Water Resources Research

RESEARCH ARTICLE
10.1029/2019WR026719

Key Points:
- Elasticities of runoff to precipitation are more consistent across basins than are runoff elasticities to potential evapotranspiration
- Complementarity in precipitation and potential evapotranspiration elasticities is better met in model- than observation-based estimates
- Observation-based PET elasticity estimates are more sensible when a complementary relationship constraint is imposed

Supporting Information:
- Supporting Information S1

Correspondence to:
D. P. Lettenmaier,
dlettenm@ucla.edu

Citation:
Xiao, M., Gao, M., Vogel, R. M., & Lettenmaier, D. P. (2020). Runoff and evapotranspiration elasticities in the Western United States: Are they consistent with Dooge's complementary relationship?. *Water Resources Research*, *56*, e2019WR026719. https://doi.org/10.1029/2019WR026719

Received 9 NOV 2019
Accepted 19 JUN 2020
Accepted article online 22 JUN 2020

Runoff and Evapotranspiration Elasticities in the Western United States: Are They Consistent With Dooge's Complementary Relationship?

Mu Xiao[1], Ming Gao[1], Richard M. Vogel[2], and Dennis P. Lettenmaier[1]

[1]Department of Geography, University of California, Los Angeles, CA, USA, [2]Department of Civil and Environmental Engineering, Tufts University, Medford, MA, USA

Abstract Many studies have examined how runoff (Q) responds to long-term changes in precipitation (P) and temperature (T), but the effects of potential evapotranspiration (PET) have received less attention. We examine observational data sets for P, T, and Q, along with PET estimated from observations, to determine the extent to which derived P and PET runoff elasticities (ε_P and ε_{PET}, the fractional changes in runoff associated with given fractional changes in precipitation and PET, respectively) meet Dooge's complementary relationship (under certain conditions, $\varepsilon_P + \varepsilon_{PET} = 1$). We apply three statistical methods and two hydrologic models to estimate ε_P and ε_{PET} in 84 headwater river basins in California, Oregon, and Washington. We find that while the estimates of ε_P are generally consistent across two statistical estimators and one model-based estimator, the estimates of ε_{PET} using the statistical methods differ considerably (generally, they are much more negative) from the model-based estimates, and some appear to be implausible. The model-based estimates show better conformance to the complementary relationship (and in the median across sites, they sum to close to 1.0). We explore several factors that might explain the failure of the observation-based estimators, including interaction between P and PET and nonclosure of the water budget at annual time scales.

1. Introduction

Streamflow, often the most readily accessible water source for human use, plays a particularly important role in agricultural and municipal water supply, industrial production, hydropower generation, and other beneficial uses of water. However, the hydrologic cycle is subject to change as climatic factors such as temperature, T, and precipitation, P, respond to a warming climate. These changes have been especially prominent in the Western United States (Karl et al., 2009). Ongoing warming can intensify drought risk and severity by increasing the probability of coincident anomalous temperature and precipitation events, even if mean precipitation and the likelihood of anomalously low-precipitation do not change (AghaKouchak et al., 2014; Diffenbaugh et al., 2014). Such changes would create challenges for water resource management, especially where water demand is high and where water resources are already heavily exploited (Zhang et al., 2014). Elevated temperature stresses aquatic ecosystems, increases the number and severity of wildfires and subsequent erosion, and makes forests more vulnerable to moisture stress and insect infestations (Cayan et al., 2010; Null et al., 2010). All of these factors motivate a better understanding of how streamflow will respond to changes in climate.

While many past studies have used hydrologic models to simulate streamflow under different climate change scenarios (e.g., Christensen et al., 2004; Hayhoe et al., 2006; Vicuna et al., 2007; Zhang et al., 2014, among many others) these approaches are somewhat unsatisfying due to the wide range of hydrologically important land surface variables in the scenarios (often produced by climate models). Furthermore, scenario analysis is inevitably accompanied by attendant complexities of downscaling from the relatively coarse spatial scale of climate models to the finer spatial scale at which hydrologic information is required by, for instance, water managers. This has motivated a parallel track, dating at least to work by Schaake (1990) and formalized in Jim Dooge's Horton Lecture (Dooge, 1992) that evaluates elasticities (defined as the fractional change in runoff Q, divided by fractional change in precipitation P or fractional change in potential evapotranspiration PET). The advantage of the elasticity concept (which is most easily applied to long-term annual averages, which avoids the complications of accounting for moisture storage changes in

10.1029/2019WR026719

the water balance equation) is its elegant simplicity. As applied to precipitation elasticities of runoff, the concept is straightforward, notwithstanding questions of estimator performance when applied to observations of precipitation and streamflow (see Sankarasubramanian et al., 2001). Inclusion of PET introduces the complexity that basin-scale PET must be estimated (e.g., from physical model output or weather model reanalysis products), and there are many methods and products of doing so (notwithstanding that direct observations are available at points). This has led to work that has investigated sensitivities (as contrasted with elasticities) of runoff to (surface air) temperature, which, in contrast to PET, is readily observable. Fu et al. (2007), for instance, expressed the runoff elasticity as a two-parameter function of P and T. However, approaches that focus on temperature rather than PET have their own drawbacks. One is that P and T usually change by different amounts at different times of year, so it is difficult to decompose the effects of their changes separately. Another, perhaps more important, shortcoming is that water balance changes in a river basin are governed by changes in precipitation and evaporative demand (and not temperature directly). On a physical basis, evaporative demand is driven by net radiation, vapor pressure deficit, wind speed, and temperature; the first two variables depend not only on temperature but also on other variables such as net solar radiation, albedo, and humidity. These attributes have been changing over time along with temperature, and they can have greater influence on evaporative demand than T (Milly & Dunne, 2020; Vano et al., 2012), whereas they are often neglected in runoff sensitivity studies. As an alternative, some studies have used Budyko-based methods (Budyko et al., 1974) to evaluate the effect of PET on runoff (Berghuijs et al., 2017; Donohue et al., 2011; Roderick & Farquhar, 2011).

Another complication pertains to hydrologic model-based approaches that use T as an input to estimate PET. It is tempting to use simple temperature-indexed PET algorithms such as Hamon (Hamon, 1961), Thornthwaite (Thornthwaite, 1948), Hargreaves (Hargreaves, 1975), and many others, which avoid the need for the additional variables (which often are not available from observations at the locations and time periods of interest). However, these methods tend to overestimate changes in the water balance associated with general warming and cause inaccuracies (generally overestimates) in resulting model-based runoff sensitivity estimates (Milly & Dunne, 2011). In short, it is the variation of evaporative demand (PET) that drives hydrologic change, whereas T is only an index embedded in PET (Vano & Lettenmaier, 2014), and it is important to understand not only how P change affects runoff in a warming climate but also how runoff responds to PET change.

Dooge (1992) demonstrated that under fairly general conditions, (annual) elasticities of Q with respect to P and PET sum to one (the complementary relationship). Although elasticities of Q with respect to P have been evaluated in many previous studies (Andreassian et al., 2016; Donohue et al., 2011; Sankarasubramanian et al., 2001; and many others), far fewer have considered PET elasticities, and the complementary relationship specifically, notwithstanding that the complementary relationship is important for evaluating climatic sensitivities of the water balance in a warming climate. Furthermore, although we are unaware of the use of the complementary relationship in this context to date, it also has potential as a diagnostic tool for coupled land-atmosphere models in terms of their ability to reproduce observed land surface dynamics. The elasticity approach and complementary relationship are elegant in their simplicity and offer some advantages as an alternative (or perhaps complement to) now widely used scenario analysis for assessment of hydrologic change. Given the above, the question we address here is: What are the runoff elasticities with respect to P and PET in headwater streams of the Pacific States (California, Oregon, and Washington), and does the complementary relationship hold when based on observations of P and commonly used estimates of PET? Below, we describe the elasticity concept and Dooge's complementary relationship in section 2. Section 3 reports the data set and methods we utilized in this study. Results and discussion are provided in section 4, with conclusions in section 5.

