Supplement of Hydrol. Earth Syst. Sci., 30, 6075–6093, 2026
https://doi.org/10.5194/hess-30-6075-2026-supplement

Hydrology and Earth System Sciences
Open Access

Supplement of

Incorporating spatial heterogeneity into evapotranspiration estimates for bioretention basins

Joshua S. Caplan et al.

Correspondence to: Joshua S. Caplan (jcaplan@temple.edu)

S1 Extended Methods

S1.1 Chamber ET flux measurements

ET measurements were made using the closed dynamic chamber approach, in which all vegetation and soil within a plot are temporarily enclosed in a fixed-volume to measure the accumulation of water vapor. Although more commonly used to measure carbon fluxes (e.g., Geoghegan et al., 2018; Norman et al. 1997), the method has been successfully applied to water flux measurement as well (Garcia et al., 2008; Hamel et al., 2015; Luo, 2015). Custom-made collars were set into the soil at the beginning of the field season, and these collars demarcated plot boundaries. The upper faces of collars were square (52 × 52 cm or 92 × 92 cm) and approximately parallel to the basin's bottom (i.e., not the basins' sloped sides). Small collars were installed at the eight plots with minimal or short-stature vegetation approximately a week before the first measurement campaign while the large collars were installed at the three plots with taller vegetation between the first and second measurement dates. Collar walls were transparent and typically extended 2-6 cm into the soil. At the start of each measurement day, any gaps that had developed between collar walls and the ground surface were filled with soil to prevent water vapor from entering or escaping during ET measurement. Vegetation that had grown in minimally-vegetated plots was also clipped to within ~0.5 cm of the soil surface at the beginning of each measurement day.

We measured ET at each plot approximately five times per measurement day, typically between 800 and 1900 h at approximately 90-minute intervals. For each measurement, first we placed a transparent chamber onto the above-mentioned collar (Fig. S1). Chambers were 52 × 52 × 52 cm or 92 × 92 × 62.5 cm; closed-cell foam was attached to the rims of open faces (i.e., at the top and bottom of chambers) so they would form airtight seals. We clamped the collar and transparent chamber together and placed at least one small, battery-powered fan within the chamber to continually mix air. We measured water vapor in the chamber once per second using an ultra-portable greenhouse gas analyzer (UGGA; Los Gatos Research, San Jose, CA, USA) that was connected to the chamber using a pair of flexible tubes (4.3 m long). Data were recorded as water vapor mole fraction (specifically μmol mol^{-1}, or ppm) for at least 2 min when using small chambers and at least 3 min when using large chambers.

Calculating rates of ET required isolating the linear portions of time series recorded while chambers were closed and performing linear regression analysis with the extracted data. The resulting rates are proportional to ET but reflect changes in water vapor concentration (ET_{ppm}) rather than the actual amount of water vapor. Conversion to the amount of water vapor per time (ET_{mmol}) was made by multiplying regression slopes (in ppm s^{-1}) by the total amount of air in the chamber, collar, and UGGA (n, in mol), with a factor of 10^{-3} accounting for the conversion from ppm to mole fraction and from mol to mmol:

$$ET_{mmol} = ET_{ppm} \times n \times 10^{-3}.$$

The total amount of air in the system was determined from the ideal gas law:

$$n = \frac{P \times V}{R \times T},$$

where P is atmospheric pressure (assumed to be 1 atm), V is the total volume of the enclosed air space (in L), R is the universal gas constant (0.082057 L atm K^{-1} mol^{-1}), and T is the air's temperature (in K). Finally, ET_{mmol} values were normalized to the sampled plot area (A, in m^2), thus yielding fluxes in units of mmol H_2O m^{-2} s^{-1}, and multiplied by a correction factor that accounts for water vapor adsorbing to interior surfaces of tubing (1.187). The correction factor was interpolated from values in Hamel et al. (2015) given the length of tubing we used (4.3 m):

$$ET = \frac{ET_{mmol} \times 1.187}{A}.$$

S1.2 Basin spatial characterization

We characterized the basin's topography and spatial heterogeneity in vegetation height using a combination of terrestrial LiDAR scanning and drone-based photogrammetry. We first used a Trimble TX5 3D terrestrial laser scanner (Trimble Inc., Sunnyvale, CA, USA) to generate a digital elevation model (DEM) of SMP A. The laser scanner was mounted on a tripod and generated point clouds extending up to 60 m from the laser scanner. We performed 13 scans from different positions within the basin to ensure line of sight with the vast majority of the ground surface (i.e., to minimize "shadows" from vegetation and other structures). At least three targets (glass sphere or paper checkerboard) were included in every scan; these used to georeference each XYZ point cloud scan to the others. We further recorded the coordinates of 12 of the target locations using a Trimble GeoXH GPS unit (Trimble Inc., Sunnyvale, CA, USA) and used their coordinates to georeference the point cloud in the universal transverse mercator (UTM) coordinate system. The GPS unit had a 10-cm accuracy after post-processing, though the laser scanner had an accuracy <1 cm. We merged, georeferenced, and trimmed areas outside of the basin using Trimble RealWorks (Trimble Inc., Sunnyvale, CA, USA). We further refined the dataset by removing points associated with vegetation and manmade structures (e.g., a metal cage around the basin's outlet) using the *lasground_new* algorithm from LAStools (rapidlasso GmbH, Gilching, Germany). We then used an adaptive triangular network algorithm from LAStools to extract the ground surface (Axelsson, 2000). Small areas where the laser could not achieve line of sight were interpolated. Finally, we converted the dataset into a 10-cm gridded raster using Quick Terrain Modeler (Applied Imagery, Chevy Chase, MD, USA).

