

1 **This is the authors' final version. Please note that it differs slightly from the**
2 **published version. Where there are differences, the published version should be**
3 **taken as authoritative.**

4

5 **This paper has been published in Oecologia. The original publication may be**
6 **obtained at <http://www.springerlink.com>. If you have electronic access to**
7 **Oecologia, you may obtain the publication at:**
8 **<http://dx.doi.org/10.1007/s00442-006-0565-2>**

9 Coupling between carbon cycling and climate in a high-elevation, subalpine forest: a
10 model-data fusion analysis

11

12 William J. Sacks^{1,2}, David S. Schimel², Russell K. Monson³

13

14 1) Center for Sustainability and the Global Environment, Nelson Institute for
15 Environmental Studies, University of Wisconsin-Madison, 1710 University
16 Avenue, Madison, WI 53726, USA

17 2) National Center for Atmospheric Research, 1850 Table Mesa Drive, Boulder
18 Colorado 80305, USA

19 3) Department of Ecology and Evolutionary Biology and Cooperative Institute for
20 Research in the Environmental Sciences, University of Colorado, Boulder,
21 Colorado 80309-0334

22

23 Corresponding author: William Sacks. Email: wsacks@wisc.edu; Ph: 608-890-0337;

24 Fax: 608-265-4113

25

26

27 **Abstract**

28 Fundamental questions exist about the effects of climate on terrestrial net
29 ecosystem CO₂ exchange (NEE), despite a rapidly growing body of flux observations.
30 One strategy to clarify ecosystem climate-carbon interactions is to partition NEE into its
31 component fluxes, gross ecosystem CO₂ exchange (GEE) and ecosystem respiration
32 (R_E), and evaluate the responses to climate of each component flux. We separated
33 observed NEE into optimized estimates of GEE and R_E using an ecosystem process
34 model combined with six years of continuous flux data from the Niwot Ridge AmeriFlux
35 site. In order to gain further insight into the processes underlying NEE, we partitioned
36 R_E into its components: heterotrophic (R_H) and autotrophic (R_A) respiration. We were
37 successful in separating GEE and R_E, but less successful in accurately partitioning R_E
38 into R_A and R_H. Our failure in partitioning R_E was due to a lack of adequate contrasts in
39 the assimilated data set to distinguish between R_A and R_H. We performed most model
40 runs at a twice-daily time step. Optimizing on daily-aggregated data severely degraded
41 the model's ability to separate GEE and R_E. However, we gained little benefit from using
42 a half-hourly time step.

43 The model-data fusion showed that most of the interannual variability in NEE was
44 due to variability in GEE, and not R_E. In contrast to several previous studies in other
45 ecosystems, we found that longer growing seasons at Niwot Ridge were correlated with
46 *less* net CO₂ uptake, due to a decrease of available snow-melt water during the late
47 springtime photosynthetic period. Warmer springtime temperatures resulted in
48 increased net CO₂ uptake only if adequate moisture was available; when warmer
49 springtime conditions led into mid-summer drought, the annual net uptake declined.

50

51 **Key words:** ecosystem respiration, eddy covariance, gross primary productivity (GPP),
52 net ecosystem exchange (NEE), parameter estimation

53

54

55 **Introduction**

56 Correlations between climate and CO₂ exchange at local to global scales provide
57 information about climate controls over terrestrial carbon cycling (Goulden et al. 1996;
58 Angert et al. 2005; Ciais et al. 2005; Sacks et al. 2006). Despite a growing body of
59 literature in these areas, fundamental questions remain about the separate and
60 interactive effects of temperature, water and other climate variables on terrestrial
61 ecosystems (Breshears et al. 2005). Particularly problematic is the fact that NEE
62 reflects the small difference between the opposing large CO₂ fluxes of gross ecosystem
63 exchange (GEE) and ecosystem respiration (R_E), fluxes that tend to respond similarly
64 but not identically to the environment. Predicting the terrestrial carbon response to
65 climate variables requires precise prediction of the differential effects of climate on GEE
66 and R_E and their interactions, together with the localized but large effects of
67 disturbance. Small errors in the climate sensitivity of the GEE and R_E fluxes, or the time
68 scale on which they respond, can cause major errors in NEE. Modeling of NEE at large
69 scales has proved challenging and models based on similar principles give contrasting
70 results. This is partly because of the difficulty of precise estimation of parameters
71 governing these fluxes in noisy and heterogeneous systems, and partly because GEE

72 and R_E are not independent, but rather are linked by substrate production and use, and
73 by their common dependence on nutrient availability (Ryan and Law 2005).

74 In theory, diagnosis of environmental controls on NEE should be feasible using
75 long-term eddy covariance measurements – these fluxes contain the relevant
76 responses of photosynthesis and respiration to the environment, and these
77 measurements are made at a high density capable of revealing fine-grained patterns in
78 a long time series. However, direct flux measurements do not solve the parametric
79 problem for prediction because the component processes (GEE and R_E) must be
80 inferred *a priori* from the flux data, i.e., before parameter values can be estimated. This
81 problem of circularity is in principle similar to other problems in the Earth Sciences
82 where understanding an observed phenomenon requires inference of real but
83 unobservable component quantities. The problem can be addressed by combining the
84 observations with a model of the underlying processes to produce an analysis of the
85 indirectly observed fluxes consistent with both theory (via the model) and observations
86 (via the data).

87 In past studies, we addressed the problems of parameter estimation using a
88 powerful model-data fusion approach (Braswell et al. 2005; Sacks et al. 2006; see also
89 Schulz et al. 2001; Wang et al. 2001; Raupach et al. 2005; Williams et al. 2005). This
90 approach addresses the interactive nature of GEE and R_E through modeled recognition
91 of their substrate-level controls. The model-data fusion yields a model whose
92 parameters are consistent with both prior knowledge of physiology and the observed
93 patterns of NEE as realized over all seasons and multiple years. Combining the model
94 and observations in this way produces a theoretically-accurate and highly-constrained

95 analysis of NEE, including its dependence on the responses of photosynthesis, and the
96 responses of heterotrophic and autotrophic respiration to seasonal and interannual
97 environmental variability.

98 In this study, we applied the model-data fusion approach to the six-year Niwot
99 Ridge AmeriFlux data record, which includes continuous eddy covariance fluxes and
100 other ecological measurements (Monson et al. 2002; Turnipseed et al. 2002, 2003;
101 Scott-Denton et al. 2003; Monson et al. 2005; Scott-Denton et al. 2006). The site is
102 characterized by snow-covered winters, a spring-time melt season and a dry summer
103 season with episodic and variable rains. Much of the net CO₂ uptake occurs during the
104 period when soil moisture is provided by snowmelt (Monson et al. 2002, 2005). From
105 the existing research record, it appears that a number of factors control the annual
106 balance of GEE and R_E. The timing and depth of winter snow regulates winter snow
107 cover and affects soil temperatures through insulation. The timing and duration of the
108 snowmelt period influences when springtime photosynthesis begins and for how long it
109 proceeds before water stress reduces it. The amount of summer rain affects both GEE
110 and R_E after the snowmelt moisture is gone. While qualitative and, to a certain extent,
111 quantitative inferences about GEE and R_E can be made from such observations, we
112 processed this record using a model-data fusion analysis to arrive at quantitative
113 estimates of the sensitivity of GEE and R_E to temperature, moisture and other variables
114 consistent with the patterns observed above.

115

116 **Materials and Methods**

117 *Study site*

118 The Niwot Ridge AmeriFlux Site is located approximately 50 miles west of Boulder,
119 Colorado (40° 1'58" N; 105° 32'47"W) at 3050 m elevation. The site is situated in the
120 subalpine forest ecosystem with the dominant species being *Abies lasiocarpa*
121 (subalpine fir), *Picea engelmannii* (Engelmann spruce) and *Pinus contorta* (lodgepole
122 pine). The forest is accumulating carbon, recovering from logging that ended about 100
123 years ago (Monson et al. 2002). The canopy at the site is relatively open with an
124 average gap fraction of 17% (Turnipseed et al. 2003). The average canopy height is
125 11.4 m and the average mid-summer leaf-area index is 4.2 m² m⁻² (Monson et al. 2002;
126 Turnipseed et al. 2002). The site has a sparse, heterogeneous ground cover, mostly
127 composed of *Vaccinium* sp., lichens and occasional moss. Soils are sandy and derived
128 from granitic moraine with a distinct, thin (< 6 cm) organic horizon. Mean annual
129 precipitation at the site is 800 mm and the mean annual temperature is 1.5 °C.

130

131 *Measurements of NEE and associated meteorological parameters*

132 From November 1998 through the present, fluxes of CO₂, sensible and latent heat,
133 and a number of meteorological variables have been measured as part of the Niwot
134 Ridge AmeriFlux program (http://urquell.colorado.edu/data_ameriflux/). For this study,
135 we used data from November 1, 1998 through December 31, 2004. CO₂ fluxes were
136 measured using the eddy covariance method. General details of the eddy covariance
137 approach are provided in Baldocchi (2003), and specific details for measurements at the
138 Niwot Ridge AmeriFlux site are provided in several previous papers (Monson et al.
139 2002; Turnipseed et al. 2002, 2003, 2004). Briefly, turbulent fluxes were measured
140 using a triaxial sonic anemometer (SWS-211/3K, Applied Technologies, Inc., Boulder

141 CO, USA) and a closed-path infrared gas analyzer (Li-Cor 6262, Li-Cor Inc, Lincoln NE,
142 USA). Measurements were aligned with the mean wind streamlines (Kaimal and
143 Finnigan 1994), and standard density corrections (Webb et al. 1980) were applied.
144 Beneath-canopy CO₂ storage was determined by vertical integration of six profile
145 stations located on the tower, and added to the eddy-flux measurement to compute the
146 overall NEE as described in Goulden et al. (1996). By convention, net CO₂ fluxes are
147 referenced to the atmosphere: negative fluxes represent CO₂ loss from the atmosphere
148 (net ecosystem photosynthesis) and positive fluxes represent CO₂ gain by the
149 atmosphere (net ecosystem respiration).