2. Background

The runoff elasticities to P and PET (all quantities are long-term means hereafter unless indicated otherwise) were formulated analytically by Dooge (1992). He showed, under two conditions, that these elasticities sum to one. The first condition is the long-term mean water balance, runoff = precipitation − evapotranspiration (with storage change assumed to be negligible). The second is the Budyko hypothesis, which has the form

10.1029/2019WR026719

$$\frac{\mathrm{AET}}{\mathrm{PET}} = \phi\left(\frac{\mathrm{P}}{\mathrm{PET}}\right) \tag{1}$$

where AET is actual evapotranspiration, the ratio of P to PET is termed the wetness index, and $ø$ is a homogenous function dependent on the wetness index. The Budyko hypothesis has been explored widely due to the increased focus on the effects of global change on water resources and has now been verified for thousands of natural watersheds around the globe (Padrón et al., 2017; Sankarasubramanian & Vogel, 2002; Williams et al., 2012). Sankarasubramanian et al. (2020) and Wang, Wang, et al. (2016) review the application of the Budyko hypothesis in hydroclimatology. The Budyko hypothesis states that over the long term, the ratio of AET to PET can be expressed as a function of the humidity index. Under these two conditions, Dooge (1992) and Kuhnel et al. (1991) showed that the following runoff elasticity equation holds:

$$\frac{\Delta \mathrm{Q}}{\mathrm{Q}} = \varepsilon_P \frac{\Delta \mathrm{P}}{\mathrm{P}} + (1 - \varepsilon_P)\frac{\Delta \mathrm{PET}}{\mathrm{PET}} \tag{2}$$

where Q is runoff and ε_P is the precipitation elasticity of runoff (Schaake & Chunzhen, 1989). Using the same nomenclature, $1 - \varepsilon_P$ is the PET elasticity (ε_{PET}) of runoff, and ε_P and ε_{PET} add to unity, over the long term (which is the complementary relationship). Dooge (1992) calculated the elasticities of runoff for different values of the humidity index based on different empirical expressions of the Budyko hypothesis and showed that higher humidity ratios, indicating wetter climates, lead to smaller elasticities (less positive ε_P and less negative ε_{PET}), with a limiting condition where ε_P approaches unity as the humidity index approaches infinity. This pattern has been evaluated and confirmed in many other studies (e.g., Chiew, 2006; Sankarasubramanian et al., 2001; Wang & He, 2017; Zheng et al., 2009).

Roderick and Farquhar (2011) and Yang and Yang (2011) employed an analytical approach based on the Budyko hypothesis to calculate ε_P and ε_{PET}. Although approaches based on Equation 2 can conveniently estimate the elasticity of runoff for different climatic regions once a Budyko formulation is determined, the results have three inherent sources of uncertainty. First, the key result that ε_P and ε_{PET} sum to unity depends on the Budyko hypothesis as the starting point, so the accuracy of the result depends on how well the Budyko formulation represents reality. Second, the results are highly variable depending on the specific formulation of the Budyko hypothesis (that is, the form of $\varnothing(\frac{P}{PET})$ in Equation 1). A third source of uncertainty arises from the assumption that in the long term, the water balance closes, that is, there is no long-term storage change. We explore the implications of this assumption further below. These sources of uncertainty can be considerable, because while Equations 1 and 2 are good first approximations to the long-term hydroclimatology of a river basin, it is now well known that the evapotranspiration ratio (AET/PET) in Equation 1 is a function of a number of variables in addition to the aridity index (P/PET), including the soil moisture holding capacity (Sankarasubramanian & Vogel, 2002), the number of precipitation events per year, and seasonality parameters (Milly, 1994). For a review of the myriad of approaches for estimation of ε_P and ε_{PET}, see Table 1 in Wang, Zou, et al. (2016).

Equation 2 depends on the long-term average values of Q, P, and PET. Most studies that have estimated ε_P and ε_{PET} from observations (as noted above, the predominant focus has been on ε_P) have used annual data. For example, Sankarasubramanian et al. (2001) and Zheng et al. (2009) tested several estimators of ε_P using annual Q, P, and PET (over the conterminous United States and the Yellow River basin, respectively), and Risbey and Entekhabi (1996) used annual Q, P, and T data to estimate ε_P and the streamflow sensitivity to temperature in the Sacramento River basin. Approaches based on annual data embed an implicit assumption that the annual quantities for each variable in each year are a surrogate for their long-term means, implying that different years are independent and that there is minimum carryover storage from one year to the next. This of course is not true for runoff, which typically has some interannual carryover effects (the assumption arguably is more defensible for P and PET). Furthermore, especially during extremely high and low precipitation and runoff years, carryover storage can be substantial, depending on the specifics of a given river basin with respect to runoff generation and the magnitude of effective subsurface storage capacity.

3. Data and Methods

We estimated ε_P and ε_{PET} for a set of headwater river basins along the U.S. West Coast using two general approaches. The first is data based, with various estimators applied to annual Q, P, and PET. The second is model based, in which we perturbed P and PET forcings to a hydrologic model and calculated the model-predicted changes in simulated runoff. The statistical methods require streamflow data for rivers that are minimally affected by anthropogenic activities upstream (e.g., reservoir impoundments and/or diversions) as well as climate data including precipitation and variables required to calculate PET. We discuss below the data sources for both methods.

3.1. Streamflow Data and Gauge Selection

We retrieved average daily streamflow data from USGS Water Data (U.S. Geological Survey, 2016) for the Nation (http://waterdata.usgs.gov/nwis/). The stream gauges we used are a subset of the GAGES-II (Geospatial Attributes of Gages for Evaluating Streamflow) data set (https://water.usgs.gov/GIS/metadata/usgswrd/XML/g agesII_Sept2011.xml), which includes 2057 reference sites across the conterminous United States that were selected (by USGS) to be minimally disturbed by human influences (Falcone et al., 2010). We selected gauges in the states of California, Oregon, and Washington from the reference sites based on the following criteria and procedures:

1. We identified and removed gauges with any regulation or diversions upstream based on the USGS remarks files;
2. We required all gauges to have at least 50 years of continuous data during the period 1920–2015. We arbitrarily defined a year of continuous data (we used water years in all cases) as having no more than 30 days of missing data nor more than 15 days in any continuous gap. Any stations that did not meet these criteria were removed from our list of candidates;
3. We excluded gauges where the streamflow data have visually unusual patterns or many discontinuities and zero values. Such cases included, for instance, sudden changes in streamflow patterns that could not be attributed to natural factors. We note that there is considerable (although not complete) overlap in our criteria and station list with those used by Cooper et al. (2018);
4. For stations that passed the screening criteria above, we summed the (water year) daily flows to annual.

The screening process resulted in 84 gauges, 24 of which are in California, 23 in Oregon, and 37 in Washington (Figure 1). The drainage areas associated with these gauges mostly are relatively small; the largest is 2,463 km^2, and the smallest is 21 km^2. About half of the river basins have drainage areas smaller than 250 km^2.

3.2. Climate Data

The climate data we used include gridded precipitation, temperature, net radiation, vapor pressure deficit, wind speed, and atmospheric pressure. Precipitation, temperature, wind speed, and atmospheric pressure are from the University of Washington's Surface Water Monitor (SWM; Wood & Lettenmaier, 2006) data set archived at UCLA (http://www.hydro.ucla.edu/SurfaceWaterGroup/monitors.php). P and T values were gridded directly from observations using the same methods as are used to produce the SWM. Net radiation and vapor pressure deficit are output of the variable infiltration capacity (VIC) macroscale hydrology model (Liang et al., 1994; Mao et al., 2015; Xiao et al., 2016) using the gridded P and T forcings and other forcings derived using methods described by Bohn et al. (2013). All data are at 1/16-degree spatial resolution; precipitation and atmospheric pressure are at 3- hourly time step, and other variables are daily.

We found that the runoff in a number of basins was underestimated, as indicated by a comparison of mean annual runoff and mean annual precipitation (12 basins, for instance, had apparent Q > P). In many of these cases, the mean seasonal cycle of runoff was too small, but the seasonal cycle plausibly matched observations. The likely reason is underestimation of precipitation in the SWM data set. Many headwater catchments in the Western United States (which are high enough in their drainages to have minimal water management effects) also have few precipitation gauges and hence are prone to errors in precipitation. Rather than removing these basins and losing their inherent information content, we upscaled the precipitation (in the above-mentioned 12 basins) to ensure that P > Q by factors that led to the highest Kling-Gupta efficiency (KGE; Gupta et al., 2009). The average annual KGE over all 84 basins is 0.59 (section 3.5 discusses the scaling approach and Figure S1 in the supporting information summarizes the scaling factors). We note that the issue is mostly one of scaling, as the catchments we analyzed are essentially all characterized by strongly winter-dominant precipitation (and accompanying snow accumulation in many cases). Winter

