On the spatial and temporal sampling of soil moisture fields

Salvatore Manfreda[1,2] and Ignacio Rodríguez-Iturbe[1]

Received 1 September 2005; revised 10 January 2006; accepted 24 January 2006; published 3 May 2006.

[1] Recent work by Isham et al. and Rodrìguez-Iturbe et al. has characterized the space-time variability of soil moisture through its analytically derived covariance function which depends on soil properties, vegetation structure, and rainfall patterns typical of a region. This paper uses such characterization to address the strategies and methodologies for the sampling of soil moisture fields. The focus is on the estimation of the long-term mean soil moisture and the daily soil moisture averaged over a given area as a function of the network geometry, number of stations, number of sampling days and landscape heterogeneity. It is found that the spatial geometry of the network has a significant impact on the sampling of the average soil moisture over an area in any particular day, while it is much less relevant for the sampling of the long-term mean daily soil moisture over the region. In the latter case, the length of the record is a commanding factor in what concerns the variance of estimation, specially for soils with shallow rooted vegetation. Spatial vegetation heterogeneity plays an important role on the variance of estimation of the soil moisture, being particularly critical for the sampling of the average soil moisture over an area for a given day.

Citation: Manfreda, S., and I. Rodríguez-Iturbe (2006), On the spatial and temporal sampling of soil moisture fields, *Water Resour. Res.*, *42*, W05409, doi:10.1029/2005WR004548.

1. Introduction

[2] Soil moisture represents a critical component of the hydrological cycle and is a central factor in climate, soil and vegetation interactions [e.g., *Albertson and Montaldo*, 2003; *Porporato et al.*, 2004; *Rodríguez-Iturbe and Porporato*, 2004]. It affects catchments response [e.g., *Eagleson*, 1978; *Manfreda et al.*, 2005] and modulates interactions between land surface and atmosphere influencing climate and weather [e.g., *Entekhabi et al.*, 1996; *Porporato et al.*, 2000]. Soil moisture availability represents a critical control on plant growth dynamics and ecological patterns in semiarid ecosystems [*Rodríguez-Iturbe*, 2000; *Meron et al.*, 2004; *Caylor et al.*, 2005; *Scanlon et al.*, 2005] and its space-time variability is crucial to formulate accurate predictions of the behavior of hydrologic systems [*Western et al.*, 2002].

[3] One of the most promising strategies for soil moisture monitoring is through remote sensing techniques [e.g., *Walker et al.*, 2001]. These provide estimates of soil moisture over large areas that need to be calibrated with ground-based measurements. Land surface models also require quantitative comparison between simulated soil moisture maps and field measurements. In this context, *Robock et al.* [2003] evaluated land surface model results from the North American Land Data Assimilation System (NLDAS) [see *Mitchell et al.*, 2004] using in situ observations over the southern Great Plains, but without considering the effects of sampling errors.

[4] An attempt to estimate sampling errors by using empirically estimated correlation functions is presented by *Vinnikov et al.* [1999] evaluating the root-mean-square error of the soil moisture networks of Illinois and Oklahoma Mesonet. Using the optimal averaging technique introduced by *Kagan* [1979], they show that the Mesonet network does not provide a significant improvement in the accuracy respect to the less dense Illinois network. *Yoo* [2001] adopted the formalism proposed by *North and Nakamoto* [1989] for the estimation of sampling errors in soil moisture fields using the soil moisture model of *Entekhabi and Rodríguez-Iturbe* [1994]. The procedure has been applied to the Washita 1992 data highlighting the inefficiency of ground-based networks for large-scale observations. Within the Soil Moisture Experiment 2002 (SMEX02), *Jacobs et al.* [2004] found that number of samples to measure the soil moisture content with a given confidence interval may change according to the field characteristics and to the soil moisture content itself. This suggests that further investigations are necessary to understand the role of spatial heterogeneity of the land surface on the sampling of soil moisture fields.

[5] In the present paper, the sampling errors are estimated using the space-time soil moisture correlation structures analytically derived by *Isham et al.* [2005] and *Rodríguez-Iturbe et al.* [2006] for a model considered appropriated for a water-limited ecosystem. The methodology explicitly accounts for soil characteristics, vegetation patterns, and rainfall dynamics neglecting topographical effects and the upper bound due to saturation. The soil moisture model can be considered representative of a relatively flat landscape under semiarid climate conditions. Both random sampling and stratified random sampling, which are the commonly used schemes over extended areas, are investigated for the case of homogeneous and heterogeneous vegetation. The study adopts the analytical framework introduced by

[1]Department of Civil and Environmental Engineering, Princeton University, Princeton, New Jersey, USA.
[2]Dipartimento di Ingegneria e Fisica dell'Ambiente, Università degli Studi della Basilicata, Potenza, Italy.

0043-1397/06/2005WR004548

Rodríguez-Iturbe and Mejía [1974] for rainfall sampling that has been extensively used for the design of rainfall networks [e.g., *Bras and Rodríguez-Iturbe*, 1976].

[6] It is important to point out that this paper will focus on the estimation of soil moisture integrated over the root zone of the vegetation. No attempt will be made to estimate soil moisture values across the depth of the root profile.

2. Space-Time Soil Moisture Dynamics

[7] Soil moisture dynamics is modeled through a water balance equation driven by stochastic space-time rainfall forcing as studied by *Isham et al.* [2005] and *Rodríguez-Iturbe et al.* [2006]. The rainfall model is the same proposed by *Cox and Isham* [1988] where rainfall occurrences are modeled by a sequence of circular rain cells that occur as a Poisson process of rate λ_R in space and time. Each cell is characterized by random radius, intensity and duration that are assumed independent of each other and exponentially distributed in order to keep as a minimum the number of parameters. The total intensity of the rainfall process, $Y(\mathbf{u}, t)$, is given by the superposition of cells that may overlap at any spatial location $\mathbf{u}$ at time t. Although the use of the term cell is customary in this type of model, one should be aware that they are a mathematical construct whose characteristics do not really match those of the usually identified as rain cells in different types of precipitation fields.

[8] The water balance is written in the simplified form

$$nZ_r \frac{dS(\mathbf{u},t)}{dt} = (1-\phi)Y(\mathbf{u},t) - VS(\mathbf{u},t), \tag{1}$$

where the dimensionless random variable $S(\mathbf{u}, t)$ represents the relative soil moisture at site $\mathbf{u}$ and time t, n is the soil porosity, and Z_r is the depth of the root zone. On the right hand side, the term $(1 - \phi)Y(\mathbf{u}, t)$ is the infiltration rate that is determined by the precipitation reduced by $(1 - \phi)$ representing the vegetation interception. The second term, $VS(\mathbf{u}, t)$, is the loss function given by the sum of evapotranspiration and leakage assumed to be a linear function of soil moisture. The water loss coefficient, V, depends on both vegetation and soil properties and the factor ϕ depends on the plant species and condition of vegetation. Typical values of ϕ range between 0.4 and 0.1 according to *Lull* [1964]. Although continuous in time, all results from equation (1) need to be interpreted at the daily timescale since we do not include any dynamics below the daily level neither in the evapotranspiration losses nor in the storm structure [*Rodríguez-Iturbe and Porporato*, 2004].

[9] For the purpose of this paper, the use of a linear approximation for the description of soil water losses is reasonable especially for nonhumid conditions. *Pan et al.* [2003] provide estimates of soil water loss coefficients obtained from the analysis of ground measurements of the Monsoon 1990 and Washita 1992 campaigns relating those values to the soil permeability and the vegetation Leaf Area Index (LAI) [*Pan et al.*, 2003]. They obtain values of V that range between 3 and 16 mm d^{-1}.

