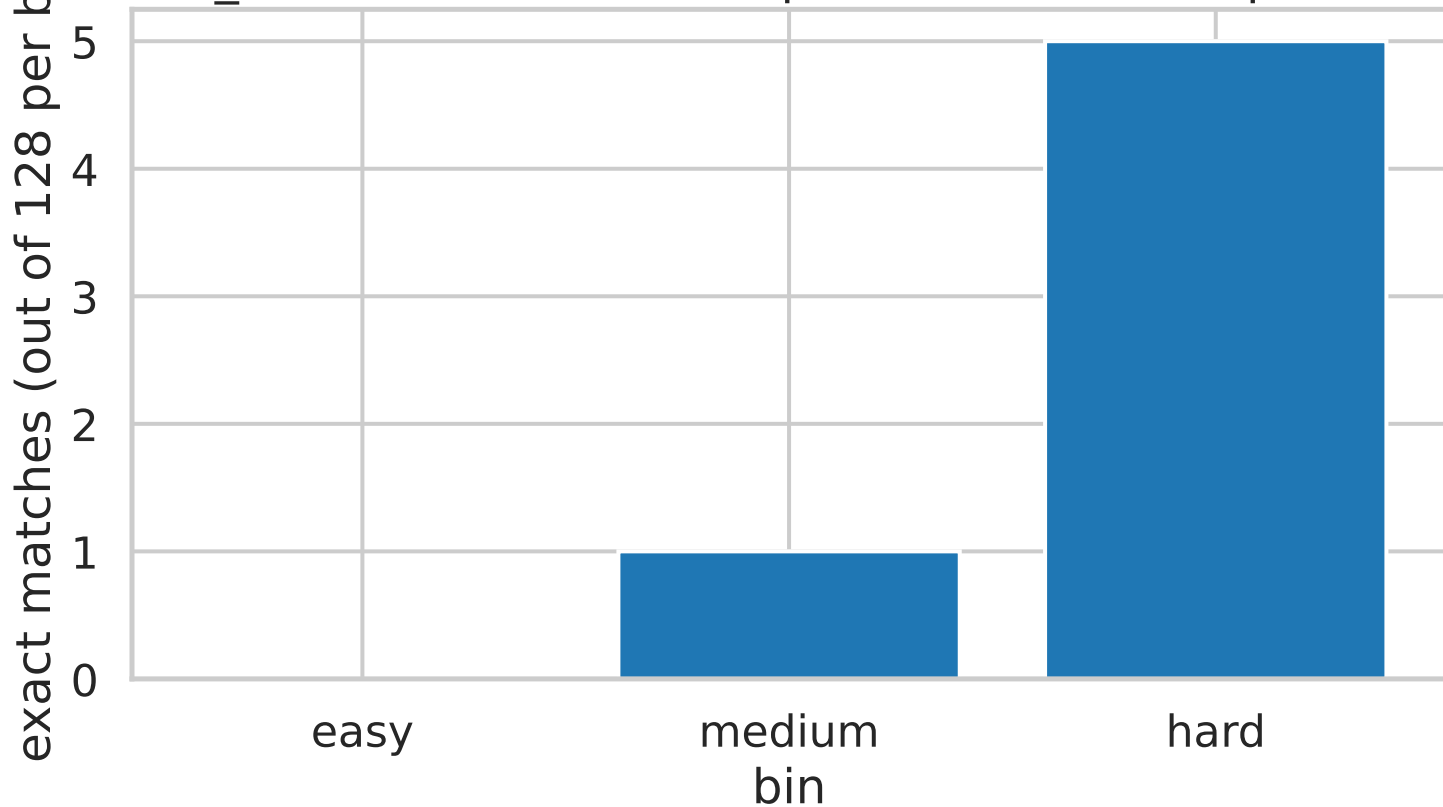
TM\_4bit — Mean tokens per bin



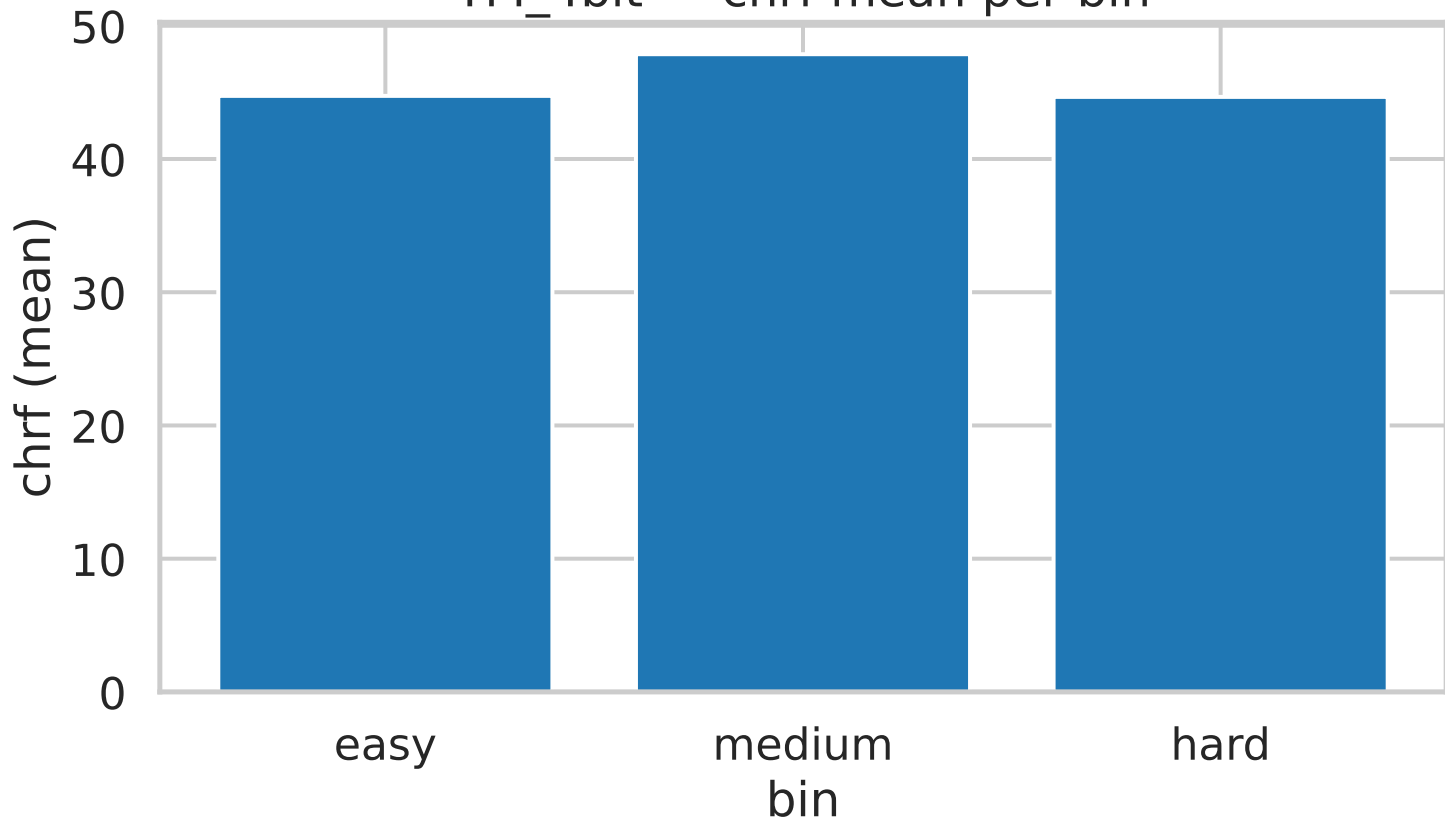
