



Rate My Data: Quantifying the Value of Ecological Data for Models of Terrestrial Carbon Cycle

Citation

Keenan, Trevor F., Eric Davidson, James W. Munger, Andrew D. Richardson. 2013. Rate my data: Quantifying the value of ecological data for models of terrestrial carbon cycle. *Ecological Applications* 23(1): 273-286.

Published Version

doi:10.1890/12-0747.1

Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:10621295>

Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Open Access Policy Articles, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#OAP>

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

**Rate my data: quantifying the value of ecological data for the development of
models of the terrestrial carbon cycle**

Trevor F. Keenan¹, Eric Davidson², James W. Munger³, Andrew D. Richardson¹

¹ Organismic and Evolutionary Biology, Harvard University, Cambridge, 02138 MA,
USA.

² Woods Hole Research Center, 149 Woods Hole Road, Falmouth, 02540 MA, USA

³ School of Engineering and Applied Sciences and Department of Earth and Planetary
Sciences, Harvard University, Cambridge, 02138 MA, USA

Corresponding author:

T.F. Keenan

Phone: +1 (617) 496-0825

Email: tkeenan@oeb.harvard.edu

Paper type: Primary research article

Abstract

Primarily driven by concern about rising levels of atmospheric CO₂, ecologists and earth system scientists are collecting vast amounts of data related to the carbon cycle. These measurements are generally time-consuming and expensive to make, and, unfortunately, we live in an era where research funding is increasingly hard to come by. Thus, important questions are: ‘Which data streams provide the most valuable information?’ and, ‘How much data do we need?’ These questions are relevant not only for model developers, who need observational data to improve, constrain and test their models, but also for experimentalists and those designing ecological observation networks.

Here we address these questions using a model-data fusion approach. We constrain a process-oriented, forest ecosystem C cycle model with seventeen different data streams from the Harvard Forest. We iteratively rank each data source according to its contribution to reducing model uncertainty. Results show the importance of some measurements commonly unavailable to carbon cycle modelers, such as estimates of turnover times from different carbon pools. Surprisingly, many data sources are relatively redundant in the presence of others, and do not lead to a significant improvement in model performance. A few select data sources lead to the largest reduction in parameter based model uncertainty. Projections of future carbon cycling were poorly constrained when only hourly net ecosystem exchange measurements were used to inform the model. They were well constrained, however, with only five of the seventeen data streams, even though many individual parameters are not constrained. The approach taken here should

42 stimulate further cooperation between modelers and measurement teams, and may be
43 useful in the context of setting research priorities and allocating research funds.

44

45 **Key Words:** process-based models, carbon fluxes, biosphere-atmosphere interaction,
46 carbon sequestration, data assimilation

47

Introduction

In recent years our ability to collect vast amounts of data related to the structure and function of the biosphere, at both high temporal and spatial frequency, has greatly increased (Luo et al. 2008). New large-scale monitoring through networks such as NEON (www.neoninc.org), ICOS (www.icos-infrastructure.eu), FLUXNET (www.fluxnet.ornl.gov), and LTER (www.lternet.edu), along with the extended satellite record, and data collation efforts such as TRY (www.try-db.org, Kattge et al. 2011), are amassing tremendous amounts of data. However, the ultimate value of the accumulating diverse data sources will depend on the extent to which the data can be used to improve our understanding of, and ability to model, the earth system.

One of the main motivations for the increase in data availability is the need to improve our understanding of terrestrial carbon cycling (IPCC, 2007). Much of the anthropogenically emitted CO₂ cycles through terrestrial ecosystems. Current estimates of CO₂ removed from the atmosphere by global photosynthesis stand at around 120PgC (Beer et al. 2010). A slightly smaller amount is respired back into the atmosphere, giving an estimated net global carbon sink in terrestrial ecosystems of ~1-2PgC (Le Quere et al. 2009, Pan et al. 2011). The main biological processes of photosynthesis and respiration that drive this cycle have long been identified. Large uncertainty remains, however, as to the mechanisms controlling the response of these processes to drivers at different spatial and temporal scales. This uncertainty is reflected in the broad range of model projections of the future of global terrestrial carbon storage (Friedlingstein et al. 2006, Heimann &

Reichstein, 2008), making the implementation of effective policy difficult at best (IPCC, 2007).

New approaches that can combine models with multiple data sources - “model-data fusion” - are emerging as a means to better understand the dominant processes controlling terrestrial carbon cycling. Such techniques can be employed both to directly inform carbon cycle models and as a tool to synthesize the growing amounts of data. The basic philosophy is that using data in a statistically rigorous manner to give the best model possible (conditional on model structure) can both highlight model deficiencies and integrate different data sources. Recent efforts have used a diverse range of data types with process-oriented models (e.g. Braswell et al. 2005, Williams et al. 2005, Sacks et al., 2007, Moore et al., 2008, Richardson et al. 2010, Weng & Luo 2011, Keenan et al. 2012b). A strength of the approach is that it can be used to assess the model against all observations simultaneously. Using multiple constraints goes beyond simple testing of a model against a single measurement type – the approach uses data both to test and inform model behavior for all aspects of the system for which observations are available. The result is a data-informed process-oriented model, which allows the researcher to quantify the degree of uncertainty in model projections.

Carbon cycle modelers typically rely on experimental and observational data that have been collected by others. One of the most common questions asked by experimentalists and (more recently) data acquisition network designers of modelers is “what data are most useful?” In response to such questions, however, modelers generally do not have a

93 better answer than what is essentially an educated guess. Indeed, from a modeling
94 perspective using more data does not always lead to a better-constrained model
95 (Richardson et al. 2010). In an environment of increasingly organized data acquisition
96 networks (Keller et al. 2008) and efforts that seek to merge models with data (Wang et al.
97 2007, Keenan et al. 2011a), it becomes imperative to develop ways of quantifying the
98 usefulness of different data sources. By identifying the next measurement that should be
99 made, which maximizes the information gained from all measurements together, the
100 efficiency and cost-effectiveness of measurement campaigns can be improved, along with
101 model projections.

102
103 Here, we develop a framework to address the question, “How useful is a particular
104 measurement for reducing uncertainty in a process-oriented model of terrestrial carbon
105 cycling?” We use multiple data sources from long-term records at the Harvard Forest, in
106 the northeastern US, in combination with a model-data fusion framework. We rank the
107 different data streams according to the incremental information that each data stream
108 provides. We do this by iteratively testing the reduction in model uncertainty achieved by
109 informing the model with each data source. At each step in the process, we assess the
110 impact of a particular measurement type on both short-term (diurnal, seasonal, annual)
111 model projections, and long-term (decadal) model responses to climate change.

114 **Materials and Methods**

Site

Hourly model simulations were run for 12 complete years (1992-2003) at the Harvard Forest Environmental Measurement Site (HFEMS) (<http://atmos.seas.harvard.edu/lab/hf/index.html>), located in the northeastern United States (42.53N 72.17W, elevation 340m) (Wofsy et al. 1993, Goulden et al. 1996, Barford et al. 2001, Urbanski et al. 2007). Measurements and simulations pertain to the area within the EMS tower footprint, which is largely comprised of deciduous trees. The area is dominated by the deciduous species red oak (*Quercus rubra*, 52% basal area), red maple (*Acer rubrum*, 22% basal area), with a small conifer component that includes eastern hemlock (*Tsuga Canadensis*, 17% basal area), and occasional white and red pine (*Pinus strobus*, *Pinus resinosa*).

Data

All data used were gathered between 1992 and 2003. We used hourly meteorological and eddy-covariance (Wofsy et al. 1993, Urbanski et al. 2007) measurements of net ecosystem exchange (NEE) (<http://atmos.seas.harvard.edu/lab/data/nigec-data.html>). Gap-filled meteorological variables used include hourly incident photosynthetically active radiation (PAR), air temperature above the canopy, soil temperature at a depth of 5 cm, vapor pressure deficit, and atmospheric CO₂ concentrations.

Quality controlled hourly eddy-covariance observations (without gap-filling) of NEE were used to constrain parameters of an ecosystem model. Gap-filled data, or model-based partitioning of NEE to respiration and photosynthesis components, were not used.

For ancillary data constraints we used 15 different data sources, which included measurements of leaf area index, soil organic carbon content, carbon in roots, carbon in wood, wood carbon annual increment, observer-based estimates of bud-burst and leaf senescence, leaf litter, woody litter, soil carbon turnover times, and three different measurement sets of soil respiration that capture spatial and methodological variability (Table 1). These data are freely available from the Harvard Forest Data Archive (<http://harvardforest.fas.harvard.edu/data/archive.html>) or the references in Table 1.

Measurement based estimates of uncertainty were used for each data stream in the optimization. Flux uncertainty estimates were taken from Richardson et al. (2006), where uncertainties were shown to follow a double-exponential distribution, with the standard deviation of the distribution specified as a linear function of the flux. Soil respiration uncertainty estimates were taken from Savage et al. (2009) and Phillips et al. (2010), where measurement uncertainty increased linearly with the magnitude of the flux. Estimates of uncertainties for the remaining data streams were based on either spatial variation or standard deviations from repeat sampling. Full details of uncertainty estimates are given in Keenan et al. (2012b).

