Errors in Output of Hydrologic Models Due to Errors in Input Potential Evapotranspiration

LESLIE H. PARMELE

USDA Agricultural Research Service, University Park, Pennsylvania 16802

Abstract. A study was conducted to give some perspective to the importance of accurate evapotranspiration (ET) input data to hydrologic models. For the analysis, computed output from three hydrologic models was considered as the true watershed response. Variations in the form of random fluctuations and fixed biases were introduced into the potential evapotranspiration (Pet) input data of the test models. By leaving the other inputs and parameters unchanged, such an analysis shows the effect of evapotranspiration on streamflow under the regulation of the other components in the model. It is shown that a constant bias of 20% in the Pet input data has a cumulative effect and results in considerable error in the computed hydrograph peaks and recession characteristics, whereas the influence of the random error on estimated streamflow was generally not measurable for the watersheds and models studied.

It appears that many of our water resources are going to be planned and coordinated on a river basin scale. To do this, the river basin commission must be able to predict the hydrologic performance of the basin, as well as to predict the effects of changes in water use or land practices. The future design and evaluation of management practices will be based on total hydrologic performance, not on singular results; thus methods are needed to predict the hydrology of a catchment on a complete model basis [*Engman*, 1970]. A classical example of this approach is the Stanford watershed model (SWM) [*Crawford and Linsley*, 1963, 1966] that takes component models for infiltration, evapotranspiration (ET), overland flow, interflow, groundwater and channel routing and combines them into a comprehensive watershed model.

The SWM and other watershed models have been developed to predict outflow hydrographs for a given drainage basin using precipitation and other selected records. Each component of the system is usually given a mathematical representation to define specific physical conditions in the hydrologic processes. One of the problems in the development and application of hydrologic models is that little is known about the sensitivity and accuracy required for the ET component and its influence on streamflow and other components of the hydrologic cycle. A discussion on the application of ET estimates in conjunction with watershed models is given in a paper by *Woolhiser et al.* [1970]. However, the authors present no data relative to the accuracy and sensitivity required of ET components.

This paper investigates the effect of random and biased errors in potential evapotranspiration (Pet) input data on the response of hydrologic models, which are assumed to be true representations of a watershed system. Because the accuracy required of a Pet estimate depends on the use to be made of the data, this information is needed for proper application of complete hydrologic models that require a measure of water loss from the system by evaporation.

It can be argued that errors in the streamflow Q computed by a hydrologic model would be directly related to errors in the evapotranspiration component by simple continuity. However, one has to consider that errors in estimating Q would also be related to errors in the measurement or estimate of precipitation P and storage S components in that over some period of time

$$Q = P - ET \pm S \qquad (1)$$

A direct relationship between steamflow and Pet might be representative of the actual hydrologic process if ET rates were always controlled by meteorologic factors and the biotic and edaphic variables could be ignored. However, this simple relationship is complicated because actual ET can be very sensitive to the state of water in the soil and plant systems.

In examining the flow diagram of a conceptual model (Figure 1), it is apparent that this homomorphism of the hydrologic cycle is a complicated system of interconnections between subsystems with numerous feedbacks. It can be seen that the ET demand influences streamflow indirectly; streamflow being modified by the three levels of soil moisture storage and the associated feedback effects on the actual rate of ET.

Such a group of systems can be considered a regulator as described by *Ashbey* [1961]. It could be conceived that the soil moisture reservoir serves as a regulator with feedback to reduce the variety of the initial ET state to a lesser variety in the response of the final state of streamflow. An analysis of these regulators would be complicated by the fact that their influence would be intermittent. For example, it is an accepted theory that actual ET will proceed at the rate defined by Pet until some limiting value is reached. After this time, the actual ET will be controlled in part by the availability of soil water. Thus it might be deduced that the influence of ET will be less on streamflow under dry as opposed to wet conditions. However, there is also the possibility that one or more of the regulators could

