Received: 20 November 2023 | Accepted: 4 March 2024 | Published online: 11 April 2024

DOI: 10.1002/saj2.20665

Soil Science Society of America Journal

ORIGINAL ARTICLE

Soil and Ecosystem Processes

Improving a nitrogen mineralization model for predicting unfertilized corn yield

Kathleen E. Arrington[1] | **Raziel A. Ordóñez**[2] | **Zoelie Rivera-Ocasio**[1] | **Madeline Luthard**[1] | **Sarah Tierney**[1] | **John Spargo**[2] | **Denise Finney**[3] | **Jason Kaye**[1] | **Charles White**[2]

[1]Department of Ecosystem Science and Management, The Pennsylvania State University, University Park, Pennsylvania, USA

[2]Department of Plant Science, The Pennsylvania State University, University Park, Pennsylvania, USA

[3]Department of Biology, Ursinus College, Collegeville, Pennsylvania, USA

Correspondence
Kathleen E. Arrington, Department of Ecosystem Science and Management, The Pennsylvania State University, University Park, PA, USA.
Email: kea106@psu.edu

Assigned to Associate Editor Steve Culman.

Funding information
Natural Resources Conversation Service, Grant/Award Number: NR20-08G010; USDA National Institute of Food and Agriculture Hatch Appropriations, Grant/Award Numbers: PEN04764A,1025969, PEN04977,7006665; USDA National Institute of Food and Agriculture Sustainable Agricultural Systems, Grant/Award Number: 2019-68012-29904; USDA National Institute of Food and Agriculture, Grant/Award Number: 2020- 51300-32178

Abstract

Crop N decision support tools are typically based on either empirical relationships that lack mechanistic underpinnings or simulation models that are too complex to use on farms with limited input data. We developed an N mineralization model for corn that lies between these endpoints; it includes a mechanistic model structure reflecting microbial and texture controls on N mineralization but requires just a few simple inputs: soil texture soil C and N concentration and cover crop N content and carbon to nitgrogen ratio (C/N). We evaluated a previous version of the model with an independent dataset to determine the accuracy in predictions of unfertilized corn (*Zea mays* L.) yield across a wider range of soil texture, cover crop, and growing season precipitation conditions. We tested three assumptions used in the original model: (1) soil C/N is equal to 10, (2) yield does not need to be adjusted for growing season precipitation, and (3) sand content controls humification efficiency (ε). The best new model used measured values for soil C/N, had a summertime precipitation adjustment, and included both sand and clay content as predictors of ε (root mean square error [RMSE] = 1.43 Mg ha^{-1}; $r^2 = 0.69$). In the new model, clay has a stronger influence than sand on ε, corresponding to lower predicted mineralization rates on fine-textured soils. The new model had a reasonable validation fit (RMSE = 1.71 Mg ha^{-1}; $r^2 = 0.56$) using an independent dataset. Our results indicate the new model is an improvement over the previous version because it predicts unfertilized corn yield for a wider range of conditions.

1 | INTRODUCTION

Microbial decomposition of soil organic matter (SOM) and plant residues can provide a source of N for crops, yet quantifying this resource requires a model that includes interactions

Abbreviations: CUE, carbon use efficiency; dY, delta yield; GDD, growing degree days; SOM, soil organic matter.

among plants, microbes, environmental conditions, and soil minerals (Daly et al., 2021). There has been a widespread call to include microorganisms more explicitly in biogeochemical models (Schimel, 2023) and in fertilizer recommendations for crops (Franzluebbers et al., 2022), but with no consensus on which microbial attributes might be required in different modeling contexts. At the same time, there has been a call to consider the role of mineral-associated nutrient pools in modeling SOM decomposition, including agricultural N fertility (Daly et al., 2021). In agroecosystems, a substantial portion of N uptake can come from indigenous nutrient pools (Cassman et al., 2002; Griesheim et al., 2023; Yan et al., 2020); thus, the ability to estimate N from organic sources (SOM and plant residues) offers an opportunity to improve fertilizer recommendations, providing economic and environmental benefits. In this paper, we advance efforts to include microbial processes and their interaction with soil minerals in a simple biogeochemical model that is the basis for an N fertilizer decision support tool for corn (*Zea mays* L.). Incorporating microbes and minerals in decision support tool models is a distinct challenge compared to mechanistic simulation models because decision support tools must be effective at a wide variety of sites with just a few inputs that are accessible to users such as farmers and agronomists.

A variety of approaches have been used to estimate the N contributions of SOM and cover crops to the following cash crop. Some of these include empirical equations based on laboratory incubations (Quemada & Cabrera, 1995; Sullivan et al., 2019; Franzluebbers et al., 2022), the presidedress soil nitrate test (White et al., 2023), and detailed crop-soil simulation models (Thapa et al., 2022; Woodruff et al., 2018). Here we focus on improving an N mineralization model (White et al., 2020) that uses simple biogeochemical equations to explicitly represent microbial decomposition and mineral-associated SOM. This model is neither an empirical relationship nor a simulation; it lies between these endpoints. It includes a system of equations based on biogeochemical theory, but with parameter and coefficient values estimated by fitting the theoretical equations to regional field data. The strengths of this model are that (1) the required inputs are accessible to many farmers and agronomists through existing commercial laboratory services, (2) it combines N mineralization and yield response to directly predict unfertilized corn yield, and (3) it can be adapted to a wide variety of sites and growing seasons, as we demonstrate here.

Our N mineralization model (White et al., 2020) includes microbial physiology based on the biogeochemical theory that N mineralization is a function of microbial carbon use efficiency (CUE) and the difference in C/N between microorganisms and their consumed substrates (Manzoni et al., 2012). CUE, the ratio of C used for growth (new microbial biomass) over C uptake from the soil, is an important control on whether N mineralization or immobilization will occur in the soil environment (Manzoni et al., 2012). The key microbial physiology term in our model is the humification efficiency (ε), a parameter representing the average CUE over a growing season, which can cover multiple generations of microbial decomposition (White et al., 2020). The ε term is modified by soil texture in the model, providing a link between byproducts of microbial decomposition and their preservation in soil via association with clay. In the previous calibration, the model-calculated values for ε led to good agreement between measured and predicted unfertilized corn yield, however if these equations are extrapolated to coarser-textured soils, the predicted value for ε would be negative for soils with more than 49% sand, possibly overestimating N mineralization and unfertilized corn yield (White et al., 2020). The focus of this study was to improve our N mineralization model so that it could be applied to sites with a broad range of soil textures.

Core Ideas

- An improved N mineralization model predicts unfertilized corn yield for a wide variety of conditions.
- The new model provides realistic estimates of microbial humification efficiency across a range of soil textures.
- Humification efficiency is affected more by soil clay content than sand content.
- The updated coefficients account for the influence of precipitation on corn yield.
- The improved model provides a foundation for site-specific N fertilizer recommendations.

The original dataset used to calibrate these equations had a moderate range of soil textures, cover crop C/N and growing season precipitation (White et al., 2020). Using experiments from both research station and commercial farms across Pennsylvania, we compiled a new dataset that includes a broader range of soil textures, a more uniform distribution of observations across the range of winterhardy cover crop C/N, and a wider range of summertime precipitation compared to the original dataset. With the new dataset, we tested three assumptions in the original model: (1) soil C/N can be set at fixed value of 10, (2) corn yield predictions do not need to be adjusted for growing season precipitation, and (3) sand content controls ε. Thus, the objectives of this study were to:

1. Assess the accuracy of the N mineralization model calibrated by White et al. (2020) to predict unfertilized corn yield using a robust new dataset for validation.
2. Reformulate and recalibrate the N mineralization model to include site-specific soil C/N and growing season precipitation using the new dataset.

TABLE 1 Experimental site-years in the new dataset, including number of observations, soil pH (1:1 in water), sand and clay content, soil C content, soil C/N, and unfertilized corn grain yield (dry basis) within each site.

Site	Year	Observations	pH	Sand (%)	Clay (%)	Soil C (%)	Soil C/N	Unfertilized corn yield[a] (Mg ha^{-1})
A	2021	4	6.8	18	32	1.59	9.5	11.1
B	2021	4	7.3	33	27	1.63	10.0	9.3
C	2019	16	6.8	14	45	1.57	9.2	5.7
D	2020	20	6.5	30	28	1.37	10.7	2.8
E1	2021	4	6.8	18	33	1.37	10.9	4.4
E2	2021	4	6.4	20	31	1.50	11.2	8.7
F	2021	4	6.7	17	30	1.82	8.8	11.6
G1	2021	4	6.9	67	12	1.43	11.1	8.1
G2	2021	4	7.2	47	19	2.35	10.9	8.0
5-N	2012	52	6.3	25	28	1.37	9.6	6.5

[a]Average unfertilized corn yield across all observations in each experiment.

