Quality Assessment of Weather Data and Micrometeorological Flux —Impacts on Evapotranspiration Calculation

Richard ALLEN[†]

(University of Idaho, USA)

Abstract

Quality assessment and control (correction) of weather data are essential to calculation of accurate and representative reference ET and to avoid biases in data. Simple visual (graphics-based) methods are presented that can be used to screen and correct large amounts of data. In the case of solar radiation, theoretical clear sky solar radiation is compared against for daily and hourly time periods. A simple least-square multiple linear regression method appears to be helpful in identifying components of eddy covariance that need adjustment to close the energy balance. The method offers an alternative to distributing error according to Bowen ratio and it provides corrections (regression coefficients) for H , LE and for soil heat flux.

Key words: Evapotranspiration, Quality assessment, Weather data.

1. Introduction

Automated weather stations (AWS) are commonly used in agricultural and environmental studies to measure primary variables affecting evapotranspiration (ET). These primary variables are typically solar radiation, air temperature, wind speed and humidity. Because modern AWS stations are electronic and frequently are in remote locations, data can sometimes be impacted by measurement biases caused by sensor malfunction, sensor aging, sensor miscalibration, sensor alignment, sensor cleanliness, datalogger programming, and sensor environment. When data are used to calculate ET, the AWS should measure air temperature, humidity and wind speed within the dynamic equilibrium boundary layer (EBL) overlying the ground surface. Properties of this boundary layer characterize the energy balance at the surface and are generally implicit to assumptions and conditions used in developing the particular ET equation used, for example, the Penman-Monteith equation (ASCE-EWRI 2005). In the case of calculating ‘reference’ ET, where the reference ET represents the rate of ET from an extensive surface of well-watered

vegetation having full-ground cover, the weather measurements should be made in a setting that includes enough well-watered vegetation within about 1 km in the upwind direction to condition the EBL to create congruency with the reference ET equation. As studies in southern Idaho by Burman *et al.* (1975) have shown, the lower level of the atmosphere (*i.e.*, EBL) changes when going from desert to a patchwork of irrigated and non-irrigated fields. Humidity, temperature and wind speed variables change when entering an irrigated field surrounded by dry or poorly irrigated fields. It is important, when making calculations of reference ET, that weather measurements are accurate and that the weather measurements reflect the environment that is defined by the reference surface.

The following sections suggest relatively simple methods for examining the accuracy of weather data collected from AWS prior to use in calculating reference ET (ET_0) and for estimating the ‘representativeness’ of the collected data for a ‘reference ET environment.’ A final section describes a simple procedure for detecting and correcting potential biases in Eddy Covariance measurement.

2. Weather Data Integrity Assessment

Weather data should be quality assessed (QA) and quality controlled (QC) before use in any ET equation, including standardized ET_0 equations, to ensure that

[†]Professor of Water Resources Engineering, Univ. Idaho, Kimberly, rallen@uidaho.edu. Presented at the International Symposium on Agricultural Meteorology (ISAM2008), Shimonoseki, Japan, 21–22 March 2008.

data are of good quality and are representative of well-watered conditions (because this is what the ET_0 equation is ‘expecting’). This is especially important with electronically collected data, since human oversight and maintenance may be limited. When weather measurements are determined to be faulty, they can frequently be adjusted or corrected using justifiable and defensible procedures, or the user may elect to replace perceived faulty data with estimates.

2.1. Weather Station Siting

When possible, meteorological data for estimating ET_0 should be measured over and downwind of vegetation that approximates the reference surface. This is important because the standardized ET_0 equations, for example the Penman-Monteith forms by FAO (Allen *et al.*, 1998; 2005) and by ASCE-EWRI (2005), were developed for use with meteorological data collected primarily over and downwind of relatively dense, fully transpiring grass or similar vegetation exhibiting behavior similar to the definition of the reference surface condition. Feedback exists between the EBL above the surface and the surface, so that the energy balance and evaporation at the surface impacts temperature and humidity of the air layer above. The preferred vegetation for a site is clipped grass. However, alfalfa or a grass-legume pasture maintained at a height of less than 0.5 m may also serve as an effective vegetation (ASCE-EWRI, 2005). Meteorological measurements made over other short, green, actively transpiring crops will approach reference measurements, provided canopy cover exceeds approximately 70%. A station may be located on the periphery of a vegetated field provided the station has the field as upwind fetch during daytime hours and that vegetation is shorter than about 0.5 m so as to not impede the wind measurement. In an ideal setting, the well-watered vegetation extends at least 100 m in all directions from the weather station. However, it is recognized that frequently such a weather station site is not available, and that often some nonvegetated areas or roadways will be present near the station.

Failure of a weather station site to meet the definition of a reference condition described above does not preclude the use of the data for estimating ET_0 . However, data from such a station should be examined carefully before use, and may, in some cases, require adjustment to make the data more representative of reference conditions. Allen *et al.* (1998) and ASCE-EWRI (2005), appendices D and E, provide guidance

for assessing whether temperature and humidity data from a weather station located in an arid or semiarid climate exhibit ‘reference-like’ characteristics. Those publications suggest methods to adjust data to resolve impacts of dryness of weather station environments.

2.2 Data Quality Control

Meteorological data sets acquired for the purposes of estimating ET_0 should be subjected to a number of quality control checks prior to use. The first and most important quality control check involves determining the following information (ASCE-EWRI, 2005):

1. Siting (environment) of the weather station (Google-Earth is useful to evaluate the “greenness” of the area within 1 to 10 km of the station).
2. Type and exposure of meteorological sensors employed at the station.
3. Procedures used to maintain and calibrate sensors.
4. Quality control procedures performed and/or data adjustments already performed on the data.
5. Availability of short interval data sets (*e.g.*, hourly) to aid the overall QC process.
6. The station operator’s experience and/or recommendations pertaining to use of the data for ET_0 assessment.

The types of sensors employed and their exposure (*e.g.*, height of installation or type of radiation shelter) provide insight into expected error levels for specific measurements, and may identify measurements requiring some form of adjustment (*e.g.*, height adjustment for wind speed).

