Received: 11 May 2020 | Revised: 5 January 2021 | Accepted: 13 January 2021

DOI: 10.1111/sum.12699

ORIGINAL ARTICLE

Soil Use and Management | BRITISH SOCIETY OF SOIL SCIENCE | WILEY

A model ensemble approach to determine the humus building efficiency of organic amendments in incubation experiments

Anton A. Gasser[1] | **Julius Diel**[1] | **Kerstin Nielsen**[2] | **Paul Mewes**[3] | **Christof Engels**[3] | **Uwe Franko**[1]

[1]Helmholtz Centre for Environmental Research GmbH – UFZ, Halle (Saale), Germany

[2]Institut für Agrar- und Stadtökologische Projekte an der Humboldt-Universität zu Berlin (IASP), Berlin, Germany

[3]Albrecht Daniel Thaer - Institut für Agrar- und Gartenbauwissenschaften der Humboldt-Universität zu Berlin, Berlin, Germany

Correspondence
Anton A. Gasser, Helmholtz Centre for Environmental Research GmbH – UFZ, Halle (Saale), Germany.
Email: anton.gasser@ufz.de

Abstract

Organic amendments are important to sustain soil organic matter (SOM) and soil functions in agricultural soils. Information about the contribution of organic amendments to SOM can be derived from incubation experiments. In this study, data from 72 incubated organic amendments including plant residues, digestates and manure were analysed. The incubation data was compiled from three experimental setups with varying incubation times, soils and incubation temperatures, in which CO_2 release was measured continuously. The analysis of the incubation data was performed with an approach relying on conceptual parts of C-TOOL, CCB, Century, ICBM, RothC and Yasso which are all well-approved first-order carbon models that differ in structure and abstraction level. All models are an approximation of reality, whereby each model differs in understanding of the processes involved in soil carbon dynamics. To accumulate the advantages from each model a model ensemble was performed for each substrate. With the ability of each carbon model to compute the distribution of carbon into specific SOM pools a new approach for evaluating organic amendments in terms of humus building efficiency is presented that, depends on the weighted model fit of each ensemble member. Depending on the organic substrate added to the soil, the time course of CO_2 release in the incubation studies was predicted with different accuracy by the individual model concepts. Averaging the output of the individual models leads to more robust prediction of SOM dynamics. The E_{HUM} value is easy to interpret and the results are in accordance with the literature.

KEYWORDS

carbon incubation, humus efficiency, model ensemble, soil carbon models

1 | INTRODUCTION

The accumulation of soil organic carbon (SOC) is discussed as a possible solution to mitigate climate change (Minasny et al., 2017). Increasing SOC can rebuild soil fertility, reduce soil erosion, and increase yield stability (Bradford et al., 2019; Harden et al., 2018). The accumulation of SOC requires a reduced decomposition of SOC and/or an increased input of organic matter (OM), where the latter depends on the amount and quality of the OM input.

For an efficient agricultural management that sustains soil organic matter (SOM) and closes the nutrient cycle, it is important to know the specific contribution from different organic materials such as manure, plant residues and recently also from digestates of biogas reactors to SOM (Larney & Angers, 2012). SOM is a composition of compounds with different turnover times in soil. Therefore it is important to assess specifically the contribution of added organic matter to the long-lasting part of SOM that is historically summarized under the term 'humus' and which demands certain attention within the debate about carbon sequestration. There are several attempts to use proxies like the C/N ratio, hemicellulose or the lignin content to evaluate the contribution of OM to SOM. But those bio-chemical proxies have restricted capacity to predict the behaviour of OM reliable enough under microbial turnover (Lashermes et al., 2009; Morvan et al., 2005), since other factors like temperature, microbial communities etc. can influence the decomposition (Dignac et al., 2017).

Incubation experiments are a research tool to assess the quality of the added substrates with regard to humification and the microbial turnover within a certain time. Their results are often analysed with statistical methods including different kinds of non-linear regressions where usually only the carbon loss is determined, whereas the transfer from the added organic material (AOM) to SOM generation is not quantitatively included (Cotrufo et al., 2013). Furthermore, there is no well-approved solution to transfer the incubation results from statistical models to the field scale with regard to environmental conditions. Soil carbon models are usually developed for field conditions and reflect a complex understanding of the carbon turnover. The general model approach comprises a network of carbon fluxes between different pools which approximates in an abstract way the microbiological turnover processes in the soil. Consequently, carbon models are able to predict the retention of carbon in soils for specific site conditions. However, soil carbon models need to be parametrized in order to compute carbon fluxes and they require information about the quality of OM (Stockmann et al., 2013). Incubation experiments, where organic matter is mixed with soil and the resulting turnover is observed from the CO_2 evolution over time may contain this information (Jha et al., 2012).

The general understanding of carbon turnover includes several 'unknowns' and its expression in models follows different concepts. Each model has distinct strengths and weaknesses to project the examined processes (Sulman et al., 2018). Therefore, it may be risky to rely on only one specific model. Model ensembles/averaging are a common method to aggregate the prediction of several models into a single prediction, which is expected to be at least as good as the prediction of a single model and also compensates partial weaknesses of a single model (Diks & Vrugt, 2010; Hagedorn et al. 2005). In addition to improving the prediction, the calculated weights of a model averaging can be transferred to further purpose. The carbon models assign specific turnover characteristics to the underlying substrates, which can be used to evaluate the efficiency of a substrate to contribute to the long-lasting SOM. Thus, the model averaging weights might be used to aggregate the humus building efficiency of a substrate into a single value depending on the performance of each ensemble member.

In this study the following questions were addressed:

- How can soil organic carbon models be applied to incubation experiments?
- How can the humus efficiency of added organic matter be expressed with a single parameter based on the results of soil organic carbon models?
- Does a model ensemble increase the reliability of the quality assessment for organic amendments?

2 | MATERIALS AND METHODS

2.1 | Incubation experiments

In this study, data were compiled from three different incubation experiments with varying incubation time, soils and incubation temperatures (Table 1). The first two data sets were obtained from the Institut für Agrar-und Stadtökologische Projekte Berlin (IASP) and are denoted by data1 and data2, the experimental setups vary in their incubation times. A data set that was already published by Sänger et al. (2014) is denoted as data3 and data4, each with the same substrates but with different soils. The third data set was derived from the Humboldt University Berlin and is denoted as data5 and data6 with different incubation periods.

In total, data from 72 incubated organic substrates including plant residues, digestates and manure were analysed (Appendix S1). The organic amendments were incubated together with soil in beakers. The moisture and the temperature were set to a constant state over the incubation time. In each experimental design, the accumulated amount of CO_2 released was measured over the incubation period. Therefore, different sampling techniques were applied, see Appendix S1. What all methods had in common was that cumulative CO_2 released from the added organic matter (AOM) was differentiated from the soil born CO_2 released by reference samples, where only soil was incubated without any organic amendments, assuming no priming effects. Based on the amount of CO_2 C evolved in each substrate, the cumulative amount of total evolved C was calculated for each observation over the entire incubation period. The CO_2 C released from the organic substrates was calculated from the difference of CO_2 C released from the samples with and without AOM.