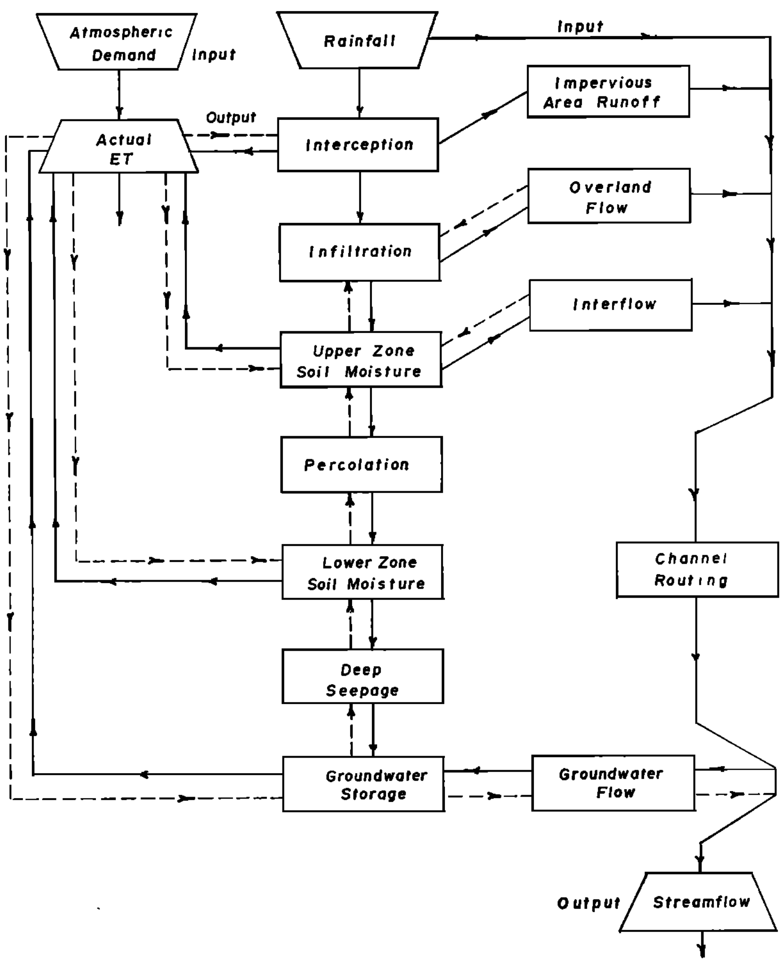

Fig. 1. Flow diagram of a conceptual hydrologic model.

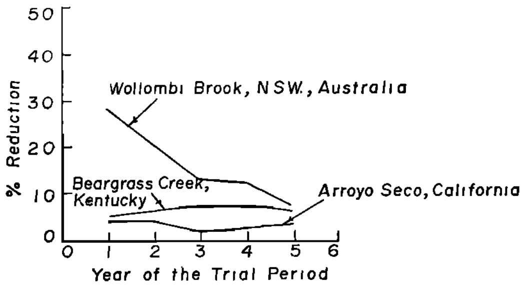

Fig. 2. Percent reduction in computed runoff caused by a 10% increase in the potential ET input data [from *Crawford and Linsley*, 1966].

amplify the relative influence of the ET demand and/or the linkage between associated 'regulators.' Thus the complexity of the conceptual ET–soil moisture runoff linkage prevents simple error and sensitivity analysis.

Crawford and Linsley [1966] report studies of the effect of changes in ET on runoff using the SWM. Figure 2 shows the result of simulations where all input data were unchanged except Pet was increased 10%. The percent reduction in runoff for each of 5 years is shown for three watersheds. For two of the watersheds, the 10% increase in Pet caused less than a 10% decrease in runoff. For an Australian watershed, however, the decrease was as great as 30%.

EXPERIMENTAL PROCEDURE

The accuracy required of Pet input data for use in hydrologic models was analyzed for simulated watershed streamflow using three conceptual hydrologic models: (1) the Hiemstra watershed model [*Sopper et al.*, 1969; *Hiemstra*, 1968] and two versions of the Stanford model [*Crawford and Linsley*, 1966] as modified by (2) [*Monro* 1971] and (3) E. A. Anderson (personal correspondence, 1970) both with the Office of Hydrology, U.S. Weather Bureau. Table 1 lists the models and watersheds tested and other pertinent data. Whereas the small sample of watersheds may limit the general applicability of the results, the basin–model combinations used will provide an indication of model response to different watershed characteristics and climatic regime.

For watersheds 1–6 in Table 1 the control parameters used were those best fit coefficients developed by the author of each particular model. Anderson and Himestra both used trial and error techniques in obtaining an optimum fit between predicted and observed streamflow. Adjustment of the model parameters was based on comparisons of predicted and observed monthly and annual water yield, storm hydrograph peak and shape, and low flow characteristics. *Monro* [1971] used a direct search optimizing technique on a digital computer to determine the model parameter values that minimized the errors between synthesized and observed streamflow.

Watersheds 7–9 in Table 1 are part of the research facility of the U.S. Department of Agriculture, Northeast Watershed Research Center. The research instrumentation on these watersheds provides detailed streamflow and precipitation measurements, and a general knowledge of the hydrologic response of the area. Optimization of model parameters for each of these three watersheds was achieved by a trial and error process in which total water yield, hydrograph shape, and low flow characteristics were matched between synthesized and recorded streamflow. Whereas it would be desirable to make the analysis over an entire year, the problems associated with snow and snowmelt runoff limited the study period on the Pennsylvania watersheds to the nonsnow periods.

In the SWM [*Crawford and Linsley*, 1963, 1966] depletion of soil water is considered limited by Pet and by evapotranspiration opportunity (Eto), which is an empirical function introduced to account for the variation of soil and water properties over the watershed. Pet results in a water loss only if water is freely available and the model attempts to satisfy the potential from interception storage and the upper soil moisture zone in that order. When near surface moisture storage is depleted, the concept of Eto is used to calculate actual ET from the lower soil moisture zone. This Eto function [*Crawford and Linsley*, 1963, 1966] defines the rate of maximum possible ET and is limited by the availability of soil moisture in the various storage zones as computed by the model.

ET is calculated for daily intervals by values of Pet, which is assumed to be equal to free water ET [*Kohler*, 1958]. Free water ET, in turn, can easily be estimated from daily pan or meteorologic data.

