Experimental Evaluation of Data Aggregation Methods Applied to Soil Moisture Measurements

Camilo Lozoya, Gilberto Mendoza, Carlos Mendoza, Valentín Torres and Miguel Grado
Instituto Tecnológico y de Estudios Superiores de Monterrey, Campus Chihuahua
31300 Chihuahua, Mexico
e-mail: camilo.lozoya@itesm.mx

***Abstract*—The implementation of precision irrigation systems represent an interesting opportunity in order to increase water efficiency for agricultural, gardens and parks areas. Precision irrigation may include a large group of sensors, one controller unit and few or only one actuator elements. Soil moisture is primarily used as the main process variable in the control loop, meanwhile wireless sensor networks are often required to communicate soil moisture data from different sensor nodes distributed over a usually large irrigation area to the central controller unit. This implies that many and different soil moisture values are simultaneously acquired during the sensing process. However a single representative and consistent soil moisture value for the complete irrigation area is required for control purposes, which can be obtained through a data aggregation process. This paper presents an experimental evaluation of different data aggregation methods applied to soil moisture measurements in a closed-loop precision irrigation system installed in an experimental field.**

I. Introduction

Precision irrigation refers to the application of the correct amount of water in the correct place at the right moment, in order to make an efficient use of this valuable natural resource. Typically, precision irrigation has been implemented, in large irrigation areas, in the form of a closed-loop control system based on soil moisture sensing with a water flow valve control as an actuator. Precision irrigation has been mostly used for agricultural purposes, as in [1] and [2]; but also in irrigation systems used in parks and gardens, within the scope of the smart cities framework [3].

In precision irrigation, soil moisture measurements may involve a large group of nodes distributed along a large piece of land, transmitting periodic data to a central controller, then the controller evaluates the data and decides to activate or not the irrigation. Usually the watering activity is conducted by using only a few or even a single water valve or pump. Therefore, there can be a large amount of soil moisture values, and in most of the cases one single actuator. In order to implement a closed control loop, a single representative value for soil moisture measurements is required.

Simple data aggregation techniques (e.g., mean, median, maximum and minimum) have been typically used to reduce the amount of data traffic, within a wireless sensor network domain. In [4], data aggregation is used to implement a fuzzy-logic control in a precision irrigation system. In [5], the application of Kalman filter algorithm for data aggregation is proposed to improve the data accuracy for greenhouse environment monitoring. In [6], a water requirement model is presented for precision irrigation by using data aggregation techniques. The work presented in this paper provides an experimental evaluation of different simple data aggregation methods applied to soil moisture measurements in a closed-loop precision irrigation system installed in an experimental field, in order to identify the level of accuracy and robustness provided by each method in different scenarios including when there are faulty sensors.

The rest of this paper is structured as follows. Section II introduces the theoretical aspects of closed-loop irrigation and data aggregation methods. Section III presents the implementation of the precision irrigation closed-loop control system. Sections IV describes the experimental setup and results. Section V concludes the paper.

II. Preliminaries

A. Closed-loop irrigation

The process dynamics of any irrigation system can be best described by using the hydrological balance model [7]. This model establishes that a change in water storage during a time period in a specific location is the result of water inflows (irrigation, rainfall, capillary rise) minus the water outflows (evaporation, plant transpiration, water runoff and deep percolation).

A closed-loop irrigation system can be implemented in order to minimize the control signal (effective irrigation) while keeping soil moisture under specific thresholds (avoiding water stress). Based on the hydrological balance, the process dynamics for an irrigation system can be described as a block element with two inputs (effective irrigation and external factors) and one output (soil moisture) where the effective irrigation is the control action, and the external factor represents the sum of the remaining elements (rainfall, capillary rise, evaporation, plant transpiration, water runoff and deep percolation).