The FöBAAR Model

We used a forest carbon cycle model that strikes a balance between parsimony and detailed process representation. Working on an hourly timescale, FöBAAR (Forest Biomass, Assimilation, Allocation and Respiration; Keenan et al. 2012b) calculates photosynthesis from two canopy layers, and respiration from eight carbon pools (leaf,

wood, roots, soil organic matter [microbial, slow and passive pools], leaf litter and [during phenological events] mobile stored carbon). Meteorological drivers considered are: canopy air temperature (T_a), 5 cm soil temperature (T_s), photosynthetic active radiation (PAR), vapor pressure deficit (VPD), and atmospheric CO_2 . Model parameters are given in Table 1.

The canopy in FöBAAR is described in two compartments representing sunlit and shaded leaves (Sinclair et al., 1976; Wang & Leuning, 1998). Canopy light penetration depends on the position of the sun, and the area of leaf exposed to the sun based on leaf angle and the canopy's ellipsoidal leaf distribution (Campbell, 1986), assuming a spherical leaf angle distribution. Assimilation rates are calculated via the Farquhar approach (Farquhar et al., 1980; De Pury & Farquhar, 1997). Stomatal conductance is calculated using the Ball–Berry model (Ball et al., 1987), coupled to photosynthetic rates through the analytical solution of the Farquhar, Ball Berry coupling (Baldocchi, 1994).

Maintenance respiration is calculated as a fraction of assimilated carbon. The remaining assimilate is allocated to different carbon pools (foliar, wood and root) on a daily time step. Root respiration is calculated hourly and coupled to photosynthesis through the direct allocation to roots. Dynamics of soil organic matter is modeled using a three-pool approach (microbial, slow, and passive pools) (Knorr et al., 2005). Decomposition in each pool is calculated hourly, with a pool specific temperature dependency. Litter decomposition is also calculated hourly, but on an air temperature basis. Litter and root

carbon are transferred to the microbial pool, then to the slow and finally to the passive pool. For further details on model structure see Keenan et al. (2012b).

Model-data fusion

An adaptive multiple constraints Markov-chain Monte Carlo approach was used to optimize the process-oriented model and explore model uncertainty. The algorithm uses the Metropolis-Hastings (M-H) approach (Metropolis and Ulam 1949, Metropolis et al. 1953, Hastings 1970) combined with simulated annealing (Press et al. 2007). Prior distributions for each parameter (Table 1) were assumed to be uniform (non-informative, in a Bayesian context).

The optimization process uses a two-step approach. In the first stage, the parameter space is explored for 100,000 iterations using the optimization algorithm. At each iteration the current step size is used as the standard deviation of random draws from a normal distribution with mean zero, by which parameters are varied around the previous accepted parameter set. This stage identifies the optimum parameter set by minimizing the cost function (see below). In the second stage, the parameter space is again explored using a Markov chain starting from the optimum identified in step 1. Acceptance of a parameter set is based on whether the cost function for each data stream (defined below) passes a χ^2 test (at 90% confidence) for acceptance/rejection, after variance normalization (e.g. Franks et al. 1999, Richardson et al. 2010).

The cost function quantifies the extent of model-data mismatch using all available data (eddy-covariance, biometric, etc.). Individual data stream cost functions, j_i , are calculated as the total uncertainty-weighted squared data-model mismatch, averaged by the number of observations for each data stream (N_i):

$$j_i = \left(\sum_{t=1}^{N_i} \left(\frac{y_i(t) - p_i(t)}{\delta_i(t)} \right)^2 \right) / N_i \quad (1)$$

where $y_i(t)$ is a data constraint at time t for data stream i and $p_i(t)$ is the corresponding model predicted value. $\delta_i(t)$ is the measurement specific uncertainty. For the aggregate multi-objective cost function we use the average of the individual cost functions, which can be written as:

$$J = \left(\sum_{i=1}^M j_i \right) / M \quad (2)$$

where M is the number of data streams used.

Each individual cost function is averaged by the number of observations for the relative data stream. The average of the cost functions from all data streams is taken as the total cost function. In this manner each data stream is given equal importance in the optimization (Franks et al. 1999, Barrett et al. 2005).

Experimental Set-up

We used a simple three-step iterative algorithm for the model experiment. The basic premise is to successively add data streams as model constraints, according to which data stream gives the best incremental reduction in model uncertainty.

1. For $i = 1$, perform model-data fusion with each measurement type in Table 1 individually.
2. Identify the single best measurement type, i.e., that which gives the minimum posterior distribution of model-data mismatch (see below).
3. For $i = 2 \dots M$, repeat steps 1 and 2 again to identify the next best measurement type (in addition to the data streams already selected). Do this until no more data streams are available.

We calculate the reduction in model uncertainty through the posterior distribution of model-data mismatch (the difference between modeled and observed variables, Eq. 2). At each iteration of Step 2 above, we calculate the model uncertainty using the entropy of the posterior distribution of model-data mismatch for each data combination. Entropy is a measure of the uncertainty associated with a random variable (Shannon 1948, Jaynes 1957, Kolmogorov 1968) and can be used to quantify the information gained by the use of a particular data source (e.g. Weng & Luo 2011). At each stage 2 in the above algorithm, we identify the best data combination as that which gives the lowest entropy (and thus lowest model uncertainty) in the posterior distribution of model data-mismatch.

Running the above algorithm took about 3 days on an 18-core computational cluster.

Climate projections to 2100

We used the extracted posterior parameter distributions to project carbon cycling and stocks to 2100 for each step in the above outlined experiment. This served as an additional means by which to quantify the incremental benefit of each additional data

251 stream. For the future climate scenario, we used downscaled data (Hayhoe et al. 2007)
252 from the regionalized projection of the GFDL-CM global coupled climate-land model
253 (Delworth et al. 2006) driven with socio-economic change scenario A1fi (IPCC 2007).
254 Mean annual temperature at Harvard forest, using this projection, is predicted to increase
255 from 7.1 to 11.9 °C, with an associated increase in atmospheric CO₂ to 969 ppm by 2100.

Results

What measurements are most important?

At each stage in the optimization process, we identified the next best measurement type by quantifying how much each data stream reduced the uncertainty in model projections (via Eq. 2). The most useful measurements were those that quantified how carbon flowed through the ecosystem at different time scales (Fig. 1). In particular, the combination of measurements on fast (net ecosystem exchange, soil respiration) and slow (soil carbon turnover rates, monthly/annual cumulative fluxes, litter from wood/leaves) carbon flows in the ecosystem lead to the largest improvement in model performance. Many measurements did not inform the model in the presence of others: for example the use of data on the size of the soil carbon pool did not lead to a large reduction in model uncertainty when soil respiration data was available along with turnover rates from the different soil carbon pools (Fig. 1). Estimates of bud-burst dates did not lead to a large reduction in model uncertainty, as they could be inferred by the model from the eddy-covariance CO₂ flux data. Observations of leaf senescence dates, on the other hand, were highly ranked. Autumn shifts in carbon cycling are driven by gradual biotic changes in canopy status, and co-occur with gradual abiotic changes in mean climate forcings. The senescence data, being biotic in nature, therefore improved the ability of the model to distinguish between autumn dynamics driven by biotic and abiotic changes. In addition to bud-burst data, litter turnover, and the proportion of autotrophic respiration in soil respiration measurements were ranked low, implying that the information contained in these measurements is also available from the higher ranked data (Fig. 1). The low

ranking of nighttime net ecosystem exchange is a good example of a situation where the information provided by a measurement is already present in another, as both annual and monthly NEE sums are constructed using night-time NEE data.

Figure 1. The iterative reduction in model posterior uncertainty.

The extent to which measurements can identify model parameters

When using all data, twenty-six of the forty parameters included could be effectively identified (parameters *a* to *y*, Table 2, Fig. 2). Here, we consider a parameter identifiable if the size of the posterior parameter distribution was half that of the prior distribution. In general, posterior parameter distributions were gradually reduced as more data streams were added to the system. Using all data together reduced the posterior parameter distributions by ~60% over all parameters (Fig. 3), when compared to the priors. The majority of the reduction in the range of posterior parameter distributions, however, was achieved with the use of relatively few data streams (Fig. 3). For example, fourteen parameters were well constrained with the use of only six different data sources (Fig. 2). The top ten parameters that were most informed by the data related to the respiration rates of the different soil carbon pools, phenology, and litterfall. Fourteen parameters were not constrained, even when using all data together (parameters *z* to *k*₂). These were predominantly related to canopy processes (e.g. leaf mass per area, dark respiration, photosynthetic potential and the fraction of photosynthesis used for maintenance respiration), and rates of transfer between soil organic matter carbon pools.

Figure 2. The iterative reduction in parameter uncertainty

Figure 3. Total parameter constraint per iteration

Equifinality and parameter interactions

When analyzing parameter posterior distributions in terms of parameter correlations, using additional data constraints increased the number of correlated parameters for the six data constraints that gave the largest reduction in model uncertainty (Fig. 4a). Using more data streams, in addition to these six, did not significantly change parameter correlations. Eight of the forty parameters optimized were strongly correlated ($r^2 \geq 0.3$) when using all data to constrain the model. For example, the extracted values for photosynthetic potential (V_{cmax} , Table 2, l_2 , Fig. 4b) were highly correlated with the proportion of photosynthate lost as maintenance respiration (parameter 8, Table 2, Fig. 4b). The strongest parameter correlations were between the basal rate and temperature dependence of root respiration (parameters 6, q , Table 2, Fig. 4b) and between different parameters governing spring phenology (parameters i , l , Table 2, Fig. 4b). Parameters that were poorly constrained ($z-K_2$, Fig. 4b) did not tend to show a better-defined correlation structure than parameters that were well constrained. This suggests that reducing correlations in the posterior parameter distributions does not imply a better-constrained model. The same is not true for parameter covariance, which was steadily reduced with the addition of each new data stream (Fig. 4c). Covariance scales the correlation by the standard deviation of the parameters, thus lowering the weight of parameters that have well constrained posterior distributions. Parameters that were not well constrained when using all available data tended to show a strong covariance structure (Fig. 4d). Well-

constrained parameters had limited covariance, even though some were highly correlated, reflecting the narrow range of variability for those parameters. This implies that using data relevant to these parameters could lead to a better-constrained model.

