Comparison of riparian evapotranspiration estimates based on a water balance approach and sap flow measurements

R.D. Mac Nish [a,*], C.L. Unkrich [b], Evelyn Smythe [a],
D.C. Goodrich [b], Thomas Maddock III [a]

[a] *Department of Hydrology and Water Resources, University of Arizona, Tucson, AZ 85721, USA*
[b] *USDA-ARS Southwest Watershed Research Center, Tucson, AZ, USA*

Abstract

Estimates of evapotranspiration (ET) from riparian vegetation along a 122 m reach of the San Pedro River using both a water balance approach and by scaling up sap flow measurements are compared. A sensitivity analysis was performed on the components of the water balance to assess the effects of measurement errors on estimates of ET using this method. It was concluded that by reducing the error in three key components to less than 5%, riparian ET could be estimated to an accuracy of 20–25% using the water balance method. The analysis also indicated that random measurement errors up to 10% in the water balance measurements would explain the difference between the water balance and sap flow ET estimates. Demonstrating agreement given reasonable error bounds provides confidence in the accuracy of both methods. © 2000 Elsevier Science B.V. All rights reserved.

Keywords: Evapotranspiration; Water balance; Riparian; Sap flow; Groundwater; San Pedro River

1. Introduction

This paper presents a preliminary attempt to estimate evapotranspiration (ET) from riparian vegetation along a 122 m reach of the San Pedro River in Cochise County, Arizona (Fig. 1). ET was estimated over the entire study area using both a water balance approach, based on measurements of streamflow and groundwater head distributions, and by scaling up sap flow measurements of individual trees. These measurements were acquired during three 32–48 h "synoptic" periods of intensive data collection in March, April and June 1997 representing different degrees of aquifer stress from transpiration. A sensitivity analysis was performed on the components of the water balance to assess the effects of measurement errors on esti-

mates of ET using this approach. The analysis also indicated the degree of component error necessary to explain the difference between the water balance and sap flow ET estimates. This work is an important step towards developing a better understanding of stream–aquifer–vegetation interactions, improving our ability to simulate those systems and our ability to scale these relationships in time and space.

The San Pedro River valley is comprised of alluvial materials deposited by streams emerging from the surrounding mountains between the middle Pliocene and the middle Pleistocene time (Huckleberry, 1996) before the San Pedro River became established as a through-flowing drainage. The Lewis Springs Research Site (LSRS) is underlain by a relatively shallow floodplain aquifer deposited by the river after first carving a valley into the earlier alluvial deposits. Approximately 165 m wide and ranging in thickness

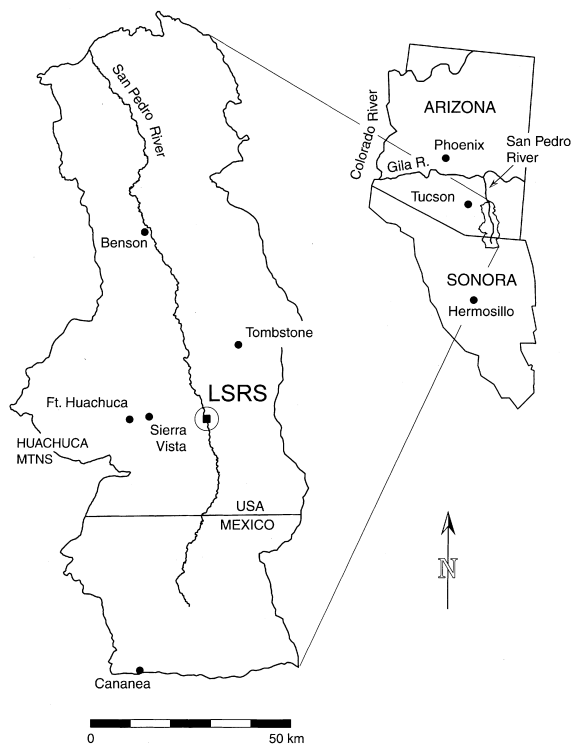---

* Corresponding author.

