

# **Benchmarking Irrigation System Performance Using Water Measurement and Water Balances**

**Proceedings from the  
1999 USCID Water Management Conference**

**San Luis Obispo, California  
March 10-13, 1999**

**Sponsored by**

**U.S. Committee on Irrigation and Drainage**

**Co-Sponsored by**

**California Polytechnic State University  
and  
Bureau of Reclamation**

**Edited by**

**Grant G. Davids  
Davids Engineering, Inc.  
and  
Susan S. Anderson  
U.S. Committee on Irrigation and Drainage**

**Published by**

**U.S. Committee on Irrigation and Drainage  
1616 Seventeenth Street, Suite 483  
Denver, CO 80202  
Telephone: 303-628-5430  
Fax: 303-628-5431  
E-mail: [stephens@uscid.org](mailto:stephens@uscid.org)  
Internet: [www.uscid.org/~uscid](http://www.uscid.org/~uscid)**

## CORRELATION BETWEEN SAMPLING INTERVAL AND VOLUME CALCULATIONS

Bryan P. Thoreson<sup>1</sup>

John Eckhardt<sup>2</sup>

Anisa Joy Divine<sup>3</sup>

### ABSTRACT

Correlation between accuracy of volume calculations and time interval selected for water elevation measurements is an important question in today's digital world. Modern techniques for digital data acquisition can record water elevations at constant or varying time intervals. Recording water elevations at constant time intervals has the advantages of providing a predictable amount of data, allowing easier data filling for missing and out-of-range records, and permitting direct use of the average function. In addition, the constant time interval supports site maintenance because data collection cessation can easily be determined. Because of these advantages, a constant time interval is often chosen; however, the length of the time interval is usually selected based on equipment requirements rather than considerations about data accuracy.

The effect of constant and varying data recording intervals on calculated daily flow was investigated by analyzing an existing data set of digitized Stevens Charts. The Stevens Chart record consists of a digitized point for each slope change in the recorded water surface level. For each site, eight sets of flow data (15-minute, hourly, every two hours, every four hours, every six hours, every eight hours, every 12 hours, and once a day at 8 a.m.) were generated. The digitized data most closely matches the original, analog flow record, and thus, is the volume standard against which the calculated volumes from the other sets are compared.

Daily volumes were calculated for each data set. The average percent difference for annual volumes ranged from -0.2 percent for 15-minute and hourly readings to -1.9 percent for the twice a day readings and 12.7 percent for daily readings at 8 a.m. An increase from one to two readings a day decreased the average percent difference from 12.7 percent to -1.9 percent. This result is for sites that have some flow nearly continuously.

---

<sup>1</sup> Project Manager, Davids Engineering, 1772 Picasso Avenue, Suite A, Davis, CA 95616-0550

<sup>2</sup> Assistant General Manager, Water, Imperial Irrigation District, 333 E. Barioni Blvd., P.O. Box 937, Imperial, CA 92251

<sup>3</sup> Water Resources Planner, Imperial Irrigation District, 333 E. Barioni Blvd., P.O. Box 937, Imperial, CA 92251

These sites with high flow variability would be expected to benefit from a detailed 15-minute record. However, the results indicated that yearly volumes calculated from an hourly record were appreciably less accurate than those calculated from a 15-minute record.

Knowing how data will be used is vital to deciding what time interval to use for data collection. This study indicates that if yearly volumes are the sole purpose of the data collection, 2-hour intervals may be enough. If the data are likely to be used for other system studies, as experience has shown is likely, hourly values are probably a good idea.

## INTRODUCTION

An analysis of the correlation between accuracy of calculated flow and the time interval selected for water elevation measurements is presented in this paper. Modern techniques for digital data acquisition can record water elevations at constant or varying time intervals. Recording water elevations at constant time intervals has three advantages -- it provides a predictable amount of data, allows easier data filling for missing and out-of-range records, and permits direct use of the average function. In addition, the constant time interval supports site maintenance because data collection cessation can easily be noted. Because of these advantages, a constant time interval is often chosen. However, the length of the time interval needed to provide sufficient data to calculate flow accurately remains to be determined.

An existing data set was analyzed to investigate the effect that constant and varying recording intervals have on calculated daily volumes. The data were developed by digitizing Stevens recorder charts for eight Imperial Irrigation District spill sites. A point was digitized whenever the slope of the recorded water surface level changed. From this record, eight time-interval (15-minute, hourly, every two hours, every four hours, every six hours, every eight hours, every 12 hours, and once a day at 8 a.m.) sets of flow data were generated. Since the digitized data most closely matches the original, analog flow record, it is the volume standard against which calculated volumes from the other sets were compared.

