Application of a dual unscented Kalman filter for simultaneous state and parameter estimation in problems of surface-atmosphere exchange

J. H. Gove[1] and D. Y. Hollinger[1]

Received 25 March 2005; revised 1 August 2005; accepted 10 December 2005; published 22 April 2006.

[1] A dual unscented Kalman filter (UKF) was used to assimilate net CO_2 exchange (NEE) data measured over a spruce-hemlock forest at the Howland AmeriFlux site in Maine, USA, into a simple physiological model for the purpose of filling gaps in an eddy flux time series. In addition to filling gaps in the measurement record, the UKF approach provides continuous estimates of model parameters and uncertainty. The process explicitly recognizes uncertainty in the measurement data and model structure, providing approximate, effectively optimal state and parameter estimates with less subjectivity than in many previous gap-filling methods. The dual UKF is a recursive predictor-corrector estimation method whereby noisy measurement data are used to continuously update nonlinear process model predictions of the desired states, in this case net ecosystem exchange, among others. Two parallel filters are run simultaneously in the dual approach, one for state and the other for parameter estimation. The unscented transformation employs a deterministic sampling of "sigma points" from the joint density that captures the first two moments of the distribution to the second order. Nonlinear process models are applied to these sigma points to propagate the joint density within the filter framework. The UKF estimate of annual NEE in 2000 at the Howland Forest totaled -296.4 ± 2.4 g carbon m^{-2} (mean ± standard deviation) using nocturnal data when the square root of the momentum flux (u^*) exceeded 0.25 m s^{-1}. This NEE value is about 9% higher than a previous estimate where gaps were filled by physiological models fitted to monthly, seasonal, and annual data. Model estimates are sensitive to the threshold set for accepting or rejecting nocturnal flux data ("u^* threshold"), and we show that uncertainty in annual estimates is dominated by the choice of u^* threshold.

Citation: Gove, J. H., and D. Y. Hollinger (2006), Application of a dual unscented Kalman filter for simultaneous state and parameter estimation in problems of surface-atmosphere exchange, *J. Geophys. Res.*, *111*, D08S07, doi:10.1029/2005JD006021.

1. Introduction

[2] Exchanges of matter and energy between the surface and atmosphere are fundamental to the operation of the Earth's climate system, and provide critical constraints on many important biogeochemical cycles. The eddy covariance method is increasingly applied to measure surface-atmosphere exchanges of latent and sensible heat, carbon dioxide, and a variety of trace gases [*Baldocchi*, 2003]. This method is robust and reliable when applied over flat, homogeneous surfaces during times of well-developed atmospheric turbulence. Eddy covariance has many advantages including high temporal resolution (data are frequently recorded at 10 or more Hz and integrated over 30–60 minutes), and good spatial integration. Unfortunately, stable atmospheric conditions often occur at night, uncoupling surface exchange from the measurement system. Additionally, eddy covariance equipment may not operate during certain weather conditions (heavy rain, ice), and may fail, like all measurement systems, for a variety of reasons. The result is that long-term flux records have many gaps. If the goal of the research requires complete records, such as for studies of evaporation or carbon sequestration, gaps must be filled. A variety of methods have been applied to address the problem of filling flux data gaps, including linear interpolation, lookup tables, simple process models, neural networks, and Kalman filter approaches [*Falge et al.*, 2001a, 2001b; *van Wijk and Bouten*, 1999; *Jarvis et al.*, 2004], with simple process models being the predominant approach.

[3] Another important use for eddy flux data is to provide parameter estimates for biological models that predict fluxes on the basis of environmental data such as solar radiation and temperature [*Luo and Reynolds*, 1999; *Reichstein et al.*, 2003]. Those using this data-based or inverse modeling approach with flux data have derived model parameters using nonlinear least squares, maximum likelihood, and Kalman filter methods [e.g., *Hollinger et al.*, 1999; *van Wijk and Bouten*, 2002; *Jarvis et al.*, 2004; *Williams et al.*, 2005].

[1]Northeastern Research Station, U.S. Department of Agriculture, Forest Service, Durham, New Hampshire, USA.

[4] The Kalman filter is an optimal, minimum mean square error estimator for linear systems. When system dynamics are intrinsically nonlinear, the extended Kalman filter has customarily been used, although other approaches such as statistical linearization have been developed [*Gelb*, 1974, p. 203; *Maybeck*, 1982, p. 243]. In general, the extended Kalman filter performs a truncated first-order Taylor linearization on the system equations about the current state, to which the linear filter equations are applied. The extended Kalman filter has been used extensively [e.g., *Gelb*, 1974; *Jazwinski*, 1970; *Lewis*, 1986], however, it does suffer from possible divergence problems because the linearization does not always capture the correct dynamics of the underlying system [e.g., *Jazwinski*, 1969; *Fitzgerald*, 1971]. As a result, several new filtering methods have recently been introduced on the basis of the Kalman filter. Rather than seeking to linearize the nonlinear dynamics of the system, these new derivativeless methods deterministically sample the joint density of the states in such a way that the mean and covariance are preserved. The full nonlinear system dynamics are then applied to these sample points in order to propagate the density through the prediction step of the filter.

[5] Kalman filters generally take the form of a two step recursion, unless these steps have been purposefully combined [e.g., *Lewis*, 1986, p. 70]. In the prediction step, the filter uses the system process equations to perform a time update on the previous filter states before any new measurements are available at the next time period. Once the measurements become available at the new time period, they are combined in an optimal adjustment to the previous prediction of the states in the measurement update step. This two-step predictor-corrector recursion is then applied at each successive time period.

[6] A formulation to the nonlinear system equations that is adopted here assumes additive noise terms plus unknown parameters. In this case, the equations can be written

$$\mathbf{x}_k = \mathbf{f}(\mathbf{x}_{k-1}; \mathbf{w}_{k-1}) + \boldsymbol{\nu}_{k-1} \tag{1}$$

$$\mathbf{y}_k = \mathbf{h}(\mathbf{x}_k; \mathbf{w}_{k-1}) + \boldsymbol{\eta}_k \tag{2}$$

where $\mathbf{x}_k$ ($n_x \times 1$) is the state of the system at time k, $\mathbf{w}_k$ ($n_w \times 1$) are parameters, which both must be estimated simultaneously from the noisy measurements $\mathbf{y}_k$ ($n_y \times 1$). In addition, the $\boldsymbol{\nu}_k$ are the zero-mean random process noises with covariance $\mathbf{Q}_k$ that drive the nonlinear system function $\mathbf{f}$. The $\boldsymbol{\eta}_k$ are the zero-mean measurement noises with covariance $\mathbf{R}_k$ that have corrupted the measurements $\mathbf{y}_k$. The system function $\mathbf{h}$ may be either linear or nonlinear. Note that no normality assumption has been placed on the noise sequences. Kalman's original derivation [*Kalman*, 1960] did not assume normality, only that the joint density of the system states could be propagated by their first two moments. In addition, many other formulations to the system equations exist; these allow for, among other things, deterministic inputs and correlated noise sequences [*Gelb*, 1974; *Maybeck*, 1982].

[7] The Kalman filter was developed for time series where the data are autocorrelated, which makes it an attractive candidate for estimating missing flux data. In this paper we demonstrate the use of the unscented Kalman filter (UKF) for estimating annual net ecosystem CO_2 exchange (NEE) using incomplete and noisy eddy covariance data recorded above a spruce-hemlock forest in the northeastern U.S. In addition to providing an objective and unbiased method for filling gaps in the data record, we show how interpretations derived from the time-varying model parameter estimates from this filter are consistent with our understanding of the ecological processes that regulate surface-atmosphere exchanges.

2. Methods

2.1. Flux Measurements

[8] Measurements were made at the Howland Forest AmeriFlux site in central Maine, USA (45°15′N, 68°44′W, 60 m asl). This site is composed of approximately 20 m tall commercial softwood forest owned by GMO Renewable Resources, LLC. Forest stands are dominated by red spruce (*Picea rubens* Sarg.) and eastern hemlock (*Tsuga canadensis* (L.) Carr.) with lesser quantities of other conifers and hardwoods. Forest live basal area measured in plots around the main Howland research tower (45.20407°N, 68.74020°W) was about 50 m^2 ha^{-1} with live biomass about 120 t C ha^{-1} [*Hollinger et al.*, 2004]. Fluxes were measured at a height of 29 m with a system consisting of model SAT-211/3K 3-axis sonic anemometer (Applied Technologies, Inc., Longmont, Colorado, USA) and model LI-6262 fast response CO_2/H_2O infrared gas analyzer (LiCor, Inc., Lincoln, Nebraska, USA), with data recorded at 5 Hz. The flux measurement system and calculations are described in detail by *Hollinger et al.* [1999, 2004]. Deficiencies in the low- and high-frequency response of the flux system were corrected by using the Horst/Massman approach of calculating a transfer function based on stability and theoretical spectra [e.g., *Massman and Lee*, 2002] to correct for missing low-frequency contributions and a ratio of filtered to unfiltered heat fluxes to account for missing high-frequency fluctuations. Half-hourly flux values were excluded from further analysis if the wind speed was below 0.5 m s^{-1}, scalar variance was excessively high or extremely low, rain or snow was falling, for incomplete half-hour sample periods, or instrument malfunction. Data from nocturnal periods were excluded when the friction velocity, u^*, was less than a threshold of 0.25 m s^{-1}. The sign convention used is that carbon flux into the ecosystem is defined as negative.

[9] The measurements were collected at 30 minute intervals k, $k = 1,\ldots, N$ in the calendar year 2000 yielding $N =$ 17,568 half-hour measurement intervals in that year. The measurements used in the subsequent analysis are assumed to be corrupted with noise, as formally shown in the filter measurement equation (2). The measurement vector $\mathbf{y}_k$ at time k is given as

$$\begin{aligned}\mathbf{y}_k &= [\mathrm{NEE} \quad \mathrm{PPFD} \quad \mathrm{T}]' \\ &= \begin{bmatrix} y_{1,k} & y_{2,k} & y_{3,k} \end{bmatrix}'\end{aligned}$$

where NEE is net ecosystem exchange of CO_2 (μmol m^{-2} s^{-1}), PPFD is photosynthetically active photon flux density (μmol photons m^{-2} s^{-1}), and T is air temperature

in °C. Diurnal measurement periods were differentiated by defining PPFD $\leq$ 5 μmol m^{-2} s^{-1} as night. There were no missing data in the photon flux and temperature measurements. However, 42 percent of the NEE measurements were missing in the year 2000 because of either instrument failure, or exclusion for one or more of the reasons described previously.

2.2. Process Model

[10] In general, the Kalman filter prediction step is driven by models of the system dynamics. The filter can accommodate a large degree of complexity in the underlying models, e.g., ranging from stochastic differential equations in the continuous time filter, to simple models that predict the state of the system on the basis of the noisy measurements, or even simple random walk formulations when the process structure is unknown. Regardless of the complexity of the underlying models, they should provide a sensible mathematical and biological description of the system. This is especially important in the case where there are large gaps of missing data that require filling because the filtered estimates for these gaps, which rely on the system models, must make sense biologically.

[11] Our goal was to use simple process models that adequately describe the system states with a minimal number of parameters. The reason for this is twofold. First, it serves to illustrate that the Kalman filter does not require overly complex system models in order to produce very good estimates. Second, because we are illustrating a new method, a simple, tractable system model will not detract from the understanding of the filtering techniques employed. As model complexity increases, it becomes increasingly likely that no single set of parameter values will be optimal, but that many model parameter sets will fit the data more or less equally well: equifinality [*Franks et al.*, 1997].

[12] The model for NEE chosen here is the frequently used "big leaf" Michaelis-Menten model, namely,

$$\mathrm{NEE} = \frac{A \times \mathrm{PPFD}}{K + \mathrm{PPFD}} + \mathrm{R} \tag{3}$$

where A is the maximum rate of uptake (μmol CO_2 m^{-2} s^{-1}), K is the half saturation constant (μmol photons m^{-2} s^{-1}), and R is ecosystem respiration (μmol CO_2 m^{-2} s^{-1}). Ecosystem respiration could be treated as either a parameter, or a system state. The latter definition is used here. Thus we chose to model respiration with the simple *Lloyd and Taylor* [1994] model

$$\mathrm{R} = R_p \exp^{\frac{-E_0}{\mathrm{T}+273.15-T_0}} \tag{4}$$

where R_p is baseline respiration rate relative to pool size (μmol CO_2 m^{-2} s^{-1}), E_0 is akin to an activation energy in °K, and $0 \leq T_0 \leq \mathrm{T} + 273.15$.

[13] The system state vector, $\mathbf{x}_k$ at time k is composed of the same first three elements as the measurements, plus R and an estimate of the integration of NEE. The system equation for the integration of NEE (INEE) is discretized for the filter as

$$\mathrm{INEE}_k = \mathrm{INEE}_{k-1} + \mathrm{NEE}_k \Delta k \tag{5}$$

where $\Delta k = 30 \times 60 = 1800$ s is the time step. The remaining two system states (PPFD and T) are modeled as a simple random walk, in lieu of a more complicated description of the process. In general, the random walk model is

$$x_k = x_{k-1} + \nu_k \tag{6}$$

where ν_k is a zero mean random noise perturbation with positive variance. The filter provides a mechanism to include similar random noise variation in each of the process equations. Therefore the final state vector is given as

$$\begin{aligned}\mathbf{x}_k &= [\mathrm{NEE} \quad \mathrm{PPFD} \quad \mathrm{T} \quad \mathrm{INEE} \quad \mathrm{R}]' \\ &= \begin{bmatrix} x_{1,k} & x_{2,k} & x_{3,k} & x_{4,k} & x_{5,k} \end{bmatrix}'\end{aligned}$$

[14] The above system process models are defined for the growing season. However, the Howland research site is characterized by a 4-month-long winter season with temperatures well below freezing and a 1 m deep snowpack. Neither the NEE nor Lloyd and Taylor models are applicable at this time of the year. The NEE model cannot be applied since the first term, photosynthesis, is inhibited by subzero temperatures [*Teskey et al.*, 1995]. The Lloyd and Taylor model fails in winter because our formulation is based solely on air temperature. During the growing season, respiration at Howland comes from both belowground and aboveground sources in a ratio of about 60:40. At this time of year, the soil acts to low-pass air temperature variations, and air temperature provides a reasonable estimate of the system state. In the winter, however, when respiration comes mostly from the soil [*Davidson et al.*, 2006], the deep snowpack uncouples soil temperature from air temperature and so air temperature does not accurately represent the temperature of the respiring material. We therefore exclude T and adjust these two process equations for the dormant season to

$$\mathrm{R} = R_p \exp^{\frac{-E_0}{273.15-T_0}} \tag{7}$$

$$\mathrm{NEE} = \begin{cases} \mathrm{R}, & \mathrm{PPFD} < 5 \mu\mathrm{mol}\ m^{-2}\ s^{-1} (\mathrm{nighttime}) \\ A, & \mathrm{otherwise (daytime)} \end{cases} \tag{8}$$

Note that the model for NEE has different day and nighttime definitions. This was necessary because there can be warm days in the fall, and during midwinter thaws, subsequent to the switch to the dormant season models, in which photosynthesis occurs. Unfortunately, however, the modified Lloyd and Taylor model (7) only allows for respiration to occur. The maximum rate of uptake parameter, A, has therefore been used during the daytime of the dormant season as a simple time-varying parameter to accommodate this potential late-season photosynthetic activity. Note that it also accommodates daytime respiration for the majority of days where there is no photosynthesis occurring.

