Improving Yasso15 soil carbon model estimates with Ensemble Adjustment Kalman Filter state data assimilation

Toni Viskari[1], Maisa Laine[1], Liisa Kulmala[1,2,3], Jarmo Mäkelä[1], Istem Fer[1] and Jari Liski[1]

[1]Finnish Meteorological Institute, Helsinki, 00101, Finland
[2]Department of Forest Sciences, University of Helsinki, P.O. Box 27, FI-00014 Helsinki, Finland.
[3]Institute for Atmospheric Sciences and Earth System Research, University of Helsinki, Helsinki, Finland

Correspondence to: Toni Viskari (toni.viskari@fmi.fi)

Abstract Model-calculated forecasts of soil organic carbon (SOC) are important for approximating global terrestrial carbon pools and assessing their change. However, the lack of detailed observations limits the reliability and applicability of these SOC projections. Here, we studied if state data assimilation (SDA) can be used to continuously update the modeled state with available total carbon measurements in order to improve future SOC estimations. We chose six fallow test sites with measurements time series spanning 30 to 80 years for this initial test. In all cases, SDA improved future projections but to varying degrees. Furthermore, already including the first few measurements impacted the state enough to reduce the error in decades long projections in by at least 1 t C ha^{-1}. Our results show the benefits of implementing SDA methods for forecasting SOC, but also highlight implementation aspects that need consideration and further research.

1. Introduction

Terrestrial soil organic carbon (SOC) pools serve a crucial role in the global carbon cycle by acting as a large long-term carbon storage for terrestrial systems and are, similarly to the other carbon cycle components, directly impacted by the changing climate and environment (Ciais et al., 2013). Local meteorological conditions drive soil temperature and moisture, which together with soil characteristics affect the microbial processes that decompose SOC (Orchard and Cook, 1983; Karhu et al, 2014; Vogel et al. 2015). SOC input is largely composed of vegetation litter and extracts with contributions from soil bacteria and mycorrhiza (Cornwell et al., 2008). Thus, when the vegetation cover is altered due to changing environmental conditions or anthropogenic activities, it will also alter the long term SOC stocks. Furthermore, the SOC response to the new surface conditions is slow and it takes years to decades, or even longer, before the peatland draining or transformation of forest to agricultural field reaches a new stable state e (Mao et al, 2019). All these factors have made it difficult to empirically assess how both local and global SOC stocks will be affected by the changing climate and environment (Sulman et al., 2018).

To address these challenges, several SOC models of varying complexity have been created over the years (e.g. CENTURY (Parton, 1996), MILLENIAL (Abramoff et al., 2017) and ORCHIDEE-SOM (Cammino-Serrano et al., 2018)) with an

increasing focus on how to better mathematically formulate the central physical soil processes (Liang et al. 2017). These models allow projecting SOC in different environments and, are important tools in approximating regional and global SOC distributions as well as how they are changing over time (Manzoni and Porporato, 2009). As such, they also serve an important role in estimating how climate change impacts the SOC stocks, which cause one of the largest uncertainties in future carbon cycle projections due to the size of pools and their direct link to ecosystem response (Hararuk et al. 2014). On a practical level, SOC models have been used to calculate soil carbon components for National Carbon Budgets or to determine carbon allocation in soils under different agricultural management conditions, when calculating carbon credit market values (Smith et al., 2020).

Despite this increasing number and variety of modelling choices, the future projections produced by them all face similar difficulties which have resulted in high uncertainties (Bradford et al, 2016). Fundamental among these challenges is the lack of observation data required to parameterize and initialize the models (Sulman et al, 2018). Relevant measurement campaigns are resource-heavy and time-costly (Jandl et al., 2014). Consequently, single measurements are used to represent SOC concentrations for wider regions despite SOC varying highly spatially, which will inherently introduce error into SOC projections. Furthermore, vast majority of the available measurements represent bulk total soil carbon contents whereas the decomposition dynamics are greatly dependent on a more nuanced representation of the organic carbon state, such as which fraction of SOC is contained by stable long-lived carbon compounds as opposed to active short-lived carbon compounds (Lehmann and Kleber, 2015). The lack of detailed measurements forces models to use less reliable methods to approximate the initial SOC state, which in turn is a major limitation in trying to estimate how the projected SOC state reacts to environmental changes (Wutzler and Reichstein, 2007; Palosuo et al., 2012).

Using observations to constrain state projections is a central question for all predictive tasks, and different approaches have been developed to address this need. State data assimilation (SDA) refers to Bayesian methods where state information from two or more sources are combined to create a more accurate estimation of the true state (Evensen, 2009). It has already been applied in several geophysical subjects (e.g., Elbern et al. 2000; Weaver et al. 2003; Viskari et al., 2012; Yang et al., 2019) and is a fundamental component that allows weather forecasts (Le Dimet and Talagrand ,1986). In recent years there have been efforts to also use SDA methods to better incorporate flux tower and satellite measurements (Viskari et al., 2015) to update ecological model projections. As one of the core advantages, SDA allows to update unobserved state variables with information from observed state variables based on the currently understood and presented process dynamics (Dietze, 2017). In SOC related systems, SDA applications have so far been limited and either focused on estimating model parameters (Trudinger et al., 2008) or constraining the drivers affecting the soil carbon fluxes (Yan et al., 2019). In Gao et al., (2011), SOC was estimated as a component of the total carbon allocated in a forest ecosystem, but even there the main focus was on the model parameter estimation.

Applying SDA methods in SOC research may address several current challenges in the field. As SDA makes it possible to continuously incorporate measurement information to update and correct the model state, it consequently both reduces the impact of initial state uncertainty and allows using multiple measurements to better constrain future SOC projections. Due to

70 SDA being able to update unobserved state variables based on observed ones, it allows use of the total carbon measurements to correct the more detailed active and stable SOC pools. These state variables are needed in models or when estimating regional carbon stocks based on local measurements. However, while the basic equations for SDA remain the same, there are practical challenges in implementing SDA that depend on the system examined, such as varying frequencies for different observations or the types of observations uncertainties (Dietze, 2017). Consequently, implementing SDA for ecosystems

75 requires addressing different issues and questions than implementing SDA in atmospheric systems (Dietze et al., 2018).

In this study, our aim was to determine if SDA is able to effectively use coarse observation data to continuously update the model SOC state and improve associated model projections. More specifically, we wanted to both determine how the total carbon measurements affect the individual model pools and how many measurement points need to be included to start

80 impacting the future predictions in a noticeable manner. The decades long SOC dataset measured at bare fallow agricultural fields around Europe (Barré et al. 2010) was used along with Yasso (Tuomi et al., 2011; https://github.com/YASSOModel), a SOC decomposition model that has been shown performing well for long-term SOC projections (Ortiz et al., 2013; Ziche et al., 2019), to test if updating the model projection with observations has an impact on future state predictions. The bare fallow sites do not include the uncertainty of litter input estimates and thus, allowed us to focus more on the impact SDA has on the

85 model projections. We applied the Ensemble Adjustment Kalman filter (EAKF; Andersson, 2001) as the SDA method in the study. In EAKF, the ensemble is created by running the model with varying initial states, which are then all updated with the information from measurements as explained in more details in the following sections. Not only is EAKF a widely established SDA method, but it is a part of the Data Assimilation Research Testbed (DART; Anderson et al, 2009) workflow.

90

2. Materials and methods

2.1. Yasso model

95

Yasso (Tuomi et al., 2011) is a soil organic carbon (SOC) model which simulates SOC decomposition by shifting C between different soil pools that represent different organic carbon forms before either releasing it back to the atmosphere as heterotrophic respiration or transforming it into inactive and slow-cycling humus. Within the model, carbon is divided into five different SOC pools: Ethanol (E), Water (W) and Acid (A) soluble pools and a non-soluble pool that is further divided in

100 to lignin-like pool (N) and a humus (H) pool having different decomposition rates. Decomposition is affected by air temperature and precipitation, which are used in the model as indicators for soil temperature and moisture. Additionally, Yasso

accounts for the size dependency for woody mass as it takes longer in those situations for the microbes to break the litter down. Model SOC can only increase by the plant litter input.

