

Spatial variability on soil pH gradient: A case study in vineyards

F. A. Scannavino Jr.^{a,b}, A. J. Perez-Kuroki^c, A. Ghobakhlou^c, P. Sallis^c,
 S. Shanmuganathan^c and P. E. Cruvinel^{a,b}

^a Agricultural Instrumentation Centre – EMBRAPA, São Carlos, Brazil

^b Physics Institute of São Carlos - IFSC, University of São Paulo, USP, São Paulo, Brazil

Emails: scannavino@usp.br; cruvinel@cnpdia.embrapa.br

^c Geoinformatics Research Centre, Auckland University of Technology, New Zealand

Emails: jagui5@yahoo.com; akbar@aut.ac.nz; philip.sallis@aut.ac.nz; subana.shanmuganathan@aut.ac.nz

Abstract: Soil pH gradient is an indicator of microorganisms and bacteria population in root zones with effects on the growth of plants, such as grapevine (Fernandez-Calviño, 2010). The pH gradient analysis can be used to determine management strategies – root development and soil quality - to reach the most suitable balance for a specific vineyard. Custom built management strategies lead to high vineyard productivity and avoid undesirable environmental impacts caused by surplus-nutrient runoff into streams, ground water reservoirs and the micro-fauna population. However, the pH gradient in agricultural soil varies in a spatial-temporal way, making its studies difficult and time-consuming. Given this scenario, the interpolation techniques have been used to build spatial maps of soil attributes by sampled locations values. Such spatial maps give the soil condition of whole agricultural field, allowing the estimation of non-sampled locations values. A common methodology of spatial interpolation is the kriging, a popular statistical method that is grounded within the geostatistics field. Exemplifying, the Figure 1 shows the 2D maps of soil macronutrients spatial variability analyzed by kriging method (Cruvinel, 1999). The objective of this study is to evaluate the spatial variability of pH gradient in soil based on the use of geostatistical mapping obtained by pH measurements and semi-variogram models. For the analysis of the pH gradient, the soil samples were collected in 58 points considering three different horizons: ~5-15 cm, ~15-25 cm, and ~25-35 cm depths. To minimize error the soil pH gradient analyses were duplicated in the vineyard soils and at laboratory. The preliminary results have shown the existence of a significant pH gradient with values in the top layers of soil lower than the bottom layers. Therefore the pH level from top layers was more acid than the bottom ones. This study is part of “Enometrica Project”, which is an ongoing research project on micro-climate monitoring and modelling in vineyards and orchards to positively influence crop management and resulting quality production.

Keywords: vineyards, soil pH, spatial variability, geostatistical maps, pH gradient analysis

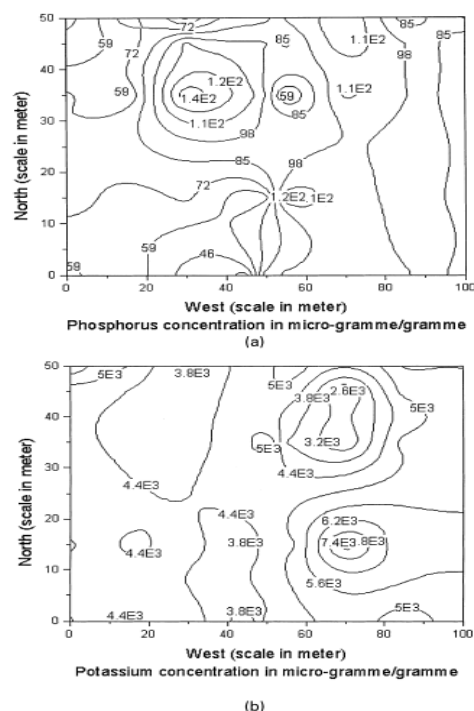


Figure 1. Illustration of spatial variability map of soil Phosphorus and Potassium concentration. Source: Cruvinel, 1999.

