A NOTE ON SAMPLING ERRORS IN THE RAINFALL AND RUNOFF DATA COLLECTED USING TIPPING BUCKET TECHNOLOGY

B. Yu, C. A. A. Ciesiolka, C. W. Rose, K. J. Coughlan

ABSTRACT. Data precision for rainfall intensity and runoff rate using tipping bucket technology is controlled by the bucket volume, V , the catchment area, A , and the sampling interval, t . The maximum absolute error is less than $V(A\Delta t)^{-1}$ and the standard error is given by $V(\sqrt{6} A\Delta t)^{-1}$, both of which are independent of the magnitude of rainfall intensity and runoff rate being measured. Therefore, the relative error can be quite large during events with low rainfall intensity or runoff rate. The magnitude of the sampling error given above can be used to assist in experiment and equipment design and in data analysis. **Keywords.** Rainfall, Runoff, Erosion, Tipping bucket, Sampling error.

Tipping bucket technology is commonly used to collect ground-level rainfall intensity data at short time intervals (WMO, 1961). For small runoff plots with relatively low runoff rates, tipping buckets are also used so that rainfall intensity and runoff rate are measured simultaneously (e.g., Pillsbury et al., 1962; Edwards et al., 1974; Barfield and Hirschi, 1986; Williams and Bonell, 1987; Ciesiolka et al., 1995a). Thus, compatible and synchronous data on rainfall intensity and runoff rate have now become available for developing infiltration-time relationships for plots and small watersheds (Amerman, 1983; Yu et al., 1997). Similarly, effectiveness of various agronomic management practices in terms of reduced runoff and soil losses can be assessed using such data with hydrologic and soil erosion models.

We are not aware, however, of other studies that have explicitly addressed the question of how to determine the magnitude and frequency distribution of the sampling errors associated with the tipping bucket technology. Analysis of sampling errors can provide useful guide for experiment and equipment design. In the context of data analysis, especially for model development and evaluation purposes, the goodness-of-fit of any model cannot be independently assessed without adequate consideration of the nature and the magnitude of measurement errors.

In this article, we derive an analytic expression for sampling errors associated with a constant rainfall intensity or runoff rate. The analytic results are then generalized to include real precipitation events when both rainfall intensity and runoff rate are highly variable. Finally, we consider the implications for experiment and equipment design and assessing model performances. In this article we concentrate on the measurement errors which are brought about by the

discrete nature of the sampling method associated with tipping bucket technology. Other types of errors in measured rainfall intensity or amount, due to wind, shelters around the rain gauge, spillage, excess tips recorded during electrical storms, and methods of correction have been extensively discussed elsewhere (Legates and Willmott, 1990; Hanna, 1995; Ciesiolka et al., 1995a).

DEFINITION OF THE PROBLEM

A tipping bucket is an over-centre device which tips after accepting a certain volume of water, V . The number of tips for a given time interval can be recorded on electronic data loggers. The observed rate of rainfall or runoff, i_o , during any time interval, Δt , is given by:

$$i_o = n_{\Delta t} \frac{V}{A \Delta t} \quad (1)$$

where $n_{\Delta t}$ is the total number of tips recorded during the time interval Δt , and A is the cross-sectional area of the pluviometer, or the runoff catchment area in the case of measuring runoff rates. It can be seen that the recorded rate of rainfall or runoff can only be multiples of a base rate, E_o , which is defined as:

$$E_o = \frac{V}{A \Delta t}$$

The ratio, V/A , can be regarded as an effective depth increment for measuring rainfall intensity or runoff rate.

Figure 1 shows recorded rainfall intensity and runoff rate during a storm event on 31 October 1993 at Goomboorian near Gympie, Australia. Total rainfall and runoff were 18.6 and 7.4 mm, respectively. Data were recorded at 1-min intervals. The bucket size for rainfall measurement was 7.061 mL with a 300 mm diameter rain gauge. The bucket size and catchment area for the runoff plots are given in table 2 and Ciesiolka et al. (1995b) described the physical characteristics of the site. It can be seen from figure 1 that both rainfall intensity and runoff

---

Article has been reviewed and approved for publication by the Soil & Water Div. of ASAE.

