Evapotranspiration information reporting: I. Factors governing measurement accuracy

Richard G. Allen
University of Idaho, RAllen@kimberly.uidaho.edu

Luis S. Pereira
Technical University of Lisbon, lspereira@isa.utl.pt

Terry A. Howell
USDA-ARS, Terry.Howell@ars.usda.gov

Marvin E. Jensen
Colorado State University, MJensen419@aol.com

Follow this and additional works at: https://digitalcommons.unl.edu/usdaarsfacpub

Part of the Agricultural Science Commons

https://digitalcommons.unl.edu/usdaarsfacpub/829

This Article is brought to you for free and open access by the U.S. Department of Agriculture: Agricultural Research Service, Lincoln, Nebraska at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Publications from USDA-ARS / UNL Faculty by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.
Evapotranspiration information reporting: I. Factors governing measurement accuracy

Richard G. Allen \textsuperscript{a,\textdagger}, Luis S. Pereira \textsuperscript{b}, Terry A. Howell \textsuperscript{c}, Marvin E. Jensen \textsuperscript{d}

\textsuperscript{a} Biological and Agricultural Engineering and Civil Engineering, Research and Extension Center, 3793 N. 3600 E., University of Idaho, Kimberly, ID, 83341, USA
\textsuperscript{b} Biosystems Engineering, Institute of Agronomy, Technical University of Lisbon, Tapada da Ajuda, 1349-017 Lisbon, Portugal
\textsuperscript{c} USDA-ARS, Bushland, TX, USA
\textsuperscript{d} Formerly USDA-ARS and Colorado State University (ret.), Ft. Collins, CO, USA

\textsuperscript{\dagger} Corresponding author. Tel.: +1 208 423 6601.
E-mail addresses: rallen@kimberly.uidaho.edu (R.G. Allen), lspereira@isa.utl.pt (L.S. Pereira), Terry.Howell@ars.usda.gov (T.A. Howell), MJensen419@aol.com (M.E. Jensen).

\begin{abstract}
More and more evapotranspiration models, evapotranspiration crop coefficients and associated measurements of evapotranspiration (ET) are being reported in the literature and used to develop, calibrate and test important ET process models. ET data are derived from a range of measurement systems including lysimeters, eddy covariance, Bowen ratio, water balance (gravimetric, neutron meter, other soil water sensing), sap flow, scintillometry and even satellite-based remote sensing and direct modeling. All of these measurement techniques require substantial experimental care and are prone to substantial biases in reported results. Reporting of data containing measurement biases causes substantial confusion and impediment to the advancement of ET models and in the establishment of irrigation water requirements, and translates into substantial economic losses caused by misinformed water management.

Basic principles of ET measuring systems are reviewed and causes of common error and biases endemic to systems are discussed. Recommendations are given for reducing error in ET retrievals. Upper limits on ET measurements and derived crop coefficients are proposed to serve as guidelines. The descriptions of errors common to measurement systems are intended to help practitioners collect better data as well as to assist reviewers of manuscripts and users of data and derived products in assessing quality, integrity, validity and representativeness of reported information. This paper is the first part of a two-part series, where the second part describes recommendations for documentation to be associated with published ET data.

© 2011 Elsevier B.V. All rights reserved.
1. Introduction

Evapotranspiration (ET) is typically modeled using weather data and algorithms that describe surface energy and aerodynamic characteristics of the vegetation. ET is typically measured using systems that require the employment of relatively complex physical principles and techniques. In many agricultural systems, plant density, height, vigor and water availability are generally uniform, and the application of estimation algorithms and the measurement of ET can be relatively straightforward, although they are still not without substantial challenge. In the case of non-agricultural systems such as forest, desert and riparian systems, the heterogeneous nature of vegetation, terrain, soils and water availability make surface energy and aerodynamic processes highly variable and poorly defined. This is especially true, for example, for riparian systems such as cottonwood, tamarisk and Russian olive in semiarid regions that can have widely varying vegetation density, tree height, stand extent and availability of water. Most information and estimates of water consumption by forest and riparian systems come from in-place measurements that have a strong empirical and local character. ET data and ET models or model calibrations reported in the literature for even ‘well-behaved’ agricultural systems can contain serious biases caused by flaws in experimental design, measurement equipment, vegetation management, data reduction, model parameterization, and interpretation of results. Therefore, it is essential that reporting of ET measurements and related products such as crop coefficients or parameterized models contain sufficient description of the procedures used to measure and derive ET information so that readers can be aware of potential flaws or shortcomings in data measurement and can be alerted to the need to question representativeness of ET presentations. ET information is more and more frequently used as a foundation for court determinations of injury among water users, for parameterization of important hydrologic and water resources planning and operation models, for operating weather and climate change forecasting models, and for water management and allocation in water-scarce regions, including the partitioning of water resources among states and nations. All too frequently the ET information used in these processes is deficient or uncertain, with too little descriptive information in the reporting to facilitate judgment of its quality.

Because of the wide range of complexities in making ET and associated weather measurements and the abundance of opportunities for biases to enter ET and weather data sets, users of ET literature need sufficient information reported in articles on ET to assess the likelihood for opportunities of bias or error to enter reported data as well as sufficient information to examine or recreate the reported data using some type of ET model. This is currently often not the case, and many journal articles do not contain sufficient information to enable readers to gauge accuracies and representativeness of information. This article is part one of a two-part series on I: ET measurement requirements and accuracies and II: ET reporting recommendations (Allen et al., 2011a). This first article describes common ET measuring systems including water balance, lysimeters, Bowen ratio, eddy covariance, scintillometry, sap flow and remote sensing. The second article lays out recommendations for the type and nature of useful documentation and description of information that should accompany ET findings reported in ET-related articles. In this first article, common errors, biases and shortcomings of common ET measuring systems are discussed to provide support for why the accompanying reporting information is needed.

Measurements of ET include a variety of methods ranging from soil water sampling to lysimeters to eddy covariance to scintillometry. Inherent to all of these methods is the reality that an improperly designed experiment or measurement can lead to highly erroneous water use estimates. Many of the erroneously high ET estimates reported in the literature violate the law of conservation of energy that governs the conversion of liquid water to vapor during the transpiration and evaporation processes. The environmental energy provided by solar radiation plus heat energy advected to the vegetation may be insufficient to explain the measurements. Relatively simple comparisons with reference ET estimates based on available energy are recommended to give cause for review of data and measurement procedures.

2. The case for limits on maximum values for ET and crop coefficients

Before addressing challenges and precautions with ET measurement systems, it is important to discuss what constitutes realistic limits on rates of ET. Evaporation constitutes the conversion of liquid water to vapor and as a result requires substantial amounts of energy. The availability of energy incident to vegetation places a constraint on the potential evaporation rate and forces adherence to the law of conservation of energy. ET rates that exceed available radiation energy (\(R_n\)) at the surface less the energy conducted as sensible heat to the ground (\(G\)), i.e., \(R_n - G\), must essentially extract that additional energy from the atmosphere via downward (negative) sensible heat flux (\(H\)) via convective transfer through the equilibrium boundary layer of air above the surface. Because increasingly negative \(H\) creates increasingly stronger density-induced stability to the equilibrium boundary layer, it becomes increasingly more difficult to transport the required \(H\) to the surface to support the conversion to ET, especially without strong mechanical mixing brought about by high wind speed (Brutsaert, 1982; De Bruin et al., 2005). As a result there is an upper limit on ET, even under extreme advection, caused by limitations on aerodynamic transport and on equilibrium forces above a vegetation canopy. That upper limit on ET is relatively well represented by the tall (alfalfa) reference that has been defined by ASCE-EWRI (2005) using a parameterized Penman–Monteith equation (Allen et al., 1989, 2007c).

The upper limit on potential crop evapotranspiration (\(ET_c\)) is readily approximated by comparing against the widely used reference ET (\(ET_{ref}\)) through a crop coefficient (\(K_c\)). \(ET_{ref}\) may refer to two types of reference crops, clipped, cool-season grass or tall alfalfa (whose common symbols are \(ET_0\) and \(ET_r\), respectively), thus crop coefficients may be expressed in relation to clipped, cool-season grass as more often used (Allen et al., 1998, 2007c) or to alfalfa; for which the symbol \(K_c\) is adopted (ASCE-EWRI, 2005; Allen et al., 2007c). An alternative and synonymous expression for \(K_c\) can be used, which is the term alfalfa reference ET fraction, \(ET_F\) (ASCE-EWRI, 2005; Allen et al., 2007a). The terms \(K\) and \(E_T\) are simply defined as the ratio of ET for a specific surface, \(ET_c\), to the ET of the standard reference surface, \(ET_{ref}\). The crop coefficient was defined in 1968 (Jensen, 1968) for use with a reference crop \(ET_{ref}\) and first used in computerized irrigation scheduling by Jensen (Jensen, 1969; Jensen et al., 1970; Jensen et al., 1971). One can express ETc...
and ETref in terms of the well-known Penman–Monteith equation (Monteith, 1965), so that the ratio of ETc to ETref, i.e., Kc (or ETc/ETF) is

$$K_c = \frac{E_{Tc}}{E_{Tref}} = \frac{[\Delta_h (R_h - G) + \rho_s e^s_c (e^s_c - e^s_h)/r_s]/[\Delta + \gamma(1 + r_s/r_a)]}{[\Delta_h (R_h - G) + \rho_s e^s_c (e^s_c - e^s_h)/r_s]/[\Delta + \gamma(1 + r_s/r_a)]},$$

(1)

where R₀ is net radiation, G is soil heat flux density, ρ₀ is air density, e^s₀ is specific heat of air, e^s_c is saturation vapor pressure at z height, e_h is actual vapor pressure at z height, r_s is aerodynamic resistance to heat and vapor transport from the surface to the surface at z height, Δ is the slope of the saturation vapor pressure–temperature curve, ρ is the psychrometric constant, and r_s is bulk surface resistance. The “c” subscripted parameters in the numerator represent characteristic values for the actual vegetation and the “r” subscripts in the denominator represent the same for the reference vegetation, grass or alfalfa (see Allen et al., 2007c). Using this ratio, one can visualize that the value for K_c (or ETc/ETF) depends on the relative roughness, leaf area, and albedo (impacting R₀) of the actual vegetative surface in relation to the same characteristics of the grass or alfalfa reference surface. The relative proportions of net radiation, wind, temperature and vapor pressure deficit above each surface all affect the value of K_c (or ETc/ETF) to some degree, and therefore, even Δ and e^s_c and e^s_h are somewhat tied to the specific surface. Clearly, the more similar a vegetative cover is to the reference condition, especially at full cover, the closer the value of K_c (or ETc/ETF) will be to 1.0 and the less variation in the value of K_c (or ETc/ETF) with changing weather conditions (see discussions in Pereira et al., 1999 and ASCE-EWRI, 2005).

Implicit to Eq. (1) is the consideration of impacts that differences between ETc and ETref have on conditioning of the equilibrium boundary layer above the surfaces and on conditioning of the surfaces themselves. This conditioning modifies values of e^s_c, e^s_h, Δ and wind speed caused by differences in cooling and humidification brought about by rates of ETc being different from ETref. This conditioning generally creates a negative feedback effect on the value for K_c (or ETc/ETF).

Two standardized calculation procedures for the ETref have adopted the Penman–Monteith (PM) approach and are defined for two reference surfaces. One, a 0.12-m tall cool-season clipped grass, follows the definition by FAO-56 (Allen et al., 1998, 2006) and uses the PM equation with specific parameterizations that now differ from the original 1998 FAO-56 definition only in the value for bulk surface resistance prescribed for hourly or shorter timestep calculations, where r_s was reduced (Allen et al., 2006) from the daily value of 70 s m⁻¹ to a new value of 50 s m⁻¹ for hourly or shorter periods during daytime. The second standardized reference surface is the ‘tall’ reference, defined to be very similar to a 0.5 m tall crop of dense, full-cover alfalfa (lucerne) having surface resistance of 30 s m⁻¹ for hourly or shorter calculation timesteps during daytime and 45 s m⁻¹ for daily calculation timesteps (ASCE-EWRI, 2005).

The tall (alfalfa) reference and its parameterizations in the PM equation tend to represent sufficiently low surface resistance and sufficiently high aerodynamic roughness to approximate near maximum rates of ET expected from large expanses of well-watered vegetation cover, even under conditions of strong regional advection (Allen et al., 1989, Pereira et al., 1999). The tall (alfalfa) reference is in many ways superior to the shorter, smoother and lower conductance clipped grass reference in that alfalfa does represent conditions for a near maximum possible ET rate. Therefore, alfalfa reference-based crop coefficients tend toward a maximum of 1.0, and an expanse of tall, well-watered vegetation that fully covers the ground will approach the maximum conversion of available energy into latent heat flux (λE), and the maximum ratio of λE for any tall vegetation to alfalfa λE will be near 1.0. This observation is born out in viewing the maximum values for K_c reported by Wright (1982) for nine annual crops in southern Idaho, where none of Wright’s K_c’s, when based on the standardized alfalfa reference (Allen and Wright, 2002), exceeded 1.03. In the case of the grass reference, values for K_c may approach 1.3 for tall, dense crops under windy arid and semi-arid conditions because of the smoother roughness and small leaf area index of the grass reference as compared to the alfalfa reference.

In humid climates, ET is dominated more by net radiation availability and less by aerodynamics and vapor deficit. Because the grass and alfalfa reference crops have about the same albedo as many crops at full cover, their energy, heat and vapor transfer rates and thus ET and K_c rates are more similar in humid and semi-humid climates than under arid conditions. Based on total energy constraints, under humid conditions, where a majority of energy for the ET process is from net radiation and regional advection is relatively minor, the K_c generally cannot exceed about 1.0–1.05 relative to the alfalfa reference for large expanses of similar vegetation and about 1.2–1.3 relative to the grass reference.

