ASSESSMENT OF THE OPTIMAL NUMBER OF OBSERVATIONS IN THE STUDY OF VINEYARD SOIL (RIGOSOL)

Djordjević, A.[1], *Životić, Lj. [1], Sivčev, B.[1], Pajić, V.[1], Ranković-Vasić, Z.[1], Radovanović, D[1].
[1]University of Belgrade, Faculty of Agriculture, Nemanjina 6, Belgrade, Zemun, Republic of Serbia, *ljubaz@agrif.bg.ac.rs

ABSTRACT
A study of soil pH on the experimental field resulted in a high variability of pH on a very small scale. This kind of heterogenity in soil pH have effects on growth of two grapevine varieties on rootstock Kober 5BB: Riesling and Pinot Noir A number of 104 soil samples were taken from an area of 1.43 ha from two depths. A goal of this experiment was to find the optimum number of samples for pH studies, and to implement the obtained results in further investigation on experimental fields. Therefore, in this paper we compared diferent deterministic interpolation techniques: inverse distance weight, splines and local polynomial interpolation, on the results of soil pH. Root mean square error (RMSE) statistitics obtained after cross validation procedure was used for the choice of appropriate exponent value for IDW, spline and local interpolation. The obtained interpolation parameters were used for mapping the field and the most accurate technique was IDW, which was further used in creation of pH maps with lower number of samples: 54, 34, 29, 24, 19 and only 14 pH samples. Maps were classified and compared by means of percentage difference in area among classes of pH in respect to classes obtained after maximum sampling. The results indicated that the criteria of 15% of change in pH area over classes could be satisfied with only on third of the samples. An obtained results will be used for further sampling of the whole experimental area.

KEYWORD
vineyard, soil, pH, interpolation, IDW, RBF, LP

INTRODUCTION

The feasibility of site-specific management relies on the understanding of temporal and spatial components of variability. Spatial variability of soil properties may vary due to natural variations or imposed sources of variability. The variability of soil properties could be vertically and horizontally along a field. The structure of variability in soil properties showed differences according to sampling spacing, soil properties, and method used in the study (Trangmar et al., 1985). Geographical Information Systems (GIS) have potential for handling information on variable soil conditions at all scales (Lark and Bolam, 1997) and the standard use of GIS implies the manipulation of data layer and generation of secondary data.

The accuracy of both geostatistical and deterministic interpolation methods was analysed in several studies (Laslet et al., 1987; Weber and Englund, 1992; Gotway et al., 1996; Kravchenko and Bullock, 1999; Li et al., 2007). Laslett et al. (1987) found splines to be better than IDW and Kriging in the analysis of soil pH, while other authors have found kriging better then IDW in the analysis of some other soil properties (Gotway et al., 1996). Weber and Englund (1992) have found IDW producing better results than kriging. Many conflincting reports were found taking into consideration the use of basic statistics to predetermine both interpolation method and their parameters, and they are reported in Kravchenko and Bullock (1999), Weber and Englund (1992) and Gotway et al. (1996).

Taking into consideration the variability of the results of previous studies the objectives of this study were to: a) assses the accuracy of well known deterministic interpolation techniques, IDW, splines and local interpolation, in mapping soil topsoil and subsoil pH through the manipulation of various parameters attributable to each technique, b) choose the appropriate technique for the examination of classes of topsoil and subsoil pH on the lower sampling density datasets obtained by eliminating samplng points from the dense grid, c) predict the optimal number of samples in mapping the soil pH on a basis of a produced maps.

MATERIALS AND METHODS

The study area is vineyard at Experimental Station Faculty of Agriculture “Radmilovac”, which is located 8 km south-east from Belgrade on hilly terrain, at an altitude of 150 m. Total area of experimental field is 1.43 ha. Experimental field has rectangular shape, with a length of 140 m, and width of 102 m. According to soil taxonomy soils are defined as Rigosols. The region is characterized with temperate continental climate; air temperature for the period 1961-2001 increase for 1°C and mean annual is 11.8°C, which is in accordance with the forecast change (Vuković et al., 2009). During vegetation period precipitation average 401.7 mm and belong to “sub-humid category”. Kober 5BB is ideal rootstock drought common and tolerance to lime-based soils. An experimental field is divided in two parts; the upper part of the field is under cv. Pinot Noir, and the downer part of the field is under cv. Riesling. Total number of 104 samples for the topsoil (0-30 cm) and subsoil (30-60 cm) were taken. The sampling space within a row was 10 m, while inter-row space was 15 m. Each sampling point and borders of the field were recorded and geo-referenced by using Trimble global positioning system. Soil pH in 1 N KCl was measured with glass electrode by using pH meter method.