TABLE 1. List of Study Watersheds and Watershed Models Used in the Analysis

	Study Watershed	Area, mi²	Period of Study	Total Precip.	Runoff, inches	Watershed Model	Computational Period, minutes	Precip. Input Data, hours	Daily Pet Input Data
1	French Broad at Rosman, North Carolina	67.9	1957* 1957* 1958* 1958*	83.6 83.6 76.9 76.9	46.6 46.6 50.7 50.7	Anderson USWB Stanford Monro USWB Stanford Hiemstra	15 60 60	1 6 6	adjusted class A pan adjusted class A pan computed by Jensen-Haise method [*Jensen*, 1966]
2	Davidson River near John Rock, North Carolina	18.3	1957*		55†	Anderson USWB Stanford	15	1	adjusted class A pan
3	West Ford French Broad	37.1	1957*		53†	Anderson USWB Stanford	15	1	adjusted class A pan
4	French Broad at Blatyre, North Carolina	296	1957*		42.8	Anderson USWB Stanford	15	1	adjusted class A pan
5	Red Moshanon, Pennsylvania	68.8	Apr.–Sept. 1964	18.8	7.9	Hiemstra	60	1	computed by Jensen-Haise method
6	Green Lick, Pennsylvania	3.1	Apr.–Sept. 1964	24.1	9.8	Hiemstra	60	1	computed by Jensen-Haise method
7	Pine Creek at Klingerstown, Pennsylvania	78	Apr.–Sept. 1969	22.2	6.6	Anderson USWB Stanford	15	1	adjusted class A pan
8	Upper Mahantango at Klingerstown, Pennsylvania	42	Apr.–Sept. 1969	23.0	5.0	Anderson USWB Stanford	15	1	adjusted class A pan
9	Mahantango Creek at Malta, Pennsylvania	160	Apr.–Sept. 1969	21.9	5.9	Anderson USWB Stanford	15	1	adjusted class A pan

* Water year.
† Estimated runoff.

In the Hiemstra watershed model [*Hiemstra*, 1968] water is stored in five soil moisture zones. ET takes place sequentially with the upper zone representing interception and depression storage. These locations are depleted first at the Pet rate. Each zone is depleted at the potential rate until soil moisture in that zone is one-fourth the wilting point value, at which time losses to ET must be satisfied from the next lower soil moisture zone. Pet is computed by the Jensen-Haise method [*Jensen and Haise*, 1963; *Jensen*, 1966] based on mean air temperature and solar radiation. An empirical crop factor is used to modify the Pet value to obtain ET from the plant–soil complex.

In the analysis of Pet input errors, synthesized output from a given model using the data on which the model parameters were optimized was considered as a true watershed response. Variations in the form of random and fixed errors were introduced into the Pet input data of each model. By leaving the other inputs and parameters unchanged, such an analysis shows the effect of the two types of errors in ET input data on streamflow predictions under the regulation of the other components in each model.

The following computational runs were made for each model and watershed combination as given in Table 1:

1. No change in the input data or coefficients.
2. Pet input data biased by a constant +10, +20, −10, and −20% on a daily basis.
3. A normally distributed random error with a maximum range of ± 50% of the daily Pet input values added onto the daily input.
4. Pet input data biased by +10, +20, −10, and −20% and a random error with a maximum range of ± 50% of the daily ET input value added onto the daily input.

The values used for the biased error component of Pet are consistent with the errors that can be expected using Pet models under other than the most ideal conditions. McGuinness and Bordine (unpublished manuscript, 1971) compared 14 methods of computing Pet daily values using meteorologic data with a value derived from a weighing lysimeter. They reported variations ranging from −35 to +50% on an annual basis. Some methods gave monthly estimates that were within ± 10 to ± 20% of the lysimeter derived values; however, some errors from other methods exceeded 50%.

The random error component was introduced to see what change in streamflow would result if the Pet input had a large deviation on a daily basis, yet was an accurate description of total Pet over a longer period. This random error was generated by a subroutine available on the IBM 360 computer used in this analysis.

RESULTS

A comparison of variations in total runoff for the three watershed models using data from the French Broad River is shown in Figure 3. It can be seen that placing a positive bias of 10% on the Pet input data decreased total streamflow by 1 to 3%, and a negative bias of 20% resulted in an increase in streamflow from 2 to 7%. There is little difference in response between the two versions of the SWM. The Hiemstra model is least sensitive to any bias in the Pet input.

The introduction of a random error with maximum limits of ± 50% of the daily input value for ET along with the bias did not significantly change the total flow from that computed with a simple fixed bias. For the models and data used, the influence of the random error component is overshadowed by the fixed error in the ET estimate, as shown by the closeness of the synthetic streamflow values with and without random error input.

Figure 4 shows a comparison of the percent change in computed mean daily flows with the Monro version of the SWM for specific flow intervals and selected error cases. It can be seen that as the volume of daily flow increases, there is an increase in the error in computed flow rates resulting from the Pet bias. The addition of a random component to the Pet bias has little effect on the results.

