A framework fusing multiple representations of same processes from different perspectives for robust modeling of plant interaction with hydrological processes

Xu Liang[1], Liuyan Hu[2], Hector W. Clavijo[2], and Jeen-Shang Lin[2]

[1]University of Pittsburgh, Pittsburgh
[2]University of Pittsburgh

November 24, 2022

Abstract

A modeling framework is presented for hydrological modeling to more accurately describe the water, energy, and carbon cycles and their interactions with participating processes. This framework extends the modeling strategy presented in Luo et al. (2013) by simultaneously using multiple plausible expressions, derived from different perspectives, in representing the same processes, and enforcing them together with an optimality rule and a semi-empirical expression for plant CO2 uptake. The objectives are to reduce unconstrained free variables, mitigate parameter or variable equifinality, reduce result uncertainties, and ultimately increase the model robustness and predictability. For demonstration, the least cost optimality theory from Prentice et al. (2014), after extended to include water-limited conditions, is combined with the updated semi-empirical Ball-Berry-Leuning formulation (Tuzet et al., 2003). These two expressions are combined with other multiple expressions adopted for hydrological modeling. This framework is incorporated into both VIC+ and a modified DHSVM hydrological models with each applied to two different sites. Numerical studies are performed that using three approaches which only differ in the stomatal conductance modeling, namely, one uses the extended Prentice, one the semi-empirical, and the new framework that uses both. Results show that although all three approaches give reasonable estimates of limited measured fluxes, the present modeling framework gives much more reasonable estimates in the stomatal conductance and in other major model variables, and it also results in giving a relationship between carboxylation and transpiration that is consistent with observations. This modeling framework is general and can be adopted for other fields of study.

A framework fusing multiple representations of same processes from different perspectives for robust modeling of plant interaction with hydrological processes

Liuyan Hu, Hector W. Clavijo, Jeen-Shang Lin, and Xu Liang‡

Department of Civil and Environmental Engineering, University of Pittsburgh, Pittsburgh, Pennsylvania, USA.

‡Corresponding author: Xu Liang (xuliang@pitt.edu)

Key Points:

- A framework is presented that combines equally plausible but different formulations for the same processes to improve model robustness.
- This framework is applied to enhance carbon and water exchange modeling of plants and their interactions with hydrological processes.
- Benefits of this framework in reducing model uncertainty and mitigating variable equifinality are demonstrated by numerical studies.

Abstract

A modeling framework is presented for hydrological modeling to more accurately describe the water, energy, and carbon cycles and their interactions with participating processes. This framework extends the modeling strategy presented in Luo et al. (2013) by simultaneously using multiple plausible expressions, derived from different perspectives, in representing the same processes, and enforcing them together with an optimality rule and a semi-empirical expression for plant CO_2 uptake. The objectives are to reduce unconstrained free variables, mitigate parameter or variable equifinality, reduce result uncertainties, and ultimately increase the model robustness and predictability. For demonstration, the least cost optimality theory from Prentice et al. (2014), after extended to include water-limited conditions, is combined with the updated semi-empirical Ball-Berry-Leuning formulation (Tuzet et al., 2003). These two expressions are combined with other multiple expressions adopted for hydrological modeling. This framework is incorporated into both VIC+ and a modified DHSVM hydrological models with each applied to two different sites. Numerical studies are performed that using three approaches which only differ in the stomatal conductance modeling, namely, one uses the extended Prentice, one the semi-empirical, and the new framework that uses both. Results show that although all three approaches give reasonable estimates of limited measured fluxes, the present modeling framework gives much more reasonable estimates in the stomatal conductance and in other major model variables, and it also results in giving a relationship between carboxylation and transpiration that is consistent with observations. This modeling framework is general and can be adopted for other fields of study.

1. Introduction

With the advance in our understanding of the soil-plant-atmosphere continuum, more and more important processes involved have been identified and introduced into various models. Naturally, efforts have been made to add these models to the land surface or hydrological modeling system to more completely and accurately describe the water, energy and carbon cycles and their interactions with the various processes. Many of these processes have been studied from different perspectives and the resulting models have their own merits and are considered equally plausible as the current understanding could not discount one against another. But as the processes involved are complex and our understanding is incomplete, these different models describing the same processes often give divergent results. It is clear that these models are not equivalent as Feynman (1967) remarked in his talk about the character of physical law that when theories are equivalent scientifically, they give exact the same consequences. How then do we use or choose among models that describe the same process but are not equivalent, especially when these models each works well under some circumstances and not all circumstances? Feynman further commented in the same talk, “But as long as physics is incomplete, and we are trying to understand the other laws then the different formulations may give clues about what might happen in other circumstances.” That is to say that different plausible models contain information that could be complementary to one another, and we believe that finding ways to put these incomplete pieces together is one key to answer the important question of how to reconcile divergent models into better insights and solutions. Additionally, another question arises because continuing bringing in new processes inevitably makes a model ever more complex and loads it with a large number of parameters, some of which may not be independent while some may be present as free or unconstrained variables. This may very well make a model unstable, inconsistent, and intractable.

We propose to address these challenges by making use of information from available perspectives simultaneously. In cases there are many plausible models, the ones with least overlap in model construct are selected. That is, we will incorporate simultaneously different plausible formulations or models of the same processes in the form of mutual constraints. The immediate consequence of doing so will lead to reduction in the number of free variables.

Solutions of the processes will also become more robust, and the solutions will be less uncertain and more versatile. We will show how this is done.

This philosophy has been successfully executed by Luo et al. (2016; 2013) into VIC+ in modeling the transpiration and carbon assimilation together with other hydrological processes such as hydraulic redistribution and its interactions with the groundwater table movement dynamics. In that work, transpiration is simultaneously considered by (1) the method of Ohm's law analogy and (2) the method of Penman-Monteith equation. For the former, transpiration is estimated based on the soil water potentials in the root zone, in leaves, and of plant storage; the hydraulic resistance from the soil to leaves, and between plant storage and leaves. For the latter, the calculation of plant transpiration is not only driven by the meteorological factors but is also directly linked to the carbon assimilation through the stomatal conductance. The carbon assimilation process involved in the plant transpiration also incorporates simultaneously two perspectives: (1) a diffusion method, and (2) the modified Farquhar biochemical model. That is, the calculation of carbon assimilation is constrained as a consequence of the interplay of the stomatal and biochemical limitations simultaneously.

In this study, we extend this modeling strategy to model leaf stomata. Plants play a pivotal role in the soil-plant-atmosphere system, in which leaf stomata is a key in balancing photosynthesis and transpiration (Bauerle & Bowden, 2011). Specifically, we simultaneously consider two drastic different modeling approaches: one is based on an optimality principle, while the other semi-empirical. The latter is used as an additional constraint to the former in the implementation. To illustrate versatility of this extended modeling strategy in improving robustness of the modeling results, we further implement them separately with two hydrological models that have very different modeling structures: the VIC+ land surface model (Luo et al., 2016; 2013) for large scale system and a modified high resolution version of the Distributed Hydrological Soil Vegetation Model (DHSVM) (Wigmosta et al., 2002; 1994) for small scale system in conducting numerical studies.

By controlling leaf stomata, plants exchange carbon dioxide and water with the atmosphere, which can regulate water loss and adapt to external CO_2 concentration by taking advantage of the biochemical and hydrological processes (Berry et al., 2010). Stomatal conductance – governing plant behavior to water stress condition and photosynthesis – is one of

the essential components affecting water and carbon exchange process of plants.

Current approaches in modeling stomatal conductance can be broadly classified into four categories (Damour et al., 2010; Miner et al., 2017). The first uses an empirical formulation, such as the Jarvis type (Jarvis, 1976), to relate the stomatal conductance in a multiplicative form to contributing factors such as solar radiation, air temperature, leaf water potential, vapor pressure, and CO_2 concentration. The second uses a semi-empirical formulation to connect the stomatal conductance to carbon assimilation (Ball et al., 1987; Collatz et al., 1992; Leuning, 1990, 1995; Tuzet et al., 2003). One widely adopted approach of this category is the simple semi-empirical model based on the Ball-Berry-Leuning (BBL) formulation. This model was first developed by Ball et al. (1987). Leuning et al. (1998) modified it to include the soil water content. Tuzet et al. (2003) further proposed a variant form, denoted as BBL-update in this study, which accounts for the leaf water potential and thus connects the root-soil-atmospheric water transfer together. The third category employs a mechanistic-based water stress response model that combines hydraulic control and abscisic acid (ABA) (Gutschick & Simonneau, 2002; Tardieu & Davies, 1993). Models related to this third category are far more complex than the first two and are not widely adopted. The fourth category employs an optimality theory. Models so developed not only generally involve fewer parameters, but also better represent plants' natural responses to the environment (Franklin et al., 2012). The optimality theory can provide certain internal correlations among different components within a complex system (Schymanski et al., 2009; Westhoff et al., 2014), and thereby reduces the number of parameters need to be estimated. Currently, the semi-empirical approach and the optimality approach are the most widely used methods for stomatal conductance modeling.

We classify the current optimality approach into three main groups following Dewar et al. (2018), which is slightly different from the classifications by Wang et al. (2020) or by Sabot et al. (2020). These three groups are: (1) maximizing carbon gain while minimizing the total loss of water over a given time period; (2) maximizing net carbon gain at every instant in time; and (3) maximizing multiple benefits for photosynthesis while minimizing associated costs at the same time.

The first group includes models applying the water use efficiency (WUE) hypothesis – the long-standing plant optimality rule by Cowan and Farquhar (Cowan, 1982; Cowan &

Farquhar, 1977) and its variants (e.g., (Lu et al., 2016; MÄKELÄ et al., 1996). Cowan and Farquhar (1977) showed that such an optimality rule leads to an optimization constraint, $\lambda=\partial A_n/\partial E_{tr}$, with λ representing a key rate of how carbon assimilation, A_n, responds to transpiration, E_{tr}. This λ has been investigated for many years, yet there is still no consensus on how to determine it under different conditions (Buckley et al., 2017; Wolf et al., 2016). Medlyn et al. (2011) followed Cowan and Farquhar optimality rule but optimized RuBP regeneration-limited photosynthesis rather than Rubisco-limited photosynthesis. Their study showed that the stomatal expression by Cowan and Farquhar's optimality was similar in the form to the BBL's semi-empirical expression when the atmospheric CO_2 concentration at leaf surface was much higher than the compensation point. Other efforts have been made to extend the Cowan and Farquhar's optimality rule in determining λ or by adding additional factors (e.g., (Katul et al., 2010; 2009; Manzoni et al., 2013). A new way to solve λ is presented later in this work.

The second group maximizes net carbon gain and includes two subgroups. One subgroup applies a penalty function to plant hydraulic behavior (Anderegg et al., 2018; Eller et al., 2018; Sperry et al., 2017; Wolf et al., 2016), while the other applies a penalty function to nonstomatal limitation (NSL) behavior, such as the carboxylation capacity (CAP) and mesophyll conductance (MES) models by Dewar et al. (2018) and the model by Hölttä et al. (2017). Mathematically, these two subgroups are similar and both maximized the carbon gain function (Dewar et al., 2018; Wang et al., 2020). Biologically, CAP and MES models would produce lower photosynthetic rate for the same leaf transpiration rate since more reduction factor is introduced into the conductance as noted in Wang et al. (2020).

The third group can be viewed as a step toward a more ideal optimality rule which would optimize over a broader base since all sources of benefits for photosynthesis are maximized, including nitrogen, light, water, while all the associated costs are minimized at the same time over multiple temporal scales as pointed out by Buckley (2017). Models belong to this group include those by Manzoni et al. (2013), Prentice et al. (2014) and Buckley et al. (2017). In the following, Prentice et al. (2014) model, called Prentice-2014 hereafter, is adopted as representative of this third group. In Prentice-2014, the optimality rule balances the tradeoff between transpiration and carboxylation capacity by minimizing the summed costs of transpiration and carboxylation. Prentice-2014 employs the coordination hypothesis that under

typical daytime conditions, when most photosynthesis takes place, its Rubisco-limited photosynthetic rate is equal to electron transport-limited photosynthetic rate. This model is realized by adjusting the ratio of CO_2 within the leaf to that outside the leaf. Its limitations are that it ignores the role of leaf water potential and plant hydraulics in the stomatal opening (Buckley et al., 2017; Dewar et al., 2018), and thus it works best under wet conditions (Prentice et al., 2014).

Among these three main groups of optimality rules, one main difference between the WUE group (first group) and the second group is that in the former multiple factors affecting the reduction of transpiration are lumped together as a total water loss through the transpiration, while in the latter the reduction of transpiration is attributed specifically to either the hydraulic factors, such as xylem water potential and canopy xylem pressure, or to the hydraulic and NSL factors together. As a result the largest differences between them occur during the dry conditions (Anderegg et al., 2018). Anderegg et al. (2018) added the water stress to the WUE group but with mixed results: For 7 species among the 41 studied, the results from the WUE group using the original λ rate provided comparable results with observations; while for the other 34 species, results of the WUE group with the rate λ being modified by soil water potential matched observations better. For all of the 41 species compared, however, they showed that results from the Wolf-Anderegg-Pacala model – of the first subgroup in the second group – led to better results than the WUE group with higher R^2 values compared to the observations. Wolf et al. (2016) showed that under some special conditions where a closed form for the stomatal conductance can be obtained, the optimal stomatal conductance using the optimality rule for the first subgroup of the second group is remarkably similar to the semi-empirical formulation of BBL, while Medlyn et al. (2011) have previously demonstrated that the stomatal conductance expression based on the optimality rule for the WUE group is the same as BBL's when the atmospheric CO_2 concentration at leaf surface was much higher than the compensation point. Dewar et al. (2018) compared models using each of the three different groups of the optimization rules and showed that the Prentice-2014 model (third group) produced similar results to those by the CAP and MES models (second subgroup of the second group), but results from the WUE group (first group) were different from the other two groups.

198 However, Dewar et al. (2018) also showed that CAP and Prentice-2014, as well as MES
199 and WUE (Medlyn et al., 2011) at low and large atmospheric CO_2 concentration, respectively, all
200 lead to the same one-parameter relationship between the ratio related to leaf CO_2 concentration
201 and vapor pressure deficit. The key difference among these three groups of the optimality models
202 is how each model estimates this one-parameter by its own optimization rule and the atmospheric
203 CO_2 concentration range considered. It is clear that considerable similarity exists among these
204 three different groups in terms of the functional forms in their CO_2 stomatal conductance
205 expression, although each of them is derived by associating itself with a different optimality rule.
206 Comparisons of the performance of these three different optimality modeling groups can be
207 found in Dewar et al. (2018), Anderegg et al. (2018), and Wang et al. (2020). Basically, the main
208 challenges of these approaches lie in how to define the penalty function associated with stomatal
209 opening used to balance the carbon gain and loss of water for plants under drought conditions.
210 There is yet no consensus on how this is best done.

211 In this study, the semi-empirical approach and the optimality-based approach are
212 considered equally plausible for stomatal conductance modeling because one cannot claim more
213 merits over the other with our current understanding. The framework presented in this paper
214 provides a rational way for stomatal conductance modeling by simultaneously incorporating both
215 to represent the relevant processes where appropriate. There is no reason to assume that one
216 optimization rule would work for all different processes under various conditions, as there is a
217 plethora of natural processes involved and contribute to the complex behavior of plants. But
218 because optimality-based formulations overlap and, under some scenarios, are identical as
219 discussed, we thus use only one among them. Specifically, we employ Prentice-2014 for two
220 main reasons. Firstly, it has been tested with measurements from various natural conditions and
221 experimental settings, and shown to correctly predict a number of related physiological
222 characteristics, such as the global pattern of the maximum carboxylation rate, V_{cmax}, in relation
223 to light, temperature and vapor pressure deficit (Smith et al., 2019), seasonal variations of V_{cmax}
224 across diverse ecosystems (Jiang et al., 2020), elevational trends in photosynthetic traits and
225 primary production (Peng et al., 2020), the trends in the ratio of leaf-internal to ambient CO_2
226 with respect to mean growth temperature, vapor pressure deficit, atmospheric CO_2, and elevation
227 (Wang et al., 2017). Secondly, compared with other optimality-based approaches, it has fewer
228 parameters need to be estimated and its parameters are more robust.

It is important to emphasize that optimization approaches should be implemented with consideration of boundary conditions (Buckley et al., 2017). This is to account for factors that (1) plants adjust their functional behaviors based not only on the resources provided by the environment and their own maximum capabilities, but also on the fact that their physical or biological properties have bounds; and (2) our descriptions and understanding of the complex eco-biological processes involved are incomplete and limited, and constraints are imposed in our optimization search. For example, under a drought condition, not only is the availability of water to plants limited, reflected by leaf water potential, but also the physical size of the minimum and maximum stomatal opening may be limited as well, leading to a constrained optimal photosynthesis process. Also, due to the complex processes involved in plants' responses to drought, our current descriptions or representations of the drought processes are likely incomplete, leading to solutions outside feasible ranges if no bounds are imposed. Therefore, upper and lower bounds posed by plant physiology should be included whenever appropriate.

Constraints and boundary conditions play a central role in fusing different perspectives of the same process (Luo et al., 2016; 2013). Constrains limit the degrees of freedom of a model caused by the large number of model parameters and their interactions. Due to the complex water transport and photosynthesis process, a large number of free parameters still remains despite the use of an optimality theory. A potential serious consequence may emerge in that similar model responses are obtained with different and even unrealistic combinations of parameters. This phenomenon is referred to as equifinality of parameter sets (Beven, 2006). Equifinality is especially pronounced when (1) the number of parameters involved is large and the available observations that can be simultaneously used to determine the parameter values are small; and (2) there are substantial errors in the data and in the model structures. Beside the presence of multiple sources of errors, one essential dominant factor leading to the equifinality pitfall is the lack of constraints (Sun et al., 2020). Equifinality is pronounced in ill-posed inverse problems which have insufficient constraints. In this study, we extend the equifinality description to variables. That is, if similar model responses are obtained with different and unrealistic combinations of values of model variables, we refer to this phenomenon as equifinality of variables. Introducing constraints based on plant physiology is an effective and rational way to reduce the "free" model variables, and thus the degree of the model's uncertainties (Prentice et al., 2015). In a broader sense the simultaneous representation of the same process using multiple

expressions from different perspectives is a form of imposing constraints. The idea behind this strategy, as stated earlier, is that each equally or quasi-equally plausible expression describes one perspective of our understanding of the whole process. Since these expressions are not equal, it implies that each of the different views gives different pieces of incomplete information about the process. When brought in together, as each perspective tells one another what the process should be in a specific view, they thus mutually constrain one another into a coherent and more complete picture. As long as all these different expressions are equally or quasi-equally plausible and not excessively overlapped, then using them at the same time would more accurately describe the reality and fill the gaps that other perspectives leave. This is the modeling framework we present in this study. With this approach, not only can one more accurately represent the plants' behaviors under different conditions, but also reduce model's uncertainties due to the removal of a large number of free variables. It is important that one is not inadvertently introducing more uncertainties when adding new expressions or constraints to a model trying to reduce the model's free variables. Therefore, one should always balance and weigh the new expressions against the existing knowledge in assessing their relevant parameters and associated uncertainties, robustness and reliability so that indeed more rational constraints, rather than more uncertainties, are added.

In this study, we simultaneously employ two formulations to represent the stomatal conductance behaviors, an optimality rule of Prentice-2014 from the third group and a semi-empirical expression of Tuzet et al. (2003), or BBL-update, to illustrate the philosophy of our framework. We choose these two even though they have identical forms of stomatal conductance, however, the slopes of their expressions are different, and that together they cover the current understanding better than other combinations. For the Prentice-2014 model, we further first extend it so that it is applicable to both wet and dry conditions, and also to conditions when the original coordination hypothesis of Rubisco-limited photosynthetic rate being equal to the electron transport-limited photosynthetic rate does not hold. We also follow Luo et al. (2013) and use both the Ohm's law analogy and the Penman-Monteith method to simultaneously represent the transpiration, and employ both the diffusion method and biochemical model to represent the carbon assimilation at the same time.

We represent the plant hydraulics considering the leaf water potential dynamics and multiple stomatal conductance formulations among transpiration, photosynthesis, and carbon assimilation. This is different from recent developments reviewed in Wang et al. (2020) in which the plant hydraulics was only associated with one optimality rule in presenting the water stress factor. The simultaneous application of an optimality rule with a semi-empirical stomatal conductance formulation used in this study is unique and different from the previous efforts (De Kauwe et al., 2015; Heroult et al., 2013; Manzoni et al., 2013).

The reminder of this paper is organized as follows: Section 2 describes the methodology of our modeling framework and the underlying insights. Section 3 presents the implementation of the modeling framework. Section 4 presents the results and analyses of the results between our modeling framework and two other approaches with two hydrological models at four locations. Conclusions are provided in Section 5.

