

2022-01
(GPT-3.5)

2022-11
(GPT-4)

2023-09
(gpt-4-1106)

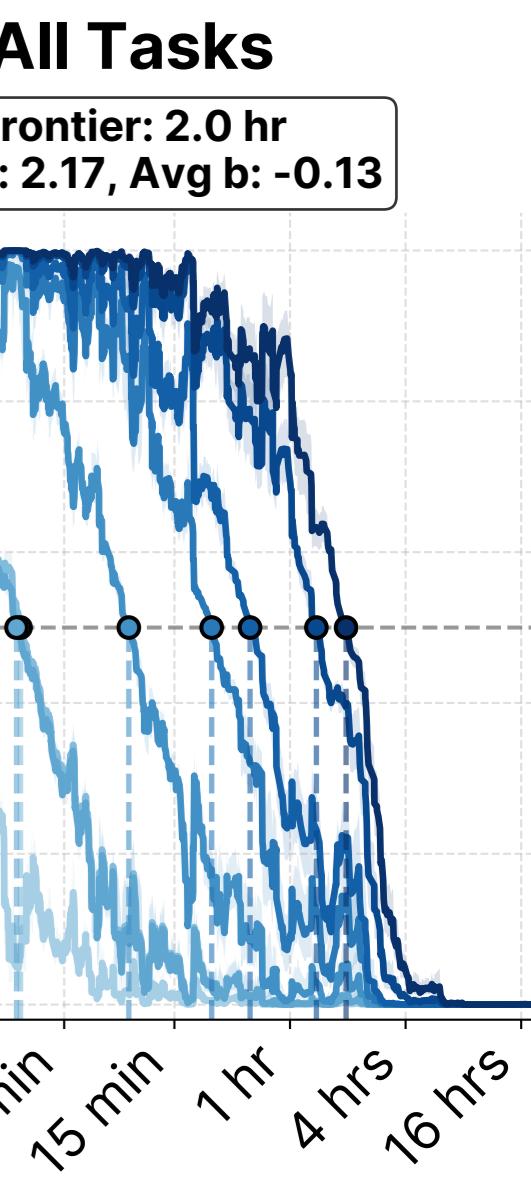