Fig. 1. Location of the Lewis Springs Research Site (LSRS) in the upper San Pedro River Basin in Cochise County, Arizona.

up to about 10 m, the floodplain aquifer feathers out against the underlying regional aquifer at its margins. Water in the regional aquifer moves toward the river from the bordering mountains, discharging upward where it meets the floodplain aquifer. Like the river, water in the floodplain aquifer flows south to north down the valley, converging slightly toward the river, causing some discharge into the stream (baseflow).

2. Data acquisition

2.1. Water balance

River flow measurements were carried out at five points distributed along the study reach. These included water level, current meter discharge measurements and sampling for dye tracer dilution analysis during the synoptic periods. Difficulties in analyzing data from the dye tracer dilution technique included

temperature-induced variation in dye injection rates and diurnal fluctuations of water level through the study reach. The presence of a large pool between the third and fourth sampling points created intractable problems with the analysis and as a result only the gains between the first and third (X1 and X3 in Fig. 2) were estimated (Table 1). Average flow rates for the March and April synoptic periods were 212 and 127 l/s, respectively. During the June study, a portable steel flume was installed near X1 to obtain accurate discharge measurements during this period of very low flow (25 l/s).

Groundwater data were collected on an hourly basis during the synoptic periods at 25 piezometers, and continuously recorded at six wells and one piezometer. The hourly measurements were made using electronic sounders accurate to within 3 mm and the continuous measurements were recorded with bubble gages also accurate to 3 mm. The details of piezometer and well construction are presented by Mac Nish et al. (1998). The piezometer network consisted of three transects roughly perpendicular to the stream, each transect having four clusters of piezometers, two on each side of the stream, with one near the channel and one about 50 m away (Fig. 2). In each cluster, one piezometer was finished just below the depth to groundwater, and one or two additional piezometers were finished at 1–1.5 m intervals below the shallowest piezometer, all within the floodplain aquifer. Depths ranged from 3.05 to 7.32 m. The clusters were identified by their location relative to the stream (west or east), their location along the study reach (north, middle, or south) and their proximity to the stream (close or far). Each piezometer within a cluster was identified by its approximate depth in feet. Thus, the WMC 14 piezometer was located on the west bank in the middle transect, close to the stream, and was 14 ft deep. Piezometer coordinates accurate to within several centimeters were obtained using an electronic total station, and level surveys were conducted to determine relative elevations to within 3 mm.

Groundwater gradients and overall groundwater level changes were computed using the average groundwater level measured in each piezometer over a 24 h interval within each synoptic period. Horizontal gradients were computed as the average gradient between the middle and south clusters, using the shallowest piezometer in each cluster, paired as shown
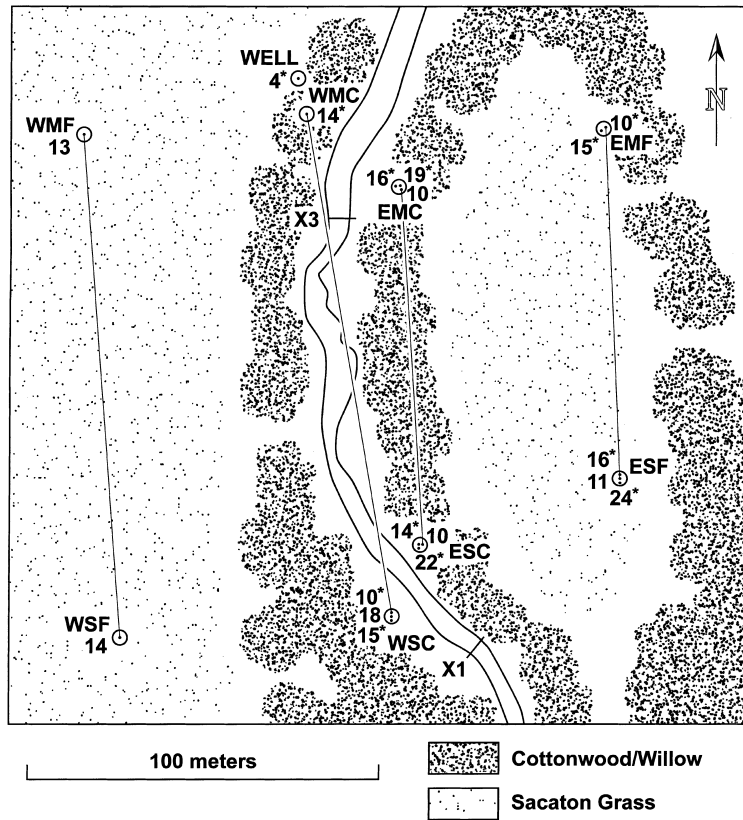
Fig. 2. Map showing the locations of piezometers used to compute horizontal and vertical gradients. Piezometer clusters are identified by their location relative to the stream (west or east), their location along the study reach (north, middle, or south) and their proximity to the stream (close or far). Each piezometer within a cluster is identified by its approximate depth in feet. Piezometers used to compute horizontal gradients are connected by lines, and those used to compute the vertical gradient at a given cluster are indicated by an asterisk. Locations where river flow measurements were taken are indicated by X1 and X2.