3. Assess the new model accuracy using the original dataset (White et al., 2020) for validation.

As part of objectives 2 and 3, we tested the following hypotheses:

(H1) Incorporating measured soil C/N values into the recalibrated equations will improve the model fit compared to assuming soil C/N is equal to 10.

(H2) Incorporating summertime precipitation will improve the model fit.

(H3) Sand content will remain the best regulator of ε, consistent with White et al. (2020).

2 | MATERIALS AND METHODS

2.1 | Experimental sites

The new dataset was generated from 10 experimental site-years conducted over four growing seasons (2012 and 2019–2021) in central and southeastern Pennsylvania (Table 1). Sites E and G each included two fields, which were subdivided into two experiments based on differences in management (Site E) and contrasting soil textures (site G). Experiments E1 and E2 were adjacent fields at a research station, with similar soil properties but different crop rotation histories: E1 had a corn–soy (*Glycine max* L.) rotation with no cover cropping and occasional tillage, whereas E2 has been managed in a corn–soy–wheat (*Triticum aestivum* L.) rotation with no-tillage and a hairy vetch (*Vicia villosa* Roth) cover crop planted between wheat and corn since 2004. Site G included two fields at the same farm managed similarly but with different soil textures, indicated as G1 and G2 in Table 1. The long-term crop rotation history among the sites involved corn, soybean, and wheat in various frequencies. The long-term manure history varied across sites from none to annual applications depending on the location, but no manure was applied in the year of the corn yield response experiment at any site. At Site 5, some of the cover crop treatments were included in the original dataset (5-O) and the rest were included in the new dataset (5-N). Soil taxonomic classes were fine, mixed, semiactive, mesic Typic Hapludalfs (Sites A, C, E, F, and 5-N); fine-loamy, mixed, semiactive, mesic Typic Hapludults (Site B); fine-loamy, mixed, mesic Typic Hapludults (Site D); coarse-loamy, mixed, mesic Fluventic Dystrochrepts (Site G1); and coarse-loamy over sandy or sandy-skeletal, mixed, mesic Fluventic Dystrochrepts (Site G2). The original dataset was based on nine experimental site-years during three growing seasons (2012–2014) with details reported in White et al. (2020).

2.2 | Weather variables

Average air temperature, cumulative growing degree days (GDD), and cumulative precipitation for the month of April to September (the time frame when most corn crops are growing within the state of Pennsylvania) were calculated for each experimental site in both the new and original datasets (Table 2). The GDD variable was calculated using the typical equation, $\mathrm{GDD} = (T\mathrm{max} + T\mathrm{min}) \times 0.5 - T\mathrm{base}$, where Tmax corresponds to maximum temperature, Tmin is equal to minimum temperature, and T*base* represents the base temperature of 8°C for corn growth. The precipitation difference from the long-term monthly average (1980–2021) across sites for June, July, and August (101.5 mm month^{-1}) was calculated for each experimental site. In a few cases, fields were located near each other, in which case historical weather records were shared across these sites (see coordinates in Table 2). The historical weather data (1980–2021) was derived from the

TABLE 2 Environmental characteristics for each experimental site in the new and original datasets.

Site	Year	Latitude, longitude (°)	Average air temperature (°C ± SD)	Growing degree days (°C day)	Cumulative precipitation (mm)	Precipitation difference from long-term monthly average (mm month^{-1})
			(Apr.–Sept.)			(June–Aug.)
New dataset						
A	2021	40.0938, −76.4669	20.1 ± 5.3	2208	810	38.1
B	2021	39.8128, −76.6713	19.6 ± 5.3	2134	698	20.3
C	2019	40.7253, −77.9260	18.8 ± 4.2	1991	566	−10.7
D	2020	40.7227, −77.9184	17.7 ± 6.1	1774	415	−42.1
E1, E2	2021	40.7127, −77.9679	18.6 ± 5.1	1940	750	17.6
F	2021	40.0933, −76.0761	17.8 ± 4.9	1802	606	−11.0
G1, G2	2021	41.2030, −76.8054	19.1 ± 5.2	2033	733	17.3
5-N	2012	40.7135, −77.9509	18.4 ± 5.0	1914	593	−15.4
Old dataset						
1, 5-O	2012	40.7135, −77.9509	18.4 ± 5.0	1914	593	−15.4
2, 3, 4	2013	40.7202, −77.9355	18.1 ± 4.7	1849	468	−0.4
6	2013	41.0493, −76.7445	18.7 ± 4.9	1504	488	−14.7
7	2014	41.0546, −76.7442	18.4 ± 4.8	1492	527	5.8
8	2013	40.0863, −76.0677	19.3 ± 5.1	2066	547	25.8
9	2014	40.0863, −76.0677	19.5 ± 5.2	2117	650	14.6

Note: Weather information includes growing season (April—September) mean air temperature (°C), cumulative growing degree days (°C day), cumulative precipitation (mm), and summertime precipitation difference (mm month^{-1}) from the long-term monthly average during June, July, and August (1980–2021). Cover crop treatments from experimental Site 5 were divided between the original (5-O) and new (5-N) datasets.

North America Land Data Assimilation System 2 (Xia et al., 2012).

2.3 | Field management

Cover crops were planted in late summer and fall prior to the corn year of each experiment. Cover crops included both monocultures and mixtures of species made up primarily of winterhardy species, which were terminated in late spring prior to planting corn as a summer cash crop. Some sites had a single cover crop management practice present, while other sites, particularly research station sites, had multiple cover crop treatments established in randomized complete block designs, such as different species or different termination timings. Each cover crop treatment had four or five replicates, leading to a total of 116 observations (Table 3). At sites with tillage-based crop production (Table 3), prior to cover crop planting, fields were tilled with a chisel plow and seedbeds prepared with a cultimulcher. At tilled sites, cover crops were terminated with a chisel plow or high speed disk, and seedbeds for corn planting were prepared with a cultimulcher. At no-till sites, weeds were controlled with herbicides prior to cover crop planting in the fall and cover crops were terminated in the spring with herbicides and residues were left to decompose on the soil surface.

Corn was planted in May with populations ranging from 74,000 to 84,000 plants ha^{-1}. Most sites used herbicides to control weeds during the growing season, except for Sites F and G, which were organic production systems on commercial farms that used three–five passes of field cultivators for weed control. All sites were nonirrigated. Each site had a range of N fertilizer levels applied in a randomized complete block design with three–five replications, however our analysis here only utilizes corn yields from the unfertilized plots, which were randomly assigned within each block. Corn grain was harvested in the fall using various methods depending on the resources available at each site, including production scale combines, small plot combines, and by hand. Harvested plot sizes ranged from 6 m × 0.76 m at the smallest to 91 m × 9.1 m at the largest. Grain from each plot was weighed at harvest, the grain moisture content was measured, and yields were adjusted to 0% moisture for data analysis. For grain harvested by the small plot combine, moisture was measured by an on-board moisture sensor in the weighing bin of the combine (Almaco). Grain subsamples from plots harvested with commercial combines were measured for moisture content using a benchtop moisture tester (GAC 2100, DICKEY-john). Moisture content on grain harvested by hand was measured gravimetrically by drying five representative ears in an oven at 65°C until a constant weight was achieved (approximately 2 weeks) and measuring the mass lost during drying. The average unfertilized corn yield for each experiment is listed in Table 1.

2.4 | Field sampling and laboratory analyses

Soil samples were collected at the start of each experiment in spring, either by plot or block, depending on the site. Soil samples were composited from six to 12 soil cores (1.8-cm diameter, 0- to 20-cm depth) and submitted to the Penn State Agricultural Analytical Services Laboratory for analysis of sand, silt, and clay fractions using the hydrometer sedimentation method (Gee & Bauder, 1986). For soil C and N analysis, samples were dried, ground, and analyzed with a dry combustion elemental analyzer, as described by White et al. (2020). Table 1 shows the soil properties for the experiments in the new dataset.

Cover crop biomass was sampled in the late fall (prior to the first frost) for winterkilled species or in the spring prior to termination of winterhardy species. For each treatment plot, aboveground biomass was clipped from two randomly placed 0.5 m × 0.5 m quadrats, separated by species, dried, weighed, ground, and analyzed for C and N content, as described by White et al. (2020). Cover crop biomass N was measured separately for winterkilled and winterhardy species, based on respective samples collected before fall frost or before spring termination (Table 3). Site 5-N included mixtures of winterkilled and winterhardy species, reflecting separate sampling of species with different growth periods.

