1. **Results** 
   1. **Comparable model predictions are observed regardless of the model scenario and stoichiometry strategy**

To test the usefulness of a lignin rate modifier, we fitted four different model scenarios to a set of publicly available data on 13C NMR compiled from the literature, using a more detailed stoichiometric litter description compared with traditional models. These four model scenarios included: i) no lignin protection and no oxidative enzyme cost (NPNE), ii) no lignin protection with invariant oxidative enzyme cost (NPWE), iii) lignin protection with invariant oxidative cost (PWOE), iv) lignin protection with variant oxidative cost (PWOV). Additionally, we modeled two microbial life history strategies to cope with stoichiometric imbalances: N retention and flexible CUE.

For the N retention strategy, we obtained comparable model calibration results of the five litter fractions (Figure 2), despite the high variability in the litter pools estimated using the molecular mixing model and NMR data (Figure S2). Many litter datasets showed an NSE > 0, indicating that the model performance is more accurate than merely using the mean of the observed values (Figure 3). Moreover, the same datasets had an RMSE close to zero, which is a good indicator of a robust model (Figure 3).

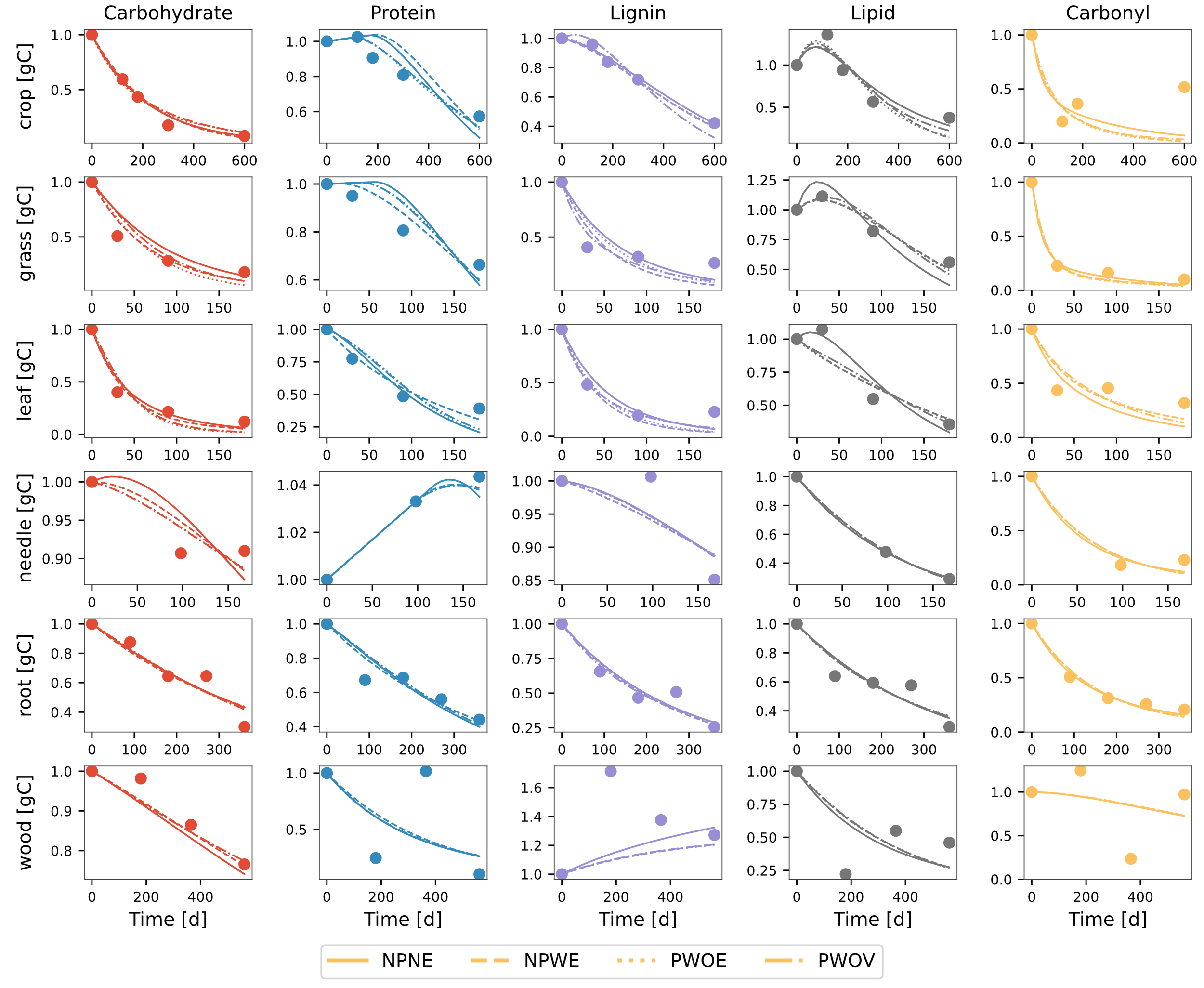


