Agricultural and Forest Meteorology 150 (2010) 219–225

Contents lists available at ScienceDirect

Agricultural and Forest Meteorology

journal homepage: www.elsevier.com/locate/agrformet

Using watershed water balance to evaluate the accuracy of eddy covariance evaporation measurements for three semiarid ecosystems

Russell L. Scott *

Southwest Watershed Research Center, USDA-ARS, 2000 E. Allen Road, Tucson, AZ 85719, USA

ARTICLE INFO

Article history:
Received 24 April 2009
Received in revised form 29 October 2009
Accepted 2 November 2009

Keywords:
Evaporation
Evapotranspiration
Eddy covariance
Water balance
Energy balance closure

ABSTRACT

The eddy covariance technique is a widely used and accepted method to quantify ecosystem-scale mass and energy fluxes. Eddy covariance measurements of evaporation, also known as evapotranspiration, are used to determine local, regional and global water budgets, calibrate and validate land surface models, and acquire understanding of ecosystem processes. This paper assesses the accuracy of eddy covariance evaporation measurements by comparing them with those derived from small watershed water balances. Comparing thirteen years of data from shrubland, grassland and savanna sites in southern Arizona USA, the two independent measures agreed to within an average of 3% annually and differed from −10 to +17% in any given year, when an assumed 5% underestimation in precipitation due to gage undercatch was considered. The agreement between the two measures was generally better in drier years and at less topographically complex sites. Despite an indication of a systematic underestimate of evaporation by a commonly used assessment of the energy balance, forcing energy balance closure on evaporation led to worse results for nine of the thirteen annual periods but improved multiyear sums at two of the three sites.

Published by Elsevier B.V.

1. Introduction

Measurements of the energy and mass exchange at the land–atmosphere interface are critical for determining local, regional and global budgets, land surface model testing, and understanding ecosystem processes. The eddy covariance technique has become the standard for monitoring energy and mass fluxes, especially water vapor and carbon dioxide (Aubinet, 2000; Baldocchi, 2003), and the global network of monitoring sites has sparked a revolution in the understanding of these exchanges (Baldocchi, 2008; Baldocchi et al., 2001). Because eddy covariance measurements are so pervasive and extensively used to quantify these fluxes, there is a continual need to assess the accuracy of these measurements as opportunities arise.

The energy balance near the land surface can be written as:

$$R_n - G - Q_b - Q_p = \lambda E + H \tag{1}$$

where R_n is the net radiation, G is the soil heat flux, Q_b is the storage of heat in the biomass and air between the measurement height and ground surface, Q_p is the amount of energy consumed by photosynthesis or released by respiration, and λE and H are the latent and sensible heat fluxes, respectively. The accuracy of the eddy covariance measurements is often assessed by comparing the left-hand side of Eq. (1), measured with radiometers, ground heat flux plates, and temperature probes, with the right-hand side of Eq. (1), measured with eddy covariance (Aubinet, 2000; Wilson et al., 2002). Eddy covariance measurements are often used to quantify other mass exchanges (most commonly, carbon dioxide) but the accuracy of these measurements is more difficult to determine (Baldocchi, 2003; Goulden et al., 1996). It is sometimes assumed that Eq. (1) can be used to assess the accuracy of these other fluxes as well even though this might not be accurate (Barr et al., 2006; Liu et al., 2006; Twine et al., 2000).

One of the persistent concerns with eddy covariance is the "closure problem", where the sum of the terms on the right-hand side of Eq. (1) is often less (typically ranging from 5 to 30%) than the sum on the left-hand side, indicating that the energy balance is not closed (Blanken et al., 1997; Oncley et al., 2007; Wilson et al., 2002). There is considerable disagreement among researchers as to whether this is indeed a problem. Numerous studies have pointed out that available energy is also difficult to quantify; often Q_b and Q_p are neglected and/or under-sampled, and the source areas for the sensors used to quantify the terms of left-hand side of Eq. (1) are fixed and often much smaller than the constantly shifting source area for the turbulent fluxes (Anthoni et al., 2000; Aston, 1985; Kustas et al., 1998, 2000; Mayocchi and Bristow, 1995; Sauer et al., 2003; Schmid, 1997). So, some have argued that when the available energy terms are duly accounted for, closure has been achieved (Haverd et al., 2007; Meyers and Hollinger, 2004) while other very comprehensive experiments could not (Oncley et al.,

* Tel.: +1 520 670 6380x113; fax: +1 520 670 5550.
E-mail address: russ.scott@ars.usda.gov.

0168-1923/$ – see front matter. Published by Elsevier B.V.
doi:10.1016/j.agrformet.2009.11.002

2007). On the other hand, others have found that when the covariances are computed over the proper time interval (which can be much longer than the typical 30 min period) or with certain techniques (e.g., frequency response corrections) energy balance is achieved (Finnigan et al., 2003; Wolf et al., 2008). Finally, other studies have suggested that problems with the widely used instrumentation itself (e.g., sonic anemometer cosine errors) may be the problem (Gash and Dolman, 2003; Van Der Molen et al., 2004).

In this study, the accuracy of evaporation (also known as, evapotranspiration) measurements by eddy covariance (E) is assessed by comparing them to a site's seasonal or annual water balance estimate of evaporation (E^*) characterized by:

$$E^* = P - R - \Delta\theta \tag{2}$$

where P is precipitation, R is runoff, and $\Delta\theta$ is the change in soil water storage. Measurements of E and E^* for thirteen site-years at three semiarid sites in southern Arizona, USA are compared. The three sites are located in or very near small (<4 ha) headwater catchments that are instrumented to measure ephemeral, surface runoff on a scale commensurate with the typical eddy covariance footprint. The watersheds have little percolation beyond the root zone so subsurface runoff can be neglected (Scott et al., 2009, 2000). Also, nearly all the dynamics of soil water redistribution occur within the upper 1 m of soil so arrays of soil moisture probes are used to quantify $\Delta\theta$.

