NIASRA

NATIONAL INSTITUTE FOR APPLIED STATISTICS RESEARCH AUSTRALIA

National Institute for Applied Statistics Research Australia

University of Wollongong, Australia

Working Paper

02-24

CQUESST: A Dynamical Stochastic Framework for Predicting Soil-Carbon Sequestration

Dan Pagendam, Noel Cressie, Jeff Baldock, David Clifford, Ryan Farquharson, Lawrence Murray, Mike Beare, and Denis Curtin

National Institute for Applied Statistics Research Australia, University of Wollongong, Wollongong NSW 2522, Australia T: +61 2 42215076. E: karink@uow.edu.au

CQUESST: A dynamical stochastic framework for predicting soil-carbon sequestration

Dan Pagendam[A], Noel Cressie[B], Jeff Baldock[C], David Clifford[D], Ryan Farquharson[C], Lawrence Murray[E], Mike Beare[F], and Denis Curtin[F]

[A]CSIRO Data61, Australia
[B]National Institute for Applied Statistics Research Australia, University of Wollongong, Australia
[C]CSIRO Agriculture and Food, Australia
[D]Formerly CSIRO
[E]Independent Researcher
[F]The New Zealand Institute for Plant and Food Research Limited, New Zealand

March 11, 2024

Abstract

A statistical framework we call CQUESST (Carbon Quantification and Uncertainty from Evolutionary Soil STochastics), which models carbon sequestration and cycling in soils, is applied to a long-running agricultural experiment that controls for crop type, tillage, and season. CQUESST embeds a dynamic stochastic model of soil carbon, motivated by the deterministic RothC soil-carbon model, within a Bayesian hierarchical statistical model. CQUESST has a coherent framework that acknowledges uncertainties in soil-carbon dynamics, in physical parameters, and in observations. The long-running experiment ran from 2000-2010 and is called the Millenium Tillage Trial; here CQUESST is used to model soil-carbon in six pools, across 42 agricultural plots, and on a monthly time-step for a decade. It is implemented efficiently in the probabilistic programming language `Stan` using its `MapReduce` parallelization. We infer the effectiveness of different experimental treatments for soil-carbon sequestration; and we show how CQUESST can be used for the analysis of designed experiments to draw statistically defensible conclusions about the dependence of soil-carbon decay rates on crop rotations and tillage treatments. These results take into account the uncertainties in the model, resulting in inferences that could inform soil-carbon sequestration decisions and policies.

1 Introduction.

Soil contains the largest store of organic carbon in the terrestrial environment, accounting for more than twice that found in vegetation [Scharlemann et al., 2014]. Agricultural soils comprise 37% of Earth's surface [Smith et al., 2008] and have been seen as potential sinks for sequestering atmospheric carbon by altering agricultural management practices. Practices that conserve or increase the mass of carbon stored in soil are important mitigation strategies to slow down climate change whilst also enhancing agricultural productivity by improving soil fertility, resilience, and sustainability [Baldock and Skjemstad, 1999, Lal, 2002, Smith et al., 2008, Lal, 2011, Baldock et al., 2012, Sommer and Bossio, 2014, Georgiou et al., 2022].

Figure S1.1: Aerial photograph of the Millenium Tillage Trial site in Lincoln New Zealand. Spring tillage is coded using the capital letters "N" (no spring tillage), "M" (minimal spring tillage), and "I" (intensive spring tillage). Autumn tillage was coded using the lower-case letters "n" (no autumn tillage), "m" (minimal autumn tillage) and "i" (intensive autumn tillage), respectively. Tillage treatments were replicated three times, as were two type of no-tillage, namely permanent pasture (PP) and permanent fallow (PF). The position of the treatments within the replicate were allocated randomly.

Long-term agricultural field trials are important sources of experimental data for studying the responses of soil-carbon stocks to different agricultural management practices. Such studies also help in the estimation of carbon-cycling parameters in models that are used for making predictions about carbon stocks at continental scales. The Millennium Tillage Trial (MTT) that took place at Lincoln, New Zealand from 2000 – 2010 was a highly strategic, decade-long field experiment, designed to identify tillage and crop-cover practices for maintaining soil organic carbon (SOC) following the conversion of long-term pasture to arable cropping.

Crops were sown in the spring and the autumn and the MTT applied different tillage treatments and crop-rotation treatments across 42 field-plots, each with dimensions of 9m $\times$ 28m. At every harvest, measurements were collected of three types of soil carbon (particulate organic carbon (POC), resistant organic carbon (ROC), and total organic carbon (TOC)) over time along with the above-ground plant biomass. The MTT design was factorial in form but not fully so (see Section 2.7). Figure 1 shows the layout of the MTT, where each treatment is described by a three character code, explained in Section 2.7. Each of the three-character treatment codes was replicated at three field-plots, and the MTT also included six no-tillage field-plots. Further details of the trial design and conduct can be found in Baldock et al. [2018].

Process-based deterministic models for describing SOC dynamics are ubiquitous in soil-science research. To predict the potential outcomes of agricultural management practices on SOC stocks, studies have employed deterministic models such as RothC [Jenkinson et al., 1990, Li et al., 2016], CENTURY [Parton et al., 1993, Nicoloso et al., 2020], DAYCENT [Del Grosso et al., 2001, Lemma

et al., 2021], APSIM [Luo et al., 2014, O'Leary et al., 2016, Mohanty et al., 2020], EPIC [Causarano et al., 2008, Le et al., 2018], DNDC [Li et al., 1994, 2016] and ICBM [Andrén and Kätterer, 1997, Bolinder et al., 2012], noting that there are others that have been developed for specific applications in other domains (e.g., in forestry; Black et al. [2014], Mao et al. [2019]). It is widely acknowledged that predictions from SOC models are most useful when accompanied by quantification of uncertainty [Ogle et al., 2003, Refsgaard et al., 2007, Post et al., 2008, Juston et al., 2010, Clifford et al., 2014]. To tackle uncertainty in models and estimates, practitioners have on occasions adopted approaches that are statistically questionable or that did not properly quantify the sources of uncertainty. For example, Wang et al. [2005] and Juston et al. [2010] employed the non-statistical Generalized Likelihood Uncertainty Estimation (GLUE) framework of Beven and Binley [1992] and Beven and Freer [2001], an approach that has been criticized for producing "incoherent and inconsistent" results [Christensen, 2004, Mantovan and Todini, 2006, Stedinger et al., 2008]. Other studies, like that of Andrén and Kätterer [1997], Post et al. [2008], and Luo et al. [2014] employed sensitivity analyses in which model parameters were sampled or perturbed in order to study the resulting variability in model output. Sensitivity analyses can identify important parameters [O'Hagan, 2012], but they are not easily adapted to quantifying uncertainties in predictions of latent (unobservable) processes.

In what is to follow, we present a statistical framework for conducting statistical of soil-carbon data collected from field trials, based on a Bayesian hierarchical statistical model (BHM) that allows prediction of the carbon fluxes (with uncertainties) cycling between various latent soil-carbon pools. Further, it allows assesment of hypotheses about the pools. and their parameters, all in a dynamical setting. We call our new framework *CQUESST* (*Carbon Quantification and Uncertainty from Evolutionary Soil STochastics*), which embeds a stochastic, dynamical version of the popular six-pool RothC model [Jenkinson et al., 1990] for soil-carbon dynamics, into a BHM.

Bayesian hierarchical modeling [e.g., Berliner, 1996, Wikle and Berliner, 2007] is a rigorous statistical framework that has gained widespread traction for modeling complex spatio-temporal phenomena in a variety of fields, including climate science [e.g., Kang et al., 2012, Katzfuss et al., 2017, Zhang and Cressie, 2020, Zammit-Mangion et al., 2022], oceanography [e.g., Wikle et al., 2013, Britten et al., 2021] and hydrology [e.g., Pagendam et al., 2014, Li et al., 2020], but its potential in soil-carbon modeling has not been fully realized. At the core of a BHM is the partitioning of a complex, joint-probability distribution into a product of conditional-probability distributions that describe uncertainties in the observed data (via a *data model*), uncertainties in the underlying scientific process (via a *process model*), and uncertainties in the parameters (via a *parameter model*).

In our case, the data model quantifies the measurement errors and the combination of the soil-carbon pools that exist in field data; the process model is a science-driven, dynamical statistical process model that describes the evolution of SOC in multiple latent soil-carbon pools through time; and the parameter model (also known as the prior distribution) captures beliefs, probabilistically, about the values and variability of the parameters in the data model and process model *prior* to observing the data.

In this article, we demonstrate how the CQUESST framework that has a BHM at its core and produces posterior distributions of all "unknowns", allows one to make inferences on the unknown soil processes and model parameters, in order to understand the soil-carbon cycle under tillage and cropping. Our study includes a range of agronomic treatments that alter the balance between inputs and losses of organic carbon from soil. The trial produced seasonal measurements of plant production, SOC stocks, and SOC composition, and we use this data to infer how decay rates (with uncertainties) of SOC stocks vary as a consequence of the tillage-cropping treatments employed. Furthermore, our study quantifies the carbon flux (with uncertainties) from the soil to the atmosphere for each treatment over this long-running field trial, providing insights into which management practices may have the greatest potential for mitigating climate change over long time-horizons.

In recent years, BHMs have been introduced to the soil-carbon-modeling community [e.g., Cable et al., 2009, Clifford et al., 2014, Kim et al., 2014, Ogle et al., 2014, Li et al., 2015, Ogle and Pendall, 2015, Davoudabadi et al., 2021, 2023]. Our earlier paper [Clifford et al., 2014] demonstrates

its application on a simple, single-pool model of soil-carbon dynamics. In what is to follow, this prototype is taken in new directions, particularly the embedding of a multi-pool model that is used in analyzing the MTT referred to above.

In Section 2, we describe the BHM at the core of CQUESST and its component statistical models used for studying the MTT. Section 3 gives a brief description of the Bayesian computational methods used in CQUESST, which are then applied to the MTT data in Section 4. Our statistical analysis compares carbon sequestration across the various treatments used in the MTT, as well as the soil-carbon decay rates as a function of factors included in the designed experiment. Section 5 discusses the results and the importance of the CQUESST framework for addressing the grand challenge of slowing climate change induced by carbon-based greenhouse gases.

2 CQUESST: A Biophysical-Statistical Model of Soil-Carbon Cycling.

This section develops the levels of the BHM that we apply to the Millennium Tillage Trial introduced in Section 1 and analyzed in Section 4.

2.1 RothC v26.3.

In this subsection, we show how biophysical knowledge of soil-carbon cycling through multiple carbon pools can be used to construct dynamical models. Many are deterministic in that identical input gives identical output. However, the uncertainty in processes that govern the cycling should be accounted for, and this we do through statistical models that build on extant biophysical knowledge. Popular models for modeling soil carbon in agricultural systems include RothC [Jenkinson et al., 1990], CENTURY [Parton et al., 1993], DAYCENT [Del Grosso et al., 2001], and APSIM [Luo et al., 2014]. Of these models, RothC focuses solely on carbon, whereas CENTURY, DAYCENT, and APSIM model carbon and nitrogen jointly. The RothC model forms the basis for the soil-carbon component of FullCAM [Richards and Evans, 2004] that is used by the Australian government to produce national greenhouse gas accounts.

RothC is a deterministic, multi-pool model of soil-carbon dynamics; in what follows we describe the current version, v.26.3. Within the RothC model, the total mass of carbon in the soil is partitioned into six 'conceptual pools' of carbon, representing substrate material that differ in chemical composition and decomposability (four pools) and biological material in microbial pools (two pools).

Recall the MTT introduced in Section 1: we now introduce some notation for the analysis that follows in this article. For month t within field-plot i, the addition of carbon to the soil from plant matter is modeled through the time-varying forcing variable, $P_{i,t}$, and the masses of carbon within the six pools are represented as a multivariate stochastic process through the time-varying process vector,

$$\mathbf{Y}_{i,t} = (D_{i,t}, R_{i,t}, F_{i,t}, S_{i,t}, H_{i,t}, I_{i,t})^\top, \tag{S1.1}$$

where $D_{i,t}$ is the carbon stock in decomposable plant material, $R_{i,t}$ is the carbon stock in resistant plant material, $F_{i,t}$ is the carbon stock in fast-decomposing biomass, $S_{i,t}$ is the carbon stock in slow-decomposing biomass, $H_{i,t}$ is the carbon stock in humidified organic matter, and $I_{i,t}$ is the carbon stock in inert organic matter. We note that $S_{i,t}$ and $F_{i,t}$ are distinct from the other pools because these correspond to organic carbon contained in biological microbial material, and that field-plot i will have a unique tillage-cropping treatment assigned to it (Figure S1.1).