Figure 2 Comparison of modeled (different line styles) and observed (circles) changes in five litter pools (carbohydrate, protein, lignin, lipid, and carbonyl) for various litter types. The data points for the crop litter are from wheat straw buried at 15cm depth in sandy loam soil from Li et al. (2020); grass and leaf litter samples are from *A. mauritanicus* and *A. unedo*, respectively, from Bonanomi et al. (2013); needle litter is from *P. radiata* from Almendros et al. (2000); root litter samples are fine roots of *M. macclurei* from Wang et al. (2013); and wood litter of Mulga twigs from Mathers et al. (2007). Different line styles differentiate the four model variants. Model parameters and initial conditions used in the simulation were [0.01, 0.01, 0.008, 0.009, 0.01] d-1, and =1e-5 gN d-1; initial fraction of carbohydrate, protein, lignin, lipid, and carbonyl pools were 0.43,0.02,0.3,0.2,0.05, respectively; and the initial mass of litter was considered to be 1g on dry weight basis.

Model fits were heavily influenced by data availability or by the uncertainty in assigning certain NMR spectra to the litter fractions. For example, fractions like protein and carbonyl performed the worst regardless of the model scenario used, and this is highly notable when looking at the model performance metrics for each litter pool per model scenario (Figure 3). The model generally performed poorly for these pools, with a negative NSE, indicating that it performed worse than simply using the mean of the observations.

When comparing the two stoichiometry strategies, the model scenarios performed similarly (Figure S4), highlighting the fact that even such a detailed description of the litter fractions is insufficient to determine when a specific stoichiometry strategy is active.