Yearly, monthly and daily volumes were calculated for each of the time-interval data sets. Differences between yearly volumes at each site are related to three measures of flow variability: 1) coefficient of variation, 2) flow range (maximum), and 3) digitized points per year. The results were expected to provide guidelines for determining which time interval to use for recording water elevations based on the flow variability at a site and the desired yearly volume accuracy. The implications of these guidelines to database design and storage requirements will be discussed.

## DATA SET DESCRIPTION

Description of Digitizing Process

Stevens charts were digitized creating depth versus time values in AutoCAD for every point. The following conventions were used during the digitization process:

1. If the stage changed linearly, only the beginning and ending points of the line were digitized.
2. Each file was a continuous record. In other words, a file was ended and a new file was begun whenever there was a break in the record.
3. When the stage dropped below the weir crest, only the times when the stage equaled the weir crest as it was decreasing and as it was increasing were digitized. The period between these two times when the stage was below the weir crest was not digitized.

An AutoLisp program dumped the values from the AutoCAD file into a comma-delimited file. The x-value represents a Julian date and time to the nearest hundredth of a day. Using the Julian date and time made it easy to calibrate the digitizer, and reference tables were provided showing the date corresponding to each whole number Julian date. The y-values are 10-times the chart stage values in feet.

An Excel macro was used to convert the comma-delimited file into a dBase 5.0 file with date, time and stage fields. The macro steps are summarized below:

1. Accept date corresponding to first Julian date and time in the file as input.
2. Check sequence of all Julian date and time values. If a value is out of sequence, force it into the correct sequence by adding one-hundredth to the last correctly sequenced value.
3. Sort records according to Julian date and time values from the most distant past to nearest present values.
4. Convert Julian date and time values into date and time fields.
5. Divide the y-values by 10 to obtain the chart stage values in feet.
6. Check for and print a list of missing days.
7. Save processed file as a dBase 5.0 file with date, time and stage fields.

The macro was checked for accuracy by graphing random weeks from the processed files and comparing the graphs to those for the corresponding week from the comma-delimited input file of digitized Stevens chart data.

Files generated from the macro were joined together to create a record for each year of data. In addition, each Stevens chart was checked for notes related to data quality. Missing record (gaps) of a day or longer was indicated by a "no record"

code (-9999) for the stage record at 5.75 hours for that day.<sup>4</sup> A no record event of less than one day that did not include a record at the 5.75 time was indicated by inserting a record with a time one-hundredth of a day beyond the last valid record with a no record code for the stage record. The end of the no record period was indicated by inserting another record with a time one-hundredth of a day before the valid record resumed.

### 15-minute Data Set Development

The 15-minute data set was developed from the digitized data set. Because the data were digitized on points of inflection (where the slope changed), a constant slope between these points was assumed. A BASIC program was written that inserted a record at each 15-minute mark of the hour. The stage value was interpolated based on the interval between the digitized points (Equation 1). As a check, the flow from the digitized data set was then compared to the flow from the 15-minute data set (Figure 1) for selected time periods at selected sites.

$$S_{15} = ((T_{15} - T_1)/(T_2 - T_1)) * (S_2 - S_1) + S_1$$

**Equation 1**

Where	$S_{15}$	=	Interpolated stage at 15-minute interval, ft
	$S_1$	=	Stage at $T_1$ , ft
	$S_2$	=	Stage at $T_2$ , ft
	$T_{15}$	=	Time of 15-minute interval for which stage is being interpolated, hr
	$T_1$	=	Last time before 15-minute interval, hr
	$T_2$	=	First time after 15-minute interval, hr

### Remaining Data Sets Development

Seven additional data sets were developed by extracting records on the hour, every two hours, every four hours, every six hours, every eight hours, every 12 hours and once a day at 8 a.m. (96, 24, 12, 6, 3, 2, and 1 record per day, respectively) from the 15-minute data set. This was accomplished by writing SQL queries on the database that returned only the records at the time required for each record. These queries simulated data collection only at the selected time intervals.

The hours used for the data sets with records every two, four, six, eight, and 12 hours were those hours that were divisible by two, four, six, eight, and 12, respectively. For each of these data sets, there are two, four, six, and eight and 12

---

<sup>4</sup> The choice of 5.75 hours (5:45 a.m.) was the result of data collection procedures in effect at the time the data were first analyzed.

possible data sets depending upon which hours are chosen. To check that the hours chosen did not affect the result, a second set of hours was also used for the every four hours set. The average percent difference and variability of the error were not significantly different for the two different sets at the four hour interval. Thus, it was concluded that the hours chosen for these sets was not critical to the errors for this data set.

Average daily flows were then calculated for each data set and summed to monthly and yearly volumes. To eliminate the effects of estimating, gaps in the record were ignored. Thus, the monthly and yearly volumes do not represent a complete record.