Figure 5 shows the annual hydrograph of mean daily streamflow for the French Broad River and compares the computed output having no change in its input with that output resulting from a constant −20% bias in the Pet input data. It can be seen that at the beginning of the water year the effect of the biased input is almost unnoticeable, but as the year advances, the input error has an accumulative effect. Soil moisture is not depleted at the accepted rate; and the soil is 'wetter' than it

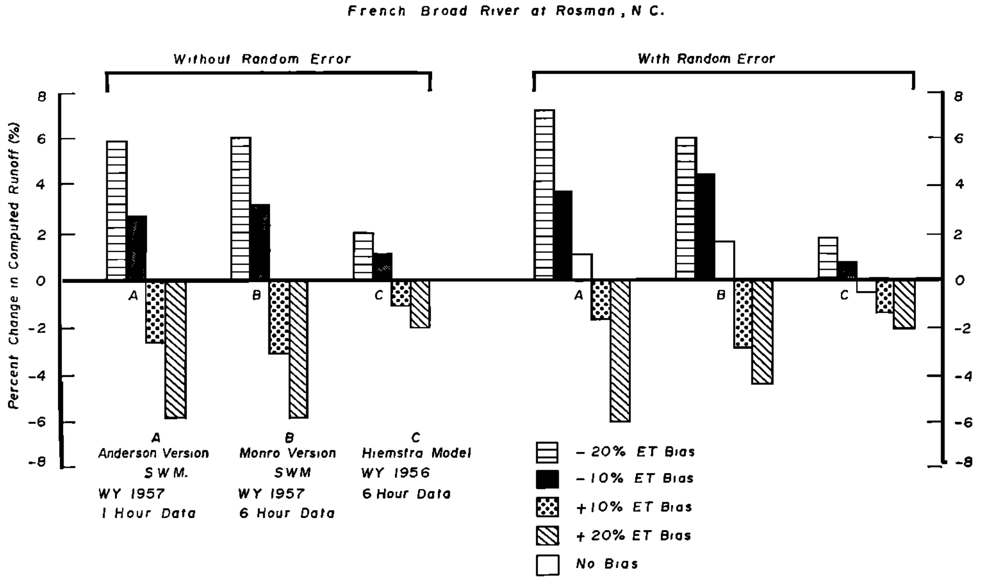

Fig. 3. Comparison of percent change in computed runoff from three hydrologic models with potential evapotranspiration (Pet) input data biased and subjected to a random error.

would be otherwise. As a result a change in the infiltration rate occurs because in all three models this variable is dependent on the available water stored in the soil profile. The change in the computed daily hydrograph peaks by as much as 200 cfs (cubic feet per second) is probably the result of this change in the infiltration rate.

A comparison of hourly calculated storm hydrograph with and without Pet input errors from the Anderson version of the SWM for the French Broad River toward the end of the

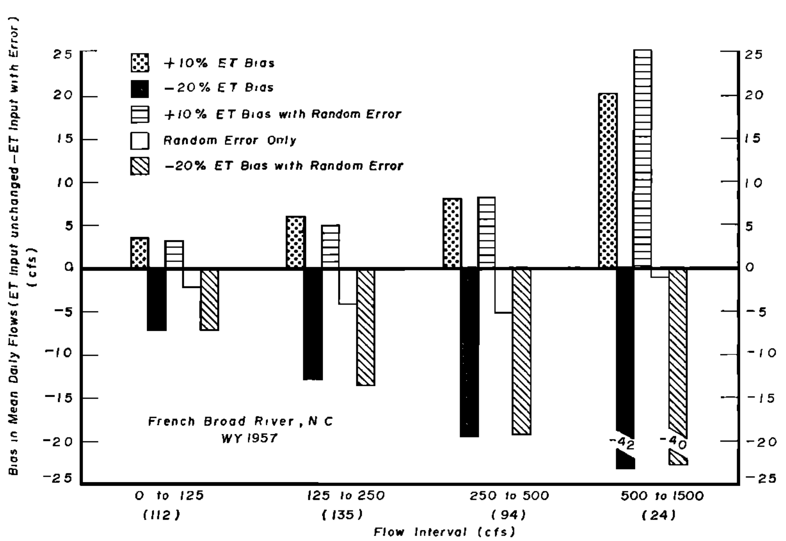

Fig. 4. Comparison of bias in computed mean daily flows with Monro version of the Stanford watershed model for specific flow intervals after Pet input biased and subjected to a random error.