Soils with a pH less than 4.5 in KCl in the topsoil could be considered toxic to most crops. The effect of pH of crops is presented by soil pH classes in Tab. 1 (Cerling, 1990). Three deterministic interpolative techniques characterized with their simplicity and easy handling are used in this study.

Tab. 1 Classes of soil pH in KCl

Group	Intensity of acidity	pH in KCl
1	Very extreme	<4.5
2	Moderately acid	4.5-5.0
3	Acid	5.0-5.5
4	Acid to neutral	5.5-6.0
5	Neutral	>6.0

Inverse distance weight (IDW) is deterministic interpolation technique which implements directly the assumption that the things that are close to one another are more alike than those that are farther apart, therefore, within this interpolator a value at an un sampled location is a weighted average of known data points within surrounding neighborhood (Shepard, 1968). Measured values closer to the prediction location will have more impact than those farther away. Therefore, an assumption is that a local influence of each measured point diminishes with distance. It is an exact interpolator. The area calculated by using IDW depends on the selection of neighborhood strategy and power parameter.

Radial basis functions (RBF) methods are a series of exact interpolation techniques. It consists of five different functions each resulting in different interpolation surface. Although being exact interpolators these methods could predict values higher and lower than measured maximum and minimum values. RBFs are formed in geostatistics around each data location. Splines consist of polynomials, which describe pieces of a line or surface, and they are fitted together so that they join smoothly (Webster and Oliver, 2001).

Local Polynomial interpolation (Schaum, 2008) is an inexact interpolator that predicts a value that is different from the measured value. It presents a kind of combination of global polynomial methods and moving average procedure. Local polynomial interpolation fits many polynomials, each within specified overlapping neighborhoods. This kind of interpolation provides surfaces that accounts for more local variation. Instead of using all data, like in global polynomial interpolation, it uses data within localized windows.

Statistical analysis and spatial predictions in this paper were conducted by using Geostatistical Wizard implemented in ArcGIS 9.2 version.

RESULTS AND DISCUSSION

Data description is achieved through basic summary statistics, including means, medians, variances and skewness. A statistical summary of the pH in 1N KCl on two depths is presented in Tab. 2.

Tab. 2 Summary statistics for topsoil (0-30 cm) and subsoil (30-60 cm) pH in KCl

Soil pH	N	Min	Max	Range	Mean	Median	Var	CV (%)	Skewness	Kurtosis
Topsoil (0-30 cm)	104	3.73	7.93	4.2	5.31	4.86	1.48	23.1	0.479	1.775
Subsoil (30-60 cm)	104	3.55	7.36	3.81	5.05	4.57	1.38	23.3	0.707	2.07

Cross-validation is commonly used to validate the accuracy of interpolation (Voltz and Webster, 1990). It is achieved by eliminating information, generally one observation at a time, estimating the value at that location with the remaining data and then computing the difference between the actual and estimated value for each data location (Davis, 1987). For the comparison of different interpolation techniques, we examined the difference between the measured and predicted data by using the mean error and the root mean squared error (Robinson and Metternicht, 2006). The best cross-validation parameters for IDW, splines and local interpolation are shown in Tab. 3.

Tab. 3 Parameters returning the lowest RMSE for IDW, Splines and Local Polynomial Interpolation for top soil (0-30 cm) and subsoil (30-60 cm)

Soil pH	Power	Neighboor	ME	RMSE
Topsoil (0-30 cm)				
IDW	4	10	-0.005068	0.3461
Splines	1	10	-0.01349	0.3846
Local	2	10	0.01377	0.3625
Subsoil (30-60 cm)				
IDW	4	10	-0.0247	0.4123
Splines	1	10	-0.02135	0.431
Local	2	10	0.01069	0.4802

In all IDW tests, the best weighting parameter was found to be four. This suggests that the weights diminish rapidly from the sample point over the chosen radius. In all cases for RBFs, the best exponent value was found to be one (completely regularized spline) suggesting that lower order polynomials were sufficient at representing the variation on the field. The same neighborhood variation was used also for local interpolation.

The interpolated maps of soil pH with the lowest RMSE from the cross-validation process are presented in Fig. 1 and Fig. 2 for each method.

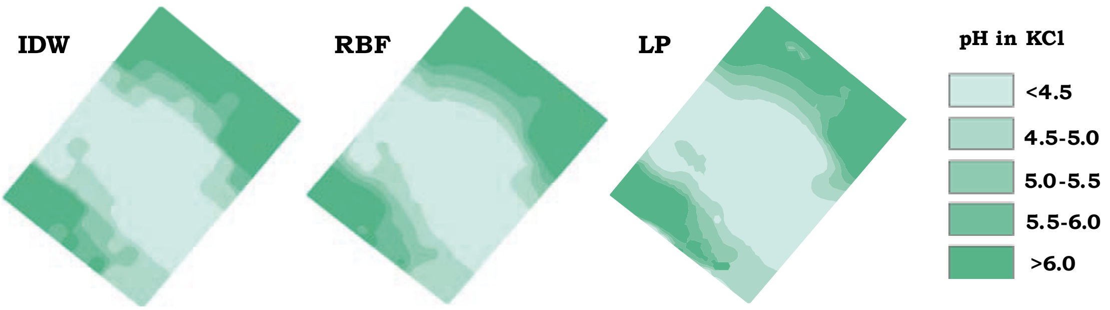

Fig. 1 Interpolated maps of pH in KCl for the depth 0-30 cm, created on a basis of the lowest RMSE obtained after cross validation by deterministic interpolative techniques.