Procedures used to maintain and calibrate meteorological sensors are important. Maintenance activities that impact sensor readings include site maintenance (*e.g.*, mowing, irrigation, and fence repair); cleaning of sensors; and leveling radiation sensors and rain gauges. On-site calibration may be performed at regular intervals by comparing sensors with calibrated sensors that are taken to the site for inter-comparison purposes. The operator of the station should present both the technical and non-technical maintenance protocols and schedule logs either on a request basis or on a public web site. ASAE (2004) describes Engineering Practice 505: “Measurement and Reporting Practices for Automatic Agricultural Weather Stations.” This standard provides specifications for sensor accuracy, resolution, placement and monitoring, as well as intervals and procedures for sensor maintenance and calibration.

It is always advisable to investigate the various QA/QC routines that have been employed on the data

set by the operator of the station. Data from weather stations operated as part of a weather network are generally subjected to some form of QC assessment (e.g., Snyder *et al.*, 1985; Stanhill, 1992; Meek and Hatfield, 1994; Snyder *et al.*, 1996; Shafer *et al.*, 2000). Common QA includes comparing incoming parameters against relevant physical extremes (e.g., relative humidity $>100\%$); using statistical techniques to identify extreme or anomalous values; and comparing data with neighboring stations. Some networks flag questionable data while others may replace questionable data with estimated values. Some networks contain rather broad data range assessments, so that application of a QA/QC procedure does not necessarily provide valid data.

Seeking the advice of the station operator regarding the fitness of a given meteorological data set for ET_0 assessment is always advisable. The operator may have considerable insight into whether the station site approaches reference conditions, and if not, may offer suggestions on how to correct or adjust the raw meteorological data. The procedures presented here for QA/QC are recommended by ASCE-EWRI (2005). These procedures are applicable in many circumstances, however, they are by no means a universal solution to all potential problems with meteorological data. Users are encouraged to seek local input regarding the subject of assessment and correction.

2.3 Radiation

2.3.1 Solar Radiation

Solar radiation (R_s) data can be screened by plotting measurements against theoretically derived solar radiation for clear sky (R_{so}) envelopes for hourly or daily timesteps. Generally, the best estimates of R_{so} apply equations that include the influence of sun angle, turbidity, atmospheric thickness, and precipitable water. For daily data sets, one can plot measured R_s and computed R_{so} against the day of the year (Fig. 1). For hourly data, one can plot measured R_s and computed R_{so} against time of day, for example, for five to ten selected “clear sky” days (Fig. 2).

After creating the R_s and R_{so} plots, the user can observe whether measured R_s “bumps” against the clear sky envelope on what the examiner concludes are cloud-free days for daily data or during cloud-free hours for hourly data. R_s will fall below the clear sky curve on cloudy or hazy days and during times when the atmosphere is more turbid than under conditions of clean air. The transmissivity of the atmosphere and thus R_{so} can shift by several percent from day to day due to changes in water vapor, particulate matter and aerosols, which are net absorbers of R_s . If the “upper” values of measured R_s lie routinely above or below the computed R_{so} curve by more than 3 to 5%, then the user should scrutinize the maintenance and calibration of the R_s sensor. Improper calibration, leveling errors,

Fig. 1. Daily Measured R_s and Calculated R_{so} for Aberdeen, Idaho, USA (lat. 42.95N, long. 112.83W, elev. 1340 m), during 1992 and 1993 prior to any adjustment (top) and following adjustment (bottom).

the presence of contaminants on the sensor (e.g., dust, salt, or bird droppings), or electrical problems can cause R_s to deviate from R_{s0} on clear days. “Abrupt” changes in the relationship between R_s and R_{s0} on clear days generally indicate: 1) accumulation or removal of contaminants from the sensor; 2) change in sensor level; 3) change in sensor calibration; 4) sensor replacement; or 5) problem with wiring or data-acquisition system. Pyranometer maintenance records, if available, may help explain changes in the relationship between R_s and R_{s0} and aid decisions related to data adjustment. Occasionally, R_s during hourly periods may exceed R_{s0} due to reflection of sunlight from nearby clouds.

Values for R_s that are consistently above or below R_{s0} on clear days can often be adjusted by dividing R_s by the average value of R_s/R_{s0} on clear days. This assumes that there is a systematic cause to the over- or undermeasurement. For example, for the data shown in Fig. 1, measured R_s on apparently clear days during summer had undermeasurement of about 15%. The reported R_s values were therefore multiplied by 1.14 for the period from day 90 to day 250 for year 1992 and by 1.16 for the period from day 90 to day 240 for year 1993, based on visual examination of the data. Occasional data points that exceeded the daily R_{s0} curve by more than about 5% were judged to be caused by faulty data measurement or processing, and were limited to R_{s0} . The adjustment of data using a multiplier should be used with appropriate caution because the procedure assumes: 1) computed values for R_{s0} are correct; 2) clear days can be effectively identified; and 3) the factor causing R_s to deviate from R_{s0} is static over time. The R_{s0} curves computed by the following equations are not “perfect.” They assume clean air and common relationships between the diffuse and beam components of short wave radiation along with typical spectral densities within the short wave band. Identification of clear days can be difficult in

cloud prone areas, especially if hourly R_s data are not available to aid in the assessment process. Finally, many of the factors causing R_s to deviate from R_{s0} , including leveling errors and contaminant accumulation (See Stanhill, 1992), may not be static over time.

The following procedure for estimating theoretical R_s under clear skies is from ASCE-EWRI (2005) and considers effects of sun angle and water vapor on absorption of short wave radiation and separates the components of beam and diffuse radiation, so that:

$$R_{s0} = (K_B + K_D) R_a \quad (1)$$

where:

K_B =the clearness index for direct beam radiation [unitless]

K_D =the transmissivity index for diffuse radiation [unitless]

R_a =extraterrestrial radiation having the same units as R_s ($\text{MJ m}^{-2} \text{d}^{-1}$ or $\text{MJ m}^{-2} \text{h}^{-1}$ or W m^{-2})

The estimation of K_B , extended from Majumdar *et al.* (1972) by Allen (1996) and Allen *et al.* (1998), is applied by ASCE-EWRI (2005) using improved coefficients developed from the ASCE Task Committee evaluation of R_s data from nearly fifty weather sites throughout the US:

$$K_B = 0.98 \exp \left[\frac{-0.00146P}{K_t \sin \beta} - 0.075 \left(\frac{W}{\sin \beta} \right)^{0.4} \right] \quad (2)$$

where:

K_t =turbidity coefficient [unitless], $0 < K_t \leq 1.0$ where $K_t=1.0$ for clean air and $K_t \leq 0.5$ for turbid, dusty or polluted air.

