

Spatio-temporal Soil Moisture Measurement with Wireless Underground Sensor Networks

Xin Dong and Mehmet C. Vuran
Department of Computer Science & Engineering
University of Nebraska-Lincoln, Lincoln, NE 68588
Email: {xdong, mcvuran}@cse.unl.edu

Abstract—In this paper, the estimation distortion of distributed soil moisture measurement using Wireless Underground Sensor Networks (WUSNs) is investigated. The main focus of this paper is to analyze the impact of the environment and network parameters on the estimation distortion of the soil moisture. More specifically, the effects of rainfall, soil porosity, and vegetation root zone are investigated by exploiting a rainfall model, in addition to the effects of sampling rate, network topology, and measurement signal noise ratio. Spatio-temporal correlation is characterized to develop a measurement distortion model with respect to these factors. The evaluations reveal that with porous soil and shallow vegetation roots, high sampling rate is required for sufficient accuracy. In addition, the impact of rainfall on the estimation distortion has also been investigated. In a storm, which carries on a large area and lasts for a long time, the estimation distortion is decreased because of the increase in spatial correlation. Moreover, only few closest sensors are needed to estimate the values of an interested location. These findings are utilized to guide the design of WUSNs for soil moisture measurement to reduce the density of network and the sampling rate of the sensors but at the same time maintain the performance of the system. Moreover, guidelines for designing WUSNs for soil moisture measurement are provided. To the best of our knowledge, this is the first work that establishes tight relations between environmental effects and distributed measurement in WUSNs.

I. INTRODUCTION

The improvements in embedded system design and low-power wireless communication techniques have made the wireless sensor networks (WSNs) an attractive solution for many application areas [1]. Among these applications, Wireless Underground Sensor Networks (WUSNs), which consist of wireless sensors buried underground, have been considered as a potential field that will enable a wide variety of novel applications in the fields of *intelligent irrigation*, *border patrol*, *assisted navigation*, *sports field maintenance*, *intruder detection*, and *infrastructure monitoring* [2]. This is possible by exploiting real-time soil conditions from a network of underground sensors and enabling localized interaction with the soil. When WUSNs are used in intelligent irrigation, the cost for maintaining a crop field can be significantly reduced by gathering soil moisture conditions through autonomously operating underground sensors. Moreover, according to recent United Nations' report [18], nearly 70% of all water usage is related to the agriculture irrigation. Considering the ever growing need for food and the diminishing water resources due to climate changes, development of irrigation systems for efficient water usage is of paramount importance. Compared to

the existing wired underground monitoring techniques, WUSN have several remarkable merits, such as concealment, ease of deployment, timeliness of data, reliability, and high coverage density [2].

Of all the properties of the soil, soil moisture constitutes the physical linkage between soil, climate, and vegetation [5]. In addition, soil moisture is the major factor that impacts the quality of communications in WUSNs. Therefore, accurate measurement of soil moisture is fundamental for agriculture irrigation management [4] and other WUSN applications. Intuitively, to make a high accurate WUSN for soil moisture measurement, the appropriate sampling rate of the system and the deployment of network, in terms of topology and sensor density, are critical. On one hand, the applications need sufficient samples over time to capture the temporal dynamics of the soil moisture. On the other hand, the sampling rate must be restricted in order to reduce energy consumption. Furthermore, since deployment cost of WUSNs is orders of magnitude higher than that of their terrestrial counterparts, determining the appropriate topology and spatial density of WUSNs to minimize the cost of the system is of paramount importance.

Fortunately, the soil moisture over an area is highly spatio-temporally correlated. This correlation can be utilized to reduce the sampling and deployment density of the system while in the same time maintain low distortion, which is defined as the error of the system estimation. It has been shown that soil moisture dynamics depend on both rainfall and the vegetation in a particular area [5]. Since both rainfall input and the vegetation are spatially correlated, spatial dependence in soil moisture process can be modeled through these phenomena. Moreover, rainfall models allow a quantitative description of the temporal variability of soil moisture [8]. As a result, these models can be exploited to develop the estimation distortion expression for the soil moisture measurement. Therefore, in this paper, we investigate the impact of the environment, namely rainfall, soil porosity, and vegetation root zone, on the distortion of measuring the soil moisture. Moreover, these findings are utilized to point out the principles and guidelines on designing WUSNs. These guidelines include the optimal choice of sensor sampling rate, the topology of the network and the density of the sensors.

The remaining of the paper is organized as follows: related work is discussed in Section II. In Section III, the models for analyzing estimation distortion of soil moisture sensors are

established. Simulation results for investigating the network and environment factors are provided in Section IV. Finally, conclusions are provided in Section V.

II. RELATED WORK

One of our contributions is the utilization of the spatio-temporal correlation of the soil moisture to design WUSNs. Several methods have been developed to analyze the spatial correlation and the temporal correlation in WSNs. In [6], variograms have been used to analyze spatial correlation in wireless sensor networks. In [17], data correlation in WSNs has been utilized to form clusters to reduce data transmission in data query. Moreover, spherical and linear models of variogram have been established to analyze real environment data and synthetic data. These studies only consider the spatial correlation in the WSNs. However, the sensor readings in a field are also correlated in the temporal domain, which can be utilized to reduce data distortion. In [11], the spatial correlation of the WSNs is exploited to allocate sampling rates to sensor nodes in order to minimize data transmission and at the same time achieve high estimation accuracy. The similar method is also used in [16]. However, the correlation model in [16] follows an exponential function, which does not capture the spatial correlation characteristics of soil moisture. Moreover, the temporal correlation is not considered.

