

Estimação de um Modelo Discriminante para Diagnóstico do Estado de Saúde de Árvores

Estimation of a Discriminant Model for Diagnosis of the State of health of Trees

Giovani Glaucio Costa^{1*}

Resumo: A avaliação da condição de árvores urbanas é de suma importância para o manejo da arborização de ruas e a gestão da floresta urbana como um todo. Este artigo visa contribuir para o estudo de metodologias que permitam avaliar em que condições se encontram as florestas urbanas, buscando a elaboração de projetos administrativos mais eficientes e fundamentar as ações preventivas de poderes públicos de desenvolvimento de espaços urbanos que proporcionem qualidade de vida às populações de grandes centros urbanos. Um levantamento reuniu de 2013 a 2014, por amostragem não probabilística, 190 árvores de espaços urbanos no estado do Rio de Janeiro. Os pesquisadores teriam que indicar 1 para a presença do item de enfermidade e 0 para a ausência do item de enfermidade. Com a utilização das técnicas multivariadas *two step cluster* e *análise discriminante*, formalizou-se o modelo classificatório, que se denominou *MODIG-SAR* ("Modelo Diagnóstico da Saúde da Árvore"). O *MODIG-SAR* será uma das contribuições sugeridas para auxiliar às tomadas de decisão relacionadas ao *check up* sistemático e necessário de árvores urbanas. A proporção de sucessos em ambos os critérios de validação indicam que se erraria muito pouco ao utilizar o modelo construído no diagnóstico da saúde de árvores urbanas. A investigação é meramente preliminar, o ideal seria uma amostra com maior representatividade, mas abre caminhos para estudos mais generalizáveis sobre a temática estudada.

Palavras-chaves: diagnóstico, saúde, árvore, two step cluster, análise discriminante.

Abstract: The assessment of the condition of urban trees is of paramount importance for the management of street tree planting and urban forest management as a whole. This article aims to contribute to the study of methodologies to assess the conditions under which urban forests, are seeking to draw up administrative projects more efficient and substantiate the preventive actions of the public authorities of development of urban spaces that provide quality of life for the population of major urban centres. A survey of 2013 to 2014, met by non-probability sampling, 190 trees of urban spaces in the State of Rio de Janeiro. Researchers would have to indicate 1 for the presence of the item of illness and 0 for the absence of the item of illness. With the use of multivariate techniques *two step cluster* and *discriminant analysis*, formalized the model classification, which is termed *MODIG-SAR* ("Health Diagnostic Model of the tree"). The *MODIG-SAR* is one of the suggested contributions to assist the decision-making related to systematic Checkup and necessary of urban trees. The proportion of successes in both validation criteria indicate that if Miss very little when using the model built in the diagnosis of the health of urban trees. The investigation is merely preliminary, ideally a sample with greater representativeness, but it opens paths to more generalizable studies on the subject studied.

Key words: diagnosis, health, tree, two step cluster, discriminant analysis.

*Autor para correspondência

Recebido para publicação em 16/04/2016; aprovado em 22/11/2016

¹Estatístico. Universidade Federal Rural do Rio de Janeiro. E-mail. giovaniglaucio@hotmail.com

INTRODUÇÃO

A condição de uma árvore urbana pode ser avaliada por meio da sua saúde, vigor, vitalidade, taxa de crescimento, imperfeições físicas, infestações e expectativa de vida, expressando o estado em que a árvore se encontra e não a condição do local onde ela está se desenvolvendo. A avaliação da condição de árvores urbanas é de suma importância para o manejo da arborização de ruas e a gestão da floresta urbana como um todo.

A necessidade de remoção de árvores, poda ou tratamentos fitossanitários pode ser detectada e quantificada ao avaliar-se a condição das árvores. Desta forma, a condição é um fator importante na metodologia de avaliação monetária de árvores que segue diversas variáveis, como o fator condição, agrupadas em uma equação matemática, e avaliação de árvores de risco que leva em consideração a condição da árvore e possíveis alvos do patrimônio público a privado que podem ser atingidos quando houver a queda da árvore ou de galhos (ARAUJO; ARAUJO, 2002).

Em 1931, a Associação de Parques Florestais de Michigan (MFPA) já demonstrava preocupações quanto a desenvolver metodologias eficientes para avaliar a condição de árvores urbanas, publicando as primeiras diretrizes, que a partir de 1956 foram revisadas e atualizadas. No ano de 1997 a MFPA passou a ser um capítulo da Sociedade Internacional de Arboricultura (ISA), realizando-se em 1982 uma revisão significativa, adequando o guia de avaliação da MFPA de acordo com as diretrizes da ISA. No Brasil, a primeira versão adaptada dessa metodologia foi apresentada no VII Encontro Nacional sobre Arborização Urbana em Belo Horizonte, MG (ARAUJO, 1997. p.69). Este método foi citado por Coltro e Miranda (2007), entretanto nenhum resultado foi apresentado utilizando esta metodologia, não tendo assim nenhum trabalho demonstrando os resultados da aplicação deste modelo no país.

Com a crescente busca de espaços verdes pela sociedade atual, a arborização de vias públicas, parques e praças vêm sendo uma grande alternativa para as cidades, pois as árvores urbanas quando bem implantadas e manejadas transmitem múltiplos benefícios e estão longe de serem caracterizadas como um problema. Os benefícios podem ser de ordem estética, relacionados principalmente a efeitos visuais que as plantas fornecem, proporcionando um contato do homem com a natureza, produzindo um efeito harmônico através de suas linhas suaves e orgânicas, formas, cores e texturas, criando paisagens específicas, proporcionando identidade às ruas. Quanto ao meio ambiente incluem benefícios na melhoria do microclima, amenização da poluição atmosférica e acústica, proteção do solo e fauna, diminuição da velocidade do vento e fornecendo sombra a população humana. Fontes benéficas de ordem física estão relacionadas ao conforto que as árvores proporcionam com a alteração do microclima, já as psíquicas trazem benefícios incalculáveis ao homem, principalmente na atenuação do estresse.

