



OPEN

DATA DESCRIPTOR

# Ensemble of CMIP6 derived reference and potential evapotranspiration with radiative and advective components

Nels Bjarke<sup>1</sup>✉, Joseph Barsugli<sup>2,3</sup> & Ben Livneh<sup>1,2</sup>

Assessing changes in future aridity requires an understanding of variations in the atmospheric demand for water. Such assessments are often driven by estimations of potential evapotranspiration ( $ET_p$ ) and/or reference evapotranspiration ( $ET_0$ ), yet no comprehensive and validated estimate of these climate metrics exists to date from the Coupled Model Intercomparison Project 6 (CMIP6). Here we describe the development and validation of a published dataset of global monthly estimates of the Penman-Monteith derived  $ET_0$ , its advective and radiation components, Priestley-Taylor derived  $ET_p$ , and vapor pressure deficit from 16 CMIP6 projections and four emissions scenarios. Historical validation of the ensemble of CMIP6 evaporative demand shows general agreement with observationally-derived baselines of  $ET_0$  and  $ET_p$  from the Climate Research Unit (CRU) and ERA5-Land reanalysis products, with GCM biases driven primarily by regional differences in modeled humidity and advective contributions to  $ET_0$ . Overall, evaporative demand is projected to increase across all emissions scenarios, with the largest increases over polar regions, and with a larger contribution from advection particularly for regions with higher baseline  $ET_0$ .

## Background & Summary

The evaporative demand of surface water to the atmosphere has been shown to be an essential driver and explanatory variable for numerous evaluations of changes in surface freshwater availability, agricultural water demand and drought risk<sup>1–9</sup>. Despite its ubiquitous use in analyses of changes in regional and global hydroclimates, there is a notable absence of an organized dataset of estimates of evaporative demand from the latest generation of Coupled Model Intercomparison Project 6 (CMIP6) General Circulation Models (GCMs). To address this need, we have developed monthly estimates of evaporative demand across a multi-model ensemble of CMIP6 GCM simulations. These 16 GCMs are hereafter referred to as the CMIP6-MME.

Debate over the utility of the differing methods available for estimating evaporative demand<sup>5,10–12</sup> and limitations in the availability of certain hydroclimate variables required for those estimations can lead to a large range of differing methodologies applied for the estimations of  $ET_p$ <sup>13</sup>,  $ET_0$ <sup>14–16</sup>, and VPD<sup>16,17</sup>. The purpose of this dataset<sup>18</sup> is to serve as a tool for future hydroclimate analyses and to promote methodological consistency across studies that utilize future estimations of evaporative demand. To this end, we provide monthly estimations using two of the most commonly used methods: the Priestley-Taylor potential evapotranspiration ( $ET_p$ )<sup>13</sup> method, and the FAO-56 Penman-Monteith reference evapotranspiration ( $ET_0$ )<sup>16</sup> method for both short grass and alfalfa vegetation land-cover. We also provide a monthly estimate of a key driving variable, vapor pressure deficit (VPD) (Sec. 2.1), that is commonly used to assess ecological impacts. All quantities are provided for the historical simulation period (1850–2014) and four future (2015–2100) CMIP6 ScenarioMIP shared socioeconomic pathways (SSPs), 1–2.6, 2–4.5, 3–7.0, and 5–8.5<sup>19,20</sup>. These estimations are generated for the CMIP6-MME, of which ensemble members are selected based on the availability of all climate variables needed to complete the multiple estimations evaporative demand described below.

<sup>1</sup>Civil, Environmental, and Architectural Engineering Dept., University of Colorado, Boulder, USA. <sup>2</sup>Cooperative Institute for Research in Environmental Sciences, Boulder, USA. <sup>3</sup>NOAA Physical Science Laboratory, Boulder, USA. ✉e-mail: [Nels.Bjarke@colorado.edu](mailto:Nels.Bjarke@colorado.edu)

Following the development of the CMIP6-MME of  $ET_p$ ,  $ET_0$ , and VPD, we perform a validation using observation-based historical baselines of both the monthly estimations of  $ET_0$  from the Climate Research Unit (CRU) dataset<sup>21</sup> and estimations of  $ET_p$  and  $ET_0$  derived from the ERA5-Land reanalysis<sup>22</sup>. Along with the expected variation in performance due to differences in GCM simulated historical hydroclimates<sup>23,24</sup>, we note a general coherence in the bias that exists across all GCMs within both  $ET_p$  and  $ET_0$  that likely arise due to the assumptions embedded within each methodology. Specifically, we evaluate differences the two major components evaporative demand, the advection and radiation components, from each GCM using the same components from ERA5-Land as the baseline to gain more insight into the sources of bias.

In sum, this data descriptor addresses the need for a comprehensive, methodologically consistent archive of future changes in multiple measures of evaporative demand from the latest suite of climate projections. We generate CMIP6-derived estimates of  $ET_p$ ,  $ET_0$ , and VPD, validate the dataset over the historical period with multiple observational datasets, and conduct a brief evaluation of projected future changes in  $ET_p$  and  $ET_0$ . This dataset<sup>18</sup> is built using methods developed for multi-disciplinary hydroclimate and agricultural analyses<sup>13,15</sup>, with the goal of providing the user community a consistent and comprehensive product useful for understanding future changes in water availability. Section 2 describes the methods used to estimate  $ET_p$ ,  $ET_0$ , VPD, and other intermediate step variables along with a description of the CMIP6-MME. Section 3 contains a description of where to access and contents of the data repository. Section 4 contains the technical validation of the multiple estimates of evaporative demand and section 5 directs users to the Python codes used to calculate all of the variables described in section 2.

## Methods

The first evaporative demand methodology described and applied herein is the FAO-56 Penman-Monteith formulation (Sec. 2.1) for estimating reference evapotranspiration ( $ET_0$ )<sup>14,16</sup> which has the most inputs relative to other methods. Within this section we describe the calculation of monthly net surface-radiation (Sec. 2.1.1), 2-meter wind-speed (Sec. 2.1.2), vapor pressure (Sec. 2.1.3), and saturated vapor pressure (Sec. 2.1.4) from both the GCMs and the ERA5-Land reanalysis. Further, we describe the iterations of  $ET_0$  estimations using two different assumptions of vegetation cover: short grass and alfalfa (Sec. 2.1.5). In Sec. 2.2, we describe the method used to determine the advective and radiation components of  $ET_0$  and how those components should be interpreted. The subsequent section contains the description of the Priestley-Taylor potential evapotranspiration formulation ( $ET_p$ ) (Sec. 2.3) and the terms used in this dataset to generate monthly estimations of  $ET_p$ . The  $ET_p$  calculated is not split into advective and radiation components due to the choice of coefficient that represents the relative contribution of each component (described further in Sec. 2.3). The final section of the methods (Sec. 2.4) describes the land masking that we apply as a last step for preparation of the data to align with the intended domain of application of the Penman-Monteith and Priestley-Taylor methodologies. Finally, in Sec. 2.5, we describe the set of GCMs were selected based on availability of data and used to generate the CMIP-MME estimates of potential and reference evapotranspiration.