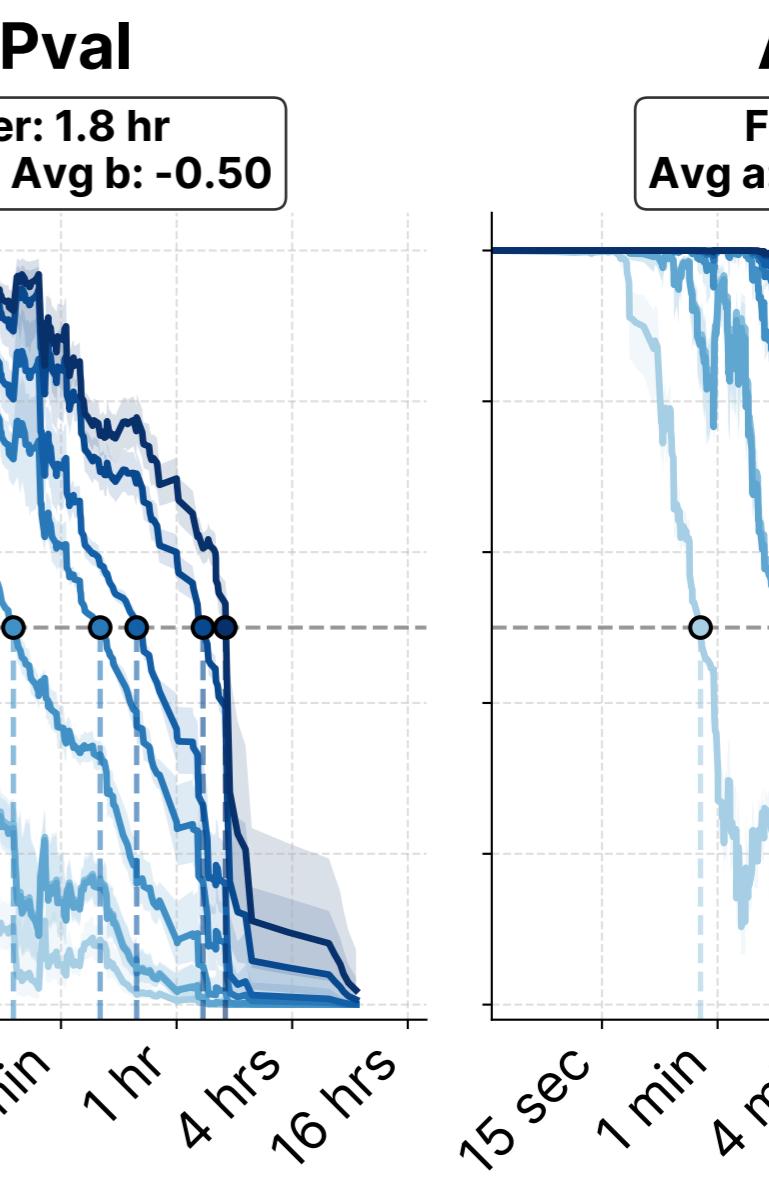
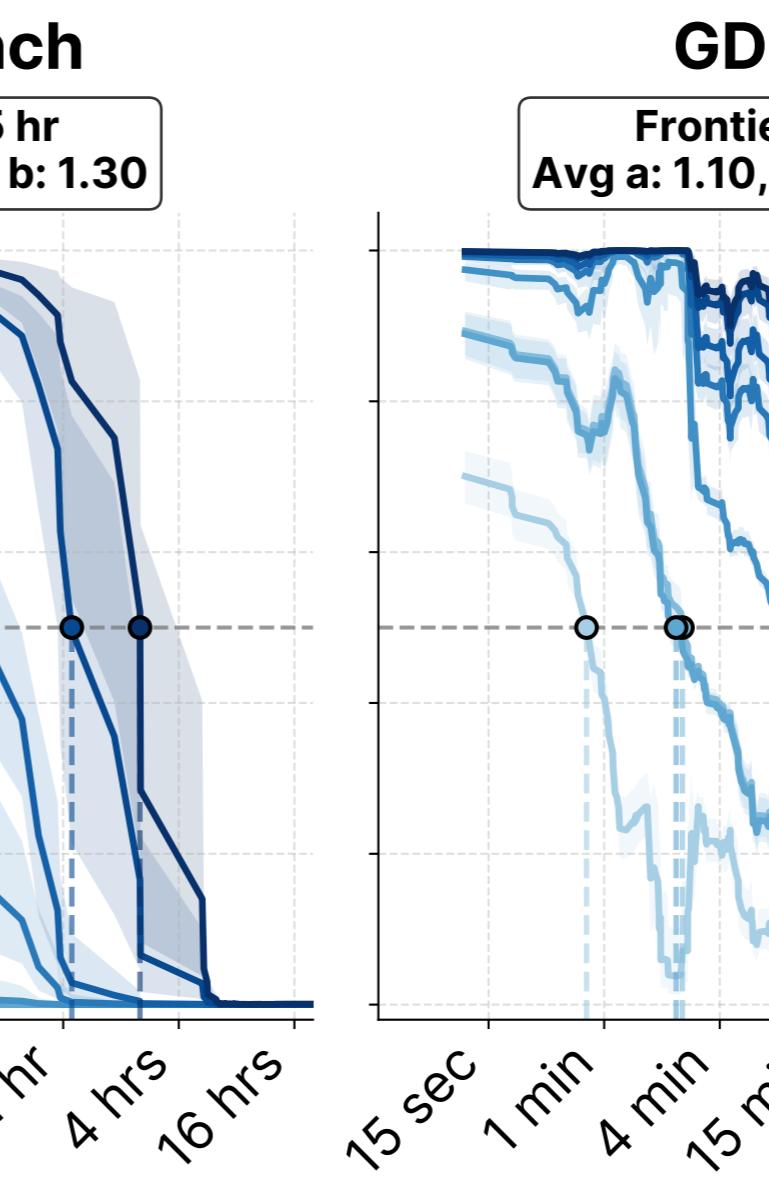
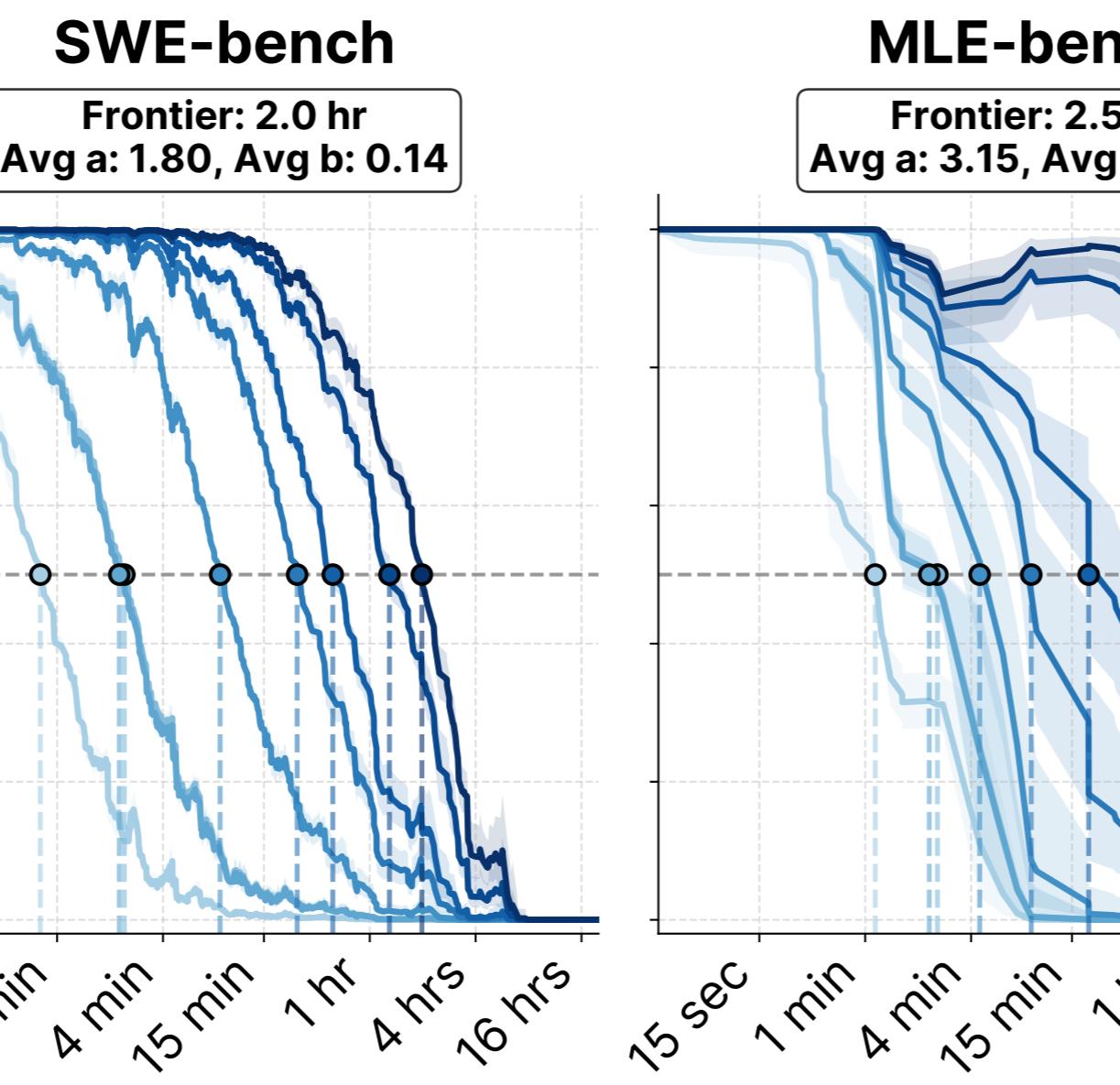
