Figure S1. Example of standard resampling method applied to instantaneous values, as well as mass-conserving resampling.

S1 Timeseries resampling

Sub-daily streamflow observations provided by USGS and WSC are provided as instantaneous values (IV). Such values are only valid for the specific time at which they were obtained, and cannot be seen as representative for a time period on either side of the observation itself. This is illustrated in Figure 1, where a synthetic example of IV are shown as black circles and
5 a simple approximation of the actual timeseries of streamflow as a black line. To accurately calculate the average flow over a particular time window, one first needs to integrate the IVs (under some assumption of how the flow changed between two consecutive IVs) to find the total volume of water that passed the gauge during the time window, and then divide this volume by the length of the time window to find the average flow rate for the time window. Such an approach is guaranteed to conserve mass (dotted blue line in Fig. 1). A potential pitfall is the direct application of tools such as the *resample('1H').mean()* method
10 available in the popular Python package *pandas*. This method does not allow for "closed" windows (meaning that observations on both ends of the time window under consideration would contribute to the mean over the time window), and only provides options to include the "left" or "right" observation but not both. The method also does not take the length of the time window and where observations are located inside the time window into account (see the 13:00-14:00 window in Fig. 1). As a result, applying this method directly to the IV data leads to unintuitive average flow rate estimates, particularly when the IVs are not
15 spaced equally in time (dashed red line in Fig. 1).

S2 Derivation of forcing variables

S2.1 Vapor pressure

ERA5 provides specific humidity and air pressure at the model level from which we download the data. We can derive vapor pressure as a function of specific humidity and air pressure as follows.
20 The ideal gas laws for dry air and water vapor are, respectively (Stull, 2017, Eq. 1.18 and 1.19):

$$P_d = \rho_d R_d T \tag{1}$$
$$e = \rho_v R_v T \tag{2}$$

Where P_d and e are dry air pressure and vapor pressure $[kPa]$, respectively; ρ_d and ρ_v are the densities of dry air and water vapor $[kg\ m^{-3}]$; R_d and R_v are the gas constants for dry air and pure water vapor $[kPa\ K^{-1}\ m^3\ kg^{-1}]$; and T is temperature $[K]$.

Starting with the definition of specific humidity (Stull, 2017, Eq. 4.6):

$$q = \frac{m_v}{m_t} = \frac{m_v}{m_d + m_v} \tag{3}$$

Where q is the specific humidity $[-]$, and m_v and m_t are the mass of water vapor and total mass of the air parcel $[kg]$, respectively. For the remainder, it makes sense to write the total air mass, m_t, as the sum of the mass of water vapor, m_v, and the mass of dry air, m_d, $[kg]$. Specific humidity can also be expressed as the ratio of densities, obtained by dividing both mass variables by their respective volumes V_v and V_d $[m^3]$:

$$q = \frac{\frac{m_v}{V_v}}{\frac{m_d}{V_d} + \frac{m_v}{V_v}} \tag{4}$$

This is equivalent to the ratio of densities ρ $[kg\ m^{-3}]$:

$$q = \frac{\rho_v}{\rho_d + \rho_v} \tag{5}$$

We can express the densities through their corresponding ideal gas laws:

$$q = \frac{\frac{e}{R_v T}}{\frac{P_d}{R_d T} + \frac{e}{R_v T}}, \tag{6}$$

and reorganize these as follows:

$$q = \frac{e}{R_v T} \frac{1}{\frac{P_d}{R_d T} + \frac{e}{R_v T}} \tag{7}$$

$$q = \frac{e}{R_v T} \frac{1}{\frac{P_d R_v T + e R_d T}{R_d R_v T^2}} \tag{8}$$

$$q = \frac{e}{R_v T} \frac{1}{\frac{P_d R_v + e R_d}{R_d R_v T}} \tag{9}$$

$$q = \frac{e}{R_v T} \frac{R_d R_v T}{P_d R_v + e R_d} \tag{10}$$

$$q = \frac{R_v T}{R_v T} \frac{e R_d}{P_d R_v + e R_d} \tag{11}$$

$$q = \frac{e R_d}{R_v \left(P_d + e \frac{R_d}{R_v} \right)} \tag{12}$$

$$q = \frac{R_d}{R_v} \frac{e}{P_d + e \frac{R_d}{R_v}} \tag{13}$$

$$\tag{14}$$

The ratio of dry air and water vapor gas constants, $\frac{R_d}{R_v}$, is often given as ϵ:

$$q = \epsilon \frac{e}{P_d + e\epsilon} \tag{15}$$

Assuming that total pressure $P = P_d + e$, we get an expression of specific humidity, q, in terms of total air pressure, P, vapor pressure, e, and gas constant ratio, ϵ (Stull, 2017, Eq. 4.7, top row):

$$q = \epsilon \frac{e}{P - e + e\epsilon} \tag{16}$$

$$q = \frac{e\epsilon}{P - e(1 - \epsilon)} \tag{17}$$

In atmospheric sciences, this equation is often simplified to:

$$q \approx \frac{e\epsilon}{P} \tag{18}$$

From which vapor pressure e can easily be obtained, under the assumption that the component $e(1 - \epsilon)$ in the denominator is small compared to P. It is however possible to obtain an exact expression of e as a function of q, ϵ, and P. Starting from Eq. 17:

$$q(P - e(1 - \epsilon)) = e\epsilon \tag{19}$$

$$qP - qe + qe\epsilon = e\epsilon \tag{20}$$

$$qe\epsilon - qe - e\epsilon = -qP \tag{21}$$

$$e(q\epsilon - q - \epsilon) = -qP \tag{22}$$

$$e = -\frac{qP}{q\epsilon - q - \epsilon} \tag{23}$$

Where ϵ [−] has a constant value of 0.622, based on gas constants $R_d = 2.871*10^{-4}$ and $R_v = 4.61*10^{-4}$ [$kPa\,K^{-1}\,m^3\,kg^{-1}$].

S2.2 Relative humidity

$$RH = \frac{e}{e_s} \tag{24}$$

Where RH is relative humidity [−], e is vapor pressure [kPa], and e_s is saturation vapor pressure at the current temperature [kPa]. e_s can be calculated using the Clausius-Clapeyron equation (Stull, 2017, Eq. 4.1a):

$$e_s \approx e_0 * exp\left[\frac{L}{R_v}\left(\frac{1}{T_0} - \frac{1}{T}\right)\right] \tag{25}$$

Where e_s is saturation vapor pressure [kPa], e_0 a known saturation vapor pressure at temperature T_0, L is a latent-heat parameter [$J\,kg^{-1}$], R_v the water-vapor gas constant, and T air temperature [K]. Typical values are $e_0 = 0.6113$ [kPa] at $T_0 = 273.15$ [K]. $R_v = 461$ [$J\,K^{-1}\,kg^{-1}$] (note the different units here compared to Appendix 2.1). The latent heat of vaporization for liquid water, $L_v = 2.5 \cdot 10^6$ [$J\,kg^{-1}$], while the latent heat of deposition for ice, $L_d = 2.83 \cdot 10^6$ [$J\,kg^{-1}$].

S2.3 Mean wind speed

ERA5 provides wind speed in the u and v directions. Mean wind speed w (in the same units as u and v) can be obtained with Pythagoras' theorem (European Centre for Medium-range Weather Forecasting, 2023):

$$w = \sqrt{u^2 + v^2} \tag{26}$$

S2.4 Wind direction

ERA5 provides wind speed in the u and v directions, from which a mean wind direction can be derived. We use here ECMWF's definition of *meteorological wind direction*, ϕ [$degrees$] which is the direction from which the wind blows, with North set at $0°$ and increasing clock-wise (European Centre for Medium-range Weather Forecasting, 2023, see also the URL associated with this reference for a helpful graphic):

$$\phi = mod\left(180 + \frac{180}{\pi}atan2(v,u), 360\right) \tag{27}$$

S2.5 Potential evapotranspiration

Daymet and RDRS provide most of the variables needed to run more complex physics-based models, but lack a potential evapotranspiration (PET) variable that would be needed for most simpler conceptual models. We derive two different PET estimates in an attempt to extract as much value out of the variables each data set does provide.