Minimum, maximum, average, standard deviation and the coefficient of variation of flow were calculated for the entire record of all eight sites for all eight data sets. These measures were compared for all the data sets. The average number of points digitized per year over the entire record of each site was also used as an indication of variability.

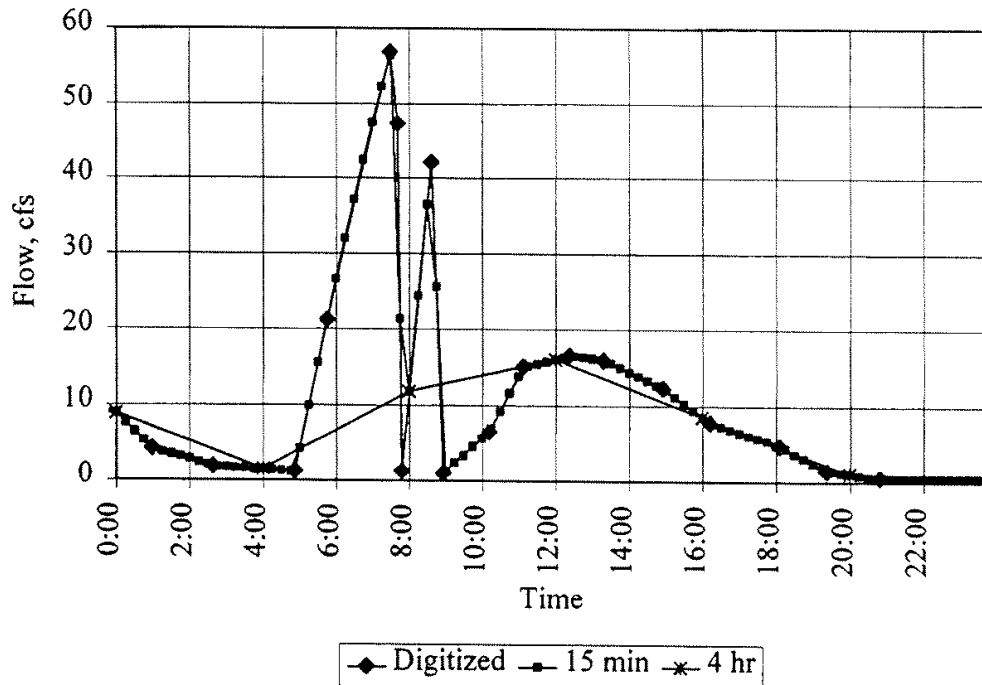


Figure 1. Comparison of One Day of Digitized Data and Data Recorded Every 15 Minutes and Every Four Hours.

The volume between two digitized points was calculated as the average flow of two consecutive points multiplied by the time between them. These volumes were summed to daily volumes which were in turn summed to monthly and yearly volumes. This volume was considered to be the most accurate volume, and

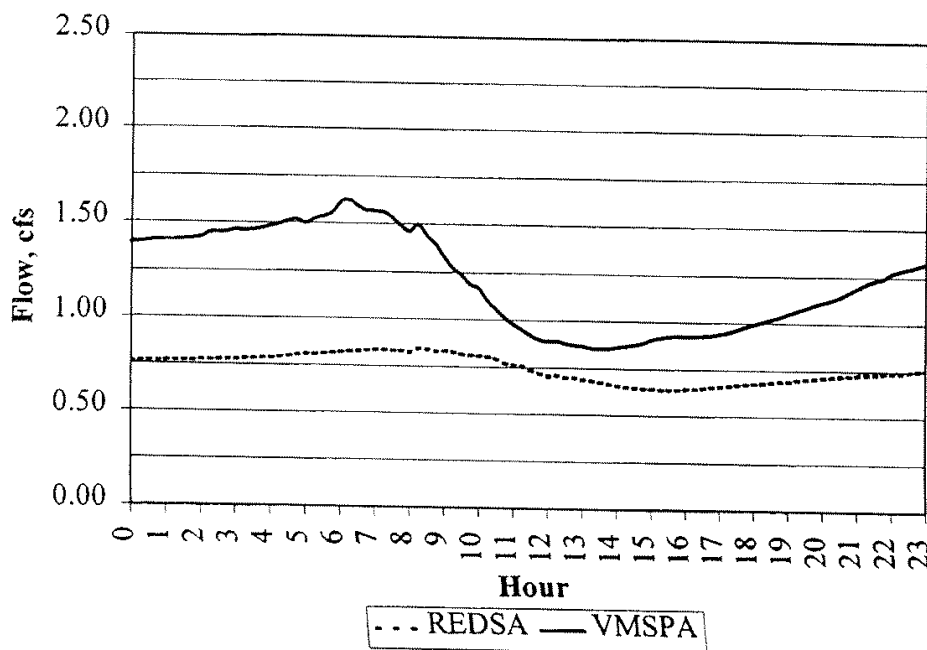
all other volumes were compared to it. The volumes for the remaining data sets were calculated by assuming that the flow value at the recorded time was the average flow value for the entire following time interval, be it 15-minutes or 24 hours.

