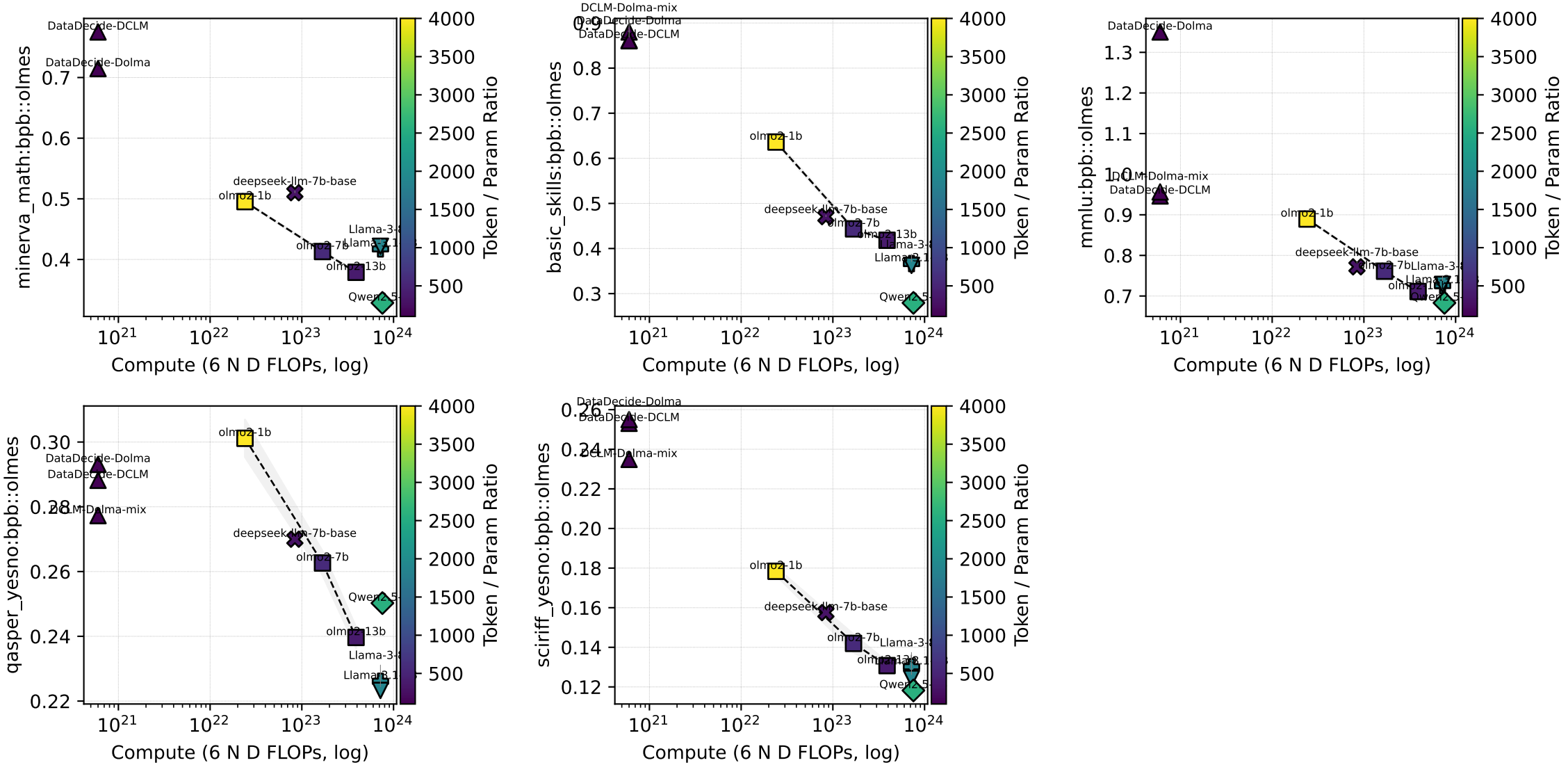
minerva_math:bpb::olmes
basic_skills:bpb::olmes
mmlu:bpb::olmes
qasper_yesno:bpb::olmes
sciriff_yesno:bpb::olmes
Compute (6 N D FLOPs, log)
Token / Param Ratio
DataDecide-DCLM
DataDecide-Dolma
DCLM-Dolma-mix
deepseek-llm-7b-base
olmo2-1b
olmo2-7b
olmo2-13b
Llama-3-8
Llama3.1
Qwen2.5