Dear Dr Weintraub,   27 April 2018

Thank you for your handling of this manuscript. We hope that these revisions address the reviewers concerns with the manuscript.

Some of the changes to the manuscript include. The exclusion of coarse woody debris from the decomposing carbon pool changed some of the exact global totals in a very minor way which did not affect the overall results of the study. Clarification on the boundaries and selection criteria for the Q10 parameter in the methods section. Clarification on the effect of soil inputs vs outputs (no effect in all but one model) as the driving flux to calculate Q10 in the Earth system model study. And finally, a brief expansion of the discussion around alternative mechanisms that need to be considered when interoperating the results of this study.

Below you will find a point by point response to the reviews as well as a mark-up draft of the changes made to the manuscript.

We thank yourself and the reviewers for their thoughtful comments which greatly improved this manuscript and hope that this satisfies the concerns raised.

Sincerely,
Kathe Todd-Brown

Reply to reviewers

This is an interesting and well written paper addressing a subject area of clear importance and relevance to the readership of Biogeosciences.

The authors clearly articulate the issue of uncertainty in future land-carbon storage, how this responds to climate change and the vital role played of soil organic carbon and its sensitivity to temperature. The results lead to a change in a central estimate, across multi-model results, of soil carbon changes by 2100, although interestingly the data constraint does not reduce the overall uncertainty. This highlights the claim made in the title of "high uncertainty" and leave this area as an outstanding issue for modellers and process community alike to address.

I recommend publication after addressing a few minor concerns as below.

Chris Jones

The paper continues the work and analysis of these authors from the Crowther et al. 2016 Nature paper. The dataset assembled of soil warming plots and changes in soil carbon is clearly very valuable for understanding this key issue of Earth System Modelling. I much prefer the approach here, by comparing with process-based models and trying to constrain their simulations, than I do the direct extrapolation to global scale as per Crowther et al. So overall I like the approach deployed here and have mainly minor comments. There are clearly assumptions and choices made which affect the results - I don't believe these should prevent publication - the authors appear to have been careful to consider these and discuss their implications.

2 specific examples are:

- the issue of timescales - the approach to diagnose Q10 relies on pseudo-equilibrium of the system. It could therefore be potentially dangerous and misleading to compare the outcomes from a Q10 diagnostic of 2-5 years warming with ESM studies of 100 years. I was pleased to see this tested explicitly and the SI shows detailed reasons to believe that this is not a large confounding issue.

- the issue of using a single-pool simple model to recreate and re-scale the responses. Again, the authors appear to have taken potential concerns seriously and present good arguments why their approach is OK. I suspect that using a two-pool model would give different answers (inevitable for modelling!) - but it's not clear that this simplification is inappropriate. Given that prominent ESMs in the past, and still some at CMIP5, still deploy a single soil carbon pool it is reasonable for a simple-model approach to do so (by definition simple models bring clarity at the expense of detail - I feel it is well argued here that the choice is appropriate).

Thank you for your contributions to this manuscript in this and previous reviews, as well as your supportive comments regarding this revision.

I have one issue with some of your methodological choices

- you sample q10 values down to 0.1, and in your ESM scaling define "typical" as going down to q10=0.5. It seems VERY odd (unphysical even) to take q10 below 1. Given this is not a linear relationship, you are not looking at taking gradients down towards 0 which represent being flat. Rather q10=1 is where you hit "flat" - i.e. no sensitivity at all to temperature. Once you go below q10=1 then you get an inverse relationship which says that respiration increase as you get COLDER. I'm not aware of any way this could be possible. So I would strongly suggest you take q10=1 as your lower boundary.

While we agree that a Q10 value as low is 0.1 is unusual we had two reasons for choosing this boundary. First given that we do not disentangle moisture effects, it was conceivable that an increase in soil temperature could result in a decrease in respiration visa-vi drier soils imposing stronger moisture limitations. Secondly from a numerical prospective choosing a boundary slightly outside the expected numerical range can demonstrate a robust convergence. We have added these justifications to the methods section and hope they satisfy your concerns.

specific minor comments:

- please be very careful discussing response to "temperature" - it is always important to specify clearly if this means air- or soil-temperature. It has long been known that the apparent q10 is much lower when calculated as a response to air temp (Raich & Potter (1995, GBC) suggest a Q10 value of 2.0 for soil temperature is roughly equivalent to 1.7 for air temperature). Please check especially the Bond-Lamberty paper/comparison - they calculate a low Q10, as you say - but isn't this for air-temp? The comparison is then perhaps misleading

The reviewer is entirely correct and we apologize for letting this slip past us from previous reviews. We have add 'soil' to each mention of temperature in the manuscript.

- be careful using cCwd - is this a distinct carbon pool? or is it a sub-component of cLitter? I believe that the total vegetation system is captured by cVeg+cLitter+cSoil (e.g. see Jones et al., 2013, J.Clim). Then, within this, cLitter is split into tier-2 variables of cCwd, cLitterAbove, cLitterBelow - these are intended to allow reporting in greater detail, but are not new pools. So I think you should remove any cCwd data from your study. I appreciate this is not well explained in the CMIP5 data request - we tried to clarify this for CMIP6 - see figure 5 of Jones et al (C4MIP documentation paper, GMD, 2016)

We removed cCwd from the carbon pools and updated the manuscript. There were minor changes to the totals in the CCSM, CESM and NorESM1 models but no significant changes to the results.

- Although HadGEM2-ES does base it's soil carbon scheme on RothC it is not identical - in particular we chose to keep a uniform q10=2 rather than the RothC temperature function (which we found to be not well behaved at very low T). Otherwise your description of HadGEM2 is correct.

Thank you! We have updated the table.

- the discussion correctly discusses the possible role of soil moisture as a rate modifier. Also to consider are vegetation cover, soil quality, soil structure and changes in input quality. Some models, like RothC, change their decomposition according to overlying vegetation. They also change their allocation (your "b" matrix in equation 1) according to vegetation type and lability of litter inputs. I don't think these are major factors, but worth mentioning its not just T and moisture which change the respiration.

We agree. We have added a sentence at the end of the discussion on how this would affect the analysis.

Anonymous Referee #2 Received and published: 12 March 2018

This is a great study, in which Todd-Brown et al. creatively combined experimental soil warming data with Earth system model outputs to illustrate the observation informed uncertainty in future responses of soil organic carbon to warming. I'd recommend publication after addressing a few issues outlined below.

We thank the reviewer for their excitement about this study and the detail with which they have treated this review. We hope that the below address the concerns raised here.

1. While it's hard to move away from steady state assumption (it leads to very convenient mathematical forms), I think it is important to at least illustrate the bias that steady state assumption can have on estimated Q10. It would be hard to do for observed SOC responses to warming without very crude assumptions about base turnover rate, but seems realistic for the ESM output. The term on the left side of equation 1 can be extracted from the ESM output and used instead of 0, and baseline k as well as Q10's can be reverse-engineered. I think the steady state bias could

be quite substantial, because the 10% (or even 1%) difference between C input and output can have a cumulative effect over time. With observation-derived Q10 estimates, I suggest excluding sites that are with high degree of certainty are not in steady state. An example is Delta Junction site in Alaska, which experienced stand replacing fire in 1999. Even though warming experiment lasted for 10 years it takes many decades for the SOC to recover after a stand replacing fire (Fu et al., 2017), so equation 4 would not be appropriate for this site.

Since the reviewer acknowledged the necessity of this assumption in the field data we'll confine our reply to the Earth system model analysis here. We would direct the reviewer to figure S3, which shows the ratio between the annual change in soil carbon stock divided by either the maximum input or output to the soil at the beginning and end of the 21st century. While there are some grid cells in most models with a relatively high imbalance, the vast majority of models have a ratio below 10% with over half of the grid cells being below 0.1%. Thus the change in soil carbon stock is at least an order of magnitude less than the input/output fluxes, allowing us to assume that the inputs approximate the outputs and apply equation 3 as stated. While the reviewer is absolutely correct that small imbalances in the input/output fluxes can have a large cumulative influence over time, we feel that we have demonstrated that most of the change in the grid-by-grid soil carbon stocks is attributed to shifts in the quasi steady state not in the relaxation of the system to a quasi-steady state. We have expanded on the implications of this figure in the results section.

2. This leads me to equation 4. I would strongly suggest not omitting the ratio of control inputs to warmed inputs. As authors rightly pointed out, if ratio is less than 1, Q10's are going to be underestimated, and if warming leads to decrease in NPP, the Q10's would be overestimated. A meta-analysis study by Wu et al. (2011) illustrated that warming increases total NPP on average by 15%, and belowground NPP by 52%, which has very substantial implications for Q10 (please see the attached image illustrating the effect of change in productivity on Q10's for Delta Junction, AK). I think looking up changes in NPP estimates for every site would make the observation-based Q10 estimates more defendable.

While we agree with the reviewer that a site by site shift in NPP would be more robust, this is beyond the current scope of the study. We would draw the reviewer to the caveats stated in the discussion around the shifts in inputs. We hope to revisit this in future studies.

3. Lastly it wasn't completely clear to me why authors chose to estimate observation- based Q10's with the method described in section 2.3 instead of directly solving for it using equation 3 or 4. Could you please touch on that in the updated version of the manuscript?

We have expanded on our reasoning in the methods section. Briefly, by selected the parameter value based on the model-data fit we have a robust way to describe the uncertainty associated with the parameter as well as clearly community the effects of that parameter on the model fit. While the reviewer is absolutely correct that we could have instead described the distribution of the site specific field-Q10 values, we assert that the method used in this study is equally valid.

Below, please see a list of minor issues:

P1L25: "stimulate"?

P1L25: We will 'stimulate' for simulate.

P1L27: "(the opposing carbon flux)" can be omitted in my view

P1L27: We respectfully disagree and have kept this in for clarity.

P3L20: It seems that the model is not accurately specified: vector C(t) of size n by 1 cannot be multiplied by n by n matrix K. The accurate specification of the negative term would be "A*(Q10*K)*C(t)"

P3L20: Thank you for catching this. We have written the equation to reflect the version in the SI where we use the form '(Q_10*K*A)*C(t)'. This does not change the results of the analysis.

P3L27: "scalar"?

P3L20: We substituted 'scalar' for scaler.

P4L2: "vice-versa"?

P4L2: We substituted 'vice-versa' for vis-versa.

P6L16: "data-driven"?

P6L16: We substituted 'data-driven' for data-drive

P6L18: "grid cell"

P6L18: We substituted cell for cells.

P8L10-11: Please revise the topic sentence, it's not very clear.

P8L10-11: We will replace this topic sentence and rework this paragraph as several reviewers found this confusing.

P8L16: did you mean Figure S3?

P8L16: The citation for Figure S4 is correct but we have added a reference to S3 as well.

P8L25: typo: "ESM-Q10"

P8L25: We fixed this typo.

P.S. here are the references used in the review:

Fu Z, Li D, Hararuk O, Schwalm C, Luo Y, Yan L, Niu S (2017) Recovery time and state change of terrestrial carbon cycle after disturbance. Environmental Research Letters, 12, 104004.

