Estimation of above- and below-ground ecosystem parameters for the DVM-DOS-TEM model using MADS: a synthetic case study

Elchin E. Jafarov[1], Hélène Genet[2], Velimir V. Vesselinov (Monty)[3], Valeria Briones[1], Aiza Kabeer[4], Andrew L. Mullen[1], Benjamin Maglio[2], Tobey Carman[2], Ruth Rutter[2], Joy Clein[2], Chu-Chun Chang[1], Dogukan Teber[1], Trevor Smith[1], Joshua M. Rady[1], Christina Schädel[1], Jennifer D. Watts[1], Brendan M. Rogers[1], Susan M. Natali[1]

[1]Woodwell Climate Research Center, Falmouth, MA, USA
[2]Institute of Arctic Biology, University of Alaska Fairbanks, Fairbanks, AK, USA
[3]EnviTrace LLC, NM, USA
[4]Program in Applied Mathematics, University of Arizona, Tucson, AZ, USA

Correspondence to: Elchin E. Jafarov (ejafarov@woodwellclimate.org

Supplementary Information

S1. Levenberg-Marquardt Algorithm

The core of the LM method is an iterative process that refines the parameter estimates through several steps:

1. Calculating Residuals:
    - Residuals represent the difference between the observed data and the values predicted by the model. By computing these residuals, we can assess how well the current parameters fit the data.
2. Computing the Jacobian Matrix:
    - The Jacobian matrix is constructed by calculating the partial derivatives of the residuals with respect to each parameter. This matrix encapsulates how small changes in parameters affect the residuals, providing a linear approximation of the system's behavior around the current parameter estimates.
3. Formulating the Hessian Approximation:
    - The Hessian matrix is a square matrix of second-order partial derivatives of a scalar-valued function. It describes the local curvature of a function of several variables, providing critical information about the function's convexity and concavity. It is approximated by multiplying the Jacobian matrix by its transpose. This approximation is key to simplifying the problem while retaining essential information about the curvature of the explored parameter space.
4. Updating the Parameters:

- To determine the next set of parameter estimates, a correction term is calculated. This term is derived by solving a linear system that incorporates both the Hessian approximation and the damping parameter. The resulting correction term is added to the current parameter estimates, nudging them towards a better fit.

5. Adjusting the Damping Parameter

The damping parameter λ plays a pivotal role in balancing the optimization approach:

- If the new parameter estimates lead to a better fit (i.e., the sum of the squared residuals decreases), the damping parameter is reduced, typically by a factor of 10. This reduction shifts the algorithm towards the Gauss-Newton method, which can converge more quickly when the solution is near.
- Conversely, if the new estimates do not improve the fit, the damping parameter is increased, also typically by a factor of 10. This increase shifts the algorithm towards gradient descent, enhancing stability and ensuring progress even in challenging regions of the parameter space.

6. Checking for Convergence

- The iterative process continues until certain convergence criteria are met. These criteria include a sufficiently small change in the parameter values or the residual sum of squares. When the algorithm converges, it means that further iterations no longer result in significant improvements, indicating that an optimal or near-optimal solution has been found.

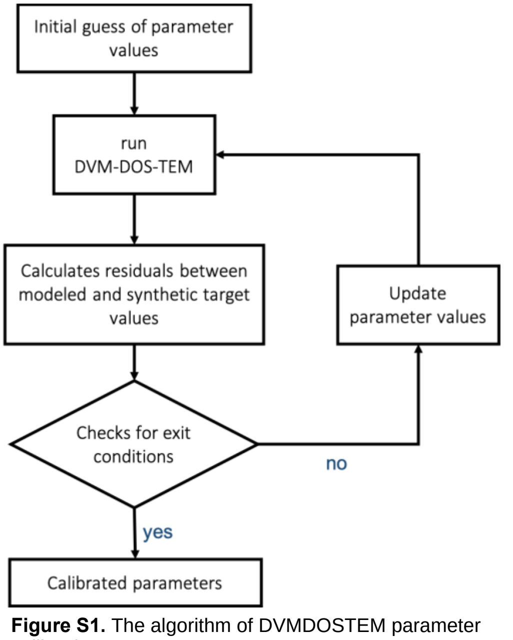

Figure S1. The algorithm of DVMDOSTEM parameter calibration.