TABLE 1 An overview of the experimental setups in this study for each data set (AOM = added organic matter)

	Data1	Data2	Data3	Data4	Data5	Data6
Incubation time [d]	139.7	251.7	41	41	301	161
Replicates	6	6	4	4	5	3
Sampling interval [d]	1/24	1/24	1–20, 22, 24, 27, 30, 34, 36, 41	1–20, 22, 24, 27, 30, 34, 36, 41	1, 3, 7, 14, 21, 35, 56, 77, 98, 120, 162, 217, 301	1, 3, 7, 14, 21, 35, 56, 77, 98, 119, 161
Temperature [°C]	20	20	25	25	22	22
Soil ph	5.7	5.7	6.5	6.5	5.9	5.9
Soil addition [g]	40	40	20	20	100	100
Counts of substrates used	7	8	10	10	22	15
Clay, silt, sand [%]	1, 9, 90	1, 9, 90	20, 75, 5	15, 39, 46	7 ,21, 72	7, 21, 72
Soil water content (% of water holding capacity)	60	60	60	60	50	50
AOM type	Digestates, manure	Digestates, manure	Digestates	Digestates	Roots, crop residues	Roots, crop residues

2.2 | Carbon models and their application to incubation data

In this study concepts of carbon models that follow first-order kinetics and are well established on a field scale, namely C-TOOL (Taghizadeh-Toosi et al., 2014), CCB (Franko et al., 2011), CENTURY (Parton et al., 1987; Parton et al. 1994), ICBM (Andrén & Kätterer, 1997), RothC (Coleman & Jenkinson, 1999) and Yasso (Tuomi et al., 2011) are applied to the incubation data. Besides their handling of AOM, they differ in the number of conceptual SOM pools, their interconnection and texture dependencies of turnover time as well as in the depiction of environmental influences on turnover. A detailed description of each model is beyond the scope of this paper, but the adaptations made to apply the models to the incubation data are described in the Appendix S1. Here it should be noted that the original model concepts are adapted to meet the requirements of modelling the incubation data. The adapted model concepts are referred to by their name with asterisk (*).

In order to apply the models to the incubation data with a unique algorithm and identical data structure, they were reduced to their core concept, as described in Appendix S1. Assuming optimal water supply during the incubation period, only the respective temperature and soil texture functions (when available) were considered. For each model, a maximum of two AOM related parameters were fitted. All other model parameters were left constant at the values according to the individual model publications. Following Sierra et al. (2012), each model was implemented as a set of ordinary differential equations in R (R Core Team, 2019) that were solved using the deSolve package (Soetaert et al., 2010). The parameter fitting was accomplished by using the Levenberg-Marquardt algorithm implemented in the nls.lm function from the R package minpack.lm (Elzhov et al. 2016). The optimized parameters are described in Table 2 for each soil carbon model for the observed cumulative net CO_2 production from AOM.

2.3 | Evaluating the substrate quality in terms of humus efficiency

For a clear differentiation, italic miniscule is used to describe the original model pool names, whereas the concept applied here with more generalized pools to determine the humus efficiency are denoted with capital letters. AOM is defined as the substrate before incubation. FOM is denoted as the part of the substrate that still has the properties of AOM before microbial turnover and humus (HUM) that integrates all SOM pools after microbial turnover for a given time t.

$$\mathbf{HUM}\ t = \sum_{i=1..p} \mathbf{SOM}_i\ t \quad (1)$$

Model	Considered F pool of the model	AOM assessment	Fitted parameters	Min/max
C-TOOL	*fom*	Distribution Transformation	f_{hum} k_{fom}	≈0/1 0/10
CCB	*aom*	Transformation	k10 k12	0/0.5 0/2.5
CENTURY	*m*	Distribution Transformation	f_{lig} k_{str}	0/999 0.24/4.8
ICBM	*y*	Transformation	k10 k12	0/0.3 0/3
RothC	*rpm*	Distribution	f_{hum}	0.252/0.98
Yasso	*w*	Distribution	p_w	0/1

TABLE 2 Overview of the models used and the fitted parameters; min/max describes the set thresholds for the parameters, the min value is numerically never truly = 0; FOM denotes fresh organic matter and AOM denotes added organic matter, for a description of the pools or the parameters see Appendix S1 or the publications

If a model does not include an explicit FOM pool, the most dynamic pool with the lowest turnover time is considered as FOM, this accounts for CENTURY, RothC and Yasso. Some substrates, especially manure, are subject to microbial turnover even before the material is added to the soil. Therefore, in several model concepts, a part of AOM is directly transferred to HUM without passing a FOM pool.

The models used to express the quality of AOM as matter transformation into SOM with two general approaches:

- Distribution: initial partitioning from AOM to one or two model pools
- Transformation: efficiency for production of HUM

In the first case, at the beginning of the incubation experiment, a fraction or the total AOM is allocated between a FOM and SOM pool according to the structure of the specific model. In the second case, SOM pools are built up continuously.

During the simulation of an incubation experiment each model predicts the evolution of AOM into a designated FOM pool, one or more SOM pools and of course the amount of mineralized carbon (CO_2).

The data used represent the difference of CO_2 evolution between a treatment (soil + substrate) and the control vessel (soil only). In order to represent the net-mineralization, priming effects were neglected and each model was initialized with empty SOM pools that will fill up according to the individual model procedures. The amount of C which is lost from the added substrate (AOM in Equation 2) during the incubation, is transferred into several SOM pools or is released as CO_2:

$$\mathbf{AOM - FOM}\,\boldsymbol{t} = \sum_{i=1..p} \mathbf{SOM}_i\,\boldsymbol{t} + \boldsymbol{C}_{\mathbf{C02}}\,\boldsymbol{t} \tag{2}$$

As mentioned above, the efficiency to build up HUM has to be quantified from two components. The first one is the quota (q) of added C that is immediately allocated to the SOM pools at the beginning of the incubation.

$$q = \frac{\mathrm{HUM}\,(0)}{\mathrm{AOM}} \tag{3}$$

This reduces the amount of FOM for further decomposition to (1-q)*AOM.

The second component represents the dynamic transformation from FOM to HUM and is calculated as the relation between the rate of HUM production (dHUM) and the rate of FOM decomposition (dFOM). The sum of both components results in the humus building efficiency parameter E_{HUM}:

$$E_{\mathrm{HUM}} = q + max\left(\frac{\mathrm{dHUM}\,(t)}{\mathrm{dFOM}\,(t)}\right) * (1 - q) \tag{4}$$

For sufficiently small time steps the changes of HUM and FOM can be calculated with a negligible loss of SOM-C to CO_2:

$$d\mathrm{HUM}\,(t) = (\mathrm{HUM}\,(t) - \mathrm{HUM}\,(t - \Delta t))\ ; \quad t > 0 \tag{5}$$

$$d\mathrm{FOM} = (\mathrm{FOM}\,(t) - \mathrm{FOM}\,(t - \Delta t))\ ; \quad t > 0 \tag{6}$$

For all 72 substrates, the variable parameters of each model were fitted to find the best agreement between observed and predicted CO_2 production. The obtained model parameters were then used to model the CO_2 mineralization for small time steps of 10^{-4} d. In this way, E_{HUM} was calculated with each model for every substrate.