The six carbon pools that appear as elements in $\mathbf{Y}_{i,t}$ cycle amongst each other and the atmosphere and are referred to as 'conceptual carbon pools' [Skjemstad et al., 2004] or 'modeled carbon pools' [Poeplau et al., 2013] by soil scientists. In what is to follow, we call them respectively, the D pool, R pool, F pool, S pool, H pool, and I pool. Soil chemists cannot directly measure the carbon

content of these latent, conceptual carbon pools. Observations are made of "measurable fractions", subcomponents of the soil that can be separated mechanically (i.e., using sieves) and chemically [e.g., Baldock et al., 2013]. In the MTT, total organic carbon (TOC), particulate organic carbon (POC), and resistant organic carbon (ROC) were measured. POC was measured as the organic carbon associated with soil particles larger than 50 μm. ROC was measured as the organic carbon associated with polyaromatic structures consistent with, but not necessarily limited to, carbon contained in charcoal using ^{13}C nuclear magnetic resonance. In terms of its six pools, RothC recognises them as combinations of the pools:

$$\begin{aligned} TOC_{i,t} &= D_{i,t} + R_{i,t} + F_{i,t} + S_{i,t} + H_{i,t} + I_{i,t} \\ POC_{i,t} &= R_{i,t} + D_{i,t} + F_{i,t} \\ ROC_{i,t} &= I_{i,t}. \end{aligned} \tag{S1.2}$$

Within RothC, there is no distinction made between the measurement and the process it is measuring; that distinction is important, and it is modeled carefully in CQUESST framework. Also, although measurements of POC have been equated solely with $R_{i,t}$ [e.g., Skjemstad et al., 2004], the actual measurable fraction that is classified as POC also includes decomposable plant matter, $D_{i,t}$, and the fast-decomposing biomass, $F_{i,t}$, which is attached to the plant-material substrate. The rationale for excluding $F_{i,t}$ and $D_{i,t}$ from $POC_{i,t}$ has been that these pools typically only represent small proportions of the POC. However, immediately after the addition of plant matter to the soil, $D_{i,t}$ can become elevated. Furthermore, despite the small contribution of $D_{i,t}$ and $F_{i,t}$ to POC, one should maintain mass-balance and link soil-carbon measurements to their respective carbon pools, as is achieved in Equation (S1.2).

At each time step, some of the carbon within each of the pools (with the exception of $I_{i,t}$, which is inert and hence remains unchanged over time) undergoes microbial decay and is either: (i) transformed into carbon belonging to one of the pools (either the same or a different type); or (ii) released into the atmosphere as carbon dioxide (CO_2) by microbial respiration. It has been estimated that microbial respiration within soils contributes the largest flux of CO_2 from terrestrial ecosystems to the atmosphere [Ogle and Pendall, 2015].

In RothC, the modeling of carbon cycling in soil is expressed mathematically according to the following set of deterministic (i.e., non-stochastic) equations, where we adopt notation that accommodates parameters varying from field-plot to field-plot (indexed by i) within the MTT. For $i = 1, \ldots, 42$,

$$\begin{aligned} D_{i,t+\Delta t} &= D_{i,t} e^{-r_{i,t+\Delta t} \frac{K_D}{12} \Delta t} + p_{P\to D} P_{i,t} + p_{M\to D} M_{i,t} \\ R_{i,t+\Delta t} &= R_{i,t} e^{-r_{i,t+\Delta t} \frac{K_R}{12} \Delta t} + (1 - p_{P\to D}) P_{i,t} + p_{M\to R} M_{i,t} \\ F_{i,t+\Delta t} &= F_{i,t} e^{-r_{i,t+\Delta t} \frac{K_F}{12} \Delta t} + p_{U\to F} U_{i,t} + p_{V\to F} V_{i,t} + p_{M\to F} M_{i,t} \\ S_{i,t+\Delta t} &= S_{i,t} e^{-r_{i,t+\Delta t} \frac{K_S}{12} \Delta t} + p_{U\to S} U_{i,t} + p_{V\to S} V_{i,t} + p_{M\to S} M_{i,t} \\ H_{i,t+\Delta t} &= H_{i,t} e^{-r_{i,t+\Delta t} \frac{K_H}{12} \Delta t} + p_{U\to H} U_{i,t} + p_{V\to H} V_{i,t} + p_{M\to H} M_{i,t} \\ I_{i,t+\Delta t} &= I_{i,t}, \end{aligned} \tag{S1.3}$$

where

$$\begin{aligned} U_{i,t} &\equiv \sum_{X \in \{D,R,F,S\}} X_{i,t} (1 - e^{-r_{i,t+\Delta t} \frac{K_X}{12} \Delta t}) \\ V_{i,t} &\equiv H_{i,t} (1 - e^{-r_{i,t+\Delta t} \frac{K_H}{12} \Delta t}), \end{aligned} \tag{S1.4}$$

and where $\Delta t = 1$ month; K_D, K_R, K_F, K_S, and K_H are respectively the annual decay-rates for carbon in the D, R, F, S, and H pools for field-plot i; $r_{i,t}$ is a decay-rate modifier that is applied to

the decay-rates as a result of changes in climate and ground cover; $U_{i,t}$ is the total carbon decayed from $D_{i,t}$, $R_{i,t}$, $F_{i,t}$, and $S_{i,t}$ pools in field-plot i, that remained in the soil from month t to $t+1$; and $V_{i,t}$ is the carbon decayed from $H_{i,t}$ in field-plot i, that remained in the soil from month t to $t+1$. The role of each of the dynamic-process model parameters is derived from soil-science considerations expanded in Supplemental Material 1.

Equations (S1.3) and (S1.4) describe a dynamical system where mass-balance is preserved through a pool of carbon released from the soil to the atmosphere in the form of the greenhouse gas CO_2. In Mg C ha^{-1} y^{-1}, the *atmospheric flux* of carbon from the ith field-plot over the interval $[0, T]$ can be written as

$$A_i = \frac{12}{T} \sum_{X \in \{D,R,F,S,H,I\}} (X_{i,0} - X_{i,T}), \tag{S1.5}$$

where $X_{i,t}$ represents the carbon stock in Mg C ha^{-1} for month t and pool X. Negative values of A_i represent soil-carbon sequestration. Tillage-cropping treatment τ is applied to three field-plots; suppose they are $i(\tau)_1, i(\tau)_2$, and $i(\tau)_3$. Then the flux of atmospheric carbon for treatment τ in the MTT is simply computed as the area-weighted average of individual fluxes:

$$A^{(\tau)} = \frac{\sum_{j=1}^{3} a_{i(\tau)_j} A_{i(\tau)_j}}{\sum_{j=1}^{3} a_{i(\tau)_j}}, \tag{S1.6}$$

where a_i is the area of field-plot i. Thus, the best tillage-cropping practices should minimize $A^{(\tau)}$; results for the MTT are given in Section 4.2.

2.2 A Dynamical Stochastic Model Based on RothC.

Rather than adopting the deterministic RothC model as an error-free, biophysical representation of soil-carbon dynamics, we developed a dynamical stochastic model based on RothC that recognizes scientific uncertainty. Stochasticity is introduced into each of the transition equations outlined in (S1.3), in the form of additive Gaussian noise on the logarithmic scale [Clifford et al., 2014]. The stochastic equations, analogous to (S1.3), for the multi-pool model considered here have multiplicative errors on the natural scale and are:

$$\begin{aligned}
D_{i,t+\Delta t} &= \exp\{\log(D_{i,t} e^{-r_{i,t+\Delta t}\frac{K_D}{12}\Delta t} + p_{P\to D} P_{i,t} + p_{M\to D} M_{i,t}) + \eta_{D,i,t}\} \\
R_{i,t+\Delta t} &= \exp\{\log(R_{i,t} e^{-r_{i,t+\Delta t}\frac{K_R}{12}\Delta t} + (1-p_{P\to D}) P_{i,t} + p_{M\to R} M_{i,t}) + \eta_{R,i,t}\} \\
F_{i,t+\Delta t} &= \exp\{\log(F_{i,t} e^{-r_{i,t+\Delta t}\frac{K_F}{12}\Delta t} + p_{U\to F} U_{i,t} + p_{V\to F} V_{i,t} + p_{M\to F} M_{i,t}) + \eta_{F,i,t}\} \\
S_{i,t+\Delta t} &= \exp\{\log(S_{i,t} e^{-r_{i,t+\Delta t}\frac{K_S}{12}\Delta t} + p_{U\to S} U_{i,t} + p_{V\to S} V_{i,t}) + p_{M\to S} M_{i,t} + \eta_{S,i,t}\} \\
H_{i,t+\Delta t} &= \exp\{\log(H_{i,t} e^{-r_{i,t+\Delta t}\frac{K_H}{12}\Delta t} + p_{U\to H} U_{i,t} + p_{V\to H} V_{i,t} + p_{M\to H} M_{i,t}) + \eta_{H,i,t}\} \\
I_{i,t+\Delta t} &= I_{i,t},
\end{aligned} \tag{S1.7}$$

where

$$\begin{aligned}
U_{i,t} &= \sum_{X \in \{D,R,F,S\}} X_{i,t}(1 - e^{-r_{i,t+\Delta t}\frac{K_X}{12}\Delta t}), \\
V_{i,t} &= H_{i,t}(1 - e^{-r_{i,t+\Delta t}\frac{K_H}{12}\Delta t}),
\end{aligned}$$

and where $\Delta t = 1$ month, $\eta_{D,i,t} \sim N(-\frac{\sigma_D^2}{2}, \sigma_D^2)$; $\eta_{R,i,t} \sim N(-\frac{\sigma_R^2}{2}, \sigma_R^2)$; $\eta_{F,i,t} \sim N(-\frac{\sigma_F^2}{2}, \sigma_F^2)$; $\eta_{S,i,t} \sim N(-\frac{\sigma_S^2}{2}, \sigma_S^2)$; and $\eta_{H,i,t} \sim N(-\frac{\sigma_H^2}{2}, \sigma_H^2)$ are independent, normally (i.e., Gaussian) distributed random errors. Here, $N(\mu, \sigma^2)$ denotes a normal distribution with mean μ and variance σ^2, and its exponential is a lognormal distribution with mean, $\exp\{\mu + \sigma^2/2\}$. Hence the exponential of

$N(-\sigma^2/2, \sigma^2)$ has a mean of 1, and consequently the means of the state equations in (S1.7) agree with the deterministic equations in (S1.3). To diagnose the assumption of independence between field-plots shown in Figure S1.1, we performed a geostatistical analysis of the MTT data. Our analysis is given in Supplemental Material 2 and shows no evidence of spatial dependence at the scale of the MTT site.

2.3 Measurement Errors in Observed Soil-Carbon Measurable Fractions.

Over the course of the MTT, laboratory measurements of three types of organic carbon were made: particulate organic carbon (POC), resistant organic carbon (ROC), and total organic carbon (TOC). The relationships between these soil-carbon measurable fractions and the latent soil-carbon pools is detailed in (S1.2). Each of these measurements is subject to measurement error, which we modeled in the following way:

$$\begin{aligned} Z_{\mathrm{POC},i,t}|\mathbf{Y}_{i,t}, \sigma^2_{\mathrm{POC}} &\sim LN(\log[D_{i,t} + R_{i,t} + F_{i,t}] - \tfrac{\sigma^2_{\mathrm{POC}}}{2}, \sigma^2_{\mathrm{POC}}) \\ Z_{\mathrm{ROC},i,t}|\mathbf{Y}_{i,t}, \sigma^2_{\mathrm{ROC}} &\sim LN(\log[I_{i,t}] - \tfrac{\sigma^2_{\mathrm{ROC}}}{2}, \sigma^2_{\mathrm{ROC}}) \\ Z_{\mathrm{TOC},i,t}|\mathbf{Y}_{i,t}, \sigma^2_{\mathrm{TOC}} &\sim LN(\log[D_{i,t} + R_{i,t} + F_{i,t} + S_{i,t} + H_{i,t} + I_{i,t}] - \tfrac{\sigma^2_{\mathrm{TOC}}}{2}, \sigma^2_{\mathrm{TOC}}), \end{aligned} \tag{S1.8}$$

where $Z_{\mathrm{POC},i,t}$, $Z_{\mathrm{ROC},i,t}$, and $Z_{\mathrm{TOC},i,t}$ are the observations of POC, ROC, and TOC in field-plot i and month t; recall $\mathbf{Y}_{i,t} \equiv (D_{i,t}, R_{i,t}, F_{i,t}, S_{i,t}, H_{i,t}, I_{i,t})^\top$; and the notation $LN(\mu, \sigma^2)$ refers to a lognormal distribution with parameters μ and σ^2. Here the (measurement) errors are again multiplicative.

As outlined in detail in Section 2.6 below, we formulated priors on measurement-error variance parameters, σ^2_{POC}, σ^2_{ROC}, and σ^2_{TOC} based on laboratory measurements of soil carbon across the field-plots, taken just before the commencement of the MTT (see Table S3.3).

2.4 Soil-Carbon Models as State-Space Models.

The process model and data model detailed in equations (S1.7) and (S1.8), respectively, specify a general class of models known as *state-space models* (e.g., Cressie and Wikle [2011], Ch. 7). A generic linear latent process is:

$$\mathbf{Y}_{t+1} = \mathbf{M}_t \mathbf{Y}_t + \mathbf{g}_t + \boldsymbol{\eta}^\star_t,$$

where $\mathbf{M}_t$ is a matrix whose elements are dictated by parameters that govern the temporal evolution of the system, $\mathbf{g}_t$ is an additive linear term that often corresponds to an external "forcing" of the process (e.g., the addition of carbon from plant material to the soil), and $\boldsymbol{\eta}^\star_t$ is a random vector with zero mean and diagonal covariance matrix that makes the process stochastic. For example, the ith element $\eta^\star_{t,i} \sim N(0, \sigma^2_{\eta,i})$, independently of the other elements. A state-space model also requires the specification of a data model that links observed quantities to the latent-process dynamics. A generic data model is:

$$\mathbf{Z}_t = \mathbf{H}_t \mathbf{Y}_t + \boldsymbol{\epsilon}^\star_t,$$

where $\mathbf{Z}_t$ is a vector of observed quantities, $\mathbf{H}_t$ is a given matrix that links each observed quantity to some linear combination of the state variables in $\mathbf{Y}_t$, and $\boldsymbol{\epsilon}^\star_t$ is a random vector with zero mean and diagonal covariance matrix that represents measurement error. For example, the ith element $\epsilon^\star_{t,i} \sim N(0, \sigma^2_{\epsilon,i})$, independently of the other elements.