Table 1
Estimated values for groundwater, surface water and sap flow components used in the water balance equations applied between X1 and X3 during the synoptic periods

Component	March	April	June
Horizontal gradient	0.0021	0.0021	0.0019
Vertical gradient	0.0076	0.0019	0.0071
Average groundwater level change (m)	0.007	0.009	0.007
Gain (l/s)	4.5	0.8	2.4
ET from sap flow (l/s)	0.0	0.4	0.6
Saturated thickness at X1 (m)	4.3	4.2	4.0
Saturated thickness at X3 (m)	3.7	3.6	3.5
Average saturated width (m)	175	165	155
Aquifer length (X1–X3) (m)	122	122	122
Saturated cross sectional area at X1 (m^2)	743	686	620
Saturated cross sectional area at X3 (m^2)	652	601	543
Saturated plan area of aquifer (m^2)	21336	20117	18898

in Fig. 2. In most cases the vertical gradients were computed using the two deepest piezometers in each cluster (in the case of WMC 14, a nearby, deeper observation well was used). Since these piezometers were finished near the bottom of the floodplain aquifer, the computed vertical flux should be approximately equal to the flux across its interface with the regional aquifer. The average vertical gradient during the April synoptic was substantially lower than in March or June (Table 1), probably due to recharge into the floodplain aquifer from elevated river levels caused by a storm several days earlier.

2.2. ET from sap flow measurements

During the synoptic periods sap flow velocity was measured at selected clusters of cottonwood and willow trees in the LSRS using heat-pulse probes logging at 30 min intervals (Schaeffer et al., 2000). Using core sample data, regression relationships between sap wood area (SWA) and the trunk diameter at breast height (DBH) were developed. The regression equations were combined with measurements of DBH for the 533 cottonwood and willow trees along the corridor between X1 and X3 to convert sap flow velocities (v_{sap}) into total volumetric ET rates leaving the corridor between X1 and X3:

$$ET = v_{sap} \cdot \sum_{X1}^{X3} SWA \tag{1}$$

These rates at 30 min intervals were then averaged over a 24 h period, giving average rates of 0.4 l/s for April and 0.6 l/s for June. During the March synoptic, sap flow velocities were zero.

The error in each estimate of ET at a given 30 min interval was assumed to be both independent and random, and was computed as

$$\varepsilon_{ET} = \sqrt{\left(\frac{\partial ET}{\partial v_{sap}}\varepsilon_v\right)^2 + \left(\frac{\partial ET}{\partial SWA_T}\varepsilon_{SWA}\right)^2} \tag{2}$$

$$\varepsilon_{ET} = \sqrt{(SWA_T \cdot \varepsilon_v)^2 + (v_{sap} \cdot \varepsilon_{SWA_T})^2} \tag{3}$$

where ε_v is the standard error of estimate from statistical analysis of the velocity data,

$$SWA_T = \sum_{X1}^{X3} SWA \tag{4}$$

is the total sap wood area and

$$\varepsilon_{SWA_T} = \sqrt{\sum_{X1}^{X3} \varepsilon_{SWA}^2} \tag{5}$$

is the error in the total sap wood area, where ε_{SWA} is the standard error from the DBH versus SWA regression. Adding and subtracting these errors yielded a range of average ET rates: 0.3–0.5 l/s for April and 0.5–0.7 l/s for June. Details on the measurement of streamflow and the estimation of riparian transpiration from sap flow measurements may be found in Mac Nish et al. (1998) and Williams et al. (1998), respectively.