TM\_4bit — Mean characters per bin



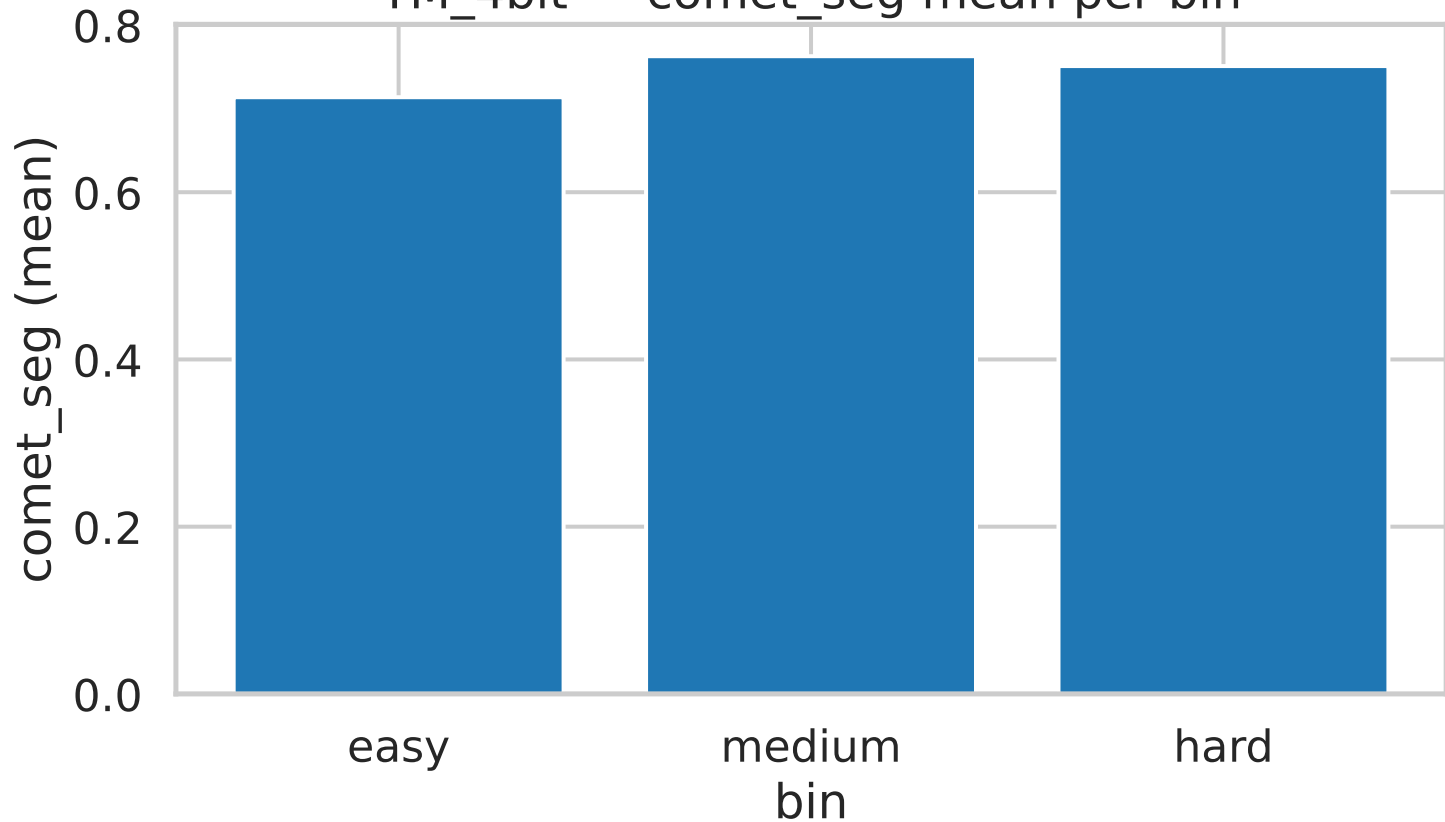
TM\_4bit — Exact matches per bin (out of 128 per bin)



