The impacts of model structure, parameter uncertainty and experimental design on Earth system model simulations of litter bag decomposition experiments

5 Daniel M. Ricciuto[1], Xiaojuan Yang[1], Dali Wang[1], and Peter E. Thornton[1]

[1]Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN, USA city,

Correspondence to: Daniel M. Ricciuto (ricciutodm@email.com)

Note: This manuscript has been authored by UT-Battelle, LLC under Contract No. DE-AC05-00OR22725 with the U.S.
20 Department of Energy. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (http://energy.gov/downloads/doe-public-access-plan).

Abstract

Accurate Earth system model simulations of the terrestrial carbon cycle and its feedbacks to climate critically
30   depend on algorithms representing the decomposition of litter and soil organic matter. Litter bag studies, in which
specific types of plant litter are subject to varying environmental conditions in the field and decomposition is
measured, serve as valuable benchmarks for model performance. Here we test the Energy Exascale Earth System
land model (ELM), which has two different structural representations of decomposition, using observations from
the Long-term Intersite Decomposition Experiment (LIDET) over six different biomes and six different leaf litter
35   types. We find that seasonal patterns in soil conditions and nutrient availability have large effects on decomposition
rates, and that it is critical to include this in the simulation design. Despite widely differing base decomposition
rates between the two different model structures, the models produce similar temporal patterns of decomposition
when nitrogen is limiting. Both models overpredict the fraction of original nitrogen present as a function of carbon
remaining when using default parameterizations. A parameter sensitivity analysis indicates strong dependence of
40   model outputs on nitrogen limitation, carbon use efficiency and decomposition rates. A large spread in model
predictions when considering an ensemble of possible parameter combinations strongly suggests parameter
uncertainty may be more influential than model structural uncertainty, and that new measurement and modelling
approaches may be necessary to constrain these uncertainties.

45

1 Introduction

50   Litter and soil organic matter (SOM) decomposition play critically important roles in carbon and nutrient cycling within terrestrial ecosystems. Environmental factors that drive this decomposition are likely to change in the future. This will impact both heterotrophic respiration, a large source of CO_2 to the atmosphere, and vegetation productivity through climate-carbon-nutrient feedbacks (Thornton et al., 2009). Developing a predictive understanding of these processes is critical to determining terrestrial carbon-climate feedbacks that play a key role

55   in the evolution of Earth's climate system. Although most coupled Earth system models do include an explicit representation of litter inputs, soil carbon stocks and decomposition, large uncertainties still exist in model predictions of carbon cycle feedbacks (Friedlingstein et al, 2014). These uncertainties can be attributed in part to large differences in the magnitudes and rates of soil carbon turnover, including litter decomposition processes (Todd-Brown et al., 2013; Tian et al, 2015). These differences stem from variations among models in specified

60   turnover rates for different pools, and from disagreement among models on how climate change-driven shifts in vegetation productivity affect the distribution of litter and SOM inputs (Koven et al., 2015). In general, current Earth system models may be underestimating soil carbon turnover times and are not able to accurately reproduce the inferred spatial patterns in turnover times from remote sensing and data synthesis studies, underscoring the need for more model development and analysis in this area (Wu et al., 2018).

65   Litter and soil carbon stocks and decomposition rates depend on many factors, such as soil environmental conditions (temperature and moisture), soil texture and mineralogy, litter input rates from vegetation, litter chemistry, microbial populations and nutrient availability. Earth System models generally represent litter and soil organic carbon as multiple pools, with transfer between pools represented by first-order kinetic functions (Luo et al, 2016). Box–model representations have been shown to reproduce patterns of observed long-term decomposition

70   (e.g. Adair et al., 2008). In these types of models, multiple litter pools are generally used to represent different tissue chemistries (e.g. labile, cellulose and lignin), and can have different decomposition rates that may be modified by environmental factors. These litter pools then decompose into soil organic matter (SOM) carbon pools. Some carbon is released to the atmosphere as CO_2 during this decomposition step, while nitrogen or phosphorus may be immobilized or mineralized depending on the size of the respiration flux, the assumed microbial

75   carbon use efficiency, and differences in carbon to nutrient ratios between the litter and SOM pools (Thornton and Rosenbloom 2005). While the general structures of these models are similar, large uncertainties remains in the conceptualization and parameterization of pools and fluxes due to lack of constraint by field observations.

Extensive observations of litter decomposition exist over a wide range of litter types, biomes and environmental conditions. For example, the Long-term Intersite Decomposition Experiment Team (LIDET) examined litter

80   decomposition rates for 10 years over 27 sites and 9 different litter types, providing a comprehensive dataset for testing global-scale models (Harmon et al., 2009). Analysis of observed variations led to quantification of the effects of climate, litter type and chemistry on decomposition in tropical systems (Cusack et al., 2009). Nitrogen release was found to depend strongly on the initial nitrogen concentration across different climates and biomes, suggesting global applicability of LIDET data to modelling frameworks (Parton et al., 2007). A number of

85   modelling groups used LIDET data as an important benchmark for evaluating models of carbon and nitrogen cycling. Adair et al. (2008) found that a simple 3-pool model performed well in simulating the observed decomposition rates in the LIDET data. Bonan et al. (2013) found that the decomposition model in CLM4.0, the converging trophic cascade (CTC) which first appeared in Biome-BGC (Thornton et al., 2002), apparently overestimated both the carbon turnover rate and immobilization of nitrogen, and that the DAYCENT modelling

90   framework provided a more accurate prediction across the sites and litter types in the LIDET experiment. However, DAYCENT did not predict improved SOM carbon stocks compared to CLM4.0 without additional changes to the model parameterization (Wieder et al., 2014). Furthermore, by isolating the decomposition model and driving it with constant environmental conditions, the Bonan et al. (2013) study did not account for changes in nutrient availability over time, which can have strong effects on litter decomposition rates (Carrerio et al., 2000).

95   One potential reason for large differences among Earth System models in soil carbon stocks and decomposition rates is design differences in model validation experiments, for example the choice of whether or not nutrient interactions are considered in evaluating the decomposition model as in Bonan et al. (2013). These choices potentially lead to different interpretations of model bias and different model parameterizations. Here, our primary objective is to evaluate an Earth system model using a framework in which the model experiment is most consistent

100   with the LIDET experiment. Our model of interest is the Energy Exascale Earth System model (E3SM) land component (ELM), which includes both the CTC decomposition model and a CENTURY scheme (CNT). We develop a modelling protocol that we believe best replicates the experimental conditions, including multiple relevant spatial scales over which processes are assumed to occur, and appropriate feedbacks within the simulation. To this end, we implement decomposition model "functional units" to represent the decomposition processes in the

105   experimental system (the litter bag, as described below) that are driven by evolving soil physical and biogeochemical conditions from an ecosystem-scale ELM simulation. We use this modelling framework to address the following science questions:

1) How do different decomposition base rates in two different decomposition models affect litter turnover times as compared to LIDET observations?

110   2) how are these turnover times affected by seasonal patterns in soil conditions and nitrogen limitation?

3) Which parameters contribute the most to uncertainty in the predictions of carbon and nitrogen content over time?

2 Methods

2.1 LIDET experiment and sites

115   LIDET (Long-term Inter-site Decomposition Experiment) was a long-term litter bag study designed to examine the effect of substrate quality and climate on carbon loss and nitrogen immobilization and mineralization during litter decomposition over a 10-year period (LIDET 1995; Gholz et al., 2000; Parton et al., 2007; Harmon et al., 2009; Currie et al., 2010). It involved transplanting 27 litter types with various litter quality at 28 sites across North America and Central America. Six core litter types were chosen to decompose at all sites. Following Parton et al.

120   (2007), we choose the core LIDET data, which includes decomposition of six core litter types at 18 upland sites. The 18 upland sites encompass a wide range of climate condition and biome type, including 2 tundra, 2 boreal forest, 5 conifer forest, 3 deciduous forest, 4 tropical forest and 2 humid grassland sites. Dry grasslands were excluded from this analysis. The sites range in climate from mean annual temperature between -8°C and 26°C, and from mean annual precipitation between 250 to over 4000 mm/yr (Table 1). The six litter types have varying litter

125   quality, with N concentration ranging from 0.38% to 1.97% and lignin content varying from 10.9% to 26.7%. (Table 2). 10g of leaf litter was confined in 10 by 10cm mesh bags and placed at the top of the native litter layer at each field site in September of 1990. Collections of bags occurred once per year, in Autumn for most sites, with more frequent collection at the tropical sites – up to 5 times per year - due to the faster decomposition at these sites.

130

Site	Location	PFT	MAT (°C)	MAP (mm/yr)
ARC	Alaska	Tundra	-8.45	285
NWT	Colorado	Tundra	0.6	1172
BNZ	Alaska	Boreal forest	-4.40	257
LVW	Colorado	Boreal forest	1.8	1120
AND	Oregon	Conifer forest	8.3	2346
BSF	California	Conifer forest	13.6	1515
JUN	Alaska	Conifer forest	3.9	1478
OLY	Washington	Conifer forest	10.2	3896
UFL	Florida	Conifer forest	21.4	1197
CWT	North Carolina	Deciduous forest	12.4	1938
HBR	New Hampshire	Deciduous forest	6.0	1315
HFR	Massachusetts	Deciduous forest	7.4	1136
BCI	Panama	Tropical forest	26.2	2853
LBS	Costa Rica	Tropical forest	25.4	4070
LUQ	Puerto Rico	Tropical forest	23.9	2256
MTV	Costa Rica	Tropical forest	25.9	4381
CDR	Minnesota	Humid grassland	6.0	817
KNZ	Kansas	Humid grassland	13.2	874

Table 1. LIDET sites used in this analysis listed with plant functional type, mean annual temperature (MAT) and mean annual precipitation (MAP).

135

Litter type	Species	Common name	C:N	% labile	% cellulose	% lignin
TRAE	*Triticum aestivum*	Wheat	133.3	10.6	73.2	16.2
PIRE	*Pinus resinosa*	Red Pine	92.7	36.2	44.6	19.2
THPL	*Thuja plicata*	Red cedar	83.1	37.4	35.9	26.7
ACSA	*Acer saccharum*	Sugar maple	61.8	56.8	27.3	15.9
QUPR	*Quercus prinus*	Chestnut oak	50.5	37.1	39.4	23.5
DRGL	*Drypetes glauca*	Tropical broadleaf	24.2	49.3	39.8	10.9

Table 2: Litter chemistry of LIDET leaf types

2.2 Model description and simulation procedure

Here we use the Energy Exascale Earth System model (E3SM) land model version 0 (ELMv0), which is based on
140 the Community Land Model version 4.5 (Oleson et al., 2013). ELMv0, like CLM4.5, includes vertically resolved
soil biogeochemistry over 10 soil layers, the explicit representation of ammonium and nitrate pools, and associated
nitrification and denitrification flux of N (Koven *et al.*, 2013). CTC and CNT, the two decomposition submodels
analyzed in this study, each have three litter pools that represent labile, cellulose, and lignin litter fractions. New
litter is allocated to these pools based on type of litter (live or dead wood, fine root, and leaf) and the plant functional
145 type (PFT). Each litter pool turns over to a separate soil organic matter (SOM) pool with a base rate that is modified
by temperature, moisture, oxygen availability, depth and competition for available nitrogen among plants and other
microbial immobilization pathways. There are three SOM pools in CNT and four in CTC, which decompose either
to other SOM pools or the atmosphere with base rates that are modified by temperature, moisture, oxygen
availability and depth. For each pool transition, there is a loss of carbon in the form of CO_2 flux to the atmosphere,
150 defined as a respiration fraction. CTC parameters are based on measured values from a series of mesocosm
experiments (Thornton and Rosenbloom, 2005), while CNT parameters derive from the CENTURY model (Parton
et al., 1988). Parameters for both decomposition models are given in Table 3.

