SCADA and Related Technologies for Irrigation District Modernization.
Proceedings from a USCID Water Management Conference, Vancouver, WA.
pp. 123-132

TEN YEARS OF SCADA DATA QUALITY CONTROL AND UTILIZATION FOR SYSTEM MANAGEMENT AND PLANNING MODERNIZATION

Bryan Thoreson[1]
Lindsay Mullins[2]
Anisa Divine[3]
Grant Davids[4]

ABSTRACT

Supervisory Control And Data Acquisition (SCADA) Systems can be used to collect the historical flow data that is needed to support good water management decisions; however, insufficient attention is typically given to data quality-control and storage. Initially driven by the unique requirement to verify water savings achieved by a water conservation program, the Imperial Irrigation District (IID) implemented a data warehouse with attendant quality-control and reporting applications as a SCADA data repository. Over the past 10 years, IID's Water Information System (WIS) has gained recognition for the benefits it now provides to ordinary operations, including reduced staff time, reduced reporting costs, improved data accuracy and overall improved water management. IID's SCADA system now includes 250 remote sites that feed data into the WIS.

This paper briefly describes the IID WIS and the primary functions it currently supports pertaining to water conservation administration, support of daily operations, strategic analysis and data reporting. The quality control procedures for data management are presented, along with recently implemented improvements. Performance of quality-control algorithms is also discussed.

[1] Project Manager, Davids Engineering, Inc.; 1772 Picasso Avenue, Suite A; Davis, CA 955616-0550; bryant@de-water.com
[2] Project Engineer, Davids Engineering, Inc.; 1772 Picasso Avenue, Suite A; Davis, CA 955616-0550; lindsay@de-water.com
[3] Senior Planner, Imperial Irrigation District, P.O. Box 937, Imperial, CA 92251
[4] Principal Engineer, Davids Engineering, Inc.; 1772 Picasso Avenue, Suite A; Davis, CA 955616-0550; grant@de-water.com

SCADA and Related Technologies for Irrigation District Modernization.
Proceedings from a USCID Water Management Conference, Vancouver, WA.
pp. 123-132

INTRODUCTION

Sound water management requires accurate flow measurement and complete historical flow records. Supervisory Control And Data Acquisition (SCADA) systems can be used in irrigation distribution systems to collect water level and flow data, as well as to automate water system structures and acquire data on timing and rate of flow. Monitoring and data records for operational analysis are often cited as advantages of installing a SCADA system. However, planning for data storage and analysis is too often an afterthought in SCADA system design.

All irrigation water purveyors use some form of monitoring to ensure that their service goals are met. Monitoring systems vary widely in scope and intensity depending on service area size, location, physical characteristics and layout among other factors. Most purveyors are under some level of pressure to improve their water management. Achieving and documenting these improvements will, in many cases, require improved monitoring systems.

Effective monitoring systems provide flow data that can be transformed into useful information. Accurate and complete data, along with well designed transformation processes (analyses), are required to provide reliable decision making information. However, the mechanical and electrical equipment components that comprise SCADA systems experience problems and failures from time to time. Experience has shown that it is necessary to regularly check incoming data to ensure prompt repair of failures.

This paper describes the experience of Imperial Irrigation District (IID) with quality control of data acquired through a SCADA system that presently has over 250 remote sites. IID's intensive quality control program was prompted by a unique requirement to verify water savings associated with the 1988 Water Transfer Agreement between IID and the Metropolitan Water District of Southern California (MWD). Meanwhile, the importance of high quality flow data for developing water balances and improving water management continues to gain recognition within IID.

First, some of the data uses and benefits of the quality-control and storage system are described. This discussion is followed by a brief description of the quality-control system and recent improvements based on ten years of experience.

DATA USE

IID has recognized the importance of placing SCADA (and other) data in a central data repository. This data repository, called the Water Information System (WIS), is a state-of-the-art, computer-based system that is used to store, process, retrieve, analyze and report high-quality, historical flow data. The WIS receives data from a number of sources, applies quality control routines to assure that the data stored are complete, consistent and accurate, and generates both standard and special information reports (Figure 1).