3. Analysis

3.1. Water budget for the floodplain aquifer

The section of floodplain aquifer between points X1 and X3 was chosen as the control volume (CV) for the water balance analysis (Fig. 3). Hydraulic properties within the CV are assumed to be constant and show only small-scale, random variation, which is consistent with the extremely small size of the control volume, from a groundwater perspective. Therefore, we assume that hydraulic conductivity and specific yield can be adequately represented by average values. Because the analysis is based only on flow across the boundaries of the CV, there is no need to account for gradients within the CV itself.

We define R to be the rate (l/s) of upward flow into the CV from the underlying regional. U and D are defined as the rate of flow (l/s) through the upstream and downstream cross sections of the CV. Within the CV, some water escapes to the stream as a gain G (l/s) in flow between X1 and X3, leaving the CV as streamflow. Estimates for gain were obtained from the dye dilution analysis and streamflow measurements. Through all three synoptic periods groundwater levels declined, as the CV contributed to streamflow and ET at the expense of water in storage, S. In April and June an additional component leaves as ET from the phreatophytic cottonwood and willow trees. Fig. 4
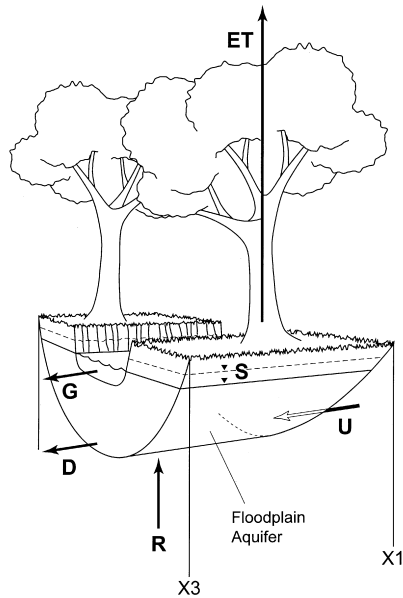
Fig. 3. Simplified schematic of the water budget for the floodplain aquifer system. ET: evapotranspiration of cottonwood and willow trees; R: inflow from the regional aquifer; U: upstream inflow within the floodplain aquifer across X1; D: downstream outflow within the floodplain aquifer across X3; G: streamflow gain between X1 and X3; S: water released from storage as water levels drop within the floodplain aquifer.

shows groundwater level changes at the EMF cluster during the three synoptic periods. While the March data shows a slight but steady decline, the traces for April and June show the diurnal signal caused by transpiration of the phreatophytes.

The water balance equation for each synoptic period can be written as

$$\varepsilon_r = U + R + S - D - G - \mathrm{ET} \tag{6}$$

where ε_r is the residual error of closure. Thus, ET can be computed as

$$\mathrm{ET} = U + R + S - D - G \tag{7}$$

if the quantities on the RHS are known. In the following analysis, the hydraulic conductivity, which appears in U, R and D, is also considered to be an unknown. We take advantage of the fact that ET was zero during March, which removes ET as an unknown for that synoptic period and allows computation of the hydraulic conductivity, which is then used in Eq. (7) for April and June.

The horizontal groundwater flow components U and D can be expressed as

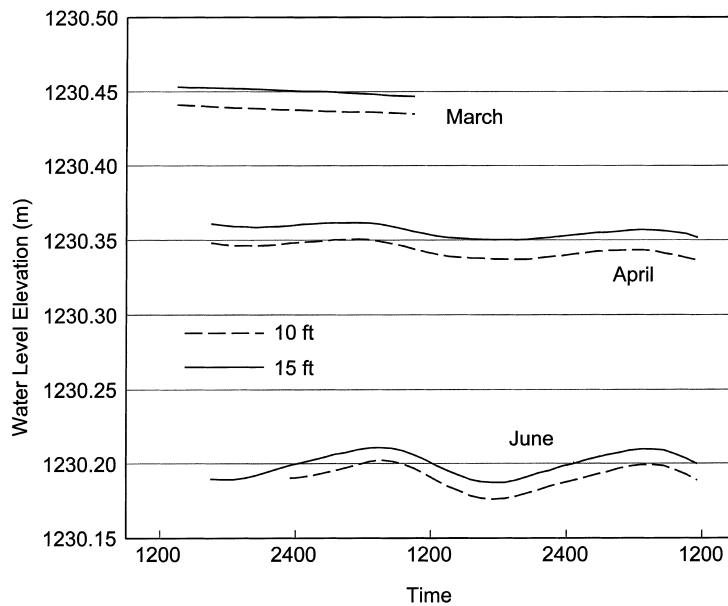$$U = grad_H \cdot area_X U \cdot K_S \cdot r_{H:V} \tag{8}$$