The change in state at time t, x_t', is represented as a matrix equation

105 $$x_t' = x_t + b \tag{1.}$$

Where x_t is a vector where each component is the amount of carbon in each pool and the litter input b is divided into the same chemical pool components as the Yasso SOC pools which added directly to the YASSO state vector. The division of b depends on the litter species. The matrix M represents the model dynamics affecting the system and is written out as

$$\mathbf{M} = \begin{bmatrix} \alpha_A & p_{WA} & p_{EA} & p_{NA} & 0 \\ p_{AW} & \alpha_W & p_{EW} & p_{NW} & 0 \\ p_{AE} & p_{WE} & \alpha_E & p_{NE} & 0 \\ p_{AN} & p_{WN} & p_{EN} & \alpha_N & 0 \\ p_H & p_H & p_H & p_H & \alpha_H \end{bmatrix}$$

110 Where α_i is the decomposition rate for pool i and p_{ij} is the fraction that is decomposed from pool i to pool j. The decomposition rate depends on environmental temperature and moisture, here indicated by the surface air temperature (T) and precipitation (P), as well as woody litter diameter according to the equation

$$\alpha_i = \frac{\alpha_{base,i}}{4} h(d)(1 - e^{\gamma_i P}) \sum_{k=1}^{4} e^{\beta_{i1}T_k + \beta_{i2}T_k^2} \tag{3.}$$

Where $\alpha_{base,i}$ is the baseline decomposition rate for pool i, $h(d)$ is the function that determines the impact of the litter size on
115 the decomposition, γ is the precipitation impact parameter and $\beta_{1,i}$ and β_2 are the temperature impact terms for pool i. The annual temperature impact here is averaged over four points of the annual temperature cycle in order to capture the change in temperature. The total annual precipitation is used here as a proxy for the soil moisture. . The parameters associated with each process were estimated with an Adaptive Metropolis MCMC (Haario et al., 2001) method based on joint information from a number of different litter decomposition data bases such as CIDET (Trofymov, 1998), LIDET (Gholz et al., 2000) and
120 Eurodeco (Berg et al, 1991a; Berg et al. 1991b). The parameter values for equations 1-3 and their associated uncertainties are shown in Supplemental Table 1.

2.2 The measurement time series and the initial carbon pool

125 Bare fallow experiments included in the study were kept vegetation-free and free of organic amendments for more than 25 years. The study sites are in Europe and selected characteristic of these are presented in Table 1. The cultivation time that lead-up to the bare fallow experiment varied from 75 years to centuries. All the SOC measurements are of the bulk soil C without details on AWEN fractions. The sites are introduced in detail by Barré *et al.* (2010) and Menichetti et al (2019). The detailed setup in Versailles is described also by van Oort et al. (2018) and in Askov by Christensen (1990) and Christensen and Johnston
130 (1997).

Yasso model requires information on the initial AWENH fractions which we estimated site-specifically based on the Net Primary Production (NPP) in the presumable native habitat and the estimated litter input during time of cultivation (Figure 1, Table 1, Kulmala and Liski, 2018). In short, we first made a general NPP estimate in the native habitat using mean temperature and precipitation as in Del Grosso et al. (2008) and divided it into non-woody, small-sized woody and large woody litter fractions based on the native ecosystem. The different litter types had individual carbon fractions based on solubility. Next, we used Yasso to determine the steady state pool of soil carbon and its fractions using the NPP, different litter fractions, their chemical fractions and mean temperature and precipitation as driver data (Fig1-a).

Before bare fallow, each field had been cultivated for 75–300 years. For that period, we simulated SOC starting from the steady state SOC and its chemical fractions achieved using the pre-agriculture litter approximating (Fig1-b), The same mean annual temperature and precipitation were used as drivers for both the pre-agriculture and agriculture SOC decomposition. The annual litter input for the cultivation period was estimated in a site-specific manner to meet the first SOC measurement after the cultivation period. The carbon fractions in that litter input were assumed to be as presented in Karhu et al. (2012) and the litter is assumed to be non-woody with a diameter of 0 cm. The resulting SOC as the starting point for the bare fallow period (Fig1-c). The AWENH distributions calculated in this manner for each site are shown in Table 1 and are used to calculate the z^f from eq. 4 for the first assimilation cycle as detailed below.

2.3 State Data Assimilation method

As there is no way to know the true state of a variable, all our information on it, be it modelled or observed, will be inherently uncertain (van Oijen, 2017). State data assimilation (SDA) is a Bayesian statistical method which combines information from multiple sources to create a statistically optimal state estimate. At each assimilation step, SDA updates a priori knowledge of the system state, almost always a model prediction, with state observations. This results in a posterior state estimate of both the expected value as well as the associated uncertainty. The posterior state estimate is considered the most reliable view on the true state given the available information in model predictions and state observations, thus being less uncertain than all the information sources used to estimate it. Each information source influences the posterior estimate in proportion to their uncertainties: higher observational uncertainty results in a posterior state estimate closer to the model prediction, and vice versa (Dietze, 2017).

In our research we used the Ensemble Adjustment Kalman filter, (EAKF; Andersson, 2001) which is based on the Kalman Filter theory (Kalman, 1960). The ensemble consists of numerous model projections started from different initial conditions which are moved forward in time independently until the next observation, and the prior state uncertainty is determined from

the ensemble spread. At the time of each observation, an update (later called as analysis) is calculated with the following equation

$$z_i^a = A^T\left(z_i^f - \acute{z}^f\right) + \acute{z}^a,$$
(4.)

where z is a joint state-observation vector, index f denotes forecast, index a denotes analysis and index i denotes each individual ensemble member. Matrix A shifts the whole ensemble so that the updated ensemble has a mean equal to

$$\bar{z}^a = P^a\left[(P^f)^{-1}\,\bar{z}^f + H^T R^{-1} y\,\right]$$
(5.)

and covariance equal to

$$P^a = \left[(P^f)^{-1} + H^T R^{-1} H\,\right]^{-1},$$
(6.)

where y is the observation vector, H denotes the observation operator, P is the model state error covariance matrix and R is the observation error covariance. In Ensemble Kalman Filter applications, each component P^f_{ij} of the forecast error covariance matrix P^f is calculated over the ensemble members

$$P_{ij}^f = \frac{\sum_{k=1}^{L}(z_{i,k}^f - \bar{z}_i^f)(z_{j,k}^f - \bar{z}_j^f)}{L-1}$$
(7.)

Where L is the size of the ensemble. It should be noted that the non-diagonal terms of the analysis error covariance matrix P^a represent the error covariances and allow the observation of a specific state also affect other members of the state variable vector.

There are practical challenges that need to be accounted for when utilizing SDA methods, filter divergence (Schlee et al, 1967) being the most relevant one regarding this study. In practice, the ensemble uncertainty does not represent all the uncertainty sources affecting the model predictions. For example, process error arises due to underrepresented model processes or in-sufficiently included model interactions. As there are no reliable ways to establish it for process-based simulators like YASSO (van Oijen, 2017), it cannot be accounted for in the assimilation process. Consequently, the modelled uncertainty does not necessarily increase enough to balance out the reduction in posterior uncertainty during the analysis phase. As a result, the updating process gives too much weight to the prior state in comparison to the observed state as the projected uncertainty decreases while the observation uncertainty remains similar. Ultimately a stage is reached where the measurements stop affecting the estimate due to the difference in uncertainties. When this happens, the forecast begins to diverge from reality. There are several methods for dealing with filter divergence (Evensen, 2004; Anderson, 2006), but as this a preliminary study, we used a simple inflation method established in Hamill (2001). In this approach, the forecast/prior covariance is multiplied with a constant factor greater than 1 before every analysis/update step to ensure that the measurement continues to affect the estimate. The practical implementation is explained in more detail in the following section.

2.4 Simulation details

The ensemble initial states were created with R language using the *rnorm* function with a condition checking that the outputs are non-negative. The initial ensemble values for each pool were determined by drawing from a normal distribution where the initial value for that site was used as a mean. A 10 % mean was used as the initial ensemble variance as there were no reliable initial state uncertainty approximation. This was found to be larger than just perturbing the litter inputs and, as such, was not likely to underestimate the initial uncertainty. For this initial study covariances between the different SOC pools were not considered. The sampling was used to create an ensemble with 50 different states. This way we can represent the uncertainty in the total amount of soil organic carbon and how it is distributed among the five pools. The initial distributions for total carbon are shown in Fig 2.

Using the first measurement to scale the initial SOC results in the first measurement to be used twice if the whole measurement time series is used in the SDA implementation. However, not scaling the initial state with the first measurement would cause large differences between the initial state and the first measurement, with the resulting SDA estimations expected to be clearly superior to the non-SDA predictions in such a situation. In other words, it would not be a fair comparison as generally in runs like these, the initial state would be constrained to some degree by available measurements. Other option could be to exclude the first observations from data assimilation. However, including the information from the first measurement in a decomposition time series in the SDA implementation is assumed to be important as the SOC state changes most drastically over the first few years which in turn would impact the initial state uncertainty propagation. In an ideal situation, there would be an independent SOC measurement that can be used to constrain the SOC initial state, but such additional data was not available here. Thus, we used the whole time series in the assimilation here as using the first measurement twice was expected to have only a very negligible effect on the overall results. We also did a comparison run where the SDA was only applied from the second measurement forward in order to be certain. These runs were set up identically except the relative error of the first measurement was used as variance to randomly draw the ensemble members.

