

Reconhecimento de Padrões de Colônias de Abelhas *Apis Mellifera* Segundo Mudanças das Estações do Ano

Felipe Anderson O. Maciel¹, Antonio Rafael Braga^{1,3}, Alisson de Lima e Silva¹,
Ticiania L. Coelho da Silva³, Breno M. Freitas², Danielo G. Gomes¹

¹Grupo de Redes de Computadores, Engenharia de Software e Sistemas (GREat)
Universidade Federal do Ceará (UFC) - Campus do Pici - Bloco 942-A

²Setor de Abelhas - Departamento de Zootecnia - UFC - Campus do Pici

³UFC - Campus Quixadá

{felipemaciel, antoniobraga, alissonsilva}@great.ufc.br

{ticianiana, freitas, danielo}@ufc.br

Abstract. *As the principal pollinating agent, bees are essential to the production of food for mankind and to the maintenance of the ecosystem. Among the crops used for human feeding, 75% rely on pollination. Aligning to a current concern with bees survival, this paper aims find out states of *Apis mellifera* colonies in order to assist the beekeeper in the management and maintenance of its hives. Our methodology consisted in the application of a clustering technique in two real datasets of hives with data of temperature, humidity and weight. From the application of the Calinski-Harabasz index and the K-means algorithm, we have identified coherent patterns associated with the transitions between the seasons. In addition, we conclude that the strongest colony is most efficient in trying to maintain the microclimate of the hive during the winter.*

Resumo. *Na qualidade de principal agente polinizador, as abelhas são essenciais à produção de alimentos para o ser humano e para manutenção dos ecossistemas. Entre as culturas agrícolas utilizadas para o consumo humano, 75% dependem de polinização. Alinhando-se a uma preocupação atual com a sobrevivência das abelhas, este artigo visa descobrir estados de colônias de *Apis mellifera* a fim de auxiliar o apicultor no manejo e na manutenção de suas colmeias. Nossa metodologia consistiu na aplicação de uma técnica de clusterização em dois datasets reais de colmeias com dados de temperatura, umidade e massa. A partir da aplicação do índice Calinski-Harabasz e do algoritmo K-means, identificamos padrões coerentes e associados às transições entre as estações do ano. Além disso, concluímos que a colônia mais forte é mais eficiente ao tentar manter o microclima da colmeia durante o inverno.*

1. Introdução

As abelhas são consideradas os principais agentes polinizadores da maioria das espécies de plantas selvagens e culturas polinizadas por insetos [Ollerton et al. 2011]. Atualmente, cerca de 75% das culturas vegetais dependem da polinização [Potts et al. 2016], o que representa um crescimento de 300% nos últimos 50-60 anos [Aizen and Harder 2009].

Diferentes estados das colônias podem ser observados durante o ciclo de vida das abelhas. Certos estados podem impactar negativamente no desenvolvimento da colônia (e.g. abandono), resultando em uma possível perda da colônia e aumento de custos para o apicultor. Outros casos, como o estado ativo de produção de cria, indicam que a colônia é capaz de se reproduzir e polinizar [Kviesis and Zacepins 2016].

A fim de determinar esses estados, o monitoramento remoto de apiários pode auxiliar apicultores agregando valiosas informações sobre o estado e o comportamento das abelhas, além de poupar as abelhas de estresses desnecessários ou outras atividades não-produtivas [Zacepins and Karasha 2012]. A literatura recente mostra a aplicação das redes de sensores sem fio, dos sistemas ciberfísicos e da Internet das Coisas (IoT) no monitoramento de colônias de abelhas [Sánchez et al. 2015, Murphy et al. 2016, Zacepins et al. 2017] para o desenvolvimento da chamada *apicultura de precisão*, cujo objetivo central é maximizar a produtividade das colmeias e diminuir o consumo de recursos [Zacepins et al. 2015].

Neste contexto, este trabalho busca reconhecer padrões comportamentais das abelhas da espécie *Apis Mellifera* para responder a perguntas como: “Quais os valores típicos de temperatura, umidade e massa de uma colmeia ao longo do ano?” e “Quais os padrões climáticos internos de uma colmeia durante e entre as estações de um ano?”. Assim, a contribuição central deste artigo é aplicar a técnica de mineração de dados clusterização para descobrir e caracterizar padrões dos estados de colônias de abelhas.

