



Identificação de padrões tecnológicos do sistema de pecuária de corte desenvolvido no Cerrado

Stanley R. de M. Oliveira¹, Urbano G P de Abreu², Maria do Carmo Ramos Fasiaben¹, Luis Gustavo Barioni¹, Helano Póvoas de Lima¹, Maxwell Merçon Tezolin Barros Almeida³, Octávio Costa de Oliveira³

¹ Embrapa Informática Agropecuária - CNPTIA, Campinas, São Paulo, Brasil
{stanley.oliveira, maria.fasiaben, luis.barioni, helano.lima}@embrapa.br

² Embrapa Pantanal - CPAP, Corumbá, Mato Grosso do Sul, Brasil
urbano.abreu@embrapa.br

³ Instituto Brasileiro de Geografia e Estatística - IBGE, Rio de Janeiro, Brasil
{maxwell.almeida, octavio.oliveira}@ibge.gov.br

RESUMO

O objetivo deste trabalho é inferir sobre os padrões de desenvolvimento tecnológico dos sistemas de produção de pecuária de corte no Cerrado. Para tanto, foram empregadas técnicas de aprendizagem de máquina e análise multivariada sobre dados do Censo Agropecuário 2006, em escala municipal. Os resultados foram considerados aderentes à realidade, permitindo distinguir diferentes tipos de exploração pecuária entre os municípios do Cerrado e seus padrões de distribuição geográfica.

PALAVRAS-CHAVE: Tipologias, Clusterização, Sistemas de produção.

ABSTRACT

The objective of this work is to infer about the technological development patterns of the production systems of beef cattle in the Cerrado. To do so, machine learning techniques and multivariate analysis were applied to the data from the 2006 Census of Agriculture, at a municipal scale. The results were considered adherent to the reality, allowing to distinguish different types of types of livestock farming among the municipalities of the Cerrado and their patterns of geographical distribution.

KEYWORDS: Typologies, Clustering, Production systems.

INTRODUÇÃO

O Censo Agropecuário brasileiro de 2006, realizado pelo Instituto Brasileiro de Geografia e Estatística (IBGE), contabilizou 62,7 milhões de bovinos no Cerrado (35,6% do rebanho nacional), o mais numeroso efetivo bovino dos biomas brasileiros. Dos 799.534 estabelecimentos agropecuários contabilizados no Cerrado, 104.288 informaram possuir mais de 50 cabeças de bovinos e ter o corte como a principal finalidade do rebanho.

O Cerrado Brasileiro, localizado na porção central do país, ocupa área contínua de 204,7 milhões de hectares, isto é, aproximadamente 24% da superfície do Brasil. Suas distintas condições ecológicas, seu diferenciado padrão espaço-temporal de ocupação e de investimentos em infraestrutura nas diferentes localidades do Bioma, bem como a conjuntura macroeconômica e os estímulos de mercado ao longo das últimas quatro décadas, determinaram um mosaico de características econômicas e sociais explicando os impactos sobre os sistemas de produção de pecuária de corte (MUELLER; MARTHA, Jr. 2008).

O número de bovinos no Cerrado aumentou em 91,7% entre 1975 e 1996, passando de 24,9 milhões para 47,8 milhões de cabeças. No período de 1990-2005, o rebanho bovino cresceu 34,5% (de 72,17 milhões para 97,07 milhões de cabeças) correspondendo a uma taxa média anual de crescimento de quase 2%. Esse crescimento, no entanto, foi mais rápido no período de 1996-2005 (2,5% ao ano), em comparação ao período de 1990 a 1996 (1,2% ao ano). Para comportar esse rápido crescimento do rebanho bovino no Cerrado, nas décadas citadas, houve aumentos significativos na área de pastagem cultivada, de maior capacidade de suporte do que as pastagens nativas (MARTHA Jr. et al., 2007).