2.4 | Model averaging and assessment

During the optimization for some models, the optimized parameters adjoin their thresholds for certain substrates. In this case, the models often insufficiently fit the data. To minimize the effect of parameters at their thresholds and to compensate individual weaknesses, the models were aggregated into an ensemble. Therefore, a model averaging method in analogy to the proposal of Bates and Granger (1969) as described by Diks and Vrugt (2010) was applied, but instead of the variance, the mean squared error (*MSE*) was used to calculate the weights ($w_{i,s}$). With this method, models with the highest squared prediction error get the lowest weight with respect to the disproportional sensitivity to larger errors. For each combination of model i and substrate s the weight $w_{i,s}$ was calculated:

$$w_{i,s} = \frac{1/\mathrm{MSE}_{i,s}}{\sum_{j=1}^{n} 1/\mathrm{MSE}_{j,s}} \quad (7)$$

where *n* denotes the number of models within the ensemble.

As performance measure of the incubation fit the root mean square error (RMSE) was used since it has the same unit. The individual model results were compared with the ensemble prediction in order to evaluate the ensemble performance to improve the prediction of the incubation data. Further on, the computed weights were applied to aggregate the model specific E_{HUM} values for each substrate to obtain $\overline{E_{HUM}}$ as a weighted ensemble value.

A model with a high weight for a substrate could provide the same information as others with lower weights, which is why another 6 ensembles were calculated, omitting one of the models each time. Afterwards the difference between $\overline{E_{HUM}}$ over all models and $\overline{E_{HUM}}$ with exclusion of the model i was calculated to get the influence of each model on the ensemble $\overline{E_{HUM}}$.

$$\text{influence}_{\ \text{model}=i} = \left| \overline{E_{\text{hum}(1,\ldots,6)}} - \overline{E_{\text{hum}\ (1,\ldots,i-1,i+1,\ldots,6)}} \right| \quad (8)$$

3 | RESULTS

3.1 | Model fitting and model performance

It is possible to adapt the concepts of the used models to the incubation data and to parameterize the underlying substrates. Nevertheless, the quality of the model fit varies between models and substrates due to the characteristics of the incubation process. However, each model performs best in terms of minimizing the RMSE for at least one substrate. Some models have explicit incubation trends (e.g. rapidly decomposable or slowly decomposable substrates) where they perform

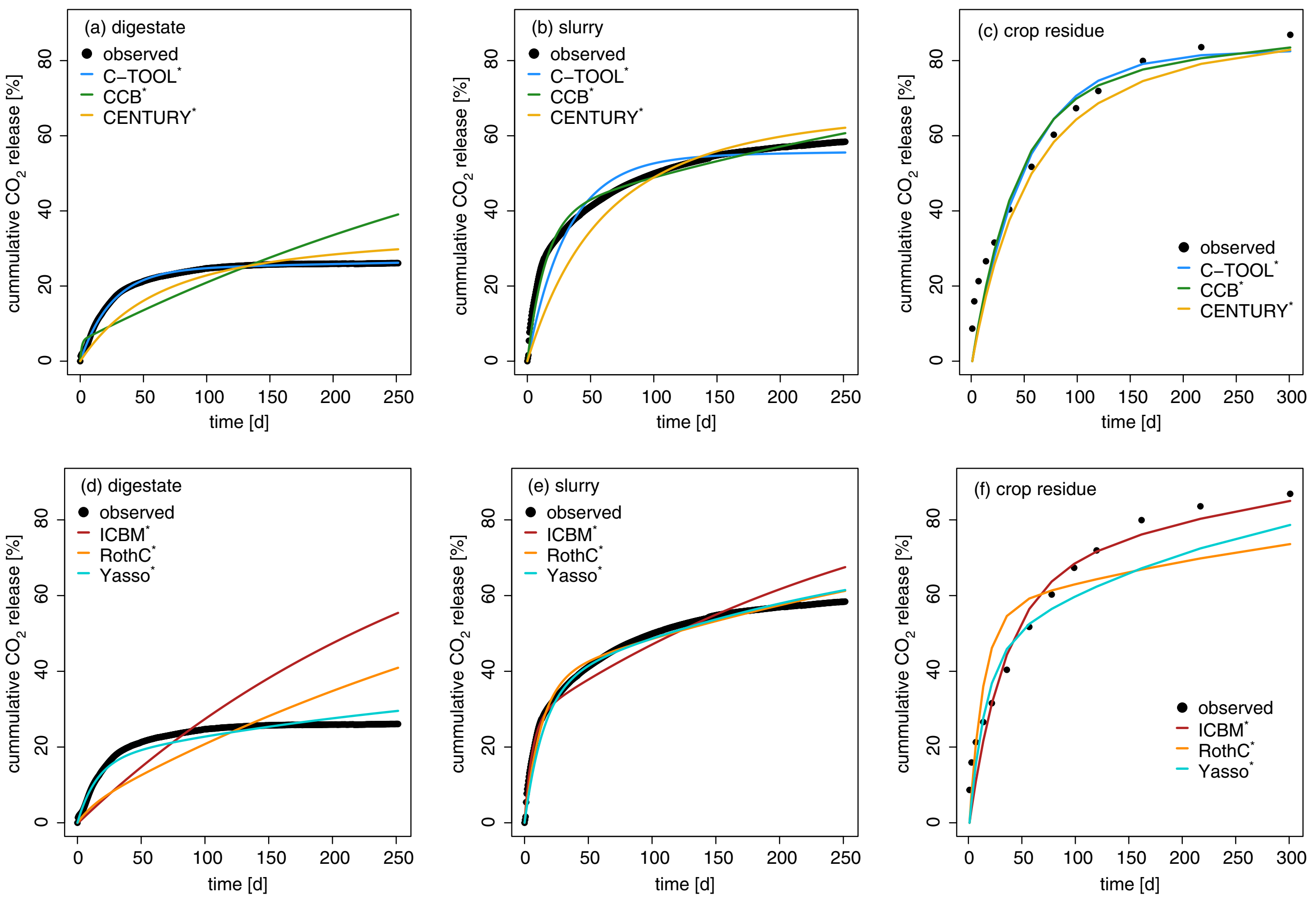

FIGURE 1 Example of model results for the cumulative CO_2 release with three different substrates, (a,d) digestate (ID: 18), (b,e) slurry (ID: 22), (c, f) crop residue of sorghum (ID: 54) , see Appendix S1 for ID