We generated a digital surface model (DSM) from drone-based photographs using a structure-from-motion approach (following Alonzo et al., 2018). We flew a Phantom 4 Pro drone (DJI Technology, Shenzhen, China) in a manual grid pattern over the basin with an off-nadir camera view to establish convergent view geometry (Alonzo 2020). This process yielded 286 photos with at least 80% front/back overlap and sidelap. The photos were processed to generate a point cloud using Pix4Dmapper 4.6.4 (Pix4D S.A., Prilly, Switzerland). This included manual georeferencing to a set of 5 visible tie points, allowing for co-registration of the final DSM with the DEM described above. From the point cloud, we created and exported a high-resolution orthomosaic and a DSM. Vegetation cover and height were established by subtracting the DEM model results from the structure-from-motion DSM. The DEM and DSM data also enabled us to manually create a set of polygons that demarcated topographic positions within the basin.

S1.3 Soil-water stress analysis

Because conventional ET estimates can optionally be adjusted to account for soil-water limitation with a water stress coefficient, we evaluated how doing so would alter agreement with empirical ET estimates. We did this for all six conventional models, even though K_S is not typically applied to some of them, to determine if it substantially improved agreement.

Water-stress coefficients (K_S) were calculated using an FAO-56-style approach (Allen et al., 1998), which essentially involved determining the degree of root-zone water depletion relative to field capacity and expressing this in terms of K_S. To do this, we first calculated mean root-zone volumetric water content (θ) from profile-based measurements at 10 and 30 cm depths (details below). We then determined VWC at field capacity (−33 kPa) and permanent wilting point (−1500 kPa) from a soil water retention curve (a.k.a. soil water characteristic curve) for the basin's media (Shakya et al., 2023). This allowed us to compute three quantities: total available water (TAW), readily available water (RAW), and root-zone depletion (D_r; constrained to the range 0 to TAW), as follows:

$$TAW = 1000(\theta_{FC} - \theta_{WP})Z_r,$$

$$D_r = 1000(\theta_{FC} - \theta)Z_r, \text{ and}$$

$$RAW = p \times TAW,$$

where θ_{FC} and θ_{WP} are VWC at field capacity and the permanent wilting point, respectively; Z_r is effective rooting depth; θ is root-zone VWC at the daily scale; and p is the depletion fraction. Based on the water retention curve, θ_{FC}= 0.444 and θ_{WP}= 0.093 $m^3 m^{-3}$. Z_r was not known and could vary throughout the basin, so we evaluated three depths spanning the plausible range of values (20, 30, and 40 cm). This range was based on observations of rooting depth in a basin adjacent to I-95 that was installed in the same year and contained the same soil media as BASIN A. For Z_r = 20 cm, θ was considered equivalent to values recorded by the sensor at 10 cm. For Z_r = 30 and 40 cm, θ was calculated as a depth-weighted mean of measurements from sensors at 10 and 30 cm, with the 10-cm sensor representing the upper 20 cm of the root zone and the 30-cm sensor representing the remaining depth. p was based on a reference value of 0.50 and adjusted daily for evaporative demand following the FAO-56 relationship p = 0.50 + 0.04(5 - ET), using daily PM-FAO estimates as the measure of evaporative demand (ET). Resulting daily values of p ranged from 0.455 to 0.674.

Daily K_S values were then calculated following the FAO-56 approach (Allen et al., 1998):

$$K_S = \begin{cases} 1, & D_r \leq RAW \\ \frac{TAW - D_r}{TAW - RAW}, & D_r > RAW \end{cases}.$$

The resulting K_S values were multiplied by the corresponding daily ET estimates from the six conventional models. We then compared unadjusted and water-stress-adjusted ET estimates with empirical ET values by examining their mean difference (or bias) and CCC. Because soil-moisture data were unavailable for three dates that did have empirical ET estimates, this analysis yielded data for 134 days. Fig. S4 presents results assuming a 30-cm effective rooting depth, though results for all three depths were used to evaluate sensitivity to this assumption.

S2 Supplemental Table

Table S1. Conventional ET model specifications. All models were implemented using the R package *Evapotranspiration* (Guo et al., 2016).