S2.5.1 Daymet

For the Daymet data set, we calculate potential evapotranspiration (PET) estimates using the Priestly-Taylor formula (Priestley and Taylor, 1972) and the necessary Daymet variables. We have selected Priestley-Taylor over other possibilities because this formulation is regularly used in rainfall-runoff modeling (e.g., Newman et al., 2015; Knoben et al., 2020), and this is supported by the assessment in McMahon et al. (2013). We envision these PET estimates to be used for similar purposes in the CAMELS-SPAT data set. Priestly-Taylor PET can be estimated as follows (based on McMahon et al. (2013), but with corrected brackets):

$$E_{PT} = \alpha_{PT}\frac{\Delta}{\Delta+\gamma}\left[\frac{R_n}{\lambda} - \frac{G}{\lambda}\right] \tag{28}$$

Where α_{PT} is the Priestley-Taylor coefficient [$-$], Δ the slope of the vapour pressure curve at air temperature [$\mathrm{kPa\,C^{-1}}$], γ the psychometric constant [$\mathrm{kPa\,C^{-1}}$], R_n the net daily radiation at the surface [$\mathrm{MJ\,m^{-2}\,day^{-1}}$], G the soil heat flux [$\mathrm{MJ\,m^{-2}\,day^{-1}}$], and λ the latent heat of vaporization [$\mathrm{MJ\,kg^{-1}}$].

Following Newman et al. (2015), we use Daymet variables where possible and derived the missing terms using the equations provided in Zotarelli et al. (2009), assuming that the ground heat flux is zero. We use a reference albedo of 0.2, based on the typical values shown in Table S3 of McMahon et al. (2013). Though this underestimates the albedo of short vegetation, and overestimates that of coniferous forests and rain forests, these errors are smaller than not accounting for surface albedo at all.

S2.5.2 RDRS

For the RDRS data set, we calculate PET estimates using the FOA Penman-Monteith formula (Allen et al., 1998) to utilize the information contained in RDRS' wind speed data. McMahon et al. (2013) did not investigate the suitability of different PET formulations at hourly time steps, but our use of Penman-Monteith for hourly data is in line with the work by Singer et al. (2021). Penman-Monteith PET can be calculated as follows (Zotarelli et al., 2009; Singer et al., 2021):

$$E_{PM} = \frac{0.408\Delta\left(R_n - G\right) + \gamma\left(\frac{900/24}{T_a+273}\right)u_2\left(e_s - e_a\right)}{\Delta + \gamma\left(1 + 0.34u_2\right)} \tag{29}$$

Where Δ the slope of the vapour pressure curve at air temperature [$\mathrm{kPa\,C^{-1}}$], R_n the net daily radiation at the surface [$\mathrm{MJ\,m^{-2}\,hr^{-1}}$], G the soil heat flux [$\mathrm{MJ\,m^{-2}\,hr^{-1}}$], γ the psychometric constant [$\mathrm{kPa\,C^{-1}}$], T_a the hourly air temperature

[C], u_2 the hourly wind speed $[\mathrm{m\,s^{-1}}]$ at 2m above the land surface, e_s the hourly saturation vapor pressure [kPa], and e_a the actual hourly vapor pressure [kPA]. Following Singer et al. (2021), we converted RDRS 10m wind speed to 2m with a logarithmic wind profile. Continuing the assumption of a short grass reference crop, u_2 can be calculated as follows (Allen et al., 1998):

$$u_2 = u_z \left(\frac{4.87}{ln\,(67.8z - 5.42)} \right) \tag{30}$$

Where z [m] is the height at which observed mean (non-directional) windspeed is available. The short grass reference crop gives us an albedo value of 0.23 (Allen et al., 1998), and an emissivity of 0.95 (Stull, 2017, Table 2-4), which are used to estimate various radiation terms. Furthermore, RDRS provides air temperature and relative humidity at 1.5m above the surface. We consider this close enough to the required 2m that no conversion was applied.

S2.5.3 WorldClim

We calculate maps of mean monthly PET estimates using the method in Oudin et al. (2005) and monthly WorldClim data (Fick and Hijmans, 2017):

$$PE = \begin{cases} \frac{R_e}{\lambda \rho} \frac{T_a + 5}{100}, & \text{if } T_a + 5 > 0 \\ 0, & \text{otherwise} \end{cases} \tag{31}$$

Where R_e is extraterrestrial radiation $[\mathrm{MJ\,m^{-2}\,day^{-1}}]$, λ the latent heat of vaporization $[\mathrm{MJ\,kg^{-1}}]$, ρ the density of water $[\mathrm{kg\,m^{-3}}]$, T_a the mean daily air temperature $[^\circ\mathrm{C}]$, and PE the potential evapotranspiration in $[\mathrm{mm\,day^{-1}}]$. This formula is intended to be used with daily time steps, and results may be inaccurate.