Uma alternativa para identificação de padrões de desenvolvimento tecnológico no Cerrado é o uso de técnicas de aprendizado não-supervisionado, notadamente clusterização, também conhecida como análise de clusters ou análise de agrupamentos (LINDEN, 2009). Um dos métodos de clusterização mais utilizados é o de particionamento, no qual forma um conjunto inicial de k partições, sendo o parâmetro k o número de partições a serem consideradas para o conjunto de dados (RAJESH, 2011). O algoritmo de particionamento mais conhecido na literatura é o K-Means (McQUEEN, 1967).

A motivação pelo uso do algoritmo K-Means para identificação de padrões de desenvolvimento tecnológico se dá pelos seguintes aspectos: a) o algoritmo é simples e escalável; b) pode ser encontrado em diversos pacotes de software (WEKA, R, SPSS, SAS, MATLAB, entre outros); c) é sensível a ruídos, isto é, agrupa todos os objetos (no caso em estudo, os objetos representam municípios) disponíveis no conjunto de dados sob análise.

O objetivo deste artigo é o de, por meio de técnicas de aprendizagem de máquina e análise multivariada, inferir sobre os padrões de desenvolvimento tecnológico dos sistemas de produção de pecuária de corte no Cerrado em escala municipal, com a utilização de dados do Censo Agropecuário 2006.

MATERIAL E MÉTODOS

A região de estudo compreende o Cerrado brasileiro, incluindo os estados da região Centro-Oeste (Goiás, Mato Grosso, Mato Grosso do Sul e Distrito Federal), além do sul do Pará e Maranhão, interior do Tocantins, oeste da Bahia e Minas Gerais, bem como o norte de São Paulo. É constituído por várias características de vegetação, é classificado em subsistemas: de campo, de cerrado, de cerradão, de matas, de matas ciliares e de veredas e ambientes alagadiços. A vegetação do Cerrado encontra-se em uma região onde o clima que predomina é o tropical, apresenta duas estações bem definidas, sendo uma chuvosa, entre outubro e abril, e outra seca, entre maio e setembro.

Para atender ao presente trabalho, foi utilizada tabulação especial do Censo Agropecuário 2006, fornecida pelo IBGE (IBGE, 2011). A tabulação consistiu de dados de agregados por município de estabelecimentos agropecuários do Bioma Cerrado com mais de 50 cabeças de bovinos e que declararam ter o corte como principal finalidade do rebanho. Por questão de sigilo estatístico foram descartados os municípios com menos de três estabelecimentos.

Após o tratamento dos dados, foi realizado o processo essencial de extração dos padrões de desenvolvimento tecnológico dados para o bioma Cerrado. A técnica escolhida foi a análise de clusters (clusterização), por meio do método de particionamento. A similaridade entre objetos (municípios) foi medida pela distância Euclidiana entre dois objetos (HAN et al., 2011). Quanto menor for a distância entre dois objetos, maior será a similaridade entre eles. O algoritmo utilizado nas análises foi o K-Means (McQUEEN, 1967), que é um método de particionamento, em que o centro de cada cluster é representado pelo valor médio dos objetos no cluster. Possui como entrada o número de clusters (k) e o conjunto de dados (D) contendo n objetos, e como saída um conjunto de k cluster.

Para determinar o número de clusters, foram testados vários valores de k ($k = 2, \dots, 16$). Em seguida, para minimizar a subjetividade na escolha do valor de k , foi utilizado o método do cotovelo (HAN et al., 2011). Este método é simples e permite uma boa aproximação de k por meio do menor erro quadrático médio, definido pelo ponto de máxima curvatura para a trajetória da dissimilaridade interna e externa. A determinação do número

apropriado de clusters foi então definido no ponto onde o gráfico apresenta uma curva acentuada, à semelhança de um cotovelo. O algoritmo K-Means e o método do cotovelo foram implementados no ambiente RStudio (RStudio, 2017), que é um software livre de ambiente de desenvolvimento integrado para R, uma linguagem de programação para gráficos e cálculos estatísticos.