Name	Type	Meteorological Inputs	Other Model Inputs	Physical Constants
Priestley-Taylor [PT] (Priestley & Taylor, 1972)	Potential ET	RH_{max} RH_{min} R_s T_{max} T_{min}	α_{PT} Elev G Lat	G_{sc} λ σ
Matt-Shuttleworth [MS] (Shuttleworth & Wallace, 2009)	Reference crop ET	RH_{max} RH_{min} R_s T_{max} T_{min} U_2 or U_z	α CH Elev Lat r_s z	C_a G_{sc} λ ρ_a σ
Hargreaves-Samani [HS] (Hargreaves & Samani, 1985)	Reference crop ET	T_{max} T_{min}	Elev Lat	G_{sc} λ
Penman-Monteith FAO [PM-FAO] (Allen et al., 1998)	Reference crop ET	RH_{max} RH_{min} R_s T_{max} T_{min} U_2 or U_z	α CH Elev G Lat r_s z	G_{sc} λ σ
Penman-Monteith ASCE [PM-ASCE] (Allen, 2005)	Reference crop ET	RH_{max} RH_{min} R_s T_{max} T_{min} U_2 or U_z	α CH Elev G Lat r_s z	G_{sc} λ σ
Granger-Gray [GG] (Granger & Gray, 1989)	Actual ET	RH_{max} RH_{min} R_s T_{max} T_{min} U_2 or U_z	α Elev G Lat z	G_{sc} λ σ

Abbreviations
α = albedo of evaporative surface (0.23)
α_{PT} = Priestley-Taylor coefficient (1.26)
C_a = specific heat of air (0.001013 MJ kg^{-1} °C^{-1})
CH = crop height (0.12 m for MS and PM-FAO; 0.50 m for PM-ASCE)
Elev = ground elevation above mean sea level
G = soil heat flux
G_{sc} = solar constant (0.0820 MJ m^{-2} min^{-1})
λ = latent heat of vaporization (2.45 MJ kg^{-1})
Lat = latitude
RH_{max} = maximum relative humidity
RH_{min} = minimum relative humidity
ρ_a = mean air density (1.20 kg m^{-3})
R_s = incoming solar radiation
r_s = surface resistance (70 s m^{-1} for MS and PM-FAO; 45 s m^{-1} for PM-ASCE)
σ = Stefan-Boltzmann constant (4.903×10^{-9} MJ K^{-4} m^{-2} d^{-1})
T_{max} = maximum temperature
T_{min} = minimum temperature
U_2 = wind speed at 2 m
U_z = wind speed at height z
z = height of wind instrument

S3 Supplemental Figures

Figure S1. Setup used for chamber flux measurements, with the larger collar (92 × 92 cm upper surface) and chamber (92 × 92 × 62.5 cm) shown. The gas analyzer (yellow) and its battery (grey) are in the foreground.