S2. Maximum rate of C assimilation (c_{max})

In the DVM-DOS-TEM model, GPP is described by the following equation:

$$GPP_{PFT} = c_{max} \cdot f(CO_2) \cdot f(PAR) \cdot f(T) \cdot f(LEAF) \cdot f(FOLIAGE) \quad \text{(S1)}$$

$$\cdot f(THAW_{pct}) \cdot f(FPC) \cdot f(NAV)$$

where $f(CO_2)$ is a function representing the effect of atmospheric CO_2 and canopy conductance on GPP, $f(PAR)$ represents the effect of photosynthetically active radiation, $f(T)$ represents the direct effect of air temperature on GPP, $f(LEAF)$ represents the effect of leaf phenology on GPP, $f(FOLIAGE)$ represents the effect of canopy development on GPP, $f(THAW_{pct})$ varies between 0 and 1 and defines the length of the growing season based on soil temperature at 10 cm, $f(FPC)$ represents the effect of competition among PFT for light, based on foliar projected cover, a function of Beer's law (McGuire et al. 1992), $f(NAV)$ is dynamically calculated to model the control of plant N status on GPP for a given PFT (see section 5). $f(PAR)$, $f(T)$, f(NAV), $f(THAW_{pct})$, $f(FOLIAGE)$ and $f(LEAF)$ are multipliers varying between 0 and 1.

S3. Maximum plant N uptake rate ($\boldsymbol{n_{max}}$)

$\boldsymbol{n_{max}}$ is the rate limiting factor of vegetation nitrogen uptake in the absence of nitrogen limitation. Vegetation N uptake is also constrained by soil moisture $f(LWC)$, and temperature $f(Ts)$, fine root biomass $f(FR)$, canopy development $f(FOLIAGE)$, available nitrogen $f(NAV)$, and plant nitrogen requirement $f(N_{require})$.

$$N_{uptake} = n_{max} \cdot f(Ts) \cdot f(LWC) \cdot f(FOLIAGE) \cdot f(NAV) \cdot f(N_{require}) \quad \text{(S2)}$$

S4. The limiting rate of maintenance respiration ($\boldsymbol{Kr_b}$)

Kr is the limiting rate of vegetation maintenance respiration (R_m) at 0ºC:

$$R_m = Kr \cdot C \quad \text{(S3)}$$

where C is vegetation C pool. Kr is itself a function of vegetation C pool:

$$Kr = exp[(Kr_a \cdot C) + Kr_b] \quad \text{(S4)}$$

Kr_a is usually set to -8.06 10^{-5}, and Kr_b is calibrated for every vegetation compartment: leaf, stem, and root. Since the relationship between biomass and maintenance respiration is not linear and the slope of the relationship decreases as biomass increases, Kr_b is a negative number.

S5. Plant-soil nitrogen feedback

Vegetation productivity is downregulated based on a comparison of the N demand to accomplish growth and the Nsupply resulting from plant N uptake, mobilisation, and resorption. If N demand is higher than N supply, vegetation productivity is reduced proportionally. In a first step, GPP is computed for every PFT, without considering N limitation (GPP^*). After computation of R_m, net primary productivity (NPP^*) and growth respiration (Rg) without nitrogen limitation are computed as follow:

$$NPP^* = (GPP^* - R_m)/(1 + frg) \quad \text{(S5)}$$

$$Rg^* = NPP^* \cdot frg \quad \text{(S6)}$$

Where frg is a parameter setting the fraction of NPP required to achieve new tissue production. This estimate of NPP^* is then used to estimate N requirement by dividing it to the parameterized $C\!:\!N$ ratio for new growth ($C\!:\!N_{even}$).

