Selection and application of spatial variability distribution function of soil moisture in farmland

L Yao¹, Y R Wang¹, Y Wang², W Gao² and Z Zhang³

¹College of Hydraulic Engineering, Tianjin Agricultural University, Tianjin, China
²Tianjin Wuqing District Water Conservancy Technology Promotion Center, Tianjin, China

E-mail: wyrf@163.com

Abstract. In order to find a method for reasonably determining the number of soil moisture monitoring points (depth 0-20, 20-40, 40-60 cm at per points) in the farmland scale, in Beijinzhuang village and Xilv village of Cui Huangkou town, Wuqing district, Tianjin city. The monitoring area (286.67ha). Monitoring stations is divided into three layers from southeast to northwest. The soil moisture content of 61 points was measured in turn. The parameters were fit analysis based on the obtained data by using P-III type distribution function and normal distribution function. The minimum sample size of the measured sequence was analyzed. The analysis results show that the P-III type distribution function (correlation coefficient $R^2=0.9954$, average relative error is 6.3%) is better than the normal distribution function (correlation coefficient $R^2=0.9918$, average relative error is 6.6%), but the two are very close. Therefore, the number of reasonable soil moisture monitoring points were determined by the normal distribution function, the number of samples of the measured sequence, by which determining the reasonable soil moisture monitoring points in the farmland scale, should not be less than 30 points.

1. Introduction

Field soil moisture monitoring and forecasting is an important for understanding the drought situation of farmland and conducting reasonable irrigation. It is also a non-engineering measure for implementing high-efficiency water-saving irrigation. According to the spatial variability of soil characteristics and the accuracy requirements of soil moisture monitoring, the scale of soil moisture monitoring can be divided into measuring point scale, farmland scale, irrigation district scale, regional scale and watershed scale. As a basis for irrigation forecasting, it is often necessary to conduct farmland-scale soil moisture monitoring [1]. In the 1990s, with the development of precision agricultural technology in developed countries, the study of spatial variability of soil properties has attracted the attention of many scholars [2]. Spatial variability refers to conditions that are generally uniform under natural conditions (for example, research on soil moisture content refers to soil quality, groundwater depth, climate, planting species, irrigation systems, etc.), Inhomogeneity of soil parameters that actually exist due to various random factors [3]. He established a prediction model of soil moisture in cotton field under drip irrigation through the experimental study on the change of soil moisture in cotton field under drip irrigation and the location of soil moisture monitoring points [4]. Lei further improved the distribution method of soil moisture monitoring in the field, and believed that the uniform distribution method, once random distribution method and twice random distribution method were all feasible and could achieve good accuracy [5]. Yang et al analyzed the spatial and
temporal dynamics of soil moisture in the field based on field test data, and explored the number and method of reasonable sampling of farmland soil moisture changes [1].

Fernandez and Ceballos et al (2003) used a time domain reflectometer (TDR) to measure the soil moisture content of 23 network moisture monitoring stations on 1285 km$^2$ in the Douro River Basin in Spain at different depths for 36 consecutive months, measured every 2 weeks. Time stability analysis of soil water content based on these observations. According to the study, from the hydrological point of view, these sites are spatially consistent with the independence of natural geography standards and soil science standards. The worst time period of soil water content stability is in the process of increasing soil moisture, and also analyzes the location and quantity of reasonable moisture monitoring stations [6].

Based on the test data of 61 soil moisture monitoring points in 286.67 ha of farmland, this paper analyzes and compares the fitting precision of P-III distribution function and normal distribution function, and gives a reasonable number of soil moisture monitoring points.