The WIS was developed as a data warehouse for storing large sets of quality-controlled, historical flow data needed for analyses to support the IID/MWD Water Conservation Program. However, it has been expanded to support other crucial IID functions and has become an integral part of routine IID operations. It enhances staff productivity, improves service and reduces operating costs. The WIS facilitates automatic report preparation, thereby increasing staff productivity, decreasing the cost of responding to information requests (such as from regulatory agencies), increasing IID analytical capabilities and, ultimately, increasing resource management opportunities for IID.

Primary WIS Functions

The WIS serves four primary functions that save the IID Water Department both time and money. The savings result from the ability to produce standard and special reports literally at the "push of a button," thereby filling data and analytical requests efficiently. More importantly, the data used for these reports and analyses are of the best possible quality, which means that IID decisions are based on highly accurate data. The four primary functions are described below.

IID/MWD Water Conservation Agreement Support: Administration of the IID/MWD water conservation agreement requires that IID prepare three major water reports annually.

1. Processed Flow Data,
2. Projected Water Conservation Savings, and
3. Systemwide Monitoring Program Annual Report.

WIS applications have been developed that allow the first two documents to be prepared with a minimum of staff time. The Processed Flow Data and Projected Water Conservation Savings reports have previously been described in Archer, et. al (1999) and Thoreson, et. al (2003b). Another WIS application that is under development will enable efficient production of the third report. Without the WIS, these reports would require months of IID staff time to prepare.

IID Daily Water Operations Support: Routine water operations require that daily and monthly reports be prepared summarizing flow and water level information at many water control sites within IID. Prior to implementation of the WIS, operators were required to transcribe all of this data by hand from computer screens for transfer to a spreadsheet, or the necessary values could be extracted by manipulating countless individual data files. Either way, the process consumed much staff time and was prone to human error. To prepare monthly reports, operators had to wait for data to be received from other operating sections and

SCADA and Related Technologies for Irrigation District Modernization. Proceedings from a USCID Water Management Conference, Vancouver, WA. pp. 123-132