2. Trabalhos relacionados

As principais grandezas físicas sensoriadas pelos sistemas de monitoramento de apiários tem sido a temperatura, umidade, massas das colmeias, gases internos, áudio, vídeo e vibrações [Meikle and Holst 2015].

Através da temperatura interna da colmeia, por exemplo, é possível traçar padrões térmicos para identificar cenários de pré-abandono das abelhas [Kridi et al. 2014, Kridi et al. 2016]. A temperatura ambiente (externa) também pode indicar alteração do comportamento de uma colônia de *Apis mellifera*. Em condições de clima temperado, por exemplo, quanto mais quente mais abelhas realizam o forrageamento, o que impacta no desenvolvimento das pupas e na divisão de trabalho das abelhas [Becher 2010].

Outra importante variável abordada em trabalhos de monitoramento na apicultura de precisão é a umidade [Kviesis et al. 2015]. A umidade está fortemente associada ao resfriamento por evaporação, um recurso fundamental para as abelhas controlarem a hipertermia da colmeia [Ostwald et al. 2016]. Juntamente com outras variáveis, como temperatura e concentração de gases (e.g. CO₂, O₂), a umidade pode ser utilizada para classificar os estados de uma colônia [Murphy et al. 2016].

A massa das colmeias (caixa + colônia) é outra métrica importante, podendo ser utilizada para estimar tanto a quantidade de mel produzido [Ruan et al. 2017] como também a evaporação de néctar e água ou a atividade de forrageamento das abelhas [Meikle and Holst 2015].

Apesar do monitoramento remoto de colônias via redes de sensores ser uma estratégia bem difundida no âmbito da apicultura de precisão, notamos uma certa carência de soluções automatizadas para extração do conhecimento dos dados das colmeias

[Zacepins et al. 2015]. Nesse sentido, este trabalho apresenta um estudo que detectou padrões comportamentais de colônias de *Apis mellifera* à partir das grandezas físicas como massa, temperatura e umidade interna da colmeia.

3. Material e Métodos

Esta seção descreve os aspectos metodológicos da pesquisa realizada em relação ao ferramental utilizado, coleta e pré-processamento dos dados, estratégias de aprendizado, bem como a análise e detecção dos estados da colônia de abelhas sob estudo.

3.1. Bases de Dados

Neste artigo utilizamos dados do portal HiveTool.net¹. Hivetool é um projeto de código aberto cujo o objetivo é produzir ferramentas de software e hardware para monitoramento e gerenciamento de colmeias de abelhas. Os dados referem-se à massa, umidade e temperatura interna de uma colmeia de *Apis mellifera*, além dos dados de temperatura e umidade ambiente. No entanto, para a análise por agrupamento, foram considerados apenas os dados internos e a data de coleta de cada amostra.

Foram analisados os conjuntos de dados de duas colmeias. O primeiro *dataset*, com 92.749 amostras, é de uma colmeia do apiário Arnas (56° 53' 17.8656" N, 9° 50' 39.5340" L), localizado na cidade de Rebild, na Dinamarca. O segundo possui 88.181 amostras de uma colmeia do apiário Emil (58° 19' 9.1596" N, 8° 32' 9.8124" L), localizado na cidade de Grimstad, na Noruega. O período de coleta foi de 01/03/2017 a 28/02/2018.

3.2. Pré-Processamento dos Dados

O pré-processamento dos dados compreendeu basicamente duas etapas: remoção de anomalias e normalização dos dados.