2. Modeling Framework: Philosophy and Construct

The objectives of our modeling framework are to combine current understandings in advancing modeling capability and reducing model uncertainties which are achieved through the following actions: First, identify processes that have different formulations but are equally or quasi-equally plausible for each activity or task of the model. In the case of describing the plant stomatal behavior, for example, there are three activities involved: photosynthesis, transpiration and carbon assimilation. Second, simultaneously combine these different formulations for each of the process identified. Third, impose boundary conditions where appropriate. Finally, solve these resulting coupled expressions.

The ideas and procedures of the modeling framework are explained herein in terms of the incorporation of the number of modeling variables related to the photosynthesis and plant transpiration processes through hydrological modeling. Conventionally, this part is formulated as a five-variable problem (Anderegg et al., 2018), and the five variables are typically chosen to be CO_2 stomatal conductance (g_{s,co_2}), leaf water potential (ψ_l), plant transpiration (E_{tr}), leaf CO_2 concentration (c_i), and carbon assimilation (A_n). Corresponding to them, our conventional

approach uses the following five equations: one optimality equation, two equations for transpiration, and two equations for carbon assimilation. As for the current framework, we have, however, six equations since we use two equations for stomatal conductance instead of one. This enables us to solve the posed problem with six variables that better represents the three activities (i.e., photosynthesis, transpiration and carbon assimilation) of the stomatal behavior. The selection of the additional variable is discussed below. For our new approach, we first extend the Prentice-2014 optimality model.

2.1 The first model for CO_2 stomatal conductance -- extended Prentice-2014 optimality model

As Prentice-2014 was originally developed for wet conditions, it is not expected to perform well under water-limited conditions. In this study, Prentice-2014 is extended to overcome this deficiency. In addition, the coordination hypothesis made in Prentice-2014 (Prentice et al., 2014; Wang et al., 2017) between Rubisco-limitation and electron transport-limitation is not required.

The Prentice-2014 model minimizes the total summed cost of carboxylation and transpiration as follows,

$$\text{Min } Cost = a \cdot E_{tr}^{'} / A_n' + b \cdot V_{cmax} / A_n' \quad (1)$$

where a is the unit cost transpiration parameter; b is the unit cost carboxylation parameter; $E_{tr}^{'}$ [mol·m^{-2}·s^{-1}] is the transpiration; $A_n^{'}$ [mol·m^{-2}·s^{-1}] is net carbon assimilation; and V_{cmax} [mol·m^{-2}·s^{-1}] is the maximum carboxylation rate. $E_{tr}^{'}$ is calculated as follows,

$$E_{tr}^{'} = 1.6 \cdot g'_{s,co_2} \cdot D' \quad (2)$$

where g'_{s,co_2} [mol·m^{-2}·s^{-1}] is the CO_2 stomatal conductance; and D' [Pa·Pa^{-1}] is the normalized leaf-to-air vapor pressure deficit calculated by $D^{'} = [e_{sat}(T_l) - e(T_a)] / p_c$, with $e_{sat}(T_l)$ being the saturated vapor pressure at the leaf temperature, T_l, and $e(T_a)$ the actual vapor pressure at the air temperature, T_a, and p_c is the surface air pressure. V_{cmax} [mol·m^{-2}·s^{-1}] depends on T_l.

Furthermore, in Prentice-2014, the Rubisco activity considered in its carbon assimilation,

through the coordination hypothesis, is related only to CO_2 concentration within the leaf (c_i) [mol·mol^{-1}], and not to leaf water potential (ψ_l) [MPa]. For the transpiration (E'_{tr}), its g'_{s,co_2} does not include leaf water potential (ψ_l) either. Thus, Eq. (1) involves only four unknows: E'_{tr}, g'_{s,co_2}, A'_n, and c_i, and they are obtained by solving Eqs. (1) and (2), together with Eqs. (3) and (4) below:

$$A'_n = g'_{s,co_2}(c_a - c_i) \tag{3}$$

$$A'_n = V_{cmax}\frac{c_i - \Gamma^*}{(c_i + K)} \tag{4}$$

where c_a is the leaf ambient mole fractions of CO_2; K is the Michaelis–Menten coefficient for Rubisco-limited photosynthesis at a pO_2 (partial pressure of oxygen) of 21 kPa.; and Γ^* [mol·mol^{-1}] is the CO_2 compensation point which also depends on T_l.

It is noted here that following Dewar et al. (2018), the c_a term in g'_{s,co_2} from Prentice-2014 can be replaced by c_i and gives,

$$g'_{s,co_2} = \frac{\xi}{\sqrt{D'}}\frac{A'_n}{c_i - \Gamma^*} \tag{5}$$

where ξ is defined by

$$\frac{c_i - \Gamma^*}{c_a - \Gamma^*} = \frac{\xi}{\xi + \sqrt{D'}} \tag{6}$$

It will become clear later that g'_{s,co_2} of Prentice-2014 (Dewar et al., 2018) which follow Eq. (5) is clearly different from that of the BBL-updated (Tuzet et al., 2003), even though their forms are the same.

We extended the preceding carbon assimilation to consider leaf water potential (ψ_l) and also relax the original coordination hypothesis that Rubisco-limitation be equal to electron transport-limitation by employing a modified Farquhar model (Farquhar et al., 1980; Daly et al., 2004). In addition, we include ψ_l in the g'_{s,co_2} calculation. By doing so, Eq. (1) is extended and become applicable to water-limited conditions, and that g'_{s,co_2} and A'_n are modified to account for

the leaf water potential (ψ_l) and other factors affecting carbon assimilation. To avoid confusion, the modified equation uses E_{tr} and A_n without prime for transpiration and carbon assimilation, respectively, and Eq. (1) becomes,

$$\text{Min } Cost = a \cdot E_{tr} / A_n + b \cdot V_{cmax} / A_n \tag{7a}$$

Following Prentice-2014 by taking the derivative of cost with respect to c_i, Eq. (7a) at optimum is given by,

$$\frac{dCost}{dc_i} = a \cdot \frac{\partial\left(\frac{E_{tr}}{A_n}\right)}{\partial c_i} + b \cdot \frac{\partial\left(\frac{V_{cmax}}{A_n}\right)}{\partial c_i} = 0 \tag{7b}$$

where, in this new cost equation, both E_{tr} and A_n are functions of leaf water potential (ψ_l) as stated, and ψ_l becomes the fifth unknown variable.

As stated, the two perspectives of transpiration are from the Penman-Monteith equation and the Ohm's law analogy, following the approach of Luo et al. (2013). They are listed below as Eqs. (8) and (9), respectively:

$$E_{tr} = \frac{\Delta(R_n - G) + \rho_a C_p D \overline{g_a}}{\rho_w \lambda_w \left(\Delta + \gamma_w + \frac{\gamma_w \overline{g_a}}{LAI\, g_s}\right)} \tag{8}$$

where E_{tr} [m·s^{-1}] is the transpiration; Δ [Pa·K^{-1}] is the rate of change of saturation vapor pressure with air temperature; R_n [W·m^{-2}] is the net radiation; G [W·m^{-2}] is the ground heat flux; ρ_a [kg·m^{-3}] is the air density; C_p [J·kg^{-1}·K^{-1}] is the specific heat capacity of air; D [Pa] is the vapor pressure deficit and $D = e_{sat}(T_a) - e(T_a)$; $\overline{g}_a$ [m·s^{-1}] is the conductance of the atmospheric boundary layer to H_2O (per unit ground area); ρ_w [kg·m^{-3}] is the water density; λ_w [J·kg^{-1}] is latent heat of water vaporization; γ_w [Pa·K^{-1}] is psychrometric constant; LAI is the leaf area index; and g_s [m·s^{-1}] is the stomatal conductance to H_2O per unit leaf area. Note that the H_2O stomatal conductance expressed as g_{s,H_2O} [mol·m^{-2}·s^{-1}] is the same as stomatal conductance g_s [$m\ s^{-1}$] but with a

different unit, and g_{s,H_2O} is equal to 1.6· g_{s,co_2} when both take the unit of [mol·m^{-2}·s^{-1}].

Equation (9) is expressed as,

$$E_{tr} = \frac{\psi_p - \psi_l}{r} + \frac{\psi_{soil} - \psi_l}{R} \quad (9)$$

where ψ_p[Pa] is the water potential of plant storage; ψ_l [Pa] is the leaf water potential; ψ_{soil} [Pa] is the lumped soil water potential in the root zone; r [Pa·s·m^{-1}] is the hydraulic resistance between plant storage and leaves; R [Pa·s·m^{-1}] is the total hydraulic resistance, a function of ψ_{soil}, from the soil to the leaves. Further details about Eq. (9) can be found in Luo et al. (2013).

The carbon assimilation (A_n) which has appeared in Eq. (7) is hereby formulated using both the modified Farquhar model (e.g., (Daly et al., 2004; Farquhar et al., 1980) and the diffusion method which are represented by Eqs. (10) and (11), respectively, as follows,

$$A_n = A_{\psi_l}(\psi_l) \times A_{\phi,c_i,T_l}(\phi, c_i, T_l) \quad (10)$$

$$A_n = g_{sba,co_2} \cdot (c_a - c_i) \quad (11)$$

where A_n [mol·m^{-2}·s^{-1}] is carbon assimilation; and in Eq. (10), $A_{\psi_l}(\psi_l)$ is a function related to leaf water potential, reflecting the reduction of carbon assimilation under water stressed conditions. $A_{\phi,c_i,T_l}(\phi, c_i, T_l)$ is the Farquhar model of biochemical carbon assimilation under well-watered condition, and it depends on photosynthetically active radiation (ϕ), CO_2 concentration (c_i) within the leaf, and leaf temperature (T_l).

The term $A_{\phi,c_i,T_l}(\phi, c_i, T_l)$ is expressed by the minimum of A_c and A_q, where A_c is the assimilation rate restricted by Rubisco activity (i.e., restricted by c_i), and A_q is the assimilation rate limited by RuBP regeneration when ϕ is low. The triose phosphate utilization (TPU) limitation and the quadratically smooth transition approach used in Daly et al. (2004) to obtain the minimum A_n from Collatz et al. (1991) are not used here as the smooth approach and TPU limitation lead to underestimation of the A_n (Rogers et al., 2021). The relevant equations for

theses parameters are:

$$A_{\psi_l}(\psi_l)=\begin{cases} 0 & (\psi_l<\psi_{l_{A0}}) \\ \frac{\psi_l-\psi_{l_{A0}}}{\psi_{l_{A1}}-\psi_{l_{A0}}} & (\psi_{l_{A0}}<\psi_l<\psi_{l_{A1}}) \\ 1 & (\psi_l>\psi_{l_{A1}}) \end{cases} \tag{12}$$

where ψ_l [Pa] is the leaf water potential; $\psi_{l_{A1}}$[Pa] is the leaf water potential value in well-watered condition; $\psi_{l_{A0}}$[Pa] is the leaf water potential value below which assimilation is reduced to zero.

$$A_c=V_{cmax}\frac{c_i-\Gamma^*}{c_i+K_c(1+o_i/K_o)} \tag{13}$$

where o_i [mol·mol^{-1}] is the oxygen concentration; K_c and K_o are the Michaelis-Menten coefficients for CO_2 and O_2, respectively, which depends on T_l.

$$A_q=\frac{J\cdot(c_i-\Gamma^*)}{4(c_i+2\Gamma^*)} \tag{14}$$

where J [mol·m^{-2}·s^{-1}] is the electron transport rate which depends on ϕ and T_l.

As for the diffusion method of Eq. (11), g_{sba,co_2} in Eq. (11) is represented by

$$g_{sba,co_2}(\psi_l)=\left(g_{s,co_2}^{-1}(\psi_l)+g_{a,co_2}^{-1}+g_{b,co_2}^{-1}\right)^{-1} \tag{15}$$

where g_{a,co_2} [mol·m^{-2}·s^{-1}] is the atmospheric conductance, and g_{b,co_2} [mol·m^{-2}·s^{-1}] is the CO_2 leaf boundary layer conductance. We note that $g_{s,co_2}(\psi_l)$ is related to leaf water potential, while g'_{s,co_2} in Eqs. (2) and (3) is not since g'_{s,co_2} is only for not water-limited condition.

This extended least cost optimality of Prentice-2014, denoted as *LC-extended* (for extended Least Cost) hereafter, is extended to water-limited conditions and to the situations where the coordination hypothesis on having $A_n = A_c = A_q$ is relaxed, allows the five unknows of E_{tr}, g_{s,co_2}, A_n, c_i, and ψ_l to be obtained by solving Eqs. (7)-(11) together. *LC-extended* as presented is shown works under both wet and water-limited conditions in the numerical study

section.

2.2 The second model for CO_2 stomatal conductance – BBL-updated formulation

Our second perspective uses BBL-updated and we make use of the fact that g_{s,co_2} is explicitly related to the CO_2 assimilation A_n, c_i, and the empirical function of $f(\psi_l)$. This BBL-updated relationship is expressed as follows,

$$g_{s,co_2} = g_0 + \frac{a' A_n}{c_i - \Gamma^*} \cdot f(\psi_l) \tag{16}$$

where g_0 [mol·m^{-2}·s^{-1}] is the stomatal conductance at the light compensation point; a' is an empirical slope coefficient which varies between [0, a'_{max}] where a'_{max} is the upper bound of a', Γ^* [mol·mol^{-1}] is the CO_2 compensation point; and $f(\psi_l)$ is an empirical function of stomatal sensitivity to leaf water potential which varies between [0, 1]. Eq. (16) is widely used and validated with observations for obtaining plant CO_2 stomatal conductance.

Eq. (16) has a slope of $a' f(\psi_l)$ which is different from the slope of $\frac{\xi}{\sqrt{D'}}$ given by Eq. (5). The slope $a' f(\psi_l)$ varies between [0, a'_{max}], while the slope of $\frac{\xi}{\sqrt{D'}}$ ranges between [0, +∞) in principle as D' might approach zero. This is what we mean that the *Semi-empirical* approach, i.e., BBL-updated, is different from the Prentice-2014, and for that matter, different from the extended Prentice et al. (2014). To implement our philosophy of combining different perspectives in our new approach, *LC-extended* and Eq. (16) of BBL-update are simultaneously considered.

That is, by adding Eq. (16), the problem now has six equations. For that, we also choose a' as the sixth variable. Does the choice of a' as a new unknown make sense? Value of a' has been assumed a constant for a given vegetation type and is predetermined via model calibration in practice. But Miner et al. (2017) have showed that a' changes under elevated CO_2 and water stressed conditions, thus, its selection here as an unknown time-varying variable incorporates their findings.

It can be shown that a' has an upper limit, a'_{max}, that is vegetation type dependent related

to the slope in the original BBL model. This slope can be obtained using field measurements. Miner et al. (2017) provided a summary table for the slope values of the original BBL model for different plant species. Thus, the maximum slope value of $a^{'}$ (i.e., $a^{'}_{max}$) in Eq. (16) can be derived through the relationship between the original BBL model and the BBL-updated model. In this way, $a^{'}$ varies over time according to the changes in water stress and CO_2 concentration level. Such a treatment on $a^{'}$ is more consistent with the observations by Miner et al. (2017) as stated.

When comparing Eq. (16) with the model of Medlyn et al., (2011), one can obtain the relationship between a' and g_1 below,

$$g_1 = \left(\frac{a' c_a}{c_i - \Gamma^*} \cdot f(\psi_l) - 1\right) \cdot \sqrt{D' * p_c} \tag{17}$$

where g_1 is the slope value in Medlyn et al. (2011) and can be computed from a'.

From g_1, one can easily obtain the values of λ. With Eq. (17), g_1 or λ associated with the WUE approach can be estimated at each time step via $a^{'}$. With this, we may have indirectly addressed the longstanding problem to some extent of how to represent the varying nature of λ at all time scales based on the WUE optimality rule as demonstrated later.

3. Numerical Studies

To investigate the effectiveness and implications of the present framework, a series of numerical studies are conducted with two hydrological models. Two hydrological models are used simply to show that the framework works in different hydrological scales. The numerical study explores three approaches which differ in how stomatal conductance is considered.

3.1 Three approaches

The three approaches are deigned as follows:

Approach 1: This, we denote as the "*New Approach*", is an implementation of the core of the present framework using two formulations for stomatal conductance as detailed in Sections 2.1 and 2.2, in which six unknows, E_{tr}, g_{s,co_2}, A_n, c_i, ψ_l, and a', are solved together (Figure 1a).

Approach 2: This uses LC-extended expression for stomatal conductance modeling, and that results in five unknows, E_{tr}, g_{s,co_2}, A_n, c_i and ψ_l with five equations (Figure 1b). We call this approach "*LC-extended*", of which a' is back-calculated from Eq. (16) after the five unknowns are solved and it is not a constant.

Approach 3: This uses the BBL-updated expression, Eq. (16), for stomatal conductance modeling and is referred to as the "*Semi-empirical*" approach. The parameter a' is treated as a constant with its values determined during model calibration. This also has five unknowns, E_{tr}, g_{s,co_2}, A_n, c_i, and ψ_l, to solve based on Eqs. (8)-(11) and (16).

With this setup, the *New Approach*, can be compared with the currently widely adopted methods of *Semi-empirical* and *LC-extended* approaches which work under water stressed conditions and for the latter the Prentice-2014 coordination hypothesis of $A_n = A_c = A_q$ is relaxed as well.

3.2 An implementation

With *New Approach* of the present modeling framework, we solve the six unknows of E_{tr}, g_{s,co_2}, A_n, c_i, ψ_l, and a' by six equations, i.e., Eqs. (7)-(11) and Eq. (16). We have assigned ranges of search during optimization computation for g_{s,co_2} and a' in the following bounds,

$$0 \leq g_{s,co_2} \leq g_{s,co_2 max} \tag{18}$$

and

$$0 \leq a' \leq a'_{max} \tag{19}$$

The upper bound of g_{s,co_2} is taken as $g_{s,co_2 max} = 0.5$ [mol·m^{-2}·s^{-1}] which is based on the general observed physical maximum value for all of the vegetation types based on the literature

reported in Nobel (1999). The value of a'_{max} depends on the type of vegetation. Slopes in the original BBL model that a' is related to vary between 2 and 250 (Miner et al., 2017). In this study, a'_{max} are obtained through regression between the original BBL and the BBL-updated for different vegetation types (Clavijo Sanabria, 2020).

For the present *New Approach*, the problem to be solved is thus a constrained optimization problem. There are multiple ways of solving these six unknows, Figure 1a describes one way of solving them. As for *LC-extended*, they are solved with Eqs. (7)-(11) as described in Figure 1b. Procedure of solving the five unknowns for the *Semi-empirical* is achieved by simply keeping *a'* constant following Luo et al. (2013).

For the optimality-based *New Approach*, the minimum cost criterion, $dCost/dc_i$, could be obtained via some optimization schemes but here we use an exhaustive search method which searches all possible c_i values at a regular interval to ascertain the global minimum is reached for in-depth discussion. The solution procedure starts with known soil water potential, ψ_{soil} , at the time step *t=1* in Figure 1a. The main steps are briefly described below:

Step 1: Follow a sequential order and use the one in the queue of the equally spaced c_i for the current time step.

Step 2: Try a value of g_{s,co_2} from the range based on Eq. (18). Calculate g_s, E_{tr}, then ψ_l.

Step 3: Calculate A_n based on Eqs. (11) and (15) and express it as A_{n1}. Calculate the CO_2 assimilation from the modified Farquhar model, i.e., Eqs. (10) and (12) – (14), and express it as A_{n2} (Figure 1a).

Step 4: Check the difference between A_{n1} and A_{n2}. If their difference, $|(A_{n2}- A_{n1})/A_{n1}|$, is greater than the threshold 10^{-10}, go back to Step 2 and select a new value for g_{s,co_2} and repeat until this threshold is met. If the tolerance is not met after the maximum number of iterations is reached, go back to Step 1 and select the next value of c_i.

Step 5: Calculate a' based on Eq. (16). If $a' > a'_{max}$ and the maximum number of iterations prescribed is not reached, go back to Step 1 and select the next value of c_i and start over. If the maximum number of iterations is reached, then stop. This study has not

encountered this scenario, however. If $a' \le a'_{max}$, calculate and store the cost defined by Eq. (7a).

Step 6: Proceed to the next c_i and repeat Step 2 – Step 5 until all the equally spaced c_i between Γ^* and c_a are checked.

The c_i that produces the least cost, together with its associated E_{tr}, g_{s,co_2}, A_n, ψ_l, and a' are the solution for the current time step.