## RESULTS

### Flow Statistics

The flow statistics discussed are for the entire time period of available data for these sites. The time periods available ranged from 2 to 10 years. All of these sites are spill sites and expected to be highly variable. Averaging the flow recorded at each 15-minute time step across the entire record at these sites reveals a diurnal cycle (Figure 2). Generally, these sites have higher flows in the early morning hours, with decreasing flows throughout the morning to the lowest flows in the early to mid afternoon hours and finally increasing flows throughout the early evening hours.

Figure 2. Daily Hydrograph Showing Typical Diurnal Cycles.



Although it is likely that these sites have weekly, monthly, seasonal and perhaps annual cycles also, these cycles were not investigated. This is because as long as the data recording intervals are once a day or more often and data is recorded

every day, cycles longer than one day in length will not affect the volume accuracy.

The flow range for each site was the greatest for the digitized data and decreased steadily as the number of points recorded per day decreased (Figure 3). The minimum flow for each of these sites was zero, so the range could be represented by the maximum flow recorded. The two sites with digitized maximum flows above 200 cfs include runoff flows entering main canals from intense desert rainfall events.

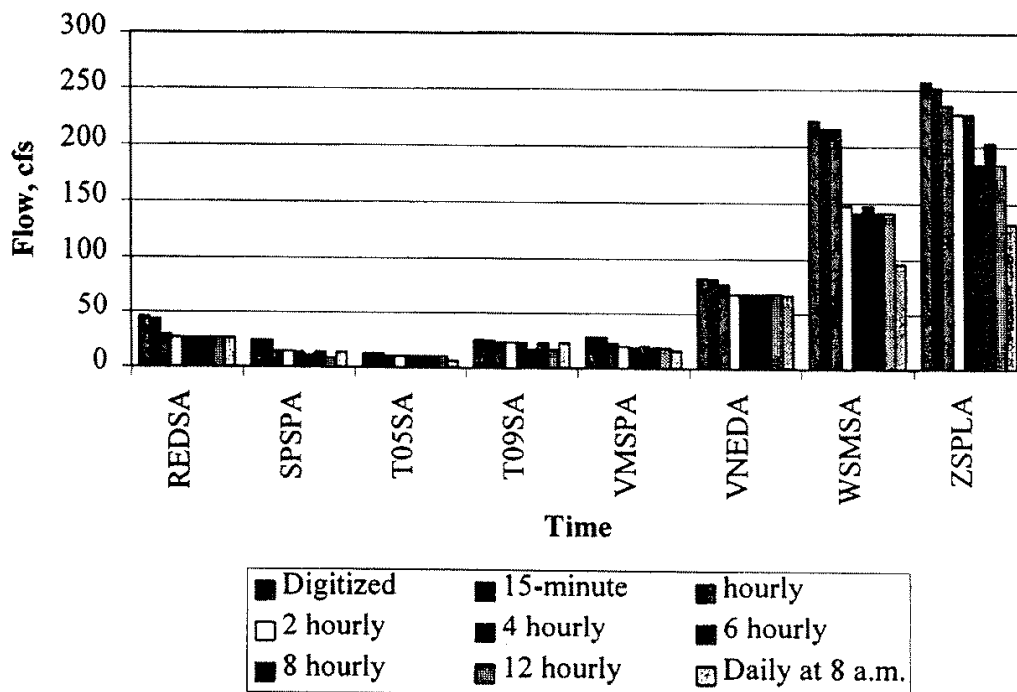


Figure 3. Comparison of Maximum Flows across Different Time Intervals.

Although knowing the maximum flow to pass the site is not particularly important in regards to annual volume calculations, it may be important for other data applications. Table 1 shows that the maximum flow from the 15-minute record averages 98 percent of the digitized maximum. This decreases to an average of 81 percent for the hourly record and an average of 73 percent for the two-hourly record.

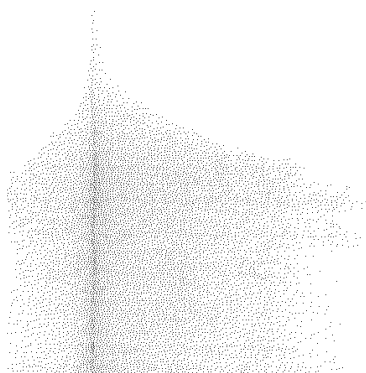




Table 1. Maximum Flows as a Percentage of the Digitized Maximum.