155

Parameter	Units	Description	CTC	CNT
k_lit1	days^{-1}	Labile pool decomposition rate	0.83	19.7
k_lit2	days^{-1}	Cellulose pool decomposition rate	13.8	74.5
k_lit3	days^{-1}	Lignin pool decomposition rate	71.4	74.5
k_som1	yr^{-1}	SOM pool 1 decomposition rate	0.0378	0.137
k_som2	yr^{-1}	SOM pool 2 decomposition rate	0.196	5.0
k_som3	yr^{-1}	SOM pool 3 decomposition rate	1.96	222
k_som4	yr^{-1}	SOM pool 4 decomposition rate	27.4	N/A
cn_som1	gC gN^{-1}	SOM pool 1 C:N ratio	12	8
cn_som2	gC gN^{-1}	SOM pool 2 C:N ratio	12	10
cn_som3	gC gN^{-1}	SOM pool 3 C:N ratio	10	10
cn_som4	gC gN^{-1}	SOM pool 4 C:N ratio	10	N/A
rf_l1s1		Respiration frac labile to SOM1	0.39	0.55
rf_l2sx		Respiration frac cellulose to SOM1(2)	0.55	0.5
rf_l3sx		Respiration frac lignin to SOM 2(3)	0.29	0.5
rf_s1s2		Respiration frac from SOM1 to SOM2	0.28	0.28
rf_s2s3		Respiration frac from SOM2 to SOM3	0.46	0.55
rf_s3s4		Respiration frac from SOM3 to SOM4	0.55	N/A
rf_s2s1		Respiration frac from SOM2 to SOM1	N/A	0.55
rf_s3s1		Respiration frac from SOM3 to SOM1	N/A	0.55

Table 3: Default parameters in the CTC and CNT decomposition models used in the analysis.

Competition for nitrogen between plants and microbes is resolved at each half-hourly model timestep using a
relative demand approach (Thornton et al., 2007). Plant nitrogen demand is driven by the portion of carbon uptake
160 that is allocated to structural pools and associated stoichiometry given fixed carbon to nitrogen ratios for those
pools. Microbial demand is driven by immobilization requirements when litter decomposes into SOM given the

differences in carbon to nitrogen ratios between these pools and the respiration fraction. The potential litter decomposition, which drives the microbial nitrogen demand, is a function of litter chemistry, pool size, soil moisture and temperature. If the total demand for nitrogen is greater than the available soil mineral nitrogen, plant
165 and microbial uptake are then scaled down by multiplying their fractions of total demand by the total available mineral nitrogen. In this way, both plant gross primary productivity (GPP) and litter decomposition are downregulated by nitrogen limitation. Effectively, nitrogen limitation serves to modify the base litter turnover rates in the decomposition model along with temperature and moisture conditions. Here we use the capability of ELMv0 to run with both the CTC and CENTURY-based (CNT) decomposition submodels with vertically resolved
170 soil biogeochemistry described in Koven et al. (2013), referred to here as ELM-CTC and ELM-CNT respectively. The vertically resolved ELM-CTC configuration is also the default soil biogeochemistry parameterization for ELMv0 and later versions.

For this study, we perform ELM-CTC and ELM-CNT simulations to calculate the environmental conditions in terms of the temperature and moisture scalars, as well as the background nutrient limitation conditions at each of
175 the LIDET experiment sites. The simulated LIDET experiment at each site then takes place in a model "functional unit", which is a standalone representation of the decomposition model (section 2.3). In ELM, each of the 18 LIDET sites is represented as a single gridcell in a multi-site ensemble simulation, similar to the model setup in ELMv0 in Ricciuto et al. (2018). Both ELM-CTC and ELM-CNT were run to steady state using pre-industrial carbon dioxide concentrations (constant year 1850), nitrogen deposition (constant year 1850) and climate forcing
180 (cycling the years 1901-1920) from the Global Soil Wetness Project 3 (GSWP3; Dirmeyer et al., 2006). The spinup procedure follows the accelerated decomposition and regular spinup techniques described by Thornton and Rosenbloom (2005). Following achievement of steady-state carbon and nutrient pools, we begin the transient simulation in model year 1850 and continue through the year 2000 using historically varying CO_2 concentrations and nitrogen deposition. In the transient simulation, we continue to cycle the 1901-1920 GSWP3 climate data until
185 the year 1921, when we switch to the historical forcing from GSWP3 for the years 1921-2000. For each site, the GSWP3 climate forcing is adjusted such that its long-term mean annual temperature and precipitation matches the reported site average given by Adair et al. (2008). ELM plant functional types are set to site-reported values following Bonan et al. (2013). Land-use change is not considered in these simulations. The LIDET experiment covers the period from October 1990 through the year 2000. During this time, soil moisture and temperature scalars
190 (W_{scalar} and T_{scalar}, respectively), and the fraction of potential immobilization (FPI) as simulated by ELM-CTC and ELM-CNT are saved for each model timestep. These scalars represent a multiplicative factor applied to the base

rates for litter decomposition (defined at 25C and ideal moisture conditions), and serve as boundary condition inputs for the decomposition functional unit.

2.3 Functional units and simulation procedure

195   In ELM, it is not possible to represent the small scale of the litter bags (10x10cm) in a simulation that includes a big-leaf representation of vegetation at canopy scales. One would have to assume the litter bag is covering a much larger spatial area, which may cause unrealistic feedbacks to vegetation (e.g. changes in plant nitrogen uptake due to the influence of the litter bag). However, the decomposition submodel is not subject to canopy-scale assumptions. Therefore, a functional representation of the ELM-CTC and ELM-CNT decomposition models is

200   developed. In this framework, the decomposition can proceed in the simulated litter bags using the environmental drivers provided by the canopy-scale ELM simulations without causing unrealistic feedbacks to the vegetation though changes in plant-available nitrogen. Unlike Bonan et al. (2013), which also used a functional unit representation of the decomposition models, we use the canopy-scale ELM-CTC and ELM-CNT simulations to provide the changing environmental drivers (soil temperature, soil moisture and nitrogen availability) as boundary

205   conditions at each timestep. This simulation configuration is intended to represent the implicit assumption of the LIDET experiment, that the introduction of the litter bag does not have an appreciable influence on the site-level ecosystem states and fluxes.

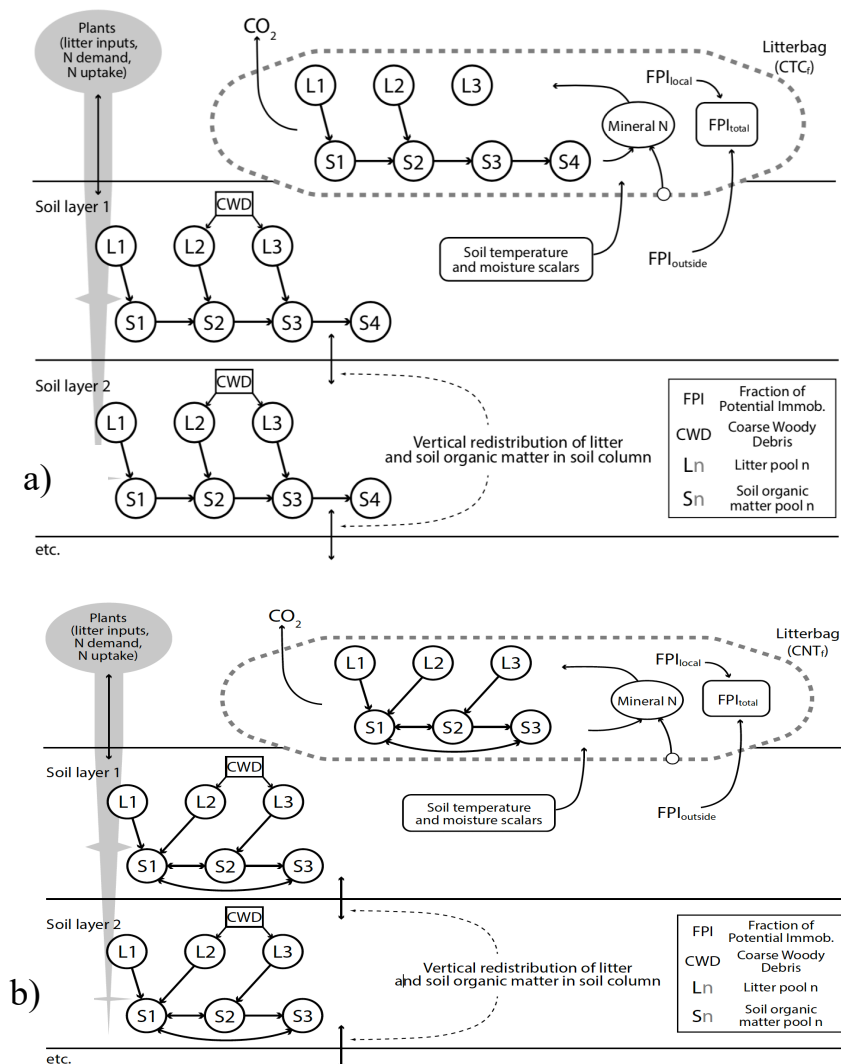
A true functional testing platform allows for the testing of specific submodels using existing ELM output and the original model subroutine code, allowing new insights about the behaviour of these submodules (Wang et al.,

210   2015). However, due to large memory and communication requirements, this platform is not computationally efficient to run over large combinations of sites and parameter values. Therefore, we developed a python-based version of the decomposition model, which is tested against the decomposition functional unit that uses the original ELM code for fidelity (Yao et al., 2019). The CTC and CNT versions of the functional unit submodel are referred to as CTC_f and CNT_f respectively. The LIDET model experiments are then performed using these functional units

215   with the site-level ELM simulations providing the soil environmental and nutrient conditions (W_{scalar}, T_{scalar} and FPI from the top soil layer) as inputs (Figure 1). For each site and each leaf litter type, CTC_f and CNT_f are initialized with the leaf litter added to litter bags at the beginning of the experiment, divided into the 3 litter pools (labile, cellulose and lignin) according to the given litter chemistry (Table 2). We assume the temperature and moisture conditions are the same in the litter bag as in the first soil layer in the canopy-scale ELM simulation. As the

220   simulations progress over a 10-year period, the litter decomposes into SOM, which also decomposes, releasing CO_2 to the atmosphere and either immobilizing or mineralizing nitrogen depending on specific pool transitions.