The correlation of the noise in each sensor is considered in [10] and linear minimum-variance-unbiased-estimator (m.v.u.e) is used to estimate the distorted and noise added physical phenomenon data. However, in this work, only the correlation of the noise is considered, the spatio-temporal correlation underlying the physical phenomenon is not exploited. In [9], the correlation in terms of entropy is measured. The entropy of a sensor is represented with respect to the entropy of its neighbors and the distances between the sensor to its neighbors. Accordingly, the total uncorrelated data of a set of nodes is calculated. However, the distortion and errors of the sensor readings are not considered in [9].

In our work, we mainly focus on the accuracy of estimating the environment phenomenon instead of maximizing the total information. In our previous work [13], [14], [15], a model that considers both the temporal and the spatial correlation is developed. By investigating correlation functions, guidelines for designing wireless sensor networks to reduce estimation distortion and enhance sensor lifetime are provided. In [15], only the effects of WSN parameters are analyzed, but the impacts of the environmental factors are not examined. In this paper, the impact of the environment on the design of WUSNs for soil moisture is the main consideration.

In [5], [7], [8], the spatio-temporal correlation in soil moisture has been investigated by using a rain fall model. The spatial and temporal sampling of soil moisture fields has also been studied in [7]. However, their work only considers the spatio-temporal correlation of the soil moisture, but the characteristics of the WUSNs are not captured. In WUSNs, the deployment of the sensors and the sampling rate are important for data accuracy. Meanwhile, the measurement signal noise ratio of the sensors cannot be ignored. To analyze

the WUSNs for soil moisture measurement, both effects of the sensor network factors and the environment factors should be investigated for the accuracy of system estimation.

III. ESTIMATION DISTORTION IN SOIL MOISTURE MEASUREMENT

In this section, we develop our estimation distortion model in soil moisture measurement, which captures both WUSN characteristics and the environment effects. The rainfall model for the soil moisture is analyzed in Section III-A, which leads to the spatio-temporal correlation model of soil moisture of two points in a field. Then the estimation distortion expression for the wireless underground sensors is developed to utilize the correlation in the spatial and temporal domain.

A. Spatio-temporal Soil Moisture Model

Soil moisture dynamics can be described by the following water balance equation [5]:

$$nZ_r \frac{dS(x, y, t)}{dt} = (1 - \phi)Y(x, y, t) - V \cdot S(x, y, t), \quad (1)$$

where $S(x, y, t)$ is the soil moisture level at time t and location (x, y) , n is the soil porosity and Z_r is the depth of the root zone; the product nZ_r represents the capacity of the soil to store water. $Y(x, y, t)$ is the intensity of the rainfall process and $(1 - \phi)$ is the net rainfall coefficient, which is determined by the plant species and the condition of the vegetation. The sum of evapotranspiration, leakage, and runoff losses is represented by $V \cdot S(x, y, t)$, in which V is the soil water loss coefficient. More specifically, the soil moisture dynamics depend on the precipitation rate and the vegetation characteristics. Assuming a space-time Poisson process of rate λ_R in space and time for rainfall process, the spatio-temporal soil moisture level process $S(x, y, t)$ can be characterized through mean, variance, and correlation function, as follows [5], [8]:

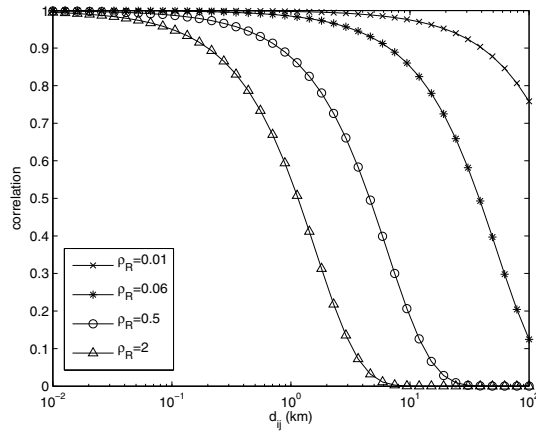
$$\mu_S = \frac{b2\pi\lambda_R}{a\rho_R^2\eta\beta}, \quad (2)$$

$$\sigma_S^2 = \frac{4\pi\lambda_R}{\eta\beta^2\rho_R^2} \frac{b^2}{a(\eta + a)}, \quad (3)$$

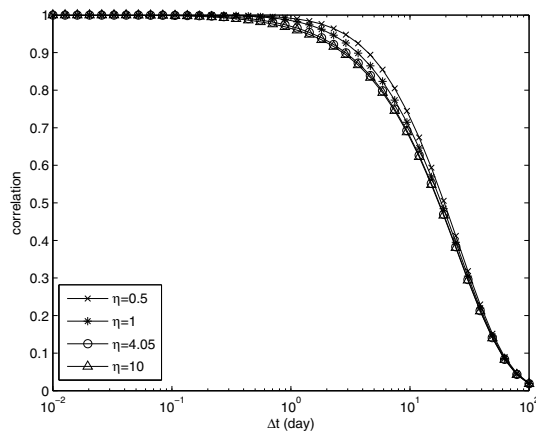
$$\begin{aligned} \rho_m(i, j, k, l) &= \rho_s(i, j) \rho_t(k, l) \\ &= \left[\left(1 + \frac{\rho_R d_{ij}}{4} \right) e^{-\rho_R \frac{d_{ij}}{2}} \right] \left[\frac{\eta e^{-a\Delta t} - a e^{-\eta\Delta t}}{\eta - a} \right] \end{aligned} \quad (4)$$

respectively, where $a = V/(nZ_r)$ is the normalized soil water loss, $b = (1 - \phi)/(nZ_r)$ is the normalized rainfall coefficient, $1/\rho_R$, $1/\eta$, and $1/\beta$ are the mean rain cell radius, storm duration, and rainfall intensity of a cell, respectively. The correlation function in (4) models the correlation between two locations (x_i, y_i) and (x_j, y_j) at distance d_{ij} at times t_k and t_l with $\Delta t = |t_k - t_l|$. Note that the correlation function in (4) can be separated into spatial and temporal components.