In arid and semi-arid climates, the constraint of K_c ~ 1.0–1.1 for the alfalfa reference holds because of the tendency toward similarity in aerodynamic exchange and leaf area between even tall vegetation and the alfalfa reference. In the case of the grass reference, however, differences in aerodynamic and surface conductances, coupled with potentially strong regional advection, may cause K_c based on the grass reference in arid and semi-arid climates to be as high as 1.2 to exceptionally high values of 1.4 for tall, well-watered vegetation.

H.L. Penman’s comment (Penman, 1948) that crop evaporation should never exceed evaporation from an open water surface may be somewhat extreme; however the exceedance of measured or reported K_c above about 1.0 for alfalfa reference or 1.2 for grass reference in subhumid regions or above about 1.0 for alfalfa reference or about 1.2–1.4 for grass reference in arid regions should give cause for intense scrutiny of the ET measurements, the weather data used to calculate ETref, and the data processing procedure. Exceeding those values should also give cause for rejecting the data.

Limiting K_c to approximately 1.0 for an alfalfa reference-base or 1.2 for a grass reference-base in humid climates and to approximately 1.0 for an alfalfa reference-base or 1.2–1.4 for a grass reference-base in arid climates applies to large expanses of vegetation (>200 m diameter) and is significant and important when evaluating field measurements of ET. Measuring ET from small expanses of vegetation should be avoided when the objective of the ET measurement is to represent general conditions of crop ET for medium to large (say >200 m) fields or clusters of small fields. When ET is measured from small expanses of vegetation, the internal boundary layer above the vegetation may not be in equilibrium with the surface and may not have developed up to the height of any meteorological or flux instrumentation. In addition, small expanses of tall vegetation surrounded by shorter cover cause a “clothesline effect” where the interchange between air and vegetation is much more efficient than with the logarithmically shaped boundary layer profiles established over large fields and that are assumed in essentially all aerodynamically based ET equations. In these cases, ET from the isolated stands, on a per unit area basis, may be significantly greater than the corresponding ETref computation and will not represent large expanses. An example of these situations would be ET from a single row of trees surrounded by short vegetation, ET from a narrow strip of cattails along a stream channel, or a vegetated lysimeter surrounded by shorter vegetation (or no vegetation). Allen et al. (1992) reported Kc values for small (6-m wide) stands of cattails and bulrushes surrounded by grass pasture equal to 1.6–1.8 during midseason, relative to an alfalfa reference. These measurements indicated a strong clothesline effect. Coefficients were only 1.15 for a cattail wetland that was 200 m in diameter (Allen et al., 1992). While not incorrect, the clothesline-influenced Kc values must be applied only in the same
clothesline-influenced context as existed for their origin and must never be used to represent or estimate ET from more extensive surfaces.

In extreme illustrations, van Bavel et al. (1963) measured ET from 1-m tall sudangrass in Arizona following cutting of the grass around the lysimeter, so that the vegetation inside the lysimeter functioned as a clothesline. After cutting, 15 mm of ET during a 24-h period was measured compared with 10 mm three days earlier before the cutting—a 50% increase. The weather data were similar for all days. In a similar situation, Allen et al. (1991b) reported measured ET from 0.6 m fescue grass to increase by 1.6 times relative to the Penman–Monteith equation when the surrounding grass was clipped to 0.1 m, but the vegetation inside the lysimeter remained at 0.6 m. The ET rate from the lysimeter under the strong clothesline condition reached 16 mm d$^{-1}$, whereas the Penman–Monteith equation estimated 11 mm d$^{-1}$ for 0.6 m grass having extensive fetch of other 0.6 m grass.

The preceding discussion indicates the importance of knowing the type of setting for which ET estimates are intended and the conditions from which measurements are acquired. If ET estimates are intended for small, isolated stands of vegetation, then the $K_c$ may be allowed to exceed the 1.0 value for an alfalfa reference and the 1.3–1.4 value for a grass reference by up to 50–100%. However, if ET estimates are to represent large expanses of vegetation or small stands of vegetation surrounded by mixtures of other vegetation having similar roughness and soil water conditions, then $K_c$'s must generally be less than or equal to 1.0 for alfalfa and less than or equal to 1.2–1.4 for grass references.

When comparing ET measurements to the ET reference to assess adherence to energy limitations, the tall reference of ASCE-EWRI (2005) is recommended over the grass reference because it is more similar, roughness and conductance-wise, to most “tall” vegetation and its near maximum ET serves as a basis to expect maximum $K_c$ and ET, $F$ values to be around 1.0.

When calculating reference ET, the global standardizations of ASCE and FAO are highly recommended for consistency and transferability of results and derived coefficients (Allen et al., 1998, 2006; ASCE-EWRI, 2005).

3. ET measurement

The following sections describe primary methods for measuring ET or indirectly deriving ET estimates. The retrieved ET data are assumed to be applied to and to represent predominately extensive vegetation systems. Advantages and disadvantages of the methods are discussed as well as recommendations for proper and best operation. The descriptions and lists of advantages and disadvantages should be used as a guide when evaluating reported data. A table near the end of this article gives estimated ranges of errors in ET derived from the various measurements. Ranges in anticipated error are given for expert and novice users. Part II of this two-paper series (Allen et al., 2011a) lists the recommended metadata that should be recorded and reported in future studies of ET.

Accuracy of described measurement systems can be high when properly applied, but in many cases, knowledge of the underlying physics of turbulence and heat and radiation transfer that govern the measurement is essential to prevent subtle biases from degrading data accuracy. Deployment of equilibrium boundary layer systems such as eddy covariance, Bowen ratio and scintillometers must adhere to fetch requirements and minimum equipment heights to produce representative and valid data. Energy balance methods such as Bowen ratio and scintillometry must incorporate representative measurements of net radiation and soil heat flux density, which generally require multiple locations for sensors, especially when in spatially non-uniform systems such as riparian, forest, and other non-agricultural systems. Physically based “corrections” must be made to eddy covariance and scintillometry measurements. ET measurement systems are sometimes deployed by individuals who do not have sufficient background or experience, and, as a result, substantial measurement biases can occur.

Lysimeters and soil water balance methods are older methods that can potentially provide dependable estimates, but only if fundamental requirements concerning representativeness of vegetation and environmental conditions are satisfied. Sap flow methods depend on empirical correction factors derived from the physiology and anatomy of the species under investigation, and on the accuracy of the scaling methods used to go from branch or tree to plant stand and biome estimates of ET.

The following sections briefly describe ET measuring systems and various issues and requirements associated with each to obtain high integrity data.

3.1. ET using change in soil water

Determining ET by measuring the change in soil water over a period of time has been used for nearly a century. Up until the early 1960s, the primary measurement was by soil sampling and gravimetric analyses to determine the soil water content. Beginning in the 1960s, the neutron soil water probe largely replaced the gravimetric procedure except for evaluating soil water content in the surface 0–0.2, or 0–0.3-m layer and for calibrating the neutron probe. Since the 1980s, new types of electromagnetic devices based on dielectric and capacitance-based, time domain reflectometry-based and neutron thermalization methods, and have described how they function. The following material addresses mostly the overall management of sensor placement and data collection.

The average rate of ET in mm d$^{-1}$ is determined in the water balance method by noting the change in soil water content over time. ET is calculated from the change in total soil water between sampling dates plus rainfall minus any known drainage or surface runoff that may have occurred. A major potential source of error in ET determined by the soil water balance method is uncertainty in drainage from the zone sampled or any upward movement of water from a lower wetter zone into the zone sampled. These errors are difficult to detect, but they can be minimized with proper precautions, for example, by employing appropriate parametric modeling to estimate deep fluxes (Liu et al., 2006; Cholpankulov et al., 2008).

When using soil water sampling to determine ET, the soil is usually sampled 1 or 2 days after a major precipitation event or irrigation, to allow time for the majority of any drainage, and again 7–15 days later or just before the next major precipitation event or irrigation. With automated electronic measuring systems, the soil water profile can be sampled nearly continuously, and frequent monitoring is useful for understanding behavior. However using a long period, for example 7–15 days, to estimate ET via change in soil water content increases the measured difference in soil water content and improves the relative accuracy and precision of the ET measurement. The period immediately following the wetting event can potentially have higher ET rates relative to climatic demand caused by evaporation from wet soil and canopy. Omission of that period can cause the measured ET during the ‘dry’ interval to understate ET over the complete wetting and drying cycle. This can sometimes be corrected by simulating ET during and immediately following wetting events using an ET process model having an evaporation component, where basic vegetation parameters are calibrated using ET from the dry period, and then including those estimates in the total ET calculation. Combined approaches of soil water observations and simulation modeling provide for appropri-
Neutron probes and some electromagnetic devices use access tubes that are left in place and the water content at the same site is measured periodically. While in-place access tubes improve the accuracy and reproducibility of measurements and reduce spatially induced biases, a major problem encountered is damage to plants around access tubes by foot traffic of technicians making the readings. Neutron probes measure water content in a sphere of soil with a radius of 0.5 m or less around the sensor, hence even minor damage to the plant community around the probe port will affect readings. Data obtained from sites with damaged plants will cause ET to be underestimated due to reduced transpiration from damaged or reduced leaf surface and due to reduced infiltration on compacted soil. Allen et al. (1993) recommended the use of a suspended platform above the measurement site to protect vegetation and soil. Another complication with neutron probes is impact of variable access tube height above the soil surface. Evett et al. (2003) developed a fixed height stand that avoids that problem. In addition, shield counts must be taken at least 1.5 m above the surface to avoid thermalization effects by water in the surface soil. Impacts by water in the human body must also be avoided. Kamgar et al. (1993) described the number of neutron probe sites needed relative to plot size.

Time domain reflectometry (TDR), time domain transmission (TDT) and capacitance-based sensors and probes have become common for measuring soil water content directly or indirectly. Advantages of these electronic sensors are in situ measurements with nearly continuous recording. Disadvantages and challenges include the relatively small volume of soil sampled (on the order of 1 cm extent from the sensor as compared to 0.1–0.5 m for neutron probes) and the possibility of a need to calibrate sensors for soil type and depth (Evett et al., 2006). The small volume of soil sampled makes the influence of gaps and lack of contact between sensor and soil problematic (Evett and Steiner, 1995).

Determining reliable ET rates by soil sampling is challenging for natural vegetation, especially trees, due to inherent spatial variability in soil water extraction as compared to the small area sampled by the measurement technique. Therefore, these methods require adequate precautions such as: (1) using six or more sampling sites representative of general field conditions; (2) selecting sites where the depth to the water table is much greater than the root zone depth; (3) using only those sampling periods where rainfall is light since values for periods of high rainfall are questionable because runoff may occur and drainage may be excessive; (4) waiting at least two days after a moderate precipitation event (or normal irrigation) before taking the first sample, and longer if heavy rainfall or irrigation is involved and when the ET rate is small, but with correction for the surface evaporation occurring during and following the wetting event; and (5) using the active root zone depth for ET computations. Detailed discussions of the problems encountered in determining ET by soil sampling have been given by Robins et al. (1954), Jensen (1967), Jensen and Wright (1978) and Hignett and Evett (2002). General problems of soil sampling were discussed by Taylor (1955), Staple and Lehane (1962), Pratt et al. (1976), Evett and Steiner (1995), Sumner (2000) and Evett et al. (2006).

Other biases in ET from soil water balances stem from inaccuracies in measuring precipitation and irrigation additions or using data that are not representative for the soil sites. Precipitation gauges should be located within a few hundred meters of the sampled area because precipitation can vary widely even over a few km distance. In addition to biases in spatial variation, precipitation amounts are often underestimated by gauges for both rain and snow amounts due to venturi effects, blow-out, splash-out and evaporation. Data should be corrected using wind functions such as those recommended by the World Meteorological Organization (Yang et al., 1996). For example, that publication recommends a correction to liquid precipitation from a standard US National Weather Service 8 (203 mm) gauge as:

$$P_{\text{corrected}} = P_{\text{gauge}} \exp\left(0.062W_{\text{gauge}}^{0.58}\right),$$

where $W$ is wind speed at gauge height, m s$^{-1}$. This correction increases the estimated capture of $P$ by 8% for wind speed of only 2 m s$^{-1}$ and by 13% when wind speed is 4 m s$^{-1}$. In addition to wind induced effects, all tipping bucket rain gauges can lose, on average, one-half tip equivalent of rain between each wetting event through evaporation from the untipped bucket.

Uniformity of irrigation application and infiltration can be low and uncertain at even the 1 m scale. Some research facilities employ only level basin irrigation systems when conducting soil water balance based ET research to increase spatial uniformity of irrigation (T.A. Howell, USDA-ARS, personal communication). With careful water metering, it may be possible to achieve reproducible accuracy of 10–30 mm over a growing season with level basins. Other methods for irrigation, including furrow, graded border, sprinkler and drip irrigation will have much greater uncertainty. In the case of sprinkler and drip irrigation, multiple spatial measurements of irrigation amounts are needed unless the average irrigation amount over the monitored area is determined by change in soil water content or by change in mass of a lysimeter.