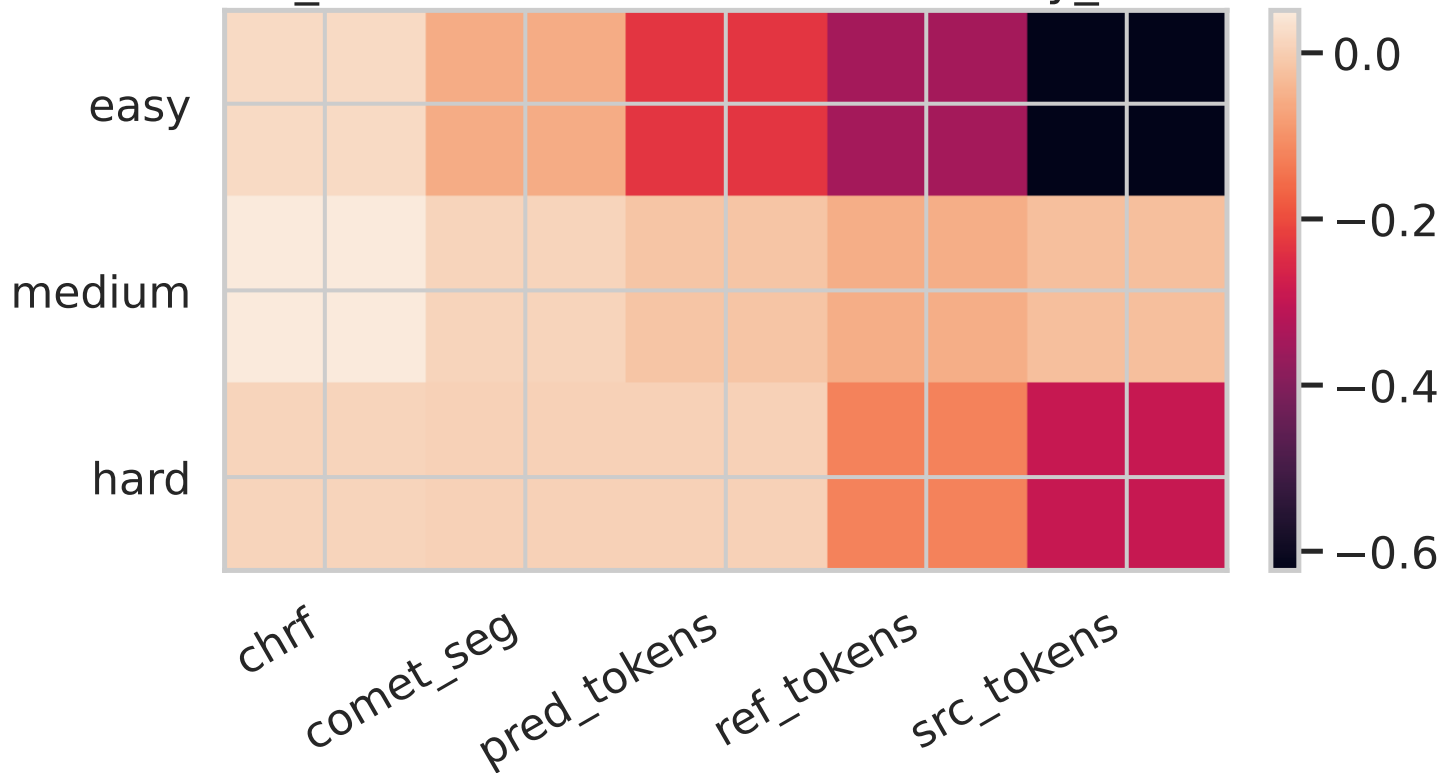
TM\_4bit — chrF mean per bin



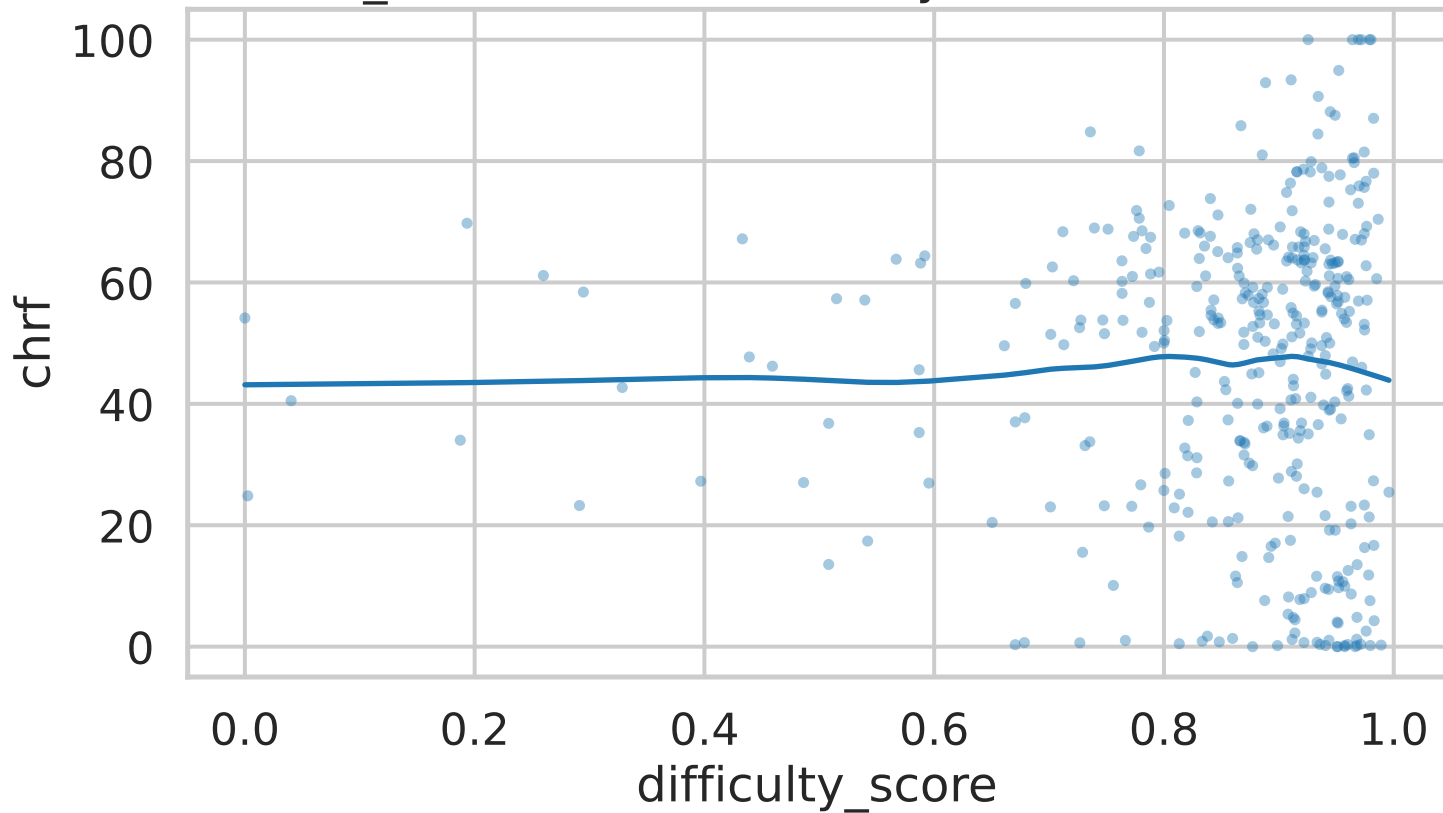