[10] The soil moisture scheme adopted above does not account for the presence of an upper bound due to soil saturation. This limitation was found to be necessary to obtain closed analytical results for the covariance structure of soil moisture fields. *Isham et al.* [2005] show that the lack of an upper bound is not a restrictive assumption in terms of its impact on the covariance when dealing with arid and semiarid environments.

[11] With the above modeling scheme, the expected value of the relative soil moisture at a point (e.g., the origin) with given vegetation is

$$E[S(0,t)] = \frac{b_0}{a_0}\frac{2\pi\lambda_R}{\rho_R^2\eta\beta}, \tag{2}$$

where $a_0 = [V/(nZ_r)]_0$ is the normalized soil water loss, $b_0 = [(1 - \phi)/(nZ_r)]_0$ is the normalized net rainfall coefficient, $\mu_R = 1/\rho_R$ is the mean rain cell radius, $\mu_D = 1/\eta$ is the mean storm duration and $\mu_X = 1/\beta$ is the mean rainfall intensity of a cell [*Isham et al.*, 2005].

[12] The expression for the space-time covariance function of the relative soil moisture conditional on the cover of the points A and B, l apart from each other, is [*Rodríguez-Iturbe et al.*, 2006]

$$\begin{aligned} C_{AB} &= \mathrm{Cov}[S(0,t), S(l,t+h)] \\ &= \frac{2\pi\lambda_R}{\eta\beta^2} b_A b_B \left(\frac{2\eta e^{-a_B h}}{(a_A+a_B)(\eta^2-a_B^2)} + \frac{e^{-\eta h}}{(a_B-\eta)(a_A+\eta)}\right) \\ &\quad \cdot \left(\frac{2}{\rho_R^2} + \frac{l}{2\rho_R}\right) e^{-\rho_R \frac{l}{2}}, \end{aligned} \tag{3}$$

where the site (B), characterized by the parameters (a_B, b_B), is "attached" at the later time and the vegetation cover controls the value of the pairs (a_A, b_A) and (a_B, b_B).

[13] In the case of uniform vegetation, the space-time covariance function of relative soil moisture becomes

$$\begin{aligned} &\mathrm{Cov}[S(0,t), S(l,t+h)] \\ &\quad = \frac{2\pi\lambda_R}{\eta\beta^2} \frac{b^2\left(\eta e^{-ah} - a e^{-\eta h}\right)}{a(\eta^2 - a^2)} \left(\frac{2}{\rho_R^2} + \frac{l}{2\rho_R}\right) e^{-\rho_R \frac{l}{2}} \end{aligned} \tag{4}$$

and the variance

$$\sigma_S^2 = \frac{4\pi\lambda_R}{\eta\beta^2\rho_R^2}\frac{b^2}{a(\eta+a)} \tag{5}$$

which have been studied in detail by *Isham et al.* [2005].

[14] The correlation function of the relative soil moisture is a separable function in its space and temporal components. For the case of uniform vegetation, the space correlation is the same of the rainfall process

$$r(l) = \left(1 + \frac{\rho_R l}{4}\right) e^{-\rho_R \frac{l}{2}} \tag{6}$$

and the temporal correlation is

$$r(h) = \frac{\eta e^{-ah} - a e^{-\eta h}}{\eta - a}. \tag{7}$$

[15] The case of heterogeneous vegetation is modeled by *Rodríguez-Iturbe et al.* [2006] by a marked point process that account for the presence of two functionally different vegetation types in the landscape (e.g., grasses and trees).

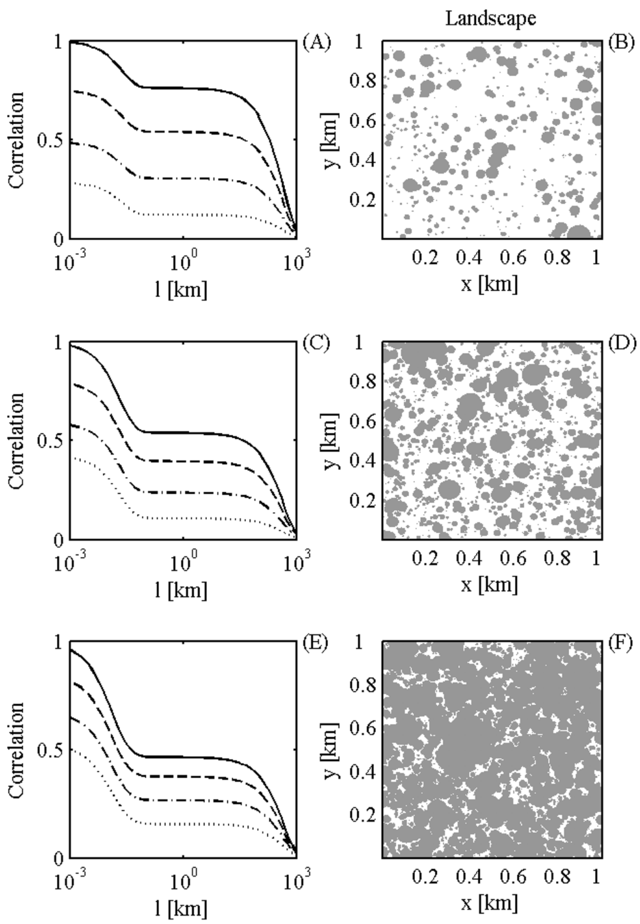

Figure 1. Examples of correlation functions of the relative soil moisture with different heterogeneous landscape obtained with ρ_T = 1/8 m^{-1} and λ_T equal to (a) 500, (c) 1500, and (e) 5000 km^{-2}. Tree coverage is (b) 18%, (d) 45%, (f) and 87%, respectively. Rainfall parameters are ρ_R = 0.0075 km^{-1}, η = 7.5 d^{-1}, β = 0.013 d mm^{-1} and λ_R = 2.9 × 10^{-6} km^{-2} d^{-1}, while the remaining parameters of vegetation are the same of Table 1. In each plot from top to bottom the lines correspond to h = 1, 10, 25, and 50 days, respectively.

The spatial structure of vegetation is described through a matrix of grass with trees located according to a Poisson process in space with rate λ_T which have random circular crowns with radii exponentially distributed with parameter ρ_T. This representation may provide a reasonable approximation for the vegetation patterns found, for instance, in savanna ecosystems (further details of the modeling scheme are reported by *Rodríguez-Iturbe et al.*, 2006).

[16] To derive the unconditional space-time soil moisture covariance, it is necessary to distinguish between points covered and not covered by a tree crown. The subscript c will be used for first ones and the subscript u for the uncovered points.

[17] The probability that a point A is not covered by a tree is

$$P_{A_u} = e^{-\lambda'_T}, \tag{8}$$

where $\lambda'_T = 2\pi\lambda_T/\rho_T^2$ represents the rate of tree crown occurrence at any point.

[18] The joint probabilities of occurrence of the different combinations of vegetation cover at two points A and B at distance l in space are given by

$$P_{A_uB_c} = P_{A_cB_u} = e^{-\lambda'_T}\left(1 - e^{-\lambda'_T + \lambda'_T\left(1+\frac{\rho_T l}{4}\right)e^{-\frac{\rho_T l}{2}}}\right), \tag{9}$$

$$P_{A_uB_u} = \exp\left[-2\lambda'_T + \lambda'_T\left(1 + \frac{\rho_T l}{4}\right)e^{-(\rho_T l/2)}\right], \tag{10}$$

$$P_{A_cB_c} = 1 - 2P_{A_uB_c} - P_{A_uB_u}. \tag{11}$$

[19] Using the above probabilities, *Rodríguez-Iturbe et al.* [2006] give the space-time covariance for the relative soil moisture between two points as

$$\begin{aligned} \mathrm{Cov}[S(0,t), S(l,t+h)] = {} & C_{A_cB_u}P_{A_cB_u} \\ & + C_{A_uB_c}P_{A_uB_c} + C_{A_uB_u}P_{A_uB_u} + C_{A_cB_c}P_{A_cB_c} \\ & + E^2[S_u]P_{A_uB_u} + 2E[S_u]E[S_c]P_{A_cB_u} \\ & + E^2[S_c]P_{A_cB_c} - (E[S_c]P_{A_c} + E[S_u]P_{A_u})^2. \end{aligned} \tag{12}$$

where point B corresponds to the later time and the conditional covariances C_{AB} are given by equation (3).