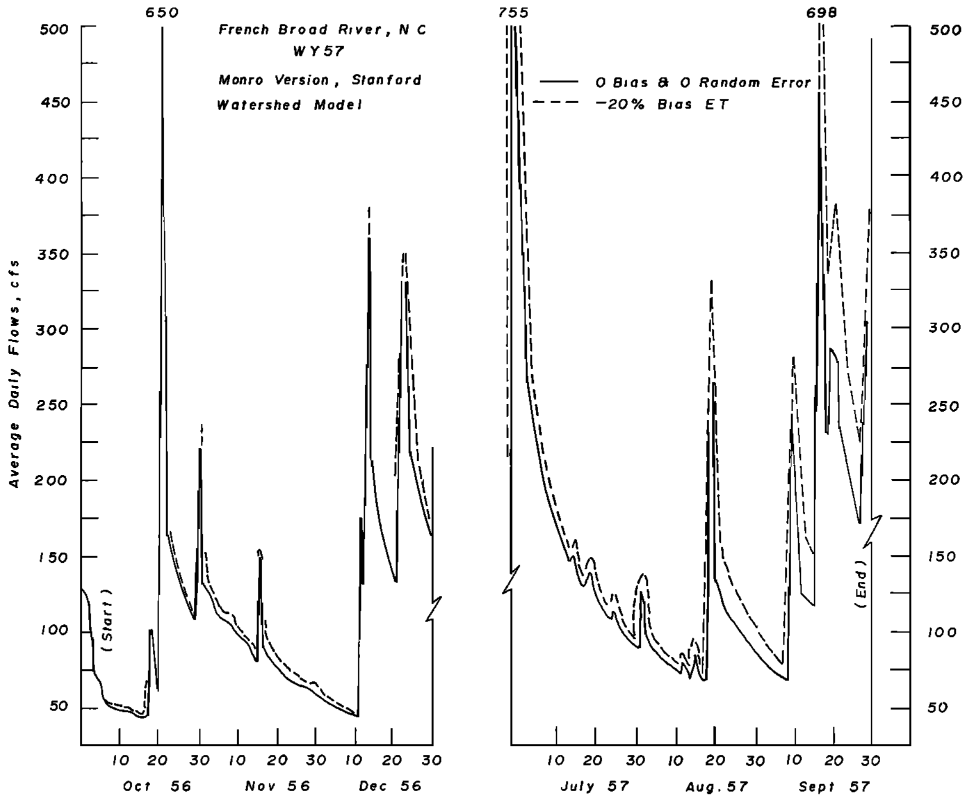

Fig. 5. Comparison of synthetic daily flow output using the Stanford watershed model with no change in Pet input data and with Pet input biased by —20%.

1957 water year is given in Figure 6. A bias of −20% without random error introduces an error of about 800 cfs or approximately a 25% overprediction in the hydrograph peak. The influence of the random error in Pet is again shown to be small when compared to the error due to a simple bias.

Figure 7 shows a graphical comparison of the calculation of total runoff from the Anderson version of the SWM for different size watersheds within the French Broad River drainage for the Pet input biases previously tested. The four watersheds range in size from 18 to 296 mi². There is practically no difference in the percent error in runoff for a fixed bias with changing watershed size. The influence of the random error in Pet input on the total runoff prediction does not appear to be measurably important.

The variation in watershed response for different locations can be seen in Figure 8, where a comparison is made of total runoff as computed by the Hiemstra model for three different watersheds. A constant Pet bias of −20% resulted in a 30% overprediction in total runoff for the Red Moshanon watershed and a 20% overprediction for the Green Lick watershed. This overprediction is in sharp contrast to results of the same watershed model on the French Broad River. However, if the comparison of watershed response were based on actual volume of runoff rather than the percent of flow, it can be shown that all three watersheds have an increased flow from 1.5 to 2.0 inches. It is again apparent that in spite of the increased sensitivity of the two Pennsylvania watersheds to a bias in the Pet input data, there is practically no response to a random error with maximum limits of ± 50% daily Pet.

Figure 9 gives the hydrograph of average daily flow for the Red Moshanon and compares the computed output having no change in its input with that output resulting from a constant −20% bias in the Pet input data. It can be observed that the relative error increases as the computation period advances. Within 4-months time, a Pet bias of −20% results in an overprediction of hydrograph peaks by 200 to 300% and an increase in estimated storm runoff volume of similar magnitude.

The only results that indicated a measurable model response to a random error component in the ET input were for the three Pennsylvania watersheds with the Anderson version of the SWM. Figure 10 shows the effect of Pet input error for three watersheds in the Mahantango Creek drainage over a 6-month period of April through September, and a similar period of computation using French Broad data for comparison. The response of the SWM is similar in magnitude to the results of the Hiemstra model (Figure 8). Computations using the SWM with the Pennsylvania data indicate a greater response to a negative bias, with a range of 18 to 36% difference in streamflow for a −20% Pet error, compared to the positive input bias in which the change in streamflow ranged from −4 to −16% for a +20% Pet error. The random error in Pet input resulted in differences between synthesized runoff volumes in some cases by upwards of 8%, the degree of response varying from one watershed to another. No trend in response to watershed size can be observed from the results.

DISCUSSION OF RESULTS

The results presented are not greatly different from those model responses reported by *Crawford and Linsley* [1966] (Figure 1); but give further insight into the importance of ac-