When processes are considered simultaneously from different perspectives, they are coupled through shared variables. There is less "freedom" for these shared variables as they must conformed to different perspectives and thereby reduces the model uncertainty. In addition, this modeling framework can reduce the required number of model parameters that need to be calibrated. A case in point: here we solve ψ_l and c_i together with the leaf water potential (ψ_l) shared in Eqs. (9), (10), (12), (15), and (16), and CO_2 concentration within leaf (c_i) in Eqs. (7), (10), (11), (13), (14), and (16).

3.3 Two hydrological models used

Two hydrological models, VIC+ and DHSVM, that have significantly different model structures are employed and each is applied to two different locations to investigate the versatility and benefits of our modeling framework.

VIC+ (Luo et al., 2016; 2013) extends the Three-Layer Variable Infiltration Capacity (VIC-3L) large-scale hydrological model (Liang et al., 1994,1996a,1996b, 2003; Liang & Xie, 2001, 2003) with important new features. The enhancement in VIC+ are as follows: First, VIC+ considers hydraulic redistribution (HR) process and its effect on the interplay between plant transpiration and groundwater dynamics under water-limited conditions. Second, it explicitly represents groundwater table movement within the soil column and its tight interactions with the HR process. Third, it explicitly represents the photosynthesis process and its interactions with transpiration process. Fourth, it introduces our strategy into hydrological modeling by simultaneously representing the same process using multiple expressions from different perspectives to constrain the model. Fifth, VIC+ considers impact of plant storage on the water,

energy, and CO_2 cycles. In addition to these new features introduced to it, VIC+ also maintains the original unique features included in the VIC-3L model (Liang et al., 1994, 1996a, 1996b, 2003, 2004; Liang and Xie, 2001, 2003; Cherkauer and Lettenmaier, 1999, 2003), such as considering subgrid spatial variability of soil and vegetation properties and precipitation, accounting for both infiltration and saturation excess runoff generation mechanisms for each modeling grid in an interactive way under the context of subgrid spatial variability associated with watershed properties (Liang and Xie, 2001; 2003). These original VIC features make the VIC model more robust and less scale dependent as compared to other land surface models as illustrated by different studies (e.g., Liang et al., 1996a, 2004; Konapala et al., 2020; Li et al., 2011).

The small-scale Distributed Hydrology Soil and Vegetation Model (DHSVM) (Wigmosta et al., 2002; 1994) was developed to numerically represent the effects of topography, soil type, and vegetation on hydrological processes, such as plant transpiration, surface and subsurface runoff, and snow process for small watersheds with high spatial resolution described by digital elevation model (DEM) data. Unlike VIC+ in which the groundwater table is computed based on the mixed form of Richards equation (Luo et al., 2013), DHSVM calculates its groundwater table based on a simple conceptual approach. Also, DHSVM does not have the hydraulic redistribution process represented either. In DHSVM, the water and energy budgets are solved for each modeling grid cell which may contain an overstory canopy and an understory or bare soil. DHSVM uses Penman-Monteith equation to calculate its plant transpiration and it does not consider photosynthesis process, nor CO_2 assimilation. Thus, we have added these processes to DHSVM in this study, and the modified DHSVM model is denoted as DHSVMm.

There are some conceptual model parameters and physically based model parameters which cannot be well determined for either VIC+ or DHSVMm due to limited available observations. Together there are eleven parameters needed to be calibrated for each model as listed in Table 1: three are common to both, and if a' is posed as an unknown as in the *New Approach* that number reduces to two. These parameters are manually calibrated for each

591 hydrological model using the *Semi-empirical* approach. The calibrated parameters are then kept
592 unchanged for use in the other two approaches.

593 3.4 Four study sites and calibration

594 The four study sites selected are all of plot-scales. The main reasons to choose the plot-
595 scale sites are because (1) the availability of observations, e.g., gross primary production (GPP),
596 latent heat flux, and soil moisture, are all of plot-scales; (2) fewer number of model parameters
597 need to be manually calibrated, and (3) the routing process and its associated routing parameters
598 do not need to be included and calibrated. This allows the major efforts of the analysis be
599 devoted to the focus of the study.

600 Since VIC+ includes the representation of hydraulic redistribution, groundwater and
601 surface water interaction, plant storage, and leaf water potential, it is applied to two forest sites
602 where impacts on fluxes due to deep roots under normal and water-limited conditions can be
603 effectively investigated. These two forest sites are the Duke Forest Loblolly Pine (US-Dk3)
604 located in North Carolina and the Blodgett Forest (US-Blo) located in California. The main
605 vegetation of the Duke forest is loblolly pine trees with different hardwood understory species.
606 Soil types are loam and clay. The mean annual precipitation is 1145 mm and the mean air
607 temperature is 15.5 ℃. The hourly forcing data of years 2004 and 2005 at the Duke site are used
608 for calibration and validation, respectively. The Blodgett site is covered by mixed-evergreen
609 conifer forest with dominant even-aged ponderosa pine. Its primary soil type is loam. The mean
610 annual precipitation is 1226 mm and the mean air temperature is 11.1 ℃. The hourly forcing data
611 of year 2004 at the Blodgett forest site are used for calibration. Since there are no complete
612 hourly forcing data available for other periods at the Blodgett site, no validation was carried out.
613 The seasonal precipitation distributions are different at the two forest sites. For the Duke forest
614 site, the dry period is short and distributed throughout the year. But for the Blodgett site, the dry
615 period is long and occurs over the summer months. Thus, plants at the Blodgett site survive the
616 summer through optimizing their behaviors and stored soil moisture to adapt to the dry climate
617 conditions.

618 The DHSVMm is applied to two grassland sites since its limitations are less severe for
619 grassland as discussed. These two are the Mather site located in Pennsylvania in USA and the

Oensingen site located in Switzerland. The Mather site is covered with alfalfa, white clover, red clover, and tall fescue grass. Its mean annual precipitation is 1148 mm with about 85 precipitation days per year, and its mean annual temperature is 10 ^{o}C. The Mather site has a temperate continental climate with warm summers. The Oensingen site is covered with mixed-grasses. The mean annual precipitation is 1100 mm and the mean annual temperature is 9 ^{o}C. Hourly forcing data of the year 2010 are used at both grassland sites.

The main reason we use the *Semi-empirical* approach to carry out the calibration and validation is that it has the additional constant *a'* in its semi-empirical stomatal conductance model which needs to be calibrated. Results from the two hydrological models are shown in Figure 2. The observed data are from AmeriFluxin for Duke and Blodgett sites, and from MODIS for Mather and Oensingen sites.

Through regression between the original BBL and the BBL-updated for different vegetation types, we obtain for both the Duke and Blodgett sites, $a'_{max} = 9$, for the Mather site, $a'_{max} = 17$, and for the Oensingen site $a'_{max} = 24$. In *Semi-empirical*, a' is considered constant and $a' = 2$ is found to have the best fit for all four sites based on the manual calibrations in both VIC+ and DHSVMm.

The relative differences based on L2 norm (LD) and the coefficient of determination (R^2) for each observed variable are computed. From Figure 2, we can see that VIC+ simulates the soil moisture best, followed by latent heat flux, then the gross primary productivity (GPP); while DHSVMm simulates the GPP quite well. At the Duke site, for soil moisture, latent heat and GPP, their R^2 and LD (in parenthesis) are 0.87 (0.13), 0.83 (0.39) and 0.7 (0.47), respectively; and at the Blodgett site, they are 0.92 (0.12), 0.62 (0.54), and 0.60 (0.81). For the two grassland sites, total evapotranspiration (ET) and GPP are measured and DHSVMm consistently gives larger ET, but better estimates of GPP. For ET and GPP, their R^2 and LD for Mather are 0.79 (1.35) and 0.88 (0.36), respectively; and 0.64 (1.21) and 0.87 (0.39) for Oensingen site. Considering the complexity of the modeling involved, these levels of relative difference with the limited measured data can be viewed as having reached reasonably good fit.

After the model calibrations, the cost function of Eq. (7) is employed to determine the ratio *a/b* needed. A sensitivity analysis between c_i/c_a and the different ratios of a/b for each of

649 the four sites is presented in Figure 3. The ratio of a/b =1/146 $\approx$ 0.0068 suggested by (Stocker
650 et al., 2020) based on data is also included in Figure 3 for comparison. All these four sites show
651 similar patterns – the value of c_i/c_a increases with a decrease in a/b. When a/b is between
652 0.001 and 1, the value of c_i/c_a is not sensitive to the a/b value at these four sites. Therefore, *a/b*
653 = 1/146, which is within the insensitive range for all these four sites, is used in this study.

654 **4. Results and analyses**

655 On the outset it is important to emphasize that for the present complex modeling problem
656 the available measurements related to plant behaviors are limited to only 3 and 2 variables,
657 respectively, for the forest sites and the grassland sites, and only one of the measurements is
658 directly related to a modeling variable, E_{tr}. Because of this limitation, a system with variables
659 incorrectly solved might still perform seemingly well with respect to the data – an equifinality
660 pitfall. Under this circumstance, what physical insights one approach can reveal over another
661 weigh more on the merits of different approaches than their goodness of fit to the few
662 measurements. This is particularly so as we are dealing with plausible expressions of the same
663 processes and, by nature, they give similar results with respect to the limited measurements in
664 order to be considered equally plausible.

665 4.1 Models versus available data

666 Results of *New Approach*, *LC-extended* and *Semi-empirical* approaches are first
667 compared with the limited available measurements. The latter two represent the current practice
668 albeit that *LC-extended* is an extension of Prentice-2014 to cover dry climates. All three
669 approaches using the same parameter values calibrated via *Semi-empirical*. Results from each
670 approach in comparison with the available observation data are shown through Figures 2, 4 and
671 5. Unsurprisingly, they all give compatible relative errors and goodness of fit to the data as the
672 results summarized in Figures 2, 4 and 5 illustrate. These plots include both daytime and
673 nighttime simulation results and observations. These results also show that the *LC-extended*

formulation developed in this study and used in *LC-extended* approach works for the dry weather condition as intended.

4.2 Comparison of the solved variables

To gain deeper insights, we compare results of the main variables A_n, c_i, E_{tr}, ψ_l, g_{s,co_2} and a' obtained from different approaches during the study period. Nighttime c_i, A_n, g_{s,co_2}, ψ_l, and E_{tr} from sunset to 8am the next day are excluded from evaluation because at nighttime c_i approaches c_a which results in the three variables of c_i, A_n, and E_{tr} having similar values among different approaches and, if included, would skew the overall differences.

Before we start our detailed analyses and discussions, it is important to note that (1) a' has not been treated as a variable in previous studies, (2) impacts of a' on other variables, A_n, c_i, E_{tr}, ψ_l, g_{s,co_2} are found significant, and (3) solutions Sof A_n, c_i, E_{tr}, ψ_l , g_{s,co_2} from *LC-extended* are the same as those from *New Approach*, when the former gives $a' < a'_{max}$, but are very different otherwise.

To show indeed each of the two different perspectives on stomatal conductance leads to different results, we first compare *Semi-empirical* and *LC-Extended.* Their results presented in Figure 6 do show significant difference.

To assess merits of the *New Approach,* pairwise comparisons are made. First, *New Approach* vs *Semi-empirical* results plotted in Figure 7 show that their g_{s,co_2}, a', c_i, and ψ_l are significantly different at all four sites studied. Relative differences presented in Figure 8 provide a sharper view of these comparisons. The largest relative differences happen in E_{tr}, ψ_l, g_{s,co_2} and a' which can be as high as 100% except for ψ_l at the two grassland sites where the relative differences can be as high as 200%, followed by c_i which can be up to 50%, while the differences in A_n, generally less than 50% at all four sites, are the smallest. The dramatic differences in a' occur because the *Semi-empirical* fails to consider the time-varying nature of a' by keeping it constant. The constant $a' = 2$, determined through calibration, is much smaller almost at all time than those obtained by *New Approach* during the study period as shown in

703 Figure 7. This implies that a' turns out to be a critical factor driving the differences in the
704 solution of other variables.

705 Comparisons between the *New Approach* and *LC-extended* are presented in Figures 8 and
706 9. For *LC-extended*, a' are back-calculated using Eq. (16) after A_n, g_{s,co_2}, c_i, and ψ_l are solved.
707 Their relative differences (Figure 8) are the lowest in A_n among the six variables at the two forest
708 sites, and are comparable between A_n and c_i at the two grassland sites. At all four sites,
709 differences in E_{tr} are similar to those between *New Approach* and *Semi-empirical* but with less
710 scatter. We note that the largest relative differences in E_{tr} between *New Approach* and *LC-*
711 *extended* are greater than those between the *New Approach* and *Semi-empirical*; furthermore,
712 13.6% (Duke), 18.4% (Blodgett), 15.9% (Mather), and 7.6% (Oensingen) of the data lie outside
713 the bounds of Figure 8. Also, E_{tr} estimates from the *New Approach* are generally greater than
714 those from *Semi-empirical* at all four sites, while mostly smaller than those from *LC-extended*
715 (Figure 8). The differences in c_i are smaller than those between *New Approach* and *Semi-*
716 *empirical* except for the summer months at the Blodgett site where the differences are larger.
717 Similar to E_{tr}, the c_i estimates from the *New Approach* are also generally greater than those from
718 *Semi-empirical* at all four sites, while mostly smaller than those from *LC-extended* (Figure 8).
719 For ψ_l, even though the largest differences between the *New Approach* and *LC-extended* are
720 larger than those between *New Approach* and *Semi-empirical*, the majority of the differences are
721 smaller as indicated by the LD metric for the two forest sites; as for the two grassland sites, the
722 differences between the *New Approach* and *LC-extended* are much smaller than those between
723 *New Approach* and *Semi-empirical* (Figures 7, 8, and 9). The relative differences between *New*
724 *Approach* and *LC-extended* that lie outside the bounds of Figure 8 for ψ_l is 9.3% for the Mather
725 site and 5.2% for the Oensingen site, respectively. On the other hand, the relative differences
726 between *New Approach* and *Semi-empirical* that lie outside the displayed bounds for ψ_l are
727 17.7% and 59.2% for the Mather and Oensingen sites, respectively. Values in ψ_l and E_{tr} from
728 *LC-extended* (Figures 8 and 9) are also closer to those from *New Approach* than from *Semi-*
729 *empirical* (Figures 78, and 9) at all four sites. Similar to E_{tr} and c_i, estimates in ψ_l from the *New*
730 *Approach* are also generally smaller than those from *LC-extended*, while they are mostly larger
731 than those from *Semi-empirical* (Figure 8) at both forest sites. For the two grassland sites,
732 however, estimates in ψ_l from the *New Approach* are generally smaller than those from both *LC-*

extended and *Semi-empirical*. These general patterns are, in fact, associated with the large differences in g_{s,co_2} and a' among the three approaches further discussed next.

In terms of the relative differences in A_n, c_i, E_{tr}, and ψ_l, (see Figure 8), the largest differences still reside in E_{tr} and ψ_l, followed by c_i, with the smallest in A_n at all four sites, although the differences in c_i and A_n are comparable at the two grassland sites. Comparing the relative differences between *New Approach* and *Semi-empirical*, those between *New Approach* and *LC-extended* are generally smaller at all four sites. However, the differences in g_{s,co_2} and a' are pronounced (Figures 8 and 9), and the trend of the differences in g_{s,co_2} and a' is reversed from that between *LC-extended* versus *Semi-empirical*. For both g_{s,co_2} and a' the *New Approach* gives much higher values than *Semi-empirical* (see Figures 7 and 8) but much lower values than *LC-extended* (see Figures 8 and 9) at all four sites. This is because a large number of a' are unrealistically high for all four sites with the *LC-extended*. For Duke, Blodgett, Mather, and Oensingen sites, there are, respectively, 45.7%, 79.1%, 49.4% and 44.9% of a' that are higher than a'_{max} (see Table 2a). These unreasonable a' values imply that using the minimum cost function, i.e., Eq. (7), together with other Eqs. (8) – (11), leaves *a'* an unmodeled free variable, and that introducing additionally Eq. (16) is necessary, even though the levels of differences in A_n, E_{tr}, c_i, and ψ_l between these two approaches are generally much smaller than those in g_{s,co_2} and a' between the *New Approach* and *LC-extended*. In fact, there are some a' values in the *LC-extended* that are higher than 200 at all four sites which are out of the display bound in Figure 9. Also, there are large percentages of relative differences between *New Approach* and *LC-extended* for g_{s,co_2} and a' whose values are outside their respective bounds displayed in Figure 8. For g_{s,co_2}, these percentages are, respectively, 29.9% (Duke), 32.3% (Blodgett), 32.6% (Mather), and 17.2% (Oensingen). Further examination on *a'* is given below. Figures 8 and 9 clearly show that the differences in A_n, E_{tr}, c_i, and ψ_l are much smaller than those in g_{s,co_2}, and a' between the *New Approach* and *LC-extended*, but these smaller differences are produced at the expense of unrealistic a'. This is an example of variable equifinality.

After studying these comparisons, it is clear that *a'* is a discriminant factor that explains the observed differences. In *LC-extended*, *a'* is an unmodeled free variable, and if $a' < a'_{max}$,

which happens 54.3% (Duke), 20.9% (Blodgett), 50.6% (Mather), and 55.1% (Oensigen) of the time, *New Approach* and *LC-extended* give same A_n, c_i, E_{tr}, ψ_l, g_{s,co_2} and a'. But for the rest of the time when $a' > a'_{max}$, results are very different. In this case, *LC-extended* gives solutions that consist of unreasonable five variable combinations—this is further detailed in the next section. As for the *Semi-empirical*, because a' is held constant, it yields even larger differences with the *New Approach.* Introducing *a'* as an additional modeled variable in the *New approach* is thus important and necessary.

Impact of *a'*, either from the unmodeled thus unconstrained *a'* in *LC-extended*, or from the inadequate treatment of a' as a constant from model calibration in *Semi-empirical*, is not known previously, and is an important investigation of this study.

4.3 Why *a'* should be a model variable

In this section we will show that consider *a'* as an additional model variable and set $a' \leq a'_{max}$ has important implications. We will further show via analyzing *LC-extended* results that *a'* being a free variable lead to a large percentage of results having $a' > a'_{max}$ as summarized in Table 2a, and that those solutions having $a' > a'_{max}$ are predominantly unreasonable in light of physical reality as shall be presented in Table 2g. In contrast, by taking up *a'* as a model variable, such a serious problem is greatly reduced (Table 2h).

Eq (16) provides a clue as to when a' could be high in *LC-extended*. The solved A_n, c_i, ψ_l, g_{s,co_2} results corresponding to $a' > a'_{max}$ are loosely parsed into four groups based on Eq. (16): (1) g_{s,co_2}>0.3 mol/m^2/s (red), (2) ψ_l < -2.2 MPa (blue), (3) A_n < 2 (green) and (4) c_i > 300 (orange). Each group by itself does not mean much, however, a daytime (defined to be from 8 AM to sunset in this study) result is deemed unrealistic if it falls into more than one of these four groupings as explained below. The intent here is not to define a rigorous boundary, but to explain how one can determine results to be unreasonable and slight variations on these boundaries would not affect the conclusions.

The *LC-extended* results having $a' > a'_{max}$ are presented in Figure 10a (forest sites) and Figure 10c (grassland sites) distinguished by the above grouping, while results that do not belong

788 to these four groups but have $a' > a'_{max}$ are also plotted (in grey). Table 2b lists the percentage of
789 each group showing that these four groups indeed contribute most to $a' > a'_{max}$. Since they are
790 not mutually exclusive from one another, the sum of them is more than 100%, indicating that
791 there are points belonging to multiple groups at the same time. This is of significance as this is
792 how we are able, guided by Figure 10, to pinpoint the problems of unreasonableness with those
793 results having $a' > a'_{max}$. Take for example the combination of $A_n < 2$ μmol/m^2/s and $c_i > 300$
794 μmol/mol (exclude those with small g_{s,co_2} which may be potentially reasonable), such
795 combination is unlikely when their corresponding g_{s,co_2} values are not small (see orange dots in
796 Figures 10b and 10d), for during daytime $c_i > 300$ μmol/mol means photosynthesis activity is
797 strong, and net carbon assimilation should be high, and thus $A_n < 2$ μmol/m^2/s is unlikely and
798 thus the set of solution should be discarded. The solutions having $\psi_l < -2.2$ with either $g_{s,co_2} > 0.3$
799 or $c_i > 300$ and $g_{s,co_2} > 0.3$ and $A_n < 2$ can be similarly discounted. The combination of $\psi_l < -2.2$
800 and $A_n < 2$ is found associated with large c_i (close to 300 μmol/mol) (see Figure 10b) which is
801 unrealistic. The combination of $g_{s,co_2} > 0.3$ mol/m^2/s and $c_i > 300$ μmol/mol in daytime is possible,
802 however, Figures 10b and 10d show that points in this group are associated mostly with small to
803 medium A_n (although $A_n > 2$) under not water stressed conditions (in contrast to larger A_n values
804 shown in Figure 9 at all four sites). Such a combination makes them unreasonable since with
805 large g_{sco2} and large c_i, the assimilation A_n should be high as well in the daytime when they are
806 not water stressed. Following such a detailed examination, we found that *LC-extended* results are
807 unreasonable when they fall into any of the two-combinations of Table 2c. It then follows, three-
808 combinations of Table 2d also are unreasonable based on similar reasoning discussed for the
809 two-combinations in Table 2c. There is no result that simultaneously satisfies the four-
810 combinations.