A arborização urbana também é positiva no fomento ao turismo, incentivando atividades lúdicas e físicas o que pode propiciar uma diminuição do sedentarismo populacional e no desenvolvimento da educação ambiental, estando assim, relacionada a valores políticos, sociais e econômicos (BIONDI; ALTHAUS, 2005). Proporciona também maior absorção da radiação ultravioleta, dióxido

de carbono e redução do impacto da água da chuva sobre o solo com diminuição de seu escoamento superficial pelo aumento da permeabilidade (MILANO; DALCIN, 2000). Contribui desta forma como fonte de recreação e lazer urbanos, estimulando a ampliação do círculo social (SOUZA, 1995).

Para produção de efeitos positivos, as árvores devem ser submetidas a adequados tratamentos silviculturais, desde a produção de mudas de alta qualidade à manutenção com podas regulares em indivíduos adultos. No Brasil, o que ocorre na maioria das vezes, salvo raras exceções, é que a arborização urbana não passa por um planejamento prévio, sendo em sua maioria executada com improvisos de técnicos não especializados na área, não havendo existência de políticas dirigidas ao setor, assim como conscientização da população de sua importância (SANTOS, 2001).

Contudo essa mentalidade está aos poucos se modificando, a notar pelos eventos da área como o Congresso Nacional de Arborização Urbana organizado pela Sociedade Brasileira de Arborização Urbana (SBAU). Resta então, a correção dos erros já constatados (escolha inadequada de espécies, plantio em locais inadequados, podas mal feitas, etc.) e a potencialização das ações notoriamente bem sucedidas, como em Porto Alegre - RS que possui um completo plano de manejo de arborização de vias públicas, Maringá - PR que é considerada uma das cidades mais arborizadas do país, Curitiba - PR - que dispõe de grande tecnologia, experiência e profissionais habilitados, obtendo ótimos resultados, e outras cidades como Vitória - ES, Recife - PE, Manaus - AM. Para tal, um inventário de análise da condição é fundamental, permitindo conhecer a situação de cada município, quais atitudes devem ter continuidade e que valores devem ser repensados (SANTOS; TEIXEIRA, 2001).

Quando a arborização é realizada sem um bom planejamento, observam-se problemas como a destruição de calçadas pelo excessivo afloramento de raízes devido à impermeabilização do solo decorrente da presença de pavimentação, disseminação de pólen alergênico na população, queda de galhos deficientes de poda, infestações por doenças e pragas como a erva-de-passarinho, gerando perdas de valores estéticos, danos a rede de fiação elétrica, dentre outros (ARAUJO; ARAUJO, 2002).

Todavia, com o recente crescimento da arboricultura no Brasil, gestores municipais têm voltado suas atenções quanto à qualidade das árvores que ocupam os espaços urbanos na busca da otimização de suas diversas funções. Desta maneira, há pesquisadores que demonstram um forte interesse no estudo de metodologias que permitem avaliar em que condições se encontram as florestas urbanas, buscando a elaboração de projetos administrativos mais eficientes. Biondi e Reissmann (1997) realizaram estudos relacionados, avaliando o vigor das espécies *Acer negundo* L. e *Tabebuia chrysotricha* (Mart ex DC.) Standl., seguindo parâmetros quantitativos de avaliação, como: diâmetro de copa, perímetro à altura do peito, peso de folhas, alongação do ramo, peso do ramo e área foliar. Gonçalves et al. (2007) preocupou-se com a criterização para análise de parâmetros que possam determinar a supressão de determinado indivíduo arbóreo, analisando diversas variáveis como a paisagística verificando a raridade da espécie no ambiente estudado; a afetividade, que diz respeito ao valor de estima da população pelo indivíduo arbóreo e o posicionamento em

relação ao contexto urbano; a ecológica, que se refere à natividade da espécie, idade e importância ecológica do indivíduo no local; fitossanitária, presença de doenças, ocorrência de pragas, outros problemas como madeira frágil e riscos avaliados pelos conflitos aéreos, subterrâneos e iminência de queda.

Este artigo visa contribuir para o estudo de metodologias que permitem avaliar em que condições se encontram as florestas urbanas, buscando a elaboração de projetos administrativos mais eficientes e fundamentar as ações preventivas de poderes públicos de desenvolvimento de espaços urbanos que proporcionem qualidade de vida às populações de grandes centros urbanos.

O objetivo deste estudo é estimar um modelo discriminante que possa classificar uma árvore urbana em um dos três grupos: *Árvore Saudável*, *Árvore Médio Saudável* e *Árvore Enferma*.

A finalidade do modelo supra especificado é realizar um diagnóstico da saúde de árvores urbanas, que possa produzir aos gestores municipais informações quanto à qualidade das árvores que ocupam os espaços urbanos na busca da otimização de suas diversas funções.

Portanto, este artigo visa criar uma ferramenta científica que possa auxiliar às tomadas de decisão relacionadas ao *check up* sistemático necessário de árvores urbanas.