Figure 4. Fully optimized parameter co-variance

The effect of improved parameterization on future projections

Reduced model uncertainty under current climate conditions (Fig. 1) translated to reduced uncertainty in modeled future projections (Fig. 5). However, uncertainty in future projections of net ecosystem exchange was most reduced by the use of the few data streams that had the largest impact on model uncertainty under current climate conditions. Parameter-based uncertainty (i.e. without consideration of process based uncertainty) as to whether the system could be projected to be a source or a sink for atmospheric carbon for the next 100 years was reduced to near zero with the use of only five of the seventeen data streams available. The use of additional data streams led to only a minor reduction in parameter-based prediction uncertainty for net ecosystem exchange (Fig. 5). This was despite the fact that fourteen model parameters remained unconstrained (Fig. 2).

Figure 5. Future projections of model uncertainty

Discussion

By iteratively testing the reduction in model uncertainty gained by the use of seventeen different data streams, we have quantified the relative value of different data for informing a carbon cycle model. By running simulations to 2100 under a climate change scenario we also assess the value of each data stream for informing future model projections. The results show that that:

1. If the appropriate data are used, relatively few data sources are needed to give a large reduction in uncertainty in both short- and long-term projections of carbon cycling.
2. The data streams that proved most effective are those that characterize the flow of carbon through the system at different time scales. In particular, turnover times from different pools, in combination with flux data, led to the largest reduction in uncertainty.
3. Parameter uncertainty was similarly reduced by the addition of a few appropriate data streams. The use of additional data streams did not lead to a significant further reduction in parameter uncertainty, though parameter covariance was reduced with each data stream added.

Short vs long-term data needs

Terrestrial carbon cycle models are usually designed and tested using data representing diurnal or seasonal time scales (e.g. Kramer et al. 2002, Morales et al. 2005, Schwalm et

al. 2010, Richardson et al. 2012, Schaefer et al. in press), and occasionally interannual (e.g. Siqueira et al. 2006, Desai 2010, Keenan et al. 2012a). Model sensitivity and uncertainty analysis is commonly performed with a focus on short-term processes (e.g. Knorr & Kattge 2005). On the other hand, such models are widely used for long-term projections (e.g. Friedlingstein et al. 2006, Sitch et al. 2008). It has previously been shown that, when using only high-frequency net ecosystem exchange data, parameter sets that give comparable fits to the observations under current climatic conditions can lead to disparate projections of future carbon cycling (Keenan et al. 2012b). Here we show that the selection of a few key data constraints, which represent both short- and long-term processes, can substantially reduce parameter-based uncertainty in future projections.

Future projections & model uncertainty

Model projections are subject to two types of uncertainty: that due to parameter misspecification, and that due to process misrepresentation (Keenan et al. 2011a). In our approach we only evaluate the affect of uncertainty stemming from model parameterization, which represents an underestimate of the true uncertainty due to factors not included in the model system (e.g. Richardson et al. 2007, Keenan et al. 2012b). Thus, the fact that long-term projections from the process-oriented model were subject to low uncertainty does not imply that we should be confident about modeled future projections. Processes that are not considered in this model (e.g. disturbances, adaptation, community dynamics, carbon–nitrogen interactions) may also affect the long-term state of the ecosystem. The relatively low uncertainty in future projections (when using adequate data constraints), however, suggests that uncertainty due to parameter

misspecification can be effectively eliminated, leaving process representation as the remaining source of uncertainty. This is highly beneficial in that a model with well-constrained parameters and narrow confidence intervals is much easier to falsify (or prove wrong) than one with poorly constrained parameters and large uncertainties. The evaluation of process error in long-term model projections is non-trivial (Medlyn et al. 2011, Keenan et al. 2011b, 2012b, Migliavacca et al. 2012), and may require observations of long-term ecosystem processes (Luo et al. 2011) in combination with manipulation experiments (Templer & Reinmann 2011, Leuzinger et al. 2011).

Parameter uncertainty

One predominant goal of studies that aim to inform models with data is to identify model parameters. Early attempts in the field of terrestrial carbon cycling reported a limited number of parameters could be identified when using only eddy-covariance data (Wang et al. 2001, 2007, Knorr & Kattge, 2005). Recent efforts using multiple constraints (Rayner 2010) report a much larger proportion of identifiable parameters. Richardson et al. (2010) reported 11 out of 12 parameters were well constrained when using 6 different data constraints with a simple model, whilst two studies (Weng & Luo, 2011, Keenan et al. 2012b) constrained roughly half of the model parameters with comparatively more complex models. Here we show that improving parameter constraint is not solely a matter of using more data, but of selecting the correct data to use. Four of the available data sets (net ecosystem exchange, soil carbon turnover, soil respiration, leaf and woody litter fall) constrained 16 (64%) of the total parameters constrained (Fig. 2). Many parameters remained unconstrained even when using all data streams together. The fact that these

parameters were not identifiable, whilst model projections were well constrained, may suggest that they are redundant in the current model structure (when accounting for parameter covariance, see below). Simplifying process representation for model aspects that cannot be parameterized could aid in reducing the complexity of current models. Invoking ‘Occam’s razor’ in this fashion (making models only as complex as justified by the data), would minimize the common problem of model over-fitting, and could be considered a necessary step to avoid the development of excessively complex models.

Equifinality and parameter co-variance

Equifinality is defined as the situation where different parameter combinations or model structures can yield similar model performance (Beven, 2006). In the case of parameters, equifinality can be detected by assessing correlation and co-variance in posterior parameter distributions. Here we find that the level of equifinality depends on the number of different measurement types used to constrain the model. When using few data constraints, large equifinality allowed for divergent future projections of carbon cycling (Fig. 5). When using sufficient constraints, however, a lower level of equifinality was reached that did not prove detrimental to model performance and did not necessarily lead to an increase in model uncertainty over time. The model parameters that were least constrained tended to be those that had higher covariance (Fig. 4d). This implies that trade offs between these parameters allowed the model to get equivalent results with varied parameter values.

Strong parameter correlations were observed for both well and poorly constrained parameters. For example, despite being very well constrained, parameters governing the basal respiration rates and temperature sensitivity of different soil organic matter layers were highly correlated (parameters a , b , c , Fig. 4b). Similarly, parameters controlling the rate of root turnover, and the size of the root carbon pool were correlated, with higher values of one compensated for by lower values in the other (parameters o , k , Fig. 4b). Eight out of fourteen parameters that were poorly constrained showed strong correlation with other parameters. The majority of these correlative pairs were with other parameters that were already relatively well constrained (i.e. all except pairings photosynthetic potential (j_2) with the fraction of photosynthesis respired for maintenance (δ), Fig. 4b). Some poorly constrained parameters were not correlated with other parameters (e.g. the soil respiration scaling parameter, k_2). In our analysis, the introduction of additional data constraints increased parameter correlations, implying that apparently uncorrelated parameters may have high-dimensional parameter relationships that are not detected by simple 1-1 correlative analysis (Richardson & Hollinger 2005, Trudinger et al. 2009, Ricciuto et al. 2011). Strong posterior parameter correlation is often interpreted as an indicator that the constraining data was not sufficient to distinguish between counteracting processes in the model (e.g. Ricciuto et al. 2011). Here we show that strong, non-detrimental correlations can persist even in a well-constrained model, and may be an inevitable consequence of model structure. This correlation is not necessarily reduced by the use of additional data. Parameter co-variance, however, was continuously reduced with the use of additional data.

What data are most useful?

Previous studies have demonstrated the success of using additional data streams in conjunction with eddy-covariance flux data to improve estimates of ecosystem carbon exchange at different time scales (e.g. Williams et al. 2005, Moore et al., 2006, Xu et al. 2006, Richardson et al., 2010, Weng & Luo 2011, Ricciuto et al. 2011, Keenan et al., 2012b). The majority of studies emphasized the combination of stocks with fluxes, though no guidance is available as to what is the most appropriate or informative data to use. Our results show highly informative measurements at both ends of the cost of acquisition spectrum (e.g. senescence dates or leaf litter fall, vs eddy covariance or soil carbon turnover times). Coarse (woody) litterfall and leaf litterfall are often-overlooked measurements, but are ranked highly here. The results also show that some measurements, which have been the focus of much interest, are of low relative importance for modeling the carbon cycle. It should be kept in mind that we have not included all measurements that can possibly be made. Other measurements could include, for example, non-structural carbohydrate reserves, nutrient stoichiometry, leaf-angle distributions, transfer rates between carbon pools, bole respiration, etc. All data sources are almost never available at the same site, but studies using synthetic data could be performed by those interested in quantifying the relative value of different data (e.g. for proposed measurement campaigns).