Fig. 4. Water level elevation changes during and between March, April, and June synoptic studies in the 10 and 15 ft deep piezometers of the east, middle and far clusters at the LSRS.

$$D = grad_H \cdot area_X D \cdot K_S \cdot r_{H:V} \qquad (9)$$

where $grad_H$ is the horizontal groundwater gradient, and $area_X U$, $area_X D$ are the saturated cross sectional areas of the aquifer at X1 and X3. Due to alternating layers of coarse and fine sediments, alluvial aquifers as a rule show pronounced anisotropy, where horizontal conductivity is much greater than vertical conductivity. Here, K_s is the *vertical* saturated hydraulic conductivity, assumed to be the same at both cross sections, and $r_{H:V}$ is the ratio of horizontal to vertical hydraulic conductivity (assumed to be approximately 10, after Freeze and Cherry (1979)). The cross sectional areas were approximated by the average saturated aquifer width multiplied by the average saturated thickness at X1 and X3 for each synoptic period (Table 1). Saturated width and thicknesses were estimated using information on aquifer thickness obtained from the piezometer bore holes and from average groundwater levels measured during the synoptic periods.

Vertical inflow from the regional aquifer can be expressed as

$$R = grad_V \cdot area_P \cdot K_S \qquad (10)$$

where $grad_V$ is the vertical groundwater gradient near the bottom of the floodplain aquifer and $area_P$ is the saturated plan area of the aquifer. The saturated area was approximated by the measured distance between X1 and X3 multiplied by the average saturated width.

The change in storage was computed as

$$S = \frac{\Delta h \cdot area_P \cdot s}{\Delta t} \qquad (11)$$

where Δh is the average groundwater level change over a 24 h interval ($\Delta t = 86\,400\,s$) during each synoptic period (Table 1) and s is the specific yield, estimated to be approximately 0.2 (Freeze and Cherry, 1979).

Two variables which remain unknown are K_s and ET. Since ET was zero during the March synoptic, K_s can be computed directly from the water balance equation for March:

$$K_s = \frac{G_{Mar} - [\Delta h \cdot area_P]_{Mar} \cdot (s/\Delta t)}{grad_H \cdot r_{H:V}[area_X U - area_X D]_{Mar} + [grad_V \cdot area_P]_{Mar}} \qquad (12)$$

The value of K_s (vertical conductivity) computed from Eq. (12) is 0.025 mm/s. The corresponding horizontal conductivity value of 0.25 mm/s is somewhat high for silty sand aquifers (Freeze and Cherry, 1979), but given the presence of sand and gravel stringers, particularly near the stream, it is a defensible value in this case.

Assuming conductivity is constant between synoptic periods, the K_s value can then be used with the average gradients and cross sectional areas for the April and June synoptic periods to compute values of average ET during those periods, where

$$\begin{aligned}
ET_{Apr,Jun} \\
= K_s\{grad_H \cdot r_{H:V}[area_X U - area_X D]_{Apr,Jun} \\
+ [grad_V \cdot area_P]_{Apr,Jun}\} + \frac{s}{\Delta t}[\Delta h \cdot area_P]_{Apr,Jun} \\
- G_{Apr,Jun}
\end{aligned} \qquad (13)$$

The resulting ET rates for April and June are 0.6 and 1.3 l/s, respectively. Neither ET values were within the assumed error range of the sap flow estimates (0.3–0.5 l/s for April and 0.5–0.7 l/s for June). This discrepancy motivated a sensitivity analysis which became the basis for an independent assessment of error in the sap flow ET estimates and for estimating errors in the ground and surface water components of the water balance.