[15] Determination of the initiation and completion of the growing season was made using a deterministic switch based

on recorded soil temperature. When the daily soil temperature reaches 0°C, in both spring and fall, the switch is made to the alternate process model formulations. For the year 2000, this happened at Julian days 84 and 328. In our current model, the switch occurs only once each in spring and fall; alternative possible formulations will be discussed later.

[16] There are five system parameters in this model formulation: A, K, R_p, E_0 and T_0. However, because of a high degree of correlation between E_0 and T_0, we fixed T_0 = 261.2°K, on the basis of a least squares fit of the June 2000 data to model (4). This left a total of four unknown parameters, which are collected into the parameter vector $\mathbf{w}_k$, namely,

$$\begin{aligned}\mathbf{w}_k &= \begin{bmatrix} A & K & E_0 & R_p \end{bmatrix}' \\ &= \begin{bmatrix} w_{1,k} & w_{2,k} & w_{3,k} & w_{4,k} \end{bmatrix}'\end{aligned}$$

[17] The vectors $\mathbf{y}_k$, $\mathbf{x}_k$ and $\mathbf{w}_k$ make up the system noisy measurements, unknown states and unknown parameters, respectively. It should be noted explicitly that the parameters are time-varying in this formulation.

2.3. Scaled Unscented Transformation

[18] The unscented transformation (UT) was first introduced by *Uhlmann* [1995]. The main idea behind the unscented transformation, alluded to earlier, is to completely capture the first two moments of the joint density in both of the filter steps with a deterministic sampling of "sigma points" from that density, and subsequently apply the nonlinear dynamics to these sampled points. It was founded on the simple observation that it should be easier to approximate a probability distribution than an arbitrary nonlinear function [*Uhlmann*, 1995; *Julier and Uhlmann*, 2004]. It should be noted that the unscented transformation is quite general in that it can be applied to any problem that requires nonlinear transformations of probability distributions, not just filtering problems. However, the main motivation for its development was to address the shortcomings of the EKF linearization approach [*Julier and Uhlmann*, 2004]. These shortcomings are manifested in two primary ways. First, derivation of the Jacobian matrices used in the extended Kalman filter, if they exist, can be difficult, and programming these often large calculations is error prone. Secondly, it has been known for some time that the extended Kalman filter can indeed diverge because the truncation of higher-order terms in the Taylor linearization of the dynamics can yield poor mean and covariance estimates of the system state [e.g., *Jazwinski*, 1969; *Fitzgerald*, 1971]. Some corrections and filter "tuning" can be made to address this second issue, such as Monte Carlo analysis and the addition of stabilizing noise. However, ameliorations must be undertaken on a case-by-case basis, as such problems tend to be problem-specific. In addition, these methods may actually result in the unintended consequence of inflating the variance in some portions of the state space, while providing no correction for the bias [*Julier and Uhlmann*, 2004]. Other specialized methods also exist, but are complex and may be applicable only to specific classes of problems, e.g., where the error distributions are Gaussian.

[19] *Julier and Uhlmann* [1997, 2004] present the derivation for the UT along with an example inherent to target tracking, illustrating that converting from polar to Cartesian coordinates using linearization is both biased and inconsistent. For comparison, the UT shows no bias and correctly captures the state covariance. The basic UT has been developed further, to incorporate higher-order moments of the distribution, among other things [e.g., *Tenne and Singh*, 2003; *Julier and Uhlmann*, 2004]. Extending the sigma point set to incorporate higher-order moments requires a larger set of points and increases the computational burden in the filter. Pragmatically, if only the first two moments of the distribution are estimated, as in the Kalman filter, using an extended sigma point set incorporating higher-order moments, such as skewness, could actually be detrimental, for example, if the skew of the sigma point set does not align with that of the true distribution. However, the scaled sigma point set can incorporate the effect of higher-order moments through a scaling constant that shrinks or expands the set as desired [*Julier and Uhlmann*, 2004]. Choosing a small scaling factor concentrates the points about the mean, and thus minimizes the effect of higher-order moments, and vice versa. An additional weighting scheme allows incorporation of higher-order moments, such as kurtosis, when known.

[20] The UT and UKF have been applied to nonlinear estimation problems in a variety of fields such as global positioning systems [*van der Merwe and Wan*, 2004], spacecraft attitude estimation [*Crassidis and Markley*, 2003], ballistic missile tracking [*Saulson and Chang*, 2004] and object tracking in image analysis systems [*Chen et al.*, 2002]. In this paper, we use a modification of the scaled unscented transformation, originally developed by *Julier and Uhlmann* [2002]; the modification is due to *van der Merwe* [2004], who presented an alternative rearrangement that allows the sigma point selection and scaling to occur in one step, thereby reducing calculations. As before, let n_x be the dimension of the state space, $\mathbf{x}$, with mean $\bar{\mathbf{x}}$ and covariance $\mathbf{P}_\mathbf{x}$. Select $2n_x + 1$ sigma points by letting $\lambda = \alpha^2(n_x + \kappa) - n_x$ and drawing

$$\begin{aligned}
\mathcal{X}_0 &= \bar{\mathbf{x}} \\
\mathcal{X}_i &= \bar{\mathbf{x}} + \left(\sqrt{(n_x+\lambda)\mathbf{P}_\mathbf{x}}\right)_i && i = 1, \ldots, n_x \\
\mathcal{X}_i &= \bar{\mathbf{x}} - \left(\sqrt{(n_x+\lambda)\mathbf{P}_\mathbf{x}}\right)_{i-n_x} && i = n_x+1, \ldots, 2n_x \\
w_0^{(m)} &= \frac{\lambda}{n_x+\lambda} && i = 0 \\
w_0^{(c)} &= \frac{\lambda}{n_x+\lambda} + \left(1-\alpha^2+\beta\right) && i = 0 \\
w_i^{(m)} &= w_i^{(c)} = \frac{1}{2(n_x+\lambda)} && i = 1, \ldots, 2n_x
\end{aligned} \tag{9}$$

The sigma point set $\mathcal{S} = \{\mathcal{X}_i, w_i^{(j)}; i = 0, \ldots, 2n_x, j \in (m, c)\}$ is composed of the sigma points $\mathcal{X}_i$ and their respective mean (m) and covariance (c) weights $w_i^{(j)}$. The weights can be positive or negative and must sum to one [*Julier and Uhlmann*, 2004]. The parameters $0 \leq \alpha \leq 1$ and $\beta \geq 0$ control the spread of the sigma points and weighting for higher-order moments; parameter $\kappa \geq 0$ is not critical and is often set to zero [*van der Merwe*, 2004, p. 56]. For α, the

smaller the value, the smaller the sigma-point spread and the less likely to pick up anomalous effects in the distribution. For Gaussian distributions, $\beta = 2$ is optimal [*Julier and Uhlmann*, 2002, 2004].

[21] Given an arbitrary nonlinear transformation on the state $\mathbf{g}$, the scaled unscented transformation consists of the following four steps [*van der Merwe*, 2004, p. 56]:

[22] 1. Choose the parameters α, β and κ as described.

[23] 2. Determine the set $\mathcal{S}$ of $2n_x + 1$ sigma points and associated weights from (9).

[24] 3. Apply the nonlinear transformation $\mathbf{g}$ to each sigma point, namely,

$$\mathcal{Y}_i = \mathbf{g}(\mathcal{X}_i), \qquad i = 0, \ldots, 2n_x$$

[25] 4. Calculate the mean, covariance and cross covariance, respectively, of the transformed sigma points as

$$\bar{\mathbf{y}} = \sum_{i=0}^{2n_x} w_i^{(m)} \mathcal{Y}_i$$

$$\mathbf{P}_{\mathbf{y}} = \sum_{i=0}^{2n_x} w_i^{(c)} (\mathcal{Y}_i - \bar{\mathbf{y}})(\mathcal{Y}_i - \bar{\mathbf{y}})'$$

$$\mathbf{P}_{\mathbf{xy}} = \sum_{i=0}^{2n_x} w_i^{(c)} (\mathcal{X}_i - \bar{\mathbf{x}})(\mathcal{Y}_i - \bar{\mathbf{y}})'$$

[26] A simple example will provide some intuition into the UT and the effect that the choice of parameter values can have in drawing the sigma points. Let the untransformed state vector for this example be composed of PPFD and T. The mean and covariance were calculated from June 2000 measurements where PPFD $\geq$ 1500 μmol photons m^{-2} s^{-1}, yielding

$$\bar{\mathbf{x}} = [1753.54 \quad 16.5]$$

$$\mathbf{P}_{\mathbf{x}} = \begin{bmatrix} 18766 & -62.3 \\ -62.3 & 16.6 \end{bmatrix}$$

The transformation $\mathbf{g}$ is from a maximum likelihood fit of models (3) and (4) to the June 2000 data

$$\mathrm{R} = 35 \exp^{\frac{-46.4}{\mathrm{T}+273.15-262.2}}$$

$$\mathrm{NEE} = \frac{-20 \times \mathrm{PPFD}}{524 + \mathrm{PPFD}} + \mathrm{R}$$

Thus the states PPFD and T are mapped into the states NEE and R through $\mathbf{g}$. The results of applying the unscented transformation on this example with $\beta = 2$ are shown in Figure 1. In Figure 1a, $\alpha = 1$, $\kappa = 0.15$ and it will be noted that the sigma points align with the 66 percent confidence ellipse. In addition, note that the weight $w_0^{(m)}$ is small and positive. In Figure 1b, $\alpha = 0.4$, $\kappa = 0$; the resulting set of points is much tighter around the mean, illustrating the effect of reducing α. As mentioned above, adjusting the spread of the sigma points allows one to include the effect of higher-order moments if desired. Finally, notice that the weight $w_0^{(m)}$ is large and negative. The bottom panel shows the effect of applying the nonlinear mapping $\mathbf{g}$. Note that the sigma points have transformed accordingly to capture the new mean and covariance of the states NEE and R. In addition, this example serves to illustrate the notable difference that can occur in the first two moments with even a simple nonlinear transformation.

[27] This simple example also illustrates that while the sigma points capture the essential information of the joint distribution, they are not a probability density function [*Julier and Uhlmann*, 2004]. Thus the weights can be positive or negative, providing they meet the constraint of summing to unity. As mentioned by *Julier and Uhlmann* [2004], there are a number of important properties of the UT. First, it is algorithmic, and can be easily programmed in a "black-box" sense regardless of the complexity of the nonlinear transformations, with computational cost on the same order of magnitude as the EKF. Relatedly, because the sigma points can straddle discontinuities, $\mathbf{g}$ need not be a continuous function set. Finally, the sigma points from the UT, including the scaled UT, calculates the projected mean and covariance to the second order (i.e., in the Taylor series expansion) and therefore implicitly includes a second-order "bias correction" that is not in the EKF.

2.4. Dual Unscented Kalman Filter

[28] A number of methods have been developed to handle uncertainty in parameter estimates in addition to state estimation. For example, the Schmidt-Kalman filter [*Jazwinski*, 1970, p. 285] accounts for the effect of uncertain parameters on the estimate of the state without explicitly estimating the parameters themselves. In addition, the state-dependent approach [*Young*, 2000] is a general method used to estimate unknown states and parameters considered to be statistically dependent, through a multistep autoregressive filtering and modeling process. Two other general methods, have been developed to simultaneously estimate the unknown states and parameters from the noisy measurements using the Kalman filter, namely, joint and dual estimation [*van der Merwe*, 2004; *Wan and van der Merwe*, 2001]. While the former methods are noteworthy, we concentrate on the latter two methods, specifically the dual filtering paradigm.

[29] The joint Kalman filter takes the approach of augmenting the state vector with the parameters to form a joint state vector $\tilde{\mathbf{x}}_k = [\mathbf{x}_k', \mathbf{w}_k']'$ [*Gelb*, 1974, p. 348; *Jazwinski*, 1970, p. 282]. This joint state is now estimated through a single Kalman filter recursion. The alternative is to run two parallel filters, one on the state and the other on the parameters; this is known as the dual approach. In the dual setting, the parameters are treated as known within the state filter at any given time, k, while the states are treated as known in the parallel parameter filter. Both the joint and dual approaches can be run within either the extended or unscented Kalman filter frameworks [*van der Merwe*, 2004; *Wan and Nelson*, 2001; *Wan and van der Merwe*, 2001]. The main difference between the two approaches, aside from the number of filters required, is that the joint filter explicitly allows for cross parameter and state dependencies. For example, let the joint covariance for the augmented state be

$$\mathbf{P}_k = \begin{bmatrix} \mathbf{P}_{\mathbf{x}_k} & \mathbf{P}_{\mathbf{x}_k \mathbf{w}_k} \\ \mathbf{P}_{\mathbf{w}_k \mathbf{x}_k} & \mathbf{P}_{\mathbf{w}_k} \end{bmatrix}$$

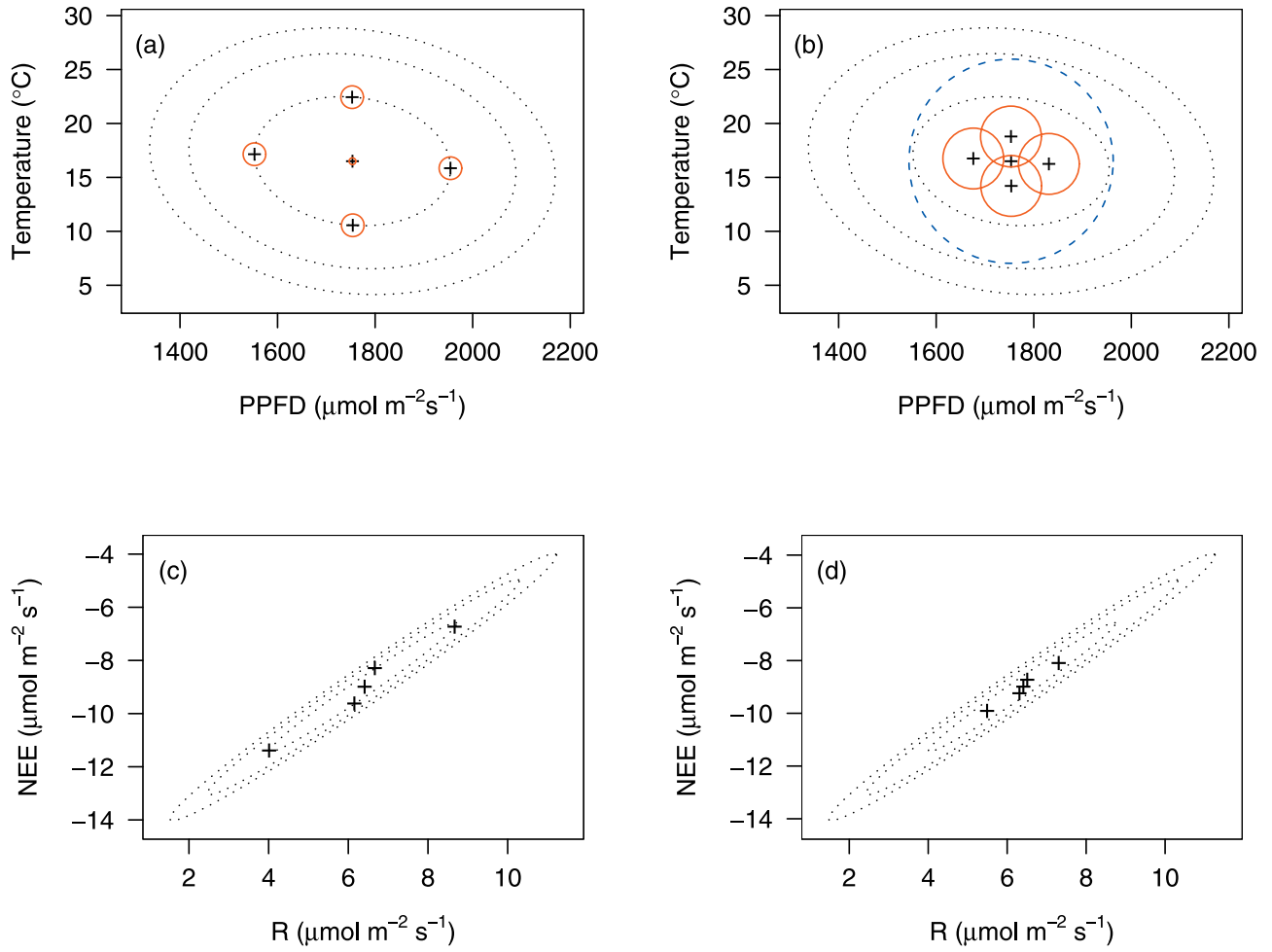

Figure 1. Bivariate sigma points (a and b) before transformation and (c and d) after transformation. The left- and right-hand plots show the effect that differing values for parameters α and κ have on the spread and weighting of the sigma points. In Figures 1a and 1b, the sigma points (pluses) are weighted proportional to $\alpha w_i^{(m)}$ as shown by the circles; solid circles imply positive weights, and dashed circles imply negative weights. The confidence ellipses are for a bivariate normal distribution at the 66, 90 and 95 percentiles. Figures 1c and 1d illustrate the transformed sigma points. Details on parameter settings are given in the text.