2.5 | Model description

We recalibrated the biogeochemical equations with the new dataset using the process described in White et al. (2020), except that instead of using fivefold cross-validation, we validated the new model with the original dataset of White et al. (2020). The recalibration process not only allowed for changes in the coefficients of the equations but also repeated the testing of which parameters were significant in the system of equations, allowing for the possibility of a new model structure. The system of equations included in the recalibration process is presented below:

Equation (1) calculates ε, allowing both sand and clay content as possible predictors:

$$\varepsilon = \text{int} + b_{\text{sand}}\,\%\text{sand} + b_{\text{clay}}\%\text{clay}. \quad (1)$$

The cover crop yield credit Y_{cc} was calculated with Equation (2) (White et al., 2016), where N_{wk} is the N content (kg N ha^{-1}) of winterkilled cover crops (measured in fall before termination by frost), N_{whcc} is the N content (kg N ha^{-1}) of winterhardy cover crops (measured in spring just before termination), $(C/N)_{whcc}$ is the C/N of winterhardy cover crops, measured in spring, and α_{wh} is the slope of the yield response to N that is potentially mineralized from winterhardy cover crop residues.

TABLE 3 Cover crop biomass N content (kg N ha^{-1}) for winterkilled species (fall) and winterhardy species (spring just prior to termination) and spring C/N for winterhardy species.

Site (tillage)	Cover crop Species[a] (Termination)[b]	Plots *n*	Cover crops Biomass N (kg N ha^{-1}) Winterkilled[c]	Winterhardy[c]	C/N (g g^{-1}) Winterhardy	Corn Yield (unfertilized) (Mg ha^{-1})
A (NT)	RY	4	–	35	16	11.1
B (NT)	CC	4	–	96	15	9.3
C (NT)	No cover[a]	4	–	5	16	7.3
	RY (Early)[b]	4	–	28	19	6.5
	RY (Medium)[b]	4	–	36	28	5.3
	RY (Late)[b]	4	–	45	33	3.9
D (NT)	No cover[a]	5	–	5	16	3.5
	RY (Early)[b]	5	–	44	23	3.5
	RY (Medium)[b]	5	–	59	33	2.6
	RY (Late)[b]	5	–	62	43	1.5
E1 (NT)	No cover	4	–	0	0	4.4
E2 (NT)	HV	4	–	152	13	8.7
F (T)	RG + CC	4	–	89	15	11.6
G1, G2 (T)	OA	4	–	21	17	8.1
5-N (T)	CA	4	–	109	24	6.2
	CA + RY + BA + RG	4	–	63	43	5.0
	CA + RY + FR + OA	4	19	60	34	5.7
	CA + RY + RC + HV	4	–	159	16	7.8
	CA + RY + SH + SB	4	5	68	38	4.9
	FR	4	30	37	19	6.8
	FR + OA + RC + HV	4	32	162	10	8.7
	FR + OA + SH + SB	4	31	8	26	6.5
	FR + OA + SS + FM	4	29	7	23	6.5
	OA	4	30	8	28	6.8
	SB	4	16	30	28	5.6
	SH	4	3	41	27	5.8
	8-MIX	4	27	162	15	8.4

Abbreviations: 8-MIX, all 8 species at Site 5-N; BA, barley (*Hordeum vulgare* L.); CA, canola (*Brassica napus* L.); CC, crimson clover (*Trifolium incarnatum* L. 'Dixie'); FM, foxtail millet (*Setaria italica* L. P. Beauv.); FR, forage radish (*Raphanus sativus* L.); HV, hairy vetch (*Vicia villosa* Roth); NT, no-till; OA, oat (*Avena sativa* L.), RC, red clover (*Trifolium pratense* L.); RG, ryegrass (*Lolium perenne* L.); RY, cereal rye (*Secale cereale* L.); SB, soybean (*Glycine max* L.); SH, sunn hemp (*Crotalaria juncea* L.); SS, sorghum sudangrass (*Sorghum bicolor var. bicolor* × *bicolor var. sudanense*); T, tillage.

[a]Weeds in the no cover treatments were sampled and treated as winterhardy biomass.

[b]Early, medium, and late refer to termination times for RY.

[c]At Site 5-N, winterkilled species were sampled in fall before frost and winterhardy species were sampled in spring before termination.

$$Y_{cc} = 0.0078 N_{wk} + \alpha_{wh} N_{whcc} + \left(1 - \frac{\varepsilon (C/N)_{whcc}}{10}\right). \quad (2)$$

Winterkilled and winterhardy cover crops are treated separately in Equation (2) due to differences in decomposition time before the corn growing season and the potential for losses of nitrogen from winterkilled cover crops through denitrification and leaching (Dean & Weil, 2009). Microbial decomposition of winterhardy cover crop residues is expected to result in N mineralization if $(C/N)_{whcc} < 10/\varepsilon$, and N immobilization otherwise, where ε is calculated from Equation (1). In Equation (2), $\alpha_{wh} = 0.026$ when N mineralization is expected and $\alpha_{wh} = 0.085$ when N immobilization is expected, based on a previous calibration of Equation (2) using a wide range of different winterkilled and winterhardy cover crop species planted in monocultures and mixtures (White et al., 2016). During the previous calibration of Equation (2), the ε term for winterkilled cover crops was zero, indicating no potential for microbial N immobilization, therefore only a single coefficient for N_{wk} (0.0078) is used and the C/N of winterkilled

cover crops is not required. For winterhardy cover crops, however, separate values of α_{wk} for N mineralizing and N immobilizing residues were identified, with N immobilization having a stronger impact on the yield response than N mineralization. The original calibration of Equation (2) also had separate values of α_{wk} for N mineralizing residues depending on tillage, with tillage-incorporated residues having a slightly greater value (0.026) than no-till residues (0.022). However, there were no observations in the calibration dataset of immobilizing residues in no-till sites, so a value of α_{wk} for that condition could not be estimated. Because differences in α_{wk} between tillage and no-tillage for mineralizing residues were slight, and because values of α_{wk} for both mineralizing and immobilizing residues were available for tillage systems, we adopted the tillage-based coefficients for use in the current study.

The yield credit from soil organic matter (Y_s) is calculated with Equation (3), where α_s is a calibrated coefficient selected for the best-fit model (representing the yield response due to mineralization of SOM), C_s is the total soil C (percent), and $(C/N)_s$ is the soil C/N.

$$Y_s = \alpha_s C_s\left(1 - \frac{\varepsilon\,(C/N)_s}{10}\right). \quad (3)$$

In White et al. (2020), soil C/N and microbial C/N were both assumed to be 10, which cancelled them out of the equation for the SOM yield credit. Here we tested the measured $(C/N)_s$ at each site as an input parameter in the recalibration process while maintaining the assumption that microbial C/N is equal to 10.

Equation (4) calculates the predicted unfertilized corn yield (Y_t) from Y_s and Y_{cc}, assuming a quadratic response to N additions. The g_{Yield} coefficient is calibrated to the training dataset. We included no intercept in this equation because the intercept term was found to be not significant in White et al. (2020) and had a high covariance with other model parameters.

$$Y_t = \left(Y_s + Y_{cc}\right) + g_{Yield}\left(Y_s + Y_{cc}\right)^2. \quad (4)$$

Since the new dataset had a wider range of precipitation (compared to the long-term average for June, July, and August), we tested adding a precipitation adjustment to Equation (4), allowing for summertime precipitation to influence the predicted unfertilized corn yield:

$$Y_t = \left(Y_s + Y_{cc}\right) + g_{Yield}\left(Y_s + Y_{cc}\right)^2 + b_{rain} \times \text{PrecipDev}, \quad (5)$$

where b_{rain} is the coefficient for the precipitation adjustment term and PrecipDev is the difference between the monthly average precipitation for June, July, and August at a site in the year of the experiment and the long-term monthly average precipitation for June, July, and August across sites and spanning from 1980 to 2021, which was 101.5 mm month^{-1}. The rationale for adding the precipitation adjustment is to ensure that when PrecipDev is set to zero, the model is calibrated to average rainfall conditions.

2.6 | Statistical analyses

We assessed the accuracy of the original N mineralization model developed by White et al. (2020) by comparing its predictions to measured unfertilized corn yield for the new dataset. Model fit was evaluated by calculating its root mean square error (RMSE) and the coefficient of determination (r^2) and comparing these statistics to those presented in White et al. (2020). Testing our three hypotheses required assessing 12 potential new model structures developed from a factorial of the various model inputs (Table 4). Four groups of models were developed to test the factorial of soil C/N = 10 versus measured soil C/N in Equation (3) (Hypothesis 1) and summertime rainfall adjustment in Equation (5) versus no rainfall adjustment in Equation (4) (Hypothesis 2). Within each group of soil C/N and rainfall adjustment levels, three potential models were tested, where ε in Equation (1) was regulated by sand alone, clay alone, or sand and clay together (Hypothesis 3). For each of the 12 potential models, we evaluated the model fit based on the Akaike information criterion (AIC) of the calibration dataset and the RMSE of both the calibration and validation datasets (White et al., 2020). Coefficients for the system of equations in each of the 12 potential models were fitted in the NLMIXED procedure of SAS 9.4, which also calculates the AIC statistic. The RMSE of each fitted model was calculated based on a tabulation of the residuals between predicted and observed unfertilized yields. Although testing for effects of tillage versus no-tillage soil management on the model performance was not the primary objective of this study, after identifying the best model from the testing described above, we evaluated how fitting separate coefficients for α_s based on tillage or no-tillage in Equation (3) of the best model affected the coefficients and prediction accuracy.