[20] The variance of the relative soil moisture in a heterogeneous landscape is obtained from equation (12) as

$$\begin{aligned} \sigma_S^2 = {} & \frac{4\pi\lambda_R}{\eta\rho_R^2\beta^2}\left(\frac{b_u^2(1-\mathrm{CF})}{a_u(\eta+a_u)} + \frac{b_c^2\mathrm{CF}}{a_c(\eta+a_c)}\right) \\ & + \frac{4\pi^2\lambda_R^2}{\eta^2\rho_R^4\beta^2}\left(\frac{b_u}{a_u} - \frac{b_c}{a_c}\right)^2(\mathrm{CF} - \mathrm{CF}^2), \end{aligned} \tag{13}$$

where CF = 1 − exp($-\lambda'_T$) represents the tree cover fraction over the landscape.

[21] Figure 1 shows an example of the correlation function of the relative soil moisture in a heterogeneous landscape. Two characteristic regimes are observed arising from the strong separation of scales between the rainfall forcing and the vegetation patterns. The first sharp reduction of correlation is related to the structure of vegetation and, as one may observe in Figure 1, to the density of tree coverage which is 18%, 45%, and 87%, respectively in this particular example. The second reduction is due to the rainfall forcing that imposes a large scale structure in the spatial correlation of the soil moisture field. The control of the atmosphere on the correlation structure of soil moisture fields at large spatial scales has been also point out by *Vinnikov et al.* [1996] and *Entin et al.* [2000].

[22] The mean value of the soil moisture in the heterogeneous vegetation case may be obtained as the sum of the means of covered soil and uncovered soil weighted by their relative probabilities

$$E[S] = \frac{2\pi\lambda_R}{\rho_R^2\eta\beta}\left(\frac{b_c}{a_c} + \left(\frac{b_u}{a_u} - \frac{b_c}{a_c}\right)e^{-\lambda'_T}\right). \tag{14}$$

When the entire landscape is covered by tree or grass, the mean assumes the same expression of the uniform vegetation case.

3. Soil Moisture Sampling in Time and Space

[23] Two descriptors of the soil moisture process are particularly relevant in hydrology: (1) the long-term mean soil moisture for a given time interval during a given season (e.g., daily soil moisture during month of June) at any point of a statistically homogeneous region, m_S, and (2) the mean soil moisture over an area $S_A = \frac{1}{A}\int_A S(x_i)da$, where $S(x_i)$ is the soil moisture at a site x_i and represents a realization of the soil moisture process over a region assumed statistically homogeneous. $S(x_i)$ may represent an instantaneous value or an average value over a particular time interval (e.g., mean soil moisture for a given day). The last one may be used for water resources management and ecohydrological analyzes. The first one is important for the calibration of remote sensing techniques and global circulation model (GCM). The estimation of m_S is described in the following, while that of S_A is presented in section 3.3. In both cases we will assume that daily soil moisture is the point variable one wishes to estimate either in a long-term basis or integrated over an area for a particular day.

[24] Assuming N sample points in space operating during T days of the same statistically homogeneous season, m_S is estimated through $\bar{S}$ as given by

$$\bar{S} = \frac{1}{NT}\sum_{t=1}^{T}\sum_{i=1}^{N} S(x_i, t). \tag{15}$$

[25] It is assumed that the individual samples, $S(x_i, t)$, adequately represent the average daily soil moisture at site x_i and day t. The goodness of the estimation is measured through the variance of $\bar{S}$,

$$\mathrm{Var}[\bar{S}] = \frac{1}{N^2T^2}E\left[\sum_{t=1}^{T}\sum_{i=1}^{N}(S(x_i, t) - m_S)\right]^2. \tag{16}$$

Equation (16) is a measures of the magnitude of the estimation errors, $\bar{S} - m_S$, that are random variables with zero mean. The variance of the estimate of m_S is equivalent to the point variance, σ_S^2, if there is only one sample in space and time ($T = 1$, $N = 1$). Any additional samples have the effect of reducing this variance, and the degree of that reduction is a subject of this paper.

[26] The variance of $\bar{S}$ is a function of the space-time correlation function of the process, the number of sampling points, the geometry of the sampling scheme and the length of operation of the stations. Following *Rodríguez-Iturbe and Mejía* [1974], the variance of $\bar{S}$ can be written as

$$\begin{aligned}\mathrm{Var}[\bar{S}] = \frac{1}{N^2T^2}E\Bigg[&\sum_{t=1}^{T}\sum_{i=1}^{N}\sum_{i'=1}^{N} f(x_i, t)f(x_{i'}, t)\\ &+2\sum_{t=1}^{T-1}\sum_{t'=t+1}^{T}\sum_{i=1}^{N}\sum_{i'=1}^{N} f(x_i, t)f(x_{i'}, t)\Bigg],\end{aligned} \tag{17}$$

where $f(x_i, t) = (S(x_i, t) - m_S)$.

[27] Equation (17) may be written as the product of two reduction factors affecting the variance of the daily soil moisture at a point, σ_S^2,

$$\mathrm{Var}[\bar{S}] = \sigma_S^2 F_1(T)F_2(N). \tag{18}$$

The time-dependent factor, $F_1(T)$, is given by

$$\begin{aligned}F_1(T) &= \frac{1}{T^2}\left[\sum_{t=1}^{T}1 + 2\sum_{t=1}^{T-1}\sum_{t'=t+1}^{T} r(t'-t)\right]\\ &= \frac{1}{T} + \frac{2}{T^2}\sum_{t=1}^{T-1}\sum_{t'=t+1}^{T} r(t'-t),\end{aligned} \tag{19}$$

and the space-dependent factor, $F_2(N)$, is

$$F_2(N) = \frac{1}{N^2}\left[\sum_{i=1}^{N}\sum_{i'=1}^{N} r(x_i - x_{i'})\right] = \frac{1}{N} + \frac{2}{N^2}\sum_{i=1}^{N-1}\sum_{i'=i+1}^{N} r(x_i - x_{i'}), \tag{20}$$

where $r(t)$ represents the correlation function in time and $r(x)$ in space [*Rodríguez-Iturbe and Mejía*, 1974].

[28] The reduction factor $F_1(T)$ can be derived explicitly from equation (19) for the case of uniform vegetation using the correlation function given in equation (7),

$$\begin{aligned}F_1(T) = \frac{1}{T} + \frac{2}{T^2(\eta - a)}&\left(\frac{\eta}{e^a - 1}\left(T - 1 - \frac{e^{-a} - e^{-aT}}{1 - e^{-a}}\right)\right.\\ &\left. - \frac{a}{e^\eta - 1}\left(T - 1 - \frac{e^{-\eta} - e^{-\eta T}}{1 - e^{-\eta}}\right)\right).\end{aligned} \tag{21}$$

In this case, the temporal reduction factor, $F_1(T)$, depends on the parameters η and a that are related to the physical characteristics of the soil (a) and the dynamics of rainfall (η). The inverse of each of these parameters represents a characteristics timescale; the former is the characteristic time of drying, equal to nZ_r/V, while the latter represents the mean storm duration, or the characteristic time of wetting. When $\eta \gg a$, the temporal correlation function of soil moisture may be simplified to $\exp(-ah)$, implying that the second term in the parenthesis on the right hand side of equation (21) becomes negligible. It is important to remark here that the lack of dependence of the correlation structure on the frequency of storm arrivals is due to the assumption of a Poisson process for rainfall occurrences.

