

# An Optimal Sampling Design to Capture the Watershed-Scale Soil Moisture Dynamic in a Tropical Agricultural Watershed of Eastern India



Gurjeet Singh  and Rabindra Kumar Panda

**Abstract** Understanding the soil moisture dynamic at the watershed-scale is essential for hydrological applications (i.e., drought monitoring, flood forecasting, irrigation, etc.) and river basin management activities. Globally, in-situ measured accurate soil moisture data at the watershed-scale is quite scarce due to the high maintenance cost of large-density sensor networks and complex/challenging conditions for soil moisture field campaigns. Characterizing the soil moisture variability at the watershed-scale requires a robust in-situ monitoring strategy at the point-scale to balance representativeness and minimization of monitoring cost. Thus, this study determined an optimal sampling design to capture the spatiotemporal variability of soil moisture at the watershed-scale. The study was conducted for the typical eastern Indian conditions of extreme seasonal variability that lead from very wet (during monsoon) to dry (during hot summer). Soil moisture measurements were carried out at 83 locations in an agricultural watershed of 500 km<sup>2</sup> for 56 days across a year. A hand-held soil moisture probe (ThetaProbe) was used for the measurements from June 2016 to July 2017. Based on the analyzes of 41,832 measurements collected during field measurements, it was found that the maximum numbers of required locations necessary to estimate watershed-mean soil moisture within  $\pm 2\%$  absolute error are 30. Moreover, the five most representative locations identified through time stability analysis were found to be sufficient for capturing the temporal pattern of watershed-mean soil moisture with a root-mean-square error of  $\pm 2.17\%$ . The findings will be helpful in providing guidelines for optimizing short-term measurement and a robust sensor network to assess the watershed-scale soil moisture dynamics.

**Keywords** Soil moisture · Spatiotemporal variability · Tropical watershed · Time stability · Watershed-scale

---

G. Singh (✉)

Department of Civil and Environmental Engineering, Michigan State University, East Lansing, MI 48824, USA

e-mail: [singhg23@msu.edu](mailto:singhg23@msu.edu)

R. K. Panda

Centre for Climate Smart Agriculture, Siksha 'O' Anusandhan University, 751030, Bhubaneswar, Odisha, India

## 1 Introduction

Surface soil moisture is a key dynamic hydrological state variable [28] that affects various hydrological process. Spatiotemporal variation of surface soil moisture significantly affects the watershed-scale soil moisture dynamics, which may result in variations of surface atmospheric feedback, runoff dynamics, groundwater recharge, and crop yield. Understanding surface soil moisture dynamics at a watershed-scale with reasonable temporal and spatial resolution is required for hydrological modeling [7, 19, 30], climate modeling and weather prediction [1, 12, 20] and agricultural modeling [2, 23, 31]. Thus, precise measurements of surface soil moisture at the watershed-scale are necessary to fulfill the requirements of various applications.

The spatiotemporal variation in geophysical characteristics such as rainfall, vegetation cover, soil properties, and topography over a watershed-scale makes soil moisture distribution highly nonlinear across time and space. Therefore, measuring soil moisture in a large watershed ( $>100 \text{ km}^2$ ) to represent the average soil moisture dynamic accurately is challenging. The conventional in-situ point scale measurement (i.e., gravimetric sampling) of soil moisture is precise. Still, gravimetric sampling is not practicable for watershed-scale soil moisture measurement, where many in-situ observations are needed at frequent intervals. The use of electronic sensors-based geophysical techniques with fixed-type automatic data-logging devices for continuous in-situ soil moisture measurement eliminates the need for time-intensive gravimetric sampling [6]. But, it is expensive to maintain a sufficiently large-density sensor network for watershed-scale soil moisture assessment. Soil moisture scaling theory and soil moisture spatiotemporal variability analysis reveals that a reliable estimate of large-scale soil moisture could be obtained using a few-point observation [3, 10, 13, 22]. Inline this context, past studies show the potential of soil moisture spatiotemporal variability analysis with a few statistical analyzes to determine the number of required samples (NRS) to estimate mean soil moisture of a large area [3–5, 14, 29]. In addition, a hypothesis on the temporal stability of soil moisture spatial pattern [27] provided an opportunity to minimize the NRS to capture the temporal pattern of the soil moisture over a large region [4, 5, 10 and 11].

Notably, these studies also suggested that areal mean soil moisture assessment using a few-point observation and scale of temporal stability must be established using dense soil moisture measurements over a large region. However, due to expenses and limited conditions experienced in soil moisture field campaigns, monitoring is complex, and soil moisture spatiotemporal variability analysis in an area greater than  $100 \text{ km}^2$  [5] is poorly understood. Over the past, studies focused on measuring soil moisture and its variability analysis either favoring the large spatial scale [8, 14, 17, 18] or long time periods [4, 15, 16, 21], but very few consider both of them [5, 9]. Besides, tropical regions are rarely studied for soil moisture variability analysis, which may hold key differences from previous studies due to very high variability in soil moisture from very dry to very wet soil conditions.

In view of the facts mentioned above, the present study aimed to investigate the long-term soil moisture spatiotemporal variability over a large tropical agricultural

watershed ( $>100 \text{ km}^2$ ) for the optimal sampling design of soil moisture measurements. For this objective, frequent in-situ measurements at various point locations were carried out in an eastern Indian watershed of  $500 \text{ km}^2$  to capture different soil wetness conditions for a year.