In the dual filtering approach the cross covariances are not explicitly estimated, so that it effectively assumes that the cross covariances $\mathbf{P}_{\mathbf{x}_k\mathbf{w}_k} = \mathbf{P}_{\mathbf{w}_k\mathbf{x}_k} = 0$. It could be argued, therefore, that if correlation is suspected between states and parameters, the joint approach would be preferred [*van der Merwe*, 2004, p. 104]. However, experiments performed by *van der Merwe* [2004] show little difference between the two approaches. The reason may be due to the fact that switching parameters and states between the dual filters, coupled with using the exact same measurements in both filters, acts as a type of constraint on the filters and implicitly develops the cross covariance terms, as our results will demonstrate.

[30] The joint and dual unscented Kalman filters have been largely developed to address problems in estimation with neural networks such as in the control of unmanned aerial vehicles [*van der Merwe*, 2004; *Wan and van der Merwe*, 2001]. In neural network problems, unknown parameters are often referred to as weights, and the estimation of the weights is often termed identification. *Sitz et al.* [2002] have also applied the joint UKF to a continuous time estimation scheme for nonlinear systems. They demonstrate the efficacy of the technique using the classical Lotka-Volterra and Lorenz systems. In addition, *VanDyke et al.* [2004] have applied the dual UKF approach to spacecraft attitude dynamics estimation problems. In the following, the dual UKF is presented in the form of the state and parameter filters. While the filtering equations themselves are similar, the differences in the system models require separate presentations.

2.4.1. State Filter

[31] *Julier and Uhlmann* [2004], *van der Merwe* [2004], and *Wan and van der Merwe* [2001] present a general system framework allowing for correlated system or process noise [see also *Gelb*, 1974, p. 78]. However, it is often the case that the noise structure is assumed to be additive,

which is the assumption here. The dual UKF system model for the state with additive stochastic noise is

$$\mathbf{x}_k = \mathbf{f}(\mathbf{x}_{k-1}; \hat{\mathbf{w}}_{k-1}) + \boldsymbol{\nu}_{k-1} \tag{10}$$

$$\mathbf{y}_k = \mathbf{h}(\mathbf{x}_k; \hat{\mathbf{w}}_{k-1}) + \boldsymbol{\eta}_k \tag{11}$$

with $\boldsymbol{\nu}_{k-1}$ and $\boldsymbol{\eta}_k$ noise sequences as defined in (1) and (2). The $\hat{\mathbf{w}}_{k-1}$ are estimates of the parameters from the previous time step in the parameter filter and are treated as constants in the state filter. The state filter recursions are given in the following steps [*Julier and Uhlmann*, 2004; *van der Merwe*, 2004; *Wan and van der Merwe*, 2001]:

[32] 1. Initialize the filter:

$$\hat{\mathbf{x}}_0 = \mathrm{E}[\mathbf{x}_0], \qquad \mathbf{P}_{\mathbf{x}_0} = \mathrm{E}\left[(\mathbf{x}_0 - \hat{\mathbf{x}}_0)(\mathbf{x}_0 - \hat{\mathbf{x}}_0)'\right]$$

[33] 2. Calculate the sigma point weights as in (9).

[34] 3. Repeat the following for $k = 1,\ldots, N$: (1) Calculate the sigma points as in (9):

$$\mathcal{X}_{k-1} = \left[\hat{\mathbf{x}}_{k-1},\ \hat{\mathbf{x}}_{k-1} + \sqrt{(n_x + \lambda)\mathbf{P}_{\mathbf{x}_{k-1}}},\ \hat{\mathbf{x}}_{k-1} - \sqrt{(n_x + \lambda)\mathbf{P}_{\mathbf{x}_{k-1}}}\ \right] \tag{12}$$

(2) filter prediction equations:

$$\mathcal{X}^*_{k|k-1} = \mathbf{f}(\mathcal{X}_{k-1}; \hat{\mathbf{w}}_{k-1}) \tag{13}$$

$$\hat{\mathbf{x}}_{k|k-1} = \sum_{i=0}^{2n_x} w_i^{(m)} \mathcal{X}^*_{i,k|k-1} \tag{14}$$

$$\mathbf{P}_{\mathbf{x}_{k|k-1}} = \sum_{i=0}^{2n_x} w_i^{(c)} \left(\mathcal{X}^*_{i,k|k-1} - \hat{\mathbf{x}}_{k|k-1}\right)\left(\mathcal{X}^*_{i,k|k-1} - \hat{\mathbf{x}}_{k|k-1}\right)' + \mathbf{Q}_k \tag{15}$$

$$\mathcal{X}_{k|k-1} = \left[\hat{\mathbf{x}}_{k|k-1},\ \hat{\mathbf{x}}_{k|k-1} + \sqrt{(n_x + \lambda)\mathbf{P}_{\mathbf{x}_{k|k-1}}},\ \hat{\mathbf{x}}_{k|k-1} - \sqrt{(n_x + \lambda)\mathbf{P}_{\mathbf{x}_{k|k-1}}}\ \right] \tag{16}$$

$$\mathcal{Y}_{k|k-1} = \mathbf{h}(\mathcal{X}_{k|k-1}; \hat{\mathbf{w}}_{k-1}) \tag{17}$$

$$\hat{\mathbf{y}}_{k|k-1} = \sum_{i=0}^{2n_x} w_i^{(m)} \mathcal{Y}_{i,k|k-1} \tag{18}$$

and (3) filter measurement equations:

$$\mathbf{P}_{\mathbf{y}_k} = \sum_{i=0}^{2n_x} w_i^{(c)} \left(\mathcal{Y}_{i,k|k-1} - \hat{\mathbf{y}}_{k|k-1}\right)\left(\mathcal{Y}_{i,k|k-1} - \hat{\mathbf{y}}_{k|k-1}\right)' + \mathbf{R}_k \tag{19}$$

$$\mathbf{P}_{\mathbf{x}_k\mathbf{y}_k} = \sum_{i=0}^{2n_x} w_i^{(c)} \left(\mathcal{X}_{i,k|k-1} - \hat{\mathbf{x}}_{k|k-1}\right)\left(\mathcal{Y}_{i,k|k-1} - \hat{\mathbf{y}}_{k|k-1}\right)' \tag{20}$$

$$\mathbf{K}_k = \mathbf{P}_{\mathbf{x}_k\mathbf{y}_k}\mathbf{P}_{\mathbf{y}_k}^{-1} \tag{21}$$

$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k\left(\mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1}\right) \tag{22}$$

$$\mathbf{P}_{\mathbf{x}_k} = \mathbf{P}_{\mathbf{x}_{k|k-1}} - \mathbf{K}_k\mathbf{P}_{\mathbf{y}_k}\mathbf{K}_k' \tag{23}$$

where $\mathbf{R}_k$ and $\mathbf{Q}_k$ are the system and process noise covariances, respectively. Note that we redraw new sigma points in (16) to incorporate the new information in the prediction density; *van der Merwe* [2004, p. 109] discusses other possible strategies.

[35] In our model formulation, the process models, $\mathbf{f}$, are given by (3)–(8). The measurement model is an identity mapping with

$$\mathbf{h} = \mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

Since this form of $\mathbf{h}$ is linear, we can substitute the linear Kalman filter measurement update recursions [e.g., *Gelb*, 1974, p. 110] for the above UKF measurement recursions, with the state update, (22), remaining the same, namely,

$$\mathbf{K}_k = \mathbf{P}_{\mathbf{x}_{k|k-1}}\mathbf{H}'\left(\mathbf{H}\mathbf{P}_{\mathbf{x}_{k|k-1}}\mathbf{H}' + \mathbf{R}_k\right)^{-1} \tag{24}$$

$$\mathbf{P}_{\mathbf{x}_k} = (\mathbf{I} - \mathbf{K}_k\mathbf{H})\mathbf{P}_{\mathbf{x}_{k|k-1}} \tag{25}$$

where $\mathbf{I}$ is the identity matrix of dimension $(n_x \times n_x)$. There are several other possible formulations and options available for the UKF as presented in the literature cited. Incidentally, the substitution of the linear form of the Kalman recursions for the update step exemplifies the flexibility of the UKF; alternatively, had our process dynamics been linear and our observation dynamics nonlinear, the UKF would have only been used in the update step [*Julier and Uhlmann*, 2004].

2.4.2. Parameter Filter

[36] It is well established [e.g., *Bell and Cathy*, 1993; *Wan and Nelson*, 2001; *Plummer*, 1995] that the EKF can be used in parameter estimation and approximates a recursive modification to Newton's method. The UKF, like the EKF, can be used for parameter estimation for both clean and noisy time series; in the latter case joint or dual estimation strategies are used. *van der Merwe* [2004] has shown that the UKF can be interpreted as an iterative stochastic Gauss-Newton method, which builds up an approximation to the inverse of the Fisher information matrix used in the Fisher scoring step (see *van der Merwe* [2004, chap. 4] for details). In addition, the UKF has been shown to outperform the

EKF in dual estimation in terms of smaller and more consistent covariance estimates, lessening the likelihood for filter divergence.

[37] The system equations for the parameter filter, which is coupled with (10) in the dual setting are

$$\mathbf{w}_k = \mathbf{w}_{k-1} + \boldsymbol{\upsilon}_{k-1} \tag{26}$$

$$\mathbf{y}_k = \mathbf{h}(\mathbf{f}(\hat{\mathbf{x}}_{k-1}; \mathbf{w}_k); \mathbf{w}_k) + \boldsymbol{\epsilon}_k \tag{27}$$

Here the $\boldsymbol{\upsilon}_k$ and $\boldsymbol{\epsilon}_k$ are the process and measurement noises, with covariances $\mathbf{Q}_{\mathbf{w}_k}$ and $\mathbf{R}_{\mathbf{w}_k}$, respectively.

[38] The unscented parameter filter recursion steps are as follows [*van der Merwe*, 2004; *Wan and van der Merwe*, 2001]:

[39] 1. Initialize the filter:

$$\hat{\mathbf{w}}_0 = \mathrm{E}[\mathbf{w}], \qquad \mathbf{P}_{\mathbf{w}_0} = \mathrm{E}\left[(\mathbf{w} - \hat{\mathbf{w}}_0)(\mathbf{w} - \hat{\mathbf{w}}_0)'\right]$$

[40] 2. Calculate the sigma point weights as in (9).

[41] 3. Repeat the following for $k = 1, \ldots, N$: (1) Filter prediction equations:

$$\hat{\mathbf{w}}_{k|k-1} = \hat{\mathbf{w}}_{k-1} \tag{28}$$

$$\mathbf{P}_{\mathbf{w}_{k|k-1}} = \mathbf{P}_{\mathbf{w}_{k-1}} + \mathbf{Q}_{\mathbf{w}_{k-1}} \tag{29}$$

(2) calculate the sigma points as in (9):

$$\mathcal{W}_{k|k-1} = \left[\hat{\mathbf{w}}_{k|k-1},\ \hat{\mathbf{w}}_{k|k-1} + \sqrt{(n_x + \lambda)\mathbf{P}_{\mathbf{w}_{k|k-1}}},\ \hat{\mathbf{w}}_{k|k-1} - \sqrt{(n_x + \lambda)\mathbf{P}_{\mathbf{w}_{k|k-1}}}\ \right] \tag{30}$$

and (3) filter measurement equations:

$$\mathcal{Y}_{k|k-1} = \mathbf{h}\left(\mathbf{f}\left(\mathcal{W}_{k|k-1}; \hat{\mathbf{x}}_{k-1}\right)\right) \tag{31}$$

$$\hat{\mathbf{y}}_{k|k-1} = \sum_{i=0}^{2n_x} w_i^{(m)} \mathcal{Y}_{i,k|k-1} \tag{32}$$

$$\mathbf{P}_{\mathbf{y}_k} = \sum_{i=0}^{2n_x} w_i^{(c)} \left(\mathcal{Y}_{i,k|k-1} - \hat{\mathbf{y}}_{k|k-1}\right)\left(\mathcal{Y}_{i,k|k-1} - \hat{\mathbf{y}}_{k|k-1}\right)' + \mathbf{R}_{\mathbf{w}_k} \tag{33}$$

$$\mathbf{P}_{\mathbf{w}_k\mathbf{y}_k} = \sum_{i=0}^{2n_x} w_i^{(c)} \left(\mathcal{W}_{i,k|k-1} - \hat{\mathbf{w}}_{k|k-1}\right)\left(\mathcal{Y}_{i,k|k-1} - \hat{\mathbf{y}}_{k|k-1}\right)' \tag{34}$$

$$\mathbf{K}_k = \mathbf{P}_{\mathbf{w}_k\mathbf{y}_k}\mathbf{P}_{\mathbf{y}_k}^{-1} \tag{35}$$

$$\hat{\mathbf{w}}_k = \hat{\mathbf{w}}_{k|k-1} + \mathbf{K}_k\left(\mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1}\right) \tag{36}$$

$$\mathbf{P}_{\mathbf{w}_k} = \mathbf{P}_{\mathbf{w}_{k|k-1}} - \mathbf{K}_k\mathbf{P}_{\mathbf{y}_k}\mathbf{K}_k' \tag{37}$$

where $\mathbf{R}_{\mathbf{w}_k}$ is set equal to $\mathbf{R}_k$ in the state filter since the same measurements are used in each filter. Note that because of the nonlinear form of (31), it is not possible to use the linear Kalman filter recursions as replacements for the UKF update step in the parameter filter, as we did in the state filter.