811 From Table 2c, data points satisfying both $A_n < 2$ μmol/m^2/s and $c_i > 300$ μmol/mol
812 constitute the largest fraction of two-group combination (except for the Oensingen Site) that
813 leads to $a' > a'_{max}$ with 32.9% for Duke site, 14.7% for Blodgett site, 35.2% for Mather site, and
814 20.8% for Oensingen site. That is to say, since a' is a free variable as in *LC-extended*, large c_i in
815 daytime is often accompanied by unrealistically low A_n in order to achieve reasonable values of
816 g_{s,co_2} (see Figures 10b and 10d) as observed from Eq. (16) as such combination would lead to the

optimal cost. With Eq. (16) added, however, A_n can no longer be very small as a' cannot be very large to compensate the very small A_n.

The grassland sites, comparing with the forest sites, are less water stressed for their ψ_l are always higher than -2.2 MPa. In less water stressed environments, the interactions between c_i and A_n and between g_{s,co_2} and c_i over the specified ranges are stronger for the combinations of g_{s,co_2}>0.3 mol/m^2/s and c_i > 300 μmol/mol and of c_i >300 mol/m^2/s and A_n < 2 account for much higher percentages at the grassland sites than at the forest sites (see Table 2c). Other than that, the four sites exhibit similar trends as can be seen in Figure 10.

Table 2c summarizes six possible 2-combinations of the 4 groupings while Table 2d shows four possible 3-combinations of the 4 groupings. Each of the rows of Table 2c plus Table 2d does not sum up to 100% leading to an important question: might the results represented by other possible combinations be unreasonable as well? Further analysis is carried out to answer this question. First, we determine the percentage of results that only falls within one group (Table 2e) and not to any other three groups listed in Table 2e, and then further divide each group into subgroups. The difference between groups shown in Table 2e and those in Table 2b is that in Table 2e, the results in one column do not appear in another column, whereas it is not the case for Table 2b. In other words, results in each group shown in Table 2e do not overlap with any other groups. For groups in Table 2e, they are further divided into subgroups shown in Table 2f. Based on the behaviors of the plants and the observed data (e.g., Deans et al., 2020; Ennahli & Earl, 2005; Joshi et al., 2020; Leuning, 1995; Schulze & Hall, 1982; Urban et al., 2014; Zhou et al., 2013), those variable values having combinations fallen into these subgroups in Table 2f are identified as unreasonable. Using these added criteria, the fraction of data points within each subgroup of the unreasonable results is summarized in Table 2f. A total summation, together with Tables 2c and 2d, gives the total fraction of unreasonable results for those with $a' > a'_{max}$ in Table 2g. The results so obtained show that at least 70.9% for the forest sites and could be as high as 98.1% for grassland sites of all $a' > a'_{max}$ results are deemed unreasonable. We note that what this analysis shows is that we are not trying to conduct an exhaustive search

for all unreasonable variable combinations associated with $a' > a'_{max}$, but to show that there is a large fraction of them that is unreasonable.

But then there is a question that has to be addressed: are the results from *New approach* where $a' \leq a'_{max}$ all reasonable? Following the same analysis we have calculated its fraction of the unreasonable results, and find that they are much lower and are in the range between 3.2% to 9.3% as shown in Table 2h. This represents a dramatic improvement. It, nonetheless, also points out that the current understanding still has room to improve.

In addition, in terms of what we have obtained with *New Approach* on the variation of *a'* values, our results (Figure 9) have shown that *a'* varied in a narrower range over time than those from *LC-extended*, and definitely not a constant. The *a'* values from *LC-extended* have a much larger variation range over the different periods of a year at all four sites. Specifically, for the *New Approach* the ranges of the hourly *a'* obtained for the two forest sites are, respectively, between 1.38 and 9 (Duke site, Figure 9g) over the two-year period and between 0.89 and 9 (Blodgett site, Figure 9s) over one-year period; for the two grassland sites, it is between 1.79 and 17 (Mather site, Figure 9ee) and between 2.15 and 24 (Oensingen site, Figure 9qq). In contrast, for the *LC-extended* approach the hourly *a'* varies between 1.36 and 800 (Duke) and between 0.8 and 2270 (Blodgett) for the forest sites; and for the grassland sites between 1.79 and 313 (Mather) and between 2.15 and 191 (Oensingen). Furthermore, *New approach* gives much narrower diurnal *a'* variation than *LC-extended* at both forest sites (see Figures 11g and 11o) and at the two grassland sites (Figure not shown)—which agrees well with the current data and understanding of *a'* (Miner et al., 2017).

The limited data available at present time only allow us to state that g_{s,co_2}, A_n, c_i, and ψ_l combinations are more reasonable with our *New Approach* based on plants general behaviors in the daytime and the observed data shown in the literature. Additionally, the time-varying *a'* from our *New Approach* removes the need of constant assumption, and from which one can easily estimate values of g_1 based on Eq. (17) which is related to $\lambda = \partial A_n / \partial E_{tr}$ used in the optimality rule for the WUE group (e.g., Medlyn et al., 2011). In other words, one no longer needs to calibrate g_1 or λ at each daily time step as is currently done in practice. Figures 11h and 11p show the comparisons of g_1's among the three approaches over a summer week at both forest sites. In

873 Figure 11, comparisons on variables of g_{s,co_2}, A_n, c_i, ψ_l, and a', together with E_{tr} and *GPP* where
874 observations are available, are also included.

875 Results presented in Figures 6 - 11 demonstrate that, for all six solved variables, E_{tr}, A_n,
876 c_i, ψ_l, a' and g_{s,co_2}, our *New Approach*, which simultaneously employs two stomatal conductance
877 models are significantly different from those using only either one. Furthermore, these results
878 demonstrate that the differences between *Semi-empirical* and *LC-extended* (Figure 6) are larger
879 than those between our *New Approach* and either *Semi-empirical* (Figure 7) or *LC-extended*
880 (Figure 9) based on both metrics, R^2 and LD. These results clearly show that our *New Approach*
881 takes the advantage of both models, *LC-extended* and *Semi-empirical*.

882 The preceding results show that only the *New Approach* gives reasonable values of a' and
883 with it reasonable solutions, and that the problem studied is better described as having six
884 variables.

885 4.4 Relationship between carboxylation and transpiration

886 The cost function, Eq. (7), represents a trade-off between carboxylation and transpiration.
887 The relationship between carboxylation and transpiration is shown in Figure 12 in which
888 carboxylation is presented by $V = V_{cmax}/A_n$, and transpiration is presented by $G = g_{sab,co_2}/A_n$.
889 Prentice et al. (2014) showed that the V~G relationship follows a hyperbola shape based on
890 observed data with the observed V in the range between 0 and 30, and the observed G in the
891 range between 0 and 0.05. Although their data are from instant measurements under certain
892 conditions while here our modeling gives hourly results over one or two years, such a hyperbola
893 relationship between V and G does provide another baseline to evaluate the results obtained.
894 Figure 12 shows that results from both *New Approach* and *LC-extended* generally follow the
895 hyperbola shape at all four sites, while the *Semi-empirical* does not, even though they have the
896 same goodness of fit to the available data (see Figures 2, 4, and 5). Between *New Approach* and
897 *LC-extended*, the former has less scatter, and a larger percentage of its results fall inside the
898 range of $0 < G \leq 0.05$ [mol·μmol^{-1}] and $0 < V \leq 30$ [mol·mol^{-1}] (see Table 3). The larger scatter in
899 the *LC-extended* in Figure 12 is due to its larger g_{s,co_2} and a' (see Figures 8 and 9) that result in
900 larger G at all four sites, since the differences in A_n from both approaches are much smaller. On

the other hand, the reason that the *Semi-empirical* approach does not show hyperbola *V~G* relationship is because g_{s,co_2} are constrained to small values (see Figure 7) by the use of a constant *a'*=2 from calibration. Not having a dynamic *a'* imposes limitations on the *Semi-empirical* approach.

5. Conclusions

In this study, we present a modeling framework which is applied to model the water and carbon exchange of plants inside hydrological models. We have demonstrated how we pose six equations, i.e., Eqs. (7) – (11) and Eq. (16), and two constraints (18) and (19) to solve a constrained optimization problem of six variables, E_{tr}, g_{s,co_2}, A_n, c_i, ψ_l, and a'. The core idea of this modeling framework is to model processes that are important and yet not completely understood with multiple equally plausible expressions from different perspectives. This represents an extension of the modeling strategy proposed in VIC+ by Luo et al. (2016; 2013). In the process, we have also extended the minimum cost optimality rule of Prentice et al. (2014) so that it applies to water-limited conditions. The extension also relaxes the coordination assumption that the Rubisco-limited photosynthetic rate be equal to electron transport-limited photosynthetic rate.

The unique strength of this modeling framework includes the following: First, not only does this modeling framework provide more constraints to the same process to reduce the modeling system's free variables and unreasonable variable value combinations, but also individual processes are more comprehensively described because they are represented from multiple perspectives. Second, the presence of various constraints makes the individual processes more tightly coupled through shared variables which are solved simultaneously.

To further shed light on the impacts of the present framework, we conduct comparative studies employing two currently widely used stomatal conductance models. Three approaches used are denoted as: *New Approach* (the present model), *LC-extended* and *Semi-empirical*. To illustrate the versatility of the framework, this study uses two different hydrological models and four different study sites: VIC+ on two forest sites and the modified DHSVMm on two grassland sites. Our results show that all three approaches give compatible results regarding the available limited observation data. But a close examination shows that the differences in the solved

variables E_{tr}, A_n, c_i, g_{s,co_2}, ψ_l and *a'* are significant. This is a well-known phenomenon called equifinality that in modeling a complex system that the "right" results may be obtained with wrong solutions when observation data are limited. In-depth investigation reveals that the differences among the variables can be attributed to the treatment of *a': Semi-empirical* treats it as a constant; *LC-extended* treats it as a free variable by not considering it; and the present *New Approach* considers it as a modeling variable. The *New Approach* with two stomatal conductance models fused together is able to solve *a'* that agrees with the current understanding obtaining nearly constant day time *a'* and how the *a'* values vary between days . This indicates that *Semi-empirical* using a constant *a'* throughout a study period is not desirable, and that *LC-extended* not considering it at all which would result in a high fluctuation of *a'* over a diurnal cycle (see Figures 11g and 11o) is not desirable either. Furthermore, results from *New Approach* give hyperbolic relationship between plant carboxylation and transpiration as field observation dictates which further indirectly provides merits to the proposed framework as it matches with the trend observed.

Our investigation into the reasonableness of results in terms of physical reality has shown that the *New approach,* albeit makes dramatic improvements in this respect, still gives some unreasonable variable combinations, albeit a much smaller fraction, in the results. This reflects that there still exists some knowledge gap in the current understanding and expressions, and more work needs to be done.

It is also worth mentioning that the *New Approach* presented is simply one implementation of our modeling framework. The core idea of including as many least overlapped equally plausible or quasi-equally plausible formulations as possible can constrain a modeling system, reduce model's free variables and mitigate equifinality, decrease result uncertainties, and ultimately increase the model robustness and predictability – very important characteristics a model should possess (Prentice et al., 2015). This strategy also enables the modeling system to have more unknowns solved simultaneously, and thus makes it possible to have the different expressions in the model interact with and depend on one another. Plausible expressions for the same process obtained via different perspectives, when properly introduced, complement one another. This is why our modeling framework can make the system more robust and stable when the unknown variables of the system are solved in such a manner. Our results with two different

960 hydrological models, VIC+ and DHSVMm, applied to four different locations, clearly
961 demonstrate these points. That is, results from our framework – the *New Approach* – provide not
962 only good estimates on E_{tr}, A_n, c_i, and ψ_l, but also more reasonable values on g_{s,co_2} and a', and
963 the hyperbolic V~G shape as well.

964 Our modeling framework could also facilitate identification of inconsistency, should it
965 exist, among the different quasi-equally plausible expressions. This is because if no reasonable
966 solutions could be obtained by solving simultaneously all of the unknowns of the system, it could
967 imply that some of these expressions were not compatible with each other. In this way, by
968 analyzing the results, one may be provided with new insight regarding what could be the terms
969 or factors that may be missing, and under what conditions these expressions may not be
970 compatible, while at other conditions they are compatible and complement each other. After all,
971 the ecosystem is such a complex system, our current understanding of the system and the
972 available observations may not yet provide us processes that could fully describe the nature. This
973 modeling framework could help us move forward in identifying the gaps.

974 Lastly, our modeling strategy is not only applicable to the hydrological models (e.g.,
975 VIC+ and DHSVMm), but also suitable to other modeling systems in other fields where
976 knowledge is incomplete and many models from different perspectives are equally or quasi-
977 equally plausible. With the advance of our understanding and knowledge of the various complex
978 natural processes, more constraints/equations/expressions will be discovered, and they can be
979 fused together in a fashion similar to what we do here in studying the soil-plant-atmosphere
980 continuum. This study shows how this modeling framework can be realized, tested, and
981 explored.

982

983 **Acknowledgements:**

984 This work was partially supported by the U.S. National Science Foundation under CBET-
985 1236403 to the University of Pittsburgh, the LEMONTREE (Land Ecosystem Models based On
986 New Theory, obseRvations and ExperimEnts) project funded through the generosity of Eric and
987 Wendy Schmidt by recommendation of the Schmidt Futures program, and by the William Kepler

Whiteford Professorship to Xu Liang from the University of Pittsburgh. We also thank the Center for Research Computing and the Swanson School of Engineering for providing the computational resources at the University of Pittsburgh. We are indebted to Professor I. Colin Prentice of Imperial College London for his generosity in sharing insights and expertise on optimality models.

Author contributions:

For this work, Liuyan Hu and Hector W. Clavijo implemented the research ideas, performed model simulation experiments with Hu for the VIC+ model and Clavijo for the DHSVMm model, respectively, conducted analysis, and contributed to the manuscript writing at the beginning. Liuyan Hu prepared all the tables and figures. Jeen-Shang Lin contributed to the result analyses and coauthored the final manuscript. Xu Liang conceived the research ideas, designed the model simulation experiments, supervised the investigation, synthesized the results, and wrote and finalized the manuscript. All contributed to the discussions of the work.

Data access:

Data utilized in this study are accessible from the Ameriflux at https://ameriflux.lbl.gov/, from Euroflux at http://www.europe-fluxdata.eu/, from NASA MODIS at https://lpdaacsvc.cr.usgs.gov/appeears/. All the data generated in this work for the figures and tables will be available through Mendeley Data (DOI: xxx). The VIC (https://vic.readthedocs.io/en/master/) and DHSVM (https://dhsvm.pnnl.gov/) models are all open sources.

References:

Anderegg, W. R., Wolf, A., Arango-Velez, A., Choat, B., Chmura, D. J., Jansen, S., et al. (2018). Woody plants optimise stomatal behaviour relative to hydraulic risk. *Ecology letters, 21*(7), 968-977.

1014 Ball, J. T., Woodrow, I. E., & Berry, J. A. (1987). A model predicting stomatal conductance and its contribution to
1015 the control of photosynthesis under different environmental conditions. In *Progress in photosynthesis*
1016 *research* (pp. 221-224): Springer.
1017 Bauerle, W. L., & Bowden, J. D. (2011). Predicting transpiration response to climate change: insights on
1018 physiological and morphological interactions that modulate water exchange from leaves to canopies.
1019 *HortScience, 46*(2), 163-166.
1020 Berry, J. A., Beerling, D. J., & Franks, P. J. (2010). Stomata: key players in the earth system, past and present.
1021 *Current opinion in plant biology, 13*(3), 232-239.
1022 Beven, K. (2006). A manifesto for the equifinality thesis. *Journal of hydrology, 320*(1-2), 18-36.
1023 Buckley, T. N., Sack, L., & Farquhar, G. D. (2017). Optimal plant water economy. *Plant, cell & environment, 40*(6),
1024 881-896.
1025 Cherkauer, K. A., & Lettenmaier, D. P. (1999). Hydrologic effects of frozen soils in the upper Mississippi River
1026 basin. *Journal of Geophysical Research: Atmospheres, 104*(D16), 19599-19610.
1027 Cherkauer, K. A., & Lettenmaier, D. P. (2003). Simulation of spatial variability in snow and frozen soil. *Journal of*
1028 *Geophysical Research: Atmospheres, 108*(D22).
1029 Clavijo Sanabria, H. W. (2020). *An Investigation on Interactions between Plant Physiological-Hydrological-*
1030 *Biogeochemical processes and Acid Mine Drainage in Coal Refuse Piles using Optimality Principle*
1031 *Theory.* University of Pittsburgh,
1032 Collatz, G. J., Ball, J. T., Grivet, C., & Berry, J. A. (1991). Physiological and environmental regulation of stomatal
1033 conductance, photosynthesis and transpiration: a model that includes a laminar boundary layer.
1034 *Agricultural and Forest meteorology, 54*(2-4), 107-136.
1035 Collatz, G. J., Ribas-Carbo, M., & Berry, J. (1992). Coupled photosynthesis-stomatal conductance model for leaves
1036 of C4 plants. *Functional Plant Biology, 19*(5), 519-538.
1037 Cowan, I. (1982). Regulation of water use in relation to carbon gain in higher plants. In *Physiological plant ecology*
1038 *II* (pp. 589-613): Springer.
1039 Cowan, I., & Farquhar, G. (1977). Stomatal function in relation to leaf metabolism and environment.
1040 Daly, E., Porporato, A., & Rodriguez-Iturbe, I. (2004). Coupled dynamics of photosynthesis, transpiration, and soil
1041 water balance. Part I: Upscaling from hourly to daily level. *Journal of Hydrometeorology, 5*(3), 546-558.
1042 Damour, G., Simonneau, T., Cochard, H., & Urban, L. (2010). An overview of models of stomatal conductance at
1043 the leaf level. *Plant, cell & environment, 33*(9), 1419-1438.
1044 De Kauwe, M. G., Zhou, S., Medlyn, B. E., Pitman, A. J., Wang, Y., Duursma, R. A., & Prentice, I. C. (2015). Do
1045 land surface models need to include differential plant species responses to drought? Examining model
1046 predictions across a mesic-xeric gradient in Europe.
1047 Deans, R. M., Brodribb, T. J., Busch, F. A., & Farquhar, G. D. (2020). Optimization can provide the fundamental
1048 link between leaf photosynthesis, gas exchange and water relations. *Nature Plants, 6*(9), 1116-1125.
1049 Dewar, R., Mauranen, A., Mäkelä, A., Hölttä, T., Medlyn, B., & Vesala, T. (2018). New insights into the covariation
1050 of stomatal, mesophyll and hydraulic conductances from optimization models incorporating nonstomatal
1051 limitations to photosynthesis. *New Phytologist, 217*(2), 571-585.
1052 Eller, C. B., Rowland, L., Oliveira, R. S., Bittencourt, P. R., Barros, F. V., da Costa, A. C., et al. (2018). Modelling
1053 tropical forest responses to drought and El Niño with a stomatal optimization model based on xylem
1054 hydraulics. *Philosophical Transactions of the Royal Society B: Biological Sciences, 373*(1760), 20170315.
1055 Ennahli, S., & Earl, H. J. (2005). Physiological limitations to photosynthetic carbon assimilation in cotton under
1056 water stress. *Crop Science, 45*(6), 2374-2382.
1057 Farquhar, G. D., von Caemmerer, S. v., & Berry, J. A. (1980). A biochemical model of photosynthetic CO 2
1058 assimilation in leaves of C 3 species. *Planta, 149*(1), 78-90.
1059 Feynman, R. (1967). *The character of physical law*: First MIT Press Paperback Edition.
1060 Franklin, O., Johansson, J., Dewar, R. C., Dieckmann, U., McMurtrie, R. E., Brännström, Å., & Dybzinski, R.
1061 (2012). Modeling carbon allocation in trees: a search for principles. *Tree Physiology, 32*(6), 648-666.
1062 Gutschick, V., & Simonneau, T. (2002). Modelling stomatal conductanceof field-grown sunflower under varying
1063 soil water content and leafenvironment: comparison of three models of stomatal response toleaf
1064 environment and coupling with an abscisic acid-based modelof stomatal response to soil drying. *Plant, Cell*
1065 *& Environment, 25*(11), 1423-1434.
1066 Heroult, A., LIN, Y. S., Bourne, A., Medlyn, B. E., & Ellsworth, D. S. (2013). Optimal stomatal conductance in
1067 relation to photosynthesis in climatically contrasting Eucalyptus species under drought. *Plant, Cell &*
1068 *Environment, 36*(2), 262-274.