Previous studies have found relatively good agreement between eddy covariance and other independent measurements of E, but they have been inconclusive in regards to adjusting the turbulent fluxes to force energy balance closure. Barr et al. (2000) used a large-scale lysimeter approach to estimate evaporation that agreed within 6% of the eddy covariance measurement over two years, but there was too much uncertainty in both estimates to warrant conclusions regarding forcing closure. Wilson et al. (2001) found that eddy covariance and watershed estimates agreed within ~50 mm annually and both averaged around 580 mm over a five year period in a mixed-deciduous forest. Schume et al. (2005) found a good agreement (<1 mm) between soil water depletion estimates and eddy covariance over a shorter 17-day period, despite complex terrain and lack of energy balance closure. Finally, Kosugi and Katsuyama (2007) determined a three-year average estimate of 735 mm using eddy covariance and 749 mm using catchment water balance, but they needed to force closure to get this agreement.

2. Site descriptions

The three sites monitored in this study were the Santa Rita Mesquite Savanna, the Lucky Hills Shrubland, and the Kendall Grassland, monitored and maintained by the USDA Agricultural Research Service (Fig. 1). The savanna site (31.821°N, 110.866°W, elevation: 1116 m) located in the Santa Rita Experimental Range, 45 km south of Tucson, AZ, is a semidesert grassland that has been converted into a savanna by the encroachment of the woody tree, *Prosopis velutina* Woot. (see Scott et al., 2009 for site details). The shrubland site (31.7438°N, 110.0522°W, 1370 m) located in the USDA-ARS Walnut Gulch Experimental Watershed (WGEW) surrounding the town of Tombstone, AZ is a desert shrubland, dominated by a diverse stand of mainly Chihuahuan desert shrub species (Scott et al., 2006). The grassland (31.737°N, 109.942°W, 1531 m), also located in WGEW, is a semidesert grassland comprised mainly of C4 grasses with a few scattered shrubs (Emmerich, 2003). Average annual rainfall ranges from 320 (shrubland) to 380 mm (savanna) with 50–60% of the rainfall

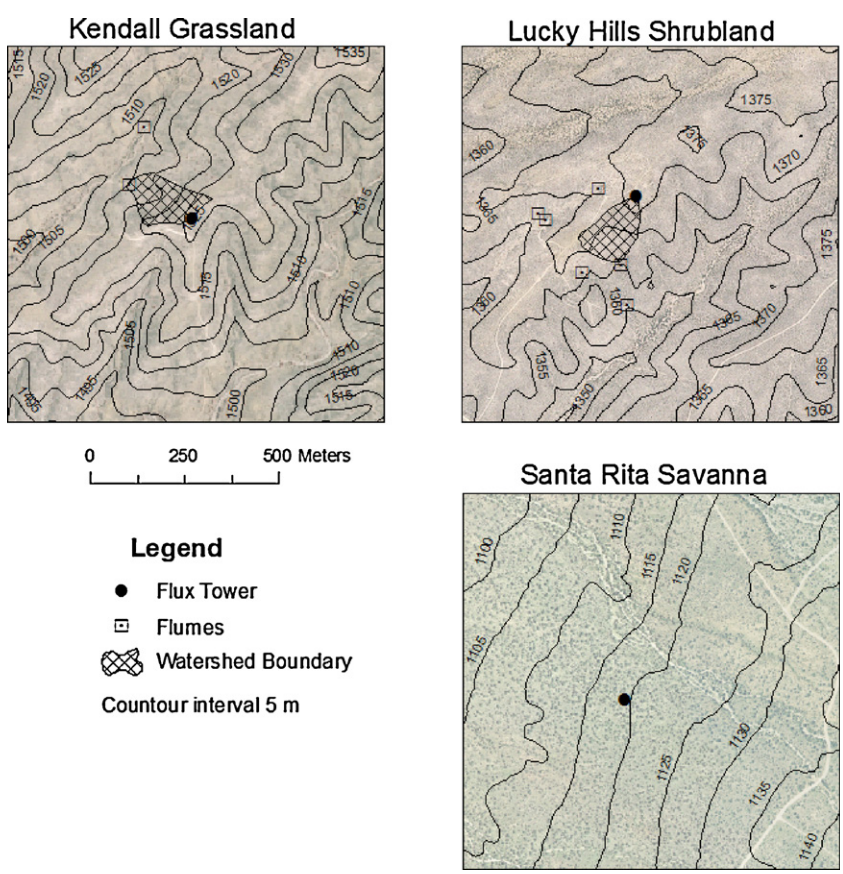

Fig. 1. Site layout and topography.

arriving around the summer months of July–September as part of the North American Monsoon (Adams and Comrie, 1997).

3. Measurements and methods

3.1. Environmental measurements

Above-canopy net radiation (R_n) was measured at ~3 m above the canopy using a 4-component radiometer (CNR1, Kipp & Zonen, Delft, The Netherlands) at the savanna and grassland site and a 2-component net radiometer (CNR2, Kipp & Zonen) at the shrubland site. Also at these heights, photosynthetically active radiation (PAR; LI-190, LI-COR, Lincoln, NE) sensors measured upwelling and downwelling fluxes. Ground heat flux was measured with soil heat flux plates (REBS Inc., Seattle, WA) installed 0.05 m below ground level under both inter-canopy and under-canopy positions (n = 8 at the savanna and n = 5 at the shrubland and grassland). Average soil temperature for 0–5 cm soil depth was determined by averaging the 2 and 4 cm thermocouples located above each of the soil heat flux plates. Measurements of the rate of change of soil temperature above the heat flux plates in combination with the soil bulk density and soil water content allowed calculation of the soil heat flux at the surface (G) by determining the changes in heat storage of the 0–5 cm soil layer. The storage or loss of heat from the biomass and between the ground and the flux measurement height (Q_b, Eq. (1)) was not considered in this study, but it was expected to be small due to the short tower heights and small amounts of biomass (Scott et al., 2009). Likewise, the energy associated with the carbon dioxide flux, Q_p, was ignored.