In Fig. 1, the spatial correlation and the temporal correlation of the soil moisture are shown, with the typical values listed in Table I. In Fig. 1(a), the rainfall cell radius of 100 km, 16.7 km, 2 km and 0.5 km are considered, respectively. It is shown



(a)



(b)

Fig. 1. Soil moisture model for (a) spatial correlation and (b) temporal correlation.

TABLE I
DEFAULT PARAMETER VALUES FOR THE SOIL MOISTURE MODEL.

Parameter	Value
V	4mm/day
Z_r	100mm
n	1

that the spatial correlation decreases as the distance between two locations increases. In addition, the correlation decreases as the rainfall cell radius decreases. Thus, in soil moisture measurement applications, the size of the typical rainfall cell radius should be considered for the deployment of the sensors at the deployment site. As long as the deployment ensures high spatial correlation of the sensor readings, this correlation can be exploited to estimate the soil moisture around a certain area.

The temporal correlation of the soil moisture is shown in Fig. 1(b). As expected, it decreases with an increase in the time difference, Δt . The temporal correlation is related to the storm duration, $1/\eta$, such that it is slightly higher when the storm duration is long (η is small).

The spatio-temporal correlation model of soil moisture

in (4) is used in our analysis of data distortion in WUSNs.

B. Data Distortion in WUSNs

The spatio-temporal model for the soil moisture governs the relationship between the effects of the environment and the soil moisture. Next, this model is utilized to analyze the data distortion of the soil moisture measurement applications. These applications can be categorized as field source applications [15], where the measurements from the sensors, $f(x, y, t)$, are independently identically distributed (i.i.d.) Gaussian random variables. In these applications, we are interested in estimating the soil moisture value at a given location (x_0, y_0) in the field over a decision interval τ . The measurement samples collected by a number of underground sensor nodes around the interested location are utilized by the estimator. These selected nodes are referred to as *representative nodes*.

The data distortion model exploiting spatio-temporal correlations is developed in [13], [14], and [15]. Next, the basic framework is briefly introduced for the completeness of this paper. Assuming the observed signal $f(x, y, t)$ is wide-sense stationary (WSS), the expectation of the signal over the decision interval τ , i.e., $S(\tau)$ can be calculated by the time average of the observed signal:

$$S(\tau) = \frac{1}{\tau} \int_0^\tau f(x_0, y_0, t) dt, \quad (5)$$

where (x_0, y_0) is the location of the signal in which we are interested. A sensor node at location (x_i, y_i) receives signal $S_i[k]$ at time t_k . Because of the spatio-temporal correlation, $S_i[k]$ is correlated to $S(\tau)$, and it is defined as

$$S_i[k] = f(x_i, y_i, t_k). \quad (6)$$

We assume $S_i[k]$'s are JGRV with $\mathcal{N}(0, \sigma_S^2)$. Two samples, $S_i[k]$ and $S_j[l]$, taken at different locations and different times, have the covariance given by (4).

Due to the noise and the associated measurement errors, we model the measurements of a sensor as a signal with additive Gaussian noise, thus

$$X_i[k] = S_i[k] + N_k[k], \quad (7)$$

where $N_i[k] \sim \mathcal{N}(0, \sigma_N^2)$. The measurement SNR φ is defined as

$$\varphi = \frac{\sigma_S^2}{\sigma_N^2}. \quad (8)$$

This measurement is transmitted to the sink through the WSN. At the sink, using minimum mean square error (MMSE) estimation technique, the estimation, $Z_i[k]$, of the event information $S_i[k]$ can be found as:

$$Z_i[k] = \frac{\varphi}{\varphi + 1} (S_i[k] + N_i[k]). \quad (9)$$

After collecting the samples of the signal in the decision interval τ from M nodes, the sink reconstructs the sensing field by interpolating the value at (x_0, y_0) . This interpolation can be expressed as:

$$\hat{S}_f(\tau, f, M) = \sum_{i=1}^M w_f(i) \left(\frac{1}{\tau f} \sum_{k=1}^{\tau f} Z_i[k] \right), \quad (10)$$

$$D_{norm}(\tau, f, M) = 1 + \frac{\kappa^2 \tau f}{\varphi M} + \frac{\kappa^2}{M^2} \sum_{i=1}^M \sum_{j=1}^M \rho_s(i, j) \sum_{k=1}^{\tau f} \sum_{l=1}^{\tau f} \rho_t(|k-l|/f) - \frac{2\kappa}{\tau M(\eta - a)} \sum_{i=1}^M \rho_s(i, s) \sum_{k=1}^{\tau f} \left[\frac{\eta}{a} \left(2 - e^{-ak/f} - e^{a(k/f-\tau)} \right) + \frac{a}{\eta} \left(2 + e^{-\eta k/f} + e^{\eta(k/f-\tau)} \right) \right] \quad (14)$$

where $\{w_f\}_{i=1\dots M}$ are the weight functions.

