“Efficiency Space” –
A Framework for Evaluating
Joint Evaporation and Runoff Behavior

Randal Koster
Global Modeling and Assimilation Office, NASA/GSFC, Greenbelt, MD

Corresponding author address:
Randal Koster
Global Modeling and Assimilation Office
Code 610.1, NASA/GSFC
Greenbelt, MD 20771
randal.d.koster@nasa.gov
Abstract

At the land surface, higher soil moisture levels generally lead to both increased evaporation for a given amount of incoming radiation (increased “evaporation efficiency”) and increased runoff for a given amount of precipitation (increased “runoff efficiency”). Evaporation efficiency and runoff efficiency can thus be said to vary with each other, motivating the development of a unique hydroclimatic analysis framework. Using a simple water balance model fitted, in different experiments, with a wide variety of functional forms for evaporation and runoff efficiency, we transform net radiation and precipitation fields into fields of streamflow that can be directly evaluated against observations. The optimal combination of the functional forms – the combination that produces the most skillful streamflow simulations – provides an indication for how evaporation and runoff efficiencies vary with each other in nature, a relationship that can be said to define the overall character of land surface hydrological processes, at least to first order. The inferred optimal relationship is represented herein as a curve in “efficiency space” and should be valuable for the evaluation and development of GCM-based land surface models, which by this measure are often found to be suboptimal.

Capsule Summary: The relationship between evaporation and runoff at the land surface is captured in a simple “efficiency space” framework.
1. Introduction: Efficiency Space

A land surface model, or LSM, is an essential component of a climate modeling system. The numerous studies that have quantified the impact of land processes on simulated climate (e.g., Shukla and Mintz 1982, Delworth and Manabe 1989, Dirmeyer 2000, Koster et al. 2000, Seneviratne et al. 2006, to name a few; see Seneviratne et al. 2010 for a review) underscore the importance of realistic treatments of land processes in climate models. Recognition of this importance has spawned multi-national and multi-institutional LSM evaluation projects such as the Project for the Intercomparison of Land-surface Parameterization Schemes, or PILPS (Henderson-Sellers et al. 1993, Chen et al. 1997, Wood et al 1998, Bowling et al. 2003), the Global Soil Wetness Project (Dirmeyer et al. 1999, Boone et al. 2004, Dirmeyer et al. 2006), and the more recent benchmarking project PLUMBER (PALS Land Surface Model Benchmarking Evaluation Project, Best et al. 2014).

One of the findings from PILPS is that in terms of LSM behavior and performance, the evaporation and runoff formulations of an LSM are inextricably linked (Koster and Milly 1997). (Note that throughout this paper, the term ‘evaporation’ is used in place of the more cumbersome ‘evapotranspiration’ to encompass all evaporation processes.) A land surface scheme’s treatment of the evaporation process has a profound impact on the runoff it generates; similarly, the scheme’s treatment of the runoff process has a first-order impact on evaporation. Perhaps counter-intuitively, an LSM with an excellent evaporation formulation will, when forced with realistic meteorology, produce poor evaporation rates if the model’s runoff formulation is poor, and it will accordingly lead to biases in any atmospheric model that is connected to it. To produce a reliable treatment of land surface processes in an Earth system model, it is arguably critical to understand the evaporation-runoff connection.
The framework introduced herein can serve as a tool for investigating this connection. Building on the work of Koster and Milly (1997) and Koster and Mahanama (2012), the framework captures the essence of the evaporation-runoff connection in a very simple and concise way: as a curve lying in “efficiency space”.

In essence, the efficiency space framework is built on the assumption that evaporation efficiency (the ratio of latent heat flux, \( \lambda E \), to net radiation, \( R_{\text{net}} \), where \( \lambda \) is the latent heat of vaporization) is a function of soil moisture, \( W \):

\[
\lambda E / R_{\text{net}} = \beta (W),
\]

(1)
and on the further assumption that runoff efficiency (the ratio of runoff production, \( Q \), over some time period to precipitation, \( P \)) is also a function of soil moisture:

\[
Q / P = F (W).
\]

(2)

Under these two assumptions, the two efficiencies (i.e., \( \lambda E / R_{\text{net}} \) and \( Q/P \)) can be related to each other, as illustrated in Figure 1a. The red curve on the right shows an arbitrary functional relationship between \( \lambda E / R_{\text{net}} \) and \( W \), and the blue curve directly below it shows a similarly arbitrary relationship between \( Q/P \) and \( W \). Note that for mathematical tractability, both evaporation efficiency and runoff efficiency are assumed throughout this paper to increase monotonically with soil moisture, reflecting the reasonable idea that wetter soils allow an easier production of both evaporation and runoff. To capture the well-known plateau for wetter soils associated with energy-limited evaporation (e.g., Manabe, 1969), the rate of increase at high soil moisture can be set very low, as in the figure.

For a soil moisture of 0.3 (degree of saturation), these functions produce an evaporation efficiency of about 0.3 and a runoff efficiency of about 0.05. For a soil moisture of 0.5, the evaporation and runoff efficiencies are about 0.5 and 0.1, respectively. These pairings of
evaporation and runoff efficiencies, along with every other possible pairing determined from the
two functions, are plotted in “efficiency space” on the left side of Figure 1a. The efficiency
space plot essentially shows how evaporation and runoff efficiencies vary with each other – how
efficient, for example, the land surface is at converting precipitation into runoff when it is able to
convert, say, 40% of net incoming radiative energy into evaporation, and how the runoff
efficiency changes when the evaporation efficiency is reduced to 30% through the drying of the
soil.

Again, the evaporation and runoff efficiency functions on the right in Figure 1a are
arbitrary; a different set of functions would lead to a different curve in efficiency space. Figure
1b shows another possible example. The number of possible efficiency function combinations,
of course, is infinite. The corresponding Q/P-vs-$\lambda E/R_{\text{net}}$ curves could lie anywhere in efficiency
space, subject to the condition that Q/P and $\lambda E/R_{\text{net}}$ increase monotonically with each other.