We used the Data Assimilation Research Testbed (DART; Anderson et al, 2009) to run our assimilation with the EAKF. The initial ensemble for each site was given to DART as a starting point and climate data measured at the sites were used as model drivers. The climate driver data is provided alongside with this article. The state vector consists of the five SOC pool stocks as presented in Yasso and the total SOC which is a sum of the five pools. The total SOC projection is compared to the measurements and the error covariances calculated by DART transfer the information to the other state vector components.

The SDA was first tested by updating the model state variables at each measurement time according to eq. (4). and using the produced state estimation was then used by Yasso to determine the next predicted state. This basic test was repeated with three different inflation factors (1, 1.25 and 1.5) in order to examine how much filter divergence affects the projections and which inflation factor range produces satisfactory predictions. The inflation factors were implemented by scaling the posterior error

covariance matrix produced by eq. 6 according to the inflation factor chosen before using it in eq. 5, and consequently, in eq. 4. Only the inflation factor results for 1. and 1.25 are shown here for the sake of clarity. In the second set of tests, only a limited number of initial measurements (first, first two, first three or first four) are used to update the state before it is then allowed to run the whole time series freely without being updated with measurement information to determine how soon the measurement information begins to noticeably impact the model projections. Only the first four measurements were used in this phase as the central question was how assimilating SOC measurements impact long term forecasts. The inflation factor of 1.25 was used in these latter tests.

All the forecasts produced with these tests were compared to both measurements and baseline Yasso SOC projection that was ran from the initial state without any SDA. In order to better assess how the SDA improved the state forecasts, we calculated the RMSE for the last four measurements at each site for all the five experiments: namely, the baseline YASSO model forecast, SDA with first observation only, SDA with first two, first three and first four observations using the forecasts that used the limited number of measurements as well as the baseline Yasso model forecast.

3. Results

Using SOC data to update the state of the model improved the model-calculated estimates compared to non-SDA model projections run from the approximated initial state (Fig. 2). While the inflated SDA predictions (Inflation factor 1.25) had larger uncertainties than the uninflated ones (Inflation factor 1.0), the predictions themselves remained close to each other with exception of the two Askov sites. There, systematic shifts occur in the observed states decades after the start of the time series, and indicative of the effect of the filter divergence, the uninflated SDA predictions did not react to these shifts while the inflated SDA predictions were adjusted to the new states. Applying the inflation value of 1.5, the analysis essentially matched the measurements (Supplemental figure 1).

Due to the multiple changes in the Askov B4 time series, the more detailed model state response is represented for it to see how the state estimate adapts to the changes there. Among the SOC pools, the humus pool changed most in response to the observations and always to their direction at Askov B4 (Fig. 3). The SDA estimate affected the AWEN pools only a little during the first half of the time series, but after approximately 10 years, these pools were also changed during the state update. Interestingly, these pools were changed to the opposite direction than the humus pool and observations. SDA affected the humus pool in a same way at other sites, and a similar difference in the behaviour between the AWEN and the humus pools was observed at Askov B3 after the systematic shift (Supplemental Figures 2-6).

At most sites incorporating information already from the first two observations had a noticeable impact on the time series prediction. Although at the Versailles site, the SDA forecast was close to the non-SDA forecasts at the end of the time series (Fig 4). At the Askov sites, the updated predictions ended up overestimating the latest measurements more than the model alone did due to the systematic shift in measurement values (At Askov B4 after 1966 and at Askov B3 after 1977). Finally, RMSE values (Table 2) show that aside from Askov, the assimilation reduced the RMSE at each site by the fourth measurement at the latest.

The comparison runs where the assimilation was only done from the second measurement forward were nearly identical for the estimated total SOC values when continuously assimilating and when only using the first few observations to constrain the predictions (Figures not shown). The more detailed examination of the state at Askov B4 (Figure 5) did show a difference as the later corrections affecting the AWEN pools are more muted than if the assimilation begun from the first measurement of the time series.

4. Discussion

This study establishes that state data assimilation (SDA) improves soil organic carbon (SOC) forecasts by continuously incorporating total carbon measurements. As such, this adds another type of measurements that SDA can use to improve future projections in addition to previously shown positive impacts of assimilating soil environmental conditions (Yan et al., 2019). At all sites assimilating already the first few measurements had a clear impact on the forecasts (Fig 4; Table 2). It should be noted that at Askov, the non-SDA forecast is closer to the measurements towards the end of the time series than the SDA forecast that assimilated the first few observations. This is due to the systematic shift in measured SOC that happens at Askov B3 around 1965 and at Askov B4 around 1975. Before that, the SDA forecasts are closer to the SOC measurements than the non-SDA forecasts. This supports previous research on the impact of initial state uncertainty on SOC projections (Todd-Brown et al. 2014; He et al. 2016).

While the inflation term does increase the uncertainty of the forecasts and thus reduces the filter divergence, the uninflated and inflated forecasts remain similar. Askov sites are the exception here as there the inflated SDA forecast reacts to the previously noted systematic shift in measurement time series. This is expected as the inflation reduces the impact of filter divergence and thus allows for the later observations affect the analysis more than they would without the uncertainty inflation. These results here indicate that it succeeds in the framework discussed in Anderson (2001) on how inflated systems should behave. However, once litter input will be introduced into the system, it will add a potentially systematic source of error as the uncertainties in the litter input affect the SOC projections. At that point, a more nuanced inflation approach or other more elaborate implementations, such as estimating the process variance from observations (Dietze, 2017), could be required.

When examining how SDA affects the model state, it is important to note that the total SOC measurements affect the model state based on the error covariances between the different pools and their total SOC. The initial uncertainties were introduced as independent of each other with the SDA calculating the error covariances between the different pools over the analysis process. The resulting error covariance between the humus pool and total SOC is strongly positive with a decrease in humus also decreasing total SOC and vice versa (Fig. 4). This is reasonable as the long-lived SOC is generally dominated by the stable humus pool (Lehmann and Kleber, 2015) and adjusting to it is crucial in capturing the decomposition without litter input. Furthermore, due to the slow decomposition rate of the H pool and the relatively high frequency of the observations, in Fig 4 the prior H value is essentially the posterior H of the previous assimilation cycle.

Error covariances between the AWEN pools and total SOC are more complicated than between SOC and H pool. Thus, it takes more analysis cycles for the method to establish them. Consequently, the analysis appears to affect the humus SOC from the start of the time series while with AWEN pools the analysis impact appears to become stronger later into the time series. Due to this, the two Askov timeseries are the only sites, where we also capture the meaningful AWEN pool impacts due to the late shift in the observed state. Even there, though, covariance is strongly affected if the uncertainty spread over the first few years of the decomposition is included (Figs 4 and 6). It is noteworthy that once the SDA properly determines the error covariance structures, the analysis adjusts the AWEN pools to the opposite direction than it does the H pool. Initially, this might appear to be counterintuitive, as SDA increases the AWEN values in response for the forecast overestimating the SOC values, but this is due to model dynamics being reflected via the error covariances. Further complicating the matter is that the active AWEN pools are affected differently by the environmental conditions than the inactive H pools (Tuomi et al, 2008). This will cause the resulting AWENH error covariances to vary between locations even if the total carbon and H error covariance appears to be consistent.

In addition to providing a valuable illustration in how the error covariances change over time and impact the later state corrections, the Askov sites also highlight both a strength and limitation of the SDA methods. As seen in the measurement time series (Fig 3), both Askov B3 and B4 have a systematic shift in measurements at different times due to reasons currently not known. The model projections alone cannot capture these developments as it is unclear if the sudden drop in observed SOC is even due to an ecological process not represented in the model or some issues relating to the measurements. In either case, the inflated SDA adapts to the new state within a few measurement cycles and produces a forecast that follows the new state well. This is clearly a strength of the SDA method that would be beneficial when forecasting SOC at locations where there are disturbances and alterations in the surface conditions. However, here the new state estimate appears questionable as there is a sudden increase in active SOC (i.e. AWEN pools) despite it being over a decade since there was any litter fall. This goes against the current understanding of the system as there is no transference from the H pool back to the AWEN pools and no litter to provide those faster decomposing carbon compounds. In this case, the late increase of the AWEN pool

concentrations at the Askov sites are an artefact of the error covariances established earlier rather than a realistic representation of what is happening in the soil. Thus, while SDA is a beneficial tool when examining changing systems, the nature of SOC model dynamics makes it important to also expertly assess how the new estimated state reacts to those changes.