10.1029/2019WR026719

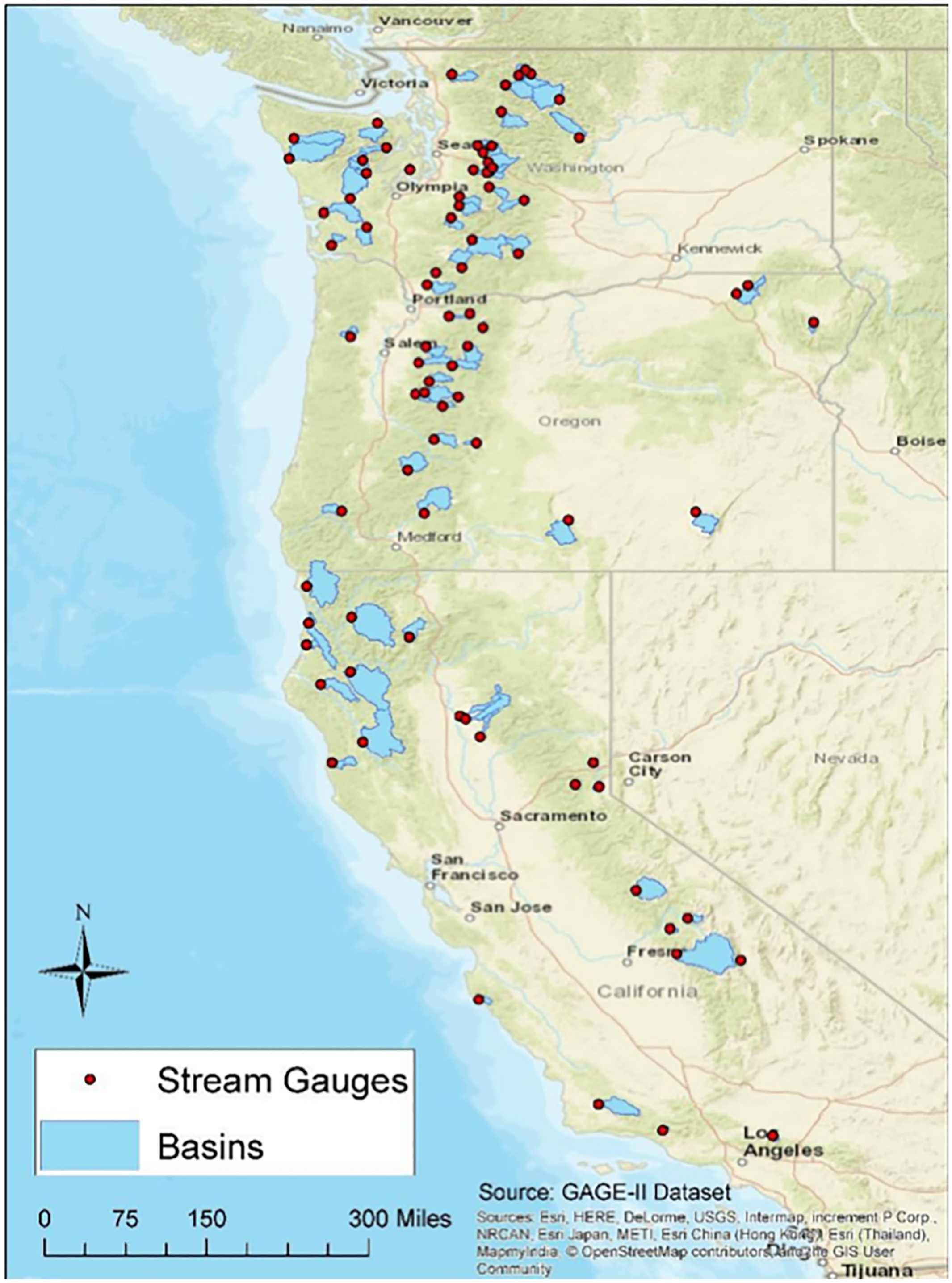

Figure 1. Locations of the 84 stream gauges.

precipitation is mostly frontal and hence has large spatial signatures, and hence relatively high spatial correlations, and therefore is amenable to rescaling to better match catchment conditions.

3.3. PET Estimation

We estimated PET as Penman-Monteith (Penman, 1948) reference ET (ET_0), which uses the temperature, net radiation, vapor pressure deficit, and wind speed as inputs, following Allen et al. (1998). Wind speed is constant (although seasonally varying) for each 1/16° forcing grid cell (as section 3.2 describes). Livneh

10.1029/2019WR026719

et al. (2013) evaluated the implications of this assumption, which are modest when averaged over long time periods. In addition, we tested four alternative (to ET_0) PET methods to evaluate the effects of the choice of the PET estimation method on our results. These include formula from Oudin et al. (2005), Priestley & Taylor, 1972, the Yang et al. (2019) equation, and net radiation scaled by the latent heat of vaporization (λ). The Yang formula considers the effect of increased CO_2 on the surface resistance in ET_0. Oudin's formula is based on temperature and clear-sky solar radiation. The Priestley and Taylor (1972) equation, like Penman-Monteith, is energy based and can be treated a simplified version of Penman-Monteith to the extent that does not include vapor pressure effects explicitly. The last approach (which multiplies net radiation by λ) is a bounding estimate of PET used in Budyko's original work (Budyko et al., 1974); it is entirely energy based.

3.4. Statistical Methods

We used two observation-based multivariate statistical methods to estimate ε_P and ε_{PET}: bivariate ordinary least squares (OLS) regression and bivariate generalized least squares (GLS) regression. We also tested the nonparametric univariate estimator employed and evaluated by Sankarasubramanian et al. (2001), but similar to Andreassian et al. (2016) and others, we found that it did not perform as well as bivariate estimators, and thus we did not further consider it.

Numerous statistical approaches to estimation of ε_P and ε_{PET} are summarized in Table 1 of Wang, Wang, et al. (2016) including the bivariate OLS and GLS regression estimators used by Andreassian et al. (2016) and Konapala and Mishra (2016). Andreassian et al. (2016) showed that both the OLS and GLS estimators performed well, and this is the basis for our choice. However, we note that even sensible use of multivariate statistical methods for sensitivity analysis can occasionally lead to nonsensical results, as was clearly shown by Wallis (1965), and which our results show as well.

The OLS and GLS estimators are based on a bivariate model of the form

$$\frac{\Delta \mathrm{Q}}{\overline{\mathrm{Q}}} = \varepsilon_P \frac{\Delta \mathrm{P}}{\overline{\mathrm{P}}} + \varepsilon_{PET} \frac{\Delta \mathrm{PET}}{\overline{\mathrm{PET}}} + \omega \tag{3}$$

where ω is a residual.

The main differences between OLS and GLS regression are that OLS assumes homoscedasticity and no spatial correlation among the annual runoff values, whereas GLS attempts to account for both. OLS assumes homoscedasticity, meaning that the variance of the error term (ω) is constant, while GLS allows the error term to have unequal variance. OLS also assumes the dependent variable is not a random variable, whereas GLS accounts for the spatial correlation of the dependent variable. For GLS, we followed the implementation outlined in Andreassian et al. (2016). We also implemented a constrained version of OLS, where the estimates of ε_P and ε_{PET} are forced to sum to 1.0, to ensure reproduction of the complementary relationship (see section 4 for more details).

For both estimators, we conducted field significance tests (Livezey & Chen, 1983) to investigate whether the null hypothesis that the sum of ε_P and ε_{PET} was equal to 1.0 was rejected at the 5% significance level. Because the degrees of freedom (number of independent sites) cannot be clearly assessed, we used a Monte-Carlo approach. First, we generated 1,000 sequences of annual data (P, PET, and Q) that have the same statistical properties and spatial correlations as the observation data. Then we applied the OLS and GLS methods to estimate the elasticities, applied the statistical test, and counted the number of rejections. By constructing the distribution of the number of rejections of the 1,000 tests, we determined the critical value (the 5% exceedance value). We constructed 95% confidence intervals for the sum of ε_P and ε_{PET} estimated at each site and counted the number of rejections.

3.5. Hydrologic Model Implementation

We implemented the USGS Thornthwaite Water Balance Model (McCabe & Markstrom, 2007; McCabe & Wolock, 1999) (which we refer to hereafter as the USGS model). The USGS model is a simplified bucket model but represents the dominant processes in runoff generation, albeit at a monthly time step (we also tested a more sophisticated, widely-used hydrological model [Sacramento, SAC; Burnash et al., 1973] with results that were quite similar to those obtained with the USGS model; therefore, we utilized the USGS monthly model as our primary tool for model-based analysis). We estimated runoff elasticities for the

10.1029/2019WR026719

USGS model by uniformly changing P or PET, calculating the change in the simulated runoff, and then determining the elasticity following Vano and Lettenmaier (2014). For example, if P is increased by 1%, and this leads to an average 2% increase in runoff, then our estimate of ε_P is 2.0.