[42] In the above, the diagonal matrix $\mathbf{Q}_{\mathbf{w}_k}$ can be set by one of several different methods [*van der Merwe*, 2004; *Nelson*, 2000]. We used a method analogous to recursive least squares

$$\mathbf{Q}_{\mathbf{w}_k} = \mathrm{diag}\left(\left(\tilde{\lambda}^{-1} - 1\right)\mathbf{P}_{\mathbf{w}_k}\right) \tag{38}$$

with $0 \leq \tilde{\lambda} \leq 1$ known as the forgetting factor and where "diag(·)" means diagonalize the resulting matrix by setting off-diagonals to zero. This scheme adjusts $\mathbf{P}_{\mathbf{w}_{k|k-1}}$ such that it is slightly larger than $\mathbf{P}_{\mathbf{w}_{k-1}}$, which has the effect of discarding older data more quickly. Smaller values of $\tilde{\lambda}$ increase this effect; however, we found that values of $\tilde{\lambda}$ close to one (0.9975) produced the most stable filter results.

2.5. Missing Observations

[43] Missing observations cause gaps in the measurement record that can sometimes be extensive. Filling gaps is a necessary component of the estimation process, whether for periodic (daily, weekly, etc.) or integrated estimates. The yearly integration of NEE is a key component to our flux analysis; therefore minimum mean square estimates of the missing values are a crucial component to obtaining this final estimate.

[44] In our measurement record, the series for PPFD and T were complete, but approximately 42 percent of the NEE observations were missing. In univariate time series, accommodating missing values is simple. The innovations, or prediction errors, are defined as the sequence

$$\mathbf{v}_k = \left(\mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1}\right) \qquad k = 1, \ldots, N \tag{39}$$

Therefore, if a measurement is missing at time k, the innovation is also missing. In the univariate case, this is handled by setting the Kalman gain to zero ($\mathbf{K}_k = 0$), or, more formally by setting $\mathbf{R}$ to infinity. The result is that $\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_{k|k-1}$ and similarly, $\mathbf{P}_k = \mathbf{P}_{k|k-1}$. In other words, the update step could simply be skipped with this assignment [*Kitagawa*, 1999]. The motivation becomes clear when regarding the state update (22), which may be rewritten as $\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k\mathbf{v}_k$. Setting $\mathbf{K}_k = 0$ in the univariate case cancels the effect of the missing innovation. Similar observations may be made for the covariance update since both the covariance ($\mathbf{P}_{\mathbf{y}_k}$) and cross covariance ($\mathbf{P}_{\mathbf{x}_k\mathbf{y}_k}$) terms will be missing. Incidentally, this alternate form of the state update also serves to illustrate the predictor-corrector structure of the Kalman recursions; applying the optimal weighting matrix, $\mathbf{K}_k$, to the innovation generates a correction to be added to the state prediction $\hat{\mathbf{x}}_{k|k-1}$, obtaining $\hat{\mathbf{x}}_k$ [*Maybeck*, 1979, p. 218].

[45] In the multivariate case, where there are gaps for some variables and not for others, there is still valuable information in the nonmissing components of the measurement vector at time k so the update step should not be skipped. Regarding the definition of our measurement

vector $\mathbf{y}_k$, NEE appears as the first measurement variable $y_{1,k}$. Setting the first row of the measurement matrix $\mathbf{H}$ and the corresponding component of the innovation vector to zero will account for the missing observation and allow the information from nonmissing components of $\mathbf{y}_k$ to be utilized in the KF update [*Shumway and Stoffer*, 2000, p. 330]. This same procedure is used in both of the dual filters and for the linear or unscented update step.

[46] The forward pass of the Kalman filter as given in the dual UKF recursions provides minimum mean square estimates for the missing observations conditional on all of the past data. In addition, a smoother can be run, which provides estimates based on all of the data, past and future. If the intent is to use periods of the individual time series for, e.g., weekly estimates, then the smoother will provide estimates with lower variance. However, because the backward recursion of the smoother is initiated with the last state and covariance updates (i.e., $\hat{\mathbf{x}}_N$ and $\mathbf{P}_{\mathbf{x}_N}$), the smoother will not provide different estimates for the final yearly integrated estimate for NEE.

[47] The fixed interval smoother used here was due originally to *Fraser and Potter* [1969], and *Wan and van der Merwe* [2001] have applied it to the UKF state filter. The smoother estimate is developed by running a UKF backward in time, yielding state prediction and covariance ($\hat{\mathbf{x}}_{k|k+1}$, $\mathbf{P}_{\mathbf{x}_{k|k+1}}$) before including the observation at time k in the backward update step. The smoothed estimates are then given by

$$\mathbf{P}^{-1}_{\mathbf{x}_{k|N}} = \mathbf{P}^{-1}_{\mathbf{x}_k} + \mathbf{P}^{-1}_{\mathbf{x}_{k|k+1}} \tag{40}$$

$$\hat{\mathbf{x}}_{k|N} = \mathbf{P}_{\mathbf{x}_{k|N}} \left(\mathbf{P}^{-1}_{\mathbf{x}_{k|k+1}} \hat{\mathbf{x}}_{k|k+1} + \mathbf{P}^{-1}_{\mathbf{x}_k} \hat{\mathbf{x}}_k \right) \tag{41}$$

Equation (40) leads to the conclusion that $\mathbf{P}_{\mathbf{x}_{k|N}} \leq \mathbf{P}_{\mathbf{x}_k}$. Therefore the smoothed estimate will be at least as precise as the filtered estimate, and generally will be better. The exception is the terminal time N when they are identical [*Maybeck*, 1982, p. 7]. We use this approach to develop dual unscented Kalman smoother (UKS) estimates for the states in gap filling.

2.6. Filter Noise Covariances and Initial Conditions

[48] In the Kalman filter, the measurement and process noise covariances are assumed known without error. Determination of the noise covariance matrices and initial conditions for the Kalman filter has been the subject of a large body of research in both engineering and statistics. Unfortunately, much of this research has been done on linear Kalman filters, often under the assumption of Gaussian noise, conditional densities, and innovations. For example, maximum likelihood is commonly used in an off-line setting to estimate the hyperparameters in the system and process noise covariances $\mathbf{R}$ and $\mathbf{Q}$ [*Harvey*, 1989, p. 140]. In addition, the initial state of the filter is normally assumed known (e.g., see the initialization step for both filters). If unknown, a diffuse prior often is used by setting the initial covariance to $\psi\mathbf{I}$, where ψ is chosen to be suitably large such that $(\psi\mathbf{I})^{-1} \approx \mathbf{0}$ [p. 121 *Harvey*, 1989]. Other methods, such as the expectation-maximization (EM) algorithm have been developed to jointly estimate the initial states and the filter covariances [*Shumway and Stoffer*, 1982]. However, each of these maximum likelihood-based methods involves Gaussian assumptions, which *Hollinger and Richardson* [2005] have shown do not hold in the case of our flux measurements. In addition, it is doubtful that such methods directly adapt to nonlinear estimation problems, especially in the case of running simultaneous dual filters. A further complication arises when the covariance matrices are time-dependent, increasing the number of parameters that must be estimated.

[49] Acknowledging these difficulties, we have used a diffuse prior approach when initial conditions are unknown or inestimable, while other quantities have been estimated either through ML or Monte Carlo methods. Because of the seasonality in environmental conditions inherent at this latitude, it seems reasonable that some measurement and state variances will differ between the dormant and growing seasons, and this has been allowed for in our formulation. In all cases, the noises are assumed uncorrelated and diag[·] means construct a diagonal matrix.

[50] The filters are started with $k = 1$ beginning 1 January 2000 using the dormant season model (7) and (8). The initial conditions for the state filter use the diffuse prior approach with $\mathbf{P}_{\mathbf{x}_0} = 100\mathbf{I}$ and $\hat{\mathbf{x}}_0 = 0$. This latter value can be loosely interpreted as a reasonable value for the state on a winter night, but with low confidence. The parameter filter was initialized with values developed from process models fitted to the April and May 2000 data, since fits to dormant season data are not estimable for these models. In model (3), maximum likelihood estimates were used on the basis of the methods developed by *Hollinger and Richardson* [2005], while the associated variance estimates were determined from Monte Carlo runs as explained by these authors. The initial state and covariances are

$$\hat{\mathbf{w}}_0 = [0.5, \quad 386.9, \quad -25.0, \quad 4.9]$$
$$\mathbf{P}_{\mathbf{w}_0} = \mathrm{diag}[0.574, \quad 4662.6, \quad 17.1, \quad 0.322]$$

In addition, models (3) and (4) were coupled so that the initial parameters were estimated jointly by ML for the models as in the current filtering approach. In the state space formulation, R is a time-varying state; therefore the estimates for these initial conditions on the parameters can be assumed only to be some time-integrated average condition. The constant parameter $T_0 = 261.2$°K in (4) was also developed from a least squares fit of the model to the June 2000 data as mentioned earlier.

[51] The measurement model covariance was determined from the measurement record and from instrument precision. The identical measurement record was used for both filters; therefore $\mathbf{R}_{\mathbf{w}_k}$ and $\mathbf{R}_k$ are also identical. The measurement variances are

$$\mathbf{R}_k = \begin{cases} \mathrm{diag}[1, \quad 219, \quad 5], & \text{dormant season} \\ \mathrm{diag}[7.29, \quad 488, \quad 5], & \text{growing season} \end{cases}$$

Note that in the dormant season, the variability in both NEE and PPFD are less than during the growing season. This is a consequence of restricted respiration and photosynthetic activity coupled with shorter day length.

[52] The state filter requires covariances for the dormant and growing seasons in the $\mathbf{Q}_k$ matrix. For those states (PPFD and T) where process equations are assumed to be random walks, the corresponding elements in $\mathbf{Q}_k$ were set equivalent to those in $\mathbf{R}_k$. The growing season variance on NEE was taken as the overall variance from the model fitted to (3) as given by *Hollinger and Richardson* [2005], and was assumed to be half this value in the dormant season. In addition, the variance for R was the estimated model variance for the Lloyd and Taylor model, (4), fitted by *Richardson and Hollinger* [2005]. The process covariance matrices for both seasons are

$$\mathbf{Q}_k = \begin{cases} \text{diag}[7.37, \quad 219, \quad 5, \quad 10.0, \quad 0.15254], & \text{dormant} \\ \text{diag}[14.74, \quad 488, \quad 5, \quad 10.0, \quad 0.15254], & \text{growing} \end{cases}$$

The state variable INEE accumulates variance from all of its components; therefore it could justifiably be set to zero. However, because the magnitude of the spatial error in the integration is uncertain, we set it to a small, finite value on the order of the estimated variance for NEE.

[53] Lastly, we set $\mathbf{P}_{\mathbf{x}_{k|k-1}}$ and $\mathbf{P}_{\mathbf{w}_{k|k-1}}$ back to their initial conditions at the switch to the growing season in response to several separate idiosyncrasies in our model formulation. First, the process dynamics are different between dormant and growing seasons, as previously explained. Second, we constrain the variance on E_0 to be constant at half its initial value during the dormant season in order to keep this parameter tight and allow excess variation to go directly into R_p, while it remains unconstrained during the growing season. Thirdly, K remains unestimated (fixed) during the dormant season and as a result, the filter variance artificially inflates during this time period and must be reset. Finally, with the exception of A, the initial conditions for the parameter filter were developed on the basis of April and May data, and therefore any evolution to their variances over the dormant season must be discounted once the switch has been made. Inasmuch as the two filters are coupled, the state variances should be reset as well.

3. Results

[54] Figure 2 presents the time courses for three of the system states T, R, and INEE. One prominent feature of the INEE trajectory is the close association between our switching time with the ecosystem's actual switch from carbon source to sink. This is seen, with a slight lag period in the spring, as ecosystem photosynthesis begins to offset respiration. In the fall, it appears that respiration begins to dominate slightly before our switching point; however, because the soil is not frozen at this time, and some daily temperatures (Figure 2, top) do indeed rise above freezing for extended periods, the trajectory rises only gradually into December. The final filtered estimate for INEE was -296.45 g C m^{-2}, with estimated standard deviation of 2.44 g C m^{-2} (which is an estimate of random uncertainty as discussed later). This estimate of Howland forest annual NEE in 2000 is about 25 g C m^{-2} (9%) greater than previously published estimates [*Hollinger et al.*, 2004].

Table 1. Summary of Correlations Between Estimated Forward Filter States and Parameters for the 2000 Howland Data

	Parameters			
States	A	K[a]	E_0	R_p
	Dormant Season			
NEE	−0.024	. . .	0.062	0.042
PPFD	0.038	. . .	−0.161	0.212
T	0.028	. . .	0.054	0.372
INEE	0.362	. . .	−0.958	0.724
R	−0.434	. . .	0.423	0.387
	Growing Season			
NEE	−0.025	0.070	−0.014	0.065
PPFD	−0.065	−0.073	−0.122	0.108
T	−0.630	0.089	−0.528	0.659
INEE	0.706	−0.805	−0.041	−0.405
R	−0.775	0.145	−0.568	0.886

[a]The parameter K is unused during the dormant season.

The previous estimate relied on monthly models identical to equation (3) to fill in missing daytime data, and Fourier series to estimate missing nocturnal values. While difficult to judge from the scale of the figure, the confidence intervals on INEE increase as the year progresses because of the integral definition of this system state. Confidence intervals on the other system states remain fairly constant or increase during periods of missing values, as would be expected. Another apparent trend is the midsummer flattening of INEE trajectory in response to increasing ecosystem respiration and associated moderation of ecosystem photosynthesis (not shown).

[55] The time-varying parameter estimates from the forward filter are presented in Figure 3. There are several points worth noting in these trajectories. First, during the dormant season, A stays close to its initialization value of 0.5 μmol m^{-2} s^{-1}. At the switch, however, it quickly acclimates to its role in the big leaf model, and reaches a maximum (because of our sign convention) in August. A is negatively correlated with R throughout the year (Table 1). The correlation during the dormant season is due to the periods where photosynthesis occurs on warmer days as mentioned above. The strong negative correlation and associated mirror image of the two trajectories during the growing season clearly shows the seasonal relationship between A and R. This result probably stems from the well-known relationship between foliage photosynthetic capacity and foliage respiration [e.g., *Reich et al.*, 1998] and the postulated relationship between current photosynthesis and root respiration [*Hogberg et al.*, 2001]. Changes in the A parameter show gradual reestablishment of forest photosynthetic capacity in the springtime and a precipitous drop in this capacity following the first hard frost in the autumn. The plateauing of A and rapid increase of R_p in June appear in association with the production of new foliage.