TM\_4bit — comet\_seg mean per bin



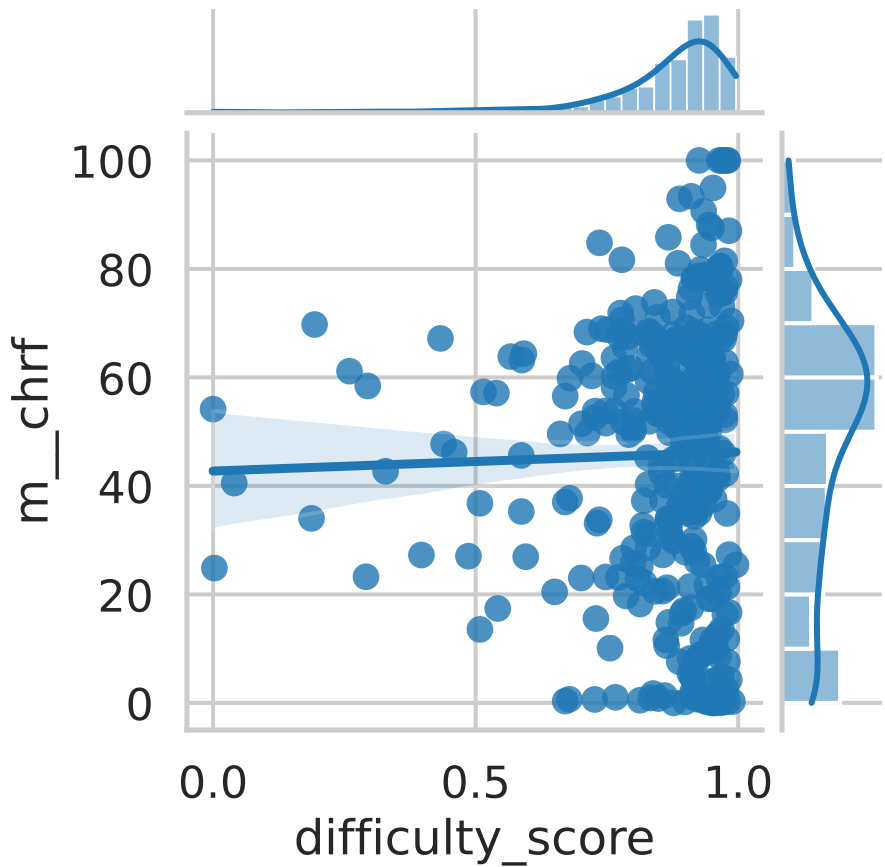
TM\_4bit — Correlations w/ difficulty\_score