**FAO 56 Penman-Monteith Reference Evapotranspiration.** The FAO-56 Penman-Monteith  $ET_0$  method was developed to estimate the evaporative demand of a well-irrigated crop land, with assumptions about the land-surface vegetation allowing for  $ET_0$  to be applied in the approximation of required water usage in agricultural operation (R. G. Allen *et al.*<sup>25</sup>; R. G. Allen & Pruitt<sup>16</sup>). It has been applied widely (e.g. Ashofteh Parisa-Sadat *et al.*<sup>26</sup>, Beven<sup>27</sup>, Kay & Davies<sup>28</sup>, Kingston *et al.*<sup>29</sup>, McKenney & Rosenberg<sup>30</sup>, Tabari & Hosseinzadeh Talaei<sup>31</sup>) and is generally considered among the most accurate methods, despite its relatively large data requirement<sup>32,33</sup>. Included in the following subsections are descriptions of the  $ET_0$  method that we applied in the generation of the subject dataset<sup>18</sup> and all of the intermediate steps required to convert the input variables derived from CMIP6-MME output to represent their form as described in the FAO-56 guidelines. To provide an estimate of the two most commonly applied crop types to the formulation, we generate estimations for both short-grass and alfalfa crops in this dataset<sup>18</sup> (described further in Sec. 2.1.5).

$$ET_0 = \frac{0.408\Delta R_n + \gamma \left( \frac{C_n}{T+273} \right) (e_s - e_a)}{\Delta + \gamma(1 + C_d u_2)} \quad (1)$$

Where  $ET_0$  is reference evapotranspiration ( $\text{mm day}^{-1}$ ),  $\Delta$  is the slope of the saturated vapor pressure curve ( $\text{kPa C}^{-1}$ ),  $R_n$  is the net radiation ( $\text{MJ day}^{-1}$ ),  $\gamma$  is the psychrometric constant ( $\text{kPa K}^{-1}$ ),  $C_n$  and  $C_d$  are empirically derived values for a given reference crop from Allen *et al.*<sup>25</sup> ( $\text{s m}^{-1}$ ),  $T$  is monthly mean air temperature ( $^{\circ}\text{C}$ ),  $(e_s - e_a)$  is the saturation vapor pressure deficit ( $\text{kPa}$ ), and  $u_2$  is the monthly average wind speed at 2 meters ( $\text{m s}^{-1}$ ).

**Net Surface-Radiation.** Net radiation ( $R_n$ ) is calculated here using sensible (hfss) and latent (hfls) heat fluxes from the GCM simulations to implicitly account for the contribution of the ground heat flux ( $G$ ) to the net radiation balance as is described in Allen *et al.*<sup>25</sup>. This approach is consistent with methods applied in previous evaluations of the surface-energy balance<sup>2</sup> as follows:

$$R_n = \text{hfss} + \text{hfls} \quad (2)$$

Though the ground heat flux has been shown in previous studies to be nearly negligible at the monthly time scale<sup>2</sup>, this approach allows us to avoid assumption of zero  $G$  in our estimation of  $ET_0$ . We followed this method of using the latent and sensible heat flux here as our analysis using the same category of GCM output data<sup>2</sup>.

**2-Meter Wind-speed.** FAO-56 guidelines call for the use of the surface windspeeds at a height of 2-meter above the land surface<sup>25</sup>, but all CMIP6-MME monthly windspeed output is documented as the 10-meter surface windspeeds. Furthermore, the monthly average 10-meter surface windspeeds we use here are directly computed outputs for each model within the CMIP6 archive (variable name: SfcWind). This approach allows for a more effective capturing of the monthly turbulent flux contribution to  $ET_0$  as opposed to computing average wind speed from north-south and east-west velocities which yields much smaller magnitude monthly wind speed values. To convert windspeeds to match the specifications of FAO-56, we apply the equation from the FAO-56 guidelines<sup>25</sup>:

$$u_2 = \frac{4.87u_z}{\ln(67.8z - 5.42)} \quad (3)$$

Where  $u_2$  is the 2-meter monthly averaged surface windspeed,  $u_z$  is the monthly averaged surface windspeed at height  $z$ , and  $z$  is the height above the land surface expressed in meters of the surface windspeed of the input data, in this case 10 meters.

**Vapor Pressure.** Following the guidance of the FAO-56 manual<sup>25</sup>, we estimate the saturated vapor pressure ( $e_s$ ) using the monthly mean air temperature from each of the GCMs. Then we utilize  $e_s$  and the monthly mean relative humidity ( $h_r$ ) to calculate a monthly vapor pressure deficit (VPD) and the slope of the saturated vapor pressure curve<sup>25</sup>. Though this variable is an intermediate step in the estimation of the evaporative demand, VPD has utility in a variety of agricultural<sup>34,35</sup> and ecological<sup>36-38</sup> applications due to its modulation of stomatal transpiration and relationship with vegetation production, so we have included this variable in the published dataset<sup>18</sup>.

**Saturated Vapor Pressure.** The saturated vapor pressure is calculated directly from the monthly mean air temperature ( $T_{avg}$ ) and represents the theoretical pressure of an air mass at saturation for the given  $T_{avg}$ . It is calculated using formulas from<sup>25,39</sup>:

$$e_s = 0.6108e^{\frac{17.27T_{avg}}{T_{avg}+237.3}} \quad (4)$$

The FAO56 guidelines suggest that the use of mean of maximum and minimum air temperatures ( $T_{max}$  and  $T_{min}$  respectively) for use in the saturated vapor pressure on daily timescales is preferable to using daily mean air temperature due to the non-linearity of the diurnal temperature variations, particularly as it can bias the estimation of relative humidity (Allen *et al.*<sup>25</sup>). However, authors tested the use of monthly  $T_{max}$  and  $T_{min}$  in the estimation of  $e_s$  and compared with the use of monthly  $T_{avg}$  and found marginal differences in the resulting  $ET_0$  (not shown). This paired with the lack of availability of  $T_{min}$  and  $T_{max}$  from the CMIP6 archive for all models within the CMIP6-MME used here, it is preferable to use the monthly  $T_{avg}$  values in this step.

**Vapor Pressure Deficit (VPD).** VPD is calculated using the relative humidity ( $h_r$ ) and  $e_s$  to estimate the difference between the actual vapor pressure of the air and the saturated vapor pressure of the air on a monthly time scale. VPD is calculated as:

$$VPD = \left(1 - \frac{h_r}{100}\right) * e_s \quad (5)$$

**Saturated Vapor Pressure Curve.** The final intermediate variable required for both the  $ET_0$  and  $ET_p$  (described below) is the slope of the saturated vapor pressure curve ( $\Delta$ ).  $\Delta$  is the instantaneous rate at which the saturation vapor pressure changes at a given temperature and is calculated as:

$$\Delta = \frac{4098e_s}{(T_{avg} + 237.3)^2} \quad (6)$$

**Vegetation Parameters.** We apply two formulations of the vegetation parameters that we calibrated for use in the Allen *et al.*<sup>16</sup> formulation. Guidance from agriculturally focused evaporative demand analyses<sup>40</sup> suggests that short-grass and alfalfa are the most widely applied crop types for estimation of  $ET_0$ , which is the justification for our selection. Parameter values for  $C_n$  and  $C_d$  (Eq. 1) are 900 and 0.34 for respectively for short grass and 1600 and 0.38 for alfalfa<sup>16</sup>. These values represent the vegetation resistance terms within the advective portion of the  $ET_0$  formulation and were empirically derived in<sup>25</sup>.

**Radiation and Advection components of  $ET_0$ .** To evaluate the relative contributions of evaporative demand components to the historical  $ET_0$  and future changes to  $ET_0$ , we split the numerator of Eq. 1 to serve as close approximations of the radiation (the  $0.408\Delta R_n$  numerator term) and advection (the  $\gamma\left(\frac{C_n}{T+273}\right)(e_s - e_a)$  numerator term) components. One can observe that both terms,  $ET_{0, Rad}$  (Eq. 7) and  $ET_{0, Adv}$  (Eq. 8) have the same denominator that contains a windspeed term, so it is not to be considered the pure separation of radiation and advection, simply a rough approximation of the relative contribution of each term. When one evaluates the ratio of the two terms, as is done in the technical validation of this manuscript, the advective component that exists in the denominators of the separated components (Eqs. 7, 8) cancel out it then becomes a comparison the relative

GCM Name	Institution	Resolution	Variant ID
ACCESS-CM2 <sup>48</sup>	Commonwealth Scientific and Industrial Research Organisation	2.8° × 2.8°	r1i1p1f1
ACCESS-ESM1-5 <sup>49</sup>			
CESM2 <sup>50</sup>	National Center for Atmospheric Research	100 km	
CMCC-CM2 <sup>51</sup>	Fondazione Centro Euro-Mediterraneo	1.25° × 0.9°	
CMCC-ESM2 <sup>52</sup>	Fondazione Centro Euro-Mediterraneo	100 km	
EC-Earth3 <sup>33</sup>	Agencia Estatal de Meteorologia	100 km	
GFDL-ESM4 <sup>54</sup>	National Oceanic and Atmospheric Administration, Geophysical Fluid Dynamics Laboratory	100 km	
INM-CM4-8 <sup>55</sup>	Institute for Numerical Mathematics	1.4° × 1.4°	
INM-CM5-0 <sup>56</sup>	Institute for Numerical Mathematics	1.4° × 1.4°	
IPSL-CM6A-LR <sup>57</sup>	Institut Pierre Simon Laplace	2° × 1.5°	
MIROC6 <sup>58</sup>	Atmosphere and Ocean Research Institute, The University of Tokyo and Japan Agency for Marine-Earth Science and Technology	2.5° × 1.3°	
MPI-ESM1-2-HR <sup>59</sup>	Max Planck Institute for Meteorology Meteorological Research Institute	250 km	
MPI-ESM1-2-LR <sup>60</sup>		250 km	
MRI-ESM2-0 <sup>61</sup>			
UKESM1-0-LL <sup>62</sup>	Met Office Hadley Centre	1.12° × 1.12°	
HadGEM-GC31-LL <sup>63</sup>	Met Office Hadley Centre	1.9° × 1.25°	r1i1p1f3