This paper has two overarching goals: (i) demonstrating that the identification of a curve
in efficiency space is tantamount to characterizing overall hydrological behavior, and (ii) finding
the curve that best represents nature and thus best serves as a target for LSM development. It is
important to emphasize here that simple functional forms such as those on the right in Figure 1a
are, at best, crude first order approximations to the much more complex behavior seen in land
surface models and (presumably) in nature itself. Nevertheless, the functional forms capture
enough of the underlying controls of soil moisture on evaporation and runoff to make the curve
on the left meaningful and to allow nature’s efficiency space curve, however it looks, to embody
in a valuable and concise way an essential aspect of surface hydrology.

Given the difficulty of measuring evaporation and soil moisture at the large scale, the
efficiency functions underlying nature’s efficiency space curve, and thus the curve itself, cannot
be derived directly from observations. This paper thus employs an indirect approach to search
for nature’s curve: the use of a simple water balance model in conjunction with streamflow,
precipitation, and radiation observations. The approach and the main results are presented in
Section 2. The paper continues in Section 3 with a discussion of the efficiency space framework
in the context of Budyko’s (1974) landmark hydroclimatic analysis, and Section 4 discusses the
use of the framework for LSM development. The summary section (Section 5) includes a
discussion of the framework’s use in the context of the new soil moisture data expected from
recent and upcoming satellite missions.

Underlying all of the discussions in the paper is the fundamental idea that it is the joint
response of evaporation and runoff production to variations in soil moisture that underlies
hydrological behavior. Focusing inordinately on one or the other in LSM development is likely
to lead to inadequate LSM performance.

2. Analysis of Efficiency Space with a Simple Water Balance Model

a. Model Structure

The tool used in this paper to evaluate different curves in efficiency space is the simple
water balance model (WBM) of Koster and Mahanama (2012). The essence of the model is
illustrated in Figure 2. In addition to a soil water holding capacity, the user chooses the model’s
functional efficiency relationships [i.e., λE/R_{net} = β(W); Q/P = F(W)]; shown in the figure are
arbitrary sample functions. The model is then driven with observations-based daily precipitation
and net radiation forcing over a long period of time, typically several decades. At a given time
step, the WBM’s β(W) function, in conjunction with the net radiation input and the current soil
moisture state, determines the evaporation for the time step. Similarly, the WBM’s F(W) function, in conjunction with the precipitation, determines the amount of runoff generated. The WBM’s soil moisture state is updated based on the prescribed precipitation rate and the computed evaporation and runoff values. Precipitation is added to a snowpack reservoir during subfreezing periods; snowpack water melts and is added to the soil (essentially as precipitation) as the prescribed temperature warms.

This WBM is undeniably much simpler than the land surface models (LSMs) typically used in climate modeling systems. The WBM lacks temperature prognostic variables and associated energy balance calculations, and it includes no specific treatment of (for example) baseflow and interception loss. Also, unless otherwise stated, most applications of the WBM here utilize the same efficiency functions in every season of the year and across large continental domains (i.e., without regard to spatial variations in topography, soil texture, vegetation type, and other features that may affect efficiency – shown in some studies to be important [e.g., Yildiz and Barros [2007]]). This simplicity, while inappropriate for an LSM working within an Earth system model, is nevertheless justified for the WBM given that it is designed to contain only the most important, first-order controls on evaporation and runoff production contained within a more complex LSM (generally implicitly in the latter – the net result of complex interacting parameterizations), thereby allowing the isolated study of these controls. The numerous features not included in the WBM are thus assumed, when acting in a full LSM, to induce mainly second-order modifications to the fluxes. While the appropriateness of such an assumption can be argued, and while more complex versions of the WBM can easily be envisioned, it must be remembered that each addition of complexity brings the WBM closer to a full LSM and thereby makes the WBM results more difficult to interpret. The WBM, despite its
simplicity, has been found in tests to reproduce successfully the first-order behavior of complex, state-of-the-art LSMs. (See Koster and Mahanama (2012) for one example.)

b. Simulation Procedure

Following precisely the procedure of Koster and Mahanama (2012), the WBM is run on a 2.5° x 2.5° grid across the conterminous United States (CONUS) over the period 1948-2000. The interannually-varying precipitation rates used to force the WBM are from the dataset of Andreidis et al. (2005), and the net radiation forcing used consists of climatological seasonal cycles derived from the Surface Radiation Budget (SRB) dataset (Gupta et al. 2006). The final 52 years of simulated streamflow data are evaluated against observed streamflow data, which consist of naturalized streamgage measurements in several large hydrological basins (see Mahanama et al. [2012] for details).

Figure 3a illustrates a sample simulation. The red and blue curves in the left panel show, respectively, the evaporation efficiency and runoff efficiency functions prescribed in the WBM. The WBM is driven across CONUS for 53 years, and the runoff rates produced in the grid cells lying within the Upper Mississippi Basin (upstream of the Grafton stream gauge site) are combined into a basin-average annual rate for each simulation year. These simulated annual rates are compared to observed streamflows in the rightmost panel of Figure 3a. This particular combination of efficiency functions is seen to lead to significantly underestimated streamflow.

Figure 3b shows a second sample simulation. The efficiency functions applied here are clearly different, particularly that for evaporation efficiency – in this second simulation, the ability of the WBM to convert net radiation energy to evaporation is, in general, reduced. As a
result, the simulated streamflows for the Upper Mississippi have increased and are now, in fact, too high.

c. Collapse of Functions onto Efficiency Space

Figure 4 illustrates a finding that underlies much of the analysis in this paper. Three pairings of efficiency functions are provided in the top row of the figure. At first glance the pairings look fundamentally different, and yet a closer look shows that all three map onto the same single curve in efficiency space (lower left panel). Furthermore, all three pairings, when implemented into the WBM, produce essentially the same hydrological behavior – the lower right panel of Figure 4 shows that the time series of runoff for the three cases is essentially the same. This example is, in fact, representative. In general, pairings of efficiency functions that produce the same curve in efficiency space generate essentially the same hydrological fluxes in the WBM.