Wu Z, Dijkstra P, Koch GW, Peñuelas J, Hungate BA (2011) Responses of terrestrial ecosystems to temperature and precipitation change: a meta-analysis of experimental manipulation. Global Change Biology, 17, 927-942.

Anonymous Referee #3 Received and published: 14 March 2018

This is a well-written and original paper with findings that are relevant to the Biogeosciences community. The code documentation is particularly great.

I have reviewed for an earlier version of this manuscript, and they have successfully addressed my previous comments, adding some discussion of how a systematic change in input rates (which I understand is not in the dataset but is nevertheless likely) would affect carbon stocks and Q10, and justifying the use of a one-pool model. From the other referee comments, these

points may warrant some additional discussion. Therefore, I recommend publication after minor revision.

We would like to thank the reviewer for this and their previous reviews. We hope that the below address the reviewers concerns raised here.

Two additional thoughts: Most versions of the Century soil model embedded in a number of the ESMs listed as 'cascade' models (e.g., the CLM family models) are feedback models where some soil C is transferred from the slow and passive to the active pool. C1

These versions of Century may well be simplified to the 'cascade' model type, but that development is not clear from the descriptions of the ESM soil models.

We have updated Table 1 and the model descriptions to clarify the models which were modified-CENTURY (which are generally cascade models) and models which are direct implementations. We also added a new model table to the SI which fully documents the nested model references to the best of our ability. We've removed "CENTURY" from the Temperature column of Table 1 and replaced it with a generic function with a citation to the CENTURY paper in the caption. We apologize for the confusion.

I know it's hot of the press, but because the authors base their analysis off of the Crowther et al. 2016 dataset, they could also mention the potential implications of the expansion on their analysis in the Van Gestel et al. comment.

van Gestel, N., Shi, Z., van Groenigen, K. J., Osenberg, C. W., Andresen, L. C., Dukes, J. S., ... & Reich, P. B. (2018). Predicting soil carbon loss with warming. Nature, 554(7693), E4.

The van Gestel manuscript is a fascinating data set that we look forward to addressing directly in future studies. We've added a comment to this effect in the discussion and to the section where we discuss the statistical power of the dataset.

RZ Abramoff rose.abramoff@gmail.com Received and published: 12 March 2018

This is a well-written and original paper with findings that are relevant to the Biogeosciences community. The code documentation is particularly great.

I have reviewed for an earlier version of this manuscript, and they have successfully ad- dressed my previous comments, adding some discussion of how a systematic change in input rates (which I understand is not in the dataset but is nevertheless likely) would affect carbon stocks and Q10, and justifying the use of a one-pool model. From the other referee comments, these points may warrant some additional discussion. There- fore, I recommend publication after minor revision.

Two additional thoughts: Most versions of the Century soil model embedded in a number of the ESMs listed as 'cascade' models (e.g., the CLM family models) are feedback models where some soil C is transferred from the slow and passive to the active pool. These versions of Century may well be simplified to the 'cascade' model type, but that development is not clear from the descriptions of the ESM soil models.

I know it's hot of the press, but because the authors base their analysis off of the Crowther et al. 2016 dataset, they could also mention the potential implications of the expansion on their analysis in the Van Gestel et al. comment.

van Gestel, N., Shi, Z., van Groenigen, K. J., Osenberg, C. W., Andresen, L. C., Dukes, J. S., ... & Reich, P. B. (2018). Predicting soil carbon loss with warming. Nature, 554(7693), E4.

Thank you for this and previous reviews of this study. I believe that your concerns are addressed in RC4.

M. Mayes (Referee) mayesma@ornl.gov Received and published: 16 March 2018

This is a very interesting study in which the newly-derived Q10 values from the Crowther et al. 2016 study are incorporated into a post-hoc simulation of 20 CMIP5 Earth system models. Surprisingly, better-constrained data does not reduce uncertainty in predicted soil organic carbon (SOC) values. Rather, predicted SOC stocks are considerably less in comparison, while retaining very large uncertainties. The study seems very well-done, relies on important new findings, uses an important suite of models, and provides important insights into both predicted SOC and model uncertainty. This paper is a logical outcome of the Crowther et al. 2016 study and will certainly be of interest to the Biogeosciences community.

The authors provide a clear description of an assumption regarding metastable SOC stocks (Eqns 2-4), and how stock differences between two soils, or control and warming, can be used in the model. Authors also completely describe the implications of the (necessary) simplifying assumption of a one-pool model. These are both rather sticky subjects and the authors do an adequate job of addressing these concerns.

Thank you for your interest and review of this study. We hope the below response to your specific comments adequately addresses the concerns you've raised.

Specific comments P1 L22: unclear meaning

P1L22 Thank you for pointing this out. We have simplified this sentence to read: "This study demonstrates that data integration should strive to capture the variation of the system, as well as mean trends."

P2 L20: probably could remove statement about a wide range of typical Q10s.

P2L20 We've chosen to keep this in, mostly to introduce the concept of a 'typical' Q10 value which we come back to later in the manuscript. However, we thank the reviewer for their feedback in this matter.

P4 L3: replace with "an inverse with all positive entries"

P4L3 Thank you for the rewording. We have made the change.

P4 L9: "constructed"

P4L9 This has been corrected. Thank you.

P6 L8: are these shown?

P6L8 Yes, we would direct the reviewer to the results section P7L5. However, several reviewers have had difficulty with this section and we have rewritten these section to clarify these points.

P8 L10: some edits needed

P8L10 Thank you for pointing out this confusing section. Several reviewers had difficulty with this section and we will rewrite this paragraph.

P10 L22: unclear what you mean by "models should increase their variability"

P10L22 Thank you for pointing this out. We have expanded on this section to attempt to clarify.

This study was unusual in using a Q10 function derived from changes in soil C pools to model C fluxes. As such, it was a novel contrast with- and useful comparison to earlier studies based on respiration metrics. The modeling rationale is solid and the simplification of the various model formulations to aggregated responses was reasonable. Overall, I'm impressed with the logic and thoroughness of this study.

This work is important for a several reasons. First, it uses a novel derivation of Q10 to address the important topic of soil C dynamics. Second, the reduction in structural complexity of several different models demonstrated how such aggregations could be done comparably and generated a range of predictions based on existing ESMs. Finally, it suggested that more attention to underlying uncertainties in factors controlling C dynamics might improve outcomes – perhaps as a logical alternative to broad data integration projects.

Regardless of the novelty of this study, the variation associated with the final output is so large that mean estimates of soil C are not different from earlier work, considering the 95% CIs. This simple fact shifts the main focus from differences between these estimates to their similarities, and as the authors noted, reasons why the variation is so large.

The authors raise several points about their underlying assumptions, some also raised by reviewers, noting that uncertainties in soil C stock data, moisture variations, the assumption of steady-state C pool dynamics, uniform temperature sensitivity of various C pools, nutrient limitations, etc., likely all contribute to variations in prediction. More- over, aggregations across time, space, and structural resolution of both the models and C pools sacrifice fine scale dynamics that are often non-linear and cannot be averaged across coarser scales, e.g., moisture responses of dry-land systems. So, it's not surprising that the variation in output was large.

I recommend publishing this article but given the large uncertainty in final predictions, I also recommend a more thorough discussion of the limitations of such broad scale approaches. I'd like to hear more from the authors about how different sources of variation could be elucidated and addressed to improve model performance.

We would like to thank the reviewer for their time and encouraging comments. We agree that other processes and variables would likely provide some predictive power in this and other studies. We have added a very limited review of mechanisms that are currently being considered for model representation in the discussion.

Field-warmed soil carbon changes imply high 21st century modeled uncertainty

Katherine Todd-Brown[1], Bin Zheng[1], Thomas Crowther[2]

[1]Pacific Northwest National Laboratory, Richland, WA, 99354, USA
[2]Institute of Integrative Biology, ETH Zürich, Univeritätstrasse 16, 8006, Zürich, Switzerland

Correspondence to: Katherine Todd-Brown (katherine.todd-brown@pnnl.gov)

Abstract. The feedback between planetary warming and soil carbon loss has been the focus of considerable scientific attention in recent decades, due to its potential to accelerate anthropogenic climate change. The soil carbon temperature sensitivity is traditionally estimated from short-term respiration measurements -- either from laboratory incubations that are artificially manipulated, or field measurements that cannot distinguish between plant and microbial respiration. To address these limitations of previous approaches, we developed a new method to estimate soil temperature sensitivity (Q_{10}) of soil carbon directly from warming-induced changes in soil carbon stocks measured in 36 field experiments across the world. Variations in warming magnitude and control organic carbon percentage explained much of field-warmed organic carbon percentage (R^2=0.96), revealing Q_{10} across sites of 2.2, [1.6, 2.7] 95% Confidence Interval (CI). When these field-derived Q_{10} values were extrapolated over the 21st century using a post-hoc correction of 20 CMIP5 Earth system model outputs, the multi-model mean soil carbon stock changes shifted from the previous value of 88 ±153 Pg-carbon (weighted mean ± 1 SD), to 19±155 Pg-carbon with a Q_{10} driven 95% CI of 248±191 to -95±209 Pg-carbon. On average, incorporating the field-derived Q_{10} values into Earth system model simulations led to reductions in the projected amount of carbon sequestered in the soil over the 21st century. However, the considerable parameter uncertainty led to extremely high variability in soil carbon stock projections within each model; intra-model uncertainty driven by the field-derived Q_{10} was as great as that between model variation. This study demonstrates that data integration should capture the variation of the system, as well as mean trends.

1 Introduction

The flux of carbon dioxide between the soil and atmosphere is a major control on atmospheric carbon dioxide concentrations. Warming temperatures, driven by increases in atmospheric carbon dioxide, have the potential to stimulate carbon decomposition, accelerating its release into the atmosphere (Davidson and Janssens, 2006). If this is not counterbalanced by an equivalent increase in primary productivity (the opposing carbon flux) then it has the potential to drive a land carbon-climate feedback that will accelerate anthropogenic climate change. Recent global compilations of data from ecosystem warming experiments lend support to this idea (Carey et al., 2016), suggesting that warming alone could drive a loss of carbon from the upper soil horizons (Crowther et al., 2016). However, these studies addressed the impact of warming in isolation, and

---

Deleted: 83
Deleted: 6
Deleted: 6
Deleted: 6
Deleted: 45
Deleted: 4
Deleted: 9
Deleted: 8
Deleted: measured
Deleted: may not improve model certainty, but instead
Deleted: strive to

1

it remains unclear how this process will interact with the variety of other global change drivers to affect the global soil carbon stock over the rest of this century. Reflective of such uncertainty, soil carbon changes projected for 2100 under business-as-usual scenario for Coupled Model Intercomparison Project Phase 5 (CMIP5) range from -70 to 250 Pg-carbon across different Earth system models (Todd-Brown et al., 2014), making the land-carbon feedback one of the largest sources of uncertainty in future climate projections (Friedlingstein et al., 2014). Improving the soil carbon component of the Earth system models is essential to predicting the future evolution of the Earth system and thus establishing meaningful greenhouse gas emissions targets.