2. Materials and methods

2.1. General situation of the test area
Beijinzhuang village and Xilv village, Cui Huangkou town, Wuqing district, Tianjin (117°1’E, 39°22’N, 8 meters above sea level), the monitoring area is 286.67 ha. The area is a temperate semi-humid continental monsoon climate with four distinct seasons. The annual average temperature is 11.6°C, the average temperature in January is -5.1°C, and the average temperature in July is 26.1°C. The annual average precipitation is 606 mm and the frost-free period is 212 days. The soil is loamy soil, and the soil is loose and fertile, which is suitable for agricultural production. The main food crops are wheat, corn. Winter wheat is watered two to three times during the growing period, the first irrigation in November, and the secondly irrigation in late March to mid-May next year. The agricultural production conditions are good. In recent years, advanced irrigation technologies such as pipe irrigation, sprinkler irrigation and drip irrigation have been used.

2.2. Soil moisture content test method and test point distribution.
According to the distribution of farmland in the soil moisture monitoring area, the soil moisture content of 61 points was measured from the southeast to the northwest (depth of 0-20, 20-40, 40-60 cm. Before the first watering), the distance between two adjacent test points is 50 m, and the whole line is 3 km. The test point distribution is shown in figure 1. The test method is the soil drilling and drying weighing method, thereby obtaining the layered soil moisture content of 61 points, and the average soil moisture content of the 0-60 cm soil layer can be obtained at the points.

![Figure 1. Test point distribution.](image_url)
2.3. Spatial distribution characteristics of farmland soil water content

2.3.1. Determination of distribution function and its parameters. In this paper, the P-III probability distribution function and the normal probability distribution function are selected to characterize the spatial distribution characteristics of soil moisture in the monitoring area.

The Pearson type III probability distribution curve is shown in equation (1).

\[ p = P\left( X \geq x_p \right) = \frac{\beta^\alpha}{\Gamma(\alpha)} \int_{x_p}^{\infty} \left( x-a_0 \right)^{\alpha-1} e^{-\beta(x-a_0)} \, dx \]  

Where \( \alpha, \beta, \) and \( a_0 \) are the three parameters of the Pearson type III distribution curve, where the \( \Gamma \) function is calculated using equation (2).

\[ \Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} \, dx \]  

\( \alpha, \beta, \) and \( a_0 \) have the following relationship with the average value \( \bar{x} \), the coefficient of variation \( C_v \), and the skew coefficient \( C_s \) of the measured sequence.

\[ \alpha = \frac{4}{C_s^2}, \quad \beta = \frac{2}{\bar{x}C_vC_s}, \quad a_0 = \bar{x}\left(1 - \frac{2C_v}{C_s}\right) \]  

Equation (3) can be used to determine three parameter, \( \alpha, \beta, \) and \( a_0 \). The \( \bar{x}, C_v \) and \( C_s \) in equation (3) can be estimated by using the sample series obtained by the test. There are various estimation methods, such as the moment method, the probability weight moment method, the weight function method, the three-point method, etc. In this paper, the weight function method is used to estimate the \( \bar{x}, C_v \) and \( C_s \).

The weight function method is using the first-order and second-order weight function moments to derive \( C_v \). The selection of the weight function should satisfy two conditions: Non-negative and

\[ \int_{a_0}^{\infty} \varphi(x) \, dx = 1 \]  

Where \( \varphi(x) \) is a weight function. The weight function is calculated according to equation (5).

\[ \varphi(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\bar{x})^2}{2\sigma^2}} \]  

In the formula, the \( \bar{x} \) of the measured series is calculated according to equation (6), and \( \sigma \) is the mean square error, and is calculated according to equation (7).

\[ \bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} = \frac{1}{n} \sum_{i=1}^{n} x_i \]  

\[ \sigma = \sqrt{\frac{1}{n} \sum (x_i - \bar{x})^2} \]  

Where \( x_1, x_2, \cdots x_n \) is the measured series of soil moisture, where \( n \) is the number of samples in the measured series.