This result can be considered in the context of the water balance equation:

\[ P = E + Q + C_w \Delta W / \Delta t, \]  

(3)

where \( C_w \) is the water holding capacity of the WBM and \( \Delta t \) is the time period considered. Using (1) and (2), this can be written as

\[ P = E + P G[\lambda E/R_{net}] + C_w \Delta W / \Delta t, \]  

(4)

where \( G \) is the functional relationship between \( \beta(W) \) and \( F(W) \) in efficiency space. Now consider the idealized limit of zero water holding capacity. In this case, given that \( P \) and \( R_{net} \) are inputs, the equation collapses into one that defines a unique value of \( E \) for a time step and,
accordingly, also a unique value for \( Q \). In other words, regardless of the evaporation function employed – regardless, for example, of which red curve in the top row of Figure 4 is used – a given curve in efficiency space (i.e., a given \( G[\lambda E/R_{\text{net}}] \) function) implies a single set of values for evaporation and runoff and thus the same overall hydrological behavior.

Naturally, the water holding capacity term is not necessarily negligible, and this explains the slight differences seen in the simulated streamflow curves in Figure 4. Again, though, the overall strong agreement between the streamflow curves in this example is found to be standard through a comprehensive series of tests – for practical purposes, pairings of functions that map to the same efficiency space curve are found to be hydrologically very similar, supporting the use of efficiency space curves as an efficient means of characterizing the joint behavior of evaporation and runoff functions.

d. Efficiency Space Curves

i. Overall ranking. The efficiency space curve in the lower left panel of Figure 4, regardless of the pairing of efficiency functions that produced it, leads to streamflow values that are too low relative to observations (lower right panel). This efficiency space curve is reproduced in Figure 5a as the yellow curve, with the yellow color chosen to reflect this level of streamflow simulation skill. (Although no actual skill numbers are shown here, as this discussion focuses only on the relative skill levels associated with different curves, the skill is quantified with a root-mean-square error, or RMSE, calculation against observed ratios of annual streamflow to annual precipitation in the basin.) A pairing of efficiency functions that leads to the blue curve in Figure 5a produces WBM streamflow values that are even more inaccurate (this is indeed the
pairing underlying Figure 3b), whereas a pairing that leads to the red curve in Figure 5a produces a relatively realistic simulation.

Using this color-coding scheme for streamflow simulation accuracy, and plotting the higher scoring curves on top of the lower scoring ones, Figure 5b shows the results of testing more than 23000 curves in efficiency space. Figure 5b (and corresponding plots for other basins) holds a substantial amount of information and is indeed a centerpiece result of this paper. It shows, for example, that any curve passing through the northwest quadrant of efficiency space will produce a poor WBM simulation of streamflow. Most importantly, given that observations are used to evaluate the skill levels, it provides an indication of the curve that effectively operates in nature across the Upper Mississippi Basin. The darkest red curves correspond to the highest skill values; any dark red curve will produce an accurate simulation and is thus a reasonable estimate of nature’s curve.

The highest scoring curve (again, not by a large margin) is highlighted with a heavy black line. Assuming that this particular curve reflects nature, we can infer the following: during dry conditions, runoff production does not begin significantly until the evaporation efficiency is about 0.3, and the runoff ratio is not especially sensitive to changes in soil moisture (i.e., relative to the sensitivity of evaporation efficiency to soil moisture) until the soil is wet enough to support an evaporation efficiency of about 0.6, at which point the runoff ratio becomes strongly sensitive to soil moisture. Evaporation efficiency, in contrast, is strongly sensitive to soil moisture variations during dry conditions and is much less sensitive to soil moisture variations during wet conditions.
ii. *Bearing of curve position on simulated streamflows.* Figure 6 is provided to illustrate how different curves in efficiency space translate to different simulations of streamflow. When the WBM is fitted with evaporation and runoff efficiency functions corresponding to the black curve in Figure 5b, the resulting simulation of streamflow is, as expected, quite accurate (Figure 6a). Again, though, any curve lying in the dark red region of the plot will produce a reasonably accurate simulation. The right panel of Figure 6b, for example, shows the WBM streamflows associated with the white curve in the left panel of Figure 6b, a curve which lies within the dark red region. The simulated streamflows are only slightly less accurate than those in Figure 6a, and given presumed uncertainties in the streamflow observations and in the meteorological forcing data used, one simulation cannot be said to be truly superior to the other.

Worth mentioning here is that the skill levels shown in Figure 5b are robust. If only the first half of the meteorological forcing and streamflow observations are used to establish the skill scores, the resulting optimal curves (not shown) are found to be very similar to those in the figure and result in similarly accurate streamflow simulations during the second half of the study period.

Figure 6c shows that when the efficiency space curve passes through the northwest quadrant, the resulting streamflow simulations are grossly overestimated. It seems safe to conclude that this curve does not come close to representing nature. On the other hand, when the efficiency space curve lies far below the black curve, as in Figure 6d, the simulated streamflows are significantly underestimated. Figure 6e shows the results obtained for an insensitivity of runoff efficiency to soil moisture. The simulated streamflows do vary slightly from year to year, but as expected, their interannual variability appears significantly underestimated relative to observations.
e. Water Holding Capacity

The two main user-defined elements of the WBM are the evaporation efficiency and runoff efficiency functions, which, as shown above, can be described together as a single curve in efficiency space. The remaining user-prescribed WBM parameter is the water holding capacity, $C_w$. In the simulations discussed above, $C_w$ is set to 169 mm, based on a chosen depth of 0.5 m, an assumed wilting point at a degree of saturation of 0.25, and a porosity of 0.45. For completeness, it is worth examining how the WBM results vary with $C_w$.

The efficiency space skill analysis in Figure 5b is repeated in Figure 7 for depths of 0.1 m, 0.5 m, 1.0 m, and 1.667 m; again assuming a porosity of 0.45 and a wilting point of 0.25, these depths correspond to $C_w$ values of 34 mm, 169 mm, 338 mm, and 563 mm. Here the skill is evaluated over all of the shaded basins in Figure 7a collectively, with the skill contribution from each basin weighted by basin area.

Two features from Figures 7b-7e stand out. First, while the optimal (black) curves differ somewhat between the four panels, the regions of high skill within efficiency space (the darker red areas) are basically the same – the highest scoring curves tend to hug the x-axis for lower values of $\lambda E/R_{net}$ and then climb sharply into the plot’s northeast quadrant for higher values of $\lambda E/R_{net}$. This basic curve shape is also seen for individual basins and other depths (not shown). The estimation of a valid relationship between evaporation efficiency and runoff efficiency functions is largely insensitive to the choice of $C_w$.