However, our component models of the BHM in CQUESST are not linear. The framework has at its core a difference equation that leads to a non-linear state-space model different from the generic one just described. Define the propagator matrix, $\mathbf{M}_t \equiv [\mathbf{M}_{1,t}, \mathbf{M}_{2,t}]$, where

$$\mathbf{M}_{1,t} = \begin{pmatrix} e^{-K_D\Delta_t} & 0 & 0 \\ 0 & e^{-K_R\Delta_t} & 0 \\ p_{U\to F}(1-e^{-K_D\Delta t}) & p_{U\to F}(1-e^{-K_R\Delta t}) & e^{-K_F\Delta_t}+p_{U\to F}(1-e^{-K_F\Delta t}) \\ p_{U\to S}(1-e^{-K_D\Delta t}) & p_{U\to S}(1-e^{-K_R\Delta t}) & p_{U\to S}(1-e^{-K_F\Delta t}) \\ p_{U\to H}(1-e^{-K_D\Delta t}) & p_{U\to H}(1-e^{-K_R\Delta t}) & p_{U\to H}(1-e^{-K_F\Delta t}) \\ 0 & 0 & 0 \end{pmatrix}$$

$$\mathbf{M}_{2,t} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ p_{U\to F}(1-e^{-K_S\Delta t}) & 0 & 0 \\ e^{-K_S\Delta_t}+p_{U\to S}(1-e^{-K_S\Delta t}) & p_{U\to S}(1-e^{-K_H\Delta t}) & 0 \\ p_{U\to H}(1-e^{-K_S\Delta t}) & e^{-K_H\Delta_t}+p_{U\to H}(1-e^{-K_H\Delta t}) & 0 \\ 0 & 0 & 1 \end{pmatrix};$$

define the vector of carbon inputs to the system as,

$$\mathbf{g}_t \equiv \begin{pmatrix} p_{P\to D}P_t + p_{M\to D}M_t \\ p_{P\to R}P_t + p_{M\to R}M_t \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix};$$

and corresponding to the data collected, define the observation matrix as,

$$\mathbf{H}_t = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}.$$

The process model and data model together specify a *hierarchical statistical model*, with which we could employ Bayesian or empirical Bayesian statistical methods for inference on parameters and prediction of latent-process dynamics. From Section 1, a BHM, has a third level of the hierarchy given by the parameter model (or prior) for parameter vector $\boldsymbol{\theta}$; see Section 2.6. Rather than estimating $\boldsymbol{\theta}$ from the data and "plugging" it into the prediction equations, CQUESST has a BHM at its core, and all inferences come from the posterior distribution of the "unknowns" given the data.

2.5 CQUESST as a State-Space Model.

On the natural scale of fluxes, the stochastic soil-carbon model specified in (S1.7) can be considered "almost" linear, but with multiplicative errors. That is,

$$\mathbf{Y}_{t+1} = (\mathbf{M}_t\mathbf{Y}_t + \mathbf{g}_t) \odot \boldsymbol{\eta}_t,$$

where $\boldsymbol{\eta}_t$ is a random vector with kth element, $\eta_{t,k} \sim LN(-\frac{\sigma^2_{\eta,k}}{2}, \sigma^2_{\eta,k})$, and $\odot$ is the Hadamard (elementwise) product of two vectors. Equivalently, this process model can be written as evolving nonlinearly on the log-scale, with additive Gaussian-process noise:

$$\log(\mathbf{Y}_{t+1}) = \log(\mathbf{M}_t\mathbf{Y}_t + \mathbf{g}_t) + \boldsymbol{\eta}^\star_t,$$

where $\boldsymbol{\eta}^\star_t$ is a random vector with kth element, $\eta_{t,k} \sim N(-\frac{\sigma^2_{\eta,k}}{2}, \sigma^2_{\eta,k})$, independently of the other elements. Similarly, the data model specified in (S1.8) is "almost" linear, but with multiplicative measurement error:

$$\mathbf{Z}_t = (\mathbf{H}_t\mathbf{Y}_t) \odot \boldsymbol{\epsilon}_t,$$

where $\boldsymbol{\epsilon}_t$ is a random vector with kth element, $\epsilon_{t,k} \sim LN(-\frac{\sigma^2_{\epsilon,k}}{2}, \sigma^2_{\epsilon,k})$. Again, this can be written on the logarithmic-scale as a non-linear model with additive Gaussian measurement error:

$$\log(\mathbf{Z}_t) = \log(\mathbf{H}_t \mathbf{Y}_t) + \boldsymbol{\epsilon}^\star_t,$$

where $\boldsymbol{\epsilon}^\star_t$ is a random vector with kth element, $\eta_{t,k} \sim N(-\frac{\sigma^2_{\epsilon,k}}{2}, \sigma^2_{\epsilon,k})$, independently of the other elements.

These models are not of a standard form, that would yield to maximum likelihood and Kalman Filtering and their closed-form solutions. Rather than developing approximate solutions with unknown inaccuracies, the soil-carbon state-space model in CQUESST is embedded into a BHM by providing a parameter model (or prior). Doing so has three substantial benefits: (i) specifying prior distributions on the parameters in $\boldsymbol{\theta}$ allows one to draw upon all sources of information available (including expert opinion and past studies; see Section 2.6); (ii) no linear approximations to the process dynamics are needed; and (iii) inference on parameters and state variables can be undertaken simultaneously from the posterior distribution, here through Markov Chain Monte Carlo (MCMC), as in Section 3.

2.6 Priors on Model Parameters and Initial Conditions.

Prior distributions were placed on the initial values in the carbon pools (at month $t = 0$) and on all of the parameters in the model outlined in (S1.7). For simplicity, we denote the complete vector of parameters as $\boldsymbol{\theta}$, which includes all of the scalar parameters listed in Tables S3.1, S3.2, and S3.3 of Supplemental Material 3. It is also necessary to specify prior distributions for the initial states of the latent state variables (soil-carbon pools in each of the field-plots), and these are listed in Table S3.4 of Supplemental Material 3. The prior distributions listed in these tables summarize our beliefs about the likely values for each parameter before observing the MTT data, and they could be broadly classified as either informative or weakly informative.

2.7 Analyzing a Designed Experiment with CQUESST: The Millennium Tillage Trial.

As outlined in Section 1, the Millennium Tillage Trial (MTT) was a long-running agricultural trial that examined the effects on soil-carbon sequestration, of spring and autumn tillage treatments and the presence or absence of winter cover crops. Specifically, the MTT used three levels of a spring tillage treatment, listed in increasing order of intensity: no spring tillage (denoted "N"), minimal spring tillage ("M"), and intensive spring tillage ("I"). The same three levels of tillage were also applied in autumn and denoted by the lower-case letters "n", "m", and "i", respectively. In addition to the spring and autumn tillages, a third treatment was used in the MTT, namely whether or not a cover crop was grown over winter; the presence or absence of the winter cover crop, was coded as a binary variable, taking the values "1" and "0", respectively. Figure S1.1 shows the layout of the MTT, where each experimental treatment is described by a three character code: for example, the two field-plots in the bottom right corner would be "Mm1" and "Mm0". The MTT did not employ a full factorial design; instead it was constrained so that the autumn tillage was at the same or lower intensity as that applied in the spring (e.g., there are no codes with "M" and "i"). This reduced the number of possible treatments from 18 down to 12. Each of the three-character treatment codes was replicated at three field-plots, resulting in a total of 36 field-plots with these types of treatments. In addition, the MTT also included six no-tillage field-plots: these comprised three plots of a permanent pasture (PP) treatment, and three plots of a permanent fallow (PF) treatment that applied herbicide to keep the plots plant-free. In all, the MTT had a total of 42 field-plots.

The CQUESST framework can be used to infer model parameters and soil-carbon trajectories for the MTT, where uncertainty in data, process, and parameters are coherently accounted for. An important scientific question surrounding this dataset is: Do tillage-cropping treatments affect the decay-rate of soil carbon? Soil scientists are aware that crop types and tillage treatments affect the amount of plant material that enters the soil (the main route for soil-carbon sequestration), but they may also affect soil microbial communities responsible for its decomposition. They may also influence the chemical composition and bioavailability of organic carbon cycling through the D, R, F, and S pools, and thus the rate at which decomposition occurs. We have particular interest in determining whether the tillage-cropping practices induced differences in the decay rates of the D, R, and H pools, a question that can be studied by allowing the decomposition rates to be functions of the treatments. We generalize the process model so that the rate parameters in (S1.7), $\{K_X : X = D, R, F, S, H\}$, depend on the tillage-cropping treatment, τ, through its assignment to field plot i, which we have denoted $i(\tau)$. We write

$$K_{X,i(\tau)} = \kappa_X \alpha_\tau, \tag{S1.9}$$

where in (S1.7), i is replaced with $i(\tau)$ and K_X is replaced with $K_{X,i(\tau)}$ for carbon $X \in \{D, R, F, S, H\}$ and treatments, $\tau \in \mathcal{T} = \{\text{PP}, \text{PF}, \text{Nn0}, \ldots, \text{Ii1}\}$. In (S1.9), κ_X is the marginal decay-rate for pool X, and $\alpha_\tau \in (0, \infty)$ is the treatment-specific modifier applied to decay rates for treatment τ. Within each treatment, the decay rates, $K_{X,\tau}$, have the desirable property that the scientifically justified ordering of decay rates for the pools is retained.

3 Bayesian Inference and Computation within the CQUESST Framework.

Here we demonstrate how CQUESST can be used to model soil-carbon data from long-term field trials and to improve our understanding of important latent biogeochemical processes that drive carbon cycling. This is demonstrated in three ways: (i) CQUESST is fitted to the data from the MTT to demonstrate its ability to infer model parameters, latent-process dynamics, and carbon fluxes under different experimental treatments; (ii) uncertainties are captured around these quantities; and (iii) the parameterization of the model can be augmented to see whether carbon-cycling varies under different treatments. In (iii), we specifically evaluate the hypothesis that soil-carbon decay rates are not homogeneous across treatments and should vary depending on the type of production used (e.g., single cropping versus double cropping).

3.1 Assembling the CQUESST Framework.

The BHM which is at the core of CQUESST, uses conditional probability distributions to simplify the complex joint probability distributions encountered in spatio-temporal modeling. In what follows, we denote the series of observations taken at various times in field-plot $i \in \{1, 2, \ldots, 42\}$ as the vectors $\mathbf{Z}_{\text{POC},i}$, $\mathbf{Z}_{\text{ROC},i}$, and $\mathbf{Z}_{\text{TOC},i}$, which we then concatenate for field-plot i into the vector $\mathbf{Z}_i \equiv (\mathbf{Z}_{\text{POC},i}^\top, \mathbf{Z}_{\text{ROC},i}^\top, \mathbf{Z}_{\text{TOC},i}^\top)^\top$. From the MTT data, we wish to obtain estimates of the latent process and parameters, which in the CQUESST framework comes from the posterior distribution obtained through an application of Bayes' Rule. The posterior density function is:

$$p(\mathbf{Y}, \boldsymbol{\theta} | \mathbf{Z}) = \frac{p(\mathbf{Y}, \boldsymbol{\theta}, \mathbf{Z})}{p(\mathbf{Z})} = \frac{p(\mathbf{Z}|\mathbf{Y}, \boldsymbol{\theta}) p(\mathbf{Y}|\boldsymbol{\theta}) p(\boldsymbol{\theta})}{p(\mathbf{Z})}, \tag{S1.10}$$

where $\mathbf{Y} \equiv (\mathbf{Y}_{0,1}^\top, \ldots, \mathbf{Y}_{T,1}^\top, \ldots, \mathbf{Y}_{0,42}^\top, \ldots, \mathbf{Y}_{T,42}^\top)^\top$ is a vector containing all six of the state variables across all 108 months between October 2000 and September 2009, and all 42 field-plots; $\mathbf{Z} \equiv (\mathbf{Z}_1^\top, \ldots, \mathbf{Z}_{42}^\top)^\top$ is the vector containing all of the soil-carbon observations across all field-plots

and observation times; and $\boldsymbol{\theta}$ is the vector of all model parameters outlined in Tables S3.1, S3.2, and S3.3. In (S1.10), $p(\mathbf{Y}, \boldsymbol{\theta}, \mathbf{Z})$ is the joint probability density function of the latent soil-carbon processes, the parameters, and the data. In the BHM, this can be written as the product of three conditional probability density functions, introduced earlier in this paper as the *data model*, the *process model*, and the *parameter model*. The data model is defined by (S1.8), the process model is defined by (S1.7), and the parameter model is given by Tables S3.1 - S3.3, with statistical independence between parameters assumed.

The right-hand side of (S1.10) has a normalizing constant $p(\mathbf{Z})$, which is the marginal probability density function of $\mathbf{Z}$ and is generally intractable. Instead, we use Markov Chain Monte Carlo (MCMC) algorithms to sample from $p(\mathbf{Y}, \boldsymbol{\theta}|\mathbf{Z})$. Specifically, we use an extension of Hamiltonian Monte Carlo (HMC), called the No-U-Turn Sampler (NUTS) [Gelman and Hoffman, 2014], which is implemented as an auto-tuning, 'turnkey' sampling algorithm in `Stan`. Further, `MapReduce` functionality is available in `Stan` for parallelizing computation of $p(\mathbf{Y}, \boldsymbol{\theta}|\mathbf{Z})$ over numerous CPUs on a high-performance computing cluster. In the present application, the HMC in CQUESST, parallelizes the posterior calculation of individual field-plots. The `Stan` code employing `MapReduce` is available at `https://github.com/dpagendam/CQUESST`. When performing Bayesian inference with an MCMC algorithm, it is important to verify that the Markov chain provides a representative set of samples from the posterior distribution $p(\mathbf{Y}, \boldsymbol{\theta}|\mathbf{Z})$, and details of the methods used are provided in Supplemental Material 4.