**Table 1.** Description of each member of the CMIP6-MME including the institution of origin, horizontal resolution of the simulation, and the variant ID used to calculate  $ET_p$ ,  $ET_0$ , and VPD. Variant IDs are designated as part of CMIP6 guidelines, where each modeling institution labels simulations systematically by realization (r), initialization method (i), physics (p) and forcing (f). The need to use GCM simulations with differing variant IDs arise primarily from data availability. Differences in the forcings (eg. r1i1p1f1 vs r1i1p1f2) of the simulations described below are derived from updated versions of the CMIP6 standardized input forcings but are quantitatively equivalent for each ScenarioMIP experiment in terms of greenhouse gas emissions.

magnitudes of the radiation and advection terms in the numerator. To be explicit about the equations used, equations of the radiation and advection components of  $ET_0$  are shown as Eqs. (7, 8) respectively.

$$ET_{0,Rad} = \frac{0.408\Delta(Rn)}{\Delta + \gamma(1 + C_d\mu_2)} \quad (7)$$

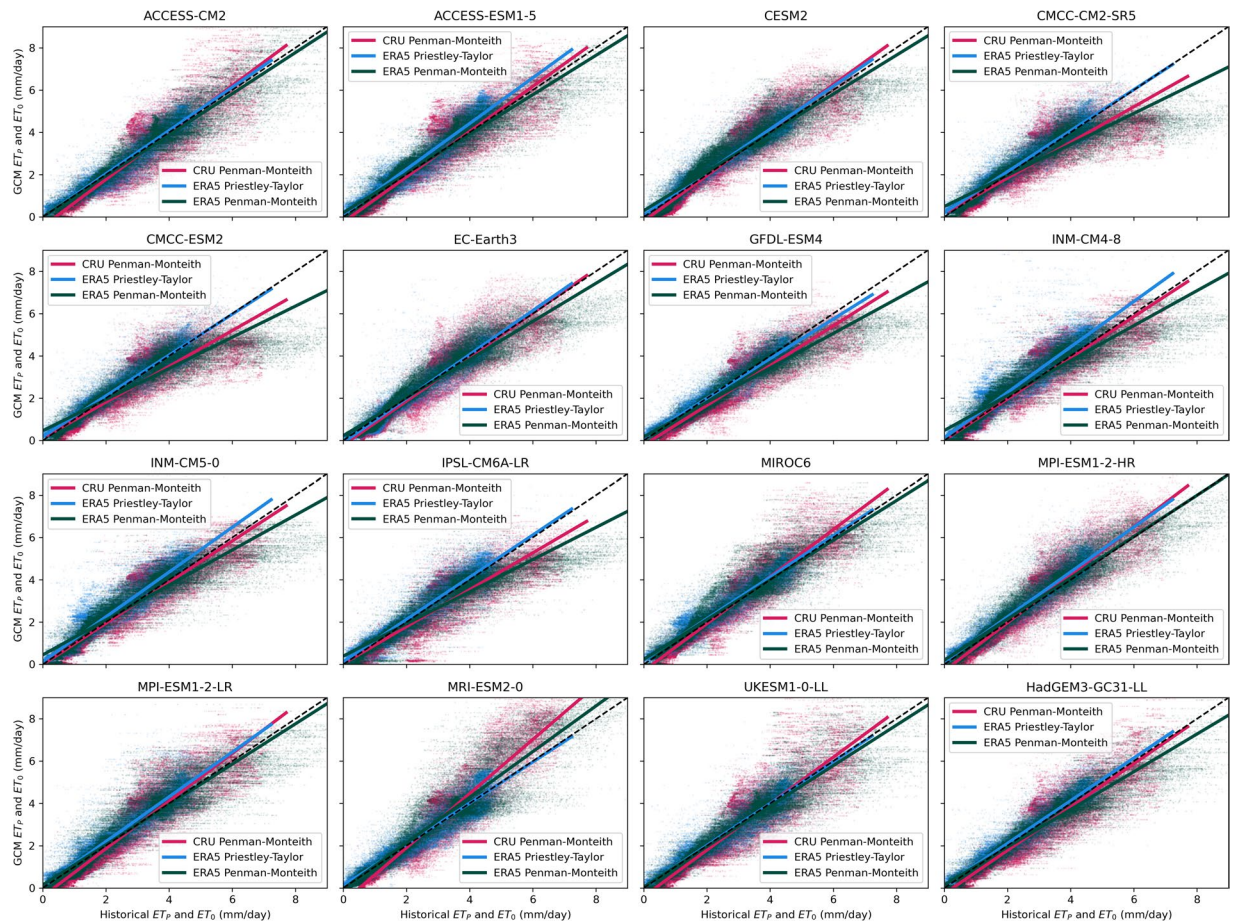
$$ET_{0,Adv} = \frac{\gamma\left(\frac{C_n}{T+273}\right)(e_s - e_a)}{\Delta + \gamma(1 + C_d\mu_2)} \quad (8)$$

**Priestley-Taylor.** The Priestley-Taylor method is applied for the quantification of monthly radiation-based  $ET_p$ <sup>13</sup>, both for comparison to the  $ET_0$  method of estimating evaporative demand and to allow for evaluation of projected changes to evaporative demand driven primarily by increased net radiation at the surface. The Priestley-Taylor methodology assumes a constant ratio of the advective and radiative components of  $ET_p$ , therefore its inclusion in this dataset<sup>18</sup> allows for evaluation of how the increases in surface energy will drive changes to evaporative demand without the inclusion of more uncertain climate parameters such as wind speed or estimates of vegetation contributions. The equation for the Priestley-Taylor formulation of  $ET_p$  is as follows:

$$ET_p = \frac{\alpha\Delta(Rn)}{Lv(\Delta + \gamma)} \quad (9)$$

Where  $\Delta$ ,  $Rn$ , and  $\gamma$  were previously defined in section 2.1 and  $\alpha$  is a constant equal to 1.26 empirically derived in<sup>13</sup>. The  $\alpha$  coefficient of 1.26, which is most commonly applied to empirically account for the non-radiative components of PET for humid regions<sup>5,13</sup>, is held constant to avoid introducing uncertainty associated with regionalized calibration of the parameter. Other studies have shown that the coefficient can be varied spatially with a higher value generally corresponding to PET estimation in more arid regions<sup>41,42</sup>. If users wish to estimate  $ET_p$  using distributed coefficients, then we recommend a simple transformation by dividing the entire dataset by 1.26 to which the desired distributed coefficients could be applied.

**Land Masking.** The methodologies presented here were explicitly developed for application to land-areas, as opposed to open water areas<sup>25</sup>. Because of this, we mask out all oceans and large water body grid cells using land-sea masks obtained for each GCM, where grid cells containing less than 5% of the total area labelled as land were screened and replaced with null values. We include an archive of the land-sea masks in the published dataset<sup>18</sup> for the convenience of data users.



**Fig. 1** Comparison of GCM  $ET_p$  and  $ET_0$  to observational datasets. The CMIP6-MME estimated 30-year mean annual  $ET_p$  and  $ET_0$  rates (mm/day) are compared to historical validation datasets (1951–1980) on a grid by grid basis. The resolution of all datasets are spatially adjusted to match the CRU dataset  $0.5^\circ \times 0.5^\circ$  resolution for validation, whereas the published datasets retain native GCM resolutions.  $ET_0$  from GCMs is compared to CRU  $ET_0$  (magenta) and ERA5-Land  $ET_0$  (green), whereas  $ET_p$  estimated from GCMs is compared to ERA5-Land  $ET_p$  (blue).