3.2. Error and sensitivity analysis

The maximum error in computing average horizontal gradients was estimated to be less than 2% based on the horizontal and vertical precision of the piezometer surveys and the average error of groundwater level measurements during the synoptic studies. Consideration of the previously discussed problems with the dye dilution measurements suggested an error of about 10% in the streamflow gain. The estimation error or uncertainty in the remaining water balance components was impossible to ascertain directly. However, if both the water balance and sap flow estimates are sufficiently close to the true values, then for reasonable error bounds on the water balance and sap flow measurements the resulting ranges of ET estimates should overlap. Therefore, to quantify the error in the water balance components, the error bounds on the water balance measurements were expanded until the resulting range of ET values overlapped those from the sap flow measurements. To obtain a

more conservative estimate of the error it was further required that the range of ET values from the water balance include the sap flow estimates themselves rather than just overlap the sap flow error range.

This was implemented in an iterative computer algorithm, using the error estimates for horizontal gradients and streamflow gain, assuming errors at first up to 5%, and then up to 10%, in the remaining variables and then adjusting them within these error ranges until the resulting ET values exactly matched the values estimated from the sap flow measurements. Vertical conductivity K_s was also constrained to values less than or equal to 0.02 mm/s as a conservative upper bound for the type of aquifer material. Since the errors were assumed to be random rather than systematic, most variables were adjusted independently not only within a given synoptic period but also between synoptic periods, i.e. the gain in March was varied independently from the gain in April. Variables were adjusted by applying a separate multiplier ρ to each variable, where each multiplier varies independently in three discrete steps. As an example, for $\pm 10\%$, ρ would vary among the three values $\{0.9, 1.0, 1.1\}$.

A general sensitivity analysis based on a 10% error range was then carried out by holding a given variable constant while allowing the others to vary indepen-dently as described above. Sensitivity was assessed using the root mean squared error (RMSE) of the computed ET values relative to the original values for both April and June, summed over all ($\sim$150 billion) possible multiplier combinations. To help quantify the benefit of striving for more accurate measurements, the drop in RMSE was examined while the error range of individual variables was reduced from 10 to 5%, starting with the most sensitive and proceeding to the least sensitive variable. The procedure was then repeated, this time reducing the range from 10 to 2%.

The adjusted horizontal to vertical conductivity ratio $r'_{H:V}$ and specific s' yield do not vary between synoptic periods:

$$r'_{H:V} = r_{H:V} \cdot \rho_1 \tag{14}$$

$$s' = s \cdot \rho_2 \tag{15}$$

The adjusted K_s was computed using the March (ET = 0) water balance

$$K'_s = \frac{\rho_3 \cdot G_{Mar} - \rho_4 \cdot \Delta h_{Mar} \cdot \rho_5 area_{P_{Mar}} \cdot (s'/\Delta t)}{\rho_6 \cdot grad_H \cdot r'_{H:V} [\rho_7 \cdot area_X U - \rho_8 \cdot area_X D]_{Mar}} \\ + \rho_9 \cdot grad_V \cdot \rho_5 \cdot area_P \tag{16}$$
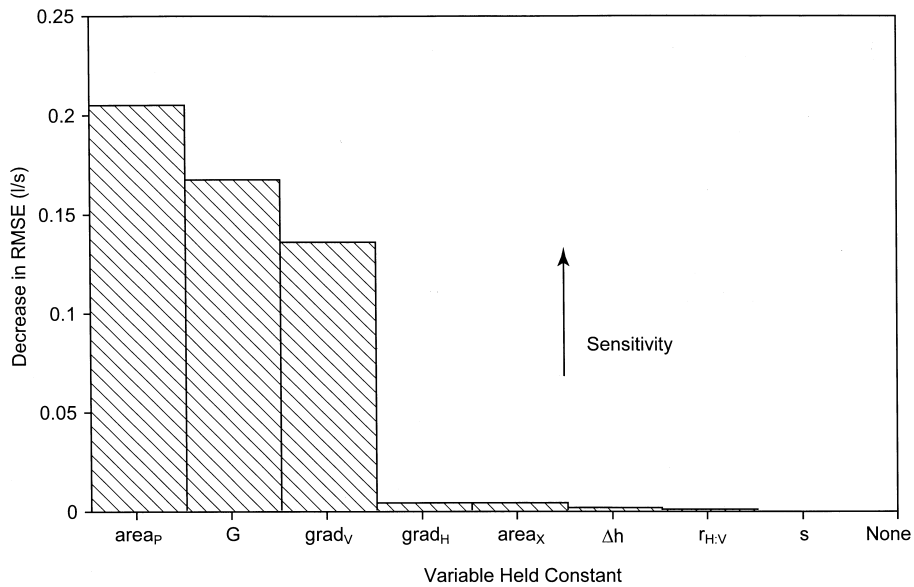