**Typical problems associated with estimating ET by root zone water balance using gravimetric, neutron scattering, capacitance, time domain reflectometry, or water potential measurement of soil water can involve inaccuracies caused by:**

- Large spatial and vertical variability of bulk density and water holding characteristics of the soil so that discrete measurements do not represent the integrated volume of soil and/or the full root zone depth; the vertical spacing of measurements must be congruent with the characteristics of the equipment used, the spatial variation in soil characteristics and the root uptake activity.
- Inaccuracies in measuring precipitation and irrigation additions or unrepresentativeness of data.
- Differential spatial wetting of soil due to local spatial variation in irrigation (or precipitation) additions, for example under tree, beneath the drip line, or between trees.
- Possible deep percolation losses or gains by capillary rise.
- Differential spatial extraction of soil water due to spatial variation in root systems, for example with trees.
- The possible need for calibration of sensors for soil type, density and depth (Evett et al., 2006).
- For some vegetation types, perhaps the most serious problem with gravimetric, neutron–scattering, and electromagnetic methods is lack of care in obtaining samples or taking readings (or installing access tubes or sensors) without significantly altering the plant cover at the sampling site. ET from trampled plants can be drastically reduced as compared to ET by plants in a pristine condition.

A related serious problem is the altering of density, aeration and infiltration characteristics of the surface soil from foot traffic, excavation or backfilling. If water intake and/or aeration at and near the sampling site are different from general field areas, then plant growth is very likely to be affected. This may reduce soil water extraction, leading to estimated ET unrepresentative of actual field ET.

Adopting a reliable soil water balance simulation model can be very helpful in identifying and overcoming most of these difficulties in estimating ET from soil water observations by providing a comparative means to estimate soil water behavior over time. In addition, results from a calibrated model may provide appropriate support for irrigation scheduling and estimation of deep fluxes.
3.2. Evapotranspiration from mass balance over large areas

The water balance technique, also referred to as the inflow-outflow or mass balance method, can be applied over large, integrated areas of land and water to develop watershed-scale estimates of ET. These types of estimates are useful for calibration or validation of watershed and remote-sensing-based models. Examples of ET by water balance are applications to large areas such as valleys where the inflow and outflow are determined from stream flow and precipitation measurements, and where the basin is confined to eliminate other significant sources of inflow or outflow such as groundwater underflow (Lowry and Johnson, 1942; Chow et al., 1988; McCabe and Markstrom, 2007; Gao et al., 2007; Dingman, 2008). Such studies generally provide only gross seasonal estimates of the average water evaporated and transpired from agricultural and non-agricultural areas within a watershed or project. The results represent ET from a combination of vegetation types and are generally applicable only to climatic, vegetation and water availability conditions similar to those existing in the study area. An evaluation of large-scale models of an experimental catchment (watershed) and a river basin in Australia indicated satisfactory agreement in monthly values between the computed and measured ET (Dunin and Aston, 1984). Wilson et al. (2001) found good agreement between ET derived from an inflow-outflow analysis of a 98 ha forested watershed and eddy covariance methods in Tennessee.

Mass balances require precipitation as the primary input and consequently the ET estimate can only be as accurate as the precipitation estimate. Fine resolution grids of precipitation on the scale of a km or less may be required in areas having substantial precipitation from convective storms and areas having substantial relief with orographic influence on precipitation amounts. As with soil water balance applications, precipitation data from standard gauges should be corrected for undercatch. Accurate measurement of basin or irrigation project outflow is necessary as is locating the measurement area boundaries to minimize ground water inflows and outflows.

3.3. Lysimetry

Lysimeters have been used extensively to provide baseline information for development, calibration, and validation of ET methods (Makkink, 1957; Jensen, 1974; Doorenbos and Pruitt, 1977; Wright, 1981, 1982; Allen et al., 1989; Jensen et al., 1990). Unfortunately, lysimeter measurements of ET are extremely sensitive to environmental factors, many of which are often poorly understood or have been ignored in practice. Lysimeter measurements are point measurements, representing measurements of ET from areas generally ranging from 0.05–40 m². However, a common usage of lysimeter measurements has been to characterize ET for large areas. Because of this extrapolation of data from small to large areas, it is paramount that vegetative and environmental conditions of lysimeter systems closely duplicate one-dimensional evapotranspiration from the larger areas. Insufficient description and documentation of environmental conditions and management of lysimeter data reported in the literature have led to uncertainty and confusion as to the quality and representativeness of reported data sets. Consequently, many poor and non-representative data sets of lysimeter ET reported in the literature have been used in important ET studies (Jensen, 1974), leading to inaccurate and misleading conclusions and with significant economic consequences. High resolution of precision lysimeter systems can give scientists and data users false senses of security concerning the quality and value of collected data. Values of 0.02–0.6 mm are commonly cited as “resolutions” or “precisions” of weighing lysimeter systems (Aboukhaled et al., 1982; Lourence and Moore, 1991; Howell et al., 1991), although, if improperly managed from an environmental context, measured ET can differ from actual ET from a one-dimensional extensive system of vegetation by as much as 50–100 percent. Machine and system precision can never substitute for, nor negate the need for, environmental, physical and physiological representativeness.

Lysimeters can be grouped into three categories: (1) non-weighing, constant water-table types that provide reliable data for weekly or longer time periods in areas where a high water table normally exists and where the water table level is maintained essentially at the same level inside as outside the lysimeter; (2) non-weighing, percolation types, in which changes in water stored in the soil are determined by sampling or neutron methods or by precision measurement of inputs, and the rainfall and percolate are measured (these types are often used in areas of high precipitation); and (3) weighing types, in which changes in soil water are determined either by weighing the entire unit with a mechanical scale, counter-balanced scale and load cell, directly suspended by load cells, or by supporting the lysimeter hydraulically. Weighing lysimeters, if well managed, will provide the most accurate data for short time periods, with accuracy depending on the range of dynamic scale, the use of counterbalancing, and resolution of load-cells; ET can be determined accurately over periods as short as 30 min with mechanical scales, but electronic load cell systems are used in nearly all modern weighing lysimeters. Hydraulically weighed lysimeters generally are not accurate for periods less than 24 h due to thermal and pressure effects. A detailed summary of the use of lysimeters for ET can be found in publications by Harrold (1966), Aboukhaled et al. (1982), Howell et al. (1985), Marek et al. (1988), and Pruitt and Lourence (1985). The proceedings of an international symposium on lysimetry organized by ASCE (Allen et al., 1991b) provides many examples of poor lysimeter systems along with guidelines for the operation of lysimeters to insure high quality data for ET and related environmental purposes.

Weighing lysimeters are isolated blocks of soil or tanks filled with soil and suspended on a weighing mechanism in which vegetation is grown under natural conditions. The amount of water lost by evaporation and transpiration is based on the change in mass of the lysimeter. This method provides a direct measurement of the soil water mass balance, which can be translated directly into ET after accounting for any managed drainage and runoff and irrigation and precipitation inputs (Howell et al., 1985; Phene et al., 1989). However, lysimeter data may not be representative of natural or field conditions. Soil conditions inside the lysimeters must be essentially the same as those outside to insure that the vegetation density, water availability, vigor, evaporation from the soil surface, and thus ET are the same. The lysimeter must be surrounded by the same vegetation that is growing in the lysimeter to insure one-dimensionality of the measurement (Allen et al., 1991a). Studies of advection at edges of fields suggest that the lysimeter should be located at least 100 m from the edge of the field or particular stand of vegetation represented by the lysimeter so that the equilibrium boundary layer of air can be considered to be fully adjusted above the lysimeter.

When lysimeter facilities are appropriate, results from drainage or percolation lysimeters can be accurate if there is accuracy in observing changes in soil water content. The use of models may help producing accurate Kc and ET results (Liu et al., 1998). In addition, observations provide for calibrating or validating a model, which may then be explored for supporting improved irrigation scheduling (Phene et al., 1989).

Lysimeters surrounded by sidewalks or gravel will not provide reliable data, nor will lysimeters planted to tall vegetation if surrounded by shorter vegetation, or planted to short vegetation and surrounded by taller vegetation (van Bavel et al., 1963). Differences in growth and maturity between the lysimeter plants and
surrounding plants and resulting ET measured in the lysimeter relative to ET measured in the surrounding area can be significant (Pruitt and Lourence, 1985). Some high crop coefficients reported in the literature may have been the result of such differences in plant growth between the lysimeter and the surrounding field or vegetation stand.

Some of the seriously flawed lysimeter results have been produced by a so-called “bloom effect” on small lysimeters where the area of exposed plant canopy has exceeded the assumed effective area of the lysimeter. Tanner (1967) provided an excellent discussion on lysimeters, and was probably one of the earliest to recognize how serious this problem can be. Coverage of this and other aspects of lysimetry was detailed by King et al. (1956), Pruitt and Lourence (1985), Meyer and Mateos (1990), Allen and Fisher (1990) and Allen et al. (1991b). Examples are frequent in the literature of crop or so-called environmental problems that beset lysimeter measurements and that cause lysimeter measurements of ET to deviate from one-dimensional ET measurements characteristic of extensive natural and agricultural environments. Allen and Fisher (1990) listed nine recommendations to help preserve the environmental integrity of a lysimeter installation. Effects of these environmental violations are significant and often invalidate lysimeter measurements for engineering or scientific use. Common environmental problems include:

- nonrepresentative gaps in vegetation along lysimeter boundaries,
- vegetation height and density differences between inside and outside lysimeters,
- different vegetation characteristics between adjacent and local fetch around the lysimeter and around the weather station,
- thermal conditions of the lysimeter soil block and differences from the surrounding profile,
- thermal conditions and advective effects due to a high, exposed rim or uninsulated tank,
- bulk density and lysimeter depth effects on root development, and
- effects of the specific moisture profile or water table within the lysimeter on evaporation and moisture extraction.

When calculating ET in units of depth, the change in lysimeter mass must be divided by the effective evaporating and transpiring area of the lysimeter. This effective area of evaporating foliage and radiation absorption can often be larger than the physical area of the lysimeter due to vegetation extending outside the lysimeter or due to micro advection of energy, via convection or thermal radiation, from immediately outside the inner lysimeter tank to inside. Incorrect estimation of the effective evaporating area of a lysimeter is probably the most common error made in computing lysimeter evapotranspiration. For example, if a lysimeter is constructed 1 m square, inside, and has a 10 cm total boundary thickness (including widths of both inner and outer tanks, if present), the average vegetative and evaporating area can be approximated as 1.1 m × 1.1 m. The 1.1 m length is the average dimension based on the inner and outer dimensions of the lysimeter, and is a reasonable approximation if vegetation from both inside and outside the lysimeter reach across the lysimeter rim and meet midway between outer and inner tanks. The ratio of vegetative area to inside area in this example is 1.1²/1.0² or 1.21. In other words, lysimeter measurements of ET, if based on the inner dimension of the lysimeter, only would be overstated by 21%. As a consequence, derived Kc values would be overstated by 21% as compared to field conditions. This type of adjustment is often needed for all types of lysimeters, including weighing and water-table types. Errors may be particularly important for lysimeters containing trees where the canopy area is often much larger than that of the lysimeter. The bias in reported ET is further exacerbated if the vegetation outside is not identical to that inside the lysimeter and if the outside vegetation does not grow right to the lysimeter edge. Otherwise, a ‘bloom’ condition can occur, where if the vegetation inside the lysimeter is taller, it will absorb even more solar radiation from horizontal interception and be subject to increased turbulent penetration of air, and therefore the effective area of ET is even larger.

The occurrence of dissimilar or missing vegetation directly outside the lysimeter is all too common in practice. In these situations, determination of the true, equivalent area of transpiring vegetation may be complicated by mismatched “leaning” of inside and outside vegetation across lysimeter borders. Outside vegetation leaning into the lysimeter “robs” evaporative energy from the inside vegetation, whereas inside vegetation leaning outside the lysimeter increases evaporative energy via solar radiation interception, and water extraction by the lysimeter system increases. One such effect is commonly referred to as the “clothesline” effect, where, due to the taller nature of lysimeter vegetation or increased leaf area, both aerodynamic and radiative transfer to the lysimeter canopy are increased, resulting in increased ET from the lysimeter area. Increased vegetation height within a lysimeter can result from dissimilar fertilization practices, thermal characteristics of the lysimeter tank, differences in soil moisture regimes inside and outside the lysimeter, or due to reduced plant growth outside the lysimeter caused by foot or vehicle traffic, soil compaction, moisture availability or plant density. Taller vegetation receives evaporative energy not only in the vertical, one-dimensional plane, but also receives side loading of radiation and substantial increases in turbulent exchange of vapor and heat. This side loading significantly alters lysimeter measurements from the one-dimensional case and eliminates them from consideration for agricultural representation. The smaller the lysimeter area, the more pronounced the effect. Small lysimeters planted with trees are very prone to errors.

One disadvantage of lysimeter systems is that the soil profiles are commonly too shallow compared to the surrounding soil profile. As a result, drainage and soil water profiles may differ. Soil profiles of lysimeters may be wetter than surrounding soils. These profiles may affect the total availability of water to plant roots and subsequently the ET process. They may also affect root development when a water table has been established both inside and outside of the lysimeter, but at different elevations. Presence of a shallow water table or excess moisture in a lysimeter used for soil evaporation studies may also affect results.