MATERIAL E MÉTODOS

Numa tentativa preliminar de contribuir para a monitoração do estado de qualidade de vida de espaços urbanos, este artigo se propõe estabelecer uma função discriminante que possibilite classificar uma determinada árvore urbana em um dos perfis de saúde (*Árvore Saudável*, *Árvore Médio Saudável* e *Árvore Enferma*) e assim se disponibilizar uma “ferramenta estatística” para diagnosticar o estado de saúde de árvores urbanas.

Para se obter o modelo diagnóstico da saúde de árvores urbanas, será necessário realizar uma pesquisa de avaliação junto a árvores do espaço urbano, baseada numa lista de indicadores de enfermidades. Esta lista consta da Quadro 1. Um questionário, baseado nesta lista, que medisse o nível de enfermidade, baseado em 12 estados de árvores enfermas, seria aplicado por técnicos de silvicultura e engenharia florestal a uma amostra aleatória simples de árvores urbanas no estado do Rio de Janeiro.

O levantamento reuniu de 2013 a 2014, por amostragem não probabilística, 190 árvores de espaços urbanos no Estado do Rio de Janeiro. Estudantes de silvicultura e engenharia florestal avaliaram a saúde das árvores baseada na lista de 12 indicadores de enfermidade de árvores. Os pesquisadores teriam que indicar 1 para a presença do item de enfermidade e 0 para a ausência do item de enfermidade.

Com a utilização de técnica multivariada de segmentação, as árvores amostradas seriam classificadas em um dos três grupos de saúde (*Árvore Saudável*, *Árvore Médio Saudável* e *Árvore Enferma*).

Após a classificação das árvores urbanas, que pasará a ser uma nova variável categórica da base de dados, ir-se-á formalizar o modelo matemático classificatório, que se denominará a partir deste artigo de **MODIG-SAR** (“*Modelo Diagnóstico da Saúde da Árvore*”).

O **MODIG-SAR** será uma das contribuições sugeridas para auxiliar às tomadas de decisão relacionadas ao *check up* sistemático necessário de árvores urbanas.

A análise de dados se estrutura da seguinte forma: Aplicação da técnica da análise multivariada Two Step Cluster para criar a variável na base de dados que classifica as árvores nos três grupos de estado de saúde ((*Árvore Saudável*, *Árvore Médio Saudável* e *Árvore Enferma*) (CHIU et al, 2001); Aplicação da técnica Análise Discriminante Múltipla para criar um modelo que possa classificar uma dada árvore em uma das categorias registradas na variável Two Step Cluster da base de dados (*Árvore Saudável*, *Árvore Médio Saudável* e *Árvore Enferma*) (JOHNSON et al, 1999).

A *Two Step Clusters* é um instrumento analítico que privilegia as situações nas quais é necessário lidar conceitualmente e metodologicamente com objetos de configuração complexa (AAKER, 2001).

A *Two Step Clusters* consiste numa abordagem multifacetada e relacional sobre o objeto em estudo indicando traços de sua configuração complexa (ESTER et al, 1996).

A *Two Step Clusters* é particularmente interessante para gerar conglomerados em grandes bases de dados, onde se afiguram variáveis com diferentes níveis de mensuração (GUHA et AL, 1998).

A *Two Step Cluster*, uma metodologia da análise multivariada, constitui num conjunto de critérios para analisar dados multivariados em adversidade de medição, com o objetivo de se criar grupos homogêneos dentro e heterogêneo entre (KHATTREE & NAIK, 2000).

A *Two Step Cluster* trata múltiplas variáveis categóricas e contínuas analisadas conjuntamente e desenvolve análises estruturais (HUANG et al, 1998).

A *Two Step Clusters* é um método de análise quantitativa para variáveis mistas e tem por objetivo atingir a solução ótima na criação de *clusters* ou “*subperfis*” (LI et al, 2002).

A *Análise Discriminante Múltipla* é uma “ferramenta” estatística utilizada para classificar um determinado elemento (E) num determinado grupo de variáveis (HAIR, et. al, 2005), entre os diversos grupos existentes $\pi_1, \pi_2, \pi_3, \dots, \pi_i$. Para tal é necessário que o elemento (E), a ser classificado, pertença realmente a um dos i grupos, e que sejam conhecidas as características dos elementos dos diversos grupos. Essas características são especificadas a partir de n variáveis aleatórias ($X_1, X_2, X_3, \dots, X_n$) (HAIR, et al, 2005). No processo de classificação consideram-se os custos decorrentes de eventuais erros de classificação, bem como as probabilidades “a priori” de que o elemento pertença a cada um dos grupos (HAIR, et. al., 2005).

Com esse procedimento, espera-se achar no espaço de análise, um conjunto de estados de saúde de árvores urbanas que possa, por conseguinte, servir de insumo representativo de obtenção de uma ferramenta de diagnóstico de apatia e saúde de espaços naturais em nossa sociedade.

A *Two Step Clusters* e a *Análise Discriminante Múltipla* são as hipóteses preliminares de recursos de análises sugeridas para se atingir o objetivo deste estudo e que serão exploradas e prestigiadas no desenvolvimento das seções a seguir.

Os parágrafos que virão tratarão do desenvolvimento das etapas do método *Two Step Clusters* e em se-

guida do uso da *Análise Discriminante Múltipla* para a obtenção do **MODIG-SEMP**. As rodadas das respectivas técnicas foram realizadas no SPSS 15.0.