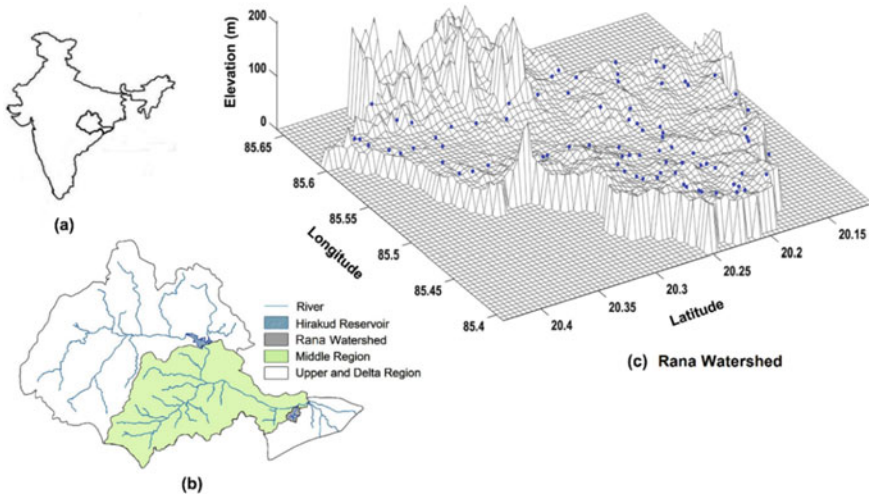
## 2 Study Area and Measurements

### 2.1 Study Area

An agricultural watershed, namely, Rana watershed of approximately  $500 \text{ km}^2$ , located in the middle region of the Mahanadi River basin of eastern India, is selected for this study (Fig. 1). The climate of the watershed is tropical, having marked seasonality in rainfall with the long dry season. The average annual rainfall of the study area is 1458 mm, mainly occurring (about 70%) in the southwest monsoon season of June to September [24]. Due to the tropical climate, the study area experiences very high temperatures during April and May. The mean annual temperature of this region is  $27.4 \text{ }^\circ\text{C}$  with maximum and minimum temperatures of  $42.2 \text{ }^\circ\text{C}$  and  $11.3 \text{ }^\circ\text{C}$ . Due to the high climate variability, the study watershed experiences varying soil wetness conditions, i.e., very dry to very wet and back to very dry conditions. Paddy is a dominant crop in the study region and is usually cultivated during the southwest monsoon season. Elevation in the watershed ranges between 22 and 299 m (see Fig. for the topography view). Based on the soil textural analysis of various locations of the watershed it was found that soil texture class of the watershed varies considerably from sandy loam to clay, where a major portion of the watershed has sandy loam, followed by sandy clay loam and clay loam.

### 2.2 Soil Moisture Measurements

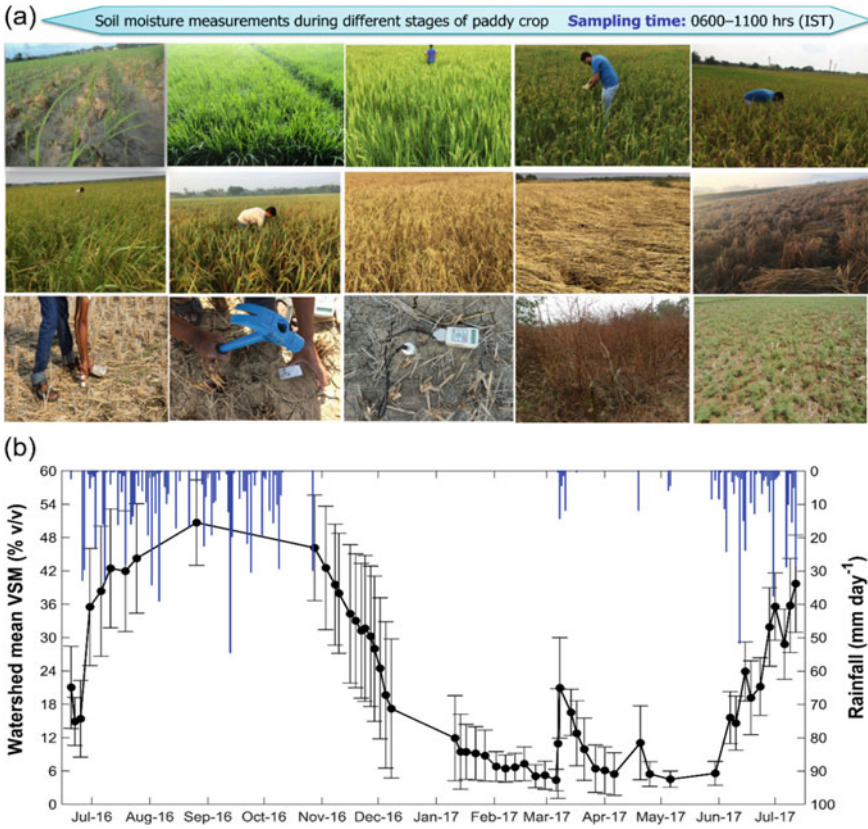
With the aim of achieving dense sampling at a large spatial scale for a long period, 83 locations (i.e., 79 agricultural fields and 4 grasslands) were selected within the watershed of  $500 \text{ km}^2$  for measuring soil moisture (see Fig. 1). In addition to the vegetation, the criteria for choosing the sampling locations were geophysical characteristics that affect the spatiotemporal variation of the soil moisture, such as soil texture and elevation. The choice of sampling periods was based on the criterion of minimum interaction with human activities, such as tillage. Therefore, soil moisture measurements were initiated in June when the paddy crop was planted in most fields. Samplings were not carried out for a one-month duration (10 December 2016–10 January 2017) due to tillage activities after the paddy crop in a few areas of the watershed for the cultivation of pulse crops (i.e., moong). Overall, soil moisture measurements were conducted for nearly one year, from 20 June 2016 to 12 July