A fundamental parameter describing soil temperature-sensitivity in soil carbon models is the Q_{10} – the factor of the change in decomposition rate associated with 10°C of warming from a reference temperature (Davidson and Janssens, 2006; Lloyd and Taylor, 1994). Traditional laboratory incubations have found a wide range of Q_{10} values, varying from 1.4 (Townsend et al., 1997) to > 3 (Davidson et al., 1998, 2006) with 2 being the most commonly accepted value. Complicating this, theoretical analyses based on chemical kinetics suggest Q_{10} is itself dependent on temperature (Davidson and Janssens, 2006; Lloyd and Taylor, 1994), though these values are typically very close to 2 in most environmental temperature ranges (Lloyd and Taylor, 1994). More recently, large-scale analyses of field respiration converge on Q_{10} estimates of 1.4 to 1.5 (Bond-Lamberty and Thomson, 2010; Hashimoto et al., 2015; Mahecha et al., 2010). Unsurprisingly, this temperature response is also critical in Earth system models, where the temperature sensitivity parameter is known to be a major driver of variation (Booth et al., 2012; Jones and Cox, 2001; Jones et al., 2006). However, it is unclear what is driving the lower Q_{10} estimates in these field-based syntheses compared to the average lab-based estimates from single-site studies, and there appears to be a relatively wide range of 'typical' Q_{10} values in the literature. Nevertheless, most Earth system models use values that range from 1.5 (Oleson et al., 2013; Raddatz et al., 2007) to 2 (Bonan, 1996; Cox, 2001).

Traditionally, these Q_{10} values have been calculated from warming-induced changes in soil respiration rates. However, this approach has two main limitations: 1) respiration rates measured under idealized laboratory conditions fail to reflect the structure, heterogeneity and variability of natural systems, whereas 2) field measurements cannot directly isolate heterotrophic soil respiration from autotrophic root respiration without substantially altering the system. Estimating Q_{10} directly from warming-induced changes in soil carbon stocks could be a valuable approach to address these limitations, but the variability and relative imprecision of soil carbon stock data necessitates a large sample size to adequately describe variation at the global scale (Bradford et al., 2016). Yet, results from a recent Earth system model meta-analysis indirectly suggests that, with enough sample coverage it may be possible to infer Q_{10} directly from changes in soil carbon stocks (Todd-Brown et al., 2014).

Here we present a new approach to estimate the global Q_{10} value from net changes in soil carbon stocks under warming, rather than soil respiration measurements, and examine the consequences of these estimates –with associated uncertainty– on CMIP5 Earth system model projections of global carbon storage over the rest of the 21st century. To do this, we use a global database of soil carbon stock data from 36 field-warming experiments around the world, each of which includes control (ambient) plots, and those which have been warmed for extended (years to decades) periods of time (Crowther et al., 2016) (Tables SI1) and outputs from 20 CMIP5 (Taylor et al., 2011) Earth system models RCP 8.5 business-as-usual experiment (Tables 1 and SI3).

2

These field data were used previously to derive Earth system model independent estimates of global soil carbon temperature sensitivity where the effect of warming was isolated from other global change drivers or the interacting climate system (Crowther et al., 2016). In this study we develop a novel approach that enables us to explore these field results in the context of the temperature sensitivity function (Q_{10}) used in integrated Earth system model. We then examine the consequences of the data-driven Q_{10} estimates, and the associated uncertainty, for CMIP5 Earth system model projections of global carbon storage over the rest of the 21st century using a novel post-hoc modification of the CMIP5 simulation outputs.

2 Materials and Methods

2.1 Field sites

The field sites were drawn from a previous analysis (Crowther et al., 2016). From this initial database of 48 paired case-control studies, we selected 36 studies that were run longer than 2 years to match the metastable state assumption articulated below. 18 of these sites were temperate grasslands, savannas, and shrublands, 10 temperate broadleaf and mixed forests, 6 tundra, 1 boreal forests or taiga, and 1 site was in a Mediterranean forest, woodland and scrub. A traditional statistical analysis of the sites is provided by Crowther *et al.* (2016). For this study, we used the increase in soil temperature due to warming, length of the study, and the percent of soil organic carbon in paired warmed and control plots (Table SI1).

2.2 Q_{10} calculations

We calculated traditional Q_{10} estimates based on these warming-induced soil carbon losses, enabling us to embed this temperature sensitivity information into a soil decomposition model framework. Traditional soil decomposition models follow a first order linear decay framework where:

$$\frac{d\boldsymbol{C}(t)}{dt} = u_{in}(t)\boldsymbol{b} - (\boldsymbol{Q_{10}}(T,t)\boldsymbol{KA})\boldsymbol{C}(t), \tag{1}$$

where the $\boldsymbol{C}$ is a vector of soil carbon pools with unique turnover times, t time, u_{in} a scaler of soil carbon inputs, $\boldsymbol{b}$ an allocation vector describing how the inputs are divided between the soil carbon pools, $\boldsymbol{K}$ is a diagonal matrix representing the decomposition rates of the pools, $\boldsymbol{Q_{10}}$ is a diagonal matrix with entries of the form $q_i^{(T(t)-T_0)/10}$ representing the temperature sensitivity factor, T a scalar describing the soil temperature and T_0 an arbitrary reference temperature, and A the transfer matrix representing movement of carbon between soil carbon pools.

The temperature sensitivity was assumed to be constant across pools. This allows us to collapse the diagonal $\boldsymbol{Q_{10}}$ matrix to be collapsed into a single scalar value of the form $Q_{10}^{(T(t)-T_0)/10}$. This constant temperature sensitivity assumption is discussed below and follows the structure of the CMIP5 Earth system models.

In general, there are three classes of pool structure for traditional models: *independent* where there was no exchange between soil carbon pools making A the identity matrix, *cascade* where pools with faster turnover times passed carbon to pools with slower turnover times making A a lower triangular matrix, and fully *feedback* models where carbon was exchanged between

3

faster and slower pools and vice-versa making A a fully dense matrix. In all cases KA is an M-matrix, implying there exists an inverse with all positive entries. For the independent and cascade pools KA is diagonalizable, implying it can be broken down into a diagonal matrix D and an invertible matrix P such that $KA=P^{-1}DP$.

For most well-developed soils, soil carbon stocks are at a metastable state where soil inputs approximately equal outputs (see Results for discussion of Earth system model outputs). Given that KA is an M-matrix and this metastable state approximation, we can describe the total soil organic carbon as follows:

$$C = \frac{u_{in}}{kQ_{10}^{(T-T_0)/10}}, \tag{2}$$

where C is the total organic carbon stock, u the sum of the soil inputs, and k a bulk decomposition rate that can be constructed from the decay matrix KA and allocation vector of the soil inputs b. For details see of this derivation the SI: Mathematical Analysis.

We can now describe the soil carbon stock difference between two soils with the same decay rate but different temperatures and inputs. This could either be two time points from a simulation where the soil output is close (within 10%) of the soil inputs, or a warmed treatment and a control:

$$C_2 = C_1 \left(\frac{u_2}{u_1} Q_{10}^{(T_1-T_2)/10} \right). \tag{3}$$

For the field sites, we assume that the relative change in inputs due to warming is negligible compared to the effect on the decomposition rate across sites and that the main driver of differences in decomposition rates between control and treatment is the warming treatment. Leading us to:

$$C_w = C_c Q_{10}^{-\Delta T/10}. \tag{4}$$

Finally, we assume that the bulk density of the soil at a given site was unaffected by the warming treatment. This allows us to use the mass percent soil organic carbon instead of the soil organic carbon density for Eq. 4.

2.3 Model-data integration: parameter fitting

Given the relatively small parameter space, we choose a brute-force model-data integration approach where we iteratively calculated the predicted change in soil carbon stock given the control soil carbon (Eq. 4) across a range of Q_{10} values from 0.1 to 5 in 0.1 increments. We set the lower bound of the Q_{10} range to 0.1 instead of 1 for two reasons. First, while it is generally accepted that warmer soil temperatures will increase soil respiration (constraining $Q_{10} > 1$), it is possible that a warmer soil would result in drier soils and suppress soil respiration. In addition, numerically we wanted to bracket the expected parameter range with our prior. Data-model fits were scored using root mean squared error (RMSE) and linear regression (R^2, slope and intercept). Q_{10} values were selected on low bias (slopes and intercepts within 2 standard deviations of 1 and 0 respectively) due to the relative insensitivity of the R^2 and RMSE metrics (see Figure 1). By selecting the parameter based on model-data fit instead of deriving a direct Q_{10} value for each site and using the distribution, we demonstrate the robustness of the model and have a clear metric to select the parameter range. To test for statistical power, we randomly sampled the data 1000 times with sample sizes from 5 to 34 sites and compared this to samples with randomly assigned control vs warming (for each study

4

the percent carbon of control and treatment has a 50% chance of being switched). These random and sample generated Q_{10} distributions, were compared using a two-sample Kolmogorov-Smirnov test to test that the distributions were statistical distinct.

2.4 Earth system model analysis

5   Earth system model simulations were drawn from CMIP5, the Coupled Model Intercomparison Project to support the 5th IPCC assessment report (Taylor et al., 2011). We downloaded simulation outputs from the RCP 8.5 scenario, representing the 'business as usual' scenario, including heterotrophic respiration (rh), soil temperature (tsl), and heterotrophic carbon stock ($cSoil$ and $cLitter$) from the CMIP5 repository on the Earth System Federation Grid. Ten-year means were taken at the beginning and end of the 21st century for each variable (corresponding to 2006-2015 and 2090-2099). Soil temperature was

10   averaged for the first 10 cm to correspond with experimental soil temperature readings. Soil carbon stock was calculated by adding all heterotrophic-respiring pools (including soil $cSoil$ and litter $cLitter$) where multiple pools were reported. Soil carbon inputs were calculated from the monthly change in soil carbon stock plus the reported heterotrophic respiration. Model variable summaries can be found in Tables SI3 and processing code is documented in SI.