The USGS model has six parameters with monthly total precipitation, monthly average PET, and monthly average temperature as inputs. The model parameters are (1) a runoff factor, which determines the percentage of "surplus water" that contributes to runoff at each (monthly) time step, the rest adds to the next month's "surplus water," (2) direct runoff factor, which represents the percentage of rainfall in a given month that goes directly to runoff, (3) rain temperature threshold, above which all precipitation occurs in the form of rainfall, (4) snow temperature threshold, below which all precipitation occurs in the form of snowfall, (5) maximum snowmelt rate, which determines the maximum percentage of snow storage that can melt during a month, and (6) soil moisture storage capacity. In our implementation, we use $ET_{0,}$ as a surrogate for PET, as described in section 3.2.

We calibrated the USGS model for each basin manually by testing alternative combinations of the parameters. We tested soil moisture storage capacities ranging from 50 to 200 mm at 50 mm intervals, runoff factor and maximum snowmelt rates ranging from 0.5 to 0.8 at 0.1 intervals, and several sets of temperature thresholds between −5°C and 5°C. We fixed the direct runoff factor at the default value (5%). We found that the parameters to which the model predictions were most sensitive were soil moisture storage capacity and snow temperature thresholds. The soil moisture storage capacity mainly affects low flows since precipitation has to fill the soil storage before runoff can be generated. Temperature thresholds affect the runoff seasonal cycle by determining how much and when snow melts. We selected parameter sets based on the KGE of the resulting model streamflow predictions. We upscaled the precipitation in the basins where long-term Q was smaller than P (as mentioned section 3.2) by factors that led to the highest KGE. The average annual KGE over all 84 basins is 0.59, ranging from 0.30 to 0.89.

4. Results and Discussion

Figure 2 shows P and PET elasticities estimated by OLS with the different PET algorithms described in section 3.3. While there are some differences in the elasticities calculated using different choices of PET, the major patterns and ranges of variation in ε_P and ε_{PET} are fairly similar to those based on ET_0. In particular, using Priestley-Taylor in comparison with scaled net radiation produces very similar results, because Priestley-Taylor is (nearly) a multiple of net radiation. Similarly, the Yang estimator is a variant of ET_0 and produces elasticity results that are quite similar. The variances among elasticities based on the USGS model with different choices of PET show the same patterns (Figure S3). Therefore, we only used ET_0 for our primary analysis.

Precipitation elasticities estimated by the OLS- and GLS-based methods generally were similar (Figure 3a), with medians (across sites) slightly larger than 1.2. ε_P values estimated using the USGS model were larger across their distribution relative to the statistical estimators, although the median (slightly less than 1.3) was only modestly different than for the statistical estimators. The estimates of ε_{PET} by the OLS- and GLS-based methods were quite similar across their distributions, and the medians differed by less than 0.1. Both GLS and OLS estimated implausible positive values of ε_{PET} for many catchments, whereas all values were negative for the USGS model (Figure 3b), with median around −0.29. All methods produced median elasticity sums that were slightly larger than 1.0. However, the OLS and GLS methods produced a few large positive complementary sums (>2.0). The interquartile range of the complementary sum of ε_P and ε_{PET} for OLS and GLS was approximately [0.9–1.6] and the interquartile range for the USGS model was about [1.1–1.2]. Thus, the USGS model was better able to reproduce the complementary relationship than the statistical methods (although, as noted below, we suspect that this may be partly by construct, i.e., the model preserves [its own] long-term water balance, and its evapotranspiration formulation follows a Budyko-like form).

The USGS model precipitation elasticities are all larger than 1.0. PET elasticities are all negative, and the sums of ε_P and ε_{PET} are all equal or greater than 1.0. For OLS and GLS, the medians of the sums of ε_P and ε_{PET} across sites are slightly larger than 1.0 (but with larger ranges than those estimated using the USGS model). The USGS model results generally have sums that are closer to 1.0 than the statistical methods.

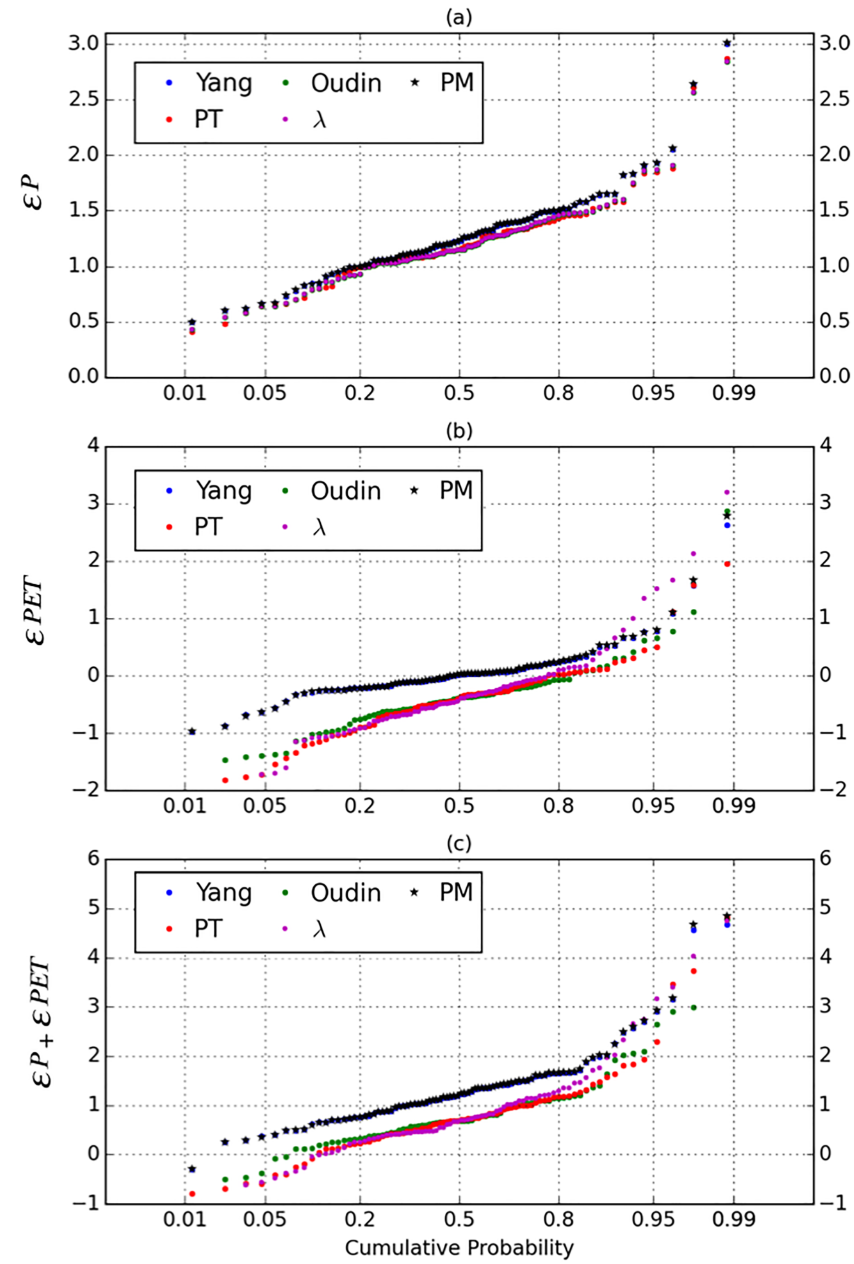

Figure 2. Comparison of (a) ε_P, (b) ε_{PET}, and (c) $\varepsilon_P + \varepsilon_{PET}$ based on the OLS estimator. The PETs are estimated from different PET methods as described in section 3.3. PT stands for Priestley-Taylor, and PM is the Penman-Monteith reference ET (ET_0). λ stands for the bounding estimate of PET based on net radiation.

In terms of field significance, the number of rejections (of the hypothesis that $\varepsilon_P + \varepsilon_{PET} = 1.0$) is 25 out of 84 basins (30%) for OLS and 23 (27%) for GLS. Thus, the OLS and GLS results indicate that complementary relationship may hold for approximately 70% of the sites. According to our Monte Carlo approach, the critical value of the field significance test was 12 for OLS and 13 for GLS. Therefore, notwithstanding the above, we reject the overall null hypothesis that the sum of elasticities at *all* sites is 1.0 using an overall 5% field significance level. On the basis of these hypothesis tests, there is considerable evidence that the statistical methods are unable to reproduce the complementary relationship across all sites, whereas, even though we were unable to perform an analogous field hypothesis test for the USGS results, Figures 3c and 3f indicate excellent reproduction of the complementary relationship by the model.