[56] The K parameter (PPFD for half-maximum photosynthesis) decreases from our initialization value to a low of about 200 μmol m^{-2} s^{-1} in April, returns to about 400 μmol m^{-2} s^{-1} in May and then gradually rises through the autumn to about 800 μmol m^{-2} s^{-1} before dropping again in November. The increase in linearity in the PPFD:canopy photosynthesis relationship late in the season may be related

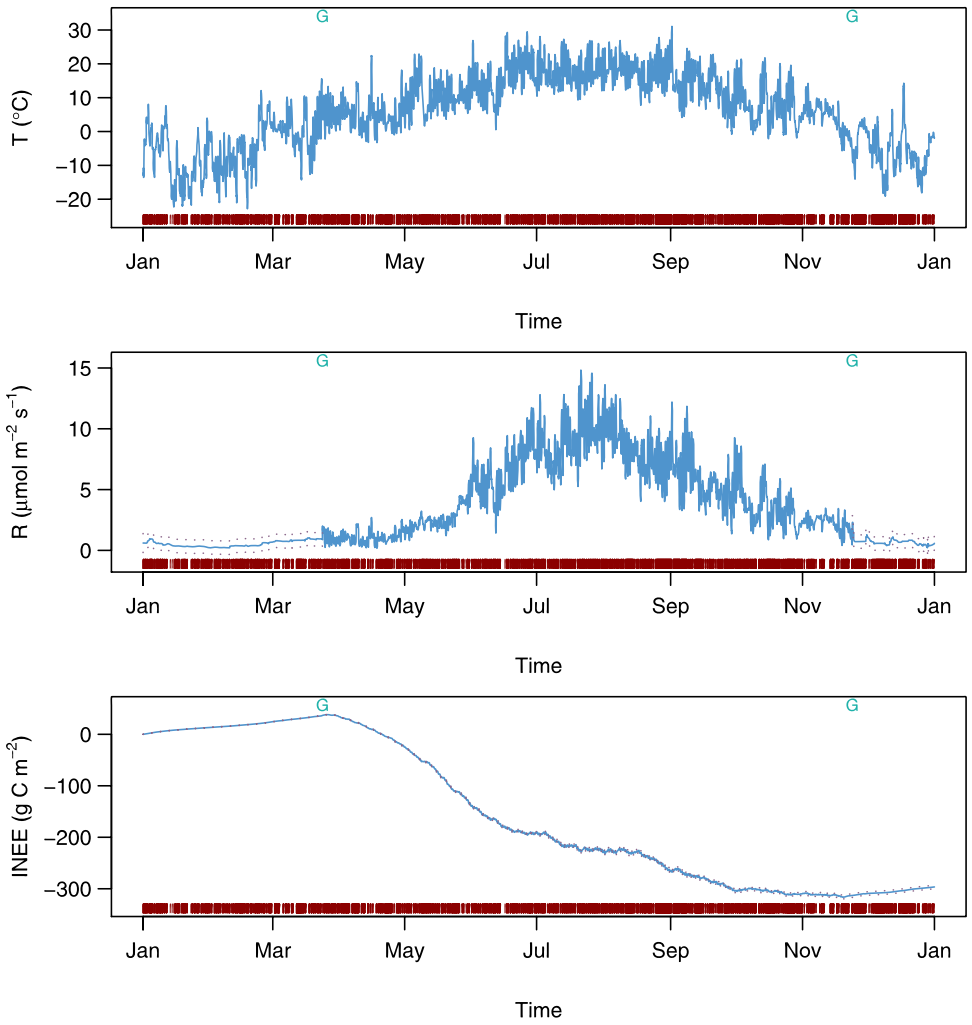

Figure 2. Kalman smoother estimates for the states T, R and INEE for the year 2000 data, with 95% confidence intervals (dashed) shown for the dormant season on R and the entire year on INEE. The vertical bars at the base of each graph show the missing values while the growing season is delineated at the top by letter G.

to the low sun angle and consequent lower PPFD intensities. Variation in leaf-level factors such as %N or canopy factors such as leaf area index could also contribute to the observed variation in the K parameter.

[57] The baseline respiration rate, R_p and ecosystem respiration are clearly highly correlated (Table 1) with very similar trajectories. This is expected from the construction of the respiration model where E_0 is a "shape factor" for the temperature response of respiration. The confidence interval widths at the switch between the dormant and growing season in March for R_p also illustrates the effect of resetting the initial conditions for the growing season at the switch. If the variances were not reset, it could allow R_p to shift more radically at this junction point. However, because we also relax the constraint on the variance of E_0 at this point, E_0 is allowed to vary freely as the full Lloyd-Taylor model comes online under the influence of temperature; these combined effects tend to keep R_p stable during this period, while allowing E_0 to absorb the variation in respiration at the onset of the growing season.

[58] It should be noted that there have been no constraints of any kind imposed upon our model at the switch from growing to dormant seasons in the fall. Yet all of the parameters seek their winter-spring dormant season values, bearing in mind the confidence of the estimates. Because the fall-winter dormant period is independent of the winter-spring season, our results suggest that the spring switch is the most sensitive of the two. Indeed, in numerous sensitivity runs, we found this to be true, resulting in the initial condition and variance switches described earlier.

[59] Figure 4 presents state trajectories for nine days in the month of August 2000. The two random walk models (PPFD and T) tend to track the measurements closely (note again that the measurement record for these states are complete). The exception is daytime PPFD on several of the more variable days (e.g., 7–9 August). Intermittent cloud cover will affect these measurements and the smoothed estimates tend to discount this higher observed variability; that is, they smooth the effect of passing clouds. This result is quite reasonable because the spatial average

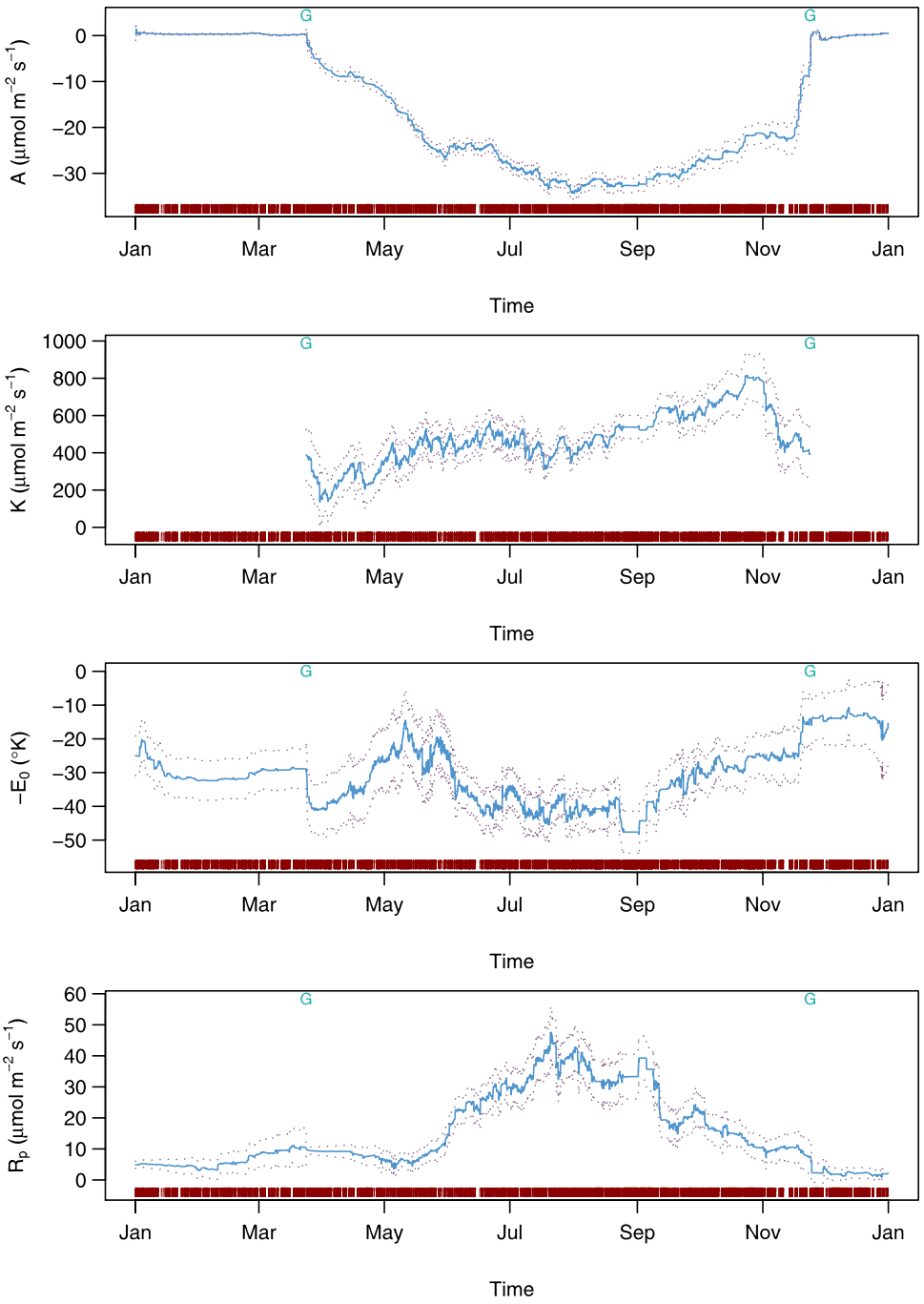

Figure 3. Kalman filter estimates for the four model parameters for the year 2000 data, with 95% confidence intervals (dashed) shown where appropriate. The vertical bars at the base of each graph show the missing values while the growing season is delineated at the top by letter G.

PPFD over the flux source (footprint) region will be less variable than the point measurement made at the flux tower. The estimates for respiration are clearly driven by temperature, the alignment shown is nearly perfect. Again, since T is the driving variable in our model, this result is also reasonable.

[60] Finally, the trajectory for NEE illustrates clearly the efficacy of the Kalman filter as a gap-filling method, even with our simple model. It is difficult to resolve the individual effects that PPFD and T have on the daytime NEE trajectory from the figures alone because of the correlation between PPFD and T (−0.54 for this period). However, correlations between NEE and PPFD for this period were −0.92, while those between NEE and T were −0.55, implying that PPFD is indeed the main driver of the system at this time of year under our model. However, where PPFD is zero at night, there is no photosynthesis and NEE is equivalent to nighttime respiration; therefore the sole driving variable at night is temperature. This effect can be discerned quite clearly on the night of 4 August where all NEE measurements are missing. Here, NEE peaks early in the twilight hours and decreases as T decreases into early

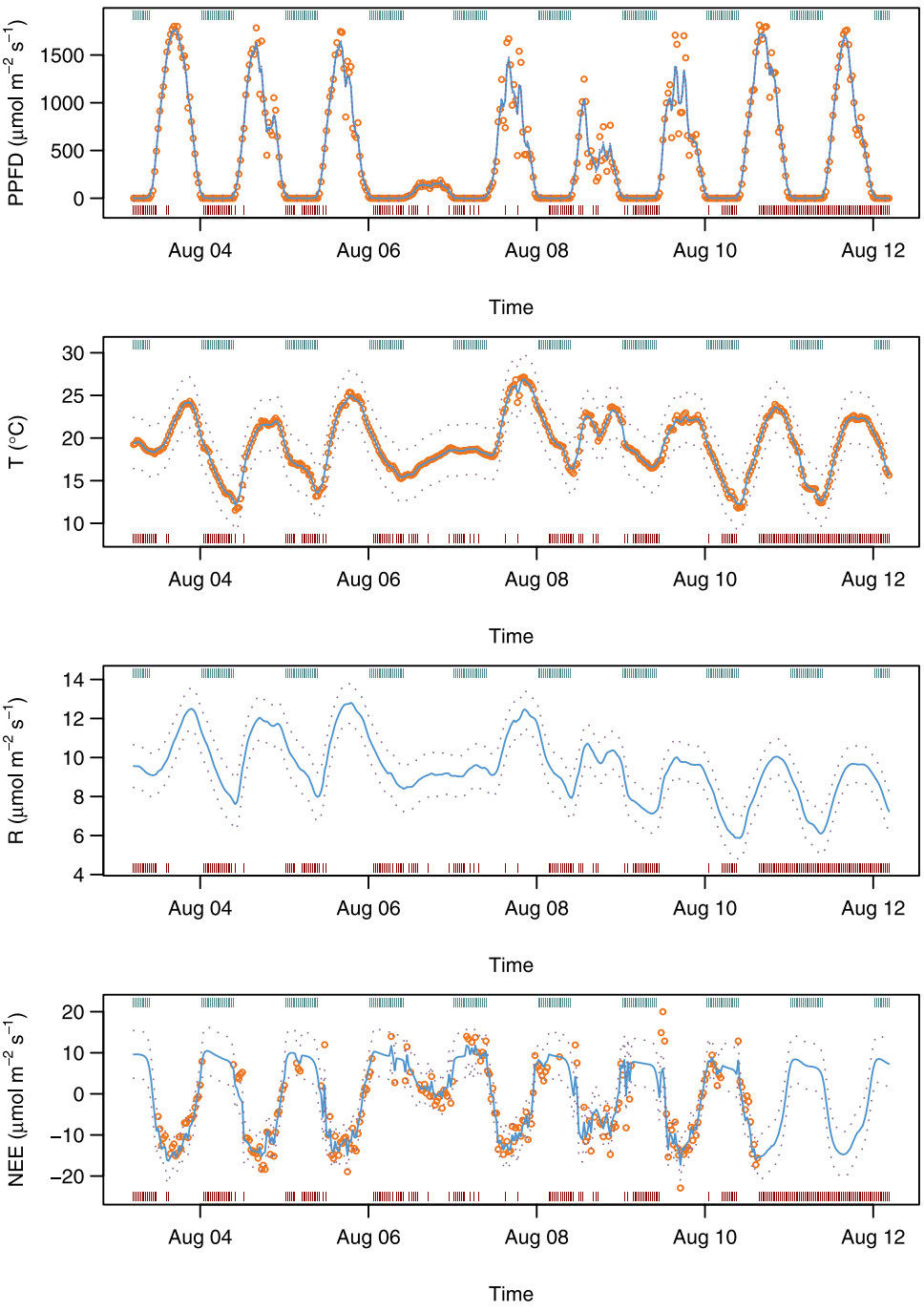

Figure 4. Kalman smoother estimates for the states PPFD, T, R and NEE for several days in August 2000, with 95% confidence intervals (dashed) and noisy observations (circles). The vertical bars at the base of each graph show the times where measurements of NEE are missing while those at the top delineate nighttime.

morning. On nights such as 5 August, where a few measurements interrupt the gaps, the smoother clearly adjusts from all model prediction to an optimal combination of model and data with corresponding downward adjustment in variance. In addition, the estimates for 7 August show how low light levels and cooler temperatures under heavy overcast, are accounted for in the estimates. Finally, there is a gap of several days at the end of this period (not all are shown) where the model predictions, driven by the complete measurement records for PPFD and T evidently do quite an adequate job at describing the trajectory for NEE during this period, when compared against similar days with complete data records. The conclusion to be drawn is that while our model is simple, it appears to be quite powerful in capturing a good degree of the system dynamics in relatively few state variables and associated few parameters.