The second interesting feature of Figures 7b-7e is the fact that higher skill levels are possible with a depth of 0.5 m or 1.0 m than with a depth of 0.1 m or 1.667 m. (The four panels
use the same color bar to illustrate skill, with levels that slightly differ from those used in Figure 5.) This result could be interpreted in terms of the depth of soil “actively” participating in the land-surface hydrological cycle – perhaps, in a gross sense, changes in the soil water storage within the top 0.5 m or 1 m of soil are most relevant to the interannual variability of surface hydrological fluxes. Curiously, this depth range is roughly consistent with nominal estimates of vegetation rooting depth. Such agreement is perhaps not a coincidence.

f. Discussion

The patterns shown in the efficiency space skill diagrams are entirely determined by the imposed observations-based precipitation and net radiation forcing and by the streamflow measurements used to evaluate the WBM output. Thus, the precipitation, net radiation, and streamflow observations, when examined jointly with the WBM, can be said to contain important and otherwise hidden information: (i) the fact that (based on the shape of the high scoring curves) evaporation’s sensitivity to soil moisture variation is greater than that of runoff in dry situations, whereas the reverse is true in wet situations; and (ii) the soil depth of greatest relevance to the generation of hydrological fluxes is perhaps of the order of 0.5 m to 1 m (Figure 7). Most importantly, the high-scoring efficiency space curves suggest how evaporation and runoff efficiencies vary with each other in nature. This joint variation is critical to defining a basin’s hydrological behavior, at least in terms of how much runoff it produces each year; any other version of the joint variation (i.e., any curve in efficiency space lying outside the high-scoring region) produces inaccurate streamflows.
The high scoring curves for individual basins (not shown) are similar to first order, and yet they do nonetheless show some differences – there is, of course, some spatial heterogeneity in how nature behaves. Preliminary analyses of the WBM results indicate, for example, that the efficiency space curves for mountainous basins tend to lie a little above those for flatter basins in the dry regime, supporting the intuitive notion that mountainous areas produce runoff more easily than flat areas. Also, preliminary analyses suggest that flatter areas can produce higher skill scores with deeper soil depths, consistent with the fact that deeper depths-to-bedrock are found in flatter areas (e.g., as seen in STATSGO data [NRCS, 2012]). While a proper analysis of the spatial heterogeneity of basin behavior is beyond the scope of this paper, the potential for such heterogeneity must be kept in mind when interpreting the results presented herein.

3. The Budyko Perspective

Budyko (1974), building on analyses by Schreiber (1904) and Ol’dekop (1911), pioneered the joint analysis of evaporation and runoff response to precipitation and net radiation forcing. Although Budyko’s analyses are focused on climatic means – how, for example, annual mean evaporation varies with annual mean precipitation and net radiation – rather than on the short time scale (daily time steps or shorter) of relevance here, a comparison of Budyko’s results with those seen in the efficiency space diagrams is of interest.

Budyko’s equation for evaporation is:

$$\frac{E}{P} = \left[ \left( \frac{R_{\text{net}}}{P\lambda} \right) \tanh \left( \frac{P\lambda}{R_{\text{net}}} \right) \left( 1 - \cosh \left( \frac{R_{\text{net}}}{P\lambda} \right) + \sinh \left( \frac{R_{\text{net}}}{P\lambda} \right) \right) \right]^{1/2}.$$  (5)
This function is plotted in Figure 8a. If we let the variable D represent the dryness index (i.e., $D = \frac{R_{\text{net}}}{\lambda P}$), then Budyko’s equation can be rewritten:

$$\frac{\lambda E}{R_{\text{net}}} = f(D);$$

(6)

that is, the climatological evaporation efficiency is some complex function of dryness index.

Now consider that for long-term averages, soil moisture variations in the water budget equation are irrelevant, so that the water balance can be written:

$$P = Q + E.$$  

(7)

Combining (6) and (7) then provides an equation for the climatological runoff efficiency:

$$\frac{Q}{P} = 1 - D f(D).$$

(8)

Together, (6) and (8) allow the equivalent of Budyko’s climatological equation to be plotted in efficiency space. The idea is simple: a given value of D leads, through the two equations, to a unique pairing of $\frac{\lambda E}{R_{\text{net}}}$ and $\frac{Q}{P}$ and thus to a unique point in efficiency space. Examining a wide range of D values leads to a wide range of points, allowing the full curve to be plotted within the space.

This curve is shown in white in Figure 8b. The color-coded skill levels appearing in the plot are taken from Figure 7c (i.e., they represent the WBM-based skill levels over the gray-colored basins in Figure 7a for a depth of 0.5 m). Two aspects of the Budyko curve in Figure 8b stand out. First, the curve has the same basic shape as the WBM-based high-scoring curves; indeed, Budyko’s curve falls within the high-scoring region of the space, though only just barely. Thus, even though Budyko’s equation is built to capture climatological efficiencies and implicitly (if empirically) accounts, for example, for the effects of seasonal phase differences in
the precipitation and net radiation forcing (Koster et al. 2006), the equation produces a curve in
efficiency space that is not grossly inconsistent with the WBM-based curves, which focus on the
short time-scale problem.

Second, the Budyko-based curve is nevertheless seen to overestimate evaporation
efficiency at the wet end, if we assume that either the optimal curve (in black) or the median
curve within the high scoring region of the space is most correct. When the soil is fully wet (at
the northeast corner of the plot), the Budyko curve implies that all of the net radiation will be
converted into latent heat flux, i.e., that the land surface provides no resistance whatsoever to
evaporation, even given the complexity of transpiration pathways through the vegetation. If such
an extreme were true for climatological means, arguably it would also need to be true at any
smaller time step, thus contradicting the median behavior of the WBM results. Some past
studies (e.g., Koster et al. 2006) have indeed found observational evidence that Budyko’s
estimates of climatological evaporation at the wet end are too high.