**Fig. 1** **a** Location of Mahanadi river basin in India **b** Location of the study area, Rana watershed in Mahanadi River basin **c** Topographic overhead view of Rana watershed with sampling locations

2017, to capture the entire range of soil moisture variability from dry to wet and wet to dry conditions. Due to unfavorable conditions, sampling was not carried out frequently from August to October because of heavy rainfall and standing water in most paddy fields. Similarly, measurements were not taken frequently in April and May because of the soil's hardness (i.e., very dry condition). Overall, a total of 56 days of soil moisture sampling was carried out in one year. A view of soil moisture measurements during field campaigns is presented in Fig. 2a. The temporal pattern of the watershed-mean soil moisture based on soil moisture sampling at 83 sampling locations is also shown in Fig. 2b.

Soil moisture was sampled using an impedance probe (ThetaProbe, type ML3 and HH2 recording device, Delta-T Devices, Cambridge, England) which consists of four sharpened, 6 cm long stainless-steel rods. For each sampling location, three-point measurements of ThetaProbe were taken at 10–15 m separation distances, and each measurement consisted of three ThetaProbe samples. A total of nine ThetaProbe samples were taken from each sampling location to reduce the uncertainty in the estimates of mean soil moisture of a sampling location. Figure 3 shows a view of the sampling design and soil moisture measurements. During each sampling day, 747 soil moisture measurements were carried out, and a total of 41,832 samples were collected in 56 sampling days. In addition, on each sampling day, 10% of the total locations were also sampled with gravimetric sampling. Gravimetric samples were taken adjacent to one of the ThetaProbe samples with a soil core sampler having a fixed volume of 137 cm<sup>3</sup> with a 6 cm depth. The measured impedance from ThetaProbe was calibrated using gravimetric-based volumetric soil moisture content. A single generalized calibration of Thetaprobe was developed [25] for the watershed and used for precise soil moisture measurements at each sampling location.



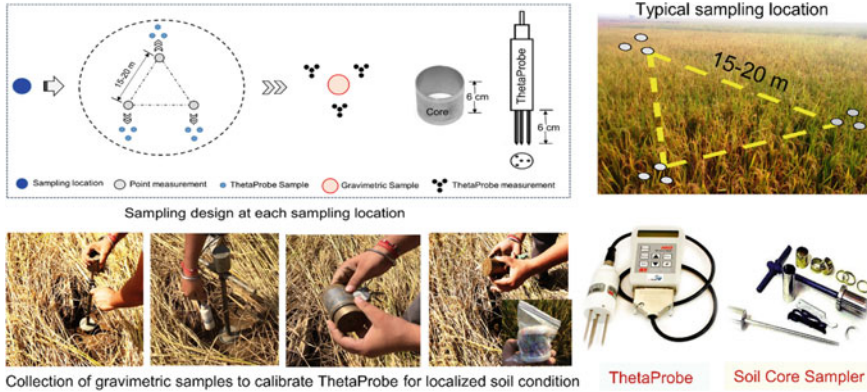
**Fig. 2** **a** A view of soil moisture measurements during different stages of the paddy crop and field conditions throughout the year. **b** Temporal pattern of the watershed-mean volumetric soil moisture (VSM), along with error bars of  $\pm 1$  standard deviation in space. The blue bars represent a variation of daily rainfall measured by a weather station in the watershed

### 3 Methodology

The statistical properties of each sampling day and whole field campaign are analyzed in terms of their variability in space and time for the soil moisture spatiotemporal variability analysis as given below:

Let  $\theta_{ijk}$  the soil moisture measured at point  $i$ , sampling location  $j$  and sampling day  $k$ , then the spatial mean of the sampling location  $k$  and sampling day,  $\overline{\theta_{jk}}$ , is given by.

$$\overline{\theta_{jk}} = \frac{1}{N_p} \sum_{i=1}^{N_p} \theta_{ijk} \tag{1}$$



**Fig. 3** Sampling design at each sampling location, where a large blue circle represents the sampling location and three-point measurements at the sampling location forming a triangle is shown with a gray circle. Small black circles show the shape of ThetaProbe measurement, and the red circle represents the position of the gravimetric sample. A view of the gravimetric sample collection in the rice field is presented through photographs

where  $N_p$  is the number of measurement points at the sapling location  $j$ . Similarly, the spatial mean of each sampling day,  $\bar{\theta}_k$ , can be defined as.

$$\bar{\theta}_k = \frac{1}{N} \sum_{j=1}^N \bar{\theta}_{jk} \tag{2}$$

where  $N$  is the number of sampling locations. Also, the temporal mean for each sampling location,  $\bar{\theta}_j$ , can be defined as:

$$\bar{\theta}_j = \frac{1}{M} \sum_{k=1}^M \bar{\theta}_{jk} \tag{3}$$

where  $M$  is the number of sampling days.

The coefficient of variation of each sampling day in space,  $CV_k$ , is calculated as follows:

$$CV_k = \frac{\sigma_k}{\bar{\theta}_k} = \frac{\sqrt{\frac{1}{N-1} \sum_{j=1}^N (\bar{\theta}_{jk} - \bar{\theta}_k)^2}}{\bar{\theta}_k} \tag{4}$$

where  $\sigma_k$  is the standard deviation in space for a sampling day.