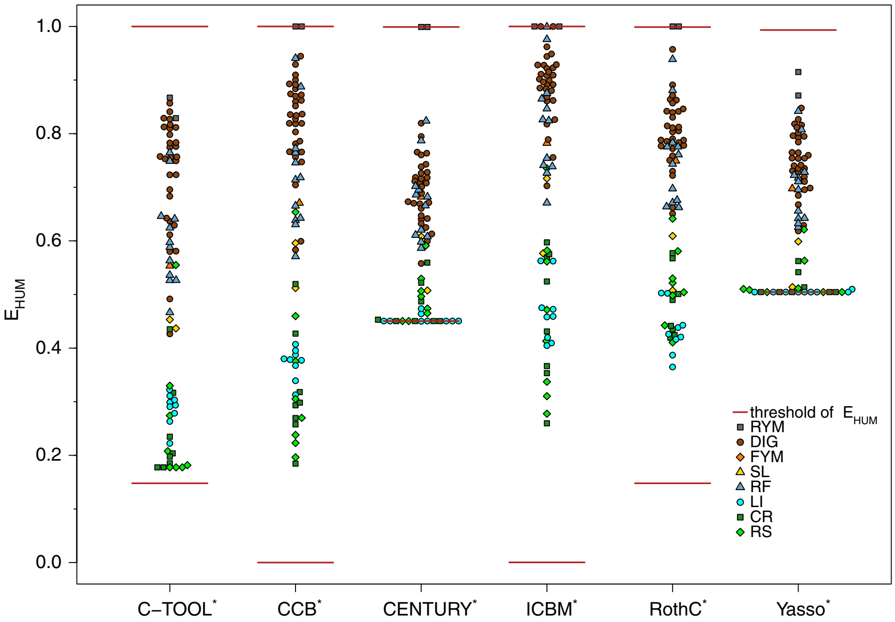

FIGURE 2 E_{HUM} values for each model and substrate, line: model specific thresholds of E_{HUM} considering texture dependency, RYM: rotten yard manure, DIG: digestate, FYM: fresh yard manure, SL: cattle slurry, RF: fine roots, LI: litter, CR: crop residue, RS: coarse roots, classes don´t consider material composition for example different plant residues

superior to other models. Figure 1 shows three substrates, a digestate with a slow carbon mineralization, slurry with an intermediate mineralization and crop residue of sorghum with a fast mineralization. It is shown that some models have difficulties with stable substrates (Figure 1a,d) while others struggle with easy decomposable substrates (Figure 1c,f) and most models fit substrates with an intermediate mineralization sufficiently.

C-TOOL* fits the incubation data best when the measured mineralization rate reaches a constant state (Figure 1a). For substrates where the mineralization still rises continuously at the end of the incubation period, C-TOOL* fits the incubation data worse compared to other models (Figure 1b). Due to the combination of a distribution and transformation approach during the AOM turnover, C-TOOL* is able to model highly decomposed organic matter like digestates and rotten yard manure.

The model performance of CCB* is also relatively robust but, the model reaches its limits with highly decomposed AOM, since its AOM turnover is solely based on an efficiency approach (Figure 1a).

CENTURY* is challenged by strong mineralization rates at the beginning of the incubation and the approximation to the steady state, but works better on easy decomposable substrates (Figure 1a,b).

ICBM*, on the other hand performs best when C mineralization is rising fast at the beginning of the incubation (Figure 1f), but also lacks the ability to simulate a stagnating C mineralization due to the efficiency approach of AOM turnover (Figure 1d).

Initially, RothC* predicted the mineralization dynamics satisfactorily, but with advanced incubation time the predictions become almost linear (Figure 1e,f). Therefore substrates that show a pronounced saturation are not well represented.

Yasso* performs well compared to other models, when applied to highly decomposed materials (Figure 1d) like rotten farm yard manure or digestates. Due to the pre-distribution of FOM into the defined HUM pool, however, Yasso* lacks the ability to simulate easily decomposable AOM (Figure 1f).

During optimization each model had reached its parameter limitation at least once. In this case, the model fits are deficient and the incubation data is not well represented (Figure 1). Relying solely on one model may, therefore, lead to an under- or overestimation of the incubation data.

3.2 | Model specific diversity of E_{HUM}

Figure 2 shows the distribution of calculated E_{HUM} values for each model. In general, it shows very well how all model concepts designate high E_{HUM} values to already pre-composed materials such as manure and digestates and low E_{HUM} values to plant residues. But the model concepts differ in their overall conception of E_{HUM}. Models such as C-TOOL*, CENTURY*, RothC*, and Yasso* have a lower threshold of E_{HUM} unequal to 0 (Figure 2), where the minimum E_{HUM} value is equivalent to the parameter describing the flux of the chosen FOM pool to CO_2. This parameter is texture dependent for C-TOOL* and RothC*, which leads to a slight shift in the possible flux from FOM to CO_2, to be observed by C-TOOL* (Figure 2), where the E_{HUM} values align at a lower threshold, although with less clay content the values would be even lower (red line).

Model concepts without an AOM distribution, like CCB*, CENTURY* and ICBM*, lack the ability to fit highly decomposed substrates adequately with fitted parameters adjoining the parameter thresholds. This leads to an insufficient representation of the incubation data and results in E_{HUM} values of 1, even though more resilient substrates theoretically exist (like peat). In this case, the max E_{HUM} is limited by the model structure.

Furthermore, the aggregation of all SOM pools into a HUM pool causes for RothC* information loss for high E_{HUM} values. The optimized parameter f_{hum} shifts C from *rpm* to *hum* for values higher than 0.772, but since both pools are considered as HUM pool E_{HUM} will reach its maximum at the parameter value $f_{hum} = 0.772$ (Dechow et al., 2019).

Only C-TOOL* and Yasso* have not adjoined the upper threshold of E_{HUM} whereas CCB*, ICBM* and RothC* have not adjoined their lower E_{HUM} threshold within the analysed substrates. This demonstrates that each single model has its strength and weaknesses with regard to the substrates analysed due to model structure and the intended purpose of the model. Nevertheless from each model concept follows a similar E_{HUM} value for a given substrate, with rotten yard manure and digestates as substrates with the highest humus efficiency and plant materials (except fine roots) with the lowest humus efficiency (Figure 2).

3.3 | Ensemble performance

The averaging method applied leads to an overall robust performance of the ensemble in displaying the incubation data. Compared with the single models, the ensemble scores 32 times the lowest RMSE for the underlying substrates and never has the highest RMSE compared to a single model. The average RMSE over all 72 substrates was calculated, which is for C-TOOL* = 2.39, CCB* = 2.48, CENTURY* = 4.59, ICBM* = 4.30, RothC* = 4.11, Yasso* = 3.62, ensemble = 1.92.

Furthermore, the calculated weights were applied to the E_{HUM} value each model supplies for a substrate. The average weight of the ensemble $\overline{E_{HUM}}$ value, is compiled out of the analysed 72 substrates is, C-TOOL* = 28.4%, CCB* = 28.7%, CENTURY* = 7.9%, ICBM* = 13.3%, RothC* = 10.3%, Yasso* = 11.4%, which demonstrates that every model used contributes to the ensemble $\overline{E_{HUM}}$ values.