Site	15-min.	hourly	2-hourly	4-hourly	6-hourly	8-hourly	12-hourly	daily
REDSA	96%	64%	59%	59%	59%	59%	59%	59%
SPSPA	100%	60%	60%	57%	57%	57%	35%	57%
T05SA	98%	82%	82%	82%	82%	82%	82%	48%
T09SA	98%	90%	90%	90%	90%	90%	65%	90%
VMSPA	99%	81%	71%	65%	65%	65%	65%	58%
VNEDA	100%	94%	83%	83%	83%	83%	83%	82%
WSMSA	97%	97%	66%	63%	63%	63%	63%	43%
average	98%	81%	73%	71%	71%	71%	64%	62%

The average flow rate remained nearly constant as points recorded per day decreased until only one point was recorded (Figure 4). When only one point was recorded, some major changes in average flow rates were observed, in particular at the group of sites with larger flow rates.

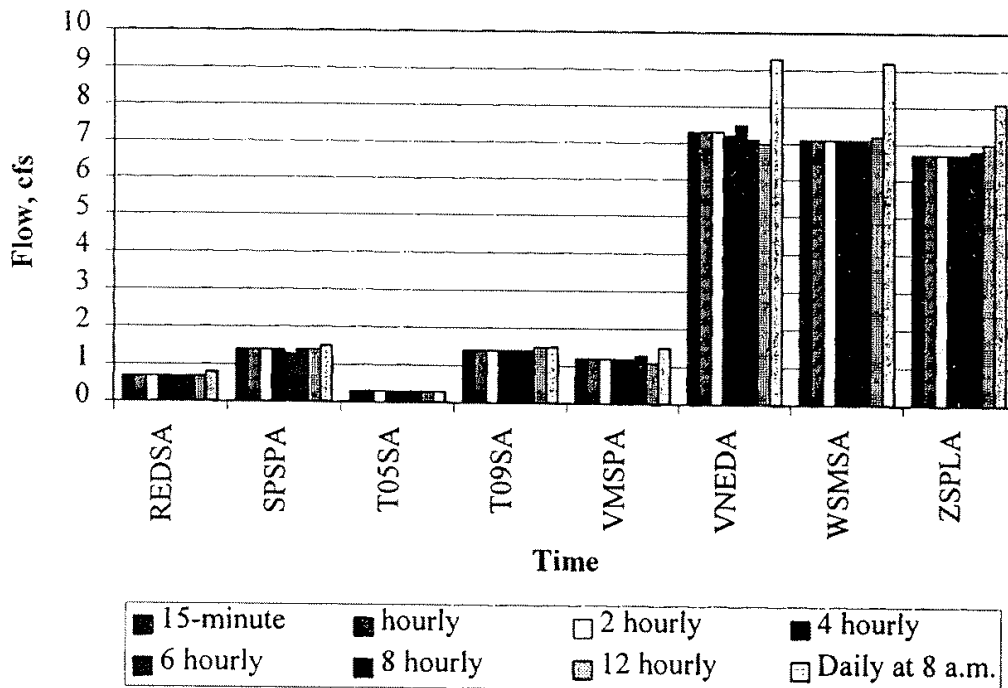


Figure 4. Comparison of Average Flow Rates at Different Data Recording Intervals (an average flow rate was not calculated for digitized data because of the varying time interval).

The coefficient of variance was nearly constant as the number points recorded in a day decreased (Figure 5). At less than 6 points per day, some changes in the coefficient of variance were observed.