S3 Data set size

One benefit of large-sample over large-scale data sets is their ease-of-use (in terms of disk space and computational effort). In terms of area, CAMELS-SPAT covers a fairly small area compared to the land extent of the North American continent (see Figure 2 in main paper). However, in terms of data set size our original selection of 1697 stations required approximately 20 TB of disk space at the end of data set development. This should be no concern in High Performance Computing environments, but such requirements (still) outstrip easily available storage on personal computing devices. Figure 2 shows an overview of how data set size changes as a function of area thresholds used to determine which basins are included in the data set. Disk space requirements are highly non-linear. Based on this diagram, we decided to exclude any basins with an area larger than $10^4\,\mathrm{km^2}$ from the final data set distribution. This reduces the number of gauges by 113, but halves the disk space requirement of the total data set. This analysis was performed without the benefits of aggregating yearly and monthly forcing netcdf files into single files (reducing file structure overhead costs): total data set size after initial attempts to optimize disk space sums to roughly 6TB.

S4 Potential evpotranspiration estimates

Figure 3 shows an overview of three variables from the four sources of meteorological data available in CAMELS-SPAT. The first column shows that long-term annual average precipitation is reasonably consistent between the four different data sources (Fig. 3a, d, g, j). There are certainly small scale differences, but the large-scale patterns and order of magnitudes match. This is less so for the potential evapotranspiration (PET) estimates shown in the second column. The large-scale patterns are mostly consistent between PET estimates from RDRS (Fig. 3b), Daymet (Fig. 3e) and WorldClim (Fig. 3h), and these three estimates in turn are similar to the large-scale patterns in Figure 1 in Singer et al. (2021). However, the order of magnitude between
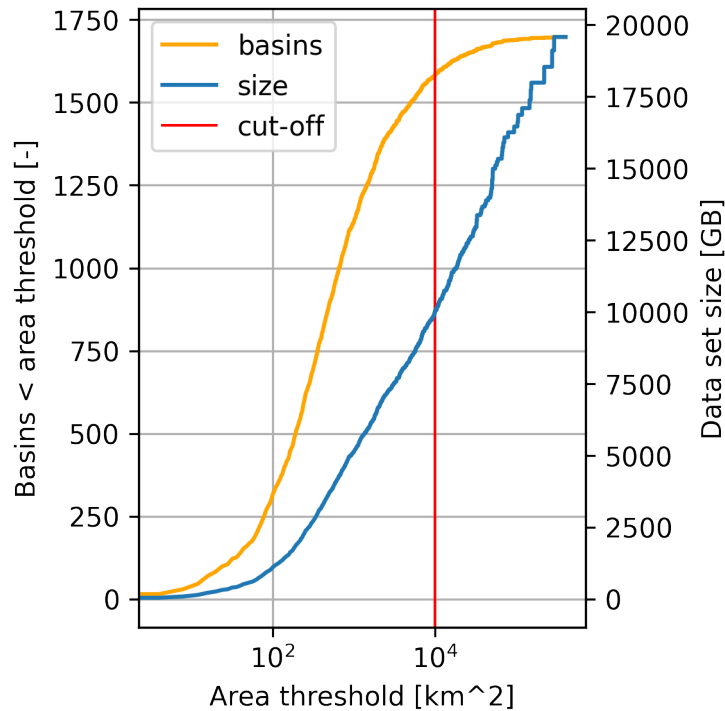
Figure S2. Overview of data set size (number of basins on the left y-axis, disk space requirements on the right y-axis) as a function of area thresholds below which basins are included in the counts. On the left, only the smallest basins are included and disk space requirements are consequentially small. On the right, all basins are included resulting in large disk space requirements.

the three data sources differs considerable. Visually, RDRS and Daymet estimates seem slightly high and close to the average
145 of the five data sources listed in Singer et al. (2021). WorldClim estimates are comparatively very low, possibly due to their
dependence on monthly mean temperature inputs. Clearly different are the PET estimates from ERA5 (Fig. 3k). The spatial
pattern shows a broad north-south gradient in PET values, but there are multiple locations where estimates seem unreasonably
low (the Appalachians for example stands out as having unexpectedly low PET totals). In fact, for 116 basins ERA5 predicts
negative annual PET amounts.