Intuitively, since the correlation is higher when two locations are closer, the sensors that are closer to the interested location should have higher weight. However, because of the high correlation of the soil moisture over space, through our simulations, we find out that this type of weight function has almost indistinguishable performance improvement over the simple average strategy. Therefore, in this paper, the weight function is defined as

$$w_f = \frac{1}{M}. \quad (11)$$

Accordingly, the distortion of the estimation is

$$D = E \left[\left(S(\tau) - \hat{S}(\tau, f, M) \right)^2 \right], \quad (12)$$

which is normalized by the variance of the signal, σ_S^2 , i.e.,

$$D_{norm} = \frac{D}{\sigma_S^2}. \quad (13)$$

The expression for the normalized distortion of the estimation at the sink is shown in (14), where

$$\kappa = \frac{1}{\tau f \left(1 + \frac{1}{\varphi} \right)}. \quad (15)$$

In (14), $\rho_s(i, j)$ is the spatial correlation between two sensors i and j , and $\rho_t(|k-l|/f)$ is the temporal correlation between two samples taken at sample indexes k and l , as shown in (4). Replacing κ with (15), the second term can be expressed as $\varphi/[\tau f M(1 + \varphi)^2]$. Accordingly, using more representative nodes (increasing M) or gathering more samples over a decision interval (increasing τf) can compensate for the distortion caused by the noise. Similarly, higher quality sensors might improve distortion for increasing φ . The third term in (14) is related to the spatio-temporal correlation between each pair of representative sensor nodes and the fourth term is related to the spatio-temporal correlation between each representative node and the target location. Note that when the measurements between two representative nodes are highly correlated, the overall distortion increases. This is because in this situation, the additional nodes only provide redundant information in terms of estimating the value at the interested location, but the errors add up due to independent measurements. Therefore, to decrease the distortion, the representative nodes should be selected such that they are highly correlated to the interested location, but not correlated to each other [14]. In this way, the sink can get more information to estimate the values at the interested location.

TABLE II
WUSN PARAMETER VALUES FOR THE EVALUATION.

Parameter	Value
M	5
K	10
f	1 sample per day
φ	80 dB
grid size	0.1 km

IV. NUMERICAL EVALUATION

In this section, numerical evaluations for exploiting spatio-temporal correlation characteristics of soil moisture is provided. A sensor network of a grid topology with 121 nodes and a grid size of 100 meters is employed for the evaluations. The network aims to estimate the soil moisture value at the center of the grid using the samples of the sensors located on the grid. Other default parameter values used in the evaluations are listed in Table II.

The impacts of the WUSN parameters on the data distortion are evaluated in Section IV-A while in Section IV-B, the impacts of the environment factors are investigated. Finally, in Section IV-C, effects of different deployment strategies are discussed.

A. Impact of Network Parameters

In Fig. 2, the normalized distortion is shown as a function of the sampling rate, f , and the number of representative nodes, M , for different cases of rainfall cell radius, $1/\rho_R$, and measurement SNR, φ . It can be observed that in all the cases, the distortion decreases with increasing sampling rate, f , as expected. On the other hand, there is a limit of improving distortion by increasing the sampling rate. This observation reveals that there is an optimal value, f_{opt} , for temporal resolution such that further increase in sampling rate, f , does not decrease the distortion. It is shown in Fig. 2 that 20 to 50 samples per day can effectively minimize the distortion. However, since sensing and transmitting samples consume energy of the sensors, to improve the lifetime of the unchargeable sensors, the sampling rate should be as low as possible. In our evaluations, the sampling rate is set at 1 sample per day, which gives a relatively low distortion. When the rainfall cell radius is 16.6 km (Fig. 2(a)), the normalized distortion at that sampling rate is 0.092.

The number of representative nodes M does not affect the distortion when rainfall cell radius is large (Fig. 2(a)). The reason is that the rainfall cell radius ($1/\rho_R = 16.6$ km) is much larger than the size of the field, all the sensors in the field are highly correlated. Therefore, the sensors observe

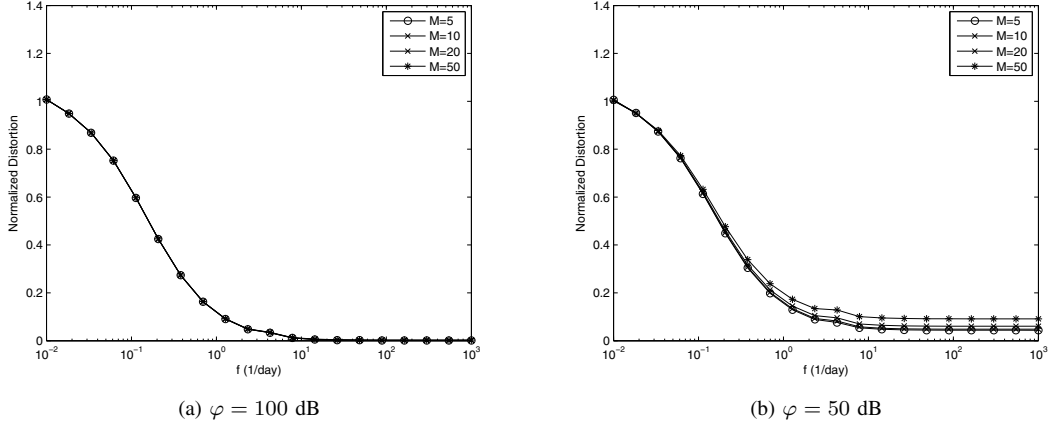


Fig. 2. Normalized distortion vs. sampling frequency, f , for different number of nodes, M when (a) rainfall cell radius is 16.6 km and (b) rainfall cell radius is 0.5 km.