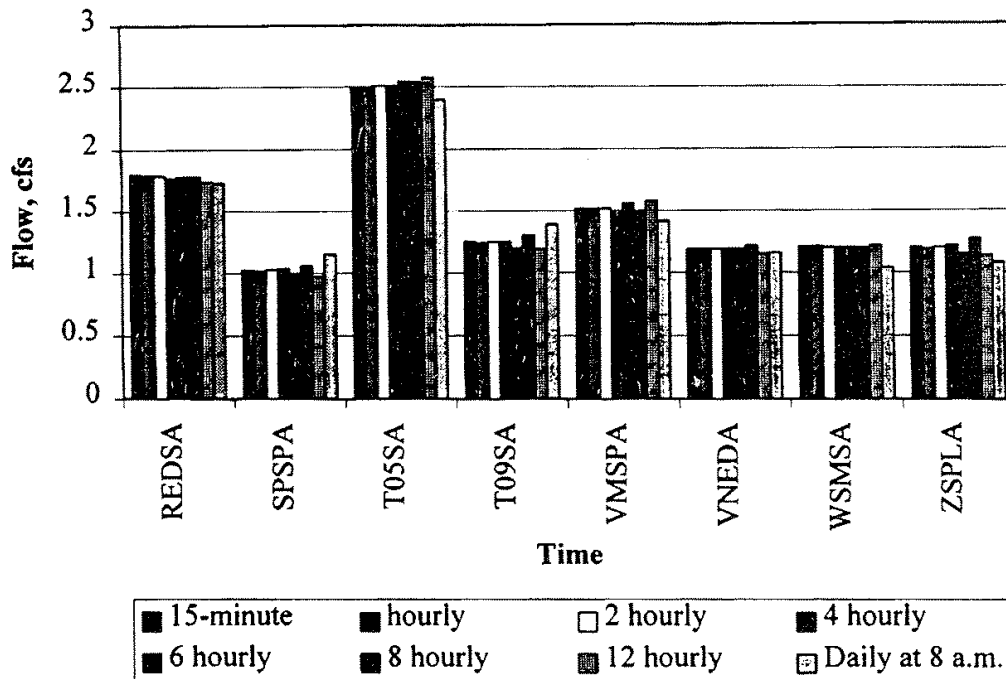


Figure 5. Comparison of the Coefficient of Variation of Flows across Different Time Intervals

Volume Comparisons

Yearly: The average percent difference across all site-year combinations (36) is essentially the same for the 15-minute, hourly, 2 hourly and 4 hourly records (Figure 6) at -0.2, -0.2, -0.1 and -0.1 percent, respectively. The average percent difference for the second 4 hourly record was -0.2 percent. The average percent difference remains less than 10 percent even up to as few as two records per day. When the record goes from two records per day (midnight and noon) to one record per day (at 8 a.m.), the average percent difference jumps from -1.9 percent to 12.7 percent.

If the average percent difference equals zero for a data set, the volume calculated for this data set has little or no systematic error as a result of the fewer records per day used to determine volume. Figure 6 shows that the average percent difference is slightly negative for all data sets except for the one record per day set, which jumps to 12.6 percent. This means that for all data sets except the daily, the number of flow peaks missed are about the same as the number of flow valleys missed and little systematic error results. Using one reading each day at 8 a.m. to calculate volume results in a positive systematic error of about 12.6 percent. Daily hydrographs based on hourly records for this site type indicates that the magnitude and direction of the systematic error depends upon the time at which the reading is taken. In fact, a graph of the average percent difference versus the

time of the day can be expected to have the same shape as the daily hydrograph shown earlier in Figure 2.

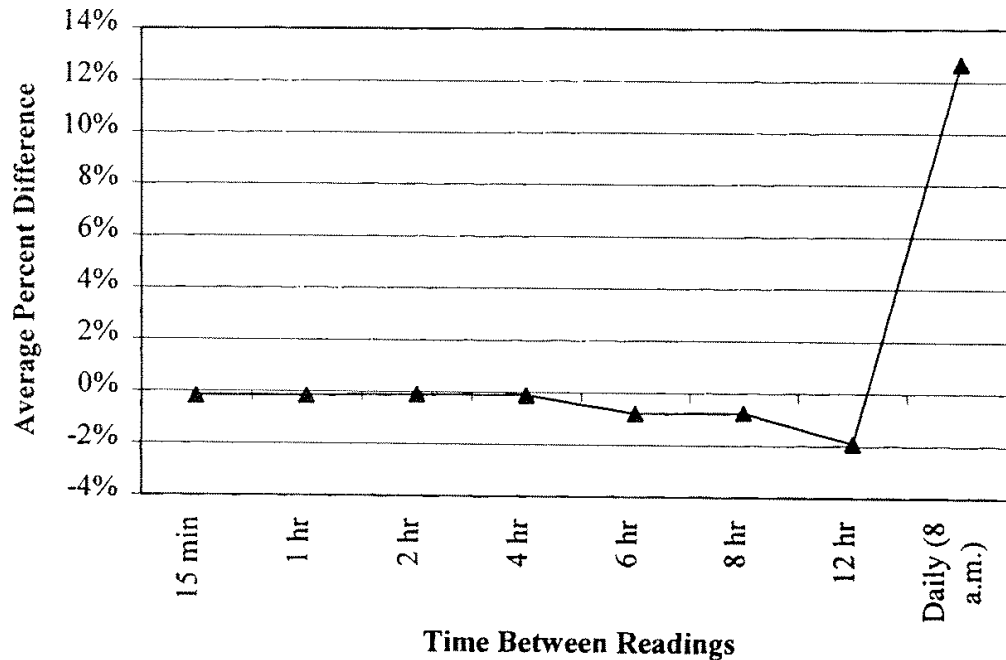


Figure 6. Yearly Volume Differences Compared to Digitized Record.

All the percent differences of each annual volume for each site (36) taken together form a population ( $n=36$ ) for each interval-based data set. The range of the population of percent differences (Table 2) is essentially the same for 15-minute, hourly and two hourly values. The range more than doubles from 13.8 to 30 percent for records every eight hours to every 12 hours and increases more than 6 times to 199.8 percent for one record per day.