The authors are Bofu Yu, and Calvin W. Rose, Faculty of Environmental Sciences, Griffith University, Nathan, Queensland, Australia; Cyril A. A. Ciesiolka, Queensland Department of Primary Industries, Toowoomba, Queensland, Australia; and Kep J. Coughlan, Australian Centre for International Agricultural Research, Canberra, Australia. Corresponding author: Bofu Yu, Faculty of Environmental Sciences, Griffith University, Nathan, Queensland 4111, Australia; tel.: 617-3875-5258; fax: 617-3875-7459; e-mail: <b.yu@ens.gu.edu.au>.

Figure 1—Rainfall intensity and runoff rate were recorded at 1-min intervals during a storm event at Goomboorian near Gympie, Australia (see Ciesiolka et al., 1995b, for details of this site).

Table 1. Observed number of tips for each time interval and equivalent rate for a hypothetical event with a rate of 25 mm/h

Time (min)	1	2	3	4	5	6	7	8	9	10	11	12	...
No. of tips	4	4	4	4	4	5	4	4	4	4	4	4	5
Observed rate (mm/h)	24	24	24	24	24	30	24	24	24	24	24	24	30
Error (mm/h)	-1	-1	-1	-1	-1	+5	-1	-1	-1	-1	-1	-1	+5

NOTE: The effective depth increment (i.e., bucket volume to catchment area ratio) is 0.1 mm and the sampling interval is 1 min.

rate fluctuate at high tipping frequencies, and such high-frequency oscillations become more apparent when the rate is low, e.g., see the periods between 100 and 150 min for runoff and 150 and 200 min for rainfall. Questions that arise from this data are:

- Do actual rainfall intensity and runoff rate indeed vary in the manner illustrated?
- How much of this observed variation in rainfall intensity and runoff rate is due to the sampling errors?
- What are the magnitude and frequency distribution of these errors?
- Under what conditions can the effects of sampling errors be safely ignored?

The problem associated with discrete data in terms of the number of tips per unit time interval becomes glaringly apparent when one considers sampling a hypothetical event with a constant rate. As an example, consider an event with a constant rate of 25 mm/h. If the bucket size to catchment area ratio is 0.1 mm and the sampling interval is 1 min, the observed number of tips would always be a sequence of five 4s followed by a single 5 (table 1). The corresponding error would be -1 mm/h and 5 mm/h, respectively. The actual rate, held constant at 5, 15, and 25 mm/h, respectively, and the corresponding rates that would have been recorded are shown graphically in figure 2. It can be seen that the error as a result of the sampling technique can be considerable and that the error is not necessarily distributed symmetrically around the actual rate. Three aspects of these sampling errors are worth noting with respect to this hypothetical event with a constant rate. First, the errors occur in a regular and periodic pattern. Secondly, there is no systematic error. Thirdly, errors of greater magnitude, 5 mm/h in this instance, occur less frequently by comparison with errors of lesser magnitude. This article develops a theory to explain these observations and derives

Figure 2—Observed rate with the actual rate held constant at 5, 15, and 25 mm/h, respectively. The sampling interval is 1-min and the V/A ratio is 0.1 mm.

an analytic expression for the sampling error and its distribution in the case of a constant rainfall intensity or runoff rate. The results are then generalized to quantify the sampling errors during a real precipitation event with variable rainfall intensity and runoff rate.