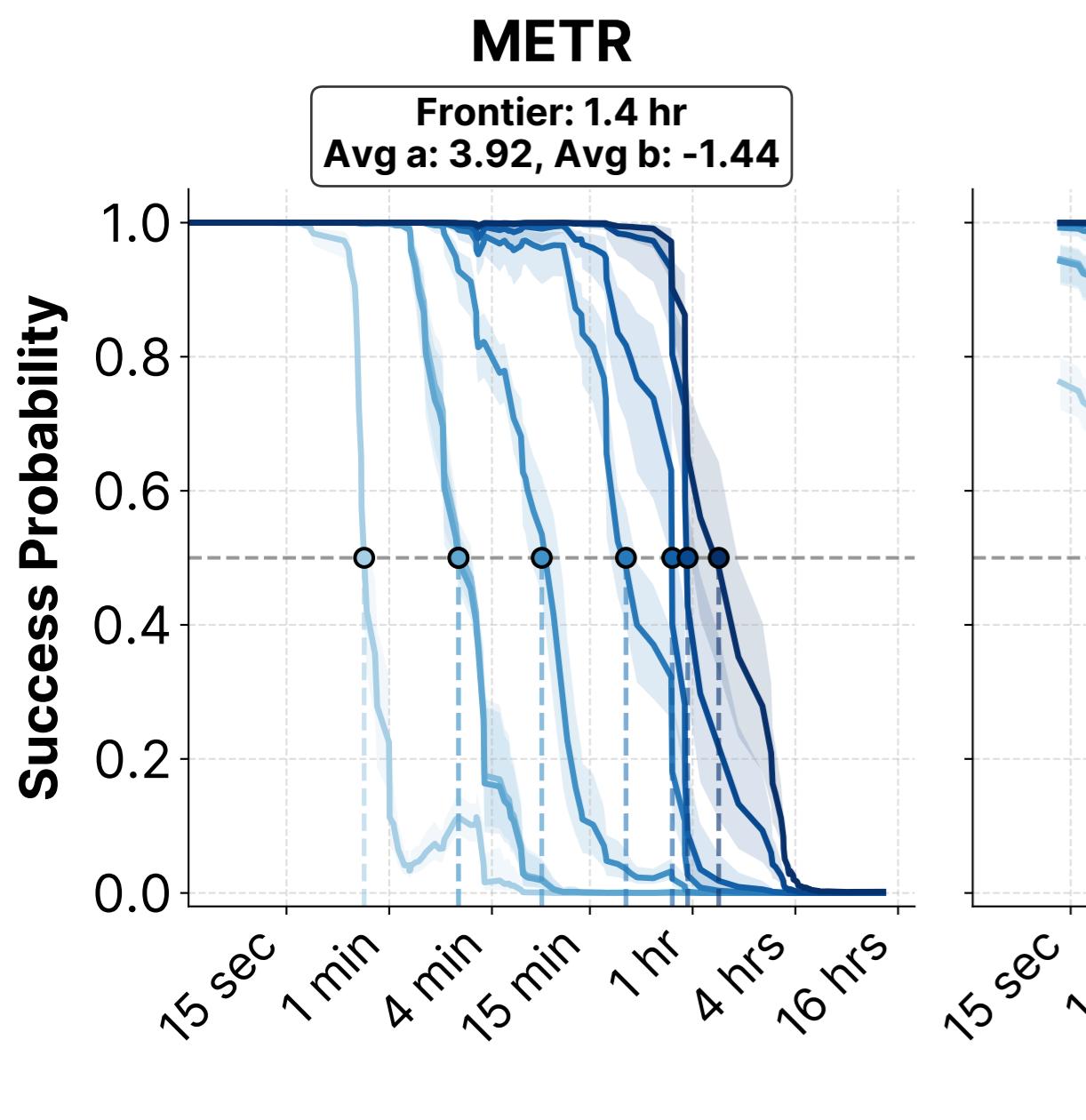
2024-02
(Claude 3.5 Sonnet)

2024-07
(o1-preview)

2024-12
(Claude 3.7 Sonnet)

2025-05
(Claude 4 Sonnet)

2025-10
(Claude 4.5 Opus)



Task Length