Development of a multi-layer canopy model for E3SM Land Model with support for heterogeneous computing

Gautam Bisht[1], William J. Riley[2], and Richard T. Mills[3]

[1]Atmospheric Sciences & Global Change Division, Pacific Northwest National Laboratory, Richland, WA 99354, USA

[2]Climate & Ecosystem Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

[3]Mathematics and Computer Science Division, Argonne National Laboratory, Lemont, IL 60439, USA

Key Points:

- A multi-layer canopy model with support for heterogeneous is developed.
- Performance of different stomatal conductance models is evaluated against observations of surface energy fluxes.
- The changes in different environmental variables impact multiple canopy processes differently.

Corresponding author: Gautam Bisht, gautam.bisht@pnnl.gov

Abstract

The vertical structure of vegetation canopies creates micro-climates, which can substantially affect ecosystem responses to climate change. However, the land components of most Earth System Models, including the Energy Exascale Earth System Model (E3SM), typically neglect vertical canopy structure by using a single layer big-leaf representation to simulate water, CO_2, and energy exchanges between the land and the atmosphere. In this study, we developed a standalone Multi-Layer Canopy Model (MLCMv1) for the E3SM Land Model (ELM) to resolve the micro-climate created by vegetation canopies. The support for the heterogeneous computation architectures is included by using the Portable Extensible Toolkit for Scientific Programming. The numerical implementation of ELM-MLCMv1 was verified against CLM-ml_v1 for a month-long simulation using data from the Ameriflux US-University of Michigan Biological Station (US-UMB) site. Model structural uncertainty was explored by performing control simulations for five stomatal conductance models (SCMs). All SCMs after calibration were able to accurately match observations of sensible and latent heat flux, though the bias of the three SCMs with plant hydrodynamics (PHD) was slightly lower than that of two SCMs without PHD. Additionally, six idealized simulations were performed to study the impact of environmental variables on canopy processes. All SCMs agreed on the direction of simulated changes in canopy processes due to the changes in these environmental variables. ELM-MLCMv1 achieves a speedup of 25-50 times when comparing performance on a GPU relative to a CPU. This study provides the first necessary model development for including the representation of vertical canopies within ELM.

Plain Language Summary

Vegetation in Earth System Models is typically represented by a sunlit and shaded *big-leaf* at a certain height above the ground surface. Such a simplified representation of vegetation ignores key terrestrial processes arising due to the vertical structure of vegetation canopies. In this work, we developed a multi-layer canopy model for the Energy Exascale Earth System Model that can efficiently use the upcoming class of supercomputers. The new model is verified against an existing benchmark model and site-level observations. Multiple configurations of the new model are able to accurately reproduce observations of surface energy fluxes.

1 Introduction

The vertical structure of a vegetation canopy creates micro-climates. The vertical variations of leaf area density (LAD) creates turbulence within vegetation canopies that lead to vertically-varying profiles of wind speed, temperature, and scalar concentrations, such as water vapor and CO_2 (Harman & Finnigan, 2007). Wind speed decreases within canopies and interior daytime temperature in dense forest canopies can be several degrees cooler than the temperature above the canopy (Lenoir et al., 2013). Under unstable atmospheric conditions and during daytime, CO_2 is higher above the canopy and decreases within the canopy due to the uptake of CO_2 by plants through photosynthesis and increases again near the soil surface due to soil respiration (Scanlon & Albertson, 2001). Under similar unstable atmospheric conditions, the water vapor in the canopy and near the soil surface is higher than in the atmosphere due to evapotranspiration during the day. Vegetation canopies intercept and scatter radiation, thereby creating heterogenous vertical light environments (Fisher et al., 2018). Vegetation canopies have vertical gradients of leaf nitrogen and phosphorous per unit leaf area with higher nutrients at the top of the canopy (Bond et al., 1999). During the day, vertical gradients of water potential lead to water transport from the soil to the leaves (Fatichi et al., 2016).

The vertical structure of canopies can affect ecosystem responses to climate change. Drought and warmer temperatures have increased forest mortality by causing low soil moisture and high atmospheric vapor pressure deficit (VPD) (Allen et al., 2010; Stovall et al., 2019). Tall trees with high LAD, low hydraulic conductivity, and strong stomatal regulation in response to leaf water potential are at a higher mortality risk in the future (McDowell & Allen, 2015). Forrest canopies can create micro-climates that buffer the understory from macro-climate temperatures (Zellweger et al., 2020). Using paired measurements, De Frenne et al. (2019) showed that vegetation cover decreases the maximum temperature within the understory, while simultaneously increasing the minimum temperature. They reported that the difference in temperature between above and within the canopy becomes larger as the above-canopy air temperature increases, and this temperature difference can be of greater magnitude than the increase in land surface temperature over the last century.

The vertical structure of canopies is neglected in most current global Land Surface Models (LSMs), although a few next-generation models are addressing this shortcom-

ing. The vegetation canopy in current LSMs is typically represented by a single *big-leaf* that, along with a surface soil layer, exchanges water, energy, and CO_2 fluxes with the atmosphere. The Monin–Obukhov similarity theory (MOST) is used to describe the vertical transport of momentum, energy, and water vapor betweeen the canopy and overlying atmosphere. The variation of light is typically accounted for by including a sunlit and shaded big-leaf, both of which are assumed to be at the same canopy height. Plant hydrodynamics (PHD) has been included in many site-level and regional ecosystem models (Sperry et al., 1998; Bohrer et al., 2005; Grant et al., 2006; Xu et al., 2016; Agee et al., 2021), but only recently PHD has been included in global LSMs (Kennedy et al., 2019; Fang et al., 2021). Recently, multi-layer canopy models (MLCMs) have been added in LSMs (G. Bonan et al., 2014; Ryder et al., 2016; G. B. Bonan et al., 2018; Longo et al., 2019). MLCMs vertically resolve radiation (solar and longwave), photosynthesis, and leaf temperatures within canopies, though the canopy air space (CAS) can be represented with one or multiple layers. The standard boundary-layer flux-gradient relationships used in the MOST turbulence scheme are violated within the canopy and in the roughness sublayer (RSL) near the top of the canopy. MLCMs employ turbulence schemes that are valid for the turbulent transport of momentum, energy, and scalars within the canopy and RSL. The turbulence scheme of Massman and Weil (1999) is used in ED2.2 (Longo et al., 2019) and ORCHIDEE-CAN v1.0 (Chen et al., 2016), while the Harman and Finnigan (2007, 2008) turbulent scheme is used in CLM-mlv0 (G. B. Bonan et al., 2018) and CLM-ml v1 (G. B. Bonan et al., 2021).

Multiple stomatal conductance models (SCMs) have been developed to account for the stomatal control of photosynthesis. SCMs use semi-empirical relationships or optimization theory and may additionally account for PHD. A commonly used SCM is the Ball-Berry (BB) model that uses a semi-empirical approach for modeling the dependence of stomatal conductance on VPD by using fractional humidity at the leaf surface (Ball et al., 1987). Leuning (1995) modified the BB model by adjusting for the CO_2 compensation point and directly included dependence on VPD. Alternatively, optimization theory has also been used to model stomatal conductance, g_s, in which stomata maximize the carbon gain relative to a cost associated with the carbon gain. The commonly used objective functions in optimization-based SCMs are water-use efficiency (WUE = $\Delta A_n/\Delta E$) and intrinsic water-use efficiency (iWUE = $\Delta A_n/\Delta g_s$), defined as the marginal carbon gain (ΔA_n) with respect to water loss (ΔE) and stomatal opening (Δg_s), respectively (Cowan

& GD, 1977; Katul et al., 2010; G. Bonan et al., 2014; Buckley et al., 2017). A few semi-empirical and iWUE/WUE SCMs additionally include the control of PHD on g_s (Manzoni et al., 2011; Christoffersen et al., 2016; Mirfenderesgi et al., 2016). G. Bonan et al. (2014) developed an optimization-based g_s model that simultaneously satisfied two constraints of iWUE and leaf water potential to be greater than a minimum threshold. An objective function other than WUE and iWUE has also been used in optimization-based SCMs. Wang et al. (2020) evaluated 10 optimization-based SCMs for seven criteria that tested for a valid mathematical solution and realistic biological stomatal responses to the environment that were consistent with empirical studies. They found that only four SCMs met all seven criteria. Two of those four models used the loss of xylem hydraulic conductivity as the marginal cost for carbon gain (Sperry et al., 2017; Wang et al., 2020), while the other two models used non-stomatal limitations on photosynthesis as the marginal cost for carbon gain (Hölttä et al., 2017; Dewar et al., 2018). Stomatal conductance models that use optimization theory and include a hydraulic constraint have shown superior skill in reproducing observations (G. Bonan et al., 2014; Eller et al., 2020; Wang et al., 2020, 2021)

Development of new process models within LSMs should include support for the upcoming heterogeneous high-performance architectures and have a flexible core to explore model structural uncertainty. High-performance computing architectures are transitioning from homogeneous CPUs-based systems to heterogeneous systems that include CPUs and GPU accelerators. Currently, there are many GPU programming models such as Kokkos (Edwards et al., 2014), RAJA (Beckingsale et al., 2019), CUDA (NVIDIA et al., 2020), ROCm (ROCm, 2019), and OpenCL (Stone et al., 2010); and it is not feasible for LSMs to include native support for all of the programming models. The profusion of these programming models has led the Exascale Computing Project (ECP) to develop *performance portable* programming models that aim to provide high-level abstractions for mapping code to appropriate hardware. The Portable Extensible Toolkit for Scientific computation (PETSc) library, supported by ECP, provides scalable numerical solvers (Balay et al., 2021) and allows application code (e.g. LSMs) to be agnostic of the underlying computer architectures (Mills et al., 2021). Additionally, PETSc provides a flexible framework to assemble individual physics problems into a single tightly coupled multiphysics system via `DMComposite` (Brown et al., 2012). The `DMComposite` has been previously used in the open-source, multi-physics problem (MPP) library to flexibly couple multiple bio-

physics problems within multiple domains e.g. transport of water through soil-root-xylem and thermal advection-diffusion in the subsurface (Bisht & Riley, 2019).

Objectives of this work include developing a model that resolves micro-climate created by the vertical structure of the vegetation canopy and providing support for heterogeneous computing architectures. In this study, a standalone MLCM is developed for ELM, hereafter referred to as ELM-MLCMv1. The new model is based on CLM-ml v1 and includes three additional optimization-based SCMs as compared to CLM-ml v1. PETSc is used in ELM-MLCMv1 to provide numerical solutions of discretized equations, flexible model coupling, and performance portability. First, the numerical implementation of the model was verified by comparing results against CLM-ml v1 for a month-long simulation performed using data from the AmeriFlux station at the University of Michicigan Biological (UMB). Second, the impact of SCM model structure on the simulation of biophysical processes was studied. Five SCMs, calibrated against site-level observations of sensible and latent heat fluxes, were then used to study the response of vegetation under three idealized climate change (CC) scenarios: (i) higher air temperature, (ii) higher atmospheric CO_2 concentrations, and (iii) drier soil moisture. Third, the performance portability of ELM-MLCMv1 was studied by performing simulations across a range of problem sizes on two GPU systems that used NVIDIA and AMD GPUs. The model developed in this study provides the foundation for ELM to undertake research to better understand the role of vertical canopy heterogeneity on the ecosystem climate sensitivity.

2 Model description

The ELM-MLCMv1 (Figure 1) accounts for sunlit and shaded leaves at each canopy level and includes the following four sub-models: (i) shortwave and longwave radiation, (ii) stomatal conductance, (iii) RSL parameterization, and (iv) transport of heat and water vapor in the CAS. The radiation model and RSL parameterization lump sunlit and shaded leaves as one at each canopy layer, while the other two sub-models explicitly account for sunlit and shaded leaves. Similar to CLM-ml v1, it is assumed that water from the soil is transported to each leaf via unconnected xylems (Figure 1d). The vertical profiles of the leaf and stem of a plant are described using a beta distribution (see Text S1 in the Supporting Information).

2.1 Shortwave and longwave radiation sub-model

The shortwave radiation sub-model is based on Norman (1979) and accounts for direct beam and diffuse radiation. At each canopy level, the sunlit and shaded leaves are first lumped into a single leaf. The lumped leaf intercepts a fraction of the beam and diffuse radiation incident on the leaf surface. The intercepted radiation is absorbed or scattered as diffuse radiation. The radiation at each canopy position thus consists of (i) non-intercepted beam radiation, (ii) non-intercepted diffuse radiation, and (iii) intercepted beam and intercepted diffuse radiation that is scattered upward and downward. The shortwave radiation model resolves radiation in the visible and near-infrared bands. For N canopy layers, the discretized model leads to a linear system of equations with $(N+1) \times 2$ unknowns corresponding to upward and downward radiation at each canopy level and the soil surface. The simulated radiation absorbed by a lumped leaf at each canopy level is disaggregated for sunlit and shaded leaves. It is assumed that shaded leaves absorb diffuse radiation, while sunlit leaves absorb beam and diffuse radiation. The details about the shortwave radiation model are provided in Text S2 in the Supporting Information.

The longwave radiation sub-model is also based on Norman (1979). Unlike the shortwave radiation model, the longwave radiation model only includes diffuse radiation. The emitted longwave radiation of sunlit and shaded leaves is weighted by their fractions at each canopy layer. The sunlight and shaded leave absorb equal longwave radiation per unit leaf area. The discretized model also leads to a linear system of equations with $(N+1) \times 2$ unknowns corresponding to upward and downward radiation at each canopy level and the soil surface. The details about the longwave radiation model are provided in Text S3 in the Supporting Information.

2.2 Stomatal conductance sub-model

Photosynthesis is modeled assuming the biological demand for CO_2 uptake is met by the diffusive transport of CO_2 from the CAS into the stomata. In such a modeling approach, there are two equations with three unknowns. The two equations describe the biological uptake and diffusive transport of CO_2, while the three unknowns include net CO_2 assimilation, CO_2 concentration within the stomata, and stomatal conductance (see Text S4 and Text S5 in the Supporting Information). Thus, in order to close the system

of equations, a third equation for stomatal conductance is needed. ELM-MLCMv1 includes five previously published SCMs that use semi-empirical relationship or optimization theory and may additionally account for PHD (see Table 1). All five SCMs result in a nonlinear equation that is solved for sunlight and shaded leaves at each canopy level. A detailed description of all supported SCMs and their numerical solutions are presented in Text S6 and S7 in the Supporting Information. The model for the transport of CO_2 from the CAS into the stomata additionally includes parameterizations for the leaf boundary layer, which is presented in Text S8 in the Supporting Information.

2.3 Turbulent transport sub-model

Unlike big-leaf canopy models, MLCMs require models to describe vertical profiles of scalars (i.e. wind, air temperature, and water vapor) within the canopy. ELM-MLCMv1 uses the model of Harman and Finnigan (2007, 2008) that provides a unified theory of turbulent transport within and above the canopy. The scalars within RSL of the canopy follow an exponential profile, while the log profile above the canopy deviates from the MOST predicted profile. Continuity of scalar values is assumed at the canopy height which leads to a system of equations that is solved iteratively. A detailed description of the RSL model is presented in Text S9 in the Supporting Information. Similar to the radiation models, the RSL model lumps sunlit and shaded leaves together at each canopy level.

2.4 Multi-layer canopy air space sub-model

ELM-MLCMv1 simulates the transport of energy and water vapor within the CAS using a one-dimensional vertical diffusion model. The conservation of energy for sunlit and shaded leaves is modeled explicitly at each canopy level. The sensible and latent heat fluxes from the leaves tightly couple CAS water and energy transport equations with leaf energy balance equations leading to a system of four tightly coupled equations. The water vapor and energy flux between the leaf and CAS take into account stomatal and leaf boundary conductances, while the vertical fluxes of water and energy account for aerodynamic turbulence within and above the canopy. The equations are solved using finite-volume spatial discretization and an implicit time integration scheme. Similar to CLM-ml v1, the non-linear source and sink terms of water and energy are linearized lead-

ing to a linear system of equations. A detailed description of the MLC sub-model is presented in Text S10 in the Supporting Information.