3 | RESULTS

3.1 | Characteristics of the new and original calibration datasets

The experimental sites and cover crop treatments included in the new calibration dataset contain a wide range of soil and cover crop properties that are well suited to calibrating N mineralization models to predict unfertilized corn yield. The new dataset had a similar range of soil C content and unfertilized corn yield as the original dataset of White et al. (2020) but included some sites with coarser textured soils (Figure 1a).

TABLE 4 Potential models developed from the new calibration dataset and their associated fit statistics: Akaike information criterion (AIC), ΔAIC, and root mean square error (RMSE).

		Equation 1		Equation 3	Equation 4/5	Model fit statistics				
		ε		Y_s	Y_t					
		Predictors				Calibration			Validation	Calibration + validation
Group	Model	Sand (%)	Clay (%)	Soil C/N	PrecipDev	AIC	ΔAIC	RMSE (Mg ha^{-1})	RMSE (Mg ha^{-1})	RMSE (Mg ha^{-1})
1	a	X	X	Measured	No	460.2	33.8	1.67	1.93	1.78
	b		X			467.8	41.4	1.74	1.78	1.76
	c	X				478.2	51.8	1.82	1.87	1.84
2	a	X	X	Assumed = 10	No	475.1	48.7	1.78	1.97	1.86
	b		X			486.9	60.5	1.89	1.77	1.84
	c	X				490.9	64.5	1.92	1.87	1.90
3	a	X	X	Measured	Yes	426.4	0	1.43	1.71	1.55
	b		X			432.7	6.3	1.48	1.69	1.57
	c	X				445.4	19	1.57	1.95	1.73
4	a	X	X	Assumed = 10	Yes	443.8	17.4	1.54	1.67	1.59
	b		X			451.9	25.5	1.61	1.63	1.62
	c	X				464.1	37.7	1.70	1.77	1.73

Note: The ΔAIC statistic is the difference between each model and the model with the smallest AIC value (Model 3a). RMSE was calculated for the calibration dataset, the validation dataset (White et al., 2020) and the two datasets combined. Humification efficiency (ε) was calculated with Equation (1), the soil organic matter credit (Y_s) was calculated with Equation (2), and unfertilized corn yield (Y_t) was calculated either with Equation (4) or (5), with or without the precipitation adjustment (PrecipDev). Group 1 compared to Group 2 tests Hypothesis 1 (soil C/N). Group 3 compared to Group 1 tests Hypothesis 2 (precipitation). Within each group, submodels a, b, and c test Hypothesis 3 using three combinations of soil texture variables (% sand and % clay individually and combined) to predict ε. For Equation 1, X's indicate which soil texture components were used to calculate ε for each submodel.

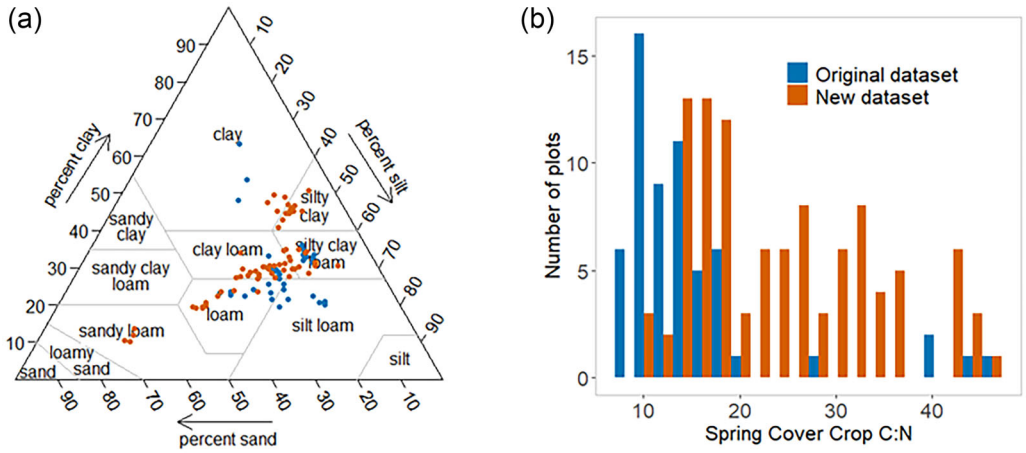

FIGURE 1 Comparison of (a) soil textures and (b) cover crop C/N for the observations in the original and new datasets. Data are shown at the plot level here, whereas in Table 1, % sand and % clay are averages for each experiment.

The new dataset also has a more uniform distribution of winterhardy cover crop C/N than the original dataset (Figure 1b). The original dataset included several species of winterhardy legumes (hairy vetch and several species of clover), resulting in most observations with cover crops (88%) having winterhardy C/N less than 20 (Figure 1b). In contrast, only 42% of observations with cover crops in the new dataset had a winterhardy C/N less than 20. In part, this difference in winterhardy C/N is due to fewer plots in the new dataset with winterhardy legumes. In addition, Sites C and D in the new dataset included three different termination dates for cereal rye, resulting in a wide range of winterhardy C/N within each experiment.

Table 3 illustrates the variability in unfertilized corn yield within sites that have similar soils but differences in cover crop species or cover crop management practices. Although soil properties were similar for the two fields in Site E (Table 1), the unfertilized corn yield for E2 was approximately twice that of E1 (8.7 vs. 4.4 Mg ha^{-1}; Table 1). Experiment E2 had a high winterhardy cover crop N content along with a low C/N from the hairy vetch cover crop, which would likely result in N mineralization as the cover crop residues decompose. Thus, the higher unfertilized corn yield in Experiment E2 potentially reflects N supplied to corn by the cover crop residues compared to Experiment E1, which had no N contribution from cover crops.

Sites C and D provide another example of the potential effects of management on unfertilized corn yield. Both experiments included a treatment with no cover crop and treatments with a cereal rye cover crop terminated at three different times in the spring, corresponding to increasing cover crop maturity and C/N. In these experiments, the average unfertilized corn yield declined with later termination dates (Table 3). The higher C/N (28–43) of the medium and late-terminated rye most likely caused N immobilization, resulting in reduced N availability to corn compared to plots with early-terminated rye or no cover crop.

Across sites and years, the average growing season temperatures ranged from 17.7 ± 6.1 to 20.2 ± 5.3°C (Table 2). The temperature variation between the hottest and coldest environment was only 2.5°C, a magnitude consistent with the typical temperature differential observed between the eastern and central regions in Pennsylvania. The cumulative growing season GDDs had a wide range across sites, from 1504 to 2244 GDDs. The cumulative precipitation during the growing season exhibited a wide range across both years and sites (Table 2), spanning from 415 to 875 mm. The driest year (2020) accumulated only half of the rainfall compared to the long-term normal precipitation. Across the six geographic locations where experiments were conducted, the long-term (1980–2021) average monthly precipitation in June, July, and August was 101.5 mm month^{-1}. The precipitation difference from the long-term monthly average for June, July, and August exhibited a broader range of variation in the new dataset as compared to the original dataset. This variation ranged from 38.1 mm month^{-1} higher (Site A) to 42.1 mm month^{-1} lower (Site D) than the long-term monthly average (Table 2).

3.2 | Mineralization models to predict unfertilized corn yield

When the original N mineralization model structure (White et al., 2020) was used to predict unfertilized corn yield for the new dataset, predictions had relatively poor accuracy (Figure 2, RMSE = 2.65 Mg ha^{-1}; r^2 = 0.14). This contrasts with the accuracy of the model calibration to the original dataset (RMSE = 1.58 Mg ha^{-1}; r^2 = 0.62) reported in White et al. (2020), which was assessed using fivefold cross-validation, rather than with an independent validation dataset.

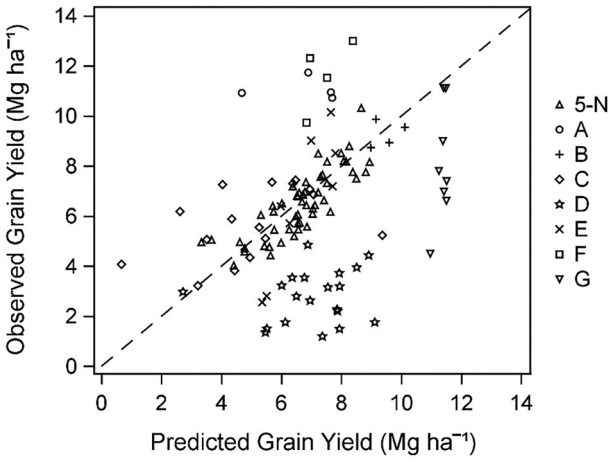

FIGURE 2 Comparison of observed and predicted unfertilized corn yield for the new dataset using the original model structure and calibration (White et al., 2020). Dashed line is the 1:1 relationship and symbols correspond to the experiment codes listed in Table 1. Site 5 plots were divided between the original and new datasets (5-O and 5-N, respectively). The accuracy of the original model assessed with this independent dataset was RMSE = 2.65 Mg ha^{-1}, $r^2 = 0.14$.