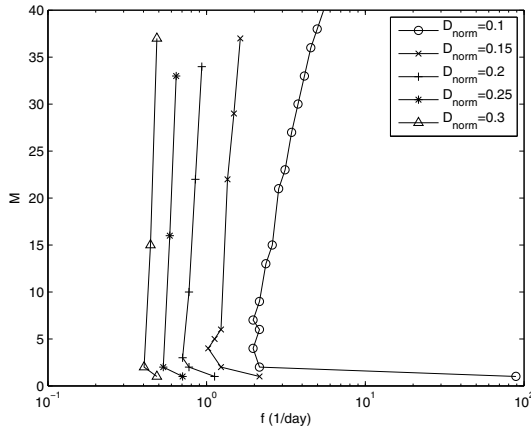


Fig. 3. The trade off of number of representative nodes, M , and the sampling rate, f .

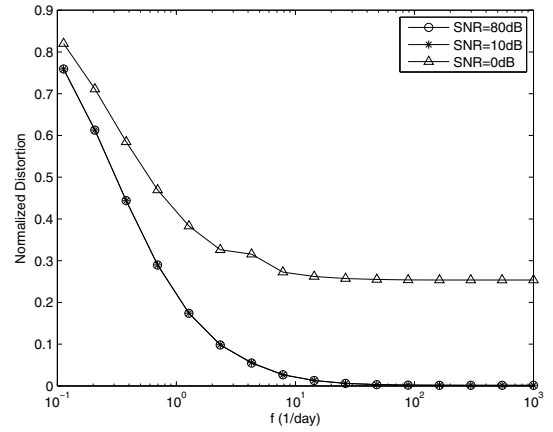


Fig. 4. Normalized distortion vs. sampling frequency for different values of φ .

very similar readings. In Fig. 2(b), the distortion is shown when the rainfall cell radius is small ($1/\rho_R = 0.5$ km). Note, in this case, the rainfall area is just enough to cover the test field ($1 \text{ km} \times 1 \text{ km}$). Smaller rainfall cell radius is not considered because it is uncommon in the nature. In this case, using fewer representative nodes to estimate the value has better performance. When measurement SNR is 80 dB, at the sampling rate of 1 sample per day, compared with using 50 sensors, using the closest 5 sensors decreases the normalized distortion by 22.6%. This is because when the rainfall cell radius is small, sensors far away are not spatially correlated to the interested location. The impact of the rain fall cell radius is further investigated in Section IV-B.

When the rainfall cell radius is small ($1/\rho_R = 0.5$ km), the relation of the number of the representative nodes, M , and the sampling rate, f , is shown in Fig. 3 for different distortion requirements. Given a distortion requirement (the maximum normalized distortion), the number of representative nodes with the minimal sampling rate that can achieve this re-

quirement is plotted in Fig. 3. In applications, the goal is to use fewer representative nodes and lower sampling rate to estimate readings. The optimal combination of M and f for each given distortion requirement is shown in Fig. 3. Notice that smaller number of representative nodes has better performance, which confirms our finding in Fig. 2. However, it is also inadvisable to use only one representative node to estimate the value at a given interested point, since the measurement error from this node cannot be reduced by the knowledge of other nodes. The optimal number of representative nodes is about 4–5 in this figure.

In addition, it is also noticed that at different measurement SNR, φ , the distortion has similar behavior against sampling rate. In Fig. 4, the impact of measurement SNR is studied with rainfall cell radius of 16.6 km and $M = 5$. In this figure, the cases where $\varphi = 10$ dB and $\varphi = 80$ dB have similar performance. In these cases, if the sampling rate is high enough (more than 30 samples per day), the system can achieve very low distortion. However, if the SNR is

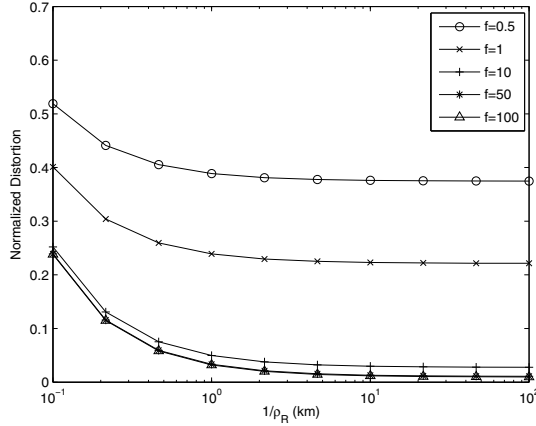


Fig. 5. Normalized distortion vs. rainfall cell radius for different values of sampling frequency.

deteriorated, the system cannot achieve low distortion even with high sampling rate. For instance, in the case where $\varphi = 0$ dB, the distortion decreases to 0.25 after f is increased to 60 samples per day. Nevertheless, the distortion does not decrease further even f is further increased.

The measurement SNR is mainly determined by the quality of the sensors. This finding reveals that for soil moisture applications, to accurately estimate a moisture value at a location, the quality of the sensors needs to achieve a certain requirement. In our evaluations, the default SNR is 80 dB. This value is chosen because if we consider the maximum allowed error as $3\sigma_N$, the maximum allowed error is 3%, which is achievable by current soil moisture sensors. For instance, the soil moisture sensor 5TE from *Decagon Devices* [19] has accuracy of $\pm 3\%$.