3 Numerical implementation

The discretized equations of ELM-MLCMv1's four sub-models result in linear and non-linear equations that are solved using PETSc's `KSP` and `SNES` solvers, respectively. Unlike CLM-ml v1, ELM-MLCMv1 allows for multiple vegetation canopies to simultaneously interact with a CAS (Figure 2). The `DMComposite` of PETSc, a special type of the `DM` ("Domain Management") abstraction in PETSc, is used to provide a flexible framework to solve equations for the transport of energy and water within CAS which are tightly coupled to the energy balance equation for sunlit and shaded leaves. Similar to the previous use of `DMComposite` in the MPP library (Bisht & Riley, 2019), the multi-layer CAS model includes four governing equation (`GE`) objects that correspond to GEs for (1) CAS temperature, (2) CAS vapor pressure, (3) sunlit leave temperature, and (4) shaded leaf temperature. The four `GE` objects make up a system of equations `SoE` object that is used to obtain the numerical solution via `DMComposite`. The detail about the use of `GE` and `SoE` is presented in Bisht and Riley (2019). A schematic representation of two vegetation canopies sharing a CAS is shown in Figure 2a, while the resulting discretized linear system, $Ax = b$, with four `GE`s is shown in Figure 2b. The entries of `GE` corresponding to air temperature and air water vapor within the x and b vector are shown in blue and magenta, respectively The entries of `GE`s corresponding to the sunlit leave (in light green) and shaded leave (in dark green) includes unknown temperature values for the two vegetation canopies. The cyan colored sub-block of matrix A indicates non-zero values. The `DMComposite` allows the flexibility to grow the linear system when multiple vegetation canopies share a CAS.

An important detail associated with using `DMComposite` has a potentially large impact on performance. To avoid very high costs due to repeated memory allocation and copying during assembly, accurate preallocation—that is, giving PETSc an estimate of how many nonzero elements will be present in each row—for sparse matrices in PETSc is required. When a `DM` is being used to manage the computational domain, PETSc can usually use the information in the `DM` to automatically perform the appropriate preallocation, but when `DMComposite` is being used to glue many different subdomains together, there is not always a good way to determine, a priori, the nonzero pattern of the

full coefficient matrix. It is important to accurately characterize this patter since underestimating the number of nonzero elements can dramatically increase the time required for matrix assembly, and overestimating can easily exhaust available memory resources. We address this issue in ELM-MLCMv1 by using two passes through the routines that calculate the matrix elements: In the first pass, we perform no actual computation of matrix entries and "assemble" a "matrix" of type `MATPREALLOCATOR`, a dummy `Mat` type that is used simply to count nonzero elements. With the accurate counts made in this first phase, we then preallocate storage for the actual coefficient matrix, which is assembled during the second pass. This simple optimization is absolutely necessary; without it, the matrix assembly is several orders of magnitude more time-consuming than it should be—so slow as to make using the code impracticable.

Besides providing data management tools such as `DMComposite` and implementations of many state-of-the-art nonlinear and linear solver algorithms, utilizing PETSc also provides a means to offload to GPU accelerator devices the most computationally expensive portion of the ELM-MLCMv1 model: the solution of the large linear system, which comprises at least half of the total execution time of the model when running executing purely on CPUs. The design of PETSc makes this transparent to the application developer; changing from CPU to GPU execution simply requires specifying the appropriate runtime options. PETSc follows a delegation pattern approach to object-oriented design, in which each PETSc object is an instance of a class whose specific data structures and functionality are determined by choosing a delegated implementation type at runtime. GPUs can be used for linear algebra operations by simply choosing appropriate delegated types for the `Vec` and `Mat` objects that use the desired GPU back-end for executing vector and matrix operations, respectively. Because higher level components of PETSc (`KSP` linear solvers, `SNES` nonlinear solvers, `TS` time steppers, etc.) employ `Vec` and `Mat` operations for the bulk of their underlying computations, most of the computation in these solvers will occur on the GPU. The current back-ends supported by PETSc's `Vec` and `Mat` classes include CUDA (NVIDIA's native model), HIP (AMD's "Heterogeneous Interface for Portability" model, which we use to target their ROCm GPU platform), and Kokkos (a performance portability framework that targets NVIDIA, AMD, and Intel GPUs as well as parallel CPU execution).

4 Methodology

4.1 Experiment design for model verification

We verified the numerical implementation of ELM-MLCMv1 by comparing results against CLM-ml v1. Simulations were performed using atmospheric data from the Ameriflux US-University of Michigan Biological Station (US-UMB), which is a temperate deciduous broadleaf forest site (Schmid et al., 2003). The simulations follow the model configuration of G. B. Bonan et al. (2021) and use 92 vertical layers with a vertical spacing of 0.5 m. The CAS is occupied by a single tree with an LAI and SAI of 4.15 m^2 m^{-2} and 0.90 m^2 m^{-2}, respectively (Figure 3). Half-hourly atmospheric forcing data from July 2006 is used that included (1) downwelling direct and diffused solar radiation in the near-infrared and visible bands, (2) downwelling longwave radiation, (3) near-surface air temperature, (4) near-surface specific humidity, (5) near-surface atmospheric pressure, (6) near-surface wind speed, (7) near-surface atmospheric CO_2 concentration, and (8) near-surface atmospheric O_2 concentration. The time series of vertically-resolved soil moisture, surface temperature, and surface albedo was prescribed as boundary conditions. Simulations were performed for the BB and Medlyn SCMs. The relative error for ELM-MLCMv1 with respect to CLM-ml v1 was computed at each model time steps to verify the numerical implementation.

4.2 Experiment design for studying the impact of environmental variables on canopy processes

After establishing confidence in the numerical implementation of ELM-MLCMv1, a series of seven simulations (Table 2) were designed to investigate the impact of environmental variables on canopy processes and how these impacts vary across the different SCMs. First, a control simulation for each SCM was first performed to calibrate model output to match observations of the monthly average diurnal cycle of sensible heat flux and latent heat fluxes using the same model setup as described in Section 4.1. The model performance was evaluated using the following metrics: (1) bias, (2) root-mean-square-error (RMSE), (3) correlation coefficient (R^2), (4) the Kling-Gupta Efficiency (KGE) (Gupta et al., 2009), and (5) the Nash–Sutcliffe efficiency (NSE) (Nash & Sutcliffe, 1970). Next, additional simulations were performed for each SCM by changing the following three environmental variables: (1) air temperature, (2) atmospheric CO_2, and (3) soil moisture.

Each environmental variable is changed twice while keeping the other variables unchanged. The simulated changes in net CO_2 assimilation, leaf temperature, and leaf water potential were studied.

4.3 Experiment design for studying the computational performance on heterogeneous computing architectures

The computational performance of the multi-layer CAS sub-model of ELM-MLCMv1 on CPU and GPU is studied by performing simulations for a range of problem sizes, using two platforms at the Oak Ridge Leadership Computing Facility (OLCF): The Summit Supercomputer, the fifth fastest supercomputer in the world as of November 2022 (`https://www.top500.org/lists/top500/2022/11/`), and the Crusher testbed system, which contains identical hardware and similar software as the DOE's Frontier Exascale supercomputer, currently ranked as the fastest supercomputer in the world. These systems were chosen both because they represent the prior and current generation of leadership-class supercomputers, and because they rely on GPUs from different vendors and with different programming models. Each compute node of Summit has two 22-core IBM POWER9 CPUs and six NVIDIA Volta V100 GPUs, which uses NVIDIA's CUDA programming model, while each compute node of Crusher has one 64-core AMD EPYC CPU and four AMD MI250X GPUs, which uses AMD's ROCm programming model. We note that each AMD MI250X GPU consists of two Graphics Compute Dies (GCDs), each of which can be viewed by the programmer as a separate GPU; for the purposes of making consistent comparisons, we will use the term "GPU" in the rest of the paper to refer to one GPU on Summit or one GCD on Crusher.

We performed *static-scaling* experiments, in which simulations are run across a range of problem sizes, using a fixed level of parallelism (in our case, using a single CPU core or a single GPU). This is in contrast to the more usual *strong-scaling* (in which the degree of parallelism is varied for a fixed problem size, and the marginal efficiency of each additional parallel resource is measured) or *weak-scaling* (in which the local problem size is fixed and the number of parallel resources is varied) analyses. Strong-scaling or weak-scaling analyses do not make sense for ELM-MLCMv1, which, although intended to be used in parallel ELM simulations, is essentially a serial code: an independent instance of the ELM-MLCMv1 model will be run by each process in a parallel ELM simulation, so the parallelism (from the viewpoint of ELM-MLCMv1) is fixed, using only a single

CPU or GPU (though the full number of GPU cores will be used by the PETSc solvers via one of the PETSc GPU back-ends). The relevant question to be answered is how a given compute platform will handle the range of problem sizes that the ELM-MLCMv1 model will be run with. By running over a wide range of problem sizes, we can answer questions such as whether there is a minimum solution time, below which the cost of the computation is dominated by GPU latency, or whether there is a problem size where suboptimal memory access becomes a dominant factor.

In this work, the model performance on a single CPU core is compared against one GPU. The problem setup used in Section 4.1 is modified such each CAS is shared by 1000 trees. Static scaling experiments were performed in which the number of CAS is varied between 1 to 100. The numerical solution of the model uses PETSc's `KSP` iterative linear solver component and the computational performance of the model in this work examines the two major steps of the solution process: (1) The `PCSetUp` event, in which the linear preconditioner data structures are set up and computed (in this case, and incomplete LU factorization with zero fill-in is computed) (2) The `KSPSolve` event, which occurs after `PCSetUp` and involves solution of the linear system by a preconditioned Krylov subspace iteration method (GMRES with a restart size of 30 here); this phase includes the application of the preconditioner (sparse triangular solve) but not the setup. The sum of the time spent in `PCSetUp` and `KSPSolve` comprises the total time spent inside PETSc solvers in an ELM-MLCMv1 simulation. We consider the time spent in PETSc solvers only for two reasons: First, we wish to evaluate how the model can be sped up by making use of GPU accelerators, and the PETSc solve is the most suitable part of the computation for GPU acceleration. Second, the time in the solvers represents fifty percent or more of the total time spent in the benchmark (when GPU acceleration is not used), and would represent an even larger fraction when ELM-MLCMv1 is deployed in an ELM simulation. In the benchmark code, the linear system is assembled and solved only once, but as part of an ELM simulation, much of the costly initial setup will be amortized across multiple solves; though the entries of the linear system matrix will be recomputed, the nonzero structure of the matrix will stay the same and the allocated data structures can be re-used.

The computational performance of the model was quantitatively evaluated using (1) speedup, and (2) work-time spectrum plots (Chang, Nakshatrala, et al., 2018; Chang, Fabien, et al., 2018). Speedup, a commonly employed metric in parallel computing, is

computed here as the ratio of the execution time (as measured for a given PETSc event) for the CPU-only case to that of the GPU-enabled run. Speedup up gives a simple measure of how much faster the computation is when the GPU is employed. The work-time spectrum plot presents a less common but very informative view of model's computational performance by comparing the computation throughput (total degrees of freedom in the problem divided by the execution time) against the execution time. (Note that the quantity being controlled–the problem size—does not appear on any axis, but it increases from left to right for each curve.) The advantage of this view is that the latency and asymptotic throughput can be read directly from the plot: the former is represented simply by the smallest time that a given curve passes through, and the latter by the maximum throughput at the top of the curve. When a static scaling experiment runs over a wide enough range of problem sizes, performance curves on a work-time spectrum plot generally follow three regimes that proceed from left to right as: (1) regime with a positive slope where execution is close to the strong-scaling limit because the problem size is too small to allow full use of all of the parallel resources, (2) regime where slope levels off to a horizontal line that indicates a region of optimal scaling in which there is sufficient work for full utilization of all resources, and (3) regime of a negative slope where parallel efficiency drops due to memory subsystem effects or other causes.

5 Results

5.1 Model verification

The ELM-MLCMv1 model is able to accurately match the spatio-temporal results from the CLM-ml_v1 simulation. The two models show excellent agreement for vertically-averaged time series of air temperature, water vapor, and net radiation for shaded and sunlit leaves for Medlyn SCM (Figure 4). The relative error in the time series of vertical profile for the aforementioned variables for Medlyn SCM is $\mathcal{O}(10^{-9})$. The relative error between the two models for SCMs other than Bonan14 is similarly very small (Table 3). While the ELM-MLCMv1 is not validated against observations in this study, these results provide a high degree of confidence in the implementation of the ML model as it can accurately reproduce results of CLM-ml_v1 that has been validated against observational datasets (G. B. Bonan et al., 2018).

5.2 Model validation

The calibrated ELM-MLCMv1 for each SCM is able to accurately simulate the monthly-averaged diurnal cycle of sensible and latent heat flux (Figure 5) with low bias and RMSE, and high R^2, KGE, and NSE (Table 4). The bias of the three SCMs that include PHD (i.e., WUE, Bonan14, and Manzoni11) is smaller than the bias of the two SCMs that do not include PHD (i.e., Medlyn and BB). However, all other evaluation metrics are comparable across the five SCMs. The model performance of all SCMs is lower at the hourly scale than compared to the monthly-averaged diurnal cycle (see metrics in parentheses in Table 4), though the KGE and NSE still remain larger than 0.69 and 0.77, respectively.

The simulated monthly-averaged vertical profiles of leaf water potential, ψ_{leaf}, air temperature, T_{air}, and air vapor pressure, q_{air}, show vertical variation within the canopy (Figure 6). The differences in the simulated variables across SCMs is small (as indicated by the cyan shading) and only limited to below the canopy height (as indicated by the dashed line). Given the strong agreement among SCMs, we only present temporally-varying, vertically-explicit profiles of variables simulated by ELM-MLCMv1 for the Medlyn SCM (Figure 7). The top of the canopy has largest net assimilation, A_{net} (Figure 7a) with corresponding lower intercellular CO_2 (Figure 7b) and high stomatal conductance (Figure 7c). The canopy buffers the variation of temperature by lowering the maximum leaf temperature of the lower canopy as compared to canopy top during mid-day and vice-versa during the night (Figure 7d). The higher A_{net} at the top of the canopy results in higher water loss through transpiration, which results in lowers ψ_{leaf} (Figure 7e). The vertical profile of T_{air} shows similar spatio-temporal pattern as that of T_{leaf} with the upper canopy buffering temperature extrema in the lower canopy (Figure 7f). The vertical gradient in q_{air} is only observed during the day with higher values at the top of canopy (Figure 7g).

5.3 Impact of climatic factors on canopy processes

All five SCMs agree on the direction of the simulated changes (i.e., an increase or a decrease) in the three vertically-averaged canopy processes (i.e., A_{net}, T_{leaf}, and ψ_{leaf}) due to the changes in the three climatic factors (i.e. ΔT_{air}, ΔCO_2, and ΔSM) (Figure 8). The changes in a climate factors impacted canopy processes differently. For example, an increase in CO_2 resulted in higher A_{net}, but T_{leaf} and ψ_{leaf} remained unchanged (Fig-

ure 8b,e,h). Only the changes in T_{air} impacts two canopy processes (i.e., A_{net} and T_{leaf}), and only A_{net} is impacted by two climatic factors (i.e., T_{air} and ΔCO_2). The spatio-temporal differences of the simulated changes in the canopy processes are studied next for the four cases (Figure 8a,b,d,i) that showed a noticeable change due to the changes in climatic factors (Figure 9). An increase in T_{air} (or a decrease in CO_2) leads to a decrease (or an increase) in A_{net} for all canopy layers (Figure 9a,b). The change in A_{net} is largest at the top of the canopy and has a well-defined diurnal pattern with higher values during midday. The changes in T_{leaf} due to ΔT_{air} vary spatially and temporally (Figure 9c). During nighttime, mid-canopy has lower ΔT_{leaf} compared to the top and the bottom of the canopy, and vice-versa during daytime. The ψ_{leaf} decreased due to ΔSM throughout the canopy with the largest changes at the top of the canopy and during late afternoon (Figure 9d).

5.4 Computational performance

The speedup plots for both the Summit and Crusher machines follow the expected trends, with speedup generally increasing as the problem size grows to be large enough to allow full utilization of the large number of GPU compute elements present (Figure 10). The speedup is most pronounced for the `PCSetUp` event, which involves a large amount of work that can saturate the GPU, whereas the `KSPSolve` event—which involves comparatively much less work and cannot keep the GPU as busy—displays much more modest speedup. On Summit, the speedup increases consistently with problem size, whereas on Crusher the speedup levels off somewhat before the maximum problem size is reached. At larger problem sizes, the speedups on Summit are consistently higher than those on Crusher, but the execution on Crusher is actually always faster (Figure 11)as the speedups in each plot are measured relative to the performance of that machine's CPU.