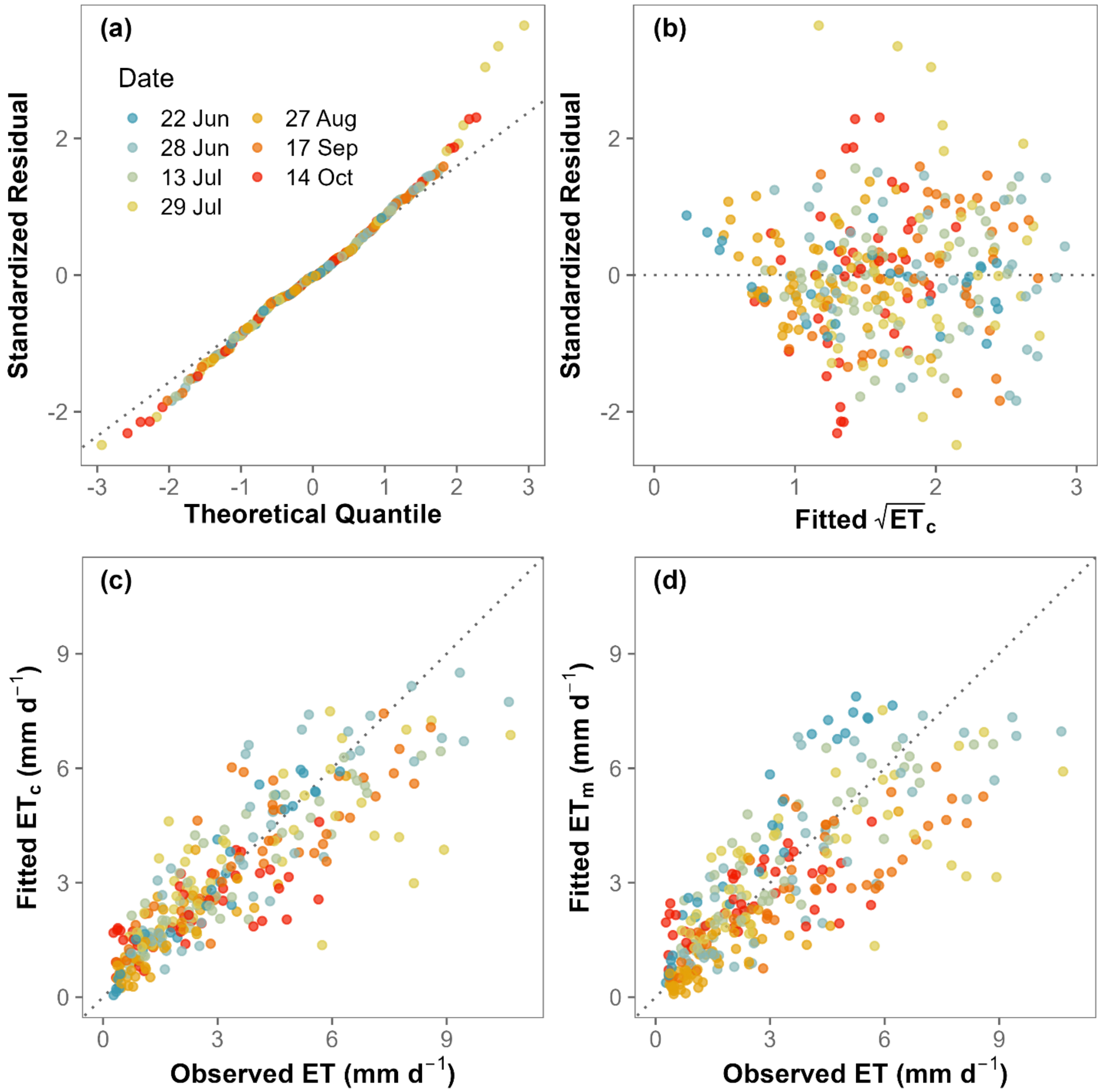

Figure S2. Model diagnostic and performance plots, with points colored by measurement date. **(a)** Normal quantile-quantile plot of standardized residuals; **(b)** standardized residuals versus conditional fitted values; **(c)** observed versus conditional fitted ET (ET_c), which incorporates fixed and random effects; and **(d)** observed versus marginal fitted ET (ET_m), which incorporates fixed effects only.