The weight assigned to each measurement potentially has a large impact on the ranking of different data. In our optimization framework, we chose to weight each data stream equally, independent of the number of observations, to ensure that the optimization did

not favor model performance for one aspect of the ecosystem over another. We also weight each data stream by its associated uncertainty to account for the quality of the information contained therein. This choice, however, could affect the ranking of data streams. Other alternatives include giving each measurement equal weight, instead of each data stream. The problem boils down to information content: theoretically, an observation should be given weight relative to the information it contributes to the optimization. When using multiple constraints, the problem of quantifying the relative information is well exemplified by, say, quantifying the contribution of one estimate of soil carbon, compared to one half hourly measurement of net ecosystem carbon exchange. This is particularly relevant when using high frequency measurements of net ecosystem exchange – given 10,000 estimates of net ecosystem exchange, one additional NEE estimate does not necessarily contribute new information, whilst one estimate of the soil carbon stock does. Our chosen approach is in keeping with the philosophy that a model should predict all observations within measurement uncertainty, independent of the number of measurements available. Clearly, a detailed assessment of the real information content of observations, and an associated scheme for adequately weighting different data streams is an area in need of much research.

Turnover times of soil carbon pools have been suggested to be of utmost importance for accurately modeling the carbon cycle (Matamala et al. 2003, Strand et al. 2008, Richardson et al. 2010, Gaudinski et al. 2010). They have been inferred by model inversion approaches (Barrett 2002, Luo et al. 2003, Xu et al. 2006, Zhou and Luo, 2008, Zhang et al. 2010), though measurements are rarely available to test different model

structures and parameterizations (but see Riley et al. 2009, Gaudinski et al. 2009). Here we show that, after net ecosystem carbon exchange, turnover rates of the different soil carbon pools have the largest impact for improving model performance. Turnover times of different soil carbon pools (e.g. Gaudinski et al. 2010) and non-structural carbohydrate reserves (Richardson et al. in press), are not commonly available for model testing and should greatly aid in generating better-informed models in the future.

Conclusions

Financial resources in the field of earth system science are highly limited, and field campaigns expensive, so it is imperative to identify what measurements are of most use for a specific question. Here we present results using a method by which to quantify the value of a diverse range of ecological data for improving models of the terrestrial carbon cycle. Using a hierarchical framework, we show that relatively few data streams contribute to the largest reduction in uncertainty in model performance. In the presence of these data streams, which are distributed across the cost of acquisition spectrum, other measurement sources become redundant. For example, bud-burst dates, and carbon stock sizes, were of relatively little value for constraining model performance in the presence of more informative measurements. Our results highlight the importance of estimates of carbon stock turnover times, in conjunction with soil respiration and net ecosystem carbon exchange measurements. These data sources should be given priority in future efforts. Using this framework together with information on the cost of measurement acquisition would help project managers to develop more efficient and effective measurement campaigns.

Acknowledgments

Carbon flux and biometric measurements at HFEMS have been supported by the Office of Science (BER), U.S. Department of Energy (DOE) and the National Science Foundation Long-Term Ecological Research Programs. TFK, ADR, and JWM acknowledge support from NOAA's Climate Program Office, Global Carbon Cycle Program, under award NA11OAR4310054. TFK and ADR acknowledge support from the Northeastern States Research Cooperative, and the Office of Science (BER), US Department of Energy, through the Northeastern Regional Center of the National Institute for Climatic Change Research. We especially thank the many participants who have sustained the long-term data collection, and in particular the summer students engaged in collecting field data, supported by NSF Research Experience for Undergraduates (REU) program, and the Harvard Forest Woods Crew for logistical and maintenance support. The computations in this paper were run on the Odyssey cluster supported by the FAS Science Division Research Computing Group at Harvard University.

References

- Barford, C. C., S. C. Wofsy, M. L. Goulden, J. W. Munger, E. H. Pyle, S. P. Urbanski, L. Hutya, S. R. Saleska, D. Fitzjarrald, and K. Moore (2001), Factors controlling long- and short-term sequestration of atmospheric CO₂ in a mid-latitude forest. *Science*, 294(5547), 1688-91, doi:10.1126/science.1062962.
- Barrett, D. J. (2002), Steady state turnover time of carbon in the Australian terrestrial biosphere, *Global Biogeochemical Cycles*, 16(4), doi:10.1029/2002GB001860
- Barrett, D. J., M. J. Hill, L. B. Hutley, J. Beringer, J. H. Xu, G. D. Cook, J. O. Carter, and R. J. Williams (2005), Prospects for improving savanna biophysical models by using multiple-constraints model-data assimilation methods, *Australian Journal of Botany*, 53(7), 689-714, doi:10.1071/bt04139.
- Beer, C. et al. (2010), Terrestrial gross carbon dioxide uptake: global distribution and covariation with climate. *Science (New York, N.Y.)*, 329(5993), 834-8, doi:10.1126/science.1184984.
- Beven, K. (2006), A manifesto for the equifinality thesis, *Journal of Hydrology*, 320(1-2), 18-36.
- Braswell, B. H., W. J. Sacks, E. Linder, and D. S. Schimel (2005), Estimating diurnal to annual ecosystem parameters by synthesis of a carbon flux model with eddy covariance net ecosystem exchange observations, *Global Change Biology*, 11(2), 335–355, doi:10.1111/j.1365-2486.2005.00897.x.

566 Shannon, C.E. (1948), A mathematical theory of communication, *Bell System Technical*
567 *Journal*, (27), 379-423.

568 Delworth, T. L. et al. (2006), GFDL's CM2 Global Coupled Climate Models. Part I :
569 Formulation and Simulation Characteristics, *Journal of Climate*, 19(5), 643–674.

570 Desai, A. R. (2010), Climatic and phenological controls on coherent regional interannual
571 variability of carbon dioxide flux in a heterogeneous landscape, *Journal of*
572 *Geophysical Research*, 115, 1-13, doi:10.1029/2010JG001423

573 Franks, S. W., K. J. Beven, and J. H. C. Gash (1999), Multi-objective conditioning of a
574 simple SVAT model, *Hydrol. Earth Syst. Sci.*, 3(4), 477-488, doi:10.5194/hess-3-
575 477-1999.

576 Friedlingstein, P. et al. (2006), Climate-carbon cycle feedback analysis: Results from the
577 (CMIP)-M-4 model intercomparison, *Journal of Climate*, 19(14), 3337-3353.

578 Gaudinski, J. B., M. S. Torn, W. J. Riley, T. E. Dawson, J. D. Joslin, and H. Majdi
579 (2010), Measuring and modeling the spectrum of fine-root turnover times in three
580 forests using isotopes, minirhizotrons, and the Radix model, *Global Biogeochemical*
581 *Cycles*, 24(3), 1-17, doi:10.1029/2009GB003649

582 Gaudinski, J. B., M. S. Torn, W. J. Riley, C. Swanston, S. E. Trumbore, J. D. Joslin, H.
583 Majdi, T. E. Dawson, and P. J. Hanson (2009), Use of stored carbon reserves in
584 growth of temperate tree roots and leaf buds: analyses using radiocarbon

585 measurements and modeling, *Global Change Biology*, 15(4), 992-1014,
586 doi:10.1111/j.1365-2486.2008.01736.x.

587 Goulden, M. L., J. W. Munger, S. M. Fan, B. C. Daube, and S. C. Wofsy (1996),
588 Measurements of carbon sequestration by long-term eddy covariance: Methods and
589 a critical evaluation of accuracy, *Global Change Biology*, 2(3), 169-182.

590 Hastings, W. K. (1970), MONTE-CARLO SAMPLING METHODS USING MARKOV
591 CHAINS AND THEIR APPLICATIONS, *Biometrika*, 57(1), 97

592 Hayhoe, K., C. Wake, B. Anderson, X.-Z. Liang, E. Maurer, J. Zhu, J. Bradbury, A.
593 DeGaetano, A. M. Stoner, and D. Wuebbles (2007), Regional climate change
594 projections for the Northeast USA, *Mitigation and Adaptation Strategies for Global*
595 *Change*, 13(5-6), 425-436, doi:10.1007/s11027-007-9133-2.

596 Heimann, M., and M. Reichstein (2008), Terrestrial ecosystem carbon dynamics and
597 climate feedbacks. *Nature*, 451(7176), 289-92, doi:10.1038/nature06591.

598 Jaynes, E. T. (1957), Information theory and statistical mechanics, *Physical Review*,
599 (106), 620-630.

600 Kattge, J. et al. (2011), TRY - a global database of plant traits, *Global Change Biology*,
601 17(9), 2905-2935, doi:10.1111/j.1365-2486.2011.02451.x.

602 Keenan, T. F. et al. (2012a), Terrestrial biosphere model performance for inter-annual
603 variability of land-atmosphere CO₂ exchange, *Global Change Biology*, 18, 1971–
604 1987, doi: 10.1111/j.1365-2486.2012.02678.x

605 Keenan, T. F., M. S. Carbone, M. Reichstein, and A. D. Richardson (2011a), The model–
 606 data fusion pitfall: assuming certainty in an uncertain world, *Oecologia*, (167), 587–
 607 597, doi:10.1007/s00442-011-2106-x.