325

Another site that shows SDA having challenges is Versailles where SDA only slightly improves the forecasts towards the end of the time series. This is probably due to either the model not representing a SOC decomposition dynamic relevant to that particular site or the input drivers are somehow lacking. This site highlights that while SDA is a valuable tool for improving forecasts, it is still limited by how well the applied model captures the local SOC dynamics. For instance, if the soil respiration

330 parameters for the H pool is poorly represented for that location, then the resulting H pool, as well as the total SOC, projections would still be more uncertain even when applying SDA methods (Gao et al., 2011). However, SDA is still useful in these situations as it can indicate sites where the forecast error is not driven by the state uncertainty. Therefore, it is easier to analyse the differences between the sites like Ultuna and Versailles. For example, it is known that soil quality affects the SOC decomposition (Chapin et al. 2011; Vogel et al. 2015), so here it would support in further researching the soil properties at

335 Versailles to determine if those dynamics have an impact there that should be acknowledged with SOC forecasts at other similar sites.

The continuous SDA forecasts from Askov sites and Versailles (Fig. 3) also indicate the complexity of the filter divergence issue in SOC systems and how it should be accounted for. As explained in section 2.2, one of the key reasons for filter

340 divergence is due to the prior state uncertainty being underestimated due to ignoring of model process error which results in the prior state being given progressively more and more weight in the assimilation phase. At more frequently measured sites, such as Askov B3 and B4, there are more assimilation steps, which would intuitively speed up the filter divergence issue. However, as can be seen in Eq. 6., the reduction in posterior uncertainty depends on the observation uncertainty, with less uncertain observations also reducing the posterior uncertainty more. Thus, at the Askov site, the measurement uncertainties

345 are large enough that it partially balances out the measurement frequency and the resulting forecast uncertainty is large enough to allow for rapid adaptation to changes in the system.

At Versailles, though, while the measurements are much less frequent, they also have small associated uncertainties, especially the first few ones. Furthermore, for long decomposition systems like this, the uncertainty propagation within the ensemble is

350 so slow that it only marginally increases the state uncertainty until the next observation point, resulting in filter divergence becoming an even more pronounced issue here. As a result, even with uncertainty inflation, the first few assimilation steps reduce the state uncertainties to the degree that the difference between projections and measurements affects the state estimate much less than at the other sites. The new observations still affected the inflated SDA, as can be seen at the last Versailles measurement in Fig 3, but it will take multiple observations with increasing difference between forecasted and measured state

355 for SDA to properly adjust to the new state. This highlights the importance to carefully consider the relationships between the

observation uncertainty, frequency and inflation in order to improve the assimilation results. This is also a general issue within the application of SDA in geosciences and, as such, there have already been attempts to mathematically address it such as Li et al. (2009).

5. Conclusion

The results here show that there are benefits in implementing SDA methods in SOC research as assimilating the first few observations was generally sufficient in improving long-term SOC projections when compared to later measurements. Furthermore, the SDA methods here successfully used coarse bulk C measurements to update the more detailed model state with the developing error covariance matrices connecting the different state variables. The work also highlights the need for additional study such as, for example, how to best address the filter divergence issue or what is driving the differences in how SDA performs at different sites. The focus here was in a very simple system where there was no litter input and on a specific SDA method with its own benefits and hindrances. Increasing the complexity of the system, such as by introducing different types of litter, using measurements from other locations to estimate local SOC or incorporating flux tower respiration measurements to constrain projected SOC changes, will raise new practical challenges that have to be addressed in future work. Still, by allowing actively incorporating multiple information sources, SDA is a crucial tool for all process-based model projections, for example approximating the amount of SOC in a forest or assessing how agricultural carbon allocation changes in response to field management.

Acknowledgments

We thank the two anonymous reviewers for the constructive feedback for improving this manuscript. Academy of Finland (grant numbers 325549, 297350, 277623) and ERA-NET FACCE ERA-GAS project FORCLIMIT are acknowledged for financial support. FACCE ERA-GAS has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 696356. We are grateful to the persons and institutions who provided data from the long-term experiments. In particular we are grateful to Vladimir Romanenkov (Kursk site, Lomonosov Moscow State University, All-Russian Institute for Agrochemistry), Folkert van Oort and Sabine Houot (the two French sites, French National Institute for Agricultural Research and Agroparistech) and Bent T. Christensen (Askov Experimental Station, Dept. Agroecology, Aarhus University). We also thank the Lawes Agricultural Trust and Rothamsted Research for data from the e-RA database. The Rothamsted Long-term Experiments National Capability (LTE-NC) is supported by the UK BBSRC (Biotechnology and Biological Sciences Research Council, BBS/E/C/000J0300) and the Lawes Agricultural Trust.

Author contributions

Dr. Toni Viskari came up with the study, planned the experiments and wrote majority of the manuscript. M.Sc. Maisa Laine did the simulations and the initial analysis of the results. Doc. Liisa Kulmala provided the data and expertise on the

390 measurement time series used in the study. Drs. Jarmo Mäkelä and Istem Fer provided mathematical insight into the methods used as well as the interpretation of the results. Prof. Jari Liski is the PI of the project this research was a part of and has created the Yasso model used here.

Data availability

395 The Yasso model and the parameters used here can be downloaded from https://github.com/YASSOmodel/YASSO15. The data and scripts used to run the model and produce the figures can be accesses at https://github.com/Viskari/Viskari_SDA_Data. The permanent version of the Yasso15 code and data used in this publication has also been uploaded to Zenodo (doi:10.5281/zenodo.4041038).

References

400 Abramoff, R., Xu, X., Hartman, M., O'Brien, S., Feng, W., Davidson, E., Finzi, A., Moorhead, D., Schimel, J., Torn, M., and Mayes, M.A.: The MILLENIAL model: in search of measurable pools and transformations for modeling soil carbon in the new century. Biogeochemistry, 137(1-2), 51-71, 2017

Anderson, J.L.: An ensemble adjustment Kalman Filter for data assimilation. Mon Weather Rev, 129, 2884-2903, 2001

Anderson, J.L.: An adaptive covariance inflation error correction algorithm for ensemble filters. Tellus A, 2, 210-224, 2006

405 Anderson, J. L., Hoar, T., Raeder, K., Liu, H., Collins, N., Torn, R. and Arellano, A.: The Data Assimilation Research Testbed: A community facility. Bull. Amer. Meteor. Soc., 90, 1283–1296, 2009

Barré, P., Eglin, T., Christensen, B.T., Ciais, P., Houot, S., Kätterer, T., van Oort, F., Peylin, P., Poulton, P.R., Romanenkov, V. and Chenu, C.: Quantifying and isolating stable soil organic carbon using long-term bare fallow experiments. Biogeosciences, 2010(7), 3839-3850. ⟨10.5194/bg-7-3839-2010⟩. ⟨bioemco-00576270⟩, 2010

410 Berg, B., Booltink, H., Breymeyer, A., Ewertsson, A., Gallardo, A., Holm, B., Johansson, M.B., Koivuoja, S., Meentemeyer, V., Nyman, P., Olofsson, J., Pettersson, A.-S., Reurslag, A., Staaf, H., Staaf, I. & Uba, L., Berg, B., Booltink, H., Breymeyer, A., Ewertsson, A., Gallardo, A., Holm, B., Johansson, M.B., Koivuoja, S., Meentemeyer, V., Nyman, P., Olofsson, J., Pettersson, A.-S., Reurslag, A., Staaf, H., Staaf, I. & Uba, L.: Data on Needle Litter Decomposition and Soil Climate as Well as Site Characteristics for Some Coniferous Forest Sites, Part I, Site Characteristics. Report 41. Swedish University of

415 Agricultural Sciences, Department of Ecology and Environmental Research, Uppsala, 1991a

Berg, B., Booltink, H., Breymeyer, A., Ewertsson, A., Gallardo, A., Holm, B., Johansson, M.B., Koivuoja, S., Meentemeyer, V., Nyman, P., Olofsson, J., Pettersson, A.S., Reurslag, A., Staaf, H., Staaf, I., Uba, L.: Data on Needle Litter Decomposition and Soil Climate as Well as Site Characteristics for Some Coniferous Forest Sites, Part II, Decomposition Data. Report 42. Swedish University of Agricultural Sciences, Department of Ecology and Environmental Research, Uppsala, 1991b