Quadro 1. Lista de Doze Indicadores de Enfermidade de Árvores

Variáveis	Descrição
X ₁	Vários ramos de menor diâmetro a saírem de um tronco?
X ₂	Há cavidades ao longo dos ramos?
X ₃	Há ramos em forquilha?
X ₄	Há troncos ou cavidades resultantes de cortes e que não cicatrizam?
X ₅	Há cavidades no tronco ou casca solta e morta?
X ₆	Há cavidades basais?
X ₇	Há copa morta ou folhagem pequena, de cor anormal?
X ₈	Há crescimento anormal de tecidos
X ₉	Há ramos com configurações diferentes das da árvore no estado natural?
X ₁₀	Há fungos?
X ₁₁	Há solo com fissuras?
X ₁₂	Há raízes cortadas?

Fonte: Método para avaliar a condição de árvores urbanas (1997).

RESULTADOS E DISCUSSÃO

Aplicação da *Two Step Cluster*

Pela observação da Quadro 2, a *two step cluster* segmentou a base de dados em três grupos ou *clusters* de árvores urbanas quanto à saúde.

Observado as Quadros de frequências dos perfis de cada *clusters*, indicadas pela *two step cluster*, podem-se traçar o perfil de cada grupo de árvores do espaço urbano.

O *cluster 1* ou o grupo 1 de árvores urbanas é formado pelas espécies com todas as presenças de itens de enfermidade, o que evidencia que são as “*Árvores Enfermas*”. Elas correspondem a 21,05% da amostra colhida.

O *cluster 2* ou o grupo 2 de árvores urbanas é formado pelas espécies com alguma presença de itens de enfermidade, o que evidencia que são as “*Árvores Médio Saudáveis*”. Elas correspondem a 42,63% da amostra colhida.

O *cluster 3* ou o grupo 3 de árvores urbanas é formado pelas espécies com nenhuma presença de itens de enfermidade, o que evidencia que são as “*Árvores Saudáveis*”. Elas correspondem a 36,32% da amostra colhida.

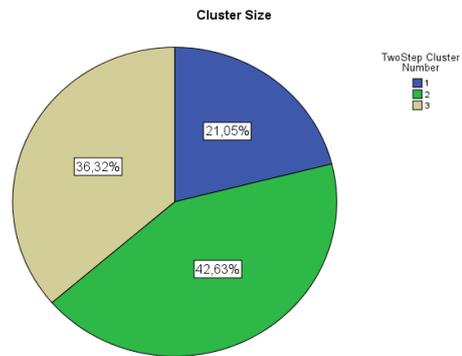
A *two step cluster* cria automaticamente na base de dados a variável que classifica cada árvore em um dos três grupos. Esta variável será renomeada para “*Diagnóstico*” e será usada pela análise discriminante para estimar o **MODIG-SAR**.

Quadro 2. Grupos Formados

Cluster Distribution			
	N	% of Combined	% of Total
Cluster 1	40	21.1%	21.1%
2	81	42.6%	42.6%
3	69	36.3%	36.3%
Combined	190	100.0%	100.0%
Total	190		100.0%

Fonte: Dados da Pesquisa (2015)

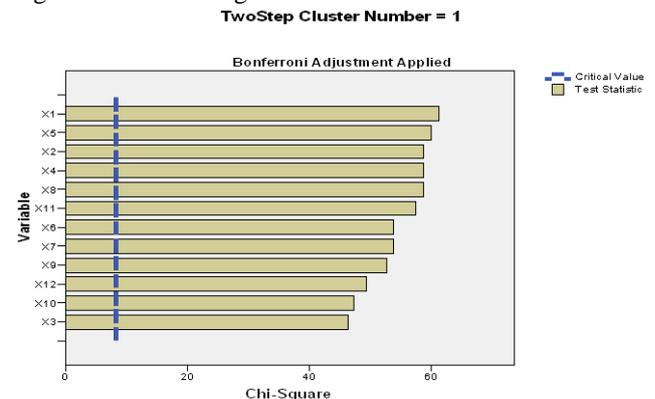
Figura 1. Tamanho dos Clusters Gerados



Fonte: Dados da Pesquisa (2015)

Após a análise de *clusters*, os dados foram submetidos à análise de significância estatística para avaliar quais as variáveis que têm importância na classificação dos grupos. Nas Figuras a seguir (Figuras 2, 3 e 4), estão ordenadas as variáveis de acordo com relevância que elas tiveram no processo de classificação dos grupos (ZHANG, 1996).

Figura 2. Teste de Significância do Cluster 1



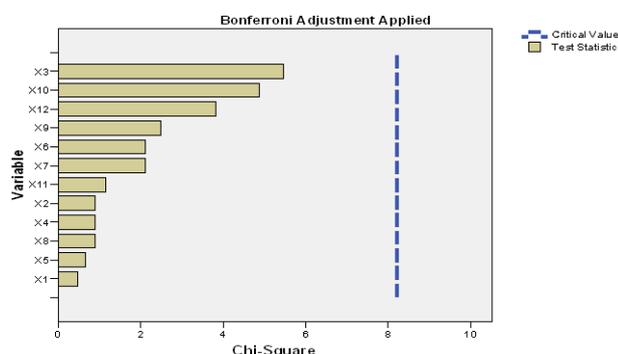
Fonte: Dados da Pesquisa (2015)

Essa ordem foi apontada através de resultados da análise de significância, não sendo, portanto, influenciada pelo autor de qualquer forma, resultando da distribuição das

respostas obtidas e do seu poder de classificação e distinção entre os grupos encontrados (HINNEBURG, 1998). A significância é apresentada em ordem decrescente.

Todas as variáveis da modelagem foram importantes para diferenciar o *cluster* 1 dos demais, com destaque para a variável X1. Portanto, o *cluster* 1 é estatisticamente diferente dos demais.