608 Keenan, T. F., E. Davidson, a. Moffat, W. Munger, and a. D. Richardson (2012b), Using
 609 model-data fusion to interpret past trends, and quantify uncertainties in future
 610 projections, of terrestrial ecosystem carbon cycling, *Global Change Biology*, 18,
 611 2555–2569, doi: 10.1111/j.1365-2486.2012.02684.x

612 Keenan, T. F., J. Maria Serra, F. Lloret, M. Ninyerola, and S. Sabate (2011b), Predicting
 613 the future of forests in the Mediterranean under climate change, with niche- and
 614 process-based models: CO2 matters! *Global Change Biology*, 17(1), 565-579,
 615 doi:10.1111/j.1365-2486.2010.02254.x.

616 Keller, M., D. S. Schimel, W. W. Hargrove, and F. M. Hoffman (2008), A continental
 617 strategy for the National Ecological Observatory Network, *Frontiers in Ecology and*
 618 *the Environment*, (6), 282-284.

619 Knorr, W., and J. Kattge (2005), Inversion of terrestrial ecosystem model parameter
 620 values against eddy covariance measurements by Monte Carlo sampling, *Global*
 621 *Change Biology*, 11(8), 1333-1351, doi:10.1111/j.1365-2486.2005.00977.x.

622 Kolmogorov, A. N. (1968), Three approaches to the quantitative definition of
 623 information, *Journal of Computer Mathematics*, 2, 157-168.

624 Kramer, K. et al. (2002), Evaluation of six process-based forest growth models using
625 eddy-covariance measurements of CO₂ and H₂O fluxes at six forest sites in Europe,
626 (8), 213-230.

627 Leuzinger, S., and R. Quinn Thomas (2011), How do we improve Earth system models?
628 Integrating Earth system models, ecosystem models, experiments and long-term
629 data, *New Phytologist*, 191, 15-18.

630 Luo, Y. et al. (2011), Coordinated approaches to quantify long-term ecosystem dynamics
631 in response to global change, *Global Change Biology*, 17(2), 843-854,
632 doi:10.1111/j.1365-2486.2010.02265.x.

633 Luo, Y., J. Clark, T. Hobbs, S. Lakshmivarahan, A. Latimer, K. Ogle, D. Schimel, and X.
634 Zhou (2008), Symposium 23. Toward Ecological Forecasting, *Bulletin of the*
635 *Ecological Society of America*, 89(4), 467-474

636 Luo, Y., L. W. White, J. G. Canadell, E. H. DeLucia, D. S. Ellsworth, A. Finzi, J. Lichter,
637 and W. H. Schlesinger (2003), Sustainability of terrestrial carbon sequestration: A
638 case study in Duke Forest with inversion approach, *Global Biogeochemical Cycles*,
639 17(1,1021), 1-13, doi:10.1029/2002GB001923.

640 Matamala, R., M. a Gonzàlez-Meler, J. D. Jastrow, R. J. Norby, and W. H. Schlesinger
641 (2003), Impacts of fine root turnover on forest NPP and soil C sequestration
642 potential., *Science (New York, N.Y.)*, 302(5649), 1385-7,
643 doi:10.1126/science.1089543.

644 Medlyn, B. E., R. a. Duursma, and M. J. B. Zeppel (2011), Forest productivity under
645 climate change: a checklist for evaluating model studies, *Wiley Interdisciplinary*
646 *Reviews: Climate Change*, 2(3), 332-355

647 Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller (1953),
648 Equation of State Calculations by Fast Computing Machines, *Journal of Chemical*
649 *Physics*, 21, 1087-1092.

650 Metropolis, N., and S. Ulam (1949), The Monte Carlo method, *Journal of the American*
651 *Statistical Association*, 44, 351-341.

652 Migliavacca, M., O. Sonnentag, T. F. Keenan, A. Cescatti, J. O’Keefe, and A.D.
653 Richardson (2012), On the uncertainty of phenological responses to climate change,
654 and implications for a terrestrial biosphere model, *Biogeosciences*, 9(6), 2063–2083,
655 doi:10.5194/bg-9-2063-2012.

656 Moore, D. J. P., J. Hu, W. J. Sacks, D. S. Schimel, and R. K. Monson (2008), Estimating
657 transpiration and the sensitivity of carbon uptake to water availability in a subalpine
658 forest using a simple ecosystem process model informed by measured net CO₂ and
659 H₂O fluxes, *Agricultural and Forest Meteorology*, 148(10), 1467–1477,
660 doi:10.1016/j.agrformet.2008.04.013.

661 Morales, P. et al. (2005), Comparing and evaluating process-based ecosystem model
662 predictions of carbon and water fluxes in major European forest biomes, *Global*
663 *Change Biology*, 11(12), 2211-2233, doi:10.1111/j.1365-2486.2005.01036.x.

664 Pan, Y. et al. (2011), A Large and Persistent Carbon Sink in the World's Forests, *Science*,
665 14(July), 1-11, doi:10.1126/science.1201609.

666 Papale, D., and R. Valentini (2003), A new assessment of European forests carbon
667 exchanges by eddy fluxes and artificial neural network spatialization, *Global*
668 *Change Biology*, 9(4), 525-535, doi:10.1046/j.1365-2486.2003.00609.x

669 Phillips, S. C., R. K. Varner, S. Frolking, J. W. Munger, J. L. Bubier, S. C. Wofsy, and P.
670 M. Crill (2010), Interannual, seasonal, and diel variation in soil respiration relative
671 to ecosystem respiration at a wetland to upland slope at Harvard Forest, *Journal of*
672 *Geophysical Research-Biogeosciences*, 115(G2), 1-18, doi:G02019
673 10.1029/2008jg000858.

674 Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery (2007), *Numerical*
675 *Recipes: The art of scientific computing*, Cambridge University Press, Cambridge
676 UK.

677 Quéré, C. L., M. R. Raupach, J. G. Canadell, and G. Marland (2009), Trends in the
678 sources and sinks of carbon dioxide, *Nature geoscience*, (November), 1-6,
679 doi:10.1038/NGEO689.

680 Le Quéré, C., M. R. Raupach, J. G. Canadell and G. Marland (2009), Trends in the
681 sources and sinks of carbon dioxide, *Nature Geoscience*, 2(12), 831–836,
682 doi:10.1038/ngeo689.

683 Rayner, P. J., E. Koffi, M. Scholze, T. Kaminski, and J.-L. Dufresne (2011), Constraining
 684 predictions of the carbon cycle using data., *Philosophical transactions. Series A,*
 685 *Mathematical, physical, and engineering sciences*, 369(1943), 1955-66,
 686 doi:10.1098/rsta.2010.0378

687 Ricciuto, D. M., A. W. King, D. Dragoni, and W. M. Post (2011), Parameter and
 688 prediction uncertainty in an optimized terrestrial carbon cycle model: Effects of
 689 constraining variables and data record length, *Journal of Geophysical Research*,
 690 116(G1), 1-17, doi:10.1029/2010JG001400.

691 Richardson, A. D. et al. (2006), A multi-site analysis of random error in tower-based
 692 measurements of carbon and energy fluxes, *Agricultural and Forest Meteorology*,
 693 136(1-2), 1-18, doi:10.1016/j.agrformet.2006.01.007.

694 Richardson, A. D. et al. (2010), Estimating parameters of a forest ecosystem C model
 695 with measurements of stocks and fluxes as joint constraints, *Oecologia*, 164(1), 25-
 696 40, doi:10.1007/s00442-010-1628-y.

697 Richardson, A. D. et al. (2012), Terrestrial biosphere models need better representation of
 698 vegetation phenology: results from the North American Carbon Program Site
 699 Synthesis, *Global Change Biology*, 18(2), 566-584, doi:10.1111/j.1365-
 700 2486.2011.02562.x.

701 Richardson, A. D., D. Y. Hollinger, J. D. Aber, S. V. Ollinger, and B. H. Braswell
 702 (2007), Environmental variation is directly responsible for short- but not long-term

703 variation in forest-atmosphere carbon exchange, *Global Change Biology*, 13(4),
704 788-803, doi:10.1111/j.1365-2486.2007.01330.x.

705 Richardson, A.D. et al. (in press) Seasonal dynamics and age of stemwood nonstructural
706 carbohydrates in temperate forest trees. *New Phytologist*.

707 Riley, W. J., J. B. Gaudinski, M. S. Torn, J. D. Joslin, and P. J. Hanson (2009), Fine-root
708 mortality rates in a temperate forest: estimates using radiocarbon data and numerical
709 modeling., *The New phytologist*, 184(2), 387-98, doi:10.1111/j.1469-
710 8137.2009.02980.x.

711 Savage, K., E. a. Davidson, A. D. Richardson, and D. Y. Hollinger (2009), Three scales
712 of temporal resolution from automated soil respiration measurements, *Agricultural
713 and Forest Meteorology*, 149(11), 2012-2021, doi:10.1016/j.agrformet.2009.07.008.

714 Schwalm, C. R. et al. (2010), A model-data intercomparison of CO₂ exchange across
715 North America: Results from the North American Carbon Program site synthesis,
716 *Journal of Geophysical Research*, 115, 22, doi:10.1029/2009JG001229.

717 Siqueira, M. B., G. G. Katul, D. a. Sampson, P. C. Stoy, J.-Y. Juang, H. R. Mccarthy, and
718 R. Oren (2006), Multiscale model intercomparisons of CO₂ and H₂O exchange rates
719 in a maturing southeastern US pine forest, *Global Change Biology*, 12(7), 1189-
720 1207, doi:10.1111/j.1365-2486.2006.01158.x.

