SPATIAL VARIABILITY OF POTATO (*SOLANUM TUBEROSUM* L.) YIELD AND QUALITY ON SAND SOIL IN THE ARID REGION OF NORTHWEST CHINA

ZHANG, R. X. – ZHANG, J. F. – LI, T.* – XIONG, S. Y.

State Key Laboratory of Eco-hydraulics in Northwest Arid Region, Xi'an University of Technology, Xi'an 710048, China
(phone: +86-029-8231-2768)

**Corresponding author*
e-mail: litao@xaut.edu.cn

(Received 3rd Dec 2021; accepted 25th Feb 2022)

Abstract. The potential for significant variability in potato (*Solanum tuberosum* L.) yield and quality caused by soil properties can be employed as a foundation for site-specific soil management. The current work investigated the relationship of soil properties, including the texture, moisture, and nitrate-nitrogen, with the variability of potato yield and quality on 28 ha center-pivot irrigated fields located in northwest China, where the soil was improved in 2014 by mixing Aeolian sandy soil with feldspathic sandstone. Geostatistic and correlation analyses were utilized to extract the relationships among soil properties and potato tuber properties. K-means clustering methods were utilized to establish critical sample points and soil factors for spatial variability of potato yield and quality. The spatial dependence of the potato tuber yield and quality varied from weak to strong, and most had unstable spatial structure considering the soil temporal variability, except reducing sugar content. More than 66% of the correlation coefficients between soil sand content and tuber variables were significant. The evaluation of the K-means clustering algorithms' effectiveness showed that θv- All, PC-All, and texture had superior prediction results for tuber yield, dry matter content, and starch, respectively. N-All and soil texture datasets exhibited the best performance for reducing sugar content. Unsatisfactory prediction results were obtained for protein and VC. The results can be utilized to realize different objectives for in-field site-specific potato management.
Keywords: *soil texture, potato quality, sand-fine mixture, geostatistics, K-means clustering*

Introduction

Potatoes (*Solanum tuberosum* L.) are considered to be the fourth most abundant food crop in China, after rice, maize, and wheat. A total of 5,815,140 ha of farmland, mostly located in the arid northwestern region of China, were used in 2016 to cultivate potatoes with an average yield of 17.04 t ha^{-1}. Potato yield and quality can change considerably within a particular field. The spatial variation of yield and quality leads to the waste of soil and water resources. Evaluating the degree and driving factors of spatial variability within-field is an important consideration to achieve site-specific management.

Potato yield temporal and spatial variability has been widely verified in the literature. Soil moisture content has been proven to be similar to the spatial distribution of crop yield or quality in many studies (Warrick et al., 1983; Wesenbeeck et al., 1988; Rockström et al., 1999; Irmak et al., 2002). Topography, soil texture, soil depth, soil organic matter, electricity conductivity and nutrients were also reported to be related to the spatial variation of potato yield (Redulla et al., 2002; Starr, 2005; Cambouris et al., 2006; Po et al., 2010; Perron et al., 2018).

According to previous reports, the coefficient of variation (CV) of potato yield ranged from 0.24 to 0.32 and the variability of potato yield was classified as moderate (Cambouris et al., 2006; Taylor et al., 2018). Geostatistical analysis was widely used in

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2003_19691989

the spatial variation of potato yield. Cambouris et al. (2006) adopted that the model of the semi-variogram of potato yield was always exponential, and the proportion of the structured variance compared to all the variance was moderate and almost constant. It showed that the variation in yield attributable to spatial variability in physico-chemical properties within the field was higher than that induced by seasonal climatic variability. Similar results has reported in previous study (Perron et al., 2018). The spherical model was also used to fit the semivariograms of yield (Taylor et al., 2018). Rosenzweig et al. (2016) demonstrated that the stable model was more suitable for potato yield in the interpolation process and the proportion of the structured variance compared to all the variance was 100%. Potatoes are usually processed into French fries, starch, whole powder. Starch, reducing sugar, protein, VC and dry matter content are very important for potato industry. Although understanding the spatial variability of potato quality is important, it is rarely mentioned.

Identifying and understanding the spatial variation of potato yield and quality is currently limited by the time and resources required to do sufficient soil monitoring. To reduce the number of sampling points, root mean square error can be used to determine the number of sampling points; rational number of sampling sites was determined by geostatistical analysis (de Souza et al., 2014; Wang et al., 2015; Li et al., 2020). Many studies estimated the mean value of soil variables by looking for stable or representative points. Temporal stability analysis (TSA) has already been applied widely to reduce the number of samples needed for estimating soil water storage in a field (Vachaud et al., 1985). However, the TSA method is based solely on empirical data, the ability to recognize why certain locations are better to sample than others is limited to the sampling points used to find the rank stable locations. To solve this problem, Van Arkel et al. (2015) adopted the K-means clustering algorithm to detect critical sampling points for estimating field-scale near-surface soil moisture and compared with the TSA method. The results showed that the clustering approach on soil and topography data resulted in field-scale average moisture estimates that were as good or better than TSA, but without the need for exhaustive presampling of soil moisture. Because of the effectiveness of this method in soil moisture estimation, it is necessary to use the principle of this method to estimate the mean value of potato yield and quality.

Based on this context, the main goals of the current work are: (i) to derive the spatial variability of potato yield and quality variables (dry matter, protein, starch, reducing sugar, vitamin C) through descriptive statistics and geostatistic analysis; (ii) to evaluate the correlation between soil attributes and potato yield and quality; (iii) to verify the effectiveness of the method which finding critical sampling points using K-means clustering algorithms in the estimation of the mean value of potato yield and quality.

Materials and methods

Experimental site

The current work was performed at the commercial crop fields, located in the border of Mu Us Sandy Land, northwest China (38°09′N, 109°00′E, 1183 m a.s.l.). The experiment was accomplished on a 28.3-ha field, cropped with potato. Before soil melioration treatment, soil texture was sandy and was not suitable for planting. In order to control desertification and increase cultivated land, Chinese government had implemented a project to mix sandy soil (Aeolian sandy soil) with another local soil

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2003_19691989

(feldspathic sandstone). *Figure 1* shows the location of the experimental site and photos before and after soil mixing.

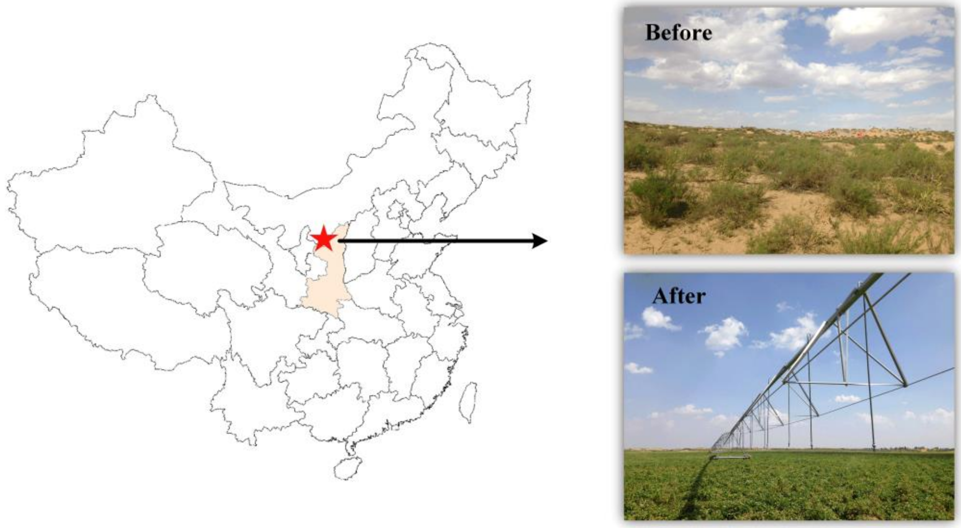

Figure 1. *Location of the experimental site in China and photos before and after soil mixing*

The physical and chemical properties of mixed soil were studied (Han et al., 2012; Zhang et al., 2021). The study fields were well mixed from a depth of 0–30 cm by Aeolian sandy soil and feldspathic sandstone in a proportion of 5 to 1 in 2014. The original aeolian sandy soil was covered with feldspathic sandstone with the mass of 850 tons per hectare. The mechanical composition of Aeolian sandy soil and feldspathic sandstone is shown in *Table 1*.

Table 1. *Composition of Aeolian sandy soil and feldspathic sandstone (%)*

Sample	Clay (<0.002mm)	Silt (0.002-0.05mm)	Sand (0.05-2mm)	Texture
Aeolian sandy soil	0.24	4.45	95.31	Sand
Feldspathic sandstone	7.06	58.09	34.85	Silt loam

After soil improvement, potato was planted within 2015, 2016 and 2017 growing seasons. *Table 2* showed soil properties, meteorological conditions and planting management in the potato field from 2015 to 2017. In addition, the variation of initial aeolian sandy soil was also measured in 2014. The sand content of the sandy soil was more than 90% and clay content was as low as 0.24%. The coefficients of variation of soil clay, silt and sand of the sandy soil were 0.53, 0.38 and 0.05 respectively.