That full chain of decomposition is assumed to occur within the litter bag. We allow for the possibility that mineral nitrogen can enter the functional unit system (the bag) over time, and this flux is assumed to originate from the adjacent native litter and soil system. There is likely to be a local influence of processes in the litter bag on the

225 available nitrogen and nitrogen limitation to decomposition. Therefore, we also calculate an internal value of FPI, referred to as FPI_{local}. Plant demand is assumed to be zero in the litter bag, and FPI_{local} is calculated in the following way at each model timestep:

230 Figure 1: Representation of the functional unit converging trophic cascade decomposition (CTC_f) model structure (a), and the CENTURY-based CNT_f model structure (b) within their respective ELM frameworks.

$$FPI_{local} = \min (f_{immob} / f_{smin}, 1.0) \tag{1}$$

Where f_{immob} is the immobilization demand in the bag and f_{smin} is the mineral nitrogen originating from
235   decomposition within the bag. Mineralized nitrogen primarily comes from the decomposition of soil organic
matter. Therefore early in the experiment, when mainly litter is present instead of SOM, FPI_{local} is near zero. In
this case, immobilization demand must be met by soil mineral N from the surrounding environment outside the
litter bag. The FPI used for litter decomposition in the functional unit, FPI_{total}, is a combination of FPI from the
site simulation ($FPI_{outside}$) and $FPI_{local,}$ weighted by the density of litter in the bag vs. the density of surrounding litter
240   as calculated from the site ELM simulations:

$$FPI_{total} = FPI_{local} * lit_{local} / (lit_{local} + lit_{site}) + FPI_{outside} * lit_{site} / (lit_{local} + lit_{site}) \tag{2}$$

Where lit_{local} is the density of remaining litter (g m^{-3}) in the litterbag as calculated by the model functional unit and
assumed to be in the first soil layer, and lit_{site} is the density of litter (g m^{-3}) in the ELM site simulation in the first
soil layer.

245   A total of 108 simulations (18 sites * 6 litter types) are performed for each functional unit. Daily output is saved
from the CTC_f and CNT_f over the course of the 10-year simulated experiment, including all litter and SOM carbon
and nitrogen pools. These outputs are aggregated to produce the fraction of initial carbon mass remaining, and the
fraction of original nitrogen present for comparison against observations.

2.4 Sensitivity analysis

250   We also use CTC_f to investigate the impacts of parameter uncertainty in the decomposition model using a global
sensitivity analysis (GSA). GSA, also known as variance-based decomposition, can attribute variations in
predictions to parameters and their interactions given ranges of possible values for the parameters. Following the
techniques used in Sargsyan et al. (2014) and Ricciuto et al. (2018), we use Polynomial Chaos surrogate models
via a Bayesian compressive sensing (BCS) approach to conduct the GSA. A surrogate model is a functional
255   representation of a model quantity of interest (QoI) that is constructed from an ensemble of simulations from the
original model (in this case CTC_f), and allows for further exploration of QoI responses over the multi-dimensional
parameter space. In this case the quantities of interest are the fraction of original carbon remaining over each of
the ten years of the experimental period, as well as fraction of original nitrogen remaining over the same time
period. This surrogate modelling approach reduces computational cost, and BCS can learn the best possible
260   surrogate model from a limited number of simulations.

We use a total of 500 CTC_f simulations to construct the surrogate model and run the GSA. 11 model parameters
are varied over uniform distributions, including properties of the input litter, and parameters in the decomposition

model (table 4). We use a simple Monte Carlo approach to randomly sample parameters within their acceptable

ranges. Because including a larger number of parameters exponentially increases the number of required

265 simulations for the sensitivity analysis, we do not consider decomposition rate parameters and respiration fraction

parameters independently. Instead, we consider a single multiplier parameter for all three litter pools which is

applied to the default base decomposition rates. Similarly, we consider a single parameter which serves as a

multiplier to all the default respiration fraction values. Parameter minimum and maximum values for litter

chemistry are determined by the range of the six litter types used in the experiment. Decomposition rate, respiration

270 fraction, and soil organic matter parameters were chosen to capture differences in values between the CTC and

CNT models. Main effect and joint sensitivity indices are calculated for each parameter. These sensitivity indices

do not provide information about whether a parameter has a positive or negative effect on a quantity of interest.

Therefore, using the same 500 model samples, we also calculate the linear correlation coefficient between each

parameter and each quantity of interest to further understand the impacts of individual parameters. The GSA is

275 performed at three sites: BCI, BNZ and HFR, representing tropical, boreal and deciduous forests respectively.

Parameter	Units	Description	Min	Max	Default
litter_CN	gC gN^{-1}	Litter C:N ratio	24	133	Input
litter_flab		Labile fraction	0.1	0.6	Input
litter_flig		Lignin fraction	0.1	0.3	input
k_lit_mult		Multiplier for all litter pools decomp base rate	0.2	2	1.0
k_som123_mult		Multiplier for SOM pools 1-3 decomp base rate	0.2	2	1.0
k_som4_mult		Multiplier for SOM pool 4 decomp base rate	0.1	2	1.0
rf_mult		Multiplier for respiration fraction	0.8	1.2	1.0
cn_som12	gC gN^{-1}	C:N ratio for SOM pools 1 and 2	8	25	12
cn_som3	gC gN^{-1}	C:N ratio for SOM pool 3	8	25	10
cn_som4	gC gN^{-1}	C:N ratio for SOM pool 4	8	25	10
FPI_outside		Background fraction of pot. Immob.	0.03	1.0	Input

Table 4: Parameter ranges used in the sensitivity analyses for the CTC model. Minimum and maximum values for litter_CN,

litter_flab and litter_flig are determined by the range of values over the different litter types; default vales for these parameters

depend on the litter types. For the sensitivity study, *FPI_outside* is held constant over the simulation. However in the forward

280 simulations, *FPI_outside* is calculated interally within ELM and varies over time.

3 Results

3.1 Comparison between CTC$_f$ and CNT$_f$ default models

When using default model parameters (table 3), both the CTC$_f$ and CNT$_f$ models generally reproduce the observed
285   temporal patterns of carbon decomposition averaged over the six litter types (Figure 2). The CTC$_f$ model displays
closer agreement with the observations for the tropical, conifer and deciduous PFT sites while the CNT$_f$ model
agrees more closely for the tundra and boreal forest PFTs. This is reflected in the lower root mean squared error
(RMSE) and bias values averaged over all litter types for CTC$_f$ or CNT$_f$ for those corresponding PFTs (table 5).
For the humid grasslands, although CNT$_f$ performs somewhat better than CTC$_f$, both models decompose the litter
290   too quickly compared to observations and are therefore biased low in their predicted carbon remaining by at least
10%. For all PFTs, CTC$_f$ shows faster decomposition rates compared to CNT$_f$, reflecting the higher base
decomposition rates in CTC$_f$ for labile, cellulose and soil organic matter pools. CTC$_f$ has a slightly lower base
decomposition rate for lignin. Base decomposition rates occur under ideal moisture conditions at 25ºC, and without
nitrogen limitation.

295

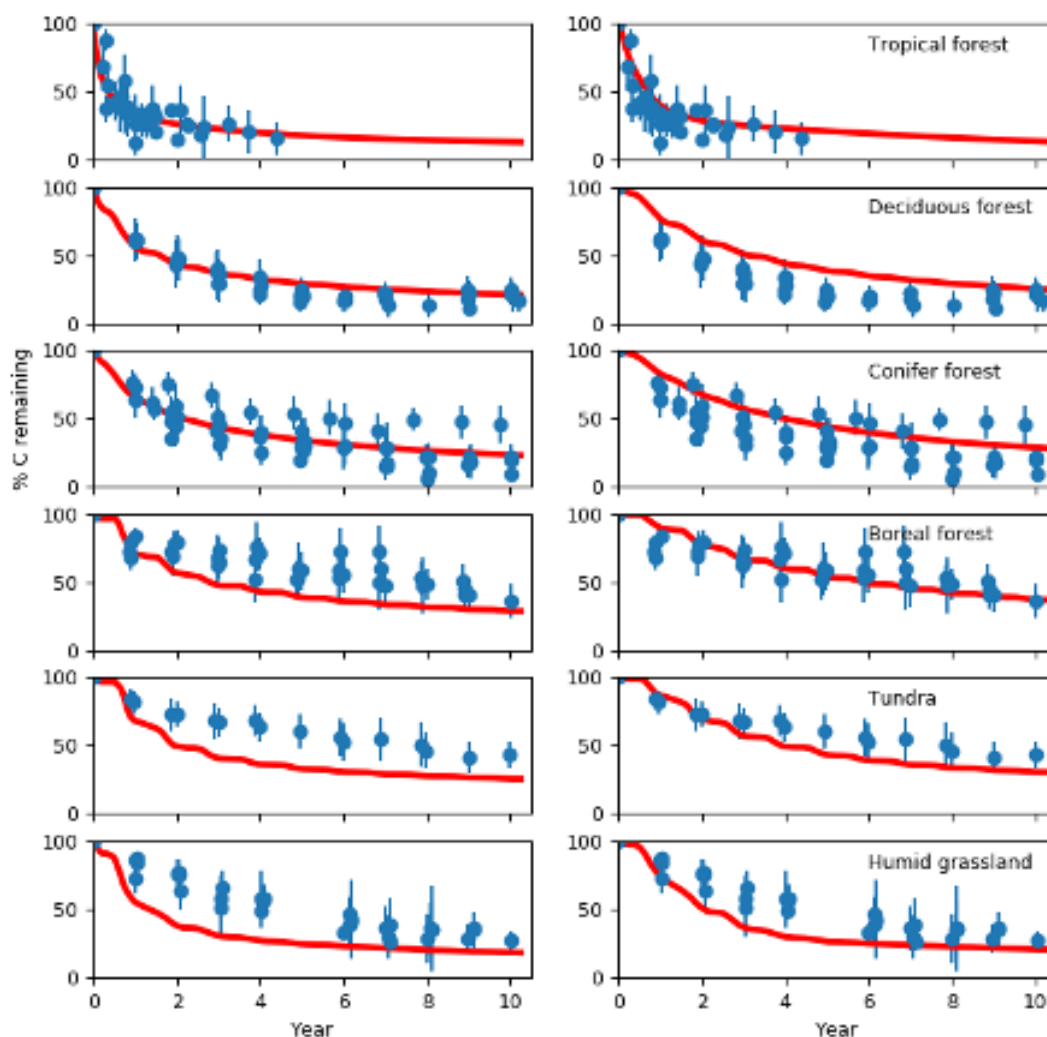



Figure 2: Simulated (red) and observed (blue) percentage of carbon remaining for six plant functional types. ELM functional
unit results from the converging trophic cascade decomposition model (CTCf) are shown in the left column, while results from
the CENTURY scheme (CNTf) shown in the right column. Results are aggregated over the six litter types used in the LIDET
study.

300

PFT	RMSE (CTC)	RMSE (CNT)	Bias (CTC)	Bias (CNT)
Tropical	14.3	15.5	-1.47	7.18
Deciduous	10.7	20.8	-3.68	17.89
Confier	12.9	16.5	-1.42	7.51
Boreal	22.8	15.3	-18.47	-3.51
Tundra	23.9	14.2	-22.45	-10.38
Humid grass	25.32	18.78	-20.45	-12.44

Table 5: Error statistics for carbon remaining (%) aggregated over all litter types by biome.