A remoção de anomalias foi realizada pelo Método de Tukey [Tukey 1977]. Neste método é definido um intervalo com limites inferior e superior dos atributos do conjunto de dados, e as amostras que estiverem fora desse intervalo são consideradas anomalias. Os limites são calculados por (1), em que Q_1 e Q_3 são respectivamente o primeiro e o terceiro quartis de um atributo do conjunto de dados.

$$[Q_1 - 1.5 \times (Q_3 - Q_1), Q_3 + 1.5 \times (Q_3 - Q_1)]. \quad (1)$$

Com os dados “limpos”, partiu-se para a etapa de normalização, que se faz necessária para que os dados estejam na mesma escala nas etapas posteriores da análise. Cada atributo X do *dataset* teve seu valor x normalizado por (2), em que μ e σ são respectivamente a média e o desvio padrão de X .

$$Z = \frac{x - \mu}{\sigma}. \quad (2)$$

¹<http://www.hivetool.net/>

3.3. Estratégias de Aprendizado

A metodologia proposta tem como objetivo caracterizar padrões de uma colônia de abelhas com o propósito de identificar possíveis estados. A caracterização foi realizada através da mineração de dados, por meio de um algoritmo de clusterização (aprendizagem não-supervisionada).

Técnicas de clusterização são responsáveis por agrupar um conjunto de objetos em classes de objetos que são similares entre si. O algoritmo de clusterização utilizado neste trabalho foi o *K-means*.

3.3.1. Clusterização

A clusterização é dita uma tarefa de aprendizagem não-supervisionada pois não se tem informação prévia sobre as classes às quais os dados pertencem. Assim, utiliza-se apenas as informações extraídas dos próprios dados, tentando agrupá-los por similaridade.

O algoritmo de clusterização utilizado neste trabalho foi o *K-means* [MacQueen 1967]. O objetivo do *K-means* é dividir o conjunto de N vetores de dados em K partições (ou *clusters*) não-superpostas ($K \ll N$), com o auxílio de K protótipos, também chamados de centroides, devidamente posicionados no espaço dos dados. Então cada vetor de dados é associado a um centroide por critério de similaridade, por exemplo, o de menor distância.

O conjunto W de K centroides é representado por (3), em que p é a dimensão do vetor de dados.

$$W = \{w_i\}_{i=1}^K \mid w_i \in \mathbb{R}^p. \quad (3)$$

O *cluster* V associado a cada centroide w é definido por (4), em que x é um vetor de atributos, e $\|x - w\|$ indica a distância euclidiana.

$$V_i = \{x \in \mathbb{R}^p \mid \|x - w_i\| < \|x - w_j\|, \forall j \neq i\}. \quad (4)$$

De maneira geral, o *K-means* se comporta da seguinte forma: inicialmente são selecionados K vetores aleatórios do espaço de dados como protótipos; o *cluster* de cada protótipo w_i é determinado por (4); então calcula-se a nova posição de cada protótipo w_i como a média dos N_i objetos do *cluster* V_i (5).

$$w_i = \frac{1}{N_i} \sum_{x \in V_i} x. \quad (5)$$

O *cluster* V_i (4) e a nova posição do protótipo w_i (5) são repetidamente recalculados até a convergência do algoritmo. Isso acontece quando a posição do protótipo w_i não muda mais ou quando atinge-se um número máximo de iterações, por exemplo. Para avaliar quantitativamente o posicionamento dos protótipos, calcula-se a soma das distâncias quadráticas (*sum of squared distance* - SSD) de um vetor de dados ao centroide mais próximo (6). Este erro ajuda a avaliar a qualidade dos clusters gerados.

$$SSD = \sum_{\forall x \in V_i} \|x - w_i\|^2. \quad (6)$$

3.3.2. Validação da clusterização

Em um problema de clusterização é comum não saber a priori o valor de K que melhor se ajusta ao conjunto de dados, e o algoritmo *K-means* requer esse valor previamente para ser executado. Uma maneira de encontrar o valor ideal de K é por meio dos índices de validação de clusterização. Normalmente esses índices avaliam dois aspectos da clusterização: (i) a coesão interna, baseada nas distâncias entre os vetores de dados de um mesmo *cluster* e que indica o quanto um agrupamento é compacto; (ii) e a separação externa, baseada nas distâncias entre os centroides dos *clusters* e que indica a distinção dos *clusters*.

