

P-LSHv2: a multi-decadal global daily evapotranspiration dataset enhanced with explicit soil moisture constraints

Jin Feng^{1,3,5}, Ke Zhang^{1,2,3,4*}, Lijun Chao^{1,3}, Huijie Zhan¹, Yunping Li^{1,3,6}

¹ National Key Laboratory of Water Disaster Prevention, and College of Hydrology and Water Resources, Hohai University, Nanjing, Jiangsu, 210024, China

² Yangtze Institute for Conservation and Development, Hohai University, Nanjing, Jiangsu, 210024, China

³ China Meteorological Administration Hydro-Meteorology Key Laboratory, Hohai University, Nanjing, Jiangsu, 210024, China

⁴ Key Laboratory of Water Big Data Technology of Ministry of Water Resources, Hohai University, Nanjing, Jiangsu, 210024, China

⁵ Hydrology Management Center of Zhejiang Province, Hangzhou, Zhejiang, 310009, China

⁶ College of Water Resources, North China University of Water Resources and Electric Power, Zhengzhou, 450046, China

Correspondence to: Ke Zhang (kzhang@hhu.edu.cn)

Abstract. Accurately quantifying the impact of soil water availability on evapotranspiration (ET) is crucial for improving ET retrieval accuracy. However, most global satellite-derived ET datasets do not explicitly incorporate soil moisture constraints, leading to significant uncertainties, particularly in water-limited regions. In this study, we propose an enhanced soil moisture constraint scheme that effectively captures soil moisture's influence on vegetation transpiration and soil evaporation using a quantile-based approach. Unlike previous methods, this scheme relies solely on soil moisture data, reducing uncertainties associated with heterogeneous soil hydraulic properties. We integrated this approach into the process-based land surface ET/heat fluxes algorithm (P-LSH, or P-LSHv1), developing an improved version, P-LSHv2. Using observations from 106 global flux towers, we calibrated biome- and climate-specific parameters and quantified moisture constraints across diverse climates and land cover types. P-LSHv2 achieves notable improvements in ET estimation, with a reduced Root Mean Square Error (RMSE) of 0.67 mm d⁻¹ and an increased Pearson correlation coefficient (R) of 0.81, indicating strong agreement with flux tower observations. As a result of these improvements, P-LSHv2 outperforming its predecessor, P-LSHv1, particularly in arid regions. Comparative analyses show that P-LSHv2 surpasses the Penman-Monteith-Leuning model and the Global Land Evaporation Amsterdam Model in capturing soil moisture anomalies' effects on ET, enhancing global ET accuracy. Employing the P-LSHv2 algorithm, we have produced a long-term global daily ET dataset spanning 1982–2023, providing a valuable resource for research on terrestrial water and energy cycles and climate change. The dataset is freely available at <https://doi.org/10.11888/Terre.tpdc.301969> (Feng and Zhang, 2025).

1 Introduction

Evapotranspiration (ET) is considered

control on stomatal conductance, thereby influencing transpiration dynamics. Importantly, root tips are capable of sensing subtle gradients in water potential (as low as 0.5 MPa mm^{-1}), exhibiting behaviors such as hydrotropism to actively seek water in heterogeneous soil profiles (Dexter, 1986; Gregory, 2006). Such sensory responses provide a physiological basis for root foraging behavior, which is particularly important for sustaining transpiration under drought conditions. Therefore, soil water content not only serves as the direct water pool for vegetation but also provides the essential hydraulic foundation that supports plant physiological functions. It ultimately regulates both the capacity and the demand side of transpiration through its dual role in water supply and physiological control. In previous remote sensing ET algorithms, the vapor pressure deficit (VPD) was generally adopted to reflect the associated moisture constraint on canopy conductance (Mu et al., 2011; Zhang et al., 2009; Zhang et al., 2010; Zhang et al., 2019). However, several studies have shown that water availability is the primary limiting factor for ET in arid areas (Yang et al., 2019; Zhang et al., 2015), namely the so-called water-limited regions, where soil moisture constraints on canopy conductance increase with climate dryness and drought severity (Han et al., 2020; Novick et al., 2016). In arid and semi-arid regions, even under high atmospheric demand (i.e., high VPD), the actual ET is often constrained by soil water availability and root uptake capacity. As the root–soil interface becomes hydraulically disconnected under drought, transpiration may decline despite strong evaporative demand. Although vapor pressure deficit and soil moisture are generally connected through land-atmosphere interactions (Liu et al., 2020a; Zhou et al., 2019), their anomalies may be temporally lagged (Zhang et al., 2022b), indicating a partial decoupling. Therefore, relying solely on VPD can misrepresent transpiration dynamics, especially under conditions of prolonged soil dryness or climate extremes. Incorporating explicit soil moisture constraints is essential for improving ET estimation at finer temporal scales (Fu et al., 2022a), and this need is expected to intensify under global warming scenarios (Liu et al., 2020b), where soil–atmosphere coupling may become even more unstable.