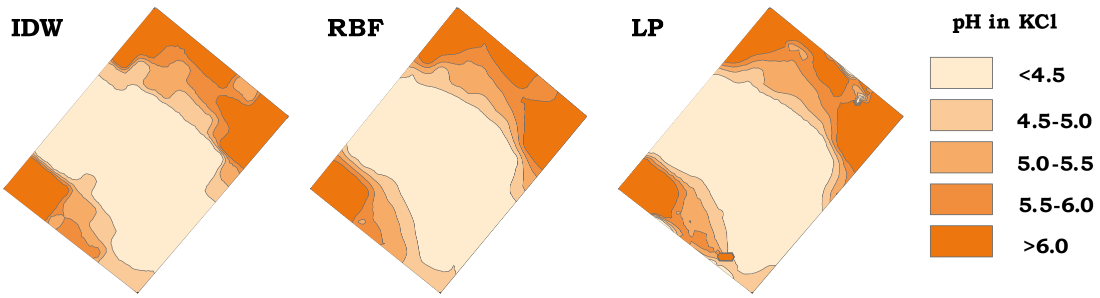

Fig. 2 Interpolated maps of pH in KCl transformed with square root transformation, for the depth 30-60 cm, created on a basis of the lowest RMSE obtained after cross validation by deterministic interpolative techniques.

Obtained results in Tab. 3 suggest IDW as the most accurate technique among three techniques used for both topsoil and subsoil pH analysis.

In further analysis, our goal was to obtain the maps of soil pH with lower number of sampling points by use of this technique. In the Tab. 4 and Tab. 5 are presented the areas for each pH class obtained after interpolation with diminished number of sampling points. Maps were produced with 54, 34, 29, 24, 19 and 14 sampling points. The simple comparison of the area of pH classes obtained with lower number of points with those pH classes obtained with maximum observation points is given as a percentage difference.

Tab. 4 The area of different pH classes in topsoil (0-30 cm) obtained with IDW for lower number of sampling points used, also expressed as a percentage difference

pH class	A104 (m^2)	A54 (m^2)	%	A34 (m^2)	%	A29 (m^2)	%	A24 (m^2)	%	A19 (m^2)	%	A14 (m^2)	%
<4.5	5783.2	5924.3	2.4	5617.3	2.9	5893.9	1.9	5431.3	6.1	5514.4	4.6	5161.0	10.8
4.5-5	2151.1	2350.7	9.3	2540.4	18.1	2331.8	8.4	2633.2	22.4	3321.8	54.4	4470.3	107.8
5-5.5	1157.4	998.5	13.7	936.5	19.1	985.5	14.9	1330.0	14.9	1357.4	17.3	733.0	36.7
5.5-6	943.8	1165.6	23.5	772.6	18.1	814.6	13.7	597.2	36.7	513.9	45.6	457.5	51.5
>6	4230.7	3827.1	9.5	4399.2	4.0	4240.2	0.2	4274.4	1.0	3558.6	15.9	3444.3	18.6

This kind of analysis was found to be not very reliable due to the fact that the obtained results expressed as a percentage of difference among classes depends on the number of sampling points included in each class, and if the number of sampling points is very low, the borders among classes will be narrower. Nevertheless, for the two major and grapevine growth and development limiting classes in our experimental field, namely, pH <4.5 and pH >6, which sum 70% of total area of topsoil, the percentage of difference seems to be quite reasonable, being lower than 10% in 54, 34, 29 and 24 sampling points maps. The map produced with 29 points is if relatively expressed more accurate than those produced with 34 and 54 points. The maps produced with 14 and 19 sampling points cannot be considered reliable.

Regarding subsoil results (Tab. 5), two major classes (pH <4.5 and pH >6), take 71% of total area, and the same behavior was obtained for 54, 34, 29 and 24 samples, meaning that 24 samples are adequate to represent these classes. The presentation of other classes is acceptable with 34 sampling points, meaning almost 1/3 of maximum.

