

Remaining Phosphorus Estimate Through Multiple Regression Analysis*¹

M. E. ALVES and A. LAVORENTI*²

Departamento de Ciências Exatas, Escola Superior de Agricultura 'Luiz de Queiroz'—ESALQ/USP, Piracicaba (SP) 13418-900 (Brazil). E-mail: mealves@esalq.usp.br.

(Received April 24, 2006; revised July 8, 2006)

ABSTRACT

The remaining phosphorus (P_{rem}), P concentration that remains in solution after shaking soil with $0.01 \text{ mol L}^{-1} \text{ CaCl}_2$ containing $60 \mu\text{g mL}^{-1} \text{ P}$, is a very useful index for studies related to the chemistry of variable charge soils. Although the P_{rem} determination is a simple procedure, the possibility of estimating accurate values of this index from easily and/or routinely determined soil properties can be very useful for practical purposes. The present research evaluated the P_{rem} estimation through multiple regression analysis in which routinely determined soil chemical data, soil clay content and soil pH measured in $1 \text{ mol L}^{-1} \text{ NaF}$ (pH_{NaF}) figured as P_{rem} predictor variables. The P_{rem} can be estimated with acceptable accuracy using the above-mentioned approach, and pH_{NaF} not only substitutes for clay content as a predictor variable but also confers more accuracy to the P_{rem} estimates.

Key Words: anionic adsorption, pH_{NaF} , pedotransfer functions, tropical soils

The remaining phosphorus (P_{rem}) consists of the P concentration that remains in solution after shaking a soil sample with $0.01 \text{ mol L}^{-1} \text{ CaCl}_2$ containing $60 \mu\text{g mL}^{-1} \text{ P}$. Although Netto and Barbosa (2005) showed that the use of a soil:solution ratio of 1:50 should be more convenient, the current method used in Brazil for P_{rem} determination consists of shaking 2.5 g of soil for 1 hour with 25 mL of $0.01 \text{ mol L}^{-1} \text{ CaCl}_2$ containing $60 \mu\text{g mL}^{-1} \text{ P}$, filtering or centrifuging the suspension and analyzing the equilibrium solution for P (Alvarez *et al.*, 2000).

The significance of P_{rem} for soils presenting low P levels and high P sorption is ascribed not only to its simple determination but mainly to its strong correlation with the less easily measurable phosphorus buffer capacity (PBC) of such soils (Novais and Smyth, 1999).

The soil PBC has an inverse relationship with the plant efficiency of P utilization and with the effectiveness of PBC-sensitive extractants such as Melich-1 for extracting elements such as P and Zn (Muniz *et al.*, 1985; Couto *et al.*, 1992). This gives rise, for instance, to different P, S and Zn critical levels in both plant and soil for a same culture when they are determined in different soils using PBC-sensitive extractants (Muniz *et al.*, 1985; Alvarez and Fonseca, 1990; Alvarez *et al.*, 2000). Therefore, it is possible to establish regression equations where the soil P and Zn critical levels determined with Mehlich-1 and the soil S critical level evaluated with $\text{Ca}(\text{H}_2\text{PO}_4)_2$ $500 \text{ mg mL}^{-1} \text{ P}$ in $2 \text{ mol L}^{-1} \text{ HOAC}$ figure as P_{rem} dependent variables in order to use such models for better interpreting the results of soil chemical analyses for fertility evaluation purposes (Alvarez *et al.*, 2000). Furthermore, as shown by Novais and Smyth (1999) P_{rem} can be used as an input variable of computational systems comprising mechanistic models that account for the several factors that define the response of a culture to fertilization in order to better recommend fertilizer rates.

The dependence of P_{rem} on the soil clay mineralogy (Alves and Lavorenti, 2004) is responsible for its great variability when determined in soils having different clay component contents. For this reason,

*¹Project supported by the State of São Paulo Research Foundation—FAPESP, Brazil (No. 98/01502-8).

*²Corresponding author. E-mail: alavoren@esalq.usp.br.