4. Relevance to Land Model Development

a. LSM Evaluation and the Model Dependence of Soil Moisture

When evaluating the realism of an LSM, a key element to consider is the manner in
which it allows variations of soil moisture to affect the magnitudes of evaporation and runoff
fluxes. Figure 9 shows the diagnosed relationship between evaporation efficiency and soil
moisture for three separate and representative LSMs (names withheld). The abscissa of each
point in each scatter plot corresponds to a monthly-averaged root zone soil moisture, and the
ordinate of each point is the ratio of monthly-averaged evaporation to monthly-averaged net
radiation; the data, taken from archived output of multidecadal offline simulations, are plotted
here for points in and about the CONUS area during warm season months (May – September).
A significant amount of scatter is seen in these plots, reflecting the fact (Section 1) that LSMs
compute evaporation and runoff using a complex set of interacting parameterizations and thus do
not generally use such efficiency functions directly. (These plots indeed demonstrate that the
WBM is substantially simpler than a standard LSM.) Nevertheless, despite the scatter, a first
order relationship between $\lambda E/R_{\text{net}}$ and soil moisture can be seen for each model.

Figure 9 shows why evaluating such relationships directly can be difficult. Within Model
A, for example, the minimum soil moisture achieved varies spatially due to spatial variations in
soil and vegetation properties, such as wilting point; for this model, there is no single relevant
evaporation efficiency function. More importantly, evaluating an LSM’s efficiency function is
difficult because the LSM’s soil moisture variable is model-specific, more of an “index of
wetness” than a quantity that can be compared directly to observations (Koster et al. 2009). Such
model dependency is reflected, for example, in the contrasting soil moisture ranges of Models B
and C. Soil moisture and evaporation observations at the large scale are notoriously sparse, but
even if they were plentiful, using them to evaluate a model’s evaporation efficiency relationship
would be hampered substantially by this model dependency.

The efficiency space framework, however, provides an elegant means for avoiding this
problem. According to this framework, the model-dependent shapes of an LSM’s $\lambda E/R_{\text{net}}$-vs-$W$
and $Q/P$-vs-$W$ relationships are of secondary importance; the key relationship defining the
LSM’s hydrological behavior is that between evaporation efficiency and runoff efficiency, a
relationship that can be examined without the direct consideration of soil moisture. The LSM
can thus be evaluated by comparing its effective efficiency space curve against the
corresponding observations-based curve (e.g., the black curve in Figure 7c).

b. Using Efficiency Space to Evaluate State-of-the-Art Land Models

Using archived model output to determine directly the efficiency space curve that best
characterizes a given LSM’s behavior comes with difficulties. Chief among these is the lack of
an optimal averaging period over which to accumulate the evaporation, runoff, precipitation, and
net radiation diagnostics. If the averaging period is too short (e.g., a week or so), then the runoff
associated with baseflow, with its associated time delay, is no longer tied strongly to the
concurrent precipitation. If, on the other hand, the averaging period is too long (e.g., a season),
seasonal variability in soil moisture will confound the curve’s derivation – most of the runoff, for
example, may be produced during a month of snowmelt, when soil moisture is high, whereas
most of the evaporation may be produced a couple of months later, when the soil moisture is
lower. Monthly diagnostics are a reasonable compromise, but even these, to some extent, are
subject to such problems.

To compute the LSM-based efficiency space curves shown in Figure 10, a different
approach is employed, one that avoids these problems and has the additional advantage of
producing a simple, smooth curve in the space. In essence, the approach used to compute the
highest scoring (black) curve in efficiency space in Figures 5b and 7 is repeated, this time using
the full LSM’s time series of streamflow values (rather than the observations) as the “truth” upon
which to base the RMSE metric. The precipitation and net radiation forcing applied to the WBM
is derived from that used originally by the LSM in question, when it was generating its archived data. Thus, whereas the black curves in Figures 5b and 7c can be said to be in line with how nature works, the curve produced using the LSM-based “streamflow truth” can be said to be in line with how the LSM works – it can be said to characterize, in a gross sense, the LSM’s underlying relationship between evaporation and runoff efficiencies.

The left panel of Figure 10a shows, in white, the optimal efficiency space curve established for a specific LSM (name withheld) based on its streamflow generation in the gray-colored basins of Figure 7a; plotted around it are the corresponding observations-based color-coded skill curves from Figure 7c. The curve for the LSM clearly lies above the high-scoring region of the space (where “nature’s curve” lies), especially toward the dry end. Based on the discussion of Figure 6c, this LSM can be expected to compute too much runoff. Sure enough, the right two panels in Figure 10a show that the annual runoff ratios computed by the full LSM, as determined from the archived diagnostics, are too large compared to observations, with the largest errors seen toward the western half of CONUS.

Note that a traditional interpretation of the comparison in the rightmost panels of Figure 10a might be that the runoff formulation in this LSM is deficient. This may be true, but the efficiency space framework allows the deficiency to be stated in a potentially more helpful way: at the drier end, this LSM’s runoff efficiency is too large relative to its evaporation efficiency.

Figure 10b shows the optimal curve obtained for a different LSM. Here the LSM-based curve falls below the high-scoring region, and, consistent with the discussion of Figure 6d, the resulting streamflows are underestimated relative to observations (rightmost panels). Again, the efficiency space framework allows us to describe this LSM’s deficiency in terms of its joint
evaporation and runoff behavior rather than solely in terms of its runoff formulation. Correcting this LSM – bringing the white curve into the high-scoring region of the plot – could be a matter of improving its runoff formulation (increasing Q/P for a given soil moisture), but alternatively, and perhaps just as effectively, it could also result from improvements in the evaporation formulation (decreasing $\lambda E/R_{net}$ for a given soil moisture).

Six other state-of-the-art LSMs were examined in the same way. The results (not shown) indicate that most of these LSMs are similar to Model Y in Figure 10b – most of the LSMs are characterized by efficiency space curves that fall below the high scoring region.

c. Potential for Simple Land Model Tuning

Model Y in Figure 10b is in fact the Catchment-CN LSM (Koster et al. 2014), a merger of the energy and water balance framework of the Catchment LSM (Koster et al. 2000) and the dynamic phenology (prognostic biogeochemistry) components of the NCAR/DOE CLM4 dynamic vegetation model (Oleson et al. 2010). As suggested by the figure, this relatively new LSM could benefit from further development or tuning of its evaporation and runoff formulations.