The work-time spectrum plots for the CPU-only runs are close to horizontal (dashed lines in Figure 11), indicating that, for this range of problem sizes, the CPU resources are saturated and no memory effects come into play. For the GPU-enabled runs (solid lines in Figure 11), we see that the `PCSetUp` phase reaches the optimal scaling regime on Crusher, and is perhaps approaching it on Summit. The `KSPSolve` phase contains less work and is expected to require a larger problem size to reach the optimal scaling regime. We observe that on Crusher `KSPSolve` appears to be approaching the optimal scaling phase at the largest problem sizes, whereas on Summit the `KSPSolve` seems to still be near the strong-scaling limit. On both machines, the GPU performance beats the CPU

performance in all cases except for the KSPSolve with the smallest problem size (a single CAS); for this case, the amount of work done in the KSPSolve is too small to mask high kernel launch latencies for the GPUs, but because this does not hold for the `PCSetUp`, the total time spent in the PETSc solvers is still lower for the GPU. As expected, because the CPUs and GPUs in Summit are several years older than in Crusher, the Crusher GPU and GPU performance consistently beats that of its counterparts on Summit.

6 Caveats and Future work

This study is a first step to including the representation of vertical canopy structure within ELM and we now briefly discuss possible future developments in the new model required for implementation in the ELM *big-leaf* scheme. First, the prescribed time-series of soil moisture, surface temperature, and surface albedo as a boundary condition should instead use ELM's prognostic soil hydrology and soil temperature models. Second, the evaporation of canopy intercepted water needs to be accounted for in the energy balance within the CAS. Third, the radiation sub-model could be extended to simulate solar-induced induced chlorophyll fluorescence, which has emerged as a proxy for terrestrial photosynthesis and can be observed from satellites (Li et al., 2022). Fourth, a more realistic vertical structure of sunlit and shaded leaves should be included in which leaves at each canopy level are not directly connected to the soil via individual xylems but are connected to a single xylem (Mirfenderesgi et al., 2016). Fifth, multi-year and multi-site simulations need to be performed to evaluate the performance of the model against observations. Sixth, a computational performance analysis of the model should be performed for a global ELM simulation. These additional model developments will significantly improve ELM's capability to resolve micro-climate due to vertical canopy structure and enable support for upcoming heterogeneous high-performance architectures.

7 Summary

In this study, we developed a standalone ELM-MLCMv1 that resolves the micro-climate created by the vertical structure of the vegetation canopies. The new model includes sub-models for shortwave and longwave radiation, stomatal conductance, turbulence scheme for modeling flow within the vegetation canopy, and transport of heat and water vapor within CAS. ELM-MLCMv1 uses PETSc to provide a numerically robust solution of discretized equations and includes support for heterogeneous comput-

ing architectures. The numerical implementation of ELM-MLCMv1 was verified by comparing results against CLM-ml_v1, an existing MLC model, for a month-long simulation using data from the Ameriflux US-UMB site.

We found that simulations using SCMs of various degrees of complexity are each able to accurately simulate observed sensible and latent heat fluxes upon model calibration. A set of idealized climate change simulations showed that all SCMs predict a consistent change in the canopy processes due to changes in environmental factors (i.e., air temperature, atmospheric CO_2, and soil moisture). The differences in the canopy processes including net assimilation, leaf temperature, and leaf water potential vary vertically within the canopy in response to changes in environmental factors. While net assimilation was impacted by the changes in both air temperature and atmospheric CO_2; however, leaf temperature and leaf water potential were only impacted by changes in air temperature and soil moisture, respectively. The use of PETSc enabled the same ELM-MLCMv1 source code to run on CPUs and GPUs without requiring any code modifications, as well as includes support of GPUs from multiple vendors (i.e., NVIDIA and AMD). Static scaling experiments for an idealized problem showed ELM-MLCMv1 achieved a speedup of 25-50 times on a GPU relative to a CPU. This work provides the first necessary model development to include the representation of vertical canopies in ELM and additionally includes the support for heterogeneous computing architectures.

8 Availability Statement

The ELM-MLCMv1 is publicly available at `https://zenodo.org/record/7809207` and scripts to perform simulations presented in this study are available at `https://github.com/MPP-LSM/mlcm-simulation`.

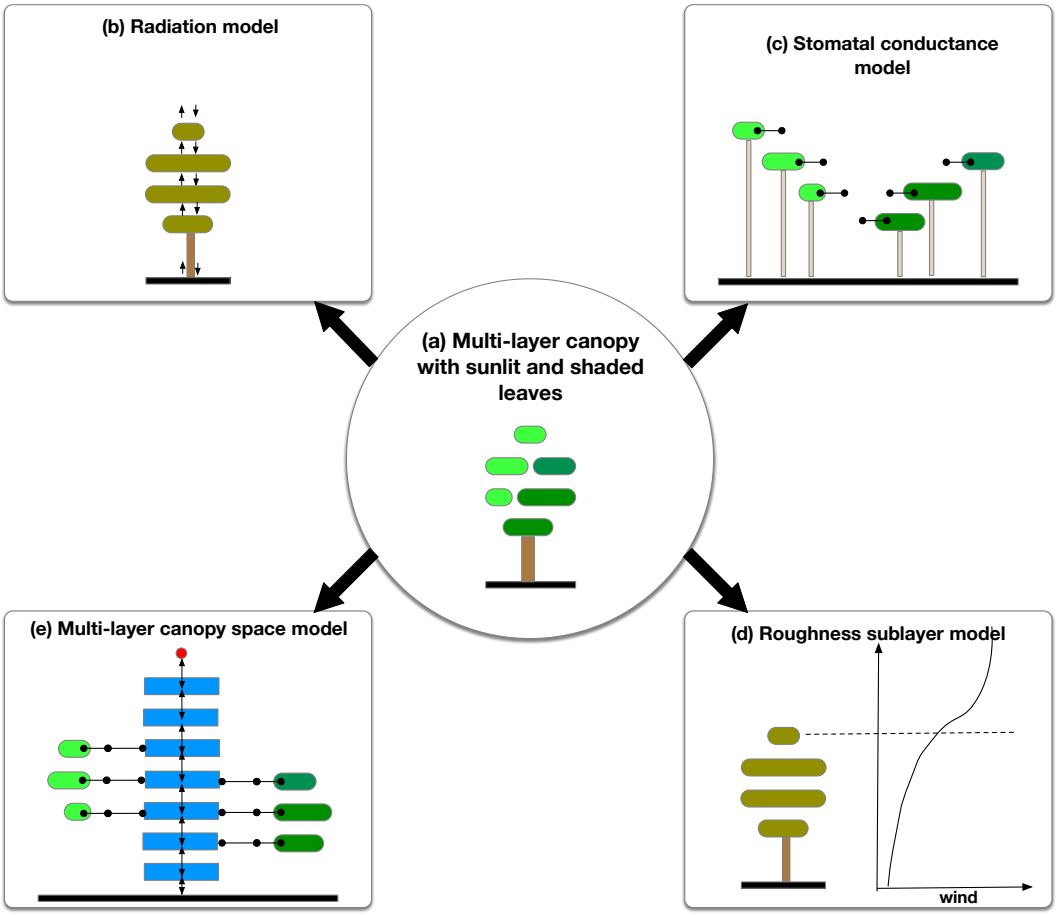


Figure 1. Schematic representation of ELM-MLCMv1 that includes (a) leaves at multiple layers that are sunlit (light green) and shaded, (b) vertically-resolved models for shortwave and longwave radiation that combine sunlit and shaded leaves into a single leaf (brown) at each canopy level, (c) model for stomatal conductance of sunlit and shaded leaves, (d) roughness sublayer parameterization of turbulent flow within and above vegetation canopy, and (e) model for the transport of heat and water vapor through the canopy air space.

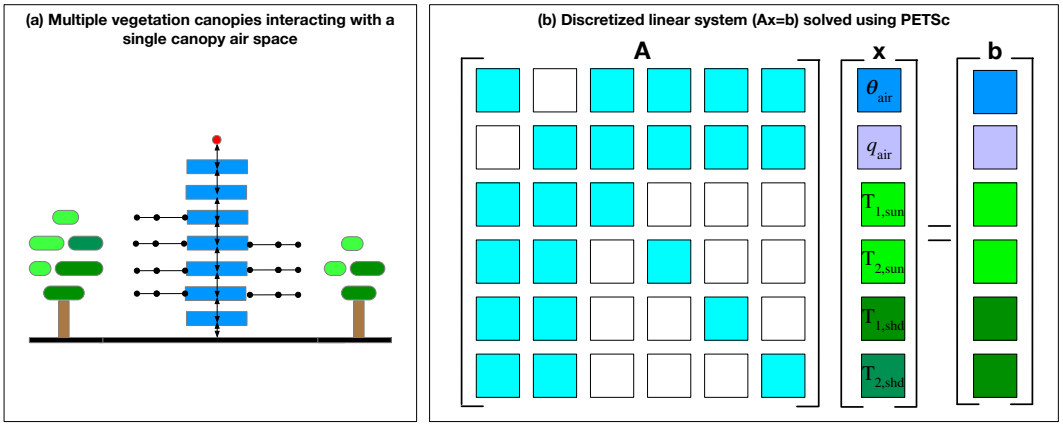


Figure 2. (a) Schematic representation of two vegetation canopies that are interacting with a single canopy air space (CAS). The model configuration comprises of four governing equations and comprises of six unknown variables that include CAS temperature (θ_{air}), CAS water vapor (q_{air}), sunlit leaf temperatures ($T_{1,sun}$ and $T_{2,sun}$ corresponding to two trees), and shaded leaf temperatures ($T_{1,shd}$ and $T_{2,shd}$ corresponding to two trees). (b) The discretized linear system is flexibly assembled using PETSc's `DMComposite` and solved using PETSc's `KSP` solver.

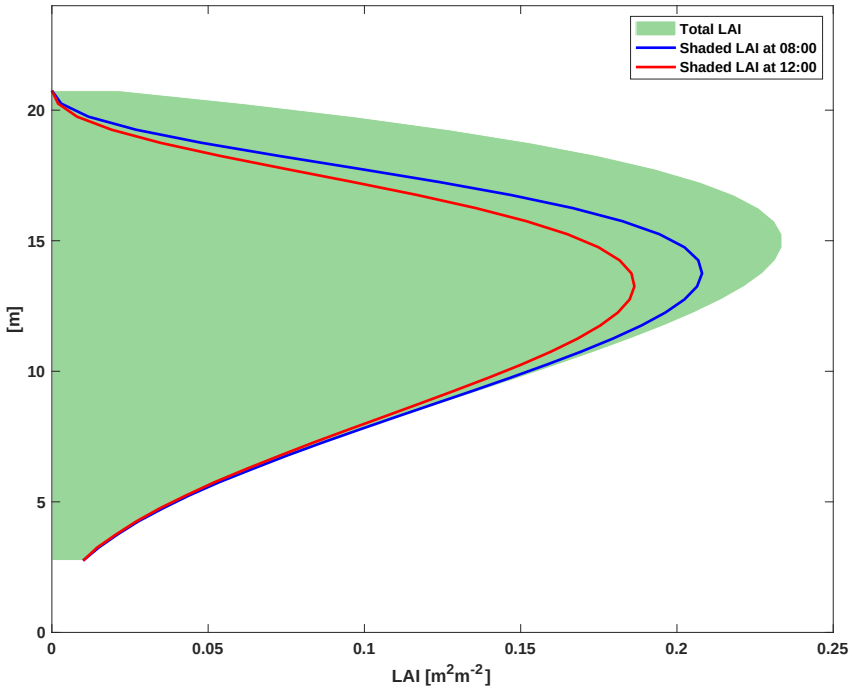


Figure 3. The vertical profile of LAI (shaded region in green) along with LAI profiles of shaded leaves at 8:00 am (in blue) and 12:00 pm (in red).

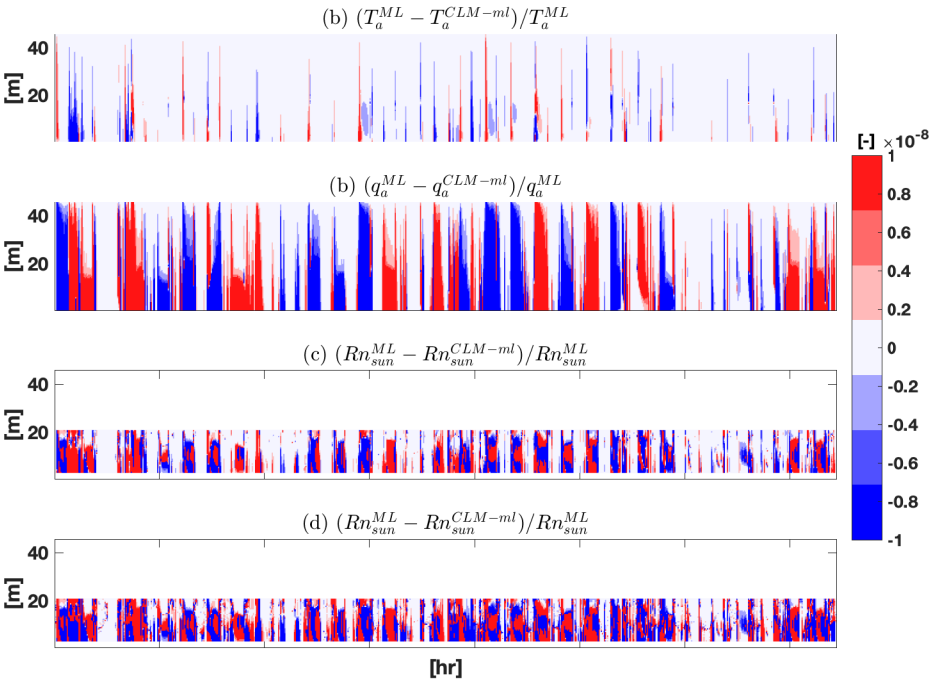


Figure 4. Time series of the vertical profile of relative error for Medlyn SCM with respect to CLM-ml v1 for (a) air temperature, (b) water vapor, (c) net radiation for sunlit leaf, and (d) net radiation for shaded leaf. simulated net radiation, R_{net}, The results are for the Bonan14 SCM.

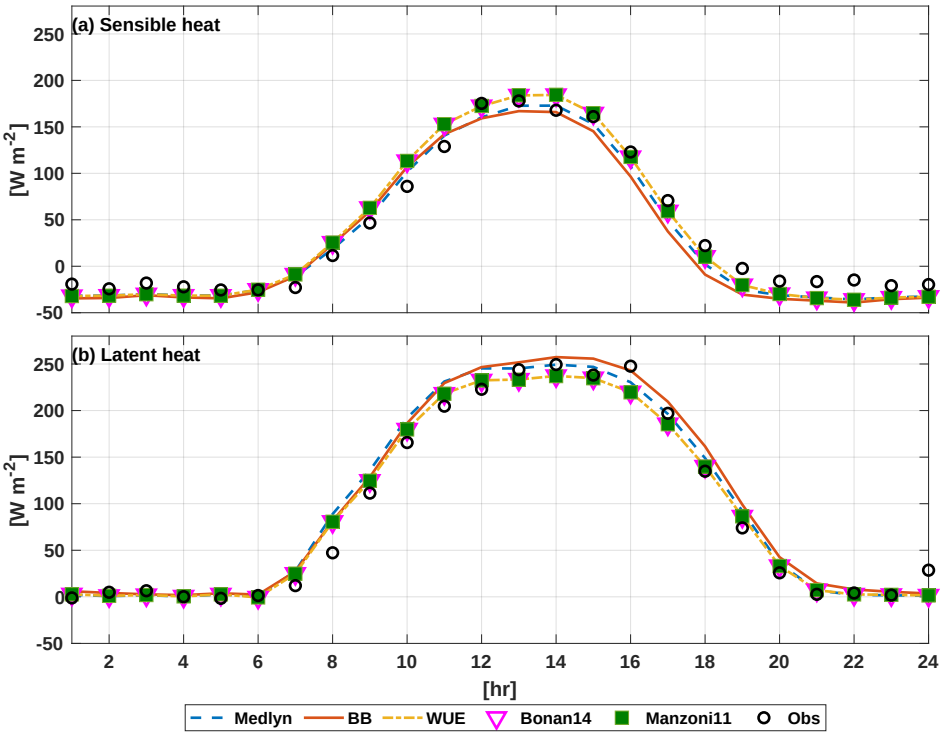


Figure 5. Simulated monthly average diurnal cycle of (a) sensible heat flux, H, (b) latent flux, LH, and (c) air temperature, T_{air}, for five SCMs. The observations are shown in circle symbols.