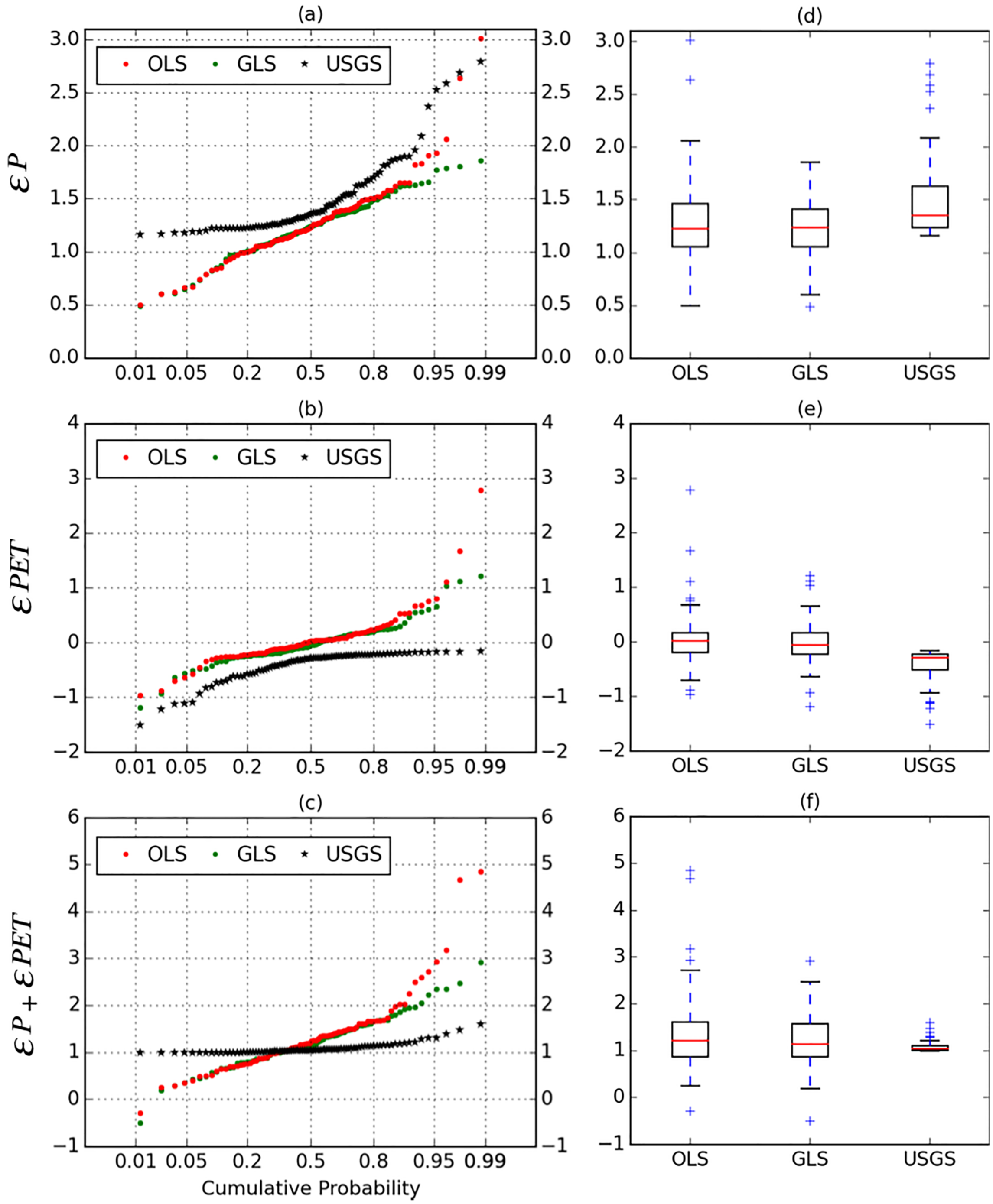

Figure 3. ε_P (a, d), ε_{PET} (b, e), and the sum (c, f), estimated by using the two statistical methods and the USGS model. The right column shows the results as box-plots.

While ε_P values computed using the statistical and model-based methods are roughly similar (at least in their medians across sites), the characteristics of ε_{PET} estimated using the statistical methods are peculiar. About 30% (26) of the basins have ε_{PET} estimated using OLS and GLS that are both positive, which is both counter-intuitive and infeasible. One fundamental difference between hydrologic model-based and statistical estimates is that the change of climate variables is controllable in the models but not in observations. Another fundamental difference between the hydrologic model-based and statistical estimates is that the hydrologic model has built into it a Budyko-like relationship, whereas the statistical models do not. We

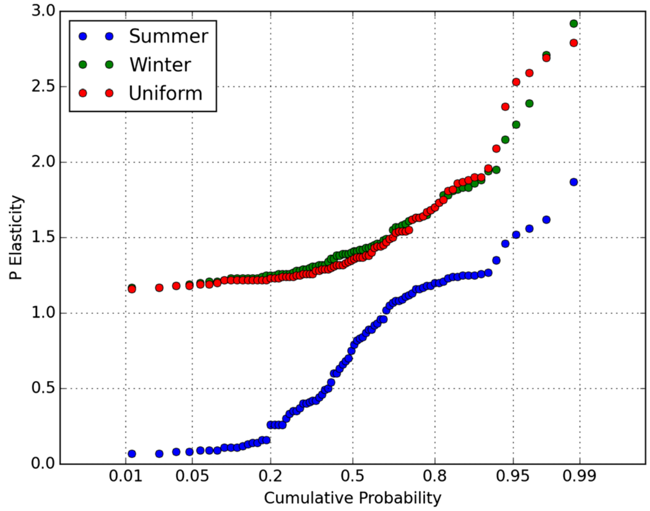

Figure 4. ε_P estimated using the USGS model. "Uniform" results from a 1% change of precipitation uniformly distributed across the year, "Winter" results from a precipitation change applied only for November–January, and "Summer" results from a 1% change allocated only to June–August.

discuss both of these fundamental differences between the statistical and hydrological model based approaches below.

In the models, we changed P and PET uniformly and independently, whereas when implementing statistical methods, both the time of change (e.g., if all of an anomaly in P occurs in 1 month or is spread throughout the year) and the change in magnitude (e.g., 1% increase or 10%) varies from year to year and are out of our control. In the analysis of climate elasticity of runoff by Dooge (1992) and other studies based on an analytical approach and the Budyko hypothesis, the climate is usually assumed to change uniformly from one long-term mean to another; however, some physical constraints exist. For instance, if PET is independent of precipitation, PET elasticity should always be negative since more water will be lost with higher evaporative demand. However, if a 1% total increase in PET is distributed as a 2% increase in a season when surface water is scarce and a 1% decrease in a season when surface water is abundant, the net result could be increased runoff and apparently positive ε_{PET}. By allocating the 1% change in P to the summer months in the USGS model, Figure 4 demonstrates the enormous impact that such a variation from a uniform to a seasonal change has on the resulting precipitation elasticities. With respect to the large changes in Figure 4, we note that allocation of all of the change in precipitation entirely to the summer season (when in reality, only a small fraction of the annual precipitation occurs during that season in the Pacific Coastal region) is highly implausible. The figure does suggest though that even a much more modest seasonal reallocation could make a substantial difference to $\varepsilon_{P.}$

Another complication is that in observations, P and PET interact and thus do not change independently. This means that the statistical methods need to be able to separate the complex and interacting influences of P and PET on runoff. Two problems arise. One is that the independent variables P and PET exhibit weak multicollinearity (however, this is not expected to be a problem for estimation of model coefficients using GLS and OLS regression as shown recently by Kroll & Song, 2013). The other is that annual PET variations (as a fraction of, say, the mean) are usually smaller than for P, and P is the main driving factor in runoff generation with greater influence than PET. This complicates the task of statistical methods to segregate the two effects, and it may partially explain why estimates of ε_P by different methods are similar, whereas estimates of ε_{PET} by statistical methods are, in some cases, implausible.

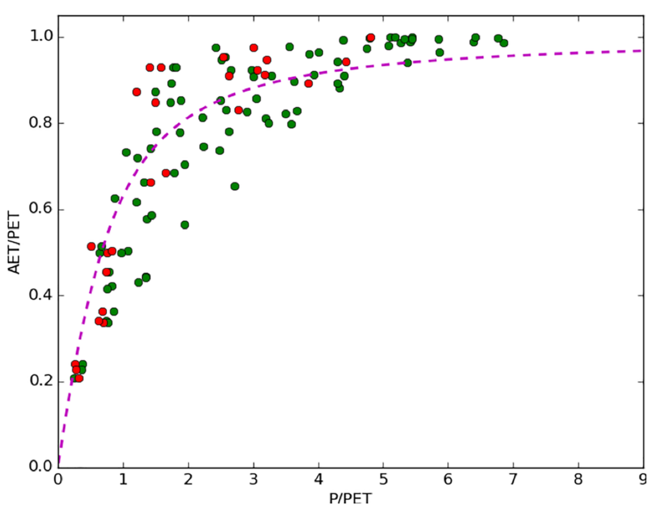

Figure 5. AET/PET versus P/PET for all 84 basins. AET is taken from the USGS model; PET is Penman-Monteith ET_0. The red dots are the basins where the null hypothesis of $\varepsilon_P + \varepsilon_{PET} = 1$ is rejected by the OLS and GLS statistical methods. The dashed line is the empirical Budyko line estimated by Equation 1 in Abatzoglou and Ficklin (2017), where the free parameter is set to 2.2.