Measuring ET from riparian systems and phreatophytes presents a special problem in lysimeter studies. Under natural conditions phreatophytes may be rooted into an aquifer that might be several meters below the ground surface, and where transpiration can be limited by soil properties such as hydraulic conductivity. In shallow lysimeters, however, where the soil profile may have been disturbed during construction, roots may be exposed to more ‘permissive’ hydraulic conductivities and improved access to water that might elevate ET well above normal field values. Black et al. (1968) and Dugas and Bland (1991) cautioned against negative impacts of thermal conduction of heat along metal lysimeter container walls that significantly warms the lysimeter soil, especially over long time periods in arid environments. This unnatural heat source can cause drying of the soil profile under bare soil conditions that exceeds that under natural conditions. It can also promote earlier and more extensive root growth and respiration within the lysimeter tank. Black et al. (1968) recommended insulation of the lysimeters. This is especially important for desert environments.
In summary, requirements of lysimeter-based measurement of ET are:

- extensive fetch around the lysimeter, for example a minimum of 50 m;
- vegetation inside and immediately outside the lysimeter must be very similar in:
  - height
  - density and ground coverage
  - leaf area
- vegetation inside and outside lysimeter should be carefully documented via photography and description;
- the effective area of the lysimeter vegetation must be accurately calculated;
- water management of the lysimeter must be precise and similar to outside conditions;
- salinity must be managed via drainage;
- thermal characteristics of the lysimeter soil must be similar to that outside;
- foot traffic should be minimized.

Advantages of lysimeters for ET measurement are:

- calibration can be validated mechanically;
- fetch requirement is smaller than for micrometeorological methods;
- lysimeters can be fully automated;
- the system is nearly fully visible for inspection.

Disadvantages of lysimeter systems are:

- lysimeter area often represents a small sample relative to a field area and therefore imperfections in vegetation or water management exacerbate measurement accuracy;
- difficult to maintain or to reconstruct original soil profile characteristics, including density, layering and structure during and following construction;
- difficult to reduce any edge effects in vegetation;
- difficult to measure for trees and plants having large spacing where often the area representing evaporation from soil is not correct;
- difficult to maintain exact field conditions for vegetation and soil water;
- difficult to reproduce rooting characteristics of the surrounding field;
- difficult to detect the presence of low levels of water stress due to subtle visual effects;
- difficult to accurately measure ET from non-potential (stressed) vegetation.

3.4. Bowen Ratio Energy Balance (BREB)

The Bowen Ratio Energy Balance (BREB) method is a practical and relatively reliable micrometeorological method. Use of the BREB concept (Bowen, 1926) enables solving the energy balance equation by measuring simple gradients of air temperature and vapor pressure in the near surface layer above the evaporating surface. The method works best when soil water is not limiting ET. As water becomes less readily available, the Bowen ratio (BR) increases, and the relative error in ET increases (Angus and Watts, 1984). Todd et al. (1998) and Devitt et al. (1998) reviewed ET studies based on BREB. Ohmura (1982) and Payero et al. (2003) described relative errors associated with the BREB method as well as techniques for data quality analysis.

The BREB equation for application to vegetation is as follows:

\[ \lambda E = \frac{R_n - G}{T + \beta} \]

where \( \lambda E \) is ET expressed in terms of energy, \( R_n \) is net radiation, \( G \) is soil heat flux density into the ground and \( \beta \) is the Bowen ratio. Given equivalent transport coefficients for sensible heat flux (\( K_h \)) and for vapor transport (\( K_v \)), \( \beta \) can be expressed in a finite difference form as:

\[ \beta = \frac{H}{\lambda E} = \frac{\rho_a c_p K_h [T_2 - T_1 + \Gamma (z_2 - z_1)]}{\rho_a c_p / \gamma K_v (e_2 - e_1) / e_2 - e_1}, \]

where \( H \) is sensible heat density, \( T_2 \) and \( e_2 \) are air temperature and vapor pressure at height \( z_2 \) and \( T_1 \) and \( e_1 \) are air temperature and vapor pressure at height \( z_1 \), \( \rho_a \) is density of air, \( c_p \) is specific heat of moist air, \( \gamma \) is the psychrometric parameter and \( \Gamma \) is the adiabatic lapse rate, generally taken as 0.01 °C m⁻¹ for non-saturated air. \( \gamma \) is proportional to barometric pressure and is influenced to some degree by air temperature. Generally the \( z_2 \) height should be at least 0.3 m above the crop canopy for a smooth, dense canopy, and should be placed further above the canopy for tall, sparse crops where microscale turbulence among individual plants and differences between source locations for heat and evaporation can disturb exponentially shaped temperature and vapor profiles. Generally, the \( z_2 \) height is 1–2 m above \( z_1 \).

Generally, \( \lambda E \) is computed each 20–30 min and summed over a daylight or 24-h period to provide daily estimates of ET. Problems with older, naturally aspirated (exposed) thermocouples included contamination by dust and spider webs that increased radiation loading on the sensor and thermal biases between \( T_1 \) and \( T_2 \). The more dependable and accurate BREB systems use interchanging arms to reduce impacts of bias in the \( T \) and \( e \) measurements (Fritschen and Fritschen, 2005). Other types of errors associated with BREB measurements are described by Sinclair et al. (1975). Whether or not correction is needed for air stability conditions in the use of the Bowen ratio–energy balance measurements is open to question. Most of the evidence suggests that correction for stability is not needed, due to nearly equal impact on both \( \lambda E \) and \( H \), so that the usual assumption of the near-equality of \( K_h \) and \( K_v \) transfer coefficients for heat and water vapor throughout a wide range of stability conditions is realistic (Cellier and Brunet, 1992). However, the major source elevations for \( \lambda E \) and \( H \) in the canopy should be similar for the transfer coefficients for the two processes to be nearly the same (Tanner, 1967). This requirement can be violated in sparse forest canopies and in row crops having incomplete ground cover.

Important advantages of the BREB method are the ability to measure ET even from non-well-watered vegetation surfaces, provided the \( \beta \) is less than about 1.0 to reduce relative error caused by measurement of small \( \Delta e \), the near elimination of the influence of turbulent transfer coefficients, and absence of surface and wind speed measurements. The disadvantages are relative sophistication and possible fragility of sensors, the numerical instability of Eq. (4) during periods of \( \beta \) near −1, and the heavy reliance of the ET estimate on the accuracy and representativeness of the \( R_n \) and \( G \) measurements. The requirement for adequate upwind fetch to establish an equilibrium boundary layer where temperature and vapor gradients are constant in horizontal space places substantial and important limits on the method. Details on BREB computations can be found in chapter 4 of the ASCE Hydrology Handbook (Allen et al., 1996) and in Fritschen and Fritschen (2005).

For tree crops and forest canopies, accurate measurement of the gradients at a height far enough above the canopy to avoid effects of individual trees is difficult due to the very small gradi-
ents involved. Nevertheless, the Bowen ratio method has served and continues to serve as one of the major methods employed for forest research (McNaughton and Black, 1973; Thom et al., 1975; McIlroy and Dunin, 1982; Denmead and Bradley, 1985; Fritschen and Simpson, 1985; Dunin et al., 1985, 1991). Accuracies of well-designed and operated BREB systems have been estimated to be approximately 10% (Sinclair et al., 1975; Seguin et al., 1982). Irmak (2010) described an extensive agricultural ET measurement network throughout Nebraska that is based on BREB systems having interchanging arms. A disadvantage of the BREB approach is the need to measure \( R_n \) and \( G \), which can be problematic under some conditions such as with sparse or heterogeneous vegetation, and over water surfaces, where estimating or measuring \( G \) has large uncertainty.

The net radiation measurement in the BREB calculation (and other energy balance systems) should be made from an elevation high enough to measure an average representative surface condition similar to that of the total contributing fetch upwind of the measurements. Because of the direct role that \( R_n \) plays in the \( \lambda E \) measurement, multiple \( R_n \) sensors are recommended for improved spatial sampling and quality control. In heterogeneous vegetation cover having vegetation type variation on the scale of tens of meters, multiple net radiation systems should always be employed (Fritschen and Fritschen, 2005). Idso and Cooley (1971, 1972) and Idso et al. (1974) provided guidance on net radiometer positioning and error analysis. Baldocchi et al. (2000) described a traversing net radiometer system for obtaining a spatially averaged \( R_n \) over forest. Anthoni et al. (2000) used ray tracing and rendering software to simulate impacts of differential surface temperature and reflectance in a juniper forest on spatially averaged hemispherical net radiation. Similar challenges exist over agricultural row crops, including orchards and vineyards where soil surface and canopy temperatures can deviate by 30 K so that the total hemispherical \( R_n \) measurement can be substantially impacted by sensor positioning. A 30 K difference in surface temperature can cause as much as 100 W m\(^{-2}\) difference in emitted long-wave radiation, and thus, 100 W m\(^{-2}\) difference in the measured \( R_n \). The user must be aware that the \( R_n \) sensor ‘sees’ in a 360° field of view both horizontally and vertically. Therefore, it will ‘see’ and be impacted by temperature and reflectance of objects within its view, including mounting tower and solar panels. Fritschen and Fritschen (2005) noted that 95% of the view of a radiometer is contained within a circle having diameter nine times the height of the instrument above the surface. A \( R_n \) sensor will not ‘magically’ provide accurate and representative \( R_n \) measurements simply because it is new and plugged into a datalogger. It must be placed in a proper representative environment. Part II of this paper series (Allen et al., 2011a) describes quality control/quality assurance approaches for \( R_n \) measurements.

Soil heat flux density is generally measured at 0.05–0.15 m below the surface using soil heat flux plates (Campbell Scientific Inc., 2003; Ham, 2001, Sauer and Horton, 2005). The 0.05–0.15 m depth range is recommended to ensure that the soil flux density is measured below the zone of soil water vaporization (Ham, 2001) and to reduce the influence of vertical conduction of heat near the soil heat flux plate (Sauer and Horton, 2005). Mayocchi and Bristow (1995) showed errors in \( G \) of 30 W m\(^{-2}\) for soil heat flux plates buried more shallow than 0.03 m. Sensible heat absorption and release above soil heat flux plates is estimated by measuring soil temperature change above the plate at multiple depths and at multiple locations and calculating corrections that must include accurate measurement or estimation of soil water content of the soil above the plate. The corrected soil heat flux density, \( G \), is computed by combining the heat flux measured by the heat flux plate system with energy change above the plates. Generally, two or more installations of soil heat flux systems are made per BREB site to reduce effects of spatial heterogeneity and to improve representation for the area. In some cases, soil heat flux plates have been placed at more shallow depths, for example at 0.01 m by Baldocchi et al. (2000) and 0.02 m by Anthoni et al. (2000) under forest. The shallow placement depths reduce the uncertainty and errors associated with correcting for soil temperature change above the plate. However, the shallow placement may introduce much larger error associated with differences in thermal conductivity between the plate and soil, which can cause the heat flux to diverge around the plate or to converge through the plate. An even larger problem is caused by the impedance of movement of soil water around the plate during infiltration, and disruption of upward flow of water to replace evaporation above the plate, which becomes more pronounced as the plate is placed closer to the surface.

Because \( G \) is quite sensitive to heating of the surface, which in turn is governed by surface wetness and surface shading, soil heat flux stations are needed in both sunlit and shaded portions of the soil. Further, stations should be located in both wet and dry environments, for example, under a canopy and in the open. In wild, heterogeneous systems having a range of vegetation types, further sampling may be needed. In forest and riparian situations, the number of soil heat flux sites needed to provide sufficient sampling of \( G \) may exceed 20.

Allen (2008) described the use of multiple linear regression to close the surface energy balance for eddy covariance systems. One application suggested 100% underestimation of the \( G \) component in an eddy covariance study in Idaho under uniform alfalfa when two soil heat flux stations at 0.10 m depth and four soil temperature stations were utilized. The regression process suggested a 200% underestimation of \( G \) under alfalfa in an Oklahoma study where the soil heat flux plates were installed at 0.01 m depth. These analyses support the need for multiple instrumented sites even under uniform vegetation and close scrutiny of installation and collected data.

Requirements of the Bowen Ratio method are as:

- uniform fetch of sufficient distance to establish an equilibrium boundary layer (EBL) deeper than the \( z_2 \) height (this is often violated);
- sufficient elevation above the canopy to avoid the roughness sublayer,
- representative measurement of \( R_n \) and \( G \),
- heterogeneous or sparse systems common to riparian vegetation generally require multiple net radiometers and soil heat flux stations.

Advantages of Bowen Ratio systems are:

- non-destructive, direct sampling of the turbulent boundary layer;
- no aerodynamic data are required;
- simple measurement of \( T \) and vapor pressure \( e \) at two heights;
- can measure ET over both potential and non-potential surfaces;
- gradient-based fluxes are averaged over a medium sized area (200–100,000 m\(^2\));
- automated.

Disadvantages of Bowen Ratio systems are:

- accuracy of ET depends substantially on the representativeness and accuracy of \( R_n \) and \( G \);
- assumes that transfer coefficients \( K_h \) and \( K_e \) for \( H \) and \( \lambda E \) are equal (therefore assumes that sources for heat and vapor are horizontally and vertically similar);
• numerically unstable when \( H \) is near zero (however, this usually causes only minor problems);
• measurements of \( T \) and \( e \) must be unbiased because \( \Delta T \) and \( \Delta e \) can be small;
• requires medium to large fetch;
• small stands of similar vegetation may be too small to use BCREB, for example, narrow riparian systems, so that the Bowen ratio system samples \( T \) and \( e \) gradient artifacts created by areas upwind of the vegetation of interest (for example, from upwind areas of pasture, desert or agriculture) due to an underdeveloped equilibrium boundary layer.

Recommendations for deployment of net radiometers:
• mount over representative vegetation away from any obstructions including:
  o away from towers;
  o away from above or below dark, hot solar panels;
• mount over representative vegetation away from any obstructions:
  o away from above or below other instruments, especially bright,
  o away from towers;
• when over complex or sparse canopies, deploy a minimum of three \( R_n \) sensors in separate locations:
  o mount to view the true proportion of sunlit vegetation, shaded vegetation, sunlit soil and shaded soil;
  o multiple towers for \( R_n \) may be needed;
  o deploy a minimum of six soil heat flux measurement systems under the same conditions;
• use some independent means for testing \( R_n \) measurements, for example, compare against remote-sensing-based \( R_n \) estimates, infrared thermometers, modeling \( R_n \) with standardized solar radiation-based algorithms, etc (Allen et al., 2011a).