B. Data aggregation

Data aggregation refers to the collection of raw data from different data sources in order to simplify them and obtain less voluminous and more refined data, but this data summarization may represent accuracy loss. If several soil moisture sensors are placed in a large irrigation area, then different measurements values are obtained when simultaneous readings are conducted at different locations, this is because irrigation may not be uniform, but also because the irrigation field may present different characteristics at different zones, i.e. different run-off, different percolation and different capillary rise. In order to implement a closed control loop, a single estimated representative value for soil moisture measurements needs to be obtained trough data aggregation when n sensors area available from measurements. Simple data (or low level) statistical based aggregations methods are considered for the evaluation, since aggregation process takes place in a microcontroller with constrained resources (memory, clock frequency) and additional computing loads (such as the execution of the control algorithm). Six different data aggregation methods for soil moisture measurements are evaluated in an experimental field, in order to find out which method provides a better accuracy and robustness under different scenarios.

- Mean value: The estimated aggregated value at instant k is obtained with the average value for n sensors,

$$\hat{\theta}_k = mean\{\theta_k^1, \theta_k^2, ..., \theta_k^n\}. \tag{1}$$

- Median value: Estimated value is calculated with the median for n sensors,

$$\hat{\theta}_k = median\{\theta_k^1, \theta_k^2, ..., \theta_k^n\}. \tag{2}$$

- Minimum value: The minimum value for the average and median value for n sensors is considered as the representative value,

$$\begin{aligned}\hat{\theta}_k = min\{&mean\{\theta_k^1, \theta_k^2, ..., \theta_k^n\},\\ &median\{\theta_k^1, \theta_k^2, ..., \theta_k^n\}\},\end{aligned} \tag{3}$$

  maximum value is not considered in the evaluation since for practical irrigation purposes is better to underestimate soil moisture than overestimate.

- Exponential Average value: This is a window average method which smoothes and averages a sequence of values along a time series,

$$\hat{\theta}_k \;=\; \alpha\ mean\{\theta_k^1, \theta_k^2, ..., \theta_k^n\} + (1-\alpha)\hat{\theta}_{k-1}, \tag{4}$$

  where α is a constant gain which defines the weight for the current value, and $1-\alpha$ defines the weight for the historical values.

- Central sensor: The readings from the central sensor, i.e. sensor located in the middle of the irrigation area, are considered as the representative values,

$$\hat{\theta}_k \;=\; \theta_k^0, \tag{5}$$

  where θ^0 is the central sensor.

- Optimal sensor: The optimal sensor, i.e. sensor with the smallest error, is selected as the representative value. For this method an a priori evaluation is required to identify the optimal sensor,

$$\hat{\theta}_k \;=\; optimal\{\theta_k^1, \theta_k^2, ..., \theta_k^n\}. \tag{6}$$

As it can be notice the last two methods (optimal sensor and central sensor) are not really aggregation methods since only one sensor is used for measurements, but they were included since it is a widespread approach in closed-loop irrigation to consider that one single sensor is enough to obtain a representative soil moisture value for the complete area, assuming a small plain piece of land with similar soil and vegetative conditions.

III. Implementation Approach

A. *Networked control system*

The closed-loop irrigation system is comprised by a group of nodes distributed over an irrigation area in order to implement a networked control system. The wireless sensor network is implemented over the protocol IEEE 802.15.4[1].

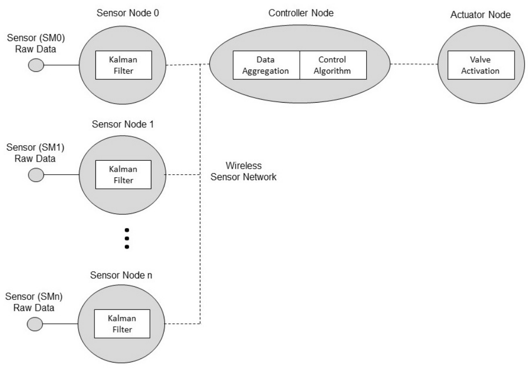


Fig. 1: Tasks and communication sequence between nodes

As depicted in Fig. 1, each sensor node obtains raw data from soil moisture sensors from a specific location in the irrigation area. Sensor nodes removes noise from reading signal by using Kalman filter algorithm. The simplified first order Kalman filter is used to estimate i-soil moisture value.

$$\begin{aligned} K &= \tfrac{P_{k-1}^i + Q}{P_{k-1}^i + Q + R}\\ \theta_k^i &= \theta_{k-1}^i + K(y_k^i - \theta_{k-1}^i)\\ P_k^i &= (P_{k-1}^i + Q)(1-K). \end{aligned} \tag{7}$$

where P^i is the estimate of the covariance error, K is the Kalman gain, y^i is the sensor reading value, R is the covariance value of the measurement noise, and Q is the covariance process noise.