Determination of standard deviation helps in the assessment of an optimal number of sampling locations (ONL) for estimating the mean soil moisture within a specified range of absolute error. The robust monitoring strategy has been optimized for

the watershed-mean soil moisture assessment through soil moisture spatiotemporal variability analysis using a statistical approach and time stability method.

### 3.1 ONL Analyzes Using a Statistical Approach

ONL, for watershed-mean soil moisture ( $\bar{\theta}_k$ ) assessment within a specific value of absolute error, has been determined using Eq. 5 given by Gilbert (1987) and is expressed as:

$$ONL = \left( t_{1-\alpha/2, ONL-1} \frac{\sigma_E}{AE} \right)^2 \tag{5}$$

where  $\sigma_E$  is a standard error,  $t_{1-\alpha/2, ONL-1}$  is the value of the student’s t-distribution at a significance level  $\alpha$ , and depending on the sample dimensions ONL; AE represents the absolute error (% v/v). If a reliable value of  $\sigma_E$  is not available, CV can be used to estimate ONL (Gilbert, 1987). For this, a functional relationship between the  $CV_k$  as well as  $\sigma_k$  and the  $\bar{\theta}_k$  is investigated. An exponential function  $CV_k = k_1 \cdot e^{-k_2 \bar{\theta}_k}$  was used to fit the  $CV_k - \bar{\theta}_k$  relationship for characterizing the variations of soil moisture, as usually employed for soil moisture campaigns [3, 14], where  $k_1$  and  $k_2$  are the model fitting parameters. Based on the  $CV_k - \bar{\theta}_k$  fitting, a relationship between  $\sigma_k$  and  $\bar{\theta}_k$  can be derived as  $\sigma_k = k_1 \cdot \bar{\theta}_k \cdot e^{-k_2 \bar{\theta}_k}$ . Further, the  $\sigma_k - \bar{\theta}_k$  relationship is used to calculate the uncertainty in watershed-mean soil moisture assessment from a certain number of soil moisture samples, including its evolution with drying or wetting.

Moreover, the exponential law described by  $CV_k - \bar{\theta}_k$  is employed to determine the ONL to achieve a specified uncertainty using  $\sigma_k - \bar{\theta}_k$ . Equation 6 is solved by an iterative procedure to estimate the ONL at a 5% significance level for different absolute errors (AE). The iterative process is repeated until  $|ONL_l - ONL_{l-1}| \leq \varepsilon$ , where  $\varepsilon$  is a control value, 0.5 in this study.

$$ONL_l = \left( t_{0.975, ONL_{(l-1)}-1} \frac{k_1 \cdot \bar{\theta}_k \cdot e^{-k_2 \bar{\theta}_k}}{AE} \right)^2, \quad l = 1, 2, 3, \dots \tag{6}$$

The statistical approach quantifies the ONL to determine the sampling size for capturing the spatial variability of soil moisture at the watershed-scale. But most hydrological applications require a temporal pattern of watershed-mean soil moisture. Besides, the exact position of the ONL in the study domain is essential to set up a soil moisture sensor network. Since the statistical approach fails to characterize the temporal pattern of the watershed-mean soil moisture, a time stability method is used to identify locations where the soil moisture can be considered “representative” of the entire area of study at the temporal scale.

### 3.2 ONL Analyzes Using Temporal Stability Analysis

The temporal stability analysis [27] identifies the sampling locations that maintain a consistent temporal relationship with the areal mean soil moisture with little variability. This study conducts temporal stability analysis based on the parametric test of the relative differences in soil moisture. For each sampling location  $j$  and total sampling days  $M$ , the mean relative difference of soil moisture,  $\bar{\delta}_j$  (% v/v), and variance of the relative difference,  $\sigma(\delta_j)^2$ , is estimated as:

$$\bar{\delta}_j = \frac{1}{M} \sum_{k=1}^M \frac{\bar{\theta}_{jk} - \bar{\theta}_k}{\bar{\theta}_k} \quad (7.1)$$

$$\sigma(\delta_j)^2 = \frac{1}{M-1} \sum_{k=1}^M \left( \frac{\bar{\theta}_{jk} - \bar{\theta}_k}{\bar{\theta}_k} - \bar{\delta}_j \right)^2 \quad (7.2)$$

$\bar{\delta}_j$ , at a sampling location computes the location's bias and helps to identify whether a particular location is wetter or drier than the areal mean. Generally, a "representative" location to capture the temporal pattern of the areal mean soil moisture can be identified by the low value of  $|\bar{\delta}_j|$  and/or standard deviation of the relative difference,  $\sigma(\delta_j)$ . [17] considered combining  $\bar{\delta}_j$  and  $\sigma(\delta_j)$  statistical metrics relative difference and presented a comprehensive evaluation criterion (CEC) to include both the bias and accuracy to locate the best time-stable locations.

$$CEC_j = \sqrt{(\bar{\delta}_j)^2 + \sigma(\delta_j)^2} \quad (8)$$

Based on the rank-ordered  $CEC_j$ , the sampling location with the highest time stability is identified as the one with the lowest  $CEC_j$  value.