1. INTRODUCTION

The correlations of physical, chemical and biological attributes of soils play a key role in the optimization of growth of plants, such as grapevines. The soil pH is an indicator of nutrient availability, and its analysis determines management strategies – soil amendments - to reach the most suitable balance for a specific vineyard. Custom built management strategies lead to high vineyard productivity and avoid undesirable environmental impact caused by surplus-nutrient runoff into streams and ground waters reservoirs. However, soil pH and other nutrients vary in a spatial-temporal manner, which makes studies in this area difficult. A good way to illustrate this variation is through soil pH mappings generated by geostatistical tools. These mappings provide an estimate of the uncertainty for the soil pH value at un-sampled locations, allowing the delineation or identification of these locations for decision-making purposes or simply monitoring (Goovaerts, 2001). Currently, stochastic simulations allow the generation of several models (images) of the spatial distribution of soil attributes that are consistent with available information.

Geostatistics also provides descriptive tools such as semivariograms to characterize the spatial pattern of soil attributes categorical and continuous (Goovaerts, 1999). Several interpolation techniques, such as kriging, capitalize on the spatial correlation between observations to predict attribute values at unsampled locations using information related to one or more attributes. Knowledge of an attribute value, as soil pH level or soil compaction, for example, is therefore of little interest unless the location, time of measurement or both are considered for the analysis. In soil science, information is usually spatial and multivariate. Pioneering studies showed that the correlation between soil variables can change from one spatial scale to another, depending on the effect of fundamental physical processes.

The spatial data analysis typically starts with the input of the data values. For both continuous and for categorical attributes, the spatial distribution of values is not random. Hence the observations close to each other, as in the soil, tend to be more similar compared to those further away. The presence of a spatial structure is a prerequisite for the application of geostatistics and its description is a preliminary step to predict spatial or stochastic simulation.

The interpolation of the spatial pattern is a value estimation procedure of unsampled locations within a study area analyzed by sampled sites. Without interpolation, the study is restricted to the sampling sites and the entire surface analysis becomes unavailable. Since it is an estimation process, there is an error intrinsic to the process, becoming important its evaluation. The main interpolation advantages are the different manners in which the data could be visualized. The calculation of a certain property in a specific site, to aid in the spatial decision making process and there are several methods of analysis available for calculating the property. Therefore, for environmental remediation measures or simply for monitoring purposes, such as soil pH or soil contamination, the spatial interpolation is very useful.

The interpolation assumes a point x_0 as a reference and start comparing attribute values of the field with the values at other location, increasing the distance from the reference point. If such attribute has a positive spatial autocorrelation, in other words, the field is smooth; a value $(Z(x))$ near to reference point $(z(x_0))$ will be slightly different. To measure the amount, it is calculated by finding the difference and square it $(Z(x) - Z(x_0))^2$. It can be done to any pair of points in the field (Longley, 2010).

With the increase in distance, it is expected that the measure will probably increase too. A semivariogram shows the degree of spatial dependence among the sampling sites, which is generally an increasing monotonic (consistent) function that reaches a plateau. The 'range' is the distance at which the semivariogram reach this plateau. The value at which the semivariogram reaches the range is called 'sill' (Figure 2). Considering the distance of two points vanishingly small, the semivariance falls but there is a concept that it never quite falls to zero, even at zero distance. In other words, these two points would give different sampled values. This is known as the 'nugget' of the semivariogram and graphically this is the point where the semivariogram model intercepts the y-axis (semivariance axis). The nugget can be defined as the variation among repeated measurements at the same point (Longley, 2010).

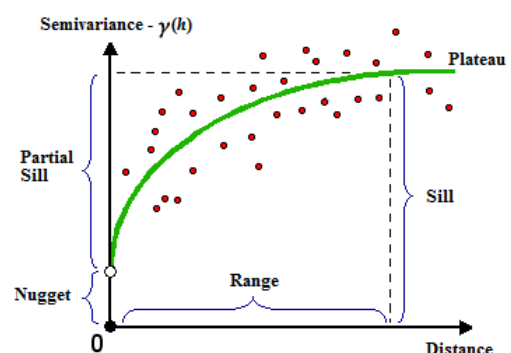


Figure 2. Graphic illustrating the definition of range, sill, and nugget components.

An interpolation problem requires some parameters of the theoretical semivariogram. Such parameters are selected from an experimental semivariogram $\gamma(h)$ which is given by:

$$\gamma(h) = \frac{1}{2N_h} \sum_{(i,j) | h_{ij}=h} [Z(j) - Z(i)]^2 \quad (1)$$

where i and j represent the location i and j , respectively, $Z(i)$ is the value of the variable at location i , h_{ij} is the vector from point i to point j , and N_h is the number of pairs of data, the locations of which are separated by h .