P_{rem} is more adequate than clay content for stratifying soils in relation to their phosphate and sulfate adsorption capacities. This aspect is very useful for defining the P and S levels to be added to soils for the establishment of P and S sorption isotherms and for the determination of soil P and S critical levels in greenhouse experiments (Alvarez and Fonseca, 1990; Alvarez *et al.*, 2000). Furthermore, P_{rem} can also be used as an ancillary variable for defining both the soil lime and gypsum requirements (Alvarez *et al.*, 2000). Nowadays, the inclusion of P_{rem} among the chemical analyses used for soil classification purposes is being considered for the next version of the Brazilian System of Soil Classification (EMBRAPA, 1999; Netto and Barbosa, 2005).

Although the laboratory determination of P_{rem} is simple, it is interesting to evaluate the possibility of estimating accurate values of it from more easily and/or routinely determined soil properties. This approach can make the P_{rem} determination faster and cheaper, which is very desirable when a large number of soil samples need to be analyzed. Therefore, the present study evaluated the possibility of estimating P_{rem} through multiple linear regression analysis using routinely determined soil chemical data, soil clay content and soil pH measured in 1 mol L⁻¹ NaF as P_{rem} predictor variables.

MATERIALS AND METHODS

Subsurface soil samples of representative soils in São Paulo State, Brazil, were collected, air-dried and passed through a 2-mm sieve (Table I). The texture (pipette method), organic carbon (Walkley-Black method) and exchangeable Al (1 mol L⁻¹ KCl extraction) were determined according to the methods described by EMBRAPA (1997). The plant-available P, K, Ca and Mg contents were evaluated by extraction with ion-exchange resin (Van Raij *et al.*, 1986). The potential acidity (H + Al) was determined using the Shoemaker-McLean-Pratt buffer solution (SMP pH) method (Quaggio *et al.*, 1985) and plant-available S content was determined by BaSO₄ turbidimetry after extraction with 0.01 mol L⁻¹ Ca(H₂PO₄)₂·H₂O (Van Raij *et al.*, 1996). Soil organic matter content was estimated by multiplying the organic carbon content by 1.72; the sum of bases (SB) was calculated as SB = Ca + K + Mg; the cation exchange capacity (CEC) at pH 7 was given by SB + (H + Al); the base saturation (V) was calculated as 100 × SB / (SB + Al) and the aluminum saturation (m) corresponded to 100 × Al / (SB + Al) (EMBRAPA, 1997). Except for soil texture, all above-mentioned determinations were carried out on a volume basis (dm⁻³) as done in most Brazilian laboratories whose analyses are directed to soil fertility evaluation.

TABLE I

Classification, parent materials, locations and sampling depths of the soils in the São Paulo State

Soil	Classification ^{a)}	Parent material	Location	Sampling depth
				cm
1	Rhodic Acrudox	Basalt	Ribeirão Preto	100–140
2	Rhodic Eutrudox	Basalt	Iracemápolis	100–110
3	Rhodic Hapludox	Basalt	Luís Antonio	80–100
4	Rhodic Acrudox	Basalt	Luís Antonio	150–170
5	Xanthic Acrustox	Basalt	Guafra	100–130
6	Typic Hapludox	Schist	Piracicaba	100–110
7	Typic Hapludox	Sandstone	Piracicaba	100–110
8	Typic Hapludox	Sandstone	São Carlos	80–100
9	Typic Hapludalf	Diabase	Piracicaba	30–40
10	Typic Hapludult	Sandstone	Pindorama	100–120
11	Typic Hapludult	Sandstone	Vera Cruz	100–120
12	Typic Hapludult	Schist	Rio Claro	70–80
13	Typic Hapludult	Basalt	Piracicaba	100–110
14	Typic Quartzipsamment	Sandstone	São Pedro	80–100

^{a)} According to U.S. Soil Taxonomy (1998).

Soil pH values were measured in water (pH_{H₂O}), 0.01 mol L⁻¹ CaCl₂ (pH_{CaCl₂}) and 1 mol L⁻¹ KCl

(pH_{KCl}) at 1:2.5 soil/water or solution ratio (EMBRAPA, 1997). The ΔpH index was calculated as $\Delta\text{pH} = \text{pH}_{\text{KCl}} - \text{pH}_{\text{H}_2\text{O}}$ (Mekaru and Uehara, 1972). Additionally, the soil pH was also measured after shaking 0.5 g of soil for 1 hour with 20 mL of 1 mol L⁻¹ NaF (pH_{NaF}) (Bolland *et al.*, 1996).