$$N_{require} = NPP^*/(C\!:\!N_{even}) \quad \text{(S7)}$$

Growth reduction associated with N limitation is computed as the ratio between N supply and N requirement. This ratio is finally used to downregulate NPP^*, Rg^* and GPP^* in the case N supply is lower than N requirement to estimate actual NPP, actual Rg and actual GPP.

$$NPP = NPP^* \cdot (N_{supply}/N_{require}) \quad \text{(S8)}$$

$$GPP = NPP + R_m + R_a \quad \text{(S9)}$$

S6. Rate of C litterfall production (c_{fall})

c_{fall} is the limiting rate of vegetation C litterfall (transferring organic carbon from vegetation to soil) and has the following equation:

$$c_{ltrfall} = c_{fall} \cdot C_{veg} \quad \text{(S10)}$$

where c_{fall} is a nondimensional term calibrated for every vegetation compartment: leaf, stem, and root.

S7. Rate of N in litter production (n_{fall})

Similarly, n_{fall} limiting rate parameter of vegetation N litterfall:

$$n_{ltrfall} = n_{fall} \cdot N_{veg}, \quad \text{(S11)}$$

where N_{veg} is the vegetation N pool, n_{fall} and is calibrated for every PFT and PFT compartment: leaf, stem, and root.

S8. Soil parameters

n_{micb}^{up} is the limiting rate of microbial N uptake per unit of detrital C respired (g/g). n_{micb}^{up} directly influences N immobilisation by decomposers, and net mineralization which controls the amount of inorganic N produced during decomposition of the soil organic matter minus N immobilised by decomposers.

kdcs are the rate-limiting parameters of soil carbon decomposition, calibrated for the four soil carbon pools: litter/raw pool, active pool, and physically and chemically resistant pools. The higher the value kdc is, the faster the turnover is. Therefore,

$$kdc_{rawC} > kdc_{soma} > kdc_{sompr} > kdc_{somcr} \quad \text{(S12)}$$

relationship has to hold.

$$Rh_l = kdc \cdot C_l \cdot f(\theta) \cdot f(T_s) \quad \text{(S13)}$$

where Rh_l , C_l, T_s, θ, are heterotrophic respiration, C stock, temperature, and moisture of soil layer l, respectively.