[61] Incidentally, while the yearly trends in parameter estimates are interesting when viewed as a whole, for short

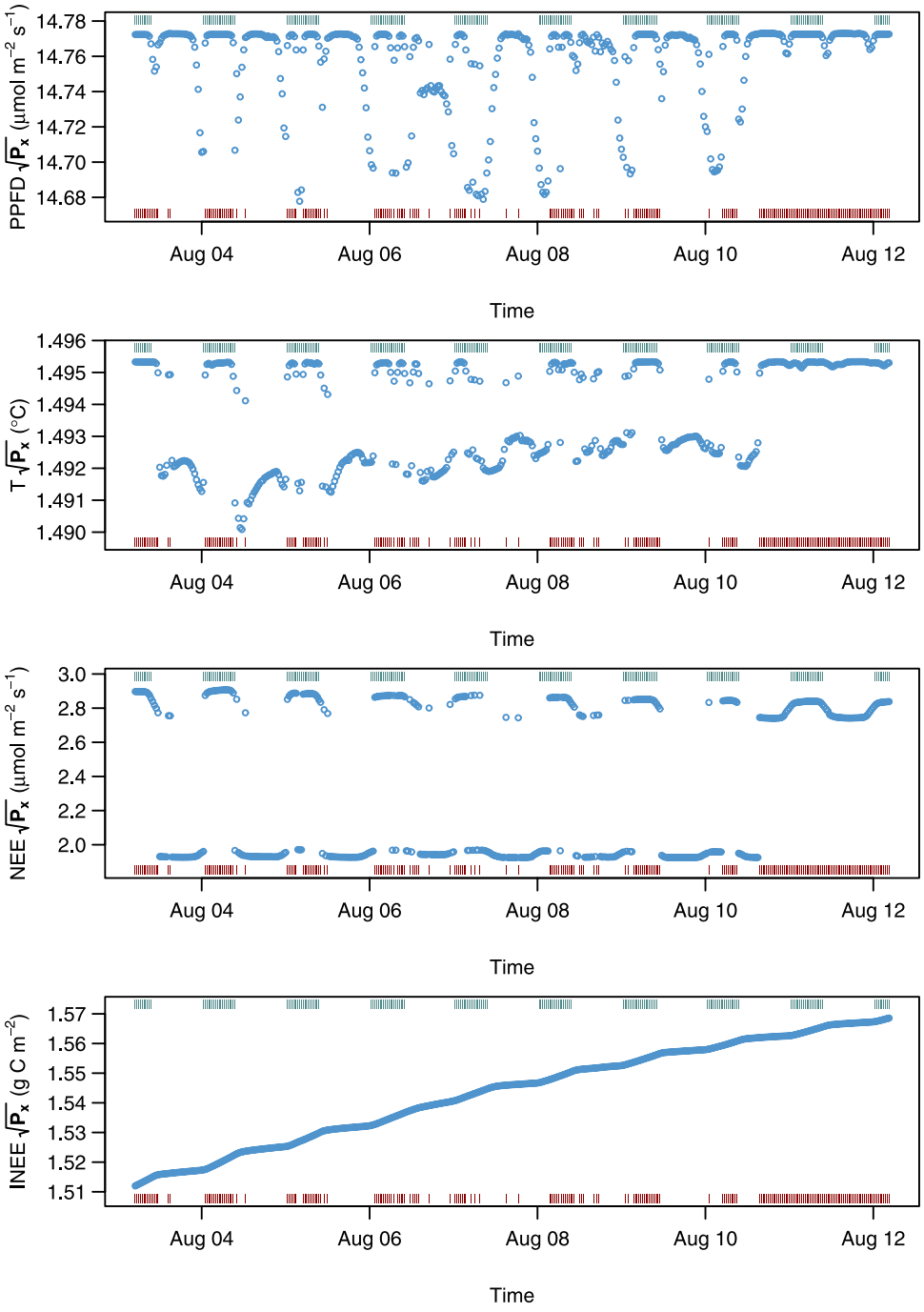

Figure 5. Kalman smoother standard deviation estimates for the states PPFD, T, NEE and INEE for several days in August 2000. The vertical bars at the base of each graph show the times where measurements of NEE are missing while those at the top delineate nighttime.

periods like our nine day illustration, they tend to be relatively constant and show only a mild degree of variation (see Figure 3). This is to be expected since the system itself is in a fairly steady state over such short periods. Large jumps in parameter values during such a period would be alerting to some exogenous inputs into the system that are not incorporated in the model.

[62] The confidence intervals on the yearly INEE trajectory and the 9-day PPFD trajectory are difficult to discern because the variances are small; their magnitude is shown in Figure 5 for the August example. Figure 5 clearly illustrates the increasing nature of the variance (as shown through the standard deviation) for INEE as mentioned earlier. In addition, it illustrates that there are essentially two distinct steady states that both the NEE and T variances tend to take, corresponding to whether NEE measurements are present or not; missing NEE values tend to raise the variances for these states to the higher level. In both cases, there is also a discernible difference in daytime and nighttime variance estimates for both NEE and T. PPFD is the exception, the

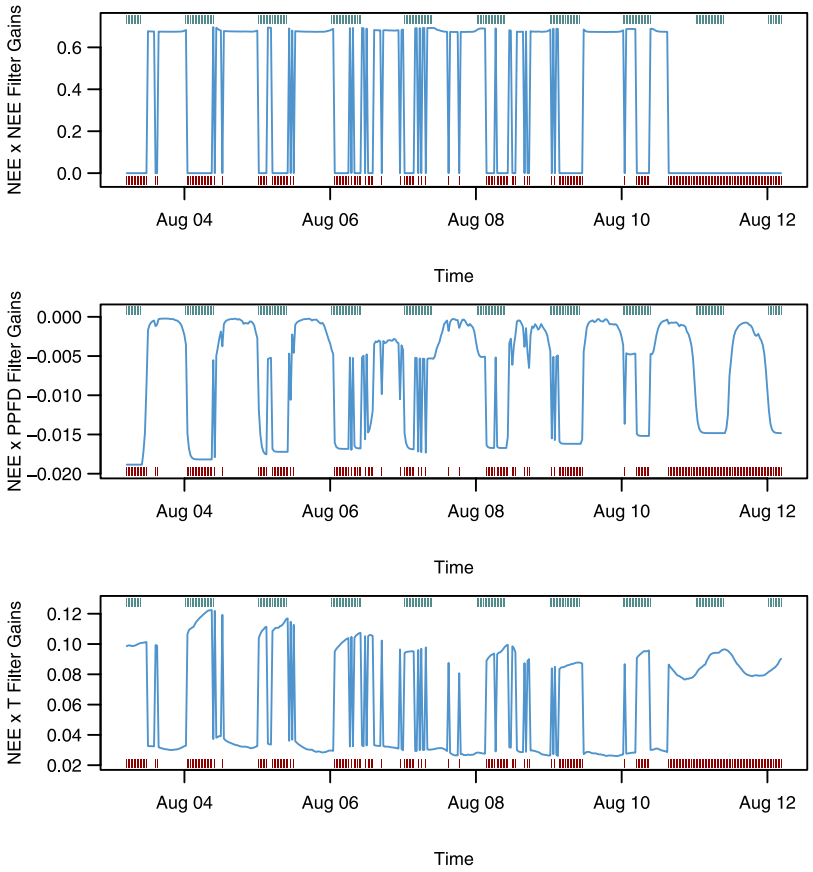

Figure 6. Forward Kalman gain estimates, $\mathbf{K}_k$, for the state NEE and measurements NEE, PPFD and T for several days in August 2000. The vertical bars at the base of each graph show the times where measurements of NEE are missing while those at the top delineate nighttime.

two levels seem to be missing and the variance values tend to fluctuate coincident with the steep downslope or upslope of the PPFD curve in the transition from day to night or vice versa. This may be at least partially an artifact of our setting all measurement values with PPFD < 5 μmol m^{-2} s^{-1} to zero in the preprocessing stage which affects the random walk model. Because the smooth estimates benefit from both forward and backward recursions over the data, this trend is established in both directions.

4. Sensitivity

[63] As mentioned earlier, one of the main assumptions leading to optimality of linear Kalman filters, or near optimality of nonlinear filters, is that the noise variances are known. We estimated these variances directly from the data. However, other methods for the estimation of these variances based on filter estimates have also been mentioned and include ML using the method of scoring on the prediction error decomposition form of the log-likelihood [*Harvey*, 1989, p. 140; *Gove and Houston*, 1996], and EM algorithms [*Shumway and Stoffer*, 1982], both of which make assumptions about the distribution of the error processes at a minimum. Because any discussion of sensitivity must be interpreted in light of the Kalman gains, which are influenced by these variances, we begin with a brief discussion of the role this matrix plays in the forward filter.

[64] The Kalman gain matrix, $\mathbf{K}_k$, is the optimal weight matrix for combining measurements and model predictions in the update step (21). It is composed of the cross covariance between the predictions (from resampled sigma points) and the measurements, $\mathbf{P}_{\mathbf{x}_k\mathbf{y}_k}$, weighted by the inverse of the measurement covariance $\mathbf{P}_{\mathbf{y}_k}$. Given our linear identity system transform, $\mathbf{h} = \mathbf{H}$, the gains are identically computed by (24). The gains can range from positive to negative depending on the cross covariances. Because of the structure of $\mathbf{H}$ in our formulation, the diagonal elements of $\mathbf{K}_k$ will be constrained to range between zero and one. When the gains are close to zero, preference is given to the model and the measurements are discounted; conversely, when the gains are near one, the measurements are preferred over the process model. In a simple univariate measurement system, this correspondence is exact; however, in a multivariate setting, the gain matrix has dimension $n_x \times n_y$ which means that even when there are missing values for one measurement variable, the gains will reflect, to some extent, the influence of the other measurements and state predictions in the resulting estimates. For example, when NEE is missing at time k, the first column of the gain matrix (as well as the residual for NEE) will be zero, but the rest of $\mathbf{K}_k$ will be nonzero. Therefore the update estimate, $\hat{\mathbf{x}}_k$, and its associated variance estimate, $\mathbf{P}_{\mathbf{x}_k}$, will be composed of the prediction step estimate plus a weighted linear combination of the other nonmissing innovations (or prediction covariances) at this time period. Similarly, negative gains act like

negative covariance terms, changing the sign of that innovation's contribution to the estimated state $\hat{\mathbf{x}}_k$ in (22).

[65] An example illustrating the influence of the gains on the estimates for NEE for our 9-day period in August can be seen in Figure 6. From Figure 6 it can be seen that when NEE is missing, the associated gains for PPFD and T are nonzero and therefore the prediction estimate for NEE will be adjusted by a weighted linear combination of these two innovation errors in the update step. When NEE is not missing, the NEE gains dominate those for PPFD and T so that most, but not all, of the updated estimate for NEE comes from its own innovation adjustment.

[66] Given this explanation, it should be clear that the Kalman update step is a highly integrated combination of information between states and measurements; in other words, there is a large degree of interaction between the variables of interest and one should not mistakenly infer that the states are estimated independently of one another. Because the Kalman filter propagates the joint density of the states through prediction and update steps, however, the cross link to the measurements through the gains can be easily overlooked.

[67] The influence of the Kalman gains are important because the state update estimates in the forward filter are the best estimates at any given time period in the forward recursions. A similar gain matrix also exists for the backward recursions, but not, per se for the smoothed estimates themselves. More importantly, the gains are influenced by the noise covariances $\mathbf{R}_k$ and $\mathbf{Q}_k$, and their affect on the gains can be assessed through sensitivity analysis in the process of filter tuning. The larger the number of states and measurements, the more difficult any sensitivity analysis is to undertake because of the interactions entering into the estimates that are partially manifest through the gains as just described. Any analysis must be undertaken conditionally, by modifying only one or two variance components at a time while holding the others constant. In addition, some criterion must be chosen for assessing the effects of the analysis. Because the final estimate of INEE and its variance are of major importance, we limit the discussion to the final forward filter estimate of these quantities.

[68] We ran numerous sensitivity trials in the tuning phase of the filter. In general, we noticed that the model we present is quite robust to changes in both initial conditions and individual (and sometimes joint) variance components. One interesting point to notice is that the standard deviation of our estimate for INEE is only 2.44 grams. Increasing the variance component for INEE in $\mathbf{Q}_k$ tenfold, produced no significant difference in either the state estimate or the estimated variance. Similarly, since it has been established that NEE (directly) and PPFD (via (3)) are the major contributors to INEE in our model, modifying their respective components of $\mathbf{R}_k$ and $\mathbf{Q}_k$ would be of interest. Increasing the measurement noise variance component for NEE in $\mathbf{R}_k$ tenfold during the dormant season and adding a fourfold increase in the measurement noise components for NEE during the growing season as well, resulted in no significant change to either the state or variance estimates for INEE. However, adding to these the same increases in the variance component for NEE in $\mathbf{R}_{\mathbf{w}_k}$ in the parameter filter reduced the overall state estimate of INEE by 24 grams and the associated estimated standard deviation by half a gram. These results suggests that the model is much less sensitive to increasing the measurement noise variance during the dormant season than during the growing season. In addition, it illustrates that the uncertainty of the estimate is quite robust to changes in NEE measurement noise variance. However, the estimate itself was decreased because of the tendency of the gains to weight more favorably in terms of the model, discounting the NEE measurements under the scenario of increased measurement noise forcing. Adding the increased variance to the parameter filter affected the growing season trajectory for E_0 while the other parameters remained largely unchanged. Similar results were found by increasing the process noise variance components for PPFD by 1000 in both the dormant and growing seasons.

[69] In order to observe any substantive change in estimated variance for the final INEE estimate, it was necessary to increase the measurement noise variance on PPFD fourfold for the growing season. This still only raised the final estimated standard deviation by 2 grams; however, it did also quite drastically increase the estimated INEE by close to 70 grams. Fortunately, this final scenario is so extreme, that it would be considered outside the possible variance domain for our model.

[70] One last point to bring out is the suggestion from the above runs that changing the variances in $\mathbf{R}_{\mathbf{w}_k}$ for the parameter filter exert a significant influence on the final estimate. However, this change in variance was in addition to parallel changes in $\mathbf{R}_k$ that had already been made. Adding these effects in a different order, it is easy to show that the change was really due to the increased variance component on NEE during the growing season for both filters. However, this raises the interesting question concerning the effects of the parameter variances on the final estimates when all initial conditions and variances are held to those in our original formulation. To help answer this question, $\mathbf{R}_{\mathbf{w}_k}$ was set equal to the identity matrix in a new filter run. The estimate for INEE was -295.94 grams with a standard deviation of 2.28 grams. All parameters effectively tracked the original estimates exactly during the dormant season and A, K and R_p were very close during the rest of the year. The main deviation was found in growing season estimates for E_0. Even then, however, the overall trajectory was very similar, just raised higher during the late spring and early summer and converging again in September. This exercise is important because it illustrates the robustness of the filter to changes in parameter measurement variance components. The parallel track taken by E_0 (with slight adjustments in R_p) could be due to equifinality, which has been observed in several other inverse modeling analyses of surface-atmosphere exchange [*Franks et al.*, 1997; *Schulz et al.*, 2001]. The main effect of decreasing the measurement variances in the parameter filter is to correspondingly decrease the estimated variances and subsequent confidence interval widths. Since the parameter filter is indeed a minimizer, this result makes a good deal of sense; it is discussed in greater detail by *van der Merwe* [2004].