3.5. Eddy covariance

Eddy covariance (EC) systems are becoming relatively widely used in ET measurement because of ease of set up, reduced costs for sensors, and the ability to co-measure \( H, \lambda E \) and \( CO_2 \) fluxes, depending on the equipment configuration. The concept of eddy covariance (also referred to as eddy correlation) draws on the statistical covariance (correlation) between vertical fluxes of vapor or sensible heat within upward and downward legs of turbulent eddies. This requires high speed measurement of \( T, w, e \) or \( q \), usually at frequencies of 5–20Hz (5–20 times per second) using quick response sensors. Ten Hz is common. Early examples of eddy instrumentation were described by Tanner (1988) and Tanner et al. (1993). Since then, many advances in instrumentation have been made and this method is now widely used as described by Wilson et al. (2002), Baldocchi (2003), Shaw and Snyder (2003) and Meyers and Baldocchi (2005). EC applications make high frequency sampling of the surface boundary layer using the statistical relationship (Swinbank, 1951):

\[
E = \rho_a Wq' \left(1 + \frac{q'}{1-q'}\right) + \frac{q'^2}{(1-q')} \left(\frac{H}{c_p T}\right),
\]

where \( \rho_a \) is density of moist air, \( P \) is atmospheric pressure, \( q' \) is the instantaneous deviation of specific humidity from mean specific humidity (\( q \)), \( e' \) is the instantaneous deviation of vapor pressure from mean vapor pressure (\( e \)), and \( w' \) is the instantaneous deviation of vertical wind velocity from mean vertical wind velocity (\( w \)). \( E \) is the evaporation rate with units of mass per unit surface area per time, for example, \( \text{kg} \cdot \text{m}^{-2} \cdot \text{s}^{-1} \). The overbar indicates means of the products of the instantaneous deviations over generally 15–30 min averaging periods. Webb et al. (1980) recommended water vapor and air density corrections to the calculation of \( E \) (Fuehrer and Friese, 2002; Leuning, 2007) that can be expressed as:

\[
E = \rho_a Wq' \left(1 + \frac{q'}{1-q'}\right) + \frac{q'^2}{(1-q')} \left(\frac{H}{c_p T}\right),
\]

where \( c_p \) is specific heat of moist air, and \( q \) and \( T \) are mean specific humidity and temperature over the averaging period.

Besides direct covariance determination of \( E, E \) can also be computed as a residual from the energy balance equation as \( \lambda E = R_n - G - H \), where sensible heat flux density is measured by eddy covariance as:

\[
H = \rho_a c_p W'T',
\]

where \( T' \) is the instantaneous deviation of air temperature from mean temperature (\( T \)). Computing \( \lambda E \) as a residual of the energy balance by measuring \( H \) via Eq. (7), so that \( \lambda E = R_n - G - H \), has the advantage of eliminating the requirement for the quick response hygrometer, which can be expensive and can create high frequency fallout caused by physical separation of the hygrometer from the sonic anemometer. The disadvantage of estimating \( \lambda E \) as a residual is the need to measure \( R_n \) and \( G \) accurately, which can be problematic under conditions such as with sparse or heterogeneous vegetation, and over water surfaces, as noted in the previous section. The energy balance residual equation can be used in an energy balance closure check on Eq. (5). Allen and Tasumi (2005) combined \( H \) determined by eddy covariance with the Bowen ratio from a BCREB system to determine \( \lambda E \) from open water to eliminate the need to measure or estimate \( G \) of water.

In general, the eddy correlation method requires personnel who are well-trained in electronics, turbulent theory, and biosciences. Instrumentation is relatively fragile and expensive. The vertical component of wind, \( w \), is generally measured using a sonic anemometer and \( T \) is measured using ultra fine wire thermocouples or using sonically determined temperature corrected for humidity effects. Specific humidity is measured using quick response hygrometers such as open path Lyman-alpha, Krypton or mid-infrared absorption hygrometers or using closed path gas analyzers (Buck, 1976; Campbell and Tanner, 1985; Tanner, 1988; Bubra and Anderson, 2008). All measurements must be made at nearly the same point in order to measure characteristics of the same eddy. For open path hygrometers, corrections must be made to \( E \) measurements to correct for temperature and humidity effects on air density of up and down portions of eddies (Webb et al., 1980; Leuning, 2007). Fuehrer and Friese (2002) provided a complete review of density- and humidity-based corrections. Other corrections are required for instrument separation, frequency response, coordinate rotation and to account for the type of hygrometer (Baldocchi et al., 1988; Tanner et al., 1993; Moncrieff et al., 1996; Villalobos, 1997; Twine et al., 2000; Aubinet et al., 2000; Paw et al., 2000; Horst, 2000; Massman, 2000, 2001; Sakai et al., 2001; Rannik, 2001; Wilson et al., 2002). Coordinate rotations include those by Tanner and Thurtell (1969), Kaimal and Finnigan (1994), and Wilczak et al. (2001). Sonic temperature correction is done following Munger and Loecher (2004), air moisture correction following Sun et al. (1995), impacts of density terms following Webb et al. (1980) and Leuning (2007), and correction for pathlength averaging, sensor separation and high frequency spectral losses following Moore (1986) and Horst (2000). Paw et al. (2000) introduced corrections for advective effects and converging or diverging flow lines. A variety of software programs for correction of EC data have been produced, including ECPack from Wageningen University, TK2 software of the University of Bayreuth (Mauder and Foken, 2004), EddySoft developed at the Max-Planck-Institute for Biogeochemistry in Jena (Kolle and Rebmann, 2007), EdiRE software from the University of Edin-
burgh, and APAK by Oregon State University (Vickers and Mahrt, 1997, 2003). These software packages include a mixture of the corrections listed. In addition to these correction ‘issues’ are issues related to flux divergence in the equilibrium boundary layer, non-uniform vegetation or moisture conditions in the source area (i.e., footprint), non-representative sensing of \( R_n \) and \( G \), improper averaging times and impacts of very large, episodic eddies and heat plumes that have large vertical components that are very efficient in heat and vapor transport, but are largely missed by the sonic sensor or covariance algorithm. Burba and Anderson (2008) give a good overview of corrections and precautions for eddy covariance.

Much literature has appeared during the past fifteen years documenting energy balance ‘closure error’ for eddy covariance data where the sum of measured \( \lambda E + H \) does not equal measured \( R_n - G \) (Baldocchi et al., 1988; Twine et al., 2000; Wilson et al., 2002). Often \( \lambda E \) and \( H \) can be undermeasured relative to \( R_n - G \) by as much as 30% (Twine et al., 2000; Wilson et al., 2002; Foken, 2008a) even after extensive ‘correction’ (De Bruin et al., 2005). Wilson et al. (2002) found an average of 20% closure error at 22 Fluxnet sites over mostly forest, but including tundra, rangeland and agriculture.

Classic, possible reasons for the lack of closure that have been pointed out are storage of heat in canopies, horizontal advection, energy used by photosynthesis, change in storage of heat in the forest, but including tundra, rangeland and agriculture.

Many users ‘close’ the energy balance by scaling \( R_n \) and \( G \) in the same proportion (Twine et al., 2000) until the sum equals \( R_n - G \). However, this places all ‘blame’ for lack of closure on \( R_n \) and \( G \). Allen (2008) suggested using multiple linear regression, where the most (perceived) dependable component (of \( R_n \), \( G \), \( H \), and \( \lambda E \)) is set as the dependent variable and all others are set as independent variables, to indicate any systematic biases in components during daytime conditions, for example. Allen (2008) gave illustrations where measurements of \( G \) were indicated by regression to be undermeasured by as much as 100–200%. The multiple regression approach does not explain why a component is biased and can only identify consistent, systematic biases in components.

Recent work by Foken et al. (2006) and Foken (2008a) has suggested that the energy balance lack of closure is in fact a scale problem related to the physical and temporal characteristics of the EC sensors. They suggested that including effects of large, low frequency eddies that can often only be measured at a landscape scale by scintillometry or airborne sensors can approximately close the balance. The relative importance of these larger eddies may be connected to landscape heterogeneity. Moncrieff et al. (2010) suggested that large, low frequency eddies can be accommodated by using long average times in (4) of up to 4 h. Burba and Anderson (2008) suggest that energy consumed by photosynthesis and change in heat storage in the canopy are needed to close the energy balance, although many have argued that these components are generally small (Twine et al., 2000). Energy balance closure cannot be considered as a data quality test, since closure error can still be present even if data acquired by the eddy covariance system are of high quality and vice versa. Additional detail on the eddy correlation method can be found in articles by Dyer (1961), Businger et al. (1967), McBean (1972), Brutsaert (1982), Weaver et al. (1986), Twine et al. (2000), Wilson et al. (2002), Shaw and Snyder (2003) and Baldocchi (2003). Meyers and Baldocchi (2005) suggested that mean uncertainty in ET derived from eddy covariance can be as good as 10% with perfectly designed, maintained and managed instrumentation. They noted errors of approximately 10–30% for ozone flux estimates via covariance as being common.

Eddy covariance has disadvantages similar to other boundary layer sampling techniques, including the need for complex instrumentation and fetch requirements. The sonic anemometer should be set parallel to the surface so that no component of wind speed parallel to the surface biases measurements of vertical deviations. If not set parallel to the surface, coordinate rotation must be applied.

Requirements of the eddy covariance method are:

- uniform fetch of sufficient distance to establish an equilibrium boundary layer (EBL) deeper than the instrument height (this is often violated);
- sufficient elevation of instruments above the canopy to reduce roughness sublayer distortions and to increase eddy size to match the sensor path length;
- a number of “corrections” are required;
- accurate, high frequency instrumentation and knowledge of physics of turbulence are required.

Advantages of the eddy covariance method are:

- non-destructive, continuous direct sampling of the turbulent boundary layer;
- can measure ET over both potential and non-potential surfaces;
- fluxes are averaged over medium sized (50–200 m) areas;
- systems are automated.

Disadvantages of the eddy covariance method are:

- a number of “corrections” are needed—these are often not well defined and are often empirical;
- may miss transport by very small and by very large eddies;
- energy balance closure error \( (R_n - G = \lambda E + H) \) can be 10–30%;
- requires substantial fetch, generally 50–100 times the height of the instrument above the zero plane displacement height;
- narrow vegetation systems may cause the EC (or Bowen ratio) system to sample ET from areas upwind of the vegetation area (for example, with a narrow riparian system, from pasture, desert or agriculture that is beyond the riparian system);
- requires consistently nearly horizontal flowlines;
- change in vertical direction of flowlines during an averaging period causes large errors (portions of the streamline can be mis-interpreted as \( w' \) (relative to mean vertical wind speed) so that \( w'^2 \) or \( w'^q \) are overestimated) (Lee et al., 2010a). This occurs when:
  - sensors are too close to roughness elements such as tree limbs;
  - the ‘surface’ of vegetation is irregular, for example, a convex shaped riparian system where the shape and vertical components of streamlines change with respect to wind direction;
  - vertical components of streamlines can easily change with wind direction as wind direction changes;

- difficult to know impacts of wind direction changes on any change in vertical components of flow lines;
- problems in eddy formation if sensors are too close to the surface or are too close to individual roughness elements;
- results may be questionable when used for understory measurements because of a changing horizontal plane orientation during an averaging period;
- mounting of net radiometers and location of soil heat flux plates is difficult for representative sampling;
- mounting arm and other nearby equipment, including the hygrometer, can impede and distort flowlines and eddy shapes for some wind directions;
- tower or scaffolding may bend and distort flowlines;
- instrumentation is relatively fragile and expensive;
- the method requires personnel who are well-trained in electronics, turbulent theory, and biophysics.

\[ R_G = \frac{R_n - \lambda E - H}{G} \]
3.6. Fetch requirements for boundary layer measurements

A general ‘rule of thumb’ regarding siting of equipment is that upwind fetch for boundary layer instrumentation should be on the order of 100 m for each m of $z_2$ above the top of the zero plane displacement height of the surface (Brutsaert, 1982). This is done to ensure that an internal equilibrium boundary layer (EBL) has been established over the $z_2$ distance that is representative of the surface energy exchange being measured. This means that there should be 100 m of upwind distance of vegetation for every meter above the ground up to the uppermost temperature and/or humidity sensor. Monteith and Unsworth (1990) also recommended that 50–100 m per meter is often adequate for micrometeorological measurement of sensible and latent heat flux. Brutsaert (1982) provided theoretical considerations of boundary layer development that can be used to estimate minimum fetch requirements as a function of surface roughness (or vegetation height). Table 1 shows estimates from Brutsaert’s equation in terms of minimum required fetch for a variety of vegetation types (Brutsaert, 1982). Brutsaert’s equation assumes near-neutral stability of the EBL. Generally, as $H$ and instability of the EBL increase, the required fetch length decreases. More modern ‘footprint’ models are routinely employed (Horst and Weil, 1992; Hsieh et al., 1997, 2000; Leclerc et al., 1997; Schmid, 2002) to define the two-dimensional spatial extent of the source areas for heat and vapor that contribute to the particular $H$ or $\lambda E$ measurement.

