The Transition from Manual to Automated Systems for Acquiring Snow Data

761-84

by Arthur G. Crook[1]/and Richard H. McCuen[2]/

INTRODUCTION

Manually observed data have been used extensively in streamflow forecasting and gaging in the western United States for decades. Automated sensor development gained momentum in the 1960's. Large data acquisition networks using onsite recorders and telemetry were relatively common by the late 1970's. Automated sensors provide accurate data more frequently than can a snow survey team, and at less cost. With the daily data provided by the snow survey telemetry (SNOTEL) system, more accurate predictions of streamflow are possible.

The Soil Conservation Service (SCS) has recently developed a procedure for operationally converting from older manually collected snow surveys to automated readings gathered by its SNOTEL system. The procedure consists of two parts--the verification of automated site data and the determination of whether automated data can reliably be substituted for manually collected data. This paper describes the current SNOTEL operations and the procedures for making the transition from manual to automated data operations.

BACKGROUND

The first snow surveys in the Western United States were conducted by Dr. James E. Church on the slopes of Mt. Rose in the Sierra Nevada in 1906. Numerous project-oriented snow course networks were established in the 1910's and 1920's. By 1935, at least nine independent and uncorrelated networks were operational in the West.

The first region-wide coordinated snow survey effort was a result of the unprecedented western drought of 1934. Congress appropriated a small amount of money for the activity and charged the Bureau of Agricultural Engineering with program responsiblity. In 1939 this responsibility was shifted to the SCS. By 1955, there were as many as 1500 to 2000 snow courses in the West being routinely measured by SCS and numerous cooperating agencies. With the data from these snow surveys, SCS and the National Weather Service forecast seasonal streamflow for some 500 gaging stations in the West.

In the late 1960's, SCS began installing prototype VHF (very high frequency) radio telemetry in several states. In 1975, Congress approved funding for a 500 site telemetry network (Barton, 1977). The network first produced data in 1977, and by October of 1980, 465 sites were regularly providing snowpack, rainfall, and temperature data from locations generally at or near existing snow courses in the West via meteor burst telemetry (Leader, 1974).

CURRENT OPERATIONS

The SNOTEL system acquires data each morning from snow-pressure pillows, precipitation storage gages, and air-temperature sensors at each remote location. Maintenance crews check each site at least once during the snow-free summer period. Other visits are made as performance problems dictate and in conjunction with regularly scheduled snow surveys.

In winter, survey teams gather data for use in validating the telemetered data. They measure the depth, water equivalent, and density of the snow around the perimeter of the pillow. They also measure fluid levels in transparent manometer tubes attached to the plumbing lines of the snow pillow and precipitation gage, and air temperature.

Paper presented Western Snow Conference, 1984.

1/ Data Collection Group Leader, Soil Conservation Service, U.S., Department of Agriculture, Portland, Oregon

2/ Professor, Department of Civil Engineering, University of Maryland, College Park, Maryland

The onsite observations are compared to electronic signals measured during the visit to determine whether the pressure transducers, sensor interfaces, and other components are functioning properly. A later office check is made to compare the onsite data with the numbers received via computer terminal link to the telemetry system.

THE TRANSITION

The primary justification for operating the SNOTEL system is the anticipated improvement in streamflow forecasting accuracy. With daily records available, the forecaster can more accurately predict the total seasonal volume, as well as prepare full hydrograph predictions based on physical process simulation models. Consequently, it is incumbent upon the data users to make the transition to use of the automated data as quickly as feasible and hydrologically justified. Concurrently, with this change is a reluctance to give up the security of the old tried and true manually collected snow course data. However, dual operations in which both networks are perpetuated are time consuming and expensive. In order to reduce costs, save staff time, and gain maximum advantage of the automated data, the transition needs to be made as soon as technically feasible.

VERIFYING THE DATA

Two stages of automated data verification are necessary prior to operational use. First is the determination of whether the sensor is "seeing" the same thing as is actually occurring onsite, and whether the value sensed is transmitted and stored correctly in the central data base. The second stage involves determining the relationship between automated data and manually collected data from the same and/or nearby location. In both stages the linear regression model $\hat{Y} = b_o + b_1X$ is used.

Stage one, verifying the accuracy of telemetered data, is in itself a two-step process—comparing the telemetered number with the onsite value and also with the ground truth. The first step consists of regressing the telemetered value against observed manometer readings. This test can be applied to both snow water equivalent and precipitation storage data. The data pairs are plotted and examined for local biases (i.e., poor fit for either high or low values, but with a good fit in general) and extreme events (i.e., an adequate range of data). In order to assure sufficient degrees of freedom, a minimum of 30 pairs of data(n) are desired. This gives some assurance that the data base will be representative of future occurrences.