## 4 Results and Discussion

The temporal pattern of measured soil moisture was found to be highly linked to rainfall (see Fig. 2b). The measured soil moisture statistical analysis shows that the spatial  $CV_k$  ranges between 0.151 and 0.901.  $CV_k$  was found to be very high during dry periods, whereas low in wet periods and, on average equal to 0.404. On the other hand, the temporal  $CV_j$  was found to have an average of 0.723, considerably higher than  $CV_k$ , and ranges between 0.578 and 0.887. This confirms that soil moisture temporal variability is more significant than spatial variability and indicates that ONL to capture the temporal pattern of the watershed-mean soil moisture can be derived in this study region. The high  $CV_k$  value follows the findings reported in the past investigation on soil moisture variability in relation to the spatial variability of



soil moisture and the dimension of the investigated area [3–5, 14]. Specifically, the spatial  $CV_k$  increases with the increase in the area, where average  $CV_k$  ranges from 0.06 to 0.20 for the area of 1 m<sup>2</sup> and 250 km<sup>2</sup>, respectively, at 0–15 cm depth in central Italy [5]. Comparing  $\sigma_k$  and  $CV_k$  values presented in [14], the values found for the eastern Indian watershed are very similar.

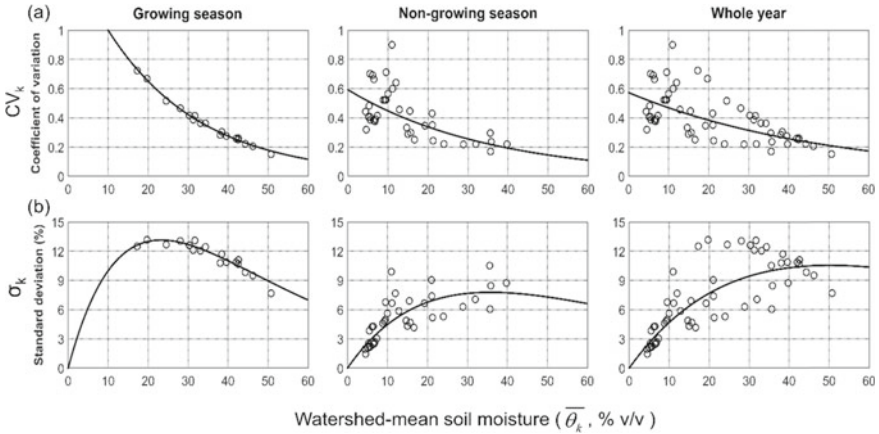
### 4.1 ONL to Capture Watershed-Mean Soil Moisture

Based on the in-situ measurements, an analytical relationship was fitted between  $\bar{\theta}_k$  and the  $CV_k$  for the growing season, non-growing season, and the whole year, as presented in Table 1, and Fig. 4a. Further, the fitted parameters of  $CV_k$  versus  $\bar{\theta}_k$  relationship were utilized to derive a relationship between  $\sigma_k$  versus  $\bar{\theta}_k$  as shown in Fig. 4b to capture the spatial variability of soil moisture. The distribution of  $\sigma_k$  versus  $\bar{\theta}_k$  shows that  $\sigma_k$  increases until  $\bar{\theta}_k$  reaches around 30% and decreases beyond that, whereas  $CV_k$  shows a rapid decreasing pattern with increasing  $\bar{\theta}_k$ . It was also observed that  $CV_k$  has a very less scattered pattern as  $\bar{\theta}_k$  decreases during the growing season but is found to be a widely scattered pattern during the non-growing season. Overall, a decreasing exponential pattern of  $CV_k - \bar{\theta}_k$  and a convex upward trend in  $\sigma_k - \bar{\theta}_k$  reported in this study is similar to those reported in previous studies across the world [3–5, 13, 14, 17] (Fig. 4).

Further, the fitted decreasing exponential pattern of  $CV_k, \bar{\theta}_k$  was used to quantify the ONL as a function of the average wetness condition and in relation to a prefixed significance level and varying absolute errors. Figure 5a. demonstrates the ONL (using Eq. 6) to capture the watershed-mean soil moisture with a 5% significance level and within an absolute error of  $\pm 2\%$ , during different seasons. The ONLs were found to be equal to 30 for the growing season and fewer for the non-growing season, equivalent to 20, whereas a maximum ONL of 25 is needed for the whole year to assess watershed-mean soil moisture. The maximum ONL found to be 30 that is during the growing season confirms the effectiveness of the sampling design adopted for this study, where 83 locations were monitored. Analysis of different absolute errors for ONL shows that fewer resources are needed on higher absolute error ( $> \pm 2\%$ ) for watershed-mean soil moisture estimation (Fig. 5b). The reported analysis reveals that ~10 to 15 sampling locations are sufficient to capture the spatial pattern of the watershed-mean having an area of 500 km<sup>2</sup> with a more relevant range of absolute

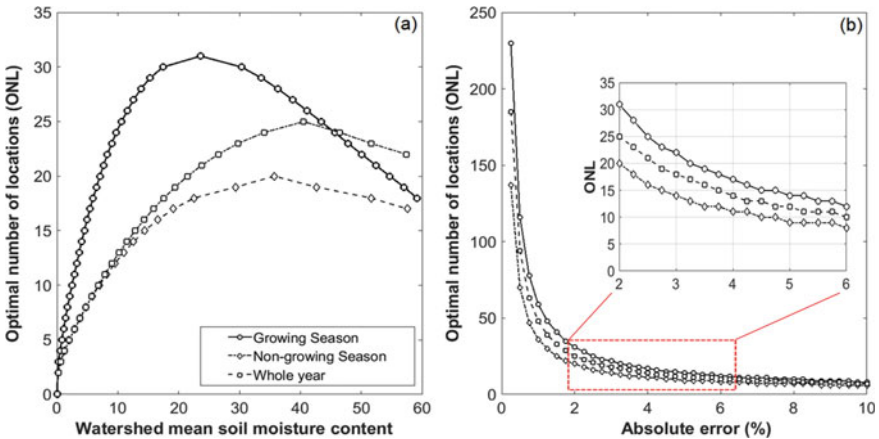
**Table 1** The exponential fitting parameters of  $CV_k$  versus  $\bar{\theta}_k$  relationship and the corresponding coefficient of determination ( $R^2$ )