B. Impact of Environment

The environment factors, such as rainfall cell radius, rainfall duration, soil porosity and vegetation, have different effects on the spatio-temporal correlation of soil moisture, which in turn impacts the data distortion of the estimation. Thus, in this section, we investigate these factors and provide some guidelines for designing the WUSNs according to the environment factors.

The effects of the environment in terms of rainfall cell radius are shown in Fig. 5, where the normalized distortion is shown as a function of the rainfall cell radius, $1/\rho_R$. As shown in the figure, the distortion decreases by only 0.75% when rainfall cell radius is increased from 5km to 100km. In these situations, rainfall area is 82 times to 3.3×10^4 times larger than the field. Thus, soil moisture is highly correlated over a large area and all the sensors in the field have highly spatially correlated readings. Further decreasing the rain cell radius ($0.1 \text{ km} \leq 1/\rho_R \leq 5 \text{ km}$) leads to higher distortion. Compared to $1/\rho_R = 10 \text{ km}$, when $1/\rho_R = 0.2 \text{ km}$, the distortion increases 17.3% with $f = 0.5$ samples per day and 872% with $f = 100$ samples per day. This is because of the decrease in spatial correlation as the rainfall cell radius decreases. On the other

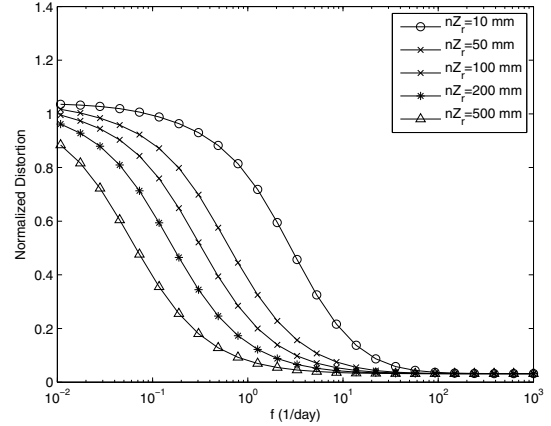


Fig. 6. Normalized distortion vs. sampling frequency for different values of soil water storage capacity, nZ_r .

hand, increasing the sampling rate can improve the distortion, but only to a limit. In Fig. 5, the distortion behaves almost the same when $f = 50$ samples per day as well as $f = 100$ samplers per day.

According to the correlation function in (4), where $a = V/(nZ_r)$ and $b = (1 - \phi)/(nZ_r)$, the soil porosity, n , and the depth of the root zone, Z_r , have the same impact on the estimation distortion. The product of the two is considered as the capacity of the soil to store water [8]. Their impact on the distortion is shown in Fig. 6. It can be observed that when the product is small, i.e., $nZ_r = 10 \text{ mm}$, high sampling rate is needed to decrease the normalized distortion. For example, to achieve normalized distortion of 0.05, when $Z_r = 500 \text{ mm}$, the sampling rate of the system can be set at 1 sample per day; but when $Z_r = 10 \text{ mm}$, the required sampling rate is 23 samples per day.

The behavior shown in Fig. 6 can be explained as follows. When the roots are shallow and the soil porosity is high, i.e., $nZ_r = 10 \text{ mm}$, the water can infiltrate into the field and also evaporate from the field very quickly. As a result, the soil moisture fluctuates fast with rainfall and water losses. Thus, in this situation, the sensors need higher sampling rate to capture the variation of the soil moisture. On the other hand, when the roots are deep and the soil density is high, i.e., $nZ_r = 500 \text{ mm}$, the process of water infiltration is slow, lower sampling rate is enough to estimate the soil moisture values.

In Fig. 7, the relationship between the distortion and storm duration, η , is depicted. It is shown that in long storms, i.e., $\eta = 0.1 \text{ day}^{-1}$, the estimation distortion is lower, especially for systems using longer decision interval. As an example, for a decision interval of 10 days, the normalized distortion when storm duration is 10 days is 0.025, on the contrary, the distortion when the storm duration is 0.1 days is 0.118. This is because during long storms, the water infiltrates to the field effectively, and the temporal correlation among the samples becomes high, which leads to more accurate estimation. However, in short storms, i.e., $\eta = 10 \text{ day}^{-1}$, the rain causes quick fluctuation of the soil moisture over time and readings from

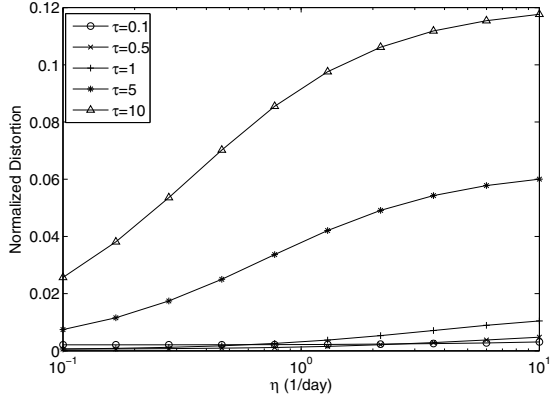


Fig. 7. Normalized distortion vs. storm duration for different values of decision interval ($M = 5$).

