Article

Understanding the Budyko Equation

Garrison Sposito 1,2
1 Department of Environmental Science, Policy and Management, University of California, Berkeley, CA 94720-3114, USA; gsposito@berkeley.edu or gsposito@lbl.gov
2 Earth & Environmental Sciences Area, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

Academic Editor: Timothy Green
Received: 2 February 2017; Accepted: 17 March 2017; Published: 25 March 2017

Abstract: The Budyko equation has achieved iconic status in hydrology for its concise and accurate representation of the relationship between annual evapotranspiration and long-term-average water and energy balance at catchment scales. Accelerating anthropogenic land-use and climate change have sparked a renewed interest in predictive applications of the Budyko equation to analyze future scenarios important to water resource management. These applications, in turn, have inspired a number of attempts to derive mathematical models of the Budyko equation from a variety of specific assumptions about the original Budyko hypothesis. Here, we show that the Budyko equation and all extant models of it can be derived rigorously from a single mathematical assumption concerning the Budyko hypothesis. The implications of this fact for parametric models of the Budyko equation also are explored.

Keywords: Budyko equation; evapotranspiration; green water; catchment hydrology; water balance

1. Introduction

1.1. Introducing Green Water

Green water is water in the vadose zone that is available for uptake by plant roots and vadose-zone biota [1] (p. 279). The principal flow of green water is by evapotranspiration, through which it is transferred from the upper portion of the vadose zone through the land surface or a vegetation canopy into the atmosphere. The hydrologic importance of evapotranspiration as a mode of water transfer becomes apparent from comparing the average annual volumes of global terrestrial evapotranspiration and runoff, which shows that currently, annual terrestrial evapotranspiration is about 54% larger than annual runoff out of the continents [2].

Both spatially and temporally, terrestrial evapotranspiration is a highly variable hydrologic process which remains challenging to quantify directly, even at a single location. Indirect measurement using combinations of ground-based, satellite-based, and model-based methods is therefore the usual approach for determining evapotranspiration at catchment to global scales [2].

The transpiration (T) component of terrestrial evapotranspiration (ET), also termed “productive green water flow” [3] because it is the ET mode directly supporting carbon assimilation and primary productivity, is an important and abiding topic in ecohydrology research which remains challenging because of the aforementioned difficulty of measuring ET directly. Schlesinger and Jasechko [4] have recently compiled best-estimate values of the ratio T/ET for nine terrestrial biomes. The average global T/ET, calculated as the sum of the biome T/ET values weighted according to either the volume of precipitation into or the volume of ET out of a given biome, is 61% ± 15%. Using this estimate and the comparison of ET with runoff, given above, one calculates that the total volume of green water flow by transpiration is approximately equal to the total volume of runoff. Remarkably, the global annual flow of green water through plant roots into the atmosphere matches the annual flow of all the rivers in the world into the oceans.

At the catchment scale, ET is also recognized as a key hydrologic process, first because of its evident connection to the primary productivity of ecosystems, both wildland and agricultural [1] (pp. 95–100), and second because the water transported by ET has not run off into streams and rivers or percolated below the vadose zone to replenish groundwater, either of which would have made it still accessible for human consumption. Thus, catchment ET is a key diagnostic process contributing to water resource management, which has become particularly important in light of stresses caused by accelerating anthropogenic land-use and climate change [5–8].

1.2. The Budyko Hypothesis

Because it involves a phase transformation, that of green water into water vapor, ET reflects the partitioning not only of water but also radiant energy at the vadose zone–atmosphere interface [9]. This perspective informs the approach to the physics of catchment ET developed by M. I. Budyko during the previous century and summarized in his iconic book, *Climate and Life* [10]. Budyko’s approach is “Darwinian,” as opposed to “Newtonian” [11,12], because it foregoes reductionist explanations based on constitutive equations [13] in favor of establishing universal relationships based solely on the mass and energy balance laws to which any physical system must conform.

Budyko’s ideas can be outlined as follows. The long-term-average annual flow of green water through the upper boundary of the vadose zone into the atmosphere is subject to the same water- and energy-balance assumptions as apply to the global flow of green water, i.e., negligible changes in subsurface water storage and negligible net heat transfer between the land surface and the vadose zone [2,12,14]. Under these two assumptions, the long-term-average annual water and energy balance at the catchment scale can be expressed [10] (pp. 16–23) and [14]:

\[ P = ET + Q, \]  
\[ R_n = L \cdot ET + H, \]

where \( P \) (L/T) is precipitation, \( ET \) (L/T) is evapotranspiration, \( Q \) (L/T) is runoff, \( R_n \) (M/T^3) is the average annual net radiative heat flux from the atmosphere to the land surface, \( L \) (M/L·T^2) is the latent heat of evaporation, the product \( L \cdot ET \) (M/T^3) then being the latent heat flux, and \( H \) (M/T^3) is the average annual heat flow from the land surface into the atmosphere, the sensible heat flux.

Equation (1) states that, over the time and spatial scales being considered, precipitation falling onto a catchment is not stored in or below the vadose zone and so must leave it, either as a flow of green water or as runoff. Equation (2) states that solar radiation impinging on a catchment also is not stored in the vadose zone, leaving it either as heat carried off by the flow of green water (L·ET) or as thermal radiation from the upper boundary of the vadose zone back into the atmosphere (H). Furthermore, because evapotranspiration and latent heat flux are equivalent processes, division of Equation (2) by \( L \) leads to an energy-balance constraint on ET analogous to the water-balance constraint expressed by Equation (1):

\[ \frac{R_n}{L} = ET + \frac{H}{L} \]  

According to Equation (3), the maximum possible ET occurs when all incoming radiation energy is consumed by the evaporation of green water from the land surface, the green water available for transfer as ET is not limiting, and there is negligible sensible heat flux [10,14]. Evidently, this maximum possible ET is numerically equal to \( \frac{R_n}{L} \), which is therefore designated as potential evapotranspiration [10] (p. 23) and given the symbol, \( ET_0 \).