Season	$CV_k$ versus $\bar{\theta}_k$ relationship		
	$k_1$	$k_2$	$R^2$
Growing season	1.537	- 0.043	0.994
Non-growing season	0.592	- 0.028	0.543
Whole year	0.573	- 0.020	0.498



**Fig. 4** Analytical relationship between watershed-mean soil moisture ( $\bar{\theta}_k$ ) and statistical characteristics **a** coefficient of variation ( $CV_k$ ), **b** standard deviation ( $\sigma_k$ ) of soil moisture measurements at watershed-scale during each sampling day for the growing season, non-growing season, and the whole year

error between  $\pm 4\%$  and  $\pm 6\%$ . The ONL obtained in this study matches with the ONL found for different hydroclimatic regions and geomorphological conditions with varying spatial scales, ranging between 15 and 40 sampling locations [4, 5, 14]. Since the computation of the “average” error on the soil moisture temporal pattern is necessary [4]), a temporal pattern of spatial mean soil moisture is analyzed to assess ONL for a reliable estimate of watershed-mean soil moisture.



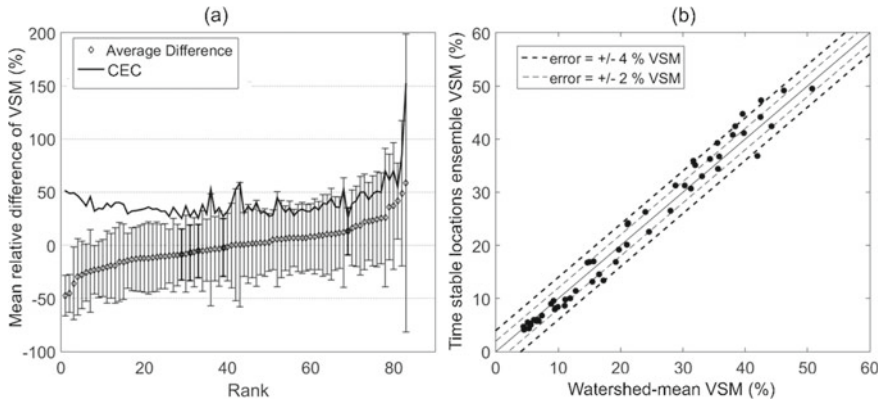
**Fig. 5** The optimal number of locations to capture the watershed-mean soil moisture at a 5% significance level during different seasons **a** with  $\pm 2\%$  absolute error and **b** at different absolute errors

## 4.2 ONL to Capture the Temporal Pattern of Watershed-Mean Soil Moisture

The most temporally stable locations that can serve as representative locations to capture the watershed-mean soil moisture were identified using time stability analysis based on the parametric test of relative differences (Eq. 7). Figure 6a shows the time stability characteristics, mean relative difference ( $\bar{\delta}_j$ ), ranked from smallest to largest along with  $\pm 1$  standard deviation. The sampling locations representing positive  $\bar{\delta}_j$  values constantly overestimate the watershed-mean soil moisture, whereas negative values of  $\bar{\delta}_j$  represents underestimation of the watershed-mean soil moisture consistently. Generally, the range of variation in  $\bar{\delta}_j$  increases with the investigated area size [5]. The CEC (Eq. 8) has the ability to effectively eliminate systematic bias as well as accurately capture the watershed-mean soil moisture at each sampling time. The selected most time-stable location based on the lowest CEC value was found to capture the watershed-mean soil moisture with a high correlation ( $R^2 = 0.852$ ) but with a high Root-Mean-Square error (RMSE) of 5.573%. Though only one representative location has the capability to capture the temporal variation pattern of the watershed-mean soil moisture with high correlation, it fails to provide a reasonable accuracy (RMSE of  $\pm 4\%$  or better). Notably, this contradicts the previous study's finding [5], where one representative site can capture the temporal pattern of areal mean soil moisture with reasonable accuracy at a regional scale of 200 km<sup>2</sup>. In further analysis, it was found that the five most time-stable locations of the study watershed can capture watershed-mean soil moisture with an excellent correlation ( $R^2 = 0.981$ ) and RMSE of  $\pm 2.17\%$ . The scatter plot in Fig. 6b shows ensemble soil moisture of the identified five best time-stable locations that can represent watershed-mean soil moisture with an error range of  $\pm 4\%$ . These analyzes imply that the five time-stable sampling locations could be utilized to capture the temporal pattern of watershed-mean soil moisture with reasonable accuracy and confirm the robust optimal sampling design for the watershed of 500 km<sup>2</sup>. The spatial distribution of the five most time-stable locations in the watershed along with other sampling locations (Fig. 7), shows that identified sampling locations are well distributed across the watershed. Remarkably it was also found that the identified time-stable locations have different elevations and soil texture conditions.