150 Air temperatures were determined from the temperature signal of the sonic
151 anemometer at the same height (21.5 m) as the turbulent flux measurements. Soil
152 temperatures were averaged over the top 10 cm of soil. Precipitation values were
153 obtained from the University of Colorado Long-term Ecological Research web site
154 (<http://culter.colorado.edu>), which in turn were obtained from standard precipitation
155 gauge measurements. The site used for the precipitation measurements is located 500
156 m north of the tower flux site.

157 Gaps occur in the flux and meteorological measurements due to instrument
158 malfunction or, in the case of fluxes, periods of high atmospheric stability (which can
159 cause an underestimate of the flux; Goulden et al. 1996; Monson et al. 2002). These
160 gaps were filled using a variety of methods, depending on the length of the gap: spline
161 fits, averages of the surrounding days, or empirical regressions based on the available
162 meteorological data (Monson et al. 2002). Because the ecosystem model was normally
163 run using two time steps per day (day and night), we aggregated the half-hourly

164 averaged data up to these time steps, using the gap-filled data where necessary. To
165 minimize the effect of fitting the model to already-modeled “data”, only twice-daily time
166 steps that contained at least 50% measured (i.e. not gap-filled) half-hourly fluxes were
167 included in the optimization (we call these “valid” data points). This led to the exclusion
168 of 23% of the twice-daily observations from the optimization. We chose not to use a
169 more conservative cut-off, because doing so would also have led to the exclusion of
170 more measured data points (but see Braswell et al. 2005 for a discussion of the
171 sensitivity of the parameter estimation to the cut-off criterion). For example, requiring all
172 half-hourly measurements making up a time step to be non-gap-filled would mean that
173 70% of the time steps would have been excluded.

174

175 *The SIPNET ecosystem model*

176 The SIPNET model was designed to simulate the most significant controls over
177 variability in CO₂ fluxes, including the diurnal cycle, while keeping the model as simple
178 as possible (Braswell et al. 2005; Sacks et al. 2006). Using a simple model reduces the
179 number of parameters, thus helping avoid the danger of over-fitting the observations
180 during the optimization, and also reducing the computational intensity of the
181 optimization. The model is, for the most part, a simplified version of the PnET
182 (Photosynthesis-EvapoTranspiration) model (Aber and Federer 1992; Aber et al. 1995;
183 Aber et al. 1996; Aber et al. 1997); acknowledging the contributions of the PnET model,
184 we refer to our version as SIPNET (for Simplified PnET).

185 SIPNET describes carbon flux dynamics in two vegetation carbon pools and an
186 aggregated soil carbon pool (Fig. 1). To model biomass partitioning and respiration, the

187 vegetation is split into leaves and wood, where “wood” refers to the combined pool of
 188 boles, branches and roots. For most model runs we used SIPNET with twice-daily
 189 daytime and nighttime time steps. The lengths of the day and night time steps vary
 190 seasonally to account for changes in day length; fluxes are appropriately scaled for
 191 these changes in the length of the time steps. In each step, six climate variables drive
 192 the flux dynamics: (1) average air temperature (T_{air}), (2) average soil temperature (T_{soil}),
 193 (3) precipitation, (4) flux density of photosynthetic photon flux density ($PPFD$), (5)
 194 relative humidity, and (6) wind speed. Atmospheric vapor pressure and vapor pressure
 195 deficit (VPD) are then computed from relative humidity and air temperature. The
 196 model’s dynamics are governed by the values of 32 parameters, 17 of which were
 197 allowed to vary in the optimization (Table 1), and 15 of which were fixed at values from
 198 the literature or from field studies at Niwot Ridge (Table 2).

199 There are six carbon fluxes in SIPNET: (1) gross ecosystem CO_2 exchange
 200 (GEE, which is synonymous with gross primary productivity and gross photosynthesis),
 201 (2) autotrophic respiration (R_A), (3) heterotrophic respiration (R_H), (4) new leaf growth,
 202 (5) formation of leaf litter, and (6) formation of wood litter. We use the same formulation
 203 for calculating GEE as is used in PnET (Aber and Federer 1992):

204

$$205 \quad GEE_{\text{pot}} = GEE_{\text{max}} D_{\text{temp}} D_{\text{VPD}} D_{\text{light}} \quad (1)$$

206

207 where GEE_{pot} is the potential gross photosynthetic rate per unit leaf mass, without water
 208 stress; GEE_{max} is the daily maximum gross photosynthetic rate, calculated from the
 209 parameters A_{max} , $F_{A_{\text{max}}}$ and K_F ; and D_{temp} , D_{VPD} and D_{light} are scalars that take on values

210 between 0 and 1 and represent decreases in photosynthesis due to temperature, vapor
 211 pressure deficit and light. D_{light} is computed by approximating the integral through the
 212 canopy, considering six vertical layers and computing $D_{\text{light},i}$ for each of these layers.

213 These three scalars are calculated as:

214

$$215 \quad D_{\text{temp}} = \text{Max} \left(\frac{(T_{\text{max}} - T_{\text{air}})(T_{\text{air}} - T_{\text{min}})}{(T_{\text{opt}} - T_{\text{min}})^2}, 0 \right) \quad (2a)$$

$$216 \quad D_{\text{VPD}} = 1 - K_{\text{VPD}} \text{VPD}^2 \quad (2b)$$

$$217 \quad D_{\text{light},i} = 1 - e^{- (I_i \ln(2) / \text{PPFD}_{1/2})} \quad (2c)$$

218

219 where T_{max} is computed as $T_{\text{opt}} + (T_{\text{opt}} - T_{\text{min}})$, I_i is the light intensity at a given layer i ,
 220 computed using Beer's Law, and other parameters are as defined in Table 1.

221 GEE_{pot} , along with the plants' water use efficiency (WUE), is then used to
 222 calculate potential transpiration (T_{pot}), as:

223

$$224 \quad WUE = \frac{K_{\text{WUE}}}{VPD} \quad (3a)$$

$$225 \quad T_{\text{pot}} = \frac{GEE_{\text{pot}}}{WUE} \quad (3b)$$

226

227 Plant-available soil water is computed using the parameter f , and actual transpiration (T)
 228 is set to the lesser of T_{pot} and plant-available water. Finally, GEE is computed using:

229

230
$$D_{\text{water}} = \frac{T}{T_{\text{pot}}} \quad (4a)$$

231
$$GEE = GEE_{\text{pot}} D_{\text{water}} \quad (4b)$$

232

233 Autotrophic respiration is the sum of foliar and wood respiration. To keep the
 234 model simple, we aggregate growth and maintenance respiration. Foliar and wood
 235 respiration are both modeled as:

236

237
$$R_x = K_x C_x Q 10_V^{T_{\text{air}}/10} \quad (5)$$

238

239 where R_x is the realized respiration rate, K_x is a base rate, C_x is the carbon in the given
 240 pool (either wood or leaves), and $Q 10_V$ is the parameter governing the increase in
 241 autotrophic respiration for every 10 degree increase in temperature.

242 Heterotrophic respiration represents the respiration of microbes in the model's
 243 single aggregated soil pool. In addition to a temperature effect similar to that governing
 244 autotrophic respiration, heterotrophic respiration also increases linearly with soil
 245 moisture:

246

247
$$R_H = K_H C_S Q 10_S^{T_{\text{soil}}/10} \left(\frac{W_S}{W_{S,c}} \right) \quad (6)$$

248

249 where C_S is soil carbon content, W_S is soil water content, and other parameters are as
 250 defined in Table 1.

251 The forest in the footprint of the flux tower is dominated by evergreen vegetation,
252 and so we assume an evergreen phenology. In each time step leaf carbon is added at a
253 rate proportional to the mean net primary productivity over the previous five days, and
254 carbon is lost from the leaves (and added to the soil) at a constant rate. Wood litter is
255 also modeled using a constant turnover rate.

256 Photosynthesis in cold climates is drastically reduced when the soil is frozen; a
257 suggested mechanism involves the shutdown of water flow in frozen soils (Hollinger et
258 al. 1999; Monson et al. 2002). Because there is little or no fresh photosynthate supplied
259 to the leaves at this time, it is likely that foliar respiration is also reduced to near-zero.
260 We incorporate these controls into SIPNET by shutting down photosynthesis and foliar
261 respiration when the soil temperature is below a given threshold. Sacks et al. (2006)
262 found that this wintertime downregulation of photosynthesis and foliar respiration
263 improved the model's ability to match observations of NEE.

264 At Niwot Ridge, water stress is one of the largest determinants of NEE (Monson et
265 al. 2002). In this study, in contrast to Sacks et al. (2006), we used a single soil water
266 pool for both evaporation and transpiration, because the CO₂ flux data do not contain
267 enough information to constrain the soil moisture dynamics of a model with two water
268 pools (Sacks et al. 2006). In fact, we found that using an additional water pool actually
269 led to a slightly degraded model-data fit.