When faced with the discrepancy in Figure 10b (i.e., the LSM’s optimal curve, in white, lying outside the plot’s high scoring region), one could envision two basic model development approaches. The first approach is simply to keep improving the evaporation and runoff formulations until the position of the white curve moves upward into the high scoring region. In principle this is the desirable approach, for if it is successful, the LSM’s curve would end up being in the right place for the right reasons. With this approach, knowledge of the relative
strengths of the evaporation and runoff formulations – knowing which formulation can be trusted more, based on the breadth of analysis underlying each – would point to the formulation that currently acts as the “weak link” in the simulation of hydrological fluxes and is thus the most appropriate target for focused development. The position of the white curve relative to the high scoring region would indicate a direction for the change in the model formulation.

Improving a land model formulation through first principles, however, is far from trivial; depending on the accuracies desired, it may encompass an entire scientific career. Given time constraints and other considerations, the second model development approach – tuning an LSM formulation using the observations-based efficiency space curves – may, at least for some applications, be advantageous.

The tuning approach in its simplest form is illustrated in Figure 11. The first step involves deciding whether the LSM’s evaporation or runoff formulation can be trusted the most. Compared to some LSMs, the Catchment-CN LSM has a relatively complex runoff calculation; still, the complexity built into the runoff formulation pales in comparison to that built into the evaporation formulation, and thus the latter is considered, for purposes of this demonstration, to be more trustworthy. Plotted in Figure 11a are monthly (May through September) values of $\lambda E/R_{\text{net}}$ versus monthly values of root zone soil moisture. The scatter here is similar to that shown in Figure 9; again, though, a first-order relationship can be seen between the two quantities. The red curve in Figure 11a, obtained by averaging the $\lambda E/R_{\text{net}}$ values over soil moisture bins, is used in this exercise to characterize the relationship.

Figure 11b shows the observations-based efficiency space curves from Figure 7c. The black dotted curve is taken to be the “target” efficiency space curve for the tuned land model;
this curve, which falls wholly within the high scoring region, is used instead of the solid black
curve because it accounts for a larger range of $\lambda E/R_{\text{net}}$ values and thus works better for the tuning
procedure. With the curves in Figures 11a and 11b in place – the first showing a relationship
between $\lambda E/R_{\text{net}}$ and $W$, and the second showing one between $\lambda E/R_{\text{net}}$ and $Q/P$ – a relationship
between $Q/P$ and $W$ is easily derived (Figure 11c).

For the next step in the demonstration, the somewhat complex runoff formulation in the
Catchment-CN LSM is completely stripped out of the model and replaced by a much simpler
formulation, one in which the total runoff for a given time step is computed by multiplying the
time step’s incident precipitation by a $Q/P$ ratio taken from Figure 11c, based on the current
value of root zone soil moisture. The results of this tuning exercise are shown in Figure 12. The
tuning, in general, leads to improved streamflow simulation. The tuned LSM (Figure 12b) –
which still, by the way, maintains most of its complexity (its diurnally-varying energy balance
calculations, its spatial variation in vegetation, and so on) – produces annual runoff ratios that
are, for the most part, more in line with the observations (Figure 12c) than those obtained with
the original LSM (Figure 12a).

This is, of course, a simple and rather heavy-handed sample demonstration. A more
refined tuning exercise, one that isolates, for example, the particular facets of the runoff
formulation that are known to be especially weak or arbitrary and then modifies only those
facets, would in fact be much more satisfying and justifiable. Also, tuning the model separately
for each basin would presumably produce better basin-by-basin results than those shown in
Figure 12b. Given the difficulty of improving LSMs through first principles, a comprehensive
tuning study may lead to a straightforward and more reasonable way of inducing an LSM to
produce a more realistic simulation of hydrology. The present example serves merely to
illustrate how the information content of the observations-based efficiency space curves (e.g., Figure 7c) can work its way into an LSM formulation.

5. Summary and Discussion

The typical land surface model, or LSM, is constructed as follows: the modeler produces the best evaporation formulation and the best runoff formulation possible and then combines them (along with formulations for various other land processes) into a single system. The modeler then tests this LSM against observations of (for example) surface turbulent fluxes and local streamflow rates and uses the test results, in conjunction with his or her knowledge of current model weaknesses, to adjust the individual LSM components. While on the surface this seems reasonable, the evaporation and runoff processes within an LSM are often developed and examined independently, thereby ignoring the fundamental truth emphasized in this paper: a first order understanding of the hydrological behavior simulated by an LSM requires the joint analysis of its evaporation and runoff formulations, so that the variations of evaporation and runoff efficiencies with respect to each other are, to first order, understood.

To explore this joint variation, and to provide a potentially useful tool for developing, evaluating, and improving LSM formulations, the present paper introduces the concept of efficiency space. A simple water balance model (WBM) used in conjunction with precipitation, net radiation, and streamflow observations produces curves in efficiency space that illustrate, in a concise and simple way, how evaporation and runoff efficiencies tend to vary with each other in nature. Considering together all of the gray basins in Figure 7a, nature’s curve in efficiency space lies somewhere within the dark red region of one of the lower panels of Figure 7. In its
own way, such a curve is conceptually powerful, as it captures, to first order, the net effect of land surface processes on annual hydrological fluxes.

Of course, such an analysis strategy does not come without its caveats. As discussed in the text, for example, the WBM’s structure is necessarily very simple, being designed to capture only the most important controls on the surface moisture fluxes. Another caveat, not yet discussed, involves the “control volume” assumed for the analysis. One holistic view of evaporation from the land surface interprets it as the net result of a variety of interacting processes in the soil, in the vegetation canopy, and in the atmospheric boundary layer, all evolving together (Betts 2004). Betts (2004) argues that addressing the evaporation problem through only a subset of these processes can have significant limitations. While this is certainly a valid and useful view, the analysis of evaporation in the present paper effectively centers on an alternative control volume, one that represents the soil and vegetation from the top of the canopy to some depth below the surface (say, a meter or two). In a sense, this control volume represents the operating world of the land surface model component of a climate model – a system that is forced from without by precipitation and incoming radiation and that must partition this incoming water and energy into a variety of fluxes and storage changes. The choice here of this alternative control volume, with its neglect of a detailed consideration of boundary layer and cloud processes, is guided – and is in fact made necessary – by the particular focus of this paper, namely, the examination of the dual role of soil moisture in controlling both evaporation and runoff (streamflow) production.