These 20 Earth system models are built from previous models which contain 10 distinct soil sub-models (Table 1). The number

15   of soil carbon pools in these ESMs varied from 1 (INM-CM4) to 8 (BCC-CSM1.1) with most models having 2 to 5 pools. None of the models reported soil carbon with depth, although GFDL documents a depth dependent model. There were three classes of pool structure for these models: *independent* where there was no exchange between soil carbon pools, *cascade* where pools with faster turnover times passed carbon to pools with slower turnover times, and fully *feedback* models where carbon was exchanged between faster and slower pools and vis-versa. In this set of models; 2 of these soil models were full feedback

20   models (HadGEM, ISPL-CM), 6 were cascade pool structure (MRI-ESM1, MIROC-ESM, MPI-ESM, CLM4.0 [CESM1, CCSM4, NorESM1], CanESM2, BCC-CSM1.1), and 2 were independent pools (GFDL-ESM2, INM-CM4). Only 2 model documented an explicit constant Q_{10} (INM-CM2 and HadGEM2, Q_{10}=2), 1 model documented a soil temperature dependent Q_{10} (CanESM2), 4 models documented a soil temperature sensitivity from Lloyd and Taylor which behaves very similar to Q_{10}=2 under moderate temperatures (Lloyd and Taylor, 1994), and the remaining 3 (ISPL-CM5, GFDL-ESM2, BCC-CM1.1)

25   all used a variation of the soil temperature sensitivity proposed in CENTURY (Parton et al., 1987, 1988) which also behaves very similar to Q_{10}=2 under moderate temperatures but declines at high temperatures (Lloyd and Taylor, 1994). The ESMs considered had a single global Q_{10}, or Q_{10}-formula dependent on soil temperature, uniformly applied to the decay pools. This documented structure should be approached with caution due to frequent lags between model development and documentation, actual values and functions may differ. For details with citations see Table 1.

30   Soil carbon stocks at the beginning of Earth system model simulations are typically documented to be spun up to close to steady state, and there is numerical support that this holds throughout the simulation (see Results and Figure S3). Thus Eq. 3 can be extended to the change in soil carbon stock over the 21st century. This leads to the following explicit calculation for a Q_{10} value at each grid cell.

Deleted: ,

Deleted: and $cCwd$

Deleted: surface

Deleted: ,

Deleted: , and coarse woody debris $cCwd$

Deleted: 1

Deleted: 1 proposed a unique temperature sensitivity (HadGEM2),

Deleted: , then

$$\ln(Q_{10}) = \left(\frac{10}{T_m - T_f}\right) \ln\left(\frac{C_f}{C_m} \frac{u_m}{u_f}\right), \tag{5}$$

where the Q_{10} value is related to the modern soil temperature T_m, future soil temperature at the end of the 21st century T_f, modern soil inputs u_m, future soil inputs u_f, modern soil carbon stock C_m and the future soil carbon stock C_f.

For soils that are very close to zero soil carbon stocks, have minimal shifts in soil temperature or have very low soil inputs, the estimated Q_{10} is not finite. Similarly soils which are not well described by their shift in soil temperature (for example, if there is a significant shift in the moisture regime) may have non-typical Q_{10} values that are either less than 0.5 or greater than 5. We examined the amount of shift in soil carbon stocks associated with the four categories of Q_{10} values (nonfinite, less than 0.5, greater than 5, or typical), as well as the spatial patterns associated with these categories.

To support the assertion that the Q_{10} value can be calculated from relatively short time scales found in the field experiments, we examined the distribution typical Q_{10} values associated with similar soil temperature steps experienced by the field experiments at 1, 5, 10, 15, 20, 50, 75, and 84 year time-scales using 10-year mean gridded values of soil carbon stocks, soil inputs, and soil temperature. It should be noted however that changes in the moisture conditions over the 21st century may complicate this analysis of the Earth system model simulations, thus it is not an exact proxy for the field experiments where the control and treatment plots experienced similar baseline climate conditions and a more or less constant offset throughout the experiment.

Finally, the Q_{10} distribution was scaled to reflect the best estimate and uncertainty from the field data. This distribution shift was done by normalizing the Q_{10} map to the mean of the distribution and multiplying it by the experimentally derived values. The Q_{10} correction was only applied to grids with typical Q_{10}'s (non-typical Q_{10}'s were considered to have predominately non-temperature driving variables and their soil carbon stocks were not altered). This normalization shifted the global Q_{10} distribution within the models to match the most common (geographically likely) Q_{10} with the data-driven Q_{10} value, yet by preserving the distribution we preserved other factors affecting changes in decomposition rate (i. e. moisture shifts) in the model. We then recalculated the change in soil carbon for each grid cell with this modified Q_{10} according to Eq. 3 and calculated the global area-weighted totals.

The full analysis script and those used to generate the figures are available in the supplemental.

3 Results

From the changes in soil carbon stocks across field studies, we find a global Q_{10} of 2.2 (90%CI 1.6, 2.7; $R^2 > 0.95$, root mean squared error < 2; Figure 1, Figure S2). The model-data fit was evaluated using a linear regression and root mean squared error (Figure 1). While the R^2 of the model-data comparison was relatively insensitive to the Q_{10} value, there was a notable improvement in the bias with Q_{10} (as defined as the slope within 2 standard deviations of 1 and intercept within two standard deviations of 0). This bias-driven criteria was used to select Q_{10} values from a prior range of (0.1, 5), see Methods for details.

Deleted: s

The Q_{10} distribution was compared with a random null distribution and was significantly distinct (Kolmogorov-Smirnov D=0.441, p < 2e-16, See Table S2, Figures 2 and S1). Randomly selecting 5 to 34 sites from the full dataset were compared to a null distribution where control vs warmed labels were randomized. The quartiles of the data subsets notably converged at a sample size of 25 where the null distribution was relatively invariant across sample size (Figure 2). The distribution of the Q_{10}

5   values under null appeared log-normal, centered around 1 demonstrating no temperature effect (Figure S1). The distribution of the Q_{10} range for the data subsets converged to around 2.2 (Figure S1).

The balance between gridded soil inputs and heterotrophic respiration at both the initial and final 10-year mean for the 21st century was within 10% for over 93% of the grid cells across all models with half of the grid cells within 0.1%. Most models had 95% and two models consistently had 100% of their grid cells within 10% -- the absolute value of the net flux was within

10   10% of the highest primary flux (Figure S3). Thus, the soil inputs are on the same order of magnitude as the soil outputs. This was reflected in very similar Q_{10} distributions regardless of whether soil inputs or heterotrophic respiration was used to derive the Q_{10} value (Figure S4). A notable exception to this was the MIROC-ESM model which did see differences in inputs and heterotrophic respiration drive different Q_{10} distributions (Figure S4).

The inferred Q_{10} values in the Earth system models derived from 10-year mean changes across different time steps (1, 5, 10,

15   15, 20, 50, 75, and 84 years) had similar distributions in most of the models (Figure S4). There were minor shifts in the mode of most models which could be attributable to changes in the moisture conditions or other (non-temperature or input) environmental variables in the simulation. Models aggregated across common land models showed marked similarity in their Q_{10} distributions (Figure S5). There was also an extremely high correlation between Q_{10} values derived from soil inputs compared to those derived for heterotrophic respiration across all models (Figure S6).

20   The inferred Q_{10} values in the Earth system models from the decadal average across the 21st century fell into four categories (nonfinite, less than 0.5, greater than 5, or typical; Figure S7), however most of the change in soil carbon stocks over the 21st century occurred in grid cells with typical Q_{10} values between 0.5 and 5 (Figure S6). A notable exception to this trend was the MRI-ESM1 model where roughly half of the change in carbon stocks occurred in grid cells with Q_{10} values greater than 5 (Figure S6). Spatially the Q_{10} categories showed strong geographical patterns (Figure S7). The GFDL-ESM2 models were

25   dominated by non-finite values in high northern latitudes (Figure S7). MIROC-ESM, CCSM4, CESM1, and NorESM1 models were dominated by Q_{10} values above 5 in the high northern latitudes (Figure S7). Unless otherwise noted, only typical Q_{10} values are addressed for the remainder of this study.

The inferred Q_{10} values for the decadal average across the 21st century, also showed strong geographic patterns (Figures 3) and was typically unimodal (Figure S6). MIROC-ESM and MIROC-ESM-CHEM showed the weakest spatial patterns with

30   high grid-to-grid variation (Figure 3). Mean Q_{10} values fell within the 90% CI of the field data Q_{10}, ranging between 1.8 (CESM1(CAM5), HadGEM2-ES, ISPL-CM5A, and MPI-ESM-MR) and 2.6 (MIROC-ESM-CHEM), with the multi-center Q_{10} values at 2.0 ± 0.2 (Tables 2).

When the inferred Q_{10} values were modified to reflect the data-driven Q_{10} range, resulting variation in the multi-center mean was almost as large as the variation across model projections (Figure 4, Table 2). Re-centering the global Q_{10} distribution to

7

Formatted: Subscript

Formatted: Subscript

Deleted:

Deleted: Differences in inputs and heterotrophic respiration drove some differences in the Q_{10} distribution in the MIROC-ESM models.

Deleted: GFDL-ESM2G

reflect the range of field-driven Q_{10} values (Figure S8) resulted in changes in soil carbon stocks over the 21st century of between -452 Pg-carbon (MPI-ESM-MR) and 525 Pg-carbon (HadGEM2-CC) with a best-estimate Q_{10} (Q_{10} = 2.2) resulting in 19 ± 155 Pg-carbon (multi-center mean ± 1 SD) and field-drive bound (Q_{10} = 1.6, 2.7) of [248 ± 191, -95 ± 209] Pg-carbon (Figure 4, Table 2).

5   **4 Discussion**

By capturing information about warming induced changes to relatively undisturbed field soil carbon stocks directly rather than inferring this from soil respiration rates, this is the first study to generate field Q_{10} estimates of soil carbon losses without needing to correct for belowground autotrophic respiration. Using a simplified version of a traditional decomposition model with a soil temperature sensitivity function, we estimate that the global Q_{10} value is 2.2 ([1.6, 2.7] 95%CI, Figure 1, S2). This

10   Q_{10} is notably higher than previous global estimates based on field soil respiration data (Q_{10} = 1.4 to 1.5 (Bond-Lamberty and Thomson, 2010; Mahecha et al., 2010)), yet well within the range of estimates from laboratory-based studies (Davidson and Janssens, 2006) as well as close to documented soil temperature sensitivity parameters (~2) of Earth system models (Table 1). This Q_{10} range is statistically significant. Resampling the 36-study data set demonstrates the need for over 25 sites to distinguish the Q_{10} range from random (Figure 2 and S1). While the Q_{10} distribution for the 34-study subset is distinct from

15   the null (Kolmogorov-Smirnov D=0.441, p < 2e-16), there appears to be some minor drift in the range suggesting that more study sites could be informative and we hope future studies will include data recently identified (van Gestel et al., 2018). Inferring a decadal-scale environmental sensitive from an annual-scale experiment is generally controversial. However, in this case, traditional model structures assume a temperature sensitivity function that is invariant across space and time and numerical trends in the Earth system model reflect this. In the traditional model structure the soil temperature sensitivity

20   function is applied as a single scaler to multi-pool models causing the relative decomposition response in both fast and slow pools to be the same (for example, (Parton et al., 1987)). Examining the inferred gridded Q_{10} values from annual means across time scales from 1 year to 84 years in Earth system models shows a strong similarity in the distribution most models (Figure S4). Similarly using soil inputs as opposed to heterotrophic respiration did not affect the distribution of the gridded Q_{10} values, with the notable exception of MIROC-ESM which is explained by unusual differences in soil inputs and outputs (Figure S3,

25   S4). Differences in the Q_{10} distribution across time scales are likely driven then by interaction with other sensitivity functions like moisture or in shifts in the allocation of dead vegetation to different pools as the plant type distribution changes over time. If soils are more sensitive to warming than previously expected, then how would this affect future soil carbon stocks over the 21st century? To address this question, we turned to the CMIP5 Earth system models run under RCP 8.5 (Taylor *et al.*, 2011). In order to modify the Earth system model output to reflect the data-driven Q_{10} we applied similar assumptions used in the

30   field-data analysis. We first examine the soil temperature sensitivity of CMIP5 Earth system model simulated soil carbon stocks. In contrast to the field data, we take into account the effect of the change in soil inputs on soil carbon stocks in the Earth system models because these coupled simulations include CO_2 fertilization and other climate effects known to influence

8

primary production (see Methods, Eq. 5). Though these inferred Q_{10} values (Q_{10} = 1.8, 2.6) fall within the uncertainty of the field derived Q_{10} values (Q_{10} = 1.6, 2.7), most ESM-Q_{10} means fell under the median data-Q_{10} value of 2.2 (Table 2) implying ESMs were, on average, less sensitive to soil temperature shifts than the field warmed data would imply. It should be noted that this inferred Q_{10} value is not exactly the parameterized Q_{10} value, and is instead a combination of the soil temperature

5   sensitivity and other environmental sensitivities. If there were, on average, an additional constraint on respiration (such as moisture) we might expect the inferred Q_{10} parameter to be lower than the model parameterized Q_{10}.