While many observations in the new dataset were close to the 1:1 relationship between predicted and observed values, unfertilized yields at several sites were consistently either over- or underpredicted. The original model overpredicted unfertilized corn yield at most plots in Site D, which had the greatest monthly precipitation deficit in the new dataset (Table 2). The original model similarly overpredicted yield at Experiments G1 and G2, which had coarser textured soils than the sites included in the original dataset. In contrast, the original model underpredicted the unfertilized corn yield for Site A, which had the greatest monthly precipitation surplus in the new dataset (Table 2). These examples illustrate how the original model, which was developed from experiments during average growing seasons and a moderate range of soil textures did not perform as well for growing seasons with higher or lower precipitation than average or for sites with coarser textured soils.

Our three hypotheses can be evaluated based on the AIC and RMSE values for each of the 12 potential models in Table 4. For Hypothesis 1 (that using measured soil C/N will improve the model accuracy), we compared these statistics for models within Groups 1 and 3 (measured soil C/N) against the models within Groups 2 and 4 (soil C/N assumed to be equal to 10). For all three combinations of predictors of ε (submodels a, b, and c), the calibration fit was improved by including measured soil C/N in the model, as indicated by lower AIC and RMSE values (Table 4). While the incorporation of measured soil C/N did not always improve the validation fit, when the accuracy of calibration and validation data were assessed together, there was a consistent pattern that including measured soil C/N in the model improved the overall accuracy across the two datasets. Therefore, Hypothesis 1 is supported: using measured soil C/N improves model fit and accuracy as compared to assuming soil C/N is equal to 10.

Hypothesis 2, that accounting for summertime precipitation will improve the model fit, was tested by comparing the AIC and RMSE values for two sets of models. These statistics for Group 2 models compared to Group 4 models indicate the effect of incorporating a precipitation adjustment while using an assumed soil C/N. Similarly, comparing these statistics for the Group 1 and Group 3 models shows how the precipitation adjustment affects the model fit when measured values are used for soil C/N. The AIC and RMSE values were lower for Group 4 compared to Group 2 models and for Group 3 compared to Group 1 models for the calibration and validation datasets individually and combined (Table 4). Therefore, Hypothesis 2 is supported: incorporating the precipitation adjustment (Equation 5 compared to Equation 4) improved the model fit.

Hypothesis 3, that sand content will remain the best regulator of ε, was tested by comparing the three submodels (a, b, and c) within each group. For the calibration dataset, submodel a, which used both sand and clay content to regulate ε, had lower AIC and RMSE values (Table 4) indicating a better model fit. In addition, the "a" submodels consistently had the lowest AIC values, indicating the improvement in model fit outweighed the complexity of an additional predictor. However, for the validation dataset, the lowest RMSE values within each group were for submodel b, which used clay content to regulate ε. When the accuracy of calibration and validation datasets combined was assessed, the lowest RMSE values were submodel b (Groups 1 and 2) and submodel a in Groups 3 and 4. Because the models in Groups 3 and 4 were more accurate due to the precipitation adjustment, we conclude that the texture submodel a (including both sand and clay content) is the preferred regulator of ε across the two datasets. Therefore, Hypothesis 3 was incorrect: sand content alone was not the best predictor of ε.

The best model emerging from the various hypothesis tests was model 3a, which included both clay and sand content, measured soil C/N, and summertime precipitation (Table 4). This model structure had the lowest AIC (426.4) and the lowest RMSE for the calibration dataset and the combined calibration and validation datasets (1.43 and 1.55, respectively, Table 4). The ΔAIC values in Table 4 show the difference between each candidate model and model 3a. Models with a ΔAIC value greater than 10 do not have support as the best model to approximate the dataset (Burnham & Anderson, 2004). The second-best model (according to the calibration dataset) is model 3b, with a ΔAIC value of 6.7. The only difference between model 3a and 3b is whether to include the % sand to calculate ε in Equation (1). Similarly, the primary difference between model 3a and the original calibration was

TABLE 5 Comparison of parameters and coefficients for Equations (1), (3), and (4) in the original calibration and in the new calibration of the best performing model (3a).

		Estimate		95% Confidence interval	
		Original	New	New	
Equation	Parameter	(White et al., 2020)	(Model 3a)	(Model 3a)	
Equation (1) (ε)	Intercept	0.774	0.0136	−0.2023	0.2294
	b_{clay}	Removed[a]	0.0111	0.0072	0.0150
	b_{sand}	−0.0157	0.0053	0.0022	0.0085
Equation (3) (Y_s)	α_S	9.85	12.76	9.88	15.66
Equations (4) and (5) (Y_t)	g_{Yield}	−0.0217	−0.0240	−0.0290	−0.0191
	b_{Rain}	Not used[b]	0.0478	0.0331	0.0625

Note: Humification efficiency (ε) is predicted with Equation (1), the yield credit due to soil organic matter (Ys) is predicted with Equation (3), and unfertilized corn yield (Yt) is predicted without a precipitation adjustment (original model; Equation 4) or with a precipitation adjustment (new model 3a; Equation 5).

[a]Not statistically significant.

[b]b_{Rain} was not included as a potential predictor in White et al. (2020).

in Equation (1), which calculates ε using % sand (Table 5). A comparison of the parameter estimates for all 12 candidate models (Table S1) highlights the importance of the choice of soil texture components in the parameter estimates for Equation (1).

The influence of clay content relative to sand content on ε is illustrated by comparing the calculated values for ε with both calibrations across the soil textures in both the new and original datasets (Table 6). Across both datasets, the values for ε calculated with the new calibration range from 0.34 to 0.73 with the new calibration, compared to a range of −0.28 to 0.55 with the original calibration. Site 2 in the original dataset has the highest clay content (55%) and has the highest value for ε (0.73) with the new calibration. With the original calibration, which was based only on sand content, Site 2 had a moderate value for ε of 0.47. Experiments G1 and G2 had the highest sand content in the new dataset (67% and 47%, respectively) and had calculated ε values of −0.28 and 0.03 with the original calibration. Under the new calibration, model 3a calculates ε for these experiments as 0.50 and 0.48.

Figure 3 shows the recalibrated model fit (Model 3a) using the calibration (new) and validation (original) datasets. Model 3a had a slightly better calibration fit (RMSE = 1.43 Mg ha^{-1}; $r^2 = 0.69$ in Figure 3a) compared to the model in White et al. (2020) (RMSE = 1.58 Mg ha^{-1}; $r^2 = 0.62$). Model 3a also performed much better on an independent validation dataset: RMSE = 1.71 Mg ha^{-1}; $r^2 = 0.56$ in Figure 3b compared to RMSE = 2.65 Mg ha^{-1}; $r^2 = 0.14$ in Figure 2.

To assess the effect of tillage system on the coefficients and performance of Model 3a, we calculated the mean bias error (MBE) of the predicted unfertilized yield for the calibration data separated by tillage type. Tilled observations had an MBE of −0.28 Mg ha^{-1}, while no-till observations had an MBE of 0.36 Mg ha^{-1}. The biases suggest the model was slightly underpredicting the yield in tilled cases and overpredicting the yield in no-till cases. However, the bias for each tillage type was small compared to the overall RMSE of 1.43 Mg ha^{-1}. Fitting separate values of α_s (Equation 3) by tillage type resulted in an α_s of 14.03 for tillage and 11.63 for no-tillage, reduced the AIC by 7.6 units and slightly reduced the RMSE to 1.37 Mg ha^{-1} (Table S2). However, the confidence intervals for the separate estimates of α_s were overlapping, indicating the differences were not statistically significant. Furthermore, all refitted parameter estimates of the tillage-specific Model 3a were within the confidence limits of the parameter estimates of the original Model 3a, suggesting that variations due to tillage effects are insignificant compared to the overall uncertainty in model coefficients. Finally, the tillage-specific Model 3a had reduced accuracy on the validation dataset (RMSE = 1.72 Mg ha^{-1}) compared to the original model 3a (RMSE = 1.71 Mg ha^{-1}). Therefore, the bulk of evidence suggests that the original model 3a is sufficient to predict unfertilized yield in both tilled and no-till systems using a single value of α_s.