4 The Millennium Tillage Trial: Inference of Key Components.

We now use the CQUESST framework to analyze the MTT soil-carbon data and demonstrate how it can be used to: (i) fit science-driven dynamics to observed data despite the fact that each carbon pool is latent and cannot be measured directly; (ii) incorporate prior knowledge around parameters of the model and the evolution of the soil-carbon trajectories through time; and (iii) put uncertainty bounds on the "unknowns" in (i) and (ii).

4.1 Posterior Inference for Model Parameters.

Figures S5.1, S5.2, and S5.3 in Supplemental Material 5 show prior and posterior densities for parameters in CQUESST. Detailed discussion of the Bayesian learning seen with these parameters is given there. Prior distributions (in gray) and posterior distributions (in purple) that are near to each other indicate that the observed MTT data did not contain information that greatly changed the prior belief about a parameter. When there is a substantial difference between the (gray) prior and (purple) posterior distributions, or when the prior distribution is so far from the posterior distribution that the prior does not appear in the figure, information in the MTT dataset has drastically changed our prior belief about the plausible values of that parameter.

In some of the cases, the posterior distributions have shifted substantially away from the prior (e.g., κ_D and κ_R). Parameters where there were less-substantial differences between prior and posterior distributions include the decay-rate parameters κ_F, κ_S, and κ_H, whilst there was no apparent learning for many of the parameters in Figure S5.3 that govern the proportions of decomposed carbon that are routed to other pools. We show that the posterior distribution for κ_D favors slightly lower values than the prior distribution, κ_R favours larger values in the posterior distribution compared to the prior distribution, and κ_F, κ_S and κ_H show little change. In the case of κ_F and κ_S this is unsurprising, since the observations, which relate to the sum of multiple pools would not be expected to offer much information about the dynamics of the relatively small (in terms of carbon stock) F and S pools.

From Figure S5.2, it is clear that posterior distributions for variance parameters (process noise and measurement errors) shifted to lower ranges of values than the priors. This indicated that measurement errors were lower than we initially expected and that the process dynamics exhibited more determinism than we had initially expected. The latter may have been due to our use of α_τ to obtain a more 'tailored' model fit to each treatment in the MTT than if we had used a global set of decay rates across all treatments.

4.2 Soil-Carbon Dynamics.

Sampled trajectories of soil-carbon pools from the posterior distribution were used to construct Figures S1.2 and S1.3. Figure S1.2 shows how CQUESST is able to make inferences on the latent soil-carbon pools given in (S1.7), even though the observations were on aggregated subsets of these given in (S1.8). In Figure S1.2, we see that the D pool rises and falls each time a crop is sown as plant material enters the soil and is rapidly decomposed. Similarly, we observe small peaks in the R pool over time that correspond to resistant plant material entering the soil and decomposing more slowly. The F and S pools start at low masses at the beginning of the MTT and rise over its duration because of the continual input of substrate for the microorganisms that comprise these pools. Overall, the F and S pools only represent a small amount of the total carbon stock. We observe that the H pool varies slowly over time, since this pool has a slow decay rate and remains relatively stable. Finally, the I pool remains completely stable over time, since this pool represents inert carbon in the soil.

From Figure S1.3, we see that the observed data agree well with the aggregated pools to which the observations correspond. POC and TOC measurements include material that is in the D and R pools and can therefore exhibit peaks associated with the annual addition of plant material from crops into the soil. In contrast, the measurements of ROC are more stable, since these are solely measurements of the inert I pool. Note that the trajectory for TOC is the sum of all trajectories in Figure S1.2. This can be used to make inferences about how much carbon was accumulated (negative flux) or lost to the atmosphere (positive flux) over the course of the MTT. Plots of the posterior distribution of the total soil-carbon flux, $A^{(\tau)}$, for each of the 14 tillage-cropping treatments in the MTT are shown in Figure S1.4, which is discussed in more detail below.

4.3 Answering Scientific Questions with CQUESST: Do the decay rates of conceptual carbon pools vary as functions of cropping and tillage?

Posterior distributions for multiplicative treatment effects, α_τ, defined in (S1.9), are shown in Figure S1.5. We note that there is substantial variability between the treatments, indicating that soil-carbon decay-rates are not uniform under all tillage-cropping management practices. This is in contrast to how soil-carbon dynamics are typically modeled in practice and indeed are modeled in the deterministic RothC discussed in Section 1. Of particular note in our results are the no-tillage PF and PP treatments, which are both at the lower range of values for α_τ, indicating that decay rates under these two treatments are low. Figure S1.4 shows that this multiplicative effect does not necessarily dictate the carbon-sequestration potential of a treatment, since sequestration also depends upon the amount of plant material cultivated, with root material (and potentially stubble) entering the soil. We observe that for the permanent pasture (PP) treatment, there is a high probability of negative carbon flux (i.e., low probability of losing soil carbon to the atmosphere). In contrast, permanent fallow (PF) which is a chemical fallow with no addition of plant material to the soil, showed very strong evidence of positive carbon flux to the atmosphere (i.e., poor carbon-sequestration). Of the other MTT treatments, the spring intensive-tillage treatments, In0, In1, and Im0, showed the next highest probabilities of soil-carbon sequestration; and consistently, no-tillage (Nn0 and Nn1) and minimum-tillage (Mm0 and Mm1) treatments showed low probability of sequestration. Figure S1.4 suggests that active tillage treatments that employ a rotation of intensive tillage in the spring, have

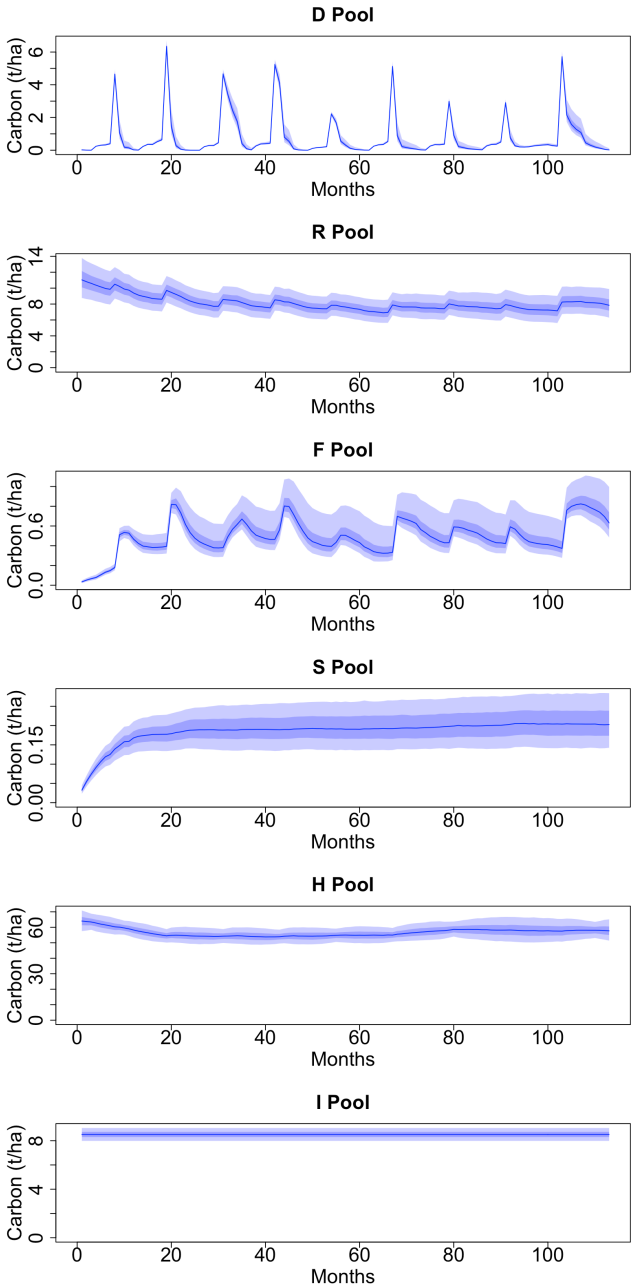

Figure S1.2: Sampled trajectories of latent soil-carbon pools in the second field-plot with tillage treatment Mn0. Central blue lines show the posterior median, dark blue ribbons span the 50% posterior two-sided intervals, and light blue ribbons span the 90% posterior two-sided intervals. The six soil-carbon pools are: decomposable plant matter (D), resistant plant matter (R), fast (F) and slow (S) decomposers, humus (H), and inert (I) material.

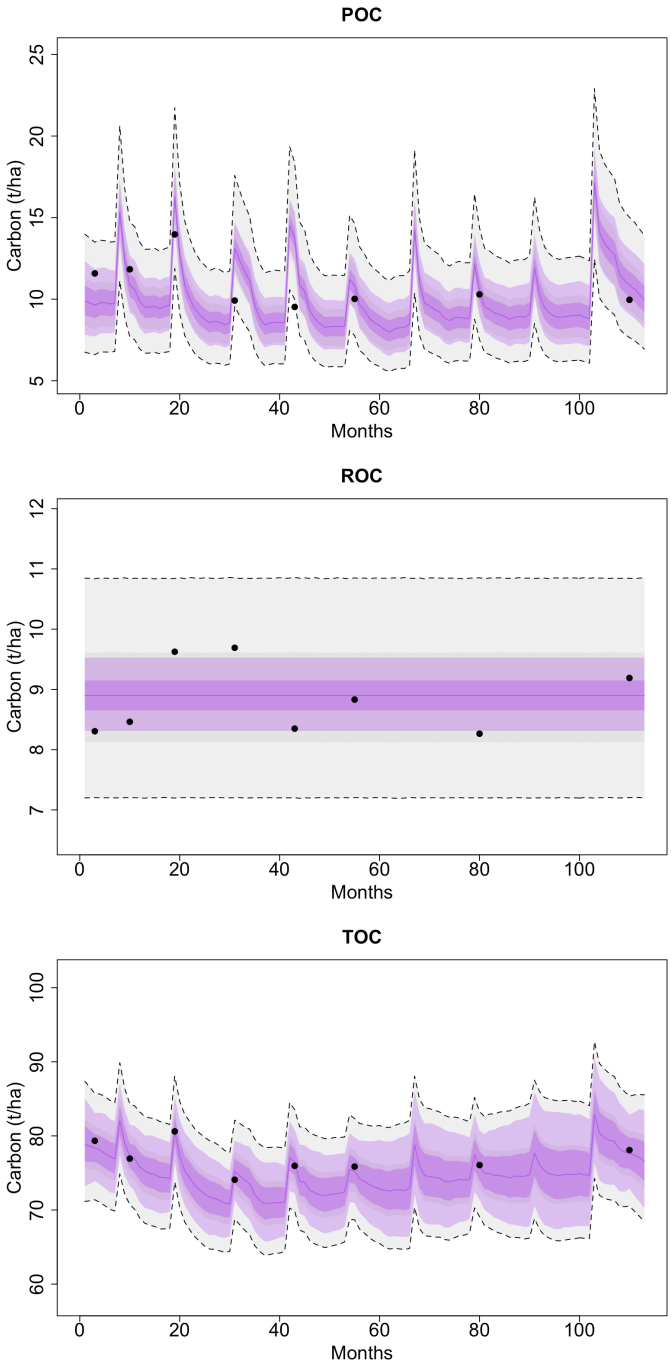

Figure S1.3: Sampled trajectories for the combined, latent soil-carbon pools for each observation type (see equation (S1.2)) in the second field-plot with tillage treatment Mn0. Black circles show observations, central purple lines show the posterior median of the combined latent pools, dark purple ribbons span the 50% posterior two-sided intervals of the combined latent pools, and light purple ribbons span the 90% posterior two-sided intervals of the combined latent pools. Dashed lines show the 90% posterior two-sided intervals for the observations with the presence of measurement error taken into account.

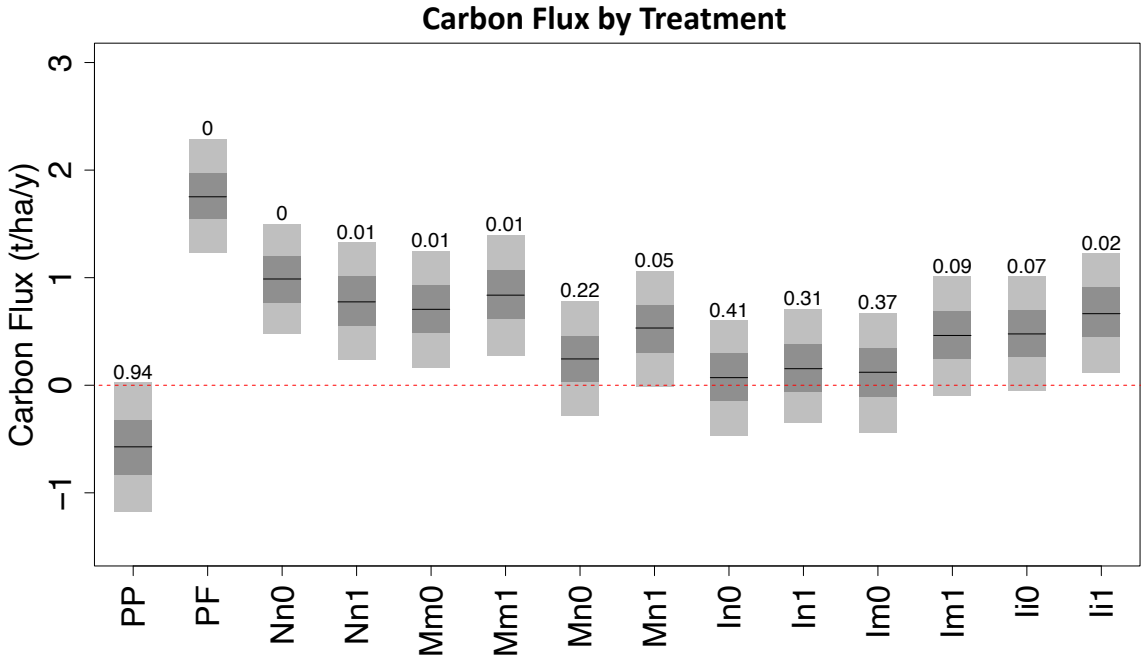

Figure S1.4: Posterior distribution of fluxes of soil-carbon between the start and end of the Millennium Tillage Trial, averaged over three field-plots for each treatment (see equation (S1.6)). Negative values indicate carbon sequestration. Darker rectangles span the 50% posterior two-sided interval, and horizontal lines show the posterior median. Lighter coloured rectangles span the 90% posterior two-sided interval. Numbers above rectangles give the posterior probability that the atmospheric carbon flux for each treatment was less than zero (i.e., soil-carbon sequestration). Three character codes for treatments are outlined in Figure S1.1.

greater soil-carbon sequestration potential than those that employ no-tillage or minimum tillage in the spring.