**CMIP6-MME.** The CMIP6-MME used to generate this dataset<sup>18</sup> was selected based on the availability of data needed to generate monthly estimations of  $ET_p$  and  $ET_0$ . Input variables were downloaded from multiple ESGF data nodes at monthly resolution for each of the CMIP6-MME described in Table 1. All data were checked for quality control to screen missing data or data corruption for all input variables. Only one realization from each GCM is used in this dataset<sup>18</sup>, primarily due to the lack of consistency in the number of individual GCM realizations from each modelling organization. The Python code used to generate this dataset<sup>18</sup> is provided for those who wish to extend this analysis to multiple realizations from individual models.

For each of the 16 GCMs in Table 1, monthly estimations of  $ET_p$ ,  $ET_0$ , components of  $ET_0$ , and VPD were generated for the historical CMIP6 simulation period (1850–2014) and for the future period of 2015–2099 using four future emission scenarios (SSPs 1–2.6, 2–4.5, 3–7.0 and 5–8.5). The only exception is the lack of a SSP 3–7.0 scenario simulation for the HadGEM3-GC31-LL due to a lack of publication of that specific simulation being published on the ESGF data node. All GCM derived variables contained in the dataset are produced in the native resolution of the parent GCM simulation.

### Data Records

This dataset<sup>18</sup> contains NetCDF files of estimated  $ET_p$ , multiple estimations of  $ET_0$  using vegetation parameters for both short-grass and alfalfa, the advective and radiative components of  $ET_0$ , and estimations of VPD that were calculated in the process of estimating  $ET_0$ . All NetCDF files contain information on the date of generation of the dataset, authors of the dataset, variable IDs, input variable information, and attributes of the parent GCM simulations used to generate the estimations of all variables described above. The link to the archived dataset can be found here: <https://doi.org/10.5281/zenodo.7789759><sup>18</sup>.

Each NetCDF file contains one variable from one GCM for either the historical simulation or one of the four future emission scenarios that were selected for use in this dataset<sup>18</sup>. The naming convention used for each NetCDF file is consistent with the CMIP6 ESGF naming conventions including variable, frequency, GCM name,

	CRU ET <sub>0</sub>	ERA5-Land ET <sub>0</sub>	ERA5-Land ET <sub>p</sub>
ACCESS-CM2	-14.3	-2.4	-7.5
ACCESS-ESM1-5	-9.8	18.2	-1.0
CESM2	-8.5	-0.9	0.2
CMCC-CM2-SR5	-13.4	3.4	-4.5
CMCC-ESM2	-14.7	1.2	-6.0
EC-Earth3	-10.7	-7.0	-2.2
GFDL-ESM4	-19.8	-1.1	-12.7
INM-CM4-8	-6.8	12.2	3.1
INM-CM5-0	-7.4	13.4	2.6
IPSL-CM6A-LR	-13.4	11.8	-5.6
MIROC6	-6.6	5.3	3.5
MPI-ESM1-2-HR	-5.2	18.2	3.5
MPI-ESM1-2-LR	-11.4	8.0	-3.3
MRI-ESM2-0	-5.8	8.1	3.4
UKESM1-0-LL	-13.7	-3.8	-6.1
HadGEM3-GC31-LL	-13.8	0.3	-5.6

**Table 2.** Percent bias of GCM-based ET<sub>0</sub> and ET<sub>p</sub> relative to observational datasets. The 30-year mean annual percent bias is calculated from the historical baselines for both ET<sub>0</sub> and ET<sub>p</sub> over period 1951–1980 of each of the historical CMIP6-MME. For each GCM, the percent bias reported is the mean global percent bias of either ET<sub>0</sub> or ET<sub>p</sub> from all land-area grid cells. ET<sub>0</sub> is compared to both the CRU dataset and the ERA5-Land estimate of ET<sub>0</sub> generated for this validation.

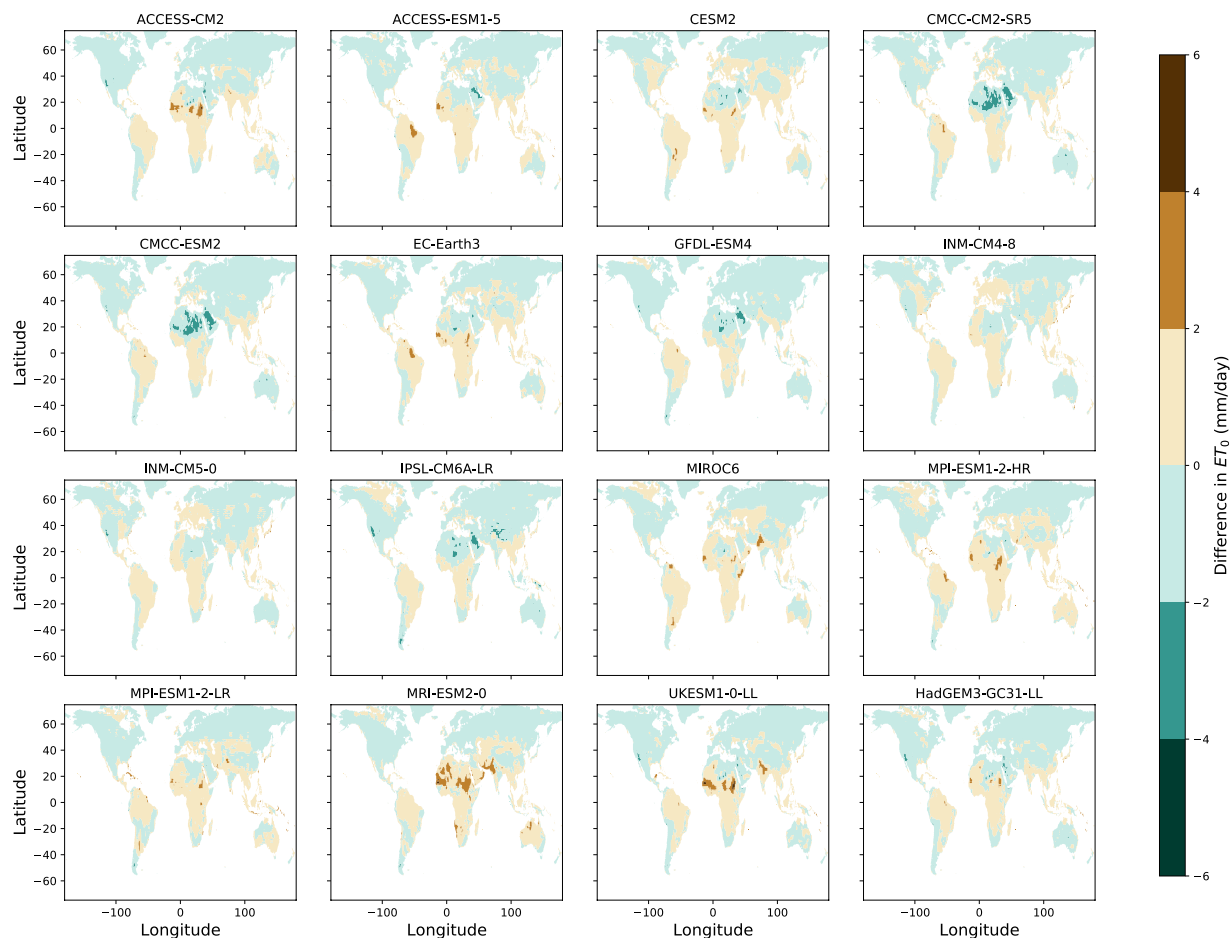
	CRU ET <sub>0</sub>	ERA5-Land ET <sub>0</sub>	ERA5-Land ET <sub>p</sub>
ACCESS-CM2	0.92	0.95	0.97
ACCESS-ESM1-5	0.93	0.95	0.97
CESM2	0.93	0.94	0.97
CMCC-CM2-SR5	0.89	0.91	0.96
CMCC-ESM2	0.89	0.92	0.97
EC-Earth3	0.93	0.95	0.96
GFDL-ESM4	0.93	0.95	0.98
INM-CM4-8	0.92	0.93	0.94
INM-CM5-0	0.93	0.93	0.95
IPSL-CM6A-LR	0.9	0.93	0.96
MIROC6	0.95	0.94	0.97
MPI-ESM1-2-HR	0.94	0.95	0.97
MPI-ESM1-2-LR	0.93	0.93	0.96
MRI-ESM2-0	0.95	0.94	0.95
UKESM1-0-LL	0.92	0.94	0.96
HadGEM3-GC31-LL	0.93	0.95	0.97