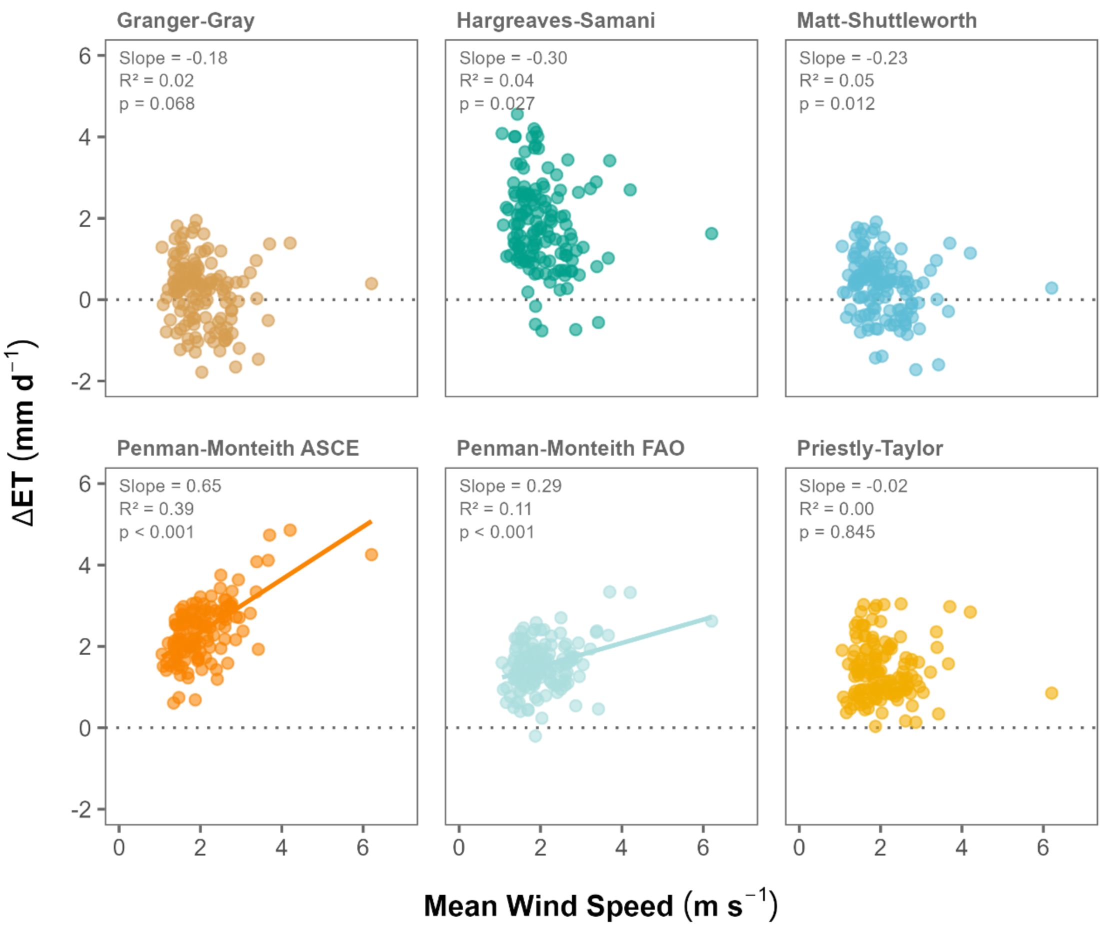

Figure S3. Difference in daily ET estimates from conventional vs. empirical models as a function of daily mean wind speed (n = 137). Positive ΔET values indicate that the conventional model estimated greater ET than the empirical model. Regression lines are shown for PM models, for which wind speed explained a non-negligible portion of the variation in ΔET.

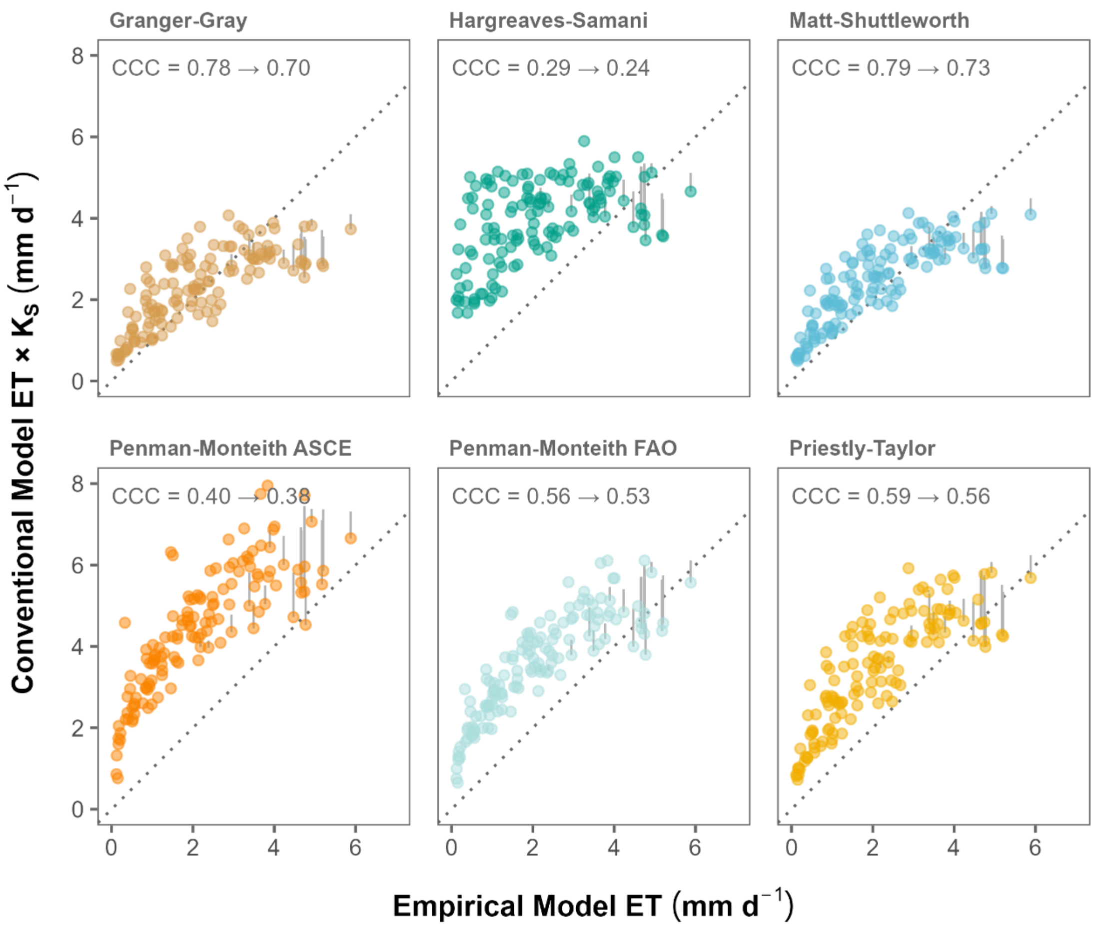

Figure S4. Daily ET calculated from our basin-scale, empirically-based model vs. ET from six conventional models after adjustment using the FAO-56 water-stress coefficient (K_S), assuming an effective rooting depth of 30 cm. Grey lines depict changes in conventional model estimates when including K_S. Dotted lines depict 1:1 relationships. CCC = concordance correlation coefficient; arrows indicate CCCs before vs. after adjusting for water stress. Unadjusted CCC values differ from those in Fig. 7 because soil moisture data were not available for three dates (n = 134).