305 The actual decomposition rates are heavily influenced by the nutrient and environmental scalars (Figure 3). Environmental scalars representing temperature and moisture effects on decomposition are similar between the two models, although slightly higher in CTC$_f$ for all PFTs. Average annual scalar values range between 0.1 and 0.75 in CTC$_f$, reflecting over a sevenfold difference in litter and SOM turnover rates depending on soil conditions (Table 6). A similar range of scalar values is observed for CNT$_f$. The slightly lower values in CNT$_f$ reflect the use of an

310 arctangent temperature function in CNT$_f$ rather than the Q_{10} function that is used in CTC$_f$. For all PFTs, the nitrogen scalar is lower in CTC$_f$ than in CNT$_f$, indicating higher nutrient limitation in CTC$_f$ for litter decomposition. In ELM, plants and microbes compete for soil mineral nitrogen using a relative demand approach at every model timestep. Potential litter decomposition rates are influenced by the environmental conditions but do not reflect nitrogen limitation. Net primary productivity is similar between the two models, with differences within 5% at 14

315 sites and within 10% at the remaining 4 sites (Table 6). Therefore, litter inputs and plant demand are similar and the higher nitrogen limitation factor in CTC$_f$ results primarily from higher potential immobilization caused by the higher base decomposition rates. This higher nitrogen limitation thus reduces the effective difference in actual decomposition rates between the two models.

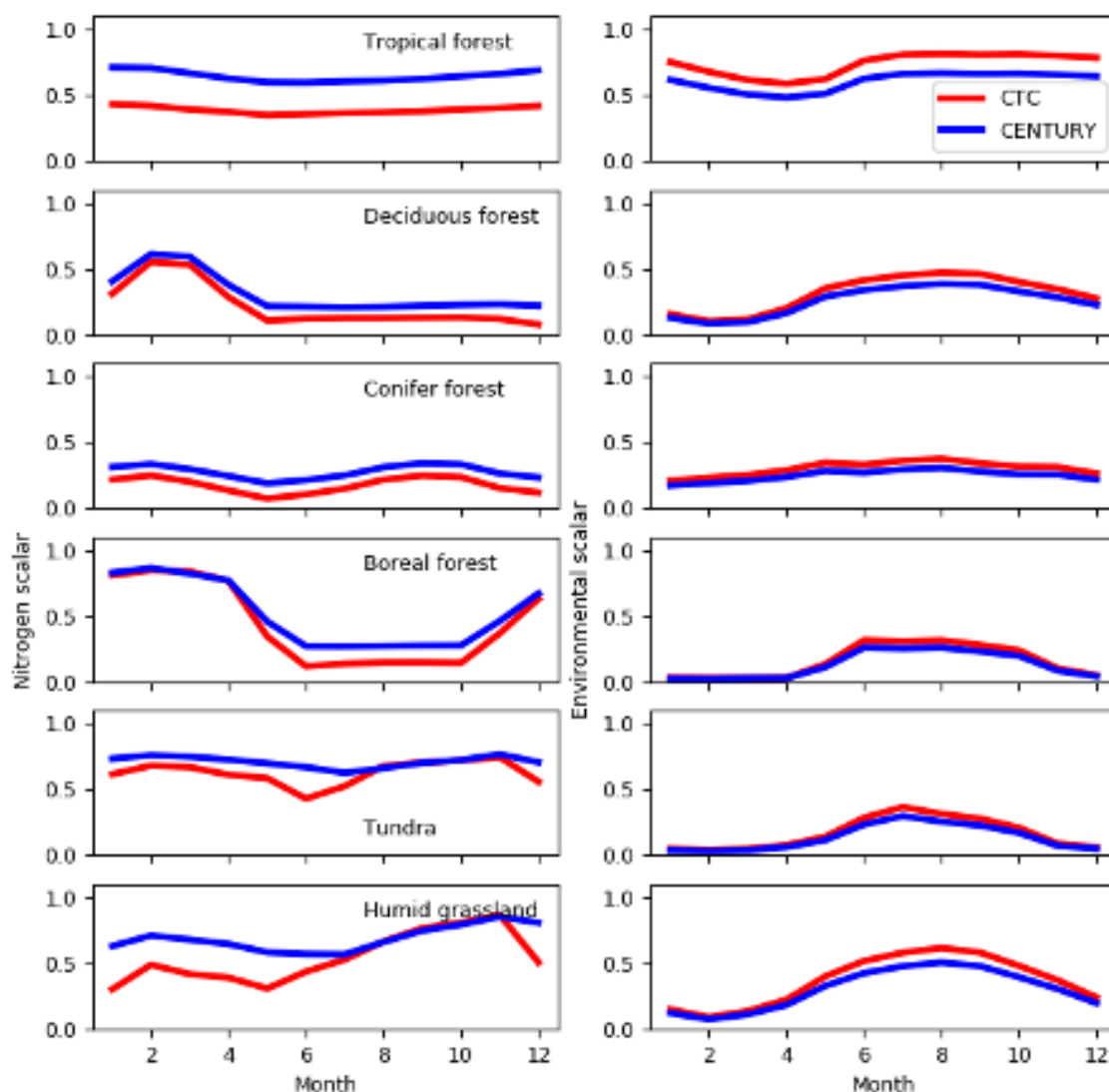



Figure 3: Simulated average seasonal cycles of the fraction of potential immobilization of nitrogen calculated by ELM
(*FPI_outside*) and environmental scalars calculated by ELM that impact litter decomposition. Lower values of FPI$_{outside}$
indicate increased nitrogen limitation. The environmental scalar includes both temperature and soil moisture effects. Higher
values indicate faster decomposition. These ELM-calculated scalars are used as inputs to CTC$_f$ and CNT$_f$.

325

Site	PFT	NPP (CTC)	NPP (CNT)	FPI (CTC)	FPI (CNT)	Scalar (CTC)	Scalar (CNT)
ARC	Tundra	259	267	0.936	0.979	0.101	0.083
NWT	Tundra	301	309	0.313	0.439	0.206	0.168
BNZ	Boreal forest	174	174	0.571	0.664	0.113	0.093
LVW	Boreal forest	359	369	0.278	0.350	0.165	0.139
AND	Conifer forest	303	341	0.059	0.128	0.296	0.241
BSF	Conifer forest	319	348	0.339	0.420	0.275	0.225
JUN	Conifer forest	232	245	0.311	0.443	0.219	0.180
OLY	Conifer forest	1204	1232	0.056	0.115	0.335	0.274
UFL	Conifer forest	645	669	0.092	0.267	0.424	0.345
CWT	Deciduous forest	703	761	0.103	0.219	0.401	0.335
HBR	Deciduous forest	542	520	0.298	0.422	0.237	0.195
HFR	Deciduous forest	546	519	0.391	0.456	0.256	0.201
BCI	Tropical forest	1267	1294	0.402	0.706	0.724	0.594
LBS	Tropical forest	1313	1334	0.426	0.676	0.749	0.614
LUQ	Tropical forest	1367	1404	0.300	0.489	0.710	0.583
MTV	Tropical forest	1264	1292	0.401	0.682	0.762	0.624
CDR	Humid grassland	399	382	0.608	0.784	0.290	0.238
KNZ	Humid grassland	450	449	0.447	0.569	0.412	0.338

Table 6: Net primary productivity (NPP). Fraction of potential immobilization (FPI) and environmental scalars (representing
temperature and moisture effects) for each site and model calculated by ELM. Units of NPP are gC/m^2/yr. All values are
averaged over the 10 year period of the experiment.

The percentage of the original litter nitrogen is plotted as a function of the percentage of original carbon remaining (Figure 4). While the percentage of carbon remaining is always declining over time, the percentage of nitrogen

335 remaining may increase due to immobilization of soil mineral nitrogen from external sources outside of the litter bag. This is most evident in the litter types with high carbon to nitrogen ratios, for example PIRE and THPL. These types incur higher immobilization demand when decomposing from litter to SOM. In the functional unit representation, the litter bag is assumed to have no influence on its surrounding environment (i.e., it does not impact plant or immobilization demand outside of the bag). Therefore, immobilization demand is assumed to be met by

340 external sources of nitrogen but is limited according to the calculated value of FPI from ELM and locally within the bag (equation 2). For both CTC_f and CNT_f, the models strongly overpredict the amount of nitrogen present in the litter bags for the PIRE, THPL, ACSA, and QUPR litter types. For the DRGL type, the models better predict the behaviour because the carbon to nitrogen ratio is low and there is little immobilization. Despite the different litter and SOM decomposition base rates, the models behave similarly because they have similar carbon to nitrogen

345 ratios for SOM and similar respiration fractions (Table 3), meaning that the total immobilization demand integrated over time is similar.

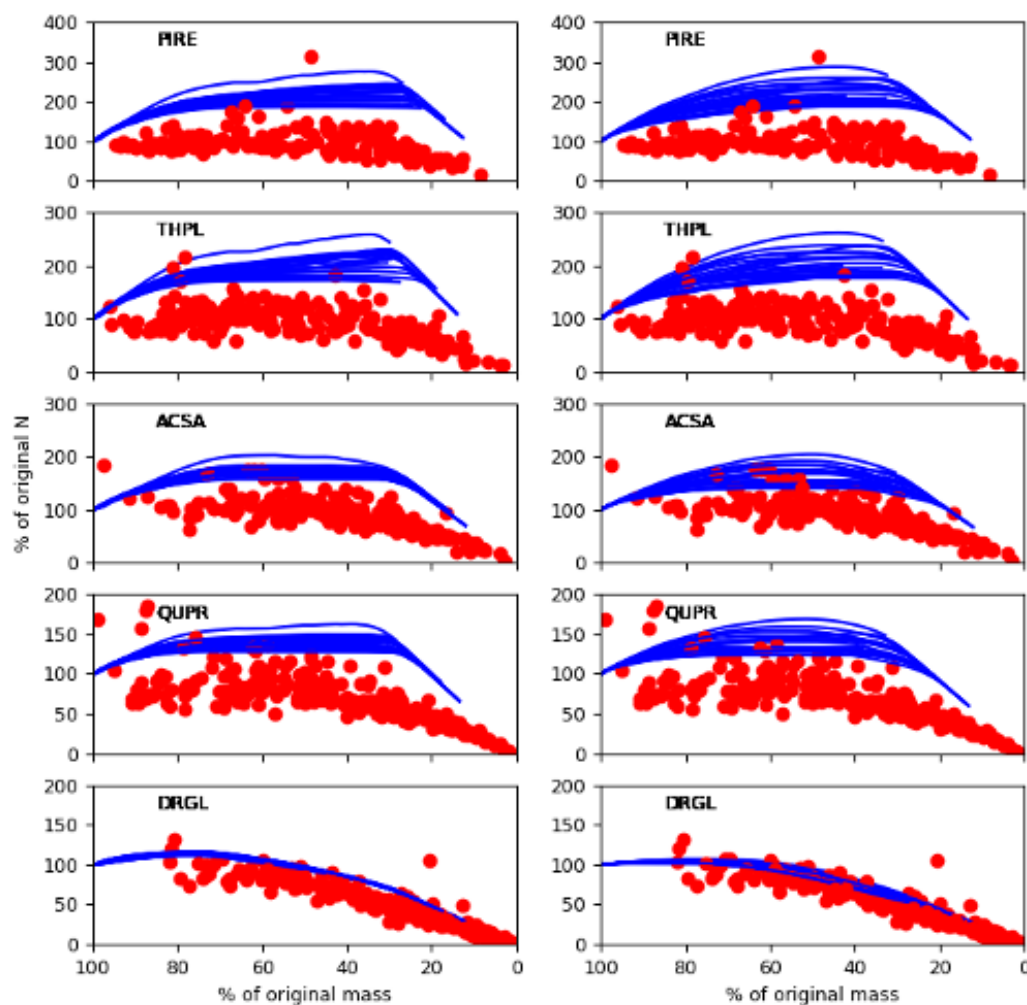



Figure 4: The modeled (blue) compared to observed (filled red circles) percentage of original nitrogen as a function of the percentage of original litter mass for 5 of the litter types. CTC$_f$ results are shown in the left column, while CNT$_f$ results are shown in the right column. Results are aggregated over all sites used in the LIDET study.