The controller node periodically (every minute) receives filtered soil moisture data from the sensor node, Then the controller combines the multiple simultaneous measurements in order to obtain a single representative value for the complete irrigation area by executing the corresponding data aggregation algorithm. Based on this value, the controller executes the control algorithm and decides to open or close the irrigation valve by sending a control message to the actuator node.

B. *Hardware implementation*

The sensors and actuator nodes are implemented with low cost boards Arduino Uno based in the microcontroller ATmega328 (http://www.arduino.cc). Soil measurements are conducted by using a Decagon (Decagon Devices, http://www.decagon.com) EC-5 volumetric water content sensor. The sensor is located in a depth of $20cm$ and it has a measurement range from 0% to 60% of volumetric water content with a resolution of 0.1% when calibrated. A Rain Bird irrigation valve (Rain Bird Corporation, http://www.rainbird.com)

[1]http://standards.ieee.org/about/get/802/802.15.html

is used to activate or deactivate the field irrigation. Sprinkler irrigation is implemented since it is a widespread method in parks and in some type of crops. The controller node is implemented with a Microchip dsPIC33 microcontroller (Microchip Technology Inc., http://www.microchip.com) mounted in the Explorer 16 board. The control tasks are executed on the Erika real-time kernel (Erika Enterprise, http://erika.tuxfamily.org). The real-time kernel provides to the microcontroller the capability to schedule several periodical tasks.

C. *Experimental field*

The experimental field corresponds to an irrigation area of $164m^2$ approximately, where sensors are deployed according to a regular hexagonal pattern, as proposed by [8], since this pattern increases the coverage of the network. The area is small and has a regular plain surface covered only with grass in order to have similar soil and vegetative conditions for every sensor, however sprinkler irrigation does not provide uniform watering. As shown in Fig. 2, six sensors were installed to cover one layer of hexagonal areas (sensors SM1 to SM6) and one sensor is installed in the center location(sensor SM0), where the length of each side of the regular hexagon is 3 meters. Decagon EC-5 soil moisture sensors have a very small area of influence ($32cm^2$ approximately), so in comparison with the hexagon area they can be considered as a punctual sensor. Smaller hexagon areas may be considered but either more sensors are required for the same overall area or the overall sensor coverage is reduced.

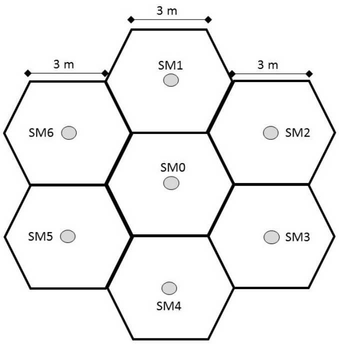


Fig. 2: Sensor deployment pattern over experimental field

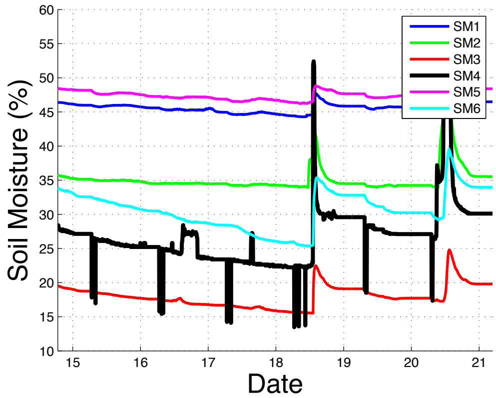
IV. Experimental evaluation

Experimental evaluation was conducted considering three scenarios: (1) typical irrigation when every sensor is working correctly, (2) there is one sensor (SM4) that produces sporadic incorrect readings, see sub-Fig. 3-a, (3) there is one faulty sensor (SM1) that produces random values for a period of time, see sub-Fig. 3-b.