**Table 3.** Spatial correlation R of ET<sub>0</sub> and ET<sub>p</sub> estimated from GCMs compared to observational datasets. The spatial Pearson correlation, R, of the 30-year mean annual evaporative demand rate is calculated from the historical baselines for both ET<sub>0</sub> and ET<sub>p</sub> over period 1951–1980 of each of the historical GCM simulations. For each GCM, the R reported is the mean global correlation of either ET<sub>0</sub> or ET<sub>p</sub> from all land-area grid cells. ET<sub>0</sub> is compared to both the CRU dataset<sup>21</sup> and the ERA5-Land estimate of ET<sub>0</sub> generated for this validation.

scenario, variant ID, grid ID, and date range. Each file for the derived variables within this dataset contains data on the native grid of the parent GCM that was used to generate the estimations described above. Additionally, we include a single land-mask for each individual GCM to clarify which grid cells were selected as land-surface grid cells. Further description of the file structure and organization of data within the dataset<sup>18</sup> can be found in the “dataset\_README.txt” file.

### Technical Validation

For the validation of the ET<sub>p</sub> and ET<sub>0</sub> estimations across the ensemble of GCMs, we use 30-year means of monthly estimated evaporative demand rates and 30-year mean annual total evaporative demand depth on a grid cell basis for both the GCMs and the historical baseline datasets. The use of long-term means on a grid cell basis permits a validation of the climatic evaporative demand and an evaluation of the spatial coherence of the simulated climate within the ensemble of GCMs when compared to the multiple historical baselines. For this



**Fig. 2** Differences between GCM and CRU  $ET_0$ . CMIP6-MME derived estimations of historical 30-year mean annual short-grass  $ET_0$  (mm/day) are compared to the CRU historical baseline of  $ET_0$  over the period (1951–1980) on a grid cell by grid cell basis over the land area. Differences here are in units of (mm/day).

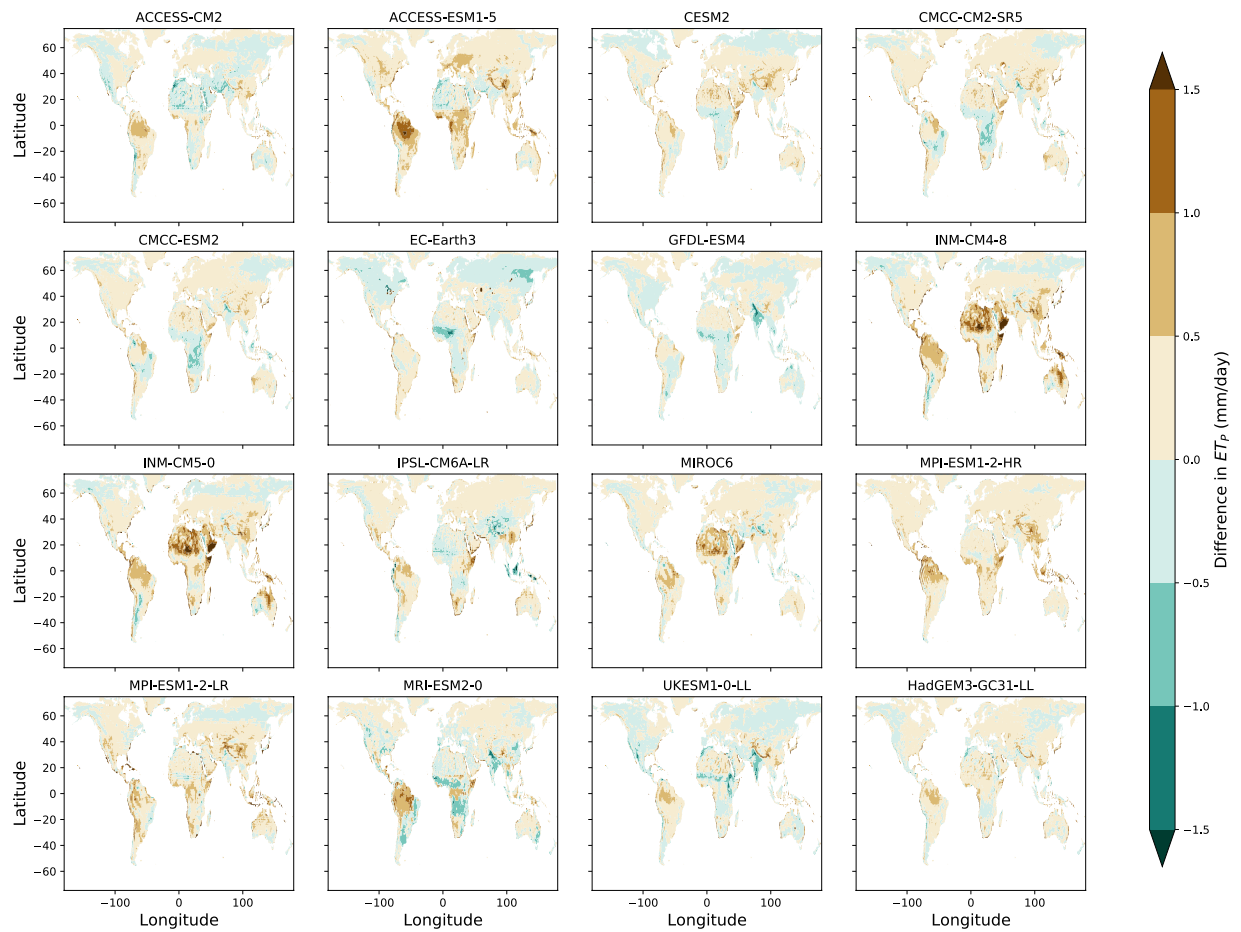
analysis, we independently generate monthly estimates of  $ET_p$  and  $ET_0$  from ERA5-Land in order to ensure consistency with our methods, though published hourly estimates of  $ET_0$ <sup>43</sup> could also serve as a baseline for observationally based historical  $ET_0$ .

The validation period of 1951–1980 coincides with the earliest available data from the ERA5-Land reanalysis and an overlapping period within the historical simulations of the GCMs with minimal acceleration of greenhouse gas emission impacts. We acknowledge that differences in the sensitivities of GCMs within CMIP6<sup>23,24</sup> could result in the inclusion of GCMs with significant warming even within the period of 1951–1980, but the impact of any warming trends in this early historical period is likely to be minimal because we are considering the long-term mean of  $ET_p$  and  $ET_0$  rates.

To compute the spatial correlation and percent bias between estimated fluxes within the GCMs and the observational datasets we align the GCM-simulated evaporative demand fluxes to the  $0.5^\circ \times 0.5^\circ$  grid cell resolution of the CRU<sup>21</sup> and ERA5-Land reanalysis fluxes using a conservative remapping algorithm<sup>44</sup>. The aligning of the GCM derived  $ET_p$  and  $ET_0$  is used only for the validation of the dataset<sup>18</sup> during the historical period and is not a feature of the dataset as published; all GCM estimated evaporative demand terms in the published dataset<sup>18</sup> remain at the native resolution of the respective GCM simulation.

**Historical Simulations.** Comparisons of the 30-year mean annual rate of  $ET_p$  and  $ET_0$  (mm/day) reveal a strong one-to-one relationship between the GCM simulations and historical baselines during the period 1951–1980 (Fig. 1). The global performance of the GCM derived estimations of  $ET_0$  and  $ET_p$  is evaluated based on the global average of the 30-year mean annual percent bias (Table 2) and the spatial Pearson correlation<sup>45</sup> between the 30-year mean annual evaporative demand rates of all land-surface grid cells (Table 3). Validation shown for  $ET_0$  is based on the short-grass vegetation parameterization within the  $ET_0$  formulation for brevity, but results for alfalfa vegetation are not qualitatively different from those shown for the short-grass vegetation.