The values for minimum fetch in Table 1 tend to follow the 100:1 rule for the moderately smooth vegetated surface of grass. Fetch requirements are greater than 100:1 for smoother surfaces and less for rougher surfaces. As an example of application, for a Bowen ratio installation having $z_2$ height at 2 m above the ground surface and 0.1 m tall vegetation, the recommended minimum fetch requirement under neutral conditions is approximately 190 m. For 10 m tall trees, even with the upper Bowen arm as close as 2 m above the tree height, $h$, a minimum of 320 m of upwind fetch having the same type and height as the trees is recommended based on an estimated zero plane displacement height of 0.6$h$. The fetch requirement, which increases as vegetation height increases, makes it difficult to use the BREC, eddy covariance and other boundary layer sampling systems to measure $\lambda E$ from narrow or isolated stands of vegetation.

Additional guidelines on fetch requirements for BREC measurement were given by Heilman et al. (1989) and a number of publications describe a variety of models for estimating footprints for both scalar and flux measurement. The ‘footprint’ of a flux or scalar measurement represents the upwind surface area that is statistically responsible for the conditioning of the measurement (Hsieh et al., 2000; Foken and Leclerc, 2004). Schmid (2002) presented a good review of various models, their mechanics and limitations. An expression for estimating the fraction of $\lambda E$ sensed at a specific instrument height that is generated from a specific distance of upwind fetch was presented by Gash (1986), Schuepp et al. (1990), and Shuttleworth (1992) and repeated by Allen et al. (1996), Horst and Weil (1992), Hsieh et al. (1997, 2000), Leclerc et al. (1997), and Schmid (2002) applied Lagrangian stochastic and large eddy simulation (LES) strategies along with Gaussian or non-Gaussian diffusion assumptions to estimate three-dimensional distribution of point source or line source fluxes. Gockede et al. (2008) applied a flux data quality approach to analytical footprint modeling, using a sophisticated Lagrangian stochastic footprint algorithm, as part of site evaluation for a large number of locations. The complex Lagrangian models have been applied under various stability conditions to demonstrate sensitivities of flux and scalar measurements to footprint conditions and to discontinuities in fetch at various distances from the sensors.

### Table 1

<table>
<thead>
<tr>
<th>Height and type of surface cover</th>
<th>Height of the eddy covariance or upper Bowen ratio measurement above the ground</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$z = 1\text{ m}$</td>
</tr>
<tr>
<td>Water ($d = 0$, $z_{om} = 0.0001\text{ m}$)</td>
<td>180</td>
</tr>
<tr>
<td>0.12 m Grass</td>
<td>80</td>
</tr>
<tr>
<td>0.5 m Alfalfa</td>
<td>45</td>
</tr>
<tr>
<td>1.5 m Cattails</td>
<td>–</td>
</tr>
<tr>
<td>10 m dense trees ($d = 6\text{ m}$)</td>
<td>–</td>
</tr>
</tbody>
</table>

Source of data: Allen et al. (1996).

3.7. Scintillometers

A scintillometer is an optical device that measures small fluctuations in the refractive index of air caused by temperature, humidity, and pressure induced variations in density. A system for measurement of sensible heat flux consists of an optical transmitter and a receiver at the ends of an atmospheric propagation path. The receiver detects and evaluates the intensity fluctuations of the transmitted signal, called scintillations. The magnitude of the fluctuations in the refractive index is usually measured in terms of the structure parameter $C_n^2$, which is the spectral amplitude of refractive index fluctuations in the inertial subrange of turbulence. Scintillometers measure sensible heat flux, $H$, by relating the structure parameter to a temperature structure parameter and the Monin–Obukhov stability parameters. Detailed descriptions are available from Meijninger and De Bruin (2000), Meijninger et al. (2002), Hartogensis et al. (2003) and De Bruin (2008).

The distinct advantage of scintillometry is the ability to derive sensible heat flux that is integrated over a long transect, up to several km in length. This integration is especially useful for measuring $H$ over complex natural vegetation where spatial variation can have scales of hundreds of meters. Scintillometer measurements do require assumptions related to Monin–Obukhov stability functions, the estimation of mean shear stress (or friction velocity) and some empirical corrections related to the frequency spectrum. To obtain ET, measurement of $R_{H}$ and $G$ are needed, which can be problematic and biased in heterogeneous and sparse vegetation systems, as discussed in the Bowen ratio section. Insufficient representativeness of $R_{H}$ and $G$ measurement over a transect can greatly reduce the value of the scintillometer-based integration of $H$. $R_{H}$ measurements may need augmentation by satellite-based images of albedo and surface temperature. Surface temperature can vary by as much as 30 K from wet to dry surfaces along a scintillometer transect, thereby varying $R_{H}$ by as much as 100 W m$^{-2}$ through differences in emitted long wave from the surface. Soil heat flux densities can vary in the same manner along a mixed transect.

Other requirements of scintillometer-based derivation of $H$ is the need for accurate measurement or estimation of friction velocity, $u^*$, which may require the deployment of sonic anemometers along the transect and uniform enough terrain along the transect so that all assumptions in Monin–Obukhov similarity theory governing ‘shape’ and distribution of turbulent structures are valid. Otherwise, $u^*$ can be estimated using horizontal wind speed and surface roughness, with some reduced accuracy in the $H$ product.

Advantages of scintillometers include:

- integration of sensible heat flux over large distances, with weighting toward the center of the transect (Meijninger et al., 2002);
- relatively simple operation and maintenance;
- apparent good consistency in application (Kleissl et al., 2008).
Disadvantages of scintillometers include:

- measurement of the magnitude of $H$ only (the direction of $H$ must be determined via other means);
- ET is derived by energy balance residual ($R_n - G - H$), requiring spatial representativeness of the $R_n$ and $G$ measurements, and where any biases in $R_n$ and $G$ transfer into the ET estimate;
- can require some post-processing correction (De Bruin, 2008);
- need measurement or estimation of friction velocity if below the free convective layer;
- relatively expensive.

Fig. 1 shows comparisons of 30-min estimates of sensible heat flux, $H$, derived from two adjacent eddy covariance systems and from a large aperture scintillometer ($H_{LAS}$) positioned over sagebrush in southern Idaho over a five day period in April, 2010. Also shown is ET retrieved directly from the EC system and ET calculated as $E = R_n - G - H$. The two EC systems were made by different manufacturers and were located 1 m apart. Agreement between $H$ from the two EC systems is good, however, differences averaged 10–30 W m$^{-2}$ or about 5–15%, after application of standard ‘corrections’ noted in the EC section. Agreement between the scintillometer-derived $H$ and EC-derived $H$ was also good, with some deviation during nighttime when $H$ was negative. $E$ from the EC compared relatively well with $E$ from energy balance, with differences ranging from 0 to about 20%. The April period had relatively high amounts of evaporation from moist soil, especially following a rain on April 24.

3.8. Sap flow methods

Sap flow methods introduce a source of low grade heat into the trunk or branch of a plant, and measure the flow of water in the xylem by either the velocity of a heat pulse carried away from the heat source in the transpiration stream or by the dissipation of heat energy in the stem due to convection in the transpiration stream. These methods do not measure the direct evaporation component of ET, but only the plant transpiration component. This can be an advantage if the goal is to measure plant water use, but a disadvantage if the goal is to project total evaporation over a heterogeneous surface. Three main methods are employed today: the heat pulse-sap velocity method (Green and Clothier, 1988; Green et al., 2003); the Granier heat dissipation method (Granier, 1985, 1987); and the tissue-heat balance method (Sakuratani, 1981; Valancogne and Nasr, 1989, 1993; Kjelgaard et al., 1997). Swanson (1994) suggested that no one set of theory and instrumentation is applicable to all sizes or species of trees. As outlined below, each method has important limitations that need to be considered when interpreting ET rates based on sap flow methods.

Heat pulse velocity measurements are appropriate for determining transpiration in forests and orchards, since the sensors can be applied to any woody stem larger than about 40 mm in diameter, are inexpensive, easy to install, and suited to automated data collection (Green et al., 2003). Commercial heat pulse units are available (Steppe et al., 2010). Heat pulse methods typically introduce a 1–2 s pulse of heat into the conducting layer of stem tissue (the sapwood layer) and measure the time lag between introduction of the heat pulse and its reception at a heat-sensing probe (thermistor) embedded in the sapwood layer downstream from the heat source. Velocity of the heat pulse is due to conduction as well as convection, and the measured velocity must be corrected for conduction velocity to calculate sap flow velocity (Green et al., 2003). The most commonly used correction method is the compensation heat pulse method. It uses two temperature-sensing probes, one installed (typically) 5 mm upstream from the heating needle, and the other placed 10 mm downstream from the needle (Steppe et al., 2010). Following introduction of a heat pulse, the closer, upstream sensor registers the heat pulse first due to conduction, while the downstream sensor quickly reaches thermal equilibrium with the upstream sensor due to conduction plus convection; the
time required for the sensors to reach thermal equilibrium is equal to the time required for the heat pulse to move convectively to a midpoint between the two sensors (i.e., 2.5 mm downstream from the heater). The less-commonly used T-max method uses a single temperature sensor upstream from the heater, and conduction velocity is calculated from velocities measured at night when convection is assumed to be zero (Steppe et al., 2010). Once convective velocity is determined, sap flux density (SFD, mm\(^3\) mm\(^{-2}\) h\(^{-1}\)) is calculated by multiplying velocity times the volume of water in the conducting tissue, estimated from the depth of the conducting layer and the proportion of the conducting layer that is xylem sap (as opposed to woody matrix material and gas). This requires a knowledge of the wood anatomy of the species being measured (to determine sapwood area) and measurement of the moisture content of the sapwood layer by oven drying freshly excised sections.

Velocity data must be corrected for tissue wound effects of conducting velocities due to insertion of the heating needle, for differences in the radial distribution of conducting tissue around the stem, and for differences in conduction as a function of depth within the sapwood layer. These factors can be evaluated by placing multiple sensors at different depths within the sapwood layer. Green et al. (2003) suggested that heat pulse methods can provide accurate measurements of sap flow in plant stems provided a reliable, independent calibration procedure is used to relate the measured heat pulse velocity to the actual sap flow.

The Granier heat dissipation method (Granier, 1985, 1987) is also suited to transpiration measurements in trees, since it also involves inserting a heating needle into the sapwood layer of a large trunk or branch, similar to the heat pulse method. In the heat dissipation method, heat is applied at a constant rate rather than as a pulse. In the most commonly employed commercial unit, two 30-mm-long stainless steel needles are inserted into the sapwood layer, one 40 mm above the other (Steppe et al., 2010). Needle lengths and distances between needles can be varied according to the species under measurement (Hultine et al., 2010). The downstream needle contains a line heater that introduces a source of heat into the stem and a thermocouple junction referenced to the upstream needle, and the difference in temperature (\(\Delta T\)) between the two needles is measured. The temperature difference between the needles is independent on sap flux density (SFD); the temperature difference decreases with increasing sap flux because the downstream needle is “cooled” by the dissipation of heat upstream from the downstream needle in the transpiration stream. Based on measurements of three different tree types, Granier (1985) developed an empirical formula relating SFD to \(\Delta T\):

\[
SFD = 0.0119K^{1.231}
\]

with \(K\) defined as:

\[
K = \frac{(\Delta T_o - \Delta T)}{\Delta T}
\]

where \(\Delta T_o\) is \(\Delta T\) in the absence of convective flow (the maximum temperature difference between the two needles measured in the predawn hours). Although Granier (1985) suggested that this formula could be applied across different tree species, the empirical calibration has little physical basis (Steppe et al., 2010). Smith and Allen (1996) concluded that calibration coefficients are needed for each tree species. The same types of caveats based on wood anatomy that apply to the heat pulse methods apply also to the thermal dissipation method because the empirical constants in the Granier formula assume that the line heater is in contact with the sapwood along its entire length and that sap flow is radially uniform around the trunk.

The tissue heat balance method is the most direct sap flux method with the fewest assumptions (Kjølgaard et al., 1997). It involves wrapping a heating wire around an entire branch segment and measuring conductive heat losses upstream \(q_u\), downstream \(q_d\), and radially \(q_r\) away from the heat source by thermocouples referenced to the temperature at the heating wire. The \(\Delta T\) values are used to calculate conductive heat losses, and convective heat loss (transpiration) is calculated as total heat input minus conductive heat losses:

\[
q_r = P_w - q_u - q_d - q_r.
\]

where \(q_r\) is convective heat loss, \(P_w\) is power (heat flux) applied to the heater, and \(q_u\), \(q_d\) and \(q_r\) are the conductive heat fluxes away from the heat source, and all terms are in units of power (W). Since the energy balance is conducted on the entire stem segment, no knowledge of the xylem anatomy or area of conducting sapwood is needed. Unfortunately, this is the most difficult method to implement and can only be applied to branches under 50 mm in diameter, because the heat source must be able to uniformly heat the entire segment of branch around which it is wrapped to produce valid results (Grime et al., 1995). In practice, temperatures are measured in the woody tissue at the heat source, and typically 5 mm upstream and downstream from the edge of the heating wire, and on the outside surface of the foam insulating layer wrapped around the stem outside of the heating wire (Kjølgaard et al., 1997). Conductive heat losses upstream and downstream in the stem are calculated by \(\Delta T\) values referenced to the temperature at the heating source, multiplied by the cross sectional area of the stem and the known thermal conductivity of wood (assumed to be constant). Thermal conductivity through the foam layer (sheath conductance) must be determined experimentally for each sensor as the contact between the foam and the wood and therefore heat conductance is irregular. Sheath conductance is derived from the radial \(\Delta T\) determined at night under assumed zero flow conditions, when convective flow is zero, from the energy balance equation. Whereas heat pulse and heat dissipation sensors placed in a tree trunk can measure transpiration through the entire tree, numerous tissue heat balance sensors placed on individual branches would be needed to sample sap flux on a single tree, and a scaling method would then be needed to extrapolate from the branches to the whole tree (e.g., Nagler et al., 2009).