Since the E_{HUM} value varies between models, the influence of a model on the ensemble $\overline{E_{HUM}}$ for each substrate was calculated. As a result one model was omitted and a new ensemble $\overline{E_{HUM}}$ value was calculated, without the regarding model. The effect that the omission of one model has on the ensemble $\overline{E_{HUM}}$ calculation is shown in Figure 3. The most influential model for the ensemble $\overline{E_{HUM}}$ value compilation is C-TOOL*, followed by CCB*, ICBM*, CENTURY*, Yasso* and RothC*. Rather than the fitting performance, as described before, the information that one model provides to the $\overline{E_{HUM}}$ value and the information that one model concept provides which is redundant and can be compensated by other models is shown in Figure 3. The most influential model for the ensemble $\overline{E_{HUM}}$ value compilation is C-TOOL*, followed by CCB*, ICBM*, CENTURY*, Yasso* and RothC*. The average divergence of $\overline{E_{HUM}}$ for one substrate, occurring when one model is left out of the ensemble $\overline{E_{HUM}}$ calculation, is for C-TOOL* = 2.91%, CCB* = 1.42%, CENTURY* = 0.749%, ICBM* = 1.131%, RothC* = 0.429%, Yasso* = 0.659%, indicating that this is a rather robust concept.

3.4 | Humus efficiency of organic substrates

For a practical application, the ensemble $\overline{E_{HUM}}$ concept with model averaging was applied to substrate classes. The classes were aggregated by their material origin, while differences

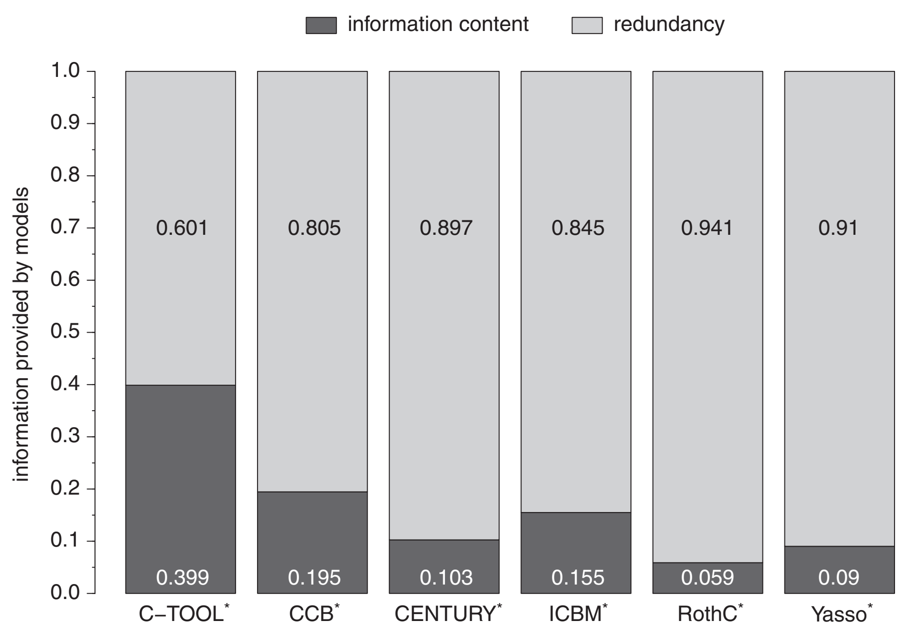

FIGURE 3 The relative change of $\overline{E_{HUM}}$ for 72 substrates without the considered model, high redundancy means that the E_{HUM} values of a model is similar to the E_{HUM} values of other models and model fits to the incubation data are worse than other models

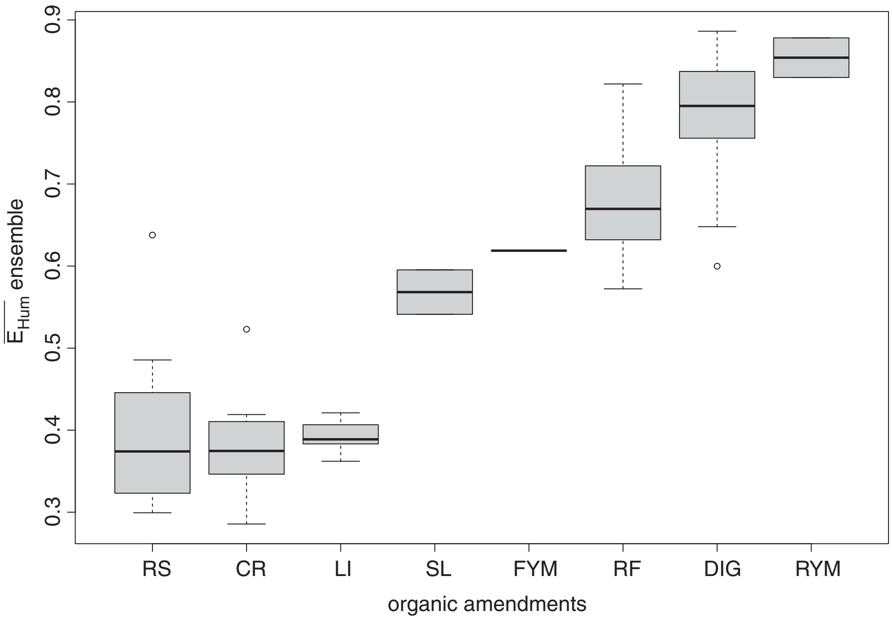

FIGURE 4 Ensemble $\overline{E_{HUM}}$ values for substrates by class; RS: root stock (coarse root), $N = 8$; CR: crop residues, $N = 8$; LI: litter, $N = 9$; SL: cattle slurry, $N = 2$; FYM: fresh yard manure, $N = 1$; RF fine roots, $N = 12$; DIG: digestates, $N = 30$; RYM: rotten yard manure; $N = 2$, N: count of group members

in species for plant material and composition among digestates were not taken into account. Figure 4 demonstrates the weighted $\overline{E_{HUM}}$ values for different substrate types. Rotten yard manure and digestates show the highest $\overline{E_{HUM}}$ values followed by fine roots. Untreated animal faeces like slurry and yard manure have medium $\overline{E_{HUM}}$ values whereas plant materials like coarse roots, crop residues and litter show the lowest $\overline{E_{HUM}}$ values and cannot be distinguished statistically.

4 | DISCUSSION

4.1 | Carbon model adaptation to incubation data

Despite being developed for field application, the chosen model concepts were applied successfully to the incubation data while preserving the core concept of the models under the application of the same algorithm and data structure. Nevertheless, some models can be better adapted to the incubation data than others due to model complexity and the model structure. This especially accounts for the handling of AOM, where a higher flexibility was required to deal with a bigger variety of substrates as by some models intended.