Regarding the extent to which estimates of ε_P and ε_{PET} are complementary, our results show better conformance by hydrologic model-based than observation-based estimates, and the analysis above suggests that one reason has to do with our use of observation-based estimates of ε_{PET}. Of course, another reason is that the USGS model (like most physically based hydrologic models) has a Budkyo-like hypothesis built into its model structure. Thus an obvious question is whether the lack of closure of the complementary hypothesis concerning ε_P and ε_{PET} is due to problems with the modeling approach (i.e., estimators) or with the hypothesis.

As we note in section 2, complementarity comes about as a result of closure of the long-term water balance (in terms of the means of Q, P, and ET) and the Budyko hypothesis (in its general, rather than any specific, form). We note that because hydrologic models (including the USGS model) balance water by construct, and generally have Budyko-like behavior in their evapotranspiration parameterizations, it should not be surprising that the USGS model results generally reproduce the complementary relationship. There are no such physical constraints on the statistical methods, so we further explored the extent to which the long-term

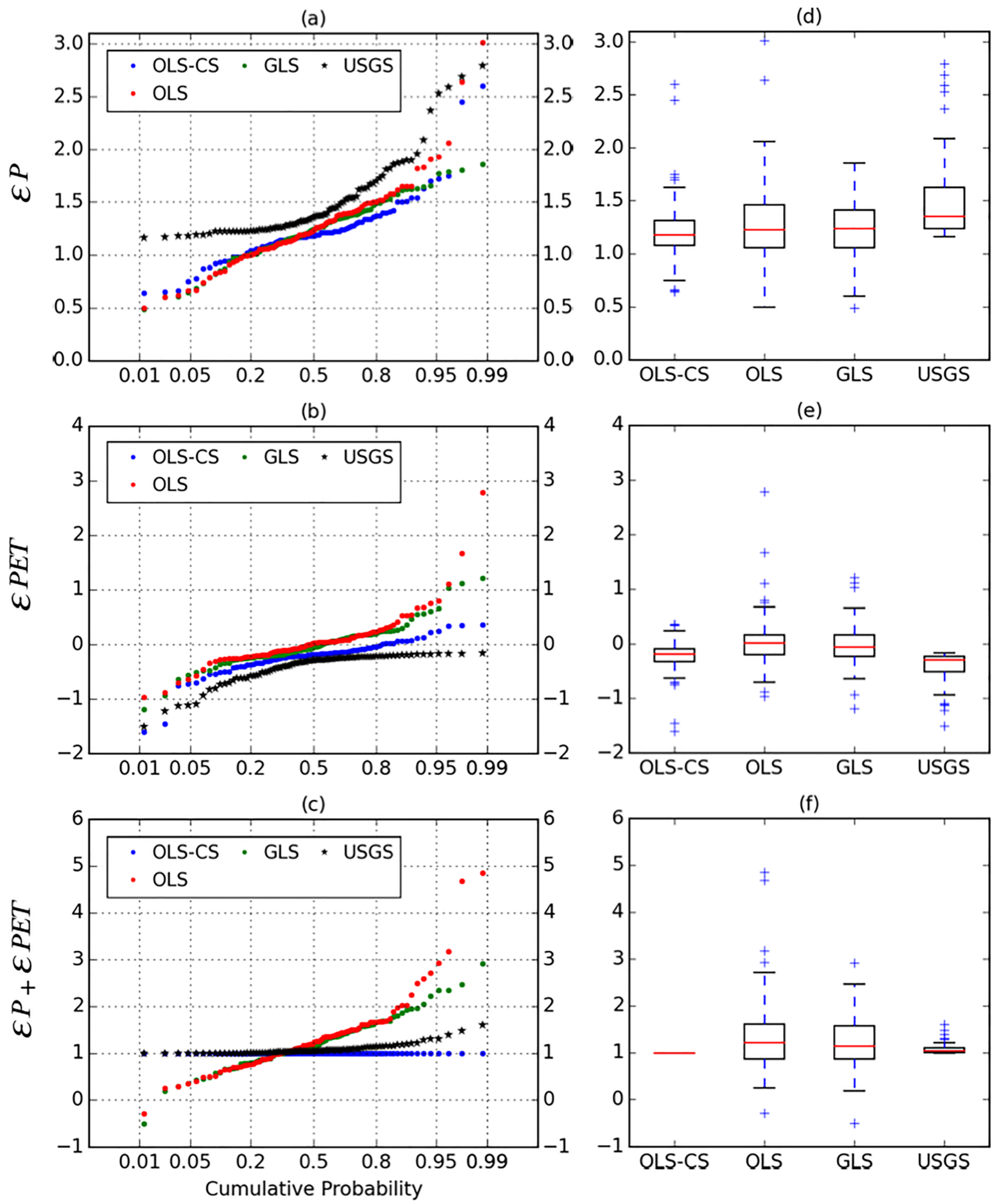

Figure 6. ε_P (a), ε_{PET} (b), and the sum (c) estimate by constrained linear regression method (OLS-CS). The other results are identical as in Figure 3 and shown for comparison here.

water balance is reproduced by the observations. Here we evaluate how ΔP/P, ΔPET/PET, and ΔQ/Q correlated with ΔStorage/Storage by calculating the corresponding R^2. Out of the total 84 basins, only eight basins have one or more variables highly correlated to storage changes (defined as at least one of the correlations greater than 0.3). This implies the influence of storage changes at the interannual time scale should not change our main findings and conclusions.

The degree to which the Budyko hypothesis holds for the 84 basins is evaluated in Figure 5, which illustrates the relationship between AET/PET and P/PET (where AET is the output of USGS model). Figure 5 suggests that there does seem to be a Budyko-like form when taken across all 84 river basins. Red and green circles in Figure 5 indicate those basins in which the complementary hypothesis was rejected or not, respectively, when using the OLS and GLS estimators. The degree to which the complementary hypothesis is reproduced by the statistical estimators does not appear to be related to either the degree to which the Budyko hypothesis holds or the hydroclimatology of the basins. Figure 5 illustrates a very broad range of hydroclimatologic regimes, because according to the climatic classification system introduced by Ponce et al. (2000, Table 1), the values of P/PET reported in Figure 5 range from approximately [0.2, 7.2] corresponding to hydroclimatic conditions ranging from arid and semiarid to subhumid and even humid. It is important to note that neither AET nor PET are observed values, which might compromise inferences from Figure 5.

The multivariate statistical estimators perform poorly compared to the physical-based model approach in terms of their ability to reproduce the complementary relationship as was shown in Figure 3c. This result should not be surprising given the warnings of Wallis (1965). If additional constraints are applied to the statistical methods, the complementary relationship can be well produced, and naturally, both the P and PET elasticities will be affected. In order to explore possible improvements of the statistical method, we constrained $\varepsilon_P + \varepsilon_{PET} = 1$ in the OLS method, which we denote OLS-CS. The resulting distribution of the elasticities is shown in Figure 6. Distributions from other methods in Figure 3 are also plotted for comparison. The ranges of ε_{PET} are generally similar to the USGS model results and importantly are more realistic than either the unconstrained OLS or GLS estimators. However, even when constrained to reproduce the complementary relationship, the OLS-CS method still produces some (albeit few) positive values of ε_{PET}. Importantly, estimates of both ε_P and ε_{PET} obtained from the constrained OLS-CS method are much less variable than those derived from the unconstrained OLS and GLS methods.