5. Discussion

[71] In this study the dual approach to state and parameter estimation was used. However, as noted earlier, a related joint approach could also have been used. Theoretically, the

difference between the two is that the joint approach assumes explicit statistical dependence between states and parameters by allowing estimation of the cross state-parameter covariance terms, whereas in the dual approach these terms do not exist, effectively tantamount to assuming independence. However, even though the states and parameters were estimated in this quasi-independent manner under the dual filter (i.e., the cross covariances were implicitly zero), the correlations between states and parameters from our filter estimates (Table 1) show that the state-parameter switching at each time period does indeed introduce some implicit dependence. It would therefore be an interesting experiment, given this dependence, to see whether a joint modeling approach might be fruitful within our filter model formulation and whether the results obtained would differ significantly. This might be undertaken as part of a larger Monte Carlo simulation study. It is anticipated, however, that the estimate of INEE calculated via the joint approach would be similar to the present estimates.

[72] Two other Kalman-based methods have been developed recently that are very similar to the UKF. *Ito and Xiong* [2000] developed the central difference filter, while *Nørgaard et al.* [2000] independently developed the divided difference filter. Both filters use Sterling's interpolation methods to propagate the first two moments through the filter using a derivativeless deterministic sampling approach as in the UKF. Because the algorithms are so similar, these two filters can be considered the same for all practical purposes. In comparisons with the UKF, *van der Merwe* [2004] has found that they both yield very similar results.

[73] The literature on optimal estimation is vast, and numerous different filtering algorithms have been developed over the years to address different shortcomings to the Kalman filter when applied outside its original intent (i.e., linear propagation of the first two moments of the joint density). Methods for parameter estimation, nonlinearities, estimation of initial conditions, Bayesian extensions and many others have been developed; however, the basic set of recursions developed by Kalman still lay at the heart of most of these extensions. This is also true of two methods that have recently been applied to eddy flux data. In the first study, the state-dependent parameter (SDP) estimation approach [*Young*, 2000] was applied to eddy covariance data on an hourly time step [*Jarvis et al.*, 2004]. The SDP approach makes no a priori assumptions about the process models; rather it uses a multistep approach culminating in an empirical process formulation. In the SDP approach, the Kalman filter and smoother are run on nontemporally sorted data with time-varying parameters following a random walk model, facilitating simultaneous gap filling in this sorted state space. Final model identification was subsequently conducted from the final Kalman smoother estimates integrated over a daily time step to capture seasonal variation, producing site-specific empirical process models with constant parameter estimates.

[74] In a slightly more recent study, *Williams et al.* [2005] have employed a very powerful method known as the ensemble Kalman filter (EnKF), originally developed for high-dimensional modeling of ocean systems [*Evensen*, 2003]. The EnKF is a Markov Chain Monte Carlo approach that runs a parallel ensemble of filters together in time. The ensemble nature of this filter obviates the need to propagate forward high-dimensional covariance estimates as part of the filter algorithms, while also handling nonlinear dynamics without resorting to linearization. *Williams et al.* [2005] applied the EnKF to data that had been previously filled and integrated to a daily time step to accommodate their process models. Their application embedded the filter within a minimizer (sums of squared innovations) to estimate constant parameter values and initial conditions. *Harvey* [1989, p. 129] discusses the conditions under which such a criterion function is asymptotically equivalent to a maximum likelihood approach assuming linear Gaussian noise. This approach results in fixed initial conditions and constant parameter estimates for use in the process model dynamics in the filter.

[75] The SDP and EnKF approaches differ from our approach in several respects. First, we have made no assumption about the distributional properties of the noise terms, measurements, or states other than that the joint densities can be captured by their first two moments. Because Gaussianity may not be appropriate for our measurement data [*Hollinger and Richardson*, 2005], traditional Gaussian-based ML and EM approaches for hyperparameter and initial condition estimation, while robust, may not be entirely appropriate, especially with nonlinear system dynamics. Instead, we opted for estimating initial conditions and noise variances prior to filtering, while letting the process model parameter estimates evolve in time contemporaneously with the states, yielding a time-varying parameter trajectory that is correlated with the state estimates. This result is appealing for simple process models like ours, because the time-varying parameters may help to account for unmodeled system processes. This result is also in contrast to the other approaches, where the final parameter estimates are constants. Finally, in our approach gap-filling occurs coincident to estimation, whereas in the EnKF gaps were filled prior to estimation, or in the case of SDP estimation, while the data were sorted nontemporally.

[76] In light of the new applications of SDP and EnKF methods, eddy flux scientists now are beginning to have a number of choices for efficient data-model fusion methods for estimation of nonlinear systems that would include the UKF and its close variants. We have already suggested several more salient ways in which our approach could be improved, in the following we provide some discussion of the more subtle components of our model and analysis, with an eye toward future enhancements.

5.1. Time-Varying Parameters

[77] Several studies have recently explored the idea that carbon flux model parameters vary with time. We have already discussed the SDP approach taken by *Jarvis et al.* [2004], for example. In addition, *Raupach et al.* [2005], in discussing a joint filtering approach with state vector augmentation, note that allowing parameters to drift through time is a potential advantage of such techniques. However, such observations are not limited to filtering approaches. *Wang et al.* [2003] demonstrate how model parameters change seasonally and yearly using process models similar to ours, when fitted by traditional least squares. *Hollinger et al.* [2004] use ML to fit (3) and (4) to monthly 1996 eddy flux data from Howland. They observed monthly variation

in the parameter estimates with a high degree of correlation between A and K. These authors suggest that parameter variability may be due at least in part to the simplicity of the model, and conjectured that incorporating other factors might allow at least one of the parameters to stabilize.

[78] Our estimation approach yields time-varying model parameters (Figure 3), and while this appears to be useful in settings such as ours, there are still some unresolved questions. First, if parameter variability is due in part to overly simplistic models, one might propose using more complicated models that would include a more complete range of the state processes captured by the measurements. Indeed, one can imagine a complex model that almost completely accounts for ecosystem behavior (and would thus have at least some time-invariant parameters) that is far beyond our simple process model. Compared to other common models of ecosystem C exchange, we ignore temperature effects on photosynthesis, saturation deficit, soil water stress, ratio of direct to diffuse PPFD, seasonal phenological effects such as foliage production, and other factors. The influences of these other factors are, to some extent, swept into the four free parameters of the present model, and when their influence is strong, such as the impact of the first hard frost in the fall (i.e., several days prior to the switch), the effect on the model parameters can be severe. In addition, while more complicated process models can easily be implemented with the UKF approach, doing so introduces more states and parameters, increasing the dimensionality of the estimation problem. In such systems, correlations between parameters could become more of a problem as the limited information content in the flux measurements [*Schulz et al.*, 2001] does not change commensurate with the increased model dimensionality. Indeed the extent to which free parameters vary in a model-data fusion such as this could be diagnostic of the "completeness" of the model.

[79] Second, it is unclear at this time how the extramodel variation in the measurements is being partitioned between noise covariances and parameter variability in the dual approach. In the usual application of the Kalman filter to state estimation (parameters assumed known or possibly estimated by embedding the filter in likelihood iteration), the role of the noise covariance matrices are quite well understood in relation to their effects on the gains, and thus the optimal weighting of measurements and model predictions in the update step. For example, in univariate filters, the ratio of process to measurement noise variances is often used, e.g., to determine the rate of convergence to steady state conditions, while in multivariate settings it is the ratio of the eigenvalues of the noise covariance matrices [*Maybeck*, 1979, p. 224]. However, we noted in the sensitivity analysis to our approach that often fairly large (several-fold) changes could be made to components of the noise covariance matrices without changing the final state or variance estimates of INEE significantly. This can be corroborated by noting that increasing the process noise variances for the random walk components of the model tenfold, changed the final estimate of INEE by less than a gram, while the estimated variance remained virtually unchanged. This brings up the interesting, and unresolved question as to whether the ability of time-varying parameters in such systems to absorb extramodel noise in the observations mediates to some extent the classic roles of the noise covariance terms.

5.2. Switching Models

[80] The deterministic switch in our model can produce a fairly abrupt change in some estimated model parameter trajectories. This can be partially explained by noting that in the dormant season, NEE is composed mostly of respiration through R, but on warm days some photosynthesis can occur in A. While the simple dormant season models in (7) and (8) are certainly useful, they impart little knowledge of the process dynamics. At the winter-spring switching point, two related events happen. First, NEE is partitioned into R plus photosynthesis components, rather than being composed almost entirely of R. Second, a switch is made from models (7) and (8) to the process models for NEE and R. These two "shocks" to the system dynamics in a filter that has been in a fairly steady state, awaken it, as it were, to the growth that is beginning in the ecosystem. Thus any observed jumps in parameter trajectories are a manifestation of our models and the sharp switch between the seasons in spring. This same explanation can be used in the fall-winter season and applies to the drop in R_p that happens at the switch point. While R_p has been declining steadily with the decline in R, we effectively turn off its relationship to temperature and again force it to be equated to R and NEE. Simultaneously, E_0 is adjusted more heavily because of its larger variance, accommodating the adjustment required in R_p.

[81] In general, a hard deterministic switch is undoubtedly not the best method to use in the transition between dormant and growing seasons because in many places such transitions tend to be gradual. Other possibilities include precomputing running temperature averages or including soil temperature as a measurement and state to facilitate model refinements that would allow for multiple switch points during the transition period. In addition, the concepts used in stochastic switching models [*Hamilton*, 1993], might also be fruitfully employed.

5.3. Estimation Uncertainty

[82] The standard deviation in the UKF estimation of annual NEE (2.44 g m^{-2}) is less than 1% of estimated annual NEE. Compared to previous estimates of uncertainty in annual NEE [*Goulden et al.*, 1996b; *Falge et al.*, 2001a; *Griffis et al.*, 2003], this is a low value and warrants additional discussion. First, when NEE data are missing, the filter assumes that the model parameters are fixed, having no new information with which to update them, only adjusting parameters to new values when measured data again become available. With locally fixed parameters and only environmental data available, model uncertainty for these gaps will thus be low. On the basis of the observed parameter variability (Figure 3) this is reasonable for short gaps (a few days) but not over longer periods (weeks). In light of this, gaps are also more problematic in the beginning and end of the growing season when model parameters are changing rapidly.

[83] Second, our present results show that, while there is some change in the variance of the estimates for NEE where gaps occur, the variance does not increase substantially with time, even over many consecutive days of missing obser-

vations. This seems somewhat counterintuitive as one would expect there to be less certainty as we progress in time in the absence of measurements. However, it is well known that the linear Kalman filter covariances (both prediction and update) and gains, will reach a steady state in circumstances where the system matrices **F** and **H**, as well as the noise covariances **Q** and **R** are constant [*Gelb*, 1974, p. 142]. Indeed, under such conditions the steady state may be reached after only a few time steps [e.g., *Maybeck*, 1979, pp. 223–226]. This phenomenon occurs because the filter covariances and gains do not depend upon the measurements directly, even though the conditional means (state estimates) do. This is actually an appealing situation, since it allows filter designers to precompute the gain and filter variances offline, before any measurements are collected; these can then be used for optimizing filter design [*Maybeck*, 1979, p. 222]. In our model, **H**, **Q** and **R** are all constant matrices as in the linear case. In addition, our system model, **f**, is only mildly nonlinear. Since the UKF performs a statistical linearization at the current state, the mild nonlinearity evidently is discounted and the system becomes effectively linear.

[84] It can be seen from Figure 5 that the smoothed standard deviations for NEE are essentially constant at two levels corresponding to missing and nonmissing data (with slight diurnal variation). In addition, so are the corresponding NEE $\times$ NEE gains in Figure 6. Mathematically, the update variance contribution in our model for the nonmissing time periods is $(1 - k_{11})p_{11} - k_{12}p_{21} - k_{13}p_{31}$ (where the k_{ij} are elements of the gain matrix and the p_{ij} are from the prediction covariance matrix at some fixed time) and for those with missing data, $p_{11} - k_{12}p_{21} - k_{13}p_{31}$, since the gain is set to zero for NEE (i.e., $k_{11} = 0$). Thus, when the covariance matrices are in steady state (i.e., the p_{ij} and k_{ij} are constant), the variances simply fluctuate between these two different values depending upon whether NEE measurements are present or not. In more complicated models where the system or noise covariance matrices are allowed to vary in time, this interpretation would not necessarily hold.

[85] Lastly, flux data are also characterized by several systematic errors (biases) that are more difficult to quantify and were excluded from the present analysis. These errors include incorrectly specified spectral correction models to account for high- and low-frequency losses by real measurement systems using finite integration intervals, as well as possible errors in instrument calibration, data acquisition, and processing. Generally, it is believed that the largest systematic errors in CO_2 flux measurements are related to problems of insufficient nocturnal turbulence. The eddy covariance method depends upon scalar turbulent transport and if the nocturnal atmosphere becomes overly stable, transport between the ecosystem and overlying instrumentation can become temporarily uncoupled with the result that the measured flux is less than the true flux. Stable nocturnal conditions favor the development of katabatic flows that may carry CO_2 out of the ecosystem in surface flows that are never perceived by the flux instrumentation. To screen out artificially low nocturnal flux measurements, many researchers establish minimum turbulence criteria and only accept measurements if turbulent mixing exceeds the criteria threshold. *Goulden et al.* [1996a, 1996b] pioneered this approach and used a mean half-hourly friction velocity (u^*, the square root of the momentum flux) of 0.2 m s^{-1} as their threshold. The exact u^* cutoff is somewhat arbitrary and may depend upon the site. For the Howland data, *Hollinger et al.* [2004] used a u^* threshold of 0.25 m s^{-1} but also calculated NEE with u^* thresholds of 0.2 and 0.3 m s^{-1}. These values led to annual NEE estimates in 2000 that ranged between -287 and -251 g C m^{-2} y^{-1}. In the current analysis, we reran the filter with these same u^* thresholds of 0.2 and 0.3 m s^{-1}. The resulting estimates were -311 and -276 g C m^{-2} y^{-1} with estimated standard deviations of 2.4 and 2.5 grams, respectively. It is interesting to note that the range in both cases is almost identical.

[86] In addition, the u^* data are also noisy measurements; however, they have been used as true, uncorrected values without the benefit of filter estimation, as part of the prefiltering data step. With this in mind, the results of our analysis can be considered conditional on the u^* threshold chosen. Likewise, we note again that the results are also conditional on the values of the a priori estimates of unknown filter components such as the initial conditions and noise covariances. This considered, it is clear that the systematic NEE uncertainty is far larger than the random uncertainty derived from the data-model fusion. Therefore our resulting estimate of uncertainty should not be taken as reflecting the true uncertainty in ecosystem C exchange.