Budyko [10] (p.323) hypothesized that Equations (1) and (3) imply that a functional relationship exists between ET and the two climate variables, \( P \) and \( ET_0 \), a relationship which should apply to all catchments as an average over long-enough time scales:

\[ ET = f(P, ET_0). \]
He further stated, on the basis of his synthesis of available data, that this functional relationship is subject to the limiting conditions:

\begin{align}
    \text{ET} \rightarrow \text{ET}_0 \text{ as } P \uparrow \infty \\
    \text{ET} \rightarrow P \text{ as } \text{ET}_0 \uparrow \infty 
\end{align}

Equation (5) describes a “wet condition” of relatively high green water availability and relatively low sensible heat flux, whereas Equation (6) describes a “dry condition” of relatively low green water availability and relatively high sensible heat flux. Thus, under the wet condition, ET is energy-limited by \( \text{ET}_0 \), whereas under the dry condition, ET is water-limited by \( P \) [12].

Equation (4) and the limits imposed on it in Equations (5) and (6) may be termed the Budkyo hypothesis concerning the long-term-average annual ET from a catchment. Using his deep physical insight into catchment hydrology and meteorology, Budyko [10] (p. 323) went a step further and proposed a mathematical representation of Equation (4) which may be termed the Budyko equation. In addition to empirical validation studies involving statistical curve-fitting [12], the Budyko equation has been the subject of a number of recent attempts to validate it with a firm mathematical foundation. These studies rely on \textit{ad hoc} assumptions about the Budyko hypothesis in order to derive the Budyko equation along with parametric models to represent it. The typical approach [15,16] invokes dimensional analysis to justify a change of the dependent and independent variables in the Budyko hypothesis to \( \text{ET}/P \) and \( \text{ET}_0/P \), respectively, then goes on to develop lengthy mathematical arguments based on the assumption that the differential form of Equation (4) is exact. The model parameter then emerges as an arbitrary constant related to separation of variables. It is the purpose of this paper to show, however, that such an approach is not necessary, that instead making a single, very general assumption concerning Equation (4) suffices to derive the Budyko equation rigorously from the Budyko hypothesis, as well as provide a unified mathematical framework for all extant models of the Budyko equation, both parametric and non-parametric.

2. Deriving the Budyko Equation

The Budyko hypothesis, as represented in Equation (4), can be developed further by postulating that \( f(\text{ET}_0, P) \) is a homogeneous function of its arguments, which means that, if the independent variables \( \text{ET}_0, P \) are each multiplied by an arbitrary factor, \( \lambda \), the value of \( f(\text{ET}_0, P) \) is multiplied by the same factor, \( \lambda \) [17]:

\begin{equation}
    f(\lambda \text{ET}_0, \lambda P) = \lambda \text{ET} 
\end{equation}

Hankey and Stanley [17] have proved that the property of homogeneity, as expressed in Equation (7), is mathematically equivalent to the representation:

\begin{equation}
    \text{ET} = P F\left( \frac{\text{ET}_0}{P} \right) 
\end{equation}

which they also show to be a valid alternative to Equation (7) for defining a homogeneous function of \( P \) and \( \text{ET}_0 \). Equivalency can be demonstrated by replacing \( \text{ET}, P, \text{ET}_0 \) everywhere in Equation (8) with \( \lambda \text{ET}, \lambda P, \lambda \text{ET}_0 \), respectively, then noting that, after cancellation of \( \lambda \) in the argument of \( F(\cdot) \), Equation (8) satisfies the definition of homogeneity. That Equation (8) is a valid alternative to Equation (7) is shown by setting \( \lambda = 1/P \) in Equation (7) to reproduce Equation (8).

After division of both sides by \( P \), Equation (8) becomes the form of Equation (4) originally proposed by Budyko [10] (p. 323), who, as noted above, derived it using physical reasoning based on trends he deduced from catchment water- and energy-balance data. Following [14], the ratio \( \text{ET}_0/P \) is termed the aridity index and given the symbol \( \phi \). Values of \( \phi < 1 \) thus indicate a humid climate, whereas \( \phi > 1 \) indicates an arid climate [14]. Within this convention, the limiting conditions on the Budyko equation,

\begin{equation}
    \frac{\text{ET}}{P} = F(\phi),
\end{equation}
that stem from Equations (5) and (6) are:

\[ \frac{ET}{P} \downarrow \phi \downarrow 0, \text{ (ET is energy-limited)} \quad (10) \]

\[ \frac{ET}{P} \uparrow 1 \text{ as } \phi \uparrow \infty. \text{ (ET is water-limited)} \quad (11) \]

It follows that a graph of the Budyko equation must begin as a curve tangent to the 1:1 line, \( \frac{ET}{P} = \frac{ET_0}{P} \), then turn concave to the \( \phi \)-axis and asymptotically approach the horizontal line defined by \( \frac{ET}{P} = 1 \) \[5,10,12,14–16\].