There were notable regional patterns across all but 2 of the Earth system model inferred Q_{10} (Figure 3, S7). High northern latitudes tended to have either large or non-finite Q_{10} values suggesting that something other than soil temperature and input shifts were driving changes in soil carbon stock. This alternative driver could be a shift in moisture regimes or dynamics driven

10   by thaw thresholds which could similarly affect the analysis of the field-data. Additional drivers of soil decomposition dynamics, beyond temperature and inputs considered here, have the potential for explaining some of the variation in the Q_{10} range and new model structures are being explored to take some of these mechanisms into account (Luo et al., 2015; Wieder et al., 2015a). This remains an active area of research.

Propagating this field-Q_{10} range into the ESM projections resulted in greater carbon losses from the soil by the end of the 21st

15   century (multi-center means of the soil carbon change, shifted from 88 to 19 Pg-carbon) with large uncertainties; ESM multi-center standard deviation was initially 152 Pg-carbon which is half of the range in multi-model mean attributed to Q_{10} 95% CI [248, -95] Pg-carbon (Figure 4, Table 2). To calculate these modified projections, means of the model specific Q_{10} distributions were re-centered to reflect the best estimate Q_{10} and associated 95% CI from the field data analysis. By preserving the distribution within the model, we attempted to propagate soil moisture sensitivities and other model-specific effects into the

20   modified projections. We also did not modify grid cells with non-typical Q_{10} values (non-finite, below 0.5 or above 5) since those grids likely governed other non-temperature drivers. The large range of carbon shifts in each ESMs driven by this Q_{10} CI confirms the importance of considering parameter uncertainty in the land carbon component of Earth system model projections. The post-hoc correction that we present provides an innovative way to account for this parameter variation without the computational burden of additional ensemble runs.

25   This analysis includes several basic assumptions and caveats. Specifically, we assume that the difference between treatment and control is driven entirely by the soil warming effect, and those warming effects are uniform across soil carbon quality. Though warming-induced changes in soil inputs are, on average, relatively small, they are have been shown to be highly variable in similar sites (Lu et al., 2013). The analysis of field data could be extended to account for these changes in inputs in follow-up studies (Eq. 3). A large increase in soil inputs would cause an underestimation of the Q_{10} value, while a decrease in

30   soils inputs would cause an overestimation of the Q_{10} value (see Eq. 3). While there is some evidence to support soil temperature sensitivity dependency on soil carbon quality (Knorr et al., 2005), there is also evidence for a uniform soil temperature sensitivity (Hicks Pries et al., 2017). as is represented in the Earth system models considered in this study (Table 1). A quality dependent Q_{10} would not be separable from the bulk decay term and thus a one pool model would be inappropriate

9

in this case (see SI). In addition, the dataset has acknowledged biases (see Crowther *et al.*, 2016), which are typical of field studies.

4.1 One pool simplification

We find that multi-year soil carbon dynamics can be well-described by a one pool model at a specific time scale in both the
5   Earth system models and field experiment. If we restrict the decomposition models to those with either independent or cascade pool structures (that is no carbon passed from the slow to the fast pools) then the temporal dynamics of the total soil carbon of the system at a specific time scale can be approximated by a single pool due to the fact that the lower triangular decomposition-transfer matrix is diagonalizable (see SI for details). While this diagonalizable property does not hold for full feedback models where carbon is transferred from the slower to faster carbon pools, all decomposition-transfer matrices are M-matrices. If we
10 combine the positive-inverse properties resulting from this M-matrix structure and assume that the soils are close to metastable state (that is soil inputs are roughly equal to the heterotrophic respiration outputs, as we show in Figure S3 for the Earth system models considered and would expect for soils from intact systems). Then the total soil carbon can be described by a bulk decay rate that is a linear combination of the transfer coefficients, decay rates, and input allocations of the component pools (see SI for analytical details). This provides analytical support for the one pool simplification seen numerically in the Earth system
15 models in the CMIP5 project (Todd-Brown *et al.*, 2013, 2014).

The one pool simplifications described above are controversial assertions. The one pool model has proven inadequate to describe laboratory incubations where heterotrophic respiration over time is compared to the soil carbon stock (Thornton, 1998; Weng and Luo, 2011). This is due to the multiple time scales considered (daily, monthly and annual) and, more importantly, the fact that these laboratory incubations are by their nature not at steady state since any inputs to the system are
20 generally removed. Thus this analysis would not be expected to hold for laboratory incubation and we would further expect the bulk decay rate to change with time scales for sites undergoing rapid changes in inputs (in other words, the bulk decay rate inferred at a 1 year time step would not match the 100 year time step at a site undergoing transition from grassland to forest). Another key assumption is that soil organic carbon of different quality response the same to warming. However, the scalar multiplier representing environmental sensitivities are independent of pools in most models (ex (Parton et al., 1988)). These
25 scalar multiplies (like the Q_{10} temperature sensitivity examined in this study) would be invariant to time scale if this modling assumption is applied to the field analysis. Finally shifts in the allocation of dead vegetation to the different soil pools would shift the bulk decay rate of the one pool approximation (see SI: Analytical proofs). With these caveats in mind, we feel that the one pool approximation is extremely valuable in analyzing soil carbon models and data.

5 Conclusion

30 It is still unclear how the terrestrial carbon cycle in general, and soils in particular, will respond to climate change over the 21st century. The CMIP5 models, representing our best coupled climate models to date, have a wide range of soil carbon responses

10

over the 21st century (Todd-Brown et al., 2014). While it would be nice to have all the models agree on a tightly bound answer, the question we should be asking scientifically is: Does the variation in the models reflect our best scientific understanding? Models must capture not only mean trends but also system variance and accurately represent scientific uncertainty.

Post-hoc correction of simulation results can provide some insight into known gaps in Earth system models without the computational hurdle of re-running simulation results. Previous studies have applied post-hoc corrections to address nutrient limitations on net primary production (Wieder et al., 2015b) and this study demonstrates the high level of uncertainty that can be driven by the soil temperature response parameter. This study suggests that soil carbon response to warming is highly variable in the field and ESMs should increase their variability to reflect this field-variation. Future studies increasing the number of field-warmed studies (van Gestel et al., 2018), as well as extending the field data to include changes in plant productivity in response to warming would inform the field-derived Q_{10} analysis explored here. In addition, explaining field moisture and applying that understanding to a post-hoc Earth system model analysis is a logical next step.

6 Acknowledgements

The authors have no conflict of interest to declare.

Deleted: model studies

Deleted: capture this

Formatted: Subscript

Deleted: also

11

7 Author Contributions

KTB and TWC both contributed to the writing of this manuscript. KTB was responsible for the analysis. TWC coordinated the field data set. BZ was responsible for the mathematical analysis and made contributions to the revisions of the manuscript.

References

5 Adachi, Y., Yukimoto, S., Deushi, M., Obata, A., Nakano, H., Tanaka, T. Y., Hosaka, M., Sakami, T., Yoshimura, H., Hirabara, M., Shindo, E., Tsujino, H., Mizuta, R., Yabu, S., Koshiro, T., Ose, T. and Kitoh, A.: Basic performance of a new earth system model of the Meteorological Research Institute (MRI-ESM1), Pap. Meteorol. Geophys., 64, 1–19, doi:10.2467/mripapers.64.1, 2013.

Arora, V. K.: Simulating energy and carbon fluxes over winter wheat using coupled land surface and terrestrial ecosystem

10 models, Agric. For. Meteorol., 118, 21–47, doi:10.1016/S0168-1923(03)00073-X, 2003.

Arora, V. K. and Boer, G. J.: Uncertainties in the 20th century carbon budget associated with land use change, Glob. Chang. Biol., 16(12), 3327–3348, doi:10.1111/j.1365-2486.2010.02202.x, 2010.

Arora, V. K., Scinocca, J. F., Boer, G. J., Christian, J. R., Denman, K. L., Flato, G. M., Kharin, V. V, Lee, W. G. and Merryfield, W. J.: Carbon emission limits required to satisfy future representative concentration pathways of greenhouse gases, Geophys.

15 Res. Lett., 38, L05805, doi:10.1029/2010GL046270, 2011.

Bonan, G. B.: A land surface model (LSM version 1.0) for ecological, hydrological and atmospheric studies: technical discription and user's guide, NCAR Tech. note NCAR/TN-417+STR [online] Available from: http://opensky.library.ucar.edu/collections/TECH-NOTE-000-000-000-229, 1996.

Bond-Lamberty, B. and Thomson, A.: Temperature-associated increases in the global soil respiration record, Nature,

20 464(7288), 579–582, doi:10.1038/nature08930, 2010.

Booth, B. B. B., Jones, C. D., Collins, M., Totterdell, I. J., Cox, P. M., Sitch, S., Huntingford, C., Betts, R. A., Harris, G. R. and Lloyd, J.: High sensitivity of future global warming to land carbon cycle processes, Environ. Res. Lett., 7(2), 24002, doi:10.1088/1748-9326/7/2/024002, 2012.

Bradford, M. A., Wieder, W. R., Bonan, G. B., Fierer, N., Raymond, P. A. and Crowther, T. W.: Managing uncertainty in soil

25 carbon feedbacks to climate change, Nat. Clim. Chang., 6(8), 751–758, doi:10.1038/nclimate3071, 2016.

Cao, M. and Woodward, F. I.: Net primary and ecosystem production and carbon stocks of terrestrial ecosystems and their responses to climate change, Glob. Chang. Biol., 4, 185–198, doi:10.1046/j.1365-2486.1998.00125.x, 1998.