The heat balance method is direct and requires no knowledge of the xylem cross sectional area or moisture content. It does require measurement of radial heat loss under zero flow conditions which is complicated if there is nocturnal transpiration; in this case, sheath conductance can be estimated at the end of the experiment by cutting the branches above the sensors and continuing to measure heat losses for several hours (Nagler et al., 2009). However, numerous studies have shown that the heat pulse and thermal dissipation methods must be calibrated for each new application, ideally through the use of weighing lysimeters or by pulling water through stem sections of the plants of interest while measuring velocity and heat dissipation parameters with sap flux sensors (Green et al., 2003; Steppe et al., 2010; Hultine et al., 2010). Steppe et al. (2010) compared sap flux measurements with gravimetric determinations of water flow through freshly cut stem segments of the Fagus grandifolia trees, and found that the heat pulse method underestimated actual sap flow by 35%, while the thermal dissipation method underestimated flow by 60%. There were no simple corrections that could be applied to the data, since each method was subject to multiple sources of error that compounded to reduce the accuracy of the methods. However, when empirical calibration factors were calculated from lab results and applied in the field, the methods gave similar values for sap flux, indicating that if properly calibrated, the sap flux methods can produce relatively accurate estimates of plant transpiration.

Steinberg et al. (1990) described a commercial sap flow system that provided daily transpiration measurements within 5% of directly weighed \(T\) measurements. Shackel et al. (1992), on
the other hand, found substantial departure between lysimeter measurements of ET from a peach tree and a heat balance technique for measuring sap flow. Lundblad et al. (2001) used a tissue heat-balance method and the heat-dissipation method in Scots pine and Norway spruce. The latter system measured up to 50% lower sap flow than the tissue heat-balance system under high sap flow rates. They found that natural temperature gradients can cause large errors in measurements made by the heat-dissipation method. Kjelgaard et al. (1997) reported differences between variable heat input and constant heat input systems and improved accuracy when readings were integrated over 24-h periods. Ewers and Oren (2000) found sap flow velocity in xylem 20–40 mm from the cambium to be 50 and 40% of sap flow velocity in the outer 20-mm band of xylem in slow- and fast-growing trees, respectively. Wilson et al. (2001) used commercially available thermal dissipation probes operated on the constant power principle developed by Granier (1987) in a forested Tennessee watershed, where four red maple, four loblolly pine, two chestnut oak, one red oak and two yellow-poplar were instrumented. Water use by common understory saplings (one red maple, one dogwood and one beech) was also measured. They found total seasonal ET estimated by the sap flow method to be 50% of that determined by both eddy covariance and inflow-outflow watershed balance.

F.G. Allen et al. / Agricultural Water Management 98 (2011) 899–920

The cambium to be 50 and 40% of sap flow velocity in the outer 20-mm band of xylem in slow- and fast-growing trees, respectively. Wilson et al. (2001) used commercially available thermal dissipation probes operated on the constant power principle developed by Granier (1987) in a forested Tennessee watershed, where four red maple, four loblolly pine, two chestnut oak, one red oak and two yellow-poplar were instrumented. Water use by common understory saplings (one red maple, one dogwood and one beech) was also measured. They found total seasonal ET estimated by the sap flow method to be 50% of that determined by both eddy covariance and inflow-outflow watershed balance.

Advantages of sap flow measurement are:

- Direct measurement of transpiration, allowing calculation of stomatal and canopy conductance for individual plants;
- Coupled with micrometeorological systems (i.e., eddy covariance, etc.), sap flow measurement can help separate T of overstory vegetation from evaporation from soil or T from understory.

Disadvantages of sap flow measurement are:

- Probe spacings and stem geometry are the most significant sources of error;
- Various wound responses to probe implantation may cause heat ratios to vary over time;
- Implanting sensors can cause mechanical damage and interrupt flow by occlusion or blocking of the plant’s vascular tissues;
- Area of conductive tissue must be estimated accurately;
- To obtain estimates of ET:
  - T of individual branches of plant(s) must be scaled up, which introduces uncertainty due to variation in exposure to solar radiation and aerodynamic turbulence;
  - Evaporation from soil and any understory vegetation must be estimated or measured (these can have large uncertainties);
  - Tree(s) monitored need to be representative of the area;
  - Relationships and calibrations may change with soil moisture, LAI, age, and disease.

3.9. Remote sensing energy balance

Techniques using satellite imagery have been developed since about 1990 to estimate E and ET from large areas using energy balance (Bastiaanssen et al., 1998a,b, 2005; Kustas and Norman, 1999; Moran, 2000; Kustas et al., 2003; Allen et al., 2007a,b; Irmak et al., 2011). The emerging technology of energy balance by satellite shows substantial promise for application over large areas and over a wide range of vegetation types and water availability. The approach has been used to quantify and illustrate population variance in ET from the same vegetation type and to refine Kc or ET,F curves (Tasumi et al., 2005a; Anderson et al., 1997, 2005; Kustas et al., 2003; Kustas and Norman, 1999; Li et al., 2005; Norman et al., 1995, 2000, 2003; Allen et al., 2007a,b; Tasumi and Allen, 2007; Singh and Irmak, 2009). Remotely sensed energy balance techniques are useful for identifying areas experiencing water stress

Angles are needed to account for leaf and branch angle effects on LAI. Unfortunately, individual tree canopies, or irregular canopies with gaps between plants, seriously violate the assumption of a uniform canopy and there is no simple way to correct the instrument readings except by combining them with leaf harvest methods, a step which is often omitted. Nagler et al. (2004) found that Licor LAI 2000 measurements of LAI on individual cottonwood and willow canopies were low by a factor of three when compared to leaf harvesting methods.

Scaling ET from individual limbs to full plants and from full plants to large areas is uncertain due to differences among tree structure, radiation interception per tree, and water availability. Shaded limbs tend to have lower rates of transpiration than sunlit limbs, and limbs lower in a canopy tend to have less aerodynamic exchange of energy and vapor. Besides scaling of transpiration from specific species of trees, additional uncertainty in a real ET is caused by stands of multiple species, presence of understory vegetation, and evaporation from bare soil. Mackay et al. (2002) used a two-source evaporation model with sap flow data to produce total ET. It seems clear that large uncertainties can exist with determining accurate transpiration measurements with various sap flow methods and determining quantitative ET estimates.

Advantages of sap flow measurement are:

- Direct measurement of transpiration, allowing calculation of stomatal and canopy conductance for individual plants;
- Coupled with micrometeorological systems (i.e., eddy covariance, etc.), sap flow measurement can help separate T of overstory vegetation from evaporation from soil or T from understory.

Disadvantages of sap flow measurement are:

- Probe spacings and stem geometry are the most significant sources of error;
- Various wound responses to probe implantation may cause heat ratios to vary over time;
- Implanting sensors can cause mechanical damage and interrupt flow by occlusion or blocking of the plant’s vascular tissues;
- Area of conductive tissue must be estimated accurately;
- To obtain estimates of ET:
  - T of individual branches of plant(s) must be scaled up, which introduces uncertainty due to variation in exposure to solar radiation and aerodynamic turbulence;
  - Evaporation from soil and any understory vegetation must be estimated or measured (these can have large uncertainties);
  - Tree(s) monitored need to be representative of the area;
  - Relationships and calibrations may change with soil moisture, LAI, age, and disease.

3.9. Remote sensing energy balance

Techniques using satellite imagery have been developed since about 1990 to estimate E and ET from large areas using energy balance (Bastiaanssen et al., 1998a,b, 2005; Kustas and Norman, 1999; Moran, 2000; Kustas et al., 2003; Allen et al., 2007a,b; Irmak et al., 2011). The emerging technology of energy balance by satellite shows substantial promise for application over large areas and over a wide range of vegetation types and water availability. The approach has been used to quantify and illustrate population variance in ET from the same vegetation type and to refine Kc or ET,F curves (Tasumi et al., 2005a; Anderson et al., 1997, 2005; Kustas et al., 2003; Kustas and Norman, 1999; Li et al., 2005; Norman et al., 1995, 2000, 2003; Allen et al., 2007a,b; Tasumi and Allen, 2007; Singh and Irmak, 2009). Remotely sensed energy balance techniques are useful for identifying areas experiencing water stress
and corresponding reductions in ET and to populate hydrologic models (Irmak and Kamble, 2009; Kamble and Irmak, 2009). Users of this information must bear in mind that satellite-based ET data are simply retrievals, or best estimates, of an aerodynamic and radiative process, as viewed from space, and cannot be considered to be “measurements.” Nevertheless, remotely sensed energy balance is discussed here because of its increasing use to estimate $K_r$ and ET over large areas. Estimation by remotely sensed energy balance should be expected to adhere to the same limitations and physics as other measurement methods.

Some ‘operational’ satellite-based energy balance models such as SEBAL (Bastiaanssen et al., 1998a,b, 2005) and METRIC (Allen et al., 2007a,b) employ an internal calibration technique referred to as CIMEC (‘calibration using inverse modeling at extreme conditions’) (Allen et al., 2008). The CIMEC technique involves the inverse calibration of the energy balance process via the sensible heat flux, $H$, computed by specifying ET, and thus the energy balance equation, at two extreme conditions (dry and wet) in the satellite image. ET is estimated at the two conditions based on knowledge of available energy and surface conditions, usually with ties to ground-based weather data. The outcome of the CIMEC calibration is to imbed all estimation biases in $K_r$, $G$ and other intermediate components that are endemic to nearly all satellite-based calculations into the estimate and calibration of $H$. These biases are in turn removed when ET is computed for the millions of pixels in a typical image as ET = $\Delta E / \Delta (\rho \omega)$, where the calculation of $H$ is a function of surface temperature $\rho$ and $w$ is density of water. Other more regional-scale models such as the ALEXI model (Anderson et al., 2005) use inversion based on radiosonde profilings of $T$ and $q$ over time to estimate large-scale heat flux and evaporation.

SEBAL, METRIC and similar models are ‘working’, operational models whose objective is to produce consistent and dependable maps of ET over areas of about 20,000 km$^2$ (the size of a Landsat scene). The $H$ estimates, because they tend to contain artifacts of biases in estimates for $K_r$ and $G$ that are entrained during calibration, are generally not considered to be absolutely accurate. However, when the model objective is the production of spatial estimates for ET rather than $H$, bias-entraining calibration techniques such as CIMEC can produce relatively accurate estimates for ET. Ordinarily, surface energy balance models that do not apply endpoint calibration will be impacted by the various biases that can plague satellite-based component estimation.

Advantages of remote sensing-based energy balance (RSEB) models:

- the energy balance yields actual ET;
- the process covers large areas, enabling sampling and integration over diverse areas;
- RSEB procedures are generally more economic than point measurements;
- RSEB procedures can be used to extend or extrapolate high intensity point measurements of ET to large areas;
- products can have high spatial resolution (30 m for satellite, 2–5 m for aerial);
- RSEB procedures are valuable for determining spatial variation in ET for highly variable systems such as riparian or forest.

Disadvantages of remote sensing-based energy balance models:

- time gaps exist between estimates of ET for many satellite systems, especially those having high spatial resolution, where images are obtained only periodically for a specific location, for example, every 16 days for a Landsat satellite; therefore, effects of evaporation from precipitation events occurring in between satellite overpasses may be missed, or processing of ‘wet’ images from recent precipitation events may bias seasonal estimates (Allen et al., 2007a);
- aerial data collection can be expensive;
- satellite pixels over narrow vegetation systems such as riparian systems and small agricultural fields may overlay broad mixtures of vegetation types and densities so that surface temperature signals are mixed and the ET retrievals are difficult to interpret;
- uncertainty in estimating aerodynamic components and surface temperature retrievals may require ‘inversion’ techniques to calibrate;
- most remote-sensing energy balance processes assume 1-D aerodynamics that may not hold true for narrow, tall systems such as riparian:

  - aerodynamic exchanges in narrow vegetation systems may be three-dimensional and therefore flowlines are poorly behaved;
  - Klaassen et al. (2002) found that horizontal penetration of flowlines and heat transfer into the leading edge of a forest canopy increased available energy by 15% over first 400 m of canopy;
  - Klaassen and Sogachev (2006) concluded that flux measurements over riparian systems should be corrected for impacts of horizontal variations in turbulence downwind of the edge;
  - the problem of narrowness of stands and edge effects impacts nearly all measurement systems that use aerodynamics or energy balance (EC, scintillometry, remote-sensing-based EB), even sap flow due to change in transpiration with distance into the system;
  - satellite view angles can impact reflectance (albedo) and surface temperature estimates:

    - the Landsat satellites have a view angle that is nearly nadir (directly overhead);
    - the MODIS satellites have a large scan angle, with the view angle varying from $-55$ to $+55^\circ$; this viewing angle variation impacts reflectance measurements and requires correction;
    - the satellite (and aerial) measurement of reflectance is ‘bidirectional’ whereas—the reflectance needed in energy balance (for $R_m$) is directional-hemispherical;
    - bidirectional reflectance may be lower than directional-hemispherical reflectance for tall canopies containing shadows, especially at lower sun angles and for nadir looking satellites;
    - potential bias in albedo and surface temperature for tall canopies may cause overstatement of ET by 5–10% (this impact needs further investigation);

- potential biases exist in retrieved albedo and surface temperature ($T_s$) from satellite and airborne systems (thus CIMEC or other calibration is needed).