The P_{rem} was determined according to Alvarez *et al.* (2000) by shaking triplicate 2.5 g of soil with 25 mL of 0.01 mol L⁻¹ CaCl₂ containing 60 mg mL⁻¹ P, filtering the solution through a Whatman 42 filter and analyzing the solutions for P using the Murphy and Riley (1962) method.

Multiple linear regression analysis was carried out with the Statistical Analysis System (SAS) software for establishing a function in which P_{rem} figured as dependent variable on the other determined soil properties. The predictor variables used in this initial model were defined employing the maximum R^2 improvement (MAXR) selection procedure (SAS, 1989), which selected the soil properties that together with clay content and pH_{NaF} maximized the coefficient of determination (R^2) of the regression model. After fitting the initial model, the influences of clay content and pH_{NaF} on the accuracy of the P_{rem} estimates were verified through the gradual removal of these variables from the regression equation. In addition to R^2 value, the F -test value calculated in the analysis of variance of the regression model, the residue analysis and the comparison of measurable and predicted values were considered in the choice of a better equation for estimating P_{rem} .

RESULTS AND DISCUSSION

Soil properties

The soils (Table II) are mostly clayey (clay > 350 g kg⁻¹), acidic ($\text{pH}_{\text{CaCl}_2} \leq 5$) (EMBRAPA, 1999), poor in P ($P < 6 \text{ mg dm}^{-3}$) and rich in S ($S > 10 \text{ mg dm}^{-3}$) (Van Raij *et al.*, 1996). All soils, except soils 1, 4 and 5, had $\Delta\text{pH} < 0$ which indicates that they have a negative net surface charge (Mekaru and Uehara, 1972).

TABLE II

Selected properties of the soils studied

Soil	Clay	pH^{a}		$\Delta\text{pH}^{\text{b}}$	$P_{\text{rem}}^{\text{c}}$	OM ^{d)}	P	S	K	Ca	Mg	Al	H+Al ^{e)}	SB ^{f)}	CEC ^{g)}	$m^{\text{h)}$	$V^{\text{i)}$
		CaCl ₂	NaF														
	g kg ⁻¹				g mL ⁻¹	g dm ⁻³	mg dm ⁻³						cmol _c dm ⁻³				%
1	529	5.0	10.4	0.11	1.0	10.4	50	5	0.05	0.4	0.1	0	3.1	0.55	3.7	0	15
2	481	5.5	10.3	-0.77	3.1	13.4	50	37	0.02	4.3	0.4	0	2	4.72	6.7	0	70
3	543	4.3	10.3	-0.08	4.1	7.8	30	2	0.03	0.3	0.1	0.1	3.8	0.43	4.2	14.6	10
4	647	5.4	10.7	0.36	0.3	8.4	30	1	0.11	0.4	0.1	0	2.5	0.61	3.1	0	20
5	324	5.6	10.4	0.87	0.1	3.5	30	17	0.21	0.4	0.1	0	1.8	0.71	2.5	0	28
6	774	4.9	10.3	-0.83	3.8	12.8	20	109	0.05	1.5	0.8	0	2.8	2.35	5.2	0.3	46
7	244	4.8	9.3	-0.94	36.1	3.6	20	35	0.15	0.8	0.2	0	1.6	1.15	2.8	0	42
8	421	4.8	10.3	-0.42	7.7	4.1	30	78	0.01	0.7	0.3	0	2.2	1.01	3.2	0	31
9	597	5.0	9.8	-0.63	5.3	3.1	40	102	0.03	4.0	0.3	0	2.2	4.33	6.5	0	66
10	347	4.0	9.5	-1.12	21.8	2.6	20	15	0.12	1.0	1.3	0.8	4.7	2.42	7.1	25.3	34
11	205	3.7	8.8	-0.76	36.1	3.6	20	22	0.09	0.4	0.2	0.7	2.8	0.69	3.5	48.9	20
12	475	5.2	9.7	-0.81	2.1	5.2	30	7	0.05	6.3	1.7	0	2.5	8.05	10.6	0	76
13	632	3.8	10.0	-1.12	1.5	4.1	20	137	0.09	0.7	0.4	2.5	7.2	1.19	8.4	68.1	14
14	11	4.0	9.0	-0.47	43.0	1.0	40	47	0.03	0.6	0.1	0.1	2.2	0.73	2.9	10.9	25