Precipitation (P) was quantified at the sites by using a 0.2 m-orifice, tipping-bucket rain gage (TE525, Campbell Scientific, Inc.) mounted at 1 m above the ground at the savanna and weighing gages consisting of a Belfort 0.2 m-orifice weighing-recording gages installed so that the top edge of the gage orifice was nominally 0.9 m above the ground at the shrubland and grassland (Goodrich et al., 2008). At all the sites there was either an additional precipitation gage at or near the site (<0.5 km), and these were used as a check on the accuracy of the primary precipitation gage. The difference between the primary and check gage annual totals averaged less than 5 mm.

A combination of flumes and weirs measured surface runoff (R) from the small headwater catchments (Stone et al., 2008). The flux tower at the shrubland is located on the upper, northern edge of WGEW catchment #102 (1.5 ha, Fig. 1). The flux tower at the grassland is positioned on the upper southern end of WGEW catchment #112 (1.9 ha). At the savanna, runoff is quantified for four small, 1–4 ha catchments located 1.5 km east the site (http://www.tucson.ars.ag.gov/dap). Watershed #6 (3.1 ha) was used because its loamy sand soils are the most similar to those surrounding the savanna site (Lane and Kidwell, 2003). This data revealed an unusual step change in runoff starting in 2004 and continuing through 2008 where the average runoff ratio (R/P) for this period was 8.5%, but the 1974–2003 average runoff ratio equaled to 1.5%. Because there was no evidence to tie this unusual step change in the runoff ratio to an actual change in the runoff processes around the tower location, and the fact that the deep, loamy upland soils at the savanna produce very little runoff at the small plot scale (Emmerich and Cox, 1992), the runoff measurements for the savanna were suspect and may be overestimated.

Volumetric soil water content (θ) in the upper ~1 m of soil was quantified using all the available soil moisture probes located at each site to determine the change in near-surface soil moisture storage ($\Delta\theta$). None of the sites showed changes in soil moisture below 1 m so it was appropriate to focus on the upper 1 m of soil. At all the sites, probe output was converted to θ using site-specific calibration curves produced in the laboratory with site soils packed in a column to site-measured bulk densities. For the savanna, θ was determined using a canopy-weighted-average from 5, 10, 20, 30, 50, 70, 100 cm depths located in a profile in an open area and another beneath a tree canopy using commercial TDR-type probes (CS616, Campbell Scientific). For the shrubland, θ was measured using a TDR-array of probes at 5, 15, 30, 50, 75 and 100 cm at multiple profiles located at mid-slope in the watershed. For the grassland averages of three different hilltop locations were used for 5 and 15 cm depths and values from one TDR profile on a nearby, south-facing hillslope were used for the 30, 50, 70 and 100 cm depths.

3.2. Eddy covariance

Three-dimensional, sonic anemometers (CSAT-3; Campbell Scientific) and open-path infrared gas analyzers (LI-7500, LI-COR) were mounted at ~5 m above the height of the vegetation to measure the three components of the wind velocity vector, sonic temperature and concentrations of water vapor and carbon dioxide. Data were sampled at 10 Hz. Whenever mean afternoon gas concentrations drifted a sufficient amount (i.e., [CO_2] differed more than ±10 ppm of expected seasonal values or [H_2O] was farther than ±~1 g m^{-3} from that determined by the temperature/relative-humidity probe), the IRGA windows were cleaned. Often, these drifts in gas concentrations were due to a dirty window, but if cleaning the windows did not correct the problem, the IRGAs were zero- and span-calibrated (typically every few months) using a CO_2/H_2O–free gas, a standard [CO_2] gas, and a dew point generator. Covariances were calculated by first filtering spikes and then using a 30-min block average. The coordinate frame was rotated using the planar-fit method (Lee et al., 2004), frequency domain corrections for path length averaging and sensor separation were applied (Massman, 2001), and density fluctuations were accounted for (Webb et al., 1980) in the calculation of the fluxes. The sonic temperature was used to calculate sensible heat flux using the method suggested by (Paw et al., 2000) which accounts for a missing energy balance term associated with the expansion of air during evaporation under constant pressure. Fluxes measured when the wind was blowing from the direction within ± 5° of the back of the anemometer (~2% of the data) were omitted due to possible interference from the anemometer support and the IRGA mounted behind the anemometer. These sites are a part of the Ameriflux network, and their instrumentation, processing software, and techniques have been tested and verified with the network's "gold-standard" (http://public.ornl.gov/ameriflux/).

The fetch at all of the sites is representative of the cover immediately surrounding the towers over several kilometers. The savanna site is fairly flat, but broadly sloping at ~3.5% from southeast to northwest (Fig. 1). Both the shrubland and the grassland sites have more topographical variation and the towers sit upon ridge tops of the small headwater catchments (Fig. 1). The shrubland slopes away from the tower at 5% to the south, 3% to the west and is fairly flat to the north and east. The grassland slopes away from the tower at ~10% to the northwest and southeast, at 9% to the southwest and is flat to the northeast over a horizontal distance of ~200 m. There was likely a greater mismatch between the source area of the available energy ($R_n - G$) that was derived from these ridge-top locations and the fluctuating turbulent flux source areas (Schmid, 1994) which were often located on sloped surfaces with different aspects. Using the model of Kljun et al. (2004) with typical daytime conditions, peak location and along-wind extent of the 90% cross-wind integrated flux footprint ranged narrowly from 76–95 to 210–230 m, respectively, for all three sites.