Meteorological conditions

The region where the experiment was installed generally has a typical semi-arid climate. In this area, the potato typically grows from May to October. *Figure 2* compares the monthly mean temperature, the rainfall, and the irrigation for all potato cropping seasons with their corresponding long-time average values (1957–2014). The rainfalls during May to October in 2016 and 2017 were 772 mm, and 634 mm, respectively, which

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2003_19691989

were higher than that in 2015 (379 mm). Besides, these values were also higher than the long-term 57-year rainfall mean (423 mm). The potato growth season's average temperatures were a little greater than the long-term mean from 1957 to 2014. Moreover, the average temperature (℃) was higher in 2017 by 1.1℃, 2.3℃, and 1.6℃, than that of 2015, 2016, and 57-year period, especially within the critical stage of the onset of flowering and Tuber formation (July-August).

Table 2. *Soil properties, meteorological conditions and planting management in potato field*

Category	Property	2014	2015	2016	2017
Soil properties	Soil texture	sandy	sandy loam	sandy loam	sandy
	Clay (＜0.002mm) (soil depth:0-40cm) (%)	0.24	3.09	4.30	0.60
	Silt (0.002-0.05mm) (soil depth:0-40cm) (%)	4.45	32.01	27.24	9.40
	Sand (0.05-2mm) (soil depth:0-40cm) (%)	95.31	64.90	68.45	90.00
	Bulk dendity	\	1.42	\	1.50
	Electricity conductivity(μs/cm)	\	103.45	\	594.54
	Soil organic matter（g/kg）	0.08	1.04	1.82	1.85
	Soil nitrate-nitrogen （mg/kg）	0.012	26.69	30	27.68
	Soil available phosphorus（mg/kg）	2.68	3.15	3.25	4.28
	Soil available potassium（mg/kg）	84	73.07	89	113
Plant management	Plant area(ha)	0	28.3	28.3	28.3
	Cultivar	\	Shepody	Favorita	Shepody
	Growing season	\	5/1-9/25	5/1-9/1	5/1-9/22
	Intra-row spacing between tuber seed pieces (cm)	\	18	18	18
	Width of the ridge (cm)	\	30	30	30
	Inter-row spacing (cm)	\	90	90	90
	Irrigation amount (mm)	\	418	429	473
Meteorological conditions	Rainfall (in growing stage)	\	379	772	634
	Average temperature (℃)	\	20.0	19.9	20.8

\: no measurement was made

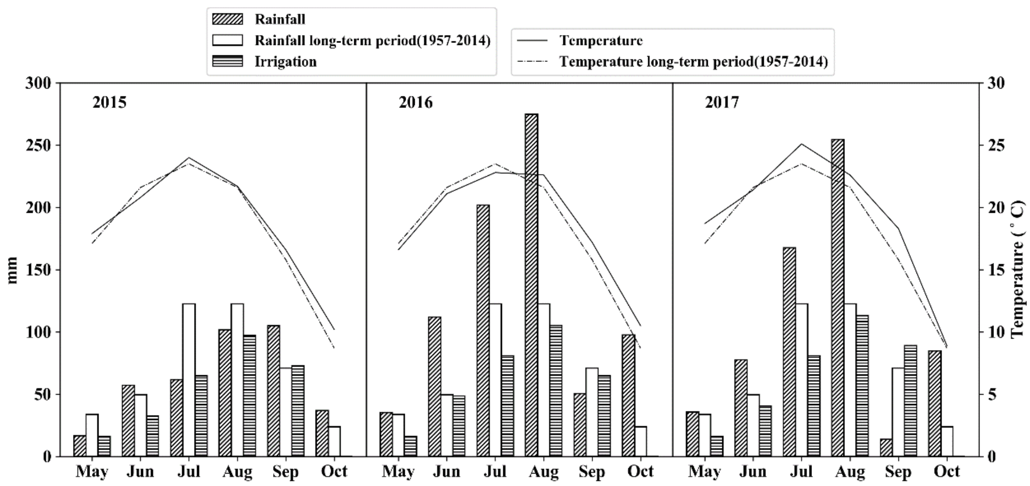

Figure 2. *The comparison between average monthly temperature, rainfall, and irrigation at the work site for all growing seasons and their corresponding long-term values (1957-2014)*

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2003_19691989

Plant management

The potato cultivar ‘Shepody’, a widely grown cultivar processed for chip products globally, was planted in 2015 and 2017, while another potato cultivar ‘Favorita’, was planted in 2016. This field was a potato monoculture system with ridge tillage. The potato was sowed in early May and grown continuously under a single pass of a valley center pivot irrigation system (E2060-G, Reinke Manufacturing Company Inc. USA) equipped with Nelson sprinkler nozzles (D3000, Nelson Irrigation Corporation, USA). The irrigation amount was 11 mm for each irrigation event, and the total irrigation amount was 418 mm, 429 mm and 473 mm during the entire growing period in 2015, 2016 and 2017, respectively. The detailed irrigation schedules are shown in *Table 3*.

Table 3. *Irrigation schedules from 2015 to 2017*

Growth Stage	Irrigation quota （mm）	2015		2016		2017	
		Irrigation events	Irrigation amount (mm)	Irrigation events	Irrigation amount (mm)	Irrigation events	Irrigation amount (mm)
Germination	11	2	22	2	22	3	33
Seeding	11	5	55	6	66	6	66
Tuber formation	11	10	110	10	110	10	110
Tuber development	11	11	121	13	143	14	154
Tuber ripe	11	10	110	8	88	10	110
Whole	11	0	418	39	429	43	473

It is necessary to note that in 2015, there was no local power supply, and the sprinkler irrigation machine was powered by diesel. In order to save the cost, the irrigation frequency was lower than that in the next two years. Fertilizer application rates were adjusted using field mean nutrient amounts followed by present local agricultural authority guidelines. For each year, before planting, 4500 kg organic fertilizer, 15 kg urea, 25-30 kg diammonium phosphate and 20-25 kg potassium sulfate were applied, at the budding stage, 20 kg special fertilizer for potato were applied by center pivot irrigation system. The same plant protection traits were applied from 2015 to 2017. To avoid pests and diseases, 3 kg of 5% phoxim particles per hectare were applied before planting in each year. In order to prevent and control 28 lady beetles, 2.5% Deltamethrin and 70% Cyano (3-phenoxyphenyl) methyl 4-chloro-α-(1-methylethyl) benzeneacetate were used. Cyhalothrin and confidor were used to control aphids. In order to prevent and control the late blight, it has been prevented and treated every 7-10 days since July 15, with a total of 6 times. The drugs were Antracol, 10% Cyanostazole SC, Amesida and DuPont Yibao, Famoxate, Redomir(Gold,MZ) and Fluazinam.

Sampling points

The samplings of soil and potato were done in a grid. *Figure 3* shows sample location and the kriging map of soil sand content in 2015. One hundred and sixteen georeferenced regions were chosen from the field. The breakdown distance among the sampled regions was typically 25 m. Financial limitations and estimation precision requirements determine the whole number of samples. Soil moisture, nitrate-nitrogen (NO_3–N), and

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2003_19691989

texture components were used to find critical sampling points in this study. Soil moisture and nitrogen were often used to delineate management zones because of the correlation with yield (Vrindts et al., 2005; Perron et al., 2018). The effect of soil texture on the variation of potato cannot be ignored due to its dramatic changes in time and space in the region where the experiment was installed.

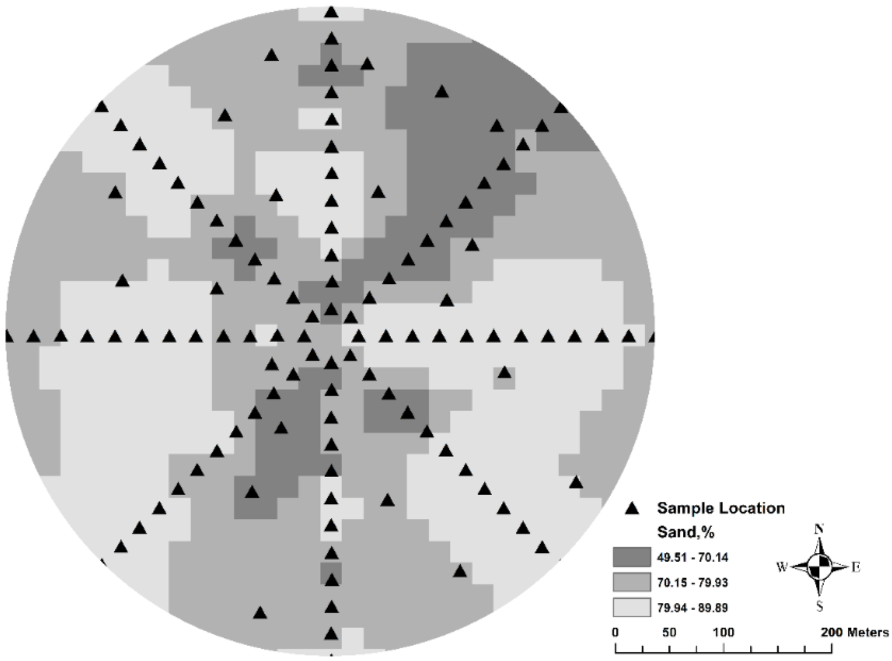

Figure 3. *Kriging map of soil sand content after mixed in 2015 and sample location*

Soil texture

At each points, soil samples were collected for soil texture measuring before planting in each growing season (2015, 2016, 2017). In order to reduce the measurement error caused by testers, testers A and B were arranged to collect soil samples independently. Three soil core samples (70 mm diameter, 52 mm height) were collected at the depth of 20 cm within a 1-m radius of each grid point by teater A and tester B, respectively. Samples were kept cool until submitted to the appropriate laboratory and air-dried, ground, and sieved through a 2 mm sieve. Soil mechanical composition of three soil core samples in each point were performed by teater A and tester B respectively using laser particle analyzer (MasterSizer 2000, Malvern Panalytical Ltd, USA). It is necessary to evaluate the reliability of measurements using the intraclass correlation coefficient (ICC). The ICC can theoretically vary between 0 and 1.0, where an ICC of 0 indicates no reliability, whereas an ICC of 1.0 indicates perfect reliability (Weir, 2005).