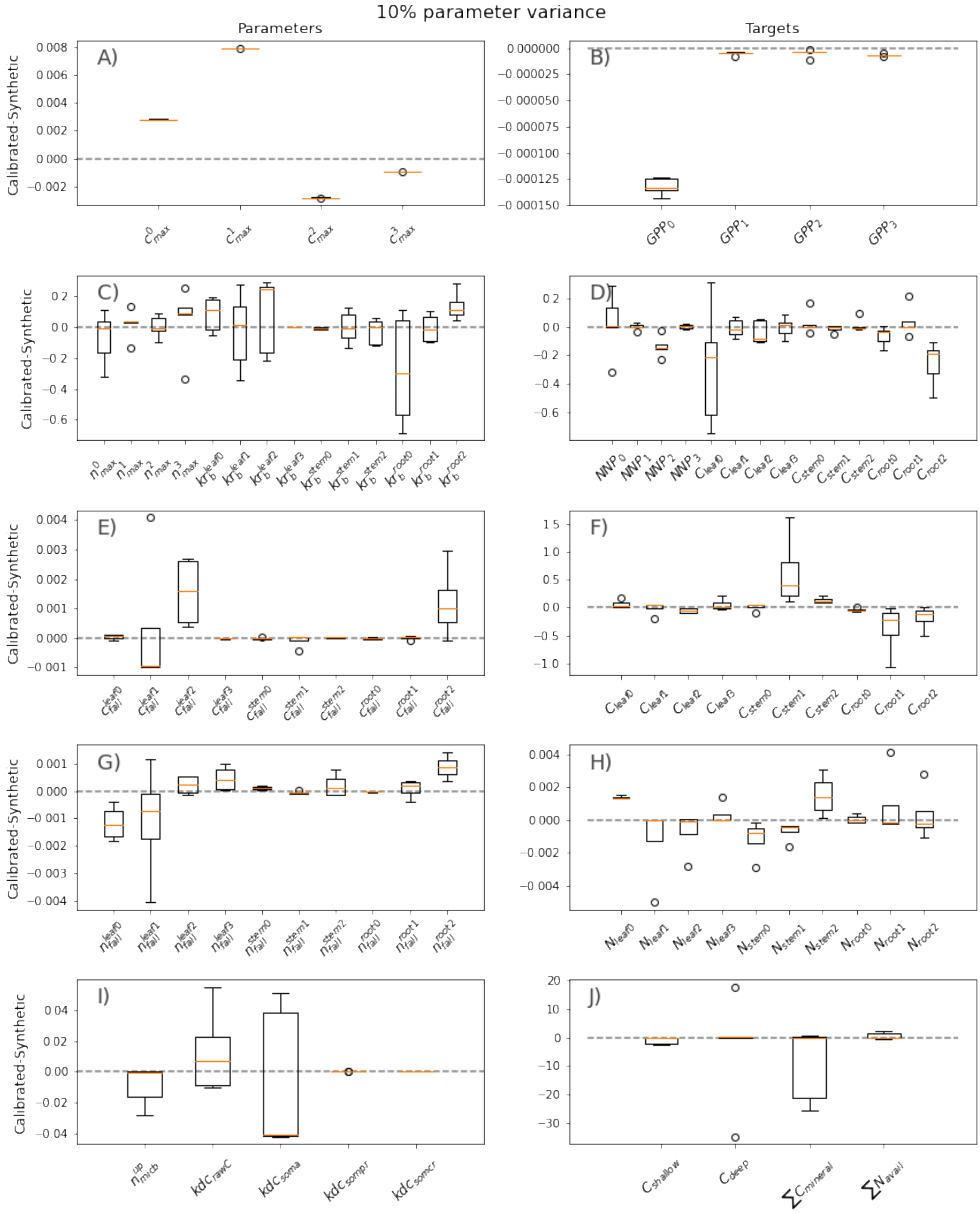

Figure S2. Box plots illustrating the differences between the five best-calibrated parameters and synthetic parameters (A, C, E, G, I), and the differences between calibrated and synthetic targets (B, D, F, H, J) for the 10% parameter variance. The subscripts represent the following plant functional types: 0 - Evergreen Tree, 1 - Deciduous Shrubs, 2 - Deciduous Tree, and 3 - Moss.

