



Supplement of

Optimal model complexity for terrestrial carbon cycle prediction

Caroline A. Famiglietti et al.

Correspondence to: Caroline A. Famiglietti (cfamigli@stanford.edu)

The copyright of individual parts of the supplement might differ from the article licence.

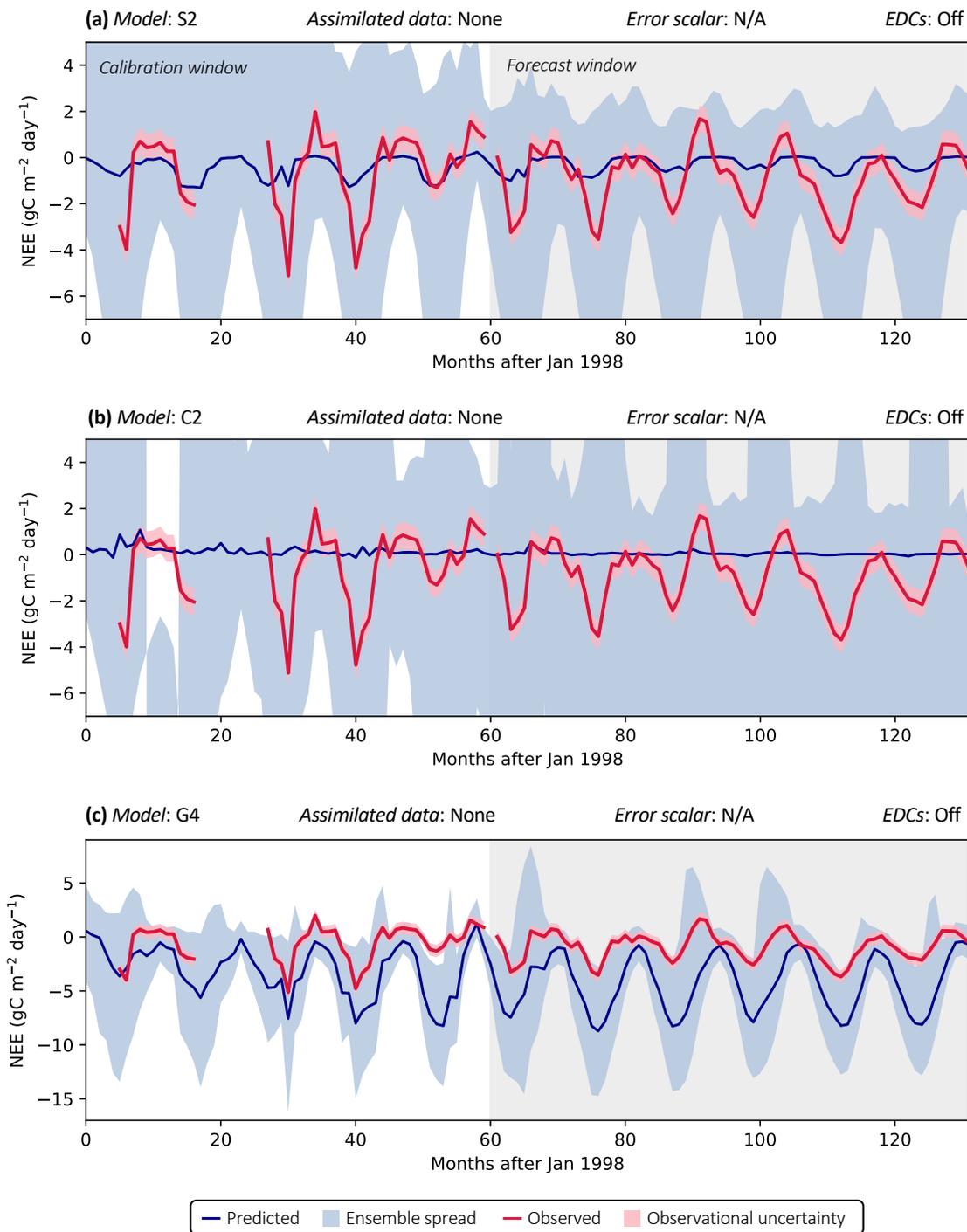
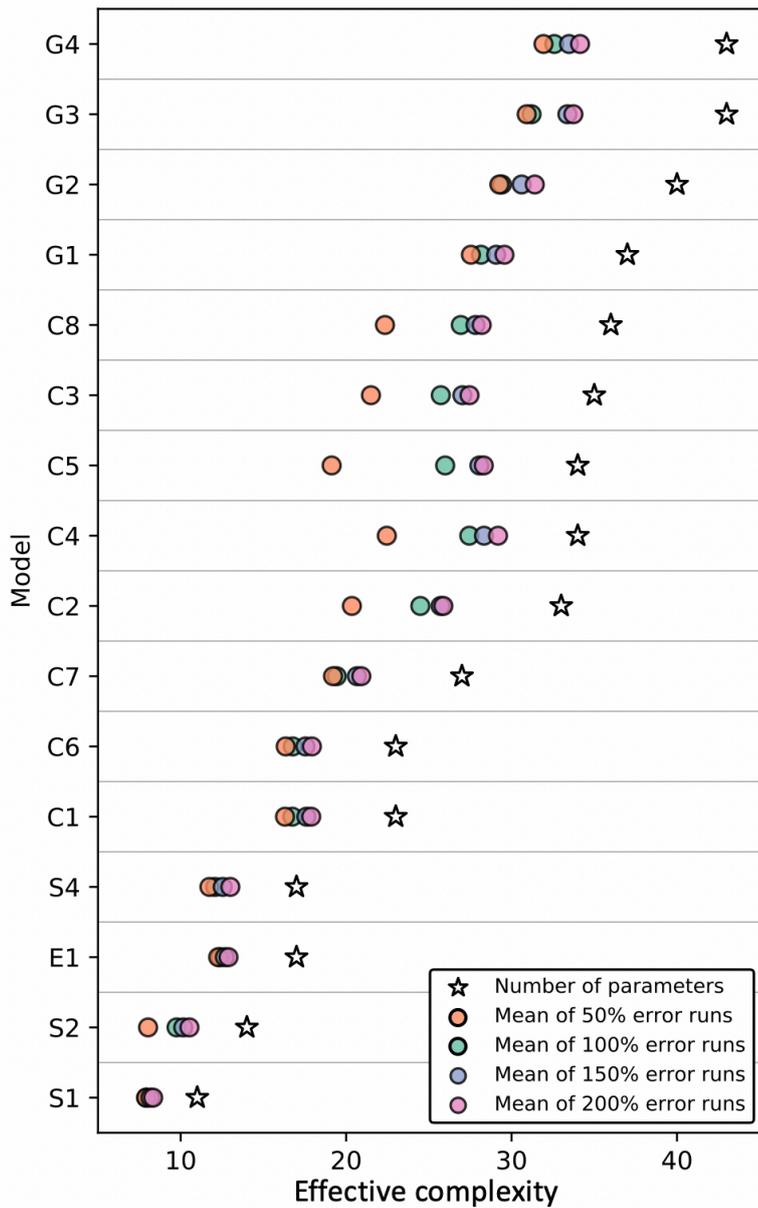