3.2 Model sensitivity analysis

We conducted a global sensitivity analysis of two model output variables to 11 CTC$_f$ parameters (Table 4). For the fraction of carbon remaining, the sensitivities are plotted for each of the 10 years of the experiment (Figure 5). In the stacked bar plots, the height of each bar represents the main effect sensitivity index for a particular parameter. Within each stacked bar, if the sensitivity is greater than 0.05, a + or – symbol is included to show whether positive

perturbations of that parameter result in positive or negative effects on the output of interest (i.e, positive or negative correlations). For the tropical site BCI (Figure 5a), the fraction of carbon remaining is sensitive in year 1 to five parameters: the fraction of potential immobilization outside the litter bag (*FPI_outside*), the respiration fraction

360   multiplier (*rf_mult*), the multiplier for SOM decomposition rates in pools 1-3 (*k_som123_mult*), the multiplier for litter decomposition rates (*k_lit_mult*) and the lignin fraction of input litter (*litter_flig*). For the first four parameters, there are negative correlations with the output variable. For example, higher values of *FPI_outside* indicate less nitrogen limitation and therefore faster decomposition of litter, decreasing the remaining carbon in year 1. Higher values of *k_som123_mult* and *k_lit_mult* directly drive faster decomposition. For *rf_mult*, a higher

365   respiration fraction means that more carbon is lost to the atmosphere during litter decomposition instead of transitioning to SOM. This directly reduces the carbon remaining in the system. In addition, because there is less litter transitioning to SOM, higher *rf_mult* also reduces immobilization demand, results in less nitrogen limitation and therefore faster litter decomposition. On the other hand, higher *litter_flig* means an increased proportion of litter with the slowest turnover rate out of the three litter pools. This results in increased retention of carbon in the

370   system.

Over the 10 years of the simulations, significant changes occur in the parameter sensitivities. At BCI, we see that the relative importance of *FPI_outside* decreases quickly and is no longer sensitive after three years. At this stage, a large fraction of the litter has already transitioned to SOM. Therefore, there is less immobilization demand, and additionally some of that demand may now be met by mineralization occurring within the bag. Also because of

375   litter transitioning to SOM, *k_lit_mult* becomes less sensitive while *k_som123_mult* become more sensitive. *rf_mult* becomes increasingly important during the first three years and remains sensitive through the simulation period. The respiration fraction parameter continues to be important as it applies not only to litter to SOM transition, but cascading SOM transitions as well (e.g., SOM1 to SOM2 and SOM2 to SOM3). Towards the end of the simulation, especially in the last year, the carbon remaining becomes marginally sensitive to *k_som4_mult*,

380   which is the longest-lived SOM pool with a base turnover rate of 27 years.

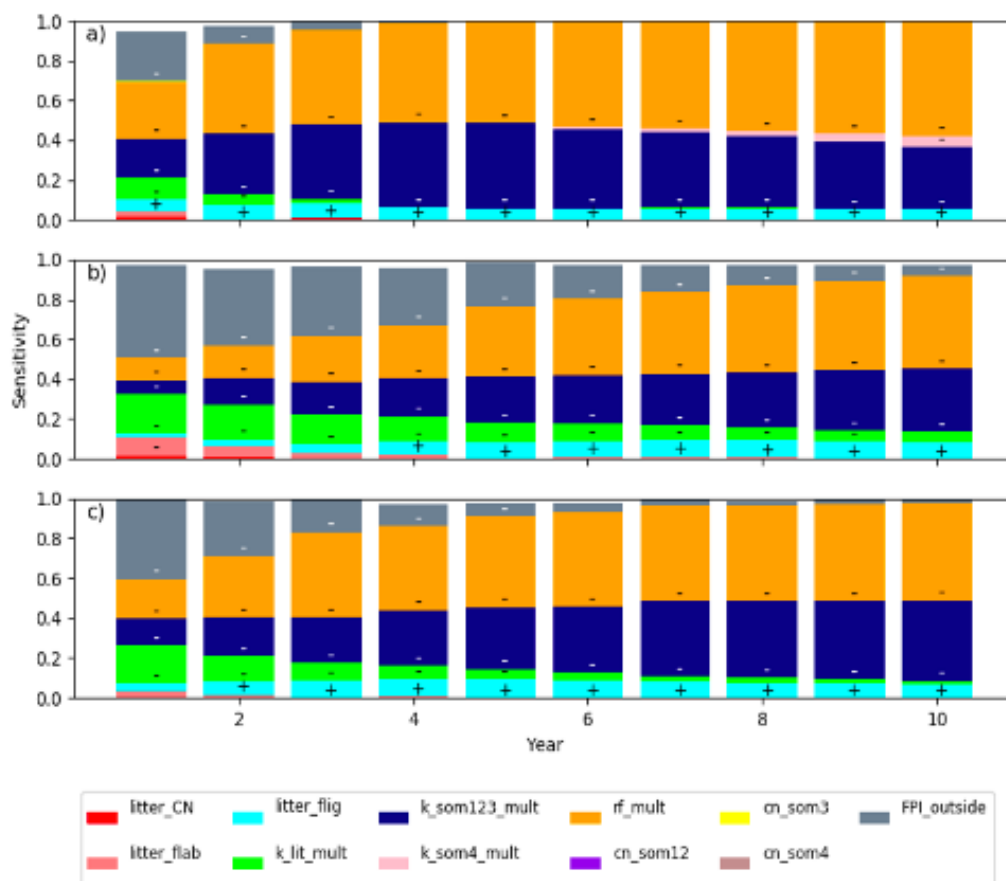



Figure 5. Sensitivity analysis for the fraction of carbon remaining for a) Barro Colorado Island (BCI); b). Bonanza Creek (BNZ), and c) Harvard Forest (HFR). First-order sensitivity indices for each parameter are indicated by the height of the colored bar. The "+" or "-" symbols indicate whether a positive perturbation of the parameter results in a positive or negative deviation in the model output, respectively.

Large differences in parameter sensitivities are also seen across sites. The contrast in sensitivities between BCI (Figure 5a) and BNZ (Figure 5b) reflects the extreme difference in environmental conditions between the tropical and boreal sites. *FPI_outside* is much more sensitive at BNZ and remains highly sensitive for a longer period of time. Therefore, litter decomposition in boreal systems may be more sensitive to the conditions in the surrounding ecosystem. There is also increased initial sensitivity to *k_lit_mult*, and a longer time to transition to higher sensitivity of *k_som_mult*, reflecting the slower decomposition rates at BNZ. The fraction of litter that is labile (*litter_flab*) is also more sensitive initially at BNZ than at BCI. The default base rate for the labile litter pool is

395 less than one day, so that at BCI after one year there is likely very little labile litter left. However, at much colder and drier BNZ, some labile material may remain especially in simulations with low values of *FPI_outside* (as low as 0.03) and low *k_lit_mult* (as low as 0.2) that could combine to produce an effective decomposition rate 150 times slower than the default parameterization with no nutrient limitation. While the magnitudes vary, the temporal patterns of sensitivity and correlations with the carbon remaining are similar between the two sites. For the

400 deciduous site HFR (Figure 5c), the sensitivities generally fall in between BNZ and BCI with similar correlations and temporal patterns.

We also calculate the sensitivity of the fraction of nitrogen remaining as a function of time for the same three sites (Figure 6). This quantity of interest is the most sensitive to the litter carbon to nitrogen ratio (*litter_CN*) for all sites and years. This parameter is the most important in determining how much immobilization occurs over the

405 simulation. Higher carbon to nitrogen ratios lead to increased immobilization, and greater amounts of nitrogen remaining. This high sensitivity is also reflected in the large differences across litter types seen in the default model (Figure 4). At all three sites and years, *rf_mult*, *k_som123_mult* and *cn_som3* are also sensitive parameters. All three of these parameters are negatively correlated with the nitrogen remaining. Higher values of *rf_mult* lead to more carbon leaving the system and lower immobilization demand. Higher values of *k_som123_mult* lead to

410 increased SOM decomposition and nitrogen mineralization, reducing the demand for external nitrogen. Higher values of *cn_som3* also lead to decreased immobilization demand. At the boreal site, *cn_som12* and *FPI_outside* are sensitive early in the simulation, reflecting the slower turnover times of the labile and cellulose pools and their associated SOM pools. At the tropical site, *cn_som4* is sensitive late in the simulation due to faster turnover and cascading through the SOM pools.

415

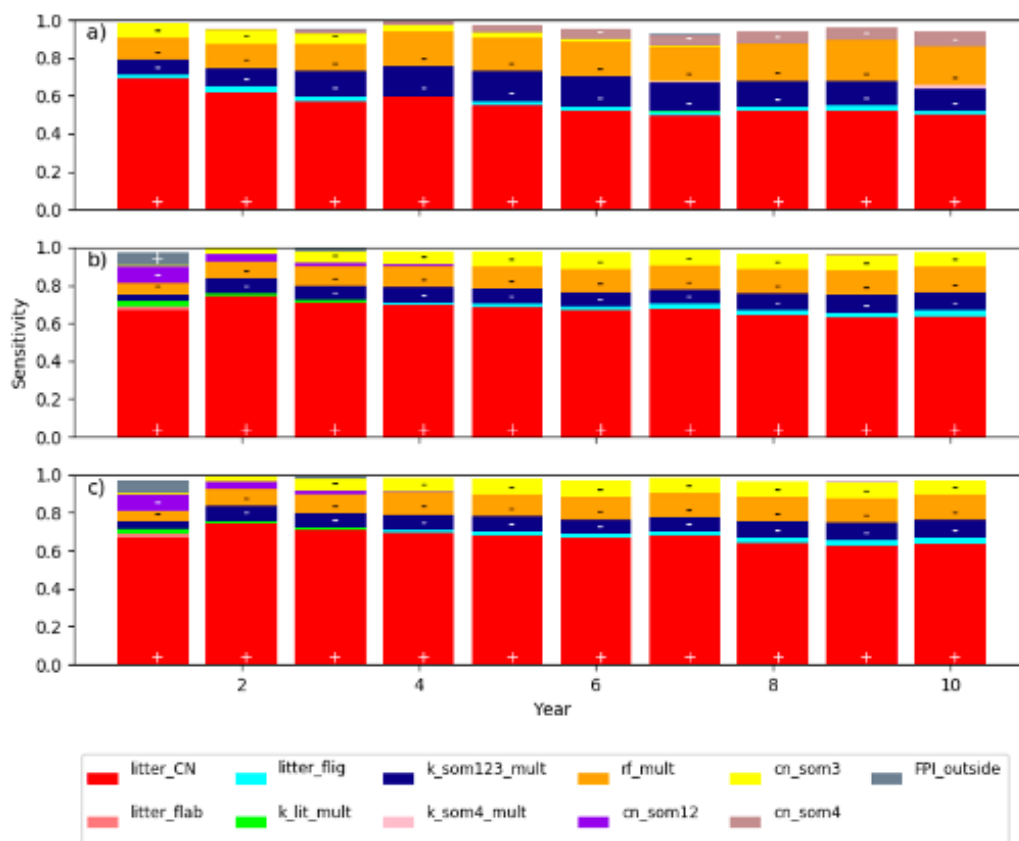



Fig 6. Sensitivity analysis for the fraction of nitrogen remaining for a) Barro Colorado Island (BCI); b). Bonanza Creek (BNZ), and c) Harvard Forest (HFR). First-order sensitivity indices for each parameter are indicated by the height of the colored bar. The "+" or "-" symbols indicate whether a positive perturbation of the parameter results in a positive or negative deviation in the model output, respectively