[29] Figure 2 shows two examples of the reduction factor $F_1(T)$ as a function of the number of days of sampling, T, ranging from 1 to 200 days. Although the duration of a season through which the atmospheric forcing may be considered statistically homogeneous is likely to be less than 90 days, one could always sample repeatedly over that season throughout many years as long as the initial condition at the beginning of the season is not too close to saturation. In such a case, we should account for the saturation bound and the actual model cannot be applied. The graphs show the improvement in the estimation of m_S achieved by increasing the number of sampling days, T, for different values of nZ_r and η. Figure 2a shows that the increase of daily measurements drastically improves the estimation of m_S. Shallow root soils that are more susceptible to temporal fluctuations induced by rainfall and soil losses will quickly benefit from the temporal sampling and 100 days of measurements will accomplish a dramatic

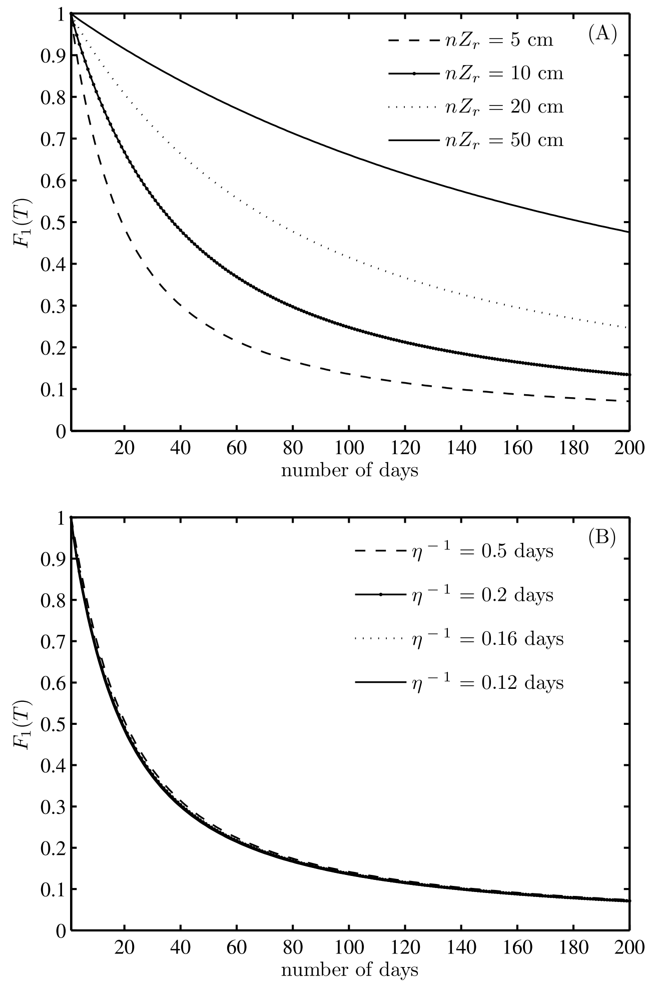

Figure 2. Variance reduction factor due to the temporal sampling, $F_1(T)$, for the case of the long-term mean daily soil moisture with homogeneous vegetation as a function of the number of days of measurement (a) for different value of nZ_r with fixed $V = 7$ mm d^{-1} and $\eta^{-1} = 0.16$ days and (b) for different values of η with fixed $V = 7$ mm d^{-1} and $nZ_r = 50$ mm.

reduction of the variance of the long-term mean daily soil moisture with respect to the point variance of the process. On the other hand, soils with deeper roots require much longer daily data to approach comparable reductions in the variance of the long-term mean daily soil moisture. One can also observe that $F_1(T)$ is not sensitive to the changes on the mean duration of the rain cells (see Figure 2b).

[30] With regard to the spatial factor, $F_2(N)$, its calculation depends on the spatial geometry of the sampling scheme. The two cases considered herein (random spatial sampling and stratified spatial random sampling) are discussed in the following.

3.1. Spatial Random Sampling

[31] In this case $F_2(N)$ (equation (20)) is the only reduction factor playing a role and it is further derived by *Rodríguez-Iturbe and Mejía* [1974] as

$$F_2(N) = \frac{1}{N^2}(N + N(N-1)E[r(x_i - x_{i'})|A]), \qquad (22)$$

where $E[r(x_i - x_{i'})|A]$ is the expected value of the correlation between two points randomly located in the area considered. For the case of a rectangular region, *Rodríguez-Iturbe and Mejía* [1974] obtain $E[r(x_i - x_{i'})|A]$ as

$$E[r(x_i - x_{i'})|A] = \int_0^R r(l)f(l)dl, \qquad (23)$$

where $f(l)$ is the probability density function of the distance between two randomly chosen points in the region (see Appendix A) and R represents the diagonal of the rectangular area. For the results of this paper we will assume the areas of interest are approximated by a square region.

[32] Since the spatial correlation of the relative soil moisture in the case of homogeneous vegetation (equation (6)) is solely a function of the dimensionless product $l\rho_R$, one can obtain the factor $F_2(N)$ as function only of $A\rho_R^2$ as shown in Figure 3 where it is plotted for different values of N. $F_2(N)$ decreases monotonically with N and $A\rho_R^2$. The decrease in $F_2(N)$ is more pronounced for increases in N when the number of stations is still small (e.g., from 1 to 5 stations). Also depending on the values of $A\rho_R^2$ one may obtain a relatively small or a very large decrease of $F_2(N)$ when increasing the number of stations. From a practical point of view, the scales of greater interest for hydrologists are frequently those with $A\rho_R^2 \leq 1$ (see Figure 3 inset). A value $A\rho_R^2 = 1$ generally implies an area of less than 10^4 km^2 given that typical values of ρ_R range between 10^{-1} and 10^{-2} km^{-1}. At those scales, we observe relatively minor changes in the function $F_2(N)$ whose role is clearly less important that the one played by the temporal reduction factor $F_1(T)$. Thus for the estimation of the long-term mean daily soil moisture in a region with homogeneous vegetation the length of operation of the network appears to be the commanding factor and trade of time versus space

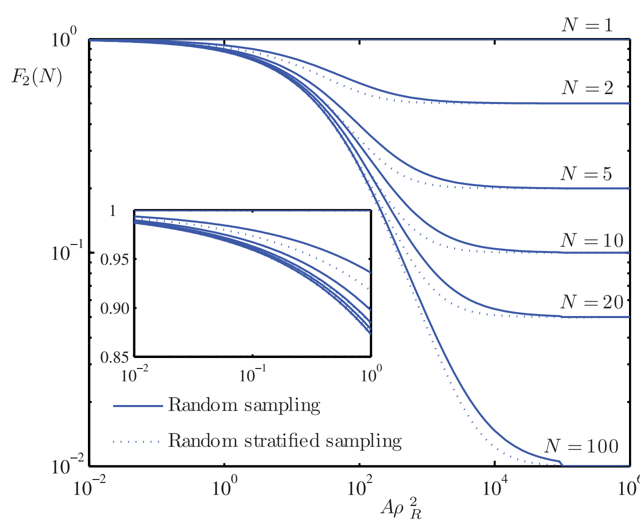

Figure 3. Variance reduction factor due to the spatial sampling, $F_2(N)$, for the case of the long-term mean daily soil moisture as function of $A\rho_R{}^2$ with homogeneous vegetation. Both spatial random sampling (solid line) and stratified sampling (dotted line) area are shown.