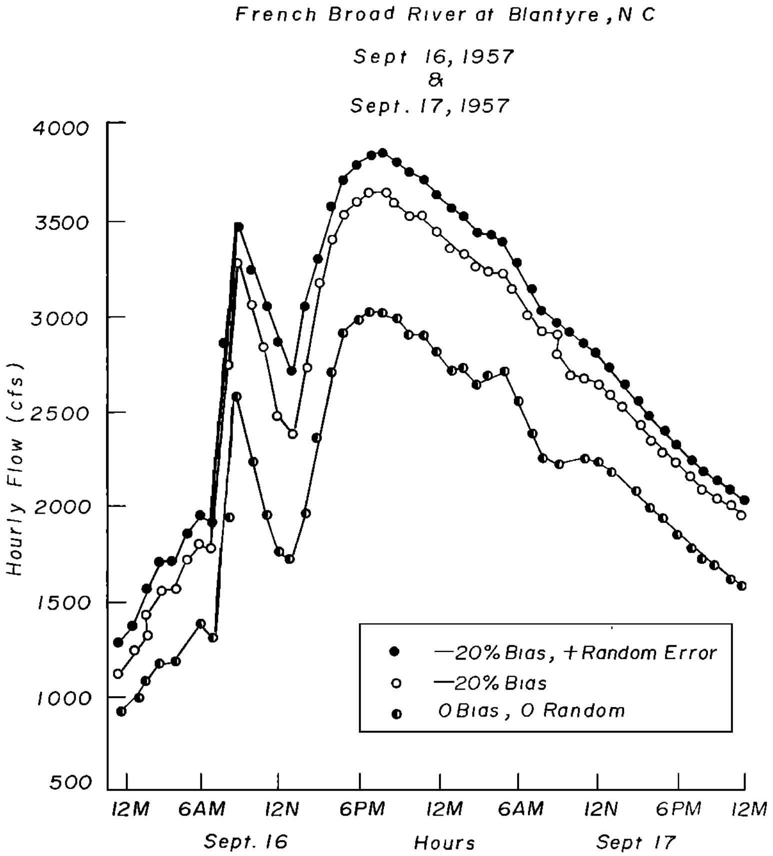

Fig. 6. Comparison of hourly flow output from the Anderson version of the Stanford watershed model for three potential ET input conditions.

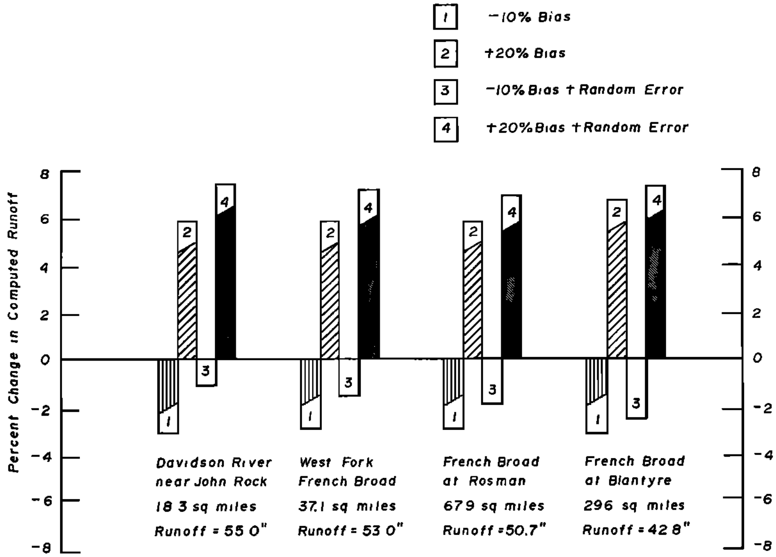

Fig. 7. Comparison of percent change in computed runoff with the Anderson version of the Stanford watershed model for watersheds of different sizes in the French Board drainage for water year 1957.

curate Pet input for use by watershed models. In particular, it is shown that even a hydrologic model with optimized parameter values may not have the desired predictive reliability for a watershed if the Pet estimates are not representative. Where records of runoff are synthesized to study many of the problems in engineering hydrology, predictions of individual storm hydrographs and possibly water yield computations could be greatly in error. In fitting a hydrologic model to a specific watershed, errors in the Pet input data would

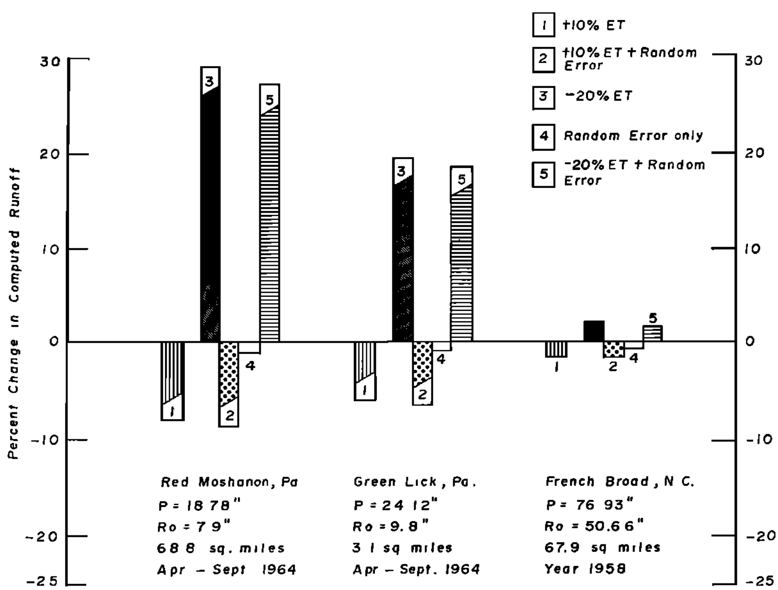

Fig. 8. Comparison of percent change in the Hiemstra model output for three watersheds after Pet input biased and subjected to a random error.

