

MODELAGEM DE EQUAÇÕES ESTRUTURAIS APLICADA AO  
ESTUDO DO CICLO DO P EM SOLOS TROPICAIS

**MARCUS VINICIUS DA SILVA SALES**

UNIVERSIDADE ESTADUAL DO NORTE FLUMINENSE  
DARCY RIBEIRO

CAMPOS DOS GOYTACAZES – RJ  
JULHO – 2015



MODELAGEM DE EQUAÇÕES ESTRUTURAIS APLICADA AO  
ESTUDO DO CICLO DO P EM SOLOS TROPICAIS

**MARCUS VINICIUS DA SILVA SALES**

Tese apresentada ao Centro de Ciências e Tecnologias Agropecuárias da Universidade Estadual do Norte Fluminense Darcy Ribeiro, como parte das exigências para obtenção do título de Doutor em Produção Vegetal.

Orientador: Prof. Antonio Carlos da Gama-Rodrigues

CAMPOS DOS GOYTACAZES – RJ  
JULHO – 2015

I. FICHA CATALOGRÁFICA

Preparada pela Biblioteca do **CCTA / UENF** 194/2015

Sales, Marcus Vinicius da Silva

Modelagem de equações estruturais aplicada ao estudo do ciclo do P em solos tropicais / Marcus Vinicius da Silva Sales. – 2015.

110 f. : il.

Orientador: Antonio Carlos da Gama-Rodrigues.

Tese (Doutorado - Produção Vegetal) – Universidade Estadual do Norte Fluminense Darcy Ribeiro, Centro de Ciências e Tecnologias Agropecuárias. Campos dos Goytacazes, RJ, 2015.

Bibliografia: f. 85 – 90.

1. Modelos estruturais 2. Variáveis latentes 3. Método de Hedley 4. *Pools* de P 5. Método de Bowman 6. Solos intemperizados I. Universidade Estadual do Norte Fluminense Darcy Ribeiro. Centro de Ciências e Tecnologias Agropecuárias. II. Título.

CDD–  
631.4

MODELAGEM DE EQUAÇÕES ESTRUTURAIS APLICADA AO  
ESTUDO DO CICLO DO P EM SOLOS TROPICAIS

**MARCUS VINICIUS DA SILVA SALES**

Tese apresentada ao Centro de Ciências e Tecnologias Agropecuárias da Universidade Estadual do Norte Fluminense Darcy Ribeiro, como parte das exigências para obtenção do título de Doutor em Produção Vegetal.

Aprovada em 29 de julho de 2015

Comissão Examinadora

---

Prof<sup>a</sup>. Emanuela Forestieri da Gama-Rodrigues (Ph.D., Ciência do Solo) – UENF  
(Coorientadora)

---

Prof. José Ramon Arica Chavez (D.Sc., Engenharia de Sistemas e Computação)  
– UENF

---

Prof. Júlio César Lima Neves (D.Sc., Produção Vegetal) – UFV

---

Prof. Antonio Carlos da Gama-Rodrigues (D.Sc., Ciência do Solo) – UENF  
(Orientador)

A Deus;

Aos meus pais José Antônio e Nilça, pelo amor e apoio em todos os momentos de minha vida;

A minhas irmãs, cunhados e sobrinhos, pelo apoio e pela paciência;

Aos meus amigos e familiares.

Dedico

## AGRADECIMENTOS

A Universidade Estadual do Norte Fluminense Darcy Ribeiro, ao Programa de Pós-Graduação em Produção Vegetal e ao Laboratório de Solos, pela oportunidade de realização deste curso;

A meus pais, minhas irmãs e familiares que contribuíram direta ou indiretamente para a realização deste trabalho, com apoio, estímulo e amor;

A meu orientador Antonio Carlos da Gama-Rodrigues, pela orientação e pelo apoio ao longo deste curso;

À prof<sup>a</sup>. Emanuela Forestieri da Gama-Rodrigues, pelos ensinamentos, pelas colaborações e pelos incentivos prestados;

Ao prof. Paulo Sérgio Dias da Silva, pelos incentivos, pelos ensinamentos, e pelas colaborações prestados;

Aos meus amigos e parceiros de laboratório, Marlon, Priscila, Renato, Seldon e Tarciana, que estiveram ao longo deste curso colaborando para a realização deste trabalho, pelo apoio e por compartilharem de seus conhecimentos;

Aos amigos do LSOL;

Aos meus amigos adquiridos ao longo desta empreitada, pelos momentos juntos, pelo apoio e pela amizade;

A todas as pessoas que não foram mencionadas e que de alguma forma contribuíram para a realização deste trabalho, meus agradecimentos.

## BIOGRAFIA

MARCUS VINICIUS DA SILVA SALES, filho de José Antônio Ferreira Sales e Nilça da Silva Sales, nasceu em 16 de janeiro de 1986, em Campos dos Goytacazes, Estado do Rio de Janeiro.

Em dezembro de 2008, diplomou-se em Licenciatura em Matemática pela Universidade Estadual do Norte Fluminense Darcy Ribeiro (UENF), em Campos dos Goytacazes, RJ. Durante a graduação foi bolsista de Extensão e Iniciação Científica pela UENF.

Em março de 2009, iniciou o Curso de Mestrado em Engenharia de Produção (Área de concentração: Pesquisa Operacional) e, em 20 de maio de 2011, diplomou-se Mestre em Engenharia de Produção pela Universidade Estadual do Norte Fluminense Darcy Ribeiro (UENF).

Em agosto de 2011, iniciou o Curso de Doutorado em Produção Vegetal (Área de concentração: Solos e Nutrição de Plantas) e, em 29 de julho de 2015 defendeu esta tese, requisito indispensável para a obtenção do título de “**Doctor Scientiae**”.

## SUMÁRIO

LISTAS DE TABELAS.....	vii
LISTA DE FIGURAS DO PRIMEIRO TRABALHO.....	ix
LISTA DE FIGURAS DO SEGUNDO TRABALHO.....	xi
RESUMO .....	xiii
ABSTRACT .....	xvi
1. INTRODUÇÃO .....	1
2. REVISÃO DE LITERATURA .....	4
2.1 Modelagem de Equações Estruturais. ....	4
2.2 Aplicações da modelagem de equações estruturais em ciências naturais..	6
2.3 Estágios para construção de modelos de equações estruturais. ....	10
2.3.1 Desenvolvimento do modelo teórico. ....	10
2.3.2 Construção do diagrama de caminho e das relações causais. ....	11
2.3.3 Conversão do diagrama de caminhos em equações estruturais e de mensuração. ....	13
2.3.4 Definição da matriz de entrada de dados e estimação do modelo proposto. ....	14
2.3.5 Identificação do modelo estrutural. ....	17
2.3.6 Avaliação do ajuste do modelo. ....	19
2.3.7 Interpretação e modificação do modelo. ....	22

2.4 Programa Aмос. ....	23
3. TRABALHOS .....	25
3.1 REESPECIFICAÇÃO DE MODELOS DE EQUAÇÕES ESTRUTURAIIS PARA O CICLO DO P EM SOLOS TROPICAIS. ....	25
RESUMO.....	25
ABSTRACT .....	26
INTRODUÇÃO .....	27
MATERIAL E MÉTODOS .....	29
RESULTADOS .....	34
DISCUSSÃO .....	44
CONCLUSÕES .....	47
REFERÊNCIAS BIBLIOGRÁFICAS.....	47
3.2 INTERAÇÃO DO CICLO DO P COM PROPRIEDADES DO SOLO USANDO MODELAGEM DE EQUAÇÕES ESTRUTURAIIS. ....	51
RESUMO .....	51
ABSTRACT .....	52
INTRODUÇÃO .....	53
MATERIAL E MÉTODOS. ....	56
RESULTADOS .....	68
DISCUSSÃO .....	77
CONCLUSÕES .....	79
REFERÊNCIAS BIBLIOGRÁFICAS .....	80
4. RESUMO E CONCLUSÕES.....	83
5. REFERÊNCIAS BIBLIOGRÁFICAS. ....	85

## LISTAS DE TABELAS

### LISTA DE TABELAS DO PRIMEIRO TRABALHO

Tabela 1. Índices de ajuste de modelos para comparação dos modelos ( <i>A</i> e <i>B</i> ) apresentados por Gama-Rodrigues et al. (2014) e modelos reespecificados. ....	36
--	----

### LISTA DE TABELAS DO SEGUNDO TRABALHO

Tabela 1. Estudo da extração sequencial de fósforo em solos tropicais sob diferentes usos da terra.....	59
Tabela 2. Estatística descritiva dos bancos de dados utilizados na modelagem do clico de P no solo.....	61
Tabela 3. Coeficientes de correlação de Pearson com os dados dos solos minerais de baixo teor de matéria orgânica, normalizados pela função Log10, N = 68.....	62
Tabela 4. Coeficientes de correlação de Pearson com os dados originais do trabalho de Oliveira et al, (2014), pelo método de extração de Bowman (Organossolos + Chernossolos + Latossolos, N = 38).....	63
Tabela 5. Coeficientes de correlação de Pearson com os dados originais do trabalho de Oliveira et al, (2014), pelo método extração de Bowman (Chernossolos + Latossolos, N = 28).....	64

Tabela 6. Índices de ajuste dos modelos para análise do ciclo do P pelo método de extração sequencial de Bowman.....	70
--	----

## LISTAS DE FIGURAS

### LISTA DE FIGURAS DO PRIMEIRO TRABALHO

Figura 1. Modelo <i>A</i> , modelo estrutural hipotético para o ciclo do P do solo por Gama-Rodrigues et al. (2014).....	28
Figura 2. Média e erro padrão das frações de fósforo no solo (Fonte: Gama-Rodrigues et al. 2014).....	30
Figura 3. Modelo <i>B</i> , modelo estrutural hipotético para o ciclo do P no solo por Gama-Rodrigues et al. (2014).....	32
Figura 4. Modelo <i>A1</i> , modelo de equações estruturais para o ciclo do P do solo. Todas as variáveis mensuradas (em retângulos) estão representadas como indicadores de efeitos associados com as variáveis latentes (em círculos ou elipses). Os valores correspondem aos parâmetros estimados padronizados ( $P < 0,001$ ) e o valor do $R^2$ (número em negrito). ° significativo a $P < 0,1$ . Valores das variáveis de erros ( $\delta_1 - \delta_8, \varepsilon_1, \zeta_1$ ) estão padronizados. Modelo $\chi^2 = 46,305, df = 17, p < 0.0001$ .....	35
Figura 5. Modelo <i>B1</i> , modelo de equações estruturais para o ciclo do P do solo. Todas as variáveis mensuradas (em retângulos) estão representadas como indicadores de efeitos associados com as variáveis latentes (em círculos ou elipses). Os valores correspondem aos parâmetros estimados padronizados ( $P < 0,001$ ) e o valor do $R^2$ (número em negrito). Valores das	

variáveis de erros ( $\delta_1 - \delta_5, \epsilon_1, \zeta_1$ ) estão padronizados. Modelo  $\chi^2 = 5,93$ ,  $df = 5$ ,  $p = 0,313$ ..... 38

Figura 6. Modelo C, modelo de equações estruturais para o ciclo do P do solo. Todas as variáveis mensuradas (em retângulos) estão representadas como indicadores de efeitos associados com as variáveis latentes (em círculos ou elipses). Os valores correspondem aos parâmetros estimados padronizados ( $P < 0,001$ ) e o valor do  $R^2$  (número em negrito). Valores das variáveis de erros ( $\delta_1 - \delta_5, \delta_8, \epsilon_1, \zeta_1$ ) estão padronizados. Modelo  $\chi^2 = 10,31$ ,  $df = 8$ ,  $p = 0,244$ ..... 39

Figura 7. Modelo C1, modelo de equações estruturais para o ciclo do P do solo. Todas as variáveis mensuradas (em retângulos) estão representadas como indicadores de efeitos associados com as variáveis latentes (em círculos ou elipses). Os valores correspondem aos parâmetros estimados padronizados ( $P < 0,001$ ) e o valor do  $R^2$  (número em negrito). \*\* significativo a  $P < 0,01$ . Valores das variáveis de erros ( $\delta_1 - \delta_5, \delta_8, \epsilon_1, \zeta_1$ ) estão padronizados. Modelo  $\chi^2 = 7,02$ ,  $df = 7$ ,  $p = 0,427$ ..... 42

Figura 8. Modelo C2, modelo de equações estruturais para o ciclo do P do solo. Todas as variáveis mensuradas (em retângulos) estão representadas como indicadores de efeitos associados com as variáveis latentes (em círculos ou elipses). Os valores correspondem aos parâmetros estimados padronizados ( $P < 0,001$ ) e o valor do  $R^2$  (número em negrito).  $^{\circ}$  e \*\* significativo a  $P < 0,1$  e  $0,01$ , respectivamente. Valores das variáveis de erros ( $\epsilon_1 - \epsilon_6, \delta_8, \zeta_1 - \zeta_3$ ) estão padronizados. Modelo  $\chi^2 = 7,02$ ,  $df = 7$ ,  $p = 0,427$ ..... 43

## LISTA DE FIGURAS DO SEGUNDO TRABALHO

- Figura 1. Modelo estrutural hipotético para o ciclo do P do solo com a inclusão de propriedades do solo..... 56
- Figura 2. Modelo A, modelo estrutural hipotético testado para avaliar o ciclo do P em solos tropicais com baixo teor de matéria orgânica, N = 68, com inclusão de propriedades do solo. Os indicadores de efeito (retângulos) estão associados com as variáveis latentes (elipses). Os valores numéricos correspondem a cargas fatoriais padronizadas com  $P < 0,001$ . Valores em negrito correspondem ao coeficiente de determinação ( $R^2$ ). “ns – não significativo”. Variâncias de erro ( $\varepsilon_1 - \varepsilon_7$  e  $\delta_1 - \delta_4$ ) estão padronizadas. Os valores de  $\chi^2 = 190,375$ ,  $df = 40$ ,  $p < 0,0001$ ..... 67
- Figura 3. Modelo B, modelo estrutural testado para avaliar o ciclo de P em solos tropicais com baixo teor de matéria orgânica, N = 68, com inclusão de propriedades do solo. Os indicadores de efeito (retângulos) estão associados com as variáveis latentes (elipses). Os valores numéricos correspondem a cargas fatoriais padronizadas com  $P < 0,001$ . Valores em negrito correspondem ao coeficiente de determinação ( $R^2$ ). “ns – não significativo”. Variâncias de erro ( $\varepsilon_1, \varepsilon_4- \varepsilon_7, \delta_5$  e  $\delta_6$ ) estão padronizadas. Os valores de  $\chi^2 = 14,374$ ,  $df = 11$ ,  $p = 0,213$ ..... 71
- Figura 4. Modelo B1, modelo estrutural testado para avaliar o ciclo do P em solos tropicais com alto teor de matéria orgânica (Chernossolos + Latossolos, N = 28), com a inclusão de propriedades do solo. Os indicadores de efeito (retângulos) estão associados com as variáveis latentes (elipses). Os valores numéricos correspondem a cargas fatoriais padronizadas com  $P < 0,001$ . Valores em negrito correspondem ao coeficiente de determinação ( $R^2$ ). “ns – não significativo, \* significativo a  $P < 0.05$ , \*\* Significativo a  $P < 0.01$ ”. Variâncias de erro ( $\varepsilon_1, \varepsilon_4- \varepsilon_7, \delta_5$  e  $\delta_6$ ) estão

padronizadas. Os valores de  $\chi^2 = 6,004$ ,  $df = 10$ ,  $p = 0,815$ ..... 72

Figura 5. Modelo C, modelo estrutural testado para avaliar o ciclo do P em solos tropicais com alto teor de matéria orgânica, N = 28, com a inclusão de propriedades do solo. Os indicadores de efeito (retângulos) estão associados com as variáveis latentes (elipses). Os valores numéricos correspondem a cargas fatoriais padronizadas com  $P < 0,001$ . Valores em negrito correspondem ao coeficiente de determinação ( $R^2$ ). “ns – não significativo, \* significativo a  $P < 0.05$ , \*\* Significativo a  $P < 0.01$ ”. Variâncias de erro ( $\varepsilon_1 - \varepsilon_5$ ,  $\delta_1 - \delta_4$ ) estão padronizadas. Os valores de  $\chi^2 = 24,788$ ,  $df = 23$ ,  $p = 0,361$ ..... 74

Figura 6. Modelo C1, modelo estrutural testado para avaliar o ciclo do P em solos tropicais com alto teor de matéria orgânica, N = 28, com a inclusão de propriedades do solo. Os indicadores de efeito (retângulos) estão associados com as variáveis latentes (elipses). Os valores numéricos correspondem a cargas fatoriais padronizadas com  $P < 0,001$ . Valores em negrito correspondem ao coeficiente de determinação ( $R^2$ ). “ns – não significativo, \* significativo a  $P < 0.05$ , \*\* Significativo a  $P < 0.01$ ”. Variâncias de erro ( $\varepsilon_1 - \varepsilon_5$ ,  $\delta_1 - \delta_3$ ) estão padronizadas. Os valores de  $\chi^2 = 15,308$ ,  $df = 16$ ,  $p = 0,502$ ..... 75

Figura 7. Modelo C2, modelo estrutural testado para avaliar o ciclo do P em solos tropicais com alto teor de matéria orgânica, N = 38, com a inclusão de propriedades do solo. Os indicadores de efeito (retângulos) estão associados com as variáveis latentes (elipses). Os valores numéricos correspondem a cargas fatoriais padronizadas com  $P < 0,001$ . Valores em negrito correspondem ao coeficiente de determinação ( $R^2$ ). “\* significativo a  $P < 0,05$ , \*\* Significativo a  $P < 0,01$ ”. Variâncias de erro ( $\varepsilon_1 - \varepsilon_5$ ,  $\delta_1 - \delta_3$ ) estão padronizadas. Os valores de  $\chi^2 = 18,952$ ,  $df = 16$ ,  $p = 0,271$ ..... 76

## RESUMO

SALES, Marcus Vinicius da Silva, D.Sc., Universidade Estadual do Norte Fluminense Darcy Ribeiro. Julho de 2015. **Transformações do P em solos tropicais usando modelagem de equações estruturais**. Orientador: Prof. Dr. Antônio Carlos da Gama-Rodrigues. Coorientadores: Prof<sup>a</sup>. Dra. Emanuela Forestieri da Gama-Rodrigues e Prof. Dr. Paulo Sérgio Dias da Silva.

O fósforo (P) é um elemento essencial para o metabolismo das plantas e limitações na sua disponibilidade, tanto no início do ciclo vegetativo quanto no ciclo produtivo, podem resultar em grandes prejuízos na produtividade agrícola. Em solos tropicais, especialmente aqueles que são altamente intemperizados, o P geralmente é encontrado abaixo dos níveis exigidos pelas culturas, devido à sua alta capacidade de adsorção no solo. Nos solos não fertilizados, a disponibilidade deste macronutriente é dependente dos mecanismos com que as raízes das plantas usam as fontes de P no solo que estão apresentadas na forma orgânica (Po) e inorgânica (Pi). Deste modo, a compreensão do ciclo do P no solo torna-se importante para fim de estabelecer estratégias de manejo que visem aumentar a disponibilidade desse nutriente às plantas em sistemas de produção de baixos insumos. A utilização da Modelagem de Equações Estruturais, metodologia utilizada no estudo, proporcionou a investigação dos múltiplos relacionamentos interdependentes existentes no estudo do ciclo do P, através de dados empíricos e conceitos teóricos com a utilização de variáveis latentes. Deste modo, trabalhos estão sendo desenvolvidos na área de ciência do solo para compreender como as

mudanças em um *pool* de P (variáveis latentes) podem afetar outros *pools* de P do modelo. O objetivo deste estudo foi desenvolver modelos de equações estruturais para analisar as interações entre diferentes *pools* de P. Neste sentido, foram utilizadas distintas metodologias de extração sequencial de P (o método de extração sequencial de Hedley e o método de extração de Bowman) divididos em dois artigos para uma melhor compreensão do ciclo do P em solos tropicais.

Mediante dados do método de extração sequencial de Hedley, foi possível testar, através da reespecificação de modelos estruturais, o grau com que o *pool* de P disponível foi controlado por diferentes *pools* de P no solo (*pools* de P orgânico, inorgânico, ocluso e mineral primário). O modelo estrutural que apresentou melhores resultados, incluindo variáveis latentes com múltiplos indicadores, mostrou ser consistente com o conceito teórico estudado. Neste modelo, não existiu apenas relações diretas entre os *pools* mineral primário, orgânico e ocluso sobre o *pool* de P disponível, mas também relações indiretas dos *pools* de P ocluso e mineral primário via *pool* orgânico que foram teórica e estatisticamente adequadas. Deste modo, o modelo apresentou as inter-relações existentes sobre o *pool* de P disponível, corroborando com a hipótese da dependência do *pool* de P disponível com o *pool* de P orgânico em solos tropicais não fertilizados.

Com a utilização dos dados do método de extração sequencial de Bowman foi possível estudar o ciclo do P analisando as influências dos conceitos teóricos dos *pools* de P estável ( $P_i$  e  $P_o$ ) e de P lábil ( $P_i$  e  $P_o$ ) sobre o *pool* de P disponível. Além do ciclo do P, foram incluídos no modelo básico os *pools* matéria orgânica (C orgânico e N total) e atributos físico-químicos (óxidos de Al e Fe, argila e pH) do solo com a finalidade de verificar as relações existentes destes atributos sobre a estabilidade e labilidade do P. Neste trabalho foram utilizados dois bancos de dados distintos. O primeiro refere-se a solos minerais com baixo teor de matéria orgânica, que configurou um modelo estrutural ajustado e o segundo refere-se a solos minerais com alto teor de matéria orgânica, que foi utilizado para validação do modelo ajustado. Com a utilização do primeiro banco de dados, o modelo estrutural que apresentou melhores resultados foi ajustado sem o *pool* de propriedades físico-químicas, mostrando uma relação de dependência do ciclo do P com o *pool* matéria orgânica. Com o segundo banco de dados, além da validação do modelo hipotético, foi gerado um modelo

alternativo, ajustado sem o *pool* de matéria orgânica, com a recuperação do *pool* de propriedades físico-químicas. Em geral, nos modelos estruturais ajustados, os *pools* de P estável e lábil se apresentaram como fontes de P sobre o *pool* de P disponível, confirmando as hipóteses de dependência entre os *pools* de P.

Palavras-chave: Modelos estruturais, Variáveis latentes, Método de Hedley, Método de Bowman, *Pools* de P, Solos intemperizados.

## ABSTRACT

SALES, Marcus Vinicius da Silva, D.Sc., Universidade Estadual do Norte Fluminense Darcy Ribeiro. July, 2015. **Transformation of P in tropical soils using structural equation modeling.** Advisor: Prof. Dr. Antonio Carlos de Gama-Rodrigues. Co-advisor: Prof<sup>a</sup>. Dra. Emanuela Forestieri de Gama-Rodrigues and Prof. Dr. Paulo Sérgio Dias da Silva.

Phosphorus (P) is an essential element for the metabolism of plants and limitations in their availability, both at the beginning of the season as the production cycle, can result in large losses in agricultural productivity. In tropical soils, especially those who are highly weathered, P is usually found below the levels required by crops, due to its high adsorption capacity of the soil. In unfertilized soils, the availability of this macronutrient is dependent on the mechanisms in which the plant roots use the sources of P in soil that is presented in organic (Po) and inorganic form (Pi). Thus, understanding of the phosphorus cycle in soil becomes important to establish management strategies aimed at increasing the availability of that nutrient to plants in low-input production systems. The use of structural equation modeling, methodology of the study, provided the investigation of multiple interdependent relationships existing in the study of the cycle P, through empirical data and theoretical concepts using latent variables. The objective of this study was to develop structural equation models to analyze the interactions between different pools of P. In this sense, we used different methodologies of sequential extraction P (sequential extraction method of Hedley

and extraction method of Bowman) divided into two articles for a better understanding of the cycle of P in tropical soils.

Upon data from sequential extraction method Hedley, it was possible to test by re-specification structural models, the degree to which the available P pool was controlled by different pools of P (organic, inorganic, occluded and mineral). The structural model showed better results, including latent variables with multiple indicators, was consistent with the theoretical concept studied. In this model, there existed only direct relationships between the primary mineral pools, organic and occluded on available P pool, but also indirect relations of pools P occluded and primary mineral via organic P pool which were theoretically and statistically appropriate. Thus, the model showed existing interrelationships on available P pool, supporting the hypothesis of P pool addition available with organic phosphorus pool in tropical soils unfertilized.

Using data from the sequential extraction method of Bowman was possible to study the P cycle analyzing the influences of the theoretical concepts of stable P pools ( $P_i$  and  $P_o$ ) and P labile ( $P_i$  and  $P_o$ ) on available P pool. In addition to the P cycle were included in the basic model pools organic matter (organic C and total N) and physicochemical properties (Al and Fe oxides, clay and pH) of the soil in order to verify the relationship of these attributes on the stability and lability of P. In this work we used two different data banks. The first relates to mineral soils with low organic matter content, which set an adjusted structural model and the second refers to mineral soils with high organic matter content, which was used for validation of the fitted model. Using the first database, the structural model showed better results without the pool was adjusted in physico-chemical properties, showing a dependence relationship of the phosphorus cycle with the organic matter pool. With the second database, in addition to validate the hypothetical model, an alternative model has been generated, adjusted without the pool of organic matter, the recovery of the physico-chemical properties pool. Overall, the adjusted structural models, stable and labile P pools were presented as sources of P on the available P pool, confirming the hypothesis of dependence between pools of P.

Keywords: Structural models, Latent variables, Hedley method, Bowman method, P pools, Weathered soils.

## 1. INTRODUÇÃO

O fósforo (P) é o macronutriente primário absorvido pelas plantas em menor quantidade, contudo a sua deficiência no solo causa danos ao crescimento e à produção das plantas. Em solos tropicais, especialmente aqueles que são altamente intemperizados, o P é geralmente encontrado abaixo dos níveis exigidos pelas culturas devido à sua alta capacidade de adsorção ao solo (Novais e Smyth, 1999). Nos solos não fertilizados, a disponibilidade deste macronutriente é dependente dos mecanismos com que as raízes das plantas usam as fontes de P orgânico (Po) e inorgânico (Pi) (Gama-Rodrigues et al., 2014).

A distinção destas formas de P é analisada principalmente nos trabalhos desenvolvidos por Hedley et al. (1982) e por Bowman (1989) com a extração sequencial de P, onde as frações (Po e Pi) são analisadas e classificadas baseadas na solubilidade. Costa (2015) avaliou as transformações de P em solos tropicais através dos métodos de extração de Hedley e Bowman. Neste trabalho foi verificada a importância das frações lábeis como fonte de P disponível. Resultados semelhantes, com a utilização do método de Bowman, foram encontrados por Duda (2000), Zaia et al. (2008a) e Oliveira et al. (2014). Estes autores verificaram que o P orgânico assume um papel importante na conservação de P no solo através da decomposição e mineralização do Po lábil. No entanto, somente através destas análises não foi possível estudar as interações entre estas frações de P no solo. Neste sentido, trabalhos com o uso da modelagem de equações estruturais (MEE) têm sido desenvolvidos para estudar as interações destas frações de P em solos tropicais (Gama-Rodrigues et al., 2014; Sales et al., 2015).

A modelagem de equações estruturais é uma metodologia adequada para estudos entre relacionamentos complexos. Com esta metodologia, pesquisadores podem analisar tanto efeitos diretos quanto indiretos entre as variáveis independentes de um modelo, calculados através de equações de regressão que são resolvidas simultaneamente e de forma independente (Mitchell, 1992; Malaeb et al., 2000). Além da análise simultânea das várias equações formadas, a MEE ainda promove a inclusão de variáveis não observáveis (latentes) nestas equações. As variáveis latentes são grandezas que não podem ser medidas diretamente, logo são incluídas no modelo para representar conceitos que explicam as variáveis observáveis (Grace, 2006).

A grande vantagem da utilização da MEE em relação a outras metodologias é a possibilidade de descrever e comparar diferentes modelos mostrando os relacionamentos hipotéticos entre as variáveis. Logo, o objetivo deste trabalho foi analisar o ciclo do P em solos tropicais altamente intemperizados através da MEE, com variáveis latentes, contribuindo com a disseminação desta técnica na área de ciências naturais no Brasil, especificamente na área de solos e ciclagem de nutrientes. Para o objetivo deste estudo são analisadas as interações entre os *pools* de P e verificados qual dos *pools* de P age como fonte ou dreno de P para a solução do solo com a utilização de dados obtidos pelos métodos de extração de Hedley e de Bowman. Como hipótese do trabalho tem-se que a MEE é uma metodologia científica capaz de representar e analisar relações de fenômenos complexos tais como as transformações do P em solos tropicais altamente intemperizados, explorando e confirmando modelos predeterminados, estabelecendo desta maneira estratégias para o manejo da terra de modo a aumentar a disponibilidade deste nutriente às plantas.

O trabalho está estruturado em Introdução, Revisão de Literatura, Trabalhos Científicos e Resumo e Conclusões.

Na revisão de literatura, são apresentados os conceitos básicos da modelagem de equações estruturais, alguns trabalhos científicos relacionados à ciência natural, ao método de construção dos modelos de equações estruturais através de sete etapas essenciais e por último foi feito uma breve revisão sobre o programa AMOS, utilizado para a modelagem do ciclo do P em solos tropicais.

No terceiro tópico são apresentados dois trabalhos científicos. No primeiro trabalho foi estudado o ciclo do P através da MEE com dados referentes à extração sequencial de P desenvolvida por Hedley. Neste, foi verificada a influência dos *pools* de P orgânico, inorgânico, ocluso e mineral primário sobre o *pool* de P disponível. No segundo trabalho científico também foi analisado o ciclo do P para solos tropicais usando a metodologia de MEE, porém os dados são exclusivos do método de extração ácido-base desenvolvido por Bowman. Neste trabalho, além do estudo das relações existentes entre os *pools* de P estável e lábil sobre o *pool* de P disponível, também foram analisadas as relações de propriedades do solo sobre o ciclo do P.

No último tópico são apresentados os resultados gerais do trabalho assim como as considerações finais e perspectivas de trabalhos futuros.