721 Sitch, S. et al. (2008), Evaluation of the terrestrial carbon cycle, future plant geography
722 and climate-carbon cycle feedbacks using five Dynamic Global Vegetation Models

723 (DGVMs), *Global Change Biology*, 14(9), 2015-2039, doi:10.1111/j.1365-
 724 2486.2008.01626.x.

725 Solomon, S., D. Qin, M. Manning, Z. Chen, and M. Marq (Eds.) (2007), *IPCC 2007*
 726 *Summary for policymakers. In Climate change 2007: the physical science basis.*
 727 *Contribution of working group I to the fourth assessment report of the*
 728 *intergovernmental panel on climate change.*

729 Strand, A. E., S. G. Pritchard, M. L. McCormack, M. a Davis, and R. Oren (2008),
 730 Irreconcilable differences: fine-root life spans and soil carbon persistence., *Science*
 731 *(New York, N.Y.)*, 319(5862), 456-8, doi:10.1126/science.1151382.

732 Templer, P., and A. Reinmann (2011), Multi-factor global change experiments: what
 733 have we learned about terrestrial carbon storage and exchange? *New Phytologist*,
 734 192, 797-800.

735 Urbanski, S., C. Barford, S. Wofsy, C. Kucharik, E. Pyle, J. Budney, K. McKain, D.
 736 Fitzjarrald, M. Czikowsky, and J. W. Munger (2007), Factors controlling CO₂
 737 exchange on timescales from hourly to decadal at Harvard Forest, *Journal of*
 738 *Geophysical Research*, 112(G2), 1-25, doi:10.1029/2006JG000293.

739 WANG, Y. P., D. BALDOCCHI, R. LEUNING, E. FALGE, and T. VESALA (2007),
 740 Estimating parameters in a land-surface model by applying nonlinear inversion to
 741 eddy covariance flux measurements from eight FLUXNET sites, *Global Change*
 742 *Biology*, 13(3), 652-670, doi:10.1111/j.1365-2486.2006.01225.x

743 Wang, Y. P., D. Baldocchi, R. Leuning, E. Falge, and T. Vesala (2007), Estimating
 744 parameters in a land-surface model by applying nonlinear inversion to eddy
 745 covariance flux measurements from eight FLUXNET sites, *Global Change Biology*,
 746 *13*(3), 652-670, doi:10.1111/j.1365-2486.2006.01225.x.

747 Wang, Y.-P., R. Leuning, H. a. Cleugh, and P. a. Coppin (2001), Parameter estimation in
 748 surface exchange models using nonlinear inversion: how many parameters can we
 749 estimate and which measurements are most useful?, *Global Change Biology*, *7*(5),
 750 495-510, doi:10.1046/j.1365-2486.2001.00434.x.

751 Wang, Y.-P. P., C. M. Trudinger, and I. G. Enting (2009), A review of applications of
 752 model-data fusion to studies of terrestrial carbon fluxes at different scales,
 753 *Agricultural and Forest Meteorology*, *149*(11), 1829-1842,
 754 doi:10.1016/j.agrformet.2009.07.009.

755 Weng, E., and Y. Luo (2011), Relative information contributions of model vs. data to
 756 short- and long-term forecasts of forest carbon dynamics. *Ecological applications* :
 757 *a publication of the Ecological Society of America*, *21*(5), 1490-505

758 Williams, M., P. a. Schwarz, B. E. Law, J. Irvine, and M. R. Kurpius (2005), An
 759 improved analysis of forest carbon dynamics using data assimilation, *Global*
 760 *Change Biology*, *11*(1), 89-105, doi:10.1111/j.1365-2486.2004.00891.x.

761 Wofsy, S. C., M. L. Goulden, J. W. Munger, S. M. Fan, P. S. Bakwin, B. C. Daube, S. L.
 762 Bassow, and F. a Bazzaz (1993), Net Exchange of CO₂ in a Mid-Latitude Forest.,
 763 *Science (New York, N.Y.)*, *260*(5112), 1314-7, doi:10.1126/science.260.5112.1314.

764 Xiao, J., J. Chen, K. J. Davis, and M. Reichstein (2012), Advances in upscaling of eddy
 765 covariance measurements of carbon and water fluxes, *Journal of Geophysical*
 766 *Research*, 117(November 2011), 2-5, doi:10.1029/2011JG001889.

767 Xu, T., L. White, D. Hui, and Y. Luo (2006), Probabilistic inversion of a terrestrial
 768 ecosystem model: Analysis of uncertainty in parameter estimation and model
 769 prediction, *Global Biogeochemical Cycles*, 20(2), 1-15,
 770 doi:10.1029/2005GB002468.

771 Zhang, L. L., Y. Luo, and G. Yu (2010), Estimated carbon residence times in three forest
 772 ecosystems of eastern China: Applications of probabilistic inversion, *J. Geophys.*
 773 *Res.*, 115(G1), G01010, doi:10.1029/2009jg001004

774 Zhou, T., and Y. Luo (2008), Spatial patterns of ecosystem carbon residence time and
 775 NPP-driven carbon uptake in the conterminous United States, *Global*
 776 *Biogeochemical Cycles*, 22(3), 1-15, doi:10.1029/2007GB002939.

777

779 **Table 1.** Data sets used in this study

Data set no.	Measurement	Acronym	Frequency	# data points	Reference
1	Eddy-covariance	NEE	Hourly	73,198	Urbanski et al., and ¹
2	Soil Respiration 1	Rsoil	Hourly	26,430	Savage et al., 2009
3	Soil Respiration 2	Rsoil	Hourly	19,030	Phillips et al., 2010
4	Soil Respiration 3	Rsoil	Weekly	498	²
5	Leaf area index	LAI	Monthly	51	Norman, 1993; Urbanski et al., and ¹
6	Leaf litter fall	Lfleaf	Yearly	10	Urbanski et al., and ¹
7	Woody biomass	Wood C	Yearly	15	Jenkins et al., 2004. Urbanski et al., and ¹
8	Woody litterfall	Lfwood	Yearly	8	Urbanski et al., and ¹
9	Fine root biomass	Root C	One Year	1	DIRT project ¹
10	Forest floor carbon	Lit C	One Year	1	Gaudinski et al., 2000
11	Budburst	Budburst	Yearly	15	O’Keefe, 2000 ¹
12	Leaf Drop	Senescence	Yearly	14	O’Keefe, 2000 ¹
13	Soil carbon pools	Soil C	Three years	3	Gaudinski et al., 2000 Magill et al., 2000

					Bowden et al., 2009
14	Soil carbon turnover	Soil C TO	One	1	Gaudinski et al., 2000
15	Proportion of heterotrophic respiration in soil	% Root Resp.	One	1	Gaudinski et al., 2000 Bowden et al., 1993
16	Litter Turnover	Litter TO	One	1	Gaudinski et al., 2000

780

781 ¹ See data download page: <http://harvardforest.fas.harvard.edu/data/archive.html>

782 ² ftp://ftp.as.harvard.edu/pub/nigec/HU_Wofsy/hf_data/ecological_data/soilR/

783

Table 2. FöBAAR model parameters and pools. Both parameters and initial pool sizes were optimized conditional on the data constraints. Parameters are arranged in descending order of constraint (i.e. best constrained parameters first, to worst constrained parameters last) to relate to Figure 2.

Id.	Name	Definition	Min	Max
a	SOM _C Pd	Passive SOMC respiration rate (Log)	-10	-1
b	SOM _C SdT	Fast cycling SOMC temperature dependence	0.01	0.1
c	SOM _c Fd	Fast cycling SOMC respiration rate (Log)	-6	-1
d	AirTs	Leaf senescence onset mean air temperature (°C)	0	15
e	Lf _f	Litterfall from foliage (Log)	-6	-1
f	SOM _C Sd	Slow cycling SOMC respiration rate (Log)	-6	-1
g	Lf _w	Litterfall from wood (Log)	-6	-1
h	Fc	Fraction of Cf not transferred to mobile carbon	0.4	0.7
i	GDD ₀	Day of year for growing degree day initiation	50	150
j	Lit2SOM	Litter to fast SOMC transfer rate (Log)	-6	-1
k	Lf _r	Litterfall from roots (Log)	-6	-1
l	Af	Fraction of GPP allocated to foliage	0.5	1
m	Lit _d T	Litter respiration temperature dependence	0.01	0.1
n	Lit _C	Carbon in litter	10	1000
o	R _C	Carbon in roots	20	500
p	Lit _d	Litter respiration rate (Log)	-6	-1
q	R _{root} d	Root respiration rate (Log)	-6	-1

r	Mob _C Tr	Fraction of mobile transfers respired	0	0.01
s	Rsoil ₁	Soil respiration scaling co-efficient (data set 1)	0.5	1.5
t	W _C	Carbon in wood	8000	14000
u	Ar	Fraction of NPP allocated to roots	0.5	1
v	Lit ₂ SOMT	Litter to fast SOMC temperature dependence	0.03	0.5
w	Rsoil ₂	Soil respiration scaling co-efficient (data set 2)	0.5	1.5
x	SOM _C P	Carbon in passive cycling SOM layer	2000	12000
y	SOM _C S	Carbon in slow cycling SOM layer	2000	12000
z	Mob _C R	Mobile stored carbon respiration rate (Log)	-6	-1
1	GDD ₁	Growing degree days for spring onset	150	300
2	SOM _C F ₂ SOM _C S	Fast SOMC to slow SOMC rate	0.03	0.5
3	SOM _C S	Carbon in slow cycling SOM layer	2000	12000
4	SOM _C S ₂ SOM _C P	Transfer rate from slow to passive SOM	0.001	0.4
5	SOM _C S ₂ SOM _C PT	Fast SOMC to slow SOMC temp. dependence	0.03	0.5
6	R _{root} dT	Root respiration rate temperature dependence	0.01	0.2
7	GDD ₂	Spring photosynthetic GDD maximum	500	1000
8	MaintR	Fraction of GPP respired for maintenance	0.1	0.4
9	LMA	Leaf mass per area (gC m ⁻²)	50	90
i2	Rd	Rate of dark respiration	0.001	0.1
j2	Vcmax	Velocity of carboxylation (umol mol ⁻¹)	60	150
l2	Mob _C	Mobile carbon	75	200
f2	Q ₁₀ Rd	Temperature dependence of Rd	0.5	2.5
k2	Rsoil ₃	Soil respiration scaling co-efficient (data set 3)	0.5	1.5