Hölttä, T., Lintunen, A., Chan, T., Mäkelä, A., & Nikinmaa, E. (2017). A steady-state stomatal model of balanced leaf gas exchange, hydraulics and maximal source–sink flux. *Tree physiology*, *37*(7), 851-868.

Jarvis, P. (1976). The interpretation of the variations in leaf water potential and stomatal conductance found in canopies in the field. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences, 273*(927), 593-610.

Joshi, J., Stocker, B. D., Hofhansl, F., Zhou, S., Dieckmann, U., & Prentice, I. C. (2020). Towards a unified theory of plant photosynthesis and hydraulics. *BioRxiv*.

Katul, G., Manzoni, S., Palmroth, S., & Oren, R. (2010). A stomatal optimization theory to describe the effects of atmospheric CO2 on leaf photosynthesis and transpiration. *Annals of Botany*, *105*(3), 431-442.

Katul, G., Palmroth, S., & Oren, R. (2009). Leaf stomatal responses to vapour pressure deficit under current and CO2-enriched atmosphere explained by the economics of gas exchange. *Plant, Cell & Environment*, *32*(8), 968-979.

Konapala, G., Kao, S. C., & Addor, N. (2020). Exploring Hydrologic Model Process Connectivity at the Continental Scale Through an Information Theory Approach. *Water Resources Research*, *56*(10), e2020WR027340.

Leuning, R. (1990). Modelling stomatal behaviour and and photosynthesis of Eucalyptus grandis. *Functional Plant Biology*, *17*(2), 159-175.

Leuning, R. (1995). A critical appraisal of a combined stomatal-photosynthesis model for C3 plants. *Plant, Cell & Environment, 18*(4), 339-355.

Leuning, R., Dunin, F., & Wang, Y.-P. (1998). A two-leaf model for canopy conductance, photosynthesis and partitioning of available energy. II. Comparison with measurements. *Agricultural and Forest Meteorology, 91*(1-2), 113-125.

Li, H., Huang, M., Wigmosta, M. S., Ke, Y., Coleman, A. M., Leung, L. R., et al. (2011). Evaluating runoff simulations from the Community Land Model 4.0 using observations from flux towers and a mountainous watershed. *Journal of Geophysical Research: Atmospheres*, *116*(D24).

Liang, X., Guo, J., & Leung, L. R. (2004). Assessment of the effects of spatial resolutions on daily water flux simulations. *Journal of Hydrology*, *298*(1-4), 287-310.

Liang, X., Lettenmaier, D. P., & Wood, E. F. (1996). One-dimensional statistical dynamic representation of subgrid spatial variability of precipitation in the two-layer variable infiltration capacity model. *Journal of Geophysical Research: Atmospheres*, *101*(D16), 21403-21422.

Liang, X., Lettenmaier, D. P., Wood, E. F., & Burges, S. J. (1994). A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *Journal of Geophysical Research: Atmospheres*, *99*(D7), 14415-14428.

Liang, X., Wood, E. F., & Lettenmaier, D. P. (1996). Surface soil moisture parameterization of the VIC-2L model: Evaluation and modification. *Global and Planetary Change*, *13*(1-4), 195-206.

Liang, X., & Xie, Z. (2001). A new surface runoff parameterization with subgrid-scale soil heterogeneity for land surface models. *Advances in Water Resources*, *24*(9-10), 1173-1193.

Liang, X., & Xie, Z. (2003). Important factors in land–atmosphere interactions: surface runoff generations and interactions between surface and groundwater. *Global and Planetary Change*, *38*(1-2), 101-114.

Liang, X., Xie, Z., & Huang, M. (2003). A new parameterization for surface and groundwater interactions and its impact on water budgets with the variable infiltration capacity (VIC) land surface model. *Journal of Geophysical Research: Atmospheres*, *108*(D16).

Lu, Y., Duursma, R. A., & Medlyn, B. E. (2016). Optimal stomatal behaviour under stochastic rainfall. *Journal of theoretical biology*, *394*, 160-171.

Luo, X., Liang, X., & Lin, J. S. (2016). Plant transpiration and groundwater dynamics in water-limited climates: Impacts of hydraulic redistribution. *Water Resources Research*, *52*(6), 4416-4437.

Luo, X., Liang, X., & McCarthy, H. R. (2013). VIC+ for water-limited conditions: A study of biological and hydrological processes and their interactions in soil-plant-atmosphere continuum. *Water Resources Research*, *49*(11), 7711-7732.

Manzoni, S., Vico, G., Palmroth, S., Porporato, A., & Katul, G. (2013). Optimization of stomatal conductance for maximum carbon gain under dynamic soil moisture. *Advances in Water Resources*, *62*, 90-105.

Medlyn, B. E., Duursma, R. A., Eamus, D., Ellsworth, D. S., Prentice, I. C., Barton, C. V., et al. (2011). Reconciling the optimal and empirical approaches to modelling stomatal conductance. *Global Change Biology*, *17*(6), 2134-2144.

Miner, G. L., Bauerle, W. L., & Baldocchi, D. D. (2017). Estimating the sensitivity of stomatal conductance to photosynthesis: a review. *Plant, cell & environment*, *40*(7), 1214-1238.

MÄKELÄ, A., BERNINGER, F., & HARI, P. (1996). Optimal control of gas exchange during drought: theoretical analysis. *Annals of Botany*, *77*(5), 461-468.

Nobel, P. (1999). Leaves and fluxes. *Physicochemical and environmental plant physiology*, *2*, 293-349.

Prentice, I. C., Dong, N., Gleason, S. M., Maire, V., & Wright, I. J. (2014). Balancing the costs of carbon gain and water transport: testing a new theoretical framework for plant functional ecology. *Ecology letters*, *17*(1), 82-91.

Prentice, I. C., Liang, X., Medlyn, B. E., & Wang, Y.-P. (2015). Reliable, robust and realistic: the three R's of next-generation land-surface modelling. *Atmospheric Chemistry and Physics*, *15*(10), 5987-6005.

Rogers, A., Kumarathunge, D. P., Lombardozzi, D. L., Medlyn, B. E., Serbin, S. P., & Walker, A. P. (2021). Triose phosphate utilization limitation: an unnecessary complexity in terrestrial biosphere model representation of photosynthesis. *New Phytologist*, *230*(1), 17-22.

Schulze, E.-D., & Hall, A. (1982). Stomatal responses, water loss and CO 2 assimilation rates of plants in contrasting environments. In *Physiological plant ecology II* (pp. 181-230): Springer.

Schymanski, S. J., Sivapalan, M., Roderick, M., Hutley, L. B., & Beringer, J. (2009). An optimality-based model of the dynamic feedbacks between natural vegetation and the water balance. *Water Resources Research*, *45*(1).

Sperry, J. S., Venturas, M. D., Anderegg, W. R., Mencuccini, M., Mackay, D. S., Wang, Y., & Love, D. M. (2017). Predicting stomatal responses to the environment from the optimization of photosynthetic gain and hydraulic cost. *Plant, cell & environment*, *40*(6), 816-830.

Stocker, B. D., Wang, H., Smith, N. G., Harrison, S. P., Keenan, T. F., Sandoval, D., et al. (2020). P-model v1. 0: an optimality-based light use efficiency model for simulating ecosystem gross primary production. *Geoscientific Model Development*, *13*(3), 1545-1581.

Sun, R., Hernández, F., Liang, X., & Yuan, H. (2020). A Calibration Framework for High-Resolution Hydrological Models Using a Multiresolution and Heterogeneous Strategy. *Water Resources Research*, *56*(8), e2019WR026541.

Tardieu, F., & Davies, W. (1993). Integration of hydraulic and chemical signalling in the control of stomatal conductance and water status of droughted plants. *Plant, Cell & Environment*, *16*(4), 341-349.

Tuzet, A., Perrier, A., & Leuning, R. (2003). A coupled model of stomatal conductance, photosynthesis and transpiration. *Plant, Cell & Environment*, *26*(7), 1097-1116.

Urban, O., Klem, K., Holišová, P., Šigut, L., Šprtová, M., Teslová-Navrátilová, P., et al. (2014). Impact of elevated CO2 concentration on dynamics of leaf photosynthesis in Fagus sylvatica is modulated by sky conditions. *Environmental Pollution*, *185*, 271-280.

Wang, Y., Sperry, J. S., Anderegg, W. R., Venturas, M. D., & Trugman, A. T. (2020). A theoretical and empirical assessment of stomatal optimization modeling. *New Phytologist*, *227*(2), 311-325.

Westhoff, M. C., Zehe, E., & Schymanski, S. J. (2014). Importance of temporal variability for hydrological predictions based on the maximum entropy production principle. *Geophysical Research Letters*, *41*(1), 67-73.

Wigmosta, M. S., Nijssen, B., Storck, P., & Lettenmaier, D. (2002). The distributed hydrology soil vegetation model. *Mathematical models of small watershed hydrology and applications*, 7-42.

Wigmosta, M. S., Vail, L. W., & Lettenmaier, D. P. (1994). A distributed hydrology-vegetation model for complex terrain. *Water resources research*, *30*(6), 1665-1679.

Wolf, A., Anderegg, W. R., & Pacala, S. W. (2016). Optimal stomatal behavior with competition for water and risk of hydraulic impairment. *Proceedings of the National Academy of Sciences*, *113*(46), E7222-E7230.

Zhou, S., Duursma, R. A., Medlyn, B. E., Kelly, J. W., & Prentice, I. C. (2013). How should we model plant responses to drought? An analysis of stomatal and non-stomatal responses to water stress. *Agricultural and Forest Meteorology*, *182*, 204-214.

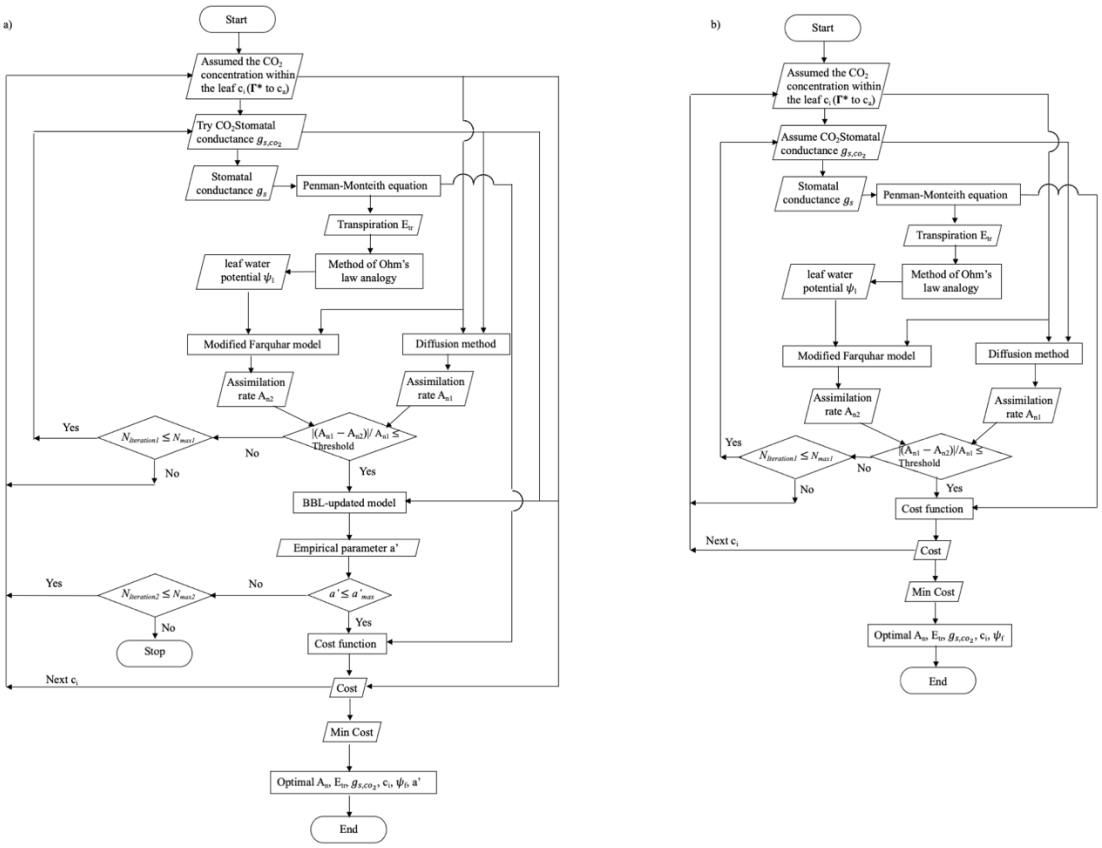

Figure 1. Flowchart of (a) The New approach, and (b) LC-extended approach.

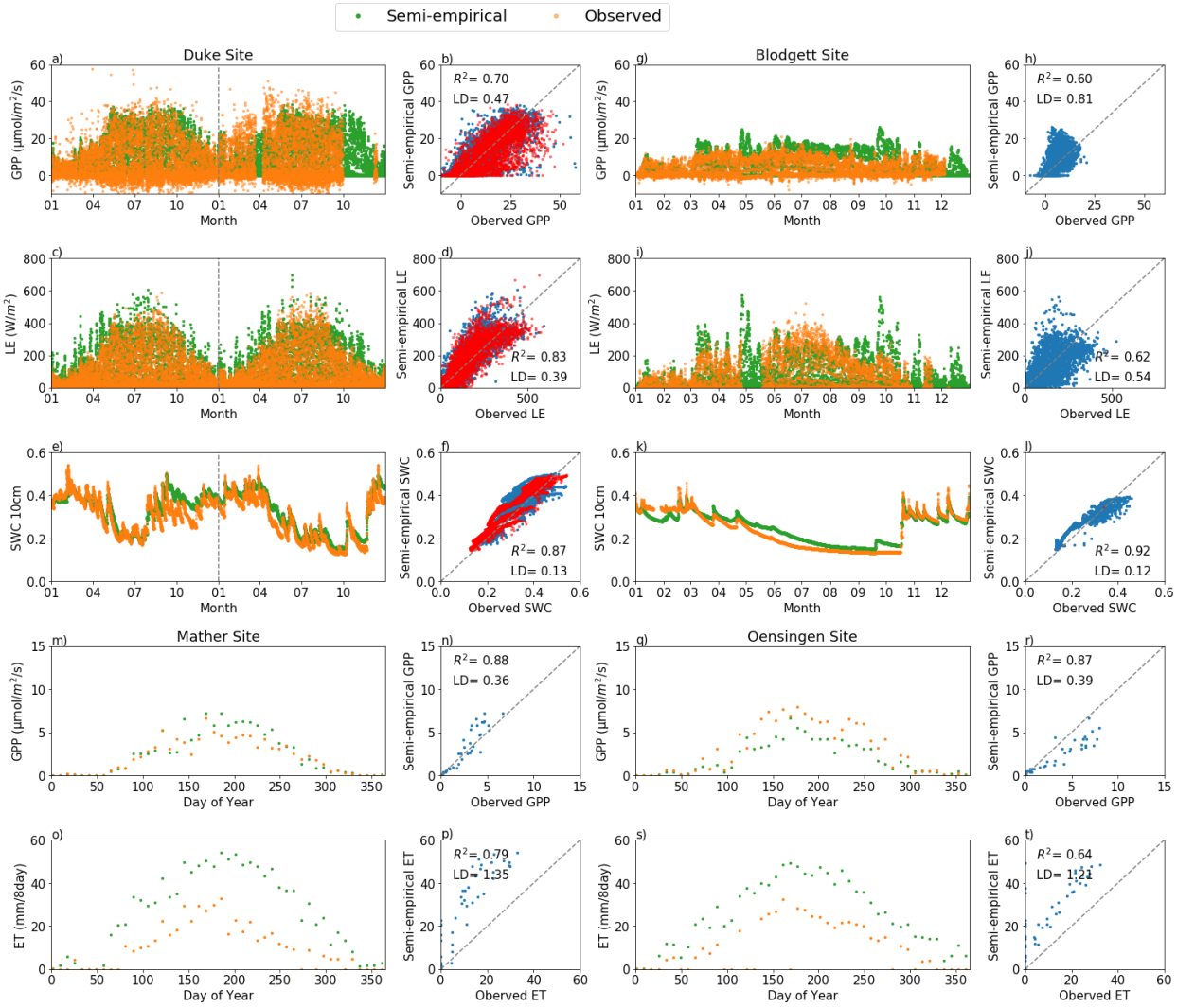

Figure 2. Comparison of the results using calibrated parameters by the Semi-empirical approach (green dots)with observations (orange dots). Plots (a)-(f): Hourly results over two years at the Duke site. Plots (g)-(*l*): Hourly results over one year at the Blodgett site. (m)-(p): 8-day results over one year at the Mather site. (q)-(t): 8-day results over one year at the Oensingen Site. (a), (b), (g), and (h) represent gross primary productivity (GPP); (c), (d), (i), and (j) latent heat flux (LE); (e), (f), (k), and (*l*) soil water content (SWC) at the depth of 10 cm; (m), (n), (q), and (r) represent gross primary productivity (GPP); and (o), (p), (s), and (t) 8-day total evapotranspiration (ET). For the Duke site, results on the left side of the vertical black dotted line in (a), (c) and (e) are for the calibration period while results on the right side are for the validation period.

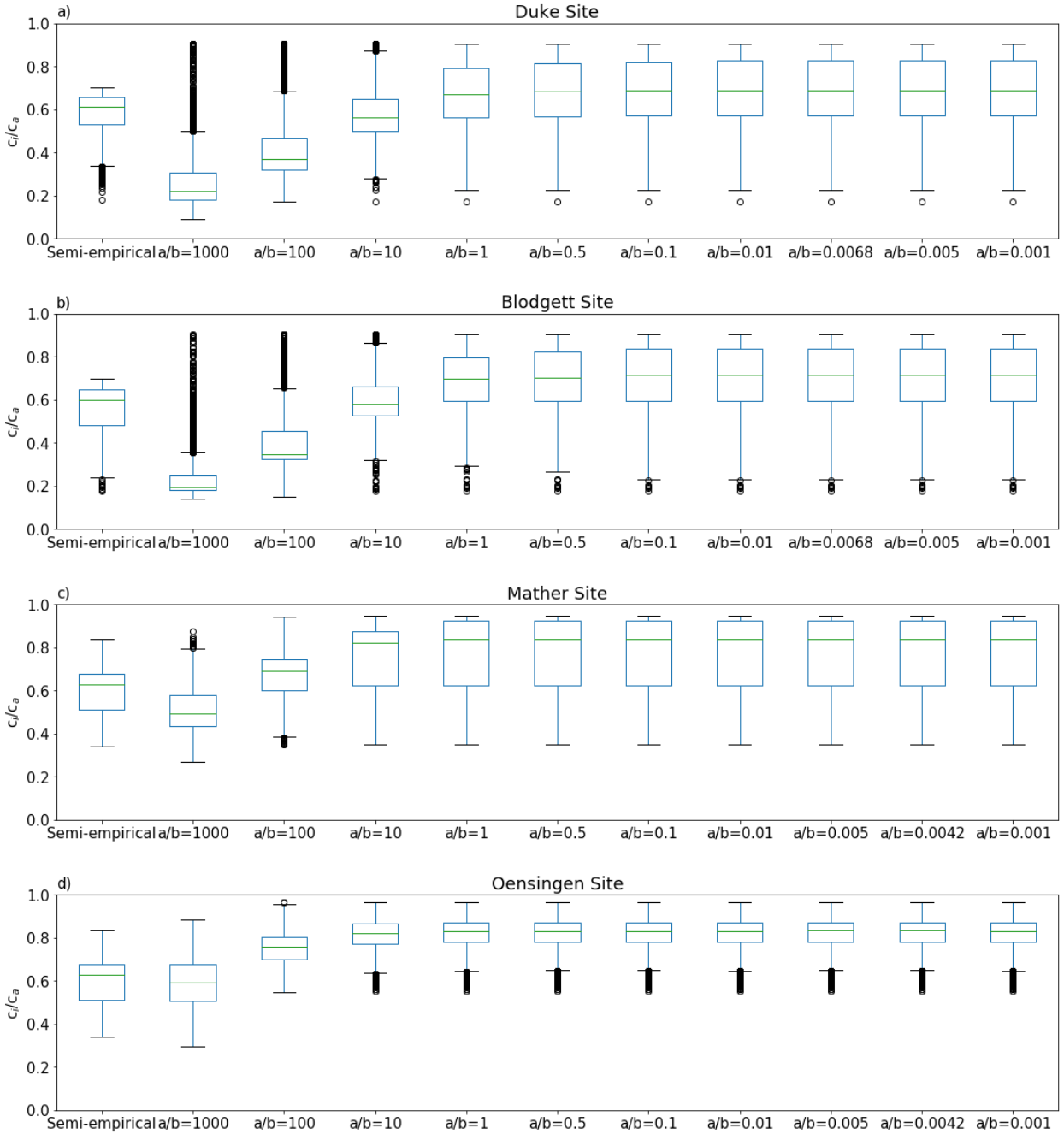

Figure 3. Sensitivity analysis between c_i/c_a and the different ratios of a/b in Eq. (7): (a) Duke site, (b) Blodgett site, (c) Mather site, and (d) Oensingen Site.