## 2. REVISÃO DE LITERATURA

### 2.1 Modelagem de equações estruturais

A modelagem de equações estruturais, segundo Klem (2006), se estruturou no ano de 1970 quando o estatístico Karl Jöreskog teve a ideia de combinar a econometria e a psicometria em um único modelo. A ideia original desta metodologia, conforme Hair et al. (2009), tem como base a combinação da análise de caminhos com análise fatorial e regressão múltipla. Desta maneira, pode-se descrever a MEE como uma estrutura capaz de modelar relações multivariadas complexas entre duas ou mais equações estruturais que envolvam influências e respostas simultâneas (Grace, 2006), usando combinações de dados estatísticos e qualitativos (Hair et al., 2009).

Esta metodologia é conhecida por diversos nomes, entre eles análise de estrutura de covariância, análise de variáveis latentes e análise fatorial confirmatória. A MEE também pode ser conhecida pelo nome de LISREL, que foi o software desenvolvido para este tipo de análise. Porém, independente do nome em que é conhecida, a MEE prevê um método direto para calcular simultaneamente múltiplos relacionamentos interdependentes com eficiência estatística, explorando-os de maneira aprofundada, gerando análises exploratórias e confirmatórias, permitindo a inclusão de variáveis não observáveis (variáveis latentes) nestes relacionamentos. As variáveis latentes são incluídas no modelo para representar conceitos (Grace, 2006) e desempenham um papel

especial na modelagem de equações estruturais, representando uma ligação entre os dados observados e a generalização teórica (Grace e Bollen, 2008).

Geralmente o pesquisador procura aplicar esta técnica em modelos predefinidos, onde se possui um conhecimento prévio do que se deseja modelar. Deste modo, a MEE é considerada como uma técnica mais confirmatória do que exploratória, e que consiste em analisar um sistema de equações de regressão lineares formado por intermédio das relações causais impostas no modelo (Kaplan, 2000; Ullman, 2007).

Este sistema de equações é assim chamado uma vez que as relações entre todas as variáveis, observáveis ou não observáveis (latentes), podem ser representadas por equações estruturais lineares ou podem ser assim transformadas (Bollen, 1989). O sistema de equações que forma a MEE pode ser dividido em dois submodelos. O modelo estrutural, composto pelas relações entre as variáveis latentes, e o modelo de mensuração, que analisa os relacionamentos existentes entre as variáveis latentes e as variáveis medidas, os indicadores (Grace, 2006).

As variáveis mensuráveis são chamadas geralmente de indicadores de efeito em MEE, uma vez que contribuem para a formação do construto (Grace et al., 2012). Estes indicadores, diferentemente do que ocorre em outras técnicas de modelagem, recebem em suas análises variáveis de erros de mensuração (Mitchell, 1992) melhorando o ajuste do modelo. Todavia, de forma similar a outras técnicas de análise de dados, a MEE exige que os dados (variáveis mensuráveis) sejam independentes, aleatórios e que seja observada a linearidade entre as relações estabelecidas.

Em um sentido amplo, a MEE representa a interpretação de uma série de relações de causa e efeito entre as variáveis, para a composição de hipóteses, que consideram os padrões de dependência estatística. Os relacionamentos dentro desta composição são descritos pela magnitude dos efeitos diretos ou indiretos entre as variáveis que compõem o modelo.

## 2.2 Desenvolvimento da modelagem de equações estruturais em ciências naturais

Trabalhos com o uso da modelagem de equações estruturais vêm sendo apresentados com maior frequência durante as últimas décadas em áreas de pesquisa como biologia, ecologia e agrossistemas. Parte destes trabalhos ainda se limita ao uso da MEE para analisar efeitos diretos e indiretos em modelos mensuráveis (Mitchell, 1992; Allison et al., 2007) fazendo distinção da análise de caminhos pelo uso das variáveis de erro de medição e pela possibilidade de haver mais de uma variável dependente no modelo. Porém, em outros trabalhos (Pugesek e Tomer, 1996; Malaeb et al., 2000; Grace e Bollen, 2006; Gama-Rodrigues et al., 2014) foi verificada essência do uso da MEE, onde foram analisadas relações diretas e indiretas em modelos com a inclusão de variáveis latentes (construtos).

A apresentação dos trabalhos a seguir tem por objetivo mostrar a versatilidade da modelagem de equações estruturais em ciências naturais e destacar que são poucos os trabalhos que desenvolvem esta metodologia na área de ciência do solo, principalmente com o uso de variáveis latentes em suas análises. Nesta seção são apresentados trabalhos relacionados a ciências naturais a partir da década de 90, período em que a MEE começou a se desenvolver nesta área de aplicação.

Na área de ecologia e biologia evolutiva, foram trabalhados por Mitchell (1992) dois modelos de equações estruturais mensuráveis concorrentes que analisam, através de caminhos diretos e indiretos, a influência das características morfológicas de uma flor sobre a frutificação da espécie por meio de abordagens de beija-flores de polinização. Com este trabalho Mitchell (1992) afirma que a MEE além de apresentar os coeficientes de caminhos para análise de efeitos diretos, indiretos e totais, proporciona a inclusão de estimativas de erro de medição, análise do  $R^2$  para todas as variáveis endógenas do modelo e análise de índices de ajuste do modelo proporcionando a escolha do modelo melhor ajustado na análise.

Pugesek e Tomer (1996) estudaram a sobrevivência de pássaros domésticos (pardais) através de suas características morfológicas distribuídas em fatores correspondentes a atributos físicos. Para este estudo foi utilizado a MEE, que permitiu identificar características morfológicas e fatores que influenciam na

sobrevivência destes pássaros. Os autores recomendam o uso da MEE para o estudo de seleção da espécie pela sua capacidade de testar e confirmar ou refutar hipóteses.

Malaeb et al. (2000) estudaram as interdependências entre quatro variáveis latentes, contaminação dos sedimentos, variabilidade natural, biodiversidade e potencial de crescimento relacionadas a onze variáveis ambientais através do uso da modelagem de equações estruturais. O objetivo destes autores foi medir os efeitos diretos, indiretos e totais da contaminação dos sedimentos e variabilidade natural sobre a biodiversidade e potencial de crescimento. Os autores ressaltam que, quando um modelo ajustado revela algo novo, seus resultados podem ser considerados confiáveis, pois os efeitos diretos e indiretos, bem como os erros de medição são confiáveis.

Arhonditsis et al. (2006) também trabalharam com a modelagem de equações estruturais para analisar os efeitos diretos, indiretos e erros de medição, explorando estruturas ecológicas de ecossistemas aquáticos eutróficos e mesotróficos. Para os autores a modelagem resultou em modelos plausíveis, onde foi observada uma proporção maior da variabilidade observada contabilizada no ambiente eutrófico.

No mesmo ano, Grace e Bollen (2006) utilizaram a MEE em dois exemplos de estudos ecológicos de sistemas vegetais. Os autores analisaram, no primeiro exemplo, os fatores limitantes como condições do solo, paisagem e concorrência das espécies sobre colonização bem-sucedida de espécies herbáceas. No segundo exemplo, os autores analisaram através da MEE a influência de micro-habitat e macro-habitat sobre a diversidade da espécie de anfíbios. Ambos os exemplos apresentados foram caracterizados pela inclusão da variável composite nos modelos, onde a composite representa variáveis conceituais distintas deixando os modelos mais complexos.

Allison et al. (2007) examinaram através da MEE o quanto a composição da comunidade microbiana do solo é alterada com a profundidade e ao longo de um gradiente sucessivo de restauração de pradaria nativa. Os autores previam que o carbono seria o fator de controle primário tanto na composição da biomassa quanto na comunidade microbiana, no entanto, verificaram com o uso da MEE que a biomassa microbiana do solo caiu fortemente com a profundidade, e aumentou com a idade da restauração, e que as mudanças na biomassa

microbiana foram, em grande parte, atribuídas a alterações nas concentrações de C e ou N no solo, juntamente com os impactos diretos e indiretos da biomassa radicular e magnésio.

Bernatchez et al. (2008) trabalharam a análise de caminhos através da metodologia da MEE para analisar as relações diretas e indiretas entre árvores, fungos micorrízicos arbusculares e fertilidade do solo em agrossistemas. Com o uso da MEE os autores conseguiram observar a existência significativa do impacto direto das árvores sobre a matéria orgânica do solo e do pH, e que estes afetam diretamente nas concentrações de N e P disponível das plantas, porém também foi observado que as relações hipotéticas entre as árvores, fungos micorrízicos arbusculares e fertilidade do solo não foram significativas como previsto.

Riginos e Grace (2008) estudaram o comportamento de animais selvagens analisando o quanto as variações da cobertura lenhosa e densidade da vegetação influenciam na composição de espécies herbáceas selvagens e domésticas, na savana africana, através do uso da MEE. Os resultados das análises apontaram que as comunidades herbáceas selvagens, em geral, buscam áreas com menor densidade de árvores em resposta a maior visibilidade de predadores.

Doblas-Miranda et al. (2009) estudaram, através da MEE, os fatores que determinam a distribuição e a estrutura do solo, e os grupos de camadas das macrofaunas que ainda permanecem pouco compreendidos. Os autores tomaram como hipótese que o papel dos fatores bióticos e abióticos difere entre grupos de solos minerais e da serapilheira em sistemas áridos e que ambos são ligados por animais que se deslocam entre estes níveis. Um dos resultados obtidos com o uso da modelagem de equações estruturais revelou que os fatores abióticos, bem como as relações entre os diferentes grupos tróficos foram fatores determinantes que afetam a densidade da fauna no sistema.

Barto et al. (2010) também analisaram a estrutura do solo com a utilização da modelagem de equações estruturais, estudando o impacto direto do uso da terra sobre agregados do solo, e a influência indireta dos fatores bióticos e abióticos sobre estes agregados. Para o estudo foram medidos os comprimentos das raízes e massa, colonização radicular por fungos micorrízicos arbusculares (FMA), comprimento extrarradical de hifas FMA, agregação e hidrofobia do solo,

além de propriedades físicas e químicas do solo. Com a MEE os autores conseguiram mostrar que o uso da terra pode aumentar a agregação do solo e que os fatores abióticos podem ser mais importantes para determinar a agregação do solo do que os fatores bióticos.

Travis e Grace (2010) aplicaram a modelagem de equações estruturais para compreender o desenvolvimento da grama do pântano comumente usada para restauração ecológica na área de distribuição natural na América do Norte. O modelo hipotético teve como base a análise dos efeitos diretos e indiretos da distância geográfica, medida pela distância de dispersão e latitude, e distância genética, determinadas pela comparação de marcadores moleculares neutros com os da população local, sobre o desempenho da espécie como diâmetro, número de caules, número de inflorescência entre outros. O modelo foi ajustado apresentando fortes relações estruturais de efeitos diretos e indiretos, sendo destacada a relação da distância genética com o desempenho da espécie.

Na área de fertilidade do solo, Gama-Rodrigues et al. (2014), aplicaram a modelagem de equações estruturais, a um nível exploratório, para testar modelos hipotéticos do ciclo do P, com variáveis latentes, usando dados do método de extração de Hedley. Em seu modelo básico foram analisadas as influências diretas das variáveis latentes *pools* de P orgânico, inorgânico, ocluso e mineral primário sobre o *pool* de P disponível. Através de modificações nos modelos, os autores chegaram a um modelo mensurável final, onde concluíram que o *pool* de P orgânico composto pelas frações bicarbonato Po, hidróxido Po e sônico Po foi a principal fonte de P para o *pool* de P disponível em solos tropicais, verificando a teoria estudada.

Sales et al. (2015) trabalharam com reespecificações de dois modelos estruturais apresentados por Gama-Rodrigues et al. (2014). A utilização da MEE em seu trabalho permitiu que, além das relações diretas previamente analisadas, fossem feitas análises de relações indiretas dos *pools* de P ocluso e mineral primário sobre o P disponível por intermédio do *pool* de P orgânico. Com esta análise foi possível verificar a dependência do *pool* de P disponível pelo *pool* de P orgânico em solos tropicais não fertilizados.

## 2.3 Estágios para a construção da modelagem de equações estruturais

Elaborar modelos de equações estruturais demanda conhecimentos técnicos que devem ser executados de forma iterativa, pois a falha na execução de uma etapa implica na não execução da etapa seguinte (Kline, 2011). Estas etapas são apresentadas por Hair et al. (2009), Bollen e Noble (2011) e Kline (2011), na forma de estágios, como forma prática de elaborar tais modelos. O fluxograma (Figura 1) foi elaborado e explicado de acordo com os estágios desenvolvidos por estes autores.

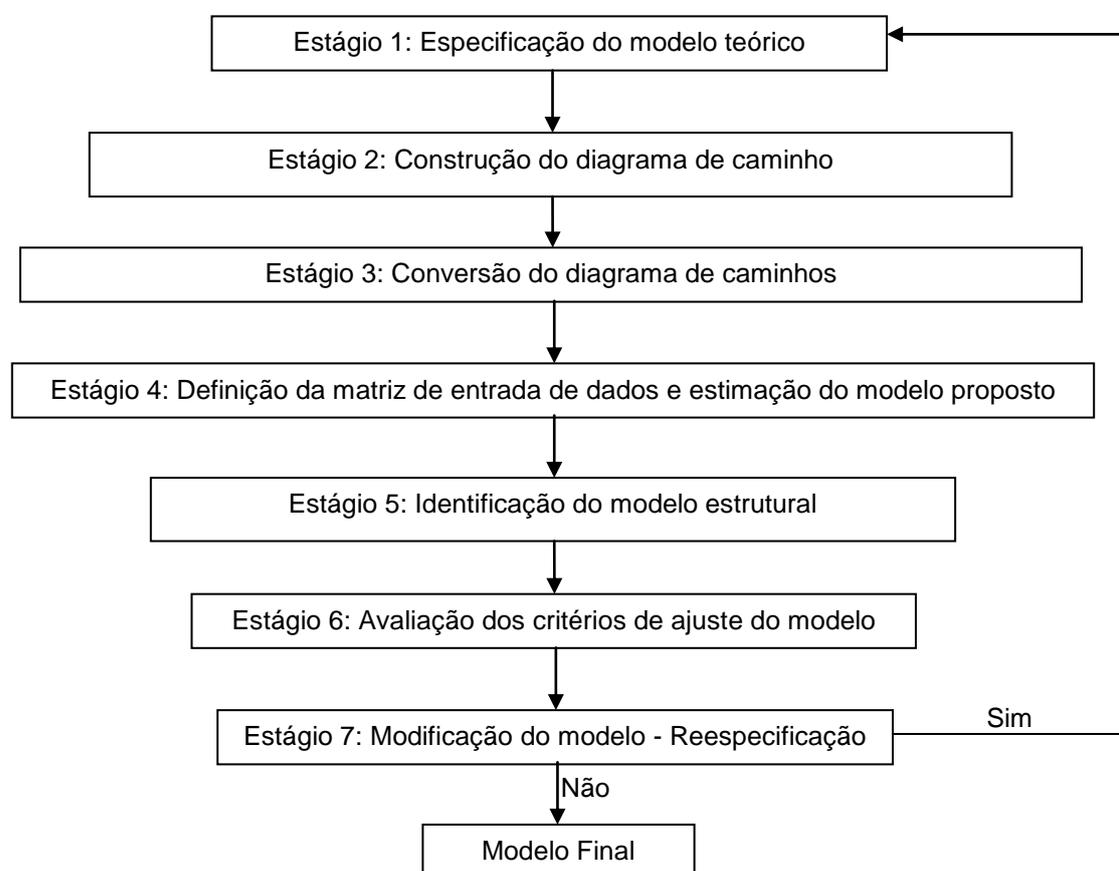


Figura 1. Fluxograma de etapas para o desenvolvimento de modelos baseados em equações estruturais.

### 2.3.1 Desenvolvimento do modelo teórico

A especificação é a etapa mais importante da modelagem, pois nesta etapa são definidas as hipóteses que são analisadas durante o processo de

modelagem (Kline, 2011). Para o desenvolvimento do modelo teórico devem ser definidos inicialmente os construtos individuais (variáveis latentes) que conformarão a estrutura do modelo. Uma boa base teórica de mensuração é necessária para obter resultados satisfatórios para a definição destes construtos (Hair et al., 2009), onde devem ser selecionados itens indicadores individuais (variáveis mensuráveis), que previamente são estipulados por forte embasamento teórico ou analisados por meio de estatísticas básicas, como análise de correlação. Cada variável latente incluída no modelo é formada por indicadores que sejam representativos, conformando modelos de mensuração. Nos modelos de mensuração as variáveis latentes ajudam a explicar as variáveis mensuráveis, pois se presume que elas os representam (Bryne, 2009). Além desta característica, cada indicador é associado a uma variável de erro de medição (Grace, 2006), que será incluída em cada equação de mensuração formada.

Não existe um limite máximo de variáveis latentes ou mensuráveis incluídas no modelo, porém estas devem ser significativas, pois a técnica de modelagem de equações estruturais tem como finalidade a confirmação e avaliação de modelos concorrentes, ou o desenvolvimento de novos modelos (Hair et al., 2009).

Com os construtos definidos inicia-se o processo de desenvolvimento do modelo teórico em que é avaliada a estratégia da modelagem confirmatória e são feitas as especificações das relações causais entre as variáveis latentes selecionadas para a pesquisa.

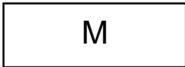
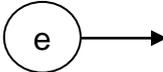
### 2.3.2 Construção do diagrama de caminhos das relações causais

Neste estágio, o diagrama de caminhos é construído utilizando as relações causais impostas no estágio anterior. A forma gráfica do modelo teórico fornece um retrato visual das relações que são assumidas entre as variáveis latentes em estudo e seus indicadores (Byrne, 2009). Estas relações e variáveis são ilustradas a partir de símbolos indicando os relacionamentos que são apresentados na Tabela 1.

Na construção do diagrama de caminhos de qualquer modelo, foram indicadas as variáveis latentes (L) com símbolos de círculos ou elipses e as

variáveis medidas (M) através de quadrados ou retângulos. Variáveis de erros (e) também são consideradas latentes, pois não são medidas diretamente. A seta de relação direcional tem papel importante na construção dos diagramas, ela indica se o construto é reflexivo, com setas partindo da variável latente para os indicadores, ou se o construto é formativo, com setas partindo dos indicadores para o construto (Grace et al., 2010). Neste trabalho todos os modelos apresentados são reflexivos.

Tabela 1. Simbologia básica do diagrama de caminhos usado na MEE.

Simbologias	Significados
	Variável latente.
	Variável mensurável.
	Relação direcional (indicador de causalidade).
	Modelo de mensuração reflexivo. Análise da variável latente sobre a variável mensurável.
	Modelo de mensuração formativo. Análise da variável mensurável sobre a variável latente.
	Modelo estrutural. Análise da variável latente exógena sobre variável latente endógena.
	Modelo estrutural. Análise de correlação/covariância entre variáveis latentes exógenas.
	Variáveis de erro de medição associado com as variáveis mensurável ou latente endógenas.

A relação direcional também diferencia as variáveis latentes entre exógenas e endógenas. Os construtos endógenos são previstos por construtos exógenos ou por outros construtos endógenos, ou seja, as setas direcionais partem das variáveis latentes exógenas ou de outras variáveis latentes endógenas para o construto em análise. As variáveis exógenas não recebem

relação causal de nenhuma outra variável do modelo, exceto quando a relação é de covariação, que só ocorre entre este tipo de variável (Byrne, 2009; Iacobucci, 2009, Bollen e Noble, 2011).

### 2.3.3 Conversão do diagrama de caminhos em equações estruturais e de mensuração

Geralmente pesquisadores iniciam a modelagem de equações estruturais pela construção de diagramas de caminhos, que permitem a visualização das diferentes relações estruturais entre as variáveis do modelo de maneira clara (Byrne, 2009). Cada diagrama de caminhos construído pode ser representado algebricamente por um sistema de equações estruturais lineares ou pode ser transformado (Bollen, 1989). Assim, o sistema que forma a MEE pode ser dividido em dois subsistemas, o modelo estrutural, constituído pelas relações entre as variáveis latentes, e o modelo de mensuração, constituído pelas relações entre variáveis latentes e seus respectivos indicadores (Byrne, 2009).

O modelo estrutural, associação entre as variáveis latentes, pode ser representado algebricamente pela equação,

$$\eta_i = \alpha_\eta + B\eta_i + \Gamma\xi_i + \zeta_i \quad (1)$$

onde os vetores  $\eta_i$  e  $\xi_i$  são as variáveis latentes endógenas e exógenas, respectivamente para unidade  $i$ ;  $\alpha_\eta$  é um vetor com interceptos de cada equação, sendo assumido neste trabalho como vetor de zeros; as matrizes  $B$  e  $\Gamma$  são matrizes de coeficientes que representam os efeitos entre as variáveis latentes; e  $\zeta_i$  representa o vetor de erros aleatórios relacionados às variáveis latentes endógenas.

O modelo mensurável pode ser representado de duas maneiras: primeiro, considerando as grandezas de medições com seus respectivos construtos endógenos, configurando o modelo de mensuração de  $y$ ,

$$y_i = \alpha_y + \Lambda_y\eta_i + \varepsilon_i \quad (2)$$

onde  $y_i$  representa o vetor de variáveis mensuráveis ou indicadores associados às variáveis latentes endógenas; o vetor  $\alpha_y$  representa cada intercepto das

equações de  $y_i$ , sendo um vetor de zeros; a matriz  $\Lambda_y$  representa os coeficientes relacionados aos indicadores das variáveis latentes endógenas, e o vetor  $\varepsilon_i$  representa os erros de mensuração para cada um dos indicadores de  $y_i$ . A segunda maneira de representar o modelo mensurável é pelo modelo de mensuração de  $x$ . Este está associado às variáveis latentes exógenas e é representado pela expressão,

$$x_i = \alpha_x + \Lambda_x \xi_i + \delta_i, \quad (3)$$

onde  $x_i$  representa o vetor de variáveis mensuráveis ou indicadores associados às variáveis latentes exógenas; o vetor  $\alpha_x$  representa cada intercepto das equações de  $x_i$ , sendo um vetor de zeros; a matriz  $\Lambda_x$  representa os coeficientes relacionados aos indicadores das variáveis latentes exógenas; e o vetor  $\delta_x$  representa os erros de mensuração para cada um dos indicadores de  $x_i$  (Bollen e Noble, 2011). Deste modo, todo e qualquer diagrama de caminhos pode ser transformado em um sistema de equações estruturais.

#### 2.3.4 Definição da matriz de entrada de dados e estimação do modelo proposto

Tendo convertido o diagrama de caminhos em modelos estrutural e mensuráveis o pesquisador deve definir o tipo de matriz de entrada de dados para a estimação dos parâmetros do modelo. As matrizes a serem definidas são de covariância ou correlação, onde a escolha entre estas matrizes determina o tipo de análise interpretativa a ser feita (Grace, 2006, Hair et al., 2009).

O desenvolvimento original da MEE se formou com a utilização da matriz de covariância e por este motivo a MEE também pode ser chamada de análise de estrutura de covariância (Hair et al., 2009). A vantagem da utilização da matriz de covariância sobre a matriz de correlação é que ela fornece comparações válidas para métricas ou amostras distintas, sendo utilizada para expressar valores brutos (Grace, 2006). No entanto, o uso da matriz de correlações facilita na interpretação das relações entre os construtos do modelo, aonde a grande vantagem da sua utilização vem do fato de que as estimativas paramétricas são padronizadas, com valores estimados no intervalo de  $-1,0$  a  $+1,0$ , e estas não são dependentes de métricas (Hair et al. 2009). Também tem sido utilizado por alguns programas de

MEE o uso de planilhas com os dados originais sem que o pesquisador calcule uma das matrizes citadas, deixando a critério a utilização ou não de resultados padronizados.

Assim como é necessário definir a matriz de entrada de dados para a análise dos parâmetros do modelo, deve ser escolhido a técnica de estimação para estes parâmetros e o programa computacional para implementação do modelo. A técnica de estimação mais utilizada é a Máxima Verossimilhança (ML) (Bistaffa, 2010, Iacobucci, 2010), que fornece resultados válidos com tamanho amostral tão pequeno quanto 50, mas que para garantir soluções mais consistentes são recomendados tamanhos mínimos de 100 a 150 e idealmente de 200 (Hair et al., 2009).

A função de estimação baseada em máxima verossimilhança (ML) é dependente da suposição de que os indicadores possuam uma distribuição normal multivariada, e mesmo para amostras pequenas, os estimadores são assintoticamente não viciados, consistentes, assintoticamente eficientes, exigindo que as observações sejam independentes (Grace et al., 2012). Seu cálculo pode ser executado através da formulação matemática,

$$F_{ML} = \log|\hat{\Sigma}| + tr(S\hat{\Sigma}^{-1}) - \log|S| - (p + q) \quad (4)$$

em que a matriz  $\hat{\Sigma}$  representa as covariâncias estimadas, e a matriz  $S$  representa as covariâncias observadas entre cada variável presente no modelo;  $p$  representa o número de indicadores das variáveis latentes endógenas e  $q$  o número de indicadores das variáveis latentes exógenas (Grace, 2006). A hipótese geral em MEE é que as matrizes de covariância estimada e observada sejam estatisticamente iguais.

O método ML é um procedimento iterativo que melhora as estimativas de parâmetros maximizando a probabilidade de os dados representarem a população analisada (Kline, 2011). À medida que o tamanho amostral aumenta, considerando que estes estimadores são normalmente distribuídos, é possível realizar testes de significância estatística para cada parâmetro calculado. Além destas análises individuais, também é possível realizar testes de ajuste global do modelo, onde a hipótese de nulidade geral do modelo,  $H_0$ , representa a igualdade estatística entre as matrizes de covariâncias  $S$  e  $\hat{\Sigma}$ ,

$$H_0 : S = \hat{\Sigma}.$$

É assumido que o modelo tem um bom ajuste quando a hipótese de nulidade não for rejeitada. Outras técnicas de estimação de parâmetros como mínimos quadrados generalizados (GLS), mínimos quadrados não-ponderados (ULS) e mínimos quadrados ponderados (WLS) também são utilizadas na análise de MEE.

Os estimadores GLS têm por finalidade minimizar a diferença entre os valores das matrizes de covariâncias  $S$  e  $\hat{\Sigma}$ , possuindo as mesmas propriedades dos estimadores de ML, executando os mesmos testes individuais para os parâmetros calculados e ajuste global. Já os estimadores de ULS minimizam a soma de quadrados de cada elemento da matriz residual ( $S - \hat{\Sigma}$ ) e têm como principal característica não requerer que as variáveis observadas tenham uma distribuição em particular, além de produzirem estimativas consistentes. A única exigência destes estimadores é que os dados observados sejam contínuos. Entretanto, o método ULS não possui estimadores assintoticamente eficientes e não pode ser avaliado pelo teste  $X^2$  (Bistaffa, 2010).

O método de WLS também é comumente conhecido como estimador assintoticamente livre de distribuição ou função de distribuição arbitrária (ADF). A vantagem para se utilizar estes estimadores é que são feitas suposições mínimas sobre a distribuição dos dados. Este estimador é eficiente, fornecendo uma matriz de estimativas de parâmetros e teste de  $X^2$  para ajuste global do modelo. Contudo, exige muito esforço computacional para estimar o modelo e necessita de tamanho amostral superior ao dos outros métodos apresentados (Hair et al., 2009).

Após a definição da matriz de entrada de dados e do método de estimação dos parâmetros, deve ser escolhido um software que calcule estes parâmetros. Os softwares mais utilizados em MEE são: AMOS (Análise de Estrutura de Momentos), LISREL (Linear structural relations) e EQS (Equações). Estes três pacotes possuem em comum a utilização de modelos de regressão linear na sua análise. O programa AMOS se destaca dentre os programas citados pela sua interface gráfica que é de fácil manuseio, incluindo ferramentas de construção dos diagramas (Bryne, 2009). O LISREL é o programa originalmente

desenvolvido para estimar os parâmetros dos modelos de equações estruturais, desenvolvido por Karl Jöreskog em 1973. E o programa EQS além de testar modelos estruturais, resolve modelos de regressão e de análise fatorial (Hair et al. 2009). Neste trabalho foi utilizado o programa AMOS pelas características que são apresentadas no próximo tópico.

### 2.3.5 Identificação do modelo estrutural

Neste estágio é verificada a ocorrência de limitações explanatórias do modelo para gerar soluções únicas. O problema de identificação é apresentado frequentemente em trabalhos de MEE, sendo fortemente relacionado à estimação dos parâmetros e ao número de indicadores do modelo (Iacobucci, 2010). A identificação pode ser visualizada previamente quando se constrói o sistema de equações estruturais verificando o excesso ou a falta de parâmetros (incógnitas) a serem estimados em relação ao número de equações construídas. O ideal é que para cada coeficiente estimado existisse uma única equação, no entanto, à medida que a complexidade do modelo aumenta essa suposição não é mais garantida (Hair et al., 2009).

Grace (2006) destaca que uma condição necessária, porém não suficiente, para a identificação de um modelo é que o número de correlações entre as grandezas de medição seja maior ou igual ao número de parâmetros causais a serem estimados. Kline (2011) acrescenta que para um modelo ser identificado, um requisito também necessário, porém, insuficiente é que o modelo possua no mínimo graus de liberdade igual ou superior a zero,  $df \geq 0$ , fornecendo informação suficiente para que o conjunto de equações estruturais possua uma solução única (Hair et a., 2009). Os graus de liberdade podem ser calculados facilmente pela expressão

$$df = \frac{1}{2}[p(p + 1)] - k \quad (5)$$

em que  $p$  representa o número de indicadores do modelo e  $k$  os parâmetros a serem estimados.

De acordo com o valor dos graus de liberdade os modelos de equações estruturais são classificados como não-identificado, exatamente identificado e superidentificado (Figura 2).

Um modelo é não-identificado quando o número de parâmetros a serem estimados é maior do que o número de covariâncias ou correlações dos itens medidos, ocasionando  $df < 0$ . Um modelo é exatamente identificado ou saturado quando existe o número suficiente de graus de liberdade para estimar todos os parâmetros livres do modelo, ou seja,  $df = 0$ . E o modelo é considerado superidentificado quando há excesso de covariâncias ou correlações calculadas em relação ao número de parâmetros livres do modelo,  $df > 0$  (Grace, 2006).

Quando um modelo é não-identificado, porém apresenta  $df > 0$ , podem ser fixados parâmetros neste modelo até que o problema de identificação seja resolvido. Geralmente são fixadas cargas fatoriais de construtos que possuem três ou mais indicadores com o valor de 1, (Hair et al., 2009). E para construtos que possuem indicador único, a variância de erro pode ser fixada com o valor 0,1 (10%) correspondendo ao erro de medição (Grace e Bollen, 2008).

