Similarity among eucalyptus planted areas based on leaf-cutting ant nest sizes

Daniela Maria Lemos Barbato Jacobovitz[1] , Joissy Mayara de Almeida Andrade[2] , Jean-Jacques Georges Soares De Groote[3]*

[1]Educacros/Barbato Estudos de Engenharia, Rua Isis Fernandes, 255, CEP13561-40, São Carlos, SP, Brasil

[2]Eficiente Soluções Florestais Ltda, Rua Alice Além Saadi, 814, Nova Ribeirânia, CEP 14096-570, Ribeirão Preto, SP, Brasil

[3]Centro Universitário Moura Lacerda, Avenida Doutor Oscar de Moura Lacerda, 1520, Independência, CEP 14076-510, Ribeirão Preto, SP, Brasil

Abstract - Techniques for leaf-cutting ant control have been investigated in literature due to the importance of the damage they cause to agriculture. Plantations with large territorial extensions, which can be contiguous or not, are usually subdivided into local administration to collect data to determine the frequencies and area occupied by ant nests. The objective of this work was to build a relationship of similarities among different geographical regions using the frequency data and size of nests by applying Information Bottleneck method and principal component analysis. The approach was applied to data of leaf-cutting ants of the genus *Atta* in cultivated *Eucalyptus* spp. forests in São Paulo State, Brazil. The results showed similar regions that are not geographically close, regarding the occurrence of nests. With this information, administrators will be able to coordinate more precisely the allocation of bait, material and workers in regions with vast territories, avoiding resources waste.

Similaridades entre áreas plantadas de eucalipto com base no tamanho dos ninhos de formigas cortadeiras

Resumo - Técnicas para controle de formigas cortadeiras têm sido investigadas na literatura devido à importância dos danos que causam à agricultura. Plantações com grandes extensões territoriais, contíguas ou não, geralmente são subdivididas em administrações locais, que coletam dados para determinar as frequências e tamanho das áreas ocupadas pelos ninhos de formigas. O objetivo deste trabalho foi construir uma relação de semelhanças entre diferentes regiões geográficas, utilizando dados de frequência da ocorrência e do tamanho dos ninhos, aplicando o método *Information Bottleneck* e a análise das componentes principais. A abordagem foi aplicada a dados de formigas cortadeiras do gênero *Atta* em florestas cultivadas de *Eucalyptus* spp. em São Paulo, Brasil. Os resultados mostraram regiões semelhantes com relação à ocorrência de ninhos, que não estão geograficamente próximas. Com posse desta informação, os administradores poderão coordenar, com maior precisão, a alocação de iscas, material e pessoal em regiões com vastos territórios, evitando o desperdício de recursos.

Introduction

Leaf-cutting ants are important pests of several crops, pastures and planted forests, such as commercial eucalyptus forest. Ants belonging to the *Attini* tribe live in the Americas, from Southern United States to Argentina (Pérez et al., 2011). Brazil has the largest number of leaf-cutting ant species, classified into two genus, *Acromyrmex* Mayr and *Atta* Fabricius (Hymenoptera: Formicidae). Both attack cultivated plants and use the vegetables to nurture a fungal culture they feed on (Zanetti et al., 2014).

Damage caused by the ants are relevant since they continuously cut tender leaves and branches of plants at any stage of development, leading to intense defoliation which can even destroy them altogether (Souza et al., 2011). For planted forests, they are considered the worst pests (Vilela, 2019), being responsible for significant losses, with expenses in their control reaching up to 30% of forest cost by the end of the third cutting cycle.

In Brazil, planted forests were estimate over 7.8 million ha in 2018, represented by 78.1% of *Eucalyptus* L'Hér. (Myrtaceae) and 21.9% of *Pinus* L. (Pinaceae) (IBÁ, 2019). The highest concentration of eucalyptus and pine plantations occurs in the South and Southeast of Brazil (72%), where the niche market of their products is located. The economic interest of eucalyptus is mainly to paper and cellulose industry (72.5% of the planted area), followed by steel and coal (19.5%), wood panels (7.3%) and independent producers (0.7%) (Abraf, 2013).

Economic loss caused by leaf-cutting ants in crops have led to investigations of efficiency and impact of different insecticides (Oi et al., 2000; Choate & Francis, 2013), new baits carriers (Buczkowski et al., 2014) and ant colony mapping to understand and predict their potential expansion (Morrison et al., 2004; Pitt et al., 2009).

The importance of insect control has also encouraged researches focused on statistical modeling of the spatial distribution of nests (Doncaster, 1981; Croft & Hoyt, 1983; Caldeira et al., 2005; Nickele et al., 2010; Gao, 2013; Sousa-Souto et al., 2013). It allows the development of better ant bait distribution strategies through more appropriate sampling plans than common basic standardization applied in most Brazilian plantations (Nickele et al., 2013; Bollazzi et al., 2014; Reis et al., 2016).