Fig. 5. Sensitivity of water balance estimates of ET to measurement uncertainty in individual components as indicated by the drop in RMSE when a given variable is held constant.
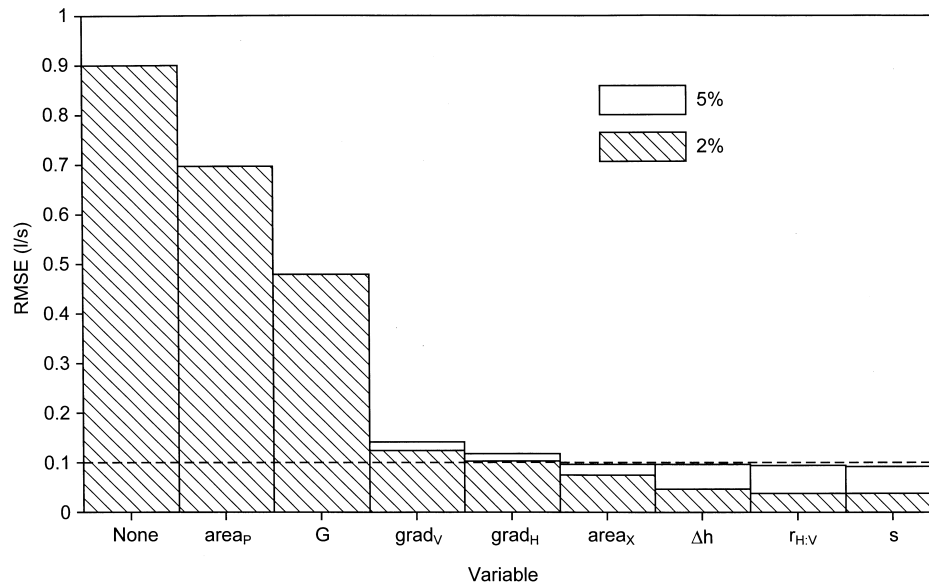
Fig. 6. Effect of successively reducing measurement uncertainty in individual water balance components, from 10 to 5% and from 10 to 2%, on uncertainty in ET estimates (none = all variables at 10%). For area$_P$ and G, the difference between 5 and 2% is indistinguishable.

Using the adjusted K_s computed from March, the adjusted ET values for April and June were computed as

$$
\begin{aligned}
\mathrm{ET'_{Apr,Jun}} =\ & K'_s\{\rho_{10} \cdot grad_H \cdot r'_{H:V}[\rho_{11} \cdot area_X U \\
& -\rho_{12} \cdot area_X D]_{Apr,Jun} + \rho_{13} \cdot grad_{V_{Apr,Jun}} \\
& \cdot \rho_{14} \cdot area_{P_{Apr,Jun}}\} \\
& +\frac{s'}{\Delta t}\rho_{15} \cdot \Delta h_{Apr,Jun} \cdot \rho_{14} \cdot area_{P_{Apr,Jun}} \\
& -\rho_{16} \cdot G_{Apr,Jun}
\end{aligned}
\tag{17}
$$

Adjustments up to 5% did not allow convergence to the sap flow-derived ET values, but adjusting up to 10% did, suggesting that errors were probably greater than 5% and likely in the neighborhood of 10%, which is quite remarkable given the rough estimation of some of the variables in question. This demonstrates basic agreement between the water balance and sap flow ET estimates, given a reasonable magnitude of error in the water balance, and reinforces the validity of the ET estimates from each approach.