5 Discussion.

The Bayesian hierarchical statistical model that we have implemented through the CQUESST framework provides a powerful tool to better understand a phenomenon of global importance, namely the cycling of carbon stocks in agricultural soils. CQUESST provides soil scientists with a tool that can be used for combining data, expert knowledge, and biogeochemical-process dynamics in a statistically rigorous way. In addition, we have shown how CQUESST can also harness all of these sources of information when analyzing designed experiments such as the MTT. An important attribute of using Bayesian hierarchical statistical modelling is that it allows these inferences to be carried out, in the presence of uncertainty in observed data, in our understanding of the process dynamics, and in the model's parameters. These uncertainties are critical to acknowledge in order to draw statistically rigorous scientific conclusions.

Analysis of the MTT using CQUESST involved a high-dimensional state-space model whose state vector was of length 252 (42 fields, each with 6 pools) and where dynamics evolved over 108 time steps (months). Our implementation of CQUESST harnesses the computational power of `Stan`, and it is noteworthy that the inclusion of the `MapReduce` functionality in `Stan` made much of our

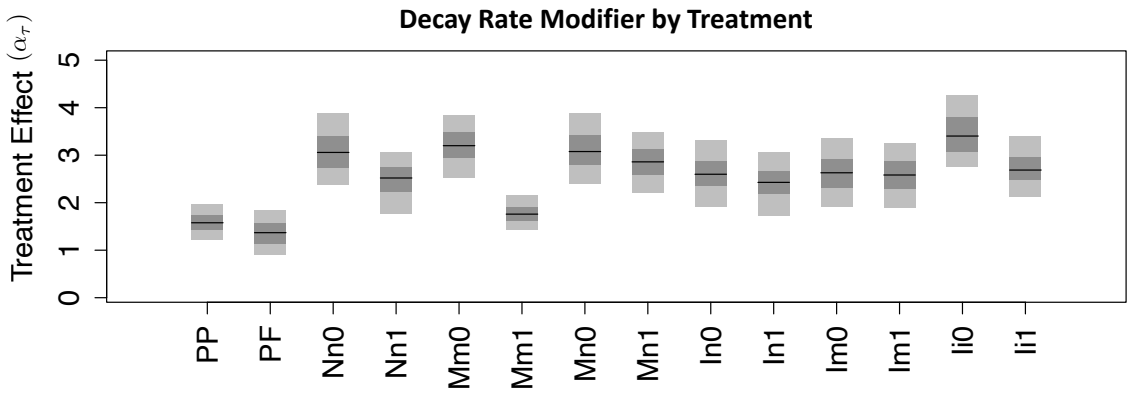

Figure S1.5: Posterior distributions for the multiplicative decay-rate modifier (α_τ) for each treatment in the Millennium Tillage Trial. Darker rectangles span the 50% posterior two-sided interval, and horizontal lines show the posterior median. Lighter coloured rectangles span the 90% posterior two-sided interval.

statistical analysis parallelizable and hence very efficient. `Stan` is a key enabler of analyses of the type presented here, where data have been collected at many independent sites but share some underlying parameters governing process dynamics, which would allow detailed analyses of carbon stocks to be made at broad geographical scales (e.g., national soil carbon accounting).

The inferences we made about soil-carbon cycling from the MTT dataset using CQUESST provide valuable insights about soil-organic-carbon stocks and how agricultural practices can affect them. Of particular note were the large carbon fluxes to the atmosphere using PF compared to other treatments, which is consistent with the results previously reported by Curtin et al. [2022]. This lends strength to the observations made by other researchers [e.g., Halvorson et al., 2002] that the use of crop-fallow rotations may lead to net carbon losses from the soil. Furthermore, we observed that PP (i.e., permanent-pasture, no tillage) had a continual integration of plant organic carbon into the soil with large negative carbon fluxes (i.e., maximum sequestration of carbon). Apart from PP and PF, we observed the highest posterior probabilities of soil carbon sequestration for treatments In0, In1 and Im0 that employed one round of intensive tillage per year. Related to this, the work of Cai et al. [2022] suggests that over time scales less than about 14 years, no tillage agricultural treatments may show lower carbon stocks compared to those that employ conventional tillage, but that differences diminish beyond 14 years. Therefore, it is possible that had the MTT run for longer, it would have shown less difference in the carbon flux between these treatments.

Our use of CQUESST to perform the analysis of a designed experiment (the MTT), allowed us to model and examine cropping-specific decay rates in the different carbon pools. Our results indicate that the decay rates of carbon pools vary as a function of management practices (i.e., tillage intensity and winter-cover-crop use), rather than being static parameters consistent over all farms. This is an important finding that we hope will guide how soil carbon is modeled in the future. Despite treatment effects suggesting that decay rates were higher in treatments without cover crops, we noted higher probabilities of negative carbon flux when comparing each treatment without cover crop to that with cover crop. This may suggests the presence of additional process dynamics that are not yet present in the deterministic mechanisms of RothC, but that are captured through the stochastic analogue of RothC employed in CQUESST. Further experimental work is required to better understand the specifics of such mechanisms.

Understanding the complex biogeochemical cycles that take place in agricultural soils is critical

to finding strategies to sequester carbon on the 37% of Earth's surface that is used for agricultural production. This will be advantageous in mitigating climate change as well as improving the productivity of agricultural soils. CQUESST provides a framework with Bayesian hierarchical modeling at its core to model complex soil-carbon dynamics in agricultural systems. We envisage that this will enable scientists to extract important information, with uncertainties quantified, from agricultural experiments to help nations take steps towards net-zero carbon emissions.

Acknowledgements

Pagendam was supported by CSIRO's Digiscape Future Science Platform and Machine Learning and Artificial Intelligence Future Science Platform. Cressie was supported by the Australian Research Council (ARC) Discovery Project DP190100180. Further, this material is based on work supported by the Air Force Office of Scientific Research under award number FA2386-23-1-4100. The MTT data set was collected with scientific and technical support from Trish Fraser, Richard Gillespie, Weiwen Qiu, Chris Dunlop, Peg Gosden and Sarah Glasson. The MTT trial was completed under the New Zealand Institute for Plant and Food Research Limited's Sustainable Agro-ecosystems (SAE) programme, with funding from the Strategic Science Investment Fund (SSIF), contract C11X1702. Helpful comments from Gerard Heuvelink of the University of Wageningen are gratefully acknowledged.

References

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

J. Baldock and J. Skjemstad. Soil organic carbon / soil organic matter. In K. I. Peverill, L. A. Sparrow, and D. J. Reuter, editors, *Soil Analysis: An Interpretation Manual*, pages 159–170. CSIRO Publishing, 1999.

J. A. Baldock, I. Wheeler, N. McKenzie, and A. McBrateny. Soils and climate change: potential impacts on carbon stocks and greenhouse gas emissions, and future research for Australian agriculture. *Crop and Pasture Science*, 63(3):269–283, 2012.

J. A. Baldock, J. Sanderman, L. M. Macdonald, A. Puccini, S. S., and J. McGowan. Quantifying the allocation of soil organic carbon to biologically significant fractions. *Soil Research*, 51:561–576, 2013.

J. A. Baldock, M. Beare, D. Curtin, and B. Hawke. Stocks, composition and vulnerability to loss of soil organic carbon predicted using mid-infrared spectroscopy. *Soil Research*, 56(5):468 – 480, 2018.

L. Berliner. Hierarchical Bayesian time series models. In K. Hanson and R. Silver, editors, *Maximum Entropy and Bayesian Methods*, volume 79 of *Fundamental Theories of Physics*, pages 15–22. Springer Netherlands, 1996. ISBN 978-94-010-6284-8.

K. Beven and A. Binley. The future of distributed models: model calibration and uncertainty prediction. *Hydrological Processes*, 6(3):279–298, 1992.

K. Beven and J. Freer. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *Journal of Hydrology*, 249(1-4):11 – 29, 2001.

K. Black, R. E. Creamer, G. Xenakis, and S. Cook. Improving forest soil carbon models using spatial data and geostatistical approaches. *Geoderma*, 232-234:487–499, 2014.

M. A. Bolinder, T. Kätterer, O. Andrén, and L. E. Parent. Estimating carbon inputs to soil in forage-based crop rotations and modeling the effects on soil carbon dynamics in a Swedish long-term field experiment. *Canadian Journal of Soil Science*, 92(6):821–833, 2012.

G. L. Britten, Y. Mohajerani, L. Primeau, M. Aydin, C. Garcia, W.-L. Wang, B. Pasquier, B. B. Cael, and F. W. Primeau. Evaluating the benefits of Bayesian hierarchical methods for analyzing heterogeneous environmental datasets: A case study of marine organic carbon fluxes. *Frontiers in Environmental Science*, 9, 2021. ISSN 2296-665X. doi: 10.3389/fenvs.2021.491636. URL `https://www.frontiersin.org/articles/10.3389/fenvs.2021.491636`.

J. Cable, K. Ogle, A. Tyler, M. Pavao-Zuckerman, and T. Huxman. Woody plant encroachment impacts on soil carbon and microbial processes: results from a hierarchical Bayesian analysis of soil incubation data. *Plant and Soil*, 320(1-2):153–167, 2009.

A. Cai, T. Han, T. Ren, J. Sanderman, Y. Rui, B. Wang, P. Smith, M. Xu, and Y. Li. Declines in soil carbon storage under no tillage can be alleviated in the long run. *Geoderma*, 425:116028, 2022.

H. J. Causarano, P. C. Doraiswamy, G. W. McCarty, J. L. Hatfield, S. Milak, and A. J. Stern. EPIC modeling of soil organic carbon sequestration in croplands of Iowa. *Journal of Environmental Quality*, 37:1345–1353, 2008.

S. Christensen. A synthetic groundwater modelling study of the accuracy of GLUE uncertainty intervals. *Nordic Hydrology*, 35:45–59, 2004.

D. Clifford, D. Pagendam, J. Baldock, N. Cressie, R. Farquharson, M. Farrell, L. Macdonald, and L. Murray. Rethinking soil carbon modelling: a stochastic approach to quantify uncertainties. *Environmetrics*, 25(4):265–278, 2014.

N. Cressie and C. K. Wikle. *Statistics for Spatio-Temporal Data*. John Wiley & Sons, Hoboken, New Jersey, USA, 2011.

D. Curtin, M. Beare, and W. Qiu. Hot water extractable carbon in whole soil and particle size fractions isolated from soils under contrasting land-use treatments. *Soil Research*, 60(8):772 – 781, 2022.

M. J. Davoudabadi, D. E. Pagendam, C. Drovandi, J. Baldock, and G. White. Advanced Bayesian approaches for state-space models with a case study on soil carbon sequestration. *Environmental Modelling and Software*, 136:104919, 2021.

M. J. Davoudabadi, D. Pagendam, C. Drovandi, J. Baldock, and G. White. The effect of biologically mediated decay rates on modelling soil carbon sequestration in agricultural settings. *Environmental Modelling & Software*, 168:105786, 2023.

S. Del Grosso, W. Parton, A. Mosier, M. Hartman, J. Brenner, D. Ojima, and D. Schimel. Simulated interactions of carbon dynamics and nitrogen trace gas fluxes using the DAYCENT model. In M. Shaffer, L. Ma, and S. Hansen, editors, *Modeling Carbon and Nitrogen Dynamics for Soil Management*, pages 3–24. Lewis Publishers, Boca Roton, Florida, USA, 2001.

A. Gelman and M. Hoffman. The No-U-Turn Sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research*, 15:1593–1623, 2014.

K. Georgiou, R. B. Jackson, O. Vindušková, R. Z. Abramoff, A. Ahlström, W. Feng, J. W. Harden, A. F. A. Pellegrini, H. W. Polley, J. L. Soong, W. J. Riley, and M. S. Torn. Global stocks and capacity of mineral associated soil organic carbon. *Nature Communications*, 13:3797, 2022.

A. D. Halvorson, B. J. Wienhold, and A. L. Black. Tillage, nitrogen, and cropping system effects on soil carbon sequestration. *Soil Science Society of America Journal*, 66(3):906–912, 2002. doi: https://doi.org/10.2136/sssaj2002.9060. URL `https://acsess.onlinelibrary.wiley.com/doi/abs/10.2136/sssaj2002.9060`.

D. S. Jenkinson, S. P. S. Andrew, J. M. Lynch, M. J. Goss, and P. B. Tinker. The turnover of organic carbon and nitrogen in soil [with discussion]. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 329(1255):361–368, 1990.