In terms of insect spatial arrangements there are three main discrete statistical models, defined as random, described by Poisson distributions, contagious, negative binomial distributions, and regular or uniform, positive binomial distributions (Taylor, 1984; Begon et al., 1996). By modeling the distribution, it is expected that better planning will help to decrease the use of bait, avoiding considerable expenses and also control deleterious effects to the environment resulting from excessive use of insecticides (Laranjeiro, 1994; Nickele et al., 2013).

Plantations with large territorial extensions, which can be contiguous or not, are usually subdivided into local administration (geographical sub-regions) that collects data to determine the frequency and area size occupied by the ant nests.

For such cases, identification of similarities among different geographical sub-regions based on nest distribution can also be relevant information to determine strategies to reduce loss and also avoid unnecessary environmental contamination.

Standard procedure for determining nest size consists in dividing regions into plots (stands) that are subdivided to identify and estimate the nest area using the greatest length and width (Caldeira et al., 2005).

In cases of great variability in the sizes of nests, the amount of data collected and the number of nest size intervals can be large, thus, making it difficult to infer the correlation between regions. For such cases, it becomes necessary to investigate statistical methods developed for multivariate data analysis.

A usual way to analyze a large volume of data is to group them in clusters according to their similarity. However, clustering can lead to loss of information. Different data clustering algorithms are used in information retrieval problems, image segmentation, pattern classification, phylogenetic inference and microarray gene expression. For most of them, it is necessary to pre-set parameters as the distance between pairs of points or distortion rate between point and class. These parameters are arbitrary and influence on the result. For certain problems, the definition of an appropriate measure is a difficult task. Information Bottleneck (IB) is an unsupervised clustering method based on information theory, which allows information compression while preserving its relevance. Given two random variables through the IB, each group is composed in a way that is maximally informative about its elements. IB method has been applied in different areas

such as speech recognition (Hecht et al., 2009), galaxy spectra classification (Slonim et al., 2001), psychometric properties of questionnaires (Barbato & De Groote, 2018) and unsupervised document clustering, essential for problems information retrieval (Slonim & Tishby, 2000), where similarity measure between documents is the similarity between their conditional distributions of words. Variants for geometric clustering have also been developed (Strouse & Schwab, 2019). Currently IB has been pointed out as a method that can help to understand deep learning (Tishby & Zaslavsky, 2015).

The objective of this work is to apply the theory of information approach IB to provide a direct relationship between different geographical regions based on the frequency of nests size occurrence. IB method was applied in conjunction with the principal component analysis, which is used for multivariate data analysis to uncorrelated variables and is useful as a tool to reduce the dimension size to facilitate a visual data structure analysis.

Material and methods

Data

We used the ant nests data presented by Andrade et al. (2014), obtained from seven regions of eucalyptus plantations in São Paulo State, $\{R_i, i=1..7\}$. The regions begin in the Itararé town towards São Simão, SP, comprising about 43,000 ha of plantations. The authors pointed out that these regions stand out for the infestation of ants, obtaining indications that the similarity between distributions may be related to a different form of management, as well as the altitude, suggesting data collections in order to prove the hypothesis.

The average size of the nests found in each region of eucalyptus plantations was applied to construct the

Table 1. Nests occurrence frequency from different regions (R_i).

Nests area (m²)	R_1	R_2	R_3	R_4	R_5	R_6	R_7
0-5	167	60	146	55	145	91	62
5-10	4	18	4	24	10	13	41
10-15	0	12	3	16	5	9	32
15-20	1	5	2	16	3	8	26
20-25	0	11	3	2	3	2	15
25-30	1	6	1	9	2	7	7
30-35	1	8	0	9	4	4	0
35-40	2	5	4	9	1	2	0
40-45	1	4	0	5	1	4	0
45-50	0	6	0	9	3	3	0
50-55	0	4	1	5	0	5	0
55-60	1	9	2	3	0	4	0
60-65	0	2	0	4	0	3	0
65-70	1	1	2	1	1	1	0
70-75	1	4	1	1	1	5	0
75-80	0	1	3	0	0	2	0
80-85	0	2	0	0	0	1	0
85-90	1	4	0	2	1	4	0
90-95	0	1	0	2	0	1	0
95-100	0	1	0	3	0	2	0
100-105	1	0	0	1	0	1	0
105-110	0	1	0	0	0	4	0
110-115	0	2	0	1	0	0	0

Cont.

D. M. L. B. Jacobovitz et al.

Table 1. Cont.