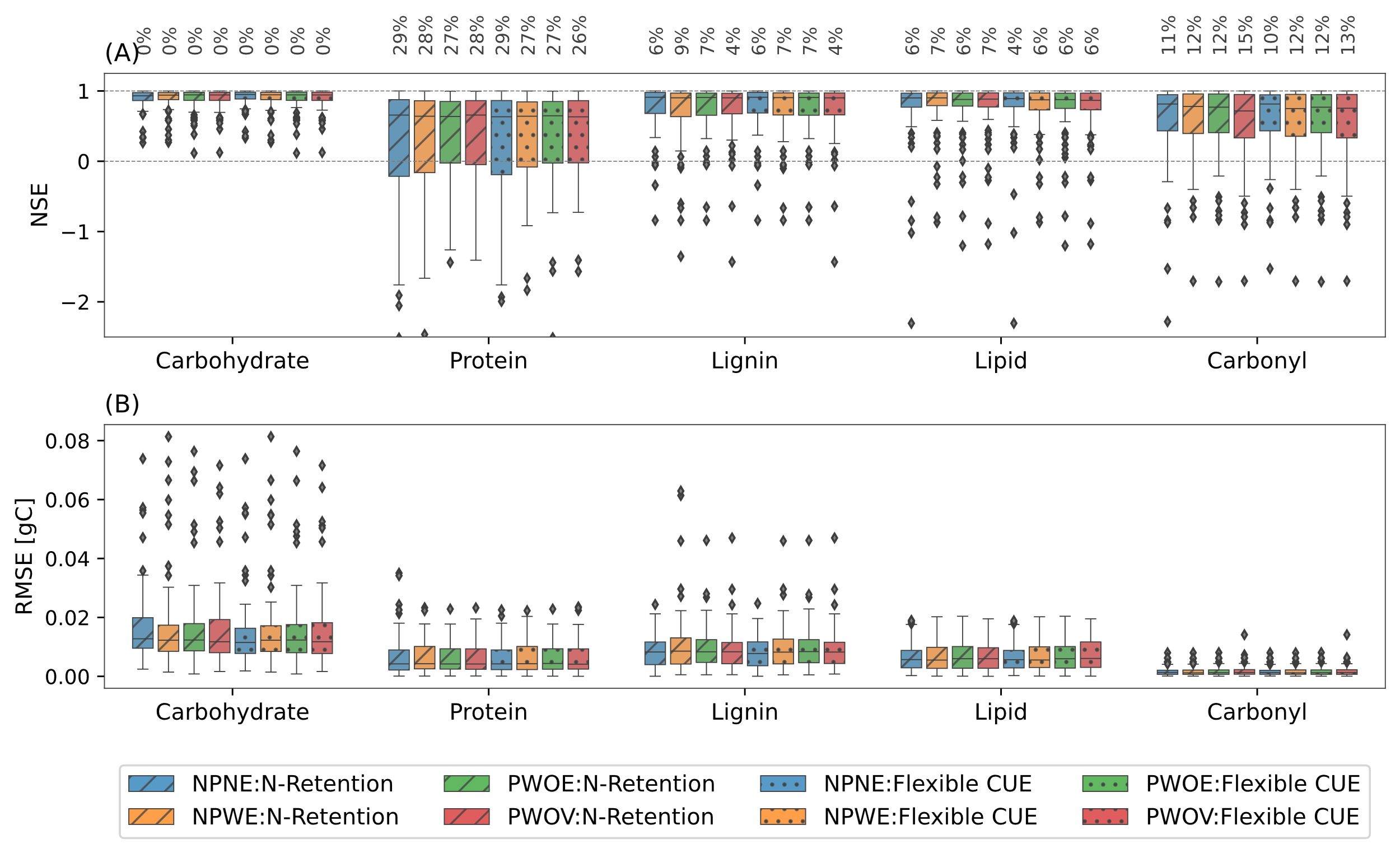


Figure 3 Boxplot of Nash–Sutcliffe modeling efficiency coefficient (NSE) (A) and root mean square error (RMSE in gC) (B) for five litter fractions for four model scenarios. . The boxplots show the interquartile range (IQR), with the median marked by a horizontal line. Whiskers extend to values within 1.5 times the IQR from the quartiles and outliers beyond the whiskers are indicated by individual markers. The percentage numbers in panel A show the % of litter dataset having NSE<0.

* 1. **Detailed description of litter fractions in models improves the constraining of kinetic model parameters**

One persistent issue common to all biogeochemical models is equifinality, or the inability to properly constrain model parameters, especially kinetic parameters. By adding more realism to the litter fractions compared to traditional models, we tested whether our proposed models could aid in constraining kinetic parameters. Overall, our model scenarios under the two stoichiometry strategies do not improve the uncertainty in the estimation of the decay rates of the litter fractions compared to more traditional models (Figure 3). The estimated parameters varied across many orders of magnitude, and these uncertainties in the model parameters will propagate to the model outputs as well. Following the parsimony principle, and given the similarity of parameter estimations across models, including simpler ones, our results suggest that models with fewer free parameters should be preferred.