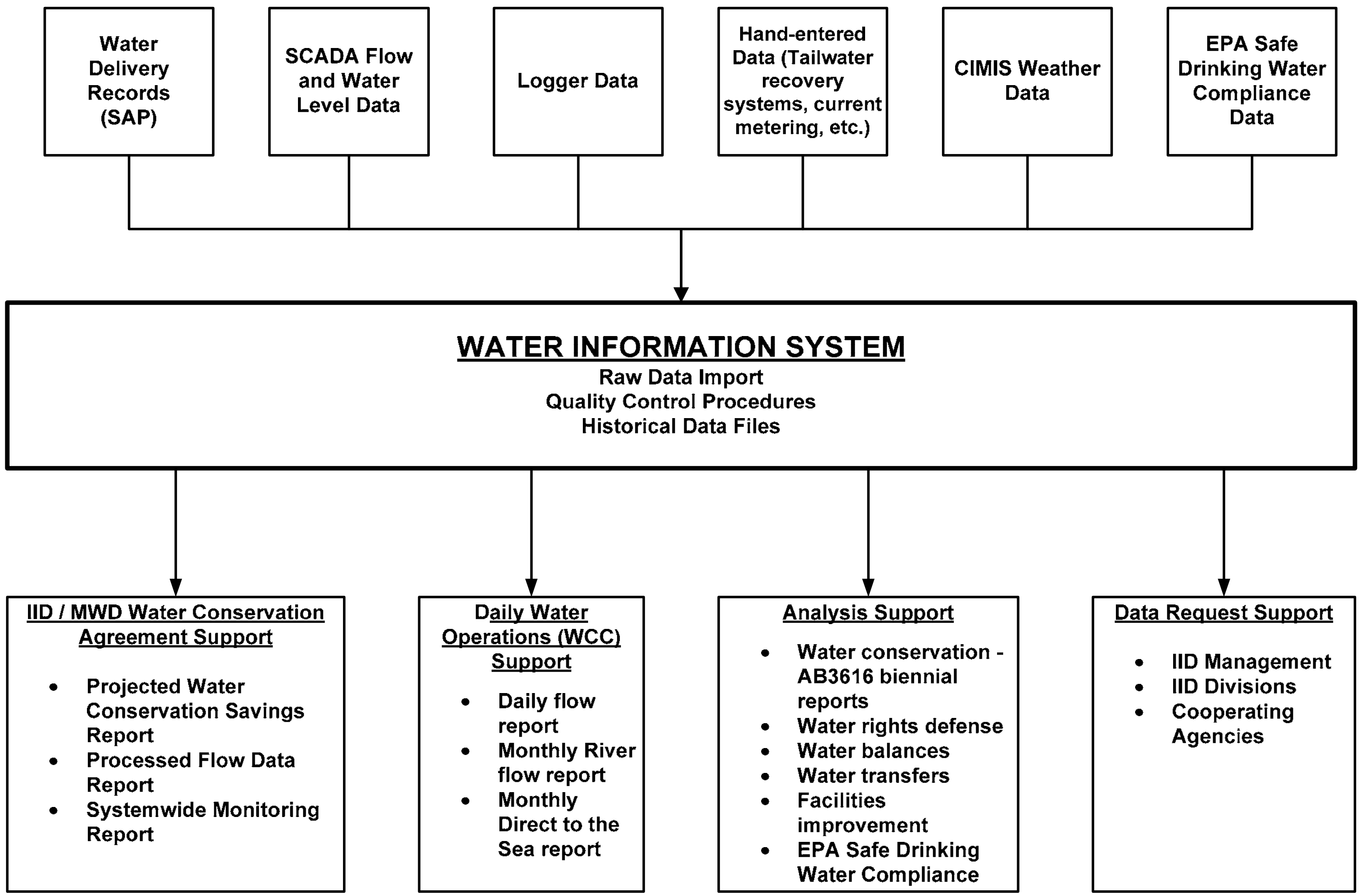

Figure 1. Schematic Diagram of IID's WIS

then manually enter it into spreadsheets. With the WIS, many of these reports are now prepared automatically, reducing IID's labor requirements and improving the timeliness, consistency and accuracy of reporting.

Water Operations historical flow and crop acreage data have been stored electronically in IID's business systems since 1986. Implementation of SAP (an enterprise-wide computer business process software package) in 2000, however, did not include these historical data. Hence, WIS applications have been developed to collect these data from SAP on a regular basis and add them to the pre-existing historical base of information, all residing on the WIS. The WIS and the quality-controlled SCADA data it stores are currently being integrated into a broader Water Management System that will incorporate water order entry and scheduling of main canal flows and customer delivery (Young, et. al, 2005).

Analysis Support: IID is constantly faced with the need to analyze its system operation and water use. Some of these analyses support historical analysis and planning efforts, while others respond to or support IID's efforts to comply with regulatory or quasi-regulatory requirements, such as those stemming from the federal Clean Water Act and the State Efficient Water Management Practices Act (AB3616). Furthermore, operations data are analyzed to support routine as well as strategic planning of facilities improvements and water conservation measures. The data analysis and reporting features of the WIS support all of these activities.

Four examples of operations analyses that have benefited greatly from the availability of quality controlled flow data in an easily accessible format are: 1) the Vail Canal operations decision-support system initial study and development (Thoreson, et al, 2003a), 2) the lower Westside Main Canal operations study, and 3) a system improvement study that is currently underway. Planning for system improvements in response to the 2003 Quantification Settlement Agreement, which includes water transfer to San Diego, is expected to be benefit greatly by this quality controlled data set.

Data Requests: The data available on the WIS allow IID to respond to both internal and external data requests in a timely manner with minimum interruption to IID staff regular duties. Importantly, identical data requests result in identical responses.

DATA QUALITY CONTROL

History and Basic Procedures

The IID SCADA system initially had three types of sites: 1) remote monitoring sites, 2) small canal sites and 3) major sites (Villalón and Korinetz, 1998). The remote monitoring sites provide level and flow information via telemetry and are not used for operations. The small canal sites provide level and flow information via telemetry and local control. These sites are not monitored by operations personnel. The major sites enhance operation of the main canals by providing remote monitoring and control capabilities for operators in addition to level and flow data via telemetry.