^{a)} Soil pH values were measured in CaCl₂ ($\text{pH}_{\text{CaCl}_2}$) and NaF (pH_{NaF}) solutions, respectively; ^{b)} $\Delta\text{pH} = \text{pH}_{\text{KCl}} - \text{pH}_{\text{H}_2\text{O}}$; ^{c)} Remaining phosphorus; ^{d)} Organic matter; ^{e)} Potential acidity; ^{f)} Sum of bases, SB = K + Ca + Mg; ^{g)} Cation exchange capacity (CEC) at pH 7, CEC = K + Ca + Mg + (H + Al); ^{h)} Aluminum saturation (m), $m = 100 \times \text{Al}/(\text{K} + \text{Ca} + \text{Mg} + \text{Al})$; ⁱ⁾ Base saturation (V), $V = 100 \times (\text{K} + \text{Ca} + \text{Mg})/(\text{K} + \text{Ca} + \text{Mg} + \text{Al})$.

The pH_{NaF} values have a coefficient of variation (CV) of nearly 6 % in the evaluated soil set, ranging from 8.8 (soil 11) to 10.7 (soil 4). On the other hand, the P_{rem} values have a wide variation characterized by a CV of 130 %. This high variability observed for P_{rem} can be ascribed to the differences of their

clay component contents as discussed by Alves and Lavorenti (2004).

Remaining phosphorus estimate from multiple linear regression analysis

The results of the multiple regression analyses are presented in Table III. The first equation (Eq. 1) used the following variables as P_{rem} predictors: pH_{CaCl_2} , organic matter content (OM), potential acidity (H + Al), aluminum saturation (m), base saturation (V), available S (S), clay content (CLAY), available P (P), exchangeable aluminum (Al), cation exchange capacity at pH 7 (CEC), ΔpH and pH_{NaF} . From this equation three others were fitted: Eq. 2) removing CLAY; Eq. 3) removing pH_{NaF} and Eq. 4) removing both CLAY content and pH_{NaF} .

TABLE III

Results of the multiple linear regression analyses relating remaining phosphorus to soil properties

Parameters	Equation ^{a)}			
	1	2	3	4
a_0	64.2042	83.4299	111.6418	109.2295
a_1	24.6685	20.6917	-14.1579	-13.7079
a_2	-0.9701	-0.9135	-0.4135	-0.4199
a_3	9.4610	7.7090	-6.7810	-6.5810
a_4	0.6069	0.4923	-0.2154	-0.2020
a_5	-0.0466	-0.0271	0.1927	0.1906
a_6	-0.0468	-0.0609	-0.1477	-0.1461
a_7	-0.0039	0	-0.6006	0
a_8	-0.6846	-0.5359	0.0005	-0.6197
a_9	-24.666	-19.320	17.233	16.611
a_{10}	-5.0890	-5.0220	-3.7900	-3.7960
a_{11}	-23.0136	-21.7292	-13.5513	-13.7018
a_{12}	-17.1478	-16.9246	0	0
$R^{2b)}$	0.9862	0.9919	0.8917	0.9278
$F^c)$	78.2	146.0	10.7	17.7

^{a)} $P_{rem} = a_0 + a_1 pH_{CaCl_2} + a_2 OM + a_3 (H + Al) + a_4 m + a_5 V + a_6 S + a_7 CLAY + a_8 P + a_9 Al + a_{10} CEC + a_{11} \Delta pH + a_{12} pH_{NaF}$ where: P_{rem} = remaining phosphorus ($\mu g mL^{-1}$); pH_{CaCl_2} = pH measured in $0.01 mol L^{-1} CaCl_2$; OM = organic matter ($g dm^{-3}$); (H + Al) = potential acidity ($cmol_c dm^{-3}$); m = aluminum saturation (%); V = bases saturation (%); S = available sulfur ($mg dm^{-3}$); CLAY = clay content ($g kg^{-1}$); P = available phosphorus ($mg dm^{-3}$); Al = extractable aluminum ($cmol_c dm^{-3}$); CEC = cation exchange capacity at pH 7 ($cmol_c dm^{-3}$); $\Delta pH = pH_{KCl} - pH_{H_2O}$ where pH_{KCl} and pH_{H_2O} are pH values measured in $1 mol L^{-1} KCl$ and in H_2O , respectively; $pH_{NaF} = pH$ measured in $1 mol L^{-1} NaF$.