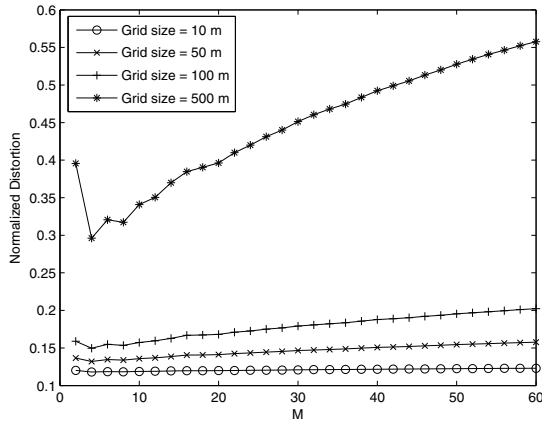


Fig. 8. Normalized distortion vs. the number of representative nodes for different grid size ($f = 1$ sample per day).

sensor nodes become less temporally correlated. Accordingly, the overall distortion increases. It can also be observed that if the decision interval is short, i.e., $\tau = 0.1$ day, the overall distortion can be kept low even in short storms. This is because short decision interval captures the fluctuation of soil moisture over time.

C. Impact of the Topology

In the previous evaluations, only grid topology is considered. However, in real applications, because of the limitation caused by topography, sometimes grid topology deployment is not available. In this section, the effects of the network topology on the overall distortion is analyzed. The grid size in the grid topology, which determines the density of the sensor nodes in WUSNs, is investigated first. Moreover, the effects of random deployment are also studied.

The grid size in the grid topology is defined as the shortest distance between two neighbor nodes. When the rainfall cell radius is large ($1/\rho_R = 16.6$ km), the grid size in the ranges of 10 m to 500 m does not have an impact on the data distortion. However, if the rainfall cell radius is small, the impact of the grid size becomes important. In Fig. 8, the

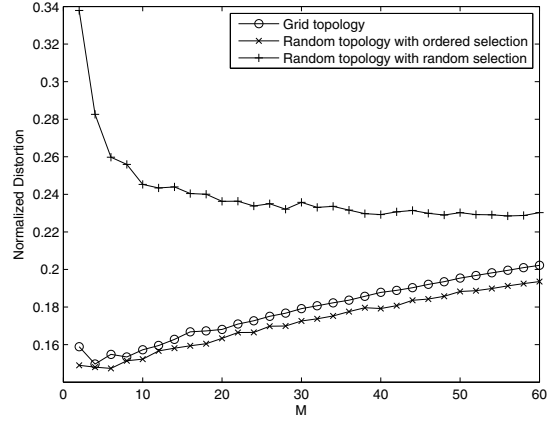


Fig. 9. Normalized distortion vs. the number of representative nodes for different network topology and node selection procedures ($f = 50$ samples per day).

normalized distortion over the number of representative nodes for different grid sizes is shown for the rainfall cell radius of 0.5 km. As expected, the network with smaller grid size (< 0.1 km) has better performance, since when the grid size becomes larger, the readings from the nodes become spatially uncorrelated. However, there is an optimal value for the grid size, beyond which the performance has limited improvement. For our cases, the optimal grid size for the grid topology is about 100 m. Compared to the minimum distortion when grid size is 10 m, the minimum distortion when the grid size is 100 m increases from 0.12 to 0.15.

The ideal number of the representative nodes depends on the grid size. On one hand, since the closer nodes have higher correlation and similar readings, it is desired to depend on these closer nodes to make the estimation. On the other hand, because of the measurement noise, the system needs sufficient readings to reduce the measurement distortion. Therefore, when the grid size is small (10 m), since all the sensors are highly correlated, more representative nodes can further decrease the distortion as the measurement noise can be reduced. In Fig. 8, for the case when grid size is 10 m, the ideal number is 8. However, for larger grid size, the spatial correlation factor dominates the distortion, thus using smaller number of representative nodes, which are closer to the interested location, can achieve lower distortion. In Fig. 8, the ideal number of representative nodes for the grid size of 50 m and 100 m is about 4.

In practical soil moisture measurement applications, because of the limitation of the topography of the deployment site, it may be difficult to deploy a network with grid topology. In some cases, a random deployment of sensors is required. Hence, the evaluation is extended to also consider the random deployment. For random topology, there are two strategies to select the representative nodes. When the sink knows the location of each sensor node, it can choose the nodes which are the closest M nodes to the interested location as the representative nodes. This strategy is called *random topology*

with ordered selection. However, if the sink does not have the knowledge of the locations of sensor nodes, it cannot find the closest nodes to the interested location. In this situation, the sink has to randomly select M nodes as the representative nodes. We call this strategy *random topology with random selection*.

The performance of these two strategies, as well as the grid topology, is shown in Fig. 9. In each evaluation, the rainfall cell radius is 0.5 km, and 121 nodes are deployed in a 1 km \times 1 km area. For the random topology, 100 instances are generated and the results are averaged. The random topology with ordered selection results in a slightly better performance than the grid topology. When $M = 8$, the grid topology and the random topology with ordered selection both have the minimal distortion, however, the random topology with ordered selection decreases distortion by 1.5%. This is because in random topology, the representative nodes may have closer distance to the interested location. Thus, they can better exploit the spatial correlation. As expected, the random topology with random selection has the worst performance due to the fact that the representative nodes may include less correlated nodes.

Another interesting observation is that for grid topology and random topology with ordered selection, there is an optimal number of representative nodes which is small. However, in the random topology with random selection case, the normalized distortion decreases as the number of representative nodes increases. This reveals that for soil moisture measurement applications, irrespective of the topology, sensors closer to the interested location should be exploited for estimation. Therefore, it is desired to know the locations of each sensor nodes, which can achieve lower distortion and save energy since only small number of nodes are needed to achieve the optimal performance. If the sink is not aware of the location of the event, a larger number of representative nodes are required to reduce the distortion, which consumes more energy and also could cause congestion in the network.