TM\_4bit — chrF vs difficulty (LOWESS=True)

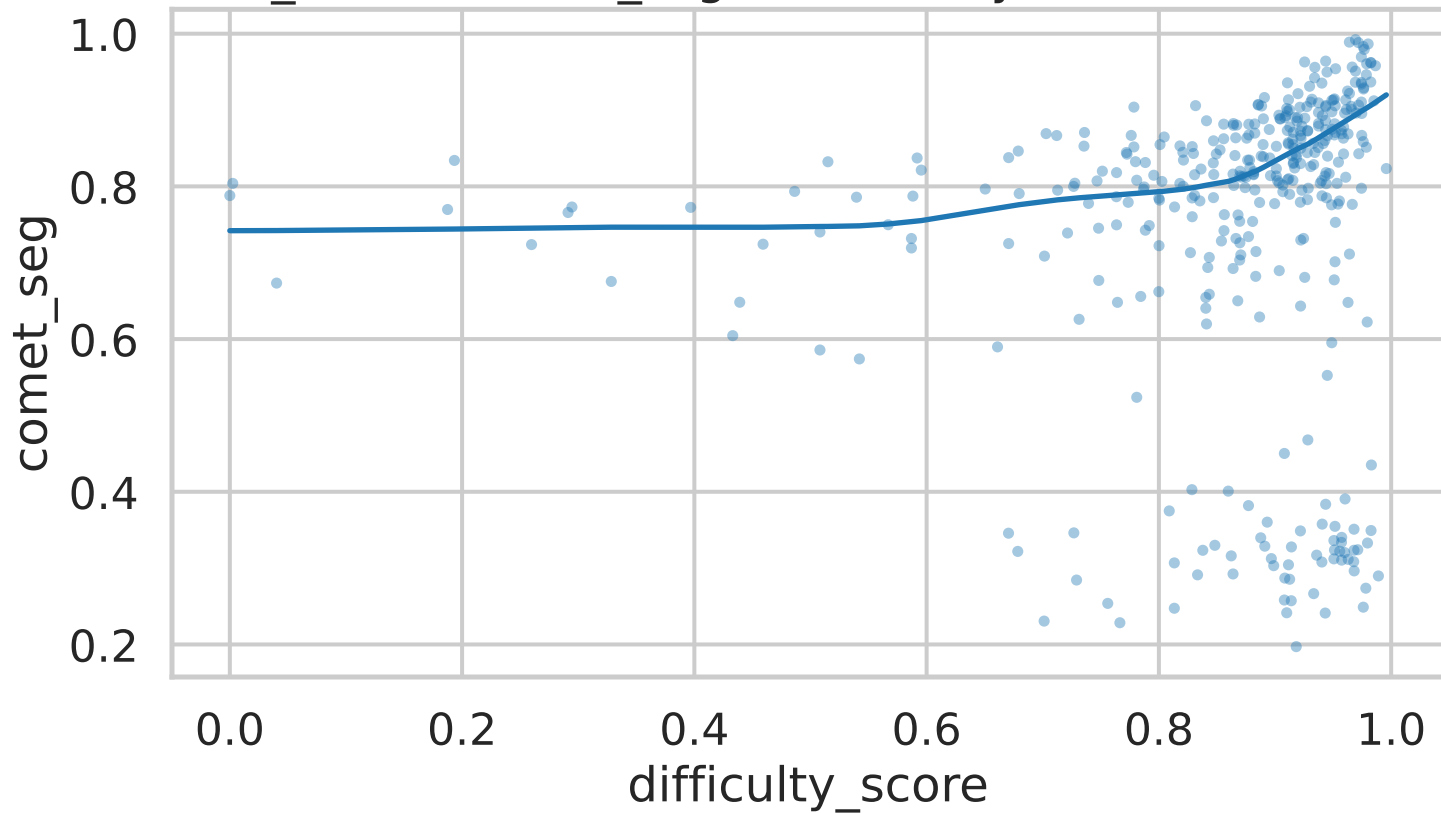


l\_4bit = joint chr1 vs difficulty\_score