Table 4 showed the mean value of intraclass correlation coefficient among three years between different measurements of soil texture components. It can be seen that ICC of different measurements were higher than 0.85 indicating good reliability. The average value in each year of measurements performed by teater A and tester B were calculated and shown in *Table 4*. The USDA/FAO textural classifications systems were used to evaluate different textural classes.

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2003_19691989

Table 4. *The mean value of intraclass correlation coefficient(ICC) among three years between different measurements of soil texture components*

Soil texture components	**Intraclass correlation coefficient(ICC)**		
	M_{A3}	**M_{B3}**	**A*B**
Clay	0.987	0.991	0.893
Silt	0.989	0.987	0.856
Sand	0.978	0.989	0.879

M_{A3}: Three measurements collected by tester A; M_{B3}: Three measurements collected by tester B; A*B:The mean value of three measurements collected by tester A and the mean value of three measurements collected by tester B

Soil nitrate-nitrogen(NO_3–N)

From each sampling location, soil composite sample was collected 3 times in June, July and August in each growing season (2015, 2016, 2017) for soil NO_3–N measuring, which was composited by three soil core samples (70 mm diameter, 52 mm height) collected at the depth of 20 cm within a 1-m radius of each grid point. NO_3–N was determined according to Cawse (1967). N_1, N_2, N_3 represented NO_3–N in June, July and August in each year.

Soil moisture

The volumetric soil water content was measured at each sampling location at depths of 0–40 cm every 4 days during the growing season using a time domain reflectometer sensor (Xi'an Bi Shui RV1, China), and the measured values of the volumetric soil water content were calibrated by a gravimetric method during the experimental period of each year. Soil moisture content was measured 35, 32 and 35 times in 2015, 2016 and 2017 respectively. Θ_{v1}, Θ_{v2}, Θ_{v3} represented the mean value of soil moisture content in June, July and August in each year.

Soil properties

Three soil core samples (70 mm diameter, 52 mm height) were collected at three random points at the depth of 20cm within a 1-m radius before planting in each growing season. One of samples was used for measuring soil bulk density determined according to cutting ring method. One of samples was used for measuring soil organic matter, available phosphorus, available potassium. Soil available potassium was extracted with ammonium acetate and determined by flame photometer (SHUMAN et al., 1990). Soil available phosphorus was extracted with sodium bicarbonate solution and determined by spectrophotometer (V-T3, Yipu Instrument Manufacturing Co., LTD, Shanghai, China) (Olsen, 1954). Soil organic matter was determined according to Yeomans et al. (1988). Another sample was used to measure soil electricity conductivity determined by conductivity meter (DDSJ-318T, Shanghai Yi Electrical Scientific Instrument Co., LTD, Shanghai, China). The average value of each year were shown in *Table 1*.

Yield and quality

Before the field's commercial harvest, a typical yield digs of 1 m-rows was carried out by hand at all sample points and weighed by an electronic scale (measuring range: 0.05-50 kg, measuring accuracy: 0.01 kg, Zhuoshangqi Co. Ltd., China). Three fresh

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2003_19691989

potatoes were collected in each points and submitted to the laboratory. One fresh potato in each point was used for VC measurement and the rest were used for other quality measurement. Tuber dry matter content was determined according to oven drying method, slicing the three potatoes and mixing the processed potatoes together. Fragmented tubers were dried at 105 °C until the weight became constant. The starch content, protein, reducing sugar, vitamin C in tubers were determined according to iodine colorimetry method (McGrance et al., 1998), biuret reagent method (Yan et al., 2006), 3,5-dinitrosalicylic acid (DNS) colorimetric method (Zhao et al., 2008), moly-blue colorimetric method (Rietjens et al., 2002), respectively.

Data analyses

Statistical analysis, correlation analysis and geostatistical analysis were used to evaluate the spatial variability of potato yield and quality. ArcGIS10.3 (ESRI, Redlands, CA, USA) was adopted to construct semivariograms and kriged surface maps. The linear correlation coefficients ($P < 0.05$) were obtained through the Pearson's test by SPSS10.0 (International Business Machines Corporation, USA). The normal distribution estimation was performed on a skewness basis, i.e., for a skewness range between -1 and 1, a normal distribution was chosen for the data. If the data does not accord with normal distribution, the original data should to be log-transformed.

Semivariogram parameters, including the nugget (C_0), sill (C_1), and range (a), were utilized to define the spatial framework of all variables. Nugget describes the distance zero variance or the experimental error; sill defines the semivariance amount where the semivariogram attains the upper bound after its primary growth. It represents the maximum variance for this type of semivariogram and indicates the overall (a priori) semivariance of the selected region; the range defines the amount(x-axis) where a variable becomes spatially independent or the lag-distance where the semivariogram becomes smooth. The nugget to sill ratio determines the random part significance and estimates the spatial dependence quantitatively. Nugget/sill ratios can be divided into three categories (Cambardella et al., 1994): (i) <25%, strong spatial dependence; (ii) 25–75%, moderate spatial dependence; (iii) >75% ,spatially independent or pure nugget (i.e., when semivariograms' slopes are around zero).

Spatial variation can be described by various models (spherical, circular, etc.) fitting the semivariograms. The best-fitting model can be selected using the highest determination coefficient and confirmed through a visual inspection. The adopted lag distance was between 5 and 12, considering the variable. Cross-validation and ordinary kriging were utilized to extrapolate the amounts of unsampled field components.

K-means cluster analysis

K-means cluster algorithm aims to find categories, or clusters, containing objects with similar features. The distance measures over the different dimensions in the dataset are utilized to extract the similarity among objects. In the initial step of this algorithm, all data locations are randomly allocated to one of the k clusters. The n-dimensional centroid location, where n indicates the number of attributes in the input vector for any point, is obtained for all k clusters. The minimum Euclidean distance between the input vector and the centroid vector is derived to obtain the distance from any point to every centroid. Then, all points are reassigned to a cluster with the nearest centroid. This procedure continues iteratively until cluster membership is kept unchanged (Van Arkel et al., 2015).

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2003_19691989

In order to reduce the cost of sampling, it is necessary to find optimal sampling locations with fewer soil variable types and fewer measurement times. Therefore, the data involved in clustering were divided into 8 groups. Firstly, the single measurement result of a single factor was clustered in order to analyze whether good results can be obtained at the least cost, including sand dataset, N-1 dataset, Θ_v-1dataset. Sand dataset were composed of soil sand contents in 116 sample points representing soil texture. N-1 dataset and Θ_v-1 dataset were composed of nitrogen and mean value of soil water contents in all points in August, respectively. According to the research of Tian et al. (2011), the water requirement of potato was higher from flowering stage to tuber expansion stage, and lower at seedling stage and before harvest. Secondly, multiple measurement results of single factor were clustered, including texture dataset, N-All dataset, Θ_v-All dataset. The texture dataset contains clay, silt, sand in all sampling points. N-All dataset was composed of three measurements of nitrogen, N1, N2, N3. Θv- All dataset was composed of Θv1, Θv2, Θv3, the mean value of soil moisture content in June, July and August. In addition, the principal components extracted by multi-factor were clustered. The first principal component and all principal components were clustered respectively. PC-1, as the first principal component extracted from these attributes, contains clay, silt, sand, N_1, N_2, N_3, Θ_{v1}, Θ_{v2}, Θ_{v3}. PC-All includes principal components whose eigenvalues were extracted from these attributes and are higher than 1. Eigenvectors and cumulative contribution rates of principal components in each year were shown in *Table 5*. There were 4, 3 and 4 principal components in 2015, 2016 and 2017, contained 81.58%, 82.34% and 75.65% soil information respectively.