The flux data were filtered for spikes, instrument malfunctions, and poor quality (representing ~10% of the measurements). The rejection criteria used to screen data were: rain events, out-of-range signals, and spikes with the standard deviation of [H_2O] or

$[CO_2]$ greater than two standard deviations from the mean determined on a yearly basis. A u^* filter, which is commonly used for CO_2 fluxes (Malhi et al., 1998) was not applied to H and λE. However, the findings of Barr et al. (2006) that suggest the need to reject H and λE for low u^* values were considered. Similar to their results, energy closure improved for increasing values of u^* with no clear daytime threshold beyond which it plateaued, but the improvement was much smaller (<5% over full range of threshold) than their results, perhaps due to the much shorter tower heights in this study. Annual gap-filled evaporation changed less than 5 mm when a u^* filter of 0.15 m s^{-1} was evaluated (threshold previously used for CO_2 fluxes at the savanna by Scott et al., 2009).

Daily evaporation was calculated by first filling the gaps in the 30 min data using 14-day moving-average look-up tables of E and incoming PAR, averaged over 100 μmoles m^{-2} s^{-1} intervals (Falge et al., 2001) and separated into morning or afternoon periods. Varying the 14-day gap-filling period from 7 to 28 days resulted in very little changes in annual sums because under or over estimates of filled gaps tend to sum to the average flux over the year. Also, the amount of gaps filled was small (~10%). Linear interpolation was used to fill any gaps in daily E that occurred when a site lost power.

4. Water balance

Eddy covariance measurements started in 2004 and continued till present at the savanna and grassland sites. Shrubland flux measurements also began in 2004, discontinued in February 2005, and resumed in mid-2007.

Below the eddy covariance and water balance estimates of E are compared by looking at the difference between the terms:

$$\varepsilon = E^* - E \tag{3}$$

on a yearly basis and for consecutive multiyear periods. When doing this, the recommendation of Twine et al. (2000) is also evaluated. They suggest that forcing closure is justified when available energy is known and errors in its measurement modest. In this study, errors in available energy were likely similar to those of Twine et al. (2000), who estimated an uncertainty of ~10%. To force closure, E was divided by the average of $(\lambda E + H)/(R_n - G)$ computed daily. This results in closing the energy balance on a daily basis, rather than for every 30 min period. This was preferable because energy storage was unmeasured but likely averages out to near zero on a daily basis.

In addition to the closure adjustment suggested by Twine et al. (2000), the effect of precipitation gage undercatch was considered. Undercatch is a well-documented effect that results from a systematic underestimation of rainfall when using rain gages that are exposed to the wind (Larson and Peck, 1974). Wind effects around the orifice have the effect of reducing the amount of rain that falls into the orifice and depends on the gage setup, wind speed and drop-size distribution (Sieck et al., 2007). For the type and setup of weighing rain gages used at the shrubland and grassland, undercatch is estimated to be 5% (Duchon and Essenberg, 2001). While less is generally known about the systematic errors of the tipping-bucket gage used at the savanna (Fankhauser, 1997; La Barbera et al., 2002), the same 5% error was used because of the size and height of the orifice was similar to the gages at the other sites.

Table 1
Energy balance assessment for the three sites using 30 min values of R_n, G, λE and H.

Site	$\lambda E + H = m(R_n - G) + b$			$\frac{\sum \lambda E + H}{\sum R_n - G}$
	m	b	R^2	
Savanna	0.83	12.73	0.96	0.96
Shrubland	0.81	13.12	0.94	0.97
Grassland	0.73	23.80	0.92	1.04

5. Results

The 30-min data were used to compute both the slope and the intercept of the regression line $\lambda E + H = m[R_n - G] + b$ and the ratio of the sum of the turbulent fluxes over the available energy $(\sum \lambda E + H / \sum R_n - G)$ to provide two different measures of energy balance closure (Table 1). The slope of the regression line in all cases was less-than-unity and the intercepts were all positive, which resulted in ratios of the sum of fluxes that are higher than the slopes of the regression lines. The ratio of $(\sum \lambda E + H / \sum R_n - G)$ computed on a daily basis was generally lower (~0.8–0.9) in the summer months and higher (~1.0–1.1) in the winter (data not shown).

Precipitation (P) and E dominated the water balance at each site and for all years with R and $\Delta\theta$ being much lower (Tables 2–4). Fig. 2 presents an example of these fluxes for 2008 at the grassland. The other years and other sites (not shown) followed the same general pattern of cumulative E usually exceeding P and depleting θ after the typical, fore-summer drought (April–June) and then with E lagging behind P, and θ increasing, during the summer monsoon (July–September). In fall (October–December) E declines as the vegetation down-regulates and soil evaporation declines in response to decreasing soil moisture, colder temperatures, and reduced light.

At the savanna, E was nearly equal to P in most years so the difference (ε, Eq. (3)) was always negative implying an over-

Table 2
Savanna water balance (all units in mm). In italics, percent error as a ratio of the residual over the precipitation (− indicates losses > $\underline{P}$).

	2004*	2005	2006	2007	2008	All
Precipitation (P)	285	335	289	330	402	1640
Evaporation (E)	293	338	290	321	408	1650
Closure-forced evaporation (E_f)	321	357	311	347	441	1778
Runoff (R)	2	42	34	24	39	141
Change in soil moisture from 0 to 1 m ($\Delta\theta$)	−1	−3	9	23	−8	20
$P - E - R - \Delta\theta$	−9	−43	−44	−38	−37	−171
	−3%	*−13%*	*−15%*	*−12%*	*−9%*	*−10%*
$1.05P - E - R - \Delta\theta$	5	−26	−30	−22	−17	−89
	2%	*−7%*	*−10%*	*−6%*	*−4%*	*−5%*
$P - E_f - R - \Delta\theta$	−37	−62	−66	−64	−70	−299
	−13%	*−19%*	*−23%*	*−19%*	*−17%*	*−18%*
$1.05P - E_f - R - \Delta\theta$	−23	−45	−51	−47	−50	−217
	−8%	*−13%*	*−17%*	*−14%*	*−12%*	*−13%*

* Jan. 4, 2004–Dec. 31, 2004.

Table 3
Shrubland water balance (all units in mm). In italics, percent error as a ratio of the residual over the precipitation (− indicates losses > P).