P =atmospheric pressure, which can be estimated from the site elevation [kPa]

β =angle of the sun above the horizon [rad]

W =precipitable water in the atmosphere [mm]

The value for K_t may vary with time of year and with

Fig. 2. Hourly R_s and R_{s0} data for Kimberly, Idaho (lat. 42.55N, long. 114.35W, elev. 1195 m) in 2006 (Data from Garcia, 2007).

cleansing of the atmosphere by precipitation. General values for K_t for a region can be determined using a pristine pyranometer that has a calibration traceable to the national or international solar standard. In general, for routine prediction of R_n and R_{so} envelopes, $K_t=1.0$ is recommended. The $\sin \beta$ in Eq. 2 should be limited to ≥ 0.01 for computational stability.

Precipitable water, for purposes of applying Eq. (2), is estimated as:

$$W = 0.14e_a P + 2.1 \quad (3)$$

where:

W =precipitable water in the atmosphere [mm]

e_a =actual vapor pressure of the air (at approximately 2 m) [kPa]

P =atmospheric pressure at the site elevation [kPa]

The diffuse radiation index is estimated from K_B (following Allen, 1996):

$$K_D = 0.35 - 0.36K_B \quad \text{for } K_B \geq 0.15$$

$$K_D = 0.18 + 0.82K_B \quad \text{for } K_B < 0.15$$

(4)

For daily (24-hour) time periods, the average value of β , weighted according to R_a , can be approximated from Allen (1996) as:

$$(\sin \beta)_{24} = \sin \left[0.85 + 0.3\varphi \sin \left(\frac{2\pi}{365} J - 1.39 \right) - 0.42\varphi^2 \right] \quad (5)$$

where:

$(\sin \beta)_{24}$ =average $\sin \beta$ during the daylight period, weighted according to R_a [rad]

φ =latitude [rad]

J =day of the year [unitless]

The $(\sin \beta)_{24}$ variable is to be used in place of $\sin \beta$ in Eq. (2) and represents the weighted average sun angle during daylight hours. The value for $(\sin \beta)_{24}$ should be limited to ≥ 0 . For hourly or shorter periods the sun angle β is calculated as:

$$\sin \beta = \sin \varphi \sin \delta + \cos \varphi \cos \delta \cos \omega \quad (6)$$

where:

φ =latitude [rad]

δ =solar declination ("delta") [rad]

ω =solar time angle at the midpoint of the hourly or shorter period [rad]

Calculation of δ and ω is described in Allen *et al.* (1998), ASCE-EWRI (2005) and elsewhere. Plotting hourly measured R_s against the theoretical R_{so} can be

helpful in detecting errors or shifts in the reported times associated with the data set (*i.e.*, errors in datalogger time clocks). Shifts in time and lack of instrument level can both cause measured R_s to plot out of phase with the theoretical R_{so} curve. Fig. 2 shows hourly data plotted for five days in 2006 for Kimberly, Idaho, where measured R_s agrees well with estimated R_{so} . The good alignment of hourly values between R_s and R_{so} indicates that the pyranometer had good levelness and that the datalogger time clock was accurately set.

2.3.2. Net Radiation

When net radiation (R_n) data are measured, values can be compared with R_n estimated from standard equations that are based on R_s , for example those presented in FAO-56 (Allen *et al.*, 1998) and elsewhere, as a means of integrity assessment. One should not expect measured R_n to exactly agree with estimated R_n . However, significant variation between the two should be cause for a closer investigation of the measured data. Some net radiometers do not accurately measure the long wave component of net radiation. In addition, when compared to computed standardized R_n for grass, the R_n measurement should be made over a well-watered surface of clipped grass so that albedo is similar to that defined for ET_o (albedo~0.23). A shift in the relationship between measured and estimated R_n may reflect a change in the quality or condition of the surface at the measurement site. Other measurement related factors that can shift the relationship between measured and estimated R_n include scratched or dirty radiometer domes, an off-level sensor, or condensation of moisture inside domes of the R_n sensor. Allen (1996) and ASCE-EWRI (2005) show comparisons between measured and computed R_n .

2.4. Humidity and Air Temperature

2.4.1. Relative Humidity

As with R_s data, humidity and air temperature data can be visually screened, relatively rapidly, to identify questionable data. A portion of the screening process involves the user having a sense of what are reasonable and unreasonable values. For example, relative humidity (RH) values chronically lower than 5 to 10% in arid regions or lower than 30% in subhumid regions are generally not expected under 'reference conditions' and therefore may indicate problems with a sensor. Arid, semiarid and subhumid climates can be loosely defined as regions where annual precipitation is less than 200 mm (arid), between 200 and 500 mm (semiarid) and between 500 and 1000 mm (subhumid).

Fig. 3. Hourly measured relative humidity from three locations in central New Mexico, USA for a seven day period during February 2007 (Alb. Bosque —lat. 35.26N, long. 106.60W, elev. 1503 m; Jarales —lat. 34.61N, long. 106.75W, elev. 1460 m; Angostura— lat. 35.38N, long. 105.50W, elev. 1550 m) (data from Garcia, 2008a).

Similarly, RH values in excess of 100% do not occur in the natural environment and may indicate that the sensor is out of calibration. The accuracy of many modern-day electronic RH sensors is generally within $\pm 5\%$ RH (ASAE, 2004); thus, recorded RH values in excess of 105% provide good evidence that the sensor is out of calibration.

One should use RH data sets containing values in excess of 100% with caution. All RH values in excess of 100% should be set equal to 100% prior to use in the ET_0 computation process. Use of this simple adjustment procedure, however, does not alleviate sensor calibration errors in recorded RH data that lie below 100%. Some type of proportional adjustment to all data may be warranted. Furthermore, RH values in excess of 100%, if not accompanied by a QC flag, may indicate that the data set has not been subjected to rigorous QC.