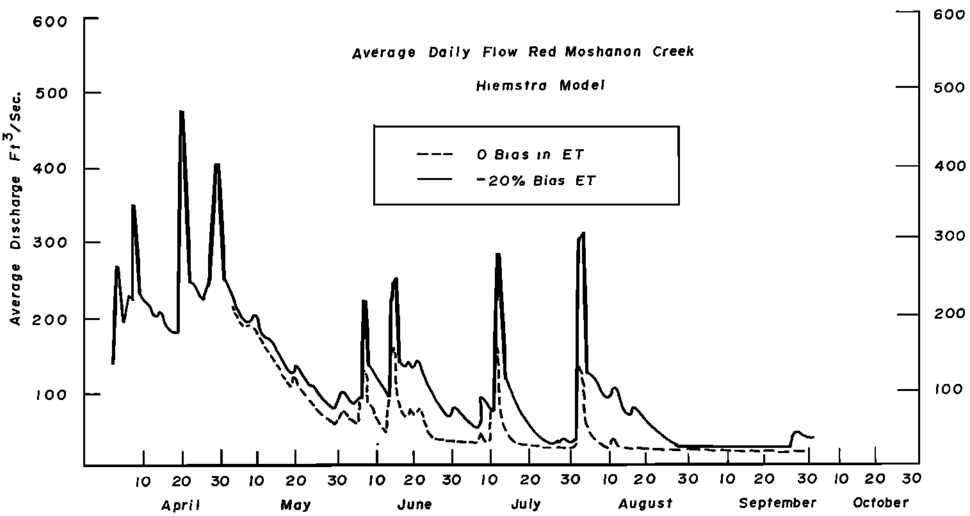

Fig. 9. Comparison of synthetic output for the Red Moshanon Creek, Pennsylvania, using the Hiemstra model with no change in Pet input data and with Pet input data biased by —20%.

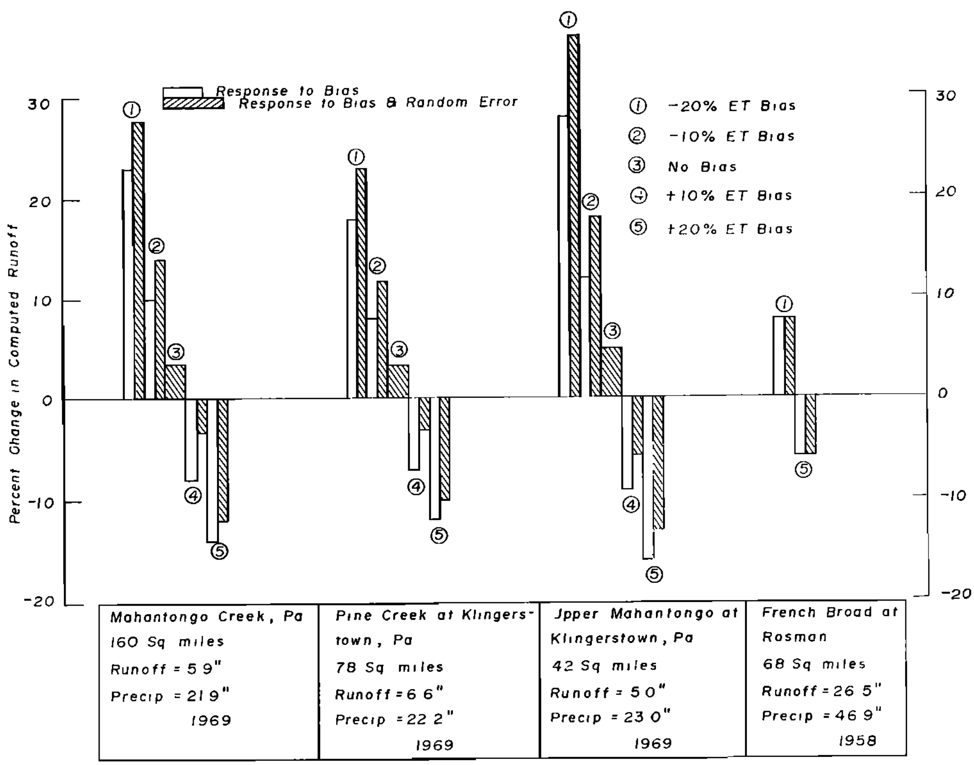

Fig. 10. Comparison of percent change in Stanford model output for a 6-month summer period at four watersheds.

conceivably result in nonrepresentative parameter optimization to match synthesized and recorded streamflows.

It should be pointed out that the results are discussed under the assumption that three synthetic watershed models are representative of the hydrologic response of the various catchments. It is recognized that differences in ET response might be balanced out or amplified by errors in other recorded data, by the structure of the particular model, or in the nonuniqueness of the model parameters [*Amorocho and Hart*, 1964]. Although arguments of this type may be valid, particularly when considering the sample size, the watershed models chosen were believed to be accurate homomorphisms of the hydrologic cycle and only those model–watershed combinations listed in Table 1 were considered to be optimized to the degree necessary for reliable predictions. The results were generally comparable between the three models tested, which could be an indication that the variations in response were the result of differences in the Pet data and not due to a nonrepresentative model structure.

From the standpoint of fitting a hydrologic model to a given watershed, it appears that an accurate estimate of Pet to better than $\pm$ 20% may not be necessary when total water yield for a climatic regime, such as that exemplified by the French Broad River with 40 to 50 inches of annual runoff, is considered. The greatest amount of error on the French Broad River with any of the models tested is probably within the accuracy of the measurements of streamflow and areal extent of precipitation. However, for climatic regions more typical of the eastern United States with considerably less runoff, as exemplified by the two watershed models using Pennsylvania data, it may be necessary to estimate Pet to better than $\pm$ 10% to achieve good fit between simulated and observed water yields.

