Formulating the next generation of models of the terrestrial carbon cycle in the conterminous US

TECO-R-QP Model

Indices/Sets

$i \in I$	geographic sites. i is a particular site. I is the set of all sites
$j \in J$	biomes. $j_1 \ldots j_8$ are particular biomes. J is the set of all biomes j_1=evergreen needleleaf forest j_2=deciduous broadleaf forest j_3=mixed forest j_4=woodland j_5=wooded grassland j_6=shrubland j_7=grassland j_8=cropland
$k \in K$	data sets. k is a particular data set that contains observed data on solar radiation, precipitation, etc. in geographical sites (for example, the ORNL DAAC data set). K is the set of all of the data sets

The following Data/Parameters are set by the user prior to solving TECO-R-QP. The values of the parameters are typically determined via literature review, analysis of remote-sensed data, or expert judgment. For the proposed work, the Data/Parameters will be obtained from existing publicly-available data sets including ORNL DAAC.

Data/Parameters

w_k	weight for data set k. If one data set is more reliable than the others, this parameter can be used to weight it more heavily. $0 \leq w_k \leq 1$
$NDVI_{ijk}$	Normalized difference vegetation index in site i in biome j in data set k. NDVI is a measure of vegetation that is approximately equal to 1 if live green vegetation is present and approximately -1 otherwise
SR_{ijk}	annual solar radiation at site i in biome j in data set k
PPT_{ijk}	annual mean precipitation at site i in biome j in data set k
TMP_{ijk}	annual mean temperature at site i in biome j in data set k
ST_{ijk}	soil texture at site i in biome j in data set k
x_{ijk}	$=vec(NDVI_{ijk}, SR_{ijk}, PPT_{ijk}, TMP_{ijk}, ST_{ijk})$. x_{ijk} is a vector of the forcing parameters that affect the state variables at site i in biome j in data set k (see below)
y_{ijk}	observed data point at site i in biome j in data set k
T_ϵ	temperature scalar used to calculate net primary productivity (NPP)
W_ϵ	moisture scalar used to calculate NPP
$ADAR_{ijk}$	absorbed photosynthetically active radiation in site i in biome j in data set k

The following state variables are variables whose values are estimated by solving TECO-R-QP.

State variables

ϵ^*_{ijk}	light use efficiency in site i in biome j in data set k
α^L_{ijk}	proportion of NPP allocated to leaves in site i in biome j in data set k
α^W_{ijk}	proportion of NPP allocated to wood in site i in biome j in data set k
α^R_{ijk}	proportion of NPP allocated to roots in site i in biome j in data set k
ξ^{R1}_{ijk}	proportion of NPP allocated to roots of depth 0-20 cm in site i in biome j in data set k
ξ^{R2}_{ijk}	proportion of NPP allocated to roots of depth 20-50 cm in site i in biome j in data set k
ξ^{R3}_{ijk}	proportion of NPP allocated to roots of depth 50-100 cm in site i in biome j in data set k
θ^F_{ijk}	C partitioning coefficient of the fine litter pool in site i in biome j in data set k
θ^C_{ijk}	C partitioning coefficient of the coarse litter pool in site i in biome j in data set k
θ^{S1}_{ijk}	C partitioning coefficient of soil organic C; depth: 0-20 cm
θ^{S2}_{ijk}	C partitioning coefficient of soil organic C; depth: 20-50 cm
η_{ijk}	fraction of mechanical breakdown for the coarse litter pool
τ^L_{ijk}	C residence time of leaves in site i in biome j in data set k
τ^W_{ijk}	C residence time of wood in site i in biome j in data set k
τ^{R1}_{ijk}	C residence time of roots (depth: 0-20 cm) in site i in biome j in data set k
τ^{R2}_{ijk}	C residence time of roots (depth: 20-50 cm) in site i in biome j in data set k
τ^{R3}_{ijk}	C residence time of roots (depth: 50-100 cm) in site i in biome j in data set k
τ^F_{ijk}	moisture and temperature corrected residence time of fine litter in site i in biome j in data set k
τ^C_{ijk}	moisture and temperature corrected residence time of coarse litter in site i in biome j in data set k
τ^{S1}_{ijk}	moisture and temperature corrected residence time of soil organic C (depth: 0-20 cm)
τ^{S2}_{ijk}	moisture and temperature corrected residence time of soil organic C (depth: 20-50 cm)
τ^{S3}_{ijk}	moisture and temperature corrected residence time of soil organic C (depth: 50-10 cm)
NPP_{ijk}	net primary productivity at site i in biome j in data set k
a_{ijk}	$=vec(\epsilon^*_{ijk},\alpha^L_{ijk},\alpha^W_{ijk},\alpha^R_{ijk},\xi^{R1}_{ijk},\xi^{R2}_{ijk},\xi^{R3}_{ijk},\tau^L_{ijk},\tau^W_{ijk},\theta^F_{ijk},\theta^C_{ijk},\eta_{ijk},\tau^{R1}_{ijk},\tau^{R2}_{ijk},\tau^{R3}_{ijk},$ $\tau^F_{ijk},\tau^C_{ijk},\tau^{S1}_{ijk},\tau^{S2}_{ijk},\tau^{S3}_{ijk},\theta^{S1}_{ijk},\theta^{S2}_{ijk})$. a is a vector of the state variables that describe C residence times and other aspects of the C cycle
$\widehat{y}_{ijk}(x_{ijk}, a_{ijk})$	$\widehat{y}$ is a linear function that maps the forcing parameters x to the vector of state variables a