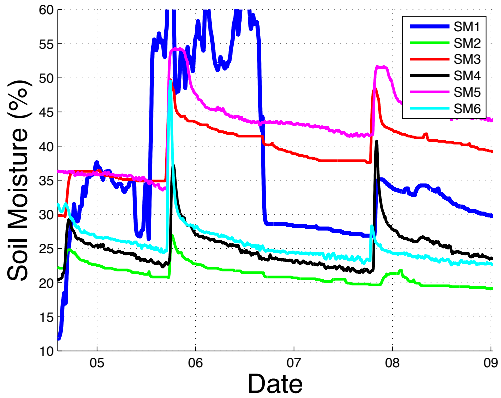
The performance for each data aggregation method is measured with the mean squared error (MSE) defined as,

$$MSE = \frac{1}{m}\sum_{k=0}^{m-1}(\hat{\theta}_k - \theta_k^{ref})^2 \tag{8}$$

where m is the number of data, $\hat{\theta}$ is the estimated value obtained by the data aggregation method, and θ^{ref} is the



(a) SM4 with sporadic incorrect readings



(b) Faulty SM1 with random readings

Fig. 3: Sensors response for (a) scenario 2 and (b) scenario 3

reference value for the complete irrigation area. The reference value is the average value for the seven sensor installed (SM0, SM1,... ,SM6). For scenarios 2 and 3 where there are incorrect readings in a specific sensor, the correct readings are used to calculate the reference value.

An a priori evaluation was conducted in order to quantify the impact of the available sensors for data aggregation. Figure 4, shows the MSE for the cases of 1 to 6 available sensors when using the mean aggregation method. As it can be observed, there is an exponential error increment when the number of available sensor are reduced.

For evaluation purposes, the results presented for each scenario considers that 3 and 4 sensors are available for measurements. These numbers of sensors were selected since the error level is acceptable, and also these numbers are not very close to the 7 sensors used to calculate the reference value. In order to conduct a fair evaluation every possible combination of the sensors from the outside layer (SM1 to SM6) is included in the presented results, just for the central method the inside sensor (SM0) was considered. Therefore 15 sensor location combinations (C_6^4) were evaluated for the 4 sensor case, and 20 sensor location combinations were (C_6^3) evaluated for the 3 sensor case.

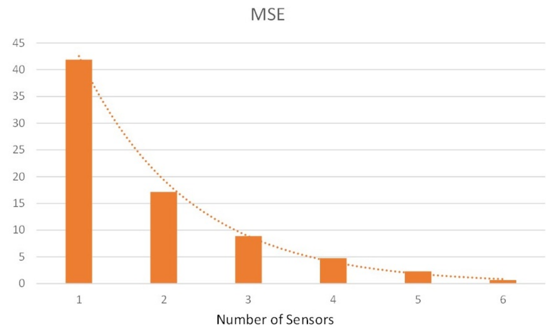


Fig. 4: MSE according to the number of sensors available for aggregation

TABLE I: Data aggregation methods performance (MSE)

Method	Scenario 1	Scenario 2	Scenario 3
3-sensors			
Mean	5.7188	22.2921	21.9336
Median	13.2016	44.9121	43.6235
Minimum	8.7089	28.3615	30.7932
Exponential Average	5.6017	21.4973	21.6320
Central Sensor	28.0382	107.7123	91.5130
Optimal Sensor	12.3708	20.3337	44.5158
4-sensors			
Mean	2.8922	11.4233	13.1710
Median	5.8914	20.0367	20.2991
Minimum	3.8493	14.6593	17.1161
Exponential Average	2.8009	10.7645	12.9086
Central Sensor	28.0382	107.7123	91.5130
Optimal Sensor	10.2568	8.9167	37.7556

As it can be observed from Table I, the central sensor approach produces the worst performance since only one sensor (SM0) is used to obtain the representative soil moisture value. The optimal sensor approach also uses only one sensor, but in this method the sensor is selected not by the location but for its performance, even though the MSE is considerably larger in comparison with the approaches that use several sensors (mean, median, minimum, exponential average). The exponential average method provides the best overall performance and robustness, as it can be seen in Fig. 5 when compared with the optimal sensor approach. Typically median approach has been used to filter out bad readings when one sensor is faulty, however for soil moisture measurements average approaches obtain better results since bad readings have a random behavior.