A discriminante canônica é uma técnica para auxiliar na classificação ou na pré-classificação de um elemento em determinado grupo, além de testar a hipótese de que os centróides do grupo de um conjunto de variáveis independentes para os outros grupos são iguais. O teste para a significância da função discriminante é uma medida generalizada da distância entre os centróides dos grupos (HAIR et al., 2009). O método foi inicialmente desenvolvido por Fisher (1936), com objetivo de maximizar a distância entre grupos, não levando em consideração o problema de minimização das probabilidades de erro no processo discriminatório. A função discriminante linear de Fisher (FDLF) consiste em separar objetos, sendo separados ou classificados com base nas medidas de p variáveis aleatórias.

O objetivo foi a criação dessa regra de reconhecimento de padrões e classificação. Assim as observações multivariadas X 's foram transformadas nas observações univariadas Y 's, tal que os Y 's dos agrupamentos π_1 e π_2 fossem separados em relação às médias dos dois grupos tanto quanto possível. A ideia de estimar combinações lineares de X para criar os Y 's, dado que essas combinações são funções de X . Seja $\mu_1 Y$ a média dos Y 's obtidos dos X 's pertencentes a π_1 , e $\mu_2 Y$ a média dos Y 's obtidos dos X 's pertencentes a π_2 e, considerando a matriz de covariância, Σ , como sendo a mesma para ambos grupos, selecionou-se a combinação linear que maximiza a razão entre o “quadrado da distância entre as médias” e a “variância de Y ”, ou seja: Essa razão é maximizada por: $C = \Sigma^{-1} (\mu_1 - \mu_2)$, assim: $Y = C'X = (\mu_1 - \mu_2)' \Sigma^{-1} X$, que é conhecida como FDLF (MARQUES; LIMA, 2002).

Os agrupamentos dos municípios formados pelo método de Wrapper foram classificados, e em seguida submetidos à análise de discriminante canônica com objetivo de verificar se a classificação dos municípios foi significativamente diferente. Seis variáveis foram utilizadas:

- 1) número de estabelecimento de cria/número de estabelecimento de gado de corte (NECNGC);
- 2) número de estabelecimento de cria e cria /número de estabelecimento de gado de corte (NECRNGC);
- 3) número de estabelecimentos que fizeram suplementação com ração mais estabelecimentos que utilizaram ração e sal mineral / número de estabelecimento de gado de corte (NERNGC);
- 4) número de estabelecimento que receberam orientação técnica/ número de estabelecimento de gado de corte (NEONGC);
- 5) somatório das áreas de

lavoura permanente mais áreas com lavoura temporárias / somatório das áreas de lavoura permanente; áreas com lavoura temporárias; áreas com forrageiras para corte; área com viveiro de mudas; áreas de pastagens; áreas de silvicultura e áreas com sistemas agroflorestais (SATSALP); e 6) valor da produção dos bovinos / valor total da produção (VPBVT), em reais.

RESULTADOS E DISCUSSÃO

Para determinar o número de clusters, foram testados vários valores de k ($k = 2, \dots, 16$). A escolha do número apropriado de clusters foi definido pelo método do cotovelo, indicando $k = 6$ para o banco de dados de desenvolvimento tecnológico da pecuária de corte para o Cerrado brasileiro (Figura 1).