270 Evaporation from the soil (E_S) and sublimation from the snow pack (E_P) are
271 modeled as in the SiB2 model (Sellers et al. 1996):

272

$$273 \quad \lambda E_S = \frac{e^*(T_{\text{soil}}) - e_a}{r_{\text{soil}} + r_d} \frac{\rho c_p}{\gamma} \quad (7a)$$

274
$$\lambda_s E_p = \frac{e^*(0^\circ\text{C}) - e_a}{r_d} \frac{\rho c_p}{\gamma} \quad (7b)$$

275

276 where λ is the latent heat of vaporization, λ_s is the latent heat of sublimation, $e^*(T)$ is the
 277 saturation vapor pressure at temperature T , e_a is the atmospheric vapor pressure, ρ is
 278 the density of air, c_p is the specific heat of air, γ is the psychrometric constant, r_d is the
 279 aerodynamic resistance between the ground and the canopy air space, which
 280 decreases linearly with wind speed, and r_{soil} is a soil resistance term, computed as in
 281 Sellers et al. (1992):

282

283
$$r_{\text{soil}} = e^{R_{\text{soil},1} - R_{\text{soil},2} \cdot (W_s / W_{s,c})} \quad (8)$$

284

285 Finally, transpiration is the smaller of potential transpiration and plant-available soil
 286 water, as described above.

287

288 *Parameter optimization*

289 We used a version of the Metropolis simulated annealing algorithm (Metropolis et
 290 al. 1953) to optimize about half of the parameters governing the initial state and time
 291 evolution of SIPNET (Table 1). The remaining parameters (Table 2) were held constant
 292 at values from the literature or from field studies at Niwot Ridge. These latter
 293 parameters were ones that we found, through initial experiments, could not be
 294 estimated well either because of correlations with other parameters, or because they
 295 had only a weak effect on modeled NEE (Schulz et al. 2001; Wang et al. 2001). For

296 example, most initial pool values could not be estimated independently of the various
297 rate constants because of high correlations between these two sets of parameters. The
298 procedure used for this optimization was essentially the same as that described by
299 Braswell et al. (2005).

300 The optimization algorithm generates not only a single best parameter set, but also
301 a range of parameter sets that represent approximately equally good matches to the
302 data. This allows the generation of confidence intervals on the optimal parameter
303 values, and the determination of correlations between parameters. In addition, by
304 running the model forward on the retrieved ensemble of parameter sets, confidence
305 intervals can be generated on the optimized model output. The full deployment of these
306 features of the optimization is described in Braswell et al. (2005), and not repeated
307 here. Additionally, Braswell et al. (2005) present an analysis of the parameter
308 optimization given a synthetic data set. This optimization resulted in a generally good
309 match between the retrieved parameters and the true parameters. Experiments with
310 synthetic data demonstrate that the Metropolis algorithm takes the system to the global
311 constrained minimum for this problem.

312 Initial values and boundaries for the parameters were defined through a
313 combination of literature values, best guesses and actual measurements. Searching the
314 entire parameter space is infeasible, as this space is theoretically infinite. Thus, we
315 bounded each parameter within a range (the “prior distribution”) that is biologically or
316 physically possible based on previous knowledge about the process, therefore
317 eliminating solutions that violate prior knowledge. In general, we specified relatively
318 broad parameter boundaries, and in some cases iteratively increased the boundary

319 widths until the optimized parameter value fell well within the allowable range rather
320 than being on its edge. However, we attempted to avoid making the boundaries so wide
321 that unrealistic estimates for one parameter forced other parameters to take on less
322 realistic values (so, for example, we did not further widen the ranges of Q10_V and
323 Q10_S). Thus the ranges used in the optimization represent an attempt at a compromise:
324 wide enough that the optimum values for most parameters fall within the specified
325 range, but narrow enough that the ranges exclude parameter values that are grossly
326 unrealistic.

327 The optimization consists of performing a quasi-random walk through the multi-
328 dimensional parameter space to find the parameter set that causes the model to
329 generate the best match of predicted NEE (defined as R_A + R_H – GEE) with observed
330 NEE. Although it is possible to use multiple constraints in the optimization (e.g. Wang et
331 al. 2001; Williams et al. 2005), for this study we used NEE alone (but see Sacks et al.
332 2006 for the effects of including water vapor fluxes as an additional constraint). We
333 conducted the analysis across all six years of the available NEE data, and for daytime
334 and nighttime points simultaneously, so the optimized parameter set reflects that
335 determined for the entire data set. The “best match” is defined as the model output that
336 maximizes likelihood (*L*):

337

$$338 \quad L = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma} e^{-(x_i - \mu_i)^2 / 2\sigma^2} \quad (9)$$

339

340 where *n* is the number of valid data points, *x_i* is the eddy covariance-derived NEE
341 summed over time step *i* (in g C m⁻²), *μ_i* is the modeled NEE in time step *i* (in g C m⁻²),

342 and σ is the error (one standard deviation) on each data point. Here σ represents data
343 error *relative to the given model structure*, and thus represents a combination of
344 measurement error and process representation error. Note that we assume that each
345 data point has the same uncertainty and that the deviations from the model predictions
346 are independent over time. In practice, log likelihood is used in place of likelihood
347 because it is computationally easier to work with. Details such as step size and stopping
348 point are discussed in Braswell et al. (2005).

349 Because we did not have measured values of σ , we treated σ as a parameter to
350 be estimated at each step of the optimization (Hurtt and Armstrong 1996). For a given
351 model output (that is, a given set of μ_i values), the value of σ that maximizes L , which
352 we denote σ_e , is given by:

353

$$354 \quad \sigma_e = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \mu_i)^2} \quad (10)$$

355

356 We then used σ_e in place of σ in the calculation of L . Although this procedure may
357 impact the width of the parameter confidence intervals, it should not significantly impact
358 the retrieved optimum parameter set.

359 At each step in the optimization, a parameter is chosen at random and its value
360 changed. The optimization then evaluates the likelihood at the new point (by running the
361 model using the newly-generated parameter set) and compares it to the likelihood at the
362 old point. If $L(\text{new}) \geq L(\text{old})$, then the algorithm accepts the new parameter value. If
363 $L(\text{new}) < L(\text{old})$, then the probability that the algorithm accepts the new parameter value
364 is given by the ratio of the likelihoods. If the new value is accepted, the algorithm takes

365 the next step from this new point. If the new value is rejected, then the algorithm takes
366 another random step using the old value. The occasional acceptance of worse points in
367 the parameter space helps prevent the optimization from getting stuck in local, but non-
368 global, optima. After a spin-up period (similar to that described in Braswell et al. 2005),
369 the set of accepted points is an estimator of the posterior distribution of the parameters,
370 and the estimated parameter set that yielded the maximum likelihood across the entire
371 optimization is taken to be the optimal parameter set. We ran the optimization long
372 enough to generate about 150,000 accepted points.

373

374 **Results**

375 *Optimized fluxes*

376 The optimized parameter set is given in Table 3. The posterior distributions of the
377 parameters are also given; the narrow standard deviations of the estimated parameter
378 means show that the parameters were well constrained by the observations. Braswell et
379 al. (2005) and Sacks et al. (2006) discuss the issue of correlations between the different
380 SIPNET parameters; for the present study we fixed parameters in the optimization if
381 they were highly correlated with another parameter.

382 The model, post-optimization, matched the general diurnal and seasonal patterns
383 in the observations well (Figs. 2a; 3a, b), and also matched the observed cumulative
384 CO₂ uptake in most years (Fig. 2b). There are a few areas of obvious misfit between
385 model and observations. First, the model underestimated peak nighttime respiration
386 (Fig. 2a), possibly because of microbial community and substrate dynamics that are not
387 represented in SIPNET. Second, the model underestimated peak summertime uptake

388 (Fig. 2a). Third, the model failed to fully capture periods of drought-induced decreases
389 in NEE, such as in mid-summer 2002, when the observations indicate a period of net
390 CO₂ loss from the ecosystem to the atmosphere, yet the model predicted continued CO₂
391 uptake (Fig. 3a, b). These last two errors may have arisen because the model was
392 unable to capture the proper soil moisture limitation at this site, so had to make a
393 compromise in photosynthetic rates between times of little water stress (e.g. late spring)
394 and times of significant water stress (e.g. mid-summer 2002) (Sacks et al. 2006).
395 However, the overall fit is satisfactory, with an RMS error between model and valid
396 observations of 0.530 g C m⁻² (daytime points: 0.652 g C m⁻²; nighttime points: 0.300 g
397 C m⁻²).

398 Using the optimized model, we partitioned NEE into its component fluxes, GEE
399 and R_E (Figs. 3c, d; 4a, b). The model predicted mean fluxes (over 1999-2004) of -652
400 g C m⁻² yr⁻¹ for GEE, and 579 g C m⁻² yr⁻¹ for R_E, with a resulting NEE of -73 g C m⁻² yr⁻¹
401 (the observed mean NEE over this time period was -62 g C m⁻² yr⁻¹). The model
402 predicted that the seasonal cycle of NEE is driven primarily by variability in GEE, which
403 had about twice the seasonal range as variability in R_E (Fig. 4a). On average, a key
404 switch in the seasonal cycle occurred in May, when a large increase in GEE occurred
405 and the monthly NEE became negative, indicating net CO₂ uptake. GEE continued to
406 increase between May and June, which caused NEE in June to become even more
407 negative. Month-to-month variability in summertime NEE, however, was predicted to be
408 due more to variability in R_E. Predicted R_E reached a maximum in July, whereas GEE
409 remained roughly constant from June through August; this resulted in net CO₂ uptake
410 being somewhat depressed in July relative to the net uptake in other months of the

411 growing season. NEE variability in the fall, like that in the spring, was governed more
412 strongly by variability in GEE. A large decrease in GEE in November, relative to that in
413 October, caused NEE to become positive in this month. In the winter, there was little
414 month-to-month variability in either the net or gross fluxes.