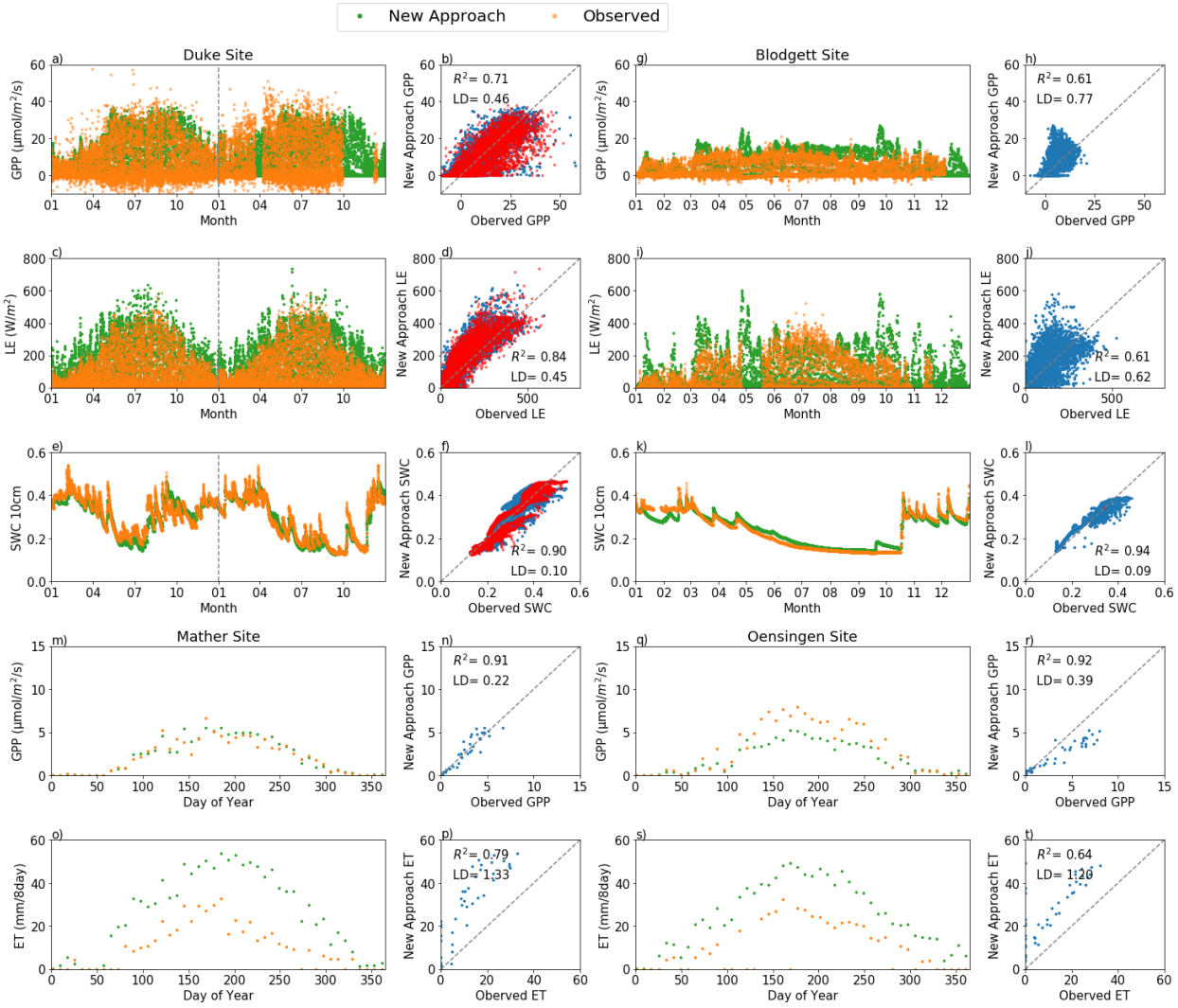

Figure 4. Comparison of the results using the same calibrated parameters by the New approach (green dots) with observations (orange dots). The notations used are the same as those in Fig. 2.

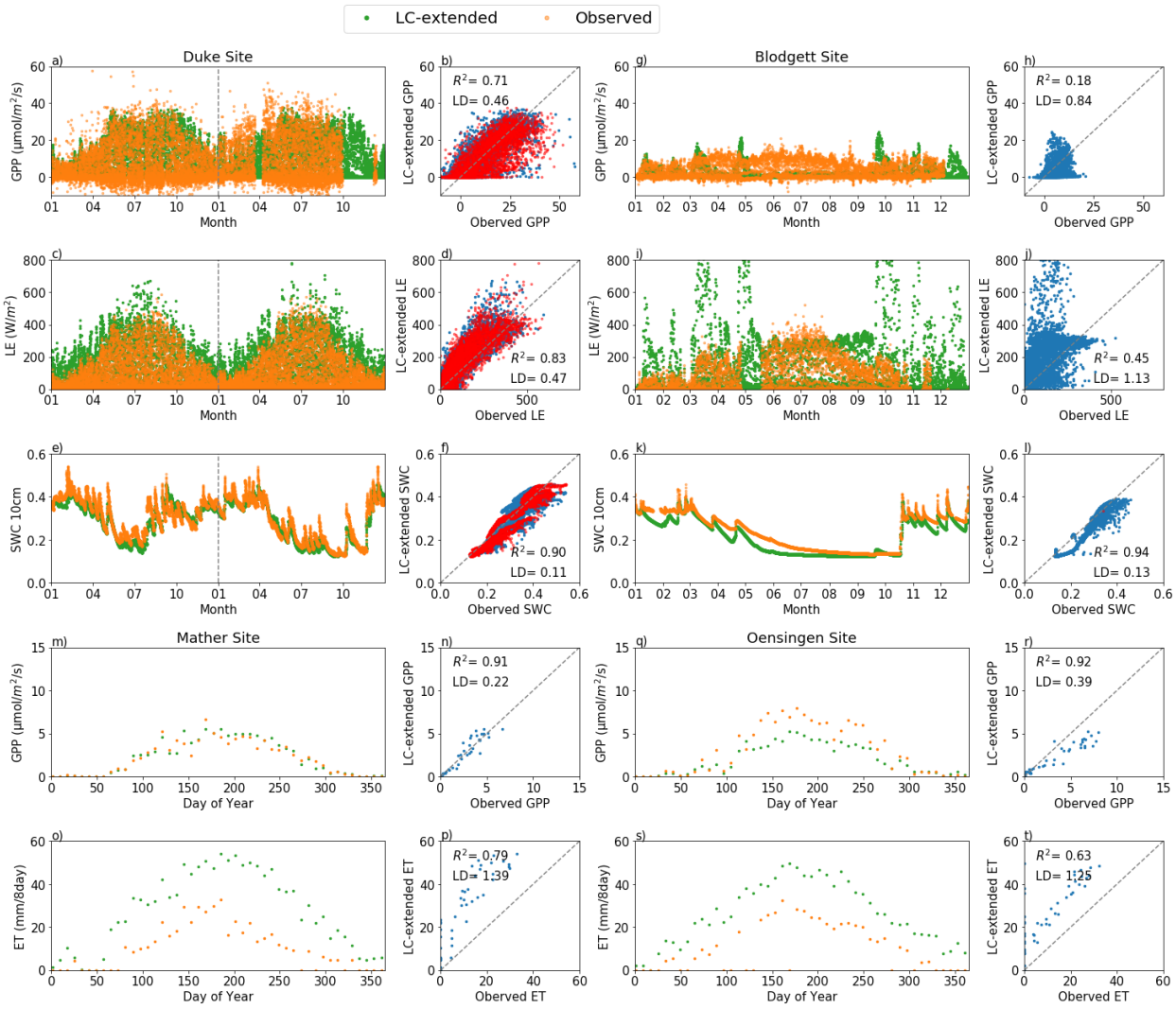

Figure 5. Comparison of the results using the same calibrated parameters by the LC-extended (green dots) with observations (orange dots). The notations used are the same as those in Fig. 2.

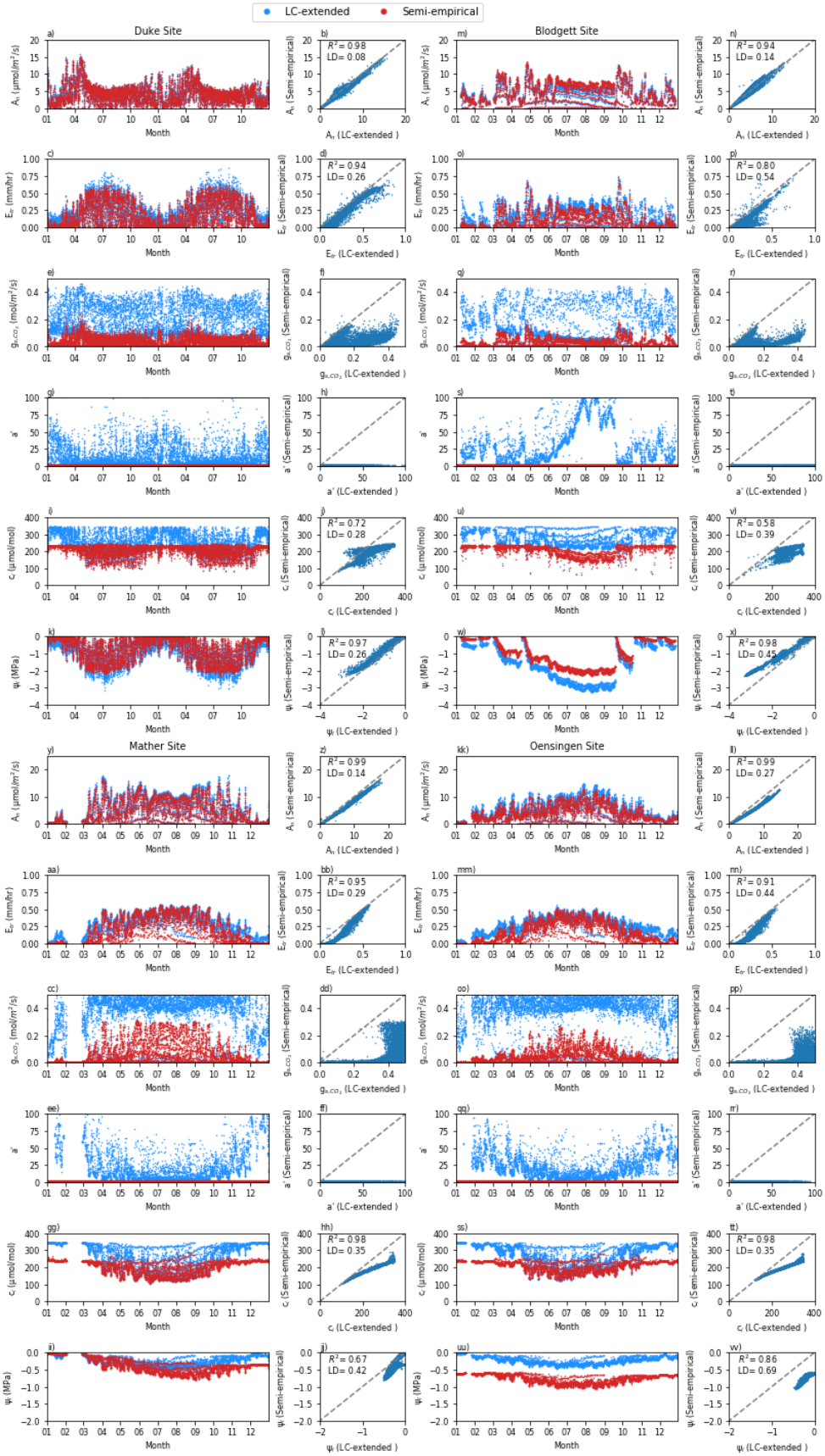
LC-extended
Semi-empirical
Duke Site
Blodgett Site
Mather Site
Oensingen Site
Month
R^2= 0.98
LD= 0.08
R^2= 0.94
LD= 0.14
R^2= 0.94
LD= 0.26
R^2= 0.80
LD= 0.54
R^2= 0.72
LD= 0.28
R^2= 0.58
LD= 0.39
R^2= 0.97
LD= 0.26
R^2= 0.98
LD= 0.45
R^2= 0.99
LD= 0.14
R^2= 0.99
LD= 0.27
R^2= 0.95
LD= 0.29
R^2= 0.91
LD= 0.44
R^2= 0.98
LD= 0.35
R^2= 0.98
LD= 0.35
R^2= 0.67
LD= 0.42
R^2= 0.86
LD= 0.69

Figure 6. Comparison of results between the LC-extended approach (green dots) and Semi-empirical (red dots) over 8AM to sunset. (a)-(l): Hourly results over two years at Duke site; (m)-(x): Hourly results over one year at Blodgett site; (y)-(jj): Hourly results over one year at Mather site; (kk)-(vv): Hourly results over one year at Oensingen Site; (a), (b), (m), (n), (y), (z), (kk), and (ll) represent carbon assimilation (A_n); (c), (d), (o), (p), (aa), (bb), (mm), and (nn) plant transpiration (E_{tr}); (e), (f), (q), (r), (cc), (dd), (oo), and (pp) CO_2 stomatal conductance (g_{s,co_2}); (g), (h), (s), (t), (ee), (ff), (qq), and (rr) empirical coefficient ($a^{'}$); (i), (j), (u), (v), (gg), (hh), (ss), and (tt) leaf CO_2 concentration (c_i); and (k), (l), (w), (x), (ii), (jj), (uu), and (vv) leaf water potential (ψ_l). The LD in the figure are calculated with a reference to the values from the semi-empirical approach.

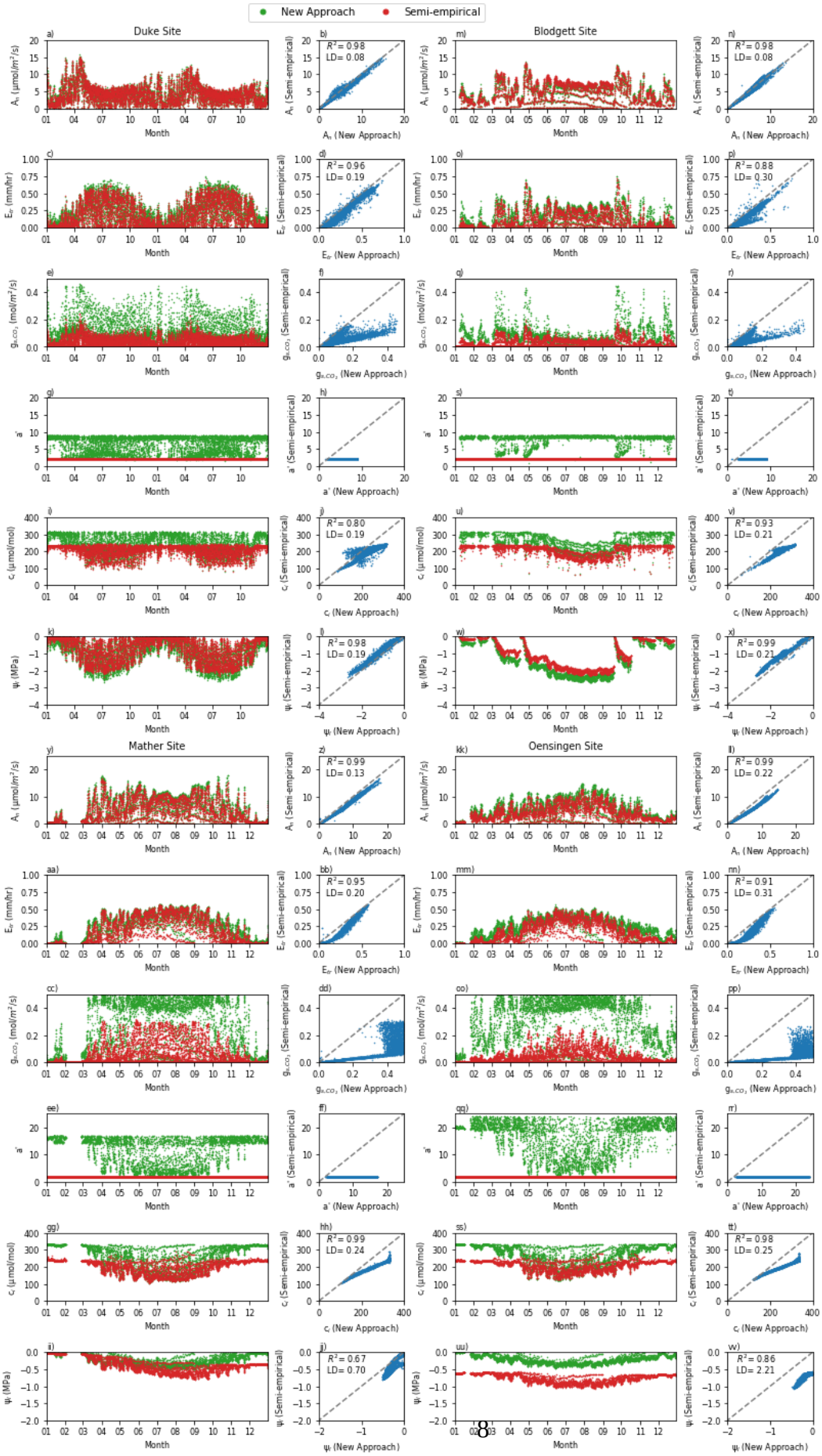
New Approach
Semi-empirical
Duke Site
Blodgett Site
Mather Site
Oensingen Site
Month
A_n (New Approach)
E_{cr} (New Approach)
g_{s,CO_2} (New Approach)
a' (New Approach)
c_i (New Approach)
ψ_l (New Approach)
R^2= 0.98
LD= 0.08
R^2= 0.96
LD= 0.19
R^2= 0.88
LD= 0.30
R^2= 0.80
LD= 0.19
R^2= 0.93
LD= 0.21
R^2= 0.98
LD= 0.19
R^2= 0.99
LD= 0.21
R^2= 0.99
LD= 0.13
R^2= 0.99
LD= 0.22
R^2= 0.95
LD= 0.20
R^2= 0.91
LD= 0.31
R^2= 0.99
LD= 0.24
R^2= 0.98
LD= 0.25
R^2= 0.67
LD= 0.70
R^2= 0.86
LD= 2.21

Figure 7. Comparison of the results between the New approach (green dots) and the Semi-empirical approach (red dots) over 8AM to sunset. (a)-(l): Hourly results over two years at Duke site; (m)-(x): Hourly results over one year at Blodgett site; (y)-(jj): Hourly results over one year at Mather site; (kk)-(vv): Hourly results over one year at Oensingen Site; (a), (b), (m), (n), (y), (z), (kk), and (ll) represent carbon assimilation (A_n); (c), (d), (o), (p), (aa), (bb), (mm), and (nn) plant transpiration (E_{tr}); (e), (f), (q), (r), (cc), (dd), (oo), and (pp) CO_2 stomatal conductance (g_{s,co_2}); (g), (h), (s), (t), (ee), (ff), (qq), and (rr) empirical coefficient ($a^{'}$); (i), (j), (u), (v), (gg), (hh), (ss), and (tt) leaf CO_2 concentration (c_i); and (k), (l), (w), (x), (ii), (jj), (uu), and (vv) leaf water potential (ψ_l). The LD in the figure are calculated with a reference to the values from the New approach.