	2004*	2005	2006	2007**	2008	2007–2008
Precipitation (P)	140			209	265	474
Evaporation (E)	139			161	247	409
Closure-forced evaporation (E_f)	145			182	270	452
Runoff (R)	0			25	32	57
Change in soil moisture from 0 to 1 m ($\Delta\theta$)	5			−4	−4	−8
$P-E-R-\Delta\theta$	−4			27	−10	17
	−3%			*13%*	*−4%*	*4%*
$1.05P-E-R-\Delta\theta$	3			37	3	40
	2%			*17%*	*1%*	*8%*
$P-E_f-R-\Delta\theta$	−10			6	−32	−27
	−7%			*3%*	*−12%*	*−6%*
$1.05P-E_f-R-\Delta\theta$	−3			16	−19	−3
	−2%			*7%*	*−7%*	*−1%*

* May 5, 2004–Feb. 1, 2005.
** June 29, 2007–Dec. 31, 2007.

Table 4
Grassland water balance (all units in mm). In italics, percent error as a ratio of the residual over the precipitation (− indicates losses > P).

	2004*	2005	2006	2007	2008	All
Precipitation (P)	191	162	274	313	312	1252
Evaporation (E)	197	192	204	276	262	1131
Closure-forced evaporation (E_f)	224	189	220	299	280	1212
Runoff (R)	2	0	43	17	31	93
Change in soil moisture from 0 to 1 m ($\Delta\theta$)	−5	−28	14	5	−12	−26
$P-E-R-\Delta\theta$	−2	−1	13	15	31	55
	−1%	*−1%*	*5%*	*5%*	*10%*	*4%*
$1.05P-E-R-\Delta\theta$	7	7	27	31	47	118
	4%	*4%*	*9%*	*9%*	*14%*	*9%*
$P-E_f-R-\Delta\theta$	−29	2	−3	−8	13	−26
	−15%	*1%*	*−1%*	*−3%*	*4%*	*−2%*
$1.05P-E_f-R-\Delta\theta$	−20	10	10	7	29	37
	−10%	*6%*	*4%*	*2%*	*9%*	*3%*

* May 8, 2004–Dec. 31, 2004.

estimate of water leaving the system and/or an underestimate of incoming water with results for ε ranging from −15 to −1% of P (Table 2). When considering a probable 5% undercatch error in P, the results improved to within −10 to 2%. Forcing closure on E (E_f) increased the annual evaporation by 6 to 14% and generally made ε more negative (further out of balance). The five-year totals also showed a better agreement between the watershed and eddy covariance estimates of evaporation for the non-forced case.

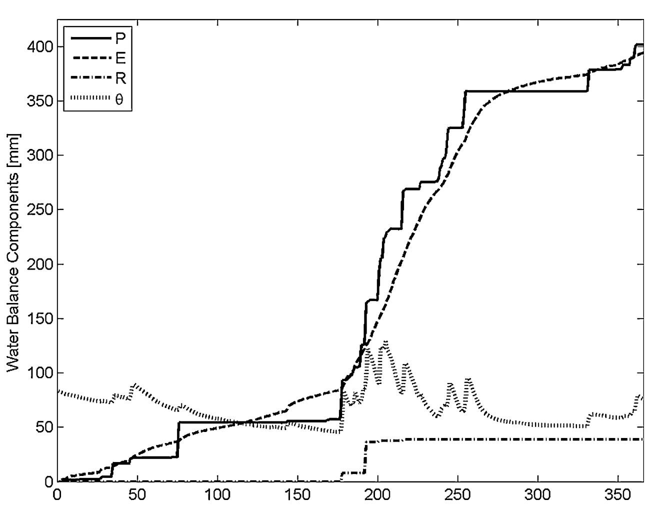

Fig. 2. Example of water balance components (Eq. (2)) for 2008 at the grassland site showing cumulative fluxes of precipitation (P), evaporation (E), and runoff (R) along with total storage of soil moisture (θ) in the top 1 m of soil.

At the shrubland, E was nearly equal to P in the driest year and less than P in the other two years that had runoff (Table 3). The ratio ε/P was small and negative in 2004 (−3%) and 2008 (−4%) and larger (13%) and positive in 2007. The residual in 2004 and 2008 was still small after including the 5% undercatch factor (<3%), but it increased to 17% in 2007. Forcing closure at this site increased E by about 4% in 2004 and by 11% in 2007 and 2008. This had the effect of reducing ε to approximately 7% of P in 2007 (1.05P case), but little improvement was made for the water balances in 2004 and 2008. For the 2004 period, both the height of the eddy covariance system and the net radiometer were mounted closer to the surface (by ∼3 and 1 m, respectively). Considering the two consecutive years together, forcing closure improved the results (within 2% for 1.05P case) relative to not changing evaporation (8% difference).

At the grassland ε was small in 2004 and 2005 and more positive ($E < P - R - \Delta\theta$) in 2006–2008, so when accounting for gage undercatch it increased to values ranging from 4 to 14% of P (Table 4). Forcing closure changed E by −2 to +14% which increased ε in 2004 and 2005 and decreased it for 2006–2008. When forcing closure and accounting for undercatch, ε ranged from −10 to +9% of P. Forcing closure considerably improved the results for the five-year total.

6. Discussion

Measures of energy balance closure for the three sites used in this study (Table 1) were in the range of those measured at many other sites (Wilson et al., 2002). The less-than-unity slopes and the positive intercepts indicate that $\lambda E + H$ was usually less than $R_n - G$ for high values of $R_n - G$ and often greater than $R_n - G$ for low values. Taken at face value, these slopes seem to imply that E was systematically underestimated by ~17–27% at the sites because the majority of the evaporation occurs in daytime and in the time of year when $R_n - G$ is high. On the other hand, the summed closure ratios ($\sum \lambda E + H / \sum R_n - G$) were much closer to unity and may imply some measurement problems with the available energy on both a daily and seasonal basis. Using daily averages, the slopes of $R_n - G$ versus $\lambda E + H$ were closer to 0.9 (data not shown) rather than the ~0.8 slopes on a 30-min basis. This could be because the other storage terms (Q_b and Q_p) were not included and/or G was underestimated, as the storage terms generally sum to near zero over the course of the day. The annual cycle in the ratio of daily closure ratio, which was typically higher in the winter and lower in the summer, also contributed to the well-balanced ratios of the annual sums (Table 1). This seasonality appears opposite of the behavior found in survey of sites by Wilson et al. (2002) and its cause could not be determined.