415 The model predicted that the majority of R_E was autotrophic (R_A), and that only a
416 small fraction was heterotrophic (R_H) (Fig. 4c). The optimized base soil respiration rate
417 (K_H , $0.007 \text{ g C g}^{-1} \text{ C yr}^{-1}$) was almost a factor of ten lower than the initial guess, derived
418 from soil chamber measurements at the site ($0.06 \text{ g C g}^{-1} \text{ C yr}^{-1}$; Scott-Denton et al.
419 2003). While we have reason to believe that the initial guess was too high – and it is for
420 this reason that we specified a wide prior distribution for this parameter – it is unlikely
421 that the initial value was a factor of ten too high. Furthermore, the optimized soil
422 respiration Q_{10} (Q_{10S} , 4.96) was significantly higher than the commonly accepted value
423 of 2. The net effect of these two parameter anomalies was that the model predicted
424 virtually no soil respiration in the cold winter months, and higher, but still relatively low,
425 soil respiration in the summer months (Fig. 4c). In the model, compensation for these
426 low values of R_H occurred through the prediction of relatively high values of R_A . Overall,
427 the model predicted mean respiration fluxes (over 1999-2004) of $94 \text{ g C m}^{-2} \text{ yr}^{-1}$ R_H and
428 $486 \text{ g C m}^{-2} \text{ yr}^{-1}$ R_A .

429 In order to probe possible biases in the way that the model treats R_H , we
430 performed an experiment in which we held the K_H parameter fixed at two different
431 values: $0.016 \text{ g C g}^{-1} \text{ C yr}^{-1}$ and $0.032 \text{ g C g}^{-1} \text{ C yr}^{-1}$. These values were chosen to make
432 annual R_H approximately equal to observed aboveground litterfall and twice observed
433 aboveground litterfall, respectively. (Total aboveground litterfall at this site averaged 184

434 g C m⁻² yr⁻¹ over the period 2000-2001; L. Scott-Denton, University of Colorado,
435 unpublished data.) The higher values of K_H caused increases in annual R_H (Table 4).
436 However, the increases in R_H were less than the increases in K_H because the
437 optimization caused a decrease in Q10_S when forced with higher initial values of K_H:
438 even when K_H was increased by as much as a factor of five, annual R_H increased by
439 less than a factor of two. The optimization protocol seemed to “favor” keeping R_H
440 relatively low, so that even when forced with the initial condition of annual R_H equal to
441 twice the observed aboveground litterfall carbon content, the optimized
442 parameterization produced a final solution with R_H slightly less than observed
443 aboveground litterfall.

444 Despite large perturbations to a key model parameter (i.e., R_H), the GEE/R_E ratio
445 remained approximately the same in the three runs performed for this experiment.
446 Modeled annual R_A decreased to compensate for increases in annual R_H (Table 4). An
447 experiment in which we fixed A_{max}, the maximum photosynthetic rate, at its initial guess
448 value – about 70% greater than its optimized value – gave similar results (data not
449 shown). Here, too, the change in the average GEE/R_E partitioning was negligible, but
450 there was a significant change in the R_A/R_H partitioning, with annual average R_H about
451 50% higher, and R_A slightly lower than in the control run.

452

453 *Optimization with different time steps*

454 We performed two experiments to test the optimization’s sensitivity to the time
455 scale of the model-data comparisons. First, we investigated how much information is
456 lost by performing model-data comparisons once per day, instead of twice per day as in

457 the standard setup. Second, we investigated how much information is gained by
458 performing model-data comparisons on the original half-hourly fluxes.

459 Because SIPNET was structured to assimilate daytime and nighttime fluxes
460 separately, and then is optimized on both simultaneously, the optimized model matched
461 the observed diurnal cycle fairly well (Fig. 5). The model slightly underestimated
462 daytime uptake early in the growing season, underestimated nighttime respiration in the
463 summer, and tended to overestimate both daytime uptake and nighttime respiration in
464 November, the key month for switching from net monthly CO₂ uptake to net monthly
465 CO₂ loss (Fig. 5a). This resulted in a consistent negative bias in the model's prediction
466 of annual nighttime respiration, but little bias in the model's prediction of annual daytime
467 CO₂ uptake; overall, this resulted in a slight overestimate of net CO₂ uptake relative to
468 the observations (Fig. 5b). However, the biases that exist are small, and this lends
469 support to the model's separation of GEE and R_E.

470 We evaluated the model structure with regard to diurnal information on parameter
471 retrieval by aggregating modeled and observed fluxes to a daily time step before
472 performing the model-data comparison. The model was still run at the same twice-daily
473 resolution, so the driving meteorological data remained the same. With this aggregation,
474 the model still predicted total NEE well, but significantly underestimated both daytime
475 uptake and nighttime respiration during the growing season (Fig. 6; cf. Fig. 5a).

476 We performed an additional optimization in which we used the half-hourly
477 meteorological and flux data, to test the possibility that some aspects of model-data
478 mismatch were caused by temporal resolution. This half-hourly optimization led to a few
479 significant differences in retrieved parameter values (Table 3). The most significant

480 difference was in the $PPFD_{1/2}$ parameter. The increase in $PPFD_{1/2}$ also forced an
481 increase in A_{max} in order to keep average realized GEE from changing too much; this in
482 turn forced a decrease in K_F . The K_{WUE} parameter, governing plant water use efficiency,
483 was also much higher in the half-hourly optimization. However, because the modeled
484 water fluxes were not directly constrained by observations, the parameters governing
485 these fluxes could vary widely in the optimization (Wang et al. 2001; Sacks et al. 2006).

486 Although the half-hourly optimization had the capacity to capture more
487 nonlinearities in the ecosystem's response to environmental drivers, it actually led to
488 about a 5% *increase* in the RMS model-data error (after aggregating from the half-
489 hourly time steps to a twice-daily time step). Furthermore, the half-hourly optimization
490 led to only small improvements in the major areas of model-data misfit, estimating
491 slightly greater peak summertime uptake and peak nighttime respiration than the twice-
492 daily optimization. The flux partitioning was similar in the half-hourly optimization, with
493 about a 5% increase in the estimated values of both GEE and R_E . The estimated value
494 of R_H increased by 33%, but was still too low.

495

496 **Discussion**

497 Model-data fusion is a relatively new approach to data and model analysis that
498 provides a high level of empirical constraint over model predictions, while at the same
499 time probing model structure and parameter relationships for the source of model-data
500 mismatch (e.g. Schulz et al. 2001; Wang et al. 2001; Braswell et al. 2005; Raupach et
501 al. 2005; Williams et al. 2005; Sacks et al. 2006). Model-data fusion holds great promise
502 as a tool to partition the photosynthetic and respiratory components of NEE, and study

503 their separate responses to environmental control. The assimilation system that we
504 used successfully separated GEE and R_E ; the separation into component fluxes is
505 based on variations in the observations themselves and process-level understanding,
506 as embodied in the model. There was good agreement between the model predictions
507 and observations of NEE at both the seasonal and interannual time scales.

508 When the model was fit to total daily NEE, as opposed to separate day and night
509 NEE, it predicted seasonal and interannual patterns of NEE equally well, but for the
510 wrong reasons. While modeled daily NEE remained within the uncertainty of the
511 observations, it arose from greatly underestimated fluxes of both GEE and R_E . Since the
512 total diurnal and separated day/night model structures were identical, differences in their
513 predictions of GEE and R_E were due solely to differences in parameter values.

514 The results of the above experiment have implications for using daily or longer
515 estimates of NEE to validate ecosystem models. Eddy covariance data are often
516 aggregated to longer time scales to produce estimates that are less uncertain than the
517 individual half-hourly measurements; however, aggregation to the daily, weekly, and
518 longer time scales reduces the information content of the data. As an alternative,
519 separate aggregation of the daytime and nighttime hours to daily and longer time scales
520 should be considered (Figs. 5 and 6). This produces robust time-integrals while allowing
521 a degree of independent assessment of GEE and R_E .

522 However, little advantage was gained by using half-hourly observations to run and
523 optimize the model. Most model processes are sensitive primarily to environmental
524 variations on time scales greater than thirty minutes. The main reason to use these

525 short time steps would be to estimate parameters associated with photosynthetic
526 physiology or transpiration, which are sensitive to diurnal variation in light and humidity.

527 There is a trend toward adopting carbon models that use short time steps (seconds
528 to minutes); the reductionist assumption driving this trend is that shorter time steps will
529 permit greater mechanistic realism. The principal challenge in moving models toward
530 these shorter time steps is that the seconds-to-minutes environmental information
531 required to drive the models is difficult to obtain. However, even with the necessary
532 meteorological drivers, it is not always advantageous to run the model at these shorter
533 time steps. As the difference in the retrieved values of the $PPFD_{1/2}$ parameter illustrates
534 (Table 3), an ecosystem's sensitivity to environmental conditions can vary with the time
535 scale under consideration. A model that is designed to perform well on a short time
536 scale will not necessarily perform well on longer time scales.

537 The modeled separation of heterotrophic and autotrophic respiration was far less
538 satisfactory than the overall separation of NEE into GEE and R_E . While the day-night
539 contrast allows some independent estimation of GEE and R_E , flux data contain less
540 embedded information capable of constraining the respiratory separation. This
541 separation relies primarily on subtle differences between the correlation of NEE with air
542 temperature and its correlation with soil temperature. If the system is close to steady
543 state, then heterotrophic respiration should be close to the sum of aboveground litterfall
544 and belowground root turnover. Although this system is not in steady state, it is almost
545 certainly closer to steady state than is suggested by this analysis.