5. Conclusions

We applied three statistical methods and the USGS hydrologic model to estimate P and PET elasticities of runoff for 84 basins in California, Oregon, and Washington and evaluated the degree to which those methods are able to reproduce the complementary relationship. We conclude that

1. Complementarity is generally observed (at least in the central tendency of the distribution across the 84 basins) in the hydrologic model-based estimates, but only for a portion of the observation-based estimates based on the statistical methods.
2. The estimates of ε_P using different methods are generally consistent. Deviations from complementarity are mostly attributable to the ability to estimate ε_{PET} and in particular to the counterintuitive and highly positive estimates for some of the basins. The problem appears to be a combination of (a) relatively smaller scales of variation in interannual variability of PET, which is a reflection of PET being a weaker factor in runoff generation than P, and (b) correlation between P and PET (both with an understanding that for most of the basins, P is winter dominant and PET is summer dominant).
3. Hydrologic model-based relationships between AET/PET and P/PET displayed a typical Budyko form across a very broad range of hydroclimatic conditions ranging from arid to humid conditions. Although the OLS and GLS statistical methods led to strong deviations from the complementary relationship at some basins, we could not discern any significant departures from the Budyko hypothesis for those basins.
4. Estimates of P and PET elasticities derived from multivariate statistical methods such as the OLS and GLS approaches often led to questionable results, especially when compared to the results based on a hydrologic model. This result is consistent with the warnings given by Wallis (1965). However, our results indicate that the statistical approaches may be improved considerably by including the reproduction of the complementary relationship as a constraint in the fitting process. Nonetheless, the most realistic estimates were generally obtained from a hydrologic model, which, as we note above, has a general structure that meets the two key assumptions underlying complementarity—water balance closure and a Budyko-like ET parameterization.

Data Availability Statement

All the data used in this study are archived at: https://doi.org/10.6084/m9.figshare.10278089.

References

Abatzoglou, J. T., & Ficklin, D. L. (2017). Climatic and physiographic controls of spatial variability in surface water balance over the contiguous United States using the Budyko relationship. *Water Resources Research*, *53*, 7630–7643. https://doi.org/10.1002/2017WR020843

AghaKouchak, A., Cheng, L., Mazdiyasni, O., & Farahmand, A. (2014). Global warming and changes in risk of concurrent climate extremes: Insights from the 2014 California drought. *Geophysical Research Letters*, *41*, 8847–8852. https://doi.org/10.1002/2014GL062308

Allen, R. G., Pereira, L. S., Raes, D., & Smith, M. (1998). *Crop evapotranspiration—Guidelines for computing crop water requirements*, *FAO Irrigation & Drainage Paper* (Vol. 56). Rome: FAO.

Andreassian, V., Coron, L., Lerat, J., & Moine, N. L. (2016). Climate elasticity of streamflow revisited—An elasticity index based on long-term hydrometeorological records. *Hydrology and Earth System Sciences*, *20*(11), 4503–4524. https://doi.org/10.5194/hess-20-4503-2016

Berghuijs, W. R., Larsen, J. R., van Emmerik, T. H. M., & Woods, R. A. (2017). A global assessment of runoff sensitivity to changes in precipitation, potential evaporation, and other factors. *Water Resources Research*, *53*, 8475–8486. https://doi.org/10.1002/2017WR021593

Bohn, T. J., Livneh, B., Oyler, J. W., Running, S. W., Nijssen, B., & Lettenmaier, D. P. (2013). Global evaluation of MTCLIM and related algorithms for forcing of ecological and hydrological models. *Agricultural and Forest Meteorology*, *176*, 38–49.

Budyko, M. I., Miller, D. H., & Miller, D. H. (1974). *Climate and life* (Vol. 508). New York: Academic Press.

Burnash, R.J.C., Ferral, R.L., & McGuire, R.A. (1973). A generalized streamflow simulation system—Conceptual modeling for digital computers. Technical Report, United States Department of Commerce, National Weather Service and State of California, Department of Water Resources, Sacramento, Calif.

Cayan, D. R., Das, T., Pierce, D. W., Barnett, T. P., Tyree, M., & Gershunov, A. (2010). Future dryness in the southwest US and the hydrologic of the early 21st century drought. *PNAS*, *107*(50), 21,271–21,276. https://doi.org/10.1073/pnas.0912391107

Chiew, F. H. S. (2006). Estimation of rainfall elasticity of streamflow in Australia. *Hydrological Sciences Journal*, *51*(4), 613–625. https://doi.org/10.1623/hysj.51.4.613

Christensen, N. S., Wood, A. W., Voisin, N., Lettenmaier, D. P., & Palmer, R. N. (2004). The Effects of Climate Change on the Hydrology and Water Resources of the Colorado River Basin. *Climatic Change*, *62*(1-3), 337–363.

Cooper, M. G., Schaperow, J. R., Cooley, S. W., Alam, S., Smith, L. C., & Lettenmaier, D. P. (2018). Climate elasticity of low flows in the maritime western US mountains. *Water Resources Research*, *54*, 5602–5619. https://doi.org/10.1029/2018WR022816

Diffenbaugh, N. S., Swain, D. L., & Touma, D. (2014). Anthropogenic warming has increased drought risk in California. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(13), 3931–3936. https://doi.org/10.1073/pnas.142235112

Donohue, R. J., Roderick, M. L., & McVicar, T. R. (2011). Assessing the differences in sensitivities of runoff to changes in climatic conditions across a large basin. *Journal of Hydrology*, *406*(3-4), 234–244. https://doi.org/10.1016/j.jhydrol.2011.07.003

Dooge, J. C. I. (1992). Sensitivity of runoff to climate change: A Hortonian approach. *Bulletin of the American Meteorological Society*, *73*(12), 2013–2024. https://doi.org/10.1175/1520-0477(1992)073<2013:SORTCC>2.0.CO;2

Falcone, J. A., Carlisle, D. M., Wolock, D. M., & Meador, M. R. (2010). GAGES: A stream gage database for evaluating natural and altered flow conditions in the conterminous United States. *Ecology*, *91*(2), 621–621. https://doi.org/10.1890/09-0889.1

Fu, G., Charles, S. P., & Chiew, F. H. S. (2007). A two-parameter climate elasticity of streamflow index to assess climate change effects on annual streamflow. *Water Resources Research*, *43*, W11419. https://doi.org/10.1029/2007WR005890

Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, *377*(1–2), 80–91. https://doi.org/10.1016/j.jhydrol.2009.08.003

Hamon, W. R. (1961). Estimating Potential Evapotranspiration. *Journal of the Hydraulics Division, Proceedings of the American Society of Civil Engineers*, *87*, 107–120.

Hargreaves, G. H. (1975). Moisture availability and crop production. *Transactions of the American Society of Agricultural Engineers*, *18*(5), 980–984. https://doi.org/10.13031/2013.36722

Hayhoe, K., Wake, C. P., Huntington, T. G., Luo, L., Schwartz, M. D., Sheffield, J., et al. (2006). Past and future changes in climate and hydrological indicators in the U.S. Northeast. *Climate Dynamics*, *28*(4), 381–407. https://doi.org/10.1007/s00382-006-0187-8

Karl, T. R., Melillo, J. M., & Peterson, T. C. (2009). *Global climate change impacts in the United States*. New York, NY: Cambridge University Press.

Konapala, G., & Mishra, A. K. (2016). Three-parameter-based streamflow elasticity model: Application to MOPEX basins in the USA at annual and seasonal scales. *Hydrology and Earth System Sciences*, *20*(6), 2545–2556. Retrieved from. https://www.hydrol-earth-syst-sci.net/20/2545/2016/hess-20-2545-2016-relations.html, https://doi.org/10.5194/hess-20-2545-2016

Kroll, C. N., & Song, P. (2013). Impact of multicollinearity on small sample hydrologic regression models. *Water Resources Research*, *49*(6), 3756–3769.

Kuhnel, V., Dooge, J. C. I., O'Kane, J. P. J., & Romanowicz, R. J. (1991). Partial analysis applied to scale problems in surface moisture fluxes. *Surveys in Geophysics*, *12*(1-3), 221–247. https://doi.org/10.1007/BF01903420

Liang, X., Lettenmaier, D. P., Wood, E. F., & Burges, S. J. (1994). A simple hydrologically based model of land surface water and energy fluxes for GSMs. *Journal of Geophysical Research*, *99*(D7), 14,415–14,,428. https://doi.org/10.1029/94JD00483

Livezey, R. E., & Chen, W. Y. (1983). Statistical field significance and its determination by Monte Carlo techniques. *Monthly Weather Review*, *111*(1), 46–59. https://doi.org/10.1175/1520-0493(1983)111<0046:SFSAID>2.0.CO;2

Livneh, B., Rosenberg, E. A., Lin, C., Nijssen, B., Mishra, V., Andreadis, K. M., et al. (2013). A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States: Update and extensions. *Journal of Climate*, *26*(3), 9384–9392.

Mao, Y., Nijssen, B., & Lettenmaier, D. P. (2015). Is climate change implicated in the 2013-2014 California drought? A hydrologic perspective. *Geophysical Research Letters*, *42*, 2805–2813. https://doi.org/10.1002/2015GL063456

McCabe, G. J., & Wolock, D. M. (1999). Future snowpack conditions in the western United States derived from general circulation model climate simulations. *Journal of the American Water Resources Association*, *35*(6), 1473–1484. https://doi.org/10.1111/j.1752-1688.1999.tb04231.x

McCabe, G. L., & Markstrom, S. L. (2007). A monthly water-balance model driven by a graphical user interface. U.S. Geological Survey Open-File Rep. 2007–1088, 6 pp.