Formulation

$$\min_{a,\widehat{y}} \sum_{k\in K} w_k \sum_{j\in J} \sum_{i\in I} [y_{ijk} - \widehat{y}_{ijk}(x,a)]^2 \tag{A1}$$

$$\text{s.t.}\alpha^L_{ijk} \geq \alpha^W_{ijk} \qquad i\in I, j\in J, k\in K \tag{A2}$$

$$\alpha^W_{ijk} = 0 \qquad i\in I, j\in \{j_7, j_8\}, k\in K \tag{A3}$$

$$\alpha^L_{ijk} + \alpha^W_{ijk} + \alpha^R_{ijk} = 1 \qquad i\in I, j\in J, k\in K \tag{A4}$$

$$\xi^{R1}_{ijk} \geq \xi^{R2}_{ijk} \qquad i\in I, j\in J, k\in K \tag{A5}$$

$$\xi^{R2}_{ijk} \geq \xi^{R3}_{ijk} \qquad i\in I, j\in J, k\in K \tag{A6}$$

$$\xi^{R1}_{ijk} + \xi^{R2}_{ijk} + \xi^{R3}_{ijk} = 1 \qquad i\in I, j\in J, k\in K \tag{A7}$$

$$\theta^C_{ijk} = 0 \qquad i\in I, j = j_7, k\in K \tag{A8}$$

$$0 \leq \tau^L_{ijk} \leq 1 \qquad i\in I, j = j_2, k\in K \tag{A9}$$

$$\tau^W_{ijk} \geq \tau^L_{ijk} \qquad i\in I, j\in J\backslash\{j_7, j_8\}, k\in K \tag{A10}$$

$$\tau^{R1}_{ijk} \leq \tau^{R2}_{ijk} \qquad i\in I, j\in J, k\in K \tag{A11}$$

$$\tau^{R2}_{ijk} \leq \tau^{R3}_{ijk} \qquad i\in I, j\in J, k\in K \tag{A12}$$

$$\tau^{R2}_{ijk} \leq 5 \qquad i\in I, j\in \{j_7, j_8\}, k\in K \tag{A13}$$

$$\tau^{R3}_{ijk} \leq 5 \qquad i\in I, j\in \{j_7, j_8\}, k\in K \tag{A14}$$

$$\tau^C_{ijk} \geq \tau^F_{ijk} \qquad i\in I, j\in J\backslash\{j_7, j_8\}, k\in K \tag{A15}$$