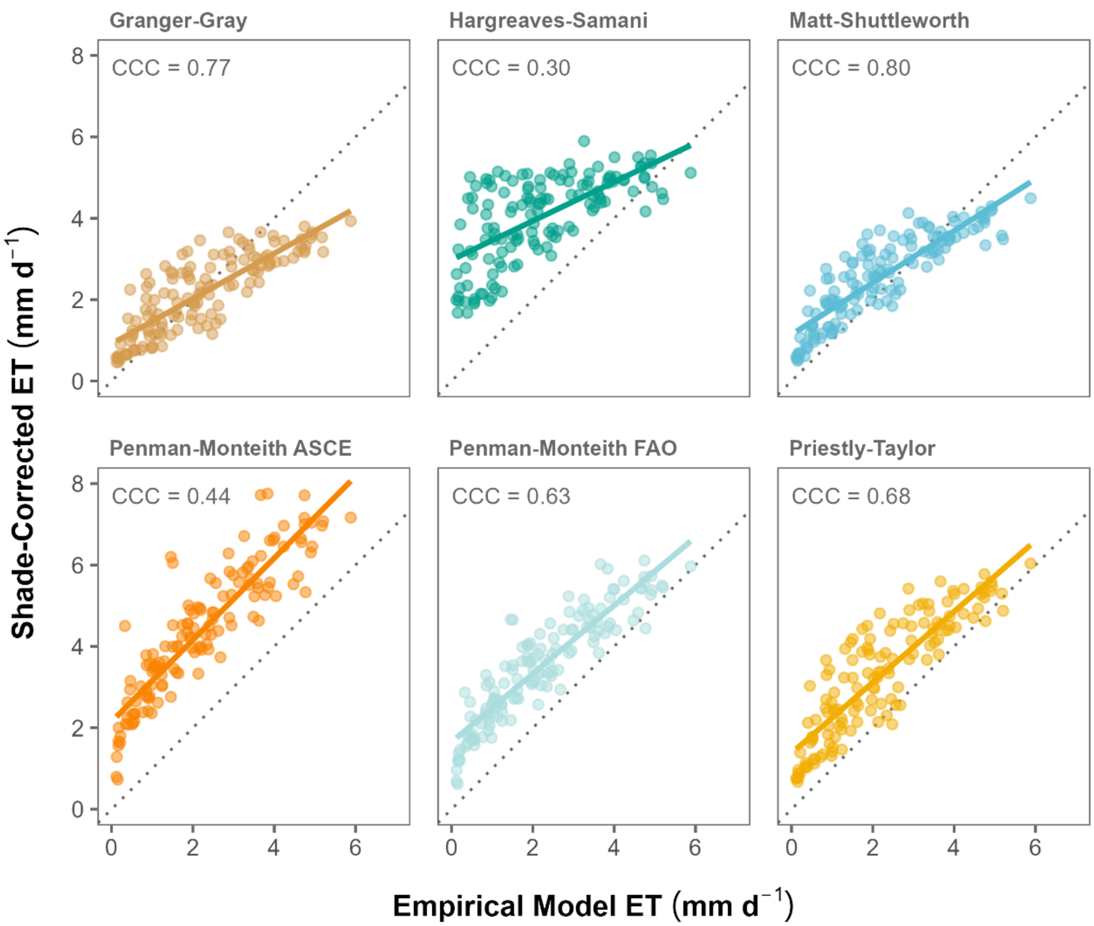

Figure S5. Daily ET calculated from our basin-scale, empirically-based model vs. ET from six conventional models whose solar radiation input data were modified to account for shade from nearby structures. Dotted lines depict 1:1 relationships. CCC = concordance correlation coefficient.