The experimental semivariogram analysis verifies if the available data match the assumptions for interpolation with analysis model. In this case, the chosen theoretical semivariogram model is fitted to experimental semivariogram samples generated by Equation (1).

The semivariogram is obtained without considering the directions between the points that form a pair. As such, it is said to be an isotropic variogram. Sometimes there is considerable variation in behaviour in different directions and anisotropic semivariograms are created for different ranges of direction.

Of all the common interpolation method, the kriging makes the estimation more convincing because it is grounded on the geostatistical principles. To make estimations using kriging method, it is required to reduce the semivariogram to a mathematical function, allowing the semivariance be evaluated at any distance and not only at intermediate points within distance ranges (buckets). Finally, the adjusted semivariogram is used to estimate the field values at points of interest.

The pH levels have an effect on grapevine growth and nutrition (Bates et al., 2002) and, for this reason, it is important to analyze the soil pH gradient through the layers. This paper presents a case study to assess the spatial variability of pH gradient in vineyard soil through kriging-based geostatistical interpolation techniques. This study is part of "Enometrica" Project (Ghobakhlou et al., 2010) which is a research project on micro-climate monitoring and modelling, in vineyards and orchards to positively influence crop management and resulting quality production. Monitoring is conducted through field Wireless Sensor Network (WSN) and data stored in a central database that populate computational models to optimise the vineyards production.

2. MATERIALS AND METHODS

The soil samples were collected from three different depths in 58 field locations of Kumeu River Winery, New Zealand. These depths were denominated: Horizon A (5 ~ 15 cm), Horizon B1 (15 ~ 25 cm) and Horizon B2 (25 ~ 35 cm). The analysis network was planned with a spacing between the sample points of 25±5 metres. The soil samples were extracted using a handmade auger designed by project members, with a diameter of 45 mm and to reach up to 50 cm depth. In addition, the pH level of each sample in field was measured using a portable pH meter - *Field Scout pH 110 Meter Data Logger*. For accurate locations a Garmin GPS60 was used to geo-reference these locations. After this, the samples were packed in plastic bags for later laboratory analysis. The methodology used in laboratory followed the procedures provided by the Regional Council of New Zealand for soil and land monitoring (Hill and Sparkling, 2009).

Using the field and laboratory pH database acquired, each set of data is fitted by different empirical semivariogram models. The data analysis was performed in four stages. Firstly, choose a desired mathematical semivariogram model. In this work, the models chosen were spherical, exponential and K-Bessel. These models influence the unknown values and the shape of semivariogram fitting curve. The geostatistical interpolation techniques used were ordinary kriging for the soil layers. Secondly, adjust the best anisotropy parameters to obtain the best fitting. This action is required because pH levels change not only with the distance but also with the direction. For comparison purposes, these parameter values were used for all three semivariogram models: spherical, exponential and K-Bessel. Third, evaluate the quality of the uncertainty models comparing statistical parameter values such as root-mean-square (RMS), average standard error (ASE), mean standard (MS), root-mean-square standardized (RMSS). Finally, interpolate the pH mapping with the best evaluated model.

The kriging method responds to both the proximity of the data points as well as to their direction. It also has a number of options that require the user greater attention. These features make the kriging method not a 'black box', in other words, it can not be run automatically, it requires user involvement in the estimation process. Therefore, kriging is a multistep process because it includes exploratory statistical analysis of the data, variogram modeling, creating the surface, and exploring a variance surface optionally. Its general formula is composed by a weighted sum of the data (Oliver and Webster, 1990):

$$\hat{Z}(s_0) = \sum_{i=1}^N \lambda_i Z(s_i) \quad (2)$$

where $Z(s_0)$ is the measured value at the i th location, λ_i is an unknown weight for the measured value at the i th location, s_0 is the prediction location, and N is the number of measured values. With the kriging method, the weights are based on the overall spatial arrangement of the measured points and not on the distance between the

measured points and the prediction location alone. The most common kriging types are the ordinary, the simple, the universal, the indicator and co-kriging. Due to all these features and tools, the kriging is the most used for applications in soil science and geology, mainly because in these areas there is a spatial correlation with distance or directional bias in the data. In this work, the application is to analyze the spatial variability of soil pH gradient.