Para determinar o valor de K , neste trabalho foi utilizado o índice Calinski-Harabasz [Calinski and Harabasz 1974]. O cálculo desse índice é definido por (7), em que B_K e W_K são, respectivamente, as matrizes de dispersão entre grupos e intragrupo, das quais são calculados os traços (operador $tr()$).

$$CH(K) = \frac{tr(B_K)/(K-1)}{tr(W_K)/(N-K)}. \quad (7)$$

O índice CH deve ser calculado para vários valores de K , e como é desejado um valor alto de B_K (*clusters* bem distintos) e um valor baixo de W_K (*clusters* compactos), o maior valor de CH indica o K mais adequado para o conjunto de dados.

4. Avaliação Experimental

Os conjuntos de dados descritos na Seção 3 foram divididos em dois períodos de 6 meses: (i) de março a agosto de 2017, e (ii) de setembro de 2017 a fevereiro de 2018. O primeiro período corresponde às estações primavera e verão, e o segundo corresponde às estações outono e inverno no hemisfério norte. O procedimento para os experimentos se deu pela aplicação das técnicas de mineração de dados a cada período dos conjuntos de dados.

Após a limpeza e normalização dos dados, investigou-se qual seria a quantidade de partições mais adequada ao conjunto de dados. Para isso, o índice Calinski-Harabasz foi calculado 100 vezes para valores de K entre 2 e 15. Tomou-se como valor ideal o K que apresentou o maior índice Calinski-Harabasz por mais vezes.

Em seguida, o algoritmo *K-means* foi aplicado até sua convergência. Como os centroides iniciais são escolhidos aleatoriamente, a posição final dos mesmos pode variar a cada execução do algoritmo. Para se definir os melhores centroides, o SSD foi utilizado como critério de decisão. Um menor valor de SSD significa menores distâncias entre os dados dos *clusters* e os seus respectivos centroides. Dessa forma, o algoritmo foi executado 20 vezes para a escolha dos centroides que produziram o menor SSD .

Após a execução do algoritmo *K-means* com os dados normalizados, cada amostra recebeu um rótulo referente ao *cluster* do qual ela faz parte. Esses rótulos foram então aplicados às amostras dos dados não-normalizados, de forma que a correspondência entre as posições das amostras nos dois conjuntos de dados fosse obedecida. Isso foi feito para se obter as estatísticas descritivas dos atributos isolados de cada *cluster* na escala dos dados originais.

5. Resultados

Os centroides dos *clusters* obtidos nos experimentos podem ser vistos na Tabela 1.

Tabela 1. Centroides dos *clusters*

	Arnas 1º período			Arnas 2º período			Emil 1º período			Emil 2º período		
	°C	%	kg	°C	%	kg	°C	%	kg	°C	%	kg
C0	14.3	89.2	25.3	10.3	78.1	31.0	26.5	55.0	12.4	13.5	79.9	15.6
C1	29.9	53.3	23.3	13.4	93.3	31.2	29.4	69.1	14.3	7.2	82.1	14.5
C2	30.8	54.0	11.1	7.7	96.0	29.5	10.9	75.8	14.0	7.0	80.6	16.5
C3	33.6	58.4	37.3							5.9	74.4	14.8
C4	16.4	85.9	24.6							7.9	78.4	15.2
C5										16.3	74.2	16.9

Para o primeiro período, de março a agosto de 2017, a metodologia utilizada retornou como resultado 5 *clusters* para a colmeia do apiário Arnas. Seus centroides são apresentados na Tabela 1, nas colunas sob a descrição “Arnas 1º período”.

Dois *clusters* apresentam centroides semelhantes: o *Cluster 0* e o *Cluster 4*. Eles podem ser entendidos como o estado da colônia durante a transição entre as épocas fria e quente do ano (do inverno para a primavera e do verão para o outono). Chega-se a essa conclusão pelos valores da temperatura (os mais baixos) e da umidade (os mais altos). Isso se dá pela influência das épocas frias (e mais úmidas, vide Figura 1) sobre as colônias.