If hourly data are available, one can examine the diurnal variation of RH on selected days to ensure that RH approaches maximum and minimum levels during the coolest and warmest portions of the day, respectively. Hourly time series of RH should also be examined for the presence of spikes and spurious values of RH that may indicate sensor malfunction. Finally, one should check RH data on several days having heavy and/or sustained precipitation events or when dew or fog events are known to have occurred. Relative humidity should approach 90–100% during a sustained precipitation, fog, or dew event, and should approach 100% in the evening hours following a heavy rain event. For example, hourly RH data in Fig. 3 for seven days in 2007 for three locations in New Mexico, USA, show RH to approach 100% on two of the seven days, which is indicative of well-watered environments. In addition, there is relatively good

consistency in measurements among three stations that are compared, which is an indicator of consistency of sensor operation at all locations.

2.4.2. Dew-point Temperature

Dew-point temperature (T_{dew}), defined as the air temperature at which air reaches saturation, can be used to represent and assess humidity information. T_{dew} can be calculated directly from vapor pressure (e_a). Any errors in measured RH will affect the calculation of vapor pressure, when e_a is calculated as $e_a = e_s(T) RH/100$, where $e_s(T)$ is the saturation vapor pressure function, and therefore, error in RH will affect the computed T_{dew} . Values for daily average or early morning T_{dew} can be compared to daily minimum temperature (T_{min}). In humid regions, the T_{dew} measurement will approach T_{min} on many days. Exceptions occur on days that feature a change in air mass (e.g., frontal passage), or on days that have high winds and/or cloudiness at night. T_{dew} may approach T_{min} in arid and semiarid environments if nighttime winds are light and measurements are made over a surface exhibiting behavior similar to the reference definition (i.e., sufficient evaporation to cause evaporative cooling) (ASCE-EWRI 2005). It is not uncommon in arid and semiarid regions to have T_{dew} 2 to 5°C lower than T_{min} under reference conditions (see discussion below) but well below T_{min} if the measurement site is subjected to local and regional dryness (lack of well-watered vegetation). If daily average T_{dew} regularly and consistently exceeds T_{min} , then the humidity sensor may be out of calibration during summer. Such data should be examined closely and possibly adjusted prior to use. T_{dew} can be calculated from RH or e_a using standard equations, such as those in Allen *et al.* (1998) or ASCE-EWRI (2005).

Fig. 4 illustrates the use of comparisons between T_{min} and T_{dew} for reported daily AWS data from

Fig. 4. Daily minimum air temperature and average daily dewpoint temperature for year 2006 near Kimberly, Idaho (lat. 42.55N, long. 114.35W, elev. 1195 m), which has a semiarid climate (data from Garcia, 2007).

Fig. 5. Hourly measured T_{air} and T_{dew} for Rexburg, Idaho (lat. 43.85N, long. 111.77W, elev. 1486 m) for the period July 11th to 13th 2002. Data from Garcia (2008b).

Kimberly, Idaho during 2006. In this example, the weather data were collected over irrigated grass in an area surrounded by 150,000 ha of irrigated agriculture. Although a semiarid environment, the proximity of T_{dew} to T_{min} reflects the impact of evaporation from the irrigated area on cooling the EBL and adding humidity. Often, the close relation between T_{dew} and T_{min} can be used to estimate T_{dew} for use in estimating reference ET_o by employing the relationship:

$$T_{dew} = T_{min} - K_o \quad (7)$$

where K_o is the average difference between T_{min} and T_{dew} . Allen *et al.* (2007) derived monthly values for K_o for southern Idaho that were used to estimate T_{dew} for locations that measured only air temperature. In Fig. 4, the value for K_o averages about 2°C during the growing season.

Plots of hourly or shorter period T_{dew} data may assist in identifying problems in T_{dew} data. Dew point and vapor pressure are relatively conservative parameters and often exhibit little change over a day, especially in humid regions. Often, T_{dew} will increase somewhat during midmorning due to evaporation of water and increased capacity for the air to contain vapor. Dew point will then stabilize or decline slightly during

the mid-day hours as the vapor near the surface gets mixed into a progressively deeper and potentially drier boundary layer. Hourly variation in T_{dew} is greater in semiarid and arid settings, especially in areas prone to strong regional advection, where T_{dew} can drop substantially during afternoon as warm, dry air from outside an irrigated area “breaks into” the boundary layer established over the irrigated area. However, large changes in T_{dew} during the day, except under circumstances such as a change in air mass (*e.g.*, frontal activity or sea/land breeze) or large change in wind direction, could signal an error or bias in the T_{dew} measurement. Fig. 5 illustrates these characteristics for hourly T_{air} and T_{dew} over a three day period for an AWS in an irrigated region having a semi arid climate near Rexburg, Idaho, USA.

2.4.3. Air Temperature

In general, air temperature (T_{air}) is the simplest and most consistent weather parameter to measure and the parameter most likely to be of highest quality. However, for use in calculating ET_o , T_{air} should be measured in a reference-type of environment so that the EBL is conditioned by the evaporating surface. Air temperature extremes in data sets should be compared to historical record extremes, if such data are avail-

able. Temperatures that routinely exceed the recorded extremes for a region indicate a problem with either the sensor or with the radiation shield used to house the sensor. Sensors mounted in non-aspirated radiation shields may produce erroneously high temperatures on days having light winds due to solar heating of the shield (Gill, 1983). Consistently hot temperatures from a sensor mounted in an aspirated radiation shield may indicate problems with the ventilation system. An effective check for spuriously high or low temperature extremes is to compare the average of the daily extremes ($T_{\max}$ and $T_{\min}$) with the mean daily temperature as averaged by the data logger for the day.