420 Bradford, M.A., Wieder, W.R., Bonan, G.B., Fierer, N., Raymond, P.A. and Crowther, T.W.: Managing uncertainty in soil carbon feedbacks to climate change. Nat Clim Change, 6, 751-758, 2016

Camino-Serrano, M., Guenet, B., Luyssaert, S., Ciais, P., Bastrikov, V., De Vos, B., Gielen, B., Gleixner, G, Jornet-Puig, A., Kaiser, K., Kothawala, D., Lauerwald, R., Penuelas, J., Schrumpf, M., Vicca, S., Vuichard, N., Walmsley, D., and Janssens, I.A.: ORCHIDEE-SOM: Modeling soil organic carbon (SOC) and dissolved organic carbon (DOC) dynamics along vertical

425 soil profiles in Europe. Geoscientific Model Development, 11(3), 937-957. doi:10.5194/gmd-11-937-2018, 2018

Ciais, P., Sabine, C., Bala, G., Bopp, L., Brovkin, V., Canadell, J., Chhabra, A., DeFries, R. et al.: Carbon and Other Biogeochemical Cycles. In: Stocker, T.F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S.K., Boschung, J., Nauels, A., Xia, Y., Bex, V., & Midgley, P.M. (eds.) Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press,

430 Cambridge, United Kingdom and New York, NY, USA, 2013

Chapin, F.S., Matson, P.A. and Vitousek, P.: Principles of Terrestrial Ecosystem Ecology, 2nd edn. Springer-Verlag, New York, USA, 2011

Christensen, B.T.: Effect of cropping system on the soil organic matter content II, Field experiments on a sandy loam 1956-1985. Tidsskrift for Planteavl 94, 161-169. (in Danish with English summary), 1990

435   Christensen, B.T. and Johnston, A.E.: Soil organic matter and soil quality - lessons learned from long-term experiments at Askov and Rothamsted. Dev. Soil Sci. 25, 399-430, 1997

Cornwell, W.K., Cornelissen, J.H.C., Amatangelo, K., Dorrepaal, E. Eviner, V.T., Godoy, O., Hobbie, S.E., Hoorens, B., Kurokawa, H., Perez-Harguindeguy, N., Quested, H.M., Santiage, L.S., Wardle, D.A., Wright, I.J., Aerts, R., Allison, S.D., van Bodegom, P., Brovkin, V., Chatain, A., Callaghan, T.V., Diaz, S., Garnier, E., Gurvich, D.E., Kazakou, E., Klein, J.A.,

440   Read, J., Reich, P.B., Soudzilovskaia, N.A., Vaieretti, M.V., and Westoby, M.: Plant species traits are the predominant control on litter decomposition rates within biomes worldwide. Ecol Lett, 11(10), 1065-1071, 2008

Del Grosso, S., Parton, W., Stohlgren, T., Zheng, D., Bachelet, D., Prince, S., Hibbard, K. and Olson, R.: Global potential Net Primary Production predicted from vegetation class, precipitation and temperature. Ecology, 89(8), 2117-2126, 2008

Dietze M.C.: Ecological Forecasting. Princeton Univ Press, Princeton, 2017

445   Dietze, M.C., Fox, A., Beck-Johnson, L.M., Betancourt, J.L., Hooten, M.B., Jarnevich, C.S., Keitt, T.H., Kenney, M.A., Laney, C.M., Larsen, L.G., Loescher, H.W., Lunch, C.K., Pijanowski, B.C., Randerson, J.T., Read, E.K., Tredennick, A.T., Vargas, R., Weathers, K.C., and White, E.P.: Iterative near-term ecological forecasting, Proceedings of the National Academy of Sciences , 115 (7) 1424-1432; DOI: 10.1073/pnas.1710231115, 2018

Elbern, H., Schimdt, H., Talagrand, O. and Ebel, A.: 4D-variational data assimilation with an adjoint air quality model for

450   emission analysis. Environ. Mod. Softw., 15, 539-548, 2000

Evensen, G.: Sampling strategies and square root analysis schemes for EnKF. Ocean Dyn, 54(6), 539-560, 2004

Evensen, G.: Data Assimilation: The Ensemble Kalman Filter. Springer, New York, 2009

Gao, C., Wang, H., Weng, E., Lakshmivaran, S., Zhang, Y. and Luo, Y.: Assimilation of multiple data sets with Kalman filter to improve forecasts of forest carbon dynamics. Ecological Applications: A Publication of the Ecological Society of America,

455   21(5), 1461–1473, https://doi.org/10.1890/09-1234.1, 2011

Gholz, H.L., Wedin, D.A., Smitherman, S.M., Harmon, M.E. and Parton, W.J.: Long- term dynamics of pine and hardwood litter in contrasting environments: toward a global model of decomposition. Glob Change Biol, 6, 751e765, 2000

Haario, H., Saksman, E. and Tamminen, J.: An adaptive Metropolis algorithm. Bernoulli, 7(2), 223-242, 2001

Hamill, T.M., Whitaker, J.S. and Snyder, C.: Distance-dependent filtering of background error covariance estimates in an

460   ensemble Kalman filter. Mon. Wea. Rev., 129, 2776–2790, 2001

Hararuk, O., Xia, J., and Luo, Y.: Evaluation and improvement of a global land model against soil carbon data using a Bayesian Markov chain Monte Carlo method. J Geophys Res Biogeosci, 119(3), 403–417, https://doi.org/10.1002/2013JG002535, 2014

He, Y., Trumbore, S.E., Torn, M.S., Harden, J.W., Vaughn, L.J.S., Allison, S.D. and Randerson, J.T.: Radiocarbon constraints imply reduced carbon uptake by soils during the 21st century. Science, 353(6306), 1419-1424, 2016

465 Jandl, R., Rodeghiero, M., Martinez, C., Cotrufo, M.F., Bampa, F., van Wesemael, Harrison, R.B., Guerrini, I.A., Richter jr, D.,dB., Rustad, L., Lorenz, K., Chabbi, A. and Miglietta, F.: Current status, uncertainty and future needs in soil organic carbon monitoring. Sci. Total Environ.*, 468-469, 376-383, 2014

Kalman, R.E.: A new approach to linear filtering and prediction problems. J. Basic Eng.*, 82(1), 35-45, 1960

Karhu, K., Gärdenäs, A.I., Heikkinen, J., Vanhala, P., Tuomi, M. and Liski, J.: Impacts of organic amendments on carbon
470 stock of an agricultural soil – Comparison of model-simulations to measurements. Geoderma, 189-190, 606-616, 2012

Karhu, K., Auffret, M.D., Dungait, J.A.J., Hopkins, D.W., Prosser, J.I., Singh, B.K., Subke, J.-A., Wookey, P.A., Ågren, G.I., Sebastia, M.-T., Gouriveau, F., Bergkvist, G., Meir, P., Nottingham, A.T., Salinas, N. and Hartley, I.P.: Temperature sensitivity of soil respiration rates enhanced by microbial community response. Nature, 513, 81–84, 2014

Kulmala, L. and Liski J.: Bare fallow experiments highlight the importance of long-term history on soil carbon decomposition
475 rate on agricultural lands. Report series in aerosol science N:o 215, p. 225-227. Available at: http://www.atm.helsinki.fi/FAAR/reportseries/rs-215.pdf, 2018

Le Dimet, F.-X. and O. Talagrang: Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. Tellus, 38A, 97–110, 1986

Lehmann, J. and Kleber, M.: The contentious nature of soil organic matter. Nature, 528, 60-68, 2015
480 Li, H., Kalnay, E. and Miyoshi, T.: Simultaneous estimation of covariance inflation and observation errors within an ensemble Kalman filter. Q J R Meterol Soc, 123(639), 523-533, 2009

Manzoni, S.P. and A. Porporato, A.: Soil carbon and nitrogen mineralization: Theory and models across scales. Soil Biol Biochem, 41(7), 1355-1379, 2009

Mao, Z., Derrien, D., Didion, M., Liski, J., Eglin, T., Nicolas, M., Jonard, M. and Saint-Andre, L.: Modeling soil organic
485 carbon dynamics in temperate forests with Yasso07. Biogeosciences, 16(9), 1955-1973, 2019

Orchard, V.A., and Cook, F.J.: Relationship between soil respiration and soil moisture. Soil Biol Biochem, 15(4), 447-453, 1983

Menichetti, L., Ågren, G.I., Barré, P. Moyano F., Kätterer T.: Generic parameters of first-order kinetics accurately describe soil organic matter decay in bare fallow soils over a wide edaphic and climatic range. Sci Rep 9**,** 20319, 2019
490 Ortiz, C.A., Liski, J., Gärdenäs, A.I., Lehtonen, A., Lundblad, M., Stendahl, J., Ågren, G.I. and Karltun, E.: Soil organic carbon stock changes in Swedish forest soils—A comparison of uncertainties and their sources through a national inventory and two simulation models. Ecol. Model.*, 251, 221-231, 2013