RESULTS

For a constant rate, i , the number of tips up to time t , n_t , is $\text{int}(iAt/V)$ where $\text{int}(\cdot)$ denotes the integer part of a real number. It follows from equation 1 that the observed rainfall intensity, i_o , for the time interval from t to $t + \Delta t$ is given by:

$$i_o = n_{\Delta t} E_o \\ = [\text{int}(u + \Delta u) - \text{int}(u)] E_o$$

where u and Δu are dimensionless continuous variables given by:

$$u = \frac{i A t}{V} \text{ and } \Delta u = \frac{i A \Delta t}{V}$$

Since $\text{int}(u + \Delta u + 1) - \text{int}(u + 1) = \text{int}(u + \Delta u) - \text{int}(u) = n_{\Delta t}$, $n_{\Delta t}$ is a periodic function of u with 1 being the fundamental period. Therefore, it is sufficient to consider $n_{\Delta t}$ with respect to u in the range from 0 to 1 only. In particular, we have:

$$n_{\Delta t} = \text{int}(\Delta u) \text{ when } u = 0 \text{ or } u = 1$$

Let Δu be written in the form:

$$\Delta u = m + \beta$$

with $0 < \beta < 1$ and m is an integer. Note that the actual rate i equals $(m + \beta)E_o$. The number of tips in the time interval, Δt , then becomes:

$$\begin{aligned}
n_{\Delta t} &= \text{int}(u + m + \beta) - \text{int}(u) \\
&= m + \text{int}(u + \beta) - \text{int}(u) \\
&= \begin{cases} m & u < 1 - \beta \\ m + 1 & u \geq 1 - \beta \end{cases}
\end{aligned}$$

Error in the observed rate is by definition:

$$\begin{aligned}
e &= i_o - i \\
&= (n_{\Delta t} - m - \beta)E_o \\
&= \begin{cases} -\beta E_o & u < 1 - \beta \\ (1 - \beta) E_o & u \geq 1 - \beta \end{cases}
\end{aligned}$$

It can be seen that the magnitude of the sampling error is bounded by E_o , i.e.:

$$\text{Max } |e| < E_o \quad (2)$$

since $0 < \beta < 1$. For given bucket volume, V , pluviometer or runoff catchment area, A , and sampling interval, Δt , E_o would be fixed while $u = \text{int}(u)$ varies periodically between 0 and 1. Thus a sampling error with a magnitude of $-\beta E_o$ would occur for $(1 - \beta) \times 100\%$ of time, while another sampling error with a magnitude of $(1 - \beta)E_o$ would occur for $\beta \times 100\%$ of time. The magnitude and frequency of occurrence of the sampling error are therefore complimentary. Large errors occur less frequently. The sampling error vanishes when and only when $\beta = 0$ or $iA\Delta t/V$ is an integer exactly.

It can easily be shown that the mean of the sampling error is zero, indicating that the expression (1) is an unbiased estimator of the actual rate i . The variance of the sampling error is:

$$\begin{aligned}
\text{Var}(e) &= [\beta^2(1 - \beta) + (1 - \beta)^2\beta] E_o^2 \\
&= \beta(1 - \beta) E_o^2 \quad (3)
\end{aligned}$$

Thus the standard error varies between 0 when $\beta = 0$ to $E_o/2$ when $\beta = 1/2$ depending on the actual rate being measured.

During a real precipitation event, both rainfall intensity and runoff rate vary over time. The inequality (eq. 2) still applies and the estimator (eq. 1) remains unbiased. β , however, can assume any value in the range between 0 and 1 depending on the actual rate. If a wide range of rainfall intensity and runoff rate is experienced during a real event, it is reasonable to assume that β is independently and uniformly distributed in the range from 0 to 1; then the standard error, σ_e , during the event with a varying rainfall intensity and runoff rate is therefore related to the error distribution under a constant rate by:

$$\begin{aligned}
\sigma_e^2 &= \int_0^1 \text{Var}(e) d\beta \\
&= \frac{E_o^2}{6}
\end{aligned}$$

or

$$\sigma_e = \frac{E_o}{\sqrt{6}} \quad (4)$$

Both equations 3 and 4 suggest that the sampling error is independent of the rate being measured and it can become predominant when the actual rate is low.