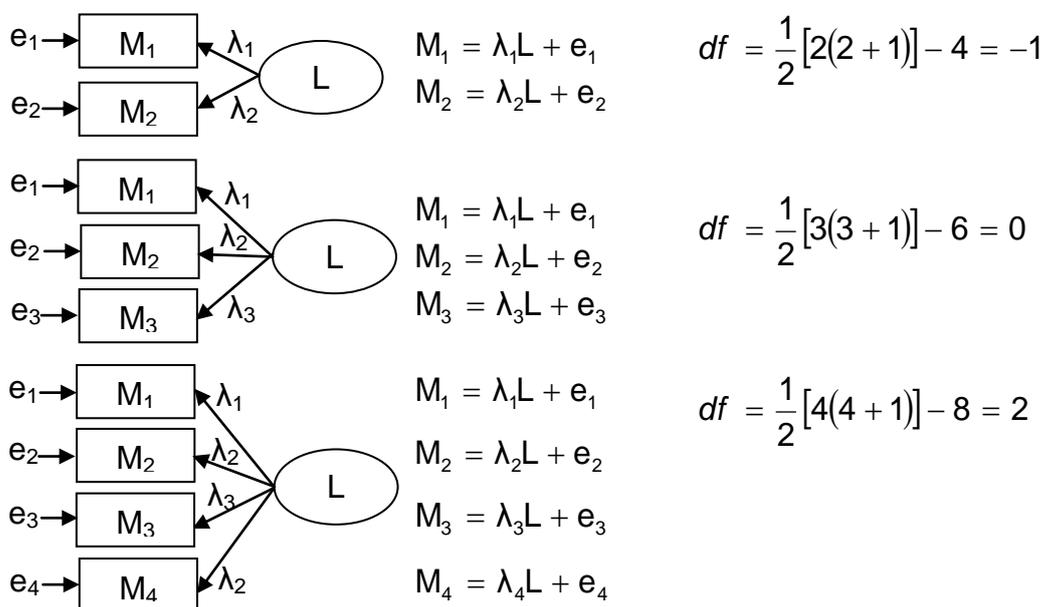


Figura 2. Caracterização da identificação de modelos de equações estruturais. As variáveis M's nos retângulos são indicadores e os e's e  $\lambda$ 's são os parâmetros livres para a estimação. O primeiro modelo é não-identificado, o segundo modelo é exatamente identificado e o terceiro modelo é superidentificado.

### 2.3.6 Avaliação do ajuste do modelo

Neste estágio o modelo é avaliado segundo alguns critérios de ajuste, que definem a adequação e o ajuste geral do modelo. São avaliadas as relações mensuráveis e a intensidade entre as relações estruturais. Segundo Bistaffa (2010), esta etapa pode ser dividida em três momentos: avaliação dos parâmetros estimados; ajuste geral do modelo; e medidas alternativas de ajuste.

Para a avaliação dos parâmetros estimados devem ser observados inicialmente os sinais e a magnitude das relações estimadas. Caso exista alguma estimativa fora do valor esperado, alguma hipótese do modelo testado deve estar mal formulada. Bistaffa (2010) destaca que alguns inconvenientes como soluções impróprias de estimativas, variâncias negativas ou não significantes e matrizes de covariância e ou correlações não positivas definidas, correlação entre duas variáveis fora do intervalo  $[-1, 1]$ , e erro padrão muito grande associado a qualquer um dos coeficientes, são causados pela má especificação dos modelos. Estes inconvenientes também podem ser atribuídos ao fato da amostra coletada para análise não corresponder à população estudada e também à presença de *outliers* nos dados amostrados. *Outliers* são valores discrepantes dentro de uma amostra que prejudicam a normalidade multivariada.

Quanto ao ajuste geral do modelo, esta é avaliada com qualidade do ajuste (GOF), indicando o quão bem o modelo especificado corresponde a matriz de covariância entre os indicadores, ou seja, é analisada a similaridade entre as matrizes de covariâncias  $S$  e  $\hat{\Sigma}$  (Byrne, 2009; Hair et al. 2009), onde quanto mais próximos forem os valores exibidos entre estas matrizes, melhor será o ajuste do modelo e maior a possibilidade de aceitar a hipótese de nulidade,  $H_0$ . Para verificar este ajuste é utilizado o teste qui-quadrado,  $X^2$

$$X^2 = (n - 1)F_{ML} \quad (5)$$

onde  $n$  refere ao tamanho da amostra e  $F_{ML}$  à função de máxima verossimilhança (Grace, 2006). Para Hair et al. (2009), este é o índice de ajuste mais utilizado em MEE e espera-se que o valor desta estatística seja o menor possível associado ao grau de liberdade ( $df$ ) e ao valor- $p$  não significativo ( $p > 0,05$ ), diferentemente da estatística básica, que busca por valores de  $X^2$  altos e valores de

probabilidade significativos. Se o valor do teste qui-quadrado estiver associado ao valor- $p$  significativo ( $p < 0,05$ ) e o tamanho da amostra for muito grande (superior a 150), outras medidas de ajuste podem indicar que o modelo é aceitável. Segundo Byrne (2009), a razão  $X^2/df$  também pode ser analisada para indicar ajuste do modelo. Valores que não excedam a 5 indicam ajuste. Para modelos exatamente identificados não é possível fazer este tipo de análise.

Devido à sensibilidade do teste qui-quadrado ao tamanho amostral, outros índices de ajuste absoluto, índices de ajuste incremental e índices de parcimônia são tomados como alternativas para a verificação do ajuste global do modelo.

**Índice de qualidade de ajuste (GFI)** – O GFI foi o índice de ajuste absoluto pioneiro na tentativa de produzir uma estatística de ajuste que fosse menos sensível a tamanho amostral ( $N$ ). Mesmo que  $N$  não seja incluído na fórmula, esta estatística continua indiretamente sensível ao tamanho da amostra por conta do efeito de  $N$  sobre distribuições amostrais. Valores de GFI podem ser calculados através da função,

$$GFI_{ML} = 1 - \frac{tr\left[\left(\hat{\Sigma}^{-1}S - I\right)^2\right]}{tr\left[\left(\hat{\Sigma}^{-1}S\right)^2\right]}, \quad (6)$$

onde a matriz  $I$  corresponde a uma matriz identidade  $p \times p$ , em que  $p$  representa o número de indicadores do modelo (Bistaffa, 2010). O intervalo admissível de valores para o GFI é de 0 a 1, com valores maiores indicando melhor ajuste (Hair et al., 2009). Neste trabalho, pelo número de variáveis e tamanho amostral, espera-se por valores de  $GFI \geq 0,95$ .

**Índice de ajuste comparativo (CFI)** – Outro índice bastante utilizado para análise de modelo de MEE é o índice de ajuste incremental CFI, índices de ajuste comparativo. Ele avalia o quão bem um modelo especificado se ajusta relativamente a algum modelo alternativo de referência, conhecido como modelo nulo, gerado dentro do programa. Seu cálculo pode ser efetuado pela seguinte formulação:

$$CFI = 1 - \frac{\left(X_k^2 - df_k^2\right)}{\left(X_n^2 - df_n^2\right)}, \quad (7)$$

onde, neste caso, os índices  $k$  e  $n$  representam respectivamente os valores associados ao modelo especificado pelo pesquisador e valores associados ao modelo estatístico nulo. Este é um índice normado, de forma que seus valores variam entre 0 e 1, com valores mais altos indicando melhor ajuste. Pelo fato do CFI ter muitas propriedades desejáveis, incluindo sua insensibilidade relativa, mas não completa, em relação à complexidade do modelo, ele está entre os índices mais usados. Valores CFI abaixo de 0,90 indicam, geralmente, que o modelo e os dados amostrados não se ajustam bem (Hair et al., 2009).

**Raiz do erro quadrático médio de aproximação (RMSEA)** – Outra medida que tenta corrigir a tendência da estatística  $X^2$  a rejeitar modelos é o índice RMSEA (Hair et al., 2009). Este índice representa o quão bem um modelo se ajusta a uma população e não apenas a uma amostra usada para estimação. O seu cálculo pode ser realizado a partir da formulação,

$$RMSEA = \sqrt{\frac{(X_k^2 - df_k^2)}{(N-1)}}, \quad (8)$$

Byrne (2009) indica que valores de RMSEA abaixo de 0,08 são satisfatórios e que o modelo é aceitável. Este índice é considerado de má qualidade de ajuste, pois apresentando valores altos há fortes indícios de diferença significativa entre valores observados e estimados pelas matrizes de covariância.

Outros índices como AGFI (GFI ajustado para os graus de liberdade), NFI (índice de ajuste normalizado), TLI (índice de Tucker Lewis) e AIC (critério de informação de Akaike) também são índices que ajudam a definir o ajuste do modelo.

**Índice de qualidade de ajuste corrigido (AGFI)** – tenta levar em conta diferentes graus de complexidade do modelo. Este índice ajusta GFI por uma proporção entre os graus de liberdade usados em um modelo e o número total de graus de liberdade disponíveis. O AGFI penaliza modelos mais complexos e favorece aqueles com um número mínimo de caminhos livres. Seus valores são tipicamente menores do que os valores do GFI em proporção à complexidade do modelo.

**Índice de ajuste normado (NFI)** – Para Hair et al. (2009), o NFI é um dos índices de ajuste incremental. Ele é uma proporção da diferença no valor  $X^2$  para

o modelo ajustado e um modelo nulo. Seu valor varia entre 0 e 1, e valores mais próximos de 1 indicam melhor ajuste.

**Índice de Tucker Lewis (TLI)** – O TLI não é normado, seu valor pode ficar abaixo de 0 ou acima de 1. No entanto, tipicamente, modelos com bom ajuste têm valores que se aproximam de 1. Para valores maiores do que 1 este índice é considerado superajustado, entretanto, para valores muito menores do que 1 este índice indica má especificação do modelo. Hair et al. (2009) afirmam que na prática, o TLI e o CFI geralmente fornecem valores muito parecidos.

**Critério de informação de Akaike (AIC)** – Esta medida é utilizada na comparação de modelos concorrentes. É recomendado escolher o modelo que apresentar o menor resultado do índice AIC.

Outro valor que também pode ser considerado é o valor do coeficiente de determinação ( $R^2$ ) que é calculado para qualquer variável endógena. Mas, assim como na análise de regressão múltipla, estes resultados podem ser afetados pela multicolinearidade existente entre os indicadores do modelo. Hair et al. (2009) recomendam que valores de correlações, entre duas variáveis independentes, superiores a 0,9 em valor absoluto sejam examinadas, pois possivelmente comprometerão os resultados obtidos na estimação do modelo.

### 2.3.7 Interpretação e modificação do modelo

Com todos os resultados obtidos, neste estágio são feitas as interpretações e possíveis modificações do modelo. Na verificação e considerações finais sobre o modelo são comparados os resultados obtidos no desenvolvimento do modelo com os objetivos e as hipóteses estabelecidas no estudo.

Neste estágio deve ser verificada a significância dos relacionamentos estabelecidos a partir da teoria, e caso haja necessidade de modificação do modelo, podem ser ainda estabelecidos parâmetros para a reespecificação do modelo, sendo necessário retornar ao estágio inicial para fazer uma nova especificação. O objetivo da reespecificação é aumentar o ajuste ou a sua correspondência com a teoria. Bollen e Noble (2011) definem a reespecificação

com o processo de inclusão ou exclusão de parâmetros a serem estimados no modelo teórico original.

Tendo visto que o modelo encontra-se aceitável, deve ser escolhido entre uso de soluções padronizadas ou não-padronizadas para interpretação do fenômeno. O pressuposto é que a MEE seja baseada em análise de covariâncias, logo os resultados obtidos são na forma não-padronizada, porém cabe ao pesquisador solicitar ou não resultados padronizados. Saídas padronizadas ignoram escala de medição das variáveis e trabalham com a mesma escala para diferentes relacionamentos. Em geral, pesquisadores utilizam valores padronizados para análises de variância dos fatores do modelo (Grace, 2006).

## 2.4 Programa AMOS

O programa utilizado neste trabalho foi o programa AMOS (*analysis of moment structural*) versão 22, desenvolvido pela IBM® (Internacional Business Machines), que executa facilmente a modelagem de equações estruturais.

A sua estrutura permite especificar, estimar, avaliar e apresentar modelos mostrando relacionamentos hipotéticos entre as variáveis. Este software permite que o usuário construa modelos de forma mais precisa do que com técnicas estatísticas multivariadas convencionais, sendo possível integrar uma interface gráfica com um método computacional avançado (Arbuckle, 2009; Byrne, 2009).

Com a interface gráfica o usuário dispõe de várias ferramentas para construção do diagrama de caminhos, que fornecem uma representação clara dos modelos construídos sem a necessidade de construir o sistema de equações. Os símbolos apresentados na Tabela (1) fazem parte de uma variedade disponível no pacote gráfico do AMOS.

Os métodos numéricos implementados no AMOS estão entre os mais eficazes e confiáveis (Arbuckle, 2009). Neste software são encontrados os métodos de estimação de máxima verossimilhança (ML), mínimos quadrados generalizados (GLS), mínimos quadrados não-ponderados (ULS) e mínimos quadrados ponderados (WLS). Além do ser possível utilizar análise bayesiana para melhorar as estimativas de parâmetros de modelo.

O programa AMOS ainda permite que os dados sejam inseridos de diferentes maneiras, através de matrizes de covariância ou correlação e através de planilhas contendo as amostras. Como saídas de resultados o programa calcula erros padrão e intervalos de confiança para todas as estimativas de parâmetros, estimativas de efeitos de trilha, estimativas padronizadas, médias, variância, covariância e correlações. Para maiores informações recomenda-se ler Arbuckle (2009) e Bryne (2009).

### 3. TRABALHOS

#### 3.1 REESPECIFICAÇÃO DE MODELOS DE EQUAÇÕES ESTRUTURAIS PARA O CICLO DO P EM SOLOS TROPICAIS

##### RESUMO

Reespecificação de modelos estruturais permite avaliar novas hipóteses e aumenta a compreensão de como as mudanças em um *pool* podem afetar os outros *pools* no modelo. O objetivo deste estudo foi avaliar o conceito teórico de *pools* de P (variáveis não mensuráveis - latentes), utilizando dados do método de extração sequencial de Hedley, para uma melhor compreensão do ciclo do P em solos tropicais. O conceito teórico testado foi o grau em que P disponível é controlado por diferentes *pools* de P no solo, conforme medido pela técnica de extração. Os modelos reespecificados foram adequados e capazes de representar uma generalização do ciclo P nos solos. O melhor modelo estrutural, incluindo variáveis latentes com múltiplos indicadores, foi consistente com o conceito teórico de que os indicadores dos *pools* de P do solo são as frações de P determinadas pelo método de extração sequencial. Neste modelo, não só existe uma relação direta entre os *pools* mineral primário, orgânico e ocluso com o *pool* de P disponível, mas relações indiretas via *pool* orgânico que foram teórica e

estatisticamente adequadas. Assim, o modelo mostrou as inter-relações de processos geoquímicos e biológicos sobre o P disponível e que corrobora com a hipótese da dependência do *pool* de P disponível com o *pool* orgânico em solos tropicais não fertilizados.

Palavras-chave: *Pool* de P orgânico, *Pool* de P ocluso, *Pool* de P mineral primário, Modelagem em equações estruturais.

## ABSTRACT

### RESPECIFICATION OF STRUCTURAL EQUATION MODELS FOR THE P CYCLE IN TROPICAL SOILS

Respecification of structural models allows evaluating new hypothesis and enhances understanding of how changes in one pool can affect the other pools in the model. The aim of this study was to evaluate the theoretical concept of P pools (unmeasurable variables – latent) using data from Hedley's sequential extraction method for a better understanding of the P cycle in tropical soils. The theoretical concept being tested is the degree to which available P is controlled by different soil P pools as measured by extraction techniques. The respecified models were adequate and able to represent a generalization of P cycling in soils. The best structural model including latent variables with multiple indicators was consistent to the theoretical concept that the indicators of soil P pools are P fractions determined by the sequential extraction method. In this model, not only is there a direct relationship between the organic, occluded and primary mineral pools and the available P pool, but the indirect relationships via the organic pool were theoretically and statistically adequate. Thus, the model showed the interrelationships of geochemical and biological processes on the available P and it corroborated the hypothesis of the dependence of the available P pool to the organic pool in unfertilized tropical soils.

Keywords: Organic P pool, Occluded P pool, Primary mineral P pool, Structural equation modeling.

## INTRODUÇÃO

Modelagem de equações estruturais (MEE) é uma metodologia científica que aspira a fazer uma forte e explícita conexão entre dados empíricos e conceitos teóricos utilizando variáveis latentes (Grace et al., 2010). Variáveis latentes são variáveis conceituais que formam e explicam um conjunto de variáveis observadas através de uma teoria e, por essa razão, elas são incorporadas em modelos de equações estruturais para tornar o modelo teoricamente mais consistente e testar hipóteses sobre vários processos que operam no sistema.

As inter-relações em MEE são representadas por dois tipos de modelos: o modelo estrutural, ou a associação de variáveis latentes (construtos), e o modelo mensurável, onde há uma relação entre as grandezas de medição e os seus respectivos construtos. Estes modelos são descritos pelas seguintes equações fundamentais (Bollen e Noble, 2011):

$$\eta_i = \alpha_\eta + B\eta_i + \Gamma\xi_i + \zeta_i \quad (\text{Modelo estrutural})$$

$$y_i = \alpha_y + \Lambda_y\eta_i + \varepsilon_i \quad (\text{Modelo de mensuração de } y)$$

$$x_i = \alpha_x + \Lambda_x\xi_i + \delta_i \quad (\text{Modelo de mensuração de } x),$$

onde os vetores  $\xi_i$  e  $\eta_i$  são, respectivamente, as variáveis latentes exógenas e endógenas; os vetores  $x_i$  e  $y_i$  são vetores das variáveis mensuráveis ou indicadores; os vetores  $\alpha_\eta$ ,  $\alpha_y$  e  $\alpha_x$  são vetores formados por intercepto para cada equação; as matrizes  $\Lambda_x$  e  $\Lambda_y$  são formadas pelos coeficientes relacionados aos indicadores das variáveis latentes; as matrizes B e  $\Gamma$  são matrizes de coeficientes representando os efeitos das variáveis latentes endógenas e exógenas sobre a variável latente endógena; e os vetores  $\zeta_i$ ,  $\varepsilon_i$  e  $\delta_i$  correspondentes aos erros aleatórios de cada equação construída. Em uma análise geral, as equações descrevem uma estrutura em que as causas são

vistas como variáveis latentes e as variáveis observáveis são as manifestações dos processos latentes (Grace et al., 2010).

A utilização de MEE é comum nas ciências naturais, especialmente em ecologia e biologia evolutiva (Mitchell, 1992; Pugesek et al., 2003; Grace et al., 2012). Na ciência dos solos, a MEE foi aplicada em parte ou na totalidade em estudos de transformação de fósforo (P) no solo.

Inicialmente, os modelos de caminhos foram utilizados para analisar a transformação de diferentes frações de P em solos temperados (Tiessen et al., 1984; Zheng et al., 2002) e tropicais (Beck e Sanchez, 1994). Recentemente, Gama-Rodrigues et al. (2014) aplicaram a modelagem de equações estruturais com variáveis latentes para testar modelos hipotéticos do ciclo do P em solos tropicais. No entanto, esses autores não encontraram uma adaptação adequada ao modelo estrutural básico com base na relação entre as cinco variáveis latentes (*pools* de P) com indicadores múltiplos (frações P), em que as hipóteses testadas foram os efeitos diretos do *pool* orgânico, *pool* inorgânico, *pool* ocluso e *pool* mineral sobre o *pool* de P disponível (Figura 1 – Modelo A).

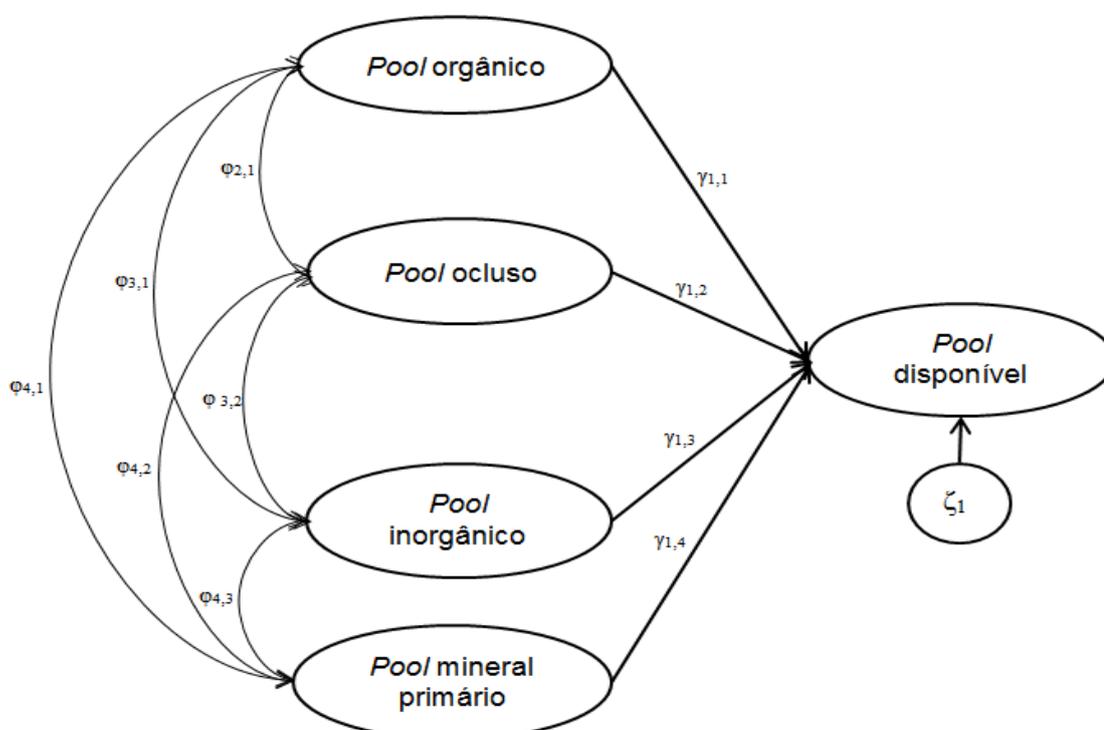


Figura 1. Modelo A, modelo estrutural hipotético para o ciclo do P do solo por Gama-Rodrigues et al. (2014).

Entretanto, o modelo ajustado mostrou o *pool* de P disponível como uma variável latente endógena diretamente dependente das frações de P (variáveis mensuráveis). Os resultados relatados por Gama-Rodrigues et al. (2014), portanto, enfatizaram o problema descrito por Gijsman et al. (1996) para os estudos de modelagem de P nos solos tropicais: " ... um problema na validação do modelo é atingir a congruência entre frações de P que podem ser medidas no solo e os *pools* conceituais (variáveis não mensuráveis - latentes) que representam tais frações no modelo".

Neste contexto, a modelagem de equações estruturais é uma técnica que permite a avaliação de hipóteses alternativas em modelos desenvolvidos, e considerando que um determinado modelo usado para representar um conjunto de dados pode ser melhorado a qualquer momento, os modelos estruturais em Gama-Rodrigues et al. (2014) foram reavaliados. Isto foi feito para testar a possibilidade de melhorar os parâmetros de ajuste e estimativas de índice para estes modelos, bem como para apresentar novos modelos criados através da reespecificação e reconstrução de modelos estruturais existentes. Assim, o objetivo deste estudo foi validar o conceito teórico dos *pools* de P usando dados do método de extração sequencial de Hedley para uma melhor compreensão do ciclo do P em solos tropicais. O conceito teórico a ser testado é o grau em que o P disponível é controlado por diferentes *pools* de P no solo, conforme medido por técnicas de extração.

## MATERIAL E MÉTODOS

Dados de P do solo foram coletados de Gama-Rodrigues et al. (2014) e corresponderam a estudos que utilizaram a técnica de fracionamento sequencial desenvolvida por Hedley et al. (1982) e modificada por Tiessen e Moir (1993). O banco de dados é composto por 81 observações que abrangem uma grande variação de tipos de solo não fertilizados (de Cambissolos e Argissolos a Luvisolos e Latossolos) sob diferentes sistemas de uso da terra na região tropical (floresta nativa, floresta secundária, agroflorestal, savana e pastagem).

As frações de P utilizadas foram resina P inorgânico (Pi), bicarbonato de P [Pi e fósforo orgânico (Po)], hidróxido P (Pi e Po), sônico P (Po-

NaOH+Ultrassom), HCl Pi, HCl P<sub>hot</sub> e P residual (Figura 2 ). A fração resina Pi é um índice de P trocável a que as plantas têm acesso removendo Pi a partir da solução do solo através da absorção (Sato e Comerford, 2006). As frações bicarbonato Pi e bicarbonato Po representam P que pode ser liberado por troca de ligantes com o íon bicarbonato (Sato e Comerford, 2006). Os hidróxidos Po e Pi, sônico Pi, sônico Po, HCl Pi, HCl P<sub>hot</sub> e frações de P residuais são formas de P que são muito mais difíceis de transportar para a solução do solo (Cross e Schlesinger, 1995; Tiessen e Moir, 1993).

Como as frações de P do solo baseadas no método de extração sequencial, não são variáveis verdadeiramente independentes, uma vez que todas elas têm que ser adicionadas para se obter o P total, então estas extrações não formam *pools* discretos. Deste modo, assumiu-se que os *pools* funcionais de P (variáveis não mensuráveis - latentes) derivados das várias frações de P são classificados em função do seu grau de disponibilidade para as plantas: formas prontamente disponíveis, compostas pelas frações resina Pi e bicarbonato (Pi e Po); formas com disponibilidade restrita, compostas das frações hidróxido (Pi e Po), sônico Po e HCl Pi; e formas que não estão disponíveis em curto prazo ou são recalcitrantes, compostas pelo HCl P<sub>hot</sub> e frações de P residuais.

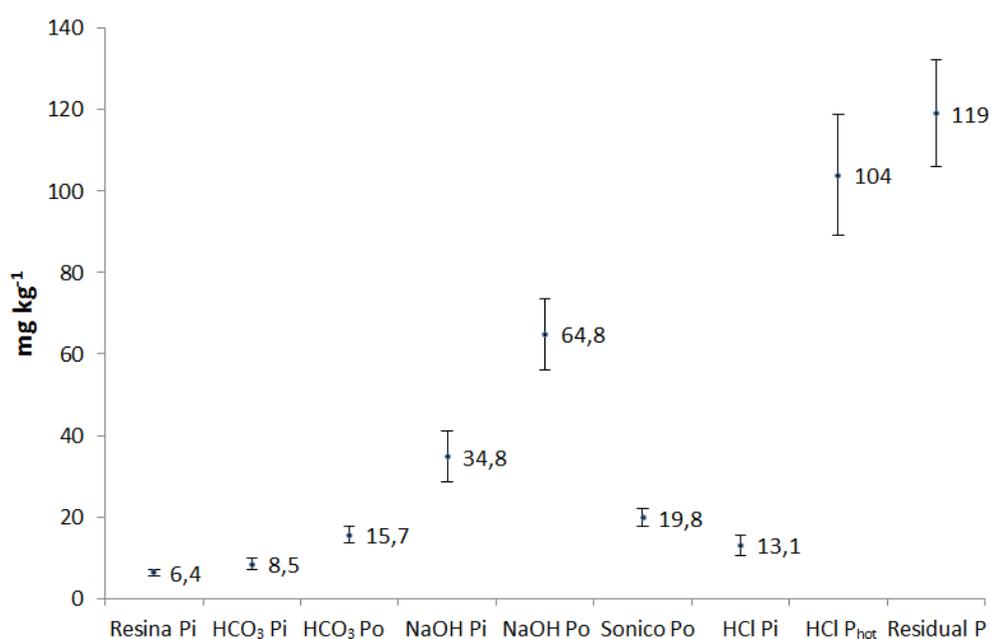


Figura 2. Média e erro padrão das frações de fósforo no solo (Fonte: Gama-Rodrigues et al. 2014).

O modelo estrutural básico para o ciclo do P (Figura 1 - Modelo A) proposto por Gama-Rodrigues et al. (2014) teve um alto valor de Chi-quadrado ( $\chi^2 = 165,61$ ) com 17 graus de liberdade e  $P < 0,0001$ , e os coeficientes de regressão para as relações entre os vários *pools* de P (variáveis latentes exógenas) e o *pool* de P disponível (variável latente endógena) não foram significativos.

Para a reespecificação do Modelo A, as seguintes hipóteses foram testadas:

H<sub>1</sub>: O *pool* de P orgânico atua diretamente como fonte de P para o *pool* de P disponível;

H<sub>2</sub>: O *pool* de P ocluso atua diretamente como dreno de P sobre o *pool* de P disponível;

H<sub>3</sub>: O *pool* de P inorgânico atua diretamente como fonte de P para o *pool* de P disponível; e

H<sub>4</sub>: O *pool* de P mineral primário atua diretamente como fonte de P para o *pool* de P disponível.

Outro modelo estrutural com ajuste geral insatisfatório, proposto por Gama-Rodrigues et al. (2014), foi também reespecificado (Figura 3 - Modelo B). Este modelo é uma versão simplificada do Modelo A, no qual os *pools* de P mineral e inorgânicos foram suprimidos e apenas o *pool* de P orgânico atuou diretamente como fonte de P para o *pool* de P disponível. As reespecificações dos Modelos A e B consistiram na inclusão ou não de uma ou mais variáveis latentes e suas relações, ou na fixação de parâmetros (Bollen e Noble, 2011). Estas reespecificações resultaram em cinco novos modelos estruturais, onde foi verificado que o Modelo C2 obteve o melhor ajuste.

Em primeiro lugar, foram feitas algumas alterações no modelo. A variância de erro da variável mensurável (indicador único) para o *pool* mineral, HCL Pi, foi assumida como 10% e este valor foi definido como uma quantidade fixa no modelo. Além disso, a variância do erro do indicador único, resina Pi, do *pool* de P disponível, foi definida como zero, assumindo que essa variável foi medida sem erro de medição.