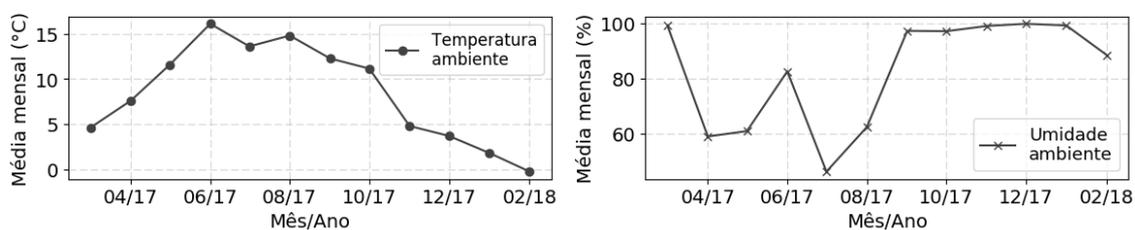


Figura 1. Médias mensais da temperatura e umidade ambientes em Arnas.

O *Cluster 1* pode ser entendido como o estado da colônia em meados da primavera. Isso é evidenciado principalmente pela baixa umidade, já que nessa estação as precipitações, e conseqüentemente a umidade, apresentam médias baixas, como podemos observar na Figura 1. O *Cluster 2* pode ser interpretado como o estado da colônia no verão, não só pela baixa umidade (similar a observada no *Cluster 1*), mas também pela pouca massa. O baixo valor da massa pode indicar a coleta pelo apicultor do mel produzido pelas abelhas durante a primavera. O *Cluster 3* pode ser entendido como o estado da colônia no final da primavera. Nesse momento ocorre o ápice da produção de mel, que se reflete no alto valor da massa do centroide.

Esses resultados podem ser comprovados pela Figura 2(a), que mostra a proporção da quantidade de amostras de cada *cluster* durante os meses do primeiro período. Nota-se que as amostras do *Cluster 0* estão distribuídas principalmente no mês de março, e as amostras do *Cluster 4* no mês de agosto. Esses meses correspondem, respectivamente, ao início da primavera (após o inverno) e ao final do verão (antes do outono). Já as

amostras do *Cluster 1* estão distribuídas principalmente entre os meses de março, abril e maio, que correspondem à primavera. As amostras do *Cluster 2* estão distribuídas entre os meses de junho, julho e agosto, que correspondem ao verão. Por fim, as amostras do *Cluster 3* estão distribuídas principalmente no mês de maio, o último mês da primavera no hemisfério norte.

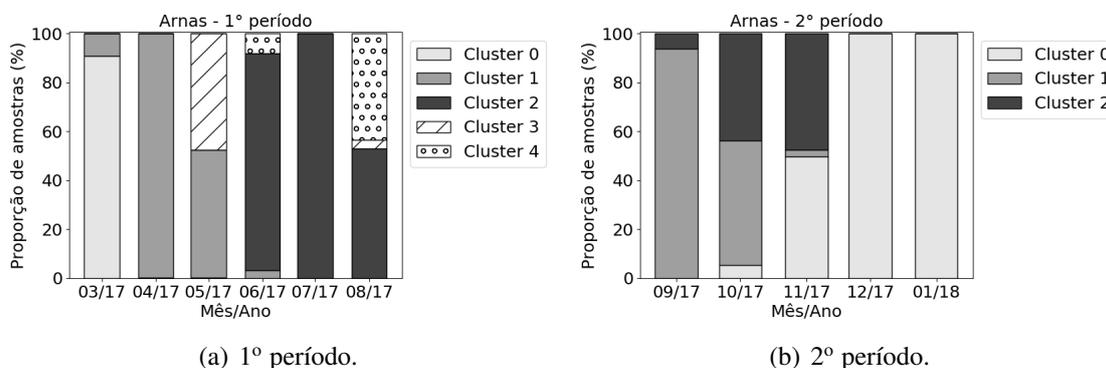


Figura 2. Proporção da quantidade de amostras de cada *cluster* por mês - Arnas.

Para o segundo período, de setembro/2017 a fevereiro/2018, a metodologia utilizada retornou como resultado 3 *clusters* para a colmeia do apiário Arnas, cujos centroides são apresentados na Tabela 1, nas colunas sob a descrição “Arnas 2º período”.