Fig. S1: Example model runs parameterized strictly using prior distributions at the FR-LBr site. For comparison, panels correspond directly to the models shown in Fig. 3. The calibration window—the first 5 years of the record—is shown in white and the forecast window is shaded gray. The ensemble spread (blue shading) encapsulates the 5th-95th percentile of runs.



25 **Fig. S2:** Effect of observational error scalar on effective complexity. Models are ordered from fewest (S1) to greatest (G4) number of parameters.

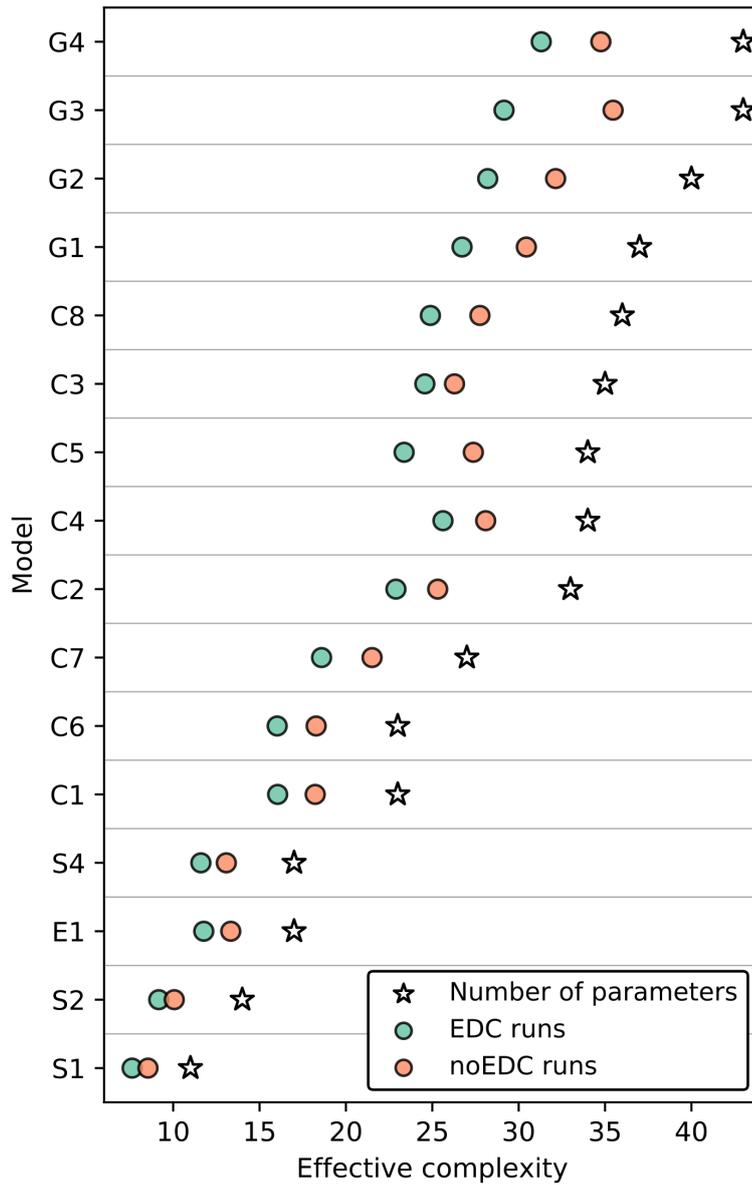
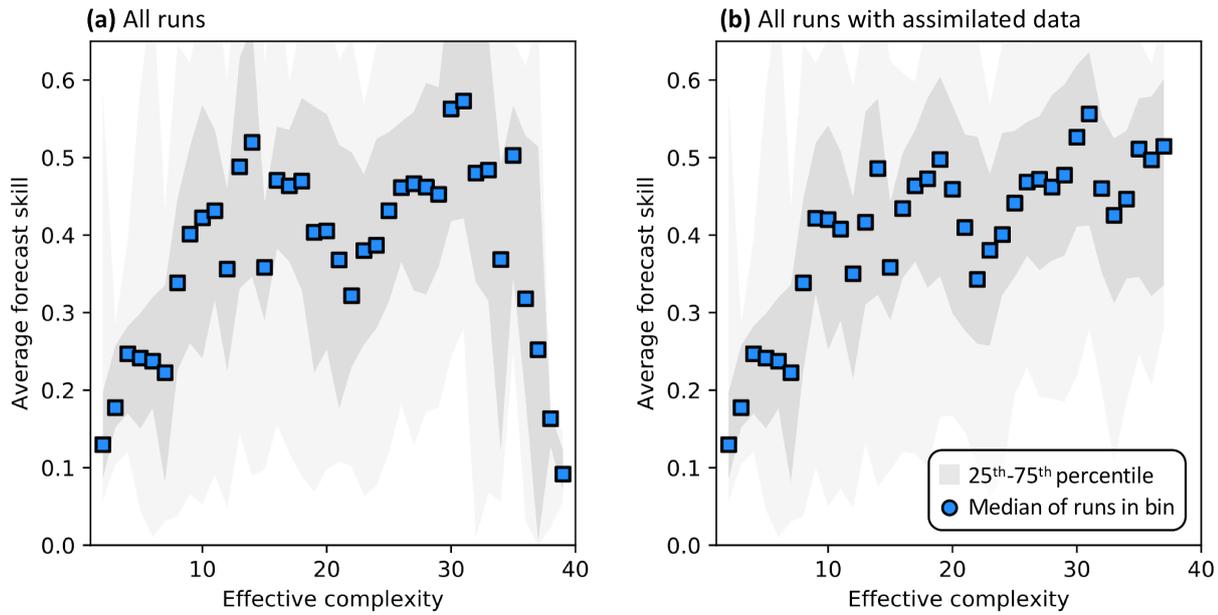
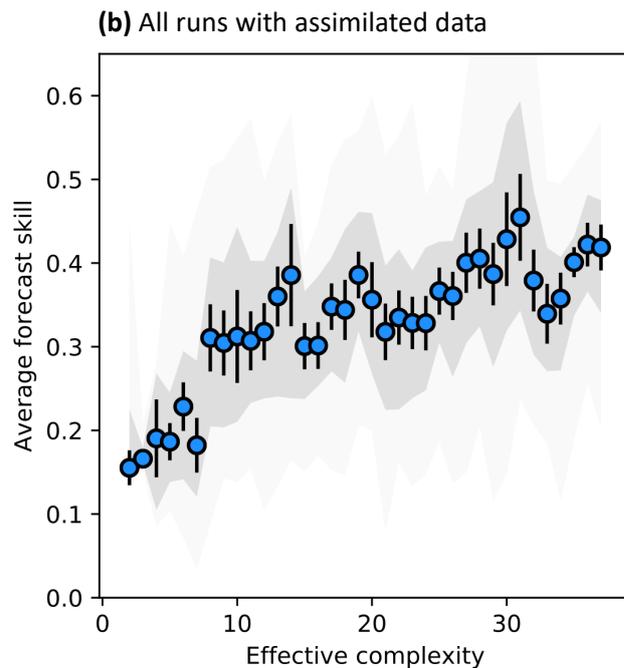
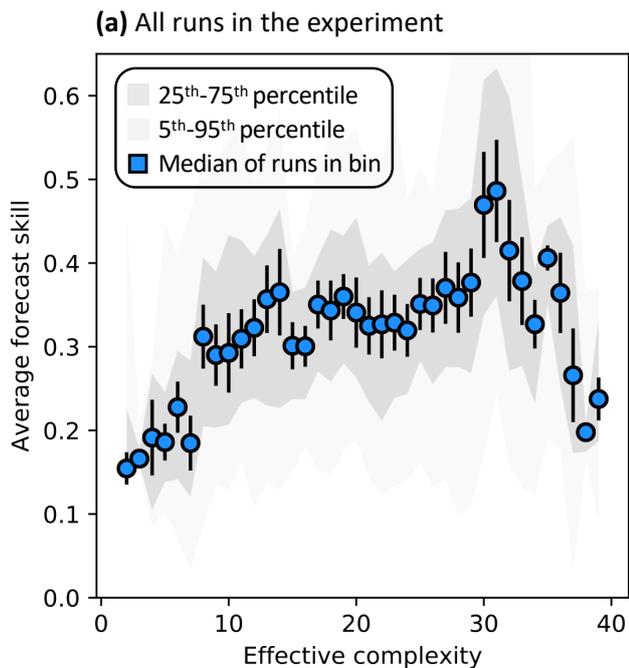