^{b)} R^2 = coefficient of determination.

^{c)} F = F -test value (relationship between the P_{rem} variations explained and not explained by the regression model).

Considering the F and R^2 values, Eq. 2 was the most efficient for estimating accurate P_{rem} values. Thus, due to its easier and cheaper determination, pH_{NaF} can effectively substitute for clay content as a P_{rem} predictor variable in a multiple linear regression model. The P_{rem} estimates given by Eq. 2 were not only similar to the measured values (Fig. 1) but also showed the same variation as the latter ones (CV = 130%). The poor influence of clay content on P_{rem} can also be observed from Fig. 1c and d, which are almost identical. In addition, the residue analysis (Fig. 2) showed that the clay content could be substituted by pH_{NaF} as predictor variable in multiple regression models fitted for soils presenting different clay mineralogies in order to obtain more accurate P_{rem} estimates. When the residues (P_{rem} measured - P_{rem} estimated) obtained with Eqs. 2 and 3 (Fig. 2b and c) are compared, it can be seen that the estimates given by the first one were much more accurate. Again, the poor influence of clay content on P_{rem} was revealed by the similarity of the Fig. 2c and d.

The greater capacity shown by pH_{NaF} in relation to the clay content for estimating P_{rem} is basically due to the consistent correlation between pH_{NaF} and P_{rem} , which is caused by their very similar dependencies on the mineralogical clay composition (Alves and Lavorenti, 2004). Therefore, despite the

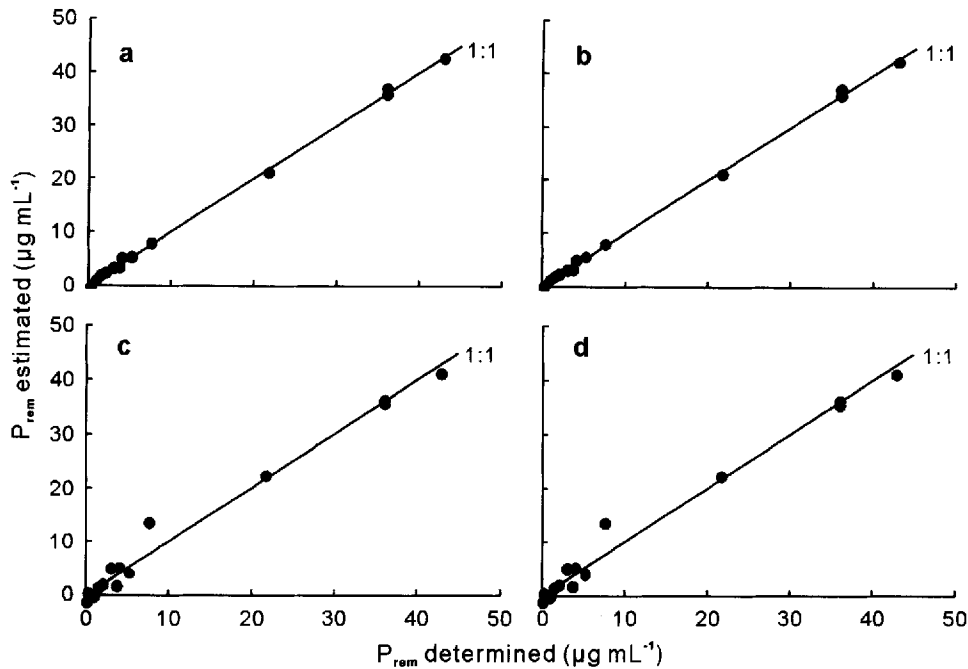


Fig. 1 Relationships between the determined and estimated remaining phosphorus (P_{rem}) values. (a) Estimates obtained with all predictor variables in the initial model; (b) estimates obtained with all predictor variables in the initial model except clay content; (c) estimates obtained with all predictor variables in the initial model except pH_{NaF} ; (d) estimates obtained with all predictor variables in the initial model except clay content and pH_{NaF} .