Figura 3. Teste de Significância do *Cluster* 2
TwoStep Cluster Number = 2

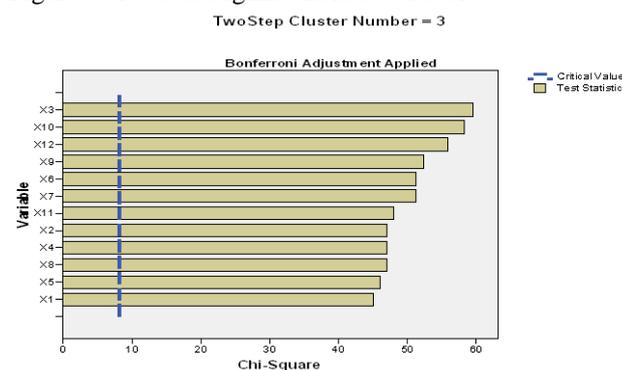


Fonte: Dados da Pesquisa (2015)

Nenhuma variável da análise foi suficientemente determinante para diferenciar o grupo 2 dos demais. Portanto, o *cluster* 2 não é estatisticamente diferente dos demais. Contudo, o estudo vai considerar a legitimidade da distinção do *cluster* formado.

Todas as variáveis da modelagem foram importantes para diferenciar o *cluster* 3 dos demais, com destaque para a variável X3. Portanto, o *cluster* 3 é estatisticamente diferente dos demais.

Figura 4. Teste de Significância do *Cluster* 3



Dados da Pesquisa (2015)

Aplicação da Análise Discriminante

O método de estimação é o método direto e com a variável dependente com 3 níveis e com 12 variáveis independentes, teremos duas funções discriminantes geradas.

A Quadro 3 mostra que nenhuma árvore avaliada foi excluída da análise. Todas as árvores observadas foram consideradas no estudo.

Quadro 3. Resumo dos Resultados

Analysis Case Processing Summary			N	Percent
Unweighted Cases				
Valid			190	100.0
Excluded	Missing or out-of-range group codes		0	.0
	At least one missing discriminating variable		0	.0
	Both missing or out-of-range group codes and at least one missing discriminating variable		0	.0
	Total		0	.0
	Total		190	100.0

Fonte: Dados da Pesquisa

Quadro 4. Teste de Existência de Multicolinearidade. Matriz de Correlações

Pooled Within-Groups Matrices												
Correlation	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12
X1	1.000	.001	.064	.001	-.078	.114	.114	.001	.042	-.117	.079	.078
X2	.001	1.000	.064	.005	.177	.013	.063	-.194	.040	-.074	.082	-.078
X3	.064	.064	1.000	.155	.034	-.115	-.065	-.047	-.196	.008	.075	.017
X4	.001	.005	.155	1.000	-.122	.063	.013	.005	.040	-.024	.032	.222
X5	-.078	.177	.034	-.122	1.000	-.011	.088	-.022	-.182	-.095	-.094	.150
X6	.114	.013	-.115	.063	-.011	1.000	.012	.113	.086	-.040	-.111	-.138
X7	.114	.063	-.065	.013	.088	.012	1.000	-.086	-.112	-.240	.236	.110
X8	.001	-.194	-.047	.005	-.022	.113	-.086	1.000	.040	.077	.032	.172
X9	.042	.040	-.196	.040	-.182	.086	-.112	.040	1.000	.032	-.035	.033
X10	-.117	-.074	.008	-.024	-.095	-.040	-.240	.077	.032	1.000	-.253	-.106
X11	.079	.082	.075	.032	-.094	-.111	.236	.032	-.035	-.253	1.000	.194
X12	.078	-.078	.017	.222	.150	-.138	.110	.172	.033	-.106	.194	1.000

Fonte: Dados da Pesquisa (2015)

Existem correlações apresentadas na matriz da Quadro 4 que estão dentro dos limites aceitáveis para a modelagem, o que significa que as variáveis explicativas são correlacionadas adequadamente e a multicolinearidade não deve apresentar problemas na análise.

Quadro 5. Poder de Explicação do Modelo Gerado. Quadro dos Autovalores

Eigenvalues				
Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	16.581 ^a	99.3	99.3	.971
2	.110 ^a	.7	100.0	.314

a. First 2 canonical discriminant functions were used in the analysis.

Fonte: Dados da Pesquisa (2015)

A correlação canônica da primeira função discriminante é predominantemente maior que a correlação canônica da segunda função discriminante. Neste primeiro momento, a função discriminante 1 tem vantagem interpretativa.

Pela Quadro 6, ao nível de 5% de significância, somente a função discriminante 1 é significativa estatisticamente. O λ de Wilks da primeira função é baixo e o da segunda função muito alto. Este resultado evidencia que a função discriminante 1 é a que melhor diferencia os grupos. Vamos, então, adotar a primeira função discriminante como modelo diagnóstico de saúde de árvores urbanas.

Quadro 6. Teste de Significância da Função Discriminante. λ de Wilks

Wilks' Lambda				
Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1 through 2	.051	539.188	24	.000
2	.901	18.863	11	.064

Fonte: Dados da Pesquisa (2015)

A Quadro 7 fornece os coeficientes de regressão canônicos das funções discriminantes. Estes coeficientes também são chamados de pesos brutos. Os pesos brutos dão uma primeira ideia, embora com cautela, das variáveis que foram mais importantes para a discriminação dos grupos (X_3 , X_5 , X_9 e X_{10}).