Once the remote monitoring and small canal sites had been installed and commissioned, they continued to collect data. After some months of data collection during the first year of operation, text files of data from these sites were manually reviewed. The review showed significant gaps and anomalies in the data due to various system failures. The primary failures included: 1) radio communications outages, 2) computer downtime on the receiving end, 3) battery failure at the remote location and 4) sensor failure at the remote location.

These problems resulted in the development of a FORTRAN program that was run weekly to check the data as part of a weekly data review process. Due to the number of sites, this process generally took about one-half of a person-day. It was also found that, with weekly data review, a site could be inoperable for up to two weeks before a problem was discovered and a crew dispatched to repair it. Thus, data gaps of longer than a week could occur. The WIS was developed to handle the large amount of data that was being collected and apply quality control checks more frequently, thereby eliminating long data gaps and reducing staff time requirements.

The WIS was implemented in an Oracle database, and the FORTRAN program was reprogrammed as a PL/SQL procedure in the Oracle database. The SCADA system puts data in text files that are loaded into a table on the WIS early each morning. The PL/SQL procedure checks the data, moves them to a quality-controlled table and provides a report listing sites with more than five percent of the data missing or otherwise coded.

The daily programmatic quality control is followed by a monthly manual (human) review. Upon completion of the monthly review, the 15-minute data are used to calculate hourly and daily volumes in the WIS. Daily volumes include estimates for missing 15-minute values and, thus, provide a complete record for use in analyses and reporting.

Results

The WIS identifies days when more than 50 percent of the flow volume on that day was estimated. The FORTRAN program was used to quality control data from 1993 through 1995. The PL/SQL program was used beginning in 1996 and is simply a translation of the FORTRAN program. In other words, the same quality control checks were completed based on the same criteria. Thus, the percent of total days with estimates can be tracked from 1993 through 2004 (Figure 2).

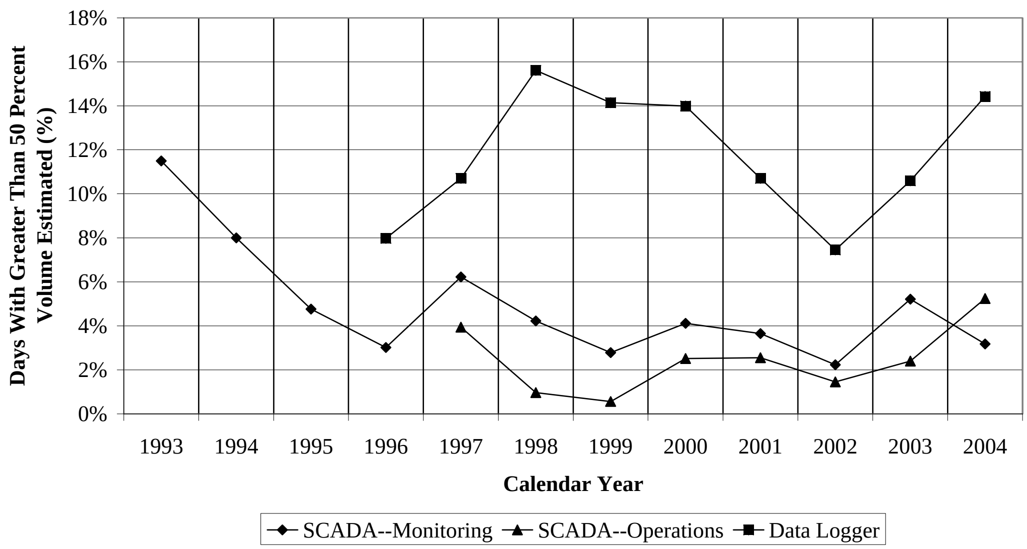

Figure 2. Percent Days with more than 50 Percent of Volume Estimated for Sites with SCADA Telemetry and Data Loggers.