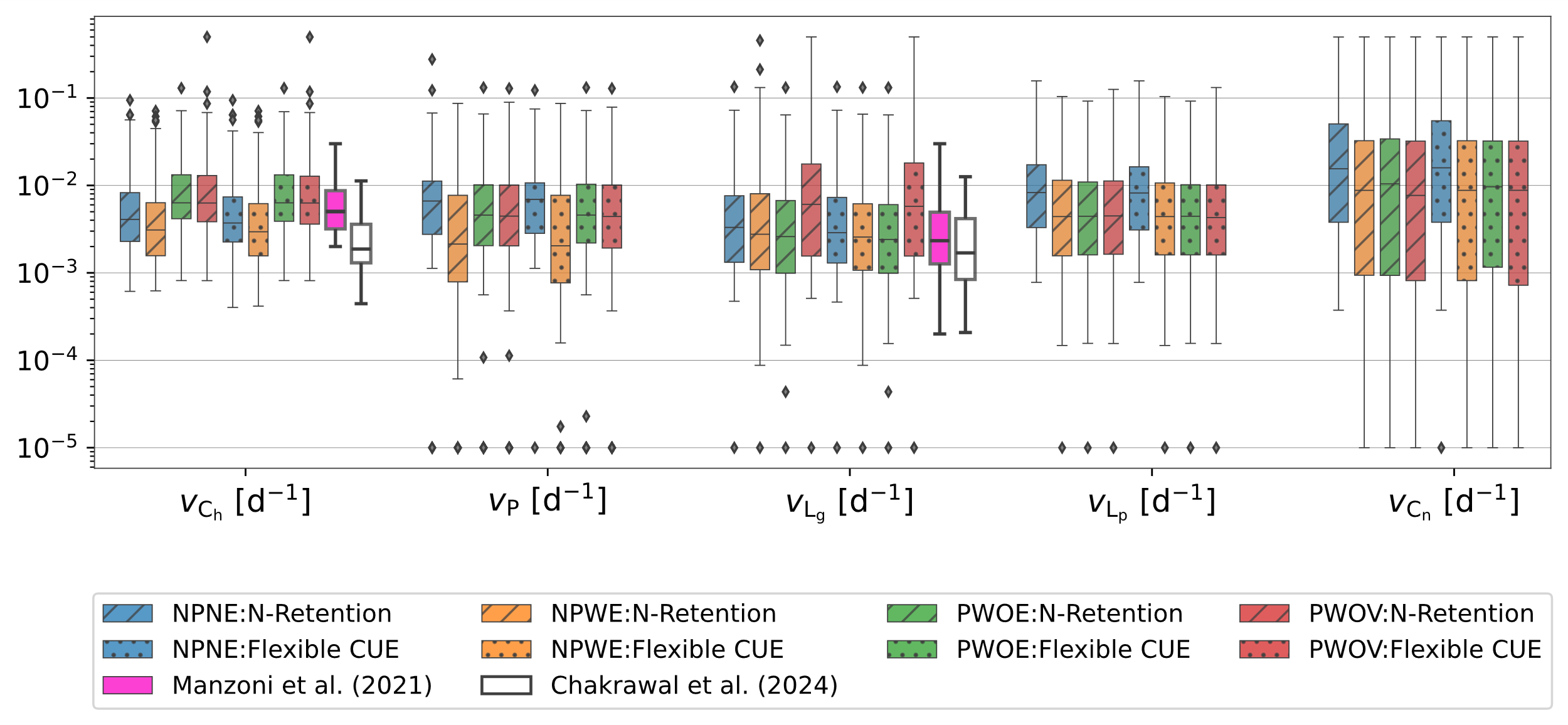


Figure 4. Distribution plot of estimated model parameters, i.e., the uptake rate constant for (A) carbohydrate, (B) protein, (C) lignin, (D) lipid, and (E) carbonyl for three model scenarios as different line styles. Note the log transformed values on the X-axis for each variable. Distribution plots illustrated by white and purple lines in panels A and C are the rates of carbohydrates, proteins, and oxidizable (representative of lignin) pools, respectively, taken from Manzoni et al. (2021), and nonaromatic (representative of carbohydrates and proteins) and aromatic (representing lignin and other aromatic compounds) pools from Chakrawal et al. (2024a).

* 1. **Model exploration: Mechanisms underlying litter decomposition under model scenarios and stoichiometry strategies**

We explored the dynamics of the different litter pools over time to unravel the mechanisms underpinning such dynamics under the four model scenarios and two stoichiometry strategies, namely, N retention (Figure 5) and flexible CUE (Figure S3). C pool dynamics did not differ under the proposed model scenarios, but the stoichiometry strategy did play a role in determining the mechanisms driving C pool dynamics, especially when accounting for lignin protection and the cost of oxidative enzymatic production. Under the N retention strategy, and without lignin protection and costs for oxidative enzyme production, microbes can take up C from each litter fraction without restriction, including lignin, at a maximum rate (Figure 5H, I, and J), independently of the lignin content in the system. Because uptake is at its highest, N retention in the form of biomass remains the longest compared to the other model scenarios (Figure 5G), and this is reflected by an increment in the protein pool (Figure 5B). These processes did not affect microbial CUE, as this remained constant under both stoichiometry strategies (Figure 5G and S3G). On the other hand, when enzymatic costs are accounted for, we observed a strong reduction in CUE (Figure 5G), reducing microbial biomass formation and consequently N retention as necromass (Figure 5F), which is in turn observable as losses in the protein pool (Figure 5B). Comparable results were also observed under a flexible CUE strategy (Figure S3).