IMPLICATIONS

Inequality (eq. 2) indicates that the bucket volume, V , the catchment area, A , and the sampling interval, Δt , determine the maximum absolute error that could occur in the observed rainfall intensity or runoff rate. Likewise, the quantity $E_o = V(A\Delta t)^{-1}$ controls the magnitude of the standard error.

For design of tipping bucket rain gauges, using a small bucket or a large diameter rain gauge would reduce the sampling error because there would be more tips per time interval for given rainfall intensity. With respect to tipping buckets for runoff measurement, the catchment area (plot size) is pre-determined. The inequality for the maximum absolute error can be used to determine the bucket size for design purposes. For example, for a runoff plot of 60 m² and a sampling interval of 1-min, if the required relative error, r_e , is 1% when the runoff rate, Q , is 200 mm/h, the bucket volume should be no more than:

$$V < r_e Q A \Delta t = 2 L$$

Although it is possible to reduce the sampling errors further by using smaller buckets, the amount of time it takes to complete one tip may become significant relative to the amount of time it takes to fill up each bucket.

Once rate data are collected, the sampling errors can be reduced considerably by using a larger sampling interval than that used to collect the data. Data at large sampling intervals can easily be obtained by accumulating the number of tips or simply averaging the observed rate over a longer time interval. Equation 4 suggests that doubling the sampling interval would reduce the sampling error by half. Table 2 shows the standard error associated with tipping buckets installed at six sites in Australia and Southeast Asia as part of an ACIAR (the Australian Centre for International Agricultural Research) project (Number 8551) (Ciesiolka et al., 1995a). The standard error for rainfall intensity is 2.4 mm/h and that for runoff rate varies from 0.6 to 3.8 mm/h depending on the plot and bucket size. The peak rainfall intensity averaged over all recorded precipitation events varies between 40 and 70 mm/h. The

Table 2. Bucket volume (V), catchment area (A), standard error (σ_e), number of events recorded (n), mean peak rainfall intensity (P_m) and runoff rate (Q_m), and the corresponding relative error (RE) for each of the six sites in Australia and southeast Asia

	V (L)	A (m ²)	V/A (mm)	σ_e (mm/h)	n	P_m (mm/h)	RE (%)	Q_m (mm/h)	RE (%)
1) Rainfall	0.007	0.07	0.10	2.4					
2) Runoff									
Goomboorian, Australia	5	108	0.046	1.1	95	45.4	5	23.4	5
Imbil, Australia	5	32	0.16	3.8	341	40.3	6	9.0	43
Kemaman, Malaysia	2.4	20	0.12	2.9	258	61.8	4	45.6	6
Los Baños, Philippines	5	96	0.052	1.3	302	46.1	5	10.7	12
VISCA, Philippines	5	71.4	0.070	1.7	85	68.8	4	4.4	39
Nan, Thailand	5	216	0.023	0.6	66	43.3	6	7.6	7

NOTE: The sampling interval was 1 min for both rainfall intensity and runoff rate at all sites.

sampling error (using one standard deviation) relative to this peak rainfall intensity is therefore between 4 and 6%. The average peak runoff rate varies much more widely from 4.4 mm/h at the Visayas State College of Agriculture (VISCA) in the Philippines to 45.6 mm/h at Kemaman in Malaysia. The lower the flow rate the higher the relative error. As a result, the relative error is in general much higher in the runoff data than in the rainfall data.