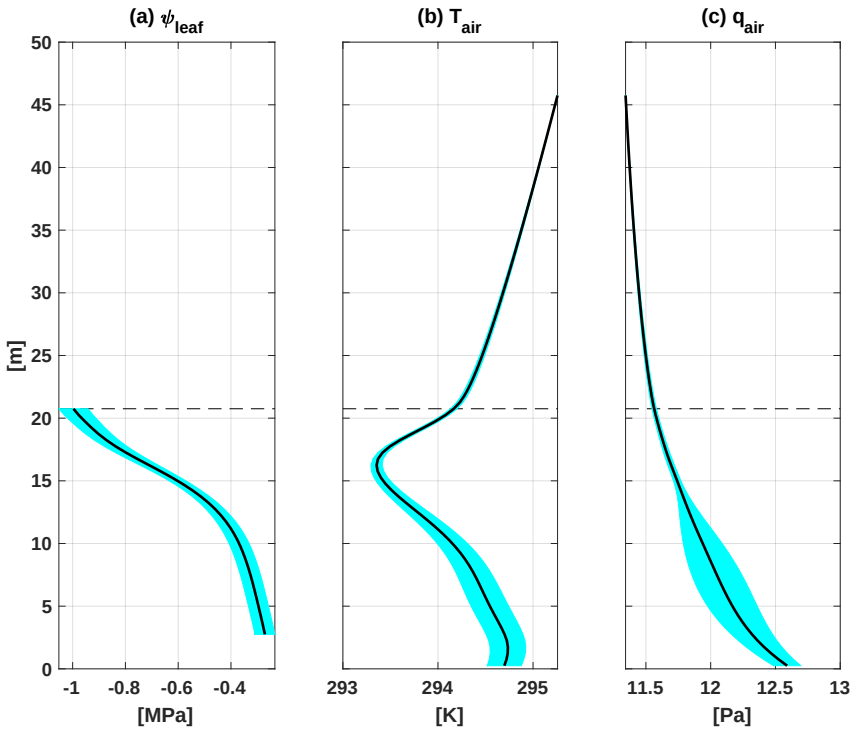


Figure 6. Simulated monthly average vertical profile of (a) leaf water potential, ψ_{leaf}, (b) air temperature, T_{air}, and (c) air vapor pressure, q_{air}. The black line indicates the average of five SCMs, while the shading denotes the standard deviation among five SCMs.

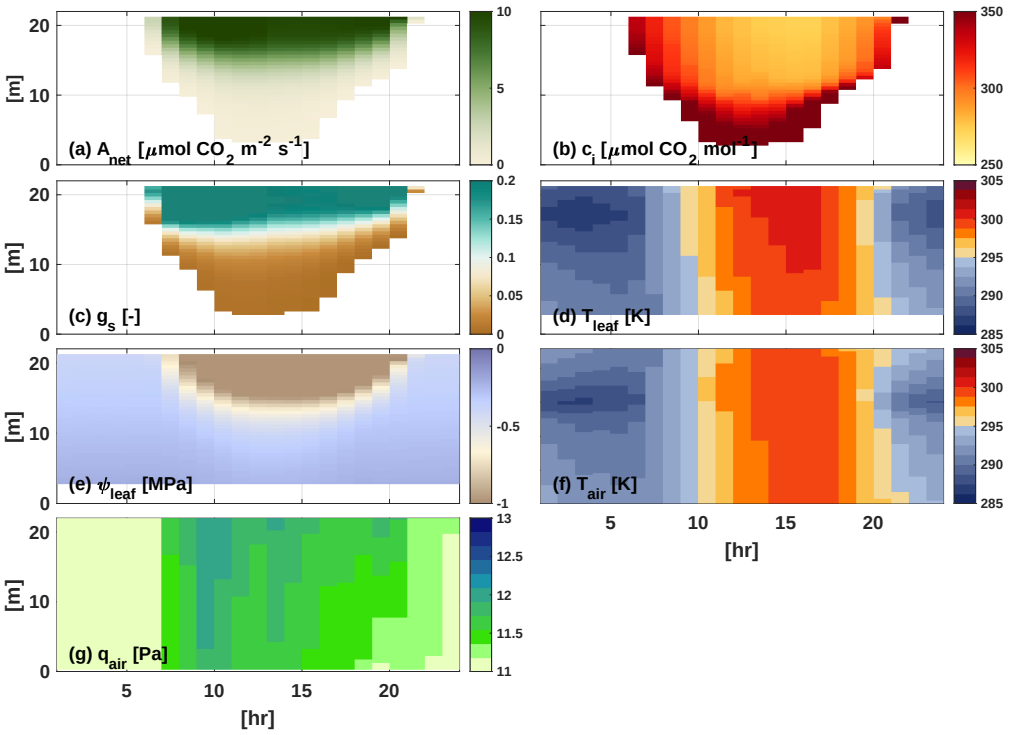


Figure 7. Simulated vertical profile of average diurnal cycle of (a) net assimilation, A_{net}, (b) intercellular CO_2, c_i, (c) stomatal conductance, g_s, (d) leaf temperature, T_{leaf}, (e) leaf water potential, ψ_{leaf}, (f) air temperature, T_{air}, and (g) water vapor, q_{air}, for the Medlyn SCM.

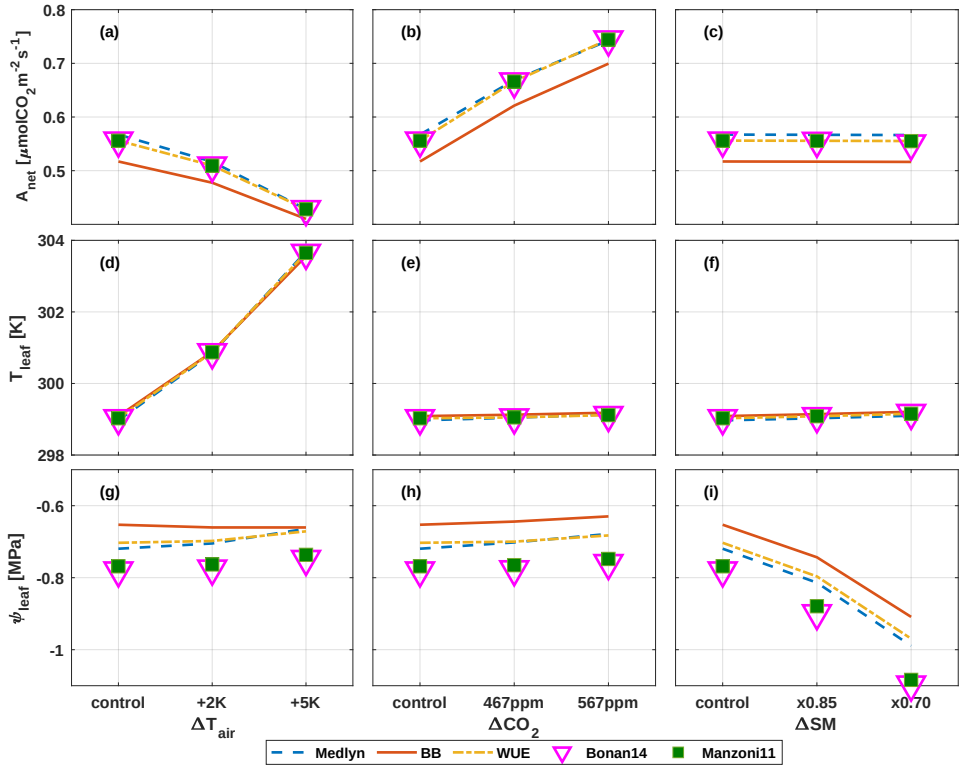


Figure 8. Monthly-average net assimilation, A_{net}, due to changes in three environmental variables: (a) increase in temperature, ΔT_{air}, (b) increase in CO_2, ΔCO_2 , and (c) decrease in soil moisture, ΔSM. Impact of the changes in the same three environmental variables on (d-f) monthly average leaf temperature, T_{leaf}, and (g-j) monthly average leaf water potential ψ_{leaf}. The results are presented for five different SCMs.

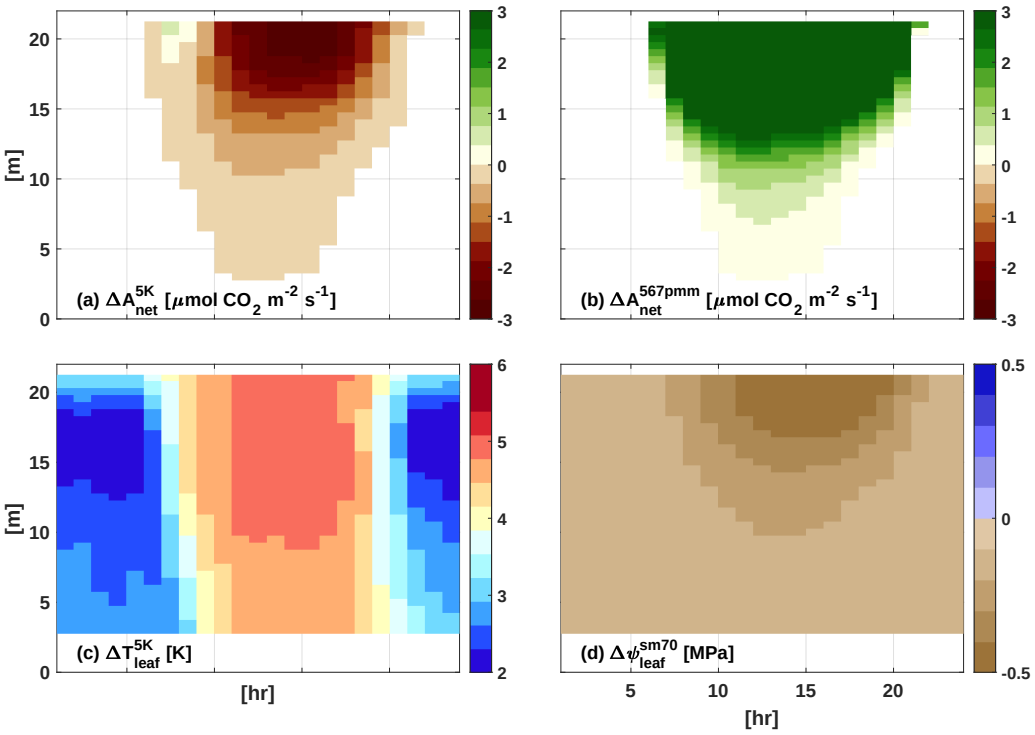


Figure 9. (a) Vertical profile of average diurnal cycle of change in net CO_2 assimilation, ΔA_{net}. The anomalies in A_{net} with respect to the control case for the case with (b) 5K increase in air temperature, (c) higher atmospheric CO_2 at 567 [ppm], and (d) lower soil moisture at 0.75× the control value. for the control case using the Medlyn SCM.

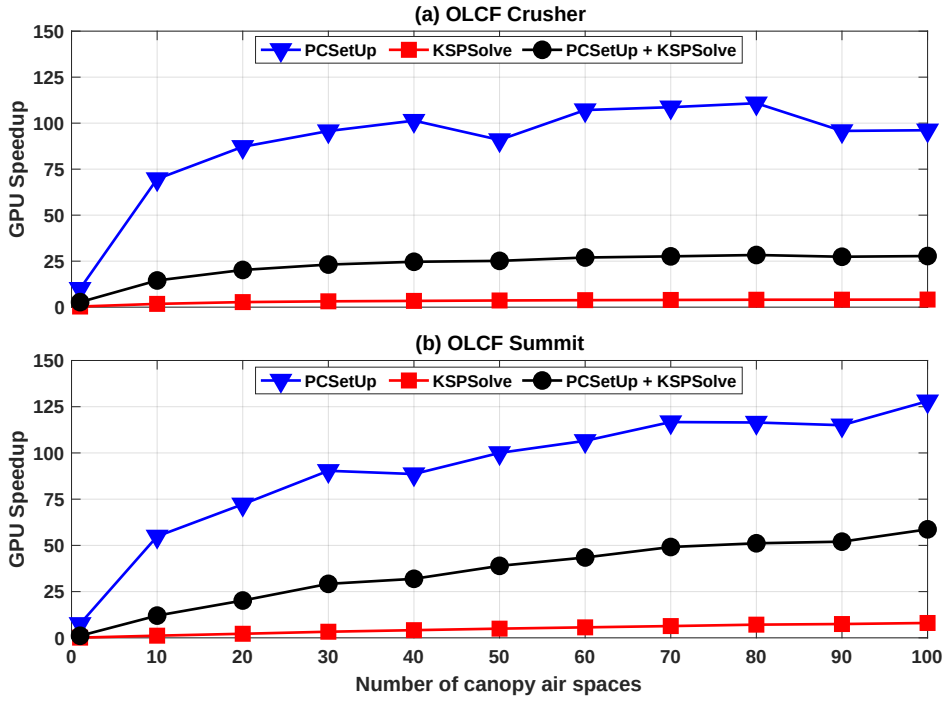


Figure 10. Speedup of the PETSc solvers as a function of problem size (determined by number of CAS) when running on one GPU vs. one CPU on a node of the OLCF Summit (left) and Crusher (right machines), broken down into time spent in `PCSetUp`, `KSPSolve`, and the combination thereof.

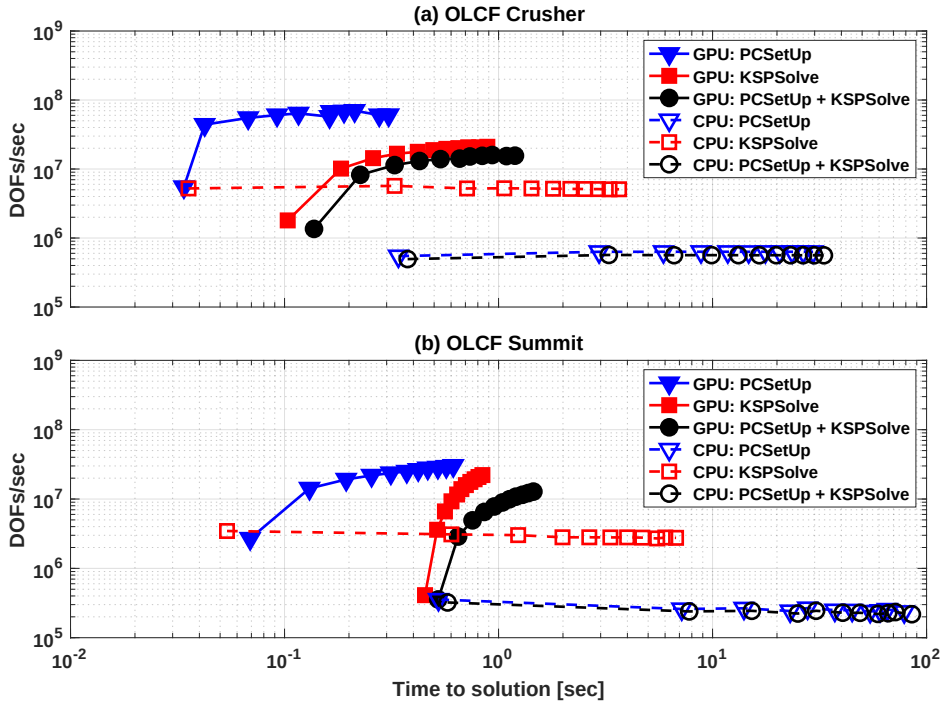


Figure 11. A work-time spectrum view of the performance of PETSc solvers for our computational performance benchmark runs on the OLCF Summit and Crusher computers. The plot depicts the performance measurements for all of the different runs, though the quantity being controlled (the problem size, i.e., the number of CAS) is not plotted on any axis. This view has the advantage of allowing both latency and asymptotic throughput to be read directly from the plot.

Table 1. Description of SCMs supported in ELM-MLCMv1.