The standard deviation of the percent difference approximately doubles as the number of records in a day decreases from four to three. Thus, more volumes are far from the average percent error. The negative skewness in all populations of the percent differences shows that the populations have an asymmetric tail extending toward values that are more negative. This indicates that slightly more flow peaks than flow valleys are missed by the sampling scheme. The positive kurtosis shows that all these distributions are more peaked than the normal distribution, or more of the percent errors are near than mean than expected in a normal population.

Table 2. Percent Difference, Statistics of Year-Site Volumes (n=36).

	15 min	Hourly	2 hours	4 hours	6 hours	8 hours	12 hours	Daily at 8 a.m.
Minimum	-4.5	-5.0	-4.7	-8.3	-10.8	-25.9	-23.4	-100.2
Maximum	4.6	4.1	3.8	4.0	3.0	4.1	8.4	99.6
Range	9.1	9.1	8.5	12.3	13.8	30.0	31.8	199.8
Average	-0.2	-0.2	-0.1	-0.1	-0.8	-0.8	-1.9	12.7
Std. Dev.	2.1	2.1	2.1	2.4	2.8	5.2	5.3	36.3
Skewness	-0.40	-0.60	-0.68	-1.18	-1.47	-3.39	-1.62	-0.29
Kurtosis	0.07	0.02	-0.05	2.51	3.36	14.85	6.41	2.12

Figure 7 is a histogram of the percent difference population of the 15-minute data set. The cumulative distribution of the data and the normal population are plotted and showing that this population approximates a normal population.

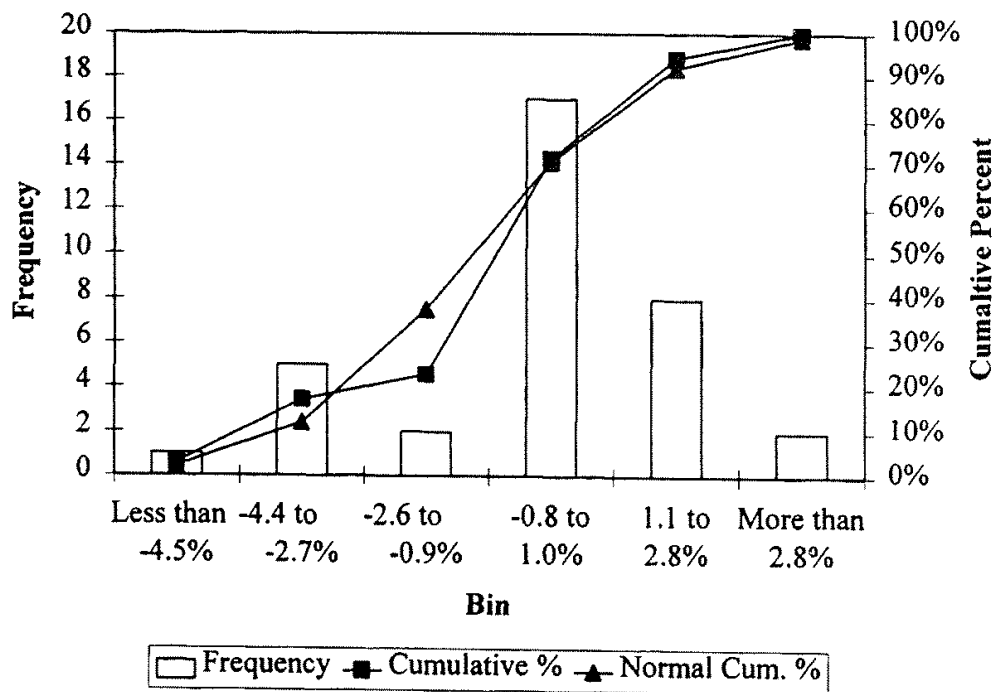


Figure 7. Histogram of the Population of Yearly Volume Percent Differences for the 15-Minute Data Set.

**Monthly and Daily:** The average percent differences for monthly volumes is greater than that for yearly volumes and shows the same pattern. The daily volume average percent differences show the same pattern, but are a negative four to five percent different. As for the yearly, the monthly and daily average percent volume differences show little improvement as the number of records per day

increases. However, the variability of the percent differences increases significantly.

#### Flow Variability versus Average Percent Difference

The relationship of flow variability to the average percent difference as a function of the coefficient of variability (Figure 8) was very weak for all data recording intervals. The average percent difference in yearly volume increased as the coefficient of variability increased, however, the best correlation coefficient was only 0.25.