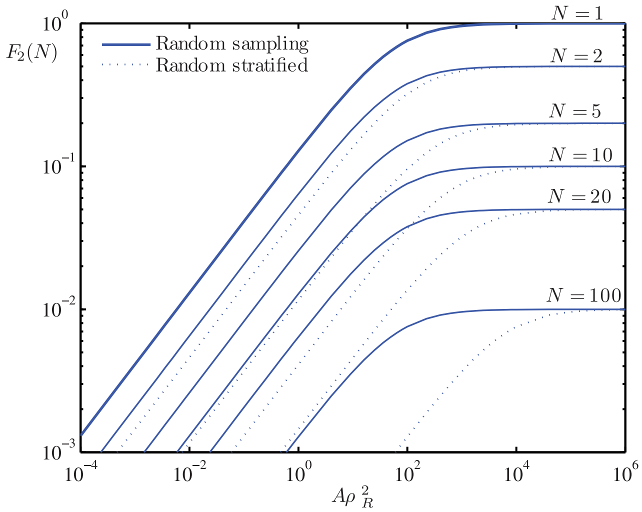

Figure 4. Variance reduction factor due to the spatial sampling, $F_2(N)$, for the case of the mean soil moisture over an area for a particular day with homogeneous vegetation cover. Both spatial random sampling (solid line) and stratified sampling (dotted line) area are shown.

using more stations over a shorter period of measurements is not a realistic option.

3.2. Stratified Spatial Random Sampling

[33] The variance reduction factor due to spatial sampling, $F_2(N)$, is now obtained by dividing the area A in non overlapping strata of area a. Following *Rodríguez-Iturbe and Mejía* [1974], $F_2(N)$ may be written as

$$F_2(N) = \frac{1}{N^2}\left(N + A^2\frac{E[r(x_i - x_{i'})|A]}{a^2} - NE[r(x_i - x_{i'})|(A/N)]\right). \tag{24}$$

[34] When there is only one station for each stratum $F_2(N)$ is

$$F_2(N) = \frac{1}{N^2}\left(N + N^2E[r(x_i - x_{i'})|A] - NE[r(x_i - x_{i'})|(A/N)]\right). \tag{25}$$

Equation (25) is plotted with a dotted line in Figure 3 where it can be observed that the use of stratified spatial random sampling does not provide significant improvements for the estimation of m_S respect to those obtained through a simple random sampling.

3.3. Mean Soil Moisture Over an Area for Any Given Day

[35] Estimates of the daily soil moisture averaged over a given area, S_A, are required in numerous cases like those aimed to provide input or validation of land surface models. Moreover, the use of remotely sensed images provides spatial and temporal description of soil moisture fields, but such techniques also need to be calibrated with ground based approaches.

[36] Following *Rodríguez-Iturbe and Mejía* [1974], in the case of a spatial random sampling, the performance of a network for the estimation of $S_A = \frac{1}{A}\int_A S(x_i)da$ may be described by the variance of the sample mean, $\widehat{S}$, using N points in space as

$$E\left[\left(\widehat{S} - S_A\right)^2\right] = \sigma_S^2\frac{(1 - E[r(x_i - x_{i'})|A])}{N} = \sigma_S^2F_2(N), \tag{26}$$

while in the case of stratified random sampling with the assumption of one station per strata, the variance reduction factor due to the spatial sampling is

$$F_2(N) = \frac{\{1 - E[r(x_i - x_{i'})(A/N)]\}}{N}. \tag{27}$$

[37] In contrast to m_S, the variance of the estimate of S_A will only be as large as the point variance, σ_S^2, if one has a single sample over an infinitely large area. Thus, for smaller areas (e.g., especially areas smaller than the correlation scale of the rain cells), the variance of the estimate will be less than the point variance. If the area is very small (i.e., smaller than the vegetation and rainfall correlation scales), then a single measurement will be adequate to determine the spatial average. Also, if the averaging area is very large (much larger than any correlation scales), then the variance of the estimate of S_A decreases as $1/N$ as confirmed by Figure 4.

[38] Figure 4 shows the reduction factor $F_2(N)$ computed using equations (26) and (27) with the correlation function of soil moisture for the case of a region with homogeneous vegetation. It is seen that the stratified random sampling may produce significant reductions on the errors of estimation of the areal soil moisture with respect to those occurring with a random design. The difference between the two sampling schemes becomes more significant with the increase of N. For cases where $A\rho_R^2 \leq 1$, the reduction factor $F_2(N)$ assumes relatively small values even with only few stations located in the region.

3.4. Soil Moisture Fields With Heterogeneous Vegetation

[39] The effects of vegetation heterogeneity on the soil moisture sampling are now addressed analyzing the impact of different vegetation covers resulting from varying the rate of occurrence of tree centers in space (λ_T).

[40] We will deal first with the estimation of the long-term mean soil moisture, m_S. In the case of heterogeneous vegetation cover, the algebraic expression for $F_1(T)$ is more complex and is not reported herein, but its numerical computation is straightforward using the correlation structure obtained from equations (12) and (13) in equation (19). Some examples are reported in Figure 5 that shows the variance reduction factor due to temporal sampling, $F_1(T)$, for different heterogeneous vegetation cover and with rainfall parameters estimated, during the period May/August, using a data set of 18 stations belonging to the state of Oklahoma (USA). In this graph, the solid lines represent the two extreme cases one with uniform grass vegetation (e.g., $\lambda_T = 0$) and the other for full coverage of trees. In general, the presence of a heterogeneous vegetation tends to increase the values of $F_1(T)$ with respect to those corresponding to a homogeneous cover. In the other hand, spatial sampling, as will be explained in the following, provides a stronger reduction of the variance of estimation than that obtained in the homogeneous vegetation case under comparable

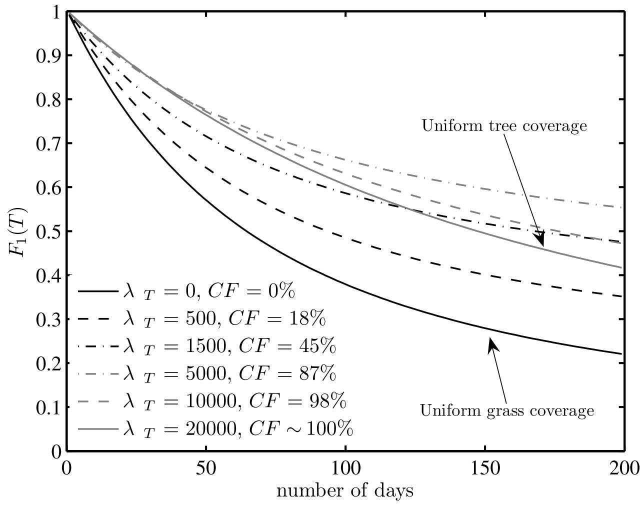

Figure 5. Variance reduction factor due to temporal sampling, $F_1(T)$, for the case of the long-term mean daily soil moisture in a region with heterogeneous vegetation for different values of the rate of tree occurrence, λ_T. Remaining parameters of vegetation are given in Table 1, and rainfall parameters are $\rho_R = 0.0075$ km^{-1}; $\eta = 7.5$ d^{-1}; $\beta = 0.013$ d mm^{-1}, and $\lambda_R = 2.9 \times 10^{-6}$ km^{-2} d^{-1}.

conditions. A larger value of $F_1(T)$ and a smaller $F_2(N)$ make trade of time versus space more appealing for the case of heterogeneous vegetation than for an homogeneous vegetation cover when estimating the long-term mean daily soil moisture over a region.