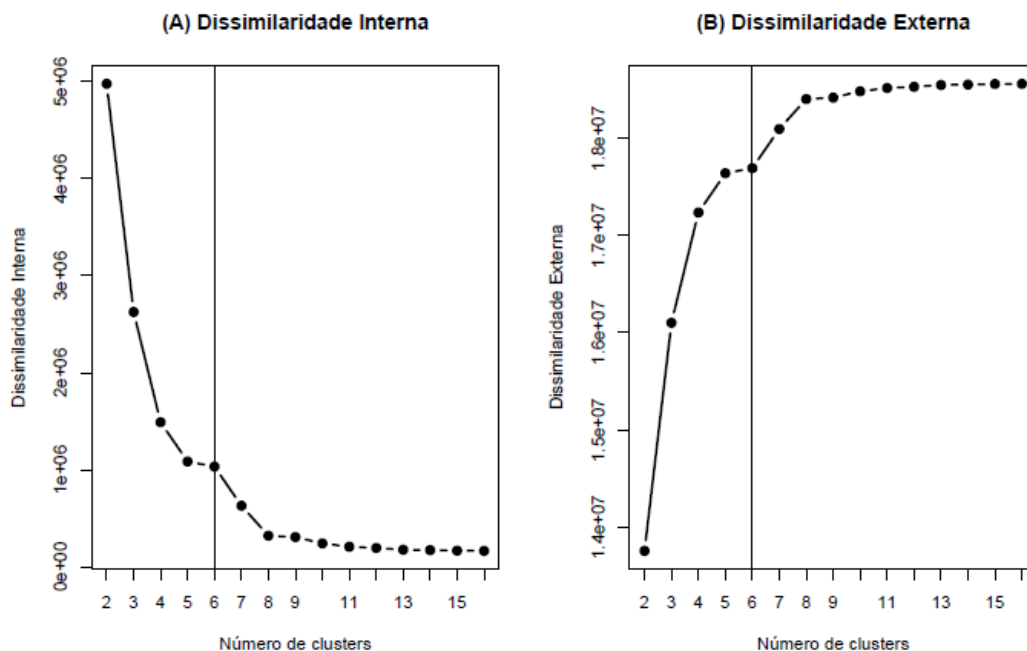


Figura 1 – Dissimilaridade interna e externa em função do número de clusters para o banco de dados de desenvolvimento tecnológico da pecuária de corte para o Cerrado brasileiro.

Na Tabela 1 são mostradas as médias dos percentuais das variáveis de cada cluster, em escala municipal. Observamos que os sistemas com predominância de sistemas de cria (NECNGC) são predominantes nos Clusters 0 e 1, sendo interessante notar que a diferença entre eles é o maior percentual de sistemas de cria e recria (NECRNGC) no Cluster 1. Por outro lado, os Clusters 1 e 2 apresentaram maior valor da produção de bovinos em relação ao valor de produção total (VPBVT), e em paralelo maior presença de área com pastagens (SAPSAT). O Cluster 3 também apresentou o mesmo perfil, entretanto com percentual menor nas duas variáveis. A utilização de suplementos (NERNGC) e orientação técnica (NEONGC),

percentualmente foram mais presentes nos grupos 4 e 5, provavelmente em função da necessidade da orientação técnica para fornecimento adequado de suplementos.

Após a classificação dos municípios os grupos foram submetidos à técnica multivariada de análise discriminante canônica (FDLF). Na Figura 2, são apresentadas as duas primeiras variáveis canônicas dos seis grupos. Todos os grupos apresentaram a distância de Mahalanobis entre suas médias canônicas significativamente diferentes ($p < 0,01$), o que nos permite inferir que os grupos formados possuem diferentes padrões tecnológicos no desenvolvimento da atividade de pecuária de corte.

Tabela 1 – Média e erro padrão da média, das variáveis selecionadas para a ‘clusterização’ dos municípios do Cerrado, e número de municípios em cada cluster observado.