Budyko \[10\] (pp. 325–327) tested Equation (9) successfully using hydrologic and climatological data for more than 1000 catchments encompassing a variety of biomes, showing also that the limiting conditions in Equations (10) and (11) were met. Since then, a large number of successful tests of the Budyko equation has been reported based on experimental measurements \[5,18,19\] and model simulations \[14,20\]. Ye et al. \[19\] performed a particularly comprehensive test of Equation (7b) based on data collected daily for 50 years (1951–2000) in more than 250 catchments in the United States which span a range of climatic zones and physiographic regions. As noted by Arora \[14\], although scatter in plots of \( \frac{ET}{P} \) vs. \( \phi \) based on experimental data always occurs, these successful tests demonstrate “the primary control of precipitation and available energy in determining the ratio of annual evapotranspiration to precipitation” \[14\] (p. 167).

Because the physical assumptions underlying the derivation of Equations (1) and (2) are reflected in Equations (10) and (11), data conforming to the Budyko equation should never lie above either the 1:1 line defined by \( \frac{ET}{P} = \frac{ET_0}{P} \) or the horizontal line defined by \( \frac{ET}{P} = 1 \). Individual data points, however, sometimes are found to violate one of these constraints (see, e.g., Figure 5 in \[5\]), which implies either measurement errors or violations of the assumptions made concerning subsurface storage that underlie the long-term-average, large-spatial-scale conditions requisite to the Budyko hypothesis \[21,22\].

Besides Equation (8), two other important properties of Equation (4) follow from the homogeneity postulate given by Equation (7). The first of these properties is based on the fact that all homogeneous functions must satisfy the Euler relation \[23\] (pp. 59–60), which is expressed in the present case by:

\[ ET = \left( \frac{\partial ET}{\partial P} \right)_{ET_0} P + \left( \frac{\partial ET}{\partial ET_0} \right)_P ET_0, \quad (12) \]

Equation (12) is derived from Equation (7) by differentiating both sides of the latter with respect to \( \lambda \), then setting \( \lambda = 1 \) \[23\] (p. 59). On physical grounds, both of the partial derivatives in Equation (12) are positive-valued and subject to boundary conditions stemming from Equations (5), (6), (10), and (11) \[16\]:

\[ \left( \frac{\partial ET}{\partial P} \right)_{ET_0} \downarrow 0 \text{ as } \phi \downarrow 0, \quad (13) \]

\[ \left( \frac{\partial ET}{\partial P} \right)_{ET_0} \uparrow 1 \text{ as } \phi \uparrow \infty, \quad (14) \]

\[ \left( \frac{\partial ET}{\partial ET_0} \right)_P \uparrow 1 \text{ as } \phi \downarrow 0, \quad (15) \]

\[ \left( \frac{\partial ET}{\partial ET_0} \right)_P \downarrow 0 \text{ as } \phi \uparrow \infty. \quad (16) \]

Thus, both partial derivatives take on values in the closed range \([0, 1]\). Graphs illustrating these boundary conditions are given in Figure 6 of \[16\] based on a parametric model of Equation (9). (See also Figure 6 in \[20\] and Figure 3 in \[14\] for illustrative plots of \( \frac{\partial ET}{\partial P} \) vs. \( \phi \)). Equations (13) and (15) reflect the wet condition described by Equation (10), whereas Equations (14) and (16) reflect the dry condition described by Equation (11). Put another way, Equations (13) and (14) express the evident
fact that precipitation is mainly transformed into runoff under the wet condition, but is transformed into evapotranspiration under the dry condition. Similarly, Equations (15) and (16) reflect the high sensitivity of ET to changes in potential evapotranspiration under the wet (energy-limited) condition and its low sensitivity to changes in ET_0 under the dry (water-limited) condition.

The second important property of Equation (4) that derives from the homogeneity postulate is the Gibbs–Duhem relation, expressed in the present case by:

\[ P \frac{d}{dP} \left( \frac{\partial ET}{\partial P} \right)_{ET_0} + ET_0 \frac{d}{dET_0} \left( \frac{\partial ET}{\partial ET_0} \right)_P = 0. \] (17)

This equation gets its name from a well-known relation among the differentials of intensive variables that arises in equilibrium thermodynamics [23] (pp. 60–62). Equation (17) is derived by calculating the total differential of ET using the Euler relation, then comparing the result to the total differential of ET based on Equation (4).

The Euler relation implies that knowledge of the \( \phi \)-dependence of the two partial derivatives suffices to determine the function \( F(\phi) \) in the Budyko equation. The Gibbs–Duhem relation shows further that variations of these two partial derivatives in response to changes in \( \phi \) are not independent. As will become evident in the following section, these two properties, true of all homogeneous functions, play key roles in the development of explicit mathematical models of \( F(\phi) \).

3. Modeling the Budyko Equation

Budyko [10] (p. 325) himself proposed and tested a mathematical model of Equation (9) which was a non-parametric interpolation formula connecting the two limiting conditions expressed in Equations (10) and (11). His model has seen widespread successful application, but parametric models of Equation (9) also have been proposed, a number of which are described and compared in a recent comprehensive review by Wang et al. [12]. The two principal models in current use, the Fu model [15] and the Mezentsev–Choudhury–Yang (MCY) model [16,24], are one-parameter models based on specific assumptions about the \( \phi \)-dependence of the two partial derivatives appearing in Equation (17). Zhou et al. [24] have shown that these two models are the only ones extant that satisfy the uniqueness requirement, that any point in the two-dimensional field spanned by \( (F, \phi) \) belongs to a single curve as defined by a chosen value of the model parameter. Zhou et al. [24] also show that both models can be derived from specific assumptions about the \( \phi \)-dependence of the ratio, \( \frac{\partial ET}{\partial P}/\phi \frac{\partial ET}{\partial ET_0} \), which they accordingly term a “generating function” for models of the Budyko equation. In the following discussion, it will be shown that all of the modeling approaches in the literature fall within a rigorous mathematical framework based solely on the homogeneity postulate in Equation (7). In this more general approach, for example, the constraint on \( \frac{\partial ET}{\partial P} \) and \( \frac{\partial ET}{\partial ET_0} \) employed by Zhou et al. [24] to deduce mathematical properties of their generating function is revealed to be just a rearranged form of the Euler equation.