Table 5. *Eigenvectors and cumulative contribution rates of principal components*

Principal components	**2015**		**2016**		**2017**	
	Eigenvectors	**Cumulative contribution rates (%)**	**Eigenvectors**	**Cumulative contribution rates (%)**	**Eigenvectors**	**Cumulative contribution rates (%)**
F1	3.69	40.98	3.28	36.45	2.66	29.54
F2	1.35	55.94	2.35	62.53	1.98	51.54
F3	1.28	70.12	1.78	82.34	1.15	64.32
F4	1.03	81.58	\	\	1.02	75.65

F1: the first principal component; F2: the second principal component; F3: the third principal component; F4: the fourth principal component

Then, the centroid vectors for all clusters in all datasets were determined. The input vector elements were sorted from small to large by the Euclidean distance from each cluster centroid. The smallest point was extracted as the typical sampling point for each cluster. The following weighted mean can be calculated from the typical sampling points and the number of points in the corresponding cluster to determine the estimated average of the field yield or potato quality from the sampling points detected through the clustering algorithm (*Eq.1*):

$$Y_j^{-est} = \frac{\sum_i^k Y_{BM_{ij}} * n_i}{N} \quad \text{(Eq.1)}$$

where Y_j^{-est} is the estimated mean yield or potato quality on the jth year, $Y_{BM_{ij}}$ is the yield or potato quality value on the jth year for typical sampling points of the ith cluster, where

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2003_19691989

their mean is denoted by $Y_{BM_{ij}}$, n_i describes the number of sampling points in the ith cluster, N indicates the whole number of sampling points, and k denotes the number of clusters.

The sum of typical sampling points is determined by multiplying the number of clusters by the number of typical sampling points for each cluster, affecting the prediction results. In the current work, we verify the choice between two to twelve clusters, including about 10% of the observed data locations, with one to three points per cluster.

Davies-Bouldin index (DBI) and the estimation coefficient of determination (R^2) were computed to compare k-means clustering validation, and the precisions of the estimated field mean yield or potato quality from various approaches. The estimated field means are compared to the corresponding “true” field means, as the arithmetic means of the whole observations for the corresponding year.

Clustering validation, assessing clustering results' performance, is necessary to evaluate the quality of clustering algorithms. The Davies-Bouldin Index, as the most popular internal clustering validation index, attempts to maximize the intra-cluster distance and minimize the inter-cluster distances. DBI can be calculated as follows (Gao et al., 2018). For any cluster C, its similarities to the other clusters are obtained, and the maximum similarity is considered as the cluster similarity for C. The mean of these cluster similarities is considered the DBI index. The smaller DBI value means superior clustering performance. The DBI should be minimized to obtain maximum separation between clusters and attain the optimum partition.

Estimation R^2 determines to what extent the algorithm is superior to a simple employment of the field average; a positive amount demonstrates the superiority of the algorithm to the three-year field mean, whereas a negative one demonstrates its weakness comparing with the three-year field mean. Estimation R^2 can be obtained from the following sums of squared errors (*Eq.2*):

$$R^2 = 1 - \frac{\sum_j^J \left(\bar{Y}_J - Y_j^{-est}\right)^2}{\sum_j^J \left(\bar{Y}_J - \bar{Y}\right)^2} \tag{Eq.2}$$

where $\bar{Y}_J$ indicates the mathematical mean of observed yield or potato quality on the jth year across 116 sampling points, Y_j^{-est} is the estimated mean yield or potato quality on the jth year, $\bar{Y}$ is the mean of all the three-year observed yield or potato quality. By considering invalid prediction number ($R^2 < 0$) as the abscissa axis and the average value of estimation R^2 as the ordinate axis, six scatter plots of the predictive validity of different K-Means methods for selecting critical sampling points were plotted here to compare the prediction of potato yield and quality based on different K-Means methods (*Figure 5*). The dotted lines in *Figure 5* show the mean values of scatter points.

Results and discussion

Descriptive statistics

Table 6 gives the descriptive statistics in tuber yield and quality within-field for the three years. Most variables had a normal distribution. A few contrasts of both yield and quality were notable. Yield in 2015, 2016, and 2017 were 54168 kg ha^{-1}, 54487 kg ha^{-1} and 70879 kg ha^{-1}, respectively. Although the same potato cultivar ‘Shepody’ was planted

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2003_19691989

in 2015 and 2017, the yield varied considerably. The higher rainfall within May to October was obtained in 2016 (772 mm) and 2017 (634 mm), compared with that in 2015 (379 mm). This may be due to the lower rainfall and poor irrigation management in the first year after mixing soil with Aeolian sandy soil and feldspathic sandstone. Dry matter content, protein, and starch in 2015 and 2017 were both higher than the corresponding ones in 2016. Differences in potato variety and climate can produce very different tuber yield and quality values from year to year.

Table 6. *Descriptive statistics for potato tuber yield and quality*

Year	Property	Mean	Min	Max	SD	Skewness	Kurtosis	CV	Num
2015	Yield（kg ha^{-1}）	54168	19048	103001	22379	0.57	-0.47	0.41	116
	DMC(%)	25.28	15.19	34.09	3.16	0.78	-0.33	0.11	116
	Protein(%)	0.62	0.21	1.33	0.00	-0.76	0.25	0.39	116
	Starch(%)	16.37	4.88	29.64	5.84	-0.70	0.07	0.36	116
	Reducing Sugar (%)	0.15	0.09	0.25	0.00	-0.01	0.70	0.24	116
	VC(mg $100g^{-1}$）	29.27	11.00	48.00	10.54	-0.08	-0.20	0.36	116
2016	Yield（kg ha^{-1}）	54487	29943	78411	12041	0.22	-0.71	0.22	116
	DMC(%)	21.36	14.41	28.21	3.04	0.37	0.10	0.13	116
	Protein(%)	0.54	0.25	0.90	0.15	0.43	0.48	0.32	116
	Starch(%)	11.59	2.44	22.75	4.09	0.21	0.44	0.35	116
	Reducing Sugar (%)	0.15	0.09	0.21	0.03	-0.28	-0.26	0.19	116
	VC(mg $100g^{-1}$）	27.12	5.00	57.50	9.02	1.29	0.13	0.33	116
2017	Yield（kg ha^{-1}）	70879	43679	107448	11675	0.10	0.22	0.16	116
	DMC(%)	23.97	18.28	27.18	2.66	0.54	0.05	0.08	116
	Protein(%)	0.73	0.34	1.34	0.00	0.37	0.37	0.28	116
	Starch(%)	15.91	6.47	35.78	5.14	0.82	0.86	0.32	116
	Reducing Sugar (%)	0.10	0.04	0.27	0.00	0.87	1.39	0.32	116
	VC(mg $100g^{-1}$）	24.00	11.00	41.00	6.00	0.54	0.73	0.25	116

DMC, Dry matter content, VC, Vitamin C, CV: Coefficient of variation

According to the coefficient of variation (CV) as a measure of field stability with years, reducing sugar and dry matter content (DMC) were more stable than others. Spatial variation was remarkable for most variables. The CV of yield in 2015 was as high as 0.41, while the minimum yield was only 19048 kg ha^{-1}, but the highest was 103001 kg ha^{-1}. The wide yield range suggested that soil variability may significantly influence the tuber. Therefore, the mean tuber production can be considerably lower than its capability. The CV of yield tended to diminish among years, ranging from 0.41 to 0.16. In contrast, according to Cambouris studies, the potato yield variability had an approximately fixed value (from 0.24 to 0.27) in each year. Although the potato was also planted in sandy loam soil, it has a lower variation of soil sand content (CV=0.03) (Cambouris et al., 2006). A similar trend also can be seen in Protein and VC.

Descriptive statistics of soil properties are summarized in *Table 7*. Regardless of the small region of the experimental plot, the spatial and temporal variabilities of most variables were considerable. After mixed, the study fields contrast in soil texture with sandy loam to sandy among three years. Since the clay proportion remained low (lower than 5%), fluctuating from 0.6% to 4.3% through three years, soil texture was determined by silt and sand. After the first year, silt reduced moderately from 32.02% to 27.04%, then tumbled to 9.4% in the third year. In contrast to silt, sand increased slightly from

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2003_19691989

64.9% to 68.45% and then raised rapidly to 90.00% through three years. This may due to instability of mixed soil and the downward migration of silt. Recent evidence suggested that continuous monoculture systems could degrade soil aggregate stability due to low organic inputs and tillage practice disturbances (Acosta-Martinez et al., 2004; Ma et al., 2016). It was reported that soil aggregate stability decreased significantly in the potato monoculture system after two years of continuous cropping. In addition, the soil was just mixed, and the soil aggregate was not stable in the process of formation. Therefore, silt may migrate downward with high-frequency irrigation and rainfall. In this study, since the rainfall in 2016 (772 mm) is higher than in 2015 (379 mm), this phenomenon was more noticeable after two cultivation years. However, soil texture in deeper than 40 cm had not been analyzed. The reasons for the drastic change of soil texture in depth of 0-40 cm need to be further studied.