	Cluster 0	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
	143 municípios	241 municípios	251 municípios	139 municípios	88 municípios	194 municípios
NECNGC (%)	0,2857± 0,0140	0,3442±0,0113	0,1334 ± 0,0053	0,1917±0,0102	0,1288±0,0131	0,1296± 0,0066
NECRNGC (%)	0,1604± 0,0105	0,2601± 0,0100	0,1720 ± 0,0063	0,0646±0,0049	0,1259±0,01091	0,1478±0,0079
NERNGC (%)	0,2600± 0,0138	0,2017±0,0097	0,4141± 0,0110	0,1973±0,013	0,5347±0,0261	0,4359±0,0115
NEONGC (%)	0,2603±0,0116	0,2386±0,0095	0,5119± 0,0106	0,2232±0,0119	0,7491±0,0209	0,6482± 0,0110
SAPSAT (%)	0,7529±0,0138	0,9007±0,0069	0,9071±0,0065	0,7991±0,0148	0,3406±0,0161	0,7547±0,0084
VPBVT (%)	0,2728±0,0134	0,8009±0,0091	0,8136±0,0085	0,7839±0,0129	0,1731±0,0191	0,2859±0,0117

Fonte: Dados do Censo Agropecuário 2006 trabalhados pelos autores.

Nota: NECNGC- número de estabelecimento de cria/número de estabelecimento de gado de corte, NECRNGC- número de estabelecimento de cria e recria /número de estabelecimento de gado de corte, NERNGC-número de estabelecimentos que fizeram suplementação com ração mais estabelecimentos que utilizaram ração e sal mineral / número de estabelecimento de gado de corte, NEONGC-número de estabelecimento que receberam orientação técnica/ número de estabelecimento de gado de corte, SAPSAT- somatório das áreas de pastagem natural mais áreas com pastagens degradadas e não degradadas / somatório das áreas com lavoura permanente e não permanente, áreas com forrageiras para corte, cultivo de flores, casa de vegetação e viveiro de mudas, áreas com silvicultura, sistemas florestais, pastagem natural mais áreas com pastagens degradadas e não degradadas, VPBVT-valor da produção dos bovinos / valor total da produção em reais.

Após a classificação dos municípios os grupos foram submetidos a técnica multivariada de análise discriminante canônica (FDLF). Na Figura 2, são apresentadas as duas primeiras variáveis canônicas dos seis grupos. Todos os grupos apresentaram a distância de Mahalanobis entre suas médias canônicas significativamente diferentes ($p < 0,01$), o que nos permite inferir que os grupos formados possuem diferentes padrões tecnológicos no desenvolvimento da atividade de pecuária de corte.

Apresenta-se a seguir a interpretação e descrição dos diferentes clusters e a Figura 3 mostra a sua localização geográfica.

Cluster 0 - “Municípios com predomínio da fase de cria, produção extensiva”

Abrange 14% dos municípios (143 municípios) e 13% dos estabelecimentos com bovinos de corte no Bioma Cerrado, respondendo por 10% do total de bovinos e 12% da área total de pastagens. Nestes municípios, o rebanho médio é de 332 cabeças, a área média de pastagens é de 381 ha e a área média dos estabelecimentos que têm bovinos é de 1537 ha. Da área explorada pelos estabelecimentos desse cluster, 75% são destinados a pastagens e 11%, a lavouras. No Cluster 0, em 26% dos estabelecimentos se declarou usar ração e em 7% adubar as pastagens. O percentual de animais confinados não atingiu 2%. Somente 26% dos

estabelecimentos produtores de bovinos de corte dos municípios desse cluster declararam receber orientação técnica. Embora a maior parte da área explorada seja destinada a pastagens, nesse cluster apenas 27% do valor total da produção provêm da bovinocultura. A taxa de lotação é de 0,69 UA/hectare, a menor entre todos os clusters. Os municípios do Cluster 0 se encontram especialmente numa faixa que abrange o Sul do Maranhão, Sudoeste do Piauí, Oeste da Bahia e Centro-Norte de Minas Gerais.