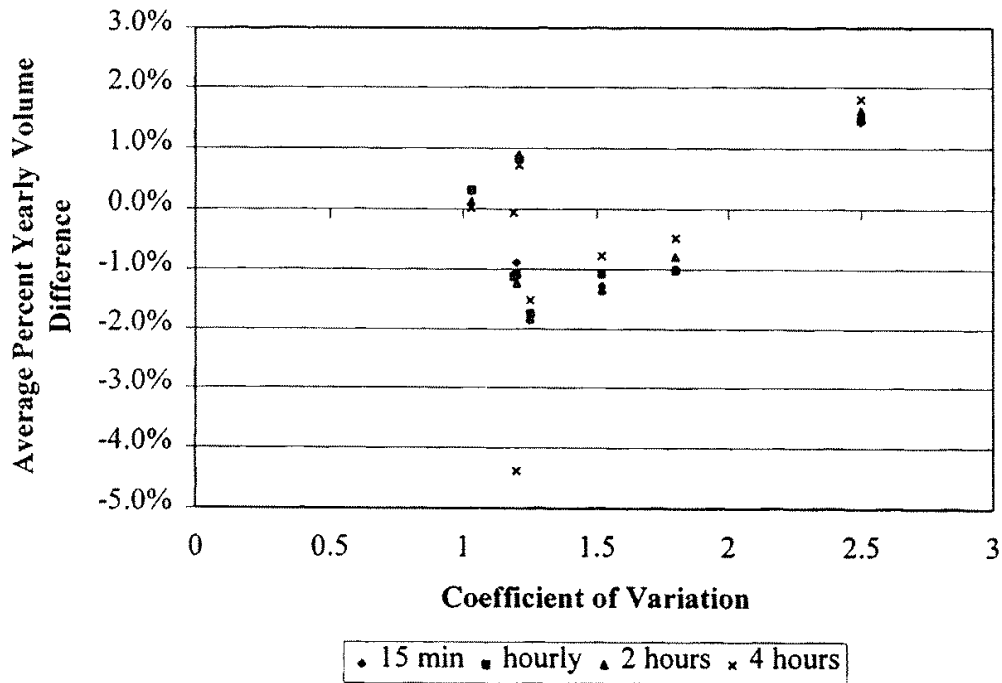


Figure 8. Percent Difference in Yearly Volume of Changing Recording Intervals as a Function of Flow Variance

The relationships between the maximum flow and the average points digitized per year were very weak even for the 15-minute and hourly data recording intervals. The maximum flow (flow range) and the average points digitized per year provide no information about the average accuracy that can be expected for yearly volumes.

## DISCUSSION

This study only looked at eight sites and 36 station-years of data. However, the sites selected were sites with high variability. These types of sites with high

variability are the ones we would expect to benefit from a detailed 15-minute record. However, the results indicate that at the daily, monthly and yearly volumes calculated from an hourly record were not appreciably less accurate than those calculated from the 15-minute record. For stream flow and canal headings sites that can be expected to have flows that vary less than these sites, hourly values may be sufficient.

Knowing the ultimate purpose of the data is vital to deciding what time interval to use for data collection. This study indicates that if yearly volumes are the sole purpose of the data collection, 2 hour intervals (12 points a day) may be enough. However, if the data may be used for other system studies, as experience has shown is likely, hourly values are probably a good idea.

This study shows that as long as the recording interval is evenly spaced and maintains at least two points per day, an accurate total yearly volume (all sites combined) can be obtained for this group of sites. This is because over an entire year and many sites, the flow peaks and valleys missed by the sampling scheme have a chance to cancel themselves out. However, more studies are needed to determine if this can be generalized to other sites.

However, if one is interested in the yearly volume at an individual site, the range of percent difference becomes important. The increase in range of percent difference as the points recorded per day decreased from 12 points (values recorded every two hours) to 12 points (values recorded every two hours) seen in this study indicates that the sampling interval should be at least one point every two hours.

Storing hourly values rather than 15-minute values is a 75 percent decrease in storage requirements. Perhaps more important in this era of cheap disk storage, is the increase in performance of databases when queried for analysis. If knowing the maximum flow at a site is important, perhaps recording 15-minute values is justified. For the sites in this study, even maximum flows from an hourly record were only 81 percent of the maximum digitized flows.

## CONCLUSIONS

1. Two-hour and one-hour intervals were as accurate as 15-minute intervals for determining yearly volume totals.
2. The greater the interval between data points, the greater the variability in the yearly, monthly and daily volume percent differences.
3. The maximum flow from the 15-minute record averaged 98 percent of the digitized maximum flow.

4. The maximum flow from the hourly and two-hourly record averaged 81 and 73 percent, respectively, of the digitized maximum flow.
5. An hourly interval may be best for developing daily hydrographs and other analyses.
6. A 15-minute record may be necessary if accurately determining the maximum flow at a site is important.