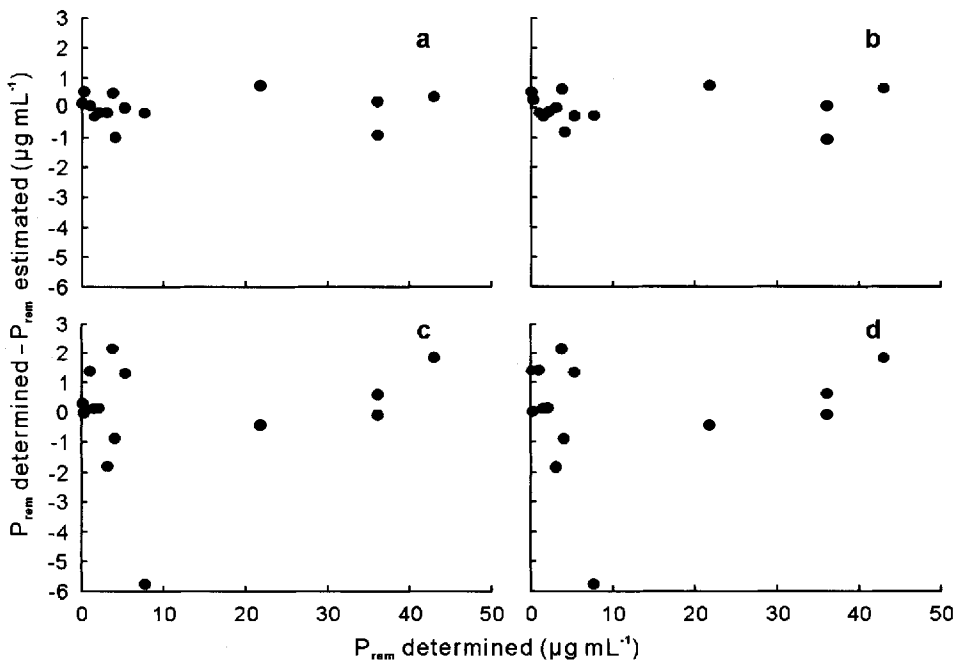


Fig. 2 Relationships between the determined values of the remaining phosphorus (P_{rem}) and the differences between determined and estimated P_{rem} values. (a) Estimates obtained with all predictor variables in the initial model; (b) estimates obtained with all predictor variables in the initial model except clay content; (c) estimates obtained with all predictor variables in the initial model except pH_{NaF} ; (d) estimates obtained with all predictor variables in the initial model except clay content and pH_{NaF} .

logarithmic scale of pH_{NaF} limits its utility for stratifying soils for their phosphate and sulfate sorption capacities, the higher sensitivity of pH_{NaF} to the clay mineralogy as well as its faster and cheaper laboratory determination justify its use as a predictor variable for the indirect P_{rem} determination from multiple linear regression models.

The present study has shown that the fit of a multiple linear regression model based on easily/routinely determined soil properties is a viable approach for P_{rem} estimates and that pH_{NaF} is a key variable for obtaining accurate P_{rem} values; however additional factors need to be considered before its large scale use. Although pedotransfer functions could be developed using the same predictor variables used in Eq. 2 (Table III), it is imperative that a larger number of soil samples collected from a wide range of soil classes under different management practices be considered in the development of more comprehensive models. Also the possibility of measuring pH_{NaF} values after shaking the soil with 1 mol L^{-1} NaF for only 2 minutes (Fieldes and Perrott, 1966; Singh and Gilkes, 1991) should be investigated. These aspects should be evaluated in future research.

CONCLUSIONS

Accurate P_{rem} values can be estimated quickly and with lower costs with multiple linear regression models fitted using routinely determined soil chemical data and the easily measurable pH_{NaF} that not only substitutes for clay content in such models but also confers more accuracy to the P_{rem} estimates.