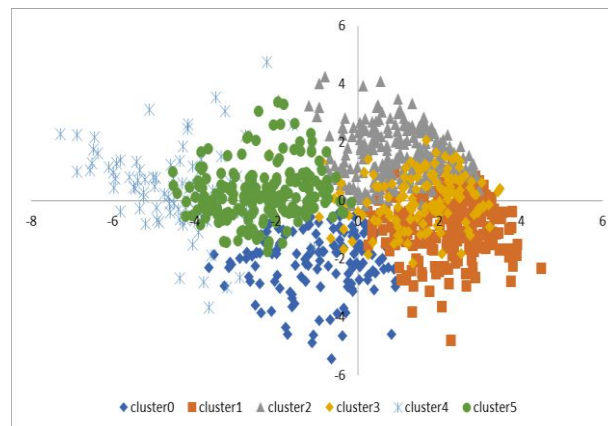


Figura 2 – Variáveis canônicas discriminadas nos seis grupos de municípios do Cerrado, conforme perfil tecnológico do sistema de produção de gado de corte.

Cluster 1 - “Municípios com importância das fases de cria e recria”

Representa 23% dos municípios (241 municípios), 29% dos estabelecimentos, 22% do total de bovinos e 25% do total de pastagens dos municípios do Bioma Cerrado (produtores de bovinos de corte). Nestes municípios, o tamanho médio do rebanho é de 338 cabeças, a área média de pastagens é de 374 ha e a área total média dos estabelecimentos, 934 ha. Aqui, 90% da área explorada pelos estabelecimentos com bovinos de corte são destinadas a pastagens. No Cluster 1, 20% dos estabelecimentos declararam usar ração e 6% adubar as pastagens. O percentual de animais confinados foi de 1,6% (o menor entre todos os clusters). Nos municípios que compõem o Cluster 1, 24% dos estabelecimentos declararam receber orientação técnica. A taxa de lotação é de 0,75 UA/hectare. Nesse cluster, 80% do valor da produção correspondem à bovinocultura de corte. Os municípios que fazem parte do Cluster 1 estão localizados predominantemente nos estados de Goiás, Tocantins e Maranhão e leste do Mato Grosso.

Cluster 2 - “Municípios com importância da bovinocultura de corte de ciclo completo”

Neste cluster enquadram-se 24% dos municípios do Bioma Cerrado produtores de bovinos de corte (251 municípios), 32% dos estabelecimentos, 49% do total de bovinos e 45% do total de pastagens. A área média dos estabelecimentos com bovinos dos municípios que

compõem o Cluster 2 é de 1093 ha, com uma média de 600 ha de pastagens. O número médio de bovinos por estabelecimento é de 668 cabeças. Da área explorada com agropecuária pelos estabelecimentos, 91% são destinados a pastagens. Consta-se, neste cluster, uma maior importância da produção de bovinos de corte a ciclo completo (cria, recria e engorda). Nos municípios que compõem esse cluster, 41% dos estabelecimentos declararam usar ração e perto de 15% fazer adubação de pastagens. A porcentagem de animais confinados chegou a 7%. Pouco mais da metade dos estabelecimentos produtores de bovinos de corte dos municípios desse cluster declararam receber assistência técnica (51%). A maior parte do valor da produção deste grupo provém da produção de bovinos (81%). O cluster apresenta uma taxa de lotação média de 0,98 UA/hectare. Os municípios que compõem o Cluster 2 distribuem-se especialmente nos estados do Mato Grosso do Sul, Goiás e interior de Minas Gerais.

Cluster 3 - “Municípios com importância da produção a ciclo completo e da fase de engorda a pasto”

Representa 13% dos municípios (139 municípios), 8% dos estabelecimentos, 5% do total de bovinos e 5% do total de pastagens. A área média dos estabelecimentos que têm bovinos é de 1028 ha, a área média de pastagens/estabelecimento é de 253 ha e o número médio de bovinos por estabelecimento é de 264 cabeças. Cerca de 80% da área explorada pelos estabelecimentos são destinados a pastagens. Referindo-se à tecnologia empregada na bovinocultura, 20% dos estabelecimentos dos municípios que fazem parte do Cluster 3 declararam usar ração, e 4% informaram adubar as pastagens. O percentual de animais confinados foi de pouco mais de 3%. Apenas 22% dos estabelecimentos produtores de bovinos de corte dos municípios declararam receber orientação técnica. Neste cluster, 78% do valor da produção provém da bovinocultura de corte. A taxa de lotação é de 0,86 cabeças/hectare. Os municípios que pertencem ao Cluster 3 concentram-se na porção mais ao norte do Cerrado brasileiro, em especial no estado do Maranhão.