J. Juston, O. Andrén, T. Kätterer, and P.-E. Jansson. Uncertainty analyses for calibrating a soil carbon balance model to agricultural field trial data in Sweden and Kenya. *Ecological Modelling*, 221(16):1880 – 1888, 2010.

E. L. Kang, N. Cressie, and S. R. Sain. Combining outputs from the North American Regional Climate Change Assessment Program by using a Bayesian hierarchical model. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 61(2):291–313, 2012.

M. Katzfuss, D. Hammerling, and R. L. Smith. A Bayesian hierarchical model for climate change detection and attribution. *Geophysical Research Letters*, 44(11):5720–5728, 2017.

Y. Kim, K. Nishina, N. Chae, S. J. Park, Y. J. Yoon, and B. Y. Lee. Constraint of soil moisture on CO_2 efflux from tundra lichen, moss, and tussock in Council, Alaska, using a hierarchical Bayesian model. *Biogeosciences*, 11(19):5567–5579, 2014.

R. Lal. Soil carbon dynamics in cropland and rangeland. *Environmental Pollution*, 116(3):353 – 362, 2002.

R. Lal. Soil health and climate change: an overview. In B. P. Singh, A. L. Cowie, and K. Y. Chan, editors, *Soil Health and Climate Change*, volume 29 of *Soil Biology*, pages 3–24. Springer, Berlin, Germany, 2011. ISBN 978-3-642-20255-1.

K. N. Le, M. K. Jha, M. R. Reyes, J. Jeong, L. Doro, P. W. Gassman, L. Hok, J. C. de Moraes Sá, and S. Boulakia. Evaluating carbon sequestration for conservation agriculture and tillage systems in Cambodia using the EPIC model. *Agriculture, Ecosystems and Environment*, 251:37–47, 2018. ISSN 0167-8809. doi: https://doi.org/10.1016/j.agee.2017.09.009. URL `https://www.sciencedirect.com/science/article/pii/S0167880917304097`.

B. Lemma, S. Williams, and K. Paustian. Long term soil carbon sequestration potential of smallholder croplands in southern Ethiopia with DAYCENT model. *Journal of Environmental Management*, 294:112893, 2021. ISSN 0301-4797. doi: https://doi.org/10.1016/j.jenvman.2021.112893. URL `https://www.sciencedirect.com/science/article/pii/S0301479721009555`.

C. Li, S. Frolking, and R. Harriss. Modeling carbon biogeochemistry in agricultural soils. *Global Biogeochemical Cycles*, 8(3):237–254, 1994.

J. Li, Q. Zhou, and W. W.-G. Yeh. A Bayesian hierarchical model for estimating the statistical parameters in a three-parameter log-normal distribution for monthly average streamflows. *Journal of Hydrology*, 591:125265, 2020. ISSN 0022-1694. doi: https://doi.org/10.1016/j.jhydrol.2020.125265. URL `https://www.sciencedirect.com/science/article/pii/S0022169420307253`.

S. Li, J. Li, C. Li, S. Huang, X. Li, S. Li, and Y. Ma. Testing the RothC and DNDC models against long-term dynamics of soil organic carbon stock observed at cropping field soils in North China. *Soil and Tillage Research*, 163:290–297, 2016. ISSN 0167-1987. doi: https://doi.org/10.1016/j.still.2016.07.001. URL https://www.sciencedirect.com/science/article/pii/S0167198716301210.

X. Li, K. Ishikura, C. Wang, J. Yeluripati, and R. Hatano. Hierarchical Bayesian models for soil CO_2 flux using soil texture: a case study in central Hokkaido, Japan. *Soil Science and Plant Nutrition*, 61(1):116–132, 2015.

Z. Luo, E. Wang, I. R. Fillery, L. M. Macdonald, N. Huth, and J. Baldock. Modelling soil carbon and nitrogen dynamics using measurable and conceptual soil organic matter pools in APSIM. *Agriculture, Ecosystems & Environment*, 186:94 – 104, 2014.

P. Mantovan and E. Todini. Hydrological forecasting uncertainty assessment: Incoherence of the GLUE methodology. *Journal of Hydrology*, 330(1-2):368 – 381, 2006.

Z. Mao, D. Derrien, M. Didion, J. Liski, T. Eglin, M. Nicolas, M. Jonard, and L. Saint-André. Modeling soil organic carbon dynamics in temperate forests with yasso07. *Biogeosciences*, 16(9): 1955–1973, 2019.

M. Mohanty, N. K. Sinha, J. Somasundaram, S. S. McDermid, A. K. Patra, M. Singh, A. Dwivedi, K. S. Reddy, C. S. Rao, M. Prabhakar, K. Hati, P. Jha, R. Singh, R. Chaudhary, S. N. Kumar, P. Tripathi, R. C. Dalal, D. S. Gaydon, and S. Chaudhari. Soil carbon sequestration potential in a vertisol in central India - results from a 43-year long-term experiment and APSIM modeling. *Agricultural Systems*, 184: 102906, 2020. ISSN 0308-521X. doi: https://doi.org/10.1016/j.agsy.2020.102906. URL https://www.sciencedirect.com/science/article/pii/S0308521X20307678.

R. S. Nicoloso, T. J. C. Amado, and C. W. Rice. Assessing strategies to enhance soil carbon sequestration with the DSSAT-CENTURY model. *European Journal of Soil Science*, 71(6):1034–1049, 2020.

K. Ogle and E. Pendall. Isotope partitioning of soil respiration: A Bayesian solution to accommodate multiple sources of variability. *Journal of Geophysical Research: Biogeosciences*, 120(2):221–236, 2015.

K. Ogle, C. Tucker, and J. M. Cable. Beyond simple linear mixing models: process-based isotope partitioning of ecological processes. *Ecological Applications*, 24(1):181–195, 2014.

S. M. Ogle, F. Jay Breidt, M. D. Eve, and K. Paustian. Uncertainty in estimating land use and management impacts on soil organic carbon storage for US agricultural lands between 1982 and 1997. *Global Change Biology*, 9(11):1521–1542, 2003.

A. O'Hagan. Probabilistic uncertainty specification: overview, elaboration techniques and their application to a mechanistic model of carbon flux. *Environmental Modelling & Software*, 36:35 – 48, 2012.

G. J. O'Leary, D. L. Liu, Y. Ma, F. Y. Li, M. McCaskill, M. Conyers, R. Dalal, S. Reeves, K. Page, Y. P. Dang, and F. Robertson. Modelling soil organic carbon 1. Performance of APSIM crop and pasture modules against long-term experimental data. *Geoderma*, 264: 227–237, 2016. ISSN 0016-7061. doi: https://doi.org/10.1016/j.geoderma.2015.11.004. URL https://www.sciencedirect.com/science/article/pii/S0016706115301208.

D. E. Pagendam, P. M. Kuhnert, W. B. Leeds, C. K. Wikle, R. Bartley, and E. E. Peterson. Assimilating catchment processes with monitoring data to estimate sediment loads to the Great Barrier Reef. *Environmetrics*, 25(4):214–229, 2014.

W. J. Parton, J. M. O. Scurlock, D. S. Ojima, T. G. Gilmanov, R. J. Scholes, D. S. Schimel, T. Kirchner, J.-C. Menaut, T. Seastedt, E. Garcia Moya, A. Kamnalrut, and J. I. Kinyamario. Observations and modeling of biomass and soil organic matter dynamics for the grassland biome worldwide. *Global Biogeochemical Cycles*, 7(4):785–809, 1993.

C. Poeplau, A. Don, M. Dondini, J. Leifeld, R. Nemo, J. Schumacher, N. Senapati, and M. Wiesmeier. Reproducibility of a soil organic carbon fractionation method to derive rothc carbon pools. *European Journal of Soil Science*, 64:735–746, 2013.

J. Post, F. F. Hattermann, V. Krysanova, and F. Suckow. Parameter and input data uncertainty estimation for the assessment of long-term soil organic carbon dynamics. *Environmental Modelling & Software*, 23(2):125 – 138, 2008.

J. C. Refsgaard, J. P. van der Sluijs, A. L. Højberg, and P. A. Vanrolleghem. Uncertainty in the environmental modelling process - a framework and guidance. *Environmental Modelling & Software*, 22(11):1543–1556, 2007.

G. P. Richards and D. M. W. Evans. Development of a carbon accounting model (FullCAM vers. 1.0) for the Australian continent. *Australian Forestry*, 67(4):277–283, 2004. doi: 10.1080/00049158.2004.10674947. URL `https://doi.org/10.1080/00049158.2004.10674947`.

J. P. Scharlemann, E. V. Tanner, R. Hiederer, and V. Kapos. Global soil carbon: understanding and managing the largest terrestrial carbon pool. *Carbon Management*, 5(1):81–91, 2014.

J. O. Skjemstad, L. R. Spouncer, B. Cowrie, and R. S. Swift. Calibration of the Rothamsted organic carbon turnover model (RothC ver. 26.3), using measurable soil organic carbon pools. *Australian Journal of Soil Research*, 42:79–88, 2004.

P. Smith, D. Martino, Z. Cai, D. Gwary, H. Janzen, P. Kumar, B. McCarl, S. Ogle, F. O'Mara, C. Rice, B. Scholes, O. Sirotenko, M. Howden, T. McAllister, G. Pan, V. Romanenkov, U. Schneider, S. Towprayoon, M. Wattenbach, and J. Smith. Greenhouse gas mitigation in agriculture. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 363(1492):789–813, 2008.

R. Sommer and D. Bossio. Dynamics and climate change mitigation potential of soil organic carbon sequestration. *Journal of Environmental Management*, 144:83 – 87, 2014.

J. R. Stedinger, R. M. Vogel, S. U. Lee, and R. Batchelder. Appraisal of the generalized likelihood uncertainty estimation (GLUE) method. *Water Resources Research*, 44(12):W00B06, 2008.

X. Wang, X. He, J. Williams, R. Izaurralde, and J. Atwood. Sensitivity and uncertainty analyses of crop yields and soil organic carbon simulated with EPIC. *Transactions of the American Society of Agricultural Engineers*, 48(3):1041 – 1054, 2005.

C. K. Wikle and L. M. Berliner. A Bayesian tutorial for data assimilation. *Physica D: Nonlinear Phenomena*, 230(1-2):1 – 16, 2007.

C. K. Wikle, R. F. Milliff, R. Herbei, and W. B. Leeds. Modern statistical methods in oceanography: A hierarchical perspective. *Statistical Science*, 28:466–486, 2013.

A. Zammit-Mangion, M. Bertolacci, J. Fisher, A. Stavert, M. Rigby, Y. Cao, and N. Cressie. WOMBAT v1.0: a fully Bayesian global flux-inversion framework. *Geoscientific Model Development*, 15(1):45–73, 2022. doi: 10.5194/gmd-15-45-2022. URL `https://gmd.copernicus.org/articles/15/45/2022/`.

B. Zhang and N. Cressie. Bayesian inference of spatio-temporal changes of Arctic sea ice. *Bayesian Analysis*, 15(2):605–631, 2020.

Supplemental Material 1: Dynamical Process Model Parameters

The models defined in (3) and (7) of the main paper share two main classes of parameters: (i) decay rates, denoted by K_X ($X \in \{D, R, F, S, H\}$); and (ii) proportions of carbon routed between pools. These are the deterministic parameters used in the deterministic model RothC. For the former class, we generalize the decay rates (see (9)) such that they are modeled as the product of κ_X and α_τ, and the parameter model used for these is provided in Table S3.1. For the parameters in the latter class, $p_{P\to D}$ describes the proportion of plant matter entering the D pool with the remainder entering the R pool; $p_{U\to F}$, $p_{U\to S}$, and $p_{U\to H}$ describe the proportions of decayed carbon $U_{i,t}$ that move into the F, S, and H pools respectively; and $p_{V\to F}$, $p_{V\to S}$, and $p_{V\to H}$ describe the proportions of decayed carbon $V_{i,t}$ that moves into the F, S, and H pools respectively. The parameter model for this latter class is also given in Table S3.1.

Several other parameters in (3) and (7) of the main paper are in fact *derived* parameters from a smaller set of primary parameters. In particular, $p_{X\to F}$, $p_{H\to S}$, p_{clay}, $\pi_{M\to Y}$, and $r_{DPM/RPM}$ are primary parameters governing respectively, the proportion of solid soil-carbon that enters pool F from pool $X \in \{D, R, F, S\}$, the proportion of solid soil-carbon that enters pool S from pool H, the proportion of clay in the soil, the proportion of manure entering pool $Y \in \{D, R, F, S, H\}$, and the ratio of DPM to RPM in plant matter that enters the soil. From these primary parameters, one obtains [Jenkinson and Rayner, 1977]:

$$\begin{aligned}
p_{P\to D} &= \frac{r_{DPM/RPM}}{1 + r_{DPM/RPM}} \\
r_{CO2/Solid} &= 1.67(1.85 + 1.6e^{-7.86p_{clay}}) \\
p_{U\to F} &= \frac{p_{X\to F}}{1 + r_{CO2/Solid}} \\
p_{U\to S} &= 0.0 \\
p_{U\to H} &= \frac{(1 - p_{X\to F})}{1 + r_{CO2/Solid}} \\
p_{V\to F} &= 0.0 \\
p_{V\to S} &= \frac{p_{H\to S}}{1 + r_{CO2/Solid}} \\
p_{V\to H} &= \frac{1 - p_{H\to S}}{1 + r_{CO2/Solid}} \\
p_{M\to Y} &= \frac{\pi_{M\to Y}}{\sum_{Q\in\{D,R,F,S,H\}} \pi_{M\to Q}}.
\end{aligned} \tag{S1.1}$$

In equation (S1.1) above, $r_{CO2/Solid}$ is the mass of CO_2 lost to the atmosphere for every unit of carbon mass that decomposes from a pool. Decomposed carbon that is not lost from the soil as carbon dioxide is cycled to other carbon pools. We note that $r_{CO2/Solid}$ contains a number of empirically derived constants that model this parameter as a function of a parameter p_{clay}, where the latter parameter is the proportion of the soil mass that can be considered clay material. For

the MTT site, p_{clay} was estimated to be 0.16. All of the parameters denoted as $p_{\cdot\rightarrow\cdot}$ are assumed to be identical across all field-plots and constant in time for the duration of the MTT. To simplify dynamics, the parameters $p_{U\rightarrow S}$ and $p_{V\rightarrow F}$ are both put equal to zero in RothC v26.3, as do we in (7); these two biological pools are typically only very small relative to the total soil carbon.