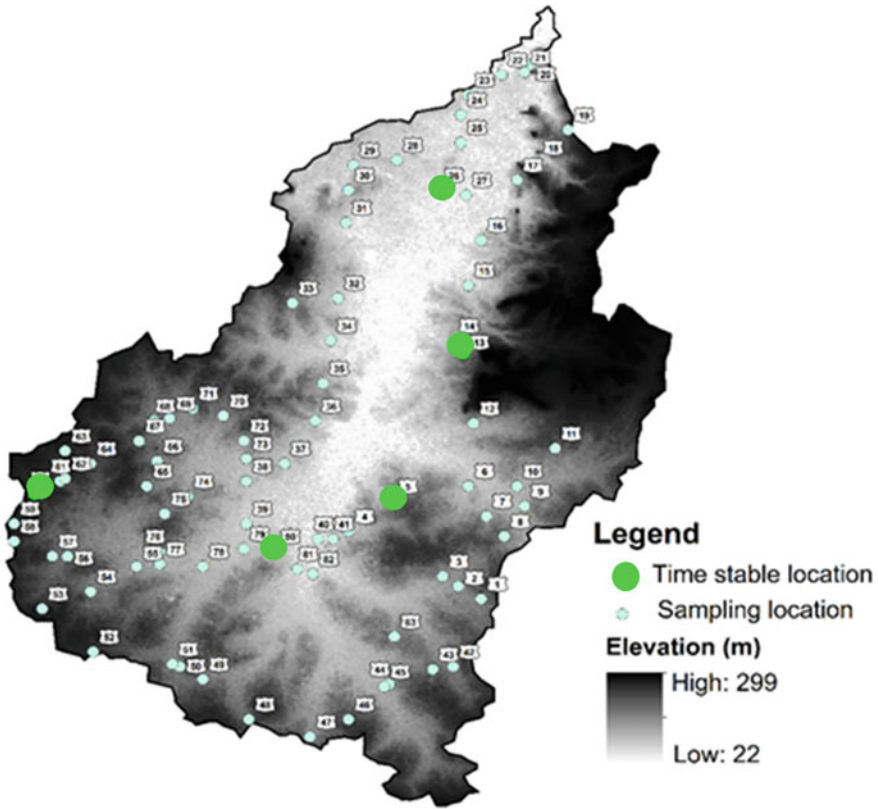
## 5 Conclusion

On the analyzes of 41,832 in-situ soil moisture samples collected during one year at 83 sampling locations in a watershed of 500 km<sup>2</sup>, it can be concluded that the dense in-situ measurements help to characterize the soil moisture spatiotemporal variability at the watershed-scale. The characterization of soil moisture spatiotemporal variability reveals that sampling at a few locations (~30 locations) is sufficient for capturing the spatial pattern of the soil moisture at a watershed of 500 km<sup>2</sup>. In



**Fig. 6** **a** Rank ordered mean relative difference of soil moisture at each sampling location with  $\pm 1$  standard deviation error bars. The solid line is a comprehensive evaluation criterion (CEC) to identify the best time-stable locations. The thick bars are the five best time-stable locations in the watershed based on the CEC. **b** Comparison of watershed-mean VSM with the ensemble VSM of the identified five best time-stable locations to achieve a reasonable accuracy (RMSE of  $\pm 4\%$  or better)

comparison, very fewer resources, and nearly five representative or time-stable locations are required to capture the temporal variation pattern of the watershed-mean soil moisture within an RMSE of  $\sim \pm 2\%$ . However, the optimal sampling design strategy must be transferred to ungauged regions to avoid the monitoring cost of dense measurements. The selection of representative locations a priori in ungauged areas to capture the temporal variation pattern of the areal mean soil moisture is possible with various geophysical characteristics information. A few past studies [17, 26], reveal that soil properties and topography are significant geophysical parameters that jointly control spatiotemporal persistence. However, more detailed investigations are needed towards transferring optimal sampling design strategy to the ungauged area by assessing the effects of heterogeneities of similar or different geophysical properties in other regions and for different space and time scales. In this study, a large soil moisture dataset at a watershed-scale has been generated through several intensive field campaigns for tropical regions where soil moisture information is sparse. The optimal sampling design based on long-term intensive sampling can be utilized as a guideline for designing a robust sensor-based network at the watershed-scale and could be helpful in planning sampling for satellite soil moisture product validation. The data used for optimal sampling design focuses only on one watershed in the tropical region of India. Therefore, further analysis is needed using in-situ soil moisture measurements of other watersheds in tropical climates and possibly other climate regions to assess the applicability of such an approach.



**Fig. 7** The elevation map of the watershed shows 83 soil moisture sampling locations. Locations marked with a circle are the five most representative or time-stable locations of the watershed

## References

1. Albergel C, Calvet JC, De Rosnay P, Balsamo G, Wagner W, Hasenauer S, Mahfouf JF (2010) Cross-evaluation of modelled and remotely sensed surface soil moisture with in situ data in Southwestern France. *Hydrol Earth Syst Sci* 14(11):2177–2191
2. Bolten JD, Crow WT, Jackson TJ, Zhan X, Reynolds CA (2010) Evaluating the utility of remotely sensed soil moisture retrievals for operational agricultural drought monitoring. *IEEE J Sel Top Appl Earth Obs Remote Sens* 3(1):57–66
3. Brocca L, Morbidelli R, Melone F, Moramarco T (2007) Soil moisture spatial variability in experimental areas of central Italy. *J Hydrol* 333(2–4):356–373
4. Brocca L, Melone F, Moramarco T, Morbidelli R (2010) Spatial-temporal variability of soil moisture and its estimation across scales. *Water Resour Res* 46(2):1–14
5. Brocca L, Tullo T, Melone F, Moramarco T, Morbidelli R (2012) Catchment scale soil moisture spatial-temporal variability. *J Hydrol* 422–423:63–75
6. Bogena HR, Herbst M, Huisman JA, Rosenbaum U, Weuthen A, Vereecken H (2010) Potential of wireless sensor networks for measuring soil water content variability. *Vadose Zone J* 9(4):1002–1013