One of the main challenges of incorporating soil moisture constraints into remote sensing ET algorithms is the development of a robust and globally applicable moisture constraint scheme for quantifying water availability. Although soil moisture constraints have been applied to estimate ET at local (Xu et al., 2018; Zhu et al., 2014), regional, and global scales (Brust et al., 2021; Purdy et al., 2018), long-term available datasets are limited by the lack of universal constraint schemes and soil moisture datasets. Furthermore, existing soil moisture constraint schemes generally rely on soil hydraulic properties as thresholds (Martens et al., 2017; Zheng et al., 2022). However, global soil data present significant uncertainties, which substantially affect ET estimation (Dennis and Berbery, 2021). On the one hand, gridded datasets derived from field investigation are prone to great uncertainties, particularly in high-altitude and other harsh regions (Dai et al., 2019). On the other hand, the pedotransfer functions for generating gridded soil properties still lack proper extrapolation methods from local to global scales (Ma et al., 2021; Van Looy et al., 2017).

In response to these challenges, we integrated a new soil moisture constraint scheme into the previously developed Process-based Land Surface evapotranspiration/Heat fluxes algorithm (P-LSH, or P-LSHv1) (Zhang et al., 2010; Zhang et al., 2015), creating the second version (P-LSHv2). The objectives of this study are to (1) develop a globally universal soil moisture constraint scheme, calibrate key parameters, and quantify moisture levels across diverse vegetation types and climate zones,

(2) upgrade ET algorithm by incorporating soil moisture constraints and evaluate its robustness from sites to basins, and (3) generate a long-term, reliable global daily ET dataset spanning from 1982 to 2023.

2 Methodology

2.1 Baseline description of the P-LSH algorithm

105 The P-LSH algorithm separately estimates three components of the ground surface evapotranspiration: vegetation transpiration ($\lambda E_{\text{Canopy}}$: W m^{-2}), soil evaporation (λE_{Soil} : W m^{-2}), and open water evaporation (λE_{Water} : W m^{-2}). These components are calculated based on the pixel's land cover type. In vegetation pixels, the ET is partitioned into vegetation transpiration and soil evaporation by partitioning available energy according to the fractional vegetation cover. In barren and open water pixels, ET consists solely of soil evaporation and open water evaporation, respectively.

110 The Penman-Monteith equation is used to calculate vegetation transpiration:

$$\lambda E_{\text{Canopy}} = \frac{\Delta A_c + \rho C_p \text{VPD} g_{ac}}{\Delta + \gamma(1 + g_{ac}/g_c)}, \quad (1)$$

where λ (J kg^{-1}) is the latent heat of vaporization, Δ (Pa K^{-1}) is the slope of the saturated water vapor pressure curve as a function of temperature; A_c (W m^{-2}) is the available energy allocated to the canopy; ρ (kg m^{-3}) is air density; C_p ($\text{J kg}^{-1} \text{K}^{-1}$) is the specific heat capacity of air; VPD (Pa) is vapor pressure deficit; g_{ac} (m s^{-1}) is the aerodynamic conductance of the canopy;