The \( C_s \) is calculated according to equation (8).
\[ C_s = -4\sigma \frac{E}{G} = -4\bar{C}_v \frac{E}{G} \]  

(8)

Where \( E \) and \( G \) are calculated according to equations (9).

\[ E = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x}) \varphi(x_i) \quad G = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2 \varphi(x_i) \]  

(9)

First, the weight function method is used to calculate the \( \bar{x}, C_v \) and \( C_s \) of the measured sequence, and then the parameters \( \alpha, \beta \), and \( \alpha_0 \) are calculated by the equation (3), and the numerical integration is used to calculate \( \Gamma(a) \), and then the numerical simulation is performed according to the equation (1) to obtain the simulated frequency.

The normal distribution function is a widely used, as shown in equation (10).

\[ F(x) = \frac{1}{\sigma \sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{(x-\bar{x})^2}{2\sigma^2}} dx \quad (+\infty < x < -\infty) \]  

(10)

The meaning of symbols in the formula is the same as before.

Then the frequency is simulated by numerical integration method and compared with the measured value of the frequency.

2.3.2. Determine the number of monitoring reasonable points. The number of reasonable samples \( N \) must be conformed to the,

\[ P = \left\{ \left| \bar{x}_N - \mu \right| \leq \Delta \right\} = P_1 \]  

(11)

In the formula, \( \bar{x}_N \) it is the sample average value, \( \mu \) is the population average value, and \( P_1 \) is the confidence level. In this study, \( P_1 = 95\% \).

3. Results

3.1. Parameter fitting analysis of soil water content distribution function

According to the measured sequence of soil moisture content, the frequency can be obtained \( (P = m/(n+1)\times100\%) \), \( m \) is the serial number of the measured sequence, \( n \) is the number of samples. In order to compare with the simulated frequency, the frequency is called the measured frequency in this paper. The results are shown in figures 2 and 3. Statistical parameters, the results of the \( \bar{x}, C_v \) and \( C_s \) are shown in table 1. According to formula (3), three parameters of the P-III distribution function can be obtained, as shown in table 1. Calculating \( \Gamma(a) \) according to the formula (2), where “\( +\infty \)” is taken as 500, dividing the integration interval into corresponding parts, and \( \Gamma(a)=1.8971\times10^{34} \) is obtained.

The simulated frequency is calculated based on numerical integration according to equations (1) and (3). During the simulation, the measured sequences are arranged from small to large, starting from the wilting point moisture, as the lower limit of the integral, the minimum water content is the upper limit of the integral in the measured series, and the integral interval is equally divided into 100, and the frequency corresponding to the minimum water content is obtained; the frequency difference between two adjacent water contents, the frequency corresponding to the previous moisture content can be accumulated to obtain the frequency corresponding to the upper limit moisture content of the integral; the maximum water content in the measured sequence is taken as the lower limit of final integral, and the saturated water content is used instead of “\( +\infty \)” as the upper limit of the integral, and the frequency
corresponding to $\infty$ can be obtained. P-III distribution curve and a normal distribution curve are given in figures 2 and 3. On the basis of the obtained simulated frequency, the analysis gives the relative errors between simulated frequency and measured frequency, as shown in Table 1.

![Figure 2. P-III distribution curve.](image)

![Figure 3. Normal distribution curve.](image)

| Distribution function parameter | Normal distribution parameter | P-III distribution function parameter |
|---------------------------------|------------------------------|--------------------------------------|
| | soil moisture monitoring points | soil moisture monitoring points |
| $x$ | 0.22 | 0.22 | 0.23 | $\alpha$ | 32.24 | 72.70 | 31.27 |
| $\sigma$ | 0.02 | 0.02 | 0.03 | $\beta$ | 250.08 | 393.84 | 375.61 |
| $C_v$ | 0.10 | 0.10 | 0.13 | $\sigma_{al}$ | 0.09 | 0.03 | 0.14 |
| $C_s$ | 0.22 | 1.95 | 0.47 | $R^2$ | 0.9954 | 0.9872 | 0.8109 |
| $R$ | 0.9918 | 0.9834 | 0.9754 | $e_{ave}$ | 64 | 92 | 38.1 |
| $e_{min}$ | 6.6 | 9.1 | 14.4 | $e_{max}$ | 624 | 360 | 1000 |
| $e_{max}$ | 25.1 | 71.3 | 61.4 | $e_{min}$ | 0.1 | 0.4 | 1.0 |
| $e_{min}$ | 0.0 | 0.1 | 0.3 |