Name	Description	Reference
BB	Semi-empirical approach without accounting for PHD	Ball et al. (1987)
Medlyn	Semi-empirical approach without accounting for PHD	Medlyn et al. (2011)
WUE	Optimization theory without accounting for PHD	Buckley et al. (2017)
Bonan14	Co-optimization of two constraints including a PHD constraint	G. Bonan et al. (2014)
Manzoni11	WUE-based approach with down regulation based on PHD	Manzoni et al. (2011)

Table 2. Simulation configurations for studying the impact on canopy processes from three environmental variables that include air temperature, atmospheric CO_2, and soil moisture.

	Name	**Air temperature**	**Atmospheric CO_2**	**Soil moisture**
1	Control	Default	367ppm	Default
2	2K	Default + 2K	367ppm	Default
3	5K	Default + 5K	367ppm	Default
4	467ppm	Default	467ppm	Default
5	567ppm	Default	567ppm	Default
6	sm85	Default	367ppm	0.85 × Default
7	sm70	Default	367ppm	0.70 × Default

Table 3. Mean relative error between CLM-ml_v1 and ML for net radiation of sunlit ($Rn_{\ell sun}$) and shaded ($Rn_{\ell sha}$) leaves, stomatal conductance of sunlit ($gs_{\ell sun}$) and shaded ($gs_{\ell sha}$) leaves, air temperature (T_a), and water vapor (q_a).

SCM	$\Delta Rn_{\ell sun}/Rn^{ML}_{\ell sun}$	$\Delta Rn_{\ell sha}/Rn^{ML}_{\ell sha}$	$\Delta gs_{\ell sun}/gs^{ML}_{\ell sun}$	$\Delta gs_{\ell sha}/gs^{ML}_{\ell sha}$	$\Delta T_a/T^{ML}_a$	$\Delta q_a/q^{ML}_a$
BB	+3.06e-09	+6.96e-09	+7.69e-06	-4.12e-09	-1.17e-11	+3.47e-10
Medlyn	+1.25e-07	-1.69e-05	+1.35e-07	+2.85e-07	-3.30e-10	+2.83e-09
Bonan14	+1.73e-06	+1.65e-06	+1.38e-07	+7.90e-08	-9.31e-11	-7.78e-09

Table 4. Evaluation metrics for the control simulation for five SCMs. The metrics in the top row are for simulated monthly-averaged diurnal cycle of sensible and latent heat flux, while the metrics in the bottom row within parentheses are for simulated hourly heat fluxes.

SCM	Sensible heat flux					Latent heat flux				
	Bias	**RMSE**	**R^2**	**KGE**	**NSE**	**Bias**	**RMSE**	**R^2**	**KGE**	**NSE**
Medlyn	-7.3	13.5	0.99	0.80	0.97	6.8	16.2	0.99	0.92	0.97
	(-9.1)	(49.3)	(0.87)	(0.73)	(0.75)	(9.2)	(53.3)	(0.89)	(0.85)	(0.78)
BB	-10.3	18.0	0.98	0.73	0.94	10.4	16.4	0.99	0.88	0.97
	(-11.5)	(47.9)	(0.89)	(0.69)	(0.77)	(12.1)	(51.3)	(0.90)	(0.83)	(0.80)
WUE	-2.3	14.0	0.99	0.88	0.96	1.3	13.2	0.99	0.97	0.98
	(-3.8)	(47.3)	(0.88)	(0.85)	(0.77)	(3.3)	(50.6)	(0.90)	(0.86)	(0.80)
Bonan14	-2.3	14.0	0.99	0.88	0.96	1.3	13.2	0.99	0.97	0.98
	(-3.8)	(47.3)	(0.88)	(0.85)	(0.77)	(3.3)	(50.6)	(0.90)	(0.86)	(0.80)

Acknowledgments

This research was supported as part of the Energy Exascale Earth System Model (E3SM) project, funded by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research. The Pacific Northwest National Laboratory is operated by Battelle for the US Department of Energy under Contract DE-AC05-76RLO1830.

References

Agee, E., He, L., Bisht, G., Couvreur, V., Shahbaz, P., Meunier, F., ... Ivanov, V. (2021). Root lateral interactions drive water uptake patterns under water limitation. *Advances in Water Resources*, *151*, 103896.

Allen, C. D., Macalady, A. K., Chenchouni, H., Bachelet, D., McDowell, N., Vennetier, M., ... others (2010). A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. *Forest ecology and management*, *259*(4), 660–684.

Balay, S., Abhyankar, S., Adams, M. F., Benson, S., Brown, J., Brune, P., ... Zhang, J. (2021). *PETSc/TAO users manual* (Tech. Rep. No. ANL-21/39 - Revision 3.16). Argonne National Laboratory.

Ball, J. T., Woodrow, I. E., & Berry, J. A. (1987). A model predicting stomatal conductance and its contribution to the control of photosynthesis under different environmental conditions. In *Progress in photosynthesis research* (pp. 221–224). Springer.

Beckingsale, D. A., Burmark, J., Hornung, R., Jones, H., Killian, W., Kunen, A. J., ... Scogland, T. R. (2019). Raja: Portable performance for large-scale scientific applications. In *2019 ieee/acm international workshop on performance, portability and productivity in hpc (p3hpc)* (pp. 71–81).

Bisht, G., & Riley, W. J. (2019). Development and verification of a numerical library for solving global terrestrial multiphysics problems. *Journal of Advances in Modeling Earth Systems*, *11*(6), 1516–1542.

Bohrer, G., Mourad, H., Laursen, T. A., Drewry, D., Avissar, R., Poggi, D., ... Katul, G. G. (2005). Finite element tree crown hydrodynamics model (fetch) using porous media flow within branching elements: A new representation of tree hydrodynamics. *Water Resources Research*, *41*(11).

Bonan, G., Williams, M., Fisher, R., & Oleson, K. (2014). Modeling stomatal conductance in the earth system: linking leaf water-use efficiency and water transport along the soil–plant–atmosphere continuum. *Geoscientific Model Development*, *7*(5), 2193–2222.

Bonan, G. B., Patton, E. G., Finnigan, J. J., Baldocchi, D. D., & Harman, I. N. (2021). Moving beyond the incorrect but useful paradigm: reevaluating big-leaf and multilayer plant canopies to model biosphere-atmosphere fluxes–a review.

Agricultural and Forest Meteorology, *306*, 108435.

Bonan, G. B., Patton, E. G., Harman, I. N., Oleson, K. W., Finnigan, J. J., Lu, Y., & Burakowski, E. A. (2018). Modeling canopy-induced turbulence in the earth system: a unified parameterization of turbulent exchange within plant canopies and the roughness sublayer (clm-ml v0). *Geoscientific Model Development*, *11*(4), 1467–1496.

Bond, B. J., Farnsworth, B. T., Coulombe, R. A., & Winner, W. E. (1999). Foliage physiology and biochemistry in response to light gradients in conifers with varying shade tolerance. *Oecologia*, *120*(2), 183–192.

Brown, J., Knepley, M. G., May, D. A., McInnes, L. C., & Smith, B. F. (2012, 06/2012). Composable linear solvers for multiphysics. In *11th international symposium on parallel and distributed computing*. Munich, Germany.

Buckley, T. N., Sack, L., & Farquhar, G. D. (2017). Optimal plant water economy. *Plant, cell & environment*, *40*(6), 881–896.

Chang, J., Fabien, M. S., Knepley, M. G., & Mills, R. T. (2018). Comparative study of finite element methods using the time-accuracy-size (TAS) spectrum analysis. *SIAM Journal on Scientific Computing*, *40*(6), C779–C802.

Chang, J., Nakshatrala, K., Knepley, M. G., & Johnsson, L. (2018). A performance spectrum for parallel computational frameworks that solve PDEs. *Concurrency and Computation: Practice and Experience*, *30*(11), e4401.

Chen, Y., Ryder, J., Bastrikov, V., McGrath, M. J., Naudts, K., Otto, J., ... others (2016). Evaluating the performance of land surface model orchidee-can v1. 0 on water and energy flux estimation with a single-and multi-layer energy budget scheme. *Geoscientific Model Development*, *9*(9), 2951–2972.

Christoffersen, B. O., Gloor, M., Fauset, S., Fyllas, N. M., Galbraith, D. R., Baker, T. R., ... others (2016). Linking hydraulic traits to tropical forest function in a size-structured and trait-driven model (tfs v. 1-hydro). *Geoscientific Model Development*, *9*(11), 4227–4255.

Cowan, I., & GD, F. (1977). Stomatal function in relation to leaf metabolism and environment.

De Frenne, P., Zellweger, F., Rodriguez-Sanchez, F., Scheffers, B. R., Hylander, K., Luoto, M., ... Lenoir, J. (2019). Global buffering of temperatures under forest canopies. *Nature Ecology & Evolution*, *3*(5), 744–749.

Dewar, R., Mauranen, A., Mäkelä, A., Hölttä, T., Medlyn, B., & Vesala, T. (2018). New insights into the covariation of stomatal, mesophyll and hydraulic conductances from optimization models incorporating nonstomatal limitations to photosynthesis. *New Phytologist*, *217*(2), 571–585.

Edwards, H. C., Trott, C. R., & Sunderland, D. (2014). Kokkos: Enabling manycore performance portability through polymorphic memory access patterns. *Journal of parallel and distributed computing*, *74*(12), 3202–3216.

Eller, C. B., Rowland, L., Mencuccini, M., Rosas, T., Williams, K., Harper, A., ... others (2020). Stomatal optimization based on xylem hydraulics (sox) improves land surface model simulation of vegetation responses to climate. *New Phytologist*, *226*(6), 1622–1637.

Fang, Y., Leung, L. R., Wolfe, B. T., Detto, M., Knox, R. G., McDowell, N. G., ... others (2021). Disentangling the effects of vapor pressure deficit and soil water availability on canopy conductance in a seasonal tropical forest during the 2015 el niño drought. *Journal of Geophysical Research: Atmospheres*, *126*(10), e2021JD035004.

Fatichi, S., Pappas, C., & Ivanov, V. Y. (2016). Modeling plant–water interactions: an ecohydrological overview from the cell to the global scale. *Wiley Interdisciplinary Reviews: Water*, *3*(3), 327–368.

Fisher, R. A., Koven, C. D., Anderegg, W. R., Christoffersen, B. O., Dietze, M. C., Farrior, C. E., ... others (2018). Vegetation demographics in earth system models: A review of progress and priorities. *Global change biology*, *24*(1), 35–54.

Grant, R., Black, T., Gaumont-Guay, D., Klujn, N., Barr, A., Morgenstern, K., & Nesic, Z. (2006). Net ecosystem productivity of boreal aspen forests under drought and climate change: Mathematical modelling with ecosys. *Agricultural and Forest meteorology*, *140*(1-4), 152–170.

Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and nse performance criteria: Implications for improving hydrological modelling. *Journal of hydrology*, *377*(1-2), 80–91.

Harman, I. N., & Finnigan, J. J. (2007). A simple unified theory for flow in the canopy and roughness sublayer. *Boundary-layer meteorology*, *123*(2), 339–363.

Harman, I. N., & Finnigan, J. J. (2008). Scalar concentration profiles in the canopy and roughness sublayer. *Boundary-layer meteorology*, *129*(3), 323–351.

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.

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.

Kennedy, D., Swenson, S., Oleson, K. W., Lawrence, D. M., Fisher, R., Lola da Costa, A. C., & Gentine, P. (2019). Implementing plant hydraulics in the community land model, version 5. *Journal of Advances in Modeling Earth Systems*, *11*(2), 485–513.

Lenoir, J., Graae, B. J., Aarrestad, P. A., Alsos, I. G., Armbruster, W. S., Austrheim, G., ... Svenning, J.-C. (2013). Local temperatures inferred from plant communities suggest strong spatial buffering of climate warming across northern europe. *Global Change Biology*, *19*(5), 1470-1481. Retrieved from `https://onlinelibrary.wiley.com/doi/abs/10.1111/gcb.12129` doi: https://doi.org/10.1111/gcb.12129

Leuning, R. (1995). A critical appraisal of a combined stomatal-photosynthesis model for c3 plants. *Plant, Cell & Environment*, *18*(4), 339–355.

Li, R., Lombardozzi, D., Shi, M., Frankenberg, C., Parazoo, N. C., Köhler, P., ... Yang, X. (2022). Representation of leaf-to-canopy radiative transfer processes improves simulation of far-red solar-induced chlorophyll fluorescence in the community land model version 5. *Journal of Advances in Modeling Earth Systems*, *14*(3), e2021MS002747.

Longo, M., Knox, R. G., Medvigy, D. M., Levine, N. M., Dietze, M. C., Kim, Y., ... Moorcroft, P. R. (2019). The biophysics, ecology, and biogeochemistry of functionally diverse, vertically and horizontally heterogeneous ecosystems: the ecosystem demography model, version 2.2 – part 1: Model description. *Geoscientific Model Development*, *12*(10), 4309–4346. Retrieved from `https://gmd.copernicus.org/articles/12/4309/2019/` doi: 10.5194/gmd-12-4309-2019

Manzoni, S., Vico, G., Katul, G., Fay, P. A., Polley, W., Palmroth, S., & Porporato, A. (2011). Optimizing stomatal conductance for maximum carbon gain under water stress: a meta-analysis across plant functional types and climates.

Functional Ecology, *25*(3), 456–467.

Massman, W., & Weil, J. (1999). An analytical one-dimensional second-order closure model of turbulence statistics and the lagrangian time scale within and above plant canopies of arbitrary structure. *Boundary-Layer Meteorology*, *91*(1), 81–107.

McDowell, N. G., & Allen, C. D. (2015). Darcy's law predicts widespread forest mortality under climate warming. *Nature Climate Change*, *5*(7), 669–672.

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

Mills, R. T., Adams, M. F., Balay, S., Brown, J., Dener, A., Knepley, M., ... others (2021). Toward performance-portable petsc for gpu-based exascale systems. *Parallel Computing*, 102831.

Mirfenderesgi, G., Bohrer, G., Matheny, A. M., Fatichi, S., de Moraes Frasson, R. P., & Schäfer, K. V. (2016). Tree level hydrodynamic approach for resolving aboveground water storage and stomatal conductance and modeling the effects of tree hydraulic strategy. *Journal of Geophysical Research: Biogeosciences*, *121*(7), 1792–1813.

Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part i—a discussion of principles. *Journal of hydrology*, *10*(3), 282–290.

Norman, J. (1979). Modeling the complete crop canopy, modification of the aerial environment of plants, asae monogr. *BJ Barfield, JF Gerber*, 249–277.

NVIDIA, Vingelmann, P., & Fitzek, F. H. (2020). *Cuda, release: 10.2.89.* Retrieved from `https://developer.nvidia.com/cuda-toolkit`

ROCm, A. (2019). *a new era in open gpu computing*.

Ryder, J., Polcher, J., Peylin, P., Ottlé, C., Chen, Y., van Gorsel, E., ... others (2016). A multi-layer land surface energy budget model for implicit coupling with global atmospheric simulations. *Geoscientific Model Development*, *9*(1), 223–245.

Scanlon, T. M., & Albertson, J. D. (2001). Turbulent transport of carbon dioxide and water vapor within a vegetation canopy during unstable conditions: Identification of episodes using wavelet analysis. *Journal of Geophysical Research: Atmospheres*, *106*(D7), 7251-7262. Retrieved from `https ://`

agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2000JD900662 doi: https://doi.org/10.1029/2000JD900662

Schmid, H., Su, H.-B., Vogel, C., & Curtis, P. (2003). Ecosystem-atmosphere exchange of carbon dioxide over a mixed hardwood forest in northern lower michigan. *Journal of Geophysical Research: Atmospheres*, *108*(D14).

Sperry, J. S., Adler, F., Campbell, G., & Comstock, J. (1998). Limitation of plant water use by rhizosphere and xylem conductance: results from a model. *Plant, Cell & Environment*, *21*(4), 347–359.

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.

Stone, J. E., Gohara, D., & Shi, G. (2010). Opencl: A parallel programming standard for heterogeneous computing systems. *Computing in science & engineering*, *12*(3), 66.

Stovall, A. E., Shugart, H., & Yang, X. (2019). Tree height explains mortality risk during an intense drought. *Nature Communications*, *10*(1), 1–6.