115 γ (-) is the psychrometric constant; g_c (m s^{-1}) is the canopy conductance calculated using the Jarvis-Stewart-type model:

$$g_c = g_0(\text{NDVI}) \times m(T_{\text{day}}) \times m(\text{VPD}) \times m(C_{\text{CO}_2}), \quad (2)$$

$$g_0(\text{NDVI}) = 1/[b_1 + b_2 \times \exp(-b_3 \times \text{NDVI})] - 1/(b_1 + b_2), \quad (3)$$

where g_0 (m s^{-1}) is the maximum value of g_c based on NDVI; $m(T_{\text{day}})$, $m(\text{VPD})$, and $m(C_{\text{CO}_2})$ are stress factors associated with daylight temperature T_{day} ($^{\circ}\text{C}$), VPD (Pa), and CO_2 concentration (ppm), respectively; b_1 , <

where $f(-)$ is the moisture constraint estimated by the complementary relationship hypothesis here; RH (-) is the relative humidity; k (Pa) is a parameter to fit the complementary relationship and reflect the sensitivity of VPD, we regard it as a parameter for different land cover types and climates. The λE_{POT} ($W m^{-2}$) is the potential soil evaporation; A_s ($W m^{-2}$) is the available energy component allocated to the soil; g_{as} ($m s^{-1}$) is the aerodynamic conductance of the soil surface and is defined as the sum of the conductance associated with convective heat transfer (g_{ch} : $m s^{-1}$) and the conductance to radiative heat transfer (g_{rh} : $m s^{-1}$). The g_{ch} term is expressed in its resistance form r_c , a biome-and-climate-specific parameter, while the g_{rh} term is calculated by daylight temperature. The g_{totc} ($m s^{-1}$) is the corrected value of total aerodynamic conductance g_{tot} ($m s^{-1}$) and the correction is based on actual air temperature and pressure. In this study, the g_{tot} term is also expressed in the form of total aerodynamic resistance r_{tot} , which is regarded as a parameter determined by land cover types and climates.

The open water evaporation is calculated using a modified Penman equation that considers the effects of surface wind speed on aerodynamic conductance:

$$\lambda E_{Water} = \frac{\Delta A + \rho C_p VPD g_{aw}}{\Delta + \gamma}, \quad (7)$$

where A ($W m^{-2}$) is the available energy for open water; g_{aw} ($m s^{-1}$) is the aerodynamic conductance of the open water estimated by wind speed:

$$g_{aw} = \frac{1 + 0.536 U_2}{4.72 [\ln(z_m/z_0)]^2}, \quad (8)$$

where U_2 ($m s^{-1}$) is the wind speed at 2 m height converted from measurement height z_m (m) using vertical wind speed function; z_0 (m) is the aerodynamic roughness of the water surface and set to 0.00137. Further details of the P-LSH algorithm are available in Zhang et al. (2010) and Feng et al. (2022).

2.2 Improvements on the P-LSH algorithm

In this study, a series of improvements are implemented to evolve P-LSHv1 into the P-LSHv2 algorithm, incorporating soil moisture constraints on both vegetation transpiration and soil evaporation, as well as a Bayesian and Sobol' uncertainty analysis framework. These improvements allow for a more explicit quantification of water supply limitations on ET and offer a structured approach to parameter uncertainty across diverse land covers and climates.

(1) Soil moisture constraint scheme for vegetation transpiration

Previous studies (Liu et al., 2020b) have revealed that atmospheric moisture and soil moisture are generally decoupled in arid areas, which introduces significant uncertainty in quantifying soil water supply on ET. In this study, vegetation transpiration response to soil moisture is represented by the typical Jarvis-Stewart model (Jarvis, 1976; Stewart, 1988), in which a revised soil moisture constraint scheme is applied to constrain canopy conductance:

$$g_c =$$