It can be seen from figures 2 and 3 that the cumulative probability curves of the P-III type distribution function and the normal distribution function are both S-shaped, and the simulated frequency complex correlation coefficient ($R^2$) of the P-III type distribution function is 0.9954. The complex correlation coefficient ($R^2$) of the state distribution function is 0.9918, both of which are
higher than 0.99, and the degree of fitting is good; As seen from table 1, the average relative error of the simulated frequencies is very close, and the maximum relative error of the normal distribution function is significantly smaller than the relative error of the P-III type distribution function, indicating that the spatial distribution of soil moisture content in the field can be approximated as a normal distribution. This is consistent with the related research by Lei et al [5].

3.2. Determining the reasonable number of monitoring points
The Reasonable number of monitoring Points are calculated under the different confidence levels and the accuracy of the estimation according to equations (12), (13) and (14). In order to analyze and determine the number of samples of farmland soil moisture content measured during the process of reasonable soil moisture monitoring points, 33 points and 16 points were taken from 61 test points, the distances are 100 m and 150 m respectively, according to the same method. the reasonable number of monitoring points is shown in table 2.

Table 2. the reasonable number of monitoring points(N) under different confidence levels and estimation accuracy.

| Confidence level | k  | N under different measured sequence lengths |
|------------------|----|--------------------------------------------|
|                  | 5% | 61  | 33  | 16  |
| 95%              | 5% | 16  | 16  | 25  |
|                  | 10%| 4   | 4   | 6   |
|                  | 15%| 2   | 2   | 3   |
| 90%              | 5% | 11  | 11  | 25  |
|                  | 10%| 3   | 3   | 6   |
|                  | 15%| 1   | 1   | 3   |

The sampling is independent, and the number of samples is sufficient. According to the central limit theorem, the reasonable number of samples can be obtained by the following formula.

$$P = \left[ \frac{x_N - \mu}{\sigma/\sqrt{N}} \right] \leq 1.960 \right] = 95\% \tag{12}$$

The number of monitoring points available from equations (13) when the confidence level is 95%.

$$N = 1.960^2 \left( \frac{\sigma}{\Delta} \right)^2 = 3.84 \left( \frac{\sigma}{\Delta} \right)^2 \quad (P = 95\%) \tag{13}$$

If the accuracy requirement $\Delta = k\mu$ (k is 5%, 10%, 15%), then equation (14) can be written as:

$$N = 3.84 \left( \frac{\sigma}{k} \right)^2 \quad (P = 95\%) \tag{14}$$

The required confidence level is 95%, the relative error is less than 5%, when the actual measured points are 61, the reasonable points are 16, as can be seen from table 2; when the actual measured points are 33, the reasonable points are 15; when the actual measured points are 16, the reasonable number of points is 25, which indicates that the shorter the measured sequence, the more points needed for the reasonable arrangement of monitoring soil moisture. When the measured sequence is too short, the reasonable number of monitoring points will be too large, even larger than the measured sequence. The required confidence level is 95%, when measured sequence is same, the lower the accuracy of the estimation, the less of the reasonable number of monitoring points. When the
corresponding estimation accuracy is 5%, 10%, 15%, the number of reasonable monitoring points are 16, 4 and 2 respectively; when the accuracy is the same, the higher the confidence level, the more monitoring points need to be arranged. When determining a reasonable number, the number of samples of the measured sequence should be no less than 30 points.