V. CONCLUSIONS

In this paper, the spatial-temporal correlation is utilized to analyze the estimation distortion of soil moisture estimation. The impacts of environment factors, such as rainfall, soil porosity and vegetation root zone, as well as the impacts of the network parameters, on estimation distortion are analyzed. The simulations reveal that in large rainfall cell areas, estimation distortion can be kept low because of the high spatial correlation. Also, for short storms, estimation interval must also be short to capture the fluctuation of the soil moisture. Moreover, in situations where soil is porous and the roots are shallow, high sampling rate is needed to achieve low distortion. For the network parameters, the evaluations show that there is an optimal sampling rate for the application and only a few closest sensors are needed to estimate the values at a given location.

We have recently shown that soil moisture also influences the wireless underground communication success [3], [12]. Therefore, our future work will integrate the spatio-temporal characteristics of the soil moisture into a communication

protocol for WUSNs to exploit the correlation between communication media and soil moisture properties.

VI. ACKNOWLEDGEMENTS

This work is supported by U.S. Geological Survey Award 2010NE209B, UNL Water Center, and the UNL Research Council Maude Hammond Fling Faculty Research Fellowship.

REFERENCES

- [1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam and E. Cayirci, "Wireless Sensor Networks: A Survey," *Computer Networks Journal (Elsevier)*, vol. 38, no. 4, pp. 393-422, March 2002.
- [2] I. F. Akyildiz and E. P. Stuntebeck, "Wireless underground sensor networks: Research challenges," *Ad Hoc Networks Journal (Elsevier)*, vol. 4, pp. 669-686, July 2006.
- [3] I. F. Akyildiz, M. C. Vuran, and Z. Sun, "Channel modeling for Wireless Underground Communication in Soil, Mines and Tunnels," to appear in *Physical Communication Journal (Elsevier)*, 2009.
- [4] S. Irmak et al., "Watermark Granular Matrix Sensor to Measure Soil Matric Potential for Irrigation Management," in http://lancaster.unl.edu/ag/Crops/Watermark_Sensor.pdf.
- [5] V. Isham, et.al., "Representation of Space-time Variability of Soil Moisture," in *Proc. Royal Society*, vol. 461, no. 2064, pp. 4035-4055, 2005.
- [6] A. Jindal and K. Psounis, "Modeling spatially correlated data in sensor networks," *ACM Trans. Sen. Netw.*, vol. 2, no. 4, pp. 466-499, 2006.
- [7] S. Manfreda and I. Rodriguez-Iturbe, "On the spatial and temporal sampling of soil moisture fields" *Water Resources Research*, vol. 42, 2006.
- [8] I. Rodriguez-Iturbe, V. Isham, D. R. Cox, S. Manfreda, and A. Porporato, "Space-time Modeling of Soil Moisture: Stochastic Rainfall Forcing with Heterogeneous Vegetation," *Water Resources Research*, vol. 42, 2006.
- [9] S. Patten, B. Krishnamachari, and R. Govindan, "The impact of spatial correlation on routing with compression in wireless sensor networks," *ACM Trans. Sen. Netw.*, vol. 4, no. 4, pp. 1-33, 2008.
- [10] Z. Quan, W. J. Kaiser, and A. H. Sayed, "A spatial sampling scheme based on innovations diffusion in sensor networks," in *Proc. ACM IPSN '07*, Cambridge, Massachusetts, USA, April 2007.
- [11] Y. Sankarasubramaniam, O.B. Akan and I.F. Akyildiz, "ESRT: Event-to-Sink Reliable Transport in Wireless Sensor Networks," in *Proc. of the ACM MobiHoc '03*, Annapolis, Maryland, June 2003.
- [12] A. R. Silva and M. C. Vuran, "Empirical Evaluation of Wireless Underground-to-Underground Communication in wireless underground sensor networks," in *Proc. DCOSS '09*, 2009.
- [13] M. C. Vuran, O. B. Akan, and I. F. Akyildiz, "Spatio-Temporal Correlation: Theory and Applications for Wireless Sensor Networks," *Computer Networks Journal (Elsevier)*, Vol. 45, No. 3, pp. 245-261, June 2004.
- [14] M. C. Vuran and I. F. Akyildiz, "Spatial Correlation-based Collaborative Medium Access Control in Wireless Sensor Networks," *IEEE/ACM Transactions on Networking*, vol. 14, no. 2, pp. 316-329, April 2006.
- [15] M. C. Vuran and O. B. Akan, "Spatio-temporal Characteristics of Point and Field Sources in Wireless Sensor Networks," in *Proc. IEEE ICC '06*, Istanbul, Turkey, June 2006.
- [16] K. Yuen, B. Liang, and B. Li, "A distributed framework for correlated data gathering in sensor networks," *IEEE Trans. Vehic. Tech.*, vol. 57, no. 1, pp. 578-593, 2008.
- [17] S. Yoon and C. Shahabi, "The clustered aggregation (cag) technique leveraging spatial and temporal correlations in wireless sensor networks," *ACM Trans. Sen. Netw.*, vol. 3, no. 1, p. 3, 2007.
- [18] United Nations - World Water Day 2007 Report, "Coping with Water Scarcity - Challenge of the Twenty-First Century," in <http://www.worldwaterday07.org>.
- [19] http://www.decagon.com/soil_moisture/