Parton, W.J.: The CENTURY model. In: Powlson, D.S., Smith, P., Smith, J.U. (eds) Evaluation of Soil Organic Matter Models. NATO ASI Series (Series I: Global Environmental Change), vol 38, Springer, Berlin, Heidelberg, 1996
495 Palosuo, T., Foereid, B., Svensson, M., Shurpali, N., Lehtonen, A.,,, Herbst, M., Linkosalo, T., Ortiz, C., Todorovic, G.R., Marcinkonis, S., Li, C. and Jandl, R.: A multi-model comparison of soil carbon assessment of a coniferous forest stand. Environ Modell Softw, 35, 38-49, 2012

Schlee, F.H., Standish, C.J., and Toda, N.F.: Divergence in the Kalman filter. AIAA J, 5(6), 1114-1120, 1967

Smith, P., Soussana, J.-F., Angers, D., Schipper, L., Chenu, C., Rasse, D.P. et al.: How to measure, report and verify soil carbon change to realize the potential of soil sequestration for atmospheric greenhouse gas removal. Glob Change Biol, 26(1), 219-241, 2020

Sulman, B.N., Moore, J.A.M., Abramoff, R., Averill, C., Kivlin, S., Georgiou, K., Sridhar, B., Hartmann, M.D., Wang, G., Wieder, W., Bradford, M.A., Luo, Y., Mayer, M.A., Morrison, E., Riley, W.J., Salazar, A., Schimel, J.P., Tang, J. and Classen, A.T.: Multiple models and experiments underscore large uncertainty in soil carbon dynamics. Biogeochem, 141, 109-123, 2018

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, 2014

Trofymow, J.A. and the CIDET Working Group: The Canadian Intersite Decomposition Experiment (CIDET): Project and Site Establishment Report. Information Report BC-X-378. Pacific Forestry Centre, Victoria, Canada, 1998

Tuomi, M., Vanhala, P., Karhu, K., Fritze, H., and Liski, J.: Heterotrophic soil respiration-Comparison of different models describing its temperature dependence, Ecol Modell, 211, 182–190, doi:10.1016/j.ecolmodel.2007.09.003, 2008.

Tuomi, M., Laiho, R., Repo, A. and Liski, J.: Wood decomposition model for boreal forests. *Ecol Model*, 222(3), 709-718, 2011

Trudinger, C. M., Raupach, M. R., Rayner, P. J., & Enting, I. G.: Using the Kalman filterfor parameter estimation in biogeochemical models. Environmetrics, 19(8), 849–870,https://doi.org/10.1002/env.910, 2008

van Oijen, M.: Bayesian Methods for Quantifying and Reducing Uncertainty and Error in Forest Models. *Current Forestry Reports*. **3:** 269–280. https://doi.org/10.1007/s40725-017-0069-9, 2017

van Oort, F., Paradelo, R., Proix, N., Delarue, G., Baize, D., Monna, F. 2018. Centennial fertilization-induced soil processes control trace metal dynamics. Lessons from a long-term bare fallow experiment. Soil Systems, 2, 23. Special issue 'Soil processes controlling contaminant dynamics' https://doi.org/10.3390/soilsystems2020023.

Viskari, T. Asmi, E., Virkkula, A., Kolmonen, P., Petäjä, T. and Järvinen, H.: Estimation of aerosol particle number distribution with Kalman Filtering – Part 2: Simultaneous use of DMPS, APS and nephelpmeter measurements. Atmos Chem Phys, 12, 11781-11793, 2012

Viskari, T., Hardiman, B., Desari, A.R. and Dietze M.C.: Model-data assimilation of multiple phenological observations to constrain and predict leaf area index. Ecol. Appl., 25(2), 546–558, 2015

Vogel, C., Heister, K., Buegger, F., Tanuwidjaja, I., Haug, S., Schloter, M. and Kögel-Krabner, I.: Clay mineral composition modifies decomposition and sequestration of organic carbon and nitrogen in fine soil fractions. Biol Ferti Soils, 51, 427–442**, 2015

Weaver, A.T., Vialard, J. and Anderson, D.L.T.: Three- and Four-Dimensional Variational Assimilation with a General Circulation Model of the Tropical Pacific Ocean. Part I: Formulation, Internal Diagnostics, and Consistency Checks. Mon. Wea. Rev., 131, 1360-1378, 2003

Wutzler, T. and Reichstein, M.: Soils apart from equilibrium – consequences for soil carbon balance modelling, Biogeosciences, 4, 125–136, https://doi.org/10.5194/bg-4-125-2007, 2007

535 Yang, Y., Dunham, E., Barnier, G. and Almquist, M.: Tsunami Wavefield Reconstruction and Forecasting using the Ensemble Kalman Filter. Geophys Res Lett, 46(2), 853-860, 2019

Yan, M., Li, Z., Tian, X., Zhang, L., & Zhou, Y.: Improved simulation of carbon and waterfluxes by assimilating multi-layer soil temperature and moisture into process-basedbiogeochemical model. Forest Ecosystems, 6(1), 12, https://doi.org/10.1186/s40663-019-0171-5, 2019

540 Ziche, D., Gruneberg, E., Hilbrig, L., Höhle, J., Kompa, T., Liski, J., Repo, A. and Wellbrock, N.: Comparing soil inventory with modelling: Carbon balance in central European forest soils varies among forest types. Sci. Total Environ., 647, 1573-1585, 2019

Site	Askov (B3, B4)	Grignon	Rothamsted	Ultuna	Versailles
Country	DE	FR	UK	SWE	FR
Mean temperature (°C)	7.8	10.7	9.5	5.5	10.7
Precipitation (mm)	862	649	712	533	628
History	arable	Arable	grassland	grassland	grassland
Bare fallow starting	1956	1959	1959	1956	1928
Soil type (FAO)	Luvisol	Luvisol	Luvisol	Cambisol	Luvisol
Clay-silt-sand (%)	7-11-82	30-54-16	25-62-13	36-41-23	17-57-26
Bulk density (kg dm^{-3})	1.5	1.2	0.94	1.44	1.3
Fertilization	✓	–	–	–	–
Tillage frequency	frequent	2/year	2-4/year	1/year	2/year
Weeding by hand	✓	✓	–	✓	✓
Measurement time series	1956-1985	1959-2007	1959-2008	1956-2007	1929-2008
Initial AWENH carbon pools (tC/ha)	(4.4,0.5,0.3,8.9,38.1) (3.4,0.3,0.2,6.9,36.9)	4.7,0.5,0.3,11.3,25.0	10.7,1.1,0.6,23.5,35.8	5.9,0.6,0.4,12.9,22.7	8.8,0.9,0.5,20.8,34.5

545

Table 1: The bare fallow sites used in this study. The Askov site was fertilized by 70 kg N/ha until 1973 and by 100 kg after that. Before bare fallow, Askov was cultivated since 1800, Grignon since 1875, Kursk since app. 1765, and Versailles since 17th century. Ultuna has been experimental field for agriculture for centuries. There are two different plots at Askov site (B3 and B4) with different initial state values.

550

Experiments	Non-SDA Yasso		First measurement assimilated		First two measurements assimilated		First three measurements assimilated		First four measurements assimilated	
Sites	RMSE (t C ha^{-1})	MRE (%)	RMSE (t C ha^{-1})	MRE (%)	RMSE (t C ha^{-1})	MRE (%)	RMSE (t C ha^{-1})	MRE (%)	RMSE (t C ha^{-1})	MRE (%)
Askov B3	2.5	7.0	2.6	7.3	2.5	7.0	3.3	9.4	3.5	9.8
Askov B4	4.0	12	4.1	12	5.8	17	6.0	18	5.9	18
Grignon	2.6	8.6	2.8	9.4	1.9	6.1	1.5	4.2	1.0	2.9
Rothamsted	4.9	14	4.7	14	3.7	11	2.0	6.1	2.3	7.0
Ultuna	3.2	12	3.3	12	0.8	2.8	0.5	1.5	0.8	2.9
Versailles	6.6	26	6.9	28	7.2	29	7.5	30	5.4	22

Table 2: The root mean square error (RMSE) as well as the mean relative error respective to the observation for the three last measurement at each site. The unit for the RMSE values is t C ha^{-1}