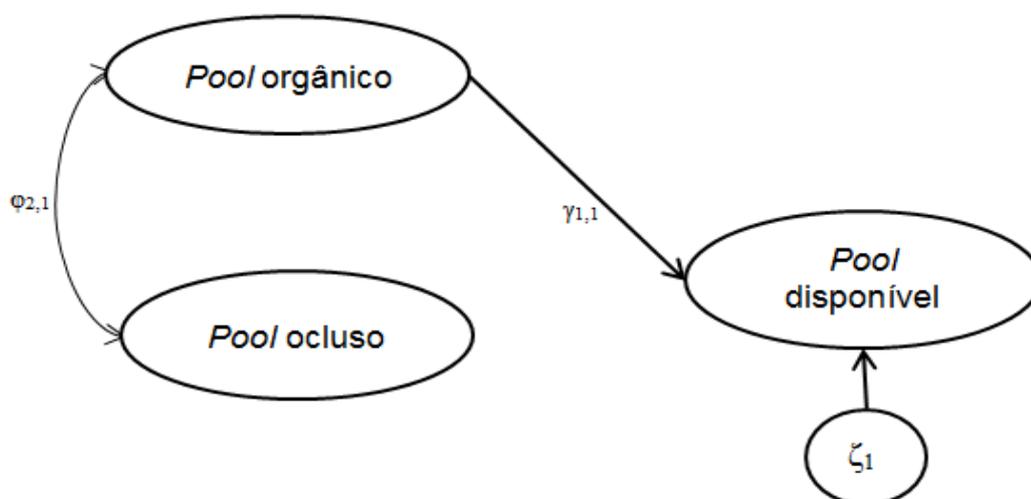


Figura 3. Modelo B, modelo estrutural hipotético para o ciclo do P no solo por Gama-Rodrigues et al. (2014).

Fixar parâmetros é um artifício utilizado para permitir o cálculo de todos os outros parâmetros livres no modelo: isto é, através da fixação de um ou mais parâmetros, o modelo pode ser identificado, gerando valores que podem então ser avaliados. Em geral, a não identificação de um modelo ocorre quando o grau de liberdade é menor do que zero; quando o número de parâmetros a ser estimado é maior do que o número de equações construídas; ou quando uma variável latente (construto) é formada com menos de três variáveis mensuráveis, isto é, um construto subidentificado que influencia todo o modelo.

Hair et al. (2009) estabelecem que, em geral, o valor da variância de erro é especificado como zero, mas que em alguns casos, um valor não nulo, tal como 1,0, pode ser utilizado para especificação e identificação futura do modelo. Dependendo da análise a ser conduzida, este valor pode ser alterado. Quando uma carga fatorial deve ser calculada e há um problema em identificar a sua construção, a variância do erro relacionada a esta carga fatorial pode ser fixada com um valor diferente de zero, pois os modelos utilizados são estatísticos e, como resultado, possuem erros de medição. Grace e Bollen (2008) recomendam o uso de 10% para a variância de erro em indicadores únicos. Assim, no presente estudo, todos os modelos testados foram reespecificados seguindo estas recomendações.

Pelo fato de Gama-Rodrigues et al. (2014) trabalharem com dados

transformados, os dados foram padronizados com média zero e desvio padrão 1. Também é importante validar a convergência dos construtos. A validade convergente de uma construção é determinada através do cálculo das cargas fatoriais. Depois de padronizar os dados, os valores das cargas fatoriais são índices importantes para a análise, porque cargas fatoriais com valores altos indicam que os indicadores convergem para um ponto comum (Hair et al., 2009). Assim, o significado dos fatores (*pools* de P) foi discutido em relação às suas respectivas variáveis medidas com base na padronização dos dados.

Ao utilizar os dados padronizados, outros parâmetros importantes, utilizados para avaliar se um construto foi bem elaborado são a variância extraída (*VE*) e confiabilidade do construto (*CR*). A *VE* indica a quantidade da variância dos indicadores que pode ser explicada pelos seus respectivos construtos, calculada como se segue:

$$VE = \frac{\sum_{i=1}^n \lambda_i^2}{\sum_{i=1}^n \lambda_i^2 + \sum_{i=1}^n \varepsilon_i}$$

onde  $\lambda_i$  representa cargas fatoriais padronizadas e  $\varepsilon_i$  representa os termos de erros aleatórios de cada indicador do construto em análise. A variância extraída é considerada adequada para valores acima de 0,5, o que sugere a convergência adequada dos seus indicadores (Hair et al., 2009).

A *CR* é também um indicador da validade convergente e pode ser calculado pela seguinte expressão:

$$CR = \frac{\left(\sum_{i=1}^n \lambda_i\right)^2}{\left(\sum_{i=1}^n \lambda_i\right)^2 + \left(\sum_{i=1}^n \varepsilon_i\right)}$$

em que os valores de confiabilidade  $> 0,7$  são considerados ideais, e valores entre 0,6 e 0,7 podem ser aceitáveis. No entanto, os valores de confiabilidade de outros construtos também devem ser bons para o modelo ser considerado adequado para análise (Hair et al., 2009).

Utilizou-se o método da máxima verossimilhança para estimar parâmetros. Para o ajuste geral dos modelos foram utilizados, o teste qui-quadrado ( $X^2$ ), graus de liberdade (*df*) e nível de probabilidade (*P*) associado ao

modelo. Devido à sensibilidade do teste do qui-quadrado ao tamanho da amostra, também foram consideradas alternativas para a medição ajuste do modelo índices de qualidade de ajuste (GFI), índice de ajuste comparativo (CFI), raiz quadrada média do erro de aproximação (RMSEA) e critério de informação de Akaike (AIC) (Grace, 2006; Iacobucci, 2010). No teste qui-quadrado, se a hipótese nula não é rejeitada ( $P > 0,05$ ), há evidências de que o modelo tem um bom ajuste. Valores da  $GFI \geq 0,95$  e  $CFI \geq 0,90$  sugerem um ajuste apropriado para o modelo (Byrne, 2009), ao passo que os valores de  $RMSEA \leq 0,10$  são indicativos de bom ajuste do modelo (Hair et al. 2009), e de forma semelhante, os valores baixos da AIC em comparação de modelos são sempre indicativos de um melhor ajuste. As análises de MEE foram realizadas com o programa AMOS, versão 22 (IBM SPSS Inc., Chicago, IL, EUA).

## RESULTADOS

O novo modelo reespecificado Modelo A1 (Figura 4) mostrou uma forte diminuição no valor qui-quadrado, mas o seu ajuste global foi insatisfatório ( $\chi^2 = 46,31$ ,  $df = 17$ ,  $P < 0,0001$ ). Além disso, o Modelo A1 teve um ajuste significativamente melhor em comparação com o Modelo A, baseado em uma diminuição significativa ( $P = 0,05$ ) no valor  $\Delta\chi^2$  entre os dois modelos ( $\Delta\chi^2 = 165,61_{(A)} - 46,31_{(A1)} = 119,30$ ).

Adicionalmente, o melhor ajuste do Modelo A1 também pode ser atribuído às relações diretas positivas entre o *pool* mineral primário e *pool* orgânico com o *pool* de P disponível, que foram significativas em 10% e 1%, respectivamente. Em contraste, as relações diretas entre os *pools* ocluso e inorgânico e do *pool* de P disponível foram negativas e não significativas. O Modelo A1, além de ter um melhor ajuste estatístico do que o Modelo A de acordo com o teste do qui-quadrado, obteve também um melhor ajuste com os índices  $GFI = 0,911$ ,  $CFI = 0,955$  e  $RMSEA = 0,154$  e mostrou uma diminuição em seu índice comparativo,  $AIC = 104$  (Tabela 1).

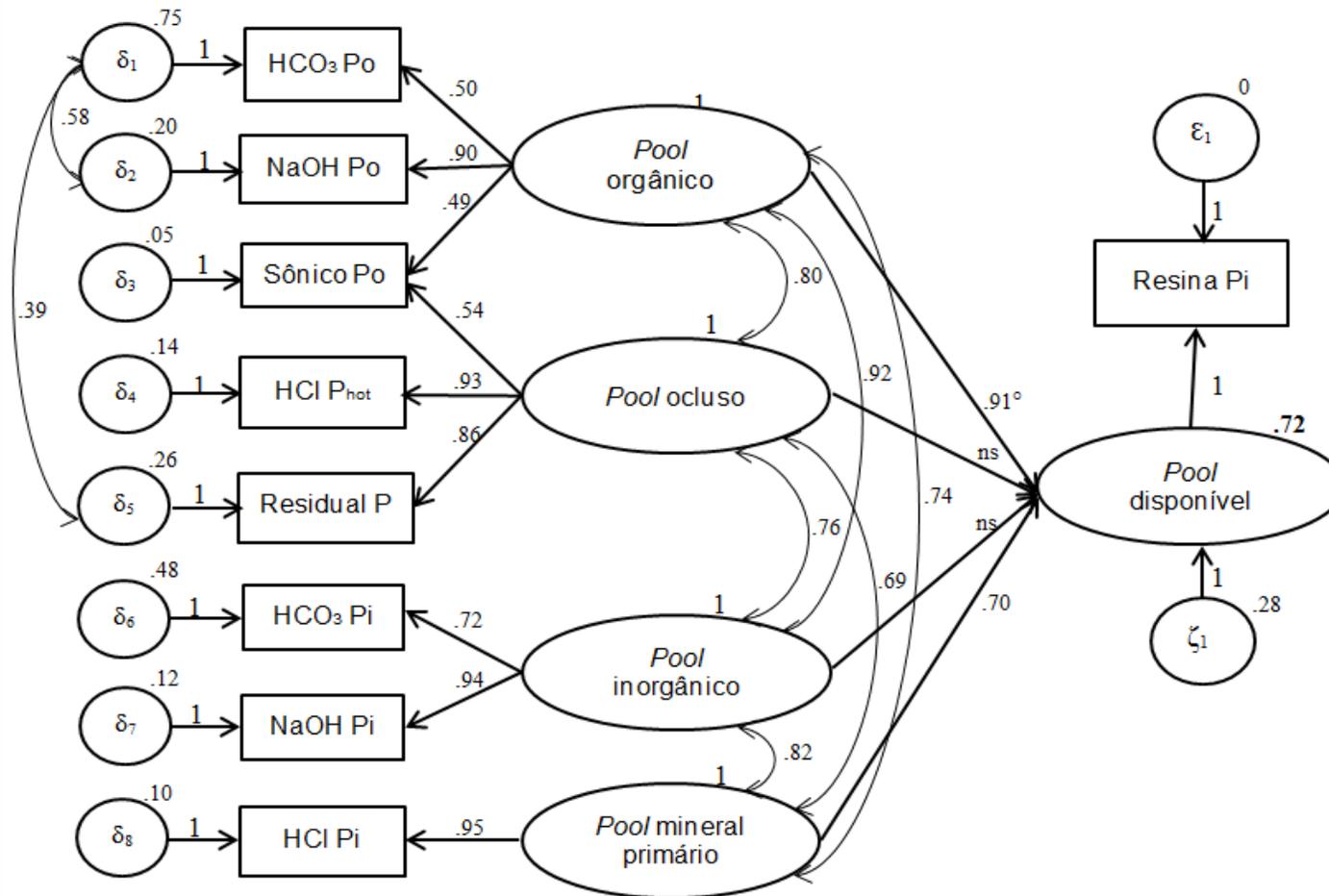


Figura 4. Modelo A1, modelo de equações estruturais para o ciclo do P do solo. Todas as variáveis mensuradas (em retângulos) estão representadas como indicadores de efeitos associados com as variáveis latentes (em círculos ou elipses). Os valores correspondem aos parâmetros estimados padronizados ( $P < 0.001$ ) e o valor do  $R^2$  (número em negrito). ° Significativo a  $P < 0,1$ . Valores das variáveis de erros ( $\delta_1 - \delta_8$ ,  $\varepsilon_1$ ,  $\zeta_1$ ) estão padronizados. Modelo  $\chi^2 = 46,305$ ,  $df = 17$ ,  $P < 0,0001$ .

Tabela 1. Índices de ajuste de modelos para comparação dos modelos (A e B) apresentados por Gama-Rodrigues et al. (2014) e modelos reespecificados.

Índices de Ajuste								
Modelo	$\chi^2$	df	P	GFI	CFI	RMSEA	AIC	$\chi^2/df$
<b>A</b>	165,61	17	***	0,828	0,777	0,331	221	9,706
<b>A1</b>	46,31	16	***	0,911	0,955	0,154	104,31	2,894
<b>B</b>	105,13	7	***	0,842	0,761	0,419	133	15,018
<b>B1</b>	5,93	5	0,313	0,976	0,998	0,048	37,93	1,186
<b>C</b>	10,31	8	0,244	0,966	0,995	0,060	50,755	1,289
<b>C1</b>	7,02	7	0,427	0,976	1,000	0,006	49,019	1,003

\*\*\*  $p < 0,0001$ .

No modelo A1, o *pool* orgânico apresentou uma VE de 56% e CR de 78,1%. O *pool* ocluso exibiu uma VE de 80,8% e CR de 92,3%. O *pool* inorgânico exibiu um VE de 70,1% e CR de 82,1%, e o *pool* mineral primário, porque continha apenas um indicador, apresentou valores de VE e CR igual a 90%, como sua variância do erro fixada em 10%. Outro resultado importante foi o de que a padronização do banco de dados gerou valores elevados significativos para os coeficientes de regressão ( $\beta$ ) entre as variáveis latentes exógenas (*pools* de P) e os respectivos indicadores (frações P) (Figura 4), enquanto no modelo A com a não padronização dos dados, os valores de  $\beta$  eram  $< 0,50$  (Gama-Rodrigues et al., 2014). Apenas a variável sônico Po, relacionada com o *pool* orgânico, apresentou um valor  $\beta < 0,50$  (Figura 4). Além disso, a variância explicada do *pool* de P disponível ( $R^2$ ) foi de 72%.

Para a construção do conjunto de equações apenas a relação direta de uma variável para outra é considerada. Os caminhos indiretos são omitidos porque eles estão necessariamente incluídos como formas diretas em uma das equações construídas. Neste modelo não existiam relações indiretas entre as variáveis latentes, todos os caminhos que levam a variável latente endógena eram caminhos diretos, ou seja, formados apenas por variáveis latentes exógenas. A variável mensurável resina Pi é dependente da variável latente endógena *pool* de P disponível no solo, que é dependente de todas as outras variáveis latentes exógenas no modelo, mostrando a relação de

interdependências entre as variáveis. Por sua vez, a variável sônico  $P_o$  se destaca porque recebe duas relações diretas de duas outras variáveis latentes exógenas, *pools* de P orgânico e ocluso.

A formulação do Modelo B1 (Figura 5) seguiu o mesmo critério adotado pela reespecificação do Modelo A1. No entanto, não há mais uma relação de três das quatro variáveis exógenas. A variável latente exógena *pool* de P ocluso ainda permanece no modelo, mas não tem efeito direto sobre o *pool* de P disponível (variável latente endógena), logo não faz parte da equação estrutural. Neste modelo, todos os índices de ajuste foram em geral satisfatórios (Tabela 1) e de todos os parâmetros estimados foram altamente significativos ( $P < 0,0001$ ), ao contrário dos resultados observados no Modelo B (Figura 3) apresentado por Gama-Rodrigues et al. (2014).

Em comparação com o Modelo A1, os valores de *VE* e *CR* para cada construto permaneceram adequados no Modelo B1 (dados não mostrados), mas houve uma diminuição de 17% do coeficiente de determinação ( $R^2 = 0,55$ ) do *pool* de P disponível (variável latente endógena) (Figura 4) e uma diminuição de 19% do coeficiente de regressão ( $\beta = 0,74$ ) da relação direta entre o *pool* orgânico e o *pool* de P disponível. A variação nos coeficientes foi causada, em parte, removendo os *pools* inorgânico e mineral primário. Portanto, este resultado justificou a recuperação de uma dessas variáveis exógenas para reestruturar o outro modelo alternativo.

Neste novo modelo reespecificado (Figura 6 - Modelo C), o *pool* de P mineral primário, que tinha uma relação direta significativa com o *pool* de P disponível no Modelo A1 (Figura 4), foi recuperado. No Modelo de C, todos os índices de ajuste geral foram satisfatórios (Tabela 1) e todos os parâmetros calculados foram altamente significativos ( $P < 0,0001$ ). Assim, o Modelo C recuperou algumas das relações do modelo estrutural completo (Modelo A - Figura 1), o aumento da variância explicada ( $R^2$ ) do *pool* de P disponível (variável latente endógeno) para 65% em relação ao Modelo B1. No entanto, os valores dos coeficientes de regressão para as relações diretas dos *pools* orgânicos e minerais, com o *pool* de P disponível no Modelo C foram menores do que nos Modelos A1 e B1 (Figuras 4, 5 e 6).

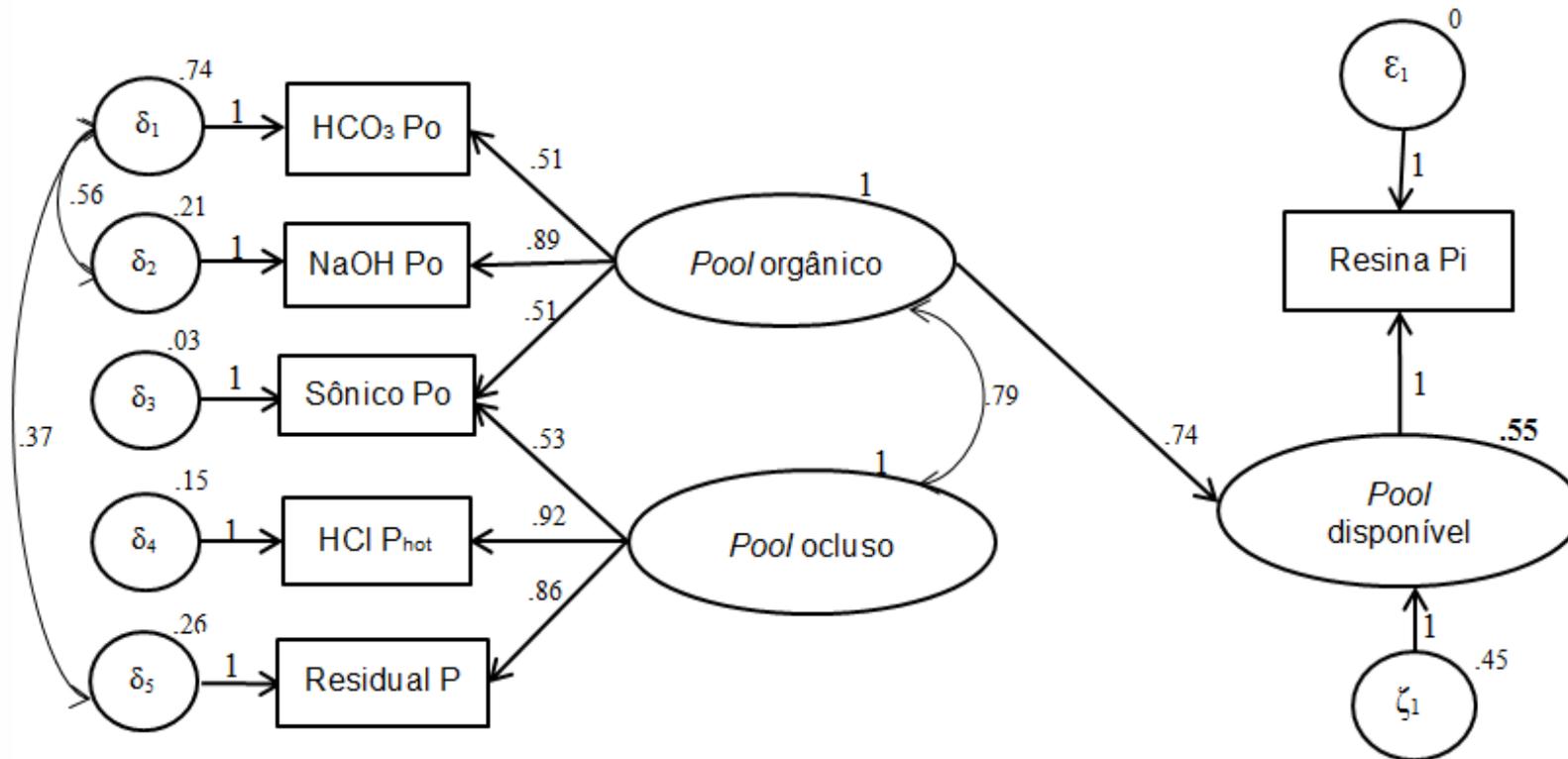


Figura 5. Modelo B1, modelo de equações estruturais para o ciclo de P do solo. Todas as variáveis mensuradas (em retângulos) estão representadas como indicadores de efeitos associados com as variáveis latentes (em círculos ou elipses). Os valores correspondem aos parâmetros estimados padronizados ( $P < 0.001$ ) e o valor do  $R^2$  (número em negrito). Valores das variáveis de erros ( $\delta_1 - \delta_5$ ,  $\varepsilon_1$ ,  $\zeta_1$ ) estão padronizados. Modelo  $\chi^2 = 5,93$   $df = 5$ ,  $P = 0,313$ .

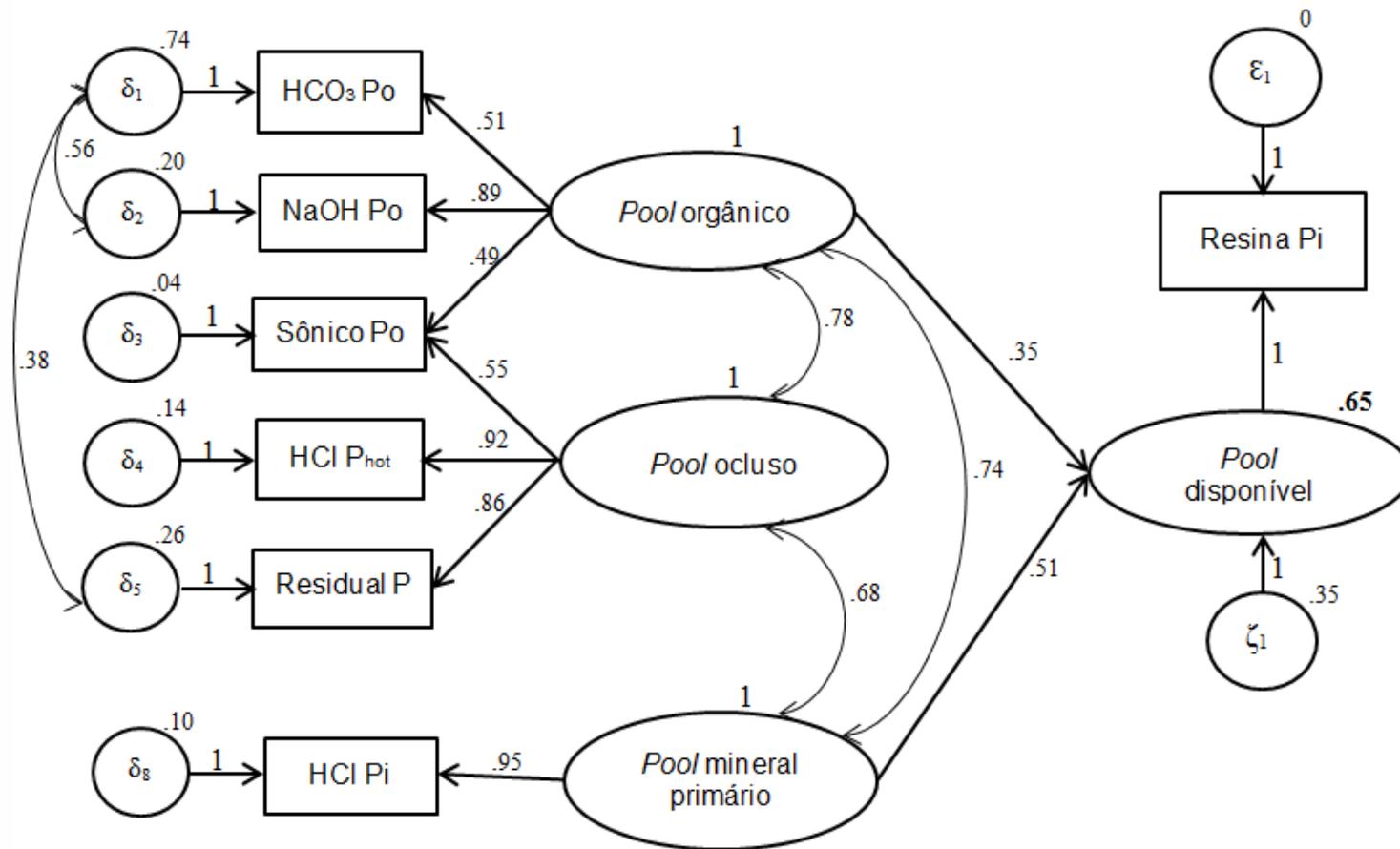


Figura 6. Modelo C, modelo de equações estruturais para o ciclo do P do solo. Todas as variáveis mensuradas (em retângulos) estão representadas como indicadores de efeitos associados com as variáveis latentes (em círculos ou elipses). Os valores correspondem aos parâmetros estimados padronizados ( $P < 0,001$ ) e o valor do  $R^2$  (número em negrito). Valores das variáveis de erros ( $\delta_1 - \delta_5$ ,  $\delta_8$ ,  $\varepsilon_1$ ,  $\zeta_1$ ) estão padronizados. Modelo  $\chi^2 = 10,31$   $df = 8$ ,  $P = 0,244$ .

O Modelo C, quando comparado com os Modelos A1 e B1, foi muito mais aproximado da configuração do modelo estrutural A1. Este modelo recupera o *pool* de P mineral primário, como variável latente exógena, e, conseqüentemente, a variável HCl Pi retorna no conjunto de equações. Este modelo ainda não está saturado, o que significa que algumas variáveis não estão diretamente relacionadas, mesmo com a inclusão da variável latente exógena *pool* de P mineral principal, pois este modelo foi construído ao longo do Modelo B1, que teve um ajuste adequado, mas faltava uma relação direta entre o *pool* de P ocluso e o *pool* de P disponível. Assim, como esta nova relação entre as variáveis latentes do Modelo C seria uma forma de reespecificação, onde o Modelo C1 foi criado (Figura 7). Foi então possível voltar a hipótese H<sub>2</sub> para o modelo, tornando este modelo mais parecido com o modelo estrutural teórico inicial (Modelo A - Figura 1).

O Modelo C1 teve um melhor ajuste do que o Modelo C para todos os índices de ajuste geral (Tabela 1), mesmo contendo um coeficiente de regressão não significativo para a relação direta entre o *pool* de P ocluso e o *pool* de P disponível. Esta relação entre esses dois *pools* conferiu maior estabilidade no Modelo C1. Neste modelo, os coeficientes de regressão para as relações diretas indicaram que os *pools* de P orgânico e P mineral agiram como fonte de P para o *pool* de P disponível e que o *pool* de ocluso agiu como um dreno de P. Os três *pools* juntos explicaram 70% da variação ( $R^2$ ) do *pool* de P disponível (Figura 7), deixando apenas 2% da variância explicada pelo *pool* inorgânico, que foi parte do modelo estrutural de base teórica.

Pelo fato do Modelo C1 ser composto de covariâncias entre as variáveis exógenas e não estar saturado, ele poderia ser melhorado através de relações diretas entre os *pools* exógenos. Assim, no modelo reespecificado C2 foram incluídos os *pools* orgânico e ocluso, como variáveis endógenas, juntamente com o *pool* de P disponível (Figura 8). Neste novo modelo, não houve alteração nos valores dos coeficientes de regressão ( $\beta$ ) para a relação direta entre os demais *pools* e o *pool* de P disponível, mas a reespecificação mostrou a relação direta entre o *pool* de P ocluso como um dreno de P (valor negativo) e o *pool* de P disponível tendo um valor  $\beta$  significativo ( $P < 0,001$ ), enquanto que os *pools* de P mineral e orgânico permaneceram fonte de P (valor positivo).

Outra mudança foi que a relação indireta entre os *pools* de P mineral,

ocluso e do *pool* de P disponível também foi positivamente significativa. O *pool* de P mineral primário teve efeitos indiretos através dos *pools* de P orgânico e ocluso, e o *pool* de P ocluso teve um efeito indireto via *pool* de P orgânico. O maior efeito direto foi a partir do *pool* orgânico ( $\beta = 0,61$ ), mas o efeito global maior foi a partir do *pool* mineral primário ( $\beta = 0,69 \times 0,57 \times 0,61 + 0,36 \times 0,61 + 0,69 \times (-0,32) + 0,53 = 0,774$ ) em contraste, o *pool* ocluso, que teve um efeito direto negativo sobre o *pool* de P disponível, teve um baixo efeito global positivo ( $\beta = 0,57 \times 0,61 - 0,32 = 0,028$ ). As modificações feitas para construir o Modelo C2 não alteraram os valores dos índices de ajuste do modelo geral (dados não publicados), o que resulta em um ajuste igual ao Modelo C1.

Ao contrário dos modelos anteriores, o Modelo C2 exibiu relações mais dependentes entre as variáveis latentes (*pools* de P). A principal equação que mede a variável latente endógena *pool* de P disponível tornou-se dependente de outras duas variáveis latentes endógenas, *pools* de P orgânico e ocluso, e da variável latente exógena *pool* de P mineral. Por sua vez, a equação que mede o *pool* de P orgânico também é dependente dos *pools* de P ocluso e mineral, e o *pool* de P ocluso é dependente da variável latente exógena *pool* de P mineral. Assim, para analisar a equação principal, é necessário encontrar o valor da variável latente endógena *pool* de P ocluso, mostrando, assim, a interdependência entre as variáveis latentes.