Carey, J. C., Tang, J., Templer, P. H., Kroeger, K. D., Crowther, T. W., Burton, A. J., Dukes, J. S., Emmett, B., Frey, S. D., Heskel, M. A., Jiang, L., Machmuller, M. B., Mohan, J., Panetta, A. M., Reich, P. B., Reinsch, S., Wang, X., Allison, S. D.,

30 Bamminger, C., Bridgham, S., Collins, S. L., de Dato, G., Eddy, W. C., Enquist, B. J., Estiarte, M., Harte, J., Henderson, A., Johnson, B. R., Larsen, K. S., Luo, Y., Marhan, S., Melillo, J. M., Peñuelas, J., Pfeifer-Meister, L., Poll, C., Rastetter, E., Reinmann, A. B., Reynolds, L. L., Schmidt, I. K., Shaver, G. R., Strong, A. L., Suseela, V. and Tietema, A.: Temperature

response of soil respiration largely unaltered with experimental warming, Proc. Natl. Acad. Sci., 113(48), 13797–13802, doi:10.1073/pnas.1605365113, 2016.

Coleman, K. and Jenkinson, D. S.: ROTHC-26.3, A model for the turnover of carbon in soils, Model Descr. Wind. user Guid. [online] Available from: http://www.rothamsted.ac.uk/ssgs/RothC/mod26_3_win.pdf, 1999.

5   Collins, W. J., Bellouin, N., Doutriaux-Boucher, M., Gedney, N., Halloran, P., Hinton, T., Hughes, J., Jones, C. D., Joshi, M., Liddicoat, S., Martin, G., O'Connor, F., Rae, J., Senior, C., Sitch, S., Totterdell, I., Wiltshire, A. and Woodward, S.: Development and evaluation of an Earth-System model – HadGEM2, Geosci. Model Dev., 4(4), 1051–1075, doi:10.5194/gmd-4-1051-2011, 2011.

Cox, P.: Discription of "TRIFFID" Dynamic Global Vegetation Model, London Road, Bracknell, Berks R12 2SY, UK. [online]
10   Available from: climate.uvic.ca/common/HCTN_24.pdf, 2001.

Crowther, T. W., Todd-Brown, K. E. O., Rowe, C. W., Wieder, W. R., Carey, J. C., Machmuller, M. B., Snoek, B. L., Fang, S., Zhou, G., Allison, S. D., Blair, J. M., Bridgham, S. D., Burton, A. J., Carrillo, Y., Reich, P. B., Clark, J. S., Classen, A. T., Dijkstra, F. A., Elberling, B., Emmett, B. A., Estiarte, M., Frey, S. D., Guo, J., Harte, J., Jiang, L., Johnson, B. R., Kröel-Dulay, G., Larsen, K. S., Laudon, H., Lavallee, J. M., Luo, Y., Lupascu, M., Ma, L. N., Marhan, S., Michelsen, A., Mohan, J.,
15   Niu, S., Pendall, E., Peñuelas, J., Pfeifer-Meister, L., Poll, C., Reinsch, S., Reynolds, L. L., Schmidt, I. K., Sistla, S., Sokol, N. W., Templer, P. H., Treseder, K. K., Welker, J. M. and Bradford, M. A.: Quantifying global soil carbon losses in response to warming, Nature, 540(7631), 104–108, doi:10.1038/nature20150, 2016.

Davidson, E. A., Belk, E. and Boone, R. D.: Soil water content and temperature as independent or confounded factors controlling soil respiration in a temperate mixed hardwood forest, Glob. Chang. Biol., 4(2), 217–227, doi:10.1046/j.1365-
20   2486.1998.00128.x, 1998.

Davidson, E. A. and Janssens, I. A.: Temperature sensitivity of soil carbon decomposition and feedbacks to climate change., Nature, 440, 165–173, doi:10.1038/nature04514, 2006.

Davidson, E. A., Janssens, I. A. and Luo, Y.: On the variability of respiration in terrestrial ecosystems: moving beyond Q10, Glob. Chang. Biol., 12(2), 154–164, doi:10.1111/j.1365-2486.2005.01065.x, 2006.

25   Dufresne, J. L., Foujols, M. A., Denvil, S., Caubel, A., Marti, O., Aumont, O., Balkanski, Y., Bekki, S., Bellenger, H., Benshila, R., Bony, S., Bopp, L., Braconnot, P., Brockmann, P., Cadule, P., Cheruy, F., Codron, F., Cozic, A., Cugnet, D., de Noblet, N., Duvel, J. P., Ethé, C., Fairhead, L., Fichefet, T., Flavoni, S., Friedlingstein, P., Grandpeix, J. Y., Guez, L., Guilyardi, E., Hauglustaine, D., Hourdin, F., Idelkadi, A., Ghattas, J., Joussaume, S., Kageyama, M., Krinner, G., Labetoulle, S., Lahellec, A., Lefebvre, M. P., Lefevre, F., Levy, C., Li, Z. X., Lloyd, J., Lott, F., Madec, G., Mancip, M., Marchand, M., Masson, S.,
30   Meurdesoif, Y., Mignot, J., Musat, I., Parouty, S., Polcher, J., Rio, C., Schulz, M., Swingedouw, D., Szopa, S., Talandier, C., Terray, P., Viovy, N. and Vuichard, N.: Climate change projections using the IPSL-CM5 Earth System Model: From CMIP3 to CMIP5, Clim. Dyn., 40(9–10), 2123–2165, doi:10.1007/s00382-012-1636-1, 2013.

Dunne, J. P., John, J. G., Shevliakova, E., Stouffer, R. J., Krasting, J. P., Malyshev, S. L., Milly, P. C. D., Sentman, L. T., Adcroft, A. J., Cooke, W., Dunne, K. A., Griffies, S. M., Hallberg, R. W., Harrison, M. J., Levy, H., Wittenberg, A. T., Phillips,

P. J. and Zadeh, N.: GFDL's ESM2 Global Coupled Climate–Carbon Earth System Models. Part II: Carbon System Formulation and Baseline Simulation Characteristics*, J. Clim., 26(7), 2247–2267, doi:10.1175/JCLI-D-12-00150.1, 2013.

Friedlingstein, P., Meinshausen, M., Arora, V. K., Jones, C. D., Anav, A., Liddicoat, S. K. and Knutti, R.: Uncertainties in CMIP5 climate projections due to carbon cycle feedbacks, J. Clim., 27, 511–526, doi:10.1175/JCLI-D-12-00579.1, 2014.

5 Gent, P. R., Danabasoglu, G., Donner, L. J., Holland, M. M., Hunke, E. C., Jayne, S. R., Lawrence, D. M., Neale, R. B., Rasch, P. J., Vertenstein, M., Worley, P. H., Yang, Z.-L. and Zhang, M.: The Community Climate System Model version 4, J. Clim., 24, 4973–4991, doi:10.1175/2011JCLI4083.1, 2011.

van Gestel, N., Shi, Z., van Groenigen, K. J., Osenberg, C. W., Andresen, L. C., Dukes, J. S., Hovenden, M. J., Luo, Y., Michelsen, A., Pendall, E., Reich, P. B., Schuur, E. A. G. and Hungate, B. A.: Predicting soil carbon loss with warming,
10 Nature, 554(7693), E4–E5, doi:10.1038/nature25745, 2018.

Giorgetta, M. A., Jungclaus, J., Reick, C. H., Legutke, S., Bader, J., Böttinger, M., Brovkin, V., Crueger, T., Esch, M., Fieg, K., Glushak, K., Gayler, V., Haak, H., Hollweg, H.-D., Ilyina, T., Kinne, S., Kornblueh, L., Matei, D., Mauritsen, T., Mikolajewicz, U., Mueller, W., Notz, D., Pithan, F., Raddatz, T., Rast, S., Redler, R., Roeckner, E., Schmidt, H., Schnur, R., Segschneider, J., Six, K. D., Stockhause, M., Timmreck, C., Wegner, J., Widmann, H., Wieners, K.-H., Claussen, M.,
15 Marotzke, J. and Stevens, B.: Climate and carbon cycle changes from 1850 to 2100 in MPI-ESM simulations for the Coupled Model Intercomparison Project phase 5, J. Adv. Model. Earth Syst., 5(3), 572–597, doi:10.1002/jame.20038, 2013.

Hashimoto, S., Carvalhais, N., Ito, A., Migliavacca, M., Nishina, K. and Reichstein, M.: Global spatiotemporal distribution of soil respiration modeled using a global database, Biogeosciences, 12(13), 4121–4132, doi:10.5194/bg-12-4121-2015, 2015.

Hicks Pries, C. E., Castanha, C., Porras, R. C. and Torn, M. S.: The whole-soil carbon flux in response to warming, Science
20 (80-.)., 355(6332), 1420–1423, doi:10.1126/science.aal1319, 2017.

Ji, J., Huang, M. and Li, K.: Prediction of carbon exchanges between China terrestrial ecosystem and atmosphere in 21st century, Sci. China Ser. D Earth Sci., 51, 885–898, doi:10.1007/s11430-008-0039-y, 2008.

Jones, C. and Cox, P. M.: Constraints on the temperature sensitivity of global soil respiration from the observed interannual variability in atmospheric CO2, Atmos. Sci. Lett., 2(1–4), 166–172, doi:10.1006/asle.2001.0041, 2001.

25 Jones, C. D., Cox, P. M. and Huntingford, C.: Climate-carbon cycle feedbacks under stabilization: uncertainty and observational constraints, Tellus B Chem. Phys. Meteorol., 58(5), 603–613, doi:10.1111/j.1600-0889.2006.00215.x, 2006.

Knorr, W., Prentice, I. C., House, J. I. and Holland, E. A.: Long-term sensitivity of soil carbon turnover to warming, Nature, 433(7023), 298–301, doi:10.1038/nature03226, 2005.

Krinner, G., Viovy, N., de Noblet-Ducoudré, N., Ogée, J., Polcher, J., Friedlingstein, P., Ciais, P., Sitch, S. and Prentice, I. C.:
30 A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system, Global Biogeochem. Cycles, 19(1), GB1015, doi:10.1029/2003GB002199, 2005.

Lawrence, D. M., Oleson, K. W., Flanner, M. G., Thornton, P. E., Swenson, S. C., Lawrence, P. J., Zeng, X., Yang, Z.-L., Levis, S., Sakaguchi, K., Bonan, G. B. and Slater, A. G.: Parameterization improvements and functional and structural advances in Version 4 of the Community Land Model, J. Adv. Model. Earth Syst., 3(3), M03001, doi:10.1029/2011MS000045,

2011.

Lloyd, J. and Taylor, J. A.: On the temperature dependence of soil respiration, Funct. Ecol., 8, 315–323 [online] Available from: http://www.jstor.org/stable/2389824, 1994.

Lu, M., Zhou, X., Yang, Q., Li, H., Luo, Y., Fang, C., Chen, J., Yang, X. and Li, B.: Responses of ecosystem carbon cycle to
5   experimental warming: a meta-analysis, Ecology, 94(3), 726–738, doi:10.1890/12-0279.1, 2013.