Keeping these caveats in mind, the curve representing nature’s behavior in Figure 7c can be used as a target for LSM development. By noting where a model’s efficiency curve lies relative to that of nature, the modeler can adjust either the evaporation or the runoff formulation.
– presumably whichever one is known to be weaker – in the right direction to bring the curves in line. The result should be an overall improvement in the simulation of streamflow and evaporation. Using the efficiency curve as a target allows the model developer to sidestep a commonly encountered problem in the evaluation of evaporation and runoff formulations: the fact that soil moisture in an LSM is more of a model-dependent index of wetness than a quantity that can be compared directly to observations.

While focusing on efficiency space takes soil moisture out of the LSM development problem in a clean way, this must not be construed to imply that measurements of soil moisture are not useful for hydrological analysis and LSM development. The advent of satellite-based soil moisture measurements, particularly L-band estimates from the SMOS (Kerr et al. 2010) and SMAP (Entekhabi et al. 2010) sensors, opens the door to new and potentially powerful research into hydrological processes, including that associated with joint evaporation and runoff efficiencies. Consider, for example, the evaporation efficiency and runoff efficiency curves in Figure 13. The shapes of the functions here are arbitrarily drawn, but together they capture the basic shape of nature’s curve in efficiency space. As suggested in the figure, evaporation-induced changes in soil moisture during non-rainy periods should be smaller and more variable in time when the soil is in the drier regime, whereas rainfall-induced changes should be smaller and more variable in time in the wetter regime. One can imagine that by sorting soil moisture changes in the new satellite data into subsets associated with rainy and non-rainy periods and then examining these changes in the context of efficiency function pairings that correspond to nature’s efficiency space curve, one could establish where on that curve nature is operating at a given time – a non-traditional and potentially very useful way to characterize the current “overall hydrological state” of the measured region. Perhaps the analysis of the remotely sensed data
could even constrain better the curve’s location. Such research, of course, would have to address a number of issues, including the limited vertical penetration (relative to the root zone) of the L-band signal; still, the possibilities are intriguing.

In closing, it is worth noting again that the joint consideration here of evaporation and runoff processes is in some ways not at all new, as it echoes, for example, the landmark work of Budyko (1971) and his predecessors. These earlier works focused on climatological means; the present paper, in a sense, discusses the ideas in the context of the short time scale and thus in the context of the operation of LSMs. The hope here is that the efficiency space framework, which succinctly captures the evaporation-runoff connection, can be used to promote LSM development. In the long run, such development should lead to improvements in overall Earth system model performance.

Acknowledgments. The author is deeply grateful to Sarith Mahanama and Greg Walker for their contributions to this work. Conversations with Max Suarez, Chris Milly, Guido Salvucci, and Siegfried Schubert about the ideas presented herein are also greatly appreciated. This work was supported by the Earth System Science Program of the NOAA/Climate Program Office and by the NASA Terrestrial Hydrology Program.

References


Twentieth-century drought in the conterminous United States. J. Hydrometeor., 6, 985-1001.


List of Figures

1. a. Illustration of how a representative evaporation efficiency function ($\lambda E/R_{\text{net}}$ vs. $W$, the red curve on the right) and runoff efficiency function ($Q/P$ vs. $W$, the blue curve on the right) combine to produce a single curve in efficiency space (the green curve on the left). b. A second sample illustration.

2. Schematic of the water balance model (WBM) used in this study. The forcing variables, imposed daily, are precipitation ($P$) and net radiation ($R_{\text{net}}$); the WBM uses the imposed evaporation and runoff relationships to compute the daily evaporation ($E$) and runoff ($Q$) as a function of its prognostic water content ($W$). (Reproduced from Koster and Mahanama [2012].)

3. a. Illustration of a typical WBM simulation. Evaporation and runoff efficiency curves (the red and blue curves on the left, respectively) are implemented into the WBM, and the WBM is driven with gridded observational data over a number of decades. In the course of the simulation, the WBM produces gridded runoff values that are spatially aggregated to hydrological basins for comparison with observed basin streamflows (right panel). b. A second representative example.

4. Demonstration that different pairings of evaporation and runoff efficiency functions lead to essentially equivalent streamflow simulations if their corresponding efficiency space curves are the same. The three pairings of efficiency functions in the top row (evaporation efficiency in red, runoff efficiency in blue) each map to the single efficiency space curve in the lower
left panel. WBM simulations with each pairing of functions correspondingly produce very similar simulations of streamflow (lower right panel).

5. a. Three efficiency space curves, color-coded according to the degree to which corresponding WBM simulations reproduce observed streamflows in the Upper Mississippi Basin. The red curve corresponds to the most skillful simulation, and the blue curve to the least skillful simulation. Skill is quantified with an RMSE metric, though numbers are not shown in order to emphasize relative skill. b. Same, but with over 23000 color-coded curves overlain on the plot. The solid black curve is the curve associated with the greatest simulation skill.

6. a. WBM simulation of streamflow in the Upper Mississippi (red curve in right panel) associated with the highest-scoring curve in efficiency space (black curve in left panel; see Figure 5b). The black curve in the right panel represents the observed streamflows for the basin. b-e. Same, but for the WBM simulation associated with the white curve in the left panel.

Figure 7. a. Basins (in gray) examined collectively in the analysis. See Mahanama et al. (2012) for details on the basins and on the streamflow observations taken therein. b. Color-coded efficiency space curves as in Figure 5b, but for an assumed 0.1 m soil depth and for the collected set of basins. The black curve represents the curve associated with the highest simulation skill. c. Same, but for an assumed 0.5 m soil depth. c. Same, but for an assumed 1 m soil depth. d. Same, but for an assumed 1.667 m soil depth.