555

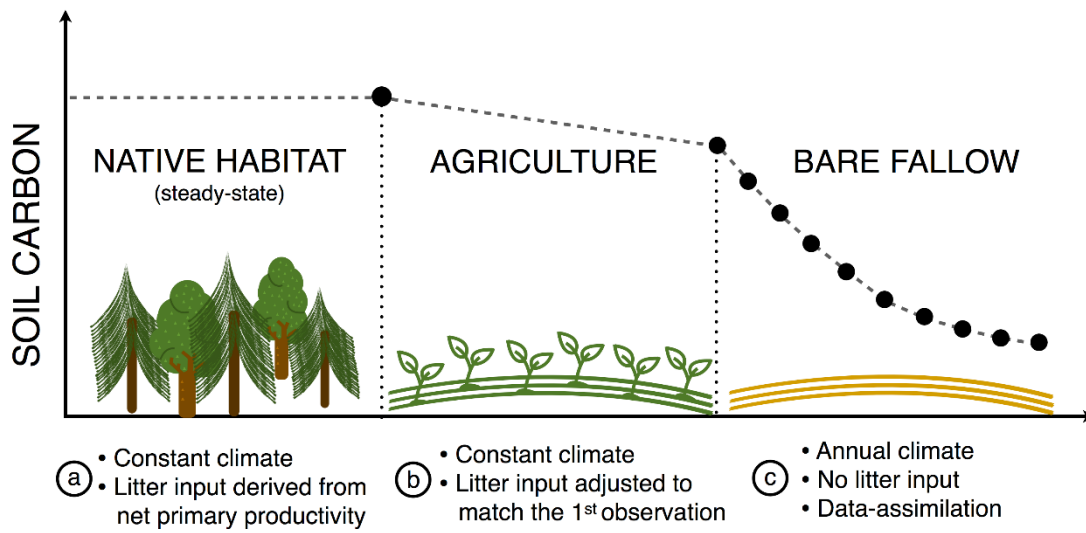



Figure 1: Conceptual figure on how the initial states for the model runs are approximated.

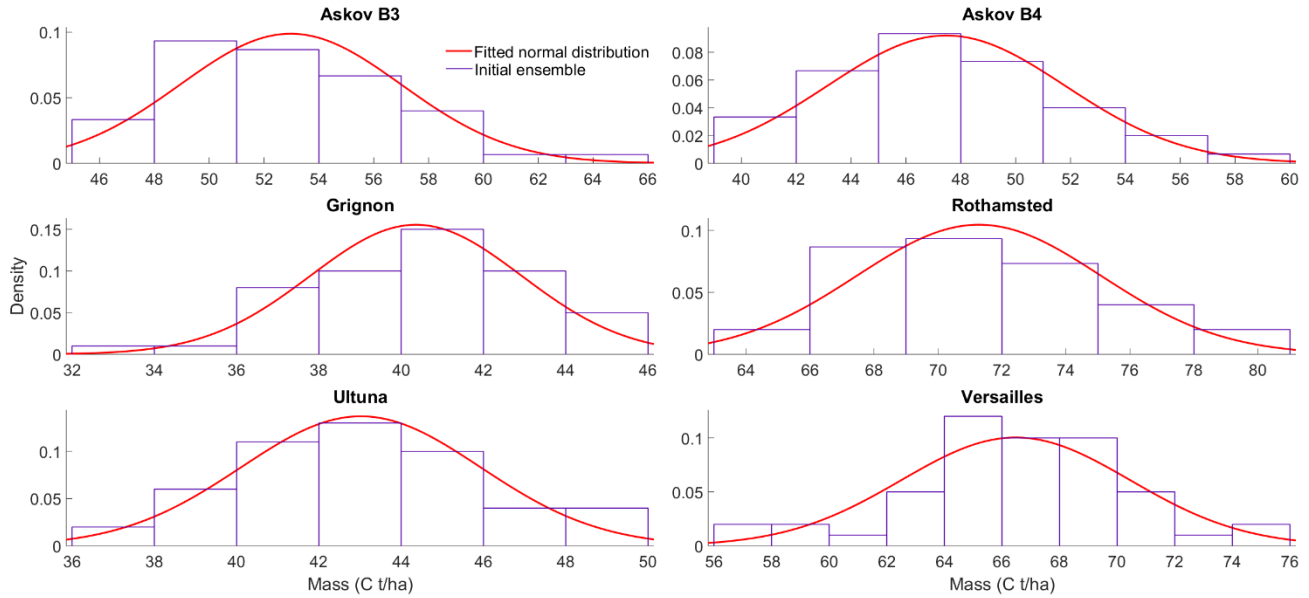



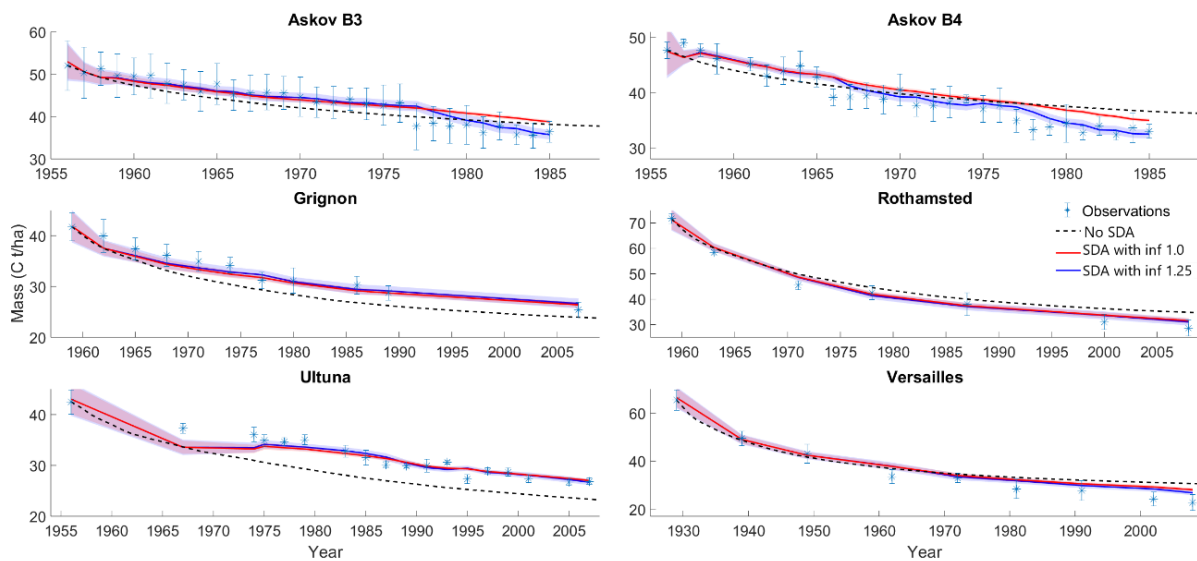
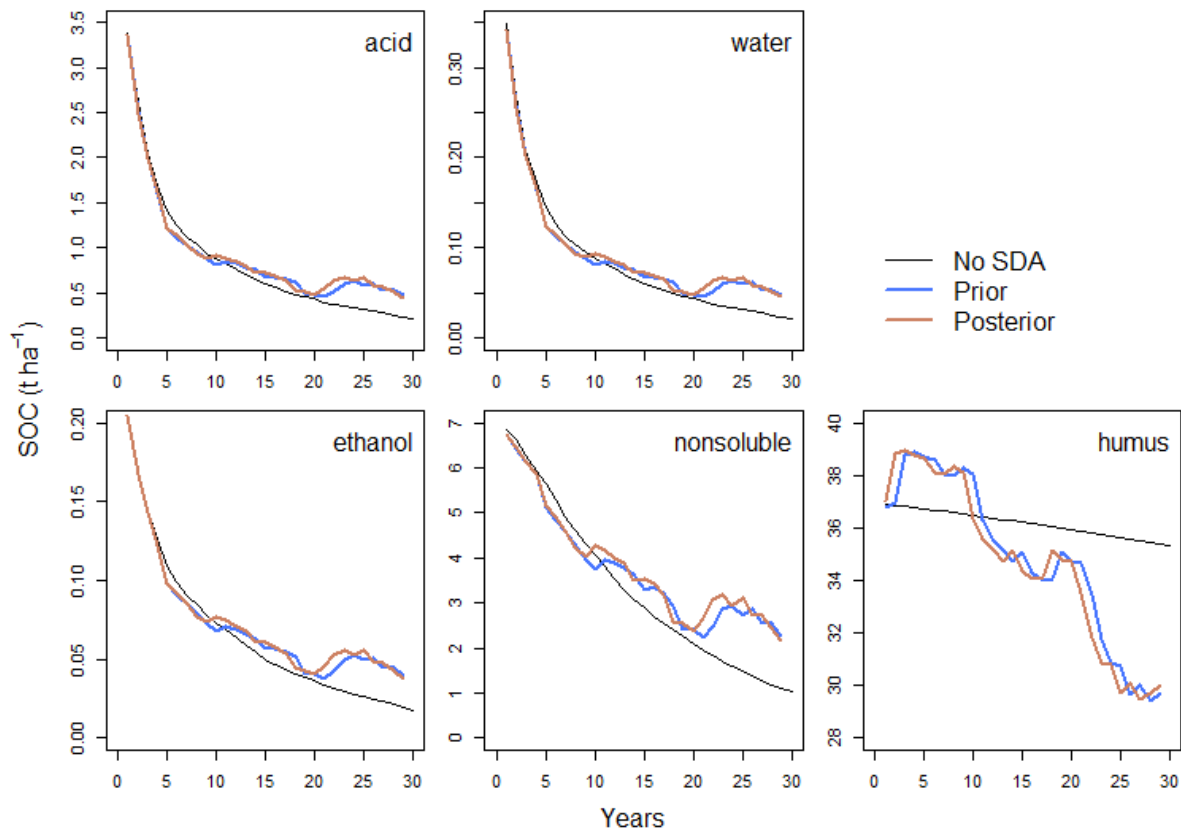
Figure 2: The initial ensemble states of total soil carbon (t/ha) at the study sites in the beginning of the fallow campaign.