4 | DISCUSSION

4.1 | Interpretation of model improvements

While our model does not fully simulate the processes affecting N mineralization/immobilization of soil and plant residues, it is consistent with frameworks that propose integrating plant, soil, and microbe interactions in C and N cycling models (Cotrufo et al., 2013; Daly et al., 2021). Our model captures several important controls on the decomposition of plant residues and SOM, including microbial ε (Equation 1) and the C/N of both cover crops and SOM (Equations 2 and 3). Some models of plant residue decomposition assume a fixed CUE, and hence a single critical C/N value

TABLE 6 Calculated humification efficiency (ε) for sites in the new and original datasets using the coefficients for Equation (1) in the original and new calibrations (Table 5).

Site	Sand (%)	Clay (%)	ε (Calculated with Equation 1)	
			Original calibration (White et al., 2020)	New calibration (Model 3a)
New dataset				
A	18	32	0.49	0.46
B	33	27	0.26	0.49
C	14	45	0.55	0.59
D	30	28	0.30	0.48
E	19	32	0.48	0.47
F	17	30	0.51	0.43
G1	67	12	-0.28	0.50
G2	47	19	0.03	0.48
5-N	25	28	0.39	0.46
Original dataset				
1	26	25	0.37	0.43
2	20	55	0.47	0.73
3	15	33	0.53	0.47
4	15	33	0.54	0.46
5-O	25	28	0.39	0.46
6	18	21	0.49	0.34
7	28	27	0.33	0.47
8	29	23	0.31	0.43
9	35	23	0.23	0.46

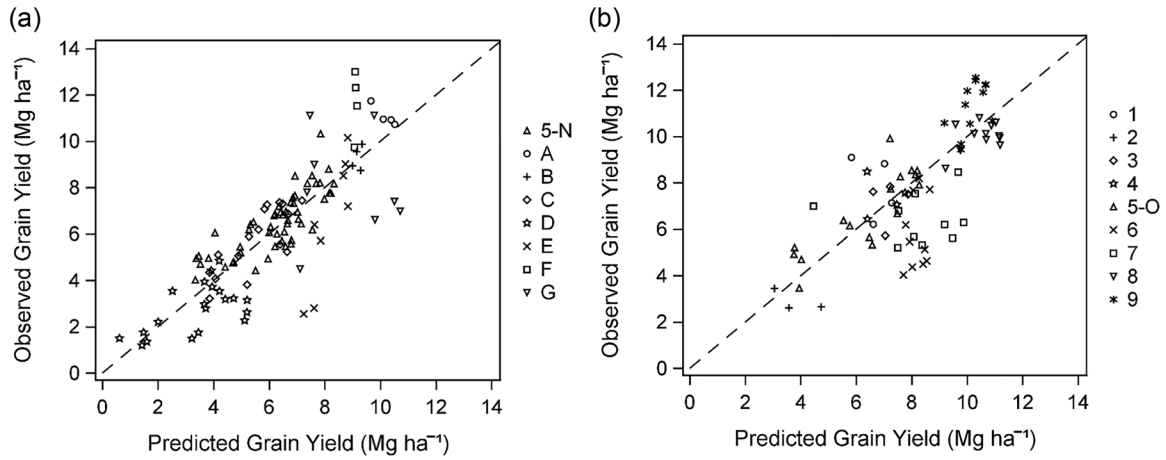

FIGURE 3 Comparison of the new model applied to (a) the new dataset (calibration fit) and (b) the original dataset (validation fit). Dashed line is the 1:1 relationship, and symbols correspond to the experiment codes listed in Table 1 for the new dataset and White et al. (2020) for the original dataset. Observations from experiment 5 were divided between the new and original datasets (5-N and 5-O, respectively). The accuracy of model 3a using the new dataset for calibration was RMSE = 1.43 Mg ha^{-1}; r^2 = 0.69 (panel a, calibration) and RMSE = 1.71 Mg ha^{-1}; r^2 = 0.56 using the original dataset for validation (panel b, validation).

for the division between N mineralization and immobilization (Thapa et al., 2022; Woodruff et al., 2018). However, our model allows ε (average CUE over the growing season) and the critical C/N to vary as a function of soil texture, explicitly accounting for the interactions among microbes, minerals, and substrates in the soil environment. The original model calibration (White et al., 2020) advanced the previous equations for predicting the N mineralization of cover crop residues (White

et al., 2016), both by adding a yield credit for SOM (Equation 3) and by allowing ε to vary as a function of soil texture (Equation 1). This study improves the original model calibration (White et al., 2020) by (1) testing the incorporation of measured soil C/N (Equation 3), (2) testing the addition of a precipitation adjustment (Equation 5), and (3) recalibrating Equation (1) using sites with a broader range of soil textures.

While we found that measured soil C/N provided an improvement in model fit for the calibration dataset and the combined datasets (Table 4), the new dataset had a relatively narrow range of soil C/N values (8.8–11.1; Table 1), which was also true of the original dataset (White et al., 2020). Our results support using a measured value of soil C/N in Equation (3), however if this information is not available, an assumed value of 10 may be used, as in White et al. (2020). Further research could perform similar testing using datasets with a broader range of soil C/N values as well as test incorporating measured values for microbial C/N based on the equation in White et al. (2020). Also, further development of the model presented here could test subdividing SOM into particulate organic matter (POM) and mineral-associated organic matter (MAOM), which are distinct components of SOM (Lavallee et al., 2020). Daly et al. (2021) present a conceptual model in which the supply of bioavailable organic N comes from a combination of POM and MAOM, depending on the POM N supply relative to the soil mineral sorption potential of organic N.

We found that the addition of the precipitation adjustment consistently improved the model fit, as measured by lower AIC and RMSE values (Table 4). For each corresponding set of models, adding a precipitation adjustment provided a greater improvement in model fit compared to adding a measured soil C/N value (Table 4). Also, a comparison of Figure 2 and Figure 3a indicates an improvement in the prediction of unfertilized yield with the new model for sites A and D, which had the greatest differences above and below normal monthly summer precipitation (Table 2). Including the precipitation adjustment ensures that our calibration of the N mineralization model to predict unfertilized yield was not biased by high or low rainfall at certain sites. Since the new model coefficients reflect the influence of precipitation, when a precipitation adjustment of zero is used, the model is truly calibrated to predict unfertilized yield for average rainfall conditions.

In the new calibration (Model 3a), both sand and clay content are significant predictors of ε, both with positive coefficients. This contrasts with the original calibration in which sand was the only significant soil texture predictor of ε, with an inverse relationship between sand content and ε (Table 5). However, the coefficient for percent sand is less than half the value of the coefficient for percent clay in Model 3a (Table 4). So, as the soil texture components sand, silt, and clay vary, if a percentage point of clay is replaced by a percentage point of sand, ε will decrease for both calibrations, although the decrease will be smaller for the new calibration. The positive relationship between clay content and ε reflects the ability of clay particles to protect organic matter from biological attack through adsorption and encapsulation within aggregates (Baldock & Skjemstad, 2000). It is also consistent with other studies that report positive relationships between clay content and soil organic N content (Ge et al., 2019; Hassink, 1994).

For soils with greater than 49% sand, the original model predicts negative values for ε (White et al., 2020). In Table 6, the two lowest ε values (−0.28 and 0.03) for site G are due to the high sand content being outside the range of soil textures included in the dataset used to calibrate the original model. Negative values of ε are biologically implausible and prevent Equation (2) from ever predicting N immobilization. Values of ε that are still positive but approach zero could also inflate predictions of N mineralization from SOM and cover crop residues. Overestimating N mineralization can lead to overprediction of unfertilized corn yield, which was the case when applying the original model (White et al., 2020) to most of the fields at site G (Figure 2). In contrast, the ε values calculated for site G with the new calibration (Table 6) resulted in a balance between fields where yield was overpredicted and fields where yield was underpredicted (Figure 3a).

Across the broad range of soil textures in both datasets, the estimates of ε by Model 3a (Table 6) fall roughly within the interquartile range of typical values for CUE in soil (~0.4–0.7; Manzoni et al., 2012). With the original calibration (White et al., 2020), the estimates of ε for both datasets (Table 6) are mostly below the median CUE values for soil (~0.55; Manzoni et al., 2012). However, considering that the best model to regulate ε has varied within this study (Table 4) and across others (White et al., 2020, 2016), potential refinements of this component of the model should be pursued through further mechanistic research.