Por terem as temperaturas mais altas, o *Cluster 0* (C0) e o *Cluster 1* (C1) podem ser entendidos como o estado da colônia durante a transição entre as épocas quente e fria do ano (do verão para o outono e do inverno para a primavera). Como o *Cluster 1* possui a umidade mais alta entre os dois, é possível presumir que ele refere-se à transição do verão para o outono, já que a média da umidade é maior no outono, como podemos observar na Figura 1.

Complementarmente, o *Cluster 2* pode ser interpretado como o estado da colônia em meados do segundo período, quando se inicia o inverno. Nessa época ocorrem as menores médias de temperatura, vide Figura 1.

É importante destacar que, se a temperatura da colônia seguisse a tendência da temperatura ambiente observada na Figura 1, o valor da temperatura do *Cluster 0* deveria ser menor. No entanto, ele é maior do que o valor da temperatura do *Cluster 2*. Isso indica a capacidade de termorregulação da colmeia, evidenciada também pelo valor alto e regular da massa, já que o controle térmico depende da presença de um maior número de abelhas operárias [Cook and Breed 2013].

Esses resultados podem ser comprovados pela Figura 2(b). Percebe-se que as amostras do *Cluster 0* estão distribuídas principalmente nos últimos meses do segundo período, que correspondem ao final do inverno (antes da primavera). Já as amostras do *Cluster 1* estão distribuídas principalmente nos meses de setembro e outubro, que correspondem ao início do outono. Por fim, as amostras do *Cluster 2* estão distribuídas principalmente entre os meses de outubro e novembro, quando se encerra o outono e se inicia o inverno.

Para a colmeia do apiário Emil, a metodologia retornou como resultado 3 *clusters* no primeiro período. Seus centroides são apresentados na Tabela 1.

Similar à avaliação feita para a colmeia de Arnas, pode-se interpretar o *Cluster 2* como o estado da colônia durante a transição entre as épocas fria e quente do ano devido ao baixo valor da temperatura. Ainda levando em consideração o valor da temperatura, pode-se interpretar o *Cluster 0* como o estado da colônia na primavera, e o *Cluster 1* no verão, já que este possui o valor mais alto.

Tais resultados podem ser comprovados pela Figura 3(a). Percebe-se que as amostras do *Cluster 2* estão distribuídas principalmente no mês de março, o primeiro mês após o inverno no hemisfério norte. Já as amostras do *Cluster 0* estão distribuídas principalmente nos meses de abril e maio, que correspondem ao final da primavera. Por fim, as amostras do *Cluster 1* estão distribuídas principalmente nos meses de junho, julho e agosto, que correspondem ao verão.

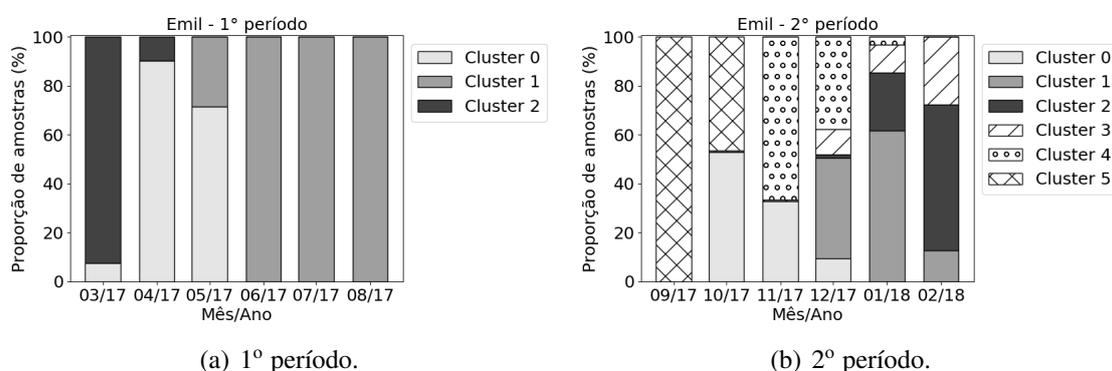


Figura 3. Proporção da quantidade de amostras de cada *cluster* por mês - Emil.