Cluster 4 - “Municípios com diversificação lavoura/pecuária e maior presença de confinamento”

O Cluster 4 é composto por 88 municípios (8%), engloba 3% dos estabelecimentos e abriga 2% do total de bovinos e 2% do total de pastagens dos seis clusters de bovinos de corte do Cerrado. Nesse cluster, a área média dos estabelecimentos é de 2789 ha, a área média de pastagens/estabelecimento é de 258 ha e o número médio de cabeças por estabelecimento é de 351. Nesse cluster, 61% da área explorada se destinam a lavouras, enquanto 34% se referem a pastagens. Aqui, 53% dos estabelecimentos declararam usar ração e 26% fazer adubação de pastagens. O percentual de animais confinados foi de 18%, e em 75% dos estabelecimentos se

declarou a presença de orientação técnica. A maior parte do valor da produção deve provir da produção vegetal, já que os bovinos representam um percentual de apenas 17% desse valor total. A taxa de lotação é a mais alta entre todos os clusters: 1,73 UA/hectare. Os municípios que integram o Cluster 4 se encontram em maior número no Sul de Minas, Norte de São Paulo e região Central de Mato Grosso.

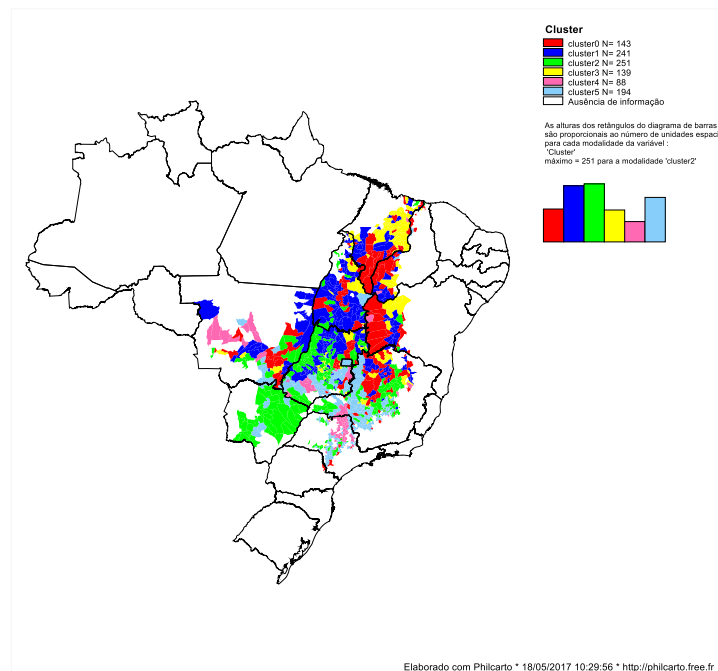


Figura 3- Localização geográfica dos clusters de municípios produtores de bovinos de corte no Cerrado em 2006.

Cluster 5 - “Municípios com importância da bovinocultura de corte de ciclo completo e de lavouras”

Abrange 18% dos municípios (194 municípios) e detém 16% dos estabelecimentos, que respondem por 13% do total de bovinos e 11% do total da área de pastagens. Nos municípios desse cluster, a área média dos estabelecimentos é de 1093 ha e a área média de pastagem/estabelecimento é de 311 ha. O tamanho médio do rebanho/estabelecimento é de 368 cabeças. Da área explorada pelos estabelecimentos desse cluster, 75% são destinados a pastagens e 19%, a lavouras. No Cluster 5, em 44% dos estabelecimentos se fez uso de ração e em 19% se adubou as pastagens. O percentual de animais confinados foi de 7%. Em 65% dos estabelecimentos nos municípios desse cluster se declarou receber orientação técnica. Nesse cluster, 29% do valor total da produção provém da bovinocultura. A taxa de lotação é de 1,06 UA/hectare, atrás apenas do Cluster 4. Os municípios do Cluster 5 aparecem especialmente no Centro-Sul de Goiás, DF, Centro-Leste de Minas Gerais e região Centro-Sul de São Paulo.