For SCADA sites, the percent of total days with more than 50 percent of the flow estimated dropped steadily from nearly 12 percent in 1993 to about three percent in 1996, remaining between two and four percent except for 1997 and 2003. Operations sites were not initially included in the WIS because they were monitored by dispatchers (distribution system operators) and a technician was dispatched almost immediately to repair these sites as needed. Beginning in 1997, theses sites were included in the WIS and they will be part of the broader integration of order entry and scheduling in the Water Management System. Because of immediate attention to problems, the number of days with more than 50 percent of the volume estimated is less for the operations sites than for SCADA or data logger sites every year except for 2004.[5]

IID also has data loggers that are visited roughly once every two weeks when data are downloaded. These data have also been loaded into the WIS since 1996 and they are also coded when more than 50 percent of the flow on a given day has been estimated. Primarily due to the bi-weekly visits compared to daily review of the SCADA data, their record has varied from just under eight percent to almost 16 percent of the days having more than 50 percent of the flow volume estimated.

[5] In 2004, a single site classified as an operations site because the data are included in operations reports had a maintenance problem for much of the year. This site measures flow in the Alamo River coming into the IID service area from Mexico, which averages around one cfs and does not affect customer service. Thus it had a low priority for repair.

Replacing data loggers with SCADA telemetry reduces the number of site visits and consequent labor cost and also improves data quality. These facts along with reduced WIS maintenance cost due the elimination of the data logger WIS applications culminated in a decision to eliminate data loggers in favor of SCADA equipment (Figure 3).

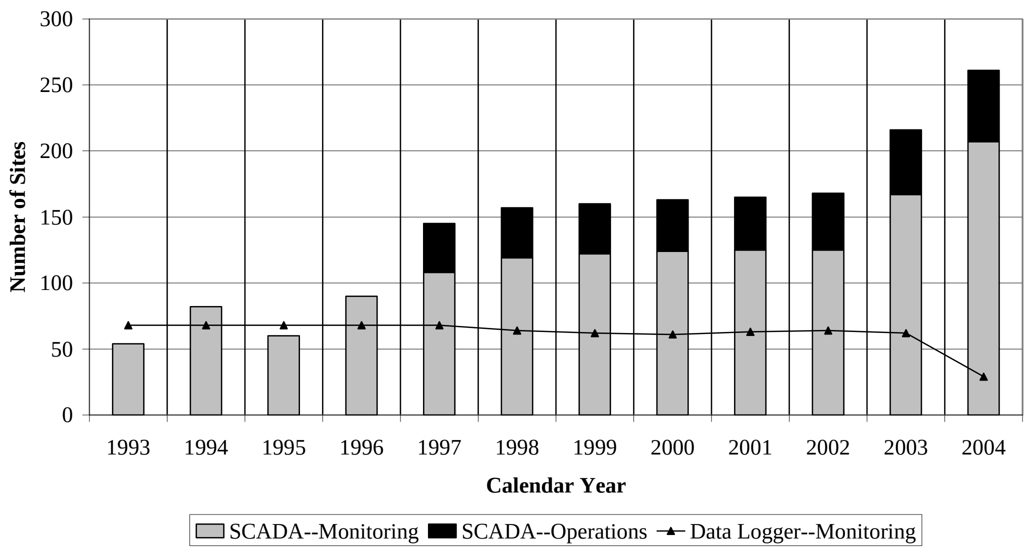

Figure 3. Annual changes in the number of SCADA Monitoring, SCADA operations and data logger monitoring sites on the WIS.

DATA QUALITY CONTROL IMPROVEMENTS

Over the last 10 years of quality control, including nine with the current PL/SQL procedure, the program has performed superbly. However, with time the need for some improvements were identified. These improvements include checking all data values prior to computation of flow, breaking the procedure into three procedures that run consecutively and correcting minor conceptual inconsistencies in the use of parameters to check reported levels and flows. In addition to these improvements, a fourth procedure was developed to consistently identify large changes in levels that occasionally occur for a duration of one or two 15-minute readings. These level changes are called "spikes" and previously were identified by reviewing graphs as part of monthly manual (human) quality control. In a typical month, between 50 and 100 spikes are identified. Although they have little effect on total flow volume, spikes raise concerns about data integrity and distract attention from more serious issues. It was decided to develop the logic to reduce human intervention in this task.