Model scenarios, including lignin protection and either static or dynamic enzymatic investment, yielded similar results in all litter fractions, with only slight differences in the protein dynamics, mainly due to reductions in CUE. In both models, the effect of lignin protection is also observable in a slower degradation of carbohydrates compared to the other model scenarios (Figure 5A). The main impact of having an invariant enzymatic cost of production was seen in a decrease in lignin decay due to a bigger delay in decomposition as lignin concentrations reduced.

A diagram of a graph

AI-generated content may be incorrect.

Figure 5. Simulated temporal variation in carbohydrate (A), protein (B), lignin (C) lipid (D), and carbonyl (E) pools, and N retention factor (F) in four model scenarios using the N-retention strategy. The bottom panels (G-J) show the variation with lignin fraction of C use efficiency (CUE) (G), rate modifier for carbohydrate () and protein () (H), lignin rate constant ( (I), and rate modifiers for lipid () and carbonyl pools () (J). In panels G–J, the red areas or lines extend the range of variation of rate modifiers with lignin fraction, and black lines are the simulated range of p vs. L in each model scenario. In panel G, the red area illustrates the plausible range of variation in CUE under different model scenarios—calculated based on a minimum and maximum degree of reduction of total litter, set at 3.8 and 4.5, respectively, to determine . Model legends are as follows, **NPNE**: **n**o **p**rotection of carbohydrate and protein, and **n**o oxidative **e**nzyme cost, **NPWE**: **n**o **p**rotection carbohydrate and protein but **w**ith oxidative **e**nzyme cost, **PWOE**: **p**rotection of carbohydrate and protein **w**ith **o**xidative **e**nzyme cost but time invariant lignin rate constant, **PWOV**: **p**rotection of carbohydrate and protein **w**ith **o**xidative enzyme cost but time **v**arying lignin rate constant. Model parameters and initial conditions used in simulation were [0.01, 0.01, 0.008, 0.009, 0.01] d-1, and =1e-5 gN d-1; initial fraction of carbohydrate, protein, lignin, lipid, and carbonyl pools were 0.43,0.02,0.3,0.2,0.05, respectively; and the initial mass of litter was considered to be 1g on dry weight basis.

Previous works have suggested that microbial CUE varies negatively proportional to the litter C:N ratio, and thus, CUE decreases as N becomes limiting in the system. This is mainly observed in a microbial community that uses flexible CUE as a strategy to cope with stoichiometric imbalances. Our four model scenarios, following a flexible CUE strategy, are also able to reproduce this relationship, CUE vs. C:N ration in litter (Figure 6B and D), when comparing both average CUE and initial CUE. However, under an N retention strategy, all models, with exception of the simplest model (neither lignin protection nor costs of oxidative enzymes), can reproduce the patterns described in a flexible CUE strategy. In the no-protection, no-enzyme-cost model scenario, CUE is higher due to reduced oxidative enzyme investment. As decomposition progresses, microbes may become less N-limited and CUE increases, reaching close to maximum CUE values under C-limited conditions. As a result, the average CUE is higher than the initial values (Figure 6A).