3.10. Satellite-based ET using vegetation indices

Satellite-based or ground-based energy balance methods generally have extensive time investment and require learned skill sets. The energy balance products can be used, however, to calibrate more simple methods that utilize general vegetation indices (VI) to estimate crop coefficients (i.e., $K_c$ or ET,F) (Tasumi et al., 2005a; Tasumi and Allen, 2007; Singh and Irmak, 2009). The estimate of $K_c$ from VI is possible because of the generally close correspondence between vegetation amount and transpiration, where, as vegetation cover increases, leaf area increases, and transpiration increases (reviewed in Glenn et al., 2007). Challenges with VI-based methods are estimating evaporation from bare soil following precipitation events and estimating reduced ET associated with soil-water shortage, because these processes are not adequately reflected in the VI. A common VI is the Normalized Difference Vegetation Index (NDVI) that is estimated from
two shortwave bands typically measured by satellites: the red band (~0.6–0.7 μm) and the near infrared band (~0.7–1.3 μm). A linear relationship between the NDVI and the crop coefficient $K_c$ was introduced by Heilman et al. (1982) and theoretically established by Choudhury et al. (1994). The resulting equation is as follows:

$$K_{cb} = 1.25 \text{NDVI} + 0.2 \quad (11)$$

where $K_{cb}$ represents the grass-based crop coefficient. Tasumi et al. (2005b) and Tasumi and Allen (2007) found a similar relationship for eight major irrigated crops in Idaho, as did Singh and Irmak (2009) for irrigated corn and soybeans in Nebraska, which, when expressed for the alfalfa reference is approximately:

$$K_{cr} = ET_r F = 1.25 \text{NDVI} \quad (12)$$

Generally, NDVI values from different satellites show close correlation (Calera-Belmonte et al., 2005). However, some differences occur due to differences in band widths. The degree and type of atmospheric correction of the image can also have an impact.

It is important to establish, with $K_c$ vs. NDVI or other VI relationships, whether the relationships are to represent the average ET that includes averaged amounts of evaporation from the soil surface caused by precipitation and irrigation or are to represent a ‘basal’ $K_{cb}$ condition where the VI-based relationship is established to represent conditions where the soil surface is dry enough to reduce evaporation from the soil surface to relatively low levels, when compared to transpiration, but transpiration still occurs (Allen et al., 2005). The basal $K_{cb}$ vs. NDVI relationship is more consistent because transpiration has a much closer association with vegetation amount than does evaporation from soil. When $K_{cb}$ vs. NDVI relationship is used, then estimates for soil evaporation are determined separately and added to produce total evaporation (Burnett et al., 2008). Allen et al. (2011b) showed that NDVI was a better basis for the $K_c$ vs. VI relationship than was the Soil Adjusted Vegetation Index (SAVI) because NDVI tends to reach a maximum value at about the same time as does $K_c$, which is at about LAI of 3.0, whereas SAVI tends to continue to increase with increasing LAI beyond 3.0. Allen et al. (2011b) also showed that the VI calculation can be based on ‘top of atmosphere reflectance’ or on ‘at-surface reflectance’ with little reduction in estimation accuracy.

Advantages of VI-based crop coefficient ($K_c$ and $K_{cb}$) estimation are as follows:

- quick analyses can be made by mid-level technicians;
- large areas can be covered;
- relationships can be calibrated using satellite-based energy balance;
- spatial resolution can be high, especially if aerial imagery is used.

Disadvantages of VI-based $K_c$ and $K_{cb}$ estimation are as follows:

- relationships may vary with type of vegetation;
  - stomatal control (and thus $K_c$ or $K_{cb}$ vs. VI relationships) can vary among types of vegetation. Therefore, single $K_c$ and $K_{cb}$ vs. VI or ET vs. greenness indices can vary;
  - trees, when short of water can exhibit more stomatal control than agricultural crops and therefore cause a shift in the $K_c$ vs. VI relationship;
- relationships tend to overestimate ET under conditions of acute water shortage;
- estimation of the evaporation (from soil) component is less certain than the transpiration component because of the lack of a direct relationship with vegetation amount;
- quality estimates of reference ET are required to transform $K_c$ into ET where the reference ET calculation requires quality weather data;
- VI’s may not identify or quantify multi-storied canopies and their effects on total ET, especially for more dense vegetation.

Nagler et al. (2009) found a reasonable correspondence between saltcedar ET estimated by sap flow sensors and satellite estimates based on a VI at the Cibola National Wildlife Refuge on the lower Colorado River. However, at any given measurement station, ground measurements showed high variability of ET due in part to differences in stomatal conductance due to stress effects, which were not captured in the remote sensing estimates.

4. Anticipated error in evapotranspiration measurements

All evapotranspiration measurements contain error. Some error is caused by systematic error associated with sensor calibration bias, improper sensor function, improper sensor operation, improper sensor placement, inaccurate sensor recording, inadequate or incorrect model associated with data interpretation or processing, unrepresentative vegetation characteristics, improper data reduction procedures, and improper use of timestep integration. In addition, measurement error includes random error components that can be associated with resolution of sensor readings, electronic noise, mechanically induced noise, thermal responses of sensors, vegetation and soil water management, as well as other random error specific to the type of measurement system. Random error components are typically dual-signed and distributed about a mean of 0. Repeated sampling over time can reduce random error, often in proportion to the square root of the number of samples. Systematic error, however, does not necessarily reduce with repeated sampling. Systematic error that is associated with a specific component of a measurement process may be additive to systematic error of another component, or may even multiply the other’s error, or may partially mitigate the other error by partial compensation in a different direction.

Because there are wide ranges in types of error and causes for error in ET measurements, it is difficult to assign estimates for average error associated with any particular type of measurement system. Estimation of expected ranges in error is also made difficult by the close association of error with human induced error including error in datalogger and data reduction programming, error in equipment assembly, error in equipment and sensor maintenance, error in managing the environment of the measurements, and error in sensor placement. Human-induced error can be even larger than other systematic error, and is often unavoidable, but is expected to reduce with operator experience, education and training. Substantial experience and understanding of the measurement process can partially offset some non-human–associated error components through proactive intervention and adjustment by cognizant operators.

Some types of ET measurement systems are more prone to the opportunity for error, both electro-mechanically and via human factors. Some types of ET measurement systems may be more sensitive to impact of error on measurement accuracy.

Table 2 gives estimated ranges of error in ET that can be expected from the various measurement systems described. These estimates are based on general discussions on error in eddy covariance systems by Meyers and Baldocchi (2005) and Burba and Anderson (2008), in Bowen ratio systems by Fritschen and Fritschen (2005), comparative discussions on error by Foken (2008b) and Lee et al.
Accuracy of ET measurement requires well-calibrated and well-maintained systems, and, in most cases, a foundational knowledge of the underlying physics of turbulence and heat and radiation transfer that govern the particular measurement. Knowledge of underlying physics is necessary to reduce the impacts of oversights of important biasing factors. In addition, an attitude of thoughtful approach to measurement and critical assessment of data are essential. Substantial measurement biases or incorrect data extrapolations can easily occur in ET data sets, many of which may not be recognized or identified and rectified prior to refereed publication. Deployment of systems that sample within the equilibrium boundary layer (eddy covariance, Bowen ratio and scintillometry) must adhere to fetch requirements and minimum equipment heights. Energy balance methods such as Bowen ratio and scintillometry that estimate ET as a residual must obtain accurate and representative measurements of net radiation and soil heat flux density to produce information on ET. For most vegetation systems, this requires multiple instrument locations within the measurement area to compensate for spatial non-uniformity of vegetation. Appropriate “corrections” must be made to eddy covariance measurements. Sap flow systems must include sufficient numbers of trees to reduce statistical error and to sample a representative population of trees. While remote sensing systems do not ‘measure’ ET, but rather deduce it via energy balance or vegetation indices, these methods are quite powerful in spatial coverage and quantification of spatial variation in ET, especially those based on thermally driven energy balance. All systems must be combined with rigorous quality assessment and quality control procedures. Precautions need to be exercised to produce accurate and defensible ET data and to recognize quality data.

Acknowledgements

The authors express their appreciation to the Technical Committee on Evapotranspiration in Irrigation and Hydrology of the American Society of Civil Engineers—Environmental and Water Resources Institute for encouraging quality assurance and integrity in evapotranspiration and weather measurement and for functioning as a sounding board for this article; to Dr. Edward Glenn of the University of Arizona and Dr. Teresa do Paço of the Technical University of Lisbon for review and contribution to the section on sap flow measurement; and to Dr. Jim Oster, editor of Agricultural Water Management, for encouraging the publication of this article.

References


Table 2

<table>
<thead>
<tr>
<th>Method</th>
<th>Typical error, %</th>
<th>Error for an experienced expert, trained and steeped in the physics of the process, %</th>
<th>Error for a novice or a person working outside their specialty area, %</th>
<th>Additional error caused by physical or equipment malfunction, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lysimeter</td>
<td>5–15</td>
<td>5</td>
<td>20–40</td>
<td>5–40</td>
</tr>
<tr>
<td>Soil water balance</td>
<td>10–30</td>
<td>10</td>
<td>20–70</td>
<td>10–40</td>
</tr>
<tr>
<td>Bowen ratio</td>
<td>10–20</td>
<td>10</td>
<td>20–50</td>
<td>5–40</td>
</tr>
<tr>
<td>Eddy covariance</td>
<td>15–30</td>
<td>10–15</td>
<td>30–50</td>
<td>10–40</td>
</tr>
<tr>
<td>Remote sensing energy balance</td>
<td>10–20</td>
<td>5–15</td>
<td>30–40</td>
<td>5–10</td>
</tr>
<tr>
<td>Remote sensing using vegetation indices</td>
<td>15–40</td>
<td>10–30</td>
<td>20–40</td>
<td>5–10</td>
</tr>
<tr>
<td>Sap flow</td>
<td>15–50</td>
<td>10–40</td>
<td>40–200</td>
<td>20–100</td>
</tr>
<tr>
<td>Scintillometersa</td>
<td>10–35</td>
<td>10–15</td>
<td>20–50</td>
<td>5–30</td>
</tr>
</tbody>
</table>

* Scintillometers measure sensible heat flux, only, and require estimating ET as a residual of the energy balance (\(\Delta E = R_n - G - H\)).

**2010b**, and the experience of the authors, who have participated in design and operation of weighing lysimeter systems, eddy covariance systems, Bowen ratio systems, soil water balance systems based on neutron probes and electromagnetic types of soil sensors, and systems employing scintillometers. The error values in Table 2 are presented with the following intentions:

i. providing reviewers of manuscripts or data sets employing ET measurements an indication of the amount of error in reported ET data to both anticipate and to perhaps tolerate (at expert levels);

ii. providing users of ET data sets or models calibrated from data sets an idea of the amount of error or bias that can be anticipated to exist within data sets or developed models; and

iii. providing perspective users of ET measurement system with some guidance on the amount of error to expect with various types of systems.

Because the human factor has large influence on amounts of error, estimated error ranges are given for both expert and novice users. The ranges given can provide guidance on interpretation of study results and to guide method selection.

Expected errors and ranges listed in Table 2 are relatively similar among types of systems and categories; with lysimeters having the smallest values, followed by Bowen ratio and remotely sensed energy balance. Eddy covariance, soil water balance systems and sap flow systems have moderately higher typical error and higher potentials for large error. This is partly due to challenges with local heterogeneity associated with soil water balance and sap flow and with measurement and error correction complications associated with eddy covariance.

Selection of the appropriate ET measurement method should not be based solely on expected error, but must also include system costs, system complexities, sensor fragility and power requirements, operator training and educational requirements and spatial sampling size.

5. Conclusions

Accuracy of ET measurement requires well-calibrated and well-maintained systems, and, in most cases, a foundational knowledge of the underlying physics of turbulence and heat and radiation transfer that govern the particular measurement. Knowledge of underlying physics is necessary to reduce the impacts of oversight of important biasing factors. In addition, an attitude of thoughtful approach to measurement and critical assessment of data are essential. Substantial measurement biases or incorrect data extrapolations can easily occur in ET data sets, many of which may not be recognized or identified and rectified prior to refereed publication. Deployment of systems that sample within the equilibrium boundary layer (eddy covariance, Bowen ratio and scintillometry) must adhere to fetch requirements and minimum equipment heights. Energy balance methods such as Bowen ratio and scintillometry that estimate ET as a residual must obtain accurate and representative measurements of net radiation and soil heat flux density to produce information on ET. For most vegetation systems, this requires multiple instrument locations within the measurement area to compensate for spatial non-uniformity of vegetation. Appropriate “corrections” must be made to eddy covariance measurements. Sap flow systems must include sufficient numbers of trees to reduce statistical error and to sample a representative population of trees. While remote sensing systems do not ‘measure’ ET, but rather deduce it via energy balance or vegetation indices, these methods are quite powerful in spatial coverage and quantification of spatial variation in ET, especially those based on thermally driven energy balance. All systems must be combined with rigorous quality assessment and quality control procedures. Precautions need to be exercised to produce accurate and defensible ET data and to recognize quality data.

Acknowledgements

The authors express their appreciation to the Technical Committee on Evapotranspiration in Irrigation and Hydrology of the American Society of Civil Engineers—Environmental and Water Resources Institute for encouraging quality assurance and integrity in evapotranspiration and weather measurement and for functioning as a sounding board for this article; to Dr. Edward Glenn of the University of Arizona and Dr. Teresa do Paço of the Technical University of Lisbon for review and contribution to the section on sap flow measurement; and to Dr. Jim Oster, editor of Agricultural Water Management, for encouraging the publication of this article.

References