The most important measure of acceptability of the degree of association between the telemetered data and the observed manometer readings is the standard error of the estimate (S_e), which is computed by:

$$S_e = S_y \sqrt{\frac{(n-1)}{(n-2)} \quad (1-R^2)}$$

in which S_e is the standard deviation of the telemetered data and R is the correlation coefficient. The acceptable maximum S_e should be a physically based measure of tolerable error. In this case, it can be related to the performance of the pressure transducer which is used to transform the snow pillow fluid pressure into an electronic signal. The static error band of the most commonly used SNOTEL transducer is $\pm$ 0.5 percent of full scale. SCS has defined the tolerable S_e at $\pm$ 1.0 percent of full scale, which allows for systematic and random errors equal in magnitude to the static error band. Hence, a 2.54 meter (100 inch) capacity transducer would have a tolerable maximum S_e of $\pm$ 2.54 cm (1.00 inch).

Other measures of the acceptability of this association are: the regression coefficient (b_1) within the range of .99 to 1.01; and the intercept of the least square regression line (b_o) within the range of $\pm$ 1 percent of the transducer full scale.

The second step in the verification process is examination of telemetered values and their relationship to actual onsite conditions—or ground truth. In the case of snow water equivalent, ground truth is determined by samples taken around the snow pillow perimeter with a Federal snow sampler. As in the first step, the data pairs are plotted and examined for local biases and range. A minimum sample size (n) of 30 is again desirable. The statistical criteria of acceptability is (R) $\leq$ 0.975, (b_o) within $\pm$ 1.0, and (b_1) within the range of 0.9 to 1.1. A comprehensive study of 5 years of SNOTEL data in Colorado and New Mexico by Schaefer and Shafer (1982) demonstrated that these criteria are reasonable and can be met or exceeded when SNOTEL sites are properly installed and operated.

The second stage, the relationship between telemetered data and manually collected data, is the final test of whether the data user can confidently rely on the real-time values for operational forecasting.

In theory, the most revealing test would examine the sensitivity of the forecast procedure that is affected by a proposed change to automated data. This test can be accompished either by extending the telemetered data back in time using bivariate or multiple regression where manual snow courses are the independent variables, or by operating both systems concurrently until a sufficient history of telemetered data is accumulated. In either case, the new record is used to prepare a new forecasting procedure, and the S_e values of the new and old procedures compared. When preparing a "synthetic" back record, and using telemetry and manual data which are well correlated, a significant difference between the S_e values of the new and old forecast procedures would not be expected. Only when the data are poorly correlated would S_e values be dissimilar. When perpetuating both systems to attain an adequate "real" data record, a lengthy period of duplicate data acquisition efforts is required. Neither approach has been deemed acceptable by the SCS because of the questionable value of comparing S_e's or the expense of duplicate systems.

If the time and expense of dual measurement systems is to be minimized, criteria can be set for acceptability of the correlation of telemetry to manual data. Sites which meet the criteria can be used to estimate data in lieu of making manual measurements. For SNOTEL and snow course data, the SCS has adopted a minimum sample size (n) of 30 encompassing at least 6 years of observation and including only snowpack accumulation phase data. The R value must equal or exceed 0.975, meaning that 95 percent of the variability is explained by the equation. When melt phase data are used in forecast procedures, a separate correlation using only data pairs from this period is required. The question of what consitutes acceptability of the correlation is currently under study by SCS.

When all of the above discussed tests have been passed, the data user can confidently begin forecasting with just telemetered data. These data can be used to estimate data for unmeasured nearby snow courses and these estimates used in existing forecasting procedures. After a sufficient record has been obtained through telemetry, a new regression equation can be developed for water supply forecasting that uses only the SNOTEL data as input.

References

Barton, Manes, "SNOTEL: Wave of the Present," Soil Conservation, March 1977.

Leader, Ray E., "Meteor Burst Communications," Proceedings of the 42nd Western Snow Conference, 1974, pp. 29-36.

Shaefer, Garry L. and Bernard A. Shafer, "A Critical Analysis of Five Years of SNOTEL Performance in the Rocky Mountains," Proceedings of the International Symposium on Hydrometeorology, Denver, Colorado, June 1982.