Table 7. *Descriptive statistics of soil features among three years*

Year	Property	Mean	Min	Max	SD	Skewness	Kurtosis	CV	Num
2015	Clay（%）	3.09	0.24	8.28	1.48	0.76	1.68	0.48	116
	Silt（%）	32.01	7.23	59.27	12.37	0.27	-0.95	0.39	116
	Sand（%）	64.90	35.10	92.54	13.44	-0.25	-0.96	0.21	116
	Θ_{v1}（%）	11.64	6.16	19.81	3.49	-0.74	0.39	0.30	116
	Θ_{v2}（%）	9.46	5.51	17.04	2.82	-0.24	0.77	0.30	116
	Θ_{v3}（%）	11.47	4.58	19.81	4.35	-0.82	0.18	0.38	116
	N_1(mg kg^{-1})	90.49	23.59	168.54	28.17	-0.08	-0.08	0.31	116
	N_2(mg kg^{-1})	46.33	7.32	91.84	16.28	-0.25	0.06	0.35	116
	N_3(mg kg^{-1})	29.69	10.48	51.10	8.27	-0.22	0.00	0.28	116
2016	Clay（%）	3.26	0.46	7.19	1.35	0.69	-0.08	0.41	116
	Silt（%）	23.09	5.75	52.89	9.06	0.41	-0.39	0.39	116
	Sand（%）	73.65	41.16	93.53	10.17	-0.33	-0.34	0.14	116
	Θ_{v1}（%）	9.19	5.51	16.19	2.26	-0.17	0.60	0.25	116
	Θ_{v2}（%）	11.31	5.85	18.94	2.73	0.27	0.58	0.24	116
	Θ_{v3}（%）	9.09	4.96	16.71	2.60	-0.26	0.64	0.29	116
	N_1(mg kg^{-1})	88.07	54.63	131.69	15.94	-0.12	0.34	0.18	116
	N_2(mg kg^{-1})	37.83	13.75	70.20	11.36	0.63	0.44	0.30	116
	N_3(mg kg^{-1})	30.00	9.58	55.56	10.58	-0.31	0.43	0.35	116
2017	Clay（%）	2.70	0.58	8.92	1.32	0.80	1.08	0.49	116
	Silt（%）	10.38	4.28	25.83	3.82	0.08	0.25	0.37	116
	Sand（%）	86.92	65.25	95.14	5.07	0.14	-0.20	0.06	116
	Θ_{v1}（%）	10.40	6.47	14.48	3.00	0.73	1.10	0.19	116
	Θ_{v2}（%）	9.89	6.80	17.18	2.95	0.71	1.25	0.21	116
	Θ_{v3}（%）	7.68	5.15	12.15	2.51	0.34	0.98	0.16	116
	N_1(mg kg^{-1})	84.06	15.01	188.78	40.63	0.27	0.62	0.48	116
	N_2(mg kg^{-1})	40.66	4.73	79.36	16.99	-0.01	0.51	0.42	116
	N_3(mg kg^{-1})	27.68	8.76	59.43	13.24	0.17	0.85	0.48	116

CV: Coefficient of Variation. Θv_1, Θv_2, Θv_3: Mean value of soil water content in June, July and August; N_1, N_2, N_3: soil nitrate-nitrogen in June, July, August

The variation range of soil moisture content was 7.68% – 10.40% in 2017, 9.09% – 11.31% and 9.46% – 11.64% in 2015 and 2016, respectively. It can be seen that the soil moisture content in 2017 was slightly lower than that in previous two years, which was related to the change of soil texture. Soil nitrate-nitrogen declined from 90.49 mg kg^{-1}

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2003_19691989

to 29.69 mg kg^{-1} within the growth period in 2015. The same trend can be seen in the next two years. Soil available nitrogen was 33 mg kg^{-1} before planting in 2015. After cultivation, soil available nitrogen in August were 29.69 mg kg^{-1}, 30 mg kg^{-1} and 27.68 mg kg^{-1} in 2015, 2016 and 2017, respectively. It can be seen that soil available nitrogen was kept approximately unchanged. The nitrogen consumption of potatoes was induced by fertilization. Meanwhile, the potato monoculture system was not conductive to nitrogen accumulation in the mixed soil.

The CV of soil sand content decreased considerably among years from 0.21 to 0.06, while the CV of clay and silt fluctuated around 0.38 and 0.48, respectively. However, according to several reported studies involving spatial variation within-field, spatial variation of soil particle distribution remained fixed year by year (Casa et al., 2008; Li et al., 2016). Meanwhile, according to the CV amounts ranging from 0.15 to 0.35, most soil chemical and physical features had moderate variability. Based on the soil survey reported by Cambouris et al. (2006), the CV of clay, silt, and sand were 0.20, 0.29, and 0.03, respectively. Similar to soil sand content, the CV of soil moisture declined during years from 0.33 to 0.16. However, for nitrogen, this was not the case. The maximum nitrogen variability was observed in 2017. In this study, fertilization was the primary source of soil nitrogen. Therefore, it is not easy to ensure the uniformity of fertilization.

Potato tuber yield and quality had between strong and weak spatial dependency, and except reducing sugar content showed non-consistent spatial variability according to soil temporal variability. This finding was compatible with Redulla's findings, demonstrating yield-driving factors variability for various fields and seasons (Redulla et al., 2002). However, the CV of yield, protein, and VC changed more dramatically among years than in previous studies. This may be due to the mixed soil's instability with the downward migration of silt.

Geostatistic analysis

The spatial variability of potato yield and quality are presented in *Figure 4*, while the best fitting model parameters are given in *Table 8*.

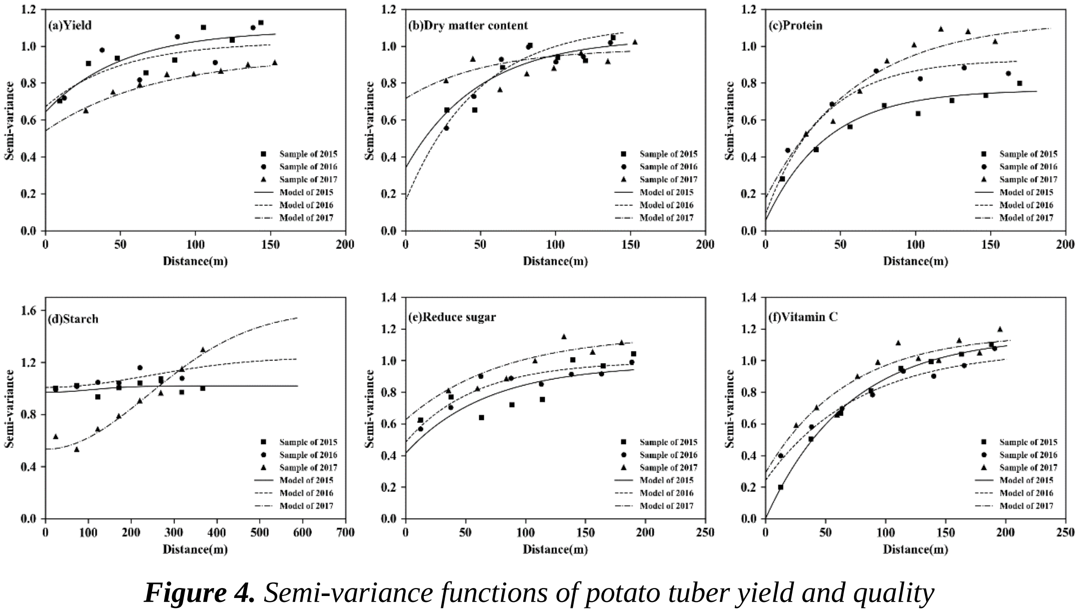

Figure 4. *Semi-variance functions of potato tuber yield and quality*

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2003_19691989

Table 8. *Semi-variance function model parameters of potato tuber yield and quality*

Year	Property	Model	Nugget	Sill	Range (m)	Nugget/sill (%)	R^2	Num
2015	Yield（kg ha^{-1})	Exponential	0.80	1.14	151	70	0.86	116
	DMC(%)	Exponential	0.33	1.03	147	32	0.83	116
	Protein(%)	Exponential	0.05	0.75	122	7	0.92	116
	Starch(%)	Exponential	0.97	1.02	254	95	0.82	116
	Reducing Sugar (%)	Exponential	0.41	0.97	202	42	0.75	116
	VC(mg $100g^{-1}$）	Exponential	0.69	0.94	324	73	0.89	116
2016	Yield（kg ha^{-1})	Exponential	0.67	1.02	148	66	0.79	116
	DMC(%)	Exponential	0.16	1.12	145	14	0.86	116
	Protein(%)	Exponential	0	0.97	103	0	0.82	116
	Starch(%)	Exponential	1	1.22	588	82	0.86	116
	Reducing Sugar (%)	Exponential	0.48	0.99	173	48	0.79	116
	VC(mg $100g^{-1}$）	Exponential	0.72	1.01	321	71	0.85	116
2017	Yield（kg ha^{-1})	Exponential	0.54	0.94	216	57	0.92	116
	DMC(%)	Exponential	0.71	0.98	160	72	0.72	116
	Protein(%)	Exponential	0.17	1.13	181	15	0.80	116
	Starch(%)	Gauss	0.53	1.58	588	34	0.85	116
	Reducing Sugar (%)	Exponential	0.62	1.16	242	53	0.81	116
	VC(mg $100g^{-1}$）	Exponential	0.29	1.17	204	25	0.82	116

DMC, Dry matter content, VC, Vitamin C

The exponential model was the most used, that approached sill value asymptotically, while the gauss model was only utilized for starch in 2017. Tuber yield showed moderate and non-consistent spatial dependence among three years, with the nugget-to-sill ratio decreasing from 70% to 57%. It had a more significant range value (216 m) and a lower nugget-to-sill ratio (57%) in 2017 than in the previous years, indicating that the tuber yield had stronger spatial dependence in 2017. According to the above results, there was higher soil sand content in 2017. It can be concluded that potato tuber yield may have more considerable spatial dependence in sandy soil. In contrast, according to Cambouris et al. (2006), consistent and unchanged variability in tuber yield from 1998 to 2000 can be found, with the nugget-to-sill ratio around 47%. The nugget-to-sill ratio values in the current work were higher than those obtained by Cambouris et al. (2006) and derived from the relatively less uniform soil texture at the mixed-soil site.