Tab. 5 The area of different pH classes (in m^2) in subsoil (30-60 cm) obtained with IDW for lower number of sampling points used, also expressed as a percentage difference

pH class	B104 (m^2)	B54 (m^2)	%	B34 (m^2)	%	B29 (m^2)	%	B24 (m^2)	%	B19 (m^2)	%	B14 (m^2)	%
<4.5	7157.1	7055.6	1.4	7081.7	1.1	7088.2	1.0	6623.3	7.5	7102.1	0.8	7612.4	6.4
4.5-5	1635.5	1912.1	16.9	1604.5	1.9	1735.7	6.1	2107.3	28.8	2220.5	35.8	2596.6	58.8
5-5.5	1354.7	1318.6	2.7	1185.4	12.5	1320.7	2.5	1380.3	1.9	1066.4	21.3	478.3	64.7
5.5-6	1110.5	1105.9	0.4	1215.3	9.4	1460.6	31.5	1199.1	8.0	504.2	54.6	450.6	59.4
>6	3008.3	2873.9	4.5	3179.1	5.7	2660.9	11.5	2956.0	1.7	3372.9	12.1	3128.2	4.0

CONCLUSIONS

In this paper, we tried to express the estimation accuracy of interpolative methods as a relative difference in the area of pH classes obtained with lower number of points to those obtained with maximum number of sampling points. A decline in number of sampling points that we used to produce maps of soil pH indicated that adequate maps could be created with 29 samples in topsoil depth and with 34 samples in subsoil depth. The higher number of sampling points used for map creation does not seem to be an accuracy advantage if expressed in relative units.

Studies that involve soil sampling following a regular grid are limited to some hectares, and a similar approach used for larger areas is time consuming and costly. A balance must be found among the scientific objective, human and monetary resources, and time. The results of this study are very important because the optimal number of samples which is obtained to be one third of maximum for both soil depths should be a guide for further analysis on the field of ES “Radmilovac” which consists of more tenths of hectares of all type of cultivation, including 12 hectares of grapevine that should be investigated in future.

ACKNOWLEDGMENTS

This research was part of project: «Organic production of grape and wine and all grapevine products» and is partially sponsored by the Republic of Serbia, Ministry of Science, Technologies and Development under grant no TR-20093.

BIBLIOGRAPHY

Cerling, V.V. (1990). Diagnostika patini seljsko-hozalistcennvih kultur. Moskva-Leningrad: Agropromizdat.

Davis, B.M. (1987). Uses and abuses of cross-validation in geostatistics. Math. Geol. 19, 241–248.

Gotway, C.A., Ferguson, R.B., Hergert,G.W. and Peterson, T.A. (1996). Comparison of kriging and inverse-distance methods for mapping soil parameters. Am. J. Soil Sci. 60, 1237–1247.

Kravchenko, A.N. and Bullock, D.G. (1999). A comparative study of interpolation methods for mapping soil properties. J. Agronomy 91, 393–400.

Lark, R.M., and Bolam, H.C. (1997). Uncertainty in prediction and interpretation of spatially variable data on soils. Geoderma 77: 263-282.

Laslett, G.M., McBratney, A.B., Pahl, P.J. and Hutchinson, M.F. (1987). Comparison of several spatial prediction methods for soil pH. J. Soil Sci. 38, 325–341.

Li Yan, Shi Zhou, Wu Ci-fang, Li Hong-yi, and Li Feng (2007). Improved Prediction and Reduction of Sampling Density for Soil Salinity by Different Geostatistical Methods Agricultural Sciences in China, Vol. 6 (7): 832-841.

Robinson, T.P. and Metternicht, G. (2006). Testing the performance of spatial interpolation techniques for mapping soil properties. Computers and Electronics in Agriculture 50: 97–108.

Shepard, D. (1968). A two-dimensional interpolation function for irregularly-spaced data, Proc. 23rd National Conference ACM, ACM, 517-524.

Schaum, A. (2008). Principles of local polynomial interpolation, aipr, pp.1-6, 37th IEEE Applied Imagery Pattern Recognition Workshop.

Trangmar, B.B., Yost,R.S., & Uehara, G. (1985). Application of geostatistics to spatial studies of soil properties. Advancement in Agronomy, 38, 45–94.

Voltz, M., Webster, R. (1990). A comparison of kriging, cubic splines and classification for predicting soil properties from sample information. J.Soil Sci. 41, 473–490.

Vuković, A., Đurđević V., Petrović N., Sivčev B., Ranković-Vasić Z. (2009). Simulation of climate changes for Europe with special analysis for important vineyard areas of Serbia, Proceeding 32nd World Congress of Vine and Wine, pp. 47-48. www.oiv2009.hr

Weber, D. and Englund, E. (1992). Evaluation and comparison of spatial interpolators. Math. Geol. 24, 381–391.

Webster, R. and Oliver, M.A. (2001). Geostatistics for Environmental Scientists. John Wiley and Sons, Brisbane, Australia.