5.4. Gap Filling

[87] *Falge et al.* [2001a] highlight the need for consistent treatment of flux data, especially with regard to the modeling (filling) of gaps in the data record prior to synthesis efforts. Although several different methodologies were compared, including mean diurnal variation, look-up tables, and nonlinear regressions, no method was judged superior. Worryingly, annual NEE could differ by over 100 g C m^{-2} (30%), depending upon methodology. For each of these methods, the results of gap filling also depend upon an arbitrary time period. This may be a moving weeklong window for constructing a lookup table or monthly or seasonal periods for establishing nonlinear physiological model parameter values. In any case, altering the length, or beginning and ending dates of the time periods will change the gap filled values and hence the annual NEE estimate. Lack of consistency in choices will lead to variation in NEE calculated at different sites that has no ecological or environmental basis. Another potential problem is the recent finding of nonnormality in eddy flux measurement error [*Hollinger and Richardson*, 2005; *Richardson et al.*, 2006]. To correctly infer the parameters of models used in the gap filling process from flux data, the ML method should be used. If data errors are normally distributed with constant variance, least squares regression techniques yield ML parameter estimates; however, since flux data diverge from this standard, parameter values calculated by least squares methods are incorrect. These authors argue that flux data uncertainty is better represented by a double exponential probability density function and use the convenient ML estimation properties of this distribution to calculate flux model parameters [*Richardson and Hollinger*, 2005]. They show that filling in eddy flux gaps with models using parameters calculated in this manner can reduce estimated

respiration and consequently increases estimated INEE by 40 g C m^{-2} relative to least squares methods.

[88] We suggest that the UKF described here (or some similar variant) may be a suitable tool for providing consistency in filling gaps in eddy flux data time series. The UKF provides an approximate, close-to-optimal solution for the missing flux estimation problem (no nonlinear method can be truly optimal in general, because the optimal Bayesian solution requires propagating the full non-Gaussian joint probability density) and the autoregressive nature of the filter means that the prediction models are continuously updated, eliminating subjective decisions about time periods or dates that underlie other methods. The UKF is not limited to assumptions of normality in the data uncertainty. The filter requires information about the measurement and model uncertainty, which can be estimated in several ways for flux towers [e.g., *Hollinger and Richardson*, 2005] and by Monte Carlo methods for models [e.g., *Richardson and Hollinger*, 2005]. As presently formulated, the UKF also requires complete (gap free) time series of PPFD and T; however, alternative formulations may not impose this requirement. Once the work of coding the UKF has been completed, it can quickly run through data from a variety of years or sites.

[89] **Acknowledgment.** The authors would like to thank the three anonymous reviewers whose insightful suggestions greatly improved this manuscript.

References

Baldocchi, D. D. (2003), Assessing the eddy covariance technique for evaluating carbon dioxide exchange rates of ecosystems: Past, present, and future, *Global Change Biol.*, *9*(4), 479–492.

Bell, B. M., and F. W. Cathy (1993), The iterated Kalman filter update as a Gauss-Newton method, *IEEE Trans. Autom. Control*, *38*(2), 294–297.

Chen, Y., T. Huang, and Y. Rui (2002), Parametric contour tracking using unscented Kalman filter, paper presented at International Conference on Image Processing, Inst. of Electr. and Electron. Eng., Rochester, N. Y.

Crassidis, J. L., and F. L. Markley (2003), Unscented filtering for spacecraft attitude estimation, *J. Guidance Control Dyn.*, *26*(4), 536–542.

Davidson, E. A., A. D. Richardson, K. E. Savage, and D. Y. Hollinger (2006), A distinct seasonal pattern of the ratio of soil respiration to total ecosystem respiration in a spruce-dominated forest, *Global Change Biol.*, *12*, 230–239.

Evensen, G. (2003), The ensemble Kalman filter: Theoretical formulation and practical implementation, *Ocean Dyn.*, *53*, 343–367, doi:10.1007/s10236-003-0036-9.

Falge, E., et al. (2001a), Gap filling strategies for defensible annual sums of net ecosystem exchange, *Agric. For. Meteorol.*, *107*(1), 43–69.

Falge, E., et al. (2001b), Gap filling strategies for long term energy flux data sets, *Agric. For. Meteorol.*, *107*(1), 71–77.

Fitzgerald, R. J. (1971), Divergence of the Kalman filter, *IEEE Trans. Autom. Control*, *16*, 736–747.

Franks, S. W., K. J. Beven, P. F. Quinn, and I. R. Wright (1997), On the sensitivity of soil-vegetation-atmosphere transfer (SVAT) schemes: Equifinality and the problem of robust calibration, *Agric. For. Meteorol.*, *86*(1–2), 63–75.

Fraser, D. C., and J. E. Potter (1969), The optimum linear smoother as a combination of two optimum linear filters, *IEEE Trans. Autom. Control*, *7*(4), 387–390.

Gelb, A. E. (Ed.) (1974), *Applied Optimal Estimation*, MIT Press, Cambridge, Mass.

Goulden, M. L., J. W. Munger, S. M. Fan, B. C. Daube, and S. C. Wofsy (1996a), CO_2 exchange by a deciduous forest: Response to interannual climate variability, *Science*, *271*(5255), 1576–1578.

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

Gove, J. H., and D. R. Houston (1996), Monitoring the growth of American beech affected by beech bark disease in Maine using the Kalman filter, *Environ. Ecol. Stat.*, *3*, 167–187.

Griffis, T. J., T. A. Black, K. Morgenstern, A. G. Barr, Z. Nesic, G. B. Drewitt, D. Gaumont-Guay, and J. H. McCaughey (2003), Ecophysiological controls on the carbon balances of three southern boreal forests, *Agric. For. Meteorol.*, *117*(1–2), 53–71.

Hamilton, J. D. (1993), Estimation, inference and forecasting of time series subject to changes in regime, in *Handbook of Statistics*, vol. 11, edited by G. S. Maddala, C. R. Rao, and H. D. Vinod, Elsevier, New York.

Harvey, A. C. (1989), *Forecasting, Structural Time Series Models and the Kalman Filter*, 1st ed., Cambridge Univ. Press, New York.

Hogberg, P., A. Nordgren, N. Buchmann, A. F. S. T. A. Ekblad, M. N. Hogberg, G. Nyberg, M. Ottosson-Lofvenius, and D. J. Read (2001), Large-scale forest girdling shows that current photosynthesis drives soil respiration, *Nature*, *6839*, 789–792.

Hollinger, D. Y., and A. D. Richardson (2005), Uncertainty in eddy covariance measurements and its application to physiological models, *Tree Physiol.*, *25*, 873–885.

Hollinger, D. Y., S. M. Goltz, E. A. Davidson, J. T. Lee, K. Tu, and H. T. Valentine (1999), Seasonal patterns and environmental control of carbon dioxide and water vapor exchange in an ecotonal boreal forest, *Global Change Biol.*, *5*(8), 891–902.

Hollinger, D. Y., et al. (2004), Spatial and temporal variability in forest-atmosphere CO_2 exchange, *Global Change Biol.*, *10*(10), 1689–1706.

Ito, K., and K. Xiong (2000), Gaussian filters for nonlinear filtering problems, *IEEE Trans. Autom. Control*, *45*(5), 910–927.

Jarvis, A. J., V. J. Stauch, K. Schulz, and P. Young (2004), The seasonal temperature dependency of photosynthesis and respiration in two deciduous forests, *Global Change Biol.*, *10*(6), 939–950.

Jazwinski, A. H. (1969), Adaptive filtering, *Automatica*, *5*, 475–485.

Jazwinski, A. H. (1970), *Stochastic Processes and Filtering Theory*, Elsevier, New York.

Julier, S. J., and J. K. Uhlmann (1997), A consistent, debiased method for converting between polar and cartesian coordinate systems, in *Proceedings SPIE AeroSense: Acquisition, Tracking and Pointing XI*, vol. 3086, pp. 110–121, Int. Soc. of Opt. Eng., Bellingham, Wash.

Julier, S. J., and J. K. Uhlmann (2002), The scaled unscented transformation, paper presented at American Control Conference, Am. Autom. Control Counc., Anchorage, Alaska.

Julier, S. J., and J. K. Uhlmann (2004), Unscented filtering and nonlinear estimation, *Proc. IEEE*, *92*(3), 410–422.

Kalman, R. E. (1960), A new approach to linear filtering and prediction problems, *J. Basic Eng., Ser. D*, *82*, 34–45.

Kitagawa, G. (1999), Processing of missing observations and outliers in time series, in *The Practice of Time Series Analysis*, edited by H. Akaike, pp. 353–365, Springer, New York.

Lewis, F. L. (1986), *Optimal Estimation With an Introduction to Stochastic Control*, 1st ed., John Wiley, Hoboken, N. J.

Lloyd, J., and J. A. Taylor (1994), On the temperature dependence of soil respiration, *Functional Ecol.*, *8*(3), 315–323.

Luo, Y., and J. F. Reynolds (1999), Validity of extrapolating field CO_2 experiments to predict carbon sequestration in natural ecosystems, *Ecology*, *80*(5), 1568–1583.

Massman, W. J., and X. Lee (2002), Eddy covariance flux corrections and uncertainties in long term studies of carbon and energy, *Agric. For. Meteorol.*, *113*(1–4), 121–144.

Maybeck, P. S. (1979), *Stochastic Models, Estimation, and Control*, vol. 1, *Mathematics in Science and Engineering*, vol. 141-1, Elsevier, New York.

Maybeck, P. S. (1982), *Stochastic Models, Estimation, and Control*, vol. 2, *Mathematics in Science and Engineering*, vol. 141-2, Elsevier, New York.

Nelson, A. T. (2000), Nonlinear estimation and modeling of noisy time-series by dual Kalman filtering methods, Ph.D. thesis, OGI Sch. of Sci. and Eng., Oreg. Health and Sci. Univ., Portland.

Nørgaard, M., N. Poulsen, and O. Ravn (2000), New developments in state estimation for nonlinear systems, *Automatica*, *36*(11), 1627–1638.

Plummer, E. S. (1995), Training neural networks using sequential-update forms of the extended Kalman filter, *Tech. Rep. LA-UR-95-422*, Los Alamos Natl. Lab., Los Alamos, N. M.

Raupach, M. R., P. J. Rayner, D. J. Barrett, R. S. Defries, M. Heimann, D. S. Ojima, S. Quegan, and C. C. Schmullius (2005), Model-data synthesis in terrestrial carbon observation: Methods, data requirements and data uncertainty specifications, *Global Change Biol.*, *11*, 378–397, doi:10.1111/j.1365-2486.2005.00917.x.

Reich, P. B., M. B. Walters, M. G. Tjoelker, D. Vanderklein, and C. Buschena (1998), Photosynthesis and respiration rates depend on leaf and root morphology and nitrogen concentration in nine boreal tree species differing in relative growth rate, *Functional Ecol.*, *12*, 395–405.

Reichstein, M., et al. (2003), Inverse modeling of seasonal drought effects on canopy CO_2/H_2O exchange in three Mediterranean ecosystems, *J. Geophys. Res.*, *108*(D23), 4726, doi:10.1029/2003JD003430.

Richardson, A. D., and D. Y. Hollinger (2005), Statistical modeling of ecosystem respiration using eddy covariance data: Maximum likelihood parameter estimation, and Monte Carlo simulation of model and parameter uncertainty, applied to three simple models, *Agric. For. Meteorol.*, *131*, 191–208.

Richardson, A. D., et al. (2006), A multi-site analysis of uncertainty in tower-based measurements of carbon and energy fluxes, *Agric. For. Meteorol.*, in press.

Saulson, B., and K. C. Chang (2004), Comparison of nonlinear estimation for ballistic missile tracking, *SPIE Opt. Eng. J.*, *43*(6), 1424–1438.

Schulz, K., A. Jarvis, K. Beven, and H. Soegaard (2001), The predictive uncertainty of land surface fluxes in response to increasing ambient carbon dioxide, *J. Clim.*, *14*(12), 2551–2562.

Shumway, R. H., and D. S. Stoffer (1982), An approach to time series smoothing and forecasting using the EM algorithm, *J. Time Ser. Anal.*, *3*(4), 253–264.

Shumway, R. H., and D. S. Stoffer (2000), *Time Series Analysis and its Applications*, Springer, New York.

Sitz, A., U. Schwarz, J. Kurths, and H. U. Voss (2002), Estimation of parameters and unobserved components for nonlinear systems from noisy time series, *Phys. Rev. E*, *66*, 016210, doi:10.1103/PhysRevE.66.016210.

Tenne, D., and T. Singh (2003), The higher order unscented filter, paper presented at American Control Conference, Am. Autom. Control Counc., Denver, Colo.

Teskey, R. O., D. W. Sheriff, D. Y. Hollinger, and R. B. Thomas (1995), External and internal factors regulating photosynthesis, in *Physiological Ecology of Coniferous Forests: A Contemporary Synthesis*, edited by W. K. Smith and T. M. Hinckley, pp. 105–140, Elsevier, New York.

Uhlmann, J. K. (1995), Dynamic map building and localization: New theoretical foundations, Ph.D. thesis, Dep. of Eng. Sci., Univ. of Oxford, Oxford, U. K.

van der Merwe, R. (2004), Sigma-point Kalman filters for probabilistic inference in dynamic state-space models, Ph.D. thesis, OGI Sch. of Sci. and Eng., Oreg. Health and Sci. Univ., Portland, Oreg.

van der Merwe, R., and E. A. Wan (2004), Sigma-point Kalman filters for integrated navigation, paper presented at 60th Annual Meeting, Inst. of Navig., Dayton, Ohio.

VanDyke, M. C., J. L. Schwartz, and C. D. Hall (2004), Unscented Kalman filtering for spacecraft attitude state and parameter estimation, in *Spaceflight Mechanics 2004: Proceedings of the 14th AAS/AIAA Space Flight Mechanics Meeting Held February 8–12, 2004, Maui, Hawaii*, vol. 119, edited by S. L. Coffey et al., pp. 217–228, Univelt, San Diego, Calif.

van Wijk, M. T., and W. Bouten (1999), Water and carbon fluxes above European coniferous forests modeled with artificial neural networks, *Ecol. Modell.*, *120*(2–3), 181–197.

van Wijk, M. T., and W. Bouten (2002), Simulating daily and half-hourly fluxes of forest carbon dioxide and water vapor exchange with a simple model of light and water use, *Ecosystems*, *5*(6), 597–610.

Wan, E. A., and A. T. Nelson (2001), Dual extended Kalman filter methods, in *Kalman Filtering and Neural Networks*, edited by S. Haykin, pp. 123–173, John Wiley, Hoboken, N. J.

Wan, E. A., and R. van der Merwe (2001), The unscented Kalman filter, in *Kalman Filtering and Neural Networks*, edited by S. Haykin, pp. 221–280, John Wiley, Hoboken, N. J.

Wang, Q., J. Tenhunen, E. Falge, C. Bernhofer, A. Granier, and T. Vesala (2003), Simulation and scaling of temporal variation in gross primary production for coniferous and deciduous temperate forests, *Global Change Biol.*, *10*, 37–51, doi:10.1046/j.1529-8817.2003.00716.x.

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

Young, P. C. (2000), Stochastic, dynamic modelling and signal processing: Time variable and state dependent parameter estimation, in *Nonlinear and Nonstationary Signal Processing*, edited by W. J. Fitzgerald et al., pp. 74–114, Cambridge Univ. Press, New York.

J. H. Gove and D. Y. Hollinger, Northeastern Research Station, USDA Forest Service, 271 Mast Road, P.O. Box 640, Durham, NH 03824, USA. (jgove@fs.fed.us; dhollinger@fs.fed.us)