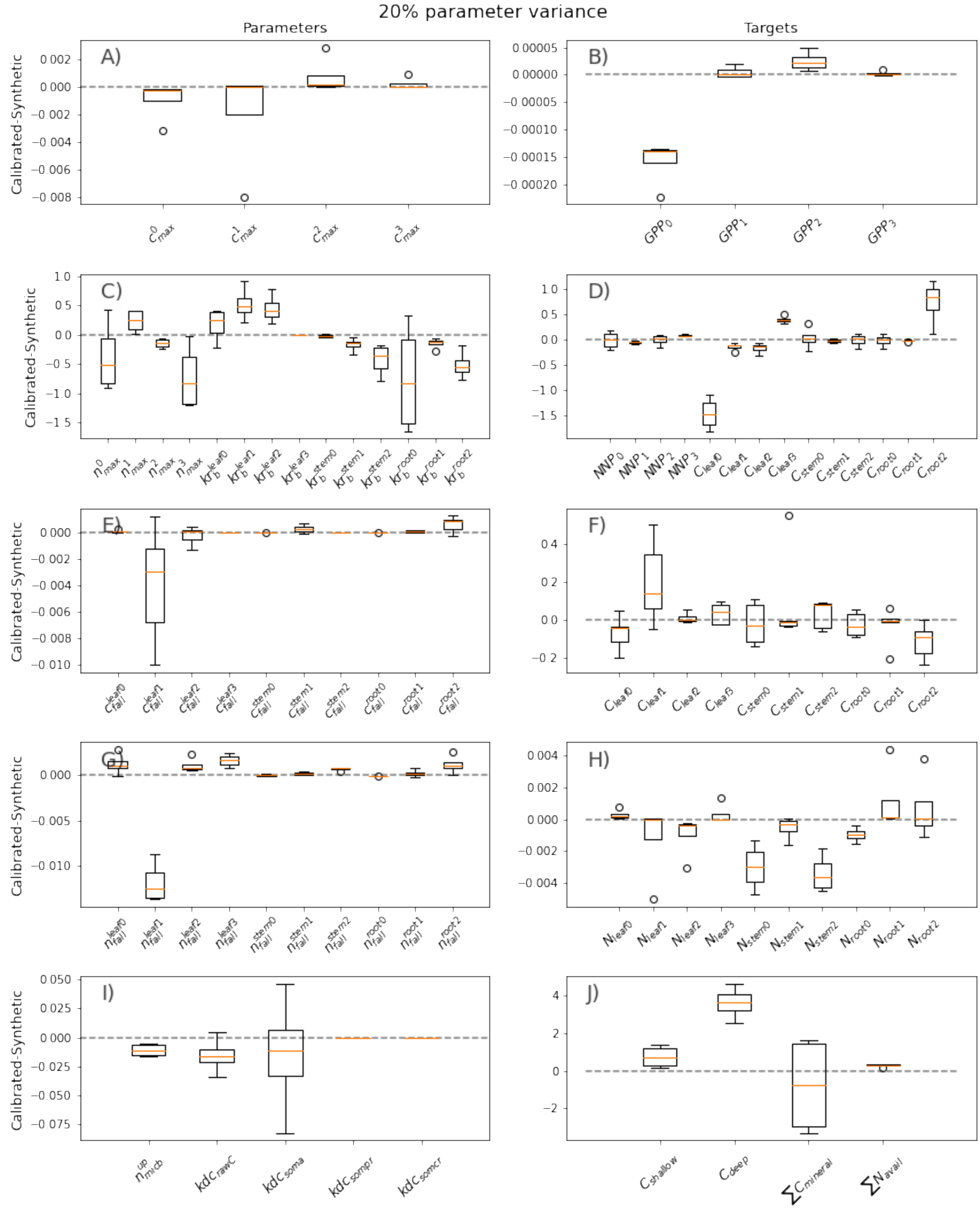

Figure S3. Box plots illustrating the differences between the five best-calibrated parameters and synthetic parameters (A, C, E, G, I), and the differences between calibrated and synthetic targets (B, D, F, H, J) for the 20% parameter variance. The subscripts represent the following plant functional types: 0 - Evergreen Tree, 1 - Deciduous Shrubs, 2 - Deciduous Tree, and 3 - Moss.