Luo, Y., Ahlström, A., Allison, S. D., Batjes, N. H., Brovkin, V., Carvalhais, N., Chappell, A., Ciais, P., Davidson, E. A., Finzi, A., Georgiou, K., Guenet, B., Hararuk, O., Harden, J. W., He, Y., Hopkins, F., Jiang, L., Koven, C., Jackson, R. B., Jones, C. D., Lara, M. J., Liang, J., McGuire, A. D., Parton, W., Peng, C., Randerson, J. T., Salazar, A., Sierra, C. A., Smith, M. J., Tian, H., Todd-Brown, K. E. O., Torn, M., van Groenigen, K. J., Wang, Y. P., West, T. O., Wei, Y., Wieder, W. R.,
10 Xia, J., Xu, X., Xu, X. and Zhou, T.: Towards More Realistic Projections of Soil Carbon Dynamics by Earth System Models, Global Biogeochem. Cycles, n/a-n/a, doi:10.1002/2015GB005239, 2015.

Mahecha, M. D., Reichstein, M., Carvalhais, N., Lasslop, G., Lange, H., Seneviratne, S. I., Vargas, R., Ammann, C., Arain, M. A., Cescatti, A., Janssens, I. A., Migliavacca, M., Montagnani, L. and Richardson, A. D.: Global convergence in the temperature sensitivity of respiration at ecosystem level, Science (80-.)., 329(5993), 838–840, doi:10.1126/science.1189587,
15 2010.

Moorcroft, P. R., Hurtt, G. C. and Pacala, S. W.: A method for scaling vegetation dynamics: The ecosystem demography model (ED), Ecol. Monogr., 71(4), 557–586, doi:10.2307/3100036, 2001.

Obata, A. and Shibata, K.: Damage of Land Biosphere due to Intense Warming by 1000-Fold Rapid Increase in Atmospheric Methane: Estimation with a Climate–Carbon Cycle Model, J. Clim., 25(24), 8524–8541, doi:10.1175/JCLI-D-11-00533.1,
20 2012.

Oleson, K. W., Lawrence, D. M., Bonan, G. B., Drewniak, B., Huang, M., Koven, C. D., Levis, S., Li, F., Riley, W. J., Subin, Z. M., Swenson, S. C., Thornton, P. E., Bozbiyik, A., Fisher, R., Heald, C. L., Kluzek, E., Lamarque, J.-F., Lawrence, P. J., Leung, L. R., Lipscomb, W., Muszala, S., Ricciuto, D. M., Sacks, W., Sun, Y., Tang, J. and Yang, Z.-L.: Technical Description of version 4.5 of the Community Land Model (CLM)., 2013.
25 Oleson, K. W., Lawrence, D. M., Bonan, G. B., Flanner, M. G., Kluzek, E., Lawrence, P. J., Levis, S., Swenson, S. C. and Thornton, P. E.: Technical Description of version 4.0 of the Community Land Model (CLM). [online] Available from: www.cesm.ucar.edu/models/cesm1.0/clm/CLM4_Tech_Note. pdf, 2010.

Parton, W. J., Schimel, D. S., Cole, C. V. and Ojima, D. S.: Analysis of factors controlling soil organic matter levels in great plains grasslands, Soil Sci. Soc. Am. J., 51, 1173–1179, doi:10.2136/sssaj1987.03615995005100050015x, 1987.
30 Parton, W. J., Stewart, J. W. B. and Cole, C. V.: Dynamics of C, N, P and S in grassland soils: a model, Biogeochemistry, 5, 109–131, doi:10.1007/BF02180320, 1988.

Raddatz, T. J., Reick, C. H., Knorr, W., Kattge, J., Roeckner, E., Schnur, R., Schnitzler, K.-G. G., Wetzel, P. and Jungclaus, J.: Will the tropical land biosphere dominate the climate-carbon cycle feedback during the twenty-first century?, Clim. Dyn., 29(6), 565–574, doi:10.1007/s00382-007-0247-8, 2007.

Sato, H., Itoh, A. and Kohyama, T.: SEIB–DGVM: A new dynamic global vegetation model using a spatially explicit individual-based approach, Ecol. Modell., 200, 279–307, doi:10.1016/j.ecolmodel.2006.09.006, 2007.

Schneck, R., Reick, C. H. and Raddatz, T.: Land contribution to natural CO 2 variability on time scales of centuries, J. Adv. Model. Earth Syst., 5(2), 354–365, doi:10.1002/jame.20029, 2013.

Shevliakova, E., Pacala, S. W., Malyshev, S., Hurtt, G. C., Milly, P. C. D., Caspersen, J. P., Sentman, L. T., Fisk, J. P., Wirth, C. and Crevoisier, C.: Carbon cycling under 300 years of land use change: Importance of the secondary vegetation sink, Global Biogeochem. Cycles, 23, GB2022, doi:10.1029/2007GB003176, 2009.

Sitch, S., Smith, B., Prentice, I. C., Arneth, A., Bondeau, A., Cramer, W., Kaplan, J. O., Levis, S., Lucht, W., Sykes, M. T., Thonicke, K. and Venevsky, S.: Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model, Glob. Chang. Biol., 9(2), 161–185, doi:10.1046/j.1365-2486.2003.00569.x, 2003.

Taylor, K. E., Stouffer, R. J. and Meehl, G. A.: An overview of CMIP5 and the experiment design, Bull. Am. Meteorol. Soc., 93, 485–498, doi:10.1175/BAMS-D-11-00094.1, 2011.

Thornton, P. E.: Regional ecosystem simulation: combining surface- and satellite-based observations to study linkages between terrestrial energy and mass budgets, Univ. of Montana, Missoula., 1998.

Thornton, P. E., Lamarque, J.-F., Rosenbloom, N. A. and Mahowald, N. M.: Influence of carbon-nitrogen cycle coupling on land model response to CO2 fertilization and climate variablity, Global Biogeochem. Cycles, 21, GB4018, doi:10.1029/2006GB002868, 2007.

Thornton, P. E. and Rosenbloom, N. A.: Ecosystem model spin-up: Estimating steady state conditions in a coupled terrestrial carbon and nitrogen cycle model, Ecol. Modell., 189, 25–48, doi:10.1016/j.ecolmodel.2005.04.008, 2005.

Tjiputra, J. F., Roelandt, C., Bentsen, M., Lawrence, D. M., Lorentzen, T., Schwinger, J., Seland, Ø. and Heinze, C.: Evaluation of the carbon cycle components in the Norwegian Earth System Model (NorESM), Geosci. Model Dev., 6(2), 301–325, doi:10.5194/gmd-6-301-2013, 2013.

Todd-Brown, K. E. O., Randerson, J. T., Hopkins, F., Arora, V., Hajima, T., Jones, C., Shevliakova, E., Tjiputra, J., Volodin, E., Wu, T., Zhang, Q. and Allison, S. D.: Changes in soil organic carbon storage predicted by Earth system models during the 21st century, Biogeosciences, 11(8), 2341–2356, doi:10.5194/bg-11-2341-2014, 2014.

Townsend, A. R., Vitousek, P. M., Desmarais, D. J. and Tharpe, A.: Soil carbon pool structure and temperature sensitivity inferred using CO2 and 13CO2 incubation fluxes from five Hawaiian soils, Biogeochemistry, 38(1), 1–17, doi:10.1023/A:1017942918708, 1997.

Volodin, E. M.: Atmosphere-ocean general circulation model with the carbon cycle, Izv. Atmos. Ocean. Phys., 43, 266–280, doi:10.1134/S0001433807030024, 2007.

Watanabe, S., Hajima, T., Sudo, K., Nagashima, T., Takemura, T., Okajima, H., Nozawa, T., Kawase, H., Abe, M., Yokohata, T., Ise, T., Sato, H., Kato, E., Takata, K., Emori, S. and Kawamiya, M.: MIROC-ESM 2010: model description and basic results of CMIP5-20c3m experiments, Geosci. Model Dev., 4, 845–872, doi:10.5194/gmd-4-845-2011, 2011.

Weng, E. and Luo, Y.: Relative information contributions of model vs. data to short- and long-term forecasts of forest carbon

dynamics, Ecol. Appl., 21(5), 1490–1505, doi:10.1890/09-1394.1, 2011.

Wieder, W. R., Allison, S. D., Davidson, E. A., Georgiou, K., Hararuk, O., He, Y., Hopkins, F., Luo, Y., Smith, M. J., Sulman, B., Todd-Brown, K., Wang, Y.-P., Xia, J. and Xu, X.: Explicitly representing soil microbial processes in Earth system models, Global Biogeochem. Cycles, n/a-n/a, doi:10.1002/2015GB005188, 2015a.

5   Wieder, W. R., Cleveland, C. C., Smith, W. K. and Todd-Brown, K.: Future productivity and carbon storage limited by terrestrial nutrient availability, Nat. Geosci., doi:10.1038/ngeo2413, 2015b.

Wu, T., Li, W., Ji, J., Xin, X., Li, L., Wang, Z., Zhang, Y., Li, J., Zhang, F., Wei, M., Shi, X., Wu, F., Zhang, L., Chu, M., Jie, W., Liu, Y., Wang, F., Liu, X., Li, Q., Dong, M., Liang, X., Gao, Y. and Zhang, J.: Global carbon budgets simulated by the Beijing Climate Center Climate System Model for the last century, J. Geophys. Res. Atmos., 118, 1–22,
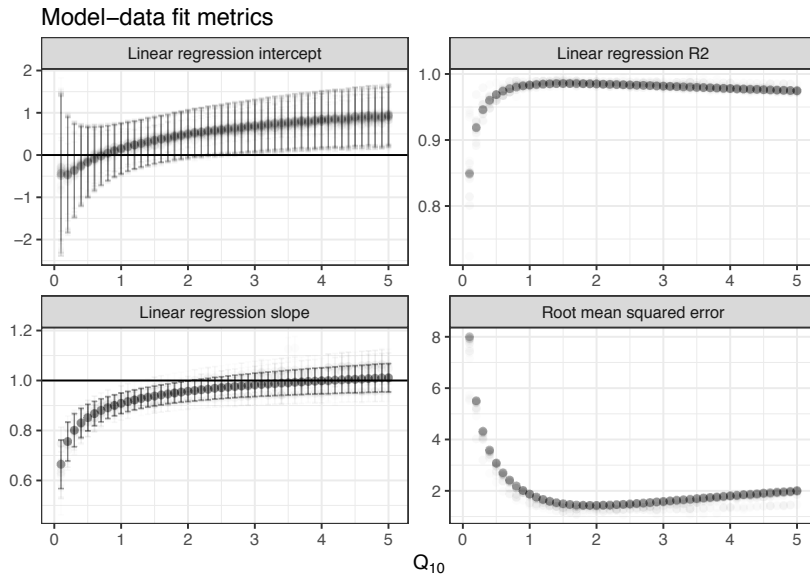10   doi:10.1002/jgrd.50320, 2013.