150 This has consequences for derived values such as aridity (third column). Aridity values for RDRS (Fig. 3c) are rather high.
Compared to Daymet, RDRS tends towards higher PET estimates and somewhat lower precipitation totals which exacerbates
aridity differences between both data sets. By comparison, WorldClim aridity estimates are rather low, mostly due to its much
lower PET estimates.

Due to the lack of uniformity in PET definitions and calculation methods (e.g. McMahon et al., 2013), it is difficult to say
155 which estimates are the most accurate. For time series, any expected systematic biases could be corrected before using the time
series as model input. Derived statistics with clear physical interpretations, such as aridity, are more difficult. A basin may be
classified as either water-limited or energy-limited solely as a consequence of the data and PET estimation method used, and
this may hinder classification and interpretation efforts. Possible ways around this may involve the use of multiple estimates of
PET-related attributes. We thus recommend caution when selecting and interpreting any PET estimates for further use.
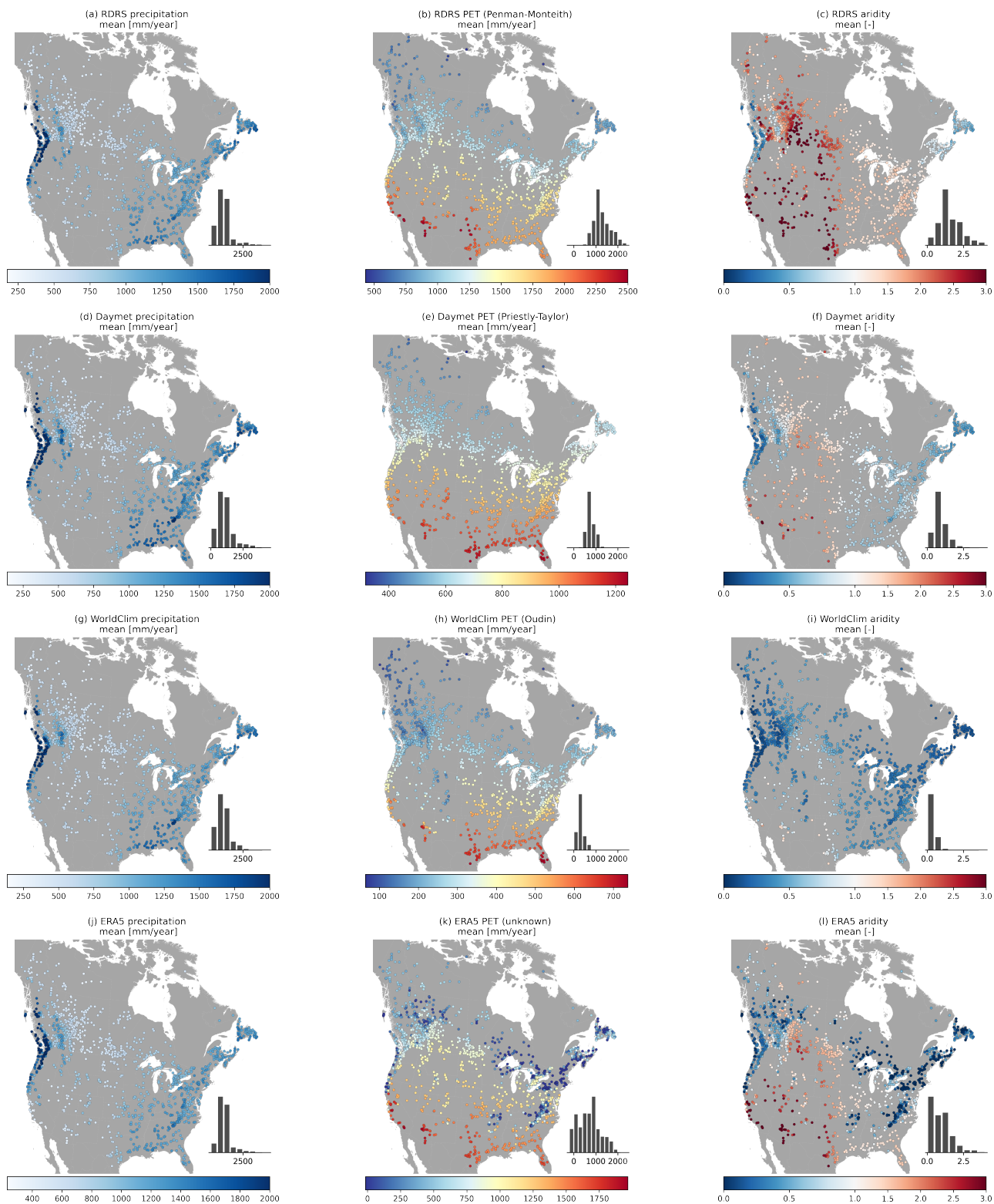
Figure S3. Overview of mean annual precipitation, potential evapotranspiration and aridity for the four sources of meteorological data used in CAMELS-SPAT.