546 At Niwot Ridge, longer growing seasons decrease carbon uptake. In this subalpine
547 ecosystem, one of the principal determinants of annual NEE is the length of the spring

548 snowmelt period, and the magnitude of NEE during this period (Monson et al. 2002,
549 2005). Early spring warming tends to be correlated with shallow late-spring snow pack,
550 and as a result, low springtime and annual net CO₂ uptake. This was confirmed in the
551 current study, in which a longer growing season (defined as the interval between the
552 spring onset of net negative NEE and the fall transition back to positive NEE: the
553 interval between zero crossings) due to an early spring warm-up was associated with a
554 decreased annual net CO₂ uptake, by an average of 1.37 g C m⁻² day⁻¹ using the
555 observed NEE (Fig. 7) (the relationship using modeled NEE was similar). The model-
556 data fusion's partitioning of NEE into its components showed that variability in GEE
557 explained 56% and 80% of the year-to-year variation in observed and modeled NEE,
558 respectively. Thus we surmise that growing season length affects NEE primarily through
559 its influence on GEE, and particularly through the relationship between the timing of
560 snowmelt and growing season water availability.

561 Churkina et al. (2005) recently showed that a strong relationship exists between
562 the carbon uptake period (defined as the number of days per year with negative NEE)
563 and annual NEE across many evergreen needle-leafed sites from the FLUXNET
564 network. The slope of that relationship, determined among many separate sites, was
565 larger than the relatively weak interannual response at our single site (data not shown).
566 This is in accordance with the prediction of Churkina et al. (2005) that the year-to-year
567 response at a single site should differ from that determined across broader spatial
568 scales. Thus, the spatial and interannual responses of NEE to growing season length
569 should not be equated.

570 Different measures of the length of the growing season or photosynthetic period
571 produce strikingly different, though ultimately consistent, conclusions. At Niwot Ridge,
572 the length of the carbon uptake period (days of negative NEE) was inversely correlated
573 with the growing season length (the interval between the spring and fall zero crossings)
574 (data not shown). Thus, years with earlier springs had more days with positive daily
575 NEE (carbon efflux) during the growing season – i.e. warm springs tend to lead into
576 summers with increased stress. All else being equal, these two measures should be
577 positively correlated, but at Niwot, the tendency is for an inverse correlation. This
578 tendency is likely a result of three factors: first, one cause of early spring is a low snow
579 pack which means low water availability; second, warm springs may alter hydrological
580 partitioning, increasing runoff and evaporation and decreasing the water available for
581 transpiration; third, climatologically warm springs have been correlated with warm
582 summers, possibly increasing stress and reducing carbon uptake.

583 Interactions between temperature and moisture are important. For example, the
584 spring of 2004 supported a relatively high rate of CO₂ uptake despite a warm
585 temperature anomaly; however, it also had a large positive precipitation anomaly.
586 Conversely, the spring of 2003 supported a high rate of CO₂ uptake despite a dry
587 precipitation anomaly; however, it also had a cold temperature anomaly. In broad terms,
588 it appears that temperature and precipitation anomalies can compensate for each other.
589 A deep spring snow pack can override the negative effects of warm temperature and
590 low spring temperatures can override the negative effects of a shallow snow pack. The
591 model-data fusion that we employed correctly identified the major carbon flux anomalies
592 associated with spring variability, agreeing with the observations that carbon uptake

593 began earliest in 2000, latest in 1999, and that 2001, 2002, 2003 and 2004 were similar
594 to each other (Fig. 3a, b).

595 Model-data fusion techniques applied to long eddy covariance time series can
596 illuminate the impact of climate on NEE, GEE and R_E . Our analysis has revealed that, in
597 general, warmer conditions during the growing season at the Niwot Ridge forest will
598 increase GEE and NEE only if adequate moisture is present. Predictions of impacts of
599 temperature changes on carbon balance will depend critically on the assumptions made
600 about the hydrological cycle. If warmer conditions result in a shorter snowmelt period,
601 then warming will likely reduce net CO_2 uptake. Conversely, if warming is accompanied
602 by an intensified hydrological cycle (Trenberth et al. 2003) with increased winter or
603 summer precipitation, GEE and net uptake could increase.

604

605 **Acknowledgements**

606 Rob Braswell provided much of the impetus for the SIPNET modeling, and
607 helped us develop many of the ideas that led to this study; we are grateful for his
608 intellectual contributions. We are also grateful for Ernst Linder's help with the statistics
609 underlying the parameter estimation. We thank Sean Burns, Andrew Turnipseed, Laura
610 Scott-Denton, and others who were instrumental in the data collection. Thanks also to
611 Galina Churkina for insights into growing season length. Finally, we thank three
612 anonymous reviewers for their constructive comments. This work was supported by
613 grants from the National Aeronautics and Space Administration (TE/02-0000-0015 &
614 NAG5-12876), the National Science Foundation (2003108), the National Oceanic and
615 Atmospheric Administration (2002192), and a grant from the South Central Section of

616 the National Institute for Global Environmental Change (NIGEC) through the US
617 Department of Energy (BER Program) (Cooperative Agreement No. DE-FC03-
618 90ER61010). The National Center for Atmospheric Research is funded by the National
619 Science Foundation.

620 The experiments performed for this study comply with the current laws of the
621 United States of America.

622
623 **References**

- 624
625 Aber JD, Federer CA (1992) A generalized, lumped-parameter model of photosynthesis,
626 evapotranspiration and net primary production in temperate and boreal forest
627 ecosystems. *Oecologia* 92:463-474
628
629 Aber JD, Ollinger SV, Driscoll CT (1997) Modeling nitrogen saturation in forest
630 ecosystems in response to land use and atmospheric deposition. *Ecol Model* 101:61-78
631
632 Aber JD, Ollinger SV, Federer CA, Reich PB, Goulden ML, Kicklighter DW, Melillo JM,
633 Lathrop RG (1995) Predicting the effects of climate change on water yield and forest
634 production in the northeastern united states. *Climate Res* 5:207-222
635
636 Aber JD, Reich PB, Goulden ML (1996) Extrapolating leaf CO₂ exchange to the
637 canopy: A generalized model of forest photosynthesis compared with measurements by
638 eddy correlation. *Oecologia* 106:257-265
639
640 Angert A, Biraud S, Bonfils C, Henning CC, Buermann W, Pinzon J, Tucker CJ, Fung I
641 (2005) Drier summers cancel out the CO₂ uptake enhancement induced by warmer
642 springs. *Proc Natl Acad Sci USA* 102:10823-10827
643
644 Baldocchi DD (2003) Assessing the eddy covariance technique for evaluating carbon
645 dioxide exchange rates of ecosystems: Past present and future. *Glob Change Biol*
646 9:479-492
647
648 Braswell BH, Sacks WJ, Linder E, Schimel DS (2005) Estimating diurnal to annual
649 ecosystem parameters by synthesis of a carbon flux model with eddy covariance net
650 ecosystem exchange observations. *Glob Change Biol* 11:335-355
651
652 Breshears DD, Cobb NS, Rich PM, Price KP, Allen CD, Balice RG, Romme WH,
653 Kastens JH, Floyd ML, Belnap J, Anderson JJ, Myers OB, Meyer CW (2005) Regional
654 vegetation die-off in response to global-change-type drought. *Proc Natl Acad Sci USA*
655 102:15144-15148

656
657 Churkina G, Schimel D, Braswell BH, Xiao XM (2005) Spatial analysis of growing
658 season length control over net ecosystem exchange. *Glob Change Biol* 11:1777-1787
659

660 Ciais P, Reichstein M, Viovy N, Granier A, Ogee J, Allard V, Aubinet M, Buchmann N,
661 Bernhofer C, Carrara A, Chevallier F, De Noblet N, Friend AD, Friedlingstein P,
662 Grunwald T, Heinesch B, Keronen P, Knohl A, Krinner G, Loustau D, Manca G,
663 Matteucci G, Miglietta F, Ourcival JM, Papale D, Pilegaard K, Rambal S, Seufert G,
664 Soussana JF, Sanz MJ, Schulze ED, Vesala T, Valentini R (2005) Europe-wide
665 reduction in primary productivity caused by the heat and drought in 2003. *Nature*
666 437:529-533
667

668 Goulden ML, Munger JW, Fan SM, Daube BC, Wofsy SC (1996) Measurements of
669 carbon sequestration by long-term eddy covariance: Methods and a critical evaluation of
670 accuracy. *Glob Change Biol* 2:169-182
671

672 Hollinger DY, Goltz SM, Davidson EA, Lee JT, Tu K, Valentine HT (1999) Seasonal
673 patterns and environmental control of carbon dioxide and water vapour exchange in an
674 ecotonal boreal forest. *Glob Change Biol* 5:891-902
675

676 Hurtt GC, Armstrong RA (1996) A pelagic ecosystem model calibrated with BATS data.
677 *Deep-Sea Res Pt II* 43:653-683
678

679 Kaimal JC, Finnigan JJ (1994) Atmospheric boundary flows: Their structure and
680 measurement. Oxford University Press, New York
681

682 Kendall MG, Ord JK (1990) Time series. Oxford University Press, New York
683

684 Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E (1953) Equation of
685 state calculations by fast computing machines. *J Chem Phys* 21:1087-1092
686

687 Monson RK, Sparks JP, Rosenstiel TN, Scott-Denton LE, Huxman TE, Harley PC,
688 Turnipseed AA, Burns SP, Backlund B, Hu J (2005) Climatic influences on net
689 ecosystem CO₂ exchange during the transition from wintertime carbon source to
690 springtime carbon sink in a high-elevation, subalpine forest. *Oecologia* 146:130-147
691