CONCLUSÕES

Os resultados demonstram que os dados obtidos pelo Censo Agropecuário 2006 são suficientes para distinguir tipos de exploração pecuária entre os municípios do Cerrado. Os agrupamentos possuem padrão geográfico definido e dentro do esperado. Na região mais ao sul do Cerrado (Sul de Minas, Mato Grosso do Sul, Sul de Minas e Sul de Goiás) predominam os clusters 2, 4 e 5, enquanto que os clusters 0, 1 e 3 predominam nos estados mais ao norte do bioma. Os biomas mais ao sul estão associados a sistemas mais tecnificados, envolvendo integração com lavouras, ciclo completo e confinamento. Os municípios mais ao norte estão associados à produção mais extensiva e às fases de cria e recria, embora haja a presença de ciclo completo e engorda, particularmente no Maranhão. O Mato Grosso foi o estado menos padronizado em termos de tipos de sistema produtivo, apresentando todos os tipos identificados.

REFERÊNCIAS

- FISHER R. A. The use of multiple measurements in taxonomic problems. **Annals of eugenics**, v. 7, p. 179-188, 1936.
- HAIR Jr; BLACK, W.C.; BABIN, B. J.; ANDERSON, R. E.; TATHAM, R. L. **Análise multivariada de dados**. 6. ed. Porto Alegre: Artmed, 2009. 688p.
- HAN, J.; KAMBER, M.; PEI, J. **Data mining: concepts and techniques**. 3rd edition. San Francisco: Morgan Kaufmann Publishers, p.770, 2011.
- IBGE. **Tabulações especiais do Censo Agropecuário 2006**. Rio de Janeiro: Instituto Brasileiro de Geografia e Estatística, 2011.
- LINDEN, R. Técnicas de Agrupamento. **Revista de Sistema de Informação da FSMA**, n. 4, 2009.
- MARQUES, J. M.; LIMA, J. D. de. A estatística multivariada na análise econômico-financeira de empresas. **Revista da FAE**, v. 5, p. 51-59, 2002.
- MARTHA JR., G. B., VILELA, L., BARCELLOS, A. O., SOUZA, D. M. G., BARIONI, L. G. Pecuária de corte no Cerrado: aspectos históricos e conjunturais. In: **CERRADO: USO EFICIENTE DE CORRETIVOS E FERTILIZANTES EM PASTAGENS**, 2007, Brasília, **Anais...** Brasília:Embrapa Cerrados, 2007. p. 17-42.
- McQUEEN, J. B. Some Methods for Classification and Analysis of Multivariate Observations. In: **BERKELEY SYMPOSIUM ON MATHEMATICAL STATISTIC AND PROBABILITY**, 5., 1967, Berkeley. Proceedings, p. 281-297.
- MUELLER, C. C.; MARTHA Jr, G. B. A agropecuária e o desenvolvimento sócio-econômico recente do Cerrado. In: **SIMPÓSIO NACIONAL CERRADO**, 9.; **SIMPÓSIO INTERNACIONAL SAVANAS TROPICAIS**, 2., 2008, Brasília. **Anais...** Brasília: Embrapa Cerrados, 2008. p. 1-41.
- RAJESH, D. Application of Spatial Data mining for Agriculture. **International Journal of Computer Applications**, v. 15, n. 2, 2011.
- RStudio – Open source and enterprise-ready professional software for R. Disponível em: <https://www.rstudio.com>. Consultado em 3 de maio de 2017.