REFERENCES

- Alvarez, V. H. V. and Fonseca, D. M. 1990. Definition of phosphorus levels for determining the maximum capacity of phosphate adsorption and for greenhouse experiments. *Brazilian Journal of Soil Science* (in Portuguese). **14**: 49–55.
- Alvarez, V. H. V., Novais, R. F., Dias, L. E. and Oliveira, J. A. 2000. Determination and use of the remaining phosphorus. *Informative Bulletin of the Brazilian Soil Science Society* (in Portuguese). **25**: 24–32.
- Alves, M. E. and Lavorenti, A. 2004. Remaining phosphorus and sodium fluoride pH in soils with different clay contents and clay mineralogies. *Brazilian Journal of Agricultural Research*. **39**: 241–246.
- Bolland, M. D. A., Gilkes, R. J., Brennan, R. F. and Allen, D. G. 1996. Comparison of seven phosphorus sorption indices. *Australian Journal of Soil Research*. **34**: 81–89.
- Couto, C., Novais, R. F., Teixeira, J. L. and Barros, N. F. 1992. Zinc critical levels in soil and plant for maize growing in soils presenting different values of capacity factor. *Brazilian Journal of Soil Science* (in Portuguese). **16**: 79–87.
- EMBRAPA. 1997. Manual of Methods of Soil Analysis (in Portuguese). Ministry of Agriculture, Rio de Janeiro. 212pp.
- EMBRAPA. 1999. Brazilian System of Soil Classification (in Portuguese). Ministry of Agriculture, Rio de Janeiro. 412pp.
- Fieldes, M. and Perrott, K. W. 1966. The nature of allophane in soils. Part 3. Rapid field and laboratory test for allophane. *New Zealand Journal of Science*. **9**: 623–629.
- Mekaru, T. and Uehara, G. 1972. Anion adsorption in ferruginous tropical soils. *Soil Science Society of America Proceedings*. **36**: 296–300.
- Muniz, A. S., Novais, R. F., Barros, N. F. and Neves, J. C. L. 1985. Critical level of phosphorus for soybean as dependent variable on the soil buffering phosphorus capacity. *Brazilian Journal of Soil Science* (in Portuguese). **9**: 237–243.
- Murphy, J. and Riley, J. P. 1962. A modified single solution method for the determination of phosphate in natural waters. *Analytica Chimica Acta*. **27**: 31–36.
- Netto, A. R. and Barbosa, J. S. F. 2005. Methodological adjustment for remaining phosphorus and for the efficient prediction of the maximum capacity of phosphate adsorption. In Ribeiro, M. R. (ed.) 30th Brazilian Congress of Soil Science (in Portuguese). Brazilian Society of Soil Science, Recife. CD ROM.
- Novais, R. F. and Smyth, T. J. 1999. Phosphorus in Soil and Plant in Tropical Conditions (in Portuguese). Federal University of Viçosa, Viçosa. 399pp.
- Quaggio, J. A., Rajj Van, B. and Malavolta, E. 1985. Alternative use of the SMP buffer solution to determine lime of soil. *Communications in Soil Science and Plant Analysis*. **16**: 245–260.
- SAS. 1989. SAS/STAT—User's Guide. Version 6. SAS Institute, Cary. 846pp.
- Singh, B. and Gilkes, R. J. 1991. Phosphorus sorption in relation to soil properties for the major soil types of South-Western Australia. *Australian Journal of Soil Research*. **29**: 603–618.
- United States Soil Survey Staff. 1998. Keys of Soil Taxonomy. Washington. 325pp.
- Van Rajj, B., Cantarella, H., Quaggio, J. A. and Furlani, A. M. C. 1996. Recommendations of Fertilizers and Lime for the São Paulo State (in Portuguese). Agronomic Institute, Campinas. 285pp.
- Van Rajj, B., Quaggio, J. A. and Silva, N. M. 1986. Extraction of phosphorus, potassium, calcium and magnesium from soils by an ion-exchange resin procedure. *Communications in Soil Science and Plant Analysis*. **17**: 547–566.