Spike Criteria Development

The WIS criterion for identifying a spike is: "The water level decreases or increases by greater than 0.2 feet from one record to the next, remaining greater

SCADA and Related Technologies for Irrigation District Modernization.
Proceedings from a USCID Water Management Conference, Vancouver, WA.
pp. 123-132

than 0.2 feet different than the initial level for four records or less, and the record immediately following the last record at the new level is within 0.05 feet of the trend line of the original reading." A procedure was developed to mimic this by comparing each new level to a running average of preceding levels. When the new level differs from the running average by a prescribed amount, the new level is identified as a spike. Using existing records with human identified spikes, the number of records in the running average and the prescribed level difference were varied. Table 2 shows the results from this work for a moving average of five records. Similar tables were developed for other moving averages. The five-record moving average and a prescribed difference from the moving average of 0.2 ft. were selected as the best criteria for identifying most of the spikes while incorrectly identifying as spikes only a very small number of good records. A second set of existing records with human-identified spikes was used to validate the spike procedure.

Table 1. Development of spike criteria based on manually identified spikes.

Level Difference, feet	Correctly Identified as Spikes, %	Incorrectly Identified as Spikes, %
0	100%	100%
0.1	97%	64%
0.2	90%	2%
0.3	85%	0%
0.4	82%	0%
0.5	78%	0%
0.6	69%	0%

Using these criteria, the logic correctly identified 99 percent of the manually identified spikes and incorrectly identified only three percent of the good records as spikes in the validation data set.

Testing on the Entire Data Set

When testing the routine on the entire data set, it was noted that procedure was occasionally incorrectly identifying actual flow changes at the heads of canals as spikes. The change of flow at a canal heading often results in a large level change from one 15-minute record to the next. Since the data are quality-controlled in 24-hour blocks, it was possible to add a forward-looking moving average to the backward-looking moving average. This eliminated the misidentification of changes in flows at headings as spikes.

SCADA and Related Technologies for Irrigation District Modernization. Proceedings from a USCID Water Management Conference, Vancouver, WA. pp. 123-132

CONCLUSIONS

Monitoring and data reporting for system control and operations analyses are often cited as advantages of a SCADA system. Experience at the IID has shown that SCADA sites used solely for monitoring must be part of a rigorous quality control program, where data are reviewed, preferably daily, and transformed into information on a regular basis for the information to be accurate and, thus, useful. The existence of a large, quality-controlled data set has found many uses and enabled IID staff and others assisting them to increase their productivity.

REFERENCES

Archer, M.C., A. J. Divine and B. P. Thoreson. 1999. Irrigation Flow Data Collection, Quality Control, Flow Computation and Site Monitoring. ICID 17th Congress, Granada, Spain. Volume 1A, p.289-299.

Thoreson, B. P.; A. J. Divine and M. C. Archer. 1999. IID Water Information System – Irrigation Database Principles and Management. In *Contemporary Challenges for Irrigation and Drainage*, Proceedings from the USCID Water Management Seminar, Denver, CO. pp. 41-54.

Thoreson, B. P.; C. Villalón, M. Kidwell and K. Dimmitt. 2003a. Initial Development of a Decision Support System for Operating a Lateral Interceptor Reservoir System. In *Water for a Sustainable World – Limited Supplies and Expanding Demand*. Proceedings from the USCID Second International Conference on Irrigation and Drainage, Phoenix, AZ. pp. 313-322.

Thoreson, B. P.; P. Kosciusko, J. Burns, G. Davids, K. Dimmitt, J. Keller and M. King. 2003b. Water Conservation: Savings Determination in Near Real-Time. In *Water for a Sustainable World – Limited Supplies and Expanding Demand*. Proceedings from the USCID Second International Conference on Irrigation and Drainage, Phoenix, AZ. pp. 747-756.

Villalón, C. Z. and J. H. Korinetz. 1999. IID/MWD Water Conservation Program – Improved Irrigation Water Management Through System Automation. In *Contemporary Challenges for Irrigation and Drainage*, Proceedings from the USCID Water Management Seminar, Denver, CO. pp. 1-28.

Young, G., B. Thoreson, A Baro and C. Villalón. 2005. Development of a Water Management System to Improve Management and Scheduling of Water Orders in Imperial Irrigation District. *Water District Management and Governance.* Proceedings from the USCID Third International Conference on Irrigation and Drainage, San Diego, CA. pp. 253-261.