8. a. Budyko’s semi-empirical relationship between the evaporation-precipitation ratio and the dryness index, defined as $R_{net}/P\lambda$. b. Efficiency space curves from Figure 7c. A transform of Budyko’s relationship is shown in white.
9. Scatter plots of evaporation efficiency (ratio of monthly-averaged $E$ to monthly-averaged $R_{net}$) versus monthly-averaged root zone soil moisture, as derived from the May-September output diagnostics produced over CONUS by three state-of-the-art LSMs.

10. a. Efficiency space curve (in white) for a specific LSM (“Model X”), overlain on the color-coded curves from Figure 7c. To the right are plots of annual runoff ratio across several basins as computed from Model X diagnostics and from observations. b. Same, but for a different LSM (“Model Y”).

11. a. Scatter plot of evaporation efficiency (ratio of monthly-averaged $E$ to monthly-averaged $R_{net}$) versus monthly-averaged root zone soil moisture, as derived from the May-September output diagnostics produced over CONUS by the Catchment-CN LSM. b. Color-coded efficiency space curves from Figure 7c, with one of the curves highlighted as a heavy dotted line. c. Resulting relationship between runoff efficiency and root zone soil moisture.

12. a. Annual runoff efficiencies (ratio of annual basin-averaged runoff to annual basin-averaged precipitation) for several CONUS basins, as produced by the Catchment-CN LSM in a multidecadal offline simulation. b. Same, but for a tuned version of the Catchment-CN LSM (see text). c. Same, but for observations.

13. Sample evaporation efficiency and runoff efficiency functions, with highlighted regimes of interest.
Figure 1. a. Illustration of how a representative evaporation efficiency function ($\lambda E/R_{net}$ vs. $W$, the red curve on the right) and runoff efficiency function ($Q/P$ vs. $W$, the blue curve on the right) combine to produce a single curve in efficiency space (the green curve on the left). b. A second sample illustration.
Figure 2. Schematic of the water balance model (WBM) used in this study. The forcing variables, imposed daily, are precipitation (P) and net radiation (R_{net}); the WBM uses the imposed evaporation and runoff relationships to compute the daily evaporation (E) and runoff (Q) as a function of its prognostic water content (W). (Reproduced from Koster and Mahanama [2012].)

Time step: daily

Integration time: ~ 50 yr

Domain: Continental U.S.
Figure 3. a. Illustration of a typical WBM simulation. Evaporation and runoff efficiency curves (the red and blue curves on the left, respectively) are implemented into the WBM, and the WBM is driven with gridded observational data over a number of decades. In the course of the simulation, the WBM produces gridded runoff values that are spatially aggregated to hydrological basins for comparison with observed basin streamflows (right panel). b. A second representative example.
Figure 4. Demonstration that different pairings of evaporation and runoff efficiency functions lead to essentially equivalent streamflow simulations if their corresponding efficiency space curves are the same. The three pairings of efficiency functions in the top row (evaporation efficiency in red, runoff efficiency in blue) each map to the single efficiency space curve in the lower left panel. WBM simulations with each pairing of functions correspondingly produce very similar simulations of streamflow (lower right panel).
Figure 5.  a. Three efficiency space curves, color-coded according to the degree to which corresponding WBM simulations reproduce observed streamflows in the Upper Mississippi Basin. The red curve corresponds to the most skillful simulation, and the blue curve to the least skillful simulation. Skill is quantified with an RMSE metric, though numbers are not shown in order to emphasize relative skill. b. Same, but with over 23000 color-coded curves overlain on the plot. The solid black curve is the curve associated with the greatest simulation skill.
Figure 6.  a. WBM simulation of streamflow in the Upper Mississippi (red curve in right panel) associated with the highest-scoring curve in efficiency space (black curve in left panel; see Figure 5b). The black curve in the right panel represents the observed streamflows for the basin.

b-e. Same, but for the WBM simulation associated with the white curve in the left panel.
Figure 7. a. Basins (in gray) examined collectively in the analysis. See Mahanama et al. (2012) for details on the basins and on the streamflow observations taken therein. b. Color-coded efficiency space curves as in Figure 5b, but for an assumed 0.1 m soil depth and for the collected set of basins. The black curve represents the curve associated with the highest simulation skill. c. Same, but for an assumed 0.5 m soil depth. c. Same, but for an assumed 1m soil depth. d. Same, but for an assumed 1.667 m soil depth.
Figure 8. a. Budyko’s semi-empirical relationship between the evaporation-precipitation ratio and the dryness index, defined as $R_{\text{net}}/P\lambda$. b. Efficiency space curves from Figure 7c. A transform of Budyko’s relationship is shown in white.
Figure 9. Scatter plots of evaporation efficiency (ratio of monthly-averaged $E$ to monthly-averaged $R_{net}$) versus monthly-averaged root zone soil moisture, as derived from the May-September output diagnostics produced over CONUS by three state-of-the-art LSMs.

Different grid cells in a given LSM may have different soil types and thus different “start values” for the curves.

Different models tend to have different soil moisture ranges.
Figure 10. a. Efficiency space curve (in white) for a specific LSM (“Model X”), overlain on the color-coded curves from Figure 7c. To the right are plots of annual runoff ratio across several basins as computed from Model X diagnostics and from observations. b. Same, but for a different LSM (“Model Y”).
Figure 11. a. Scatter plot of evaporation efficiency (ratio of monthly-averaged E to monthly-averaged $R_{\text{net}}$) versus monthly-averaged root zone soil moisture, as derived from the May-September output diagnostics produced over CONUS by the Catchment-CN LSM. b. Color-coded efficiency space curves from Figure 7c, with one of the curves highlighted as a heavy dotted line. c. Resulting relationship between runoff efficiency and root zone soil moisture.
Figure 12. a. Annual runoff efficiencies (ratio of annual basin-averaged runoff to annual basin-averaged precipitation) for several CONUS basins, as produced by the Catchment-CN LSM in a multidecadal offline simulation. b. Same, but for a tuned version of the Catchment-CN LSM (see text). c. Same, but for observations.
Figure 13. Sample evaporation efficiency and runoff efficiency functions, with highlighted regimes of interest.