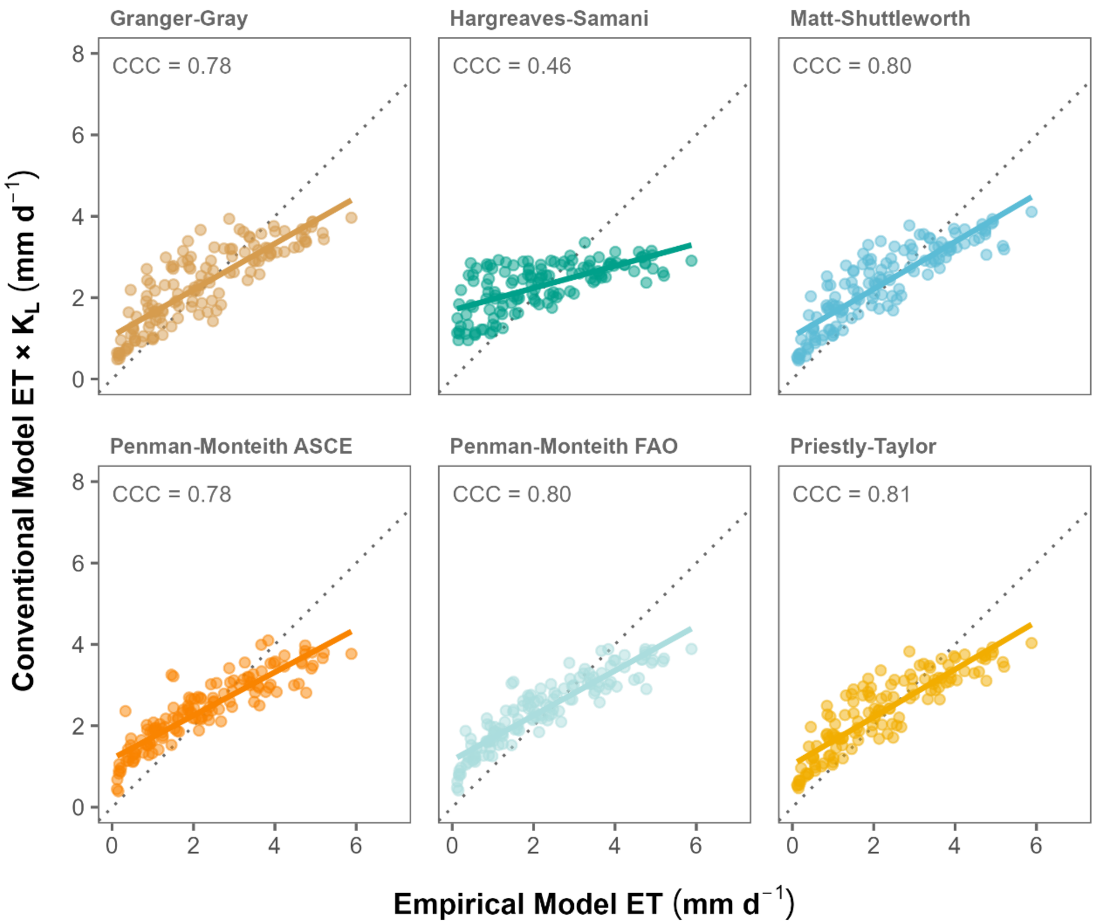

Figure S6. Daily ET calculated from our basin-scale, empirically-based model vs. ET from six conventional models that were modified with a landscape coefficient (K_L). Dotted lines depict 1:1 relationships. CCC = concordance correlation coefficient.

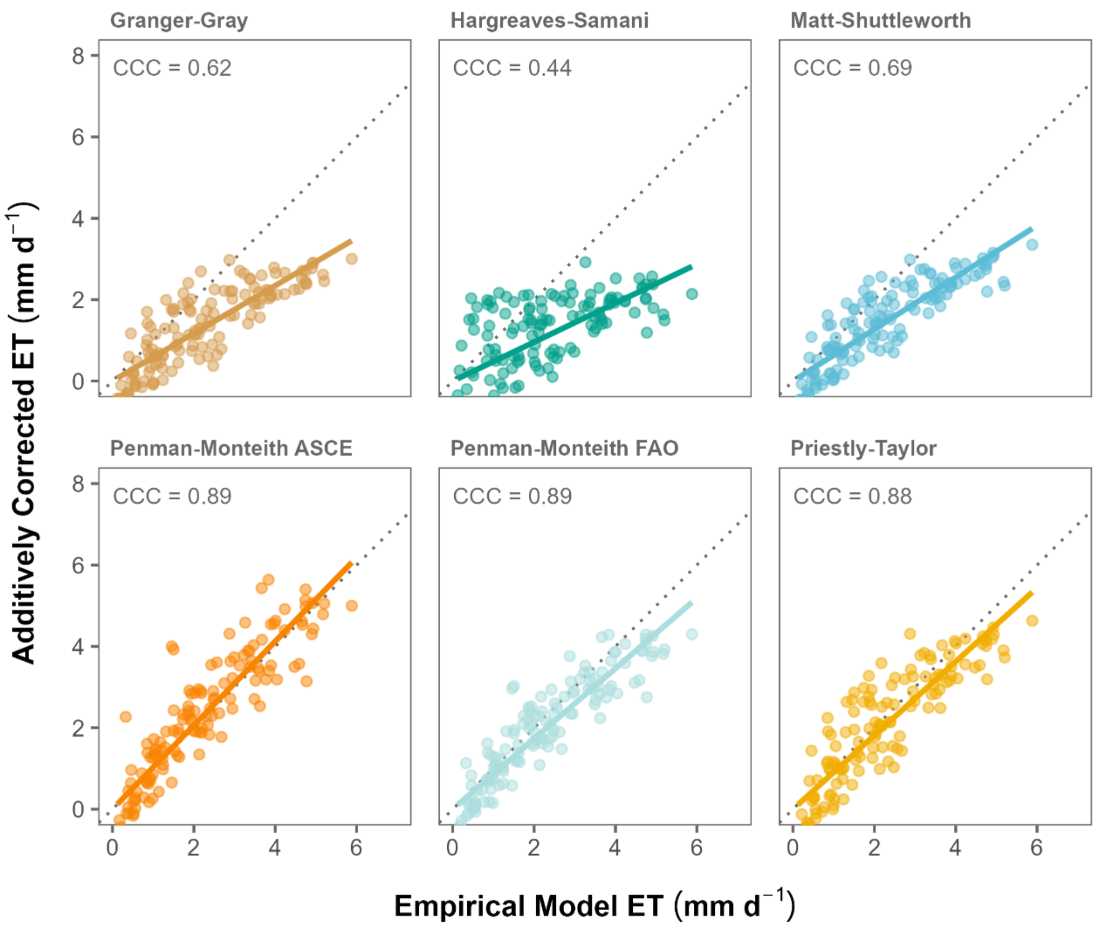

Figure S7. Daily ET calculated from our basin-scale, empirically-based model vs. ET from six conventional models modified with an additive correction factor. Dotted lines depict 1:1 relationships. CCC = concordance correlation coefficient.