Tillage is a factor that affects N mineralization of cover crops (Drinkwater et al., 2000; Rodriguez et al., 2023; Sainju & Singh, 2001) but does not appear to be a factor that needs to be explicitly accounted for with separate parameter estimates in our N mineralization model. Primary tillage at the time of cover crop termination as well as cultivation for weed control during the growing season can increase N mineralization rates due to disruption of soil aggregates, soil aeration, and mixing of cover crop residues into the soil where the microclimate is more conducive to decomposition (Finney et al., 2015; Poffenbarger et al., 2015). However, tillage also leaves soil bare, making it prone to greater evaporation losses and larger temperature swings later in the growing season (Clark et al., 2007; Tollner et al., 1984), factors that may reduce decomposition rates. Over an entire growing season, the period for which our model is calibrated, differences in early-season decomposition and N mineralization rates due to tillage

practices may even out by the end of the growing season. For instance, Poffenbarger et al. (2015) found that the N mineralization or immobilization from hairy vetch and rye cover crop residues was initially faster when residues were incorporated into the soil with tillage, but by the end of the growing season, decomposition and N loss from the residues were similar between tillage and no-tillage treatments. Although calibrating our Model 3a with separate values for $\alpha_{\text{S-tillage}}$ and $\alpha_{\text{S-no-tillage}}$ resulted in a slightly greater coefficient value for tilled observations, which supports the concept of greater N mineralization due to tillage, the difference between tillage types was not statistically significant and was both within the current uncertainty of the original model estimate. These results suggest that parameterizing separate coefficients based on tillage type would be overfitting the model given the currently available data.

While the improved model can be applied to fields across Pennsylvania, the equations would need to be recalibrated before using in other regions. Our work highlights the importance of a calibration dataset that is diverse in terms of soil textures, growing season precipitation, and cover crop C/N. The original and new datasets had a similar number of experiments (nine and 10, respectively), while the latter had more cover crop treatments within some experiments leading to more observations: 116 (new) compared to 73 (original; White et al., 2020). The greater variety of cover crops in the new dataset led to a more even distribution of cover crop C/N over the range of typical values (Figure 1b) and a greater variety of sites expanded the range of soil textures in the new dataset compared to the original dataset (Figure 1a). It should also be noted that coefficients in Equation (2) were derived from a previous calibration (White et al., 2016) with a dataset of 199 unfertilized corn yield observations in experiments where the yield following cover crops could be compared to the yield in a fallow plot in the same field. In that dataset, there was a high degree of redundancy in cover crop N content and C/N among treatments, so efficiently designed experiments could reduce the number of observations needed to calibrate such a model in the future. For instance, small plot experiments that compare a legume monoculture, a grass monoculture, a grass-legume mixture and a fallow treatment could efficiently generate a dataset with a broad range of cover crop N and C/N. Similarly, experiments that manipulate the termination timing of cover crops, such as C and D in our calibration dataset, can efficiently produce a dataset with a wide range of cover crop N contents and C/N. Multi-treatment experiments such as these paired with simple on-farm trials that strategically sample the variety of soil textures and soil C concentrations expected in a region could develop a calibration dataset with a minimum of observations required. Thus, for others interested in calibrating these equations for another region, we recommend strategically designing experiments to obtain a calibration dataset that represents the range of soil textures, growing season conditions, and cover crops present across the region.

4.2 | Potential uses

The N mineralization model was developed to predict unfertilized corn yield as part of a nitrogen decision support tool for corn in Pennsylvania that is based on inputs that can be measured through commercial lab services (White et al., 2023). This tool allows users to visualize the corn yield credits from decomposing cover crop residues (Equation 2) and SOM (Equation 3) under current conditions and possible future scenarios. The improved model presented here will allow the tool to be applied to fields across Pennsylvania with a wide range of soil textures and cover crops. The inclusion of an explicit precipitation adjustment would allow users to predict the range of unfertilized corn yields that could be expected at a site under various summertime precipitation scenarios ranging from below to above the long-term average.

The ability to predict unfertilized corn yield based on a few accessible inputs represents progress toward developing site-specific fertilizer recommendations. In the United States, N fertilizer recommendations for corn usually start with a generalized base recommendation (using either the yield goal or the maximum return to N approach), with some adjustments for site-specific management factors, which vary by state (Morris et al., 2018). An alternative site-specific approach is based on the concept of delta yield (dY), which is the difference between unfertilized corn yield and yield without N limitation (Janovicek et al., 2020; Lory & Scharf, 2003). Although dY can be determined for a site based on a comparison of unfertilized and fertilized areas of a field, several years of data are needed. The unfertilized corn yield predicted by the updated model could be combined with existing yield records to estimate dY. However, for the dY approach to become more widely used, further work is needed to establish relationships between dY and optimum N application rates (both agronomic and economic) across a wide geographic range. Also, future research could examine the relative effects of summertime precipitation on fertilized and unfertilized yields. This information would determine whether dY-based N fertilizer recommendations would be appropriate for a range of precipitation conditions or if they would need to be adjusted for expected rainfall.

5 | CONCLUSIONS

Overall, this work suggests that a system of biogeochemical equations that are mechanistically based, but calibrated with regional data, reflects a viable "middle ground" toward developing N decisions support tools. Our new N

mineralization model provides a robust prediction of unfertilized corn yields across a range of conditions. Adding a precipitation adjustment and calibrating across a broader range of soil textures were key steps to improving the model, while including measured soil C/N was less important. Our results point to the need for careful consideration of calibration datasets for developing these types of models.

A feature of the model is that it is based on site-specific inputs that can be measured through commercial lab services: soil texture and C and N content along with cover crop C and N content. The updated model presented here can be applied to a broad range of sites and is calibrated to average summertime precipitation in Pennsylvania. The cover crop credit (Equation 2) and the SOM credit (Equation 3) provide a comparison of the relative contributions of each to unfertilized corn yield. This information will allow farmers to consider how selection and management of cover crops and strategies to build SOM can contribute to soil fertility. The ability to predict unfertilized corn yield based on a few measured properties of soils and cover crops represents an advance toward developing site-specific fertilizer N recommendations for corn.

AUTHOR CONTRIBUTIONS

Kathleen E. Arrington: Data curation; formal analysis; methodology; software; visualization; writing—original draft; writing—review and editing. **Raziel A. Ordóñez**: Investigation; methodology; visualization; writing—original draft; writing—review and editing. **Zoelie Rivera-Ocasio**: Investigation; methodology; project administration; writing—review and editing. **Madeline Luthard**: Writing—review and editing. **Sarah Tierney**: Investigation; methodology; writing—review and editing. **John Spargo**: Investigation; methodology; writing—review and editing. **Denise Finney**: Investigation; methodology; project administration; writing—review and editing. **Jason Kaye**: Conceptualization; funding acquisition; writing—review and editing. **Charles White**: Conceptualization; data curation; formal analysis; funding acquisition; investigation; methodology; project administration; software; supervision; visualization; writing—original draft; writing—review and editing.

ACKNOWLEDGMENTS

The authors would like to thank Brosi Bradley and the staff at the Russell E. Larson Research Center for managing the field operations and data collection for the research station experiments. We would also like to thank the farmers who allowed us to establish experimental plots in their fields to build the dataset used in this study.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

ORCID

Kathleen E. Arrington https://orcid.org/0000-0001-6028-4701

John Spargo https://orcid.org/0000-0002-7400-4739

REFERENCES

Baldock, J. A., & Skjemstad, J. O. (2000). Role of the soil matrix and minerals in protecting natural organic materials against biological attack. *Organic Geochemistry*, *31*, 697–710. https://doi.org/10.1016/S0146-6380(00)00049-8

Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods & Research*, *33*(2), 261–304. https://doi.org/10.1177/0049124104268644

Cassman, K. G., Dobermann, A., & Walters, D. T. (2002). Agroecosystems, nitrogen-use efficiency, and nitrogen management. *AMBIO: A Journal of the Human Environment*, *31*(2), 132–140. https://doi.org/10.1579/0044-7447-31.2.132

Clark, A. J., Meisinger, J. J., Decker, A. M., & Mulford, F. R. (2007). Effects of a grass-selective herbicide in a vetch-rye cover crop system on corn grain yield and soil moisture. *Agronomy Journal*, *99*(1), 43–48. https://doi.org/10.2134/agronj2005.0362

Cotrufo, M. F., Wallenstein, M. D., Boot, C. M., Denef, K., & Paul, E. A. (2013). The microbial efficiency-matrix stabilization (mems) framework integrates plant litter decomposition with soil organic matter stabilization: Do labile plant inputs form stable soil organic matter? *Global Change Biology*, *19*(4), 988–995. https://doi.org/10.1111/gcb.12113

Daly, A. B., Jilling, A., Bowles, T. M., Buchkowski, R. W., Frey, S. D., Kallenbach, C. M., Keiluweit, M., Mooshammer, M., Schimel, J. P., & Grandy, A. S. (2021). A holistic framework integrating plant-microbe-mineral regulation of soil bioavailable nitrogen. *Biogeochemistry*, *154*, 211–229. https://doi.org/10.1007/s10533-021-00793-9

Dean, J. E., & Weil, R. R. (2009). Brassica cover crops for nitrogen retention in the Mid-Atlantic Coastal Plain. *Journal of Environmental Quality*, *38*, 520–528. https://doi.org/10.2134/jeq2008.0066

Drinkwater, L. E., Janke, R. R., & Rossoni-Longnecker, L. (2000). Effects of tillage intensity on nitrogen dynamics and productivity in legume-based grain systems. *Plant and Soil*, *227*, 99–113. https://doi.org/10.1023/A:1026569715168