420 In addition to the sensitivity analysis, we also plot the range of predictions of the CTC_f ensemble at the same three sites for three litter types (Figure 7). In these simulations, the litter chemistry parameters were held constant at the values specified for the litter type of interest (table 2) while the other eight parameters were allowed to vary over their ranges (table 4). The purpose of these simulations is to determine the relative contributions of parameter and model structural uncertainty to the predictions. The combinations of parameters produce a wide range in the

425 predictions of the fraction of carbon remaining. The *Pinus resinosa* (PIRE) type, which has the highest C to N ratio of the three types shown, has the largest spread at all three types. The predictions for the default CTC_f and CNT_f models are also shown for comparison. For all cases except one, the uncertainty range of parameters in CTC_f

leads to a range of predictions that encompasses the CNT$_f$ default. For the *Acer saccharum* (ACSA) at the deciduous
site, CNT$_f$ retains more carbon than CTC$_f$ for any combination of CTC$_f$ parameters for the first four years of the
430   simulation. This may be due in part to the relatively high labile fraction of ACSA compared to the other litter types
(table 2). Because of the very fast turnover of labile litter, and the fast cascade through the SOM pools with carbon
lost at each transition, this carbon is lost from the system quickly with any combination of the uncertain parameters.
However, in CNTf, in addition to the slower turnover, there is a more complex set of transitions between SOM
pools (Koven et al, 2013). Some of the matter from the first soil organic matter (SOM1) pool transitions directly
435   to the slowest pool (SOM3) rather than cascading to SOM2, meaning that a larger fraction of the originally labile
pool may persist. Similar behavior is seen at the boreal site BNZ; the range of uncertainty in CTC$_f$ only barely
includes the CNT$_f$ prediction.

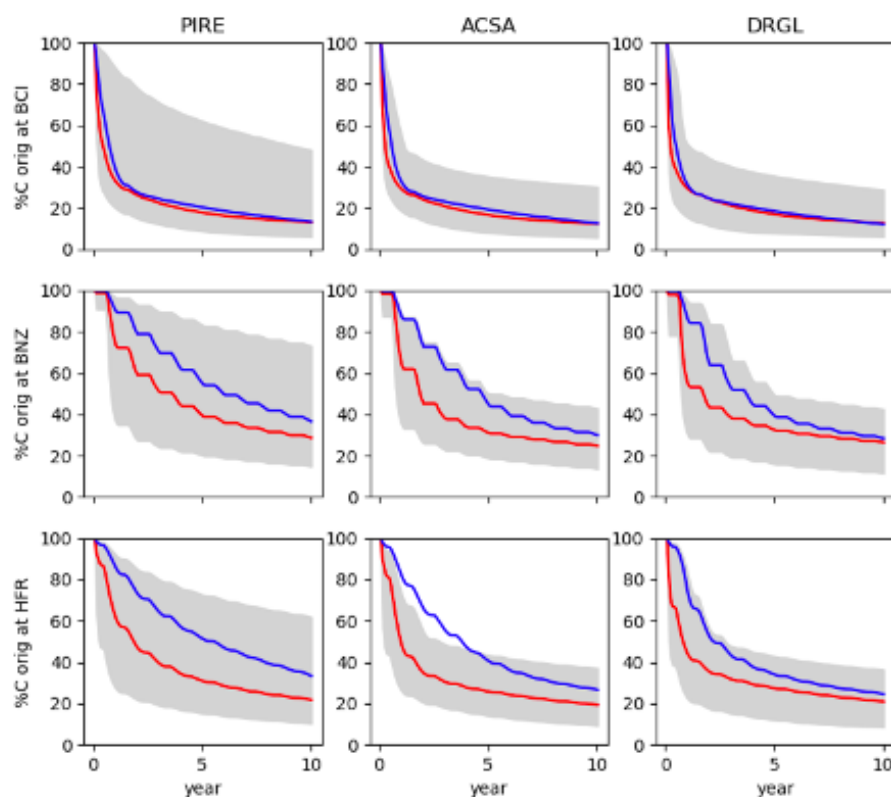



Figure 7: Spread in the predicted fraction of carbon remaining due to parameter uncertainty in CTCf, indicated by the gray
440   envelope around the default prediction (red line). The default CNFf simulation (blue line) is shown for comparison. Parameter
ranges are defined in Table 4, with the litter chemistry parameters held constant for the specific litter types. Results are plotted
for three litter types and three sites: BCI (first row), BNZ (second row), and HFR (third row).

For the percentage of original nitrogen remaining as a function of carbon remaining, the range of uncertainty in

445   CTC$_f$ always includes the CNT$_f$ prediction (Figure 8). The predictions for the PIRE and ACSA litter types have large ranges of uncertainty for all sites: For these two types, the peak amount of nitrogen is between 300% and 400% of the original in the highest cases while in the lowest cases it is barely over 100%. This large range reflects a high uncertainty in the total amount of immobilization, even when the litter chemistry parameters are specified for specific litter types. Therefore, the *rf_mult*, *cn_som* and *k_som_mult* parameters are largely responsible (Figure

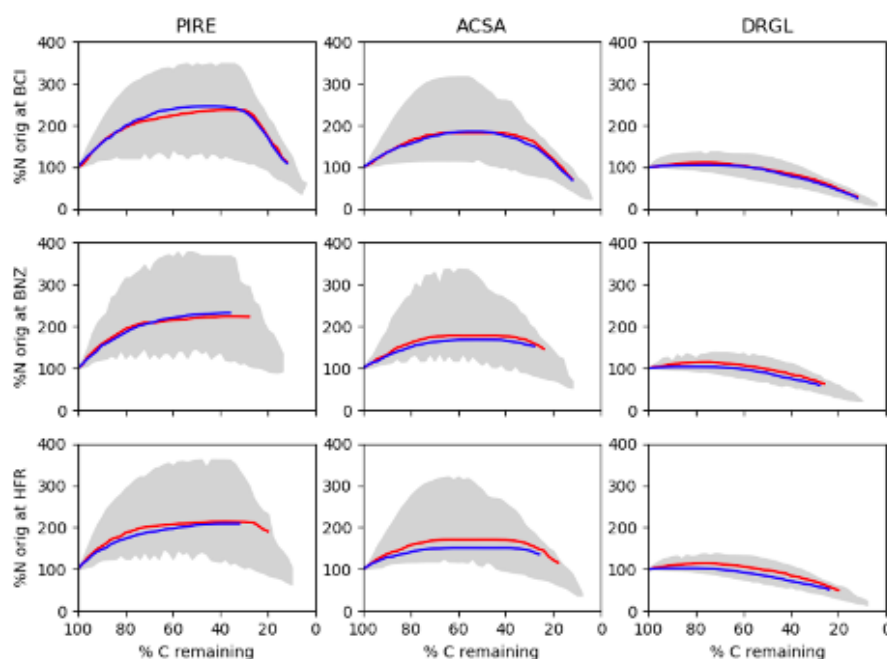
450   6). However, a substantial number of observations include values below 100% at the same levels of carbon remaining (Figure 4). This may indicate structural errors in CTC$_f$ because even with the large range of parameter uncertainty, the model predictions do not encompass the observations. A much smaller range of uncertainty is seen for the *Drypetes glauca* (DRGL) type, which has a carbon to nitrogen ratio of 24.2. For most combinations of parameters, there is little or no immobilization of nitrogen required for this litter type.

455

Figure 8: Spread in the nitrogen remaining as a function of carbon remaining due to parameter uncertainty in CTC$_f$, indicated by the gray envelope around the default prediction (red line). The default CNT$_f$ simulation (blue line) is shown for comparison. Parameter ranges are defined in Table 4, with the litter chemistry parameters held constant for the specific litter types. Results are plotted for three litter types and three sites: BCI (first row), BNZ (second row), and HFR (third row).

460

4 Discussion

We use LIDET observations to evaluate two different decomposition submodel structures within the E3SM framework. We first apply the offline E3SM land-surface model (ELM) to spin up the vegetation and soil conditions for 18 sites using both decomposition submodels. The predicted soil moisture, temperature, and nutrient

465 limitation conditions are used as boundary conditions for our functional decomposition submodel units CTC_f and CNT_f. The functional units are designed to replicate the unique conditions in the litter bag experiment. In predicting the amount of original carbon remaining over time, CTC_f outperforms CNT_f in some ecosystems while CNT_f outperform CTC_f in others. The two models behave similarly in predicting the percentage of original nitrogen remaining. Sensitivity analyses of CTC_f indicate 5-6 key parameters out of 11 drive uncertainty in the amount of

470 carbon or nitrogen remaining, but which parameters are identified depend on the site, time and quantity of interest. Our results are substantially different from those reported in Bonan et al. (2013), which report that based on the comparison to LIDET, DAYCENT (the foundation for the CNT_f model structure) is a superior decomposition model structure to CTC_f. We show that the results are inconclusive for predicting the mass remaining, that both models show significant biases in nitrogen cycling, and that parameter uncertainty is at least as important as

475 structural differences between the two models. There are several important differences in experimental design that contribute to these disparate conclusions. First, we note that Bonan et al (2013) also uses a functional unit approach similar to CTC_f and compares it to DAYCENT, a version of the CENTURY model (Parton et al., 1998). However, they did not use the Community Land Model (CLM) or any other land surface models to supply the boundary conditions for soil temperature, moisture and nutrient limitation. Instead, they applied constant environmental and

480 nutrient scalars. The climatic decomposition index (CDI) was used following Moorhead et al. (1999). The CDI values used generally agree well with the environmental scalars calculated by ELM (Table 6), but do not account for seasonal variations in these factors as we do in ELM. These seasonal variations may be quite large (Figure 3), and using constant values may result in missing important nonlinear interactions that occur between abiotic and biotic factors.