Wang, Y., Köhler, P., He, L., Doughty, R., Braghiere, R. K., Wood, J. D., & Frankenberg, C. (2021). Testing stomatal models at stand level in deciduous angiosperm and evergreen gymnosperm forests using clima land (v0.1). *Geoscientific Model Development Discussions*, *2021*, 1–35. Retrieved from https://gmd.copernicus.org/preprints/gmd-2021-154/ doi: 10.5194/gmd-2021-154

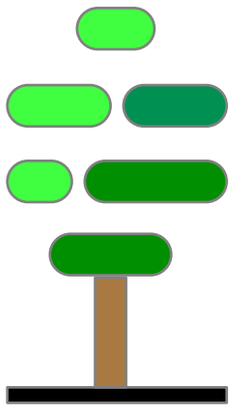
Wang, Y., Sperry, J. S., Anderegg, W. R., Venturas, M. D., & Trugman, A. T. (2020). A theoretical and empirical assessment of stomatal optimization modeling. *New Phytologist*, *227*(2), 311–325.

Xu, X., Medvigy, D., Powers, J. S., Becknell, J. M., & Guan, K. (2016). Diversity in plant hydraulic traits explains seasonal and inter-annual variations of vegetation dynamics in seasonally dry tropical forests. *New Phytologist*, *212*(1), 80–95.

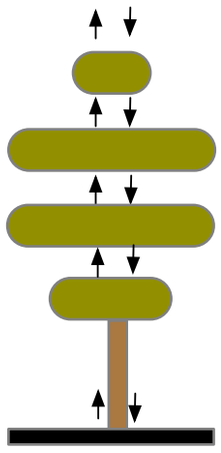
Zellweger, F., De Frenne, P., Lenoir, J., Vangansbeke, P., Verheyen, K., Bernhardt-Römermann, M., ... others (2020). Forest microclimate dynamics drive plant responses to warming. *Science*, *368*(6492), 772–775.

Figure 1.

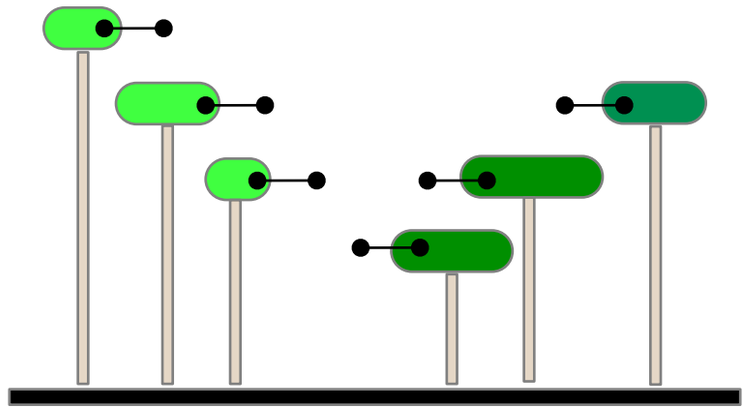
(a) Multi-layer canopy with sunlit and shaded leaves



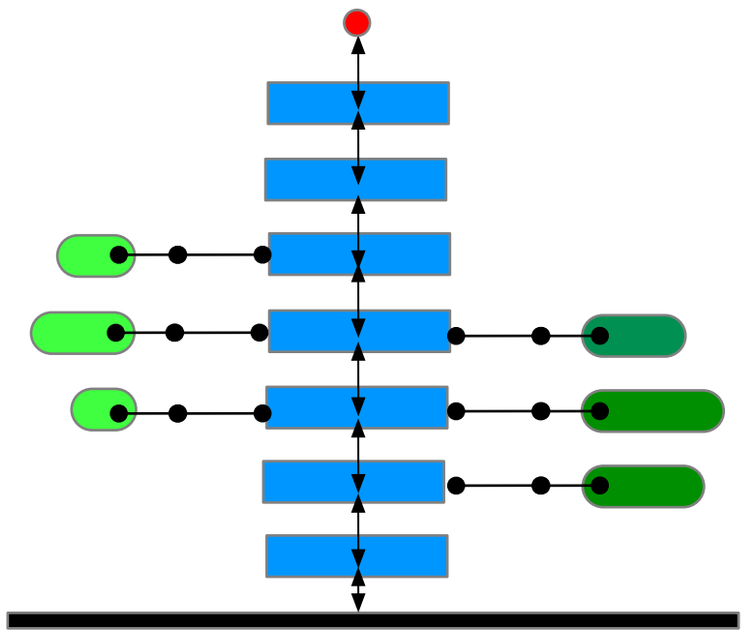
(b) Radiation model



(c) Stomatal conductance model



(e) Multi-layer canopy space model



(d) Roughness sublayer model

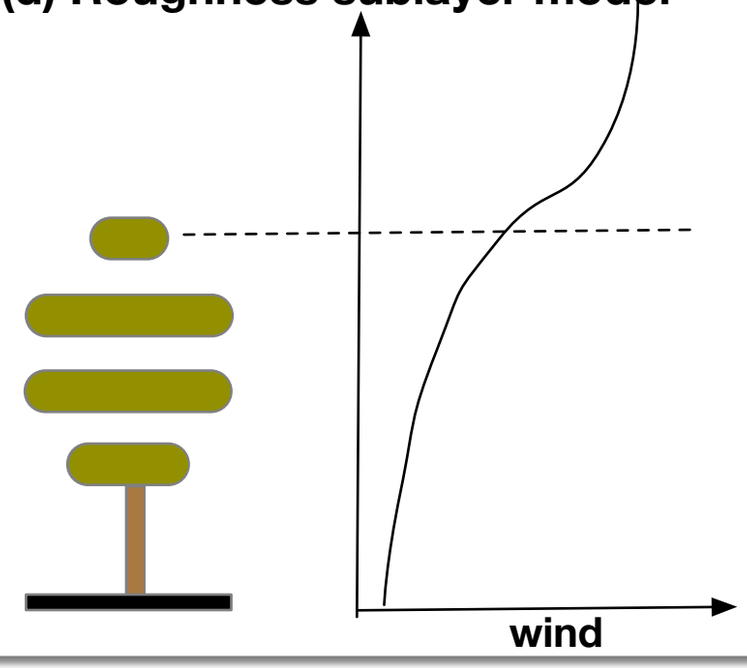
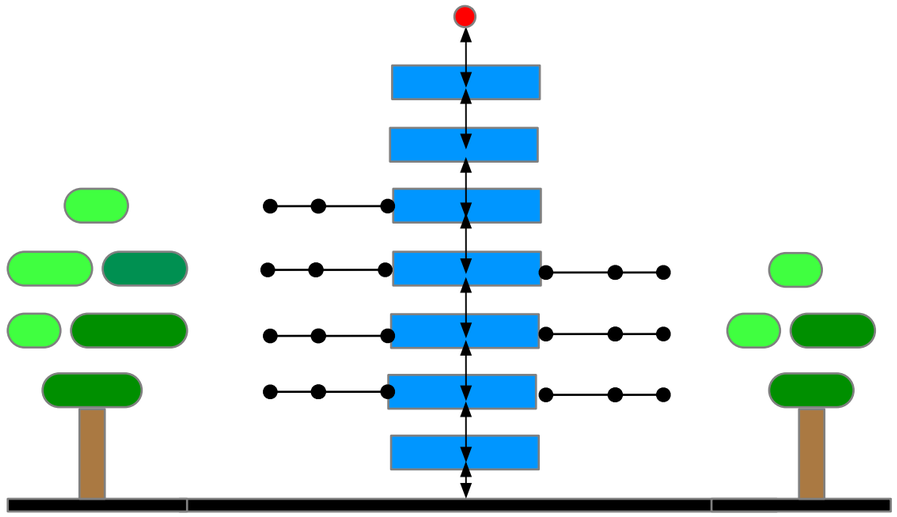


Figure 2.

(a) Multiple vegetation canopies interacting with a single canopy air space



(b) Discretized linear system (Ax=b) solved using PETSc

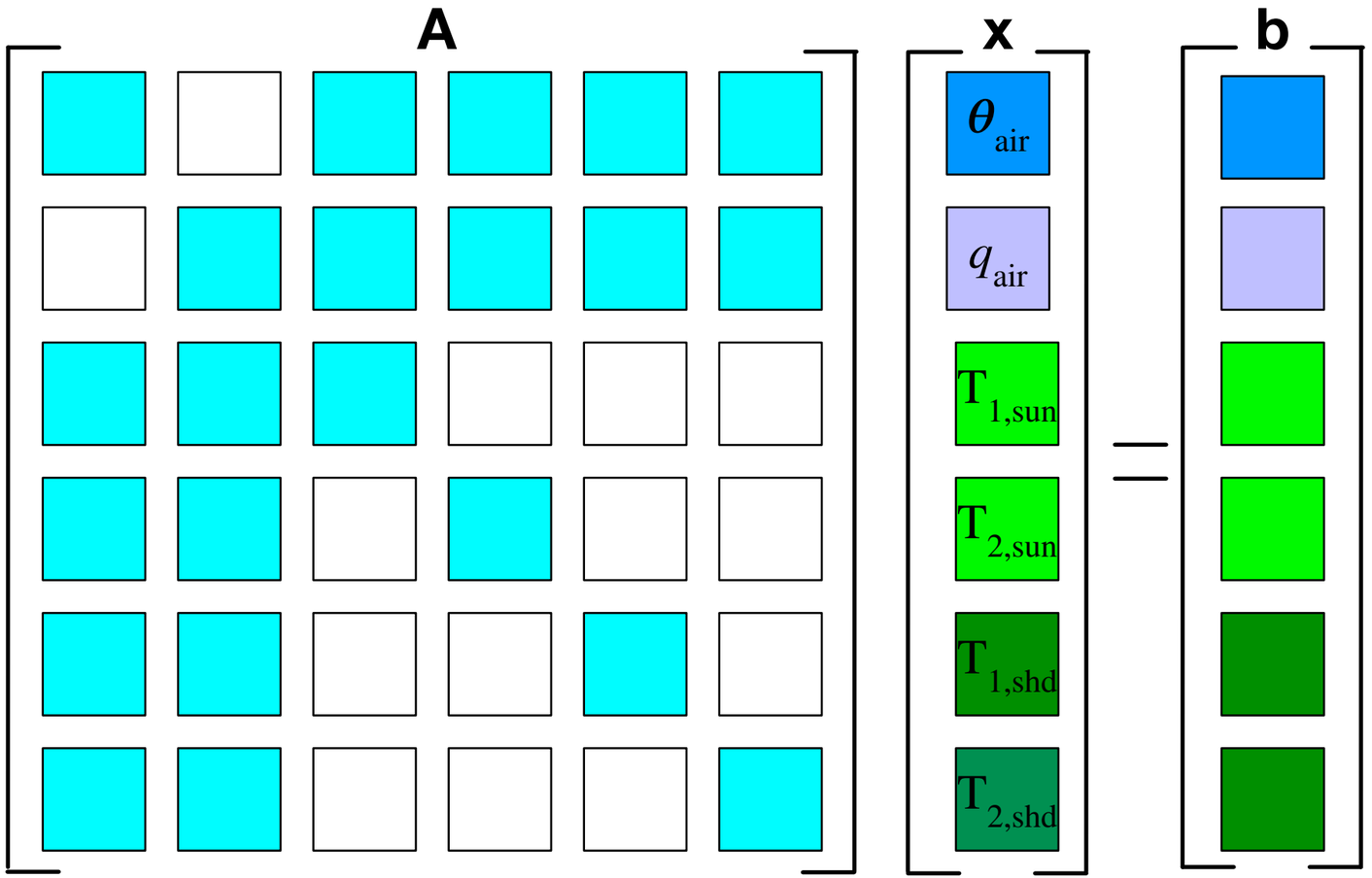


Figure 3.

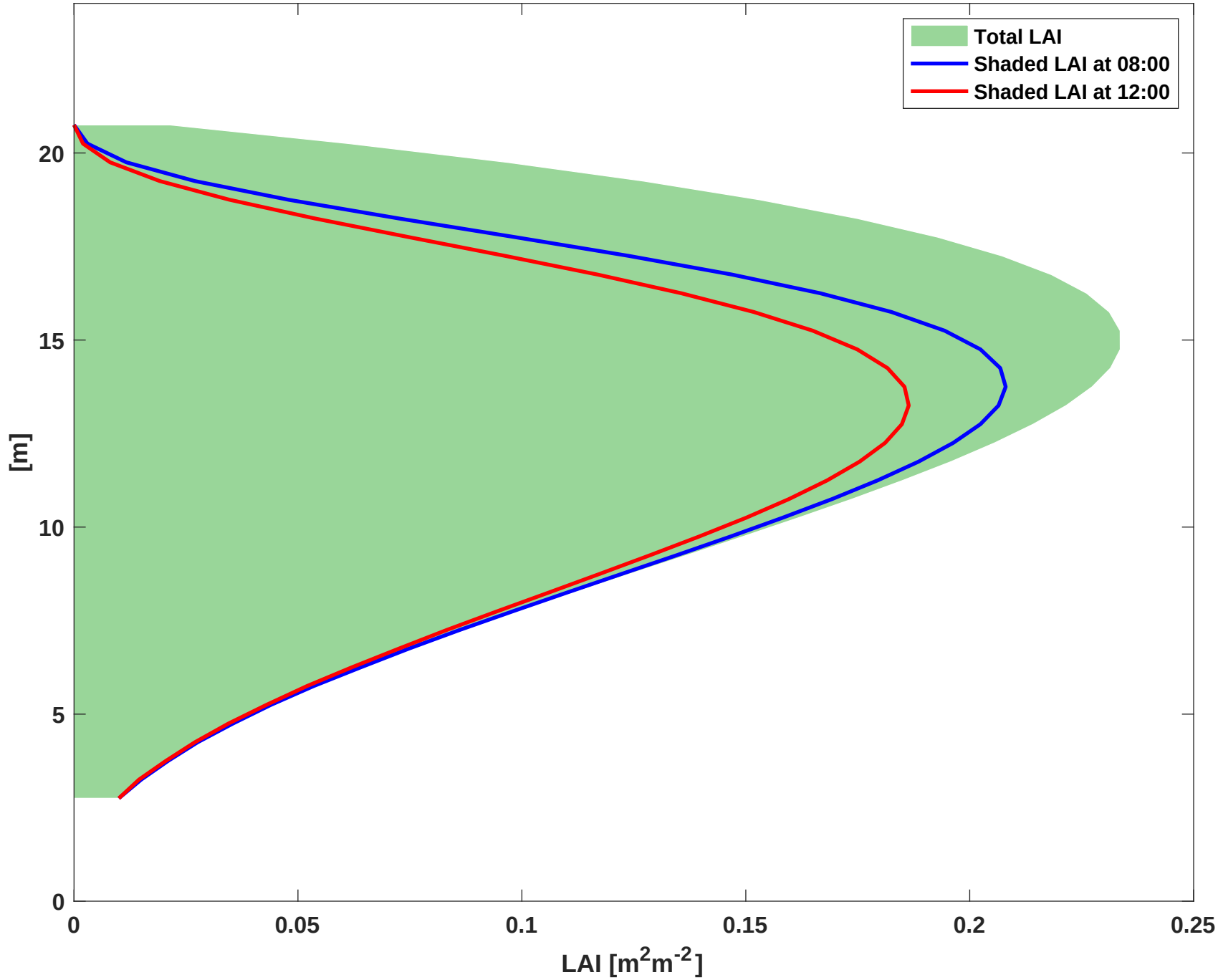

Total LAI
Shaded LAI at 08:00
Shaded LAI at 12:00
20
15
10
5
0
[m]
0
0.05
0.1
0.15
0.2
0.25
LAI [m2m-2]

Figure 4.