Both the savanna and the shrubland had very similar closure measures whereas the grassland differed. This may be due to greater topographical relief at the grassland (Fig. 1) where the available energy measurements are on top of the hill but the source area of the turbulent fluxes was more likely to be located on sloped surfaces with different aspects. For example, a daytime measurement of available energy would be exposed at higher sun angles relative to a flux footprint derived from a northwesterly facing aspect from the predominant northwest wind direction. This available energy mismatch, however, appears to be compensated in part by the oppositely signed mismatch of source-area footprints at nighttime such that the ratio of the sums of turbulent fluxes and available energy is more balanced (Table 1).

It is important to remember the additional degree of uncertainty in this analysis due to the spatial-scale mismatch between the two estimates of E. Accordingly, site topography appears to have played a role in this study as the flatter, more uniform savanna had the best agreement between E and E^* and the most complex grassland site had the greatest disagreement. Complex topography presents a challenge for both measuring turbulent fluxes and estimating watershed evaporation. For the former, the boundary layer flows are more complex and challenge basic assumptions made for eddy covariance calculations (Finnigan, 2004). Also, redistribution of water on a hillslope increases the heterogeneity of the scalar source and sinks within the flux footprint. For the water balance estimate, lateral and vertical redistribution of water on the hillslope leads to more uncertainty in the quantification of runoff and soil moisture storage terms that are representative of the turbulent flux footprint. For example, the grassland has a large swale in the middle of the watershed which intercepts much of runoff coming off the convergent hillslope before it exits the watershed at the weir (Nearing et al., 2007). Thus, the water balance of the flux source-area-integrated annual footprint is biased toward the drier hilltop location which is probably not accurately represented in the runoff and soil moisture storage measurements.

This study using the annual water balance for ten nearly complete years and three partial years at three sites is further confirmation that eddy covariance can provide accurate estimates of evaporation (Barr et al., 2000; Kosugi and Katsuyama, 2007; Schume et al., 2005; Wilson et al., 2001). Considering the most-likely case of the underestimation in rainfall being 5%, annual eddy covariance estimates of evaporation appear to be highly accurate in comparison with the watershed estimates with a mean difference of less than 3% of total precipitation. This is remarkable when considering the very intermittent and ephemeral measurements of P and R in comparison to the nearly continuous 10-Hz sampling for the eddy covariance measurements along with the associated error-checking and gap-filling that are needed to produce long-term sums.

On an annual basis, the accuracy of evaporation using eddy covariance appeared greatest at the savanna site with a yearly disagreement with the watershed evaporation less than 31 mm, or 11% of P, and a mean difference of −5%, indicating a tendency for E to exceed E^* which may have been due to a problem with the runoff estimate (see Section 3). At the shrubland, the disagreement was 3 mm, or less than 3%, in two of the years but ballooned to 37 mm, or 17%, for June–December 2007. The agreement between the two approaches at the grassland was worse than the other sites but still within 31 mm, or 10%, except for 2008 when it was about 14%. At the grassland and shrubland, there was a tendency to have better agreement in drier years. This lends support to the idea that topographical complexity may have complicated the comparison because the years with larger hydrologic fluxes would exaggerate the differences in the spatial representativeness of the water balance and eddy covariance measurements.

The slope and intercept of the regression line between available energy and the sum of the turbulent fluxes are a commonly used statistic to quantify energy balance closure. According to this measure, there was a systematic energy imbalance that varied both on a daily and annual basis at all of the sites. However, on a seasonal to yearly basis, the justification for forcing closure using the method suggested by Twine et al. (2000) was ambiguous. Nine of the thirteen years showed the same or less disagreement between the eddy covariance and watershed evaporation when closure was not forced ($|E^* - E| \leq |E^* - E_f|$). Yet, there was some evidence of a systematic underestimation of E (or overestimation of E^*) at the shrubland and grassland so forcing closure at these sites was an improvement when comparing multiyear sums.

Acknowledgements

The author would like to thank technicians W. Cable and R. Bryant for help with the establishment and maintenance of the flux sites, T. Keefer and J. Smith for keeping the high fidelity hydrologic measurements at WGEW and SRER going, and J. Shuttleworth and two anonymous reviewers and associate editor for their constructive comments that improved this manuscript. He is also grateful for the support from the rest of the dedicated staff at the USDA-ARS Southwest Watershed Research Center in Tucson and Tombstone, AZ.

References

Adams, D.K., Comrie, A.C., 1997. The North American Monsoon. Bull. Am. Meteorol. Soc. 78, 2197–2213.
Anthoni, P.M., Law, B.E., Unsworth, M.H., Vong, R.J., 2000. Variation of net radiation over heterogeneous surfaces: measurements and simulation in a juniper-sagebrush ecosystem. Agric. Forest Meteorol. 102, 275–286.
Aston, A.R., 1985. Heat storage in a young eucalypt forest. Agric. Forest Meteorol. 35, 281–297.
Aubinet, M., 2000. Estimates of the annual net carbon and water exchange of forests: the EUROFLUX methodology. Adv. Ecol. Res. 30, 113–175.
Baldocchi, D., 2008. Turner review no. 15. 'Breathing' of the terrestrial biosphere: lessons learned from a global network of carbon dioxide flux measurement systems. Aust. J. Bot. 56, 1–26.
Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer, C., Davis, K., Evans, R., Fuentes, J., Goldstein, A., Katul, G., Law, B., Lee, X., Malhi, Y., Meyers, T., Munger, W., Oechel, W., Paw, U.K.T., Pilegaard, K., Schmid, H.P., Valentini, R., Verma, S., Vesala, T., Wilson, K., Wofsy, S., 2001. FLUXNET: a new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities. Bull. Am. Meteorol. Soc. 82, 2415–2434.