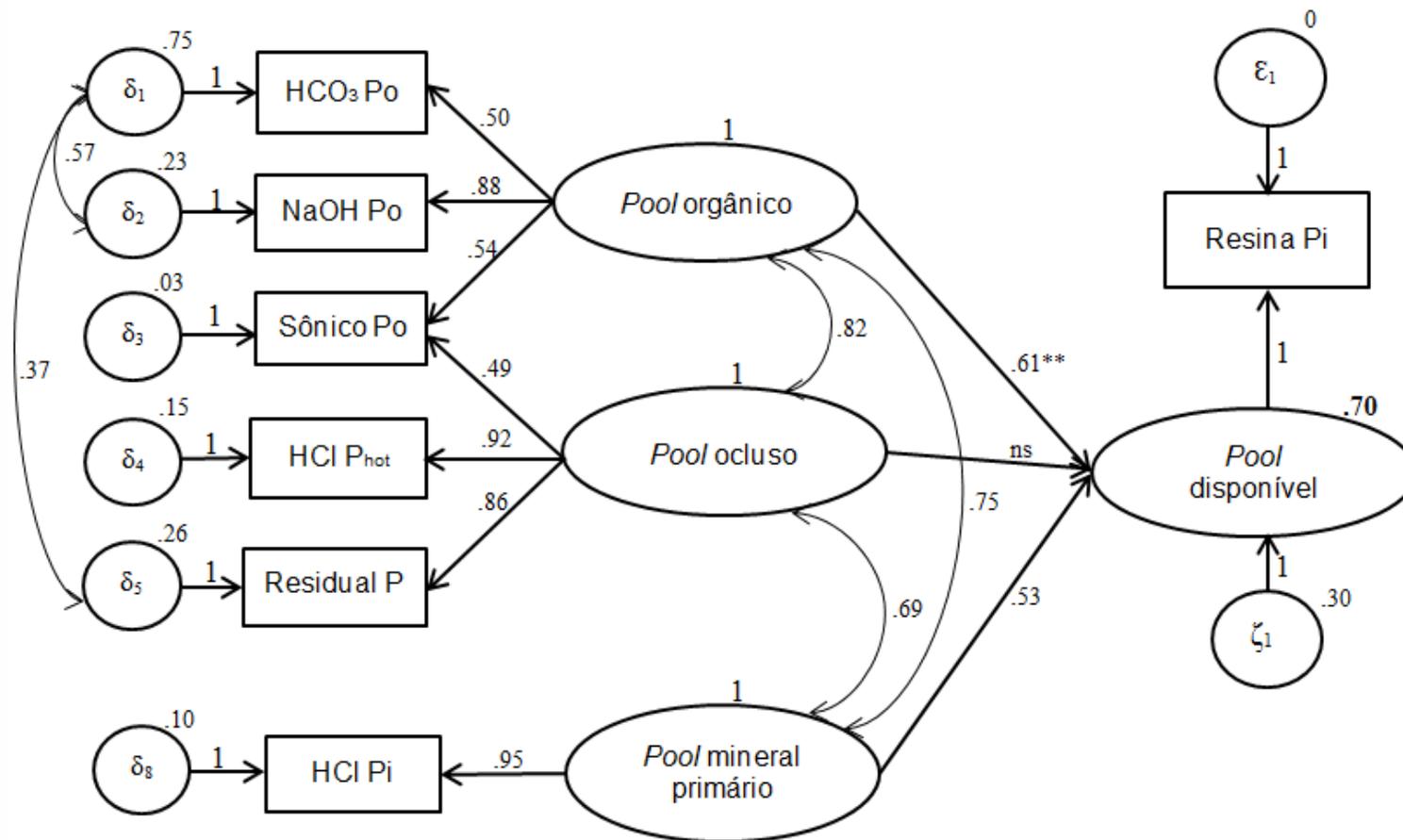


Figura 7. Modelo C1, modelo de equações estruturais para o ciclo do P do solo. Todas as variáveis mensuradas (em retângulos) estão representadas como indicadores de efeitos associados com as variáveis latentes (em círculos ou elipses). Os valores correspondem aos parâmetros estimados padronizados ( $P < 0,001$ ) e o valor do  $R^2$  (número em negrito). \*\* Significativo a  $P < 0,01$ . Valores das variáveis de erros ( $\delta_1 - \delta_5$ ,  $\delta_8$ ,  $\varepsilon_1$ ,  $\zeta_1$ ) estão padronizados. Modelo  $\chi^2 = 7,02$ ,  $df = 7$ ,  $P = 0,427$ .

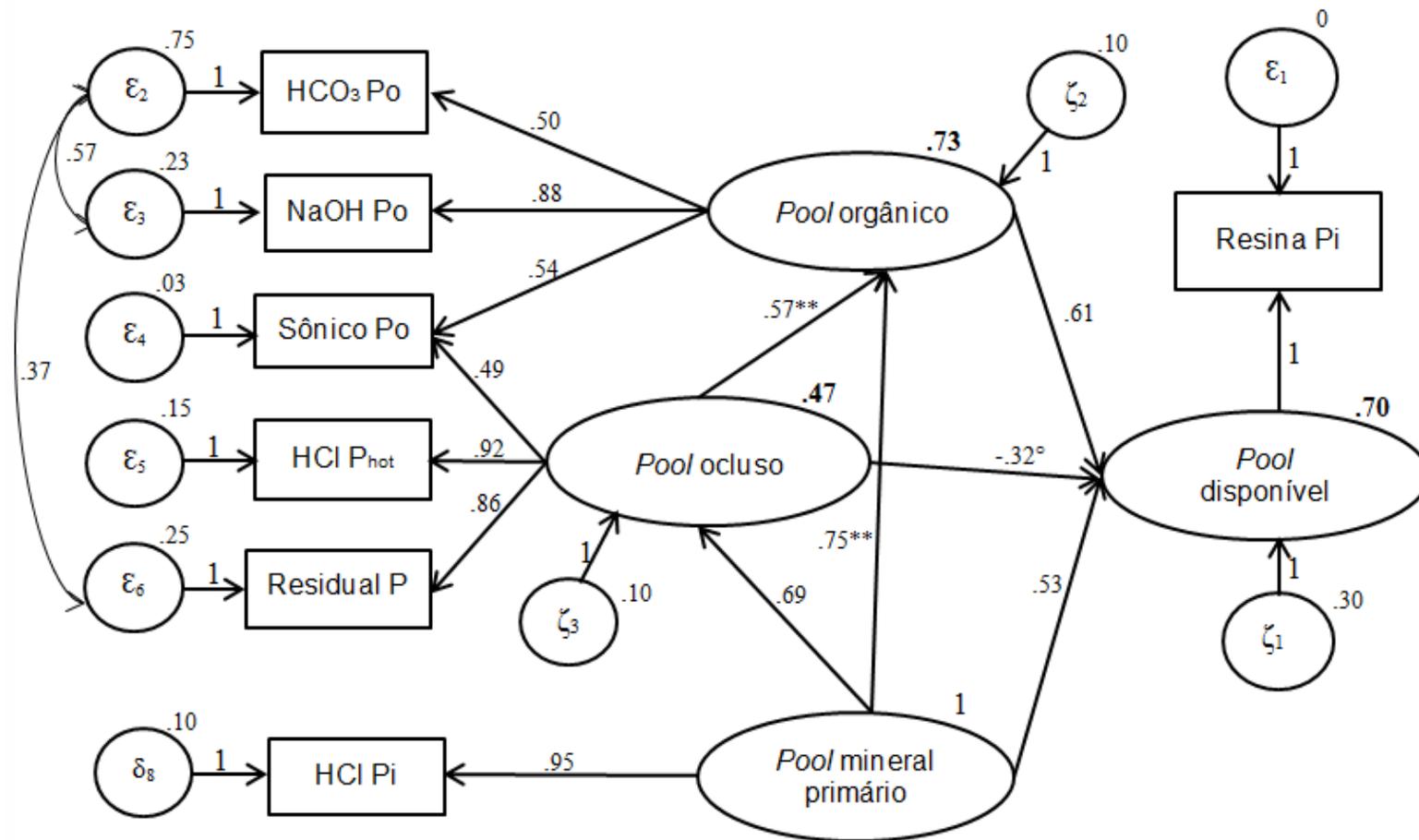


Figura 8. Modelo C2, modelo de equações estruturais para o ciclo do P do solo. Todas as variáveis mensuradas (em retângulos) estão representadas como indicadores de efeitos associados com as variáveis latentes (em círculos ou elipses). Os valores correspondem aos parâmetros estimados padronizados ( $P < 0.001$ ) e o valor do  $R^2$  (número em negrito). <sup>o</sup> e <sup>\*\*</sup> Significante a  $P < 0,1$  e  $0,01$ , respectivamente. Valores das variáveis de erros ( $\epsilon_1 - \epsilon_6$ ,  $\delta_8$ ,  $\zeta_1 - \zeta_3$ ) estão padronizados. Modelo  $\chi^2 = 7,02$ ,  $df = 7$ ,  $P = 0,427$ .

## DISCUSSÃO

Embora as teorias simples sejam atraentes, nenhuma teoria do ciclo do P pode ou deve ser representada por um único modelo de equações estruturais. Dependendo da reespecificação conduzida, se realizada corretamente e sem alterar a ideia geral do modelo hipotético, resultados muito aproximados podem ser obtidos, que só exigem então o bom senso para escolha do melhor modelo. Este julgamento pode estar relacionado com análises predefinidas ou a um conjunto de critérios, como o caso dos índices de ajuste do modelo. As modificações feitas neste estudo recuperaram hipóteses que haviam sido refutadas nos modelos apresentados por Gama-Rodrigues et al. (2014).

Os modelos construídos a partir das reespecificações exibiram uma forte base teórica e ajuste aos dados, representando bem a teoria do ciclo do P em solos tropicais. A análise de *VE* e *CR* mostrou que os indicadores de cada um dos *pools* de P foram agrupados de forma adequada. Altos valores de *VE* e de *CR* indicam a existência de consistência interna, o que significa que todas as medidas representam de forma consistente o mesmo construto latente (Hair et al. 2009).

Assim, o Modelo A1, primeira reespecificação, mostrou possibilidades de futuras modificações, apesar de ajuste insatisfatório do modelo. Neste modelo, apenas os *pools* de P mineral primário e P orgânico tornaram-se significativos em comparação com o modelo hipotético de base (Modelo A); ambos os *pools* representando fontes de P para o *pool* de P disponível. A reespecificação do Modelo B gerou o novo Modelo B1, sendo adequado em todos os testes de ajuste do modelo, e todos os parâmetros resultantes foram altamente significativos, reforçando a hipótese de que o *pool* orgânico é uma fonte direta de P para o *pool* de P disponível.

A partir dos resultados observados em Modelos A1 e B1, foi possível recuperar o *pool* de P mineral primário que havia sido refutado no Modelo A. Este modelo permitiu a construção de três modelos em sequência, todos ajustados de forma satisfatória. Modelos C1 e C2 são estatisticamente equivalentes, pois ambos têm o mesmo número de graus de liberdade e os mesmos valores para os índices de ajuste global, mas eles diferem um do outro sob a forma das equações estruturais construídas e pela presença de relações indiretas entre os *pools* de P

no Modelo C2. Estes resultados caracterizam o Modelo C2 como mais complexo e, ao mesmo tempo, mais próximo do modelo hipotético de base (Modelo A), porque existe uma base teórica para todas as relações entre os *pools* de P. Por conseguinte, este modelo apresenta as inter-relações dos processos geoquímicos e biológicos sobre o P disponível, mas este modelo necessita de confirmação utilizando dados independentes.

No Modelo C2, a função do *pool* mineral como uma fonte de P para o *pool* de P disponível sugere que alguns dos tipos de solos utilizados para a modelagem não sejam fortemente intemperizados, embora este *pool* represente, em média, apenas 3,3% do P total extraído (Figura 2 - soma de todas as frações de P). Esta fração, como um indicador do *pool* mineral primário, representa o P ligado ao Ca (apatita) e, portanto, formas mais resistentes de P. Assim, em solos tropicais que não fertilizados ou não fortemente intemperizados, a disponibilidade de P pode ser diretamente regulada pelo *pool* mineral primário (taxa de dissolução P-Ca) quando o *pool* de P orgânico não é grande o suficiente para sustentar requisitos de produção (Araujo et al., 2004). Por outro lado, os efeitos indiretos do *pool* mineral primário no *pool* de P disponível eram complexos no Modelo C2. O efeito através do *pool* ocluso foi negativo ( $\beta = 0,69 \times (-0,32) = -0,22$ ), mas os efeitos através dos *pools* ocluso e orgânico foram positivos ( $\beta = 0,69 \times 0,57 \times 0,61 + 0,36 \times 0,61 = 0,46$ ).

Estes resultados sugerem que uma relação estreita entre o *pool* mineral primário e o *pool* orgânico pode ocorrer, em parte, porque a fração de HCl Pi deve ter um teor considerável de Po estável (Po-Ca). Dieter et al. (2010) relataram a presença de Po na fração HCl Pi, o que correspondeu a 6,5% do Po total e 1,7% do total de P, em alguns solos com regimes distintos de mineralogia e intemperismo no Panamá. Em alguns solos de clima temperado, quantidades significativas de Po também foram encontradas na fração HCl Pi (He et al., 2006). Assim, parte do efeito direto do *pool* de P mineral primário como fonte de P para o *pool* de P disponível pode ser atribuída mais à mineralização de formas Po-Ca e menos para a dissolução das formas Pi-Ca. Além disso, a relação entre os *pools* de P mineral primário e ocluso mostra a natureza recalcitrante destes dois *pools* no solo.

O efeito direto do *pool* ocluso como um dreno de P sobre o *pool* de P disponível pode ser atribuído à ligação deste aos minerais de ferro e alumínio,

enquanto o seu efeito positivo indireto, através do *pool* orgânico, pode ser atribuído à presença de formas estáveis de Po nas frações HCl P<sub>hot</sub> e residual P (Tiessen e Moir, 1993; Oberson et al., 2001), em adição à própria fração sônico Po, como variável indicadora deste *pool*.

No Modelo C2, como o *pool* orgânico exibiu uma maior contribuição relativa para a variância explicada ( $R^2$ ) do *pool* de P disponível, e considerando a sua relação positiva com os *pools* de P ocluso e mineral primário devido ao conteúdo significativo Po neste último, há assim uma forte evidência de que o *pool* de P disponível seja controlado pela mineralização de todas as formas Po. Neste contexto, o estresse de P no solo estimula organismos que podem adquirir P de formas mais recalcitrantes de P orgânico do solo (Turner et al., 2014).

Plantas e microrganismos podem produzir mais enzimas fosfatases extracelulares que eliminam ligações éster-P na matéria orgânica do solo (McGill e Cole 1981; Nasto et al., 2014). Espécies de plantas diferentes podem acessar *pools* de Po discretos, devido à sua capacidade de produzir enzimas fosfatases específicas que têm como alvo uma variedade de compostos de Po (Turner, 2008). Além disso, o baixo nível de P no solo também aumenta a colonização micorrízica arbuscular que é outra estratégia de aquisição de P (Lambers et al., 2008). Portanto, essas duas estratégias de aquisição de nutrientes podem reforçar a capacidade das plantas para adquirir P do solo e acelerar o ritmo do ciclo, mas não a quantidade de P dentro dos ecossistemas (Wang et al., 2007; Vitousek et al., 2010; Marklein e Houlton, 2012).

Além disso, os ânions orgânicos de exsudatos de raízes (Richardson et al., 2011) ou derivado da degradação da matéria orgânica do solo (Guppy et al., 2005) podem aumentar a solubilidade do Po ou facilitar mineralização aumentando a solubilidade do substrato Po (Reed et al., 2011). Ânions orgânicos podem aumentar a biodisponibilidade do Pi no solo através da mobilização do *pool* P moderadamente disponível por meio de uma solução de reações de ligantes trocáveis ou eles podem também reduzir a capacidade do solo de fixar P e competir com ortofosfato (liberado a partir da matéria orgânica) por sítios de sorção em oxihidróxidos de Fe e Al (Guppy et al., 2005). Estes mecanismos, no entanto, não são diretamente inferidos no Modelo C2 devido à ausência de outros agrupamentos de solo que afetam o ciclo do P, tal como o C orgânico e biomassa microbiana (*pool* orgânico do solo), e oxihidróxidos de Fe e Al (*pool* químico do

solo). Esses resultados, portanto, sugerem a necessidade de refinar o Modelo C2 para uma investigação mais aprofundada da biodisponibilidade de diferentes formas de P, considerando que as plantas têm mecanismos diferentes para extrair P dos *pools*.

## CONCLUSÕES

Em geral, os modelos reespecificados foram adequados e capazes de representar uma generalização do ciclo P nos solos que têm características semelhantes como os dados amostrados. Eles não invalidam os modelos apresentados por Gama-Rodrigues et al. (2014), mas ao invés disso foram apresentados como alternativas para estudos futuros.

A construção do modelo estrutural C2 como um modelo contendo variáveis latentes com vários indicadores mostrou ser o mais robusto dos modelos testados porque gerou novas hipóteses e recuperou os pressupostos iniciais do modelo de base (Modelo A).

O Modelo C2 validou o conceito teórico que os indicadores de *pools* de P do solo são frações de P determinadas pelo método de extração sequencial. Neste modelo, não só existe uma relação direta entre os *pools* de P mineral primário, orgânico e ocluso sobre o *pool* de P disponível, mas as relações indiretas via *pool* orgânico foram teórica e estatisticamente adequadas.

O Modelo C2 corroborou com a hipótese da dependência do *pool* de P disponível pelo *pool* orgânico em solos tropicais não fertilizados.

## REFERÊNCIAS BIBLIOGRÁFICAS

- Beck MA, Sanchez PA (1994). Soil phosphorus fraction dynamics during 18 years of cultivation on a Typic Paleudult. *Soil Sci Soc Am J* 58:1424-1430.
- Bollen KA, Noble MD (2011). Structural equation models and the quantification of behavior. *Proceeding of the National Academy of Sciences* 108:15639-15646.

- Byrne BM (2009). Structural equation modeling with AMOS: basic concepts, applications and programming, 2ª edição. Routledge, 2009.
- Cross AF, Schlesinger WH (1995). A literature review and evaluation of the Hedley fractionation: application to the biogeochemical cycle of soil phosphorus in natural ecosystems. *Geoderma* 64:197-214.
- Dieter D, Elsenbeer H, Turner B (2010). Phosphorus fractionation in lowland tropical rainforest soils in central Panama. *Catena* 82:118-125.
- Gama-Rodrigues AC, Sales MVS, Silva PSD, Comerford NB, Cropper WP, Gama-Rodrigues EF (2014). An exploratory analysis of phosphorus transformations in tropical soils using structural equation modeling. *Biogeochemistry* 118:453-469.
- Gijssman AJ, Oberson A, Tiessen H, Friesen DK (1996). Limited applicability of the CENTURY model to highly weathered tropical soils. *Agron J* 88:894-903.
- Grace JB (2006). Structural equation modeling and natural systems. Cambridge University Press, New York, p 365.
- Grace JB, Anderson TM, Olf H, Scheiner SM (2010). On the specification of structural equation models for ecological systems. *Ecol Monogr* 80:67-87.
- Grace JB, Bollen KA (2008). Representing general theoretical concepts in structural equation models: the role of composite variables. *Environ Ecol Stat* 15:191-213.
- Grace JB, Schoolmaster DR, Guntenspergen GR, Little AM, Mitchell BR, Miller KM, Schweiger EW (2012). Guidelines for a graph-theoretic implementation of structural equation modeling. *Ecosphere* 3:1-44.
- Guppy CN, Menzies NW, Moody PW, Blamey FPC (2005). Competitive sorption reactions between phosphorus and organic matter in soil: a review. *Australian J Soil* 43:189-202.
- Hair JF, Black WC, Babin BJ, Anderson E, Tathan RL (2009). Multivariate data analysis. Prentice-Hall Inc., Upper Saddle River, p 688.
- He Z, Fortuna AM, Senwo ZN, Tazisong IA, Honeycutt CW, Griffin TS (2006). Hydrochloric fractions in Hedley fractionation may contain inorganic and organic phosphates. *Soil Sc Soc Am J* 70:893-899.
- Hedley MJ, Stewart WB, Chauhan BS (1982). Changes in inorganic and organic soil phosphorus fractions induced by cultivation practices and by laboratory incubations. *Soil Sci Soc Am J* 46:970-976.

- Iacobucci D (2010). Structural equations modeling: Fit Indices, sample size, and advanced topics. *Journal of Consumer Psychology*, 20, 90 - 98.
- Lambers H, Raven JA, Shaver GR, Smith S (2008). Plant nutrient-acquisition strategies change with soil age. *Trends in Ecology and Evolution* 23:95-103
- Marklein AR, Houlton BZ (2012). Nitrogen inputs accelerate phosphorus cycling rates across a wide variety of terrestrial ecosystems. *New Phytologist* 193:696-704.
- McGill WB, Cole CV (1981). Comparative aspects of cycling of organic C, N, S, and P through soil organic matter. *Geoderma* 26:267-286.
- Mitchell RJ (1992). Testing evolutionary and ecological hypotheses using path analysis and structural equation modeling. *Funct Ecol* 6:123-129.
- Nasto MK, Alvarez-Clare S, Lekberg Y, Sullivan BW, Townsend AR, Cleveland CC (2014). Interactions among nitrogen fixation and phosphorus acquisition strategies in lowland tropical rain forests. *Ecology Letters* 17:1282-1289.
- Oberson A, Friesen DK, Rao IM, Bühler S, Frossard E (2001). Phosphorus transformations in an Oxisol under contrasting land use systems: role of the soil microbial biomass. *Plant Soil* 237:197-210.
- Pugesek BH, Tomer A, Eye A (2003). *Structural equation modeling: Applications in ecological and evolutionary biology*. Cambridge University Press, New York, p 409.
- Reed SC, Townsend AR, Taylor PG, Cleveland CC (2011). Phosphorus cycling in tropical forests growing on highly weathered soils. In: Bünemann, EK, Oberson A, Frossard E (ed) *Phosphorus in action: Biological processes in soil phosphorus cycling*. Springer, Berlin, pp 339-369.
- Richardson AE, Lynch JP, Ryan PR, Delhaize E, Smith FA, Smith SE, Harvey PR, Ryan MH, Veneklaas EJ, Lambers H, Oberson A, Culvenor RA, Simpson RJ (2011). Plant and microbial strategies to improve the phosphorus efficiency of agriculture. *Plant Soil* 349:121-156.
- Sato S, Comerford NB (2006). Organic anions and phosphorus desorption and bioavailability in a humid Brazilian Ultisol. *Soil Sci* 171:695-705.
- Tiessen H, Moir JO (1993). Characterisation of available P by sequential extraction. In: Carter MR (ed) *Soil sampling and methods of soil analysis*. CRC Press, Boca Raton, FL, pp 75-86.

- Tiessen H, Stewart WB, Cole CV (1984). Pathways of phosphorus transformation in soils of differing pedogenesis. *Soil Sci Soc Am J* 48:853-858.
- Turner BL (2008). Resource partitioning for soil phosphorus: a hypothesis. *J Ecology* 96:698-702.
- Turner BL, Wells A, Condron LM (2014). Soil organic phosphorus transformations along a coastal dune chronosequence under New Zealand temperate rain forest. *Biogeochemistry* 121:595-611.
- Vitousek PM, Porder S, Houlton BZ, Chadwick OA (2010). Terrestrial phosphorus limitation: mechanisms, implications, and nitrogen-phosphorus interactions. *Ecological Applications* 20:5-15.
- Wang YP, Houlton BZ, Field CB (2007). A model of biogeochemical cycles of carbon, nitrogen, and phosphorus including symbiotic nitrogen fixation and phosphatase production. *Global Biogeochemical Cycles* 21: GB1018, doi:10.1029/2006GB002797.
- Zheng Z, Simard RR, Lafond J, Parent LE (2002). Pathways of soil phosphorus transformations after 8 years of cultivation under contrasting cropping practices. *Soil Sci Soc Am J* 66:999-1007.

### 3.2. INTERAÇÃO DO CICLO DO P COM PROPRIEDADES DO SOLO USANDO MODELAGEM DE EQUAÇÕES ESTRUTURAIS

#### RESUMO

O uso da modelagem de equações estruturais (MEE) proporciona a investigação de múltiplos relacionamentos interdependentes que envolvem dados empíricos e conceitos teóricos com a utilização de variáveis latentes. Com esta característica, aplicações nos estudos do ciclo do P no solo têm aumentado nos últimos anos. O modelo hipotético básico analisado neste estudo do ciclo do P no solo consistiu dos *pools* de P estável (Pi e Po), P lábil (Pi e Po) e P disponível (P-Mehlich-1) associados aos *pools* matéria orgânica (C orgânico e N total) e atributos físico-químicos (óxidos de Al e Fe, argila e pH) do solo. O objetivo deste estudo foi avaliar a influência das interações entre os *pools* de P (lábil e estável) e *pools* matéria orgânica e atributos físico-químicos do solo sobre o *pool* de P disponível em solos tropicais. Para isso, dois bancos de dados foram utilizados. Um banco de dados com 68 amostras referentes a solos minerais com baixo teor de matéria orgânica e outro com 38 amostras constituídas de solos minerais com alto teor de matéria orgânica e solos orgânicos. O modelo hipotético básico, com a utilização de solos minerais de baixo teor de matéria orgânica, foi ajustado com  $X^2 = 14,374$ ,  $df = 11$  e  $p = 0,213$ , sem a presença do *pool* de atributos físico-químicos, e para

confirmação deste modelo ajustado, foram utilizados os dados de solos minerais de alto teor de matéria orgânica, onde na confirmação do modelo foi obtido um ajuste global com  $\chi^2 = 6,004$ ,  $df = 10$  e  $p = 0,815$ . Com a utilização deste último banco de dados, foi gerado um modelo alternativo, que foi ajustado com  $\chi^2 = 18,952$ ,  $df = 16$  e  $p = 0,271$  sem o *pool* de matéria orgânica, porém com a recuperação do *pool* de atributos físico-químicos. Em geral, nos modelos estruturais ajustados, os *pools* de P estável e lábil foram considerados fontes de P sobre o *pool* de P disponível. A partir dos modelos apresentados foi possível verificar as inter-relações entre os conceitos teóricos e confirmar as hipóteses de dependência entre os *pools*.

Palavras-chave: Método Bowman, Variáveis latentes, Solos tropicais.

## ABSTRACT

### INTERACTION OF P CYCLE WITH SOIL PROPERTIES USING STRUCTURAL EQUATION MODELING

The use of structural equation modeling (SEM) provides the investigation of multiple interdependent relationships that involve empirical data and theoretical concepts using latent variables. With this feature, applications in studies of the phosphorus cycle in soil have increased in recent years. The basic hypothetical model analyzed in this study of the phosphorus cycle in the soil consisted of stable P pool (Pi and Po), labile P pool (Pi and Po) and available P pool (P-Mehlich-1) associated to pools organic matter (organic C and total N) and physical-chemical properties (Al and Fe oxides, clay and pH) of the soil. The objective of this study was to evaluate the influence of interactions between pools of P (labile and stable) and organic matter pools and physico-chemical soil on the pool of available P in tropical soils. For this, two databases are used. A database with 68 samples relating to mineral soils with low organic matter content and the other with 38

samples consisting of mineral soils high in organic matter and organic soils. The basic hypothetical model, with the use of mineral soils of low content organic matter, was adjusted with  $X^2 = 14,374$ ,  $df = 11$  and  $p = 0,213$ , in the absence of the pool of physical-chemical properties, and to confirm this adjusted model, the mineral soil data of high organic matter content were used, where in confirming the model was obtained a global setting with  $X^2 = 6,004$ ,  $df = 10$  and  $p = 0,815$ . With the use of the latter database, an alternative model was generated, which was adjusted with  $X^2 = 18,952$ ,  $df = 16$  and  $p = 0,271$ , without the pool of organic matter, but with the recovery of the pool of physical and chemical attributes. Overall, the adjusted structural models, stable and labile P pools were considered sources of P on P pool available. From the models presented it was verified the interrelationships between theoretical concepts and confirm the hypotheses of dependence between the pools.

Keywords: Bowman method, Latent variables, Tropical soils.

## INTRODUÇÃO

A deficiência de P disponível na solução do solo de regiões tropicais intemperizadas é um fator altamente prejudicial para o desenvolvimento das plantas (Oliveira et al., 2014). Esta deficiência é resultante dos baixos teores de P total encontrados no solo e/ou causada pela forte adsorção de Pi aos oxidróxidos de Al e Fe no solo (Zaia et al., 2008b). Trabalhos, utilizando o método de extração de Bowman, apontam que o P disponível por Mehlich-1, associado à matéria orgânica presente nas camadas de horizontes superficiais do solo, é controlado especialmente pelas formas de P orgânico (Po) (Guerra et al., 1996; Duda, 2000), que passam pelo processo de decomposição e mineralização da serapilheira (Zaia et al. 2008a; Rita et al., 2013). Deste Po total analisado, em geral, grande parte é constituído de Po na forma lábil (Cunha et al., 2007; Oliveira et al., 2014) que se encontra correlacionado com o P disponível e C orgânico (Zaia et al. 2008a).

No entanto, as análises estatísticas realizadas nestes trabalhos como correlações de Pearson, análises de componentes principais e regressões múltiplas são pontuais e não conseguem explicar, em conjunto, todas estas relações discutidas. Deste modo, análises mais complexas que analisam múltiplos relacionamentos devem ser feitas para estudar o ciclo do P, permitindo a discussão das relações existentes entre o P disponível e frações de P (Pi e Po) lábeis e estáveis.

Para verificar e confirmar estes relacionamentos é utilizado a metodologia de modelagem de equações estruturais (MEE). Esta metodologia científica é utilizada para construir e analisar hipóteses sobre fortes conexões causais no estudo de fenômenos complexos, com a utilização de dados empíricos e conceitos teóricos através de variáveis latentes (Grace, 2006). Variáveis latentes ou não observáveis são variáveis conceituais que formam e explicam um conjunto de variáveis observadas através de uma base teórica. Estas variáveis são incorporadas nos modelos de equações estruturais para tornar o modelo teoricamente mais consistente e testar hipóteses sobre vários processos que operam no sistema (Sales et al., 2015).

Esta ferramenta de modelagem estatística tem como base a combinação das técnicas de análises de regressão e fatorial mais análise de caminhos (Hair et al., 2009), tendo como proposta central construir e analisar modelos hipotéticos que possam refletir as associações entre os conceitos teóricos. As associações são analisadas por intermédio das relações impostas, tais como descritas pela teoria, presentes em duas ou mais equações formadas pelo modelo (Byrne, 1994).