The soil pH measurements can be understood as a scalar field which may assume different values in an agricultural field. However, it is possible to evaluate its variation into a soil volume that one may have in the vineyard field, based on its spatial variability values in a specific soil horizon taken at different depths. In fact, it is possible to verify the behaviour of such soil parameter variation by means of the information based on its gradient or gradient vector field. The gradient of a scalar field is a vector field that points in the direction of the greatest rate of increase of the scalar field, and whose magnitude is the greatest rate of change.

The scalar field associates a scalar value to every point in a space. The scalar may be a physical quantity. Scalar fields are required to be coordinate-independent, meaning that any two observers using the same units will agree on the value of the scalar field at the same point in space. On the other hand, in vector calculus, a vector field is an assignment of a vector to each point in a subset of Euclidean space. A vector field in the plane for instance can be visualized as an arrow, with a given magnitude and direction, attached to each point in the plane.

The elements of differential and integral calculus extend to vector fields in a natural way. The vector fields can usefully be thought of as representing the velocity of a moving pH flow in space, and this physical intuition leads to notions, which may represents the rate of change of the considered parameter, as well as the rotation of its flow.

In coordinates, the gradient as a vector field on a domain in n -dimensional Euclidean space can be represented as a vector-valued function that associates an n -tuple of real numbers to each point of the domain. This representation of a vector field depends on the coordinate system in which it is being considered.

Then the gradient can be obtained taking into account the available information supplied by the spatial variability maps obtained in three different depths of a soil horizon, i.e., horizontal information, for the vineyard field, as well as the virtual three dimensional solid and its faces, which is formed under the soil surface with such volume of soil delimited by the georeferenced coordination's of the agricultural field and its roots profile due the vineyard crop zones. Besides, by using the spatial variability information of the pH horizontal maps it is possible to visualize in the vertical direction oriented in parallel with the soil depth the information of the pH gradient by considering the three-dimensional rectangular coordinates as:

$$\nabla pH(\text{longitude}, \text{latitude}, \text{depth}) = \left(\frac{\partial pH(\text{longitude}, \text{latitude}, \text{depth})}{\partial \text{longitude}} \hat{i} + \frac{\partial pH(\text{longitude}, \text{latitude}, \text{depth})}{\partial \text{latitude}} \hat{j} + \frac{\partial pH(\text{longitude}, \text{latitude}, \text{depth})}{\partial \text{depth}} \hat{k} \right) \quad (3)$$

where \hat{i} , \hat{j} , and \hat{k} are the standard unit vectors.

The gradient can also be used to measure how a scalar field changes in other directions, rather than just the direction of greatest change, by taking a dot product.

Thus, the gradient function of the soil pH is defined to be the vector field whose components are its partial derivatives. That is, the gradient is written as a row vector, but it is often taken to be a column vector. When a function also depends on a parameter such as time, the gradient often refers simply to the vector of its spatial derivatives only.

3. RESULTS AND DISCUSSION

Using ordinary kriging for the layers: Horizon A, Horizon B1 and Horizon B2, the anisotropy values used in interpolation are presented in Table 1. Those angle direction values were chosen by the best model fitting and adjustment based on the best choice of the interpolation model, as shown in Figure 4. This is a feature of the ordinary kriging where it is possible to adjust the angle

Table 1. Anisotropy Parameters adjusted for the field pH mapping.

Database	Horizon	Angle direction	Angle tolerance	Bandwidth (lags)
Field	A	319°	45°	3
	B1	141°	45°	3
	B2	285°	45°	3
Laboratory	A	351°	45°	3
	B1	331°	45°	3
	B2	126°	45°	3

direction to the semivariogram when the anisotropy option was chosen. It is important to highlight that the different angle directions for the same horizons did not show similar pH values obtained by field and laboratory data collection. After performing these steps, the results were generated considering three different semivariogram models, which are presented in Table 2. The choice criterion of the most suitable mathematical model was:

- mean error near zero and the smallest RMS;

- b) the ASE nearest the RMS;
- c) the standard error (RMSS) nearest one.