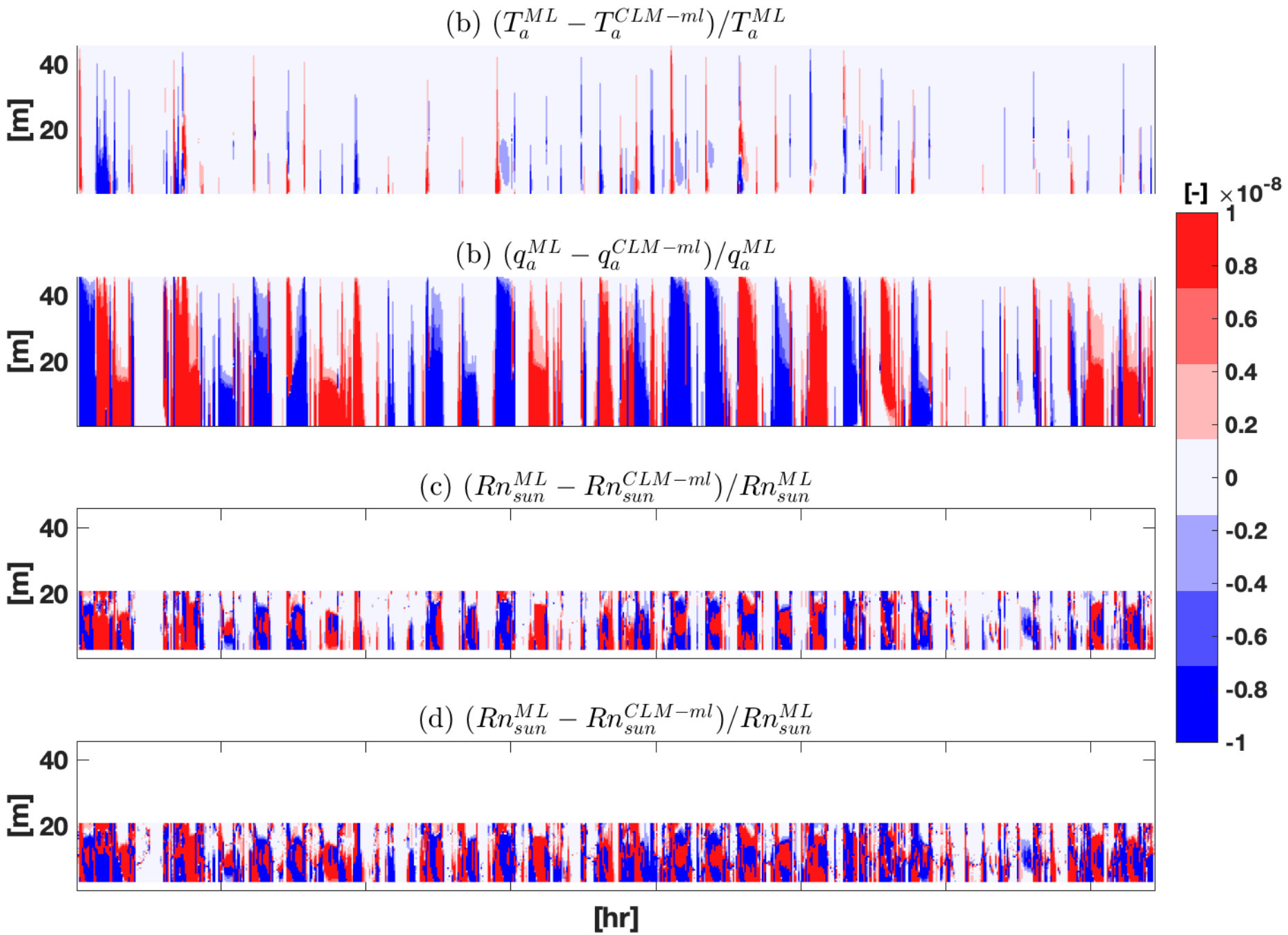

(b) $(T_a^{ML} - T_a^{CLM-ml})/T_a^{ML}$
(b) $(q_a^{ML} - q_a^{CLM-ml})/q_a^{ML}$
(c) $(Rn_{sun}^{ML} - Rn_{sun}^{CLM-ml})/Rn_{sun}^{ML}$
(d) $(Rn_{sun}^{ML} - Rn_{sun}^{CLM-ml})/Rn_{sun}^{ML}$
[m]
40
20
[hr]
[-] ×10-8
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1

Figure 5.

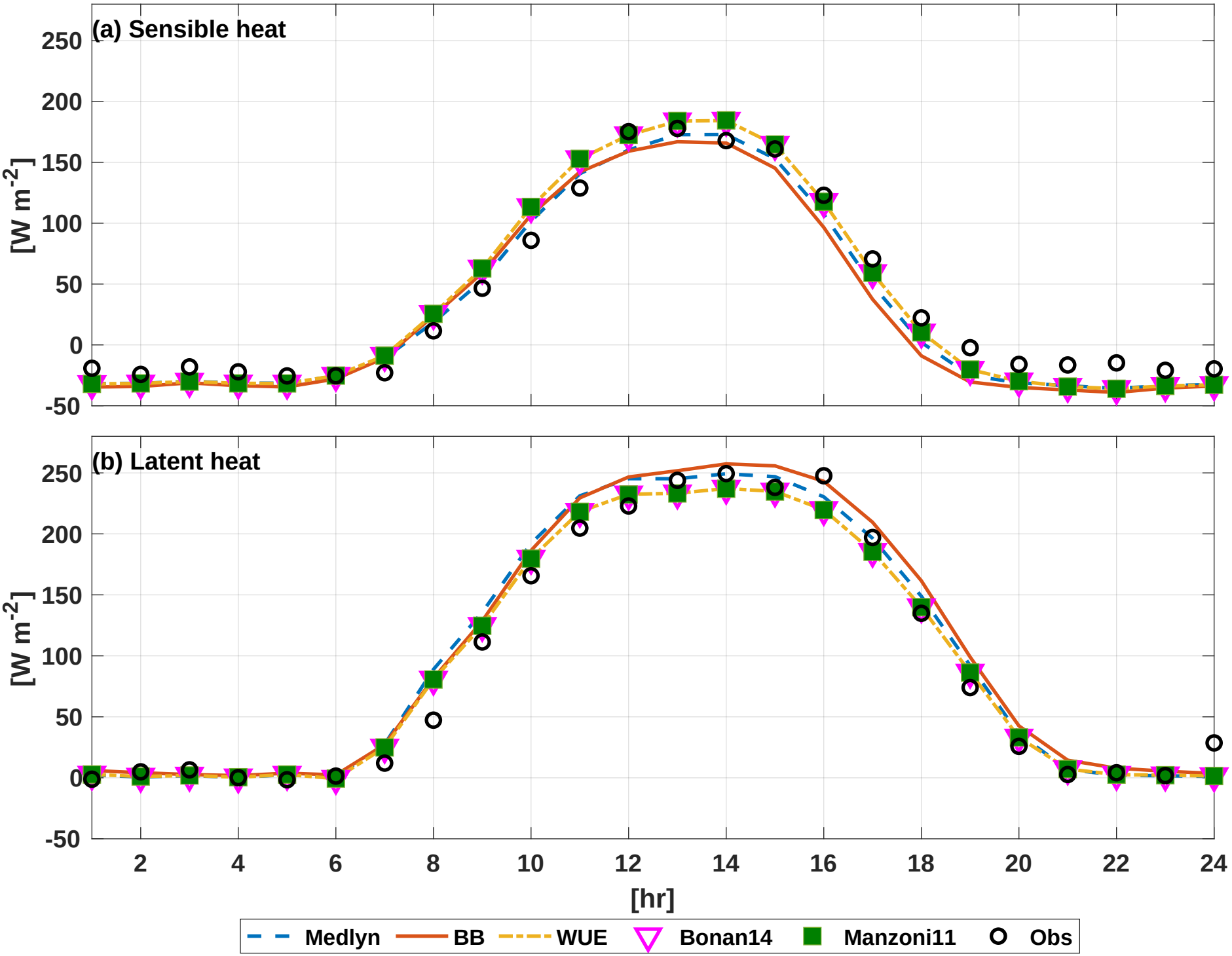

(a) Sensible heat
(b) Latent heat
[W m-2]
[W m-2]
250
200
150
100
50
0
-50
2
4
6
8
10
12
14
16
18
20
22
24
[hr]
Medlyn
BB
WUE
Bonan14
Manzoni11
Obs

Figure 6.

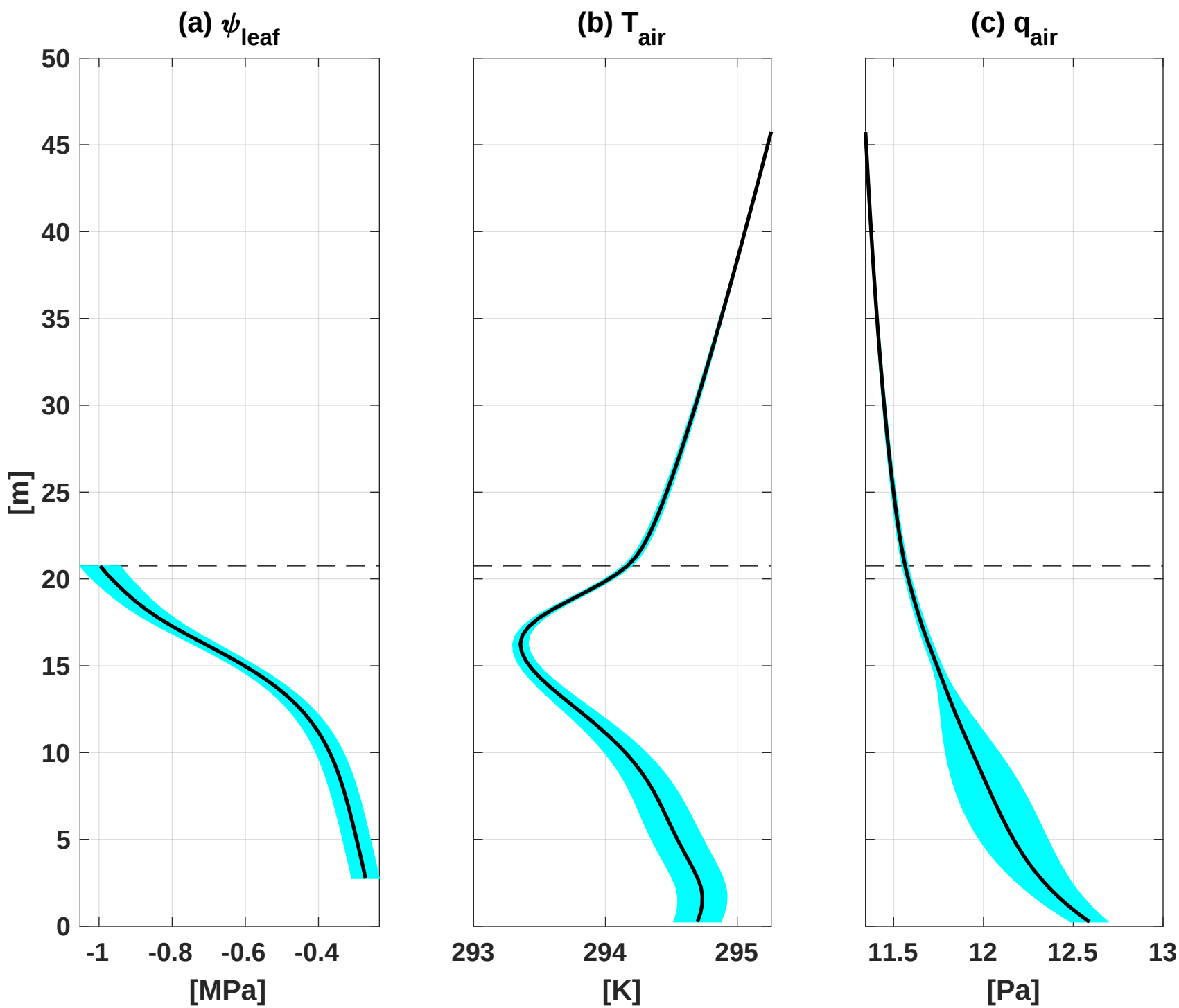

(a) ψ_{leaf}
(b) T_{air}
(c) q_{air}
[m]
0
5
10
15
20
25
30
35
40
45
50
-1
-0.8
-0.6
-0.4
[MPa]
293
294
295
[K]
11.5
12
12.5
13
[Pa]

Figure 7.

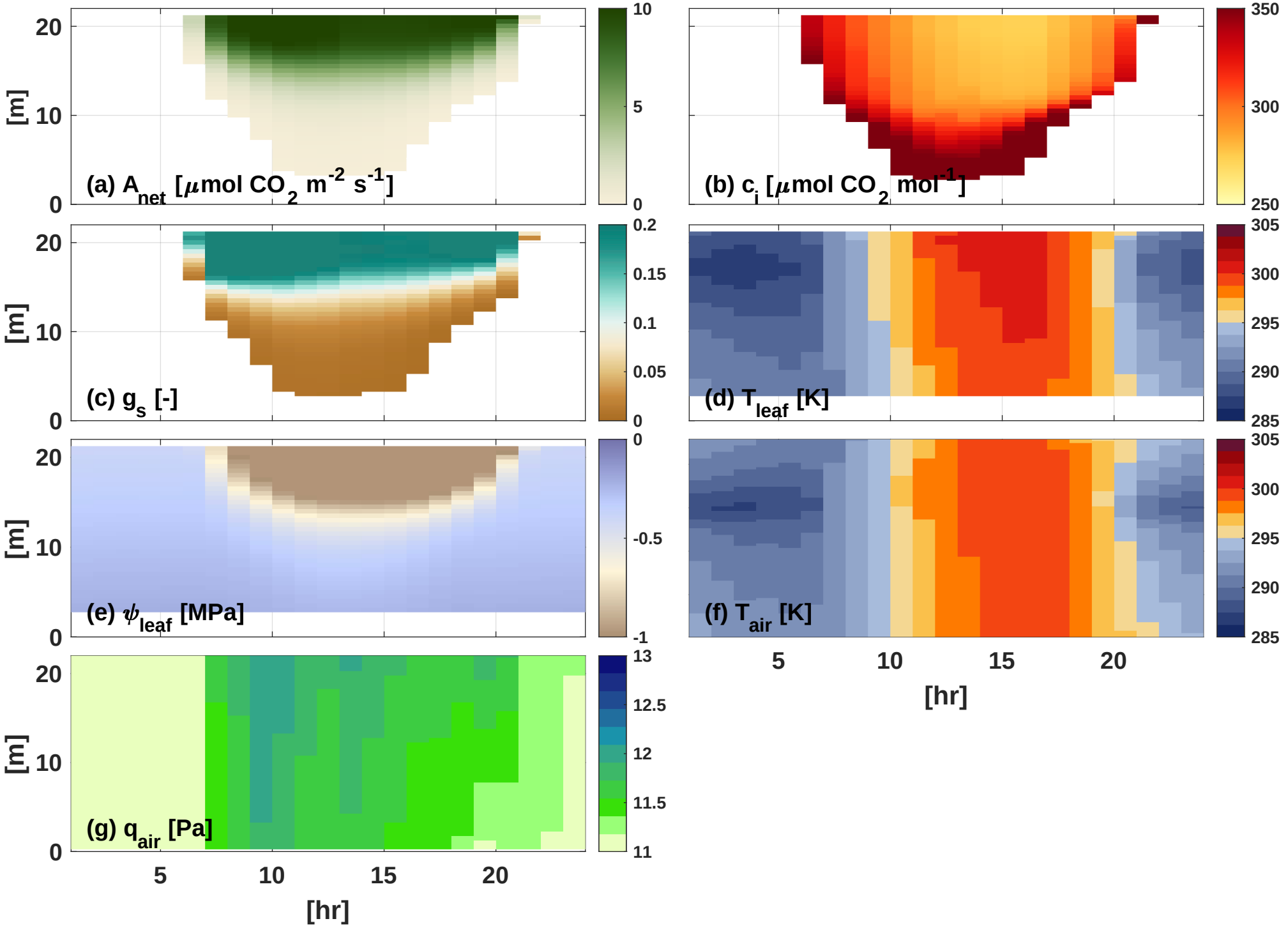

(a) A_{net} [μmol CO_2 m^{-2} s^{-1}]
(b) c_i [μmol CO_2 mol^{-1}]
(c) g_s [-]
(d) T_{leaf} [K]
(e) ψ_{leaf} [MPa]
(f) T_{air} [K]
(g) q_{air} [Pa]
[m]
[hr]
0
10
20
5
15
10
5
0
350
300
250
0.2
0.15
0.1
0.05
0
305
300
295
290
285
-0.5
-1
13
12.5
12
11.5
11

Figure 8.