Dry matter content exhibited two spatial patterns: moderate spatial dependence in 2015 and 2017 and strong spatial dependence in 2016. The lower nugget-to-sill ratio values for tuber dry matter content were obtained in 2015 (32%) and 2016 (14%). A similar trend can be seen in *Figure 3(b)*, whereas a more significant ratio can be observed in 2017 (72%). It can be concluded that the spatial structure of tuber dry matter content varied widely among years and was sensitive to the temporal variation of soil texture.

Tuber protein content had a strong spatial dependence with the nugget-to-sill ratio ranging from 0 and 15% and with the range value between 103m and 181m. Usually, strong spatial dependence may be due to inherent factors like soil water content and soil nutrient (Cambardella et al., 1994).

Starch in 2015 and 2016 both showed weak spatial dependence under the selected sampling procedure with the nugget-to-sill ratio between 95% and 82%, and showed moderate spatial dependence in 2017 with the nugget-to-sill ratio of 34%, indicated that

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2003_19691989

a severe range value could be distinguished outside the field size, or the number of samples was not enough for the spatial dependence extrapolation. A similar variation pattern for tuber to reducing sugar content was observed among three years with similar nugget-to-sill ratio and range value, demonstrating moderate spatial dependence. It demonstrated that, even though the soil texture varied widely among years, the reducing sugar content's consistent spatial distribution could be observed. Thus, soil texture could not be considered the main factor for the structure spatial variability of the reducing sugar content.

Although tuber vitamin C showed moderate spatial dependence in three years, it had a much lower nugget-to-sill ratio (25%) in 2017 than in the previous years (73% and 71%), indicating its stronger spatial dependence in 2017. The mentioned trend was analogous to the tuber yield's spatial variability.

In conclusion, potato tuber yield and quality except reducing sugar content showed non-consistent spatial variability according to soil temporal variability. These results were compatible with previous studies. The inconsistency over time of spatial variability patterns of essential crop features like yield and protein has been demonstrated in the literature (Redulla et al., 2002; Blackmore et al., 2003; Maestrini et al., 2018).

Correlation analysis

Table 9 describes the Spearman correlation among tuber yield and quality and soil features. Although significant correlation relationships were extracted among soil features and yield, a few contrasts of the correlations can be observed between soil features and potato yield and quality in various growing seasons.

Table 9. *Spearman's correlation coefficient of potato tuber yield and quality with soil properties*

Property	Year	Clay	Silt	Sand	θv_1	θv_2	θv_3	N_1	N_2	N_3
Yield	2015	NS	NS	NS	-0.31**	-0.34**	-0.43**	0.59**	-0.41**	-0.36**
	2016	NS	-0.25**	0.23*	NS	-0.27**	-0.31**	0.71**	-0.28**	-0.29**
	2017	NS	0.34*	-0.33*	NS	NS	NS	NS	NS	0.40*
DMC	2015	0.33**	0.21*	-0.24*	0.20*	NS	NS	NS	NS	-0.19*
	2016	0.27**	0.31**	-0.31**	0.36**	0.30**	0.38**	0.42**	-0.25*	-0.24*
	2017	NS	NS	NS	NS	NS	NS	NS	NS	0.38*
Protein	2015	NS	0.27**	-0.25**	NS	NS	NS	NS	NS	NS
	2016	NS	NS	NS	0.26**	NS	NS	NS	-0.23*	NS
	2017	NS	-0.33*	0.32*	-0.35*	-0.34*	-0.35*	NS	NS	-0.34*
Starch	2015	NS	0.29**	-0.28**	0.23*	0.22*	0.29**	NS	NS	NS
	2016	NS	0.23*	-0.23*	.19*	NS	NS	0.18*	NS	NS
	2017	NS	NS	NS	NS	NS	NS	NS	NS	NS
Reducing Sugar	2015	NS	-0.29**	0.28**	NS	NS	NS	NS	NS	NS
	2016	-0.21*	-0.22*	0.22*	NS	NS	NS	NS	NS	NS
	2017	NS	NS	NS	NS	NS	NS	NS	NS	NS
VC	2015	-0.20*	-0.30**	0.30**	NS	NS	NS	NS	NS	NS
	2016	NS	-0.27**	0.26**	NS	NS	NS	NS	NS	NS
	2017	-0.33*	NS	NS	-0.33*	-0.39*	NS	NS	NS	NS

DMC, Dry matter content, VC, Vitamin C, θv_1, θv_2, θv_3: Mean value of soil water content in June, July and August; N_1, N_2, N_3: soil nitrate-nitrogen in June, July, August; PC-1, The First Principal Component, NS, Not Significant. *, **: Significant at the 0.05, 0.01 probability levels, respectively

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2003_19691989

The yield was negatively and positively correlated to silt in 2016 and 2017, respectively. In contrast, it was positively and negatively correlated to sand in 2016 and 2017, respectively. This was compatible with earlier observations (Redulla et al., 2002), which showed that correlation coefficients among potato point yield and silt were positive in one field and negative in another with different soil textures. Although most soil moistures were negatively correlated to yield in the previous two growing seasons, this significant correlation relationship cannot be observed in 2017. In the first two years, the yield was negatively correlated with N_2 and N_3 while positively correlated with N_1. However, the yield was only positively correlated with N_3 in the last year. It can be concluded that soil texture was sandy loam in 2015 and 2016, while sandy in 2017. It also means in the sandy loam field, the more soil water content and nitrogen in August, the lower the potato yield. The opposite phenomenon can be observed in sandy soil. These results may be due to the less drainability and the lower soil moisture tension in the sandy clay loam soil than in sandy soil. The mentioned results were in good agreement with those observed in earlier studies. Saini (1976) studied the physical parameters of different soils and discovered the high correlation between the oxygen diffusion rate (ODR) and potato yield. Holder et al. (1984) extracted a remarkable relation between oxygen diffusion rate (ODR) and soil moisture tension (matric potential) at the 30 and 40 cm depths. These results indicated that irrigation might be excessive for potatoes in an andy clay loam field.

Similar to yield, dry matter content was negatively correlated with N_2 and N_3 while positively correlated with N_1 in 2016, indicating the higher nitrogen consumption in July and August (End of flowering-Onset of senescence and tuber expansion) compared with June (Onset of flowering and Tuber formation). The mentioned results were compatible with those of other works and suggested that the nitrogen absorption rate increased gradually after sowing to reach its maximum in the tuber expansion period (Wang et al., 2014). Therefore, it is necessary to fertilize nitrogen additionally in the mentioned period.

In general, more than 66% of the correlation coefficients among soil sand content and tuber variables were significant, with positive correlation coefficients ranging from 0.22 (Reducing Sugar in 2016) to 0.32 (Protein in 2017) and negative ones ranging from -0.23 (Starch in 2016) to -0.37 (DMQ in 2017). Redulla et al. (2002) reported that soil texture parts (sand, silt, clay) had a more significant effect on tuber yield than the soil chemical features (N, P, K, organic matter, pH). Spatial variation of soil texture may be the fundamental cause of the spatial variability of potato yield and quality in this mixed field.

Estimation analysis

Figure 5 shows DBI amount of clusters computed by k-means for all datasets. K-means gave higher DBI values for high-dimensional data (texture dataset, N-All dataset, Θ_v-All dataset, PC-All dataset) compared to one-dimensional data (sand dataset, Θ_v dataset, N-1 dataset, PC-1 dataset). For DBI, a lower criterion value represented the optimal cluster. In high-dimensional datasets, the smaller the Euclidean distance between clusters, the higher DBI was. According to similar results, the whole objects in a high-dimensional dataset usually were almost equidistant from each other and entirely mask the clusters (Parsons et al., 2004). However, it is also effective to determine the quality of clusters using DBI. According to the elbow principle, the optimal clustering number for K-Means on the PC-All dataset was 6, while this value was 2 for the other datasets.

Figure 6 shows six scatter plots of the predictive validity of different K-Means methods for selecting critical sampling points. For most tuber variables (except protein),

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2003_19691989

points of high-dimensional datasets fall on the left of points of one-dimensional datasets, indicating that high-dimensional datasets-based yield and qualities prediction was more stable than the one-dimensional ones, suggesting that these data without temporal variation information cannot be utilized to choose critical sampling points. Prior studies have noted the importance of temporal variation for yield spatial variation (Li et al., 2016).