[41] Figure 6 shows an example of variance reduction factor due to the spatial sampling, $F_2(N)$, as a function of the dimensionless parameter $A\rho_R^2$ and for different values of N. The graph displays $F_2(N)$ for two types of sampling designs and for a region with a heterogeneous vegetation cover with

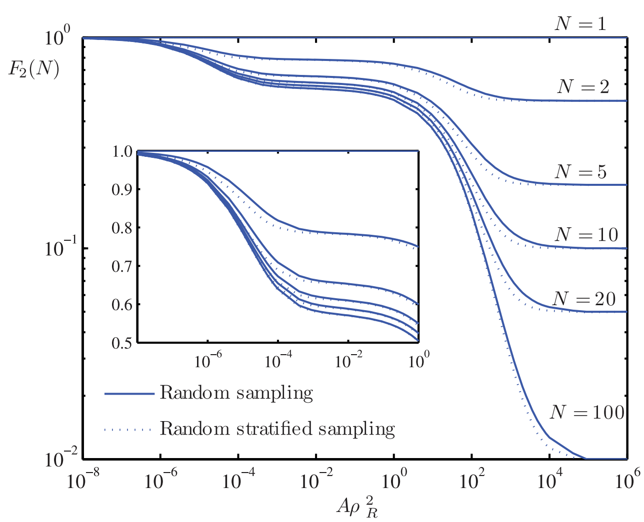

Figure 6. Variance reduction factor due to spatial sampling, $F_2(N)$, for the case of the long-term mean daily soil moisture in a region with heterogeneous vegetation. The solid line corresponds to random sampling and the dotted line to stratified random sampling. The parameters of vegetation are given in Table 1, and parameters of rainfall are the same as Figure 5.

Table 1. Parameters Adopted to Characterize the Two Functionally Different Vegetation Types[a]

Vegetation Parameter	Value
Tree	
$nZ_{r,t}$ [mm]	400
ϕ_T [dimensionless]	0.2
V_T [mm d^{-1}]	7.0
λ_T [km^{-2}]	1500
ρ_T^{-1} [km]	0.0080
Grass	
$nZ_{r,g}$ [mm]	100
ϕ_G [dimensionless]	0.05
V_G [mm d^{-1}]	4.0

[a]The percentage of tree cover is 45% in this case.

the parameters given in Table 1. This example shows the effect of heterogeneous vegetation on the parameter $F_2(N)$ that specially for the small values of $A\rho_R^2$ (see inset of Figure 6) is considerably smaller than $F_2(N)$ for the case of homogeneous vegetation shown in Figure 3. Similar to what was obtained for $F_2(N)$ for the case of the long-term mean daily soil moisture in a region with homogeneous vegetation, the stratified random sampling offers little advantage over the simple random sampling.

[42] The impact of vegetation heterogeneity on the estimation of the long-term mean daily soil moisture over a region is more clear in Figure 7, where the variance reduction factor, $F_2(N)$, for the case of $N = 10$ is given for different cases of vegetation structure. The results are a direct consequence of the characteristic shape of the correlation function of relative soil moisture (Figure 1). In particular, the presence of vegetation heterogeneity at small scales produces a reduction of the factor $F_2(N)$ for small values of $A\rho_R^2$ which is not present for the cases when λ_T is either very small or very large.

[43] We now address the impact of heterogeneous vegetation in the estimation of the mean soil moisture over an area for a particular day (S_A). Figure 8 shows the variance reduction factor, $F_2(N)$, for random and stratified random geometries. The subdivision of the domain in strata provides substantial improvement in the mean soil moisture estimation as the product $A\rho_R^2$ becomes smaller. The presence of an heterogenous vegetation cover leads to larger values of $F_2(N)$ relative to those obtained for the homogeneous vegetation case (Figure 4). This implies an increase in the number of stations in order to maintain comparable reduction factor. To evaluate the goodness of estimation, one needs to multiply the reduction factor times the variance of the soil moisture process, σ_S^2, as in equation (26). The value of σ_S^2 is larger for the heterogeneous vegetation case than for homogeneous vegetation case for comparable rainfall and soil conditions. The impact of heterogenous cover is now the opposite of the one found in $F_2(N)$ for the estimation of the long-term daily soil moisture over a region. This is particularly evident in Figure 9 where $F_2(N = 10)$ is shown for different values of the parameter λ_T.

4. Conclusions

[44] The space-time sampling of soil moisture fields has been quantitatively explored using a stochastic soil moisture

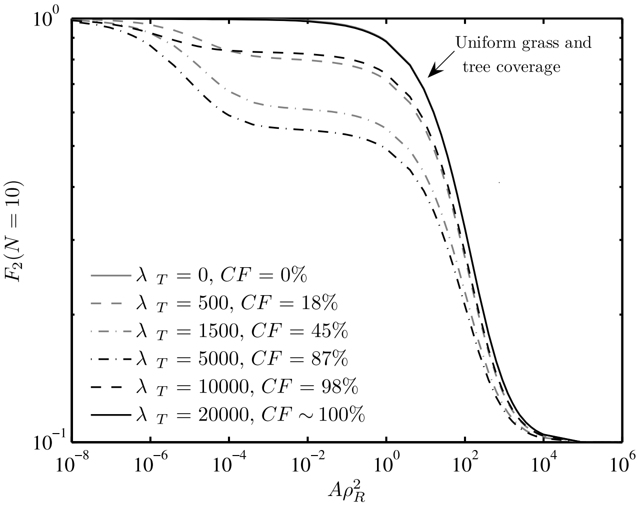

Figure 7. Variance reduction factor due to spatial sampling, $F_2(N = 10)$ for the case of the long-term mean daily soil moisture in a region with heterogeneous vegetation and different values of tree coverage. The scheme adopted is a random sampling with 10 stations. Remaining vegetation parameters are the same as in Table 1, and parameters of rainfall are the same as Figure 5.

model dependent on rainfall characteristics, soil properties and vegetation structure. The parameter $a = V/(nZ_r)$ has been identified as a key factor controlling of the temporal correlation function of the soil moisture process. Thus it strongly affects the sampling of the long-term mean soil moisture. In particular, an increase of a leads to a higher variability of the soil moisture dynamics and, as shown in Figure 2, for small values of nZ_r leads to a drastic decrease in the value of the variance reduction factor due to temporal sampling. This implies that networks with relatively short period of measurements will nevertheless provide significant reductions in the variance of estimation of the long-term mean daily soil moisture in soils with shallow roots, high permeability and high evapotranspiration rates.

[45] One may also notice, from equation (5), that the variance of the soil moisture is proportional to $(1 - \phi)^2/(nZ_rV\eta)$ when $\eta \gg a$. Thus, if from one side the variance reduction factor is smaller with smaller nZ_r, from the other side the actual variance is larger. On the contrary, the increase of the soil water losses, described through the parameter V, leads to a decrease of both the variance and $F_1(T)$. Consequently, the estimate of the long-term mean daily soil moisture will be more accurate in soils with high permeability and high evapotranspiration rates.

[46] In the case of heterogeneous vegetation cover the role of $F_1(T)$ is less commanding than in the homogeneous case but temporal sampling still remains a very important consideration for the estimation of the long-term mean daily soil moisture. The impact of the number of stations in the network depends heavily on the values of the parameter $A\rho_R^2$. In the case of homogeneous vegetation cover for $A\rho_R^2 \leq 1$, $F_2(N)$ plays a minor role in the estimation of the long-term mean daily soil moisture and there is little to gain with sizable increases in the number of stations in the network. This is not the case for regions with heterogeneous cover where $F_2(N)$ plays a more important role even for small values of $A\rho_R^2$. For both types of cover it is only important to go beyond $N = 5$ for cases when $A\rho_R^2 > 100$ showing that the greatest gain in information for the estimation of the long-term mean daily soil moisture in a region is obtained with an initial, relatively small, number of stations.