2.4.4. Aridity of Temperature and Humidity Data

Often, substituting $T_{\text{dew}} = T_{\min} - K_o$ for measured T_{dew} (i.e., using Eq. (7)), can improve estimates of daily ET_o when weather data are collected from a non-reference, arid setting. In arid and semiarid regions, it is best to check with the source of weather data to determine if K_o values have been developed for the area. For example, in Arizona the value of K_o was found to vary from 2 to 5°C over the course of a year. When local information on K_o is not available, a K_o in the range of 2 to 4°C is recommended for semiarid and arid regions (Allen, 1996; ASCE-EWRI 2005). In humid regions where T_{dew} approaches $T_{\min}$ on most nights, K_o can generally be set equal to 0°C. Some irrigated areas in very dry, advective climates can have extended periods during which $T_{\min}$ is more than 5°C above T_{dew} . High nighttime wind speed enhances the transfer of sensible heat and dry air to the surface, slowing the rate of cooling and preventing full humidification of the atmospheric boundary layer above well-watered surfaces. Caution should be exercised in the specification of K_o .

2.5 Wind Speed

Accuracy of wind speed measurements is difficult to assess unless duplicate instruments are used. Nevertheless, one should visually inspect wind records for the presence of consistently low wind speed values that may indicate a malfunctioning or failed anemometer or the presence of ice if air temperatures are near or below 0°C. Consistent and low wind speeds can indicate dirty or worn anemometer bearings that will increase the anemometer wind speed threshold and might eventually seize and stop the anemometer altogether. Wind speeds from failed anemometers will usually appear as small, constant values (less than 0.5 m s⁻¹ or the wind speed threshold for a new anemometer). With a failed anemometer, recorded maximum and mean wind speeds will often have the same values and will generally equal any numerical offset in the calibration equation.

Any appreciable period having daily mean wind speeds of less than 1.0 m s⁻¹ should be viewed with caution. Aside from exceptionally calm periods or anemometer problems, other reasons for daily wind speeds of less than 1.0 m s⁻¹ include excessive vegetation height at the station or the presence of blocking structures in the nearby landscape (e.g., solid fences or buildings).

Comparing data from a nearby station may assist in the assessment of wind speeds at a particular site. In most cases, winds at two nearby locations are related which indicates the ratio of the wind speeds at the two locations will remain nearly constant. A sudden and consistent change in the ratio often indicates a failing anemometer; a gradual change in ratio may indicate growing contamination in the bearings or too tall of growing vegetation at one site.

Daily wind speed data from two AWS stations 50 km apart is shown in Fig. 6. Both stations are located

Fig. 6. Comparison of daily average wind speed during 2002 for two weather stations (Ashton, Idaho (lat. 42.95N, long. 112.83W, elev. 1340 m) (top) and Rexburg, Idaho (lat. 43.85N, long. 111.77W, elev. 1486 m) (bottom) that are 50 km apart (data from Garcia, 2008a).

in agricultural environments in Idaho and wind was measured at 2 m above the ground. The similarity in wind speed records is apparent, indicating that both anemometers are probably functioning well. ASCE-EWRI (2005) contains an example of the impact of growing a crop of tall corn on measured wind speed. Allen and Wright (1997) have suggested procedures for adjusting wind speed data over rough crops to better represent reference conditions.

ASAE (2004) provides detailed recommendations on periodic sensor maintenance. An alternative technique for evaluating anemometers involves redundancy in instrumentation and requires placement of a second anemometer of the same design, but with fresh bearings, at the weather station for a three or four day period at least once each year, and comparing recorded values. Variations between recordings can signal a need to replace bearings, switches, or other parts.

2.6 Impact on Reference Evapotranspiration

Poor weather data quality will of course impact the quality of the ET_0 estimate. For example, hourly weather data that were QA/QC'd for an AWS in New Mexico for the first eight months of 2007 produced reference ET estimates that were as much as 50% higher during September when the data were not corrected. Most of the error was in solar radiation data. Daily reference ET estimates (alfalfa reference basis using ASCE-EWRI 2005) are shown in Fig. 7a and monthly totals are shown in Fig. 7b.

3. Quality Control of Eddy Covariance Data

Much literature has appeared during the past fifteen years documenting energy balance ‘closure error’ for eddy covariance (EC) data that indicate undermeasurement of latent heat flux (LE) and sensible heat flux

(H) by those systems to be as much as 30% (Twine *et al.* 2000; Wilson *et al.* 2002; Foken 2007). A number of corrections to EC data have been proposed to account for impacts of coordinate rotation of the 3D sonic wind speed vectors (Tanner and Thurtell 1969; Kaimal and Finnigan, 1994; Wilczak *et al.*, 2001), sonic temperature correction (Munger and Loescher 2004), air moisture correction (Sun *et al.* 1995), impacts of density terms (Webb *et al.*, 1980), and correction for pathlength averaging, sensor separation and high frequency spectral losses (Moore 1986). 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 Edinburgh, 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 EBL, nonuniform footprint, nonrepresentative sensing of R_n and G , and proper averaging times.

The software packages noted above are recommended for general corrections and should be applied as needed. Often, however, energy balance closure errors still remain in the data, where energy balance closure error (EBCE) is defined as:

$$EBCE = (R_n - G) - (LE + H) \quad (8)$$

where R_n and G (soil heat flux) comprise the ‘available energy’ and the sum of LE and H comprise the measured latent and sensible heat fluxes. The following procedure is presented as a simple means to explore

Fig. 7. Daily (left) and monthly (right) reference ET (alfalfa reference basis) for the Bosque Albuquerque weather station (lat. 35.26N, long. 106.60W, elev. 1503 m) calculated from hourly weather data before and after correction of the data using previously described QA/QC procedures (data from Garcia, 2007a).

the consistency of closure error in a data set and as an attempt to 'identify' the components of the energy balance that systematically contribute most to the EBCE on a consistent basis. This consistent error basis would be caused by a systematic calibration error or other calibration problem with a measurement sensor or estimate that is caused by sensor placement, sensor representativeness, flux divergence, etc.