788

789

790

791

792

Figure 1. The frequency distribution of model-data mismatch (log), when constraining the model with different data combinations. At each stage of the hierarchical optimization process (represented as rows in the graph), the model is constrained using a combination of different data sources, and tested against all data available. Each shaded curve represents the distribution of model-data mismatch for the model constrained using a particular data combination. The area under each curve represents the log distribution of model-data mismatch (Error) for all available data, quantified using the cost function (Eq. 1, 2) and 100,000 model runs. A value of one signifies that model estimates are on average within the error associated with the observations. Each row thus presents the posterior distribution of model uncertainty for all simulations at that stage. The data combination that gave the best model performance (shown in black) is selected for use in the next stage. Sub-optimal data combinations are shown in grey. As an example of the approach, in the first row, all data are tested together and daytime NEE is selected as giving the greatest reduction in model uncertainty. In the second row, the model is optimized again, this time with daytime NEE plus each other data stream independently. By the last column, all data streams are being used to optimize the model. Please note that the range is restricted for illustrative purposes. For the first few rows most distributions extend far beyond the restricted range.

Figure 2. The posterior parameter distributions for the best data combination at each stage in the hierarchical optimization process. Rows directly relate to the rows in Figure 1. Parameter identifiers and initial ranges are given in Table 1. The right hand column (#) gives the number of parameters well constrained at each iteration. Parameters are deemed

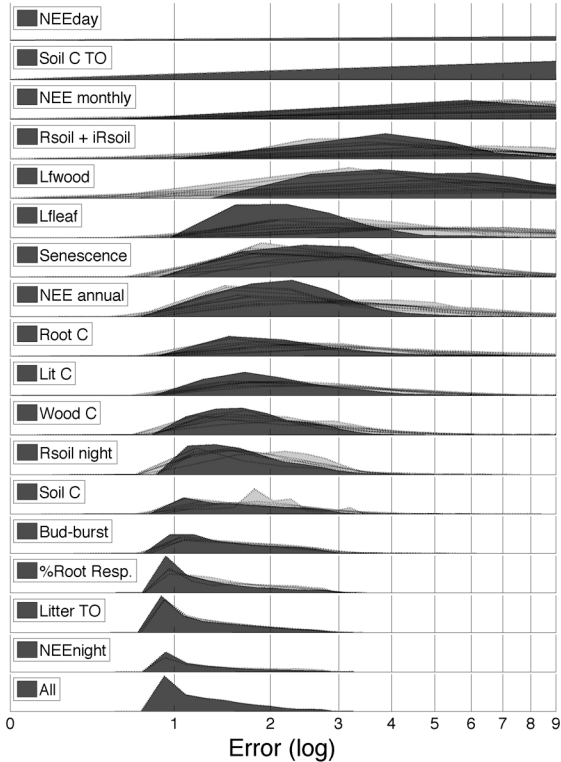
to be well constrained if their posterior distribution occupies at most half the range of the prior distribution. Grey dots represent the optimum parameter value.

Figure 3. The extent of the improvement in parameter constraint with the inclusion of additional data. Iteration numbers relate to the rows in Figure 1. The normalized parameter constraint is the mean standard deviation of all posterior parameter distributions, normalized by the standard deviation of a uniform distribution from 0 to 1 (i.e. 0.289). If all posterior parameter distributions were uniform (i.e. uninformed by the data) the normalized parameter constraint would have a value of 1. A value of zero signifies that all parameters are fully constrained.

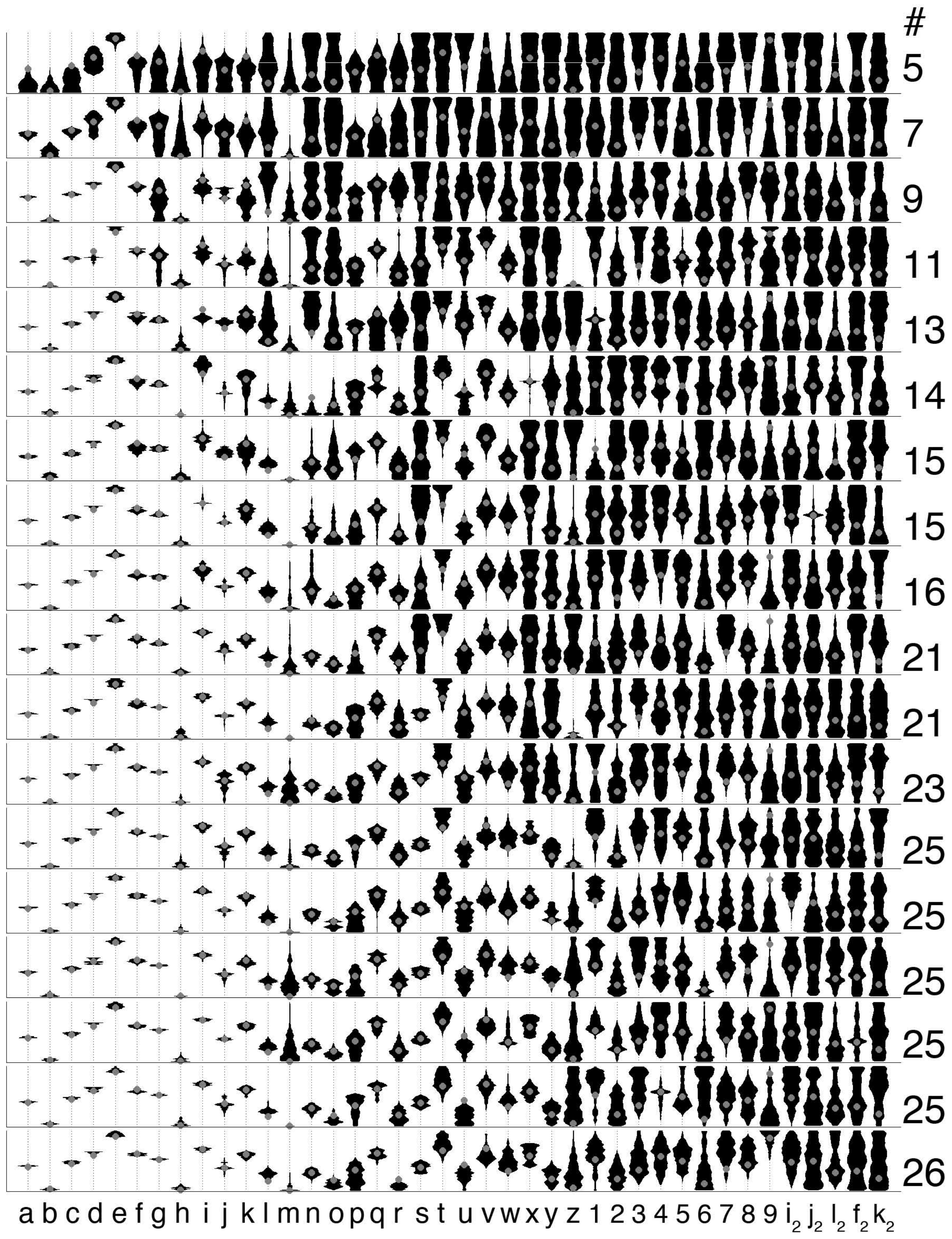
Figure 4. (a) The number of posterior parameter distributions that show significant ($p < 0.01$) correlations for different levels of correlation and different numbers of constraining data sets. Data sets 1-18 are those depicted in Fig.'s 1, 2, & 4. **(b)** The correlation matrix of model parameters for the model constrained by all available data sets. The color scale represents the r^2 correlation between each parameter. Parameters are as listed in Table 2. **(c)** The posterior parameter covariance for different numbers of constraining data sets, normalized to the maximum total covariance observed. **(d)** The covariance matrix for the for model parameters for the model constrained by all available data sets. The color code represents the covariance normalized to the maximum observed covariance value.