Milly, P. C. D. (1994). Climate, soil water storage, and the average annual water balance. *Water Resources Research*, *30*(7), 2143–2156. https://doi.org/10.1029/94WR00586

Milly, P. C. D., & Dunne, K. A. (2011). On the hydrologic adjustment of climate-model predictions: The potential pitfall of potential evapotranspiration. *Earth Interactions*, *15*(1), 1–14. https://doi.org/10.1175/2010EI363.1

Milly, P. C. D., & Dunne, K. A. (2020). Colorado River flow dwindles as warming-driven loss of reflective snow energizes evaporation. *Science*, *367*(6483), 1252–1255. https://doi.org/10.1126/science.aay9187

Null, S. E., Viers, J. H., & Mount, J. F. (2010). Hydrologic response and watershed sensitivity to climate warming in California's Sierra Nevada. *PLoS ONE*, *5*(4), e9932. https://doi.org/10.1371/journal.pone.0009932

Oudin, L., Hervieu, F., Michel, C., Perrin, C., Andréassian, V., Anctil, F., & Loumagne, C. (2005). Which potential evapotranspiration input for a lumped rainfall-runoff model? Part 2—Towards a simple and efficient potential evapotranspiration model for rainfall-runoff modelling. *Journal of Hydrology*, *303*(1–4), 290–306. https://doi.org/10.1016/j.jhydrol.2004.08.026

Padrón, R. S., Gudmundsson, L., Greve, P., & Seneviratne, S. I. (2017). Large-scale controls of the surface water balance over land: Insights from a systematic review and meta-analysis. *Water Resources Research*, *53*, 9659–9678. https://doi.org/10.1002/2017WR021215

Penman, H. L. (1948). Natural evaporation from open water, bare soil and grass. *Proceedings of the Royal Society of London. Series A*, *193*(1032), 120–145.

Ponce, V. M., Pandey, R. P., & Ercan, S. (2000). Characterization of drought across climate spectrum. *Journal of Hydrologic Engineering, American Society of Civil Engineers*, *5*(2), 222–224.

Priestley, C. H. B., & Taylor, R. J. (1972). On the assessment of surface heat flux and evaporation using large-scale parameters. *Monthly Weather Review*, *106*, 81–92.

Risbey, J. S., & Entekhabi, D. (1996). Observed Sacramento Basin streamflow response to precipitation and temperature changes and its relevance to climate impact studies. *Journal of Hydrology*, *184*(3–4), 209–223. https://doi.org/10.1016/0022-1694(95)02984-2

Roderick, M. L., & Farquhar, G. D. (2011). A simple framework for relating variations in runoff to variations in climatic conditions and catchment properties. *Water Resources Research*, *47*, W00G07. https://doi.org/10.1029/2010WR009826

Sankarasubramanian, A., & Vogel, R. M. (2002). Annual hydroclimatology of the United States. *Water Resources Research*, *38*(6), 1083. https://doi.org/10.1029/2002WR000619

Sankarasubramanian, A., Vogel, R. M., & Limbrunner, J. F. (2001). Climate elasticity of streamflow in the United States. *Water Resources Research*, *37*(6), 1771–1781. https://doi.org/10.1029/2000WR900330

Sankarasubramanian, A., Wang, D., Archfield, S., Reitz, M., Vogel, R. M., Mazrooei, A., & Mukhopadhyaya, S. (2020). HESS opinions: Beyond the long-term water balance: Evolving Budyko's legacy for the Anthropocene towards a global synthesis of land-surface fluxes under natural and human-altered watersheds. *Hydrology and Earth System Sciences Discussions*, *24*(4). https://doi.org/10.5194/hess-2019-418

Schaake, J. C. (1990). From climate to flow. In P. E. Waggoner *Climate Change and U.S. Water Resources* (pp. 177–206). Wiley.

Schaake, J. C., & Chunzhen, L. (1989). Development and application of a simple water balance models to understand the relationship between climate and water resources. In *New directions for surface water modeling. Proceedings of a Baltimore Symposium*, *IAHS Publication* (Vol. 181, pp. 343–352. Baltimore: IAHS Press. Retrieved from http://hydrologie.org/redbooks/a181/iahs_181_0343.pdf

Thornthwaite, C. W. (1948). An approach toward a rational classification of climate. *Geographical Review*, *38*(1), 55–94. https://doi.org/10.2307/210739

U.S. Geological Survey (2016). *National Water Information System data available on the World Wide Web (USGS Water Data for the Nation)*, accessed March 24, 2017. http://waterdata.usgs.gov/nwis/, https://doi.org/10.5066/F7P55KJN

Vano, J. A., Das, T., & Lettenmaier, D. P. (2012). Hydrologic sensitivities of Colorado River runoff to changes in precipitation and temperature. *Journal of Hydrometeorology*, *13*(3), 932–949. https://doi.org/10.1175/JHM-D-11-069.1

Vano, J. A., & Lettenmaier, D. P. (2014). A sensitivity-based approach to evaluating future changes in Colorado River discharge. *Climate Change*, *122*(4), 621–634. https://doi.org/10.1007/s10584-013-1023-x

Vicuna, S., Maurer, E. P., Joyce, B., Dracup, J. A., & Purkey, D. (2007). The sensitivity of California water resources to climate change scenarios. *Journal of the American Water Resources Association*, *43*(2), 482–498. https://doi.org/10.1111/j.1752-1688.2007.00038.x

Wallis, J. R. (1965). Multivariate statistical methods in hydrology: A comparison using data of known functional relationship. *Water Resources Research*, *1*(4), 447–461. https://doi.org/10.1029/WR001i004p00447

Wang, C., Wang, S., Fu, B., & Zhang, L. (2016). Advances in hydrological modelling with the Budyko framework: A review. *Progress in Physical Geography: Earth and Environment*, *40*(3), 409–430. https://doi.org/10.1177/0309133315620997

Wang, H., & He, K. (2017). Sensitivity analysis of the effects of climate change on streamflow using climate elasticity in the Lunan River basin, China. *Polish Journal of Environmental Studies*, *26*(2), 837–845. https://doi.org/10.15244/pjoes/66715

Wang, W., Zou, S., Shao, Q., Xing, W., Chen, X., Jiao, X., et al. (2016). The analytical derivation of multiple elasticities of runoff to climate change and catchment characteristics alteration. *Journal of Hydrology*, *541*, 1042–1056. https://doi.org/10.1016/j.jhydrol.2016.08.014

Williams, C. A., Reichstein, M., Buchmann, N., Baldocchi, D., Beer, C., Schwalm, C., et al. (2012). Climate and vegetation controls on the surface water balance: Synthesis of evapotranspiration measured across a global network of flux towers. *Water Resources Research*, *48*, W06523. https://doi.org/10.1029/2011WR011586

Wood, A. W., & Lettenmaier, D. P. (2006). A testbed for new seasonal hydrologic forecasting approaches in the western U.S. *Bulletin of the American Meteorological Society*, *87*(12), 1699–1712. https://doi.org/10.1175/BAMS-87-12-1699

Xiao, M., Nijssen, B., & Lettenmaier, D. P. (2016). Drought in the Pacific Northwest, 1920-2013. *Journal of Hydrometeorology*, *17*, 2391–2404. https://doi.org/10.1175/JHM-D-15-0142.1

Yang, H., & Yang, D. (2011). Derivation of climate elasticity of runoff to assess the effects of climate change on annual runoff. *Water Resources Research*, *47*, W07526. https://doi.org/10.1029/2010WR009287

Yang, Y., Roderick, M. L., Zhang, S., McVicar, T. R., & Donohue, R. J. (2019). Hydrologic implications of vegetation response to elevated CO_2 in climate projections. *Nature Climate Change*, *9*(1), 44–48. https://doi.org/10.1038/s41558-018-0361-0

Zhang, X., Tang, Q., Zhang, X., & Lettenmaier, D. P. (2014). Runoff sensitivity to global mean temperature change in the CMIP5 models. *Geophysical Research Letters*, *41*, 5492–5498. https://doi.org/10.1002/2014GL060382

Zheng, H., Zhang, L., Zhu, R., Liu, C., Sato, Y., & Fukushima, Y. (2009). Responses of streamflow to climate and land surface change in the headwaters of the Yellow River Basin. *Water Resources Research*, *45*, W00A19. https://doi.org/10.1029/2007WR006665