Quadro 7. Função Discriminante Estimada. Quadro dos Coeficientes (Pesos) Não Padronizados (Brutos)

Standardized Canonical Discriminant Function Coefficient

	Function	
	1	2
X1	.694	1.361
X2	.677	.529
X3	1.018	-1.815
X4	.636	1.486
X5	1.163	1.302
X6	.870	-.759
X7	.980	-.598
X8	.771	1.087
X9	1.153	-.386
X10	1.598	-.966
X11	.961	.635
X12	.359	-1.712
(Constant)	-4.628	.218

Unstandardized coefficients

Fonte: Dados da Pesquisa (2015)

A Quadro 8, transformando os pesos brutos em padronizados, evidencia, com mais segurança, as variáveis mais importantes para a discriminação dos grupos (X_3 , X_5 , X_9 e X_{10}).

Quadro 8. Função Discriminante Estimada. Quadro dos Coeficientes (Pesos) Padronizados. Avaliação da Importância das Variáveis Explicativas em Termos Relativos

Standardized Canonical Discriminant Function Coefficient

	Function	
	1	2
X1	.226	.444
X2	.222	.173
X3	.329	-.587
X4	.208	.487
X5	.380	.426
X6	.286	-.250
X7	.322	-.197
X8	.253	.356
X9	.379	-.127
X10	.519	-.314
X11	.316	.208
X12	.117	-.560

Fonte: Dados da Pesquisa (2015)

As cargas estruturais com "asteriscos" da função discriminante 1 indicam as variáveis mais diferenciadoras dos grupos.

A Função Classificação de Fisher permitirá classificar uma dada árvore urbana em: *Árvore Saudável*, *Árvore Médio Saudável* e *Árvore Enferma*.

Quadro 9. Quadro das Correlações Canônicas. Matriz Estruturante do Espaço de Análise

Structure Matrix

	Function	
	1	2
X10	.269*	.118
X3	.266*	.110
X5	.264*	.126
X4	.258*	-.101
X7	.253*	.076
X9	.250*	-.199
X8	.244*	.121
X13	.241*	-.113
X1	.231*	.081
X2	.215*	.145
X11	.248	.717*
X12	.249	-.386*
X14	.284	-.356*
X6	.264	-.328*

Pooled within-groups correlations between variables and standardized canonical discriminant functions. Variables ordered by absolute size of correlation with function 1.
*. Largest absolute correlation between any discriminant function

Fonte: Dados da Pesquisa (2015)

A função de designação gera um modelo de regressão para cada grupo, o que significa que entrando com 1 ou 0 da árvore nas 12 variáveis do modelo será gerado um escore discriminante. A árvore será diagnosticada no estado de saúde com maior escore discriminante.

Quadro 10. Função Classificação de Fisher. Designação aos Grupos

	Diagnóstico		
	Árvore Enferma	Árvore Médio Saudável	Árvore Saudável
X1	7.774	3.000	.000
X2	7.451	3.398	.000
X3	10.781	6.659	.000
X4	7.161	2.606	.000
X5	12.871	5.608	.000
X6	9.345	5.222	.000
X7	10.559	5.725	.000
X8	8.567	3.584	.000
X9	12.479	6.551	.000
X10	17.230	9.337	.000
X11	10.558	4.893	.000
X12	3.622	2.982	.000
(Constant)	-60.298	-16.310	-1.099

Fisher's linear discriminant functions

Fonte: Dados da Pesquisa (2015)

Gerando na planilha do *Excel* um procedimento de cálculo com os dados da *Função de Fisher*, tem-se, então, uma ferramenta automática para se diagnosticar a saúde de uma árvore urbana: **MODIG-SAR**.

Portanto, a uma árvore a que se tenha sido atribuído valor presença 1 a todas as variáveis, seria diagnosticada como muito enferma e constatado a evidência para o poder público de que está precisando de um “tratamento”.

Quadro 11- MODIG-SAR.

Variáveis	Valores	Árvore Enferma	Árvore Médio Saudável	Árvore Saudável
X ₁	1	7.774	3.000	0.000
X ₂	1	7.451	3.398	0.000
X ₃	1	10.781	6.659	0.000
X ₄	1	7.161	2.606	0.000
X ₅	1	12.871	5.608	0.000
X ₆	1	9.345	5.222	0.000
X ₇	1	10.559	5.725	0.000
X ₈	1	8.567	3.584	0.000
X ₉	1	12.479	6.551	0.000
X ₁₀	1	17.230	9.337	0.000
X ₁₁	1	10.558	4.893	0.000
X ₁₂	1	3.622	2.982	0.000
(Constant)		-60.298	-16.310	-1.099
Escores Discriminantes		54.48	40.27	-1.10

Fonte: Dados da Pesquisa (2015)

Portanto, a uma árvore a que se tenha sido atribuído valor presença 1 a todas as variáveis, seria diagnosticada como muito enferma e constatado a evidência para o poder público de que está precisando de um “tratamento”.

Como apresenta um diagnóstico de enfermidade, o poder público deve chamar imediatamente um “clínico geral”, ou seja, um “engenheiro florestal”. O engenheiro auxiliará a determinar as causas do mal e encontrar soluções que eliminem os seus efeitos.

A porcentagem do item (b) da nota de rodapé da Quadro 12 envolve todas as árvores da amostra classificados nos grupos pela função discriminante gerada e a porcentagem do item (c) somente as árvores da amostra de validação, selecionada aleatoriamente da amostra original.

A proporção de sucesso ou a porcentagem de casos classificados corretamente pela função discriminante estimada é de 98,9% para validação de grupo original e 95,3% para validação cruzada.

A proporção de sucessos em ambos os critérios de validação indicam que se erraria muito pouco ao utilizar o modelo construído no diagnóstico da saúde de árvores urbanas.