The percent bias of the 30-year global mean annual GCM estimates of  $ET_0$  range from  $-19.8$  to  $-5.2\%$  when compared to the CRU dataset and from  $-7.0$  to  $18.2\%$  compared the ERA5-Land derived estimate (Table 2).  $ET_p$  from the GCMs demonstrate a range of percent bias from  $-12.7$  to  $3.5\%$  for the global mean when compared to the ERA5-Land  $ET_p$  (Table 2). Visual inspection of the spread of points in Fig. 1. show the largest variations about the one-to-one line are in regions of with the highest historical evaporative demand.



**Fig. 3** Differences between GCM and ERA5-Land  $ET_p$ . CMIP6-MME derived estimations of historical 30-year mean annual short-grass  $ET_p$  (mm/day) are compared to the ERA5-Land historical baseline of  $ET_p$  over the period (1951–1980) on a grid cell by grid cell basis over the land area. Differences here are in units of (mm/day).

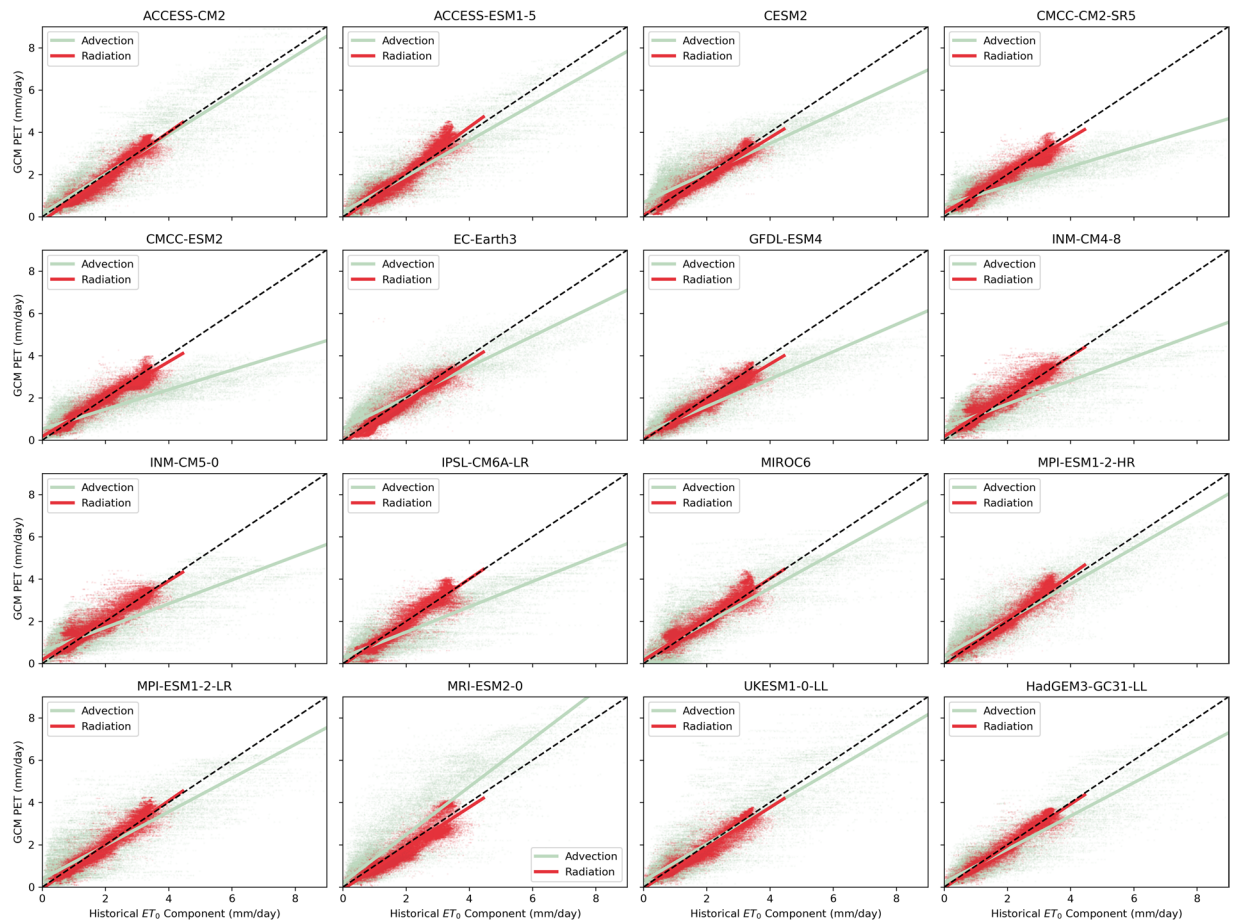
Pearson correlations ( $R$ ) of the 2-d spatial pattern of 30-year mean annual  $ET_0$  range from 0.89 to 0.95 when evaluating against the CRU dataset and 0.91 to 0.95 against the ERA5-Land estimated  $ET_0$  (Table 3). The long-term estimated  $ET_p$  from the CMIP6-MME and ERA5-Land are even more strongly correlated in the historical period, with  $R$  values ranging from 0.94–0.98 (Table 3). The strength of these correlations suggests that, despite the biases in evaporative demand estimations, the GCM simulations all generally reproduce the observed spatial pattern of  $ET_p$  and  $ET_0$  across the global land surface.

**FAO-56 Penman-Monteith.** The sign and magnitude of biases of 30-year mean annual  $ET_0$  reveal some noticeable regional patterns across GCMs. The northern hemisphere appears to have a consistent negative bias for all GCMs (Fig. 2), which is likely the source of the global mean negative  $ET_0$  bias for the historical period across all 16 GCMs (Table 2) given the majority (~68%) of the land-surface on the Earth is in the northern hemisphere.

Biases relative to the CRU  $ET_0$  are larger than those relative to the ERA5-Land derived estimates (Table 2) and the  $R$  of GCMs estimated  $ET_0$  is consistently lower when comparing to the CRU  $ET_0$  versus ERA5-Land derived estimates (Table 3). These differences in the relative agreement of the CMIP6-MME derived estimates are likely due to relatively minor differences in the methods applied to estimate the monthly time series of vapor pressure in CRU's estimates and the methods used here for both the ERA-Land derived historical baselines and model simulations. Differences in the estimation of vapor pressure deficit between the two datasets arise from differing quantification of relative humidity and different accounting of mean monthly temperatures which are required for estimation of the saturation vapor pressure. When compared the ERA5-Land derived estimates of  $ET_0$  all GCMs have high spatial correlation ( $R > 0.91$ ) (Table 3) and only two of the models produce large annual biases: ACCESS-ESM1-5 and MPI-ESM1-2-HR (both 18.2% bias during the historical simulation) (Table 2).

**Priestley-Taylor.** Scatter about the one-to-one line exists within all ensembled members of the CMIP6-MME compared to the observational datasets, but the magnitude of the global bias tends to be smaller within the  $ET_p$  estimations from GCMs compared to the biases of  $ET_0$ . The global multi-model ensemble mean percent bias is  $-2.3\%$  for  $ET_p$  from ERA5-Land, the smallest of the ensemble biases compared to the ensemble mean  $ET_0$  bias of  $-11.0\%$  for the CRU dataset and  $5.3\%$  for the ERA5-Land  $ET_0$  (Table 2). The smaller range of evaporative demand rates within the  $ET_p$  estimations compared to  $ET_0$  (Fig. 1) are consistent with the assumption of a





**Fig. 4** Comparison of advective and radiative components GCM  $ET_0$  to ERA5-Land. The advective and radiative components of short-grass  $ET_0$  for each GCM are compared to the estimated components of  $ET_0$  from ERA5-Land. Each point represents the 30-year mean monthly  $ET_0$  (mm/day) from 1951–1980 in the historical period for each grid cell across the global land-surface. The advection component (light green) of  $ET_0$  is primarily modulated by the wind speed and vapor pressure deficit, whereas the radiation component is dependent on the net radiation.

uniform  $\alpha$  coefficient of 1.26 within the  $ET_p$  formulation (see Sec. 2.3 description of the coefficient). Others have shown that arid regions should generally have a larger magnitude  $\alpha$  coefficient applied in the  $ET_p$  formula<sup>41,42</sup>, which in this case could lead to a more effective representation of the advective component of evaporative demand within those regions. While it is beyond the scope of this dataset<sup>18</sup> to calibrate spatially distributed  $\alpha$  coefficients, one could further explore this avenue by performing a transformation of the  $ET_p$  described in Section 2.3.