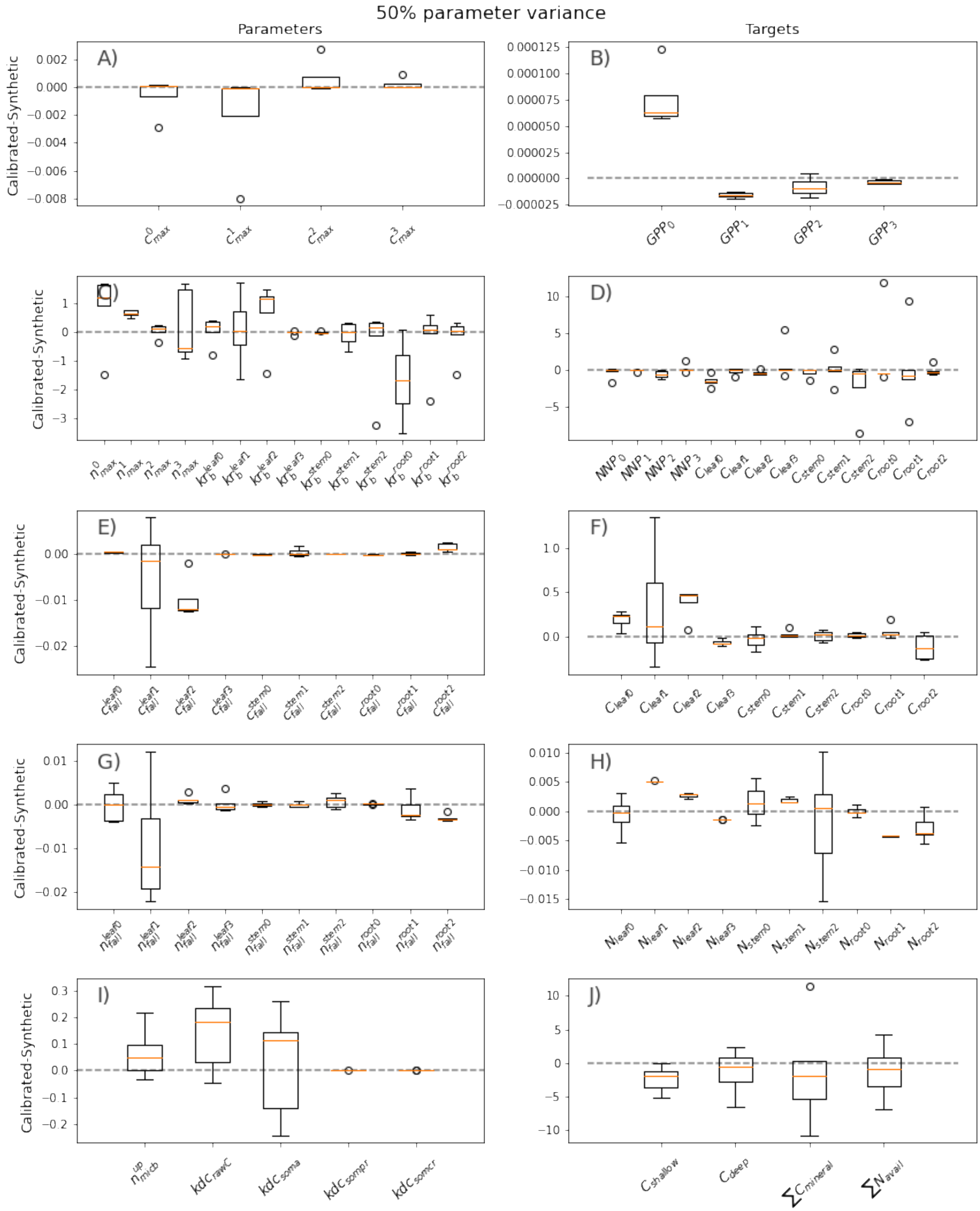

Figure S4. Box plots illustrating the differences between the five best-calibrated parameters and synthetic parameters (A, C, E, G, I), and the differences between calibrated and synthetic targets (B, D, F, H, J) for the 50% parameter variance. The subscripts represent the following plant functional types: 0 - Evergreen Tree, 1 - Deciduous Shrubs, 2 - Deciduous Tree, and 3 - Moss.