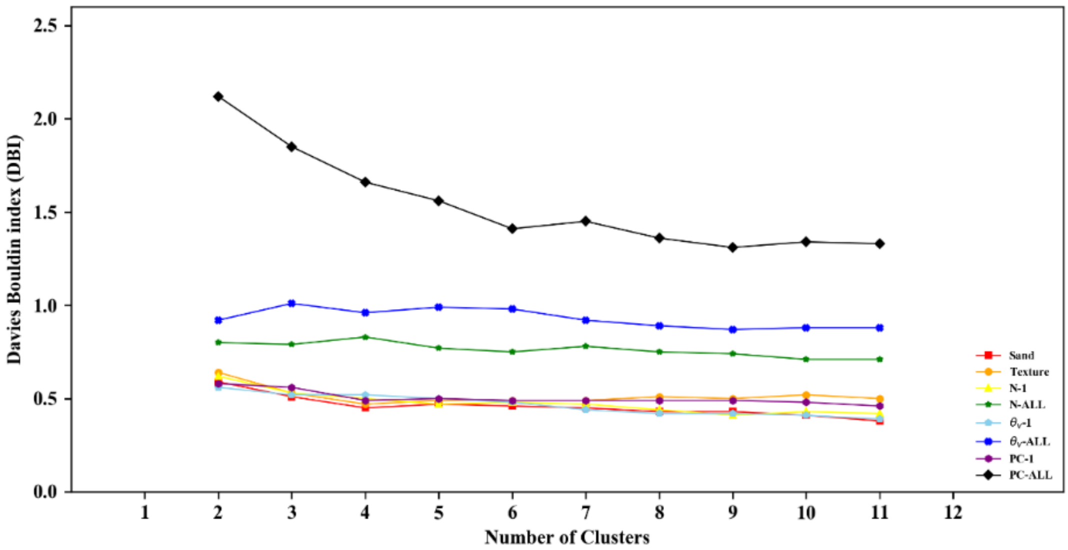

Figure 5. *DBI values of clusters*

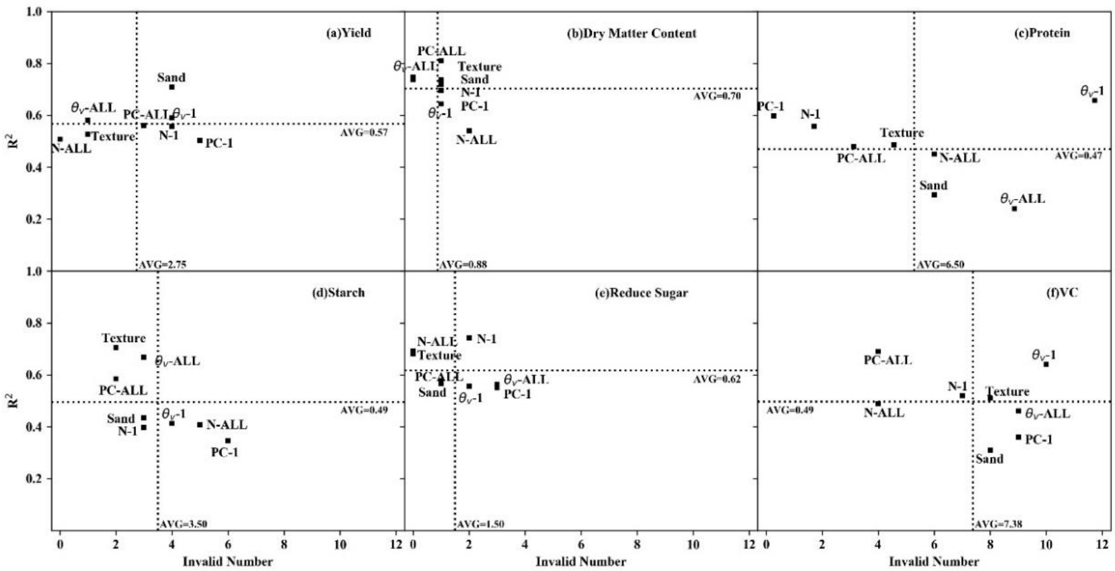

Figure 6. *Scatter plots of the predictive validity of different K-Means method*

For yield, estimation R^2 of most datasets, except sand, fluctuated moderately around the mathematical average (0.54). Although higher estimation R^2 was observed, sand dataset-based yield prediction gave inconsistent results. This result may be induced by

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2003_19691989

the inconsistency of spatial variability of yield and the relationship between soil sand content and tuber yield. ϴv-All had superior prediction results with the point fallen in the second quadrant (*Figure 6a*). It can thus be suggested that soil moisture was the primary constraint for spatial variability of potato yield in this site. Soil moisture content was detected as an essential factor influencing potato yields and tuber quality in several works (Van Loon, 1981; Fabeiro et al., 2001; Ahmadi et al., 2010). Fulton had reported that inadequate soil moisture could reduce the potato yield (Fulton, 1970).

Compared to yield and other potato quality, dry matter content had better prediction results, with a higher mean of estimation R^2 (0.72) and lower valid number (1) (*Figure 6b*). In the statistical sense, the lowest CV was obtained for the DMC. For DMC, the PC-All dataset gave the best performance with the estimation R^2 0.81, while the N-All dataset gave the lowest estimation R^2 (0.54), indicating that these vector data alone were not enough for the confident selection of critical sampling points.

In contrary to dry matter content, protein had unsatisfactory prediction results, with a lower mean of estimation R^2 (0.47) and higher valid number (6.5) (*Figure 6c*). For protein, both PC-1 and N-1 datasets exhibited relatively better performance. As shown in the first principal component's component matrix (Not listed in this paper), the correlation coefficients between the first principal component and soil particles were significant (correlation coefficients were about 0.77). As with the correlation results, there was a noticeable correlation between protein and soil particles across three years and a negative correlation among protein and nitrogen in August 2017.

For starch, K-means on texture dataset often provided satisfactory results, with higher estimation R^2 (0.71) than another approaches (*Figure 6d*). This finding confirmed correlation analysis results, indicating a significant correlation between soil particle distributions and potato starch. According to both analyses, the starch variation was mainly affected by soil particle distributions compared with soil moisture and nitrogen. Although the relationship between starch and soil texture has been studied in few works, some studies pointed that insufficient soil water could reduce starch yield (Jin et al., 2015). These results showed that soil texture might affect starch content in another way, requiring more depth experience and analysis.

The prediction effect of reducing sugar was only after the dry matter content (*Figure 6e*), which may be due to the consistent spatial distribution. The average values of estimation R^2 and valid number were 0.62 and 1.5, respectively. Across all metrics, the best performance belonged to N-All and soil texture datasets. Although a significant coefficient between nitrogen and reducing sugar was not observed, this finding suggested that soil particle distributions and nitrogen with temporal variation information were sufficient to select critical sampling points.

As shown in *Figure 6f,* VC had the most inconsistent results with the highest valid number (the mean value was 7.4) compared to all the potato quality and yield. These results may be due to the varied wildly spatial dependence among three years. The best performance belonged to the PC-All dataset, indicating that independent principal components extracted from soil moisture, nitrogen, and particle distribution could provide a little explanation of VC spatial variability for this site.

Taken together, K-means methods employed different soil attributes for identifying critical sampling points, which gave a more accurate estimation of the field mean of dry matter content and reducing sugar but generally performed erratically in the estimation of the field average of VC and protein. Interestingly, in all potato yield and quality variables of this study, dry matter content and reducing sugar had the lowest CV. In

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2003_19691989

contrast, in the geostatistical analysis sense, VC and Protein have moderate spatial dependence, which means that there is a factor affecting their structural variations. Although correlation analysis found a significant correlation between soil attributes and VC or protein, some properties like soil moisture, nitrogen, and particle distribution could not affect VC and protein directly, not the root of variability.

Conclusions

Potato tuber yield and quality had between strong and weak spatial dependency, and except reducing sugar content showed non-consistent spatial variability according to soil temporal variability.

Correlation analysis showed that more than 66% of the correlation coefficients between soil properties and tuber variables were significant, indicating that an improvement in the temporal variation of soil texture was the most excellent chance for site-specific potato management in mixed soil.

The evaluation of K-means clustering algorithms' effectiveness showed that Ɵv- All, PC-All, and texture had superior prediction results for tuber yield, dry matter content, and starch, respectively. N-All and soil texture datasets exhibited the best performance for reducing sugar. Unsatisfactory prediction results were obtained for Protein and VC. These findings can be utilized to detect the possible opportunity for in-field site-specific potato management for different objectives. The method proposed by Van Arkel was effective for estimating potato quality which had stable spatial variation structure. More research is needed for potato yield and quality with large spatial-temporal variation.

Acknowledgements. This research was supported financially by the National Natural Science Foundation of China (Grant No 51609197), CAS 'Light of West China' Program (Grant No. XAB2016AW06), Programme of Introducing Talents of Discipline to Universities (No. 104-451115012), and Scientific Research Program of Education Department in Shaan Xi Province (No. 16JS084).

REFERENCES

[1] Acosta-Martinez, V., Zobeck, T. M., Allen, V. (2004): Soil microbial, chemical and physical properties in continuous cotton and integrated crop–livestock systems. – Soil Science Society of America Journal 68(6): 216-227.

[2] Ahmadi, S. H., Andersen, M. N., Plauborg, F., Poulsen, R. T., Jensen, C. R., Sepaskhah, A. R., Hansen, S. (2010): Effects of irrigation strategies and soils on field grown potatoes: Yield and water productivity. – Agricultural Water Management 97(11): 1923-1930.

[3] Blackmore, S., Godwin, R. J., Fountas, S. (2003): The analysis of spatial and temporal trends in yield map data over six years. – Biosystems Engineering 84(4): 455-466.