As inter-relações construídas a partir dos conceitos em MEE são representadas por dois tipos de modelos descritos por Bollen e Noble (2011). O modelo estrutural, associação entre as variáveis latentes (construtos), definido pela equação:

$$\eta_i = \alpha_\eta + B\eta_i + \Gamma\xi_i + \zeta_i \quad (1)$$

onde os vetores  $\eta_i$  e  $\xi_i$  são as variáveis latentes endógenas e exógenas respectivamente;  $\alpha_\eta$  é um vetor com interceptos de cada equação, sendo assumido neste estudo como vetor de zero; as matrizes B e  $\Gamma$  são coeficientes que representam os efeitos entre as variáveis latentes; e  $\zeta_i$  representa o vetor de

erros aleatórios relacionados às variáveis latentes endógenas, e o modelo mensurável, onde há uma relação entre as grandezas de medição e os seus respectivos construtos, descritos pelas seguintes equações fundamentais :

$$y_i = \alpha_y + \Lambda_y \eta_i + \varepsilon_i \quad (2)$$

$$x_i = \alpha_x + \Lambda_x \xi_i + \delta_i, \quad (3)$$

onde  $y_i$  e  $x_i$  são vetores de variáveis mensuráveis ou indicadores relacionados às variáveis latentes endógenas e exógenas, respectivamente para cada unidade  $i$ ;  $\alpha_y$  e  $\alpha_x$  são vetores com interceptos, ambos vetores de zeros; as matrizes  $\Lambda_x$  e  $\Lambda_y$  são os coeficientes relacionados aos indicadores das variáveis latentes; e os vetores  $\varepsilon_i$  e  $\delta_i$  representam os erros de mensuração para os respectivos indicadores  $y_i$  e  $x_i$  do modelo.

Em MEE, assume-se que as variáveis são exógenas ou endógenas, onde as variáveis exógenas não possuem relação de dependência e podem ser correlacionadas com outras variáveis exógenas, e as variáveis endógenas são dependentes de pelo menos uma das variáveis do modelo (Bollen e Noble, 2011).

Dentro das ciências naturais a MEE vem se desenvolvendo e atraindo o interesse de muitos pesquisadores, especialmente na ecologia (Grace e Pugesek, 1998; Pugesek et al., 2003; Barto et al., 2010; Grace et al., 2012). Na área de ciência do solo, no estudo das transformações de P no solo, modelos de MEE foram apresentados com a utilização de dados obtidos com o método de extração sequencial de P desenvolvido por Hedley.

Recentemente, Gama-Rodrigues et al. (2014) aplicaram a modelagem de equações estruturais com variáveis latentes para testar modelos hipotéticos do ciclo do P em solos tropicais. No entanto, estes autores não encontraram no modelo estrutural básico um ajuste adequado com a utilização de cinco variáveis latentes (*pools* de P) e múltiplos indicadores (frações de P). O ajuste satisfatório somente foi obtido no modelo com apenas uma variável latente endógena, *pool* de P disponível, sendo influenciada direta e indiretamente por frações de P.

Na sequência, Sales et al. (2015) trabalharam na reespecificação dos modelos apresentados por Gama-Rodrigues et al. (2014), encontrando ajuste adequado em um modelo com quatro variáveis latentes, representantes dos *pools*

de P, que corroboraram com a hipótese da dependência do *pool* de P disponível pelo *pool* de P orgânico em solos tropicais não fertilizados.

Para este trabalho foram elaborados modelos hipotéticos de equações estruturais para analisar a disponibilidade de P, por Mehlich-1, através de ciclos alternativos de P no solo, usando dados obtidos pelo método de extração de Bowman, acrescentando dados relacionados a propriedades do solo. O modelo hipotético básico do ciclo do P do solo (Figura 1) foi constituído pelas variáveis latentes *pools* de P estável (Pi e Po), P lábil (Pi e Po) e P disponível associadas aos *pools* matéria orgânica (MO) e atributos físico-químicos do solo (AFQ). O conceito teórico a ser testado é o grau em que o *pool* de P disponível é controlado por diferentes *pools* de P e diferentes *pools* de atributos do solo.

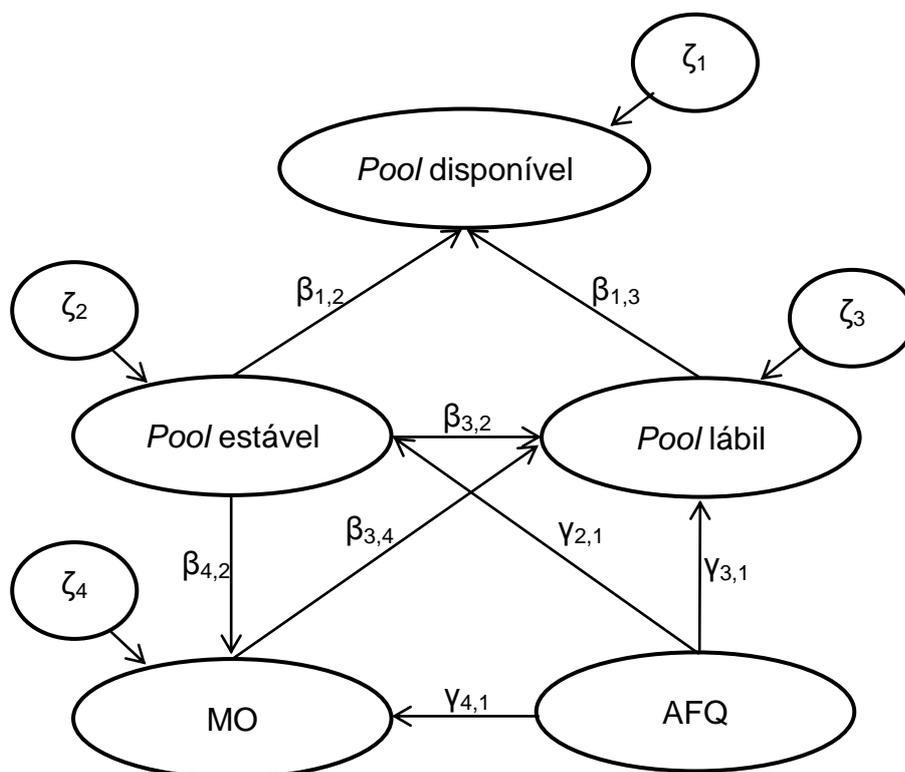


Figura 1. Modelo estrutural hipotético para o ciclo do P do solo com a inclusão de propriedades do solo.

## MATERIAL E MÉTODOS

Para construção dos modelos de equações estruturais deste trabalho foram seguidas algumas etapas essenciais recomendadas por Hair et al. (2009) e

Bollen e Noble (2011). Primeiro foi desenvolvido o modelo teórico com base no estudo do ciclo do P analisado pelo método de extração de Bowman, em seguida, foi construído o diagrama de caminhos do modelo hipotético básico estabelecendo as relações causais entre os *pools* de P do solo e os *pools* de atributos físicos e químicos do solo (Figura 1).

Com o diagrama de caminho construído foi possível convertê-lo em um conjunto de equações estruturais e mensuráveis. Ao finalizar estas etapas foi definida matriz de entrada de dados e os parâmetros foram estimados. Foi verificada a identificação do modelo estrutural e foram feitas avaliações dos modelos segundo critérios de ajuste. Com a falta de ajuste adequado em alguns modelos foi necessário fazer algumas reespecificações nos mesmos, configurando assim, modelos ajustados. Com o ajuste dos modelos foi possível fazer a discussão da modelagem elaborada, relacionando os resultados encontrados com a teoria em estudo.

Para o trabalho foram utilizados dois bancos de dados de solos não fertilizados de regiões tropicais, referentes aos estudos da disponibilidade de P no solo pelo método de extração sequencial desenvolvida por Bowman (1989). O primeiro banco de dados foi coletado da literatura (Tabela 1A) constituído por 68 observações de solos que variam de Cambissolos Háplicos a Latossolos Vermelho-Amarelo. O segundo foi coletado de Oliveira et al. (2014), constituído por 38 observações referentes a solos de diferentes horizontes superficiais como H e O hístico, A chernozêmico e A húmico (Tabela 1B). Neste último, ainda foi trabalhado uma redução de dados referentes às amostras do solo orgânico, permanecendo 28 observações de Cambissolos e Latossolos. As frações de P analisadas foram P-Mehlich-1 (P extraível), P bicarbonato (Pi e Po), P estável (Pi e Po).

A fração P-Mehlich-1, obtida pela extração de ácido-base, representa a forma de P que está prontamente disponível para a absorção da planta; as frações bicarbonato Pi e bicarbonato Po representam P que pode ser liberado pela troca de ligantes com o íon bicarbonato (Sato e Comerford, 2006); e as frações estável Pi e estável Po representam o P que se encontra na forma não-lábil, ou seja, não prontamente disponível na solução do solo. Além das frações de P do solo, o trabalho incluiu variáveis que representam propriedades do solo como carbono orgânico (C), nitrogênio total (N), óxido de alumínio ( $Al_2O_3$ ), óxido

de ferro ( $\text{Fe}_2\text{O}_3$ ), argila, pH e  $P_{\text{rem}}$  (P remanescente) para avaliar a influência destas propriedades sobre a disponibilidade de P no solo. Os elementos C orgânico e N total são indicadores do *pool* matéria orgânica (MO) do solo e as outras variáveis são indicadores do *pool* atributos físico-químicos (AFQ) do solo.

No modelo estrutural hipotético básico (Figura 1) foram verificadas as influências diretas e indiretas das variáveis latentes *pool* de P estável ( $\eta_2$ ), *pool* de P lábil ( $\eta_3$ ), *pool* MO ( $\eta_4$ ) e *pool* AFQ ( $\xi_1$ ) sobre a variável latente endógena *pool* de P disponível ( $\eta_1$ ), através do seguinte sistema de equações estruturais:

$$\eta_1 = \beta_{1,2}\eta_2 + \beta_{1,3}\eta_3 + \zeta_1 \quad (4)$$

$$\eta_2 = \gamma_{2,1}\xi_1 + \zeta_2 \quad (5)$$

$$\eta_3 = \beta_{3,2}\eta_2 + \beta_{3,4}\eta_4 + \gamma_{3,1}\xi_1 + \zeta_3 \quad (6)$$

$$\eta_4 = \beta_{4,2}\eta_2 + \gamma_{4,1}\xi_1 + \zeta_4 \quad (7)$$

Este sistema de equações reforça a ideia de inter-relacionamento entre as variáveis conceituais do modelo. Desse modo, a equação (4), equação principal, é dependente das equações (5) e (6), ao passo que a equação (6) é dependente da equação (7) e todas estas equações são dependentes da variável latente exógena  $\xi_1$ , representante dos atributos físico-químicos do solo.

As hipóteses específicas testadas do modelo estrutural hipotético básico (Figura 1) são as seguintes:

H<sub>1</sub>: O *pool* de P estável age diretamente como dreno de P sobre o *pool* de P disponível, pois representa as frações de P não lábeis no solo;

H<sub>2</sub>: O *pool* de P estável age diretamente como fonte para o *pool* de P lábil;

H<sub>3</sub>: O *pool* de P estável age positivamente sobre o *pool* MO do solo;

H<sub>4</sub>: O *pool* MO do solo age positivamente sobre o *pool* de P lábil;

H<sub>5</sub>: O *pool* AFQ do solo atua negativamente sobre o *pool* MO;

H<sub>6</sub>: O *pool* AFQ do solo atua positivamente sobre o *pool* de P estável;

H<sub>7</sub>: O *pool* AFQ do solo atua negativamente sobre o *pool* de P lábil;

H<sub>8</sub>: O *pool* lábil age diretamente como fonte para o *pool* de P disponível.

Tabela 1. Estudo da extração sequencial de fósforo em solos tropicais sob diferentes usos da terra.

<b>Banco de Dados</b>	<b>Referência</b>	<b>Cobertura</b>	<b>Tipo de solo</b>	<b>Localização</b>
<b>A</b>	Agbenin et al. (1999)	Savana.	Argissolos, Cambissolos, Luvisolos, Neossolos, Vertissolos.	Nigéria
	Cunha et al. (2007)	Floresta nativa, pastagem, <i>Eucalyptus</i> .	Cambissolo , Latossolo.	Brasil
	Duda (2000)	Floresta nativa, <i>Pinus</i> , <i>Eucalyptus</i> , pastagem.	Cambissolos, Chernossolos, Latossolos, Luvisolos, Neossolos, Nitossolos, Organossolos.	Brasil
	Guerra et al. (1996)	Pastagem.	Latossolos, Argissolos.	Brasil
	Zaia et al. (2008a)	Leguminosas florestal, pastagem, floresta nativa.	Latossolos.	Brasil
	Zaia et al. (2008b)	Agrossistema de Cacau.	Cambissolos, Latossolos.	Brasil
<b>B</b>	Oliveira et al. (2014)	Pastagem.	Chernossolos, Latossolos, Organossolos.	Brasil

**Banco de dados A** – Solos com baixo teor de matéria orgânica, N = 68;

**Banco de dados B** – Solos com alto teor de matéria orgânica, N = 38.

Para o trabalho foram aplicados testes de normalidade univariada e multivariada em todas as variáveis. No primeiro banco de dados, referente aos solos minerais com baixo teor de matéria orgânica (Tabela 1A), foram feitas normalizações dos dados pela função logaritmo de base 10 por não apresentarem normalidade multivariada exigida pela função de estimação de parâmetros do modelo. Foram feitas também tabelas com estatísticas descritivas de cada banco de dados (Tabela 2) e tabelas com correlações de Pearson entre todas as variáveis (Tabelas 3, 4 e 5) que auxiliam nos entendimentos das relações estabelecidas nos modelos hipotéticos e na verificação de colinearidade entre variáveis independentes que são altamente correlacionadas com valores  $\geq 0,90$  (Hair et al., 2009).

Na construção dos modelos, primeiramente foram feitas algumas especificações para que os mesmos pudessem ser estimados e analisados. Uma das especificações feitas foi a fixação de parâmetros, pela necessidade de transformar um modelo não-identificado em um modelo identificado, sendo possível assim estimar todos os outros parâmetros livres do modelo (Grace, 2006; Byrne, 2009).

Utilizando critérios de fixação de parâmetros apresentados em Sales et al. (2015), foram fixados parâmetros como variâncias de erro e cargas fatoriais no modelo estrutural hipotético básico. A variância de erro da variável indicadora P-Mehlich-1 foi definida com o valor de zero, assumindo que esta variável foi estimada sem erros de medição, e neste caso, sua carga fatorial também foi fixada, porém com o valor 1. As variâncias de erro das variáveis latentes *pool* de P estável, *pool* de P lábil e *pool* MO também foram fixadas, neste caso, com valores de 0,1 (10%). Estas especificações foram feitas em todos os modelos apresentados. Adicionalmente no Modelo A (Figura 2), as variâncias de erro dos indicadores  $Fe_2O_3$  e  $Al_2O_3$  também foram fixadas com o valor de 0,1 por causarem problema de identificação no modelo pela alta correlação existente entre estas variáveis independentes (Tabela 3).

Tabela 2. Estatística descritiva dos bancos de dados utilizados na modelagem do clico de P no solo.

Variáveis	Dados da literatura (Tabela 1A)				Dados da literatura (Tabela 1B)				Dados da literatura (Tabela 1B)			
	N = 68				N= 38				N= 28			
	$\bar{x}$	$\delta$	Min.	Max.	$\bar{x}$	$\Delta$	Min.	Max.	$\bar{x}$	$\delta$	Min.	Max.
<b>Mehlich-1</b> (mg kg <sup>-1</sup> )	13,8	25,7	0,1	128,0	11,4	11,8	0,4	47,7	7,8	8,2	0,4	31,8
<b>HCO<sub>3</sub> Po</b> (mg kg <sup>-1</sup> )	10,8	18,3	4,9	104,9	26,8	23,1	1,8	99,5	19,3	20,7	1,8	99,5
<b>HCO<sub>3</sub> Pi</b> (mg kg <sup>-1</sup> )	19,9	12,9	1,0	88,1	9,0	9,6	1,0	48,0	5,6	3,1	1,0	12,0
<b>Estável Po</b> (mg kg <sup>-1</sup> )	127,8	143,8	1,0	714,4	249,8	265,9	15,8	1057,3	199,1	232,5	15,8	959,6
<b>Estável Pi</b> (mg kg <sup>-1</sup> )	317,9	299,6	14,1	1261,0	369,2	332,5	36,0	1410,0	346,0	319,0	56,0	1410,0
<b>Prem</b> (mg kg <sup>-1</sup> )	-	-	-	-	25,7	18,3	1,6	58,5	26,5	16,4	1,6	56,3
<b>pH</b>	5,1	0,9	3,3	7,8	5,4	1,3	3,0	8,2	5,8	1,7	3,6	8,2
<b>C</b> (g kg <sup>-1</sup> )	26,7	16,5	3,0	86,4	98,9	140,8	10,8	541,1	31,7	17,3	10,8	66,2
<b>N</b> (g kg <sup>-1</sup> )	1,9	0,8	0,3	4,2	4,4	3,5	1,3	15,2	2,7	1,0	3,6	5,2
<b>Argila</b> (g kg <sup>-1</sup> )	300,8	191,3	10,0	830,0	322,7	130,8	33,0	597,0	322,7	153,2	33,0	597,0
<b>Fe<sub>2</sub>O<sub>3</sub></b> (g kg <sup>-1</sup> )	34,9	39,5	0,1	225,0	48,9	51,8	1,0	250,0	58,7	56,3	9,0	250,0
<b>Al<sub>2</sub>O<sub>3</sub></b> (g kg <sup>-1</sup> )	65,6	67,0	0,01	255,4	97,4	53,2	6,0	194,0	103,4	51,4	15,0	194,0

$\bar{x}$  (média);  $\delta$  (desvio padrão), **Min.** e **Max.** (valores mínimos e máximos, respectivamente).

- valores não determinados neste banco de dados.

Tabela 1A – Dados obtidos de solos minerais com baixo teor de matéria orgânica.

Tabela 1B – Dados obtidos de solos minerais com alto teor de matéria orgânica.

Tabela 3. Coeficientes de correlação de Pearson com os dados dos solos minerais de baixo teor de matéria orgânica, normalizados pela função Log10, N = 68.

	HCO <sub>3</sub> P <sub>i</sub>	HCO <sub>3</sub> P <sub>o</sub>	Estável P <sub>o</sub>	Estável P <sub>i</sub>	pH	C	N	Argila	Fe <sub>2</sub> O <sub>3</sub>	Al <sub>2</sub> O <sub>3</sub>
<b>Mehlich-1</b>	0,729***	0,428***	0,463***	0,443***	0,322**	0,246*	0,336**	-0,274*	-0,274*	-0,201
<b>HCO<sub>3</sub> Pi</b>		0,651***	0,487***	0,627***	0,031	0,314**	0,354**	-0,083	-0,219	-0,139
<b>HCO<sub>3</sub> Po</b>			0,374**	0,408***	-0,203	0,381**	0,371**	-0,113	0,039	0,143
<b>Estável Po</b>				0,687***	0,281*	0,379**	0,470***	0,139	-0,200	-0,189
<b>Estável Pi</b>					0,091	0,373**	0,440***	0,272*	-0,001	0,022
<b>pH</b>						-0,179	-0,057	-0,462***	-0,025	-0,079
<b>C</b>							0,809***	0,351**	-0,453***	-0,367**
<b>N</b>								0,357**	-0,354**	-0,275
<b>Argila</b>									-0,113	-0,114
<b>Fe<sub>2</sub>O<sub>3</sub></b>										0,967***

\* significativos com  $p < 0,05$ ;

\*\* significativos com  $p < 0,01$ ;

\*\*\* significativos com  $p < 0,001$ .

Tabela 4. Coeficientes de correlação de Pearson com os dados originais do trabalho de Oliveira et al, (2014), pelo método de extração de Bowman (Organossolos + Chernossolos + Latossolos, N = 38).

	HCO <sub>3</sub> P <sub>o</sub>	HCO <sub>3</sub> P <sub>i</sub>	Estável P <sub>o</sub>	Estável P <sub>i</sub>	Prem	pH	C	N	Argila	Fe <sub>2</sub> O <sub>3</sub>	Al <sub>2</sub> O <sub>3</sub>
<b>Mehlich-1</b>	0,556***	0,729***	0,282	0,459**	0,070	-0,321*	0,521***	0,454**	-0,103	-0,208	-0,007
<b>HCO<sub>3</sub> P<sub>o</sub></b>		0,499**	0,158	0,205	-0,063	-0,537***	0,527***	0,578***	0,201	-0,249	-0,013
<b>HCO<sub>3</sub> P<sub>i</sub></b>			0,044	0,114	0,288	-0,538***	0,773***	0,696***	0,039	-0,302	-0,320
<b>Estável P<sub>o</sub></b>				0,584***	-0,456**	-0,146	0,119	0,197	0,042	0,135	0,319
<b>Estável P<sub>i</sub></b>					-0,354*	-0,124	0,008	0,070	0,119	0,343*	0,483**
<b>Prem</b>						0,363*	0,254	0,225	-0,382*	-0,272	-0,672***
<b>pH</b>							-0,560***	-0,525***	-0,257	0,244	0,011
<b>C</b>								0,962***	0,006	-0,351*	-0,379*
<b>N</b>									0,044	-0,271	-0,353*
<b>Argila</b>										0,280	0,452**
<b>Fe<sub>2</sub>O<sub>3</sub></b>											0,549***

\* significativos com  $p < 0,05$ ;

\*\* significativos com  $p < 0,01$ ;

\*\*\* significativos com  $p < 0,001$ .

Tabela 5. Coeficientes de correlação de Pearson com os dados originais do trabalho de Oliveira et al, (2014), pelo método de extração de Bowman (Chernossolos + Latossolos, N = 28).

	HCO <sub>3</sub> P <sub>o</sub>	HCO <sub>3</sub> P <sub>i</sub>	Estável P <sub>o</sub>	Estável P <sub>i</sub>	Prem	pH	C	N	Argila	Fe <sub>2</sub> O <sub>3</sub>	Al <sub>2</sub> O <sub>3</sub>
<b>Mehlich-1</b>	0,359	0,580**	0,433*	0,607***	0,007	0,053	0,304	0,407*	-0,174	-0,007	0,100
<b>HCO<sub>3</sub> P<sub>o</sub></b>		0,588**	0,117	0,270	-0,136	-0,370	0,568**	0,638***	0,262	-0,076	0,075
<b>HCO<sub>3</sub> P<sub>i</sub></b>			0,355	0,498**	-0,094	-0,299	0,512**	0,720***	0,142	-0,120	0,052
<b>Estável P<sub>o</sub></b>				0,483**	-0,361	-0,140	0,189	0,258	0,057	0,278	0,320
<b>Estável P<sub>i</sub></b>					-0,304	-0,141	0,196	0,424*	0,145	0,437*	0,490**
<b>Prem</b>						0,7369***	-0,403*	-0,110	-0,498**	-0,305	-0,667***
<b>pH</b>							-0,648***	-0,244	-0,314	0,087	-0,248
<b>C</b>								0,738***	0,053	-0,138	0,032
<b>N</b>									0,181	0,197	0,094
<b>Argila</b>										0,302	0,548**
<b>Fe<sub>2</sub>O<sub>3</sub></b>											0,579**

\* significativos com  $p < 0,05$ ;

\*\* significativos com  $p < 0,01$ ;

\*\*\* significativos com  $p < 0,001$ .

Com a identificação do modelo estrutural hipotético foram estimados os parâmetros na forma padronizada e feitas análises da validade convergente dos construtos. A convergência de um construto indica o quão bem uma variável latente foi elaborada. Para esta análise é indispensável a utilização de cargas fatoriais estimadas na forma padronizada, pois cargas fatoriais altas, idealmente  $\geq 0,7$ , sobre um fator (construto) indicam que elas convergem para um ponto em comum (Hair et al., 2009). A convergência dos construtos foi calculada através da variância extraída (*VE*) e da confiabilidade de construto (*CR*).

A *VE* indica o quanto da variância dos indicadores pode ser explicada pelos respectivos construtos, sendo calculada do seguinte modo:

$$VE = \frac{\sum_{i=1}^n \lambda_i^2}{\sum_{i=1}^n \lambda_i^2 + \sum_{i=1}^n \varepsilon_i}$$

onde os valores de  $\lambda_i$  representam as cargas fatoriais padronizadas e os  $\varepsilon_i$  termos de variância de erro para cada construto. A variância extraída é considerada adequada quando são encontrados valores superiores a 0,5, sugerindo convergência adequada dos seus indicadores (Fornell e Larcker, 1981; Hair et al., 2009).

A *CR* também é um indicador de validade convergente e pode ser calculada pela seguinte expressão:

$$CR = \frac{\left( \sum_{i=1}^n \lambda_i \right)^2}{\left( \sum_{i=1}^n \lambda_i \right)^2 + \left( \sum_{i=1}^n \varepsilon_i \right)}$$

onde valores de confiabilidade maiores do que 0,7 são considerados ideais e entre 0,6 e 0,7 podem ser aceitáveis, porém os valores de confiabilidade de outros construtos devem ser bons para que o modelo seja considerado adequado para a análise (Fornell e Larcker, 1981; Hair et al., 2009).

Para a estimação dos parâmetros dos modelos foi utilizado o método de máxima verossimilhança (ML) e para o ajuste geral dos modelos foram usados a estatística qui-quadrado ( $X^2$ ), o grau de liberdade (*df*) e o nível de probabilidade (*p*) associados ao modelo. Com o teste  $X^2$  e valor-*p*  $> 0,05$  há evidências estatísticas de que o modelo hipotético está ajustado, e a hipótese de nulidade

não deve ser rejeitada. Em razão da sensibilidade do teste do qui-quadrado para o tamanho das amostras, os índices de qualidade de ajuste (GFI), de ajuste comparativo (CFI), de raiz do erro quadrático médio de aproximação (RMSEA) e  $\chi^2/df$  também foram considerados como alternativas de mensuração de ajuste do modelo. Os valores de GFI e CFI  $\geq 0,95$  (Hair et al. 2009) e RMSEA  $\leq 0,08$  (Byrne, 2009) indicam ajuste apropriado para o modelo. O indicador  $\chi^2/df$ , segundo Iacobucci (2010), deve ser menor do que 3, para que o modelo seja considerado ajustado.

Para os modelos reespecificados foram calculados deltas do qui-quadrado ( $\Delta\chi^2$ ), onde para valores  $> 3,84$  para um  $\Delta df = 1$ , conclui-se que, em nível de probabilidade de 0,05, o modelo alternativo foi mais ajustado. Também foi utilizado neste trabalho o critério de informação de Akaike (AIC), índice de parcimônia, utilizado na comparação de modelos distintos, onde quanto menor for o valor calculado para este índice mais bem ajustado é o modelo (Byrne, 2009; Kenny, 2014). Os modelos de MEE deste trabalho foram analisados através do software estatístico AMOS versão 22 (IBM – SPSS Inc., Chicago, IL, USA).

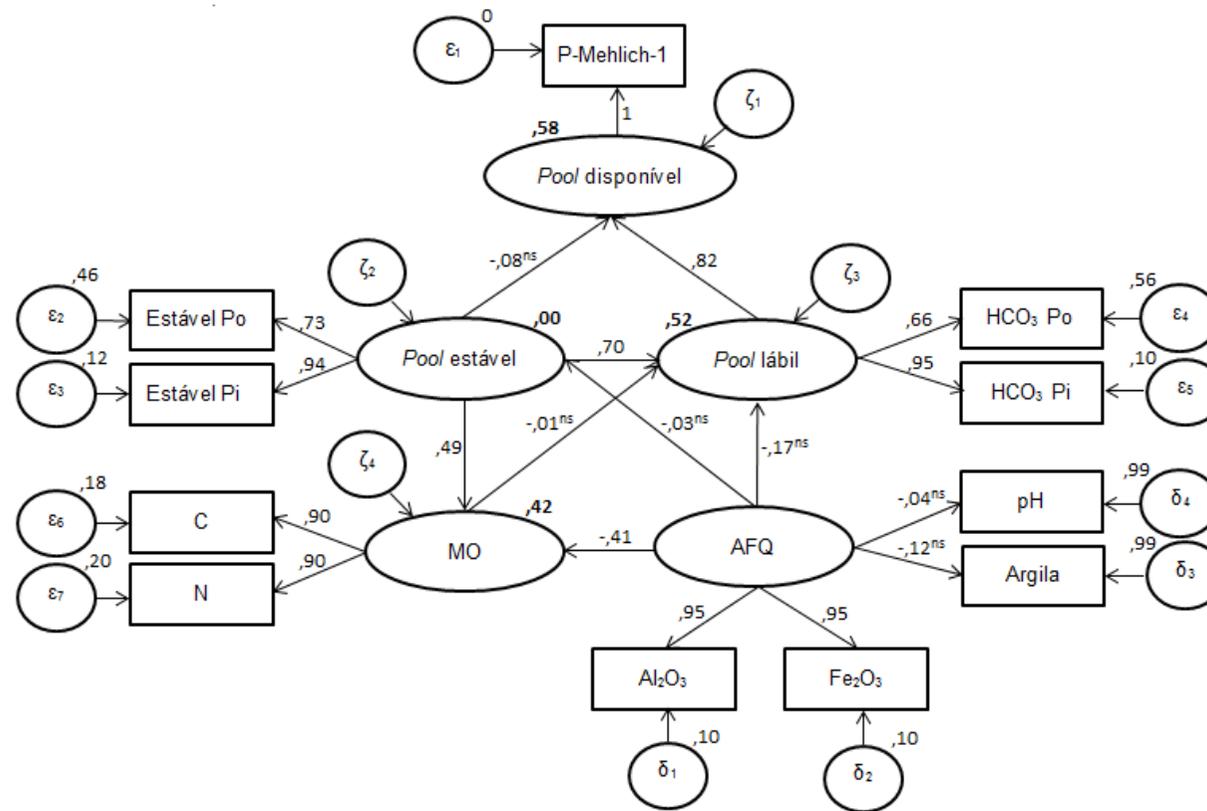


Figura 2. Modelo A, modelo estrutural hipotético testado para avaliar o ciclo do P em solos tropicais com baixo teor de matéria orgânica, N = 68, com inclusão de propriedades do solo. Os indicadores de efeito (retângulos) estão associados com as variáveis latentes (elipses). Os valores numéricos correspondem a cargas fatoriais padronizadas com  $P < 0,001$ . Valores em negrito correspondem ao coeficiente de determinação ( $R^2$ ). “<sup>ns</sup> – não significativo”. Variâncias de erro ( $\varepsilon_1 - \varepsilon_7$  e  $\delta_1 - \delta_4$ ) estão padronizadas. Os valores de  $\chi^2 = 190,375$ ,  $df = 40$ ,  $p < 0,0001$ .