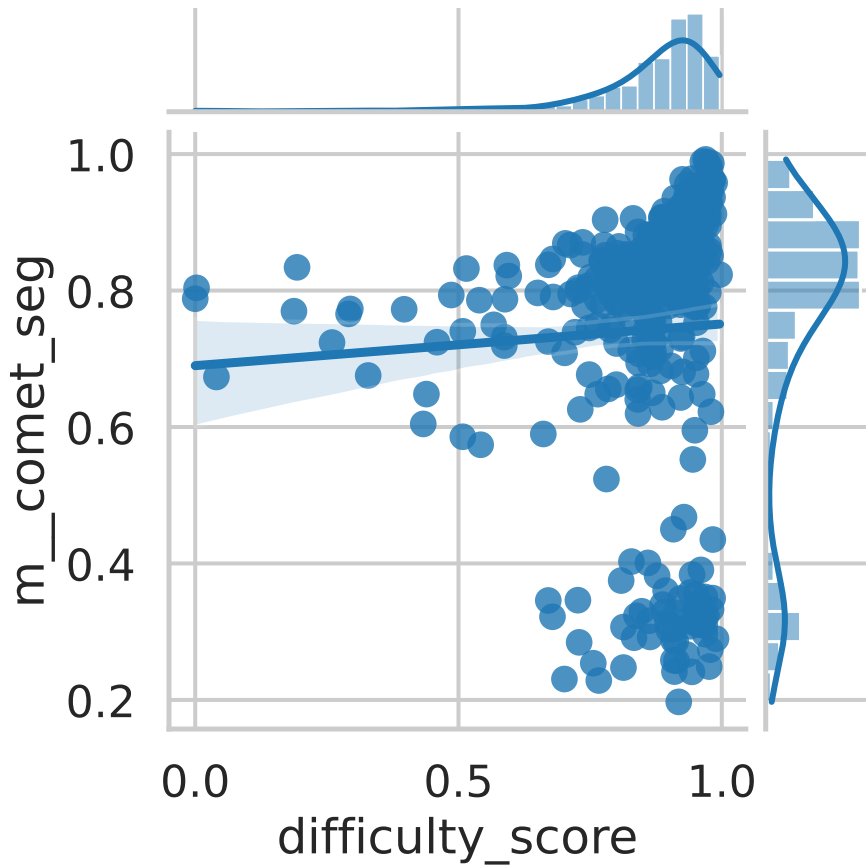




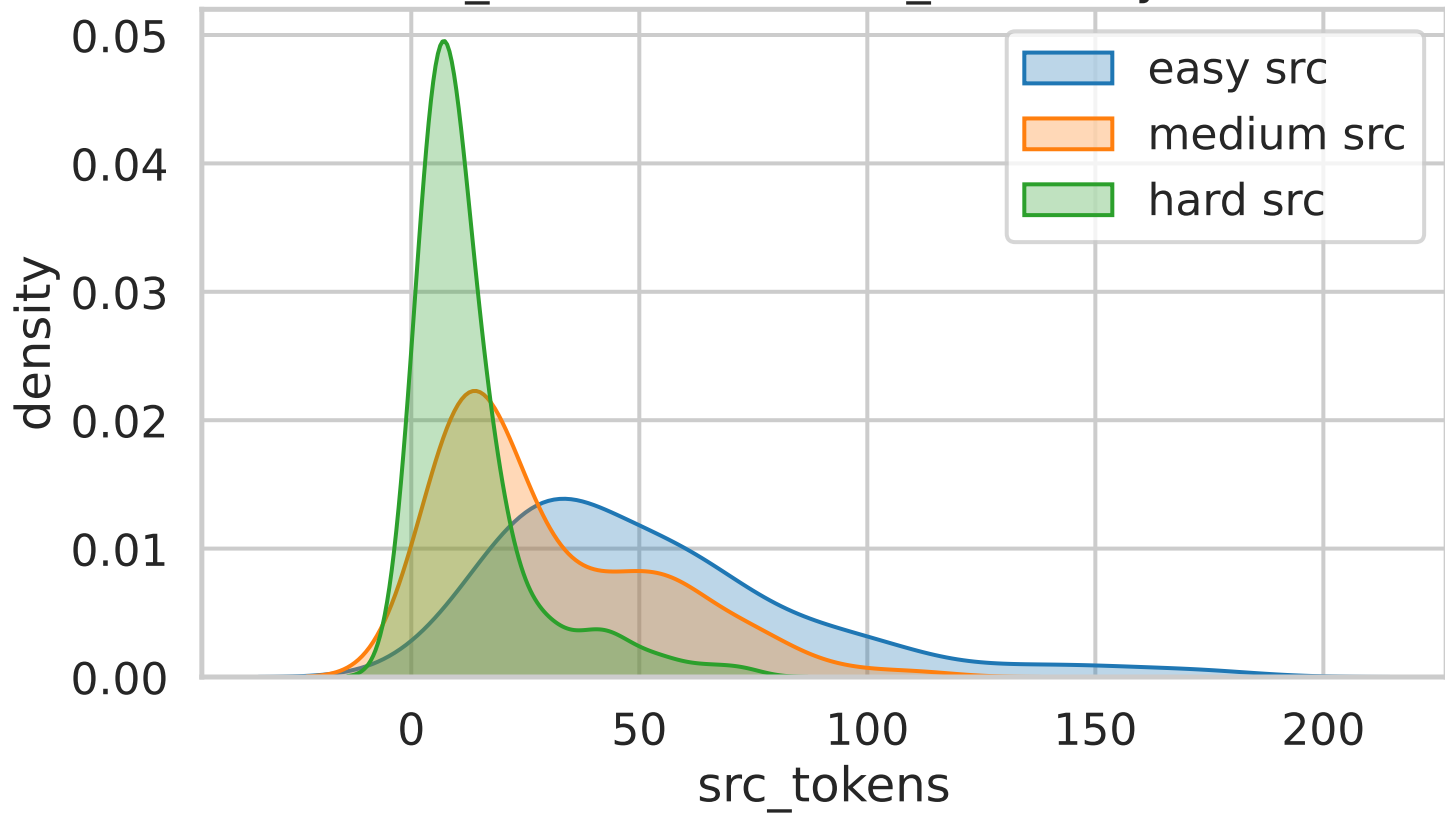
TM\_4bit — comet\_seg vs