Historical  $ET_p$  from all members of the CMIP6-MME demonstrate positive biases in humid parts of northern South America due simulated net radiation that is much larger in magnitude than in the ERA5-Land reanalysis (Fig. 3). However, this positive bias does not appear to consistently impact regions with high relative humidity climates, as only 8 out of 16 GCMs demonstrate the same positive bias within humid regions of Central Africa. Bands of small-magnitude negative  $ET_p$  bias in the highest northern latitudes are present for 11 out of 16 GCMs, consistent with known differences in the radiative balance of northern-hemisphere ice covered land-areas within the CMIP6-MME compared to observation datasets<sup>46</sup>.

**Advection and Radiation Components of  $ET_0$ .** The variations in biases of  $ET_0$  estimations between GCMs are primarily driven by differences in the biases of the advective component of the evaporative demand. Nine of the 16 GCMs produce both (1) too much advection-driven  $ET_0$  ( $ET_{0,ADV}$ ) in regions where ERA5-Land has lower magnitude ( $<3$  mm/day)  $ET_{0,ADV}$  and (2) too little  $ET_{0,ADV}$  in regions of higher magnitude ERA5-Land  $ET_{0,ADV}$  (Fig. 4). Only one of the GCMs, MRI-ESM2-0, consistently produces more  $ET_{0,ADV}$  than ERA5-Land across all regions (Fig. 4). A potential driver of this bias that exists across multiple GCMs is that the simulated VPD is too large (small) within the regions that demonstrate low (high) levels of advection within ERA5-Land. This hypothesis is supported by examination of the tendency for the nine GCMs that display this bimodal advection bias to underpredict total  $ET_0$  in more arid regions such as Northern Africa where VPD is known to be high<sup>47</sup> (Fig. 2). The radiative component of  $ET_0$  appears to be well-captured by the historical simulations across all 16 GCMs (Fig. 4), which is consistent with the performance of the  $ET_p$  estimations that are primarily driven by the net radiation (Fig. 1).

## Code availability

The Python scripts used to calculate  $ET_p$ ,  $ET_0$ ,  $ET_0$  components, and VPD can be found within the repository alongside the data described herein. Accompanying the code is a small subset of GCM data that can be used to test run the script. Python scripts utilize a small subset of libraries in the Python3 base and the xarray (v2022.11.0) library to handle calculations of the gridded datasets. The python code used to generate the dataset<sup>18</sup> described above can be found in the GitHub repository using the following link: <https://github.com/nelsbjarke/PET>.

Received: 9 January 2023; Accepted: 6 June 2023;

Published online: 27 June 2023

## References

- Berghuijs, W. R., Larsen, J. R., Emmerik, T. H. M. van & Woods, R. A. A Global Assessment of Runoff Sensitivity to Changes in Precipitation, Potential Evaporation, and Other Factors. *Water Resources Research* **53**, 8475–8486 (2017).
- Cook, B. I., Smerdon, J. E., Seager, R. & Coats, S. Global warming and 21st century drying. *Clim Dyn* **43**, 2607–2627 (2014).
- Cook, B. I. *et al.* Twenty-First Century Drought Projections in the CMIP6 Forcing Scenarios. *Earth's Future* **8**, (2020).
- Hobbins, M. T. *et al.* The Evaporative Demand Drought Index. Part I: Linking Drought Evolution to Variations in Evaporative Demand. *Journal of Hydrometeorology* **17**, 1745–1761 (2016).
- Milly, P. C. D. & Dunne, K. A. On the Hydrologic Adjustment of Climate-Model Projections: The Potential Pitfall of Potential Evapotranspiration. *Earth Interactions* **15**, 1–14 (2011).
- Mukherjee, S., Mishra, A. & Trenberth, K. E. Climate change and drought: a perspective on drought indices. *Current Climate Change Reports* **4**, 145–163 (2018).
- Scheff, J. & Frierson, D. M. Terrestrial aridity and its response to greenhouse warming across CMIP5 climate models. *Journal of Climate* **28**, 5583–5600 (2015).
- Scheff, J. & Frierson, D. M. W. Scaling potential evapotranspiration with greenhouse warming. *Journal of Climate* **27**, 1539–1558 (2013).
- Thornthwaite, C. W. An Approach toward a Rational Classification of Climate. *Geographical Review* **38**, 55–94 (1948).
- Hargreaves, G. H. & Samani, Z. A. Estimating potential evapotranspiration. *Journal of the Irrigation and Drainage Division* **108**, 225–230 (1982).
- Milly, P. C. D. & Dunne, K. A. Colorado River flow dwindles as warming-driven loss of reflective snow energizes evaporation. *Science* <https://doi.org/10.1126/science.aay9187> (2020).
- Xiang, K., Li, Y., Horton, R. & Feng, H. Similarity and difference of potential evapotranspiration and reference crop evapotranspiration – a review. *Agricultural Water Management* **232**, 106043 (2020).
- Priestley, C. H. B. & Taylor, R. J. On the Assessment of Surface Heat Flux and Evaporation Using Large-Scale Parameters. *Mon. Wea. Rev.* **100**, 81–92 (1972).
- Allen, R. G., Jensen, M. E., Wright, J. L. & Burman, R. D. Operational Estimates of Reference Evapotranspiration. *Agronomy Journal* **81**, 650–662 (1989).
- Crop evapotranspiration: guidelines for computing crop water requirements.* (Food and Agriculture Organization of the United Nations, 1998).
- Allen, R. G. & Pruitt, W. O. FAO-24 Reference Evapotranspiration Factors. *Journal of Irrigation and Drainage Engineering* **117**, 758–773 (1991).
- Anderson, D. B. Relative humidity or vapor pressure deficit. *Ecology* **17**, 277–282 (1936).
- Bjarke, N., Barsugli, J. & Livneh, B. CMIP6 derived ensemble of global vapor pressure deficit, potential evapotranspiration, and reference evapotranspiration. *Zenodo* <https://doi.org/10.5281/ZENODO.7789759> (2023).
- O'Neill, B. C. *et al.* The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6. *Geosci. Model Dev.* **9**, 3461–3482 (2016).
- Tebaldi, C. *et al.* Climate model projections from the Scenario Model Intercomparison Project (ScenarioMIP) of CMIP6. *Earth Syst. Dynam.* **12**, 253–293 (2021).
- Harris, I., Osborn, T. J., Jones, P. & Lister, D. Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Sci Data* **7**, 109 (2020).
- Copernicus Climate Change Service. ERA5-Land monthly averaged data from 2001 to present. *Copernicus Climate Change Service (C3S) Climate Data Store (CDS)*. <https://doi.org/10.24381/CDS.68D2BB30> (2019).
- Marvel, K., Schmidt, G. A., Miller, R. L. & Nazarenko, L. S. Implications for climate sensitivity from the response to individual forcings. *Nature Climate Change* **6**, 386–389 (2016).
- Sherwood, S. C. *et al.* An Assessment of Earth's Climate Sensitivity Using Multiple Lines of Evidence. *Reviews of Geophysics* **58**, e2019RG000678 (2020).
- Allen, R. G., Pereira, L. S., Raes, D. & Smith, M. Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56. *Fao, Rome* **300**, D05109 (1998).
- Parisa-Sadat, A., Bozorg, H. O. & Mariño Miguel, A. Risk Analysis of Water Demand for Agricultural Crops under Climate Change. *Journal of Hydrologic Engineering* **20**, 04014060 (2015).
- Beven, K. A sensitivity analysis of the Penman-Monteith actual evapotranspiration estimates. *Journal of Hydrology* **44**, 169–190 (1979).
- Kay, A. L. & Davies, H. N. Calculating potential evaporation from climate model data: A source of uncertainty for hydrological climate change impacts. *Journal of Hydrology* **358**, 221–239 (2008).
- Kingston, D. G., Todd, M. C., Taylor, R. G., Thompson, J. R. & Arnell, N. W. Uncertainty in the estimation of potential evapotranspiration under climate change. *Geophysical Research Letters* **36** (2009).
- McKenney, M. S. & Rosenberg, N. J. Sensitivity of some potential evapotranspiration estimation methods to climate change. *Agricultural and Forest Meteorology* **64**, 81–110 (1993).
- Tabari, H. & Hosseinzadeh Talaei, P. Sensitivity of evapotranspiration to climatic change in different climates. *Global and Planetary Change* **115**, 16–23 (2014).
- López-Urrea, R., Martín de Santa Olalla, F., Fabeiro, C. & Moratalla, A. Testing evapotranspiration equations using lysimeter observations in a semi-arid climate. *Agricultural Water Management* **85**, 15–26 (2006).
- Ortega-Farías, S., Olliso, A., Antonioletti, R. & Brisson, N. Evaluation of the Penman-Monteith model for estimating soybean evapotranspiration. *Irrigation Science* **23**, 1–9 (2004).
- Sinclair, T. R. *et al.* Limited-transpiration response to high vapor pressure deficit in crop species. *Plant Science* **260**, 109–118 (2017).
- Will, R. E., Wilson, S. M., Zou, C. B. & Hennessey, T. C. Increased vapor pressure deficit due to higher temperature leads to greater transpiration and faster mortality during drought for tree seedlings common to the forest–grassland ecotone. *New Phytologist* **200**, 366–374 (2013).
- Grossiord, C. *et al.* Plant responses to rising vapor pressure deficit. *New Phytologist* **226**, 1550–1566 (2020).
- López, J., Way, D. A. & Sadok, W. Systemic effects of rising atmospheric vapor pressure deficit on plant physiology and productivity. *Global Change Biology* **27**, 1704–1720 (2021).