## RESULTADOS

Com a identificação do modelo estrutural hipotético (Modelo A – Figura 2) foram obtidos os seguintes valores de índices de ajuste:  $\chi^2 = 190,375$ ,  $df = 40$ ,  $p < 0,0001$ . Estes índices de ajuste indicam que o modelo estrutural proposto ainda não se encontra ajustado aos dados. Outros índices para ajuste global estão apresentados na Tabela 6, confirmando a falta de ajuste deste modelo. Através dos parâmetros estimados observou-se que este modelo, mesmo não estando ajustado, apresentou algumas relações estruturais que correspondem ao esperado pela teoria apresentada nas hipóteses. Por exemplo, as relações estabelecidas entre o *pool* atributos físico-químicos do solo (AFQ), com o *pool* de P lábil e com *pool* matéria orgânica (MO) que são negativas.

Com os parâmetros estimados na forma padronizada foi possível efetuar os cálculos aproximados da variância extraída e da confiabilidade de cada construto. Para a variável latente endógena *pool* de P estável foi calculado um valor de variância extraída,  $VE = 0,71$ , e confiabilidade do construto,  $CR = 0,83$ . Para a variável latente endógena *pool* de P lábil foi calculado um valor de  $VE = 0,67$  e  $CR = 0,80$ . Em relação às propriedades do solo, foi calculada uma  $VE = 0,81$  e  $CR = 0,90$  para o *pool* MO, e para o *pool* AFQ foi obtido o valor de  $VE = 0,45$  e  $CR = 0,58$ . Estes valores apontam que a única variável latente que não se adequou matematicamente às especificações apresentadas na literatura foi a variável latente exógena *pool* AFQ. Em relação à explicação geral do modelo ( $R^2$ ), a variável latente endógena *pool* de P disponível foi explicada por 58% das variações, as variáveis latentes endógenas *pool* de P estável, *pool* de P lábil e *pool* MO possuíram, respectivamente,  $R^2$  igual a 0%, 52% e 42%.

Visto que o modelo estrutural (Figura 2) não apresentou ajuste adequado, foi retirada do modelo a variável pH, que dentre as relações estruturais e mensuráveis ( $H_1$ ,  $H_4$ ,  $H_6$ ,  $H_7$ , argila e pH), foi a que apresentou valor de  $p$  menos significativo, além de fazer parte do construto que possui  $VE$  e  $CR$  não satisfatório.

Ao se retirar esta variável do modelo foram calculados novos parâmetros. O novo modelo (Modelo A1) apresentou uma redução no qui-quadrado, porém ainda se manteve com nível de probabilidade altamente significativo. Os índices

de ajuste desde modelo foram  $X^2 = 133,471$ ,  $df = 31$ ,  $p < 0,0001$ , mostrando falta de ajuste adequado, porém com  $\Delta X^2 = (X^2_{(A)} - X^2_{(A1)}) = 56,904$  significativo em nível de probabilidade de 0,05 e  $\Delta df = 9$ . Com os parâmetros estimados do Modelo A1, foi verificado que as relações estruturais hipotéticas citadas anteriormente ainda não se encontravam significativas e apenas a carga fatorial da variável argila não estava altamente significativa, tornando-a a opção de retirada do modelo. As figuras referentes aos Modelos A1 e A2 não são apresentadas no trabalho. O Modelo A2, reespecificado com a retirada da variável argila também não foi ajustado. Os valores encontrados foram  $X^2 = 95,159$ ,  $df = 23$ ,  $p < 0,0001$ , com  $\Delta X^2 = (X^2_{(A1)} - X^2_{(A2)}) = 38,312$  com  $\Delta df = 8$ , indicando que o Modelo A2 melhorou estatisticamente em relação aos modelos anteriores. Com as reespecificações feitas, tanto a variância extraída quanto a confiabilidade das variáveis latentes *pool* de P estável, *pool* de P lábil e *pool* MO se mantiveram adequadas, com cargas fatoriais altamente significativas.

Ao analisar as variáveis mensuráveis presentes na formação da variável latente exógena *pool* AFQ, percebe-se que os óxidos de alumínio e ferro foram as únicas variáveis mensuráveis restantes neste construto. Estas variáveis são altamente correlacionadas (Tabela 3) e tiveram a variância de erro fixada com o valor de 0,1 para a identificação do modelo, tornando-as possíveis variáveis para serem excluídas. Com a retirada da variável  $Fe_2O_3$  do Modelo A2, escolha aleatória, obtiveram-se os seguintes resultados para o Modelo A3:  $X^2 = 41,548$ ,  $df = 15$ ,  $p < 0,0001$ ,  $\Delta X^2 = (X^2_{(A2)} - X^2_{(A3)}) = 53,611$  com  $\Delta df = 8$ , indicando falta de ajuste para o modelo, porém com valor do qui-quadrado reduzido a cada reespecificação.

Com a falta de ajuste do modelo foi retirada a variável latente exógena *pool* AFQ do Modelo A3, obtendo o modelo final ajustado, Modelo B (Figura 3), com  $X^2 = 14,374$ ,  $df = 11$ ,  $p = 0,213$ . Todos os índices de ajustes foram adequados, GFI = 0,945, CFI = 0,986, RMSEA = 0,068 (Tabela 6), e a variável latente endógena *pool* disponível foi explicada por 57% das variações.

Tabela 6. Índices de ajuste dos modelos para análise do ciclo do P pelo método de extração sequencial de Bowman.

Modelo	Índices de Ajuste								
	N	$\chi^2$	Df	P	GFI	CFI	RMSEA	AIC	$\chi^2/df$
<b>A</b>	68	190,375	40	***	0,707	0,729	0,237	242,375	4,759
<b>B</b>	68	14,374	11	0,213	0,945	0,986	0,068	48,374	1,307
<b>B1</b>	28	6,004	10	0,815	0,940	1,000	0,000	42,004	0,600
<b>C</b>	28	24,788	23	0,361	0,848	0,977	0,054	68,788	1,078
<b>C1</b>	28	15,308	16	0,502	0,897	1,000	0,000	55,308	0,957
<b>C2</b>	38	18,952	16	0,271	0,889	0,974	0,071	58,952	1,184

\*\*\*  $p < 0,0001$ .

Na validação do Modelo *B*, foram utilizados os dados dos solos minerais com alto teor de matéria orgânica (Organossolos + Chernossolos + Latossolos, Tabela 1B, N = 38). No entanto, ao tentar estimar os parâmetros foi verificada a não identificação do modelo com estes dados. A alta correlação entre os indicadores C e N (Tabela 4) gerou erros na identificação do modelo. Com os resultados destas análises foi utilizado o banco de dados dos solos minerais com alto teor de matéria orgânica composto apenas dos Chernossolos + Latossolos (Tabela 1B, N = 28), sem as amostras do solo orgânico. Com esta redução do banco de dados ainda foi verificado que a correlação entre as variáveis (C e N) se manteve altamente significativa, porém mais fraca  $\leq 0,90$  (Tabela 5), reduzindo, assim, a colinearidade.

Com o banco de dados (Chernossolos + Latossolos) atribuído ao Modelo *B*, foram geradas novas estimativas e índices de ajuste configurando o Modelo *B1* (Figura 4). O Modelo *B1*, identificado, apresentou índices de ajuste geral  $\chi^2 = 6,004$ ,  $df = 10$ ,  $p = 0,815$ , com índices de ajuste global GFI = 0,94, CFI = 1,00 e RMSEA = 0,00. Com estes valores foi obtido um modelo ajustado, com variável latente endógena *pool* de P disponível explicada por 60% das variações. Mas, no banco de dados coletados por Oliveira et al., (2014) existia uma variável adicional, o P remanescente (Prem), que pôde ser incluída no modelo estrutural hipotético (Figura 1) através da variável latente exógena *pool* AFQ.

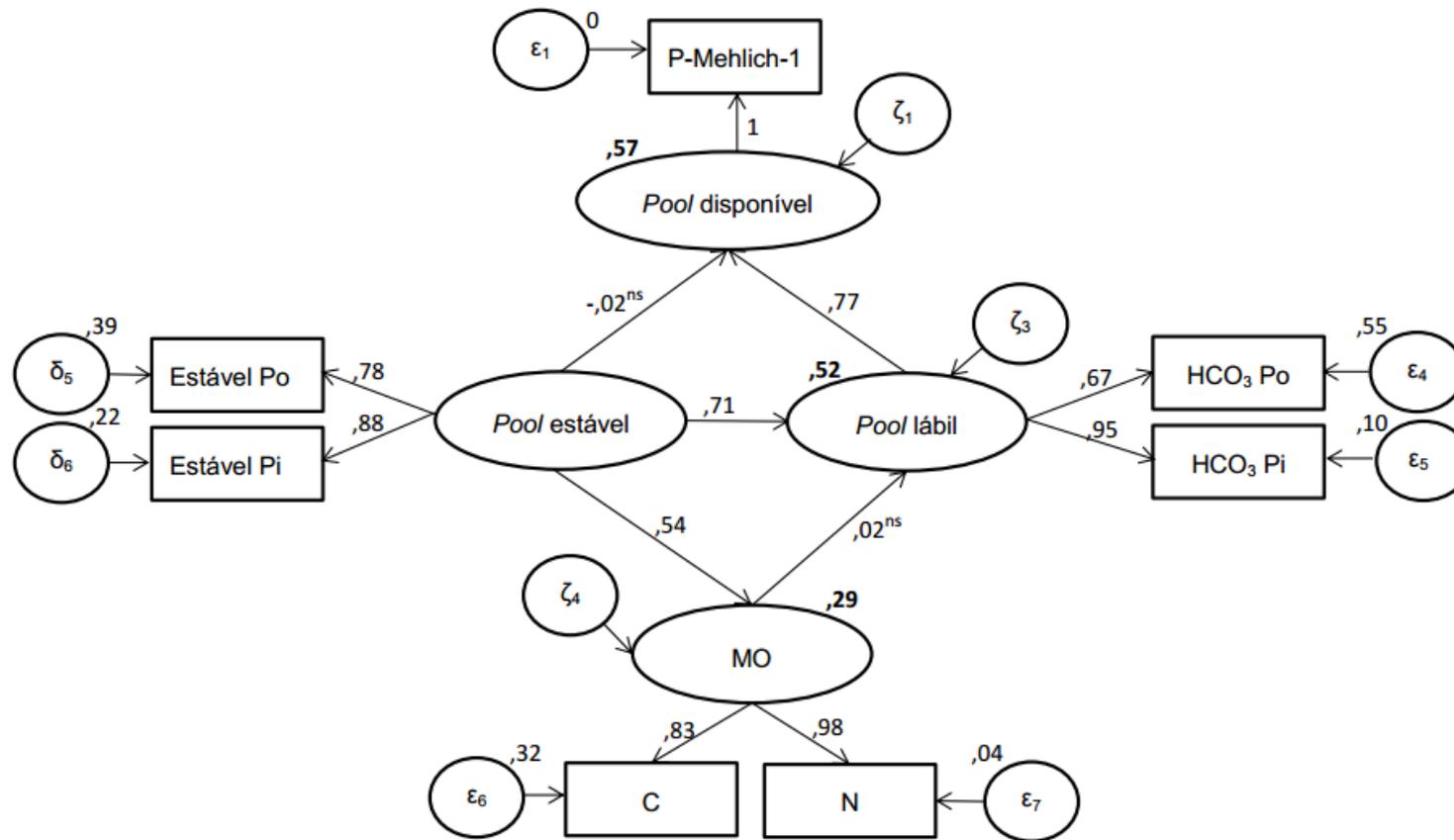


Figura 3. Modelo B, modelo estrutural testado para avaliar o ciclo de P em solos tropicais com baixo teor de matéria orgânica, N = 68, com inclusão de propriedades do solo. Os indicadores de efeito (retângulos) estão associados com as variáveis latentes (elipses). Os valores numéricos correspondem a cargas fatoriais padronizadas com  $P < 0,001$ . Valores em negrito correspondem ao coeficiente de determinação ( $R^2$ ). “<sup>ns</sup> – não significativo”. Variâncias de erro ( $\epsilon_1$ ,  $\epsilon_4$ -  $\epsilon_7$ ,  $\delta_5$  e  $\delta_6$ ) estão padronizadas. Os valores de  $\chi^2 = 14,374$ ,  $df = 11$ ,  $p = 0,213$ .

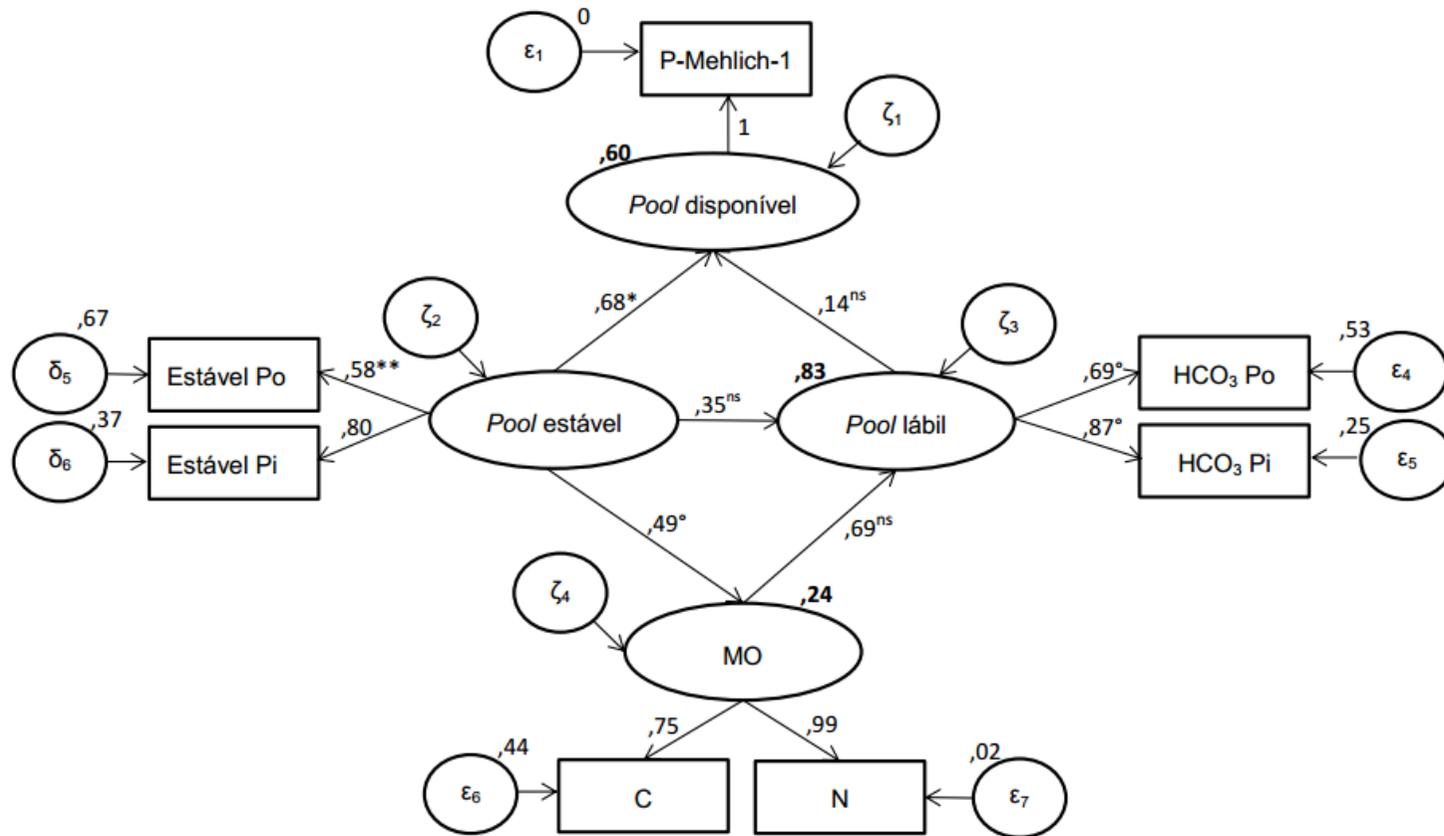


Figura 4. Modelo B1, modelo estrutural testado para avaliar o ciclo do P em solos tropicais com alto teor de matéria orgânica (Chernossolos + Latossolos, N = 28), com a inclusão de propriedades do solo. Os indicadores de efeito (retângulos) estão associados com as variáveis latentes (elipses). Os valores numéricos correspondem a cargas fatoriais padronizadas com  $P < 0,001$ . Valores em negrito correspondem ao coeficiente de determinação ( $R^2$ ). “<sup>ns</sup> – não significativo, \* significativo a  $P < 0,05$ , \*\* Significativo a  $P < 0,01$ ”. Variâncias de erro ( $\epsilon_1$ ,  $\epsilon_4$ -  $\epsilon_7$ ,  $\delta_5$  e  $\delta_6$ ) estão padronizadas. Os valores de  $\chi^2 = 6,004$ ,  $df = 10$ ,  $p = 0,815$ .

Para recuperar a variável latente exógena *pool* AFQ, existente no modelo estrutural hipotético (Figura 1), foram feitas novas análises utilizando o banco de dados da Tabela 1B com  $N = 28$ . Com a identificação do modelo foi verificado novamente que a variável mensurada pH possuiu uma relação não significativa com a variável latente *pool* AFQ, sendo excluída do modelo. O modelo reespecificado se manteve com todas as variáveis latentes iniciais e cargas fatoriais significativas, entretanto não foi ajustado. Como o Modelo B1 foi ajustado com a variável latente *pool* MO, utilizando este mesmo banco de dados, um modelo alternativo foi verificado, o Modelo C (Figura 5).

No Modelo C foram verificadas as influências diretas da variável latente *pool* AFQ, com a inclusão da variável Prem, sobre as variáveis *pool* de P estável e *pool* de P lábil. Este modelo alternativo, além de ser ajustado com  $\chi^2 = 24,788$ ,  $df = 23$ ,  $p = 0,361$ , apresentou valores de  $VE = 0,59$  e  $CR = 0,73$  para o *pool* de P estável,  $VE = 0,65$  e  $CR = 0,78$  para o *pool* de P lábil, e para o *pool* AFQ foram calculados  $VE = 0,52$  e  $CR = 0,53$ . O valor de  $R^2$  para a variável latente endógena *pool* de P disponível foi de 50%.

Para utilizar o banco de dados completo de Oliveira et al. (2014), Tabela 1B, foi verificado, a priori, o ajuste do Modelo C sem a variável mensurada argila, configurando o Modelo C1 (Figura 6). Os resultados dos parâmetros estimados do Modelo C1 não foram estatisticamente diferentes dos apresentados no Modelo C. Os índices de ajuste geral para este modelo foram  $\chi^2 = 15,308$ ,  $df = 16$ ,  $p = 0,502$ , com  $\Delta X^2 = (X^2_{(C)} - X^2_{(C1)}) = 9,48$  e  $\Delta df = 7$ . Para esta diferença de graus de liberdade o valor de delta do qui-quadrado deveria ser  $\geq 26,88$ . Porém, para ser utilizado o banco de dados completo (Organossolos + Chernossolos + Latossolos,  $N = 38$ ) se fez necessária a exclusão da variável argila.

Com a exclusão da variável argila do Modelo C e a utilização do banco de dados completo (Tabela 1B,  $N = 38$ ), os valores dos parâmetros do Modelo C2 (Figura 7) foram estimados e analisados. Os valores de  $\chi^2 = 18,952$ ,  $df = 16$ ,  $p = 0,271$  indicam ajuste geral para o modelo. Os índices de ajuste global para o Modelo C2 também se encontram ajustados (Tabela 6). O valor de  $R^2 = 0,83$  para a variável latente endógena *pool* de P disponível indica uma boa variação explicada.

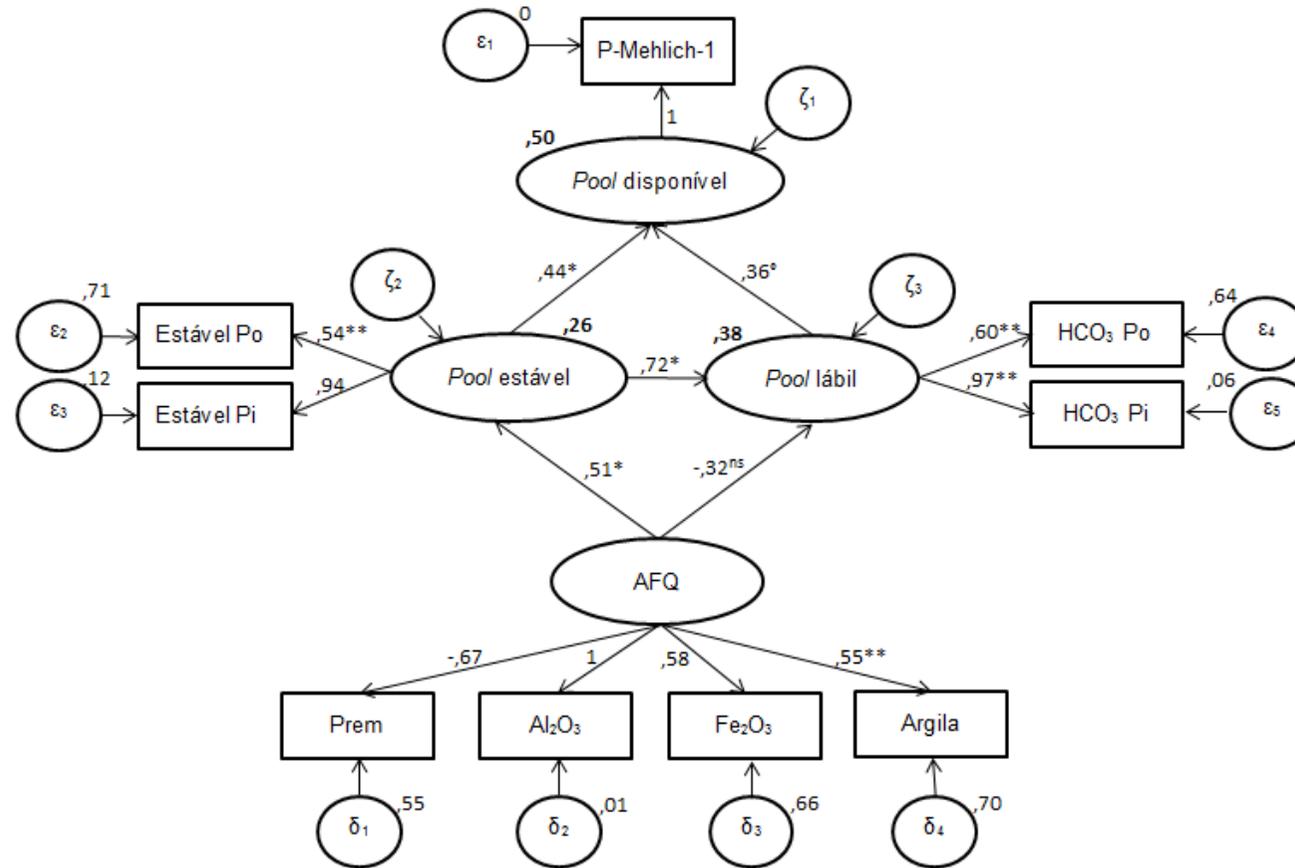


Figura 5. Modelo C, modelo estrutural testado para avaliar o ciclo do P em solos tropicais com alto teor de matéria orgânica,  $N = 28$ , com a inclusão de propriedades do solo. Os indicadores de efeito (retângulos) estão associados com as variáveis latentes (elipses). Os valores numéricos correspondem a cargas fatoriais padronizadas com  $P < 0,001$ . Valores em negrito correspondem ao coeficiente de determinação ( $R^2$ ). “ $^{ns}$  – não significativo, \* significativo a  $P < 0,05$ , \*\* Significativo a  $P < 0,01$ ”. Variâncias de erro ( $\varepsilon_1 - \varepsilon_5$ ,  $\delta_1 - \delta_4$ ) estão padronizadas. Os valores de  $\chi^2 = 24,788$ ,  $df = 23$ ,  $p = 0,361$ .

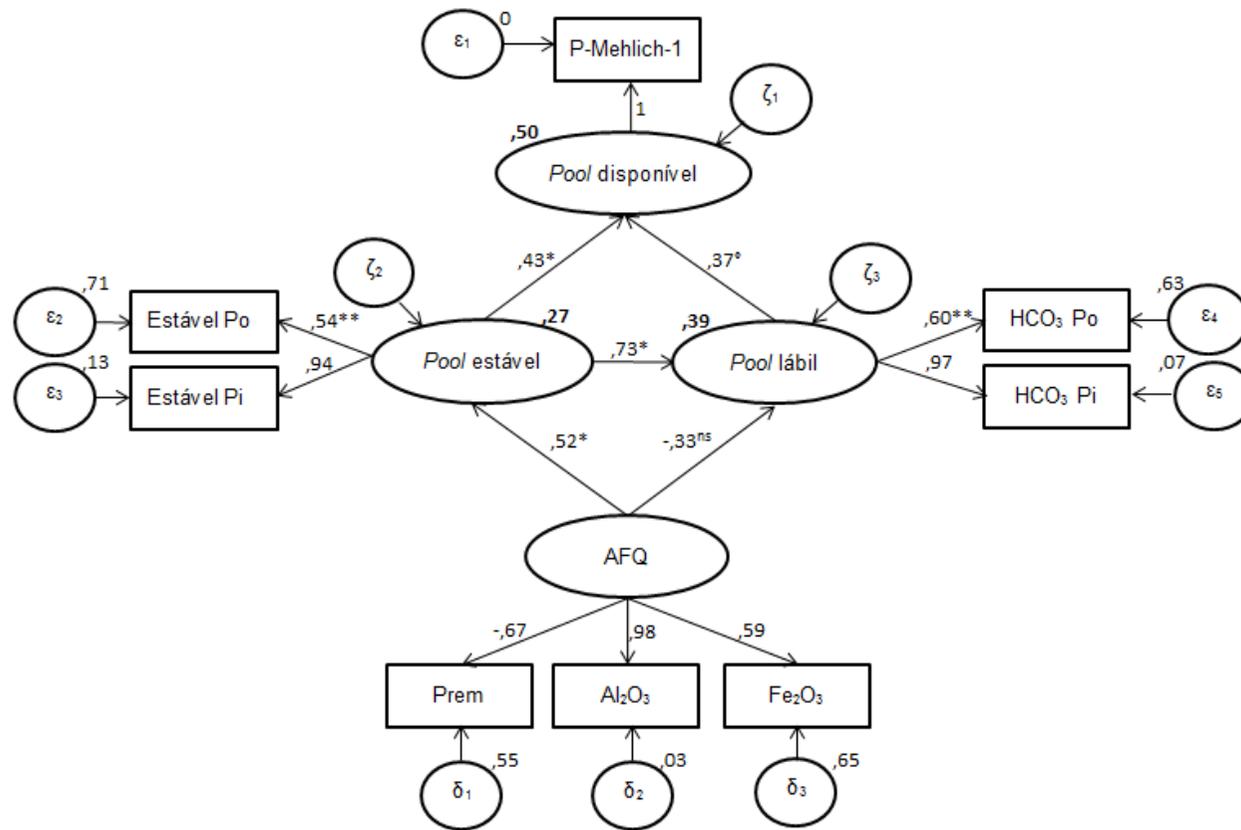


Figura 6. Modelo C1, modelo estrutural testado para avaliar o ciclo do P em solos tropicais com alto teor de matéria orgânica, N = 28, com a inclusão de propriedades do solo. Os indicadores de efeito (retângulos) estão associados com as variáveis latentes (elipses). Os valores numéricos correspondem a cargas fatoriais padronizadas com  $P < 0,001$ . Valores em negrito correspondem ao coeficiente de determinação ( $R^2$ ). “<sup>ns</sup> – não significativo, \* significativo a  $P < 0,05$ , \*\* Significativo a  $P < 0,01$ ”. Variâncias de erro ( $\varepsilon_1 - \varepsilon_5$ ,  $\delta_1 - \delta_3$ ) estão padronizadas. Os valores de  $\chi^2 = 15,308$ ,  $df = 16$ ,  $p = 0,502$ .

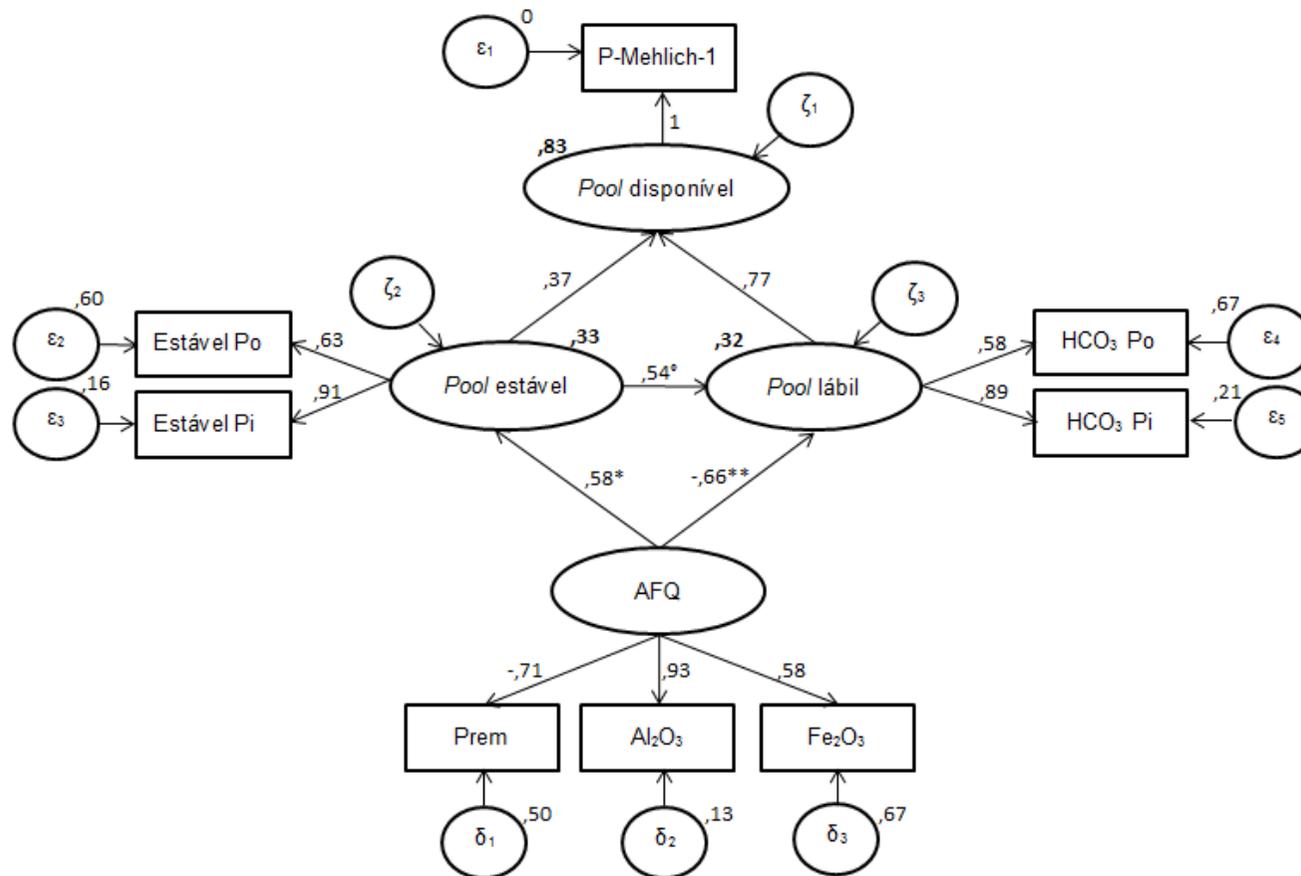


Figura 7. Modelo C2, modelo estrutural testado para avaliar o ciclo do P em solos tropicais com alto teor de matéria orgânica, N = 38, com a inclusão de propriedades do solo. Os indicadores de efeito (retângulos) estão associados com as variáveis latentes (elipses). Os valores numéricos correspondem a cargas fatoriais padronizadas com  $P < 0,001$ . Valores em negrito correspondem ao coeficiente de determinação ( $R^2$ ). “\* significativo a  $P < 0,05$ , \*\* Significativo a  $P < 0,01$ ”. Variâncias de erro ( $\varepsilon_1 - \varepsilon_5$ ,  $\delta_1 - \delta_3$ ) estão padronizadas. Os valores de  $\chi^2 = 18,952$ ,  $df = 16$ ,  $p = 0,271$ .