These criteria were based on the fact that although the minimum RMS leads to conclude that the method is great but it is not possible to affirm that it is a valid method. When an unknown point prediction is performed, there is only the estimated standard error to assess the uncertainty of that prediction. On the other hand, when the ASE is close to RMS from the cross-validation results, the prediction standard error is appropriated. Therefore, the latter had a greater weight in the decision. The method and the value chosen are represented by the red colours within Table 2.

Table 2. Statistical values from three different mathematical methods: spherical, exponential and K-Bessel.

Database	Horizon	Method	Mean	RMS	ASE	RMSS
Field	A	Spherical	0,0089	0,3365	0,3011	1,0970
		Exponential	0,0133	0,3444	0,3036	1,1120
		K-Bessel	0,0069	0,3450	0,2932	1,1580
	B1	Spherical	0,0068	0,3601	0,3638	0,9843
		Exponential	0,0093	0,3591	0,3772	0,9452
		K-Bessel	0,0066	0,3605	0,3611	0,9916
	B2	Spherical	0,0082	0,4469	0,4237	1,0510
		Exponential	0,0080	0,4461	0,4273	1,0420
		K-Bessel	0,0076	0,4478	0,4216	1,0580
Laboratory	A	Spherical	0,0006	0,1622	0,1636	0,9936
		Exponential	0,0004	0,1644	0,1687	0,9759
		K-Bessel	0,0002	0,1620	0,1614	1,0070
	B1	Spherical	0,0007	0,1795	0,1705	1,0450
		Exponential	0,0006	0,1842	0,1780	1,0290
		K-Bessel	0,0012	0,1827	0,1746	1,0330
	B2	Spherical	0,0085	0,3102	0,3281	0,9542
		Exponential	0,0095	0,3124	0,3305	0,9542
		K-Bessel	0,0094	0,3111	0,3277	0,9581

Table 2 has shown that K-Bessel method was the best for horizon B1, the Exponential method for horizon B2, the spherical for horizon A, all from field database. For laboratory database, the Spherical method was the best for horizons B1 and B2, and K-Bessel for horizon A. With the respective chosen methods, the pH mapping from field and laboratory database is illustrated by Figures 3, 5 and 6, respectively.

When comparing the results generated by field and laboratory experiments (Figures 3, 5 and 6), it is possible to note a greater pH variation from the laboratory data in relation to one obtained on the field. This behaviour was expected because on the field the measurements were collected with soil in its natural conditions, such as temperature and moisture. In addition, the measures were obtained from direct contact of the soil and the sensor while in laboratory the samples were processed on the same way (dry, grind and measure), then dissolved in water before measuring the pH level. The modelling analysis has shown to be adequate and important to give more feasibility to prediction mapping.

The next stage was to analyse graphically the soil pH gradient seen by the aspect of depth. In other words, the soil pH variation was observed on the vertical plane. For this purpose, two faces of the Kumeu vineyard in two different slices: one in the face of the volume of soil under the vineyard crop and other at the intern of such soil volume, were studied. The field data collection was chosen for the analysis due the pH value on the soil natural condition. For this initial analysis the ordinary kriging with the K-Bessel method was used. The Figures 6 and 7 illustrate the results of soil pH gradient.

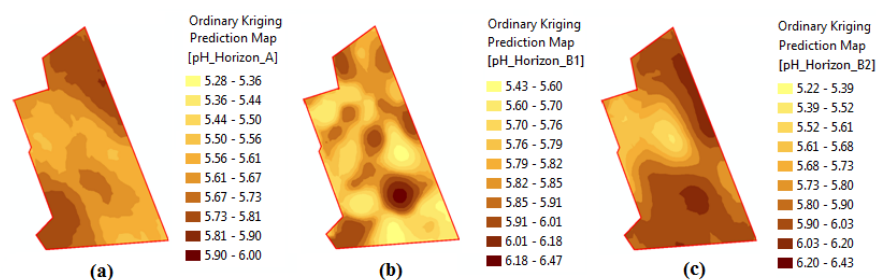


Figure 3. pH mapping from the laboratory data: (a) Horizon A; (b) Horizon B1 and (c) Horizon B2