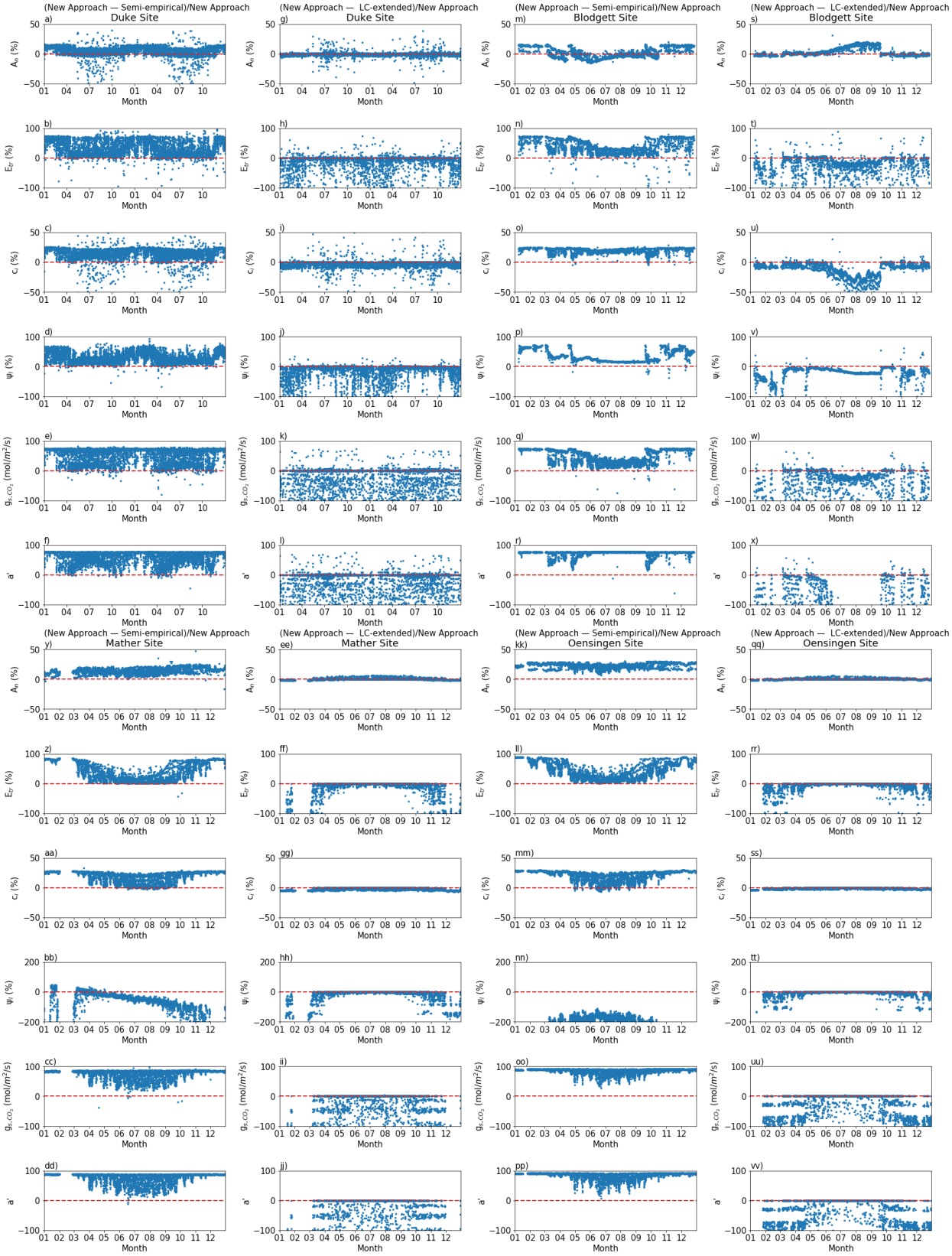
(New Approach — Semi-empirical)/New Approach
(New Approach — LC-extended)/New Approach
Duke Site
Blodgett Site
Mather Site
Oensingen Site
Month

Figure 8. Relative differences between New approach and Semi-empirical and between New approach and LC-extended over 8AM to sunset. The first 4 rows are for the two forest sites with columns 1-2 for the Duke site and columns 3-4 the Blodgett site. The bottom 4 rows are for the two grassland sites with columns 1-2 for the Mather site and columns 3-4 the Oensingen site. Some of the relative differences between New Approach and LC-Extended are outside the plot bounds and not plotted. The percentages of data outside the plots displayed are as follows: (1) for the Duke site, they are 13.6%, 5.2%, 42.2%, and 30.9% for E_{tr}, ψ_l, g_{s,co_2}, a' respectively; (2) for the Blodgett site, they are 18.4%, 7.0%, 42.4%, and 61.1%; (3) for the Mather site, they are 15.9%, 9.3%, 34.5%, and 34.4% ; and (4) for the Oensingen site, they are 7.6%, 5.2%, 18.4%, and 18.2%. For the variable ψ_l, its relative differences between New Approach and Semi-empirical have 17.7% and 59.2% data outside plot bounds for Mather and Oensingen respectively.

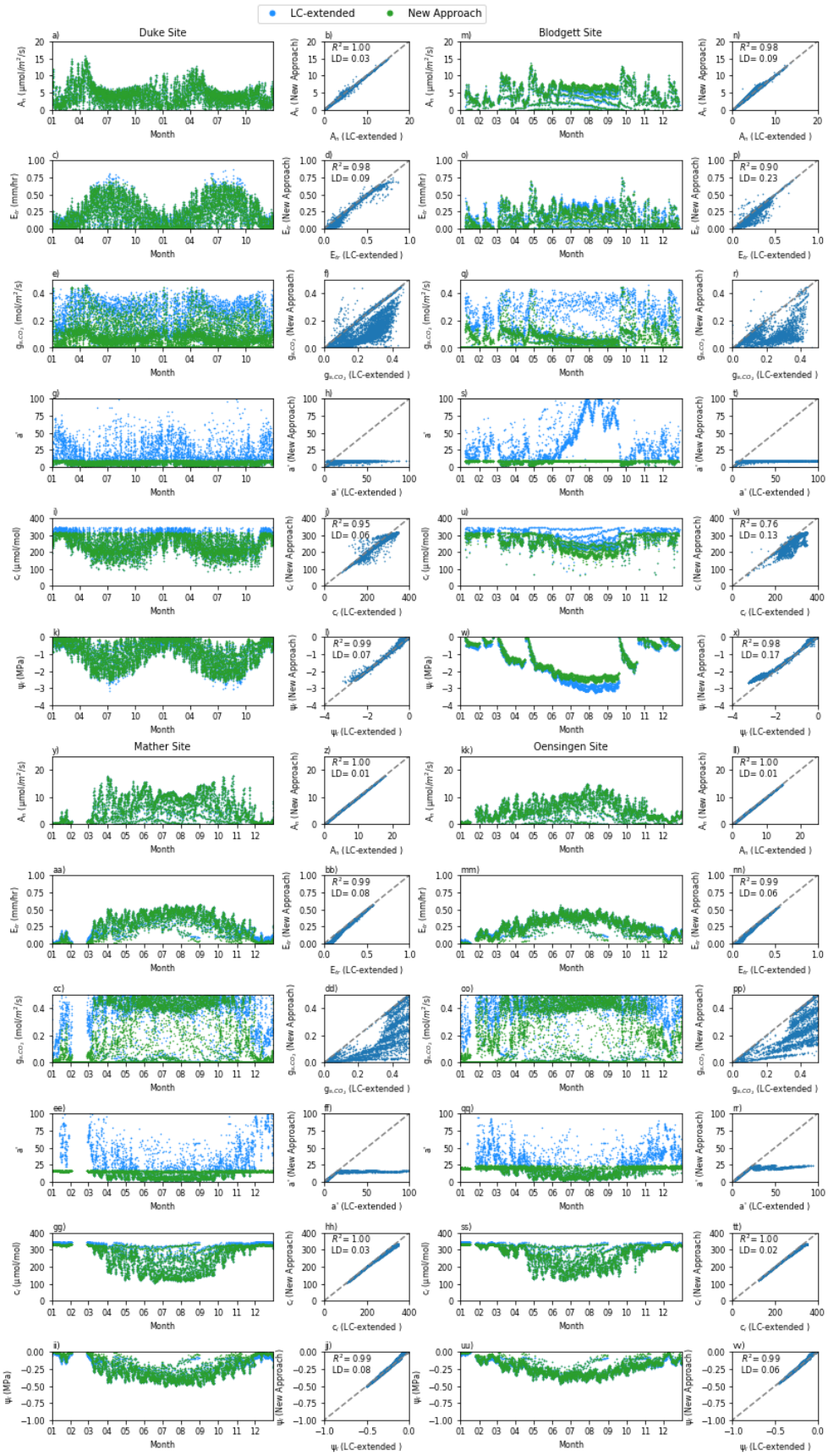
LC-extended
New Approach
Duke Site
Blodgett Site
Mather Site
Oensingen Site
Month
R^2= 1.00
LD= 0.03
R^2= 0.98
LD= 0.09
R^2= 0.98
LD= 0.09
R^2= 0.90
LD= 0.23
R^2= 0.95
LD= 0.06
R^2= 0.76
LD= 0.13
R^2= 0.99
LD= 0.07
R^2= 0.98
LD= 0.17
R^2= 1.00
LD= 0.01
R^2= 1.00
LD= 0.01
R^2= 0.99
LD= 0.08
R^2= 0.99
LD= 0.06
R^2= 1.00
LD= 0.03
R^2= 1.00
LD= 0.02
R^2= 0.99
LD= 0.08
R^2= 0.99
LD= 0.06

Figure 9. Comparison of results between the New approach (green dots) and the LC-extended approach (blue dots) over 8AM to sunset. Other notations are the same as those in Fig. 5. The LDs in the figure are calculated with a reference to the values from the New approach. Note that for *a'*, values greater than 100 from the Prentice-updated approach are not shown in the plots here. These large *a'* values account for, respectively, 1.9%, 1.1%, 9.3%, and 5.2%, for the Duke, Blodgett, Mather, and Oensingen sites.

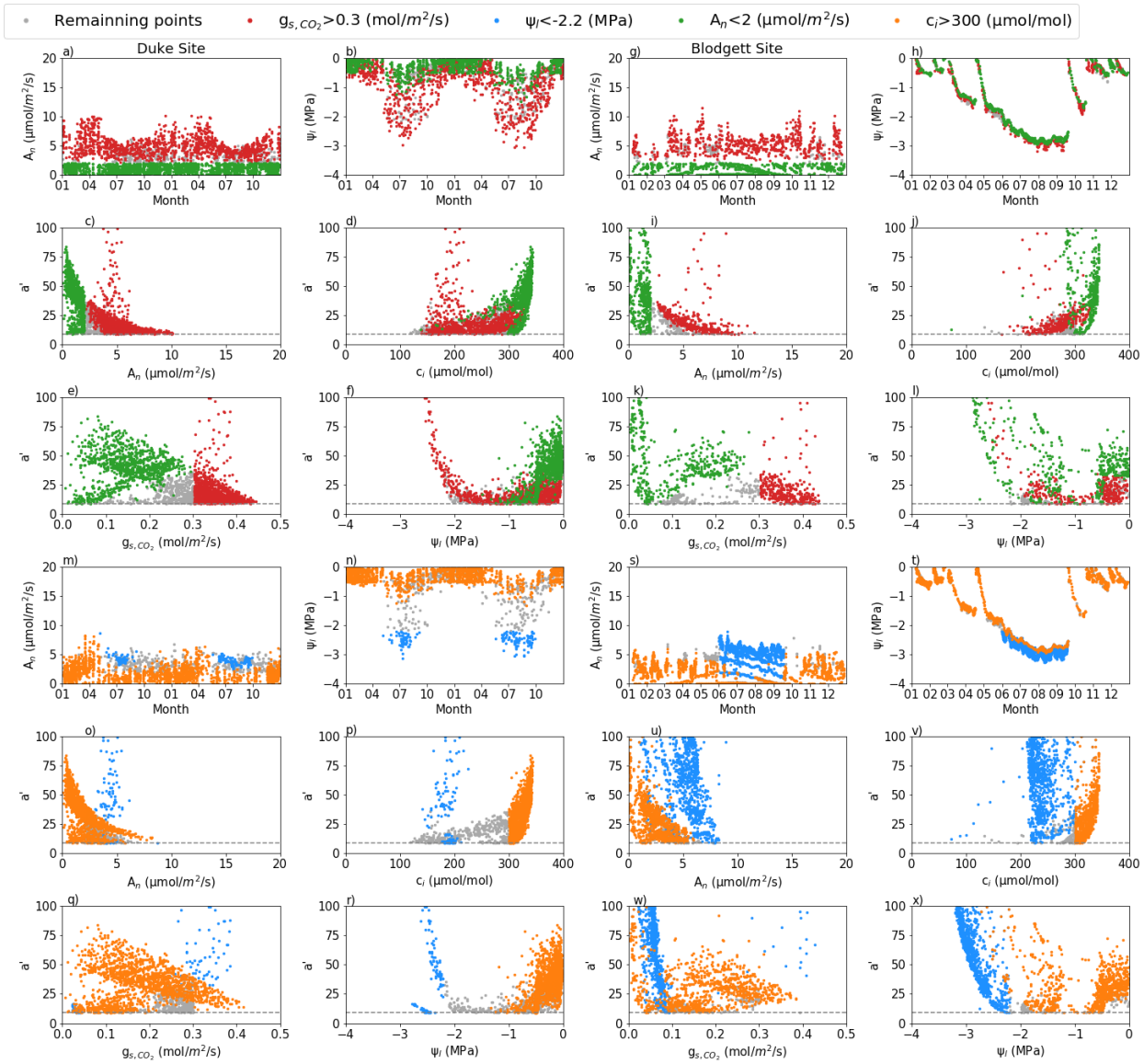

Figure 10a. Results over 8AM to sunset with the LC-extended approach at two sites in groups associated with $a' > a'_{max}$: $g_{s,co2}$ > 0.3 [mol/m^2/s] (red dots), ψ_l < -2.2 [MPa] (blue dots), A_n < 2 [μmol/m2/s] (green dots), ci > 300 (orange dots), and data points not belonging to the preceding four groups (grey dots). (a)-(f) and (m)-(r): Duke site; (g)-(l) and (s)-(x): Blodgett site; (a), (g), (m), and (s) represent carbon assimilation (A_n); (b), (h), (n), and (t) leaf water potential (ψ_l); (c), (i), (o), and (u) empirical coefficient (a') vs. carbon assimilation (A_n); (d), (j), (p), and (v) empirical coefficient (a') vs. leaf CO_2 concentration (c_i); (e), (k), (q), and (w) empirical coefficient (a') vs CO_2 stomatal conductance (g_{s,co_2}); and (f), (l), (r) and (x) empirical coefficient (a') vs. leaf water potential (ψ_l). The dotted lines in each of the a' subplots represent their respective upper bound values. Like Figure 9, a' values greater than 100 are not shown in the plots here.

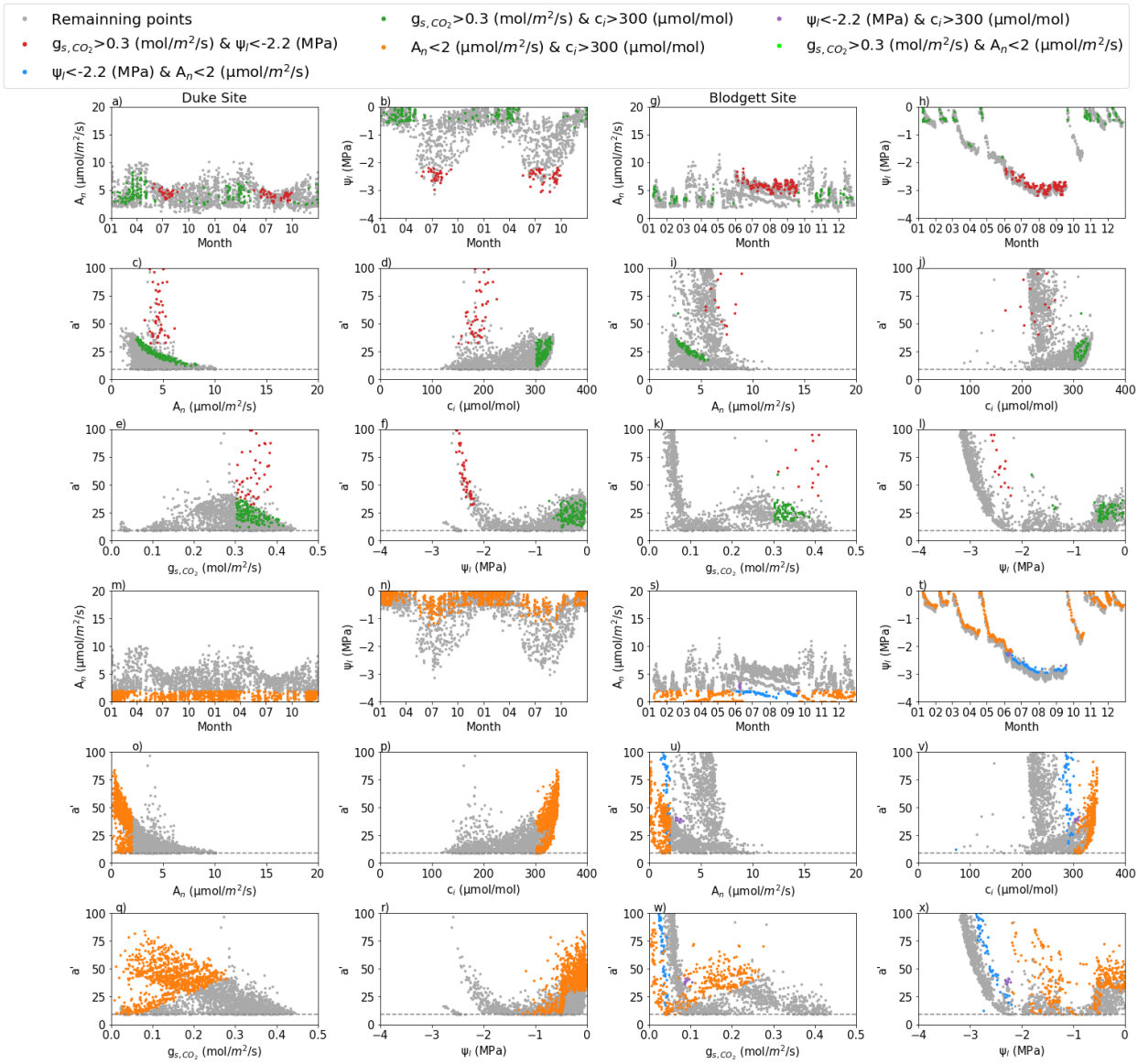

Figure 10b. The overlapping data points belonging to two groups shown in Figure 10a at the two forest sites by the LC-extended approach: overlapping between ψ_l < -2.2 [MPa] and A_n < 2 [μmol/m2/s] (blue dots), between A_n < 2 [μmol/m2/s] and c_i > 300 (orange dots), between $g_{s,co2}$ > 0.3 [mol/m^2/s] and ψ_l < -2.2 [MPa] (red dots), between $g_{s,co2}$ > 0.3 [mol/m^2/s] and c_i > 300 (green dots), between ψ_l < -2.2 [MPa] and c_i> 300 (purple dots), between g_{s,co_2} > 0.3 [mol/m^2/s] and A_n, < 2 [μmol/m2/s] (lime dots), and remaining points (grey dots). These seven groups are not overlapping with each other. (c), (i), (o), and (u) empirical coefficient (a') vs. carbon assimilation (A_n); (d), (j), (p), and (v) empirical coefficient (a') vs. leaf CO_2 concentration (c_i); (e), (k), (q), and (w) empirical coefficient (a') vs CO_2 stomatal conductance (g_{s,co_2}); and (f), (l), (r), (x) empirical coefficient (a') vs. leaf water potential (ψ_l). The dotted lines in each of the a' subplots represent their respective upper bound values (a'_{max}). Like Figure 9, a' values greater than 100 are not shown in the plots here.