$$\tau^{S1}_{ijk} \leq \tau^{S2}_{ijk} \qquad i\in I, j\in J, k\in K \tag{A16}$$

$$\tau^{S2}_{ijk} \leq \tau^{S3}_{ijk} \qquad i\in I, j\in J, k\in K \tag{A17}$$

$$0 \leq \epsilon^*_{ijk} \leq 2.76 \qquad i\in I, j\in J, k\in K \tag{A18}$$

$$0 \leq \alpha^L_{ijk} \leq 1 \qquad i\in I, j\in J, k\in K \tag{A19}$$

$$0 \leq \alpha^W_{ijk} \leq 1 \qquad i\in I, j\in J, k\in K \tag{A20}$$

$$0 \leq \alpha^R_{ijk} \leq 1 \qquad i\in I, j\in J, k\in K \tag{A21}$$

$$0 \leq \xi^{R1}_{ijk} \leq 1 \qquad i\in I, j\in J, k\in K \tag{A22}$$

$$0 \leq \xi^{R2}_{ijk} \leq 1 \qquad i\in I, j\in J, k\in K \tag{A23}$$

$$0 \leq \xi^{R3}_{ijk} \leq 1 \qquad i\in I, j\in J, k\in K \tag{A24}$$

$$0 \leq \theta^F_{ijk} \leq \frac{1}{2} \qquad i\in I, j\in J, k\in K \tag{A25}$$

$$0 \leq \theta^C_{ijk} \leq \frac{1}{2} \qquad i\in I, j\in J, k\in K \tag{A26}$$

$$0 \leq \theta^{S1}_{ijk} \leq \frac{1}{10} \qquad i\in I, j\in J, k\in K \tag{A27}$$

$$0 \leq \theta^{S2}_{ijk} \leq \frac{1}{10} \qquad i\in I, j\in J, k\in K \tag{A28}$$

$$0 \leq \eta_{ijk} \leq \frac{1}{10} \qquad i\in I, j\in J, k\in K \tag{A29}$$

$$0 \leq \tau^L_{ijk} \leq 10 \qquad i\in I, j\in J, k\in K \tag{A30}$$

$$0 \leq \tau^W_{ijk} \leq 500 \qquad i\in I, j\in J, k\in K \tag{A31}$$

$$0 \leq \tau^{R1}_{ijk} \leq 10 \qquad i\in I, j\in J, k\in K \tag{A32}$$

$$0 \leq \tau^{R2}_{ijk} \leq 20 \qquad i\in I, j\in J, k\in K \tag{A33}$$

$$0 \le \tau^{R3}_{ijk} \le 50 \qquad i \in I, j \in J, k \in K \qquad \text{(A34)}$$
$$0 \le \tau^{F}_{ijk} \le 10 \qquad i \in I, j \in J, k \in K \qquad \text{(A35)}$$
$$0 \le \tau^{C}_{ijk} \le 50 \qquad i \in I, j \in J, k \in K \qquad \text{(A36)}$$
$$0 \le \tau^{S1}_{ijk} \le 100 \qquad i \in I, j \in J, k \in K \qquad \text{(A37)}$$
$$0 \le \tau^{S2}_{ijk} \le 250 \qquad i \in I, j \in J, k \in K \qquad \text{(A38)}$$
$$0 \le \tau^{S3}_{ijk} \le 500 \qquad i \in I, j \in J, k \in K \qquad \text{(A39)}$$
$$NPP_{ijk} = APAR_{ijk} \cdot \epsilon^{*}_{ijk} \cdot T_{\epsilon} \cdot W_{\epsilon} \qquad i \in I, j \in J, k \in K \qquad \text{(A40)}$$

The objective function (A1) minimizes the deviation of the state variables $\widehat{y}$ from the observed data y. The state variables represent estimates of C residence times and other aspects of the C cycle. Minimizing this deviation is intended to ensure that the residence time estimates are as accurate as possible. TECOR-R-QP is called a quadratic program because (A1) is a quadratic function of $\widehat{y}$.