 Fig

Figure 3: Observed and modelled SOC with and without data assimilation (SDA) of all previous measurements using two different inflation factors (inf). The coloured area around the two different SDA estimates are the 95 % confidence interval. SDA with higher inflation factor improved predictions at all sites while the, the SDA with lower inflation factor was susceptible to filter divergence.

570

575 **Figure 4**: The mean of prior and posterior distribution SOC pools at Askov B4 before and after each assimilated observation. Different SOC pools showed different responses to SDA where humus pool was adjusted the most in response to the observations and always to their direction. AWEN pool dynamics responded SDA later over the course of assimilation.

580

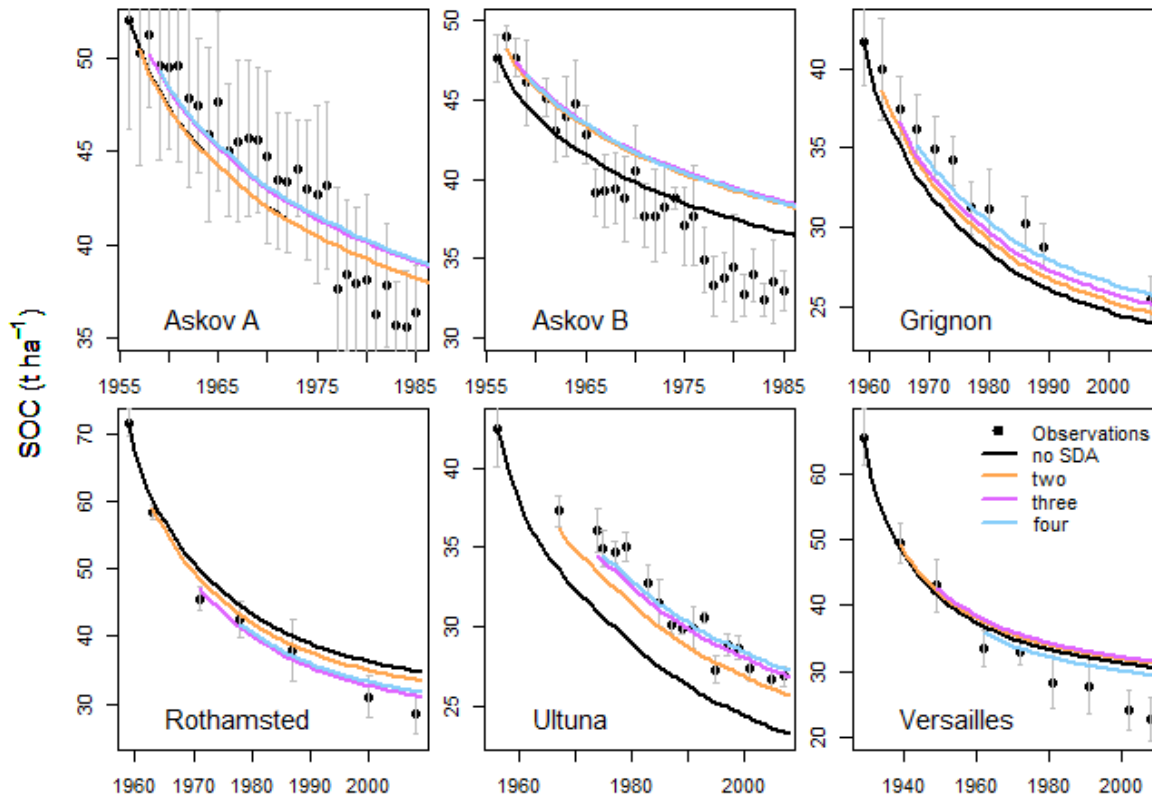



Figure 5: Observed and forecasted SOC without data assimilation (black) and with 2-4 initial observations assimilated (coloured lines). Incorporating information already from the first two observations had a noticeable impact on the time series
585   prediction.

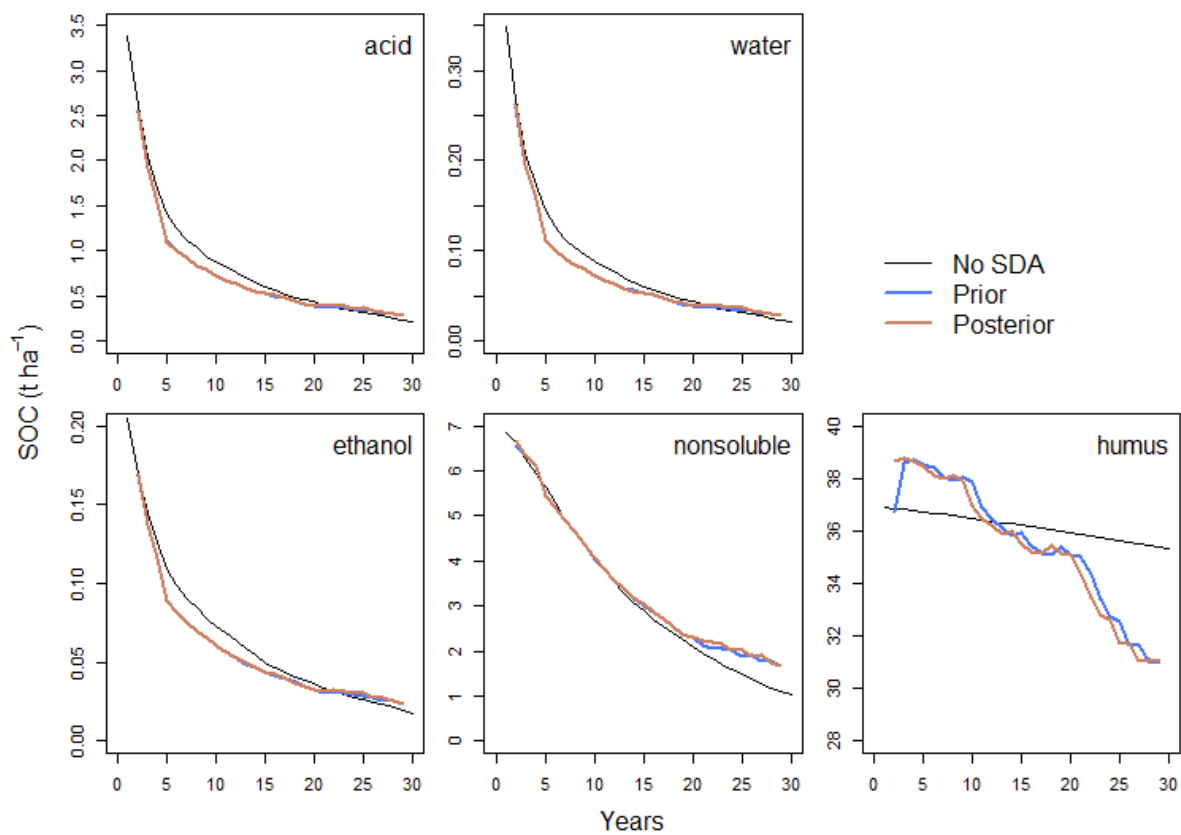



Figure 6. The mean prior and posterior distribution SOC pools at Askov B4 before and after each assimilated observation if the assimilation begins from the second observation. While the AWEN pools still show an opposite shift to the H pool later in the assimilation cycle, it is smaller than in Fig 3.