V. Conclusions

An experimental evaluation was conducted of six different statistically based data aggregation methods applied to soil moisture measurements in a closed-loop precision irrigation system installed in an experimental field. Four methods (mean, median, minimum, exponential average) use several sensor for the estimation, meanwhile two methods use a single sensor approach (optimal sensor, central sensor). Single sensor approach methods obtains the worst performance even for a small, plain and uniform irrigation area. Therefore the use of several sensors is required in order to obtain a representative soil moisture value for the irrigation area. Among the several

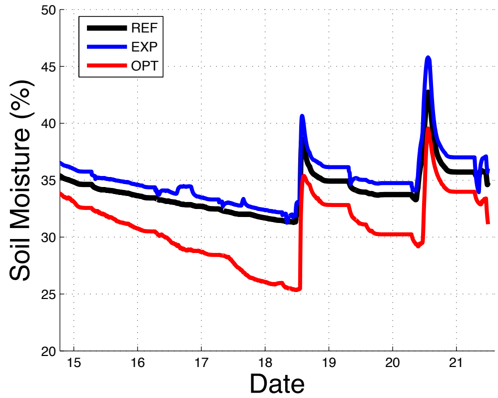


Fig. 5: Sensor response for exponential average and optimal sensor approaches against reference

sensor approach the exponential average obtains the best performance indicating that this data aggregation method is adequate in terms of accuracy and robustness. As a future work the relationship between number of sensor and size of irrigation area will be analyzed.

Acknowledgment

The authors would like to thank the “Fondo Mixto de Fomento a la Investigación Científica y Tecnológica CONACyT - Gobierno del Estado de Chihuahua, México” for funding this project.

References

[1] Kim Y.; Evans R.G. and Iversen W.M., Evaluation of Closed-Loop Site-Specific Irrigation with Wireless Sensor Network,” *Journal of Irrigation and Drainage Engineering*, Vol. 135, No. 1, pp. 25-31, January/February 2009.

[2] Gutierrez, J.; Villa-Medina, J.F.; Nieto-Garibay, A. and Porta-Gandara, M.A.,“Automated Irrigation System Using a Wireless Sensor Network and GPRS Module,” *Instrumentation and Measurement, IEEE Transactions on* , vol.63, no.1, pp.166,176, Jan. 2014.

[3] Sanchez, L.; Galache, J.A.; Gutierrez, V.; Hernandez, J.M.; Bernat, J.; Gluhak, A. and Garcia, T., “SmartSantander: The meeting point between Future Internet research and experimentation and the smart cities,” *Future Network & Mobile Summit (FutureNetw)*, pp.1,8, 15-17 June 2011.

[4] Zheng Y.; Guohuan L.; XiuLi Z.; Qingxin Z., Research and development precision irrigation control system in agricultural,” *Computer and Communication Technologies in Agriculture Engineering (CCTAE), 2010 International Conference On* , vol.3, no., pp.117,120, 12-13 June 2010.

[5] Wang Jing; Xue Heru; Jiang Xinhua, Application of Kalman filtering algorithm in greenhouse environment monitoring,” *Instrumentation and Measurement, Sensor Network and Automation (IMSNA), 2013 2nd International Symposium on*, pp.539,544, 23-24 Dec. 2013

[6] Shaohua W., Research on the Model for Crop Water Requirements in Wireless Sensor Networks,” *Management of e-Commerce and e-Government (ICMeCG), 2012 International Conference on*, vol., no., pp.234,237, 20-21 Oct. 2012

[7] Allen R.G.; Pereira L.S.; Raes D.; and Smith M., Crop evapotranspiration: Guidelines for computing crop water requirements, FAO Irrigation and Drainage Paper 56”. *FAO - Food and Agriculture Organization of the United Nations*. Rome, Italy, 1998.

[8] Hema, N. and Kant, K., Optimization of sensor deployment in WSN for precision irrigation using spatial arrangement of permanent crop,” *Contemporary Computing (IC3), 2013 Sixth International Conference on*, vol., no., pp.455,460, 8-10 Aug. 2013