Constraints (A2)–(A40) describe properties that any solution to TECO-R-QP must have. Each property models one aspect of the real-world C cycle in the eight terrestrial biomes considered here. Constraints (A2)–(A40) are derived from the TECO-R model of [1], which is based on expert judgment and a review of the C cycle literature. However, the parameter values in some of the constraints (A2)–(A40) may be open to the charge of arbitrariness. One aspect of the proposed work will be to conduct sensit Constraint (A2) states that the proportion of NPP allocated to leaves can be no greater than the proportion allocated to wood. Constraint (A3) states that no NPP can be allocated to wood in the grassland or cropland biomes because they lack woody biomass. Constraint (A4) requires that the proportions of NPP allocated to leaves, wood, and roots add up to one. Constraints (A5) and (A6) ensure that the proportion of NPP allocated to roots increases with root depth. Shallow roots (depth: 0-20 cm) are assigned less NPP than deeper roots (depth: 20-50 cm or 50-100 cm). Constraint (A7) requires that the proportions of NPP allocated to roots at all three depths sums to one. Constraint (A8) ensures that none of the C partitioning coefficient is assigned to coarse litter pool in the grassland biome because this biome lacks coarse litter. Constraint (A9) restricts the C residence time of leaves in deciduous broadleaf biome to be at most one year since trees in this biome lose their leaves each autumn. Constraint (A10) states that in all biomes other than grassland and cropland, which lack woody biomass, the residence time of C in wood must be at least as great as the C residence time in leaves. Constraints (A11) and (A12) state that the residence time of C in roots is at least as great in deep roots (depth: 20-50 cm or 50-100 cm) as in shallow roots (depth: 0-20 cm). Constraints (A13) and (A14) restrict the C residence time of deep roots to be no more than five years in the grassland and cropland biomes. Constraint (A15) states that in all biomes other than cropland and grassland, which lack coarse litter, the residence time of C in the coarse litter pool must be at least as long as the residence time in the fine litter pool. Constraints (A16) and (A17) state that C persists at least as long deep in the soil (depth: 20-50 cm or 50-100 cm) as in the topsoil (depth: 0-20 cm).

Constraint (A18) states that the maximum light use efficiency is a non-negative real number no greater than 2.76. Constraints (A19)–(A21) state that the proportion of NPP allocated to leaves, wood, and roots must each be between zero and one. Similarly, constraints (A22)–(A24) restrict the proportion of NPP allocated to roots at the three depths considered here (0-20 cm, 20-50 cm, and 50-100 cm) to be between zero and one. Constraints (A25) and (A26) state that no more than half of the C partitioning coefficient can be allocated to the coarse or fine litter pools in any biome.

Constraints (A27) and (A28) state that no more than ten percent of the C partitioning coefficient can be allocated to soil organic C at a depth of either 0-20 cm or 20-50 cm. Constraint (A29) states that no more than ten percent of the C in the coarse litter pool can be broken down mechanically into C in the fine litter pool.

Constraints (A30)–(A39) define the lower and upper bounds of C residence time in various terrestrial pools. The lower bound is zero in each case because the duration of C residence in each pool is a non-negative number. Constraints (A30) and (A31) define the C residence times of leaves and wood as no more than ten and five hundred years, respectively. Constraints (A32) and (A34) set upper bounds of ten years for the duration of C residence in roots less than 20 cm deep, twenty years for roots between 20 and 50 cm deep, and fifty years for roots between 50 and 100 cm deep. Constraints (A35) and (A36) indicate that C resides in fine litter for at most ten years and coarse litter for at most fifty years. Constraints (A37)–(A39) establish upper limits on the duration of C residence in soil at three depths: soil organic C resides for at most a century in the pool comprised of soil less than 20 cm deep, for up to 250 years in soil more than 20 cm deep and less than 50 cm deep, and for up to 500 years in soil between 50 and 100 cm deep. Finally, constraint (A40) defines NPP as the product of absorbed radiation, light use efficiency, and temperature and moisture scalars.

Constraints (A2)-(A40) do not incorporate temporal changes in the sizes of the C pools q_L, q_W, q_R, q_F, q_C, q_{S1}, q_{S2}, and q_{S3} (see Fig. 1 of the proposal). Zhou and Luo [1] formulate differential equations to describe changes in the C pool sizes over time. Future work will incorporate these differential equations into TECO-R-QP.

References

[1] T. Zhou and Y. Lou. Spatial patterns of ecosystem carbon residence time and NPP–driven carbon uptake in the conterminous United States. *Global Biogeochemical Cycles*, 22:GB3032, 2008.