Commonly in incubation studies the incubation data gets analysed with some sort of regression where the results are presented as mineralization in per cent or as amount of mineralized C over the incubation time (Sänger et al., 2014). Those regressions are not transferrable to other substrates and serve only a descriptive purpose. Rather than developing models to predict the incubation data, the here used models are already successfully used in field studies and most parameters are derived from field experiments. This gives the models a high credibility, as well as involving more sophisticated pool interactions and SOC processes. Furthermore, a maximum of two parameters describing the decay were optimized thereby minor equifinality is expected compared to other approaches (Tang & Riley, 2020).

As a further advantage of using SOM models for incubation, the results from the model fitting can be used to transfer observations from incubation experiments to field scale. Thus the calibration results for models on field scale can be validated and possibly improved.

4.2 | Model and ensemble performance

The method applied for model averaging considers the results of model fits to the cumulative CO_2 released, giving higher weights to models with a lower *MSE*. The *MSE* is sensitive to larger errors due to the squaring, therefore the influence of models with high prediction errors gets minimized.

The model averaging leads to a better prediction accuracy and the ensemble has the lowest RMSE for 32 substrates, 27 times the second lowest, 11 time the third lowest and two times the fourth lowest RMSE. Also the average RMSE over all 72 treatments demonstrates that the ensemble is more robust in predicting incubation results than a single model concept and that some model concepts don't vary as much as others in their prediction accuracy.

In hydrological and meteorological forecasting, model ensembles are a common tool, to reduce model uncertainties (Li et al., 2017) whereat this technique is not fully established in SOC modelling yet. A recent study by Riggers et al. (2019) successfully used the same model compilation like this study in an ensemble approach on a field scale. In this study the ensemble members were not weighted, rather the combination of models and different initialization processes were reduced to minimize the prediction error and to find a robust ensemble. The application of several models as an ensemble can help to balance the prediction errors of the individual models which result from the specific structure of their embedded processes, as well as from their individual parametrization or scope and the scale the models were developed for (Martre et al., 2015; Tebaldi & Knutti, 2007).

4.3 | Concept of E_{HUM}

In this study a method addressing the humus building efficiency of AOM is presented, that evaluates the CO_2 mineralization of AOM and then allows to draw conclusions about the substrate quality. Several carbon models were fitted to the cumulative mineralization and the resulting pool dynamics of each model was used to evaluate the substrate quality. The obtained E_{HUM} value describes the efficiency of a substrate to generate new humus with a time independent metric that considers the incubation temperature during calculation, which allows the comparison of incubation experiments with different time spans and different incubation temperatures. Nevertheless a certain incubation period is required for the models to predict the incubation trend. There are several other methods which describe the humification processes like the E4/E6 ratio which is determined by the optical density of humic and fulvic acids, the Humification Index (HI) and other methods which mostly rely on chemical and physical properties of the substrates (Klavins et al., 2008). Additionally there exist also field experiments in which the application of organic substrates is compared to a control plot with no application of organic substrates (Kätterer et al., 2011). The focus of the presented approach is based on the incubation curve characteristics and the behaviour of SOM pools. In contrast to chemical analysis, E_{HUM} also considers microbial turnover and incubation experiments are less costly and time consuming than field experiments. Some of the applied ensemble members even employ mineralization processes with regard to soil properties. Therefore this new concept could be a valuable addition to existing methods.

Each model employed represents a slightly different understanding of the soil processes involved in soil carbon dynamics, which is why their model structure and interpretation of SOM generation varies. Aggregating several SOM pools into a single, conceptional HUM pool influences the E_{HUM} value for each model differently. Thus, each model comes to slightly different predictions of E_{HUM}. Some models show a restricted range for E_{HUM} whereas others are theoretically able to display the complete expected scale of E_{HUM} from 0 to 1. This depends on the model structure and the selected approach of only one FOM pool, which was defined in this study to be the one with the fastest C turnover.

Based on the robust ensemble performance, the calculated weights are not applied in the first place to improve the overall prediction, but rather to evaluate the substrate quality in terms of the humus building efficiency using a model ensemble. The influence of model concepts, which are not suitable for certain substrates, is thereby minimized in the calculation of the ensemble $\overline{E_{HUM}}$ value. The averaging therefore leads to a more trustworthy prediction of the C dynamics from the incubation data and also prevents over-and underestimation of the E_{HUM} value when model parameters reach their limits during optimization.

Computing power is hardly a limiting factor and it is possible to calculate complex models within a very short time. Nevertheless it is important to know how much information a model contributes to the results of an ensemble and if a model can be left out of the ensemble formation. It was shown (Figure 3) that every model contributes to the ensemble $\overline{E_{HUM}}$ value. Within the combination of these six models, C-TOOL* has the biggest influence on the ensemble $\overline{E_{HUM}}$, followed by CCB*, ICBM*, CENTURY*, Yasso* and RothC* which can be compensated for the most part by other models. Whether a model can be omitted from the ensemble calculation is within the discretion of the user.

4.4 | Application of the ensemble $\overline{E_{HUM}}$

The 72 substrates analysed were classified by their origin in order to evaluate the E_{HUM} value for practical applications. The values for each substrate class are in an expected order where more mature substrates have a higher E_{HUM} and therefore a stronger resilience to microbial depletion (Bernal et al., 1998). Ajwa and Tabatabai (1994) found similar results for the mineralization of organic material in soil. Where C of plant material had a half-life between 39 and 54 days, animal manure had a half-life which ranged from 37 to 169 days, and for sewage sludge the half-life was 39 to 330 days.

E_{HUM} characterizes the humus efficiency of a substrate in one single value and is therefore easy to interpret. All mature substrates and animal faeces show a higher humus building efficiency compared to plant materials except fine roots (Figure 4). Digestates or animal faeces on the other hand, undergo microbial turnover either in the digestive system and/or in a bioreactor and therefore, the substrate contains more fungal and bacterial necromass as well as decomposition products, which are considered to be a main component of stable SOC (Kallenbach et al., 2015; Liang et al., 2017). An explanation for the higher humus building efficiency of fine roots compared to the other plant materials analysed, can be found in Rasse et al. (2005), who pointed out, that the high resilience of roots against carbon turnover is due to the physicochemical protection caused by the steady contact of roots to soil particles.

Alongside the chemical analysis of organic amendments like the C/N ratio, Lignin content etc. the E_{HUM} value delivers an easy to interpret assessment to evaluate the quality of the organic amendments based on their behaviour in soils under controlled conditions.

For field management that aims at retaining or increasing carbon stock the ensemble $\overline{E_{HUM}}$ value can deliver valuable information about management options concerning the choice of organic amendments. Knowing the humus efficiency, the quantity of newly applied organic amendments can be adjusted to cover the C demands and to conform to possible restrictions in terms of carbon dioxide mitigation goals. An organic amendment with lower E_{HUM} value could be replaced by one with a higher E_{HUM} value, which needs a lower application rate to reach an equal soil carbon stock. Nevertheless such a decision is complex and other nutrients like nitrogen and phosphorus have to be considered as well.