485 Similarly, constant values are used for the fraction of potential immobilization (FPI) in Bonan et al. (2013). They base the primary comparison on the assumption that N is not limiting, but they note that better agreement is obtained for CTC_f when nitrogen is assumed to severely limit decomposition. In fact, this assumption of high limitation (low FPI) is internally consistent with the predictions of ELM at most of the sites in this analysis, especially during the growing season when most decomposition is occurring (Figure 3). Both ELM and CLM4.5 include vertically

490   resolved soil organic matter and litter decomposition (Koven et al., 2013), which was a subsequent model development not tested in the Bonan et al. (2013) analysis. Due to the effects of higher inputs of litter into the top soil layers combined with the slow process of vertical diffusion that moves SOM to deeper layers, immobilization demand tends to be much higher than mineralization in these top layers. Therefore, values of FPI are much lower in the first soil layer (that is used to provide the boundary condition for the LIDET experiment) than for a column-

495   level average. It is also important to note that although representing soil carbon with a single layer led to low biases in global SOC with CTC in the CESM framework (Todd-Brown et al., 2013), using CTC in ELM and E3SM with vertically resolved SOC produces global SOC stocks more in line with observations (Burrows et al., accepted). Another discrepancy between our analysis and that of Bonan et al. (2013) is that our analysis of CNT_f differs in a few key ways from DAYCENT. First, CNT_f differs from DAYCENT because it was adapted to better match the

500   existing model structure in CLM4.5 and ELM. Although the litter and SOM pool structures, flows and turnover times in CNT_f closely match DAYCENT, the carbon to nitrogen ratios for SOM are fixed parameters in CNT_f while in DAYCENT these ratios increase with low soil mineral nitrogen. The specified values for these carbon to nitrogen ratios are on the low end of the ranges used in DAYCENT (Parton et al., 1998), and are similar to those used in CTC_f (Table 3). Although we did not test the parameter sensitivities of CNT_f, we note that the carbon to nitrogen

505   ratio and the respiration fraction are sensitive parameters for the fraction of nitrogen remaining in CTC_f. This may explain why CTC_f and CNT_f behave similarly for this variable (Figure 4), although DAYCENT performed significantly better than CTC_f in Bonan et al (2013). Adjusting these parameters or allowing them to be variable in time may improve the performance of both models. Second, setting the FPI boundary conditions for the functional units with ELM simulations also causes important differences for the comparison between the two

510   models. Because of the difference in base turnover rates, immobilization demand is generally lower in CNT_f than CTC_f, leading to higher FPI values in CNT_f and effectively reducing the difference in actual litter turnover rates between the two models. This suggests that nitrogen availability is a limiting factor on litter decomposition, which is consistent with experimental studies (Cleveland et al., 2006). Therefore, under certain conditions, the model turnover rates and structure for litter decomposition is less important than how nitrogen mineralization and external

515   nitrogen inputs are represented. Finally, we attempt to more closely represent the unique environment within the litter bag. Inside the bag at the beginning of the experiment, there is likely to be higher unmet immobilization demand because of the relatively high C:N ratios and lack of SOM. Therefore, we believe the weighted FPI (equation 2) is a more realistic representation of the nitrogen limitation inside the litter bag. Generally, this value of FPI is lower than the external FPI at the beginning of the experiment and contributes to slower decomposition.

520   A key parameter sensitivity in CTC$_f$ is the respiration fraction multiplier (*rf_mult*). This parameter sets the carbon loss for the transitions between litter and SOM and between different SOM pools. Therefore, for the LIDET simulation experiment it is an important control on the carbon remaining in the system over time and on the amount of immobilization demand caused by litter decomposition. These respiration fraction parameters are closely linked (inversely) to carbon use efficiency (CUE), which describes the metabolic efficiency of microbes in the

525   decomposition process. While the respiration fractions for specific transitions are constant in CTC$_f$ and CNT$_f$, there is strong evidence for the dependence of CUE on environmental and stoichiometric parameters (Manzoni et al., 2012; Sinsabaugh et al., 2013). Some of this dependence can be captured by the variation in respiration fractions among different pool transitions, but the model is limited in its capability to represent CUE dynamics because it does not explicitly represent microbial biomass or communities. In the absence of these mechanisms, we may

530   consider CUE (and therefore respiration fraction) as uncertain parameters to be best fit for the observations of interest. However, CUE may be positively or negatively impacted by different aspects of climate change (Manzoni et al., 2012), so that predictions of litter and soil carbon dynamics under climate change scenarios using fitted CUE parameters under present day conditions may be inaccurate. Including a more mechanistic representation of CUE in Earth system models has been increasingly recognized over the last several years an important goal.

535   A number of model development efforts are underway to represent microbes explicitly in globally relevant modelling frameworks that are being or may eventually be included in Earth system models. Such models may more accurately represent the dynamics of CUE and therefore also improve the representation of SOM. For example, the microbial-enzyme-mediated decomposition model (MEND) explicitly represents microbial biomass carbon and two key enzymes in the decomposition process (Wang et al., 2013; Wang et al., 2014). The MEND

540   model includes physically measurable pools and dynamic CUE, which can change the sign of soil organic carbon response compared to the assumption of constant CUE in a warming scenario (Wang et al., 2013). Globally, the temperature sensitivity of CUE may regulate the long-term response of soil carbon to warming (Li et al., 2014). Another SOC model, Carbon, Organisms, Rhizosphere, and Protection in the soil environment (CORPSE) can simulate root-microbe interactions (e.g., priming effects) and the formation of protected SOM pools that may occur

545   under certain conditions (Sulman et al., 2014). A global-scale comparison of CORPSE and the Microbial-Mineralization carbon stabilization model (MIMICS) model (Wider et al., 2014), which both represent microbes explicitly, shows disagreement on the magnitude of soil carbon stocks and their long-term fate (Wieder et al., 2018). Oscillatory behaviour is also observed with many of these microbially explicit approaches (Wang et al., 2014b, Li et al., 2014). More model-data integration may be necessary to constrain these increasingly mechanistic models to

550   produce consistent and credible responses over long timescales. MIMICs was also compared to selected LIDET

observations with favourable results (Wider et al., 2014). Unfortunately, the breadth of measurements made in the LIDET study is not comprehensive enough to constrain these complex models, therefore, a new LIDET type study that includes measurements of microbial biomass, communities and enzymes may be of great benefit to the community.

555   In addition to the lack of explicit microbes in CTC_f and CNT_f, other structural shortcomings may contribute to biases seen in the models. A large number of observations indicate the fraction of nitrogen remaining falling below 100% for intermediate values of carbon mass remaining during which the CTC_f predictions are always above 100% for any combination of parameters. This may indicate a mechanism for nitrogen loss that is not currently represented in ELM. Soil mineral nitrogen may be lost through denitrification and leaching, but these terms tend

560   to be very small in the first several years of the simulations due to high immobilization demand. Leaching of organic nitrogen may be an important loss pathway from the litter bags, although it is not usually considered in terrestrial nitrogen budgets (Neff et al., 2003). Including this effect in the model may improve predictions. The "home field" advantage effect may also cause biases in model predictions. This effect refers to litter types that are introduced into ecosystem where they are native decomposing faster because the microbial communities are present

565   that prefer these types of inputs, whereas non-native litter types may be less preferred. This was found to be an important factor for DRGL in the LIDET experiment (Gholz et al., 2000). Parton et al. (2007) also noted that the relationship between CDI and decomposition rates did not work well for arid grasslands as for other biomes in the LIDET study and hypothesized that ultraviolet radiation may accelerate decomposition in these systems. This is supported by other studies in arid systems (e.g. Gallo et al., 2006, Brandt et al., 2007). For this reason, we did not

570   include the arid grassland biome in our study. However, UV radiation may play a role in other ecosystems as well, including in mesic grasslands. A litter decomposition study designed to measure the role of UV photodegradation (Brandt et al., 2010) reported significant UV effects at the Cedar Creek (CDR) site, a humid grassland system also used in the LIDET study.

5 Conclusions

575   To predict carbon cycle feedbacks to climate change, it is critically important to model the decomposition of litter and soil organic matter accurately. A functional unit modelling approach was designed to simulate the LIDET decomposition study using two decomposition models in the ELM framework. We found that the converging trophic cascade (CTC) decomposition model, which was parameterized with a series of mesocosm experiments, reproduces the observed patterns of decomposition in LIDET reasonably well when driven with values of soil

580   moisture, temperature, and nutrient limitation that are internally consistent with ELM simulations. The introduction

of vertically resolved SOM in CLM4.5 and later ELM helped to correct apparent biases related to previous single-layer implementations of CTC that resulted in too rapid litter decomposition and underpredicted soil carbon stocks. When both decomposition models are implemented in ELM and tested against LIDET observations, CTC is not substantially different from the DAYCENT-based CNT model; the spread of predictions resulting from parameter 585 uncertainty is equal to or greater than the differences caused by model structure. Modelled outputs were highly sensitive to the respiration fraction, highlighting the importance of accurately simulating carbon use efficiency. In the future, microbially explicit modelling frameworks currently under development in Earth system models would benefit from litter bag experiments coupled with observations of microbial population dynamics.

590

Acknowledgements:

This material is based upon work supported by the US Department of Energy, Office of Science, Office of Biological and Environmental Research. Oak Ridge National Laboratory is managed by UT-Battelle, LLC, for the 595 US Department of Energy under contract DE-AC05-00OR22725. The views expressed in this article do not necessarily represent the views of the US Department of Energy or the United States Government. This research used resources of the Compute and Data Environment for Science (CADES) at the Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC05-00OR22725.

600

605

References

610   Adair, E. C., W. J. Parton, S. J. Del Grosso, W. L. Silver, M. E. Harmon, S. A. Hall, I. C. Burke, and S. C. Hart: Simple three-pool model accurately describes patterns of long-term litter decomposition in diverse climates, Global Change Biol, 14(11), 2636-2660, doi: 10.1111/j.1365-2486.2008.01674.x, 2008.

Bonan, G. B., M. D. Hartman, W. J. Parton, and W. R. Wieder: Evaluating litter decomposition in earth system models with long-term litterbag experiments: an example using the Community Land Model version 4
615   (CLM4), Global Change Biol, 19(3), 957-974, doi: 10.1111/gcb.12031, 2013.

Brandt, L. A., J. Y. King, S. E. Hobbie, D. G. Milchunas and R. L. Sinsabaugh: The Role of Photodegradation in Surface Litter Decomposition Across a Grassland Ecosystem Precipitation Gradient. Ecosystems **13**(5): 765-781, 2010.

Burrows, S. M., M. Maltrud, X. Yang, Q. Zhu, N. Jeffery, X. Y. Shi, D. Ricciuto, S. Wang, G. Bisht, J. Y. Tang,
620   J. Wolfe, B. E. Harrop, B. Singh, L. Brent, T. Zhou, P. Cameron-Smith, N. Keen, N. Collier, M. Xu, E. C. Hunke, S. M. Elliott, A. K. Turner, H. Li, H. Wang, J. C. Golaz, B. Bond-Lamberty, F. M. Hoffman, W. J. Riley, P. E. Thornton, K. Calvin and L. R. Leung: The DOE E3SM v1.1 biogeochemistry configuration: Description and simulated ecosystem-climate responses to historical changes in forcing. Journal of Advances in Modeling Earth Systems, 2020.

625   Carreiro, M. M., R. L. Sinsabaugh, D. A. Repert, and D. F. Parkhurst (2000), Microbial enzyme shifts explain litter decay responses to simulated nitrogen deposition, Ecology, 81(9), 2359-2365, doi: Doi 10.1890/0012-9658(2000)081[2359:Meseld]2.0.Co;2.