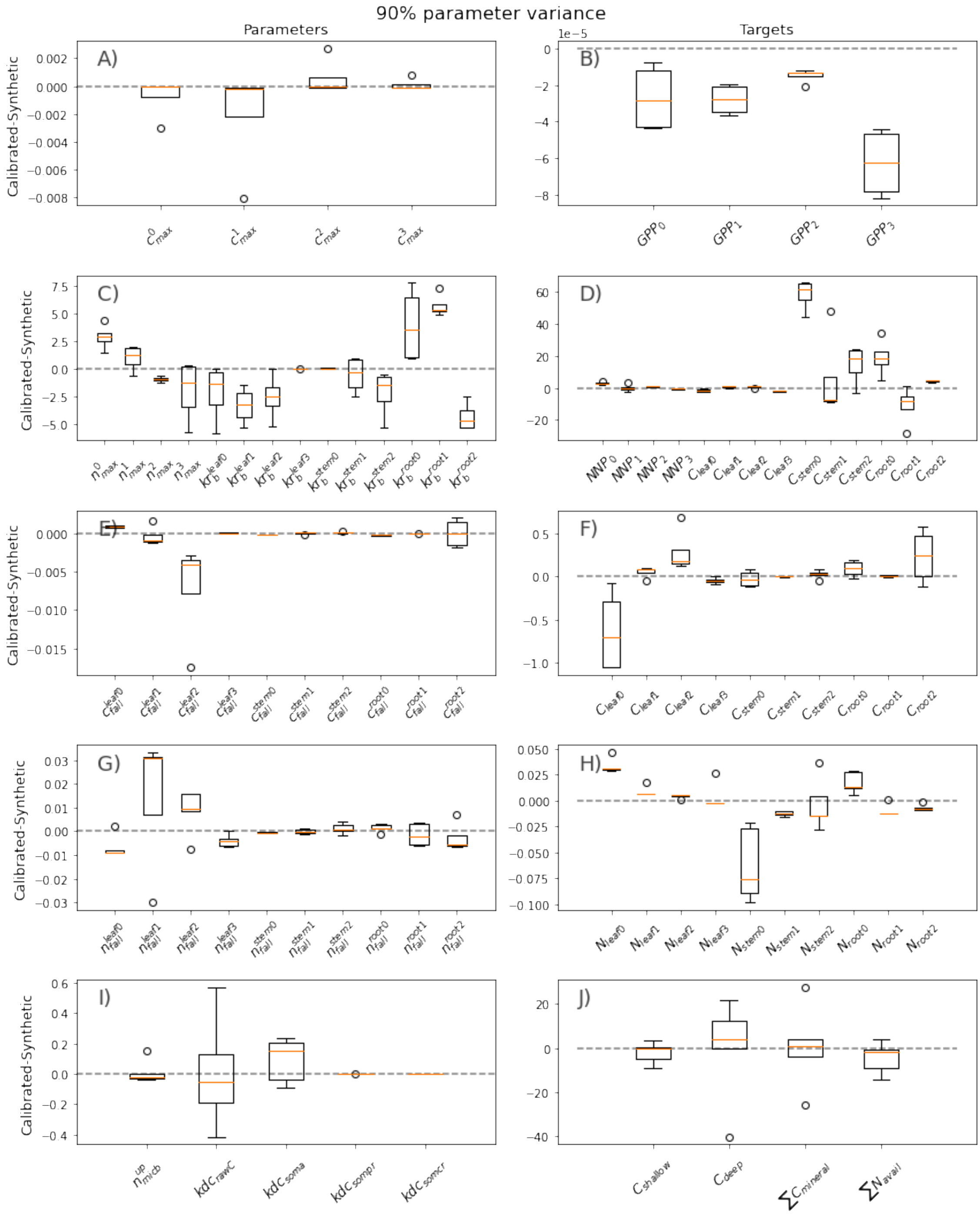

Figure S5. Box plots illustrating the differences between the five best-calibrated parameters and synthetic parameters (A, C, E, G, I), and the differences between calibrated and synthetic targets (B, D, F, H, J) for the 90% parameter variance. The subscripts represent the following plant functional types: 0 - Evergreen Tree, 1 - Deciduous Shrubs, 2 - Deciduous Tree, and 3 - Moss.