Because ET has a seasonal variation, the influence of errors in the Pet input to a model could be more significant during the late spring and summer months when the atmospheric demand is greatest. However, the availability of water in the soil profile may also influence the degree of variability in model response. Figure 10 shows a more variable response pattern during the growing season for the three Pennsylvania watersheds as compared to the consistent response obtained with the French Broad watershed, which experiences considerably more precipitation. A bias in model response under the Pennsylvania summer conditions is measurably greater when Pet is underestimated, compared to the drier soil conditions that result from an overestimate of Pet.

It is shown that a constant bias in the Pet input data has a cumulative effect and can result in considerable error in hydrograph peaks and hydrograph recession characteristics even on the French Broad watershed where total water balance did not show a highly measurable difference using biased and unbiased Pet input data. This effect on the simulated hydrograph would tend to increase unless some method of self-learning was used in the model. Over a period of time, the failure of a model that produces hydrograph errors of the magnitude reported would be obvious. A hydrologist would then change methods or change model parameters using the biased data. However, one may not correctly evaluate the reason for the biased results and thereby accept a set of parameters that are not representative of the hydrologic process being modeled.

The influence of the random error input on estimated streamflow was not measurable on the French Broad, Green Lick, and Red Moshanon watersheds. The results with and without random error on these watersheds indicate that precise daily values of Pet may not be required and that accurate weekly or monthly estimates might be satisfactory for application in hydrologic models at these locations. However, the Stanford model did noticeably respond to the random error component on the three Mahantango basin watersheds, so it may be necessary to give more consideration to this form of error under similar hydrologic conditions.

CONCLUSIONS

The magnitude of the differences in streamflow between the best fit output and that output resulting from applied ET input data errors gives some perspective to the importance of accurate Pet input in the models examined. However, the possible limitations of the individual models and the quality of the other input data used in the analysis should be kept in mind. Because of such limitations, it cannot be stated with any degree of certainty that the results reported in this paper show a specific quantitative effect of ET on actual basin runoff.

A comprehensive systems analysis into the accuracy and sensitivity of Pet estimates required for hydrologic models would seem desirable to establish limitations for model input data. There is a greater need, however, to know how to convert such numerical manipulations with conceptual hydrologic models into real world implications.

Acknowledgments. The author is indebted to Dr. L. A. V. Hiemstra, former assistant professor of Civil Engineering at The Pennsylvania State University, and E. A. Anderson and J. Monro, research hydrologists with the Office of Hydrology, U.S. Weather Bureau, for their assistance and for providing the models and data used in this analysis. Contribution from the Northeast Branch, Soil and Water Conservation Research Division, ARS-USDA, University Park, Pennsylvania, in cooperation with the Agricultural Experiment Station, The Pennsylvania State University, University Park, Pennsylvania.

REFERENCES

Amorocho, J., and W. E. Hart, A critique of current methods in hydrologic systems investigation, *Eos Trans. AGU, 45*(2), 307–321, 1964.

Ashbey, W. R., *An Introduction to Cybernetics,* Chapman and Hall, London, 1961.

Crawford, N. H., and R. K. Linsley, A conceptual model of the hydrologic cycle, *Int. Ass. Sci. Hydrol. Publ. 63,* 573–587, 1963.

Crawford, N. H., and R. K. Linsley, Digital simulation in hydrology: Stanford watershed model IV, *Tech. Rep. 39,* Dep. Civil Eng., Stanford Univ., Stanford, Calif., 1966.

Engman, E. T., Watershed models as a research tool, paper presented at Watershed Modeling Workshop, Agr. Res. Serv., Soil Conserv. Serv., Tucson, Ariz., pp. 14-9–14-10, 1970.

Hiemstra, L. A. V., Frequencies of runoff for small basins, Ph.D. dissertation, pp. 3–57, Colo State Univ., Fort Collins, 1968.

Jensen, M. E., Empirical methods of estimating or predicting evapotranspiration using radiation, paper presented at Symposium on Evapotranspiration and its Role in Water Resources Management, Amer. Soc. Agr. Eng., Chicago, Ill., December 5–6, 1966.

Jensen, M. E., and H. R. Haise, Estimating evapotranspiration from solar radiation, *J. Irrig. Drain. Div., Amer. Soc. Civil Eng., 89*(LR4), 15–41, 1963.

Kohler, M. A., Meteorological aspects of evaporation phenomena, *Trans. Int. Ass. Sci. Hydrol., 3,* 421–436, 1958.

Monro, J. C., Direct search optimization in mathematical modeling and a watershed model application, *Tech. Memo. Hydro 12,* 52 pp., Nat. Oceanic Atmos. Admin., U.S. Dep. of Commer., Silver Spring, Md., April 1971.

Sopper, W. E., et al., Hydrologic behavior of selected watersheds in the northern Appalachian region, *Res. Publ. 62,* 108 pp., Inst. for Res. on Land and Water Resour., Penn. State Univ., University Park, 1969.

Woolhiser, D. A., R. E. Smith, and C. L. Hanson, Evapotranspiration components of watershed models for the Great Plains, paper presented at Seminar on 'Evapotranspiration in the Great Plains,' Kans. State Univ., Bushland, Tex., Mar. 23–25, 1970.

(Manuscript received January 8, 1971;
revised December 31, 1971.)