Figure 1: The model-data fits across different Q$_{10}$ values for random subsets of 34 sites including the root mean squared error, and linear regression metrics R^2, slope, and intercept. The model is take from Eq. 4 ($C_w = C_c Q_{10}^{-\Delta T/10}$). Slope and intercept values are
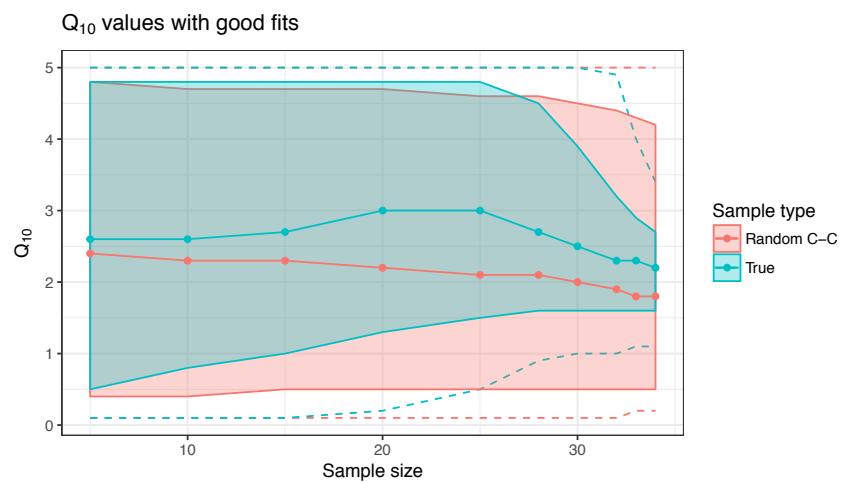15   **shown with 2 standard deviation error bars.**

Figure 2: The Q_{10} with good 1-to-1 model-data fits defined in Figure 1, at 90% confidence interval (band) with minimum and maximum values (dotted line) and median value (solid line), across 10 different sample sizes ranging from 5 to 34, for the original data set (True: blue) and randomized case-control (Random C-C: red).
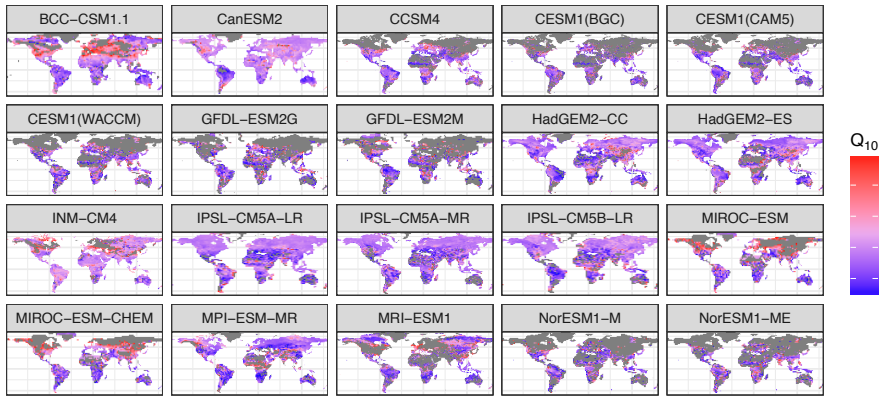
Figure 3: Inferred Q_{10} values from the Earth system models (CMIP5, RCP 8.5). The colour scheme is centered around the field-driven Q_{10} median value of 2.2. Grey indicates non-typical Q_{10} values that were either non-finite, less than 0.5 or greater then 5.
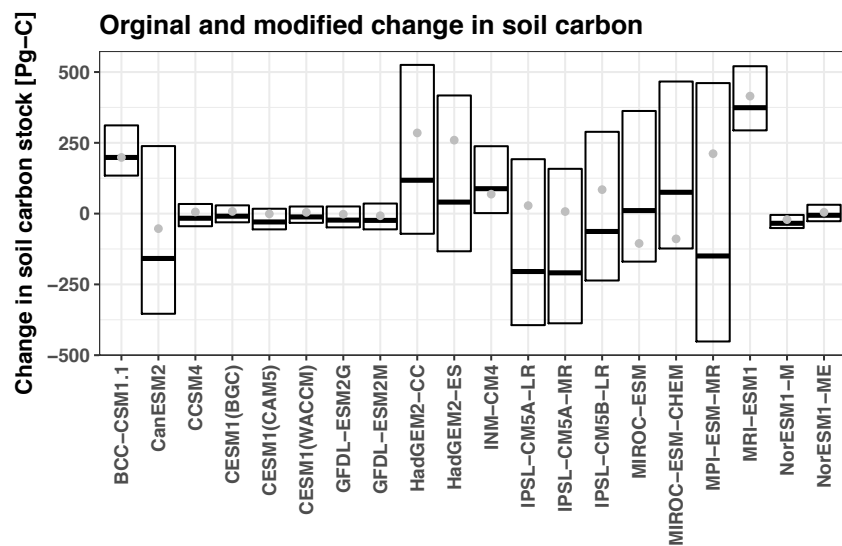
19

Figure 4: Changes in soil carbon stock (10 year means) over the 21st century from Earth system models (RCP 8.5). Grey dots are the original estimates, the open box is the soil carbon loss after the Q_{10} is rescaled using the 2.5%, 50%, and 97.5% quartiles from the field data

5

Model Center	Earth system model	Soil/land carbon sub-model	Number of soil carbon pools	Pool structure	Temperature sensitivity	Citations
BCC	BCC-CSM1.1	BCC_AVIM1.0; AVIM2; CEVSA	8	Cascade	f(T)	(Cao and Woodward, 1998; Ji et al., 2008; Wu et al., 2013)
CCCma	CanESM2	CTEM	2	Cascade	$Q_{10}(T)$	(Arora, 2003; Arora et al., 2011; Arora and Boer, 2010)
NCAR	CCSM4	CLM 4.0; CN	7	Cascade	Lloyd and Taylor	(Gent et al., 2011; Lawrence et al., 2011; Oleson et al., 2010; Thornton et al., 2007; Thornton and Rosenbloom, 2005; Tjiputra et al., 2013)
NSF-DOE-NCAR	CESM1(BGC)					
	CESM1(CAM5)					
	CESM1(WACCM)					
NCC	NorESM1-M					
	NorESM1-ME					
NOAA GFDL	GFDL-ESM2G	LM3.0; ED	2	Independent	f(T)	(Dunne et al., 2013; Moorcroft et al., 2001; Shevliakova et al., 2009)
	GFDL-ESM2M					
MOHC*	HadGEM2-CC	ROTHC	4	Feedback	$Q_{10}=2$	(Coleman and Jenkinson, 1999; Collins et al., 2011)
	HadGEM2-ES					
INM	INM-CM4	LSM 1.0	1	Independent	$Q_{10}=2$	(Bonan, 1996; Volodin, 2007)
IPSL	IPSL-CM5A-LR	ORCHIDEE	5	Feedback	f(T)	(Dufresne et al., 2013; Krinner et al., 2005)
	IPSL-CM5A-MR					
	IPSL-CM5B-LR					
MIROC	MIROC-ESM	SEIB-DGVM	3	Cascade	Lloyd and Taylor	(Sato et al., 2007; Watanabe et al., 2011)
	MIROC-ESM-CHEM					
MPI-M	MPI-ESM-MR	JSBACH	5	Cascade	Lloyd and Taylor	(Giorgetta et al., 2013; Schneck et al., 2013)
MRI	MRI-ESM1	LPJ-DGBM	3	Cascade	Lloyd and Taylor	(Adachi et al., 2013; Obata and Shibata, 2012; Sitch et al., 2003)

Deleted: CENTURY

Deleted: CENTURY

Deleted: f(T)

Deleted: CENTURY

Table 1: This is a summary of the soil decomposition sub-models for the ESMs used in this study and includes the number of pools, structure of the carbon exchange between those pools, temperature sensitivity function, and citations. Temperature sensitivity

function is either denoted as: $f(T)$: borrowed from the Century model (Parton et al., 1987, 1988), **Lloyd and Taylor: taken from the recommended form from** (Lloyd and Taylor, 1994), $Q_{10}(T)$: a temperature dependent Q_{10} as defined by (Arora, 2003), or a Q_{10} parameter for the Q_{10} function as defined in this manuscript.

Deleted: CENTURY

Deleted: f(T): a unique temperature sensitivity function defined by (Coleman & Jenkinson, 1999),

	SOC [Pg-C]	Rel. Inputs	dT [°C]	Q$_{10}$	dSOC [Pg-C]	dSOC Q$_{10}$=1.6	dSOC Q$_{10}$=2.2	dSOC Q$_{10}$=2.7
BCC-CSM1.1	1050	1.40	3.7	2.2	198	312	198	134
CanESM2	1541	1.29	7.1	2.0	-53	239	-158	-354
CCSM4	515	1.32	4.2	1.9	6	34	-16	-45
CESM1(BGC)	515	1.29	3.8	1.9	8	29	-9	-31
CESM1(CAM5)	553	1.30	4.6	1.8	-1	17	-30	-56
CESM1(WACCM)	502	1.32	3.9	1.9	5	25	-12	-33
GFDL-ESM2G	1422	1.41	5.1	1.9	-2	25	-23	-49
GFDL-ESM2M	1278	1.38	4.5	2.0	-8	36	-24	-56
HadGEM2-CC	1122	1.55	8.4	1.9	285	525	118	-71
HadGEM2-ES	1129	1.56	8.3	1.8	259	417	41	-133
INM-CM4	1688	1.27	3.3	2.3	69	238	88	2
IPSL-CM5A-LR	1361	1.48	8.2	1.8	28	192	-205	-394
IPSL-CM5A-MR	1403	1.43	7.6	1.8	7	158	-209	-387
IPSL-CM5B-LR	1274	1.41	7.6	1.9	85	289	-63	-236
MIROC-ESM	2586	1.35	7.2	2.5	-105	363	11	-170
MIROC-ESM-CHEM	2588	1.30	7.3	2.6	-89	467	75	-123
MPI-ESM-MR	3110	1.31	6.3	1.8	212	461	-150	-452
MRI-ESM1	1452	1.52	4.4	2.0	415	521	374	294
NorESM1-M	547	1.31	3.7	1.9	-21	-4	-34	-51
NorESM1-ME	553	1.32	3.6	2.0	5	31	-6	-27
Multi-center mean	1403	1.37	5.4	2.0	88	248	19	-95
Multi-center sd	793	0.09	1.8	0.2	153	191	155	209

Table 2: **Global model summary with multi-center mean and standard deviation for modern soil organic carbon (SOC) stocks [Pg-C], relative shift in soil inputs ($\frac{u_f}{u_m}$), absolute change in soil temperature (dT) [°C], inferred mean of Q$_{10}$ as calculated by grid cell (see Eqn 5), the change in soil organic carbon (dSOC) over the 21st century [Pg-C], and the change in soil organic carbon with rescaled Q$_{10}$ values (1.6, 2.2, and 2.7).**

5

Formatted Table