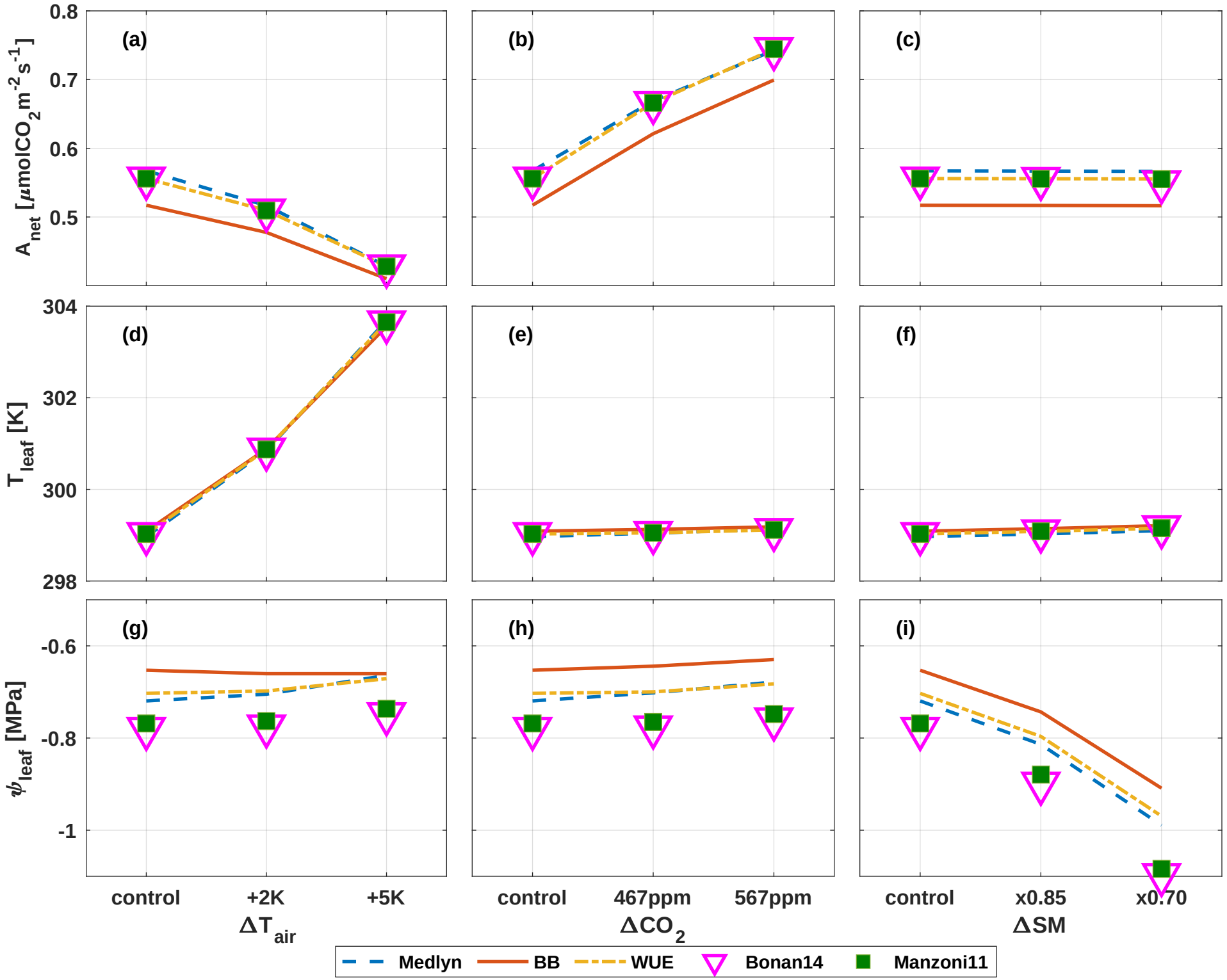

(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
(i)
A_net [μmolCO2 m-2 s-1]
T_leaf [K]
ψ_leaf [MPa]
0.8
0.7
0.6
0.5
304
302
300
298
-0.6
-0.8
-1
control
+2K
+5K
ΔT_air
control
467ppm
567ppm
ΔCO2
control
x0.85
x0.70
ΔSM
Medlyn
BB
WUE
Bonan14
Manzoni11

Figure 9.

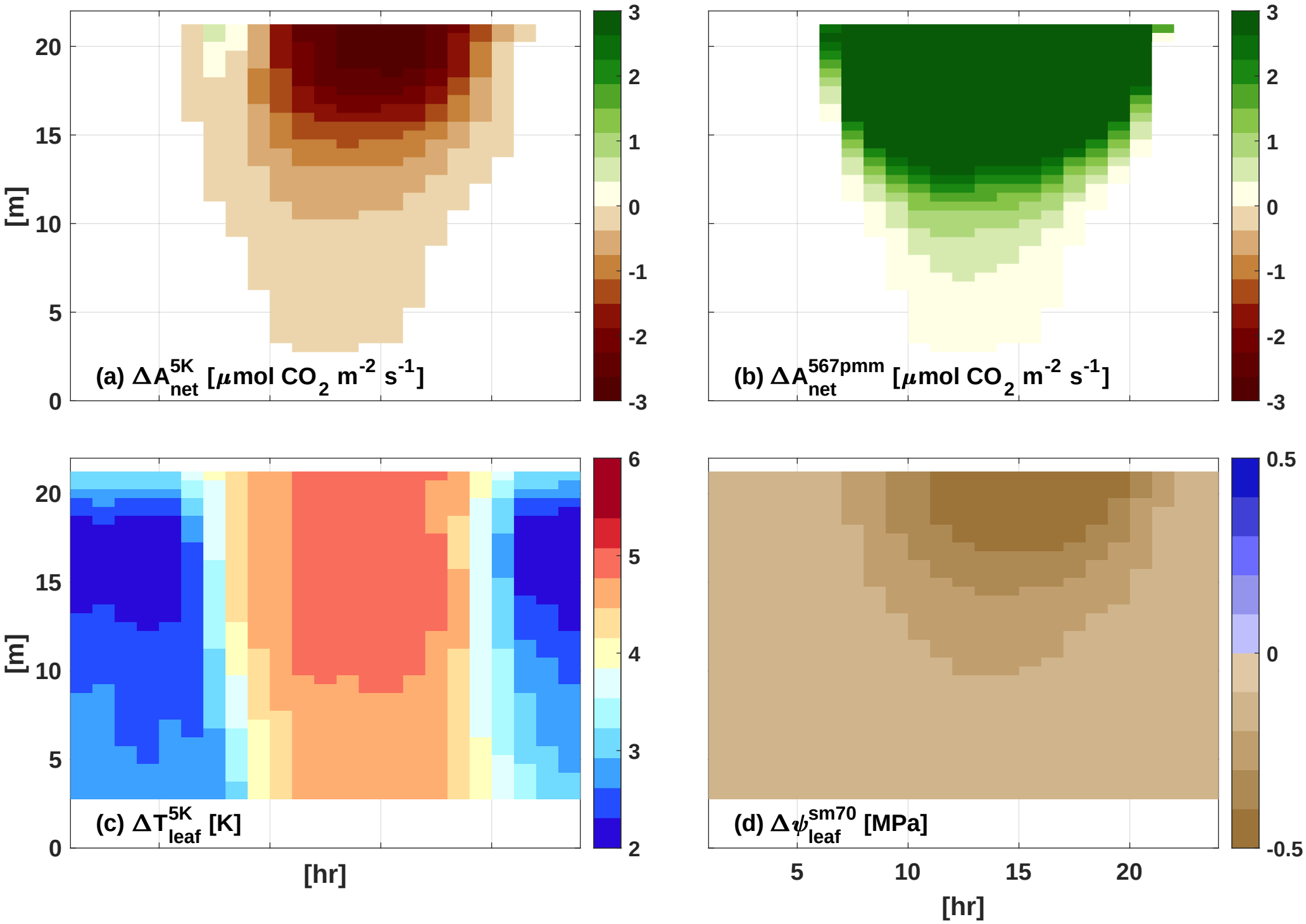

(a) ΔA_{net}^{5K} [μmol CO_2 m^{-2} s^{-1}]
(b) ΔA_{net}^{567pmm} [μmol CO_2 m^{-2} s^{-1}]
(c) ΔT_{leaf}^{5K} [K]
(d) $\Delta \psi_{leaf}^{sm70}$ [MPa]
[m]
[hr]
0
5
10
15
20
3
2
1
0
-1
-2
-3
6
5
4
3
2
0.5
0
-0.5

Figure 10.

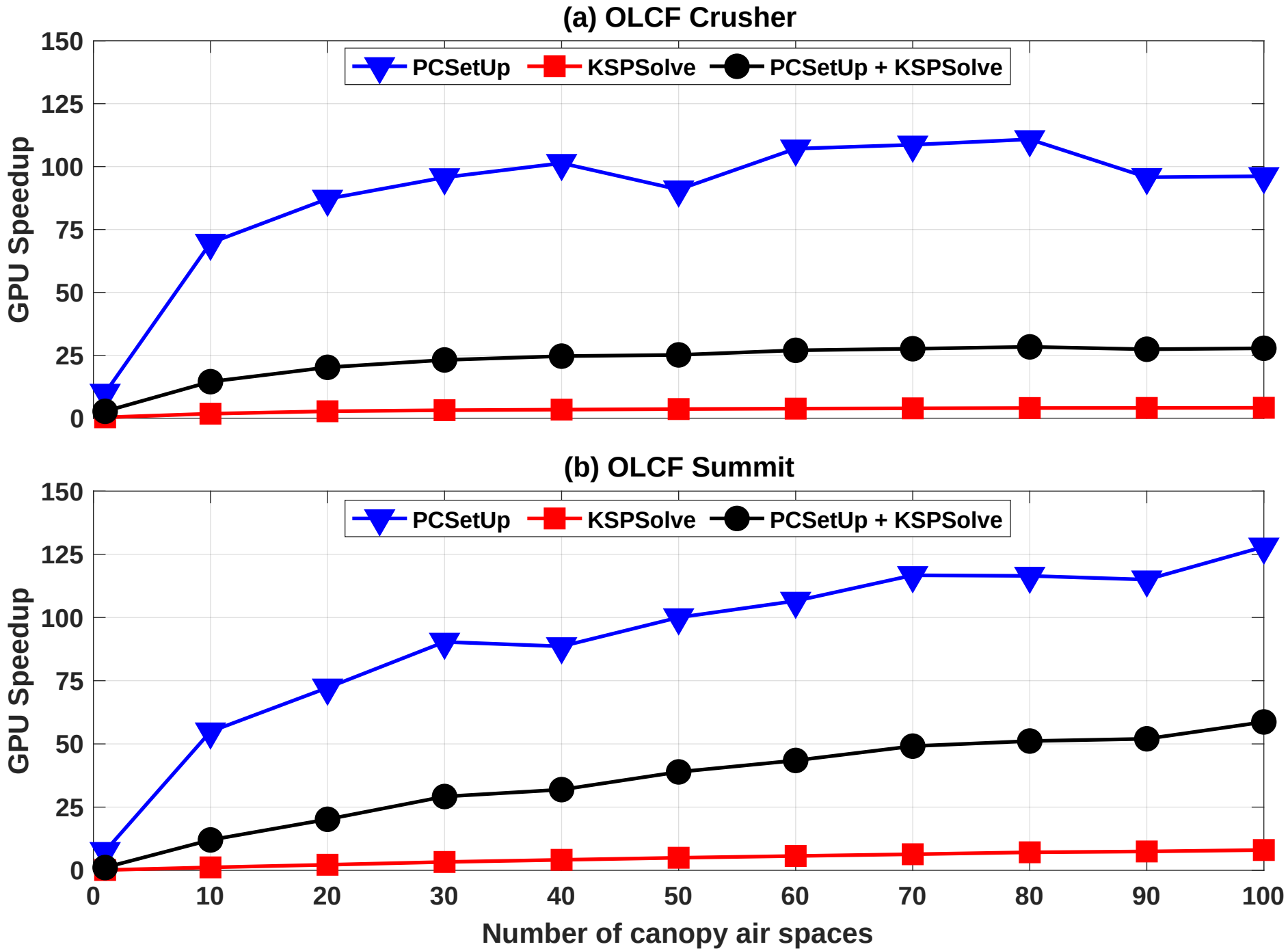
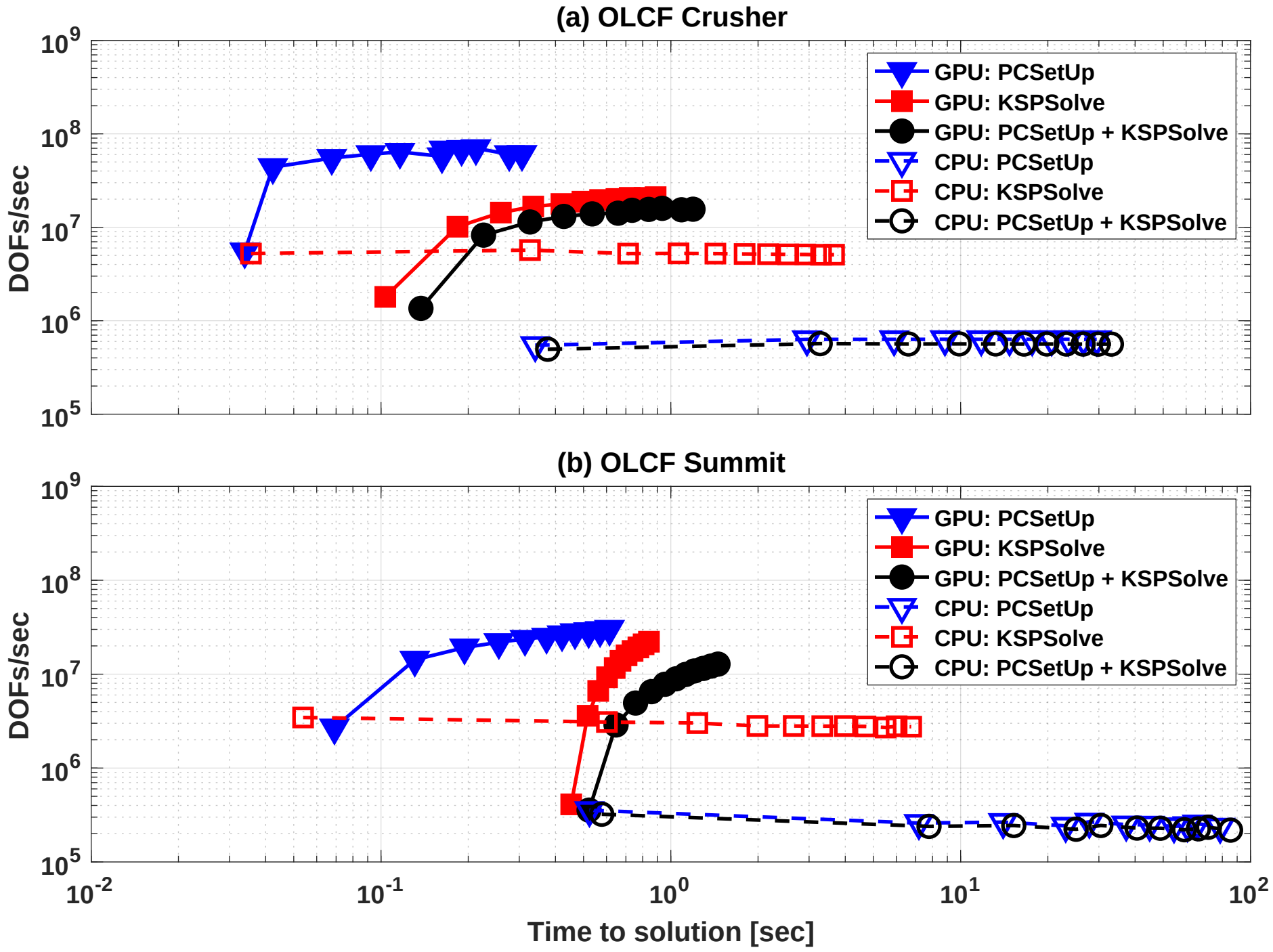

(a) OLCF Crusher
PCSetUp
KSPSolve
PCSetUp + KSPSolve
GPU Speedup
150
125
100
75
50
25
0
(b) OLCF Summit
PCSetUp
KSPSolve
PCSetUp + KSPSolve
GPU Speedup
150
125
100
75
50
25
0
0
10
20
30
40
50
60
70
80
90
100
Number of canopy air spaces

Figure 11.


(a) OLCF Crusher
(b) OLCF Summit
DOFs/sec
Time to solution [sec]
GPU: PCSetUp
GPU: KSPSolve
GPU: PCSetUp + KSPSolve
CPU: PCSetUp
CPU: KSPSolve
CPU: PCSetUp + KSPSolve
10^9
10^8
10^7
10^6
10^5
10^-2
10^-1
10^0
10^1
10^2