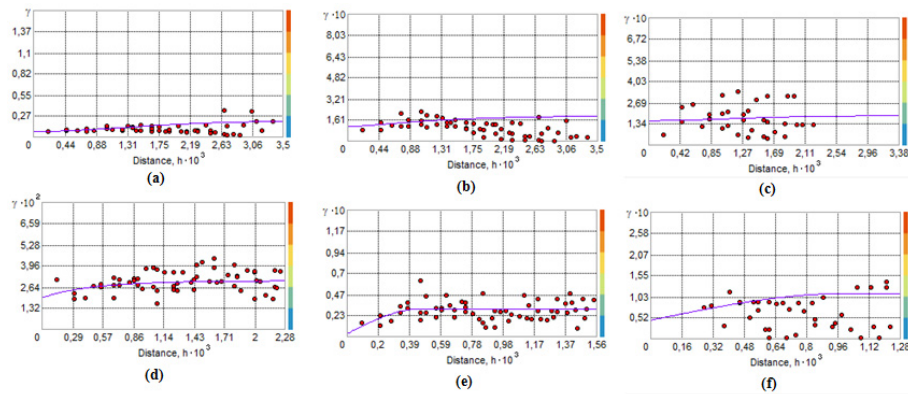


Figure 4. Kriging Semivariograms - Field data: (a) Horizon A, (b) Horizon B1, (c) Horizon B2; Laboratory data: (d) Horizon A, (e) Horizon B1, (f) Horizon B2

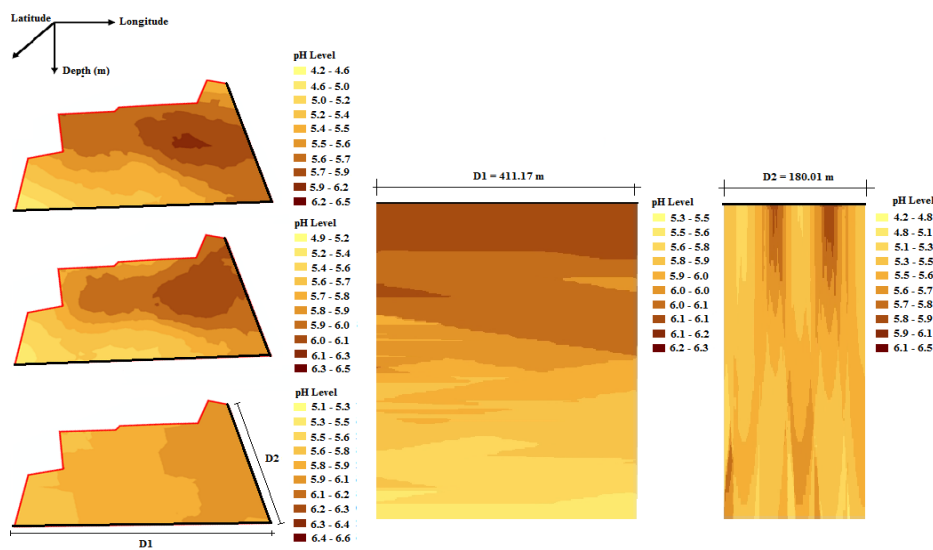


Figure 5. Kriging maps of soil pH gradient on the two Kumeu external faces. The faces are illustrated by black colour line.

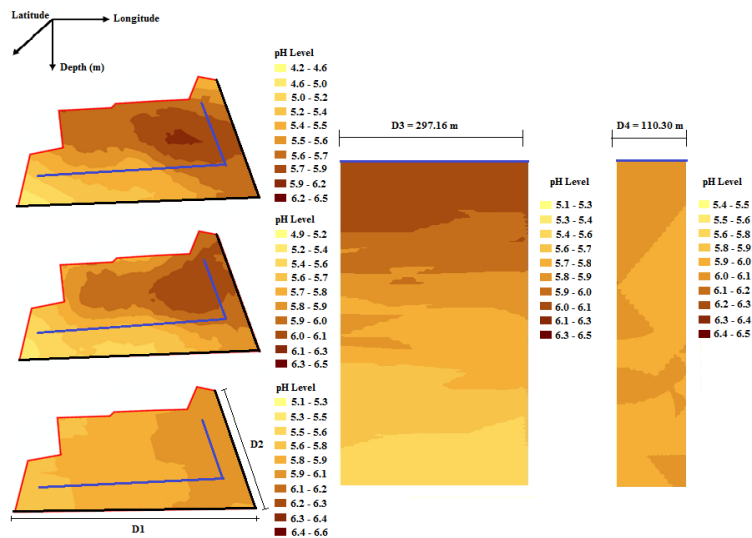


Figure 6. Kriging maps of soil pH gradient on the two Kumeu internal faces. The faces are illustrate by blue colour line