[47] The number of stations plays a crucial role, as expected, in the estimation of the average soil moisture over

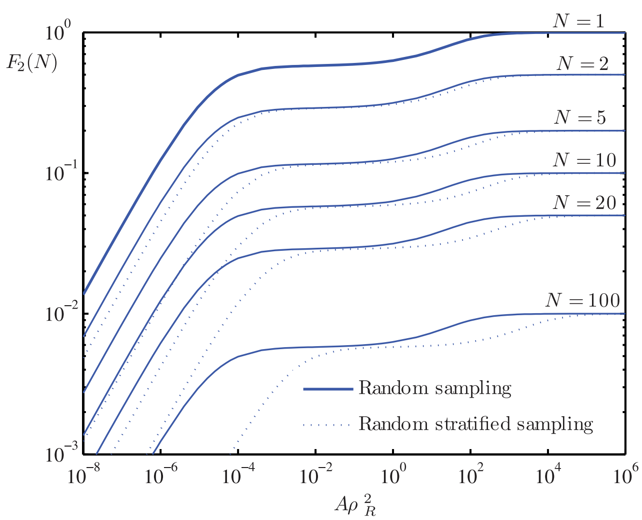

Figure 8. Variance reduction factor due to spatial sampling, $F_2(N)$ for the case of the mean soil moisture over an area for a particular day with heterogeneous vegetation cover. The solid line corresponds to random sampling, and the dotted line corresponds to stratified random sampling. The parameters of vegetation are the same of Table 1, and parameters of rainfall are the same as Figure 5.

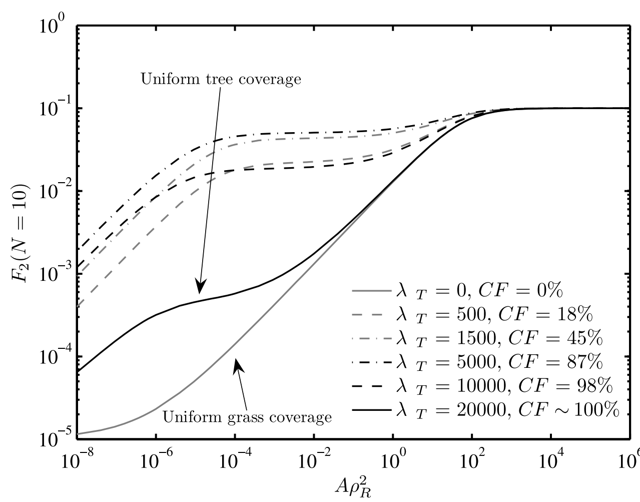

Figure 9. Variance reduction factor due to spatial sampling, $F_2(N = 10)$ for the case of the mean soil moisture over an area for a particular day with heterogeneous vegetation cover as a function of the parameter $A\rho_R{}^2$. The graph shows the effect of different vegetation patterns obtained assuming $\rho_T{}^{-1} = 0.0080$ km and varying the parameter λ_T. Other vegetation parameters are as given in Table 1, and parameters of rainfall are the same as Figure 5. The sampling geometry corresponds to random sampling with $N = 10$.

a region for a given day. This role is relatively more effective for the case of an homogeneous vegetation cover which yields values of $F_2(N)$ smaller than those obtained for the heterogeneous case in otherwise comparable conditions. This implies the need for a larger number of sampling stations over the region with heterogeneous cover to obtain a value of $F_2(N)$ similar to that of the homogeneous vegetation case.

[48] It is important to point out that although our discussion has been carried out in terms of the variance reduction factors $F_1(T)$ and $F_2(N)$, the variance of estimation of, both, the long-term mean daily soil moisture and the average soil moisture over an area in any particular day, are given by the decrease that $F_1(T)$ and $F_2(N)$ have over the point variance of the process (e.g., equations (18) and (26)). Moreover, the variance reduction factors are the only ones controlled by the length of operation and the spatial geometry of the sampling scheme, while the point variance of the process is not affected by those characteristics. The goodness of a sampling network for soil moisture fields should be judged by the values of the variance of $\overline{S}$ and $\widehat{S}$ which are a function of the point variance σ_S^2 and which itself depends on the climate, soil, and vegetation characteristics of the area under consideration.

[49] The network design has been analyzed considering two different sampling schemes: the random sampling and the stratified random sampling. It was found that the role of the network geometry is a secondary one in the estimation of the long-term mean daily soil moisture, while it is significant when one is interested on average soil moisture over an area for a particular day. The first goal is a common one in water resources management and ecohydrological studies while the second estimation is particularly useful to calibrate remote sensing techniques for soil moisture estimates and toward the validation of large-scale models (e.g., global circulation models) in the description of soil moisture dynamics.

Appendix A: Distribution of Distance of Random Points in a Given Area

[50] The probability density function (PDF) of the distance, l, between two randomly chosen points in a rectangular area has been derived by *Ghosh* [1951]. Given a rectangular area ($A = c \times d$) with sides c and d with ($c > d$), the PDF of l has the expression

$$f(l) = \frac{4l}{c^2 d^2}\psi(l), \tag{A1}$$

where

$$\psi(l) = \begin{cases} \frac{1}{2}\pi cd - l(c+d) + \frac{l^2}{2} & 0 \le l \le d, \\ cd \sin^{-1}\left(\frac{d}{l}\right) + c\sqrt{l^2 - d^2} - cl - \frac{d^2}{2} & d \le l \le c, \\ cd\left(\sin^{-1}\left(\frac{d}{l}\right) - \cos^{-1}\left(\frac{c}{l}\right) + \frac{\sqrt{l^2-d^2}}{d} + \frac{\sqrt{l^2-c^2}}{c}\right) - \frac{l^2 + c^2 + d^2}{2} & c \le l \le \sqrt{c^2 + d^2}. \end{cases} \tag{A2}$$

For a square area of side d, which is the case used in the this paper, the expression simplifies to

$$f(l) = \frac{4l}{d^4}\psi_1(l), \tag{A3}$$

where

$$\psi_1(l) = \begin{cases} \frac{1}{2}\pi d^2 - 2ld + \frac{l^2}{2} & 0 \le l \le d \\ d^2\left(\sin^{-1}\left(\frac{d}{l}\right) - \cos^{-1}\left(\frac{d}{l}\right)\right) + 2\sqrt{l^2 - d^2} + \frac{l^2 - 2d^2}{2} & d \le l \le \sqrt{2}d. \end{cases} \tag{A4}$$

More details on this topic and on the approximation of real regions with rectangular or squared shapes are given by *Rodríguez-Iturbe and Mejía* [1974].

Notation

- $a = V/(nZ_r)$ normalized soil water loss coefficient [d^{-1}].
- $b = (1 - \phi)/(nZ_r)$ normalized net rainfall coefficient that describes the effects of the vegetation interception [mm^{-1}].
- c and d sides of a rectangular area [km].
- $F_1(T)$ variance reduction factor due to the temporal sampling [dimensionless].
- $F_2(N)$ variance reduction factor due to the spatial sampling [dimensionless].
- h temporal lag [days].
- l distance between two points [km].
- m_S long-term mean soil moisture for a given time interval during a given season, [dimensionless].
- n soil porosity [dimensionless].
- R cell radius [km].
- $\overline{S}$ estimator for the long-term mean soil moisture m_S [dimensionless].
- $\widehat{S}$ estimator for the averaged soil moisture over an area in a given day [dimensionless].
- S_A mean soil moisture over an area in given day [dimensionless].
- $S(t)$ relative soil moisture at time t [dimensionless].
- V water soil loss coefficient [mm d^{-1}].
- Z_r depth of the root zone [mm].
- λ_R rate of rain cells per unit time and unit area [d^{-1} km^{-2}].
- λ_T rate of tree canopy per unit area [km^{-2}].
- ρ_T^{-1} mean value of the canopy radius [km^{-1}].
- μ_D mean value of rainstorm duration ($\eta = 1/\mu_D$) [km].
- μ_R mean cell radius ($\rho_R = 1/\mu_R$) [km].
- μ_X mean rainfall intensity ($\beta = 1/\mu_X$) [mm d^{-1}].
- ϕ interception coefficient [dimensionless].