The following method uses a simple application of least-squares linear regression on the energy balance components as:

$$R_n = c_0 + c_1 G + c_2 LE + c_3 H \quad (9)$$

to identify the component(s) having the most systematic effect on closure. R_n is placed on the left hand side of the equation under the assumption that R_n is most likely to be accurate. However, if R_n is felt to be nonrepresentative of the measurement of the footprint sensed and some other parameter, for example H , is sensed to be more dependable, then these variables can be switched in Eq. (9). The use of regression analysis to both close the energy balance, on average, and to help identify systematic biases works well because of the tendency of the four data components to vary, over the course of a day, in differing directions from one another at different times of the day, or with different extreme ranges from zero. For example, R_n reaches a positive peak during midday that is often 10 times the negative extreme that occurs at night. G reaches positive peaks during day time (positive into ground) that are of similar magnitude to negative peaks during night. LE reaches positive peaks during daytime and near zero lows during night, whereas H can vary widely from low positives or negatives during day under high levels of LE to high positives during day when LE is small. H is generally negative at night. The lack of full correlation among the timings and peaks is useful in providing means to derive relatively unique coefficients for Eq. (9).

3.1 Example Applications

Example applications of least-squares regression analysis for identifying systematic correction for EBCE in EC data are shown in Tables 1-3 for three data sets. The first data set was gathered during an Aug. 27-Sept. 17, 1999 experiment (deBruin *et al.*, 2005) near Hansen, Idaho from an irrigated alfalfa field having 2 km of fetch. Four EC systems were utilized in the study, but only one is shown. All systems exhibited EBCE of 25% during daytime hours. The equipment

were CSAT3 sonic anemometers with LiCor open path hygrometers. Three units were placed at 3 m height and one was at 10 m. R_n was measured using a CNR-1 four-component radiometer that compared closely with a collocated Swissteco S-1 net radiometer and G was measured by two sets of soil heat flux plates at 10 cm, with calorimetric computations for the layer above.

The second data set was by Kizer and Elliott (1991) and was made over alfalfa near Stillwater, Oklahoma during 1987. A 1-D CA-27 Campbell Scientific sonic anemometer and Krypton KH-20 hygrometer was used, with G measured at only 1 cm depth and R_n from a Swissteco S-1 net radiometer. Only one day of 30 minute data was evaluated in this test.

The third data set was from Sridhar *et al.* (2002) for data collected during 1999 and 2000 near Norman, OK, over grass using a CSAT3 and Krypton KH-20 hygrometer with G -plates at 5 cm and R_n measured using a CNR-1 radiometer. This data set was screened by Sridhar *et al.* (2002) to include only those 30 minute periods where EBCE was within $\pm 10 \text{ W m}^{-2}$. Therefore, full EB closure occurs for each time period and the regression analysis should have no impact.

Eq. (9) was applied to the three data sets using all data, using daytime data only, and using nighttime data only to assess impacts on regression coefficients. Results are shown in Tables 1-3. For the Idaho data set (Table 1), r^2 was highest when all data were included, even though the r^2 for nighttime was poor when nighttime periods were evaluated separately. This is possibly due to larger extremes and variations in the combined data set that reduced the mean square error. When nighttime periods were evaluated separately, the signal to noise ratio was apparently too small for most energy balance component readings to preserve any correlation among components and thus the r^2 was 0.0. No significant improvement in r^2 was achieved with the use of the Y-intercept (c_0), however, the magnitude of the intercepts is large (42 W m^{-2} for daytime and -40 W m^{-2} for nighttime). The correction on H changes from 0.95 to 1.03 as an intercept is introduced to the regression. The values for c_3 near 1 suggest that the measurements for H are valid and probably should not be corrected, and that most error resides with LE , where the value for c_2 is 1.26 and 1.36. Thus, the LE may have been undermeasured by 25 to 35% and the error in LE was somewhat constant over all daytime periods. When the c_0 intercept is used, the correction on LE reduces from 26% to 8%, which is significant,

indicating that a constant error term of 42 W m^{-2} impacts the relation of closure to LE , primarily. In general, the regression analysis suggests that soil heat flux, while relatively small under the alfalfa cover, may have been undermeasured by a factor of two ($c_1=2.1$ for daytime data) for the eddy correlation footprint. Therefore, one might conclude that on average, the H data were valid measurements with no adjustment, the LE data require an adjustment factor of 1.3 and the G data require an adjustment factor of 2.0. Data are shown in Fig. 8 (left) where $LE+H$ is plotted against R_n-G before and after correction by the regression. Following regression, the data tended around the 1:1 line. In Fig. 8 (right), LE alone is plotted against R_n-G . More scatter exists in the LE data alone, with some data above the

1:1 line, illustrating impacts of advection.

In the Kizer and Elliott data set (shown for one day only), correction by regression brought the $LE+H$ data close to a 1:1 line with R_n-G (Fig. 9, left), although some hysteresis existed at high values for LE . The regression for all data (Table 2) suggests that the soil heat flux was undermeasured by a factor of 3 and LE by about 30 to 40%. As with the Idaho data, the coefficient c_3 on H was close to 1.0, indicating that H was correctly measured during daytime. These findings and conclusions assume that R_n was correctly measured and was representative of the footprints for the H and LE fluxes. In Fig. 9 (right), LE is plotted against R_n-G , where correction before and after regression is shown. In addition, correction of LE using the strategy

Fig. 8. 30 minute EC data collected near Hansen, Idaho, USA (lat. 42.46N, long. 114.25W, elev. 1273 m) during Aug. 27-Sept. 17, 1999, before and following correction using the least-squares linear regression for $H+LE$ (left) and LE (right) vs. R_n-G .

Fig. 9. 30 minute EC data collected by Kizer (lat. 35.14N, long. 98.40W, elev. 430 m on July 9-10, 1989), before and following correction using the least-squares linear regression for $H+LE$ (left) and LE (right) vs. R_n-G , along with LE corrected according to Bowen Ratio.

Table 1. Regression coefficients for 30 minute EC data collected over alfalfa during the Idaho-WAU-RAPID experiment, Hansen, Idaho (lat. 42.46N, long. 114.25W, elev. 1273 m), Aug. 27-Sept. 17, 1999, using Eq. (9) .

Data	Intercept (c_0), W/m ²	Coef. on G (c_1)	Coef. on LE (c_2)	Coef. on H (c_3)	n	r^2	SEE, W/m ²
all	—	1.491	1.359	1.077	941	0.956	46.5
daytime	—	2.063	1.261	0.954	443	0.883	61.2
nighttime	—	1.278	0.234	0.429	498	0.000	18.2
daytime	42.5	2.066	1.085	1.030	443	0.895	57.9
nighttime	−39.6	0.431	0.310	0.351	498	0.339	12.3

Table 2. Regression coefficients for one 24 hour period of 30 minute EC data from Kizer of Oklahoma State Univ. (lat. 35.14N, long. 98.40W, elev. 430 m) (data from Sridhar, 2000, pers. commun.).