Fig. S3: Effects of EDCs on effective complexity. Models are ordered from fewest (S1) to greatest (G4) number of parameters.



30 **Fig. S4:** Complexity–skill relationship for LAI predictions. (a) All model runs included in the experiment; (b) all model runs for which data was assimilated. Dark gray shading spans the 25th to 75th percentile of runs; light gray shading spans 5th to 95th percentile; blue points are medians of complexity bins. Average forecast skill is computed using the histogram intersection metric.



35

Fig. S5: Complexity–skill relationship for NEE predictions, where a bootstrapping procedure has been performed to equalize number of runs within each complexity bin. Error bars represent one standard deviation of skill across the different bootstrap combinations.

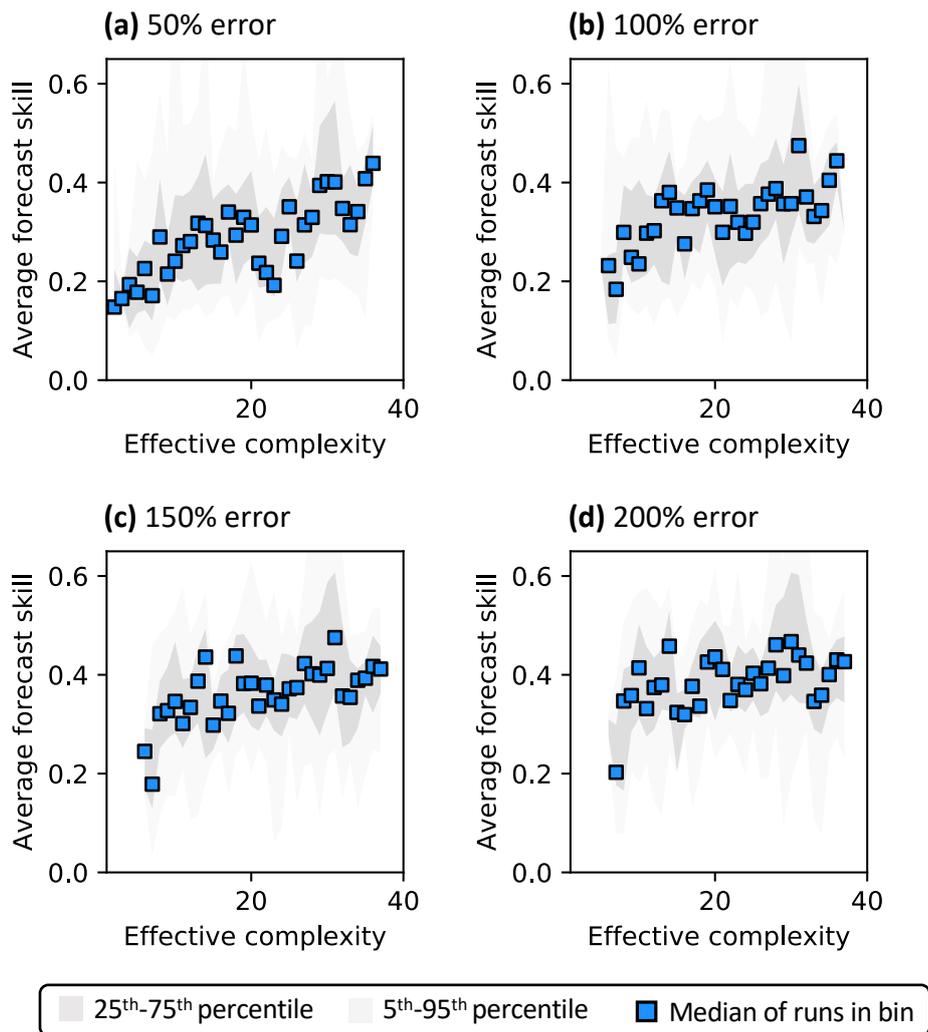


Fig. S6: Complexity–skill relationship for NEE predictions, split by observational error scalar (title of each subplot). Dark gray shading spans the 25th to 75th percentile of runs; light gray shading spans 5th to 95th percentile; blue points are medians of complexity bins. Average forecast skill is computed using the histogram intersection metric.