References

D. S. Jenkinson and J. Rayner. The turnover of soil organic matter in some of the Rothamsted classical experiments. *Soil Science*, 123(5):298–305, 1977.

Supplemental Material 2: Geostatistical Diagnostics

1 Geostatistical Exploratory Data Analysis

Here we explore whether the data from the Millennium Tillage Trial (MTT), soil-carbon measurements of POC, ROC and TOC exhibit spatial dependence. Consider the spatial statistical model,

$$\log(Z_m(\boldsymbol{s})) = \mu_m(\boldsymbol{s}) + \delta_m(\boldsymbol{s}),$$

where $\boldsymbol{s}$ is the spatial coordinate for a sample of measurement $m \in \{\text{POC}, \text{ROC}, \text{TOC}\}$, $\mu_m(\cdot)$ is a deterministic process representing the large-scale variation as a function of space, and $\delta_m(\cdot)$ is a stochastic process representing small-scale variation. Since the MTT site was arranged as a matrix of plots, each sample location was defined as $\boldsymbol{s} = (x, y)^\top$, where $(0, 0)^\top$ is the center of the plot in the $(1, 1)$ entry of the matrix, x is the distance to the plot center in the row-wise direction, and y is the distance to the plot centre in the column-wise direction. The deterministic, large-scale variation was modeled via spatial trend as follows:

$$\mu_m(\boldsymbol{s}) = \beta_{m,0} + \beta_{m,r}x + \beta_{m,c}y + \beta_{m,rc}xy,$$

where $\beta_{m,0}$, $\beta_{m,r}, \beta_{m,c}$ and $\beta_{m,rc}$ are regression parameters specific to measurement m. To establish whether there was evidence of spatial covariance in the fine-scale stochastic process $\delta_m(\cdot)$, we first used ordinary least squares to estimate regression parameters, from which we defined an estimate of the (possibly spatially dependent) small-scale variation:

$$\hat{\delta}_m(\boldsymbol{s}) = \log(Z_m(\boldsymbol{s})) - \hat{\beta}_{m,0} + \hat{\beta}_{m,r}x + \hat{\beta}_{m,c}y + \hat{\beta}_{m,rc}xy,$$

where $Z_m(\boldsymbol{s})$ is the measurement of type m, and $\hat{\beta}_{m,0}$, $\hat{\beta}_{m,r}$, $\hat{\beta}_{m,c}$ and $\hat{\beta}_{m,rc}$ are the estimated regression parameters. For each unique pair of locations $\boldsymbol{s}_i$, $\boldsymbol{s}_j$ that index the spatial dataset, $\mathcal{D}_m = \{\hat{\delta}_m(\boldsymbol{s}_1), \ldots, \hat{\delta}_m(\boldsymbol{s}_n)\}$, we computed $(\hat{\delta}_m(\boldsymbol{s}_i) - \hat{\delta}_m(\boldsymbol{s}_j))$ and the Euclidean distance $d_{i,j} = ||\boldsymbol{s}_i - \boldsymbol{s}_j||_2$. Exploratory plots of $(\hat{\delta}_m(\boldsymbol{s}_i) - \hat{\delta}_m(\boldsymbol{s}_j))^2$ against $d_{i,j}$ and $|\hat{\delta}_m(\boldsymbol{s}_i) - \hat{\delta}_m(\boldsymbol{s}_j)|^{1/2}$ (see Cressie [1993], Section 2.4) were used to seek evidence of spatially structured dependence in the small-scale stochastic process $\delta_m(\cdot)$. Those plots are given in Figure S2.1, from which we saw no evidence of spatial structure. Hence, we proceeded

Table S2.1: Parameter model: prior probability distributions on terms in the Bayesian linear regression model for $\log(Z_m(\boldsymbol{s}))$.

Parameter	Description	Probability Distribution	Type
$\beta_{\mathrm{m,0}}$	Regression intercept parameter	$\beta_{\mathrm{m,0}} \sim N(0, 1E6)$	Uniformative
$\beta_{\mathrm{m,r}}$	Regression slope parameter across field-plot rows	$\beta_{\mathrm{m,r}} \sim N(0, 1E6)$	Uniformative
$\beta_{\mathrm{m,c}}$	Regression slope parameter across field-plot columns	$\beta_{\mathrm{m,c}} \sim N(0, 1E6)$	Uniformative
$\beta_{\mathrm{m,rc}}$	Regression interaction parameter across field-plot rows and columns	$\beta_{\mathrm{m,rc}} \sim N(0, 1E6)$	Uniformative
σ^2_{m}	Residual error variance	$\sigma^2_{\mathrm{m}} \sim$ Inv.-Gam.$(0.001, 0.001)$	Uniformative

with the assumption that that $\mathrm{cov}(\delta_m(\boldsymbol{s}_i)\delta_m(\boldsymbol{s}_j)) = 0, (\boldsymbol{s}_i \neq \boldsymbol{s}_j)$, for all three measurement types.

2 Estimation of Measurement Errors in Observed Soil-Carbon Measurable Fractions

With no evidence of spatial structure in the small scale stochastic process, we fitted the model,

$$\log(Z_m(\boldsymbol{s})) = \mu_m(\boldsymbol{s}) + \delta_m(\boldsymbol{s}) \quad (\delta_m(\boldsymbol{s}) \sim N(0, \sigma^2_m)),$$

as a Bayesian linear model with independent normal and inverse-gamma conjugate prior distributions, respectively on the parameter vector $\boldsymbol{\theta} = (\beta_{m,0}, \beta_{m,r}, \beta_{m,c}, \beta_{m,rc}, \sigma^2_m)^\top$. These prior distributions are provided in Table S2.1. The estimate of σ^2_m obtained for each measurement type, served as an informative prior distribution when using the CQUESST framework to analyze the MTT data.

References

N. Cressie. *Statistics for Spatial Data (Revised Edition).* John Wiley & Sons, 1993.

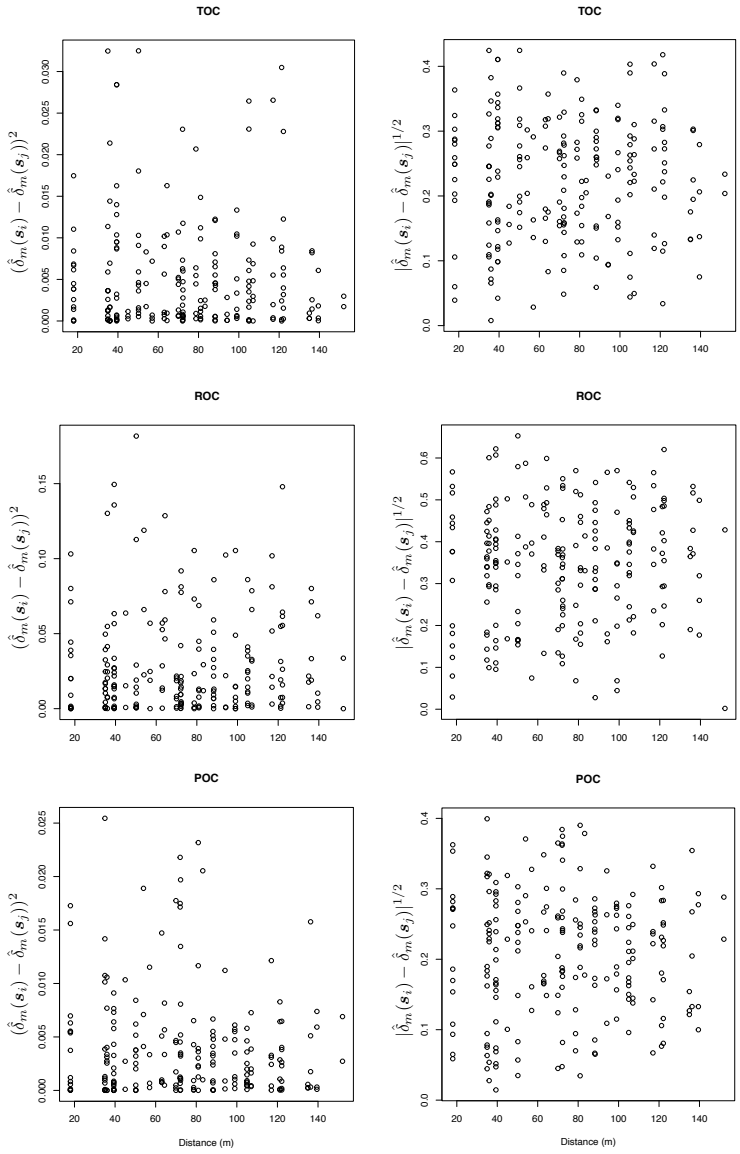

Figure S2.1: Diagnostic plots to explore how the small-scale variation $\hat{\delta}_m(\boldsymbol{s})$ varies as a function of distance for measurements of TOC, POC, and ROC.

Supplemental Material 3: Parameter Model

Here we outline the parameter model used in applying CQUESST to the Millennium Tillage Trial. Prior distributions over the parameters of the model are provided in Tables S3.1, S3.2, S3.3 and S3.4. Specifically, Table S3.1 gives the priors for the soil-carbon-cycling parameters in (7); Table S3.2 gives the priors for process-error variance parameters; Table S3.3 gives the priors on error variances for measured carbon fractions in (8); and Table S3.4 gives prior distributions for initial conditions of the soil-carbon pools defined in (7).

In these tables, TN_a^b denotes a truncated normal distribution that has been truncated on the left at a and on the right at b. We use truncated normal priors because all of the parameters in our model are bounded either from below or are within the interval $[0, 1]$, and because they provide a convenient and easily interpretable way to incorporate prior information. Other choices for priors on scale parameters in hierarchical models have been advocated [e.g., Gelman, 2006, Polson and Scott, 2012].

Finally, σ_D^2, σ_R^2, σ_F^2, σ_S^2, and σ_H^2 are not parameters that exist in the deterministic RothC model, and they are not biophysical parameters that have been estimated in past studies. Priors on these parameters were specified using inverse-Gamma distributions with shape and scale parameters chosen so that $\Pr(e^{\eta_X} < 0.6) \approx 0.01$ and $\Pr(e^{\eta_X} > 1.4) \approx 0.01$, where for $X \in \{D, R, F, S, H\}$, η_X represents the process error distributed as $\eta_X \sim N(-\frac{\sigma_X^2}{2}, \sigma_X^2)$. This results in the inverse-Gamma's shape parameter being set at 403.4 and its scale parameter at 0.318 (see Table S3.2). Since η_X is additive on the log-scale, e^{η_X} represents the corresponding multiplicative process error on the natural scale where the soil carbon is cycling. The values 0.6 and 1.4 were chosen so that the majority of the prior probability mass for σ_X^2 generated dynamics within $\pm 40\%$ of RothC's deterministic evolution of the soil carbon.

References

A. Gelman. Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis*, 1(3):515–533, 2006.

N. G. Polson and J. G. Scott. On the half-Cauchy prior for a global scale parameter. *Bayesian Analysis*, 7(4):887 – 902, 2012. doi: 10.1214/12-BA730. URL `https://doi.org/10.1214/12-BA730`.

Table S3.1: Parameter model: prior probability distributions on soil-carbon-cycling parameters in (7) and primary parameters.

Parameter	Description	Probability Distribution	Type
κ_D	Decomposition rate constant (y^{-1}) for D (decomposable plant material).	$\kappa_D \sim TN_{5.0}^{20.0}(10.0, (1.0)^2)$	Informative
κ_R	Decomposition rate constant (y^{-1}) for R (resistant plant material).	$\kappa_R \sim TN_{0.05}^{5.0}(0.07, (0.007)^2)$	Informative
κ_F	Decomposition rate constant (y^{-1}) for F (fast microbial biomass).	$\kappa_F \sim TN_{0.3}^{1.0}(0.66, 0.066^2)$	Informative
κ_S	Decomposition rate constant (y^{-1}) for S (slow microbial biomass).	$\kappa_S \sim TN_{0.3}^{1.0}(0.66, 0.066^2)$	Informative
κ_H	Decomposition rate constant (y^{-1}) for H (humus).	$\kappa_H \sim TN_{0.005}^{0.05}(0.02, (0.002)^2)$	Informative
α_τ ($\tau \in \mathcal{T}$)	Multiplicative tillage-cropping treatment effect applied to decomposition rates.	$\log(\alpha_\tau) \sim TN_{-5}^{5}(0.0, 1.0)$	Weakly Informative
$\pi_{M\to D}$	Proportion of manure to D pool	$\pi_{M\to D} \sim TN_0^1(0.49, (0.01)^2)$	Informative
$\pi_{M\to R}$	Proportion of manure to R pool	$\pi_{M\to R} \sim TN_0^1(0.49, (0.01)^2)$	Informative
$\pi_{M\to F}$	Proportion of manure to F pool	$\pi_{M\to F} \sim TN_0^1(0.0, (0.01)^2)$	Informative
$\pi_{M\to S}$	Proportion of manure to S pool	$\pi_{M\to S} \sim TN_0^1(0.0, (0.01)^2)$	Informative
$\pi_{M\to H}$	Proportion of manure to H pool	$\pi_{M\to H} \sim TN_0^1(0.02, (0.01)^2)$	Informative
$p_{X\to F}$	Proportion of soil-carbon from $X \in \{D, R, F, S\}$ to F.	$TN_0^1(0.46, (0.01)^2)$	Informative
$p_{H\to S}$	Proportion of soil-carbon from H to S.	$p_{H\to S} \sim TN_0^1(0.46, (0.01)^2)$	Informative
p_{clay}	Proportion of the soil that is clay	$p_{clay} \sim TN_0^1(0.16, (0.02)^2)$	Informative
$r_{DPM/RPM}$	Ratio of decomposable to resistant carbon in plant material.	$r_{DPM/RPM} \sim TN_0^\infty(1.44, (0.5)^2)$	Informative

Table S3.2: Parameter model: prior probability distributions on process-error variance parameters in (7).