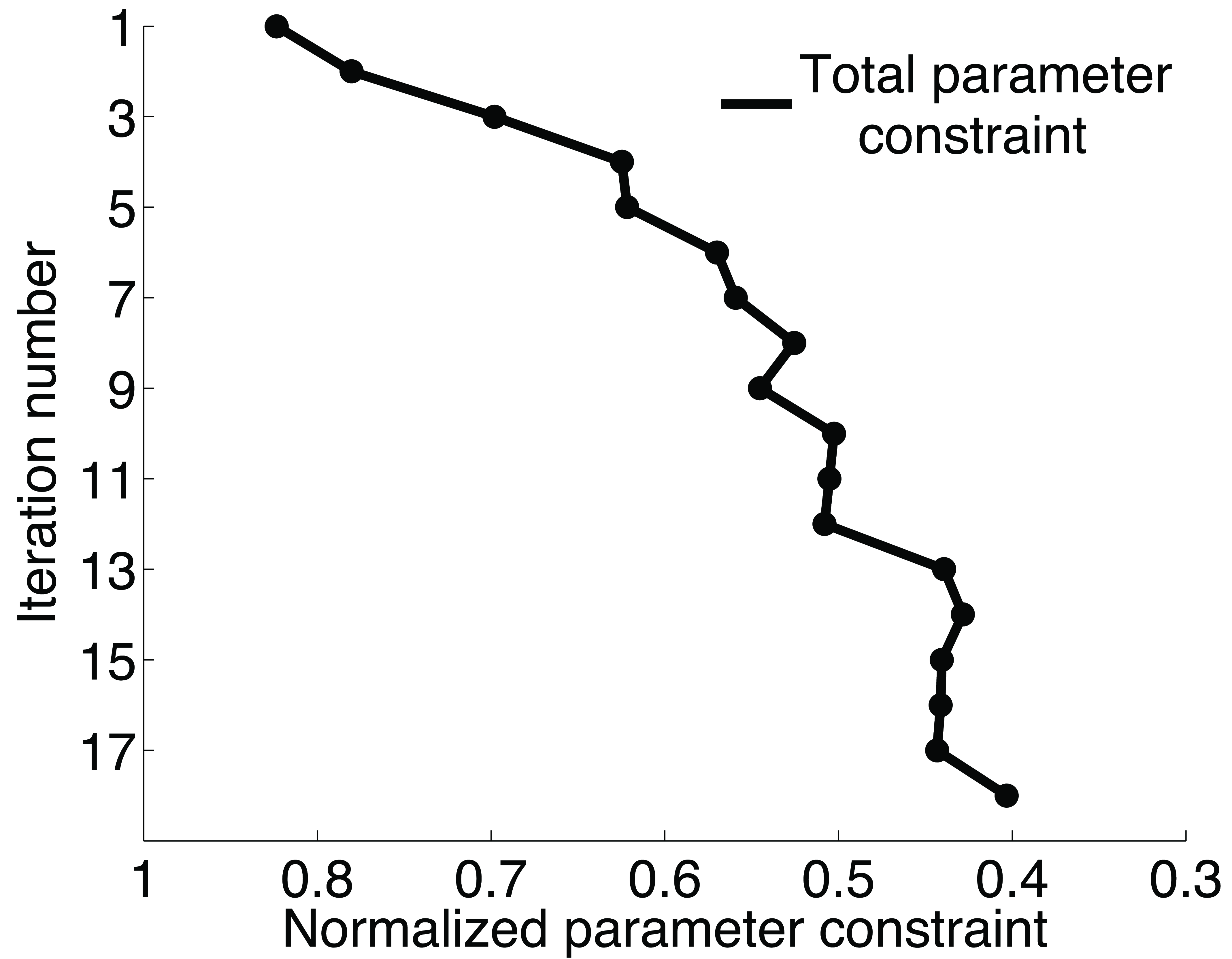
Figure 5. The range of equally plausible modeled annual net ecosystem exchange ($\text{gC m}^{-2} \text{y}^{-1}$) from 2000 to 2100 for the best data constraint combination at each stage of the hierarchical optimization process. Rows directly correspond to those of Fig. 1 and 2. The dashed line is the zero line, indicating whether the ecosystem is predicted to be either a source (>0) or a sink (<0) for CO_2 . Shaded areas represent the 95% confidence interval for model projections.

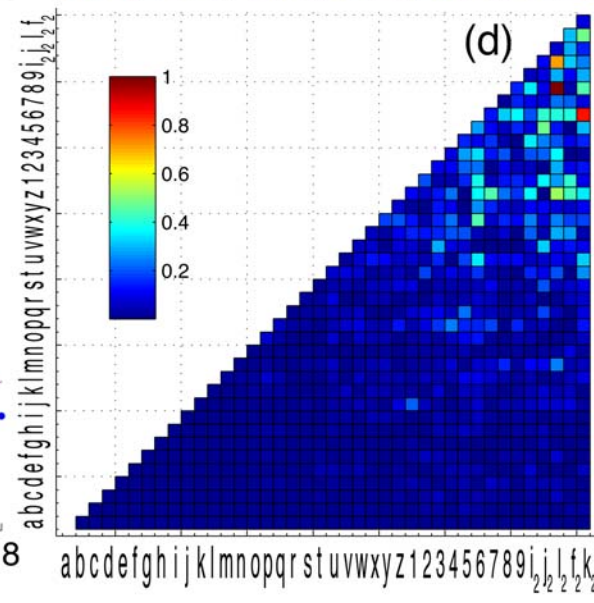
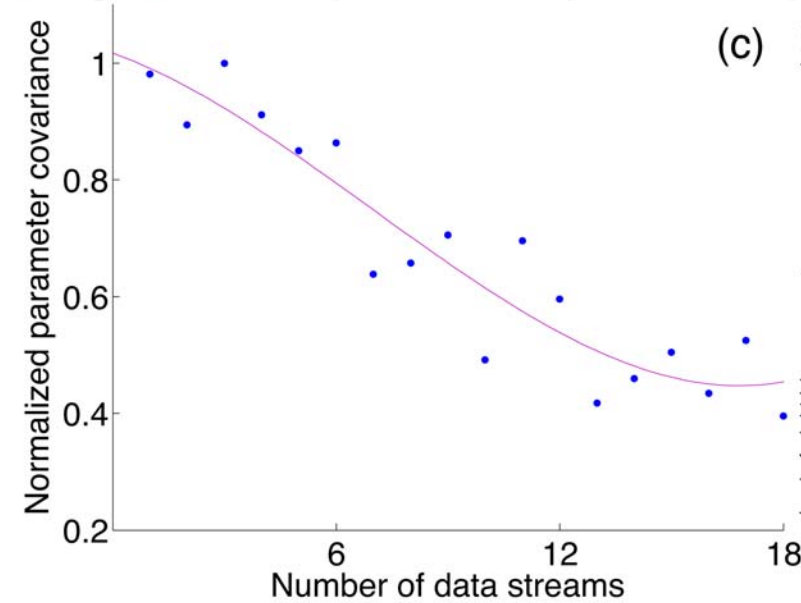
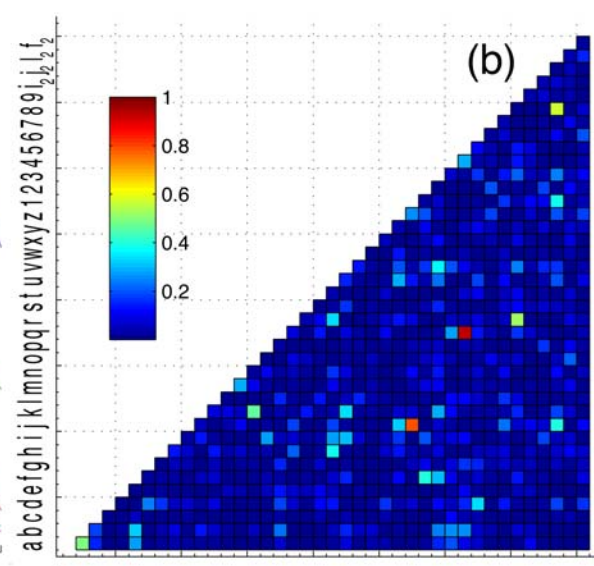
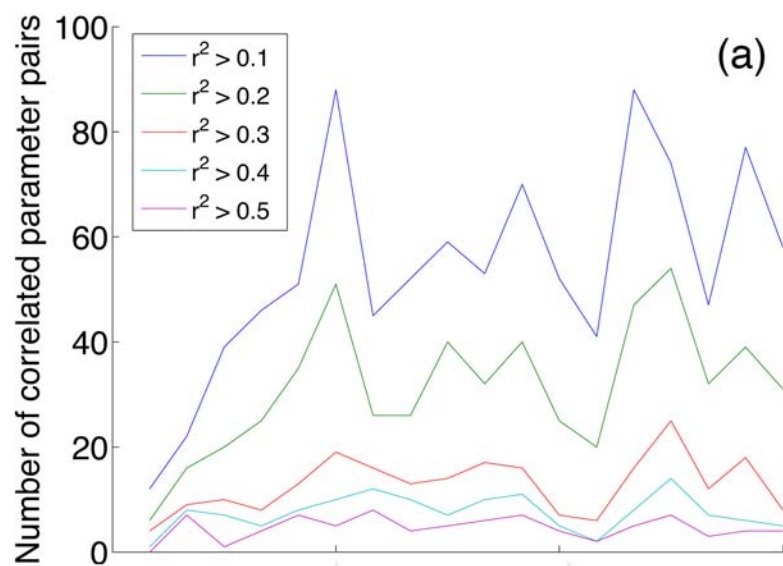
Frequency



Normalized parameter range (0-1)







Net ecosystem exchange ($\text{gC m}^{-2} \text{y}^{-1}$)