3.1. The Legendre Transformation of \( F(\phi) \)

The Legendre transformation is a well-known mathematical technique applied in equilibrium thermodynamics [23] (pp. 137–142) to convert a function, such as \( F(\phi) \), into another function, \( \psi(F') \), that is equivalent mathematically to \( F(\phi) \), where \( F' = dF/d\phi \). Explicitly, the Legendre transformation of \( F(\phi) \) is defined by the equation [23] (pp. 137–142) and [25] (pp. 63–65):

\[ \psi(F') = F - \phi F'. \] (18)

The basis for the definition in Equation (18) can be understood as follows. Given a graph of \( F(\phi) \), i.e., a curve in the \( F-\phi \) field, one can construct a line tangent to some point on the curve, the slope of
which is equal to the value of $F'$ at that point. Let the F-intercept of the line be $\psi$, i.e., let the tangent line intersect the F-axis at the point $(\psi, 0)$. The equation describing this tangent line in the F-$\phi$ field is then:

$$F = \psi + F'\phi,$$  \hspace{1cm} (19)

which can be rearranged to yield Equation (18). Thus, the Legendre transformation of $F(\phi)$, $\psi(F')$, prescribes the F-intercept of a line tangent to a curve in the F-$\phi$ field as a function of the slope of the tangent line.

Mathematical equivalence between $F(\phi)$ and $\psi(F')$ is established by the fact that every point along a curve of $F(\phi)$ is associated with a unique tangent line described by $\psi(F')$. To quote Callen [23] (p. 140) “knowledge of the intercepts $\psi$ of the tangent lines as a function of the slopes $F'$ enables us to construct the family of tangent lines and hence the curve of which they are the envelope.” Therefore, the family of tangent lines defines a curve of $F(\phi)$ just as uniquely as does the locus of points constituting the curve. In this sense, every curve has an “F-representation” as a locus of points and a “$\psi$-representation” as an envelope of tangent lines. Which representation one uses is a matter of convenience in applications, as is well known in equilibrium thermodynamics, where the Legendre transformation of the internal energy corresponds to a thermodynamic potential, such as the enthalpy, which offers a rigorous alternative to using the internal energy to describe the equilibrium states of a physical system [23] (p. 147) and [25] (p. 78).

The differential form of Equation (18),

$$d\psi = dF - F'd\phi - \phi dF'$$   \hspace{1cm} (20)

yields an expression for the first derivative of $\psi$:

$$\frac{d\psi}{dF'} = -\phi,$$   \hspace{1cm} (21)

as can be seen after introducing the definition of $F'$ into the second term on the right side of Equation (20). Therefore, given $\psi(F')$, one can calculate $F(\phi)$ by establishing the relationship between $\phi$ and $F'$ with Equation (21), then using this relationship to eliminate $\psi$ and $F'$ from Equation (19), i.e., $\psi = \psi(F')$ and $\frac{d\psi}{dF'} = -\phi$ used together yield $F(\phi)$ from $F = \psi + F'\phi$.

As an example of this inversion procedure, consider the Schreiber model, an interpolation formula cited by Budyko [10] (p. 323) as the earliest published attempt to model Equation (9) mathematically. In its $\psi$-representation, the Schreiber model has the form:

$$\psi = 1 + (\ln F' - 1) F'. $$   \hspace{1cm} (22)

Calculating the first derivative of $\psi(F')$ using Equation (22) and invoking Equation (21), one finds:

$$\frac{d\psi}{dF'} = \ln F' = -\phi,$$   \hspace{1cm} (23)

Therefore, $F' = \exp(-\phi)$, which now may be used along with Equation (22) to eliminate $F'$ and $\psi$ from Equation (19):

$$F(\phi) = \psi + F'\phi = [1 + (-\phi - 1) \exp(-\phi)] + [\exp(-\phi)\phi] = 1 - \exp(-\phi),$$   \hspace{1cm} (24)

which is the Schreiber model in its F-representation [10] (p. 323). Equations (22) and (24) thus offer equivalent mathematical representations of this model.
3.2. Physical Meaning of the Legendre Transformation

The physical significance of the Legendre transformation in the context of the Budyko equation follows from calculating the partial derivatives that appear in Equation (17) using Equation (8):

\[
\left( \frac{\partial ET}{\partial P} \right)_{ET_0} = F(\phi) + P \left( \frac{\partial F}{\partial P} \right)_{ET_0} = F(\phi) + P \left( \frac{\partial \phi}{\partial P} \right)_{ET_0} F' = F(\phi) - \phi F' = \psi, \tag{25}
\]

\[
\left( \frac{\partial ET}{\partial ET_0} \right)_P = P \left( \frac{\partial F}{\partial ET_0} \right)_P = P \left( \frac{\partial \phi}{\partial ET_0} \right)_P F' = F', \tag{26}
\]

where the chain rule for derivatives has been used to get the second step in each equation. Comparison of Equation (25) with Equation (18) shows that \( \left( \frac{\partial ET}{\partial P} \right) \) is in fact the Legendre transform of \( F(\phi) \), while reference to Equation (26) shows that \( \left( \frac{\partial ET}{\partial ET_0} \right) \) is the slope variable \( F' \) on which the Legendre transform depends, i.e., \( \psi(F') \) is in fact \( \left( \frac{\partial ET}{\partial P} \right) \) expressed as a function of \( \left( \frac{\partial ET}{\partial ET_0} \right) \), this latter function being completely equivalent mathematically to \( F(\phi) \).