[4] Cambardella, C. A., Moorman, T. B., Novak, J. M., Parkin, T. B., Konopka, A. E. (1994): Field-scale variability of soil properties in central Iowa soils. – Soil Science Society of America Journal 58(5): 1501-1511.

[5] Cambouris, A. N., Nolin, M. C., Zebarth, B. J., Laverdière, M. R. (2006): Soil management zones delineated by electrical conductivity to characterize spatial and temporal variations in potato yield and in soil properties. – American Journal of Potato Research 83(5): 381-395.

[6] Casa, R., Castrignanò, A. (2008): Analysis of spatial relationships between soil and crop variables in a durum wheat field using a multivariate geostatistical approach. – European Journal of Agronomy 28(3): 331-342.

[7] Cawse, P. (1967): The determination of nitrate in soil solutions by ultraviolet spectrophotometry. – Analyst 92(1094): 311-315.
[8] de Souza, Z. M., de Souza, G. S., Júnior, J. M., Pereira, G. T. (2014): Number of samples in geostatistical analysis and kriging maps of soil properties. – Ciencia Rural 44(2): 261-268.
[9] Fabeiro, C., Martın de Santa Olalla, F., De Juan, J. A. (2001): Yield and size of deficit irrigated potatoes. – Agricultural Water Management 48(3): 255-266.
[10] Fulton, J. M. (1970): Relationship of root extension to the soil moisture level required for maximum yield of potatoes, tomatoes and corn. – Canadian Journal of Soil Science 50(1): 92-94.
[11] Gao, X., Yang, M. (2018): Understanding and enhancement of internal clustering validation indexes for categorical data. – Algorithms 11(11): 177.
[12] Han, J., Xie, J., Zhang, Y. (2012): Potential role of feldspathic sandstone as a natural water retaining agent in mu us sandy land, northwest china. – Chinese Geographical Science 22(5): 550-555.
[13] Holder, C. B., Cary, J. W. (1984): Soil oxygen and moisture in relation to russet burbank potato yield and quality. – American Potato Journal 61(2): 67-75.
[14] Irmak, A., Batchelor, W. D., Jones, J. W., Irmak, S., Paz, J. O., Beck, H. W., Egeh, M. (2002): Relationship between plant avai lable soil water and yield for explaining soybean yield variability. – Applied Engineering in Agriculture 18(4): 471-484.
[15] Jin, G. H., Feng, Y. X., Liu, D. Q., Mao, R. (2015): Effect of different soil moistures on potato tuber yield and starch formation. – Journal of Northeast Agricultural University 46(10): 10-14.
[16] Li, T., Hao, X. M., Kang, S. Z. (2016): Spatial variability of grape yield and its association with soil water depletion within a vineyard of arid northwest China. – Agricultural Water Management 5: 4482-4491.
[17] Li, T., Zhang, J.-F., Xiong, S.-Y., Zhang, R.-X. (2020): The spatial variability of soil water content in a potato field before and after spray irrigation in arid northwestern China. – Water Supply 20(3): 860-870.
[18] Ma, K., Yang, Y. J., Ma, L., Wang, C. M., He, W. S. (2016): Effects of intercropping on soil microbial communities after long-term potato monoculture. – Acta Ecologica Sinica 36(10): 2987-2995.
[19] Maestrini, B., Basso, B. (2018): Predicting spatial patterns of within-field crop yield variability. – Field Crops Research 219: 106-112.
[20] McGrance, S. J., Cornell, H. J., Rix, C. J. (1998): A simple and rapid colorimetric method for the determination of amylose in starch products. – Starch-Stärke 50(4): 158-163.
[21] Olsen, S. R. (1954): Estimation of available phosphorus in soils by extraction with sodium bicarbonate. – US Department of Agriculture.
[22] Parsons, L., Haque, E., Liu, H. (2004): Subspace clustering for high dimensional data: A review. – ACM SIGKDD Explorations Newsletter 6(1): 90-105.
[23] Perron, I., Cambouris, A. N., Chokmani, K., Vargas Gutierrez, M. F., Zebarth, B. J., Moreau, G., Biswas, A., Adamchuk, V. (2018): Delineating soil management zones using a proximal soil sensing system in two commercial potato fields in New Brunswick, Canada. – Canadian Journal of Soil Science 98: 724-737.
[24] Po, E. A., Snapp, S. S., Kravchenko, A. (2010): Potato yield variability across the landscape. – Agronomy Journal 102(3): 885-894.
[25] Redulla, C. A., Davenport, J. R., Evans, R. G., Hattendorf, M. J., Alva, A. K., Boydston, R. A. (2002): Relating potato yield and quality to field scale variability in soil characteristics. – American Journal of Potato Research 79(5): 317-323.
[26] Rietjens, I. M., Boersma, M. G., de Haan, L., Spenkelink, B., Awad, H. M., Cnubben, N. H., van Zanden, J. J., van der Woude, H., Alink, G. M., Koeman, J. H. (2002): The pro-oxidant chemistry of the natural antioxidants vitamin C, vitamin E, carotenoids and flavonoids. – Environmental toxicology and pharmacology 11(3-4): 321-333.

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2003_19691989

[27] Rockström, J., Barron, J., Brouwer, J., Galle, S., De Rouw, A. (1999): On-farm spatial and temporal variability of soil and water in pearl millet cultivation. – Soil Science Society of America Journal 63(5): 1308-1319.
[28] Rosenzweig, N., Steere, L., Gerondale, B., Kirk, W. (2016): A geostatistical approach to visualize the diversity of soil inhabiting bacteria and edaphic qualities in potato (Solanum tuberosum) production systems. – American Journal of Potato Research 93(5): 518-532.
[29] Saini, G. R. (1976): Relationship between potato yield and oxygen diffusion rate of subsoil. – Agronomy Journal 68(5): 823-825.
[30] Shuman, L. M., Duncan, R. R. (1990): Soil exchangeable cations and aluminum measured by ammonium-chloride, potassium-chloride, and ammonium acetate. – Commun Soil Sci Plant Anal 21(13-16): 1217-1228.
[31] Starr, G. C. (2005): Assessing temporal stability and spatial variability of soil water patterns with implications for precision water management. – Agricultural Water Management 72(3): 223-243.
[32] Taylor, J. A., Chen, H., Smallwood, M., Marshall, B. (2018): Investigations into the opportunity for spatial management of the quality and quantity of production in uk potato systems. – Field Crops Research 229: 95-102.
[33] Tian, Y., Huang, Z.-G., Yu, X.-Q. (2011): Experimental study on water requirement of potato. – Modern agricultural science and technology 8: 91-92.
[34] Vachaud, G., Passerat de Silans, A., Balabanis, P., Vauclin, M. (1985): Temporal stability of spatially measured soil water probability density function. – Soil Science Society of America Journal 49(4): 822-828.
[35] Van Arkel, Z. J., Kaleita, A. L. (2015): Identifying sampling locations for field-scale soil moisture estimation using k-means clustering. – Water Resources Research 50(8): 7050-7057.
[36] Van Loon, C. D. (1981): The effect of water stress on potato growth, development, and yield. – American Potato Journal 58(1): 51-69.
[37] Vrindts, E., Mouazen, A. M., Reyniers, M., Maertens, K., Maleki, M. R., Ramon, H., Baerdemaeker, J. D. (2005): Management zones based on correlation between soil compaction, yield and crop data. – Biosystems Engineering 92(4): 419-428.
[38] Wang, H. Y., Zhou, J. M. (2014): Root-zone fertilization: A key and necessary approach to improve fertilizer use efficiency and reduce non-point pollution from the cropland. – Soils 5: 785-790.
[39] Wang, J., Yang, R., Bai, Z. (2015): Spatial variability and sampling optimization of soil organic carbon and total nitrogen for minesoils of the loess plateau using geostatistics. – Ecological Engineering 82: 159-164.
[40] Warrick, A. W., Gardner, W. R. (1983): Crop yield as affected by spatial variations of soil and irrigation. – Water Resources Research 19(1): 181-186.
[41] Weir, J. P. (2005): Quantifying test-retest reliability using the intraclass correlation coefficient and the sem. – The Journal of Strength & Conditioning Research 19(1): 231-240.
[42] Wesenbeeck, I. J. V., Kachanoski, R. G. (1988): Spatial and temporal distribution of soil water in the tilled layer under a corn crop. – Soil Science Society of America Journal 52(2): 363-368.
[43] Yan, L., Ping, Z., Yan, N., Shukun, S., Xiuqing, Z., Hong, S. (2006): Biuret method to measure the protein content in soybean whey. – Soybean Bulletin 6.
[44] Yeomans, J. C., Bremner, J. M. (1988): A rapid and precise method for routine determination of organic carbon in soil. – Communications in Soil Science and Plant Analysis 19(13): 1467-1476.
[45] Zhang, H., Cao, T., Sun, X., Xu, Y. (2021): Temporal and spatial variation characteristics of soil mechanical composition after aeolian soil improvement by soft rock in mu us sandy land. – Bangladesh Journal of Botany: 865-872.
[46] Zhao, K., Xue, P.-J., Gu, G.-Y. (2008): Study on determination of reducing sugar content using 3, 5-dinitrosalicylic acid method. – Food Science 8: 128.

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2003_19691989