692 Monson RK, Turnipseed AA, Sparks JP, Harley PC, Scott-Denton LE, Sparks K,
693 Huxman TE (2002) Carbon sequestration in a high-elevation, subalpine forest. *Glob*
694 *Change Biol* 8:459-478
695

696 Raupach MR, Rayner PJ, Barrett DJ, DeFries RS, Heimann M, Ojima DS, Quegan S,
697 Schimel CC (2005) Model-data synthesis in terrestrial carbon observation: Methods,
698 data requirements and data uncertainty specifications. *Global Change Biol* 11:378-397
699

700 Ryan MG, Law BE (2005) Interpreting, measuring, and modeling soil respiration.
701 *Biogeochemistry* 73:3-27

702
703 Sacks WJ, Schimel DS, Monson RK, Braswell BH (2006) Model-data synthesis of
704 diurnal and seasonal CO₂ fluxes at Niwot Ridge, Colorado. *Glob Change Biol* 12:240-
705 259
706
707 Schulz K, Jarvis A, Beven K, Soegaard H (2001) The predictive uncertainty of land
708 surface fluxes in response to increasing ambient carbon dioxide. *J Clim* 14:2551-2562
709
710 Scott-Denton LE, Rosenstiel TN, Monson RK (2006) Differential controls over the
711 heterotrophic and rhizospheric components of soil respiration in a high-elevation,
712 subalpine forest. *Glob Change Biol* 12:205-216
713
714 Scott-Denton LE, Sparks KL, Monson RK (2003) Spatial and temporal controls of soil
715 respiration rate in a high-elevation, subalpine forest. *Soil Biol Biochem* 35:525-534
716
717 Sellers PJ, Heiser MD, Hall FG (1992) Relationship between surface conductance and
718 spectral vegetation indices at intermediate (100 m²–15 m²) length scales. *Journal of*
719 *Geophysical Research, FIFE Special Issue* 97:19033-19060
720
721 Sellers PJ, Randall DA, Collatz GJ, Berry JA, Field CB, Dazlich DA, Zhang C, Collelo
722 GD, Bounoua L (1996) A revised land surface parameterization (SiB2) for atmospheric
723 GCMs .1. model formulation. *J Clim* 9:676-705
724
725 Trenberth KE, Dai AG, Rasmussen RM, Parsons DB (2003) The changing character of
726 precipitation. *B Am Meteorol Soc* 84:1205-1217
727
728 Turnipseed AA, Anderson DE, Blanken PD, Baugh WM, Monson RK (2003) Airflows
729 and turbulent flux measurements in mountainous terrain. Part 1: Canopy and local
730 effects. *Agr Forest Meteorol* 119:1-21
731
732 Turnipseed AA, Anderson DE, Burns S, Blanken PD, Monson RK (2004) Airflows and
733 turbulent flux measurements in mountainous terrain. Part 2: Mesoscale effects. *Agr*
734 *Forest Meteorol* 125:187-205
735
736 Turnipseed AA, Blanken PD, Anderson DE, Monson RK (2002) Energy budget above a
737 high-elevation subalpine forest in complex topography. *Agr Forest Meteorol* 110:177-
738 201
739
740 Wang YP, Leuning R, Cleugh HA, Coppin PA (2001) Parameter estimation in surface
741 exchange models using nonlinear inversion: How many parameters can we estimate
742 and which measurements are most useful? *Global Change Biol* 7:495-510
743
744 Webb EK, Pearman GI, Leuning R (1980) Correction of flux measurements for density
745 effects due to heat and water vapor transfer. *Quart J Roy Meteorol Soc* 106:85-100
746

747 Williams M, Schwarz PA, Law BE, Irvine J, Kurpius MR (2005) An improved analysis of
748 forest carbon dynamics using data assimilation. *Global Change Biol* 11:89-105

749 **Table 1.** SIPNET parameters and initial conditions that were allowed to vary in the optimization,
 750 and their allowable ranges. The ranges assume a uniform prior distribution.
 751

Symbol	Definition	Range
<i>Initial Pool Values:</i>		
$W_{S,0}$	Initial soil moisture content (fraction of $W_{S,c}$)	0 – 1
<i>Photosynthesis/Respiration Parameters:</i>		
A_{max}	Maximum net CO ₂ assimilation rate (nmol CO ₂ g ⁻¹ (leaf biomass) s ⁻¹)	0 – 34
T_{min}	Minimum temperature for photosynthesis (°C)	-8 – 8
T_{opt}	Optimum temperature for photosynthesis (°C)	5 – 30
K_{VPD}	Slope of VPD-photosynthesis relationship (kPa ⁻¹)	0.01 – 0.25
$PPFD_{1/2}$	Half saturation point of PPFD-photosynthesis relationship (mol m ⁻² day ⁻¹)	4 – 27
K_F	Foliar maintenance respiration as fraction of A_{max} (no units)	0.05 – 0.30
K_A	Wood respiration rate at 0°C (g C g ⁻¹ C yr ⁻¹)	0.0006 – 0.06
Q_{10V}	Vegetation respiration Q_{10} (no units)	1.4 – 2.6
T_s	Soil temperature at which photosynthesis and foliar respiration are shut down (°C)	-5 – 5
K_H	Soil respiration rate at 0°C and moisture-saturated soil (g C g ⁻¹ C yr ⁻¹)	0.003 – 0.6
Q_{10S}	Soil respiration Q_{10} (no units)	1.4 – 5
<i>Water-Related Parameters:</i>		
f	Fraction of soil water removable in one day (no units)	0.001 – 0.16
K_{WUE}	VPD-water use efficiency relationship (mg CO ₂ kPa g ⁻¹ H ₂ O)	0.01 – 109
$W_{S,c}$	Soil water holding capacity (cm (precipitation equivalent))	0.1 – 36
R_d	Scalar relating aerodynamic resistance to wind speed (no units) ^a	1 – 1500
<i>Tree Physiological Parameters:</i>		
NPP_L	Fraction of NPP allocated to leaf growth (no units)	0 – 1

752 (a) $r_d = R_d/u$ (r_d = aerodynamic resistance; u = wind speed)

753 **Table 2.** SIPNET parameters and initial conditions that were held constant in the optimization.
 754 Most of these parameters were highly correlated with other parameters in the model, and thus
 755 could not be estimated independently.
 756

Symbol	Definition	Value	Source ^a
Initial Pool Values:			
$C_{W,0}$	Initial plant wood C content (g C m ⁻²)	9600	RM
$C_{L,0}$	Initial leaf area index (m ² m ⁻²)	4.2	M02
$C_{S,0}$	Initial soil C content (organic layer only) (g C m ⁻²)	16000	S03
$W_{P,0}$	Initial snow pack (cm water equivalent)	0	N/A
Photosynthesis/Respiration Parameters:			
$F_{A_{max}}$	Average daily max photosynthesis as fraction of A_{max} (no units)	0.76	A96
k	Canopy PPFD extinction coefficient (no units)	0.5	A96
Water-Related Parameters:			
E	Fraction of rain immediately intercepted and evaporated (no units)	0.1	A92
F	Fraction of water entering soil that goes directly to drainage (no units)	0.1	A92
K_S	Snowmelt rate (cm (water equivalent) °C ⁻¹ day ⁻¹)	0.15	A92
$R_{soil,1}$	Scalar relating soil resistance to soil wetness (see text) (no units)	0 – 16.4	S96
$R_{soil,2}$	Scalar relating soil resistance to soil wetness (see text) (no units)	0 – 8.6	S96
Tree Physiological Parameters:			
SLW_C	C content of leaves on a per-area basis (g C m ⁻² (leaf area))	270	JS
F_C	Fractional C content of leaves (g C g ⁻¹ (leaf biomass))	0.45	A95
K_L	Turnover rate of leaf C (g C g ⁻¹ C yr ⁻¹)	0.13	LS
K_W	Turnover rate of plant wood C (g C g ⁻¹ C yr ⁻¹)	0.014	PT

757 (a) Sources are: A92 – Aber and Federer 1992; A95 – Aber *et al.* 1995; A96 – Aber *et al.* 1996;
 758 JS – J. Sparks, unpublished data; LS – L. Scott-Denton, unpublished data; M02 – Monson *et al.*
 759 2002; PT – P. Thornton, pers. comm., 1/21/05; RM – R. Monson, unpublished data; S96 – Sellers
 760 *et al.* 1996; S03 – Scott-Denton *et al.* 2003
 761

762 **Table 3.** Parameter values retrieved from the optimization (see Table 1 for parameter
763 definitions). The “Best value” column reports the parameter set that yielded the highest
764 likelihood in the standard (twice-daily) parameter estimation. The “Mean value” column reports
765 the estimated posterior mean and standard deviation of each parameter, generated from 150,081
766 parameter sets that yielded approximately equally good model-data fits. The “Half-hourly mean”
767 column reports the estimated mean and standard deviation of each parameter when the model
768 was run and optimized at a half-hourly time step. Values in bold differ from the standard
769 optimization by more than two standard deviations.
770