This relationship is implicit in the Euler relation because it may be rearranged, with the help of Equations (25) and (26), to become Equation (18), which defines the Legendre transform of \( F(\phi) \). Moreover, Equation (17) may be rearranged to become Equation (21). Thus, the Legendre transformation of \( F(\phi) \) is intimately related to the property of homogeneity underlying the derivation of the Budyko equation. Equations (25) and (26) were presented by Zhou et al. [24] as the initial step in their procedure for defining the generating function they used to derive models of the Budyko equation, but they apparently were unaware of the close connection between the two partial derivatives and the Legendre transformation of \( F(\phi) \).

Given Equations (25) and (26), the boundary conditions in Equations (13)–(16) can be combined to determine corresponding boundary conditions on \( \psi(F') \):

\[
\psi(F') \downarrow 0 \text{ as } F' \uparrow 1, \tag{27}
\]

\[
\psi(F') \uparrow 1 \text{ as } F' \downarrow 0. \tag{28}
\]

Moreover, these boundary conditions and Equation (21) show that the first derivative of \( \psi(F') \) is non-positive and that it satisfies the boundary conditions:

\[
\frac{d\psi}{dF'} \uparrow 0 \text{ as } F' \uparrow 1, \tag{29}
\]

\[
\frac{d\psi}{dF'} \downarrow -\infty \text{ as } F' \downarrow 0. \tag{30}
\]

Therefore, a graph of \( \psi(F') \) will be a curve in the \( \psi-F' \) field that is convex to the \( F' \)-axis, dropping initially from \( \psi = 1 \) parallel to the \( \psi \)-axis, where \( F' = 0 \), to end at \( F' = 1 \) running parallel to the \( F' \)-axis, where \( \psi = 0 \). Examples of this behavior can be seen in Figure 7 of [16], which is based on the MCY model. The graph in this figure shows representative curves corresponding to different values of the MCY model parameter. However, as just demonstrated, the overall characteristics of these curves are not model-dependent, but instead result from the general boundary conditions in Equations (27)–(30).

We note in passing that the concavity of \( F(\phi) \) curves with respect to the \( \phi \)-axis and the convexity of \( \psi(F') \) curves with respect to the \( F' \)-axis are well-known general properties of the relationship between the \( F \) and \( \psi \)-representations [23] (p. 149).

3.3. Understanding the MCY Model

The history of the MCY model is described by Wang et al. [12], who note its transition from an empirical data-fitting equation to an analytically-derived expression based on physical and mathematical considerations. The model has been applied often, with reasonable success, to estimate...
values of \( \frac{\partial \text{ET}}{\partial P} \) and \( \frac{\partial \text{ET}}{\partial \text{ET}_0} \), as well as those of the partial derivative of ET with respect to the model parameter, in efforts to predict the effects on runoff from changing climate (“climate elasticity”), as represented by variations of P and ET, as well as changing catchment properties, assumed to be represented by variation of the model parameter [9,16,26–29].

As noted in the Introduction, Yang et al. [16] derived the MCY model analytically through an argument based on requiring the total differential of ET in Equation (4) be exact, then applying dimensional analysis to infer that \( \frac{\partial \text{ET}}{\partial P} \) and \( \frac{\partial \text{ET}}{\partial \text{ET}_0} \) are functions of ET/P and ET/ET, that satisfy the boundary conditions in Equations (13)–(16). In the notation of the present paper, their final expressions for the two partial derivatives are:

\[
\left( \frac{\partial \text{ET}}{\partial P} \right)_{\text{ET}_0} = F(\varphi, n) \left[ 1 - \left( \frac{F(\varphi, n)}{\varphi} \right)^n \right], \tag{31}
\]

\[
\left( \frac{\partial \text{ET}}{\partial \text{ET}_0} \right)_{P} = \frac{F(\varphi, n)}{\varphi} \left[ 1 - F(\varphi, n)^n \right], \tag{32}
\]

where n, the MCY model parameter, is a positive number. Yang et al. [16] integrated this set of partial differential equations to derive the MCY model, but the same result can be obtained more simply by introducing Equations (31) and (32) directly into the Euler relation, then solving it for F(\(\varphi\), n):

\[
F(\varphi, n) = \varphi \left( 1 + \varphi^n \right)^{-1/n} \quad (n > 0), \tag{33}
\]

which is the MCY model in its F-representation [12]. This model of F(\(\varphi\)) satisfies the limiting conditions in Equations (10) and (11) while interpolating smoothly between them for any value of the model parameter n. Increasing the value of n, which functions as a curve-shape parameter, leads to a more rapid approach of F(\(\varphi\), n) to the energy-limiting condition in a humid climate (\(\varphi < 1\)) and to the water-limiting condition in an arid climate (\(\varphi > 1\)). This behavior suggested to Yang and Yang [27] that n increases with green water availability in a catchment, a large value of n thus indicating that there is sufficient green water available for ET to achieve close to its maximum value under any climatic condition.