5 | CONCLUSIONS

This study demonstrates that carbon models developed for field scale with different target environments and time scales are a suitable tool to predict carbon incubation data, derived from laboratory experiments. The pool structure of those models can be used to derive information about the efficiency of a substrate to build humus and be displayed in a single value which is easily comparable, which is not possible in the same manner with statistical models. Nevertheless some adaptations to the model concepts had to be made to cover the wide scope of organic amendments. Furthermore, substrate and model specific parameters derived from incubation experiments can be used to improve modelling of SOC turnover at field scale with regard to the modifications of the model concepts made in this approach. This could be a cost and time efficient alternative to long term field experiments and could give insights into the dynamics of organic amendments which are not fully analysed yet. The presented approach could help to close the gap between laboratory experiments under controlled conditions and field applications where much more influential factors need to be considered to evaluate the substrate quality. Furthermore, every model incorporates different mechanisms, with a different scope of application. This diversity cannot be covered by a single model and therefore, ensemble approaches can be a useful tool for future SOC modelling challenges.

ORCID

Anton A. Gasser https://orcid.org/0000-0002-4188-1217

REFERENCES

Ajwa, H. & Tabatabai, M. (1994). Decomposition of different organic materials in soils. *Biology and Fertility of Soils*, *18*(3), 175–182. https://doi.org/10.1007/BF00647664

Andrén, O. & Kätterer, T. (1997). ICBM: The introductory carbon balance model for exploration of soil carbon balances. *Ecological Applications*, *7*(4), 1226–1236.

Bates, J. M. & Granger, C. W. (1969). The combination of forecasts. *Journal of the Operational Research Society*, *20*(4), 451–468. https://doi.org/10.1057/jors.1969.103

Bernal, M., Sanchez-Monedero, M., Paredes, C. & Roig, A. (1998). Carbon mineralization from organic wastes at different composting stages during their incubation with soil. *Agriculture, Ecosystems & Environment*, *69*(3), 175–189. https://doi.org/10.1016/S0167-8809(98)00106-6

Bradford, M. A., Carey, C. J., Atwood, L., Bossio, D., Fenichel, E. P., Gennet, S., Fargione, J., Fisher, J. R. B., Fuller, E., Kane, D. A., Lehmann, J., Oldfield, E. E., Ordway, E. M., Rudek, J., Sanderman, J. & Wood, S. A. (2019). Soil carbon science for policy and practice. *Nature Sustainability*, *2*(12), 1070–1072. https://doi.org/10.1038/s41893-019-0431-y

Coleman, K. & Jenkinson, D. (1999). RothC-A model for the turnover of carbon in soil. *Model Description and Users Guide (Harpenden)*. Harpenden, UK: IACR-Rothamsted.

Cotrufo, M. F., Wallenstein, M. D., Boot, C. M., Denef, K. & Paul, E. (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

Dechow, R., Franko, U., Kätterer, T. & Kolbe, H. (2019). Evaluation of the RothC model as a prognostic tool for the prediction of SOC trends in response to management practices on arable land. *Geoderma*, *337*, 463–478. https://doi.org/10.1016/j.geoderma.2018.10.001

Dignac, M.-F., Derrien, D., Barré, P., Barot, S., Cécillon, L., Chenu, C., Chevallier, T., Freschet, G. T., Garnier, P., Guenet, B., Hedde, M., Klumpp, K., Lashermes, G., Maron, P.-A., Nunan, N., Roumet, C. & Basile-Doelsch, I. (2017). Increasing soil carbon storage: Mechanisms, effects of agricultural practices and proxies. *A Review. Agronomy for Sustainable Development*, *37*(2), 14. https://doi.org/10.1007/s13593-017-0421-2

Diks, C. G. H. & Vrugt, J. A. (2010). Comparison of point forecast accuracy of model averaging methods in hydrologic applications. *Stochastic Environmental Research and Risk Assessment*, *24*(6), 809–820. https://doi.org/10.1007/s00477-010-0378-z

Elzhov, T. V., Mullen, K. M., Spiess, A.-N. & Bolker, B. (2016). minpack.lm: R Interface to the Levenberg-Marquardt Nonlinear Least-Squares Algorithm Found in MINPACK, Plus Support for Bounds. Retrieved from https://CRAN.R-project.org/package=minpack.lm

Franko, U., Kolbe, H., Thiel, E. & Ließ, E. (2011). Multi-site validation of a soil organic matter model for arable fields based on generally available input data. *Geoderma*, *166*(1), 119–134. https://doi.org/10.1016/j.geoderma.2011.07.019

Hagedorn, R., Doblas-Reyes, F. J. & Palmer, T. (2005). The rationale behind the success of multi-model ensembles in seasonal forecasting — I. Basic concept. *Tellus A: Dynamic Meteorology and Oceanography*, *57*(3), 219–233. https://doi.org/10.3402/tellusa.v57i3.14657

Harden, J. W., Hugelius, G., Ahlström, A., Blankinship, J. C., Bond-Lamberty, B., Lawrence, C. R., Loisel, J., Malhotra, A., Jackson, R. B., Ogle, S., Phillips, C., Ryals, R., Todd-Brown, K., Vargas, R., Vergara, S. E., Cotrufo, M. F., Keiluweit, M., Heckman, K. A., Crow, S. E., … Nave, L. E. (2018). Networking our science to characterize the state, vulnerabilities, and management opportunities of soil organic matter. *Global Change Biology*, *24*(2), e705–e718. https://doi.org/10.1111/gcb.13896

Jha, P., De, A., Lakaria, B. L., Biswas, A., Singh, M., Reddy, K. & Rao, A. (2012). Soil carbon pools, mineralization and fluxes associated with land use change in Vertisols of Central India. *National Academy Science Letters*, *35*(6), 475–483. https://doi.org/10.1007/s40009-012-0082-2

Kallenbach, C. M., Grandy, A. S., Frey, S. D. & Diefendorf, A. F. (2015). Microbial physiology and necromass regulate agricultural soil carbon accumulation. *Soil Biology and Biochemistry*, *91*, 279–290. https://doi.org/10.1016/j.soilbio.2015.09.005

Kätterer, T., Bolinder, M. A., Andrén, O., Kirchmann, H. & Menichetti, L. (2011). Roots contribute more to refractory soil organic matter than above-ground crop residues, as revealed by a long-term field experiment. *Agriculture, Ecosystems & Environment*, *141*(1–2), 184–192. https://doi.org/10.1016/j.agee.2011.02.029

Klavins, M., Sire, J., Purmalis, O. & Melecis, V. (2008). Approaches to estimating humification indicators for peat. *Mires & Peat*, *3*, 8.