4. Conclusion

The results showed a slight difference between the soil layers. The analysis of locations with two different depths shows that the top layers have lower pH levels (i.e. more acidic) than bottom layers, in general the top layer is more acidity. Due the leaves emanating from pruning vines to decompose in the upper layer of the soil causes the pH of the upper soil layers have a more acidic profile in relation to the lower layers. Probably, in the course of time, due to sorption processes between the organic fertilizer and the volume of soil below surface, the acidity is being increasing and one may reached a soil pH gradient profile in the Kumeu Vineyards.

These initial results will be used in the “Enometrica Project” together with chemical soil analysis and other information as climate, plant and productivity database to complete decision making support for vineyards productivity.

ACKNOWLEDGMENTS

The authors would like to thank the Brazilian Enterprise for Agricultural Research (Embrapa), the Auckland University of Technology (AUT), and the Physics Institute of the University of São Paulo in São Carlos (IFSC) for the support in the Precision Agriculture Network (Macro Program 1, Process: 01.09.01.002.01), CNPq (Process: 306988/2007-0) and FAPESP (Process: 09/50405-1). Besides, they also thank Mr. Michael Brajkovich, owner and Master of Wine from Kumeu River Winery.

REFERENCES

- Bates T.R., Dunst R.M., Taft T. and Vercant M. (2002). The vegetative response of ‘Concord’ grapevines to soil Ph. *HortScience*, vol.37, n.6.
- Cruvinel P.E., Flocchini R.G., Artaxo P., Crestana S., Herrmann Jr P.S.P. (1999), Elemental analysis of agricultural soil samples by particle induced X-ray emission (PIXE) technique. *Nuclear Instruments and Methods in Physics Research B*, v.150, pp. 478-483.
- Fernandez-Calviño D., Martin A., Arias-Estevez M., Baath E., Diaz-Raviña M. (2010). Microbial community structure of vineyard soils with different pH and copper content. *Applied Soil Ecology*, v.46, pp. 276–282.
- Ghobakhlou A., Perera A., Sallis P. and Zandi S. (2010), Modular Sensor Nodes for Environmental Data Monitoring. *Proceedings of the Fourth International Conference on Sensing Technology*, University of Salento, Lecce, pp. 372-377 ISBN 978-0-473-16942-8, Italy June 3 – 5.
- Goovaerts P.. (2001) Geostatistical modelling of uncertainty in soil science. *Geoderma*, vol.103, pp.3-26.
- Goovaerts P.. (1999). Geostatistics in soil science: state-of-the-art and perspectives. *Geoderma*, v.89, pp. 1–45.
- Hill R. B. and Sparling G. P. (2009) Chapter 3: Soil quality monitoring. I *Land and Soil Monitoring: A guide for SoE and regional council reporting*. Published by the Land Monitoring Forum, New Zealand. [Online]. Available: http://www.envirolink.govt.nz/PageFiles/31/Land%20and%20soil%20monitoring_A_guide_for_SoE%20and%20regional%20council%20reporting.PDF.
- Longley P.A., Goodchild M.F., Maguire D.J. and Rhind D.W. (2010). *Geographic information systems & science*. Wiley Publisher, third edition.
- Oliver M.A., Webster R. (1990) Kriging: a method of interpolation for geographical information systems. *International Journal of Geographical Information Science*, v.4, pp. 313–332.
- Vieira S.R. (2000) *Geoestatística em Estudos de Variabilidade Espacial do Solo. Tópicos em Ciência do Solo*, Sociedade Brasileira de Ciência, Viçosa, MG, Brazil, pp. 1–54.