Equation 4 for the standard error can be used to select an appropriate time interval for data analysis in general, and in assessing rainfall-runoff models in particular. For example, for the event shown in figure 1, the observed peak rainfall intensity and runoff rate are 29.5 and 11.9 mm/h, respectively. The relative error based on one standard deviation for the peak rainfall intensity and runoff rate would be 8.1% and 9.2%, respectively. Suppose a relative error in peak rainfall intensity and peak runoff rate of no more than 5% is more desirable, then data at 2-min intervals should be used instead of the 1-min interval. In fact, the relative error can be reduced to any pre-set level by adjusting the sampling interval. To achieve an acceptable relative error, r_e , in the peak runoff rate, Q_p , the time interval should satisfy the following constraint:

$$\Delta t > \frac{V}{\sqrt{6} A Q_p r_e}$$

Model performance or goodness of fit can be quantified by the coefficient of efficiency (Nash and Sutcliffe, 1970):

$$E_{app} = 1 - \frac{\sum_{i=1}^N (\hat{Q}_i - Q_i)^2}{\sum_{i=1}^N (Q_i - \bar{Q})^2}$$

where Q_i are $\hat{Q}_i$ the observed and modeled runoff rate, $\bar{Q}$ is the average runoff rate for the event, N is the duration of the event, and the subscript 'app' indicates an apparent coefficient of efficiency. This coefficient of efficiency measures the fraction of total variation in the data that can be explained by the model and it is commonly used to assess model performance in hydrology (e.g., Loague and Freeze, 1985) and soil sciences (e.g., Risse et al., 1993). In Australia for example, a survey of model users suggests that for models developed to estimate monthly streamflow, the model performance was regarded acceptable when $E > 0.76$ and almost perfect when $E > 0.92$ (Chiew and McMahon, 1993). When the sampling error is relatively large, the total variation in the data is inflated, resulting in a lower apparent coefficient of efficiency, E_{app} . The total observed variation in runoff data can be written as consisting of two independent parts:

$$\sum_{i=1}^N (Q_i - \bar{Q})^2 = \left[\sum_{i=1}^N (Q_i - \bar{Q})^2 - N\sigma_e^2 \right] + N\sigma_e^2$$

The first term on the right-hand side represents the actual, yet unknown variation in the runoff rate, and the last term

represents the variation due to sampling errors. The actual coefficient of efficiency can be written as:

$$E_{act} = 1 - \frac{\sum_{i=1}^N (\hat{Q}_i - Q_i)^2}{\sum_{i=1}^N (Q_i - \bar{Q})^2 - N\sigma_e^2}$$

Algebraic manipulation results in the following relation between the actual and apparent coefficients of efficiency:

$$E_{act} = 1 - \frac{1 - E_{app}}{1 - R}$$

where

$$R = \frac{N\sigma_e^2}{\sum_{i=1}^N (Q_i - \bar{Q})^2}$$

R can be regarded as a noise to signal ratio. For perfect data with $R = 0$, E_{act} and E_{app} are identical. Since $R > 0$ for real rainfall and runoff data, it follows that $E_{act} > E_{app}$. The margin between E_{act} and E_{app} increases as the noise to signal ratio increases.

Observed variation in the runoff rate was computed for the Goomboorian site as an example where the mean peak runoff rate is relatively high (see table 2). Thirty largest storm events in terms of total rainfall were selected to calculate the standard deviation for each of three different treatments, i.e., bare soil, farmer's conventional practice, and furrows covered with mulch. Figure 3 shows that the standard deviation of 1-min runoff rates for individual events generally increases with the amount of total runoff for all treatments. The actual temporal variation in the runoff rate is therefore relatively large for large runoff events. On the other hand, the sampling error remains fixed depending only on the sampling interval for given

Figure 3—Scatter plot of the standard deviation versus total runoff from plots with three different treatments for a total of 30 largest events in terms of total rainfall. The straight line represents the standard error associated with a 1-min sampling interval.

experiment set-up. The straight line in figure 3 represents the standard error for 1-min sampling interval using equation 4. It can be seen that for small events with low runoff totals, the amount of variation in runoff is of the same order of magnitude as the amount of variation due to the sampling method alone. In order to have a fair assessment of model performance, it is sensible to reduce the noise to signal ratio by using a larger sampling interval for events with low runoff totals.