Baldocchi, D.D., 2003. Assessing the eddy covariance technique for evaluating carbon dioxide exchange rates of ecosystems: past, present and future. Global Change Biol. 9, 479–492.

Barr, A.G., Kamp, G.V.D., Schmidt, R., Black, T.A., 2000. Monitoring the moisture balance of a boreal aspen forest using a deep groundwater piezometer. Agric. Forest Meteorol. 102, 13–24.

Barr, A.G., Morgenstern, K., Black, T.A., McCaughey, J.H., Nesic, Z., 2006. Surface energy balance closure by the eddy-covariance method above three boreal forest stands and implications for the measurement of the CO_2 flux. Agric. Forest Meteorol. 140, 322–337.

Blanken, P.D., Black, T.A., Yang, P.C., Neumann, H.H., Nesic, Z., Staebler, R., Den Hartog, G., Novak, M.D., Lee, X., 1997. Energy balance and canopy conductance of a boreal aspen forest: partitioning overstory and understory components. J. Geophys. Res. [Atmos.] 102, 28915–28927.

Duchon, C.E., Essenberg, G.R., 2001. Comparative rainfall observations from pit and aboveground rain gauges with and without wind shields. Water Resour. Res. 37, 3253–3263.

Emmerich, W.E., 2003. Carbon dioxide fluxes in a semiarid environment with high carbonate soils. Agric. Forest Meteorol. 116, 91–102.

Emmerich, W.E., Cox, J.R., 1992. Hydrologic characteristics immediately after seasonal burning on introduced and native grasslands. J. Range Manag. 45, 476–479.

Falge, E., Baldocchi, D., Olson, R., Anthoni, P., Aubinet, M., Bernhofer, C., Burba, G., Ceulemans, R., Clement, R., Dolman, H., Granier, A., Gross, P., Gru?nwald, T., Hollinger, D., Jensen, N.O., Katul, G., Keronen, P., Kowalski, A., Lai, C.T., Law, B.E., Meyers, T., Moncrieff, J., Moors, E., Munger, J.W., Pilegaard, K., Rannik, U., Rebmann, C., Suyker, A., Tenhunen, J., Tu, K., Verma, S., Vesala, T., Wilson, K., Wofsy, S., 2001. Gap filling strategies for defensible annual sums of net ecosystem exchange. Agric. Forest Meteorol. 107, 43–69.

Fankhauser, R., 1997. Measurement properties of tipping bucket rain gauges and their influence on urban runoff simulation. Water Sci. Technol. 36, 7–12.

Finnigan, J., 2004. Advection and modeling. In: Lee, X., Massman, W., Law, B. (Eds.), Handbook of Micrometeorology. Kluwer Academic Publishers, Dordrecht, pp. 209–244.

Finnigan, J.J., Clement, R., Malhi, Y., Leuning, R., Cleugh, H.A., 2003. A re-evaluation of long-term flux measurement techniques: part 1. averaging and coordinate rotation. Boundary Layer Meteorol. 107, 1–48.

Gash, J.H.C., Dolman, A.J., 2003. Sonic anemometer (co)sine response and flux measurement I. The potential for (co)sine error to affect sonic anemometer-based flux measurements. Agric. Forest Meteorol. 119, 195–207.

Goodrich, D.C., Keefer, T.O., Unkrich, C.L., Nichols, M.H., Osborn, H.B., Stone, J.J., Smith, J.R., 2008. Long-term precipitation database, Walnut Gulch Experimental Watershed, Arizona, United States. Water Resour. Res. 44.

Goulden, M.L., Munger, J.W., Fan, S.M., Daube, B.C., Wofsy, S.C., 1996. Measurements of carbon sequestration by long-term eddy covariance: methods and a critical evaluation of accuracy. Global Change Biol. 2, 169–182.

Haverd, V., Cuntz, M., Leuning, R., Keith, H., 2007. Air and biomass heat storage fluxes in a forest canopy: calculation within a soil vegetation atmosphere transfer model. Agric. Forest Meteorol. 147, 125–139.

Kljun, N., Calanca, P., Rotach, M.W., Schmid, H.P., 2004. A simple parameterisation for flux footprint predictions. Boundary Layer Meteorol. 112, 503–523.

Kosugi, Y., Katsuyama, M., 2007. Evapotranspiration over a Japanese cypress forest. II. Comparison of the eddy covariance and water budget methods. J. Hydrol. 334, 305–311.

Kustas, W.P., Prueger, J.H., Hatfield, J.L., Ramalingam, K., Hipps, L.E., 2000. Variability in soil heat flux from a mesquite dune site. Agric. Forest Meteorol. 103, 249–264.

Kustas, W.P., Prueger, J.H., Hipps, L.E., Hatfield, J.L., Meek, D., 1998. Inconsistencies in net radiation estimates from use of several models of instruments in a desert environment. Agric. Forest Meteorol. 90, 257–263.

La Barbera, P., Lanza, L.G., Stagi, L., 2002. Tipping bucket mechanical errors and their influence on rainfall statistics and extremes. Water Sci. Technol. 45, 1–9.

Lane, L.J., Kidwell, M.R., 2003. Hydrology and Soil Erosion, Santa Rita Experimental Range: 100 Years (1903 to 2003) of Accomplishments and Contributions. USDA Forest Service, Rocky Mountain Research Station, Ogden, UT, p. 197.

Larson, L.W., Peck, E.L., 1974. Accuracy of precipitation measurements for hydrologic modeling. Water Resour. Res. 10, 857–863.

Lee, X., Finnigan, J., Paw, U.K., 2004. Coordinate systems and flux bias error. In: Lee, X., Massman, W., Law, B. (Eds.), Handbook of Micrometeorology. Kluwer Academic Publishers, Dordrecht, pp. 33–66.