Para o segundo período, a metodologia apresentada retornou como resultado 6 *clusters* para a colmeia do apiário Emil, cujos centroides são apresentados na Tabela 1.

É possível notar que o valor da massa dos centroides é mais baixo se comparado aos de Arnas em ambos os períodos, com exceção de um *cluster*. Pode-se, então, presumir que o controle térmico da colmeia do apiário Emil está comprometido. Isso fica visível principalmente no segundo período, que corresponde ao outono e inverno.

Percebe-se pela Figura 3(b) que a proporção dos *clusters* distribuídos pelos meses do segundo período segue a ordem decrescente da temperatura, o que corresponde à tendência da temperatura ambiente, mostrada na Figura 4. Em setembro/2017, primeiro mês após o verão, há amostras somente do *Cluster 5*, que possui o maior valor de temperatura. Já em fevereiro/2018 as amostras dos *Clusters 2* e *3*, que possuem os menores valores de temperatura, são predominantes.

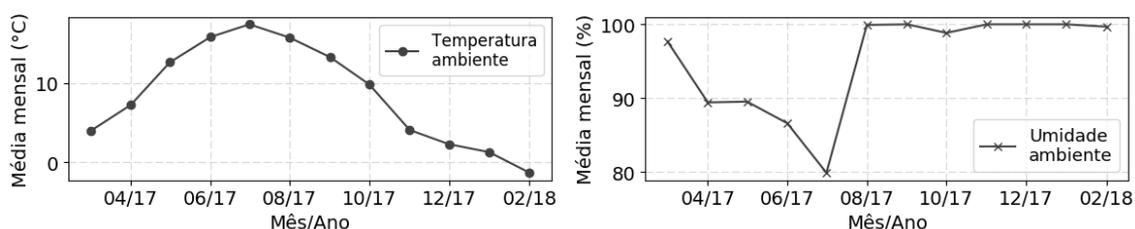


Figura 4. Médias mensais da temperatura e umidade ambientes em Emil.

6. Conclusão

Este artigo apresenta uma proposta para caracterizar os padrões de estados de colônias de abelhas *Apis mellifera* segundo a mudança das estações do ano através da aplicação técnicas de clusterização.

Utilizamos dois conjunto de dados, de uma colônia localizada na Dinamarca e outro de uma colônia localizada na Noruega com os atributos de temperatura, umidade e massa da colmeia. O cálculo do índice Calinski-Harabasz para o conjunto de dados e a aplicação do algoritmo *K-means* resultou na identificação dos padrões relacionados com as transições entre as estações. Os padrões foram reconhecidos como coerentes, já que correspondem ao que se observa em campo. Além disso, concluímos que a colmeia aparentemente mais forte consegue desempenhar com mais eficiência o controle térmico do interior da colmeia durante o inverno do norte da Europa.

Tem-se como perspectiva a aplicação da metodologia apresentada em dados de colônias de *Apis mellifera* africanizada, um polihíbrido criado no Brasil e que carece de estudos nesse sentido. Para isso pretende-se criar uma infraestrutura própria de monitoramento em diversas colmeias, para que se possa, inclusive, caracterizar outros padrões além daqueles relacionados com as estações do ano.

Agradecimentos

Os autores agradecem Leandro Alves Machado pelo apoio técnico. Breno M. Freitas agradece ao CNPq, Bolsa de Produtividade em Pesquisa #302934/2010-3. Danielo G. Gomes agradece ao CNPq, Bolsa de Produtividade em Pesquisa #311878/2016-4 e Universal #432585/2016-8.