Larney, F. J. & Angers, D. A. (2012). The role of organic amendments in soil reclamation: A review. *Canadian Journal of Soil Science*, *92*(1), 19–38. https://doi.org/10.4141/cjss2010-064

Lashermes, G., Nicolardot, B., Parnaudeau, V., Thuriès, L., Chaussod, R., Guillotin, M. L., Linères, M., Mary, B., Metzger, L., Morvan, T., Tricaud, A., Villette, C. & Houot, S. (2009). Indicator of potential residual carbon in soils after exogenous organic matter application. *European Journal of Soil Science*, *60*(2), 297–310. https://doi.org/10.1111/j.1365-2389.2008.01110.x

Li, W., Duan, Q., Miao, C., Ye, A., Gong, W. & Di, Z. (2017). A review on statistical postprocessing methods for hydrometeorological ensemble forecasting. *Wiley Interdisciplinary Reviews. Water*, *4*(6), e1246. https://doi.org/10.1002/wat2.1246

Liang, C., Schimel, J. P. & Jastrow, J. D. (2017). The importance of anabolism in microbial control over soil carbon storage. *Nature Microbiology*, *2*, 17105. https://doi.org/10.1038/nmicrobiol.2017.105

Martre, P., Wallach, D., Asseng, S., Ewert, F., Jones, J. W., Rötter, R. P., Boote, K. J., Ruane, A. C., Thorburn, P. J., Cammarano, D., Hatfield, J. L., Rosenzweig, C., Aggarwal, P. K., Angulo, C., Basso, B., Bertuzzi, P., Biernath, C., Brisson, N., Challinor, A. J., … Wolf, J. (2015). Multimodel ensembles of wheat growth: Many models are better than one. *Global Change Biology*, *21*(2), 911–925. https://doi.org/10.1111/gcb.12768

Minasny, B., Malone, B. P., McBratney, A. B., Angers, D. A., Arrouays, D., Chambers, A., Chaplot, V., Chen, Z.-S., Cheng, K., Das, B. S., Field, D. J., Gimona, A., Hedley, C. B., Hong, S. Y., Mandal, B., Marchant, B. P., Martin, M., McConkey, B. G., Mulder, V. L., … Winowiecki, L. (2017). Soil carbon 4 per mille. *Geoderma*, *292*, 59–86. https://doi.org/10.1016/j.geoderma.2017.01.002

Morvan, T., Nicolardot, B. & Péan, L. (2005). Biochemical composition and kinetics of C and N mineralization of animal wastes: A typological approach. *Biology and Fertility of Soils*, *42*(6), 513–522. https://doi.org/10.1007/s00374-005-0045-6

Parton, W. J., Ojima, D. S., Cole, C. V. & Schimel, D. S. (1994). A general model for soil organic matter dynamics: sensitivity to litter chemistry, texture and management. Quantitative modeling of soil forming processes(quantitativemod), 147–167. doi:10.2136/sssaspecpub39.c9.

Parton, W., Schimel, D. S., Cole, C. & Ojima, D. (1987). Analysis of factors controlling soil organic matter levels in Great Plains Grasslands 1. *Soil Science Society of America Journal*, *51*(5), 1173–1179. https://doi.org/10.2136/sssaj1987.03615995005100050015x

R Core Team (2019). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Rasse, D. P., Rumpel, C. & Dignac, M.-F. (2005). Is soil carbon mostly root carbon? Mechanisms for a specific stabilisation. *Plant and Soil*, *269*(1–2), 341–356. https://doi.org/10.1007/s11104-004-0907-y

Riggers, C., Poeplau, C., Don, A., Bamminger, C., Höper, H. & Dechow, R. (2019). Multi-model ensemble improved the prediction of trends in soil organic carbon stocks in German croplands. *Geoderma*, *345*, 17–30. https://doi.org/10.1016/j.geoderma.2019.03.014

Sänger, A., Geisseler, D. & Ludwig, B. (2014). C and N dynamics of a range of biogas slurries as a function of application rate and soil texture: A laboratory experiment. *Archives of Agronomy and Soil Science*, *60*(12), 1779–1794. https://doi.org/10.1080/03650340.2014.907491

Sierra, C. A., Müller, M. & Trumbore, S. E. (2012). Models of soil organic matter decomposition: The SoilR package, version 1.0. *Geoscientific Model Development*, *5*(4), 1045–1060. https://doi.org/10.5194/gmd-5-1045-2012

Soetaert, K. E., Petzoldt, T. & Setzer, R. W. (2010). Solving differential equations in R: Package deSolve. *Journal of Statistical Software*, *33*, 10.

Stockmann, U., Adams, M. A., Crawford, J. W., Field, D. J., Henakaarchchi, N., Jenkins, M., Minasny, B., McBratney, A. B., Courcelles, V. D. R. D., Singh, K., Wheeler, I., Abbott, L., Angers, D. A., Baldock, J., Bird, M., Brookes, P. C., Chenu, C., Jastrow, J. D., Lal, R., … Zimmermann, M. (2013). The knowns, known unknowns and unknowns of sequestration of soil organic carbon. *Agriculture, Ecosystems & Environment*, *164*, 80–99. https://doi.org/10.1016/j.agee.2012.10.001

Sulman, B. N., Moore, J. A. M., Abramoff, R., Averill, C., Kivlin, S., Georgiou, K., Sridhar, B., Hartman, M. D., Wang, G., Wieder, W. R., Bradford, M. A., Luo, Y., Mayes, M. A., Morrison, E., Riley, W. J., Salazar, A., Schimel, J. P., Tang, J. & Classen, A. T. (2018). Multiple models and experiments underscore large uncertainty in soil carbon dynamics. *Biogeochemistry*, *141*(2), 109–123. https://doi.org/10.1007/s10533-018-0509-z

Taghizadeh-Toosi, A., Christensen, B. T., Hutchings, N. J., Vejlin, J., Kätterer, T., Glendining, M. & Olesen, J. E. (2014). C-TOOL: A simple model for simulating whole-profile carbon storage in temperate agricultural soils. *Ecological Modelling*, *292*, 11–25. https://doi.org/10.1016/j.ecolmodel.2014.08.016

Tang, J. & Riley, W. J. (2020). Linear two-pool models are insufficient to infer soil organic matter decomposition temperature sensitivity from incubations. *Biogeochemistry*, *149*(3), 251–261. https://doi.org/10.1007/s10533-020-00678-3

Tebaldi, C. & Knutti, R. (2007). The use of the multi-model ensemble in probabilistic climate projections. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *365*(1857), 2053–2075.

Tuomi, M., Rasinmäki, J., Repo, A., Vanhala, P. & Liski, J. (2011). Soil carbon model Yasso07 graphical user interface. *Environmental Modelling & Software*, *26*(11), 1358–1362. https://doi.org/10.1016/j.envsoft.2011.05.009

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Gasser AA, Diel J, Nielsen K, Mewes P, Engels C, Franko U. A model ensemble approach to determine the humus building efficiency of organic amendments in incubation experiments. *Soil Use Manage*. 2021;00:1–12. https://doi.org/10.1111/sum.12699