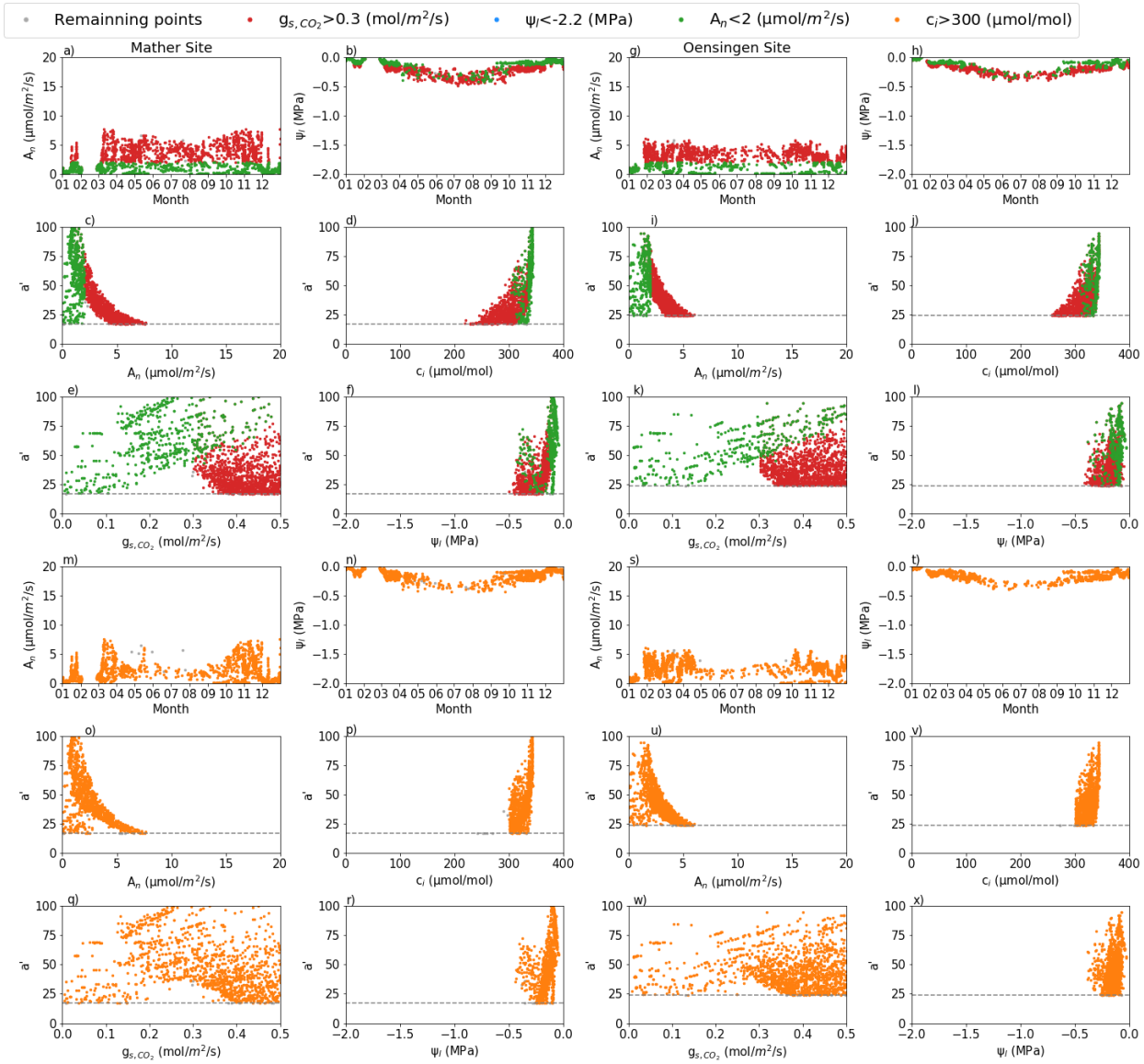

Figure 10c. Similar to Figure 10a, but for the Mather and Oensingen sites.

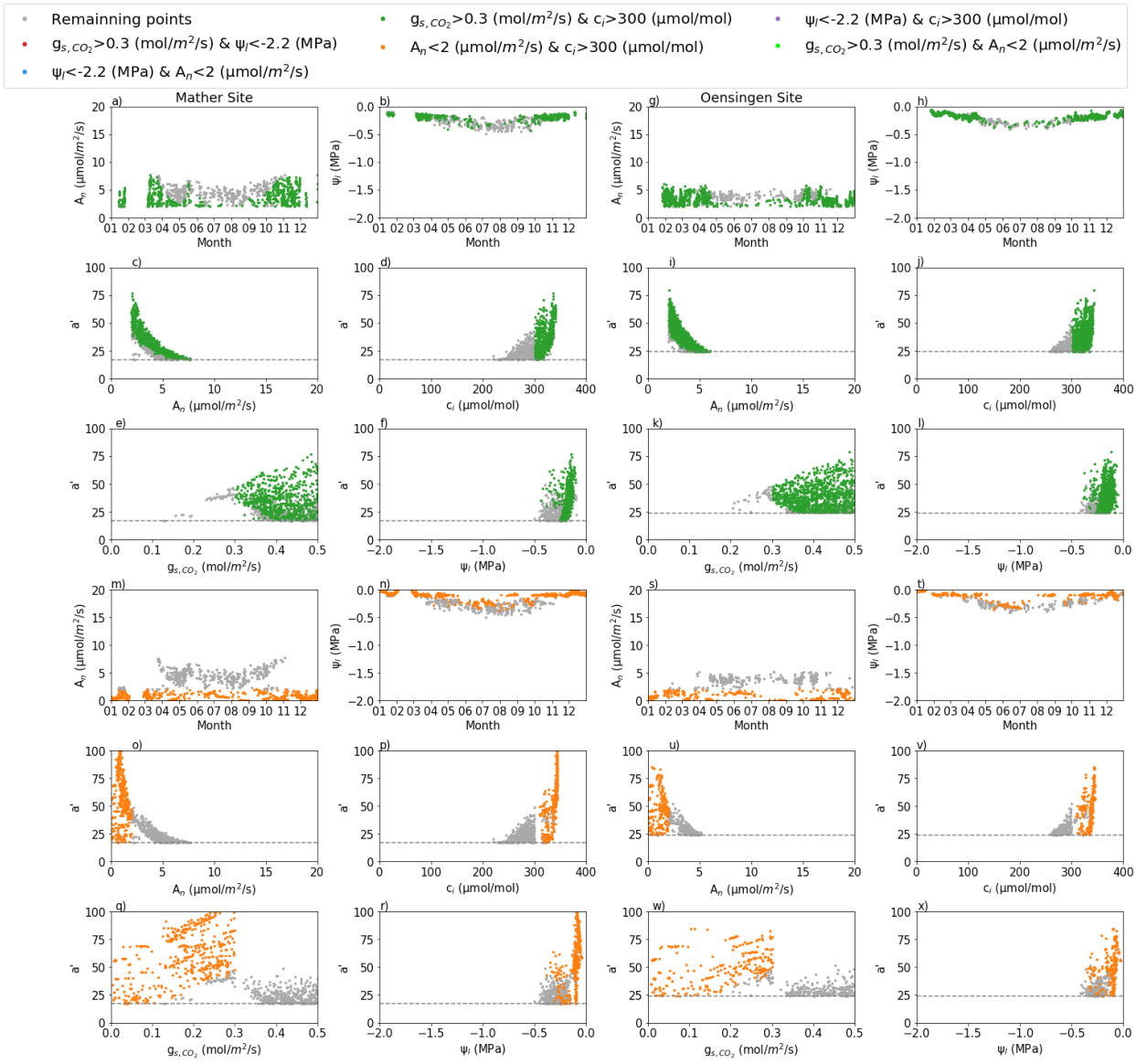

Figure 10d. Similar to Figure 10b, but for the Mather and Oensingen sites.

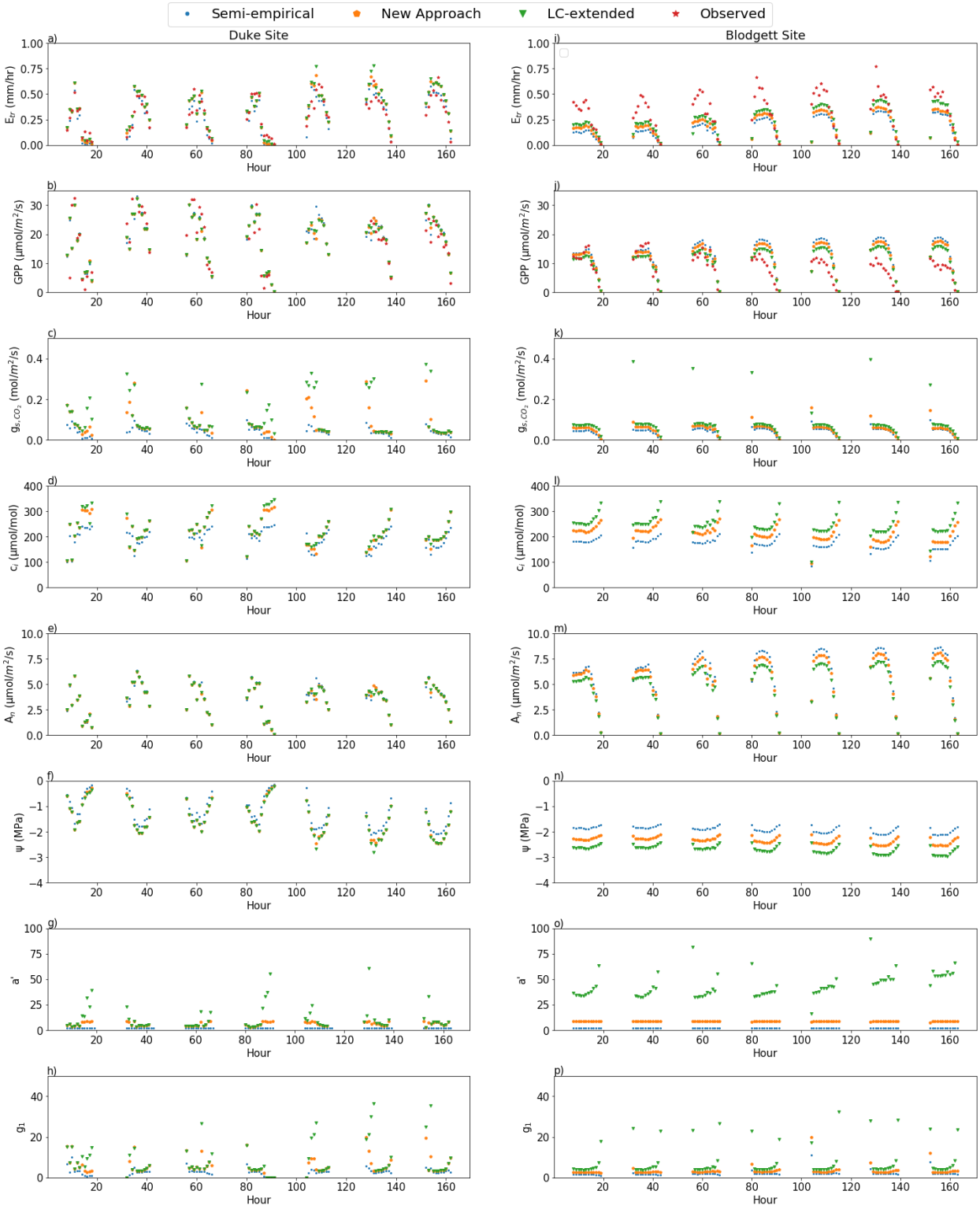

Figure 11. Comparison of one week hourly daytime results from June 30th to July 6th, 2004 among the New Approach, LC-extended, and the semi-empirical at the Duke site (a-h) and at the Blodgett site (i-p), respectively.

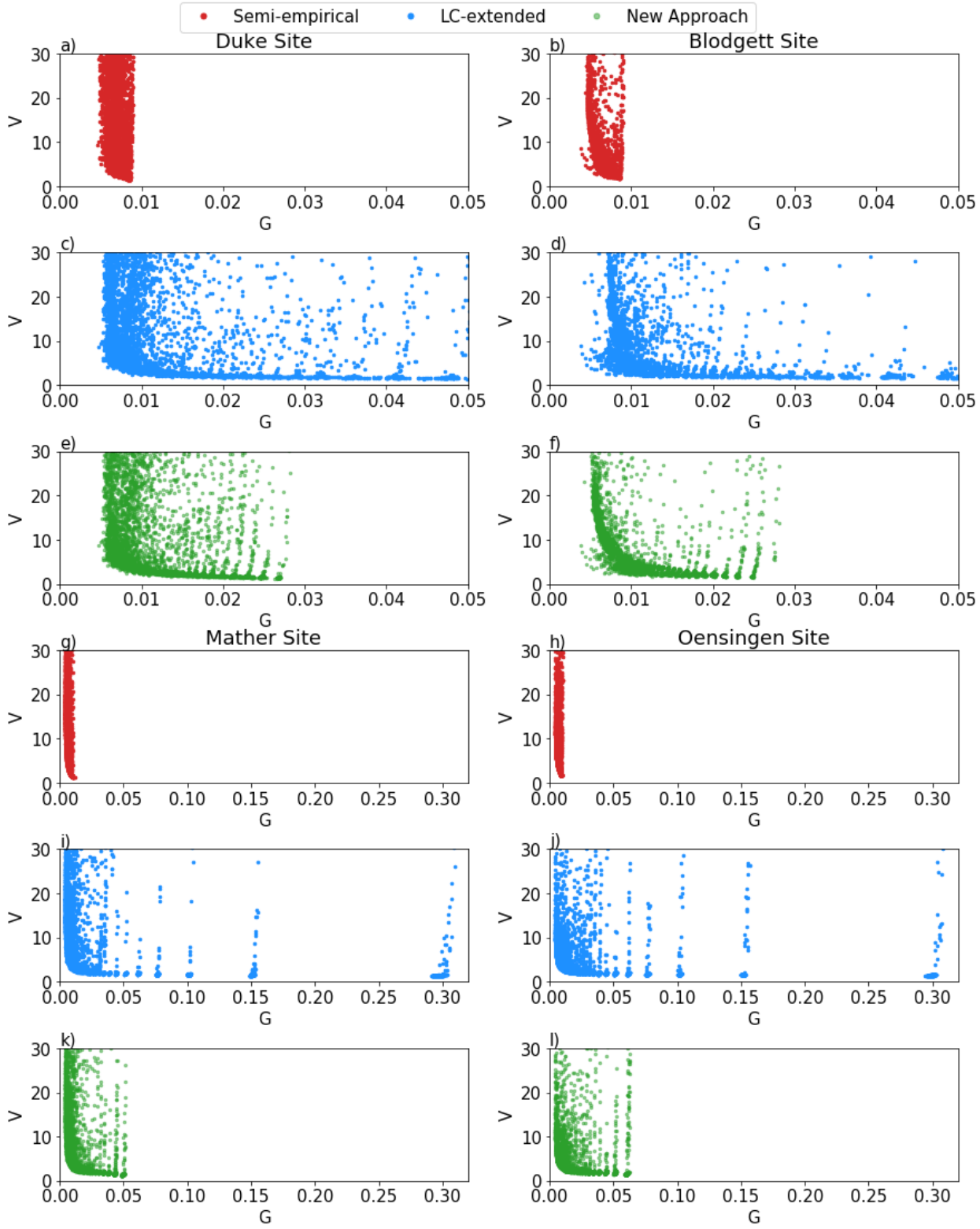

Figure 12. Comparison of the V~G G relationship among the Semi-empirical, LC-extended and New approach where $V=V_{cmax}/A_n$ and $G=g_{sab,co_2}/A_n$. (a), (c), and (e) are for the Duke site; (b), (d), and (f) the Blodgett site; (g), (i), and (k) the Mather site; (h), (j), (l) the Oensingen Site; (a), (b), (q), and (h) are from the Semi-empirical approach; (c), (d),(i), and (j) from the LC-extended approach; and (e), (f), (k), and (l) from the New approach.

Table 1. A list of model parameters calibrated for VIC+ and DHSVMm

VIC+		DHSVMm	
Parameters	Meaning	Parameters	Meaning
b	Exponent of variable infiltration capacity curve	K_{sat}	Lateral saturated hydraulic conductivity
W_s	Fraction of maximum soil moisture content of the lowest layer where nonlinear baseflow occurs	f	Exponent for change of lateral conductivity with depth (exponential decrease)
D_{smax}	Maximum velocity of baseflow	R_{omin}	Minimum stomatal resistance for the overstory
D_s	Fraction of *Dsmax* where nonlinear baseflow begins	R_{umin}	Minimum stomatal resistance for the understory
d_2	The depth of 2nd soil layer	f_c	Fraction coverage of overstory
d_3	The depth of 3rd soil layer	μ	Aerodynamic attenuation
K_{rr}	Radial hydraulic conductivity of roots per unit of root surface area	θo	Soil moisture threshold to restrict transpiration for the overstory
K_{ra}	Axial hydraulic conductivity of roots per unit area	θu	Soil moisture threshold to restrict transpiration for the understory
Common parameters to both VIC+ and DHSVMm			
R_o	Reference resistance		
C	Capacity of plant water storage		
a'	An empirical coefficient (i.e., the slope) in BBL-updated equation		

Table 2. Investigation on reasonableness of variable values: 2a-2g from *LC-extended*, 2h from *New Approach* (Only daytime, 8AM to sunset, results are included.)

2a. Percentage of results that has $a' > a'_{max}$

Duke Site	Blodgett Site	Mather Site	Oensigen Site
45.7	79.1	49.4	44.9

2b. Percentage of results in 2a satisfying four separations but overlapped groups

Site	$g_{s,co_2} > 0.3$	$\psi_l < -2.2$	$A_n < 2$	$c_i > 300$
Duke	33.3	4.8	36.0	53.6
Blodgett	16.5	46.7	22.8	42.7
Mather Site	60.1	0	47.5	81.1
Oensingen	74.5	0	33.6	86.9

Units: g_{s,co_2} in mol/m^2/s; ψ_l in MPa, A_n in μmol/m^2/s, and c_i in μmol/mol.

2c. Percentage of results in 2a that fall into non-overlapping 2-groups

Site	$g_{s,co_2} > 0.3$ $\psi_l < -2.2$	$g_{s,co_2} > 0.3$ $A_n < 2$	$g_{s,co_2} > 0.3$ $c_i > 300$	$\psi_l < -2.2$ $A_n < 2$	$\psi_l < -2.2$ $c_i > 300$	$A_n < 2$ $c_i > 300$*
Duke	2.3%	0%	5.1%	0 %	0%	32.9%
Blodgett	4.0%	0%	3.4%	2.6%	0.9%	14.7%
Mather	0 %	0%	32.6%	0%	0%	35.2%
Oensingen	0%	0%	52.4%	0%	0%	20.8%

*Note: percentages in this column exclude potentially reasonable results between 8AM and sunset.

2d. Percentage of results from 2a that belong to 3-groups

Site	$g_{s,co_2} > 0.3$, $\psi_l < -2.2$, $A_n < 2$	$g_{s,co_2} > 0.3$, $\psi_l < -2.2$, $c_i > 300$	$g_{s,co_2} > 0.3$, $A_n < 2$, $c_i > 300$	$\psi_l < -2.2$, $A_n < 2$, $c_i > 300$
Duke	0	0	0.3	0
Blodgett	0	0.1	0	5.0
Mather	0	0	9.8	0
Oensingen	0	0	10.4	0

Note: There is no result that simultaneously satisfies 4 grouping criteria.

2e. Percentage of results in 2a that belong to only one group and not to any other three groups listed in 2e

Site	g_{s,co_2}>0.3 only	ψ_l < -2.2 only	A_n < 2 only	c_i > 300 only
Duke	25.9	2.5	2.4	15.1
Blodgett	9.3	34.6	0.5	19.3
Mather	18.9	0	0	2.1
Oensingen	13.1	0	0	2.3

2f. Percentage of results after further grouping of each single group from 2e

Site	g_{s,co_2}>0.3 only		ψ_l < -2.2 only	A_n < 2 only	c_i > 300 only
	ci < 200 Etr≥ 0.1	Etr < 0.1 200 < ci≤ 300	200 < ci ≤ 300 g_{s,co_2} ≤ 0.1	200 < ci ≤ 300	5≥An≥2 0.1<gs≤0.3
Duke	30. 6	22.1	5.1	94.6	94.5
Blodgett	0.4	42.5	99.0	100	91.6
	200 < ci ≤ 300 5≥An≥2, Etr>0.2				0.2<gs≤0.3 3≥An≥ 2
Mather	64.2				83.3
Oensingen	93.5				97.7

Units: Etr in mm/hr

2g. Percentage of results in 2a that are deemed unreasonable

Site	*Unreasonable
Duke	*70.9
Blodgett	87.1
Mather	91.5
Oensingen	98.1

*= Sum of each cells in the row of Table 2c + sum of each cell of each row of Table 2d + (each cell of Table 2e) multiplied by (its corresponding unreasonable values occurrence in Table 2f)

For example, the percentage of unreasonable at the Duke site is:
* = (2.3%+5.1%+32.9%)+0.3%+25.9%×(30.6%+22.1%)+2.5%×5.1%+
2.4%×94.6%+15.1%×94.5% = 70.9%

2h. Percentage of *New Approach* results that are deemed unreasonable which all have $a' <= a'_{max}$

Site	Unreasonable
Duke	9.3
Blodgett	3.2
Mather	3.9
Oensingen	6.2

Table 3. Percentage of $0 < G \leq 0.05$ and $0 < V \leq 30$ obtained from each approach

Site	Approach	$0< G \leq 0.05$ [mol·μmol^{-1}]	$0< V \leq 30$ [mol·mol^{-1}]
Duke	New Approach	100	69
	LC-extended	91	69
	Semi-empirical	100	69
Blodgett	New Approach	100	90
	LC-extended	89	89
	Semi-empirical	100	90
Mather	New Approach	98	95
	LC-extended	75	95
	Semi-empirical	100	92
Oensingen	New Approach	86	98
	LC-extended	79	98
	Semi-empirical	100	97