38. Yuan, W. *et al.* Increased atmospheric vapor pressure deficit reduces global vegetation growth. *Science advances* **5**, eaax1396 (2019).
39. Chow, V. T., Maidment, D. & Mays, L. *Applied Hydrology*. (McGraw-Hill Book Company, 1988).
40. Irmak, S., Allen, R. G. & Whitty, E. B. Daily Grass and Alfalfa-Reference Evapotranspiration Estimates and Alfalfa-to-Grass Evapotranspiration Ratios in Florida. *Journal of Irrigation and Drainage Engineering* **129**, 360–370 (2003).
41. Aschonitis, V. G. *et al.* High-resolution global grids of revised Priestley–Taylor and Hargreaves–Samani coefficients for assessing ASCE-standardized reference crop evapotranspiration and solar radiation. *Earth Syst. Sci. Data* **9**, 615–638 (2017).
42. Cristea, N. C., Kampf, S. K. & Burges, S. J. Revised Coefficients for Priestley–Taylor and Makkink–Hansen Equations for Estimating Daily Reference Evapotranspiration. *J. Hydrol. Eng.* **18**, 1289–1300 (2013).
43. Singer, M. B. *et al.* Hourly potential evapotranspiration at 0.1° resolution for the global land surface from 1981–present. *Sci Data* **8**, 224 (2021).
44. Jones, P. W. First- and Second-Order Conservative Remapping Schemes for Grids in Spherical Coordinates. *Mon. Wea. Rev.* **127**, 2204–2210 (1999).
45. Holický, M. *Introduction to Probability and Statistics for Engineers*. <https://doi.org/10.1007/978-3-642-38300-7> (Springer Berlin Heidelberg, 2013).
46. Abe, M., Nozawa, T., Ogura, T. & Takata, K. Effect of retreating sea ice on Arctic cloud cover in simulated recent global warming. *Atmospheric Chemistry and Physics* **16**, 14343–14356 (2016).
47. Yang, Z., Zhang, Q., Hao, X. & Yue, P. Changes in Evapotranspiration Over Global Semiarid Regions 1984–2013. *J. Geophys. Res. Atmos.* **124**, 2946–2963 (2019).
48. Dix, M. *et al.* CSIRO-ARCSS ACCESS-CM2 model output prepared for CMIP6 CMIP. *Earth System Grid Federation* <https://doi.org/10.22033/ESGF/CMIP6.2281> (2019).
49. Ziehn, T. *et al.* CSIRO ACCESS-ESM1.5 model output prepared for CMIP6 CMIP. *Earth System Grid Federation* <https://doi.org/10.22033/ESGF/CMIP6.2288> (2019).
50. Danabasoglu, G. NCAR CESM2 model output prepared for CMIP6 CMIP historical. *Earth System Grid Federation* <https://doi.org/10.22033/ESGF/CMIP6.7627> (2019).
51. Lovato, T. & Peano, D. CMCC CMCC-CM2-SR5 model output prepared for CMIP6 CMIP historical. *Earth System Grid Federation* <https://doi.org/10.22033/ESGF/CMIP6.3825> (2020).
52. Lovato, T., Peano, D. & Butenschön, M. CMCC CMCC-ESM2 model output prepared for CMIP6 CMIP historical. *Earth System Grid Federation* <https://doi.org/10.22033/ESGF/CMIP6.13195> (2021).
53. EC-Earth Consortium (EC-Earth). EC-Earth-Consortium EC-Earth3 model output prepared for CMIP6 CMIP historical. *Earth System Grid Federation* <https://doi.org/10.22033/ESGF/CMIP6.4700> (2019).
54. Krasting, J. P. *et al.* NOAA-GFDL GFDL-ESM4 model output prepared for CMIP6 CMIP historical. *Earth System Grid Federation* <https://doi.org/10.22033/ESGF/CMIP6.8597> (2018).
55. Volodin, E. *et al.* INM INM-CM4-8 model output prepared for CMIP6 CMIP piControl. *Earth System Grid Federation* <https://doi.org/10.22033/ESGF/CMIP6.5080> (2019).
56. Volodin, E. *et al.* INM INM-CM5-0 model output prepared for CMIP6 CMIP piControl. *Earth System Grid Federation* <https://doi.org/10.22033/ESGF/CMIP6.5081> (2019).
57. Boucher, O. *et al.* IPSL IPSL-CM6A-LR model output prepared for CMIP6 CMIP. *Earth System Grid Federation* <https://doi.org/10.22033/ESGF/CMIP6.1534> (2018).
58. Shiogama, H., Abe, M. & Tatebe, H. MIROC MIROC6 model output prepared for CMIP6 ScenarioMIP. *Earth System Grid Federation* <https://doi.org/10.22033/ESGF/CMIP6.898> (2019).
59. von Storch, J.-S. *et al.* MPI-M MPIESM1.2-HR model output prepared for CMIP6 HighResMIP. *Earth System Grid Federation* <https://doi.org/10.22033/ESGF/CMIP6.762> (2017).
60. Wieners, K.-H. *et al.* MPI-M MPI-ESM1.2-LR model output prepared for CMIP6 CMIP historical. *Earth System Grid Federation* <https://doi.org/10.22033/ESGF/CMIP6.6595> (2019).
61. Yukimoto, S. *et al.* MRI MRI-ESM2.0 model output prepared for CMIP6 CMIP. *Earth System Grid Federation* <https://doi.org/10.22033/ESGF/CMIP6.621> (2019).
62. Tang, Y. *et al.* MOHC UKESM1.0-LL model output prepared for CMIP6 CMIP historical. *Earth System Grid Federation* <https://doi.org/10.22033/ESGF/CMIP6.6113> (2019).
63. Roberts, M. MOHC HadGEM3-GC31-LL model output prepared for CMIP6 HighResMIP. *Earth System Grid Federation* <https://doi.org/10.22033/ESGF/CMIP6.1901> (2017).

## Acknowledgements

We acknowledge the financial support provided by from the NOAA grant titled “Physically-based evaluation of CMIP6 hydrologic projections for the western United States (NA19OAR4310284) and the NOAA grant titled: “Western Water Assessment: Building resilience to compound hazards in the Intermountain West (NA21OAR4310309) that made this research possible.

## Author contributions

N.B., B.L., and J.B. wrote, edited, and reviewed the manuscript; N.B. and J.B. developed scripts to perform calculations and generation of dataset; N.B. was responsible for data access, storage, and management; J.B. and N.B. performed quality control on dataset and methodologies applied; B.L., N.B., and J.B. all coordinated the publication of the dataset in the repository and generated ideas for the description and validation of the dataset.

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to N.B.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023