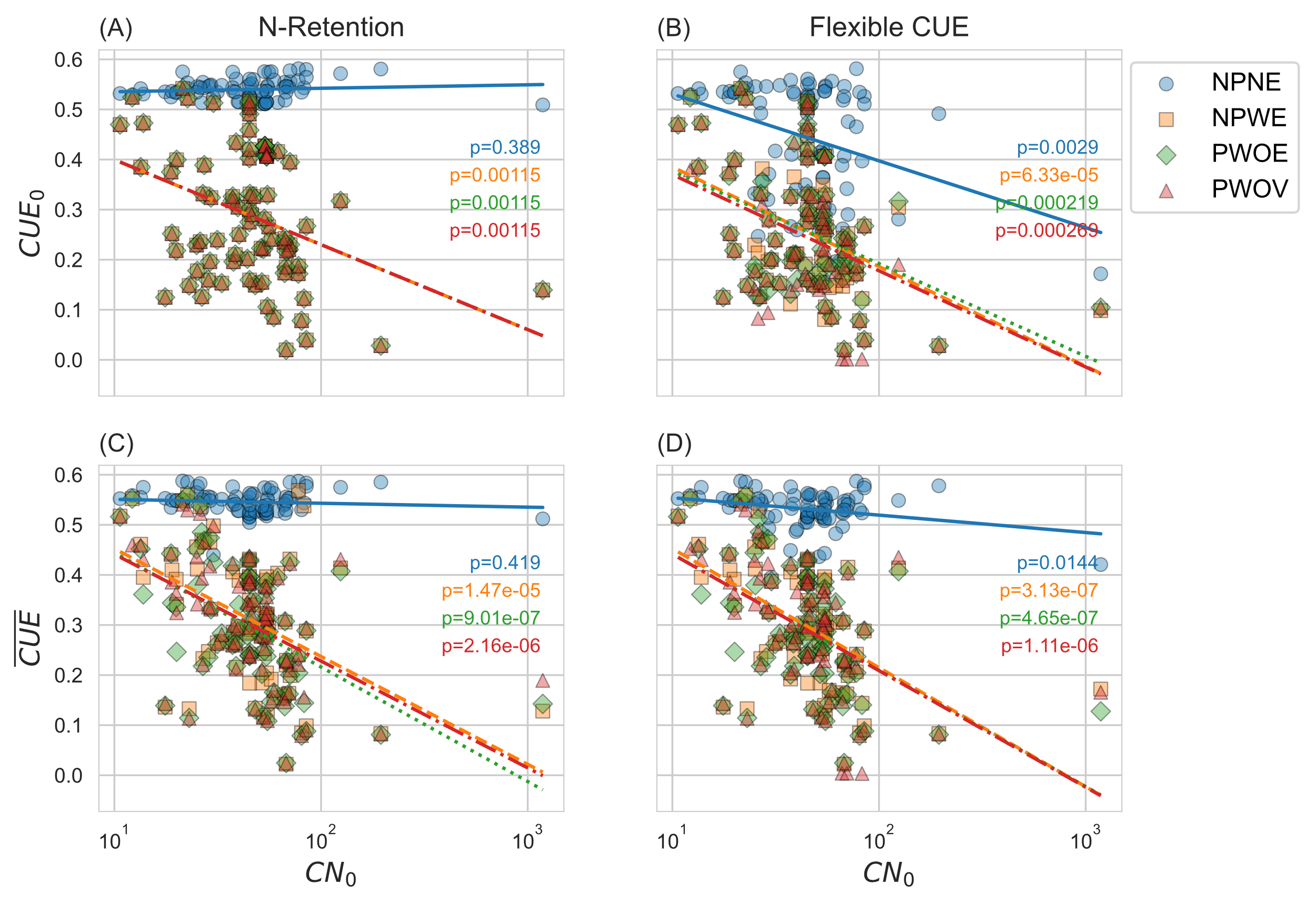


Figure 6 Variation in initial () and temporal average () carbon use efficiency under the N retention and flexible CUE adaptation strategies for four model scenarios.

1. **Discussion**

New technologies such as the Nuclear Magnetic Resonance (NMR) and Fourier Transform Ion Cyclotron Resonance mass spectroscopy (FTICR-MS) now allow us to accurately describe the composition of litter fractions and thus develop process-based models capable of incorporating more detailed process descriptors of the role of litter fractions on litter turnover rates. Taking advantage of this data, our model and model scenarios were used to understand the mechanisms of lignin protection in litter decomposition by incorporating a lignin rate modifier. We further analyze how such a rate modifier would affect microbial-driven processes related to litter turnover and their consequences and practicality in scaling them up into soil models.

* 1. **Can we use NMR data to constrain litter decomposition model parameters?**
  2. **Does the lignin rate modifier improve the calibration of the model?**
  3. **How do estimated parameters vary across three different models when constraints are used on the same dataset?**
  4. **Applications of a lignin modifier in soil models**

1. **SI**

Figure 2

Figure 5