Parameter	Best value	Mean value	Half-hourly mean
$W_{S,0}$	0.02	0.03 ± 0.01	0.020 ± 0.003
A_{\max} (nmol CO ₂ g ⁻¹ s ⁻¹)	4.79	4.77 ± 0.15	8.73 ± 0.08
T_{\min} (°C)	-2.97	-3.13 ± 0.20	-2.78 ± 0.09
T_{opt} (°C)	14.3	14.2 ± 0.5	19.4 ± 0.3
K_{VPD} (kPa ⁻¹)	0.127	0.123 ± 0.006	0.138 ± 0.001
PPFD _{1/2} (mol m ⁻² day ⁻¹)	7.4	7.5 ± 0.5	14.9 ± 0.2
K_F	0.097	0.100 ± 0.009	0.064 ± 0.002
K_A (g C g ⁻¹ C yr ⁻¹)	0.027	0.028 ± 0.001	0.027 ± 0.001
Q_{10V}	1.40	1.43 ± 0.02	1.46 ± 0.02
T_s (°C)	0.106	0.105 ± 0.002	-0.002 ± 0.002
K_H (g C g ⁻¹ C yr ⁻¹)	0.007	0.006 ± 0.001	0.0082 ± 0.0002
Q_{10S}	4.96	4.66 ± 0.28	4.99 ± 0.01
f	0.088	0.075 ± 0.019	0.059 ± 0.002
K_{WUE} (mg CO ₂ kPa g ⁻¹ H ₂ O)	47.7	59.1 ± 15.2	106.2 ± 2.6
$W_{S,c}$ (cm)	4.08	4.03 ± 0.27	3.46 ± 0.04
R_d	36.5	36.5 ± 0.9	33.7 ± 0.3
NPPL	0.40	0.41 ± 0.03	0.37 ± 0.01

771

772 **Table 4.** The effect of holding the base soil respiration rate (K_H) fixed at different values.
 773 Modeled fluxes (gross photosynthesis (GEE), total ecosystem respiration (R_E), autotrophic
 774 respiration (R_A), heterotrophic respiration (R_H), net primary productivity (NPP), and litterfall)
 775 are given as means over 1999-2004. All fluxes are in $\text{g C m}^{-2} \text{ yr}^{-1}$. Notice that the optimized soil
 776 respiration Q_{10} value (Q_{10s}) decreases as K_H increases, dampening the increase in annual R_H
 777 with increasing K_H .
 778

	Optimized K_H (0.007 $\text{g C g}^{-1} \text{ C yr}^{-1}$)	K_H fixed at 0.016 $\text{g C g}^{-1} \text{ C yr}^{-1}$	K_H fixed at 0.032 $\text{g C g}^{-1} \text{ C yr}^{-1}$
Best log likelihood ^a	-2727.0	-2748.0	-2806.6
RMS error ^b	0.530	0.534	0.543
# free parameters	17	16	16
BIC ^c	5592.6	5626.4	5743.6
Q_{10s} (optimized)	4.96	2.51	1.43
GEE	652	652	653
R_E	579	579	583
R_A	486	447	407
R_H	94	132	176
NPP	167	205	245
Litterfall (model) ^d	265	267	268
Litterfall (obs.) ^e	184	184	184

779 (a) Larger (i.e. closer to zero) numbers mean greater likelihood.
 780 (b) Root mean square error in g C m^{-2} over a single time step, for valid data points only.
 781 (c) *BIC* (Bayesian Information Criterion) = $-2 LL + K \ln(n)$, where *LL* is the log likelihood, *K* is the number of free
 782 parameters, and *n* is the number of data points used in the optimization (3474). A lower *BIC* indicates a model with
 783 greater support from the data (Kendall & Ord 1990).
 784 (d) Includes root turnover as well as aboveground litterfall. The parameters governing this flux in the model were all
 785 held fixed in the optimization, at values derived from a variety of field studies (see Table 2).
 786 (e) Aboveground litterfall only. Mean over seven plots and two years (2000-2001) (L. Scott-Denton, unpublished
 787 data).

788 **Figure Legends**

789

790 Figure 1: SIPNET pools and fluxes. The model has two vegetation carbon pools and one soil
791 carbon pool. The soil moisture sub-model includes a single soil moisture pool and a snow pack.
792 Soil moisture affects both photosynthesis and soil respiration, as described in the text.

793

794 Figure 2: Twice-daily (day and night separated) time series of modeled and observed NEE (a),
795 and cumulative NEE (b), from Nov. 1, 1998 through Dec. 31, 2004. Model results were
796 generated using the best parameter set retrieved from the optimization. Positive NEE denotes net
797 loss of CO₂ from the ecosystem to the atmosphere. Major errors between the model and
798 observations are (1) the model underestimates peak summertime CO₂ uptake, (2) the model
799 underestimates peak nighttime respiration rates, and (3) the model does not capture the mid-
800 summer NEE reduction in 2002.

801

802 Figure 3: Observed (a) and modeled (b) NEE, and modeled GEE (c) and R_E (d) for the years
803 1999-2004. By separating NEE into GEE and R_E, we are better able to diagnose causes of the
804 observed interannual variability. Fluxes are shown as running 28-day means.

805

806 Figure 4: Optimized breakdown of NEE into its component fluxes, gross photosynthesis (GEE)
807 and total ecosystem respiration (R_E), by month (a) and by year (b), and of total respiration into
808 autotrophic and heterotrophic respiration (c). November and December fluxes give means over
809 1998-2004, while other months give means over 1999-2004. Error bars show interannual
810 variability for each month (a, c), or for the entire year (b) (one standard deviation). The model
811 seems to be able to separate NEE into GEE and R_E well, but does not seem to be able to estimate
812 an accurate separation of R_E into R_A and R_H (see text).

813

814 Figure 5: Modeled and observed mean fluxes by month (a) and year (b), broken down into day
815 and night. Error bars show interannual variability for each month (a), or for the entire year (b)
816 (one standard deviation). Because we ran and optimized the model at a twice-daily (daytime and
817 nighttime) time step, the optimized model fit the observations fairly well both during the day and
818 at night. This lends support to our partitioning of NEE into GEE and R_E.

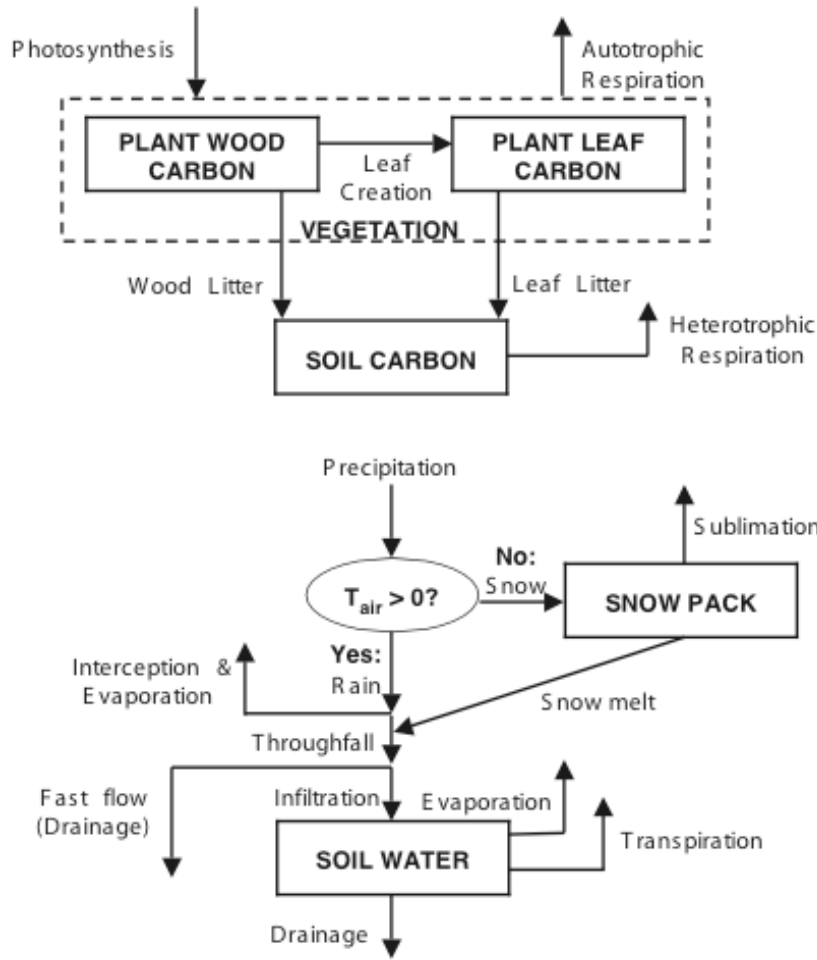
819

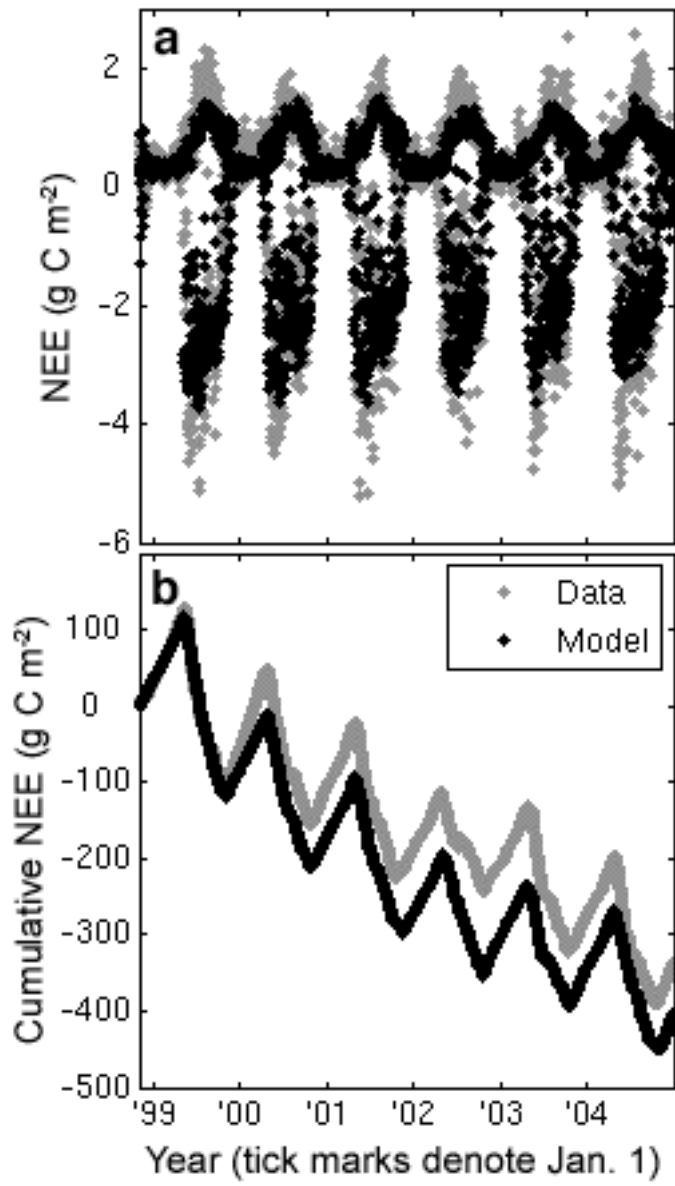
820 Figure 6: Modeled and observed mean fluxes by month, broken down into night and day, for an
821 optimization in which we aggregated the observations up to a daily time step, no longer
822 separating day and night. Error bars show interannual variability for each month (one standard
823 deviation). When such an aggregation is performed, the model is unable to match the correct
824 diurnal cycle of the observations, and thus incorrectly predicts both GEE and R_E.