Data	Intercept (c_0), W/m ²	Coef. on G (c_1)	Coef. on LE (c_2)	Coef. on H (c_3)	n	r^2	SEE, W/m ²
all	—	3.08	1.40	1.06	48	0.98	36
daytime	—	3.73	1.33	0.93	48	0.94	50
nighttime	—	1.28	1.10	1.82	48	0.84	6.3
all	109	7.35	0.47	1.12	48	0.99	27
daytime	130	6.81	0.46	1.20	48	0.97	34
nighttime	−20	0.41	1.30	1.85	48	0.85	6.3

Table 3. Regression coefficients for 60 minute EC data from June 1999–May 2000, Norman, Oklahoma (lat. 35.24N, long. 97.46W, elev. 357 m) for periods having low closure error ($-10 < \text{EBCE} < 10 \text{ W m}^{-2}$) (data from Sridhar *et al.*, 2002; Sridhar, 2000, pers. commun.).

Data	Intercept (c_0), W/m ²	Coef. on G (c_1)	Coef. on LE (c_2)	Coef. on H (c_3)	n	r^2	SEE, W/m ²
all	—	0.96	1.02	1.01	961	0.999	5.1
daytime	—	0.99	1.01	1.00	210	0.999	5.8
nighttime	—	0.95	0.98	1.00	751	0.92	4.8
all	0.33	0.96	1.02	1.01	961	0.999	5.1
daytime	0.78	0.99	1.01	0.99	210	0.999	5.8
nighttime	−3.1	0.90	0.98	0.97	751	0.92	4.7

proposed by Twine *et al.* (2000), where the EBCE is distributed between H and LE according to the Bowen ratio, is shown. In this case, the corrected LE using Bowen ratio exhibits more hysteresis.

The filtered data of Sridhar *et al.* (2002), where EBCE is less than 10 W m^{-2} , serves as a sort of ‘test’ of the regression correction method, where, because each sample in the data set needs little or no correction, all three coefficients are nearly 1.0 in value, r^2 are nearly 1.0 and the intercept, if employed, is nearly 0.0 (Table 3). This demonstrates that the regression method will not ‘manufacture’ unneeded corrections if the original data are valid and no EBCE exists.

4. Summary

QA/QC of weather data is essential before calculating

quality reference ET and to avoid any biases in data. Simple visual (graphics-based) methods are possible to screen and correct data. A simple least-squares multiple linear regression method appears to be helpful in identifying components of eddy covariance that need adjustment to close the energy balance. The method offers an alternative to distributing error according to Bowen ratio and it provides corrections (regression coefficients) for H , LE and for soil heat flux.

Acknowledgements

Eddy covariance data for Oklahoma were provided by Dr. V. Sridhar of Boise State University, formerly of Oklahoma State University. Idaho RAPID EC data were collected by the author and faculty of Wageningen Agricultural University, the Netherlands. Illustrative

graphs of weather data were constructed by Dr. Magali Garcia, visiting professor at the University of Idaho and professor of agricultural engineering at the Univ. La Pas, Bolivia.

References

- Allen, R. G., 1996: Assessing integrity of weather data for use in reference evapotranspiration estimation. *J. Irrig. Drain. Eng., ASCE*, **122**, 97–106.
- Allen, R. G., and Wright, J. L., 1997: Translating wind measurements from weather stations to agricultural crops. *J. Hydrol. Eng., ASCE*, **2**, 26–35.
- Allen, R. G., Pereira, L. S., Raes, D., and Smith, M., 1998: Crop evapotranspiration: guidelines for computing crop water requirements. *Irrig. and Drain. Paper 56*, Food and Agriculture Organization of the United Nations, Rome, 300 pp.
- Allen, R. G., Pruitt, W. O., Wright, J. L., Howell, T. A., Ventura, F., Snyder, R., Itenfisu, D., Steduto, P., Berengena, J., Baselga Yrisarry, J., Smith, M., Pereira, L.S., Raes, D., Perrier, A., Alves, I., Walter, I., and Elliott, R., 2005: A recommendation on standardized surface resistance for hourly calculation of reference ET_0 by the FAO56 Penman-Monteith method. *Agric. Water Manag.*, **81**, 1–22.
- Allen, R. G., Robison, C. W., and Wright, J. L., 2007: Updated procedures for calculating state-wide consumptive use in Idaho. In *the Role of Irrigation in a Sustainable Future*, USCID, pp. 189–212.
- ASAE, 2004: *Measurement and Reporting Practices for Automatic Agricultural Weather Stations*. Engineering Practice 505. St. Joseph, MI, 21.
- ASCE-EWRI, 2005: *The ASCE Standardized Reference Evapotranspiration Equation*. ASCE-EWRI Standardization of Reference Evapotranspiration Task Committee Report, available at <http://www.kimberly.uidaho.edu/water/asceewri/>
- Burman, R. D., Wright, J. L., and Jensen, M. E., 1975: Changes in climate and estimated evaporation across a large irrigated area in Idaho. *Trans. ASAE*, **18**, 1089–1093.
- deBruin, H. A. R., Hartogensis, O. K., Allen, R. G., and Kramer, J. W. J. L., 2005: Note on the regional advection perturbations in an irrigated desert (RAP-ID) experiment. *Theor. Appl. Climatol.* **80**, 143–152 (DOI: 10.1007/s00704-004-0095-y).
- Foken, T. 2007: The energy balance closure problem –an overview. *Ecol. Applicat.* (accepted)
- Garcia, M., 2007: *Quality Control of 2006 Weather Data for Kimberly, Idaho*. Internal report, University of Idaho, Kimberly, Idaho. 8.
- Garcia, M., 2008a: *Quality Control of Hourly AgriMet Weather Data for Year 2002, for Stations Located in Landsat Image Path 39-Rows 29 and 30*. Internal report, University of Idaho, Kimberly, Idaho. 15.
- Garcia, M. 2008b: *Quality Control of Weather Data for the Middle Rio Grande ET Toolbox*. Internal report, University of Idaho, Kimberly, Idaho. 13.
- Gill, G. C., 1983: *Comparison Testing of Selected Naturally Ventilated Solar Radiation Shields. Final Rpt Contract #NA-82-0A-A-266*. NOAA Data Buoy Off, Bay St. Louis, MS.
- Kaimal, J. C., and Finnigan, J. J., 1994: *Atmospheric boundary layer flows: Their structure and measurement*. Oxford University Press, New York, NY, 289 pp.
- Kizer, M. A., and Elliott, R. L., 1991: Eddy correlation systems for measuring evapotranspiration. *Trans. ASAE*, **34**, 387–392.
- Kolle, O. and Rebmann, C., 2007: *EddySoft-Dokumentation of a Software Package to Acquire and Process Eddy Covariance Data*. Technical Report 10, Max-Planck-Institut für Biogeochemie, Jena, 85 pp.
- Majumdar, N. C., Mathur, B. L., and Kaushik, S. B., 1972: Prediction of direct solar radiation for low atmospheric turbidity. *Solar Energy* **13**, 383–394.
- Meek, D. W., and Hatfield, J. L., 1994: Data quality checking for single station meteorological databases. *Agric. For. Meteorol.* **69**, 85–109.
- Moore, C. J., 1986: Frequency response corrections for eddy correlation systems, *Bound. -Lay. Meteorol.*, **37**, 17–35.
- Mauder, M., and Foken, T., 2004: *Documentation and Instruction Manual of the Eddy Covariance Software Package TK2*. Universität Bayreuth, Abt. Mikrometeorologie, Print, ISSN 1614–8916, 20Arbeitsergebnisse 26, 44 pp.
- Munger, J. W., and Loeschner, H. W., 2004: *Guidelines for Making Eddy Covariance Flux Measurements* (http://public.ornl.gov/ameriflux/measurement_standards_4.doc).
- Snyder, R. L., Pruitt, W. O., and Dong, A., 1985: An automated weather station network for estimation of evapotranspiration. In *Crop Water Requirements* (ed. by Perrier, A. and Riou C.). Int. Commission Irrigation and Drainage., Versailles, France, pp. 133–142.
- Snyder, R. L., Brown, P. W., Hubbard, K. G., and