Equation (33) may be used to rewrite Equations (31) and (32) in a simpler form:

\[
\left( \frac{\partial \text{ET}}{\partial P} \right)_{\text{ET}_0} = F(\varphi, n)^{n+1} = \psi_n, \tag{34}
\]

\[
\left( \frac{\partial \text{ET}}{\partial \text{ET}_0} \right)_{P} = \left( \frac{F(\varphi, n)}{\varphi} \right)^{n+1} = F_n', \tag{35}
\]

from which it follows, using Equation (33) once more to eliminate \(\varphi\), that the MCY model in its \(\psi\)-representation is:

\[
\psi_n n/(n+1) + F_n'^{n/(n+1)} = 1 \quad (n > 0). \tag{36}
\]

This power-law expression is easily seen to satisfy the boundary conditions in Equations (27) and (28). In the context of the \(\psi\)-representation, the derivation of the MCY model by Yang et al. [16] amounts to substituting the model forms in Equations (34) and (35) [or Equations (31) and (32)] into Equation (36) and solving for F(\(\varphi\), n). Alternatively, a formal inversion of Equation (36) to derive the corresponding \(\varphi\)-representation can be accomplished following the procedure used to derive the F-representation of the Schreiber model from its \(\psi\)-representation.

3.4. Understanding the Fu Model

The history of the Fu model is discussed by Zhang et al. [15], who credit its origin to the Chinese hydrologist, B. P. Fu. Their derivation of F(\(\varphi\)) is based on his 1981 study of the Budyko equation, published only in Chinese, and uses the same physical assumptions and mathematical reasoning.
Similar to the MCY model, the Fu model has been applied successfully to estimate \( \partial \frac{ET}{\partial P} \) and \( \partial \frac{ET}{\partial ET_0} \), as well as calculate the partial derivative of ET with respect to the model parameter, to predict climate elasticity as a function of the aridity index and the effects of catchment properties on ET, particularly vegetation cover and subsurface storage, which are assumed to be represented by variations of the model parameter [8,30–35].

As indicated above, Zhang et al. [15] derived the Fu model from physical assumptions about the dependence of \( \partial \frac{ET}{\partial P} \) and \( \partial \frac{ET}{\partial ET_0} \) on ET, P, and ET\(_0\), then followed the mathematical procedure based on dimensional analysis and exact differentials already described for the MCY model to solve the resulting set of partial differential equations. In the notation of the present paper, the Fu model for the two partial derivatives can be expressed:

\[
\left( \frac{\partial ET}{\partial P} \right)_{ET_0} = 1 - (1 + \phi - F(\phi, w))^{1-w}, \tag{37}
\]

\[
\left( \frac{\partial ET}{\partial ET_0} \right)_{P} = 1 - \left( 1 + \frac{1 - F(\phi, w)}{\phi} \right)^{1-w}, \tag{38}
\]

where \( w \) is the Fu model parameter, a positive number with values in the range \((1, \infty)\).

The physical assumption leading to Equation (37) is that, for a given value of ET\(_0\), \( \partial ET/\partial P \) increases with the “residual evapotranspiration,” ET\(_0\) – ET, a measure of the energy that could have been transported to the atmosphere by evapotranspiration if precipitation were not limiting [15]. In addition, \( \partial ET/\partial P \) decreases with increasing P, at fixed ET\(_0\), in accordance with Equation (10). The \( \phi \)-dependent quantity in parentheses on the right side of Equation (37), which is equal to the ratio of the residual evapotranspiration to precipitation, reflects these two assumptions. [Note that \((1 - w) < 0\).]

The physical assumption leading to Equation (38) is that, for a given value P, \( \partial ET/\partial ET_0 \) will increase with the “residual precipitation,” P – ET (i.e., runoff), the volume of water that could have been transported to the atmosphere by evapotranspiration if potential evapotranspiration were not limiting [15]. On the other hand, \( \partial ET/\partial ET_0 \) will decrease with increasing ET\(_0\), at fixed P, in accordance with Equation (16). The \( \phi \)-dependent quantity in parentheses on the right side of Equation (38), which is equal to the ratio of catchment runoff to potential evapotranspiration, reflects this assumption.

Zhang et al. [15] integrated the set of partial differential equations in Equations (37) and (38) to derive the Fu model, but, as in the case of the MCY model, the same result can be obtained more simply by introducing these two equations directly into the Euler relation and solving it for \( F(\phi, w) \) to obtain:

\[
F(\phi, w) = 1 + \phi - (1 + \phi^w)^{1/w} (w > 1), \tag{39}
\]

the Fu model in its F-representation [15]. Like the MCY model, this model of \( F(\phi) \) satisfies the limiting conditions in Equations (10) and (11) while interpolating smoothly between them for any value of the parameter \( w \). Similarly to the parameter \( n \) in the MCY model, \( w \) functions as a curve-shape parameter in the Fu model, with increasing \( w \) leading to a more rapid approach of \( F(\phi, w) \) to the energy-limiting condition in a humid climate \((\phi < 1)\) and to the water-limiting condition in an arid climate \((\phi > 1)\). [See, e.g., Figure 1 in [15].] Zhou et al. [8] associate this effect of increasing \( w \) with an increasing ability of a catchment to retain green water for evapotranspiration, similar to the interpretation of increasing \( n \) in the MCY model (Yang and Yang, 2011). Zhou et al. [8] conclude that \( 1 < w < 2 \), corresponding to low green water storage, corresponds to a lesser effect of climate on runoff relative to catchment effects. Zhang et al. [15] show further that ET/P in the Fu model is the most sensitive to the value of \( w \) at the crossover, \( \phi = 1 \), between humid and arid climates, attributing this behavior to a positive synergism between precipitation and potential evapotranspiration, each of which otherwise would tend to control ET/P separately. This peak sensitivity at \( \phi = 1 \) increases dramatically as \( w \) decreases below 2, in keeping with the conclusion of Zhou et al. [8].
Equation (39) may be used to rewrite Equations (37) and (38) in the form:

\[
\left( \frac{\partial ET}{\partial P} \right)_{ET_0} = 1 - (1 + \varphi^w)^{(1-w)/w} = \psi_w, \tag{40}
\]

\[
\left( \frac{\partial ET}{\partial ET_0} \right)_P = 1 - \left( \frac{\varphi}{1 + \varphi^w} \right)^{w-1} = F_w^*, \tag{41}
\]