7. Chen F, Crow WT, Starks PJ, Moriasi DN (2011) Improving hydrologic predictions of a catchment model via assimilation of surface soil moisture. *Adv Water Resour* 34(4):526–536
8. Choi M, Jacobs JM (2011) Spatial soil moisture scaling structure during soil moisture experiment 2005. *Hydrol Process* 25(6):926–932
9. Cosh MH, Jackson TJ, Moran S, Bindlish R (2008) Temporal persistence and stability of surface soil moisture in a semi-arid watershed. 112:304–313
10. Cosh MH, Stedinger JR, Brutsaert W (2004) Variability of surface soil moisture at the watershed scale. *Water Resour Res* 40(12):1–9
11. Cosh MH, Jackson TJ, Starks P, Heathman G (2006) Temporal stability of surface soil moisture in the Little Washita River watershed and its applications in satellite soil moisture product validation. 323:168–177
12. Entekhabi D (1995) Recent advances in land-atmosphere interaction research. *Rev Geophys* 33(S2):995–1003
13. Famiglietti J (1999) Ground-based investigation of soil moisture variability within remote sensing footprints during the Southern Great Plains 1997(SGP97) Hydrology Experiment. *Water Resour Res* 35(6):1839–1851
14. Famiglietti JS, Ryu D, Berg AA, Rodell M, Jackson TJ (2008) Field observations of soil moisture variability across scales. *Water Resour Res* 44(1):1–16
15. Hu W, Shao M, Han F, Reichardt K (2011) Spatio-temporal variability behaviour of land surface soil water content in shrub- and grassland. *Geoderma* 162(3–4):260–272
16. Hu W, Shao M, Han F, Reichardt K, Tan J (2010) Watershed scale temporal stability of soil water content. *Geoderma* 158(3–4):181–198
17. Jacobs JM, Mohanty BP, Hsu EC, Miller D (2004) SMEX02: Field scale variability, time stability and similarity of soil moisture. *Remote Sens Environ* 92(4):436–446
18. Joshi C, Mohanty BP, Jacobs JM, Ines AVM (2011) Spatiotemporal analyses of soil moisture from point to footprint scale in two different hydroclimatic regions. *Water Resour Res* 47 (1)
19. Koster RD, Mahanama SPP, Livneh B, Lettenmaier DP, Reichle RH (2010) Skill in streamflow forecasts derived from large-scale estimates of soil moisture and snow. *Nat Geosci* 3:613
20. Koster RD, Dirmeyer PA, Guo Z, Bonan G, Chan E, Cox P, Yamada T (2004) Regions of strong coupling between soil moisture and precipitation. *Science* 305(5687):1138–1140
21. De Lannoy GJM, Houser PR, Verhoest NEC, Pauwels VRN, Gish TJ (2007) Upscaling of point soil moisture measurements to field averages at the OPE3 test site. *J Hydrol* 343(1–2):1–11
22. Mohanty BP, Skaggs TH (2001) Spatio-temporal evolution and time-stable characteristics of soil moisture within remote sensing footprints with varying soil, slope, and vegetation. *Adv Water Resour* 24(9–10):1051–1067
23. Narasimhan B, Srinivasan R (2005) Development and evaluation of Soil Moisture Deficit Index (SMDI) and Evapotranspiration Deficit Index (ETDI) for agricultural drought monitoring. *Agric For Meteorol* 133(1):69–88
24. Singh G, Panda RK, Nair A (2020) Regional scale trend and variability of rainfall pattern over agro-climatic zones in the mid-Mahanadi River basin of eastern India. *J. Hydro-environ. Res.* 29:5–19
25. Singh G, Panda RK, Bisht DS (2021) Improved generalized calibration of an impedance probe for soil moisture measurement at regional scale using Bayesian neural network and soil physical properties. *J Hydrol Eng* 26(3):04020068
26. Singh G, Panda RK, Mohanty BP (2019) Spatiotemporal analysis of soil moisture and optimal sampling design for regional-scale soil moisture estimation in a tropical watershed of India. *Water Resource Research.* 55(3):2057–2078
27. Vachaud G, Passerat De Silans A, Balabanis P, Vauclin M (1985) Temporal stability of spatially measured soil water probability density function I. *Soil Sci Soc Am J* 49:822–828
28. Vinnikov KY, Robock A, Qiu S, Entin JK, Owe M, Choudhury BJ, Njoku EG (1999) Satellite remote sensing of soil moisture in Illinois, United States. *J Geophys Res: Atmos* 104(D4):4145–4168
29. Wang C, Zuo Q, Zhang R (2008) Estimating the necessary sampling size of surface soil moisture at different scales using a random combination method. *J Hydrol* 352(3–4):309–321

30. Western AW, Grayson RB, Blöschl G (2002) Scaling of soil moisture: A hydrologic perspective. *Annu Rev Earth Planet Sci* 30(1):149–180
31. de Wit AJW, van Diepen CA (2007) Crop model data assimilation with the Ensemble Kalman filter for improving regional crop yield forecasts. *Agric For Meteorol* 146(1–2):38–56