The degree of sensitivity is indicated by the decrease in variability (RMSE) when a given variable is held constant. The sensitivity of ET estimates to deviations up to 10% is shown in Fig. 5. The results of

successively dropping the deviations from 10 to 5% and from 10 to 2% are shown in Fig. 6. As we believe the error in horizontal gradients is already less than 2%, Fig. 6 suggests that reducing the error in aquifer area, vertical gradient and streamflow gain from 10 to 5% would reduce the uncertainty in ET estimates from about 0.9 to 0.1 l/s.

4. Conclusions

To indirectly assess the degree of error in the surface and groundwater measurements, their values were adjusted independently until the resulting range of ET values computed from the water balance included those from the sap flow estimates. The degree of variation required was 10%, which is equivalent to agreement between the two methods given random errors up to 10% in the water balance measurements. Demonstrating agreement by assuming a believable range of error provides confidence in the accuracy of both ET estimates. The sensitivity of the water balance to error in its various components was also explored, and it was concluded that by reducing the error in aquifer

area, streamflow gain and vertical groundwater gradient to less than 5%, riparian ET within the LSRS could be estimated to an accuracy of about 0.1 l/s or 20–25% using the water balance approach.

A detailed survey of the aquifer boundaries using accurate equipment such as an electronic total station or differential GPS would improve our estimation of the aquifer extent. Because of the small area involved, positional accuracy is important, as well as a sufficient density of survey points to capture any major irregularities when outlining the boundaries. A more detailed description of the transverse aquifer profiles at X1 and X3 would allow more accurate estimation of the cross sectional saturated area available for through-flow. At the LSRS, the regional aquifer is substantially higher in clay content than the floodplain aquifer, lending hope that measurements could be done with shallow seismic profiling or electrical resistivity equipment. Such data might also shed some light on the uniformity of material composing the floodplain aquifer, and thus on the uniformity of hydraulic conductivity.

The sensitivity analysis performed on the water balance components indicates that for estimating horizontal groundwater gradients, the existing piezometer network and measurement techniques are probably sufficient. However, decreasing the uncertainty in vertical gradient calls for more precise measurement of groundwater levels since the difference in groundwater level occurs over the relatively small differences in well depth. Ideally, high-accuracy depth sensors should replace the manual measurement methods used for this study. Additionally, the deeper piezometers in each cluster should be rehabilitated or reconstructed to better isolate the openings from overlying saturated materials.

Unfortunately, dye injection techniques were apparently not suitable for obtaining accurate estimates of streamflow gain due to temporary dye storage problems in pools and hyporheic flow zones (Bencala and Walters, 1983). Flumes or weirs equipped with recording devices are required for more accurate measurement of streamflow at the upstream and downstream boundaries.

Acknowledgements

References

Bencala, K.E., Walters, R.A., 1983. Simulation of solute transport in a mountain pool-n-riffle stream: a transient storage model. Water Resources Res. 19 (3), 718–724.

Freeze, R.A., Cherry, J.A., 1979. Groundwater. Prentice-Hall, Englewood Cliffs, NJ, 604 pp.

Huckleberry, G., 1996. Historical Channel Changes on the San Pedro River, Southeastern Arizona. Arizona Geological Survey Open-File Report 96-15, 21 pp.

Mac Nish, R.D., Peters, C.J., Schulte, M.A., Goodrich, D.C., Pool, D.R., Maddock III, T., Unkrich, C.L., Whitaker, M.P.L., Goff, B.F., 1998. Quantification of groundwater–surface water interactions in a southwestern riparian system. In: Special Symposium on Hydrology, Session 1: Integrated Observations of Semi-Arid Land-Surface–Atmosphere Interactions, Paper-P2.12, American Meteorological Society, Phoenix, Arizona, 11–16 January 1998.

Schaeffer, S.M., Williams, D.G., Goodrich, D.C., 2000. Transpiration of cottonwood/willow forest estimated from sap flux. Agric. For. Meteorol. 105, 257–270.

Williams, D.G., Brunel, J.-P., Schaeffer, S.M., Snyder, K.A., 1998. Biotic controls over the functioning of desert riparian ecosystems. In: Special Symposium on Hydrology, Session 1: Integrated Observations of Semi-Arid Land-Surface–Atmosphere Interactions, Paper-P1.11, American Meteorological Society, Phoenix, Arizona, 11–16 January 1998.