Finney, D. M., Eckert, S. E., & Kaye, J. P. (2015). Drivers of nitrogen dynamics in ecologically-based agriculture revealed by long-term, high frequency field measurements. *Ecological Applications*, *25*(8), 2210–2227. https://doi.org/10.1890/14-1357.1

Franzluebbers, A. J., Shoemaker, R., Cline, J., Lipscomb, B., Stafford, C., Farmaha, B. S., Waring, R., Lowder, N., Thomason, W. E., & Poore, M. H. (2022). Adjusting the N fertilizer factor based on soil health as indicated by soil-test biological activity. *Agricultural & Environmental Letters*, *7*, e20091. https://doi.org/10.1002/ael2.20091

Ge, N., Wei, X., Wang, X., Liu, X., Shao, M., Jia, X., Li, X., & Zhang, Q. (2019). Soil texture determines the distribution of aggregate-associated carbon, nitrogen and phosphorous under two contrasting land use types in the Loess Plateau. *Catena*, *172*, 148–157. https://doi.org/10.1016/j.catena.2018.08.021

Gee, G. W., & Bauder, J. W. (1986). Particle size analysis. In A. Klute (Ed.), *Methods of soil analysis. Part 1. Physical and mineralogical methods* (2nd ed., pp. 383–411). American Society of Agronomy.

Griesheim, K. L., Mulvaney, R. L., Smith, T. J., Nunes, V. L. N., & Hertzberger, A. J. (2023). Isotopic comparison of ammonium and nitrate sources applied in-season to corn. *Soil Science Society of America Journal*, *87*, 555–571. https://doi.org/10.1002/saj2.20531

Hassink, J. (1994). Effects of soil texture and grassland management on soil organic C and N and rates of C and N mineralization. *Soil Biology and Biochemistry*, *26*(9), 1221–1231. https://doi.org/10.1016/0038-0717(94)90147-3

Janovicek, K., Banger, K., Sulik, J., Nasielski, J., & Deen, B. (2020). Delta yield based optimal nitrogen rate estimates for corn are often economically sound. *Agronomy Journal*, *113*, 1961–1973. https://doi.org/10.1002/agj2.20521

Lavallee, J. M., Soong, J. L., & Cotrufo, M. F. (2020). Conceptualizing soil organic matter into particulate and mineral-associated forms to address global change in the 21st century. *Global Change Biology*, *26*, 261–273. https://doi.org/10.1111/gcb.14859

Lory, J. A., & Scharf, P. C. (2003). Yield goal versus delta yield for predicting fertilizer nitrogen need in corn. *Agronomy Journal*, *95*, 994–999. https://doi.org/10.2134/agronj2003.9940

Manzoni, S., Taylor, P., Richter, A., Porporato, A., & Ågren, G. (2012). Environmental and stoichiometric controls on microbial carbon-use efficiency in soils. *New Phytologist*, *196*(1), 79–91. https://doi.org/10.1111/j.1469-8137.2012.04225.x

Morris, T. F., Murrell, T. S., Beegle, D. B., Camberato, J. J., Ferguson, R. B., Grove, J., Ketterings, Q., Kyveryga, P. M., Laboski, C. A. M., McGrath, J. M., Meisinger, J. J., Melkonian, J., Moebius-Clune, B. N., Nafzinger, E. D., Osmond, D., Sawyer, J. E., Scharf, P. C., Smith, W., Spargo, J. T., … Yang, H. (2018). Strengths and limitations of nitrogen rate recommendations for corn and opportunities for improvement. *Agronomy Journal*, *110*(1), 1–37. https://doi.org/10.2134/agronj2017.02.0112

Poffenbarger, H. J., Mirsky, S. B., Weil, R. R., Kramer, M., Spargo, J. T., & Cavigelli, M. A. (2015). Legume proportion, poultry litter, and tillage effects on cover crop decomposition. *Agronomy Journal*, *107*, 2083–2096. https://doi.org/10.2134/agronj15.0065

Quemada, M., & Cabrera, M. L. (1995). Carbon and nitrogen mineralized from leaves and stems of four cover crops. *Soil Science Society of America Journal*, *59*(2), 471–477. https://doi.org/10.2136/sssaj1995.03615995005900020029x

Rodriguez, M. P., Vargas, J., Correndo, A. A., Carcedo, A. J. P., Carciochi, W. D., Sainz Rozas, H. R., Barbieri, P. A., & Ciampitti, I. A. (2023). A meta-analysis of hairy vetch as a previous cover crop for maize. *Heliyon*, *9*(12), e22621. https://doi.org/10.1016/j.heliyon.2023.e22621

Sainju, U. M., & Singh, B. P. (2001). Tillage, cover crop, and kill-planting date effects on corn yield and soil nitrogen. *Agronomy Journal*, *93*, 878–886. https://doi.org/10.2134/agronj2001.934878x

Schimel, J. (2023). Modeling ecosystem-scale carbon dynamics in soil: The microbial dimension. *Soil Biology and Biochemistry*, *178*, 108948. https://doi.org/10.1016/j.soilbio.2023.108948

Sullivan, D. M., Andrews, N., Sullivan, C., & Brewer, L. J. (2019). *OSU organic fertilizer and cover crop calculator: Predicting plant-available nitrogen* (EM 9235). Oregon State University Extension Service. https://catalog.extension.oregonstate.edu/sites/catalog/files/project/pdf/em9235.pdf

Thapa, R., Cabrera, M., Reberg-Horton, C., Dann, C., Balkcom, K. S., Fleisher, D., Gaskin, J., Hitchcock, R., Poncet, A., Schomberg, H. H., Timlin, D., & Mirsky, S. B. (2022). Modeling surface residue decomposition and N release using the cover crop nitrogen calculator (CC-NCALC). *Nutrient Cycling in Agroecosystems*, *124*, 81–99. https://doi.org/10.1007/s10705-022-10223-3

Tollner, E. W., Hargrove, W. L., & Langdale, G. W. (1984). Influence of conventional and no-till practices on soil physical properties in the southern Piedmont. *Journal of Soil and Water Conservation*, *39*(1), 73–76.

White, C., Spargo, J. T., Arrington, K., Bradley, B., Finney, D., Kaye, J., Lefever, A., Luthard, M., Ordóñez, R., Rivera-Ocasio, Z., Sanders, Z., & Tierney, S. (2023). *Soil organic matter and cover crop-based nitrogen recommendations for corn*. Penn State Extension. https://extension.psu.edu/soil-organic-matter-and-cover-crop-based-nitrogen-recommendations-for-corn

White, C. M., Finney, D. M., Kemanian, A. R., & Kaye, J. P. (2016). A model–data fusion approach for predicting cover crop nitrogen supply to corn. *Agronomy Journal*, *108*, 2527–2540. https://doi.org/10.2134/agronj2016.05.0288

White, C. M., Finney, D. M., Kemanian, A. R., & Kaye, J. P. (2020). Modeling the contributions of nitrogen mineralization to yield of corn. *Agronomy Journal*, *113*, 490–503. https://doi.org/10.1002/agj2.20474

White, C. M., Sigdel, S., Karsten, H. D., Meinen, R. J., & Spargo, J. T. (2023). Recalibrating the Pennsylvania pre-sidedress soil nitrate test recommendations for modern corn production. *Agronomy Journal*, *115*, 2600–2613. https://doi.org/10.1002/agj2.21426

Woodruff, L. K., Kissel, D. E., Cabrera, M. L., Habteselassie, M. Y., Hitchcock, R., Gaskin, J., Vigil, M., Sonon, L., Saha, U., Romano, N., & Rema, J. (2018). A web-based model of N mineralization from cover crop residue decomposition. *Soil Science Society of America Journal*, *82*(4), 983–993. https://doi.org/10.2136/sssaj2017.05.0144

Xia, Y., Mitchell, K., Ek, M., Sheffield, J., Cosgrove, B., Wood, E., Luo, L., Alonge, C., Wei, H., Meng, J., Livneh, B., Lettenmaier, D., Koren, V., Duan, Q., Mo, K., Fan, Y., & Mocko, D. (2012). Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products. *Journal of Geophysical Research: Atmospheres*, *117*, D03109. https://doi.org/10.1029/2011JD016048

Yan, M., Pan, G., Lavallee, J. M., & Conant, R. T. (2020). Rethinking sources of nitrogen to cereal crops. *Global Change Biology*, *26*(1), 191–199. https://doi.org/10.1111/gcb.14908

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Arrington, K. E., Ordóñez, R. A., Rivera-Ocasio, Z., Luthard, M., Tierney, S., Spargo, J., Finney, D., Kaye, J., & White, C. (2024). Improving a nitrogen mineralization model for predicting unfertilized corn yield. *Soil Science Society of America Journal*, *88*, 905–920. https://doi.org/10.1002/saj2.20665