Cleveland, C. C. and A. R. Townsend (2006). Nutrient additions to a tropical rain forest drive substantial soil carbon dioxide losses to the atmosphere. Proceedings of the National Academy of Sciences of the United
630   States of America **103**(27): 10316-10321.

Currie, W. S., M. E. Harmon, I. C. Burke, S. C. Hart, W. J. Parton and W. Silver (2010). "Cross-biome transplants of plant litter show decomposition models extend to a broader climatic range but lose predictability at the decadal time scale." Global Change Biology **16**(6): 1744-1761.

Cusack, D. F., W. W. Chou, W. H. Yang, M. E. Harmon, W. L. Silver, and L. Team (2009), Controls on long-term
635   root and leaf litter decomposition in neotropical forests, Global Change Biol, 15(5), 1339-1355, doi: 10.1111/j.1365-2486.2008.01781.x.

Dirmeyer, P. A., X. A. Gao, M. Zhao, Z. C. Guo, T. K. Oki, and N. Hanasaki (2006), GSWP-2 - Multimodel anlysis and implications for our perception of the land surface, B Am Meteorol Soc, 87(10), 1381-+, doi: 10.1175/Bams-87-10-1381.

640   Friedlingstein, P., M. Meinshausen, V. K. Arora, C. D. Jones, A. Anav, S. K. Liddicoat, and R. Knutti (2014), Uncertainties in CMIP5 Climate Projections due to Carbon Cycle Feedbacks, J Climate, 27(2), 511-526, doi: 10.1175/Jcli-D-12-00579.1.

Gallo, M. E., R. L. Sinsabaugh and S. E. Cabaniss (2006). The role of ultraviolet radiation in litter decomposition in and ecosystems. Applied Soil Ecology **34**(1): 82-91.

645   Gholz, H. L., D. A. Wedin, S. M. Smitherman, M. E. Harmon and W. J. Parton (2000). "Long-term dynamics of pine and hardwood litter in contrasting environments: toward a global model of decomposition." Global Change Biology **6**(7): 751-765.

Harmon, M. E., W. L. Silver, B. Fasth, H. Chen, I. C. Burke, W. J. Parton, S. C. Hart, W. S. Currie, and Lidet (2009), Long-term patterns of mass loss during the decomposition of leaf and fine root litter: an intersite
650   comparison, Global Change Biol, 15(5), 1320-1338, doi: 10.1111/j.1365-2486.2008.01837.x. Koven, C. D., J. Q. Chambers, K. Georgiou, R. Knox, R. Negron-Juarez, W. J. Riley, V. K. Arora, V. Brovkin, P. Friedlingstein, and C. D. Jones (2015), Controls on terrestrial carbon feedbacks by productivity versus turnover in the CMIP5 Earth System Models, Biogeosciences, 12(17), 5211-5228, doi: 10.5194/bg-12-5211-2015.

655   Koven, C. D., W. J. Riley, Z. M. Subin, J. Y. Tang, M. S. Torn, W. D. Collins, G. B. Bonan, D. M. Lawrence, and S. C. Swenson (2013), The effect of vertically resolved soil biogeochemistry and alternate soil C and N models on C dynamics of CLM4, Biogeosciences, 10(11), 7109-7131, doi: 10.5194/bg-10-7109-2013.

Koven, C. D., D. M. Lawrence and W. J. Riley (2015). "Permafrost carbon-climate feedback is sensitive to deep soil carbon decomposability but not deep soil nitrogen dynamics." Proceedings of the National Academy
660   of Sciences of the United States of America **112**(12): 3752-3757.

Li, J. W., G. S. Wang, S. D. Allison, M. A. Mayes and Y. Q. Luo (2014). Soil carbon sensitivity to temperature and carbon use efficiency compared across microbial-ecosystem models of varying complexity. Biogeochemistry **119**(1-3): 67-84.

Luo, Y. Q., et al. (2016), Toward more realistic projections of soil carbon dynamics by Earth system models, Global
665   Biogeochem Cy, 30(1), 40-56, doi: 10.1002/2015gb005239.

Manzoni, S., P. Taylor, A. Richter, A. Porporato and G. I. Agren (2012). Environmental and stoichiometric controls on microbial carbon-use efficiency in soils. New Phytologist **196**(1): 79-91.

Moorhead, D. L., W. S. Currie, E. B. Rastetter, W. J. Parton and M. E. Harmon (1999). Climate and litter quality controls on decomposition: An analysis of modeling approaches. Global Biogeochemical Cycles **13**(2):
670  575-589.

Neff, J. C., F. S. Chapin and P. M. Vitousek (2003). Breaks in the cycle: dissolved organic nitrogen in terrestrial ecosystems. Frontiers in Ecology and the Environment **1**(4): 205-211.

Oleson, K. W., et al. (2013), Technical description of version 4.5 of the Community Land Model (CLM), edited by N. C. f. A. Research, p. 422, Ncar Technical Note NCAR/TN-503+STR, Boulder, CO.

675 Parton, W., et al. (2007), Global-scale similarities in nitrogen release patterns during long-term decomposition, Science, 315(5810), 361-364, doi: 10.1126/science.1134853.

Parton, W. J., M. Hartman, D. Ojima and D. Schimel (1998). "DAYCENT and its land surface submodel: description and testing." Global and Planetary Change **19**(1-4): 35-48.

Parton, W. J., J. W. B. Stewart and C. V. Cole (1988). "Dynamics of C, N, P and S in Grassland Soils - a Model."
680  Biogeochemistry **5**(1): 109-131.

Ricciuto, D., K. Sargsyan, and P. Thornton (2018), The Impact of Parametric Uncertainties on Biogeochemistry in the E3SM Land Model, J Adv Model Earth Sy, 10(2), 297-319, doi: 10.1002/2017ms000962.

Sargsyan, K., C. Safta, H. N. Najm, B. J. Debusschere, D. Ricciuto and P. Thornton (2014). Dimensionality Reduction for Complex Models Via Bayesian Compressive Sensing. International Journal for Uncertainty
685  Quantification **4**(1): 63-93.

Sinsabaugh, R. L., S. Manzoni, D. L. Moorhead and A. Richter (2013). Carbon use efficiency of microbial communities: stoichiometry, methodology and modelling. Ecology Letters **16**(7): 930-939.

Sulman, B. N., R. P. Phillips, A. C. Oishi, E. Shevliakova and S. W. Pacala (2014). Microbe-driven turnover offsets mineral-mediated storage of soil carbon under elevated CO2. Nature Climate Change **4**(12): 1099-1102.

690 Thornton, P. E., S. C. Doney, K. Lindsay, J. K. Moore, N. Mahowald, J. T. Randerson, I. Fung, J. F. Lamarque, J. J. Feddema, and Y. H. Lee (2009), Carbon-nitrogen interactions regulate climate-carbon cycle feedbacks: results from an atmosphere-ocean general circulation model, Biogeosciences, 6(10), 2099-2120.

Thornton, P. E., J. F. Lamarque, N. A. Rosenbloom, and N. M. Mahowald (2007), Influence of carbon-nitrogen cycle coupling on land model response to CO2 fertilization and climate variability, Global Biogeochem
695  Cy, 21(4), doi: Artn Gb4018 10.1029/2006gb002868.

Thornton, P. E., and N. A. Rosenbloom (2005), Ecosystem model spin-up: Estimating steady state conditions in a coupled terrestrial carbon and nitrogen cycle model, Ecological Modelling, 189(1-2), 25-48, doi: 10.1016/j.ecolmodel.2005.04.0

Thornton, P. E., B. E. Law, H. L. Gholz, K. L. Clark, E. Falge, D. S. Ellsworth, A. H. Goldstein, R. K. Monson, D.
700   Hollinger, M. Falk, J. Chen and J. P. Sparks 2002. Modeling and measuring the effects of disturbance
           history and climate on carbon and water budgets in evergreen needleleaf forests. Agricultural and Forest
           Meteorology 113: 185-222

Tian, H. Q., et al. (2015), Global patterns and controls of soil organic carbon dynamics as simulated by multiple
           terrestrial biosphere models: Current status and future directions, Global Biogeochemical Cycles, 29(6),
705   775-792, doi: 10.1002/2014GB005021.

Todd-Brown, K. E. O., J. T. Randerson, W. M. Post, F. M. Hoffman, C. Tarnocai, E. A. G. Schuur, and S. D.
           Allison (2013), Causes of variation in soil carbon simulations from CMIP5 Earth system models and
           comparison with observations, Biogeosciences, 10(3), 1717-1736, doi: 10.5194/bg-10-1717-2013.

Wang, D. L., T. Janjusic, C. Iversen, P. Thornton, M. Karssovski, W. Wu, and Y. Xu (2015), A Scientific Function
710   Test Framework for Modular Environmental Model Development: Application to the Community Land
           Model, 2015 Ieee/Acm 1st International Workshop on Software Engineering for High Performance
           Computing in Science (Se4hpcs), 16-23, doi: 10.1109/SE4HPCS.2015.10.

Wang, G. S., W. M. Post and M. A. Mayes (2013). Development of microbial-enzyme-mediated decomposition
           model parameters through steady-state and dynamic analyses. Ecological Applications **23**(1): 255-272.

715 Wang, G. S., M. A. Mayes, L. H. Gu and C. W. Schadt (2014). Representation of Dormant and Active Microbial
           Dynamics for Ecosystem Modeling. Plos One **9**(2).

Wang, Y. P., B. C. Chen, W. R. Wieder, M. Leite, B. E. Medlyn, M. Rasmussen, M. J. Smith, F. B. Agusto, F.
           Hoffman and Y. Q. Luo (2014b). Oscillatory behavior of two nonlinear microbial models of soil carbon
           decomposition. Biogeosciences **11**(7): 1817-1831.

720 Wieder, W. R., J. Boehnert, and G. B. Bonan (2014), Evaluating soil biogeochemistry parameterizations in Earth
           system models with observations, Global Biogeochem Cy, 28(3), 211-222, doi: 10.1002/2013gb004665.

Wieder, W. R., et al. (2015), Explicitly representing soil microbial processes in Earth system models, Global
           Biogeochem Cy, 29(10), 1782-1800, doi: 10.1002/2015gb005188.

Wieder, W. R., M. D. Hartman, B. N. Sulman, Y. P. Wang, C. D. Koven, and G. B. Bonan (2018), Carbon cycle
725   confidence and uncertainty: Exploring variation among soil biogeochemical models, Global Change Biol,
           24(4), 1563-1579, doi: 10.1111/gcb.13979.

Wu, D. H., S. L. Piao, Y. W. Liu, P. Ciais, and Y. T. Yao (2018), Evaluation of CMIP5 Earth System Models for
           the Spatial Patterns of Biomass and Soil Carbon Turnover Times and Their Linkage with Climate, J
           Climate, 31(15), 5947-5960, doi: 10.1175/Jcli-D-17-0380.1.

730   Yao Z., Wang D., Ricciuto D., Yuan F., Fang C. (2019) Parallel Computing for Module-Based Computational Experiment. In: Rodrigues J. et al. (eds) Computational Science – ICCS 2019. ICCS 2019. Lecture Notes in Computer Science, vol 11537. Springer, Cham. https://doi.org/10.1007/978-3-030-22741-8_27

735