References

Allen, R. G., Pereira, L. S., Raes, D., and Smith, M.: Crop evapotranspiration: guidelines for computing crop water requirements, no. 56 in FAO irrigation and drainage paper, Food and Agriculture Organization of the United Nations, Rome, ISBN 978-92-5-104219-9, 1998.

European Centre for Medium-range Weather Forecasting: ERA5: How to calculate wind speed and wind direction from u and v components of the wind? - Copernicus Knowledge Base - ECMWF Confluence Wiki, https://confluence.ecmwf.int/pages/viewpage.action?pageId=133262398, accessed: 2024-01-02, 2023.

Fick, S. E. and Hijmans, R. J.: WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas, International Journal of Climatology, 37, 4302–4315, https://doi.org/10.1002/joc.5086, 2017.

Knoben, W. J. M., Freer, J. E., Peel, M. C., Fowler, K. J. A., and Woods, R. A.: A Brief Analysis of Conceptual Model Structure Uncertainty Using 36 Models and 559 Catchments, Water Resources Research, 56, e2019WR025 975, https://doi.org/10.1029/2019WR025975, 2020.

McMahon, T. A., Peel, M. C., Lowe, L., Srikanthan, R., and McVicar, T. R.: Estimating actual, potential, reference crop and pan evaporation using standard meteorological data: a pragmatic synthesis, Hydrology and Earth System Sciences, 17, 1331–1363, https://doi.org/10.5194/hess-17-1331-2013, 2013.

Newman, A. J., Clark, M. P., Sampson, K., Wood, A., Hay, L. E., Bock, A., Viger, R. J., Blodgett, D., Brekke, L., Arnold, J. R., Hopson, T., and Duan, Q.: Development of a large-sample watershed-scale hydrometeorological data set for the contiguous USA: data set characteristics and assessment of regional variability in hydrologic model performance, Hydrology and Earth System Sciences, 19, 209–223, https://doi.org/10.5194/hess-19-209-2015, 2015.

Oudin, L., Hervieu, F., Michel, C., Perrin, C., Andréassian, V., Anctil, F., and Loumagne, C.: Which potential evapotranspiration input for a lumped rainfall–runoff model?, Journal of Hydrology, 303, 290–306, https://doi.org/10.1016/j.jhydrol.2004.08.026, 2005.

Priestley, C. H. B. and Taylor, R. J.: On the Assessment of Surface Heat Flux and Evaporation Using Large-Scale Parameters, Monthly Weather Review, 100, 81–92, https://doi.org/10.1175/1520-0493(1972)100<0081:OTAOSH>2.3.CO;2, 1972.

Singer, M. B., Asfaw, D. T., Rosolem, R., Cuthbert, M. O., Miralles, D. G., MacLeod, D., Quichimbo, E. A., and Michaelides, K.: Hourly potential evapotranspiration at 0.1° resolution for the global land surface from 1981-present, Scientific Data, 8, 224, https://doi.org/10.1038/s41597-021-01003-9, 2021.

Stull, R. B.: Practical meteorology: an algebra-based survey of atmospheric science, Dept. of Earth, Ocean & Atmospheric Sciences, University of British Columbia, Vancouver, BC, Canada, version 1.02b edn., ISBN 978-0-88865-283-6, oCLC: 1054636700, 2017.

Zotarelli, L., Dukes, M. D., Romero, C. C., Migliaccio, K. W., and Morgan, K. T.: Step by Step Calculation of the Penman-Monteith Evapotranspiration (FAO-56 Method), Institute of Food and Agricultural Sciences Extension. University of Florida., AE459, 1–10, https://docslib.org/doc/12004776/penman-monteith-evapotranspiration-fao-56-method-1-lincoln-zotarelli-michael-d, 2009.