- Meyer, S. J., 1996: A guide to automated agricultural weather station networks in North America. In *Advances in Bioclimatology* (ed. by Stanhill, G.), Springer Verlag, New York, NY, pp. 1–61.
- Stanhill, G., 1992: Accuracy of global radiation measurement at unattended automated weather stations. *Agric. For. Meteorol.*, **61**, 151–156.
- Shafer, M. A., Fiebrich, C. A., and Arndt, D. S., 2000: Quality assurance procedures in the Oklahoma Mesonet. *J. Atmos. Ocean. Tech.*, **17**, 474–494.
- Sridhar, V., Elliott, R. L., Chen, F., and Brotzge, J. A., 2002: Validation of the NOAA-OSU land surface model using surface flux measurements in Oklahoma. *J. Geophys. Res.* **107**(D20), ACL 3-1:3-18.
- Sun, J., Esbensen, S. K., and Mahrt, L., 1995: Estimation of surface heat flux. *J. Atmos. Sci.* **52**, 3162–3171.
- Tanner, C. B., and Thurtell, G. W., 1969: Anemoclinometer measurements of Reynolds stress and heat transport in the atmospheric surface layer. *Final Report*, United States Army Electronic Command, Atmospheric Sciences Laboratory, Fort Huachuca, Arizona.
- Tanner, B. D., Swiatek, E., and Greene, J. P., 1993: Density fluctuations and use of the krypton hygrometer in surface flux measurements. In *Management of Irrigation and Drainage Systems* (ed. by Allen, R. G.). American Society of Civil Engineers, New York, NY, pp. 945–952.
- Twine, T. E., Kustas, W. P., Norman, J. M., Cook, D. R., Houser, P. R., Meyers, T. P., Prueger, J. H., Starks, P. J., and Wesely, M. L., 2000: Correcting eddy-covariance flux underestimates over a grassland. *Agric. For. Meteorol.* **103**, 279–300.
- Vickers, D., and Mahrt, L., 1997: Quality control and flux sampling problems for tower and aircraft data. *J. Atmos. Ocean. Tech.*, **14**, 512–526.
- Vickers, D., and Mahrt, L., 2003: The Cospectral Gap and Turbulent Flux Calculations. *J. Atmos. Ocean. Tech.*, **20**, 660–672.
- Webb, E. K., Pearman, G. I., and Leuning, R., 1980: Correction of flux measurements for density effects due to heat and water vapor transfer. *Quart. J. Roy. Meteorol. Soc.* **106**, 85–106.
- Wilson, K., Goldstein, A., Falge, E., Aubinet, M., Baldocchi, D., Berbigier, P., Bernhofer, C., Ceulemans, R., Dolman, H., Field, C., Grelle, A., Ibrom, A., Law, B. E., Kowalski, A., Meyers, T., Moncrieff, J., Monson, R., Oechel, W., Tenhunen, J., Valentini, R., and Verma, S. 2002: Energy balance closure at FLUXNET sites. *Agric. For. Meteorol.*, **113**, 223–243.
- Wilczak, J. M., Oncley, S. P., and Stage, S. A., 2001: Sonic anemometer tilt correction algorithms. *Bound.-Lay. Meteorol.*, **99**, 127–150.

蒸発量計算における気象データと微気象フラックスの品質評価について

Richard Allen

(アイダホ大学)

要 約

気象データの品質評価と補正は基準蒸発散量を正確に決定する上で重要である。本論文では多量のデータの品質を簡便に評価、選別、補正するための方法として、グラフを使った方法を解説する。たとえば日射においては実測値を晴天日における理論値と比較することにより、実測値の品質管理を行うことができる。渦関法により測定

したフラックスの補正においては、重回帰分析を使用した方法が有効なようである。これはボーエン比を使った補正法に代わる手法として顕熱・潜熱・地中伝導熱フラックスの補正に利用できる。

キーワード：蒸発散、品質管理、気象データ