It follows from Equations (40) and (41) that the Fu model in its \( \psi \)-representation is:

\[
(1 - \psi_w)^{w/(w-1)} + (1 - F_w^*)^{w/(w-1)} = 1 \quad (w > 1). \tag{42}
\]

Similar to the MCY model, the Fu model is a power-law expression in its \( \psi \)-representation which satisfies the boundary conditions in Equations (27) and (28). The derivation of the Fu model given by Zhang et al. [15] amounts to substituting the model forms in Equations (37) and (38) [or Equations (40) and (41)] into Equation (42) and solving for \( F(\varphi, w) \). Of course, a formal inversion of Equation (42) to derive the \( \varphi \)-representation can always be done following the procedure used to derive the Schreiber model from its \( \psi \)-representation.

4. Discussion

It has been shown that a single assumption concerning the Budyko hypothesis, namely that \( f(P, ET_0) \) in Equation (4) is a homogeneous function of its arguments, suffices to derive the Budyko equation [Equation (9)] and to provide a unified framework for developing mathematical models of this equation, such as the MCY and FZ models. The key elements of this simple approach to the Budyko hypothesis can be summarized as follows.

1. The assumption that \( ET \) is a homogenous function of \( P \) and \( ET_0 \) is both necessary and sufficient to derive the Budyko equation, \( ET/P = F(ET_0/P) [17] \).

2. Homogeneity also leads to the Euler relation, \( ET = \left( \frac{\partial ET}{\partial ET_0} \right)_{ET_0} P + \left( \frac{\partial ET}{\partial P} \right)_P ET_0 \), which is a well-known property of homogeneous functions. This relation can be rearranged to become an equation for the Legendre transform of \( ET/P \). The Legendre transform of any function gives a mathematically-equivalent representation of the function in terms of the y-intercepts and slopes of tangent lines to the curve representing the function as a locus of points [23] (pp. 137–142) and [25] (pp. 63–65). For the Budyko equation, the Legendre transform of \( ET/P \) is in fact \( \left( \frac{\partial ET}{\partial ET_0} \right)_P \) expressed as a function of \( \left( \frac{\partial ET}{\partial P} \right)_P \). Thus, specifying \( \left( \frac{\partial ET}{\partial P} \right)_P \) as a function of \( \left( \frac{\partial ET}{\partial ET_0} \right)_P \) is equivalent mathematically to specifying \( ET/P \) as a function of \( ET_0/P \).

3. The assumption of homogeneity also leads to the Gibbs–Duhem relation, \( P \left( \frac{\partial ET}{\partial P} \right)_{ET_0} + ET_0 \left( \frac{\partial ET}{\partial ET_0} \right)_{ET_0} = 0 \). This differential expression can be integrated to find \( ET/P \) as a function of \( ET_0/P \) if \( \left( \frac{\partial ET}{\partial P} \right)_P \) is given as a function of \( \left( \frac{\partial ET}{\partial ET_0} \right)_P \). Specific mathematical models of the Budyko equation can be developed, therefore, by specifying the dependence of \( \left( \frac{\partial ET}{\partial P} \right)_P \) on \( \left( \frac{\partial ET}{\partial ET_0} \right)_P \). Physically, these two partial derivatives represent the climate elasticity of catchment evapotranspiration. A model of the climate elasticity of \( ET \) is thus sufficient to determine a model of \( ET \) as a function of \( P \) and \( ET_0 \) through the Gibbs–Duhem relation.

4. The two leading parametric models of the Budyko equation were developed by modeling climate elasticity: \( \left( \frac{\partial ET}{\partial P} \right)_P \) and \( \left( \frac{\partial ET}{\partial ET_0} \right)_P \) were postulated as functions of \( ET/P \) and \( ET_0/P \), then the resulting set of partial differential equations was integrated to calculate an explicit form of \( F(ET_0/P) \). This approach is an alternative implicit way of specifying \( \left( \frac{\partial ET}{\partial P} \right)_P \) as a function of \( \left( \frac{\partial ET}{\partial ET_0} \right)_P \). Its rigorous mathematical justification is that it amounts to specifying the Legendre transform of \( ET/P \), then integrating the Gibbs–Duhem relation.
5. However, given \((\partial ET/\partial P)\) and \((\partial ET/\partial ET_0)\) as functions of \(ET/P\) and \(ET_0/P\), the need to solve partial differential equations can be eliminated by simply introducing the postulated relationships into the Euler relation, then rearranging it to deduce a model form of \(ET/P = F(ET_0/P)\). In the parlance of equilibrium thermodynamics, \((\partial ET/\partial P)\) and \((\partial ET/\partial ET_0)\) specified as functions of \(ET/P\) and \(ET_0/P\) are “equations of state” related to catchment ET [25] (p. 69). In equilibrium thermodynamics, the partial derivative of the internal energy with respect to the volume defines the pressure, and the pressure specified as a function of the variables on which the internal energy depends is termed an equation of state related to the internal energy (Münster, 1970). The fact that knowledge of \((\partial ET/\partial P)\) and \((\partial ET/\partial ET_0)\) as functions of \(ET/P\) and \(ET_0/P\) is sufficient to determine \(ET/P\) as a function of these two latter variables is analogous to the dictum that knowledge of all of its equations of state is equivalent to knowledge of the internal energy [25] (p. 73).