Parameter	Description	Probability Distribution	Type
σ_D^2	Variance of additive process noise of D pool.	Inv.-Gam.$(403.4, 0.318)$	Informative
σ_R^2	Variance of additive process noise of R pool.	Inv.-Gam.$(403.4, 0.318)$	Informative
σ_F^2	Variance of additive process noise of F pool.	Inv.-Gam.$(403.4, 0.318)$	Informative
σ_S^2	Variance of additive process noise of S pool.	Inv.-Gam.$(403.4, 0.318)$	Informative
σ_H^2	Variance of additive process noise of H pool.	Inv.-Gam.$(403.4, 0.318)$	Informative

Table S3.3: Parameter model: prior probability distributions on error variances for measured carbon fractions.

Parameter	Description	Probability Distribution	Type
σ_{POC}^2	Measurement error variance for log(POC).	Inv.-Gam.$(\frac{21}{2}, 0.039)$	Informative
σ_{ROC}^2	Measurement error variance for log(ROC).	Inv.-Gam.$(\frac{21}{2}, 0.290)$	Informative
σ_{TOC}^2	Measurement error variance for log(TOC).	Inv.-Gam.$(\frac{21}{2}, 0.053)$	Informative

Table S3.4: Parameter model: prior probability distributions on initial conditions of the soil-carbon pools defined in (7).

Initial Condition	Description	Probability Distribution	Type
$D_{i,0}$	The initial state of D in field-plot i (Mg/ha).	$TN_0^{\infty}(0.0, (0.1)^2)$	Informative
$R_{i,0}$	The initial state of R in field-plot i (Mg/ha).	$TN_0^{\infty}(0.0, (100.0)^2)$	Uninformative
$F_{i,0}$	The initial state of F in field-plot i (Mg/ha).	$TN_0^{\infty}(0.0, (0.01)^2)$	Informative
$S_{i,0}$	The initial state of S in field-plot i (Mg/ha).	$TN_0^{\infty}(0.0, (0.01)^2)$	Informative
$H_{i,0}$	The initial state of H in field-plot i (Mg/ha).	$TN_0^{\infty}(0.0, (100.0)^2)$	Uninformative
$I_{i,0}$	The initial state of D in field-plot i (Mg/ha).	$TN_0^{\infty}(0.0, (10.0)^2)$	Uninformative

Supplemental Material 4: MCMC Diagnostics

When performing Bayesian inference with an MCMC algorithm, it is important to verify that the Markov chain provides a representative set of samples from the posterior distribution $p(\mathbf{Y}, \boldsymbol{\theta}|\mathbf{Z})$. A common metric used to assess convergence is the $\hat{R}$ statistic [Gelman et al., 2013]. This metric assesses the convergence of each element γ_k in the vector of all sampled random variables, $\boldsymbol{\gamma} = (\boldsymbol{\theta}^\top, \mathbf{Y}^\top)^\top$, using m independent Markov chains, each containing n samples. For each γ_k, the statistic is calculated from the ratio of two variance estimators, namely:

$$\hat{R}_{\gamma_k} = \sqrt{\frac{\widehat{\mathrm{Var}}^+(\gamma_k|\mathbf{Z})}{W}},$$

where $\gamma_{i,j,k}$ denotes the ith sample from one of $j = 1, \ldots, m$ Markov chains, each consisting of n samples.

$$\widehat{\mathrm{Var}}^+(\gamma_k|\mathbf{Z}) = \frac{n-1}{n}W + \frac{1}{n}B,$$

$$B = \frac{n}{m-1}\sum_{j=1}^{m}(\bar{\gamma}_{\cdot jk} - \bar{\gamma}_{\cdot\cdot k})^2, \quad \bar{\gamma}_{\cdot jk} = \frac{1}{n}\sum_{i=1}^{n}\gamma_{ijk}, \quad \bar{\gamma}_{\cdot\cdot k} = \frac{1}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m}\gamma_{ijk},$$

$$W = \frac{1}{m}\sum_{j=1}^{m}s_{jk}^2, \quad s_{jk}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(\bar{\gamma}_{ijk} - \bar{\gamma}_{\cdot jk})^2.$$

When $\hat{R} < 1.01$ for every $\gamma_k \in \boldsymbol{\gamma}$, the aggregated set of samples from the m chains is considered to provide a set of samples from which reliable posterior inferences can be made. All inferences presented herein were performed using $m = 6$ independent Markov chains, each consisting of 100,000 samples, thinned by taking every 100th sample, so that each chain yielded $n = 1,000$ iterations (following 5,000 "warmup" iterations). For all γ_k, we computed $\hat{R}$ and assessed that posteriors were reliable.

References

A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin. *Bayesian Data Analysis (Third Edition).* Chapman and Hall/CRC, 2013.

Supplemental Material 5: Posterior Inference

The CQUESST Bayesian hierarchical model was run in `Stan`, from which we obtained posterior samples from six independent chains of the MCMC algorithm. The samples from the three chains were aggregated into a larger set of samples and kernel density estimates were plotted alongside the prior distributions used in the parameter model. These plots are provided in Figures S5.1, S5.2, and S5.3. In some cases, the prior has deviated away from the prior (known as *Bayesian learning*), indicating that the data has updated our beliefs about the distribution of the parameters. In some other cases, the prior and posterior are very similar, indicating that the data has provided little additional information about the probability distribution of the parameter beyond what was contained in the prior.

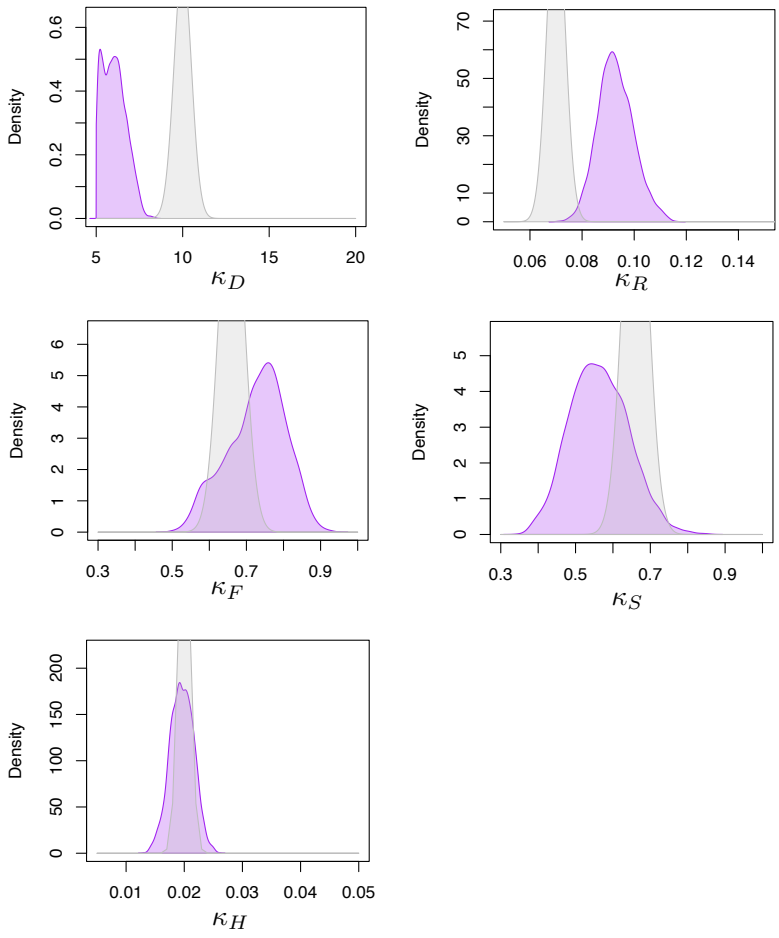

Figure S5.1: Prior (gray) and posterior (purple) distributions for decay-rate parameters of CQUESST. Plots give priority to the posterior distribution; when a prior is not visible, it indicates that it was very far from the posterior.

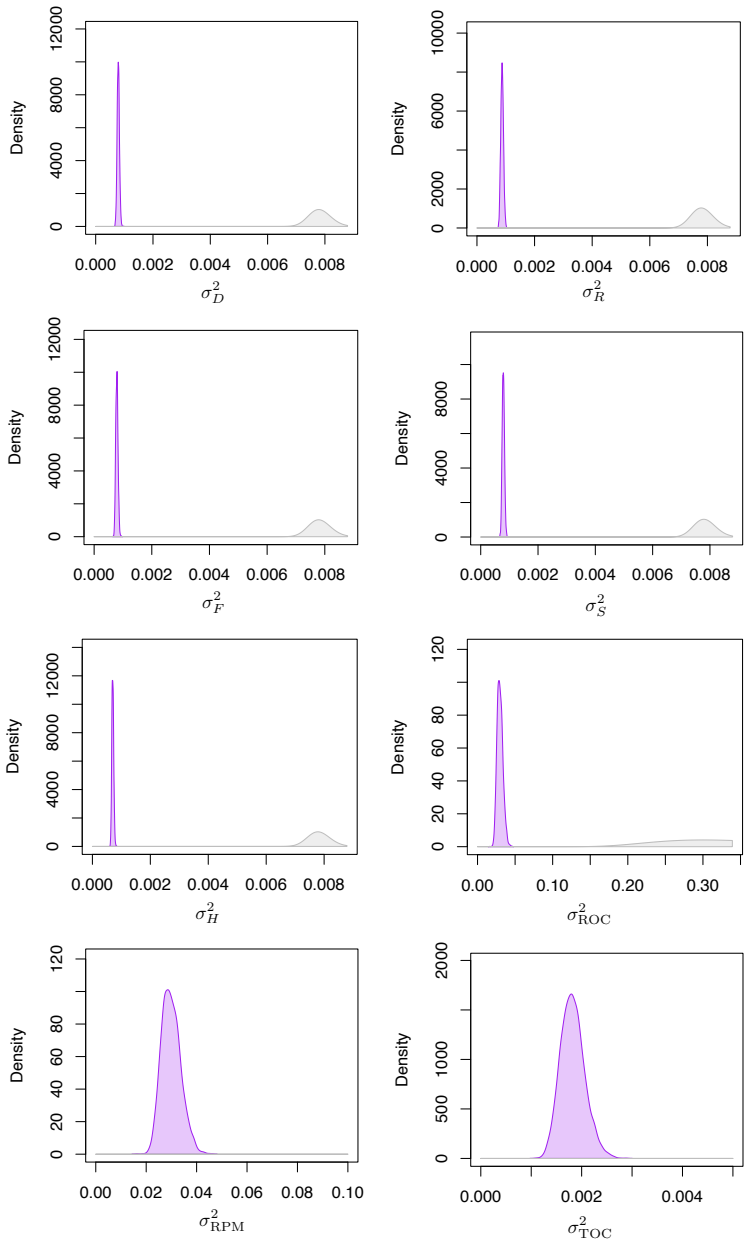

Figure S5.2: Prior (gray) and posterior (purple) distributions for variance parameters of CQUESST. Plots give priority to the posterior distribution; when a prior is not visible, it indicates that it was very far from the posterior.

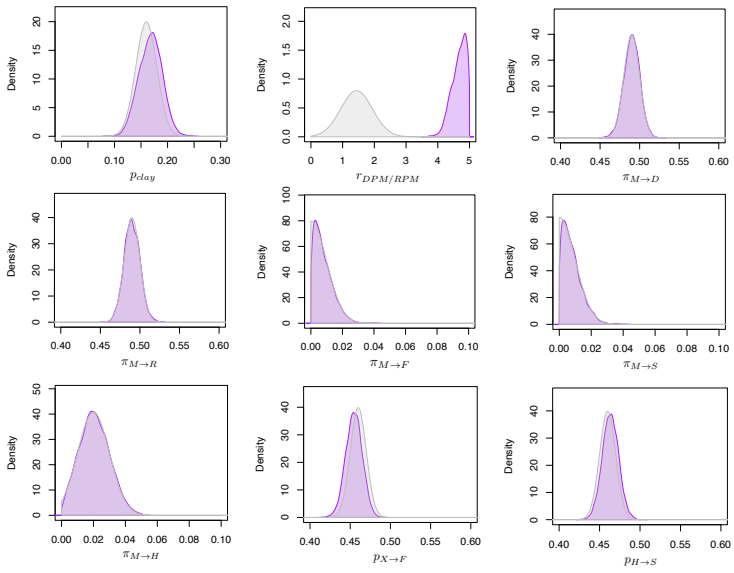

Figure S5.3: Prior (gray) and posterior (purple) distributions for a subset of the parameters of CQUESST. Plots give priority to the posterior distribution; when a prior is not visible, it indicates that it was very far from the posterior.