O que significa que o **MODIG-SAR** tem uma admirável capacidade preditiva do estado de saúde de árvores urbanas e pode ser umas das metodologias eficientes para avaliar a condição de florestas urbanas.

Quadro 12. Avaliação da Confiabilidade e Validade da Função Discriminante. Poder Preditivo da Função Discriminante. Designação aos Grupos

Classification Results^{b,c}

Diagnóstico			Predicted Group Membership			Total
			Árvore Enferma	Árvore Médio Saudável	Árvore Saudável	
Original	Count	Árvore Enferma	40	0	0	40
		Árvore Médio Saudável	1	79	1	81
		Árvore Saudável	0	0	69	69
	%	Árvore Enferma	100.0	.0	.0	100.0
		Árvore Médio Saudável	1.2	97.5	1.2	100.0
		Árvore Saudável	.0	.0	100.0	100.0
Cross-validated ^a	Count	Árvore Enferma	40	0	0	40
		Árvore Médio Saudável	7	72	2	81
		Árvore Saudável	0	0	69	69
	%	Árvore Enferma	100.0	.0	.0	100.0
		Árvore Médio Saudável	8.6	88.9	2.5	100.0
		Árvore Saudável	.0	.0	100.0	100.0

a. Cross validation is done only for those cases in the analysis. In cross validation, each case is classified by the functions derived from all cases other than that case.

b. 98.9% of original grouped cases correctly classified.

c. 95.3% of cross-validated grouped cases correctly classified.

Fonte: Dados da Pesquisa (2015)

CONCLUSÃO

O objetivo deste estudo foi estimar um modelo discriminante que pudesse classificar uma árvore em um dos três grupos: *Árvore Saudável*, *Árvore Médio Saudável* e *Árvore Enferma*.

A função estimada seria utilizada para o diagnóstico da saúde de árvores urbanas, que possam instrumentalizar as ações preventivas do poder público de desenvolvimento de políticas públicas e melhoria do espaço urbano. Portanto, este artigo visou criar uma ferramenta científica que possa auxiliar às tomadas de decisão relacionadas ao *check up* sistemático e necessário de árvores urbanas.

Para se obter o modelo diagnóstico da saúde da árvore, foi desenvolvida uma pesquisa de avaliação de 190 árvores do espaço urbano do estado do Rio de Janeiro. Um questionário que mede o nível de enfermidade, como valores 0 ou 1, baseado nos 12 indicadores de árvores enfermas, seria aplicado a uma amostra não probabilística de árvores urbanas.

Com a aplicação das técnicas *Two Step Cluster* e *Análise Discriminante*, foi formalizado o modelo matemático classificatório, o **MODIG-SAR** (*“Modelo Diagnóstico da Saúde da Árvore”*).

A função de designação gera um modelo de regressão para cada grupo, o que significa que entrando com valores 1 ou 0 nas 12 variáveis do modelo será gerado um escore discriminante. A devida árvore do espaço urbano será diagnosticada no estado de saúde com maior escore discriminante.

Gerando na planilha do *Excel* um procedimento de cálculo com os dados da *Função de Fisher*, tem-se, então,

uma ferramenta automática para se diagnosticar a saúde de uma árvore do espaço urbano.

Para uma árvore do espaço urbano que apresente um diagnóstico de enfermidade, o poder público deve chamar imediatamente um “clínico geral”, ou seja, um *“engenheiro florestal”*. O engenheiro auxiliará a determinar as causas do mal e encontrar soluções que eliminem os seus efeitos.

O **MODIG-SAR** é, então, uma das contribuições sugeridas para auxiliar às tomadas de decisão relacionadas ao *check up* sistemático necessário de árvores urbanas.

A proporção de sucessos em critérios de validação e confiabilidade indica que se erraria muito pouco ao utilizar o modelo construído no diagnóstico da saúde de árvores do meio urbano. O que significa que o **MODIG-SAR** tem uma admirável capacidade preditiva do estado de saúde de árvores.

O estudo recorreu-se a técnicas multivariadas para criar uma ferramenta para diagnóstico da saúde ou desenvolvimento de árvores. Com a consecução das análises, tem-se um *“achado”* de pesquisa e o objetivo do estudo alcançado.

Acredita-se que o mérito deste estudo foi procurar produzir conhecimento ou informação que pudesse cientificamente criar uma ferramenta de previsão do estado de saúde de árvores urbanas e assim contribuir no estudo de metodologias que permitem avaliar em que condições se encontram as florestas urbanas, buscando a elaboração de projetos administrativos mais eficientes.

A investigação é meramente preliminar, o ideal seria uma amostra com maior representatividade, mas abre caminhos para estudos mais generalizáveis sobre a temática estudada.