CONCLUSION

Due to the discrete nature of the sampling method used in tipping bucket technology, the magnitude of sampling errors is directly proportional to the bucket size used and inversely related to the catchment area and sampling interval. It is independent of the rainfall intensity or runoff rate being measured. Bucket size is an important consideration at the experiment and equipment design stage, while sampling interval should be selected carefully for the purpose of data analysis or hydrologic modeling. The sampling interval could be chosen to reduce the relative error in observed peak runoff rate or noise to signal ratio to an acceptable level.

ACKNOWLEDGMENT. The project was funded by the Australian Centre for International Agricultural Research, Project No. 9201.

REFERENCES

- Amerman, C. R. 1983. Infiltration measurement. In *Proc. Nat. Conf. on Advances in Infiltration*, 201-214, St. Joseph, Mich.: ASAE.
- Barfield, H. J. and M. C. Hirschi. 1986. Tipping bucket flow measurement on erosion plots. *Transactions of the ASAE* 29(6):1600-1604.
- Chiew, F. H. S. and T. A. McMahon. 1993. Data and rainfall-runoff modelling in Australia. In *Hydrology and Water Resour. Symp.*, 305-316. National Conf. Publ. No. 93/14. Newcastle, Australia: I.E. Aust.
- Ciesiolka, C. A. A., K. J. Coughlan, C. W. Rose, M. C. Escalante, G. M. Hasim, E. P. Paningbatin Jr. and S. Sombatpanit. 1995a. Methodology for a multi-country study of soil erosion management. *Soil Technol.* 8(3):179-192.
- Ciesiolka, C. A. A., K. J. Coughlan, C. W. Rose and G. D. Smith. 1995b. Erosion and hydrology of steepplands under commercial pineapple production. *Soil Technol.* 8(3):243-258.
- Edwards, I. J., W. D. Jackson and P. M. Fleming. 1974. Tipping bucket gauges for measuring run-off from experimental plots. *Agric. Meteorol.* 13:189-201.
- Hanna, E. 1995. How effective are tipping-bucket raingauges? A review. *Weather* 50(10):336-342.
- Legates, D. R. and C. J. Willmott. 1990. Mean seasonal and spatial variability in gauge-corrected, global precipitation. *Int. J. Clim.* 10:111-127.
- Loague, K. M. and R. A. Freeze. 1985. A comparison of rainfall-runoff modeling techniques on small upland catchments. *Water Resour. Res.* 21(2):229-248.
- Nash, J. E. and J. V. Sutcliffe. 1970. River flow forecasting through conceptual models, Part I: A discussion of principles. *J. Hydrol.* 10(3):282-290.
- Pillsbury, A. F., R. E. Pelishek, J. F. Osborn and T. E. Szuszkiewicz. 1962. Effects of vegetation manipulation on the disposition of precipitation on Chaparral-covered watersheds. *J. Geophys. Res.* 67(2): 695-703.
- Risse, L. M., M. A. Nearing, A. D. Nicks and J. M. Laflen. 1993. Error assessment in the Universal Soil Loss Equation. *SSSA J.* 57(3):825-833.
- Williams, J. and M. Bonell. 1987. Computation of infiltration properties from the surface hydrology of large surface plots. In *Proc. Int. Conf. Infiltration Dev. Appl.*, ed. Yu-Si Fok, 272-281. Honolulu, Hawaii: Water Res. Centre, University of Hawaii.
- WMO. 1961. *Guide to Meteorological Instrument and Observing Practices*. WMO-No. 8 TP3. Geneva, Switzerland: World Meteorological Organization.
- Yu, B., C. W. Rose, K. J. Coughlan and B. Fentie. 1997. Plot-scale rainfall-runoff characteristics and modeling at six sites in Australia and Southeast Asia. *Transactions of the ASAE* 40(5):1295-1303.