Referências

- Aizen, M. A. and Harder, L. D. (2009). The global stock of domesticated honey bees is growing slower than agricultural demand for pollination. *Current Biology*, 19(11):915–918.
- Becher, M. A. (2010). *The influence of developmental temperatures on division of labour in honeybee colonies*. PhD thesis, Halle (Saale), Martin-Luther-Universität Halle-Wittenberg, Diss., 2010.
- Calinski, T. and Harabasz, J. (1974). A dendrite method for cluster analysis. *Communications in Statistics*, 3(1):1–27.
- Cook, C. N. and Breed, M. D. (2013). Social context influences the initiation and threshold of thermoregulatory behaviour in honeybees. *Animal behaviour*, 86(2):323–329.
- Kridi, D. S., Carvalho, C. G. N. d., and Gomes, D. G. (2014). A predictive algorithm for mitigate swarming bees through proactive monitoring via wireless sensor networks. In *Proceedings of the 11th ACM symposium on Performance evaluation of wireless ad hoc, sensor, & ubiquitous networks*, pages 41–47. ACM.
- Kridi, D. S., de Carvalho, C. G. N., and Gomes, D. G. (2016). Application of wireless sensor networks for beehive monitoring and in-hive thermal patterns detection. *Computers and Electronics in Agriculture*, 127:221–235.

- Kviesis, A. and Zacepins, A. (2016). Application of neural networks for honey bee colony state identification. In *Carpathian Control Conference (ICCC), 2016 17th International*, pages 413–417. IEEE.
- Kviesis, A., Zacepins, A., Durgun, M., and Tekin, S. (2015). Application of wireless sensor networks in precision apiculture. *Engineering for Rural Development*, pages 440–445.
- MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, volume 1, pages 281–297. Oakland, CA, USA.
- Meikle, W. G. and Holst, N. (2015). Application of continuous monitoring of honeybee colonies. *Apidologie*, 46(1):10–22.
- Murphy, F. E., Magno, M., Whelan, P. M., O’Halloran, J., and Popovici, E. M. (2016). b+wsn: Smart beehive with preliminary decision tree analysis for agriculture and honey bee health monitoring. *Computers and Electronics in Agriculture*, 124:211–219.
- Ollerton, J., Winfree, R., and Tarrant, S. (2011). How many flowering plants are pollinated by animals? *Oikos*, 120(3):321–326.
- Ostwald, M. M., Smith, M. L., and Seeley, T. D. (2016). The behavioral regulation of thirst, water collection and water storage in honey bee colonies. *Journal of Experimental Biology*, 219(14):2156–2165.
- Potts, S. G., Imperatriz-Fonseca, V. L., Ngo, H. T., Biesmeijer, J. C., Breeze, T. D., Dicks, L. V., Garibaldi, L. A., Hill, R., Settele, J., and Vanbergen, A. J. (2016). Summary for policymakers of the assessment report of the intergovernmental science-policy platform on biodiversity and ecosystem services on pollinators, pollination and food production. Report, Bonn, Germany.
- Ruan, Z.-Y., Wang, C.-H., Lin, H.-J., Huang, C.-P., Chen, Y.-H., Yang, E.-C., Tseng, C.-L., and Jiang, J.-A. (2017). An internet of things-based weight monitoring system for honey. *World Academy of Science, Engineering and Technology, International Journal of Biological, Biomolecular, Agricultural, Food and Biotechnological Engineering*, 11(6):478–482.
- Sánchez, V., Gil, S., Flores, J. M., Quiles, F. J., Ortiz, M. A., and Luna, J. J. (2015). Implementation of an electronic system to monitor the thermoregulatory capacity of honeybee colonies in hives with open-screened bottom boards. *Computers and Electronics in Agriculture*, 119:209–216.
- Tukey, J. W. (1977). *Exploratory Data Analysis*. Addison-Wesley Publishing Company.
- Zacepins, A., Brusbardis, V., Meitalovs, J., and Stalidzans, E. (2015). Challenges in the development of precision beekeeping. *Biosystems Engineering*, 130:60–71.
- Zacepins, A. and Karasha, T. (2012). Web based system for the bee colony remote monitoring. In *Application of Information and Communication Technologies (AICT), 2012 6th International Conference on*, pages 1–4. IEEE.
- Zacepins, A., Kviesis, A., Pecka, A., and Osadcuks, V. (2017). Development of internet of things concept for precision beekeeping. In *Carpathian Control Conference (ICCC), 2017 18th International*, pages 23–27. IEEE.