## DISCUSSÃO

Os modelos construídos apresentam as relações existentes entre as variáveis latentes (*pools* de P) formadas a partir de frações de P extraídas pelo método de Bowman (1989) e as variáveis latentes que representam alguns atributos físico-químicos do solo e matéria orgânica, que influenciam de modo significativo na disponibilidade de P para o solo.

A utilização da modelagem de equações estruturais mostrou que a qualidade do solo influencia na obtenção dos resultados finais do estudo. O banco de dados (Tabela 1A) era composto por solos minerais com baixo teor de matéria orgânica, enquanto o banco de dados (Tabela 1B) era composto por solos minerais com alto teor de matéria orgânica (Organossolos + Chernossolos + Latossolos).

O Modelo A (Figura 2), com a utilização das amostras de solos de baixo teor de matéria orgânica, mostrou inicialmente que a variável latente endógena *pool* de P estável, representada pelas frações de P estáveis ( $P_i$  e  $P_o$ ), possuiu efeito direto negativo, baixo e não significativo ( $\beta = -0,08$ ), sobre o *pool* de P disponível, mostrando ser uma relação dreno de P da solução do solo. No entanto, como efeito global, esta variável atuou como fonte de P para a solução do solo ( $\beta = -0,08 + 0,70 \times 0,82 + 0,49 \times (-0,01) \times 0,82 = 0,49$ ). Este efeito positivo foi justificado pelo grau de relacionamento deste *pool* com o *pool* de P lábil, que apresentou forte associação positiva com o *pool* de P disponível ( $\beta = 0,82$ ), mostrando que, para estes tipos de solos minerais, a maior contribuição para a disponibilidade de P no solo está relacionada às frações lábeis ( $P_i$  e  $P_o$ ).

Com relação à variável latente endógena *pool* MO, ela possui somente efeito indireto via *pool* de P lábil, sendo este negativo e baixo ( $\beta = -0,01 \times 0,82 = -0,01$ ). Para a variável latente *pool* AFQ, os únicos efeitos calculados também foram indiretos, tendo como efeito global um resultado negativo ( $\beta = 0,82 \times (-0,41) \times (-0,01) + (-0,03) \times 0,7 + (-0,17) + (-0,03) \times (-0,08) = -0,15$ ). Estes resultados mostram que a variável latente *pool* AFQ atua como fixadora de P na solução do solo, retendo  $P_i$  com alta energia e estabilizando formas de  $P_o$  (Duda, 2000).

De modo semelhante foram verificados alguns resultados para o Modelo B (Figura 3). O *pool* de P estável se manteve com efeito direto negativo, baixo e

não significativo ( $\beta = -0,02$ ) sobre o *pool* de P disponível, porém o efeito global desta variável foi positivo ( $\beta = 0,54$ ). Para a variável latente endógena *pool* MO, foi verificado que seu efeito indireto é baixo, porém positivo ( $\beta = 0,02$ ), e o *pool* de P lábil se manteve fonte para a disponibilidade de P com ( $\beta = 0,77$ ), corroborando com a hipótese de que o P disponível está fortemente associado às frações de P lábeis.

Em relação ao coeficiente beta ( $\beta$ ) não significativo, encontrado entre o *pool* de P estável e o *pool* de P disponível, pode-se afirmar que a sua permanência no modelo se mostrou de extrema importância para o ajuste geral do modelo, pois a retirada desta relação aumentou o valor do qui-quadrado calculado e tornou o valor- $p$  significativo. Deste modo, a permanência desta relação é válida, e para uma possível reversão na significância do beta calculado, uma alternativa será melhorar a qualidade do banco de dados.

Para validar o Modelo *B* ajustado, foram utilizados dados independentes relacionados a solos minerais com alto teor de matéria orgânica (Organossolos + Chernossolos + Latossolos, Tabela 1B). No entanto, não foi possível fazer a validação deste modelo com esta composição de solos, pois havia no conjunto amostras de solos orgânicos com teor de carbono (C) e nitrogênio (N) altamente correlacionados (Tabela 4). Os valores máximos encontrados para C orgânico e N total neste banco de dados foram respectivamente  $541,1\text{g kg}^{-1}$  e  $15,2\text{g kg}^{-1}$  (Tabela 2, N = 38). A partir desta verificação, foi trabalhado o mesmo banco de dados, porém reduzido (Chernossolos + Latossolos, Tabela 1B), sem as amostras do solo orgânico, e os valores máximos de C orgânico e N total foram respectivamente,  $66,21\text{g kg}^{-1}$  e  $5,21\text{g kg}^{-1}$  (Tabela 2, N = 28) e a correlação entre estas duas variáveis independentes diminuiu (Tabela 5).

A partir destas características foi trabalhado com o banco reduzido (Chernossolos + Latossolos) estimando os parâmetros do Modelo *B1* (Figura 4). No Modelo *B1* ajustado, foi observado que o *pool* de P estável possui uma relação significativa direta com a variável latente endógena *pool* MO ( $\beta = 0,49$ ) e atua como fonte direta de P para a solução do solo ( $\beta = 0,68$ ). O *pool* de P lábil, nesta análise, permaneceu fonte de P, porém com um valor não significativo. As diferenças encontradas entre as análises dos modelos *B* e *B1* foram atribuídas ao uso dos dados referente aos solos com diferentes teores de matéria orgânica, porém ambos os modelos são ajustados, e deste modo o Modelo *B1* confirma

estatisticamente o ajuste do Modelo *B*. Em análise geral, são evidenciados os relacionamentos dos *pools* de P com o *pool* de matéria orgânica, pela presença da variável mensurada C orgânico que possui forte correlação com as frações de Po (Duda, 2000; Cunha et al., 2007; Zaia et al., 2008a; Oliveira et al., 2014).

Com relação ao Modelo C (Figura 5) e suas reespecificações, Modelo C1 e C2 (Figuras 6 e 7), foi observado que suas relações estruturais se mantiveram com a mesma magnitude, valores de betas estatisticamente iguais. Os *pools* de P estável e lábil atuaram como fontes diretas sobre o *pool* de P disponível, no entanto, no Modelo C e C1, com a utilização das amostras dos Chernossolos + Latossolos (N = 28), a contribuição do *pool* de P estável foi maior do que a contribuição do *pool* de P lábil, e no Modelo C2, com a utilização das amostras de Organossolos + Chernossolos + Latossolos (N = 38) esta relação se inverteu, logo os solos orgânicos influenciam diretamente nas formas de P que estão estreitamente relacionadas ao P disponível. Nestes modelos ainda foi verificado que a variável latente *pool* AFQ do solo atua contribuindo a fixação do P lábil e na manutenção de P estável corroborando com Zaia et al. (2008b).

Através dos índices de ajuste dos modelos (Tabela 6) foi observado que estes modelos se ajustaram aos dados e que o Modelo C2 confirmou o Modelo C1, uma vez que foram trabalhados com tamanho de amostras diferentes. Com a utilização do índice de ajuste AIC foi possível comparar o Modelo C com os Modelos C1 e C2, modelos alternativos para solos minerais com alto teor de matéria orgânica, verificando que estes dois modelos foram estatisticamente mais ajustados do que o Modelo C.

Na comparação dos modelos concorrentes, Modelo B1 e Modelo C1, que foram analisados com o mesmo banco de dados (Tabela 1B, Chernossolos + Latossolos, N = 28), foi possível verificar que tanto o índice AIC quanto o índice  $\chi^2/df$  foram menores para o Modelo B1, desta maneira tornando-o estatisticamente mais ajustado aos dados do que o modelo alternativo C1.

## CONCLUSÕES

Com a modelagem de equações estruturais foi possível analisar, através de efeitos diretos e indiretos, as relações existentes entre os *pools* de P estável e

lável sobre o P disponível, além de proporcionar a análise das relações existentes das variáveis latentes correspondentes às propriedades do solo sobre a disponibilidade de P na solução do solo.

Foi possível verificar também que, o uso de amostras de solos com teores orgânicos diferentes influenciou nos resultados obtidos da modelagem.

Em solos com baixo teor de matéria orgânica as principais fontes diretas de P disponível foram as frações lábeis ( $P_i$  e  $P_o$ ) e o *pool* de matéria orgânica para este modelo agiu como fonte indireta na manutenção de P no solo.

Em solos com alto teor de matéria orgânica tanto as frações lábeis quanto as frações estáveis foram fontes significativas de P disponível, e o *pool* de atributos físico-químicos corroborou a teoria, pois age diretamente como agente fixador de P lábil.

O Modelo *B1* foi o modelo que melhor se ajustou aos dados, tanto utilizando dados de solos minerais com baixo ou com alto teor de matéria orgânica, validando a análise da disponibilidade do P no solo realizada através da modelagem de equações estruturais.

## REFERÊNCIAS BIBLIOGRÁFICAS

- Barto EK, Alt F, Oelmann, Y, Wilcke W, Rilling M (2010). Contributions of biotic and abiotic factors to soil aggregation across a land use gradient. *Soil Biology & Biochemistry* 42, 2316-2324.
- Bollen KA and Noble MD (2011). Structural equation models and the quantification of behavior. *PNAS*, 108:15639–15646.
- Bowman, RA (1989). A sequential extraction procedure with concentrated sulfuric acid and diluted base for soil organic phosphorus. *Soil Science Society America Journal*, Madison, v.53, p. 326-366.
- Byrne BM (1994). *Structural equation modeling with EQS and EQS/windows: basic concepts, applications and programming*. Thousand Oaks: Sage Publications.
- Byrne BM (2009). *Structural equation modeling with AMOS: basic concepts, applications and programming*, 2ª edição. Routledge, 2009.

- Costa, MG (2015). Frações de P no solo em diferentes sítios de Eucalyptus spp. no sudeste do Brasil. Dissertação apresentada a Universidade Estadual do Norte Fluminense Darcy Ribeiro, Campos dos Goytacazes – RJ.
- Cunha GM, Gama-Rodrigues AC, Costa GS, Velloso ACX (2007). Fósforo orgânico em solos sob florestas montanas, pastagens e eucalipto no Norte Fluminense. *Revista Brasileira de Ciência do Solo*, 31, 667-672.
- Duda, GP (2000). Conteúdo de fósforo microbiano, orgânico e biodisponível em diferentes classes do solo. Tese, UFFRJ.
- Fornell C, Larcker DF (1981). Evaluating Structural Equation with unobservable variables and measurement error. *Journal of marketing Research*, vol.18, nº 1, 39-50.
- Gama-Rodrigues AC, Sales MVS, Silva PSD, Comerford NB, Cropper WP, Gama-Rodrigues EF (2014). An exploratory analysis of phosphorus transformations in tropical soils using structural equation modeling. *Biogeochemistry*, 118; 453-469.
- Grace JB, Pugsek BH (1998). On the use of path analysis and related procedures for the investigation of ecological problems. *Am. Nat.* 152, 151–159.
- Grace JB (2006). *Structural equation modeling and natural systems*. Cambridge University Press, New York, p 365.
- Grace JB, Schoolmaster DR Jr, Guntenspergen GR, Little AM, Mitchell BR, Miller KM, Schweiger W (2012). Guidelines for a graph-theoretic implementation of structural equation modeling. *Ecosphere*, 3(80).
- Guerra, J. G. M., Almeida, D. L. d., Santos, G. d. A., Fernandes, M. S. (1996). Conteúdo de fósforo orgânico em amostras de solo. *Pesquisa Agropecuária Brasileira*, 31(4), 291-299.
- Hair JF Jr, Anderson RE, Tatham RL, Black WC (2009). *Análise Multivariada de Dados*. Porto Alegre: Bookman, 6<sup>a</sup> ed.
- Iacobucci D (2010). Structural equations modeling: Fit Indices, sample size, and advanced topics. *Journal of Consumer Psychology*, 20, 90 - 98.
- Kenny DA (2014). Measuring model fit. (<http://davidakenny.net/cm/fit.htm>).
- Oliveira RI, Gama-Rodrigues AC, Gama-Rodrigues EF, Zaia FC, Pereira MG, Fontana A (2014). Organic phosphorus in diagnostic surface horizons of

different brazilian soil orders. *Revista Brasileira de Ciência do Solo*, 38, 1411-1420.

Pugesek BH, Tomer A, Avon E (2003). *Structural equation modeling*. Cambridge University Press, Cambridge.

Rita JCO, Gama-Rodrigues AC, Gama-Rodrigues EF, Zaia FC, Nunes DAD (2013). Mineralization of organic phosphorus in soil size fractions under different vegetation covers in the north of Rio de Janeiro. *Revista Brasileira de Ciência do Solo*, 37:1207-1215.

Sales MVS, Gama-Rodrigues AC, Comerford NB, Cropper WP, Gama-Rodrigues EF, Oliveira PHG (2015). Respecification of structural equation models for the P cycle in tropical soils. *Nutr Cycl Agroecosyst*, 102, 1 - 16.

Sato S, Comerford NB (2006). Organic anions and phosphorus desorption and bioavailability in a humid Brazilian Ultisol. *Soil Sci* 171:695–705.

Zaia FC, Gama-Rodrigues AC, Gama-Rodrigues EF (2008a). Formas de fósforo no solo sob leguminosas florestais, floresta secundária e pastagem no Norte Fluminense. *Revista Brasileira de Ciência do Solo*, 32, 1191-1197.

Zaia FC, Gama-Rodrigues AC, Gama-Rodrigues EF, Machado RCR (2008b). Fósforo orgânico em solos sob agrossistemas de cacau. *Revista Brasileira de Ciência do Solo*, 32, 1987-1995.

#### 4. RESUMO E CONCLUSÕES

Com o uso da modelagem de equações estruturais, com variáveis latentes, foram elaborados modelos para analisar o ciclo do P em solos tropicais. Nos modelos foram utilizados dados experimentais relacionados ao método de fracionamento de Hedley, e ao método de Bowman, que distinguem as frações de P de acordo com a sua labilidade. Nestes, foram analisadas relações causais entre os *pools* de P (variáveis latentes) formados através das frações de P extraídas em cada método.

Nos modelos com o fracionamento sequencial de Hedley, para solos tropicais, foi observada a forte relação positiva direta que o *pool* de P orgânico tem sobre o P disponível, além de servir como caminho indireto para os *pools* de P ocluso e mineral primário. Os *pools* de P orgânico e mineral primário atuaram como fontes de P disponível em toda análise, no entanto, o *pool* de P ocluso somente possui efeito indireto positivo, mostrando a sua relação dreno no efeito direto sobre o *pool* de P disponível.

Com os dados do método Bowman foram analisados dois cenários. O primeiro, com a utilização de solos minerais com baixo teor de matéria orgânica. Neste modelo *pool* de P lábil atuou como fonte direta de P disponível e o *pool* de matéria orgânica atuou diretamente na manutenção de P estável no solo.

No segundo cenário, com a utilização de solos minerais com alto teor de matéria orgânica, tanto o *pool* de P lábil quanto de P estável atuaram como fontes

de P disponível e o *pool* das propriedades físico-químicas atuou negativamente sobre o *pool* de P lábil, mostrando seu caráter drenante, atuou positivamente sobre o *pool* de P estável, mostrando sua importância na manutenção de formas de P estável no solo.

A modelagem de equações estruturais proporcionou neste estudo testes que analisam a consistência e plausibilidade dos modelos do ciclo do P, confirmando, rejeitando ou construindo novas hipóteses.

Em trabalhos futuros, modelos serão testados e comparados com o refinamento do banco de dados e inclusão de novas variáveis mensuráveis e latentes. Além dos dados obtidos pelos métodos de Hedley e de Bowman para validação de modelos, serão obtidos dados com o método de Espectroscopia de Ressonância Magnética Nuclear  $^{31}\text{P}$  para a análise de novos modelos.

## 5. REFERÊNCIAS BIBLIOGRÁFICAS

- Agbenin JO, Iwuafor ENO, Ayuba B (1999) A critical assessment of methods for determining organic phosphorus in savanna soils. *Biol Fertil Soils* 28:177–181.
- Arbuckle, JL (2009) *Amos 18 User's Guide*. Chicago, IL: SPSS.
- Allison, VJ, Yermakov, Z, Miller, RM, Jastrow, JD, Matamala, R (2007) Using landscape and depth gradients to decouple the impact of correlated environmental variables on soil microbial community composition. *Soil Biology & Biochemistry* 39:505–516.
- Arhonditsis GB, Stow CA, Steinberg LJ, Kenney MA, Lathrop RC, McBride SJ, Reckhow KH (2006) Exploring ecological patterns with structural equation modeling and Bayesian analysis. *Ecological Modelling* 192:385–409.
- Bernatchez F, Jeannotte R, Begg CBM, Hamel C, Whalen JK (2008) Soil fertility and arbuscular mycorrhizal fungi related to trees growing on smallholder farms in Senegal. *Journal of Arid Environments* 72:1247 – 1256.
- Barto EK, Alt F, Oelmann, Y, Wilcke W, Rilling M (2010) Contributions of biotic and abiotic factors to soil aggregation across a land use gradient. *Soil Biology & Biochemistry* 42, 2316-2324.
- Beck MA, Sanchez PA (1994) Soil phosphorus fraction dynamics during 18 years of cultivation on a Typic Paleudult. *Soil Sci Soc Am J* 58:1424-1430.
- Bistaffa, BC (2010) Incorporação de indicadores categóricos ordinais em modelos de equações estruturais. Dissertação. USP – São Paulo.

- Bollen KA (1989) Structural equation with latent variables. John Wiley and Sons, New York, p 514.
- Bollen KA, Noble MD (2011) Structural equation models and the quantification of behavior. *Proceeding of the National Academy of Sciences* 108:15639-15646.
- Bowman, RA (1989) A sequential extraction procedure with concentrated sulfuric acid and diluted base for soil organic phosphorus. *Soil Science Society America Journal, Madison*, v.53, p. 326-366.
- Byrne BM (1994) Structural equation modeling with EQS and EQS/windows: basic concepts, applications and programming. Thousand Oaks: Sage Publications.
- Byrne BM (2009) Structural equation modeling with AMOS: basic concepts, applications and programming, 2ª edição. Routledge, 2009.
- Costa, MG (2015). Frações de P no solo em diferentes sítios de *Eucalyptus* spp. no sudeste do Brasil. Dissertação apresentada a Universidade Estadual do Norte Fluminense Darcy Ribeiro, Campos dos Goytacazes – RJ.
- Cross AF, Schlesinger WH (1995) A literature review and evaluation of the Hedley fractionation: application to the biogeochemical cycle of soil phosphorus in natural ecosystems. *Geoderma* 64:197-214.
- Cunha GM, Gama-Rodrigues AC, Costa GS, Velloso ACX (2007) Fósforo orgânico em solos sob florestas montanas, pastagens e eucalipto no Norte Fluminense. *Revista Brasileira de Ciência do Solo*, 31, 667-672.
- Dieter D, Elsenbeer H, Turner B (2010) Phosphorus fractionation in lowland tropical rainforest soils in central Panama. *Catena* 82:118-125.
- Doblas-Miranda, E, Sánchez-Piñero, F, González-Megías, A (2009) Different structuring factors but connected dynamics shape litter and belowground soil macrofaunal food webs. *Soil Biology & Biochemistry* 41:2543–2550.
- Duda, GP (2000) Conteúdo de fósforo microbiano, orgânico e biodisponível em diferentes classes do solo. Tese, UFFRJ.
- Fornell C, Larcker DF (1981) Evaluating Structural Equation with unobservable variables and measurement error. *Journal of marketing Research*, vol.18, nº 1, 39-50.
- Gama-Rodrigues AC, Sales MVS, Silva PSD, Comerford NB, Cropper WP, Gama-Rodrigues EF (2014) An exploratory analysis of phosphorus transformations in

- tropical soils using structural equation modeling. *Biogeochemistry* 118:453-469.
- Guerra, J. G. M., Almeida, D. L. d., Santos, G. d. A., Fernandes, M. S. (1996) Conteúdo de fósforo orgânico em amostras de solo. *Pesquisa Agropecuária Brasileira*, 31(4), 291-299.
- Gijsman AJ, Oberson A, Tiessen H, Friesen DK (1996) Limited applicability of the CENTURY model to highly weathered tropical soils. *Agron J* 88:894-903.
- Grace JB, Pugesek BH (1998) On the use of path analysis and related procedures for the investigation of ecological problems. *Am. Nat.* 152, 151–159.
- Grace JB (2006) *Structural equation modeling and natural systems*. Cambridge University Press, New York, p 365.
- Grace JB, Bollen KA (2006) *The interface Between theory and data in structural equation models*: U. S. Geological Survey Open-File Report 2006-1363, 33p.
- Grace JB, Bollen KA (2008) Representing general theoretical concepts in structural equation models: the role of composite variables. *Environ Ecol Stat* 15:191-213.
- Grace JB, Anderson TM, Olf H, Scheiner SM (2010) On the specification of structural equation models for ecological systems. *Ecol Monogr* 80:67-87 .
- Grace JB, Schoolmaster DR, Guntenspergen GR, Little AM, Mitchell BR, Miller KM, Schweiger EW (2012) Guidelines for a graph-theoretic implementation of structural equation modeling. *Ecosphere* 3:1-44.
- Guppy CN, Menzies NW, Moody PW, Blamey FPC (2005) Competitive sorption reactions between phosphorus and organic matter in soil: a review. *Australian J Soil* 43:189-202.
- Hair JF, Black WC, Babin BJ, Anderson E, Tathan RL (2009) *Multivariate data analysis*. Prentice-Hall Inc., Upper Saddle River, p 688.
- He Z, Fortuna AM, Senwo ZN, Tazisong IA, Honeycutt CW, Griffin TS (2006) Hydrochloric fractions in Hedley fractionation may contain inorganic and organic phosphates. *Soil Sc Soc Am J* 70:893-899.
- Hedley MJ, Stewart WB, Chauhan BS (1982) Changes in inorganic and organic soil phosphorus fractions induced by cultivation practices and by laboratory incubations. *Soil Sci Soc Am J* 46:970-976.

- Iacobucci D (2009) Every you always wanted to know about SEM (structural equation modeling) but were afraid to ask. *Journal of Consumer Psychology*, 19, 673-680.
- Iacobucci D (2010) Structural equations modeling: Fit Indices, sample size, and advanced topics. *Journal of Consumer Psychology*, 20, 90 - 98.
- Kaplan, D (2000) *Structural Equation Modeling: foundations and extensions*. Sage: Thousand Oaks.
- Klem, L (2006) Structural equation modeling. In L.G. Grimm & P.R Yarnold (Eds.), *Reading and understanding more multivariate statistics* (pp. 227-260). Washington: American Psychological Association.
- Kline RB (2011) *Principles and Practice of Structural Equation Modeling*. Third edition.
- Kenny DA (2014) Measuring model fit. (<http://davidakenny.net/cm/fit.htm>).
- Lambers H, Raven JA, Shaver GR, Smith S (2008) Plant nutrient-acquisition strategies change with soil age. *Trends in Ecology and Evolution* 23:95-103
- Malaeb, ZA, Summers, JK, Pugsek, BH (2000) Using structural equation modeling to investigate relationships among ecological variables. *Environmental and Ecological Statistics* 7:93–111.
- Marklein AR, Houlton BZ (2012) Nitrogen inputs accelerate phosphorus cycling rates across a wide variety of terrestrial ecosystems. *New Phytologist* 193:696-704.
- McGill WB, Cole CV (1981) Comparative aspects of cycling of organic C, N, S, and P through soil organic matter. *Geoderma* 26:267-286.
- Mitchell RJ (1992) Testing evolutionary and ecological hypotheses using path analysis and structural equation modeling. *Funct Ecol* 6:123-129.
- Nasto MK, Alvarez-Clare S, Lekberg Y, Sullivan BW, Townsend AR, Cleveland CC (2014) Interactions among nitrogen fixation and phosphorus acquisition strategies in lowland tropical rain forests. *Ecology Letters* 17:1282-1289.
- Novais, R. F., Smyth, T. J. (1999) *Fósforo em solo e planta em condições tropicais* (UFV Ed.). Viçosa - MG, Brasil.
- Oberson A, Friesen DK, Rao IM, Bühler S, Frossard E (2001) Phosphorus transformations in an Oxisol under contrasting land use systems: role of the soil microbial biomass. *Plant Soil* 237:197-210.

- Oliveira RI, Gama-Rodrigues AC, Gama-Rodrigues EF, Zaia FC, Pereira MG, Fontana A (2014) Organic phosphorus in diagnostic surfasse horizons of different brazilian soil orders. *Revista Brasileira de Ciência do Solo*, 38, 1411-1420.
- Pugesek BH, Tomer A, Eye A (2003) *Structural equation modeling: Applications in ecological and evolutionary biology*. Cambridge University Press, New York, p 409.
- Rita JCO, Gama-Rodrigues AC, Gama-Rodrigues EF, Zaia FC, Nunes DAD (2013) Mineralization of organic phosphorus in soil size fractions under different vegetation covers in the north of Rio de Janeiro. *Revista Brasileira de Ciência do Solo*, 37:1207-1215.
- Reed SC, Townsend AR, Taylor PG, Cleveland CC (2011) Phosphorus cycling in tropical forests growing on highly weathered soils. In: Bünenmann, EK, Oberson A, Frossard E (ed) *Phosphorus in action: Biological processes in soil phosphorus cycling*. Springer, Berlin, pp 339-369.
- Richardson AE, Lynch JP, Ryan PR, Delhaize E, Smith FA, Smith SE, Harvey PR, Ryan MH, Veneklaas EJ, Lambers H, Oberson A, Culvenor RA, Simpson RJ (2011) Plant and microbial strategies to improve the phosphorus efficiency of agriculture. *Plant Soil* 349:121-156.
- Riginos C, Grace JB (2008) Savanna tree density, herbivores, and the herbaceous community: Bottom-UP vs. Top-Down effects. *Ecology*, 89(8), 2228-2238.
- Sales MVS, Gama-Rodrigues AC, Comerford NB, Cropper WP, Gama-Rodrigues EF, Oliveira PHG (2015) Respecification of structural equation models for the P cycle in tropical soils. *Nutr Cycl Agroecosyst*, 102, 1 - 16.
- Sato S, Comerford NB (2006) Organic anions and phosphorus desorption and bioavailability in a humid Brazilian Ultisol. *Soil Sci* 171:695-705.
- Tiessen H, Moir JO (1993) Characterisation of available P by sequential extraction. In: Carter MR (ed) *Soil sampling and methods of soil analysis*. CRC Press, Boca Raton, FL, pp 75-86.
- Tiessen H, Stewart WB, Cole CV (1984) Pathways of phosphorus transformation in soils of differing pedogenesis. *Soil Sci Soc Am J* 48:853-858.
- Travis SE, Grace JB (2010) Predicting performance for ecological restoration: a case study using *Spartina alterniflora*. *Ecological Applications*, 20(1), 192-204.

- Turner BL (2008) Resource partitioning for soil phosphorus: a hypothesis. *J Ecology* 96:698-702.
- Turner BL, Wells A, Condon LM (2014) Soil organic phosphorus transformations along a coastal dune chronosequence under New Zealand temperate rain forest. *Biogeochemistry* 121:595-611.
- Ullman, JB (2007) Structural Equation Modeling. Em B. G. Tabachnick & L. S. Fidell (Orgs.), *Using multivariate statistics* (5<sup>a</sup> ed.). Boston: Pearson Education.
- Vitousek PM, Porder S, Houlton BZ, Chadwick OA (2010) Terrestrial phosphorus limitation: mechanisms, implications, and nitrogen-phosphorus interactions. *Ecological Applications* 20:5-15.
- Wang YP, Houlton BZ, Field CB (2007) A model of biogeochemical cycles of carbon, nitrogen, and phosphorus including symbiotic nitrogen fixation and phosphatase production. *Global Biogeochemical Cycles* 21: GB1018, doi:10.1029/2006GB002797.
- Zaia FC, Gama-Rodrigues AC, Gama-Rodrigues EF (2008a) Formas de fósforo no solo sob leguminosas florestais, floresta secundária e pastagem no Norte Fluminense. *Revista Brasileira de Ciência do Solo*, 32, 1191-1197.
- Zaia FC, Gama-Rodrigues AC, Gama-Rodrigues EF, Machado RCR (2008b) Fósforo orgânico em solos sob agrossistemas de cacau. *Revista Brasileira de Ciência do Solo*, 32, 1987-1995.
- Zheng Z, Simard RR, Lafond J, Parent LE (2002) Pathways of soil phosphorus transformations after 8 years of cultivation under contrasting cropping practices. *Soil Sci Soc Am J* 66:999-1007.