REFERÊNCIAS BIBLIOGRÁFICAS

- AAKER, DAVID; KUMAR, V.; DAY, GEORGE S. **Pesquisa de Marketing**. São Paulo: Editora Atlas, 2001.
- ARAÚJO, A. J. Método para avaliar a condição de árvores urbanas. In: **Encontro Nacional sobre Arborização Urbana**, 1997, Belo Horizonte – MG. Anais. Belo Horizonte, p.69,1997.
- ARAÚJO, A. J.; ARAÚJO, M. N. **Avaliação da Condição de Árvores Urbanas: Teoria e Prática**. In: VIII Semana de Estudos Florestais, 2006, Irati – Pr. Anais. Guarapuava:UNICENTRO, p.166-172, 2002.
- BIONDI, D; ALTHAUS, M. **Árvores Urbanas de Curitiba – Cultivo e Manejo**. 1. ed. Curitiba: FUPEF, 2005. 117p.
- BIONDI, D; REISSMANN, C. B. Avaliação do vigor das árvores urbanas através de parâmetros quantitativos. **Scientia Forestalis**. Piracicaba, n.52, p. 17-29, dez. 1997.
- CHIU, T., FANG, D., CHEN, J. AND WANG, Y. **A Robust and Scalable Clustering Algorithm for Mixed Type Attributes in Large Database Environment**, Proc. 2001Int. Conf. On knowledge Discovery and Data Mining. 2001 August. San Fransico (2001).
- COLTRO, M. E; MIRANDA, G. M. Levantamento da arborização urbana pública de Irati – PR e sua influência na qualidade de vida de seus habitantes. **Revista Eletrônica Latu Sensu**.v.2, n.1, jul.2007. Disponível em: <http://web03.unicentro.br/especializacao/revista/edicao2/engenharias/EN_LevArbo.pdf>. Acesso em: 02 set. 2008.
- ESTER, M.; KRIEGEL, H. P.; SANDER, J.; XU, X. **A Density-Based Algorithm for Discovering Clusters in Large spatial databases**. Proc. of the Second International Conference on Data Mining (KDD-96), 1996 August. Portland, Oregon (1996).
- GONÇALVES, W; et al. Análise de árvores urbanas para fins de supressão. **Revista da Sociedade Brasileira de Arborização Urbana**. Piracicaba, v.2, n.4, p. 1-19, dez. 2007.
- Guha, S.; Rastogi, R.; Shim, K. **Cure An Efficient Clustering Algorithm for Large Databases**, Proc. 1998ACM-SIGMOD Int. Conf. Management of Data. 1998june. Seattle, WA (1998).
- HAIR, F.F; ANDERSON, R. L; BLACK, W.C. **Análise Multivariada de Dados**. Porto Alegre: Bookman, 2005
- HINNEBURG, A.; KEIM, D. **An Efficient Approach to Clustering in Large Multimedia Databases with Noise** Proc. 1998 Int. Conf. on Data Mining and Knowledge Discovery (KDD'98). 1998 August. New York (1998).
- HUANG, Z. Extensions to the K-Means Algorithm for Clustering Large Data Sets with Categorical Values. **Data Mining and Knowledge Discovery**, Vol. 2, pp.283_304 (1998).
- JACK, K.; WEIL, R. **Métodos de Desenvolvimento Organizacional para Executivos Fordyce**. Fundo Educativo Interamericano S.A- Caracas-Venezuela- 2001.
- JOHNSON, R. A.; WICHERN, D. W. **Applied multivariate statistical analysis**. 4th ed. Upper Saddle River, New Jersey: Prentice-Hall, 1999, 815 p.
- KHATTREE, R.; NAIK, D. N. **Multivariate data reduction and discrimination with SAS software**. Cary, NC, USA: SAS Institute Inc., 2000. 558 p.
- LI, C.; BISWAS, G. Unsupervised Learning with Mixed Numeric and Nominal Data. **IEEE Transactions on Knowledge and Data Engineering**, Vol. 14, p. 4 (2002).
- LOBADA, R. C. et al. Avaliação das áreas verdes em espaços públicos no município de Guarapuava/PR. **Revista Electrónica de Geografía y Ciencias Sociales**. Barcelona, v.9, n.194, ago. 2005. Disponível em: <<http://www.ub.es/geocrit/sn/sn-194-71.htm>>. Acesso em: 02 set. 2005.
- MILANO, M. S. **Avaliação quali-quantitativa e manejo da arborização urbana de Maringá-PR**. 1988. 120f. Tese (Doutorado em Engenharia Florestal) - Universidade Federal do Paraná, Curitiba, PR.
- MILANO, M. S.; DALCIN, E. C. **Arborização de Vias Públicas**. Rio de Janeiro, RJ: Light,2000. 226p.
- PREFEITURA Municipal De Irati. **Clima**. Disponível em:<<http://www.irati.pr.gov.br/municipio/clima.asp>>. Acesso em: 03 set. 2008.
- PREFEITURA Municipal de Manaus. **Arborização Urbana**. Disponível em: <<http://www.manaus.am.gov.br/noticias/arborizacao-urbana-6>>. Acesso em: 22 set.2008a.
- PREFEITURA Municipal de Manaus. **Como podar e cortar árvores na cidade de Manaus**. Disponível em: <http://www2.manaus.am.gov.br/portal/secretarias/semma/formularios/cartilha_requerente.pdf>. Acesso em: 07 set. 2008b.
- SANTOS, E. **Avaliação Quali-quantitativa da arborização e comparação econômica entre a poda e a substituição da rede de distribuição de energia elétrica da Região Administrativa Centro-Sul de Belo Horizonte-MG**. Viçosa, MG: UFV, 2001. 219 p. Tese (Doutorado em Ciência Florestal) – Universidade Federal de Viçosa, 2001.
- SANTOS, N. R. Z; TEIXEIRA, I. F. **Arborização de Vias Públicas: Ambiente x Vegetação**. 1. ed. Santa Cruz do Sul: Instituto Souza Cruz, 2001.
- SOUZA, M. A. de I. B. **Curso sobre Arborização Urbana**. Belo Horizonte: CEMIG/SBAU/IEF/AMIFLOR, 1995. 33 p.
- ZHANG, T., RAMAKRISHMAN, R. AND LIVNY, M., **BIRCH. An Efficient Data Clustering Method for Very Large Databases**. Proc. 1996 ACM-SIGMOD Int. Conf. Management of Data, 1996 June. Montreal, Canada (1996).