[51] **Acknowledgments.** I.R.I. and S.M. gratefully acknowledge the support of NOAA under the grant NA17RJ2612, NSF under the National Center for Earth Surface Dynamics (NCED) grant EAR0120914, and Biocomplexity grant DEB0083566. S.M. also acknowledges the support of MIUR (Italian Ministry of Instruction, University and Research) through the research program 2003084552. We are thankful to Andrew Guswa for his thoughtful comments on the manuscript.

References

Albertson, J. D., and N. Montaldo (2003), Temporal dynamics of soil moisture variability: 1. Theoretical basis, *Water Resour. Res.*, *39*(10), 1274, doi:10.1029/2002WR001616.

Bras, R. L., and I. Rodríguez-Iturbe (1976), Network design for the estimation of area mean of rainfall events, *Water Resour. Res.*, *12*, 1185–1195.

Caylor, K. K., S. Manfreda, and I. Rodrguez-Iturbe (2005), On the coupled geomorphological and ecohydrological organization of river basins, *Adv. Water Resour.*, *28*(1), 69–86.

Cox, D. R., and V. Isham (1988), A simple spatial-temporal model of rainfall, *P. R. Soc. A*, *415*, 317–328.

Eagleson, P. S. (1978), Climate, soil and vegetation, 5, A derived distribution of storm surface runoff, *Water Resour. Res.*, *14*, 740–748.

Entekhabi, D., and I. Rodrìguez-Iturbe (1994), An analytic framework for the characterization of the space-time variability of soil moisture, *Adv. Water Resour.*, *17*, 25–45.

Entekhabi, D., I. Rodríguez-Iturbe, and F. Castelli (1996), Mutual interaction of soil moisture state and atmospheric processes, *J. Hydrol.*, *184*, 3–17.

Entin, J. K., A. Robock, K. Y. Vinnikov, S. E. Hollinger, S. Liu, and A. Namkai (2000), Temporal and spatial scales of observed soil moisture variations in the extratropics, *J. Geophys. Res.*, *105*, 11,865–11,877.

Ghosh, B. (1951), Random distances within a rectangle and between two rectangles, *Bull. Calcutta Math. Soc.*, *43*(1), 17–24.

Isham, V., D. R. Cox, I. Rodríguez-Iturbe, A. Porporato, and S. Manfreda (2005), Mathematical characterization of the space-time variability of soil moisture, *P. R. Soc. A*, *461*(2064), 4035–4055, doi:10.1098/rspa.2005.1568.

Jacobs, J. M., B. P. Mohanty, E. C. Hsu, and D. Miller (2004), SMEX02: Field scale variability, time stability and similarity of soil moisture, *Remote Sens. Environ.*, *92*, 436–446.

Kagan, R. L. (1979), *Averaging Meteorological Fields* (in Russian), 212 pp., Gidrometeoizdat, St. Petersburg, Russia.

Lull, H. W. (1964), Ecological and silvicultural aspects, in *Handbook of Applied Hydrology*, edited by V. T. Chow, pp. 6-1 to 6-30, McGraw-Hill, New York.

Manfreda, S., V. Iacobellis, and M. Fiorentino (2005), DREAM: a distributed model for runoff, evapotranspiration, and antecedent soil moisture simulation, *Adv. Geosci.*, *2*, 31–39.

Meron, E., E. Gilad, J. von Hardenberg, M. Shachak, and Y. Zarmi (2004), Vegetation patterns along a rainfall gradient, *Chaos Soliton Fract.*, *19*, 367–376.

Mitchell, K. E., et al. (2004), The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system, *J. Geophys. Res.*, *109*, D07S90, doi:10.1029/2003JD003823.

North, G. R., and S. Nakamoto (1989), Formalism for comparing rain estimation designs, *J. Atmos. Oceanic Technol.*, *6*(6), 985–992.

Pan, F., C. D. Peters-Lidard, and M. J. Sale (2003), An analytical method for predicting surface soil moisture from rainfall observations, *Water Resour. Res.*, *39*(11), 1314, doi:10.1029/2003WR002142.

Porporato, A., P. D'Odorico, L. Ridolfi, and I. Rodríguez-Iturbe (2000), A spatial model for soil-atmosphere interaction: model construction and linear stability analysis, *J. Hydrometeorol.*, *1*(1), 61–74.

Porporato, A., E. Daly, and I. Rodríguez-Iturbe (2004), Soil water balance and ecosystem response to climate change, *Am. Nat.*, *164*(5), 625–633.

Robock, A., et al. (2003), Evaluation of the North American Land Data Assimilation System over the southern Great Plains during the warm season, *J. Geophys. Res.*, *108*(D22), 8846, doi:10.1029/2002JD003245.

Rodríguez-Iturbe, I. (2000), Ecohydrology: A hydrologic perspective of climate-soil-vegetation dynamics, *Water Resour. Res.*, *36*, 3–9.

Rodríguez-Iturbe, I., and J. M. Mejía (1974), The design of rainfall networks in time and space, *Water Resour. Res.*, *10*, 713–728.

Rodríguez-Iturbe, I., and A. Porporato (2004), *Ecohydrology of Water-controlled Ecosystems: Soil Moisture and Plant Dynamics*, Cambridge Univ. Press, New York.

Rodríguez-Iturbe, I., V. Isham, D. R. Cox, S. Manfreda, and A. Porporato (2006), An analytical framework for the space-time variability of soil moisture fields with heterogeneous vegetation, *Water Resour. Res.*, *42*, W06D05, doi:10.1029/2005WR004497.

Scanlon, T. M., K. K. Caylor, S. Manfreda, S. A. Levin, and I. Rodríguez-Iturbe (2005), Dynamic response of grass cover to rainfall variability: Implications the function and persistence of savanna ecosystems, *Adv. Water Resour.*, *28*, 291–302.

Vinnikov, K. Y., A. Robock, N. A. Speranskaya, and C. A. Schlosser (1996), Scales of temporal and spatial variability of midlatitude soil moisture, *J. Geophys. Res.*, *101*, 7163–7174.

Vinnikov, K. Y., A. Robock, S. Qiu, and J. K. Entin (1999), Optimal design of surface networks for observation of soil moisture, *J. Geophys. Res.*, *104*, 4145–4168.

Walker, J. P., G. R. Willgoose, and J. D. Kalma (2001), One-dimensional soil moisture profile retrieval by assimilation of near-surface observations: A comparison of retrieval algorithms, *Adv. Water Resour.*, *24*, 631–650.

Western, A. W., R. B. Grayson, and G. Blöschl (2002), Scaling of soil moisture: a hydrologic perspective, *Annu. Rev. Earth Planet. Sci.*, *30*, 149–180.

Yoo, C. (2001), Sampling of soil moisture fields and related errors: implications to the optimal sampling design, *Adv. Water Resour.*, *24*(5), 521–530.

S. Manfreda, Dipartimento di Ingegneria e Fisica dell'Ambiente, Università degli Studi della Basilicata, Potenza I-85100, Italy. (manfreda@unibas.it)

I. Rodríguez-Iturbe, Department of Civil and Environmental Engineering, Princeton University, Princeton, NJ 08540, USA.