3.3. The impact of the depth of soil moisture monitoring on the number of reasonable points.
Based on the average of three layers of 0-20 cm, 0-40 cm and 0-60 cm moisture content, reasonable number of samples with different sampling depths under different measured sequence lengths in table 3.

Table 3. Reasonable number of samples under different confidence levels and estimation accuracy N.

| Confidence level | k   | Reasonable number of samples with different sampling depths under different measured sequence lengths |
|------------------|-----|--------------------------------------------------------------------------------------------------|
|                  |     | 0-20cm | 0-40cm | 0-60cm |
| 95%              | 5%  | 39     | 26     | 19     | 16     | 16     |
|                  | 10% | 10     | 7      | 5      | 4      | 4      |
|                  | 15% | 4      | 3      | 2      | 2      | 2      |
| 90%              | 5%  | 28     | 19     | 13     | 12     | 11     |
|                  | 10% | 7      | 5      | 3      | 3      | 3      |
|                  | 15% | 3      | 2      | 1      | 1      | 1      |

Table 3 gives the reasonable monitoring numbers for different sampling depths under the conditions of 95%, 90% confidence level and different estimation accuracy (5%, 10%, 15%). The analysis shows that the confidence level is 95% and the estimation accuracy is 10%, When the sampling depth of the measured sequence is 20 cm, the number of reasonable points with a measured sequence length of 61 is 10, and the reasonable points with a measured sequence length of 33 is 7. The length of different measured sequences would affect the determination of reasonable points, and the shorter the sampled measured sequence, the fewer reasonable points. When the confidence level is 95% and the estimation accuracy is 10%, the measured sequence length is 61, when the sampling depth is 0-20 cm, the reasonable number of points determined is 10, and when the sampling depth is 0-40 cm, the reasonable number of points is 5, indicating that the shallower the sampling depth, the more reasonable points; from table 3, it can be seen that when the sampling depth is 0-40 cm and 0-60 cm, the length of the measured sequence does not affect of reasonable number of points, and the two are nearly identical.

4. Conclusion
Based on field measured data, parametric of P-III distribution function and normal distribution function were fitted. The distribution of soil water content is approximately as the normal distribution. Based on the normal distribution, the reasonable points needed to be arranged for soil moisture monitoring are determined. The results as follows. (1) The field measurement series shows that the spatial distribution of water content can be approximated as a normal distribution, which is consistent with the results of Lei [5]. (2) When determining the reasonable number of points, the number of samples of the measured sequence should be not less than 30. (3) The sampling depth of the measured sequence should not be too shallow, and the suitable depth is 0-60 cm in this paper.

Acknowledgments
This work was supported by Water saving and yield increasing mechanism based on uncertainty and limited water supply irrigation ,Integration and demonstration of precision irrigation and fertilization technology under limited water supply conditions, Tianjin Municipal Water Saving Irrigation
Technology and Equipment School and Enterprise Collaborative Innovation Laboratory Construction (Grant No. 51779174, 201701150, 17PTSYJC00110).

References
[1] Yang S X, Lei Z D, Wu W R and Duan X J 1996 Soil moisture monitoring on farmland scale *Progress in Water Sci* 1 14-9
[2] Zhu H X, Chen X M, Fang K and Du Z J 2009 Spatial variation of soil organic matter on farmland scale *Acta Agriculture Jiangsu* 25 1029-32 (in Chinese with English abstract)
[3] Yang S X and Lei Z D 1993 Spatial structure of soil moisture content in the field and determination of the number of samples *Acta Geologica Sinica* 5 447-56
[4] He S 2015 Prediction of soil moisture in cotton fields under drip irrigation (Shihezi University)
[5] Lei Z D, Yang S X and Luo Y 1996 Research on the distribution method of soil moisture monitoring in the field *Irrigation and Drainage* 3 9-15
[6] Zhu X G, Lin F C and Gao H B 2011 Research progress on moisture monitoring *Water-saving Irrigation* 11 53-55+58