As noted in Sections 3.3 and 3.4, a frequent application of parametric models of the Budyko equation has been to estimate the climate elasticity so as to predict the effects of changes in \(P\) and \(ET_0\) on catchment evapotranspiration and, therefore, runoff. This is accomplished, for example, by calculating \((\partial ET/\partial P)\) and \((\partial ET/\partial ET_0)\) using either Equations (33)–(35) or Equations (40) and (41), along with chosen values of the aridity index and the model parameter, for a selected datum state of a catchment to predict changes in ET and Q resulting from prescribed changes in \(P\) and \(ET_0\). Roderick and Farquhar [26] have illustrated this approach for the Murray-Darling Basin in southeast Australia using the MCY model. They first calibrated the model using hydrologic and climatic data for the period, 1895–2006, then used it to calculate average values of the two partial derivatives using the average values of \(ET\), \(P\), and \(ET_0\) for the period. With these estimates and the known change in \(P\) and \(ET_0\) over the period, 1997–2006, they successfully predicted the changes in ET and Q over this latter period. Evidently the same approach could be used with coupled atmosphere—ocean general circulation models (e.g., CMIP5) generating future scenarios for \(P\) and \(ET_0\) in order to predict the effects of climate change on catchment ET and Q [26].

Predictions of climate elasticity, such as the one undertaken by Roderick and Farquhar [26], are carried out with the model parameter held constant; but, as noted in Section 3.3 and 3.4, the effect of variations of the model parameter on ET also has been studied under the assumption that the parameter represents catchment properties other than \(P\) and \(ET_0\), properties such as green water storage or vegetation cover. Wang et al. [12] have reviewed these efforts, focusing on the MCY and Fu models. For example, they note that variations in the Fu model parameter among catchments have been correlated with vegetation cover (forest vs. grassland, land surface coverage) and topographic properties of catchments. Because these efforts to interpret the adjustable parameters in the MCY and Fu models are strictly empirical—the parameters \(n\) and \(w\), which appear solely as arbitrary constants in derivations of the two models [15,16], are correlated statistically with catchment properties—it is worthwhile to provide additional context for understanding these parameters by considering further the physical implications of the homogeneity assumption underlying the Budyko equation.

Borrowing again from the terminology of equilibrium thermodynamics [36] (p. 135), one can say that different catchments having the same value of \(ET_0/P\) are in “corresponding states,” and the Budyko equation, implying that different catchments having the same aridity index have the same value of \(ET/P\), can accordingly be considered an example of a “principle of corresponding states.” However, this correspondence principle cannot be upheld if a model of the Budyko equation is applied which contains an adjustable parameter, such as \(n\) or \(w\), which takes on different values for different catchments. This is because the Budyko equation implies that, in respect to \(ET\), \(P\), and \(ET_0\), any two catchments at different geographic locations, or the same catchment considered at two different time periods, are merely scaled versions of one another, with their respective values of \(ET\), \(P\), and \(ET_0\) being related by the scaling operation in Equation (6). This scaling relationship is undermined if a parametric model of the Budyko equation allows different values of the model parameter for different catchments.

Thus, because of these broad implications of homogeneity, a fundamental contradiction exists between the Budyko equation and the MCY and Fu models if their adjustable parameters are allowed...
to vary from one catchment to another, or from one time-period to another for a given catchment. Although the justification for allowing such variability has been that it represents the effects of differing catchment properties other than P and ET$_0$ [12], the only evidence for this assumption is statistical, since it has not emerged from any derivation of a parametric model for the Budyko equation. An interpretation of an adjustable model parameter as representing non-climatic catchment properties must evolve naturally from the physical reasoning underlying the development of the model, including, in the case of the MCY and Fu models, an explanation of why the non-climatic parameter should appear specifically as an exponent in a power-law expression.

It must be concluded, therefore, that the current physical interpretation of the MCY and Fu model parameters may be spurious. Gentine et al. [21], noting past successes of the non-parametric interpolation formula proposed by Budyko [10] (p. 325), showed that it provides an excellent model (< 10% deviation of experimental data from model estimates) of long-term-average annual ET for more than 300 catchments encompassing a broad range of ecological, soil, and climatic properties. Taking a view opposed to parametric modeling, they proposed that the Budyko interpolation formula could be used to understand the variability of catchment soil and vegetation parameters by constraining models of catchment water balance in which these parameters appear with the Budyko formula to infer how they vary with the aridity index and thereby reveal the ecohydrological controls on catchment water balance. The error in this line of reasoning must be demonstrated in order to substantiate claims that parametric models of the Budyko equation inherently take into account the dependence of ET on non-climatic catchment properties.

Acknowledgments: This research was supported in part by funds allocated in connection with the author’s appointment as Chancellor’s Professor, Emeritus, in the University of California at Berkeley. The author thanks Professor Juan V. Giráldez, University of Córdoba (Spain), for drawing his attention to the paper by Zhang et al. (2004) and for many useful discussions on hydrology over the past 40 years.

Conflicts of Interest: The author declares no conflict of interest. The funding sponsor had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References


© 2017 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).