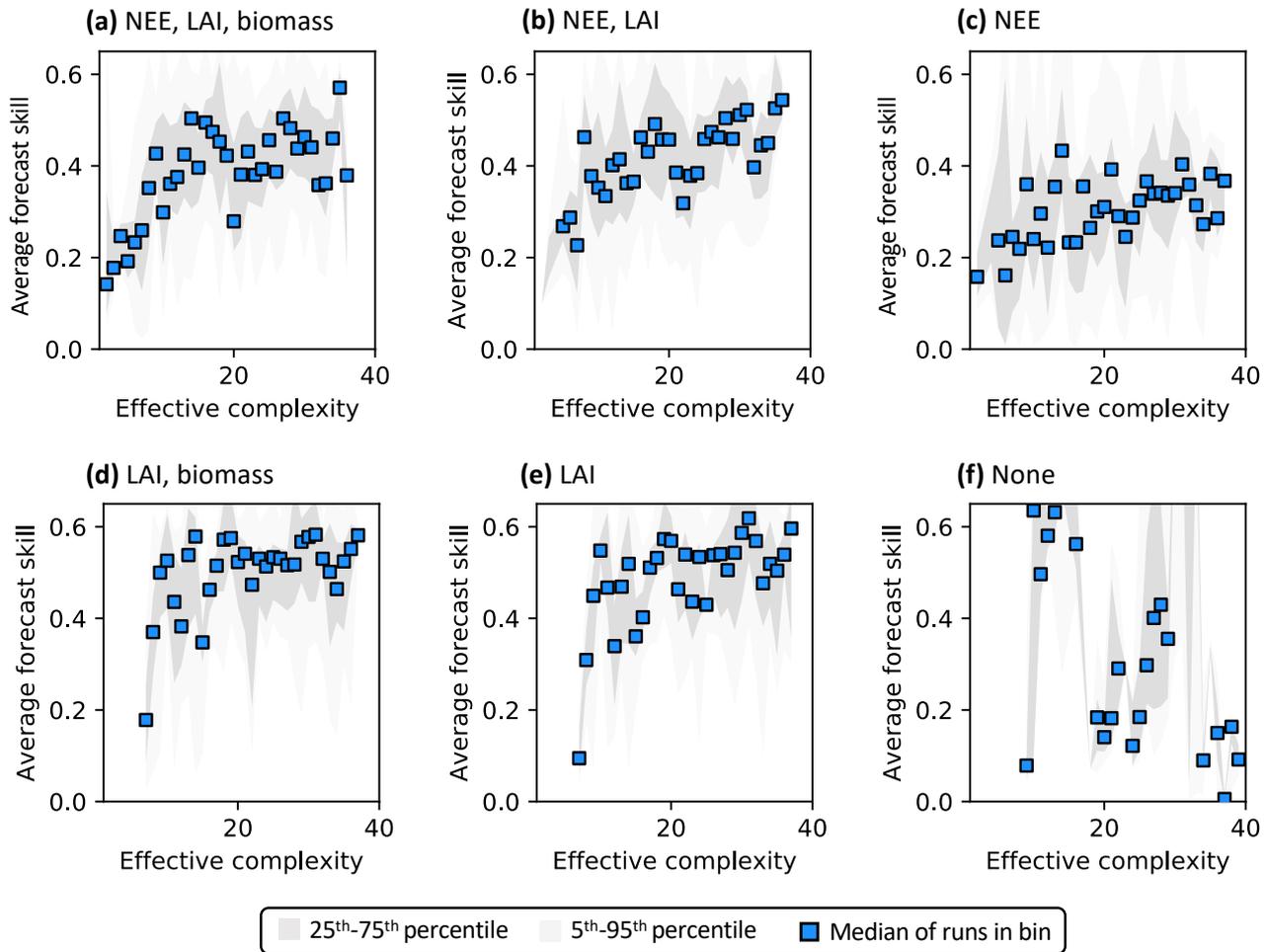


Fig. S7: Complexity–skill relationship for LAI predictions, split by assimilated data subset (title of each subplot). Dark gray shading spans the 25th to 75th percentile of runs; light gray shading spans 5th to 95th percentile; blue points are medians of complexity bins. Average forecast skill is computed using the histogram intersection metric. Ordering of subplots reflects strongest (a) to weakest (f) data constraint.