difficulty (LOWESS=True)



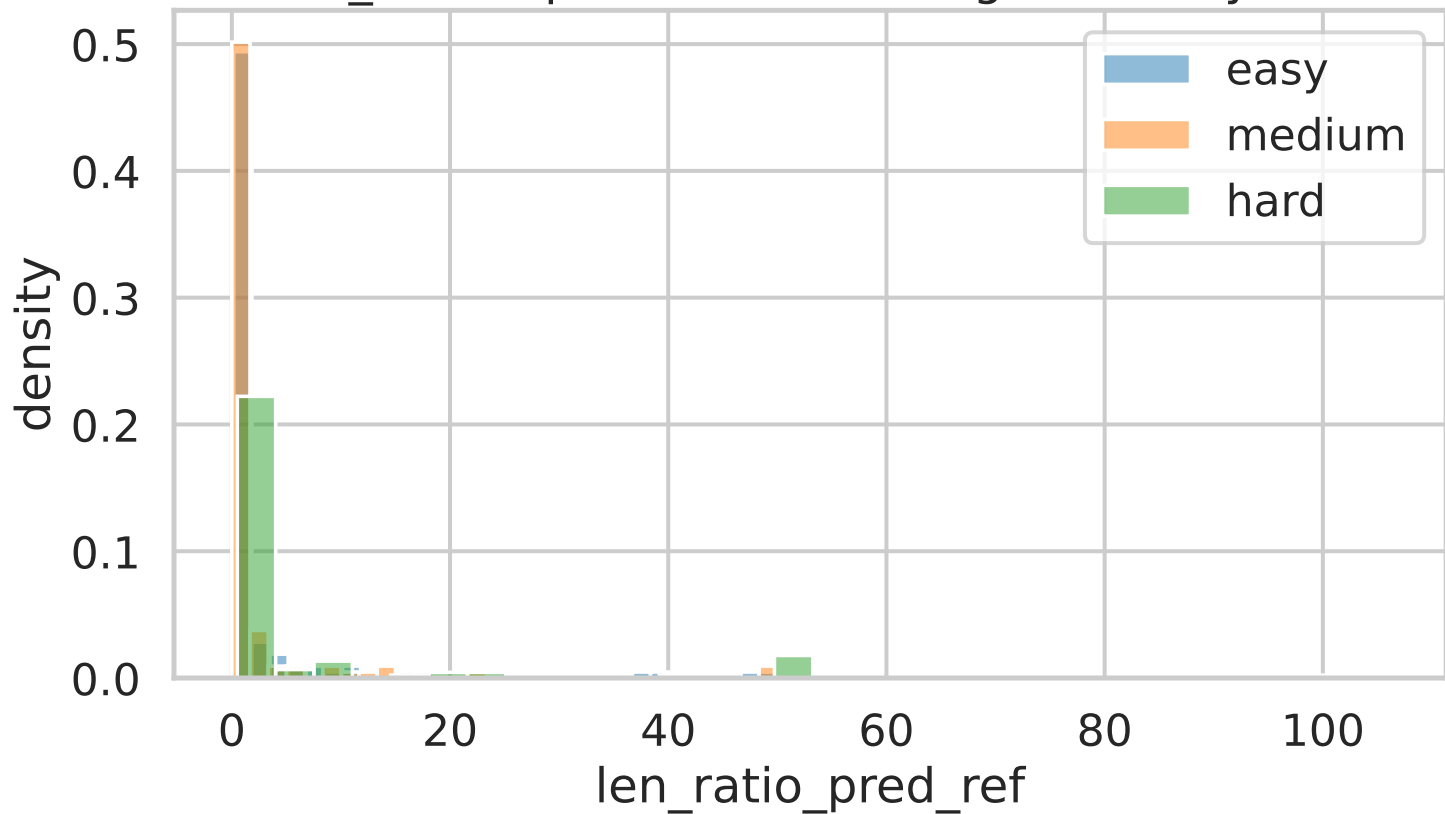
bit = joint\_comet\_seg vs difficulty\_



TM\_4bit — KDE of src\_tokens by bin



TM\_4bit — pred/ref token length ratio by bin



TM\_4bit — Number consistency (src→pred) per bin