825

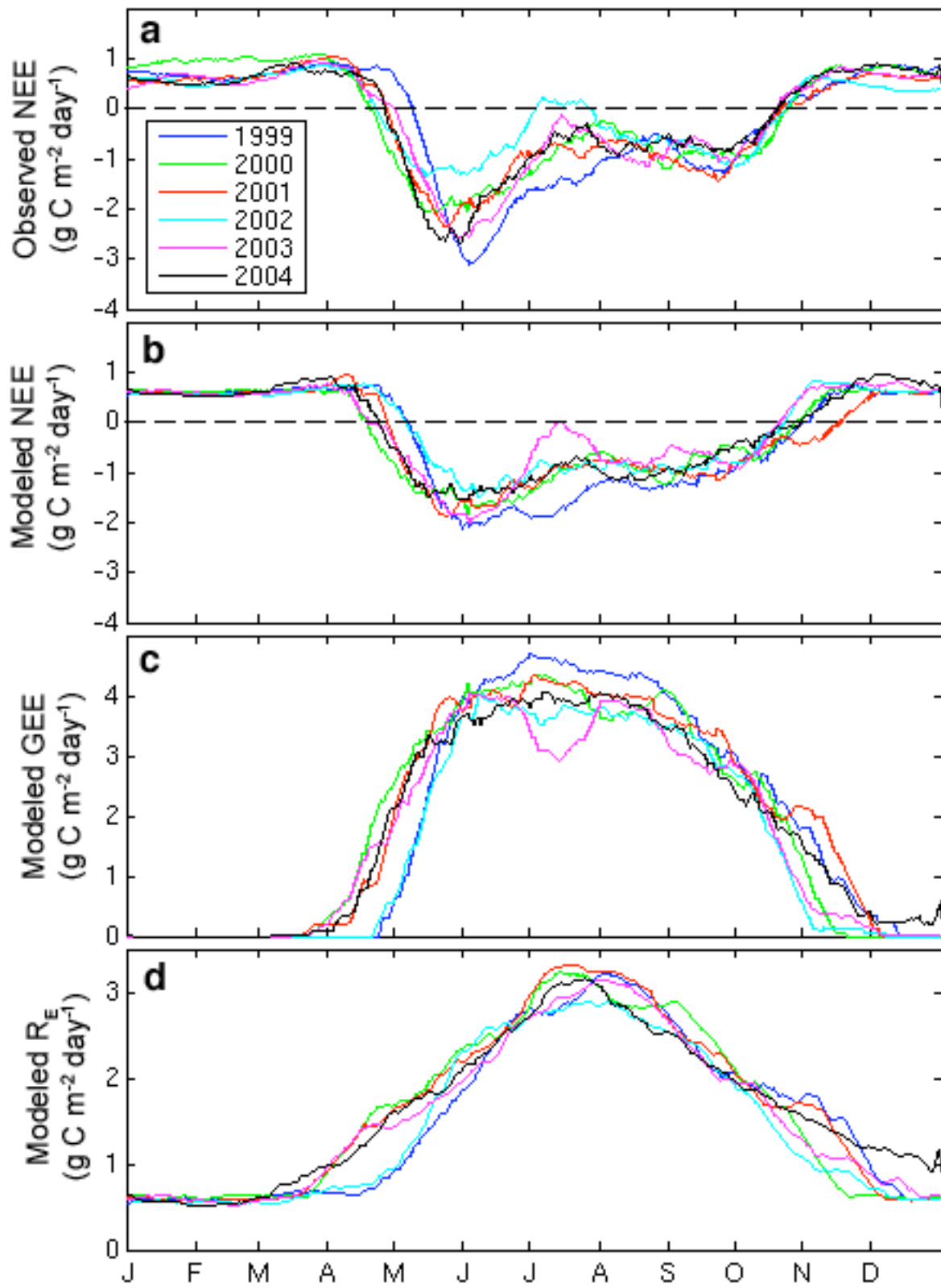
826 Figure 7: Relationship between growing season length and observed cumulative annual NEE,
827 1999 - 2004. "Growing season length" is defined here as the interval between the spring onset of
828 net negative NEE and the fall transition back to positive NEE: the interval between zero
829 crossings. Longer growing seasons are associated with *less* annual CO₂ uptake, probably because
830 of the effects of snowmelt timing on water availability. The best-fit line, shown here, is given by
831 the equation $y = 1.37x - 305$ ($R^2 = 0.45$).

832 **Figure 1**
833

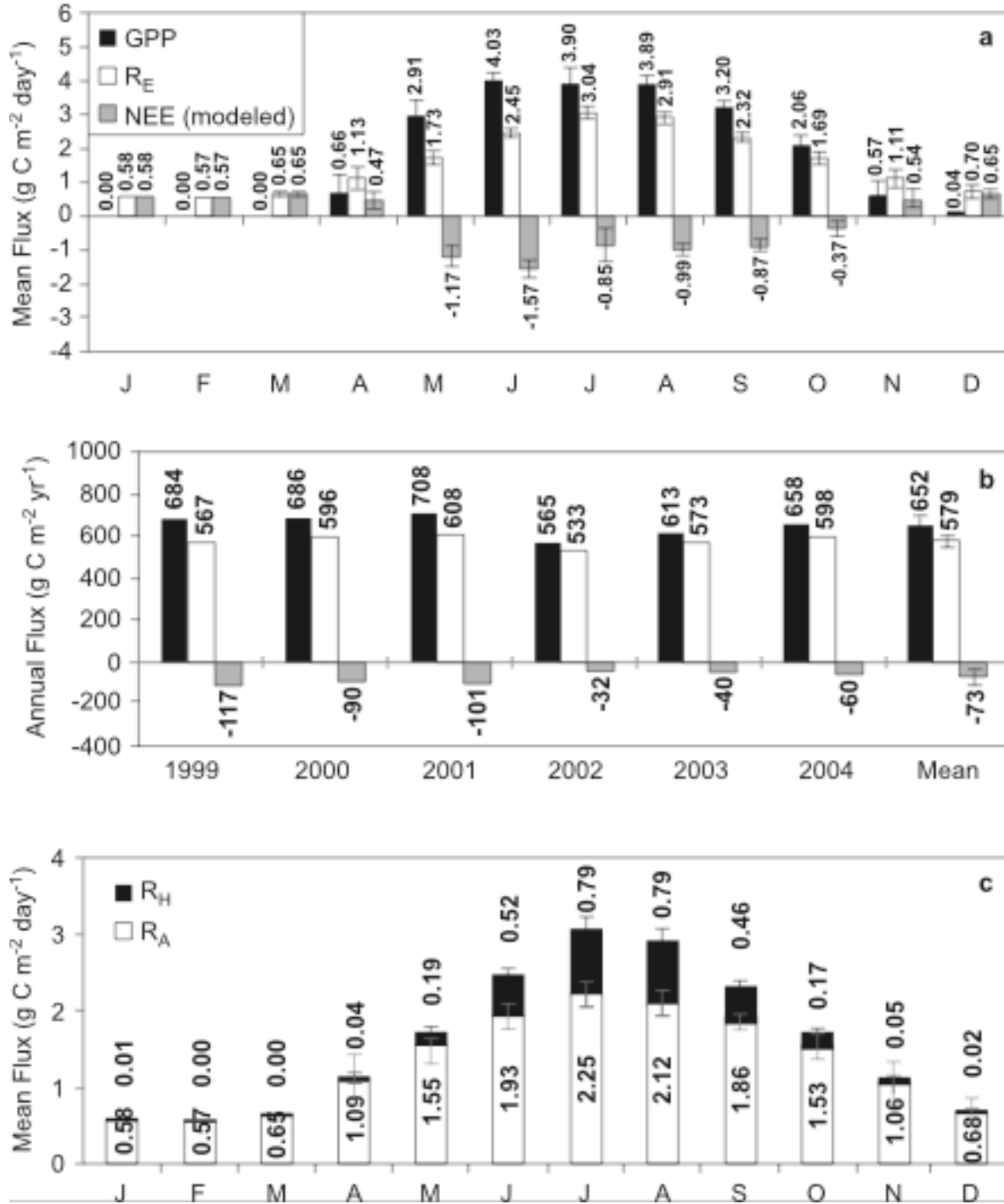




843 Figure 3



844 **Figure 4**
 845



846