Supplemental References

Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., and Süsstrunk, S.: SLIC superpixels compared to state-of-the-art superpixel methods, IEEE Trans. Pattern Anal. Mach. Intell., 34, 2274–2282, https://doi.org/10.1109/TPAMI.2012.120, 2012.

Allen, R. G., Pereira, L. S., Raes, D., and Smith, M.: Crop evapotranspiration – Guidelines for computing crop water requirements, FAO Irrigation and Drainage Paper No. 56, FAO, Rome, 1998.

Allen, R. G.: The ASCE standardized reference evapotranspiration equation, ASCE/EWRI Task Committee Report, American Society of Civil Engineers, https://doi.org/10.1061/9780784408056, 2005.

Alonzo, M., Andersen, H. E., Morton, D. C., and Cook, B. D.: Quantifying boreal forest structure and composition using UAV structure from motion, Forests, 9, 119, https://doi.org/10.3390/f9030119, 2018.

Alonzo, M., Dial, R. J., Schulz, B. K., Andersen, H.-E., Lewis-Clark, E., Cook, B. D., and Morton, D. C.: Mapping tall shrub biomass in Alaska at landscape scale using structure-from-motion photogrammetry and lidar, Remote Sens. Environ., 245, 111841, https://doi.org/10.1016/j.rse.2020.111841, 2020.

Axelsson, P.: DEM generation from laser scanner data using adaptive TIN models, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., 33, 110–117, 2000.

Garcia, A., Johnson, M. J., Andraski, B. J., Halford, K. J., and Mayers, C. J.: Portable chamber measurements of evapotranspiration at the Amargosa Desert Research Site near Beatty, Nye County, Nevada, 2003–06, U.S. Geological Survey Scientific Investigations Report 2008–5135, Reston, VA, 2008.

Geoghegan, E. K., Caplan, J. S., Leech, F. N., Weber, P. E., Bauer, C. E., and Mozdzer, T. J.: Nitrogen enrichment alters carbon fluxes in a New England salt marsh, Ecosyst. Health Sustain., 4, 277–287, https://doi.org/10.1080/20964129.2018.1532772, 2018.

Granger, R. J. and Gray, D. M.: Evaporation from natural nonsaturated surfaces, J. Hydrol., 111, 21–29, https://doi.org/10.1016/0022-1694(89)90249-7, 1989.

Guo, D., Westra, S., and Maier, H. R.: An R package for modelling actual, potential and reference evapotranspiration, Environ. Model. Softw., 78, 216–224, https://doi.org/10.1016/j.envsoft.2015.12.019, 2016.

Hamel, P., McHugh, I., Coutts, A., Daly, E., Beringer, J., and Fletcher, T. D.: Automated chamber system to measure field evapotranspiration rates, J. Hydrol. Eng., 20, 04014037, https://doi.org/10.1061/(ASCE)HE.1943-5584.0001006, 2015.

Hargreaves, G. H. and Samani, Z. A.: Reference crop evapotranspiration from temperature, Appl. Eng. Agric., 1, 96–99, https://doi.org/10.13031/2013.26773, 1985.

Luo, C., Wang, Z., Sauer, T. J., Helmers, M. J., and Horton, R.: Portable canopy chamber measurements of evapotranspiration in corn, soybean, and reconstructed prairie, Agric. Water Manag., 198, 1–9, https://doi.org/10.1016/j.agwat.2017.11.024, 2018.

Norman, J. M., Kucharik, C. J., Gower, S. T., Baldocchi, D. D., Crill, P. M., Rayment, M., Savage, K., and Striegl, R. G.: A comparison of six methods for measuring soil-surface carbon dioxide fluxes, J. Geophys. Res.-Atmos., 102, 28771–28777, https://doi.org/10.1029/97JD01440, 1997.

Priestley, C. H. B. and Taylor, R. J.: On the assessment of surface heat flux and evaporation using large-scale parameters, Mon. Weather Rev., 100, 81–92, https://doi.org/10.1175/1520-0493(1972)100<0081:OTAOSH>2.3.CO;2, 1972.

Shuttleworth, W. J. and Wallace, J. S.: Evaporation from sparse crops – An energy combination theory, Q. J. R. Meteorol. Soc., 111, 839–855, https://doi.org/10.1002/qj.49711146910, 1985.