Liu, H., Randerson, J.T., Lindfors, J., Massman, W.J., Foken, T., 2006. Consequences of incomplete surface energy balance closure for CO_2 fluxes from open-path CO_2/H_2O infrared gas analysers. Boundary Layer Meteorol. 120, 65–85.

Malhi, Y., Nobre, A.D., Grace, J., Kruijt, B., Pereira, M.G.P., Culf, A., Scott, S., 1998. Carbon dioxide transfer over a central Amazonian rain forest. J. Geophys. Res. 103, 31593–31612 [Atmos.]

Massman, W.J., 2001. Reply to comment by Rannik on "a simple method for estimating frequency response corrections for eddy covariance systems". Agric. Forest Meteorol. 107, 247–251.

Mayocchi, C.L., Bristow, K.L., 1995. Soil surface heat flux: some general questions and comments on measurements. Agric. Forest Meteorol. 75, 43–50.

Meyers, T.P., Hollinger, S.E., 2004. An assessment of storage terms in the surface energy balance of maize and soybean. Agric. Forest Meteorol. 125, 105–115.

Nearing, M.A., Nichols, M.H., Stone, J.J., Renard, K.G., Simanton, J.R., 2007. Sediment yields from unit-source semiarid watersheds at Walnut Gulch. Water Resour. Res. 43, W06426 10.1029/2006WR005692.

Oncley, S.P., Foken, T., Vogt, R., Kohsiek, W., DeBruin, H.A.R., Bernhofer, C., Christen, A., van Gorsel, E., Grantz, D., Feigenwinter, C., Lehner, I., Liebethal, C., Liu, H., Mauder, M., Pitacco, A., Ribeiro, L., Weidinger, T., 2007. The Energy Balance Experiment EBEX-2000. Part I: overview and energy balance. Boundary Layer Meteorol. 123, 1–28.

Paw, U.K.T., Baldocchi, D.D., Meyers, T.P., Wilson, K.B., 2000. Correction of eddy-covariance measurements incorporating both advective effects and density fluxes. Boundary Layer Meteorol. 97, 487–511.

Sauer, T.J., Meek, D.W., Ochsner, T.E., Harris, A.R., Horton, R., 2003. Errors in heat flux measurement by flux plates of contrasting design and thermal conductivity. Vadose Zone J. 2, 580–588.

Schmid, H.P., 1994. Source areas for scalars and scalar fluxes. Boundary Layer Meteorol. 67, 293–318.

Schmid, H.P., 1997. Experimental design for flux measurements: matching scales of observations and fluxes. Agric. Forest Meteorol. 87, 179–200.

Schume, H., Hager, H., Jost, G., 2005. Water and energy exchange above a mixed European Beech—Norway Spruce forest canopy: a comparison of eddy covariance against soil water depletion measurement. Theor. Appl. Clim. 81, 87–100.

Scott, R.L., Huxman, T.E., Cable, W.L., Emmerich, W.E., 2006. Partitioning of evapotranspiration and its relation to carbon dioxide exchange in a Chihuahuan Desert shrubland. Hydrol. Processes 20, 3227–3243.

Scott, R.L., Jenerette, G.D., Potts, D.L., Huxman, T.E., 2009. Effects of seasonal drought leads on net carbon dioxide exchange from a woody-plant-encroached semiarid grassland. J. Geophys. Res. [Biogeo.] 114, G04004, doi:10.1029/2008JG000900.

Scott, R.L., Shuttleworth, W.J., Keefer, T.O., Warrick, A.W., 2000. Modeling multiyear observations of soil moisture recharge in the semiarid American Southwest. Water Resour. Res. 36, 2233–2247.

Sieck, L.C., Burges, S.J., Steiner, M., 2007. Challenges in obtaining reliable measurements of point rainfall. Water Resour. Res. 43, W01420 10.1029/2005WR004519.

Stone, J.J., Nichols, M.H., Goodrich, D.C., Buono, J., 2008. Long-term runoff database, Walnut Gulch Experimental Watershed, Arizona, United States. Water Resour. Res. 44, W05S05 10.1029/2006WR005733.

Twine, T.E., Kustas, W.P., Norman, J.M., Cook, D.R., Houser, P.R., Meyers, T.P., Prueger, J.H., Starks, P.J., Wesely, M.L., 2000. Correcting eddy-covariance flux underestimates over a grassland. Agric. Forest Meteorol. 103, 279–300.

Van Der Molen, M.K., Gash, J.H.C., Elbers, J.A., 2004. Sonic anemometer (co)sine response and flux measurement: II. The effect of introducing an angle of attack dependent calibration. Agric. Forest Meteorol. 122, 95–109.

Webb, E.K., Pearman, G.I., Leuning, R., 1980. Correction of flux measurements for density effects due to heat and water vapour transfer. Q. J. R. Meteorol. Soc. 106, 85–100.

Wilson, K., Goldstein, A., Falge, E., Aubinet, M., Baldocchi, D., Berbigier, P., Bernhofer, C., Ceulemans, R., Dolman, H., Field, C., Grelle, A., Ibrom, A., Law, B.E., Kowalski, A., Meyers, T., Moncrieff, J., Monson, R., Oechel, W., Tenhunen, J., Valentini, R., Verma, S., 2002. Energy balance closure at FLUXNET sites. Agric. Forest Meteorol. 113, 223–243.

Wilson, K.B., Hanson, P.J., Mulholland, P.J., Baldocchi, D.D., Wullschleger, S.D., 2001. A comparison of methods for determining forest evapotranspiration and its components: sap-flow, soil water budget, eddy covariance and catchment water balance. Agric. Forest Meteorol. 106, 153–168.

Wolf, A., Saliendra, N., Akshalov, K., Johnson, D.A., Laca, E., 2008. Effects of different eddy covariance correction schemes on energy balance closure and comparisons with the modified Bowen ratio system. Agric. Forest Meteorol. 148, 942–952.