Nests area (m²)	R_1	R_2	R_3	R_4	R_5	R_6	R_7
115-120	0	3	2	0	0	0	0
120-125	0	1	1	0	0	0	0
125-130	0	1	2	2	0	1	0
130-135	0	1	0	0	0	0	0
135-140	0	0	0	0	0	2	0
140-145	0	0	0	1	0	1	0
145-150	0	2	0	0	0	0	0
150-155	0	2	0	0	1	1	0
155-160	0	1	1	0	0	0	0
160-165	0	2	0	0	0	0	0
165-170	0	3	0	1	0	0	0
170-175	0	0	0	1	0	0	0
175-180	0	0	1	0	0	0	0
180-185	0	0	0	1	0	0	0
185-190	0	0	0	0	0	1	0
195-200	0	0	0	0	0	1	0
210-215	0	0	0	0	1	0	0
225-230	0	0	1	0	0	0	0
255-260	0	0	1	0	0	0	0
270-275	0	1	0	0	0	0	0
>300	1	0	2	0	1	0	0

Data collected in 2012 (Andrade et al., 2014). Area values without nests occurrence are not shown.

occurrence frequency according to classes ranged in group intervals of 5 m² from 0 up to > 300 m², as seen in Table 1 (Andrade et al., 2014).

Principal Component Analysis

Principal component analysis (PCA) is a method used in the study of large multidimensional dataset. PCA in this work is applied as an auxiliary tool to reduce the data variables dimension allowing the visualization of results produced by IB method. This method is based on the covariance matrix and their eigenvalues and eigenvectors determination. The eigenvector with bigger eigenvalue is the principal component, related to the largest variance. Original data are projected into orthogonal eigenvector base, in such a way that they become uncorrelated. The dimension can be reduced by choosing eigenvectors with the highest eigenvalues, which explains most of the variance of the sample (Jolliffe, 2011). In the new

base, we can observe clusters, but in some cases, its configuration is not obvious, therefore other methods are necessary.

Information Bottleneck

Given the random variables X and Y and their joint probability distribution $p(x,y)$, the proposal of Information Bottleneck (IB) is to compress data from variable X into a variable T in a configuration that preserves maximum information about variable Y. The degree of compression is represented by variable $t \in T$, where $t = 1..k$, and k is the number of clusters formed. If the compression is too big, there is information loss, but if there is no compression, the analysis becomes more difficult. This method consists in the minimization of a functional based on the mutual information between X and T (Li, 1990), representing compression, and between Y and T, which represents relevancy. A functional

parameter β (inverse temperature) mediates compression and relevance.

We search a compressed representation of X, through the variable T, which preserves a significant fraction of information about Y. IB method aims to minimize the functional (Equation 1).

$$L[p(t \mid x)] = I(X,T) - \beta I(T,Y) \qquad (1)$$

Where $I(X,T)$ is the mutual information between variables X and T, and $I(Y,T)$ is the mutual information between variables Y and T. The Lagrange multiplier β controls tradeoff between compression and relevance.

The first term in Equation (1) is related to compression and the second is related to relevance. For $\beta \to 0$ we have maximum compression and information loss, for $\beta \to \infty$ only relevance is preserved. For finite β, we can find a compressed X representation without significant loss of relevance. For given T, β and $p(x,y)$, IB method provides a conditional probability distribution $p(t|x)$ that minimizes the functional L.

The IB method is characterized by the iterative algorithm (Tishby et al., 2000),

- Input: $p(x,y)$, β e T.
- Initialization: $p(t|x)$ is set with random values.
- Iterative equations (Equations 2 to 4).

$$p(t_i) = \sum_j p(x_j) p(t_i | x_j) \qquad (2)$$

$$p(y_i \mid t_j) = \sum_k p(y_i \mid x_k) p(x_k \mid t_j) \qquad (3)$$

$$p(t_i \mid x_j) = \frac{p(t_i)}{Z(x_j, \beta)} EXP\left(-\beta \sum_k p(y_k \mid x_j) \log\left(p(y_k \mid x_j) / p(y_k \mid t_i)\right)\right) \qquad (4)$$

Where Z is the partition function defined as Equation 5.

$$Z(x_j, \beta) = \sum_i p(t_i) EXP\left(-\beta \sum_k p(y_k \mid x_j) \log\left(p(y_k \mid x_j) / p(y_k \mid t_i)\right)\right) \qquad (5)$$

This algorithm finds the probability distributions $p(t|x)$ and $(y|t)$, that shows how variables are grouped, i.e., the conditional probability indicates how each element $x \in X$ and $y \in Y$ belongs to each cluster $t \in T$.

We applied the iterative algorithm for different values of IB parameters t and β with the variable X corresponding to nests classes and Y as composed by the geographical regions R. The probability $p(x)$ was obtained by normalizing the distribution of each nest area.

Results

A three-dimensional representation of the space formed by the variable R can be built when applying the principal component analysis (PCA) procedure. The first three components are associated with 99.7% explained variance indicating a very small loss of information for such representation. Eigenvalues and scores for three principal components are presented in Table 2.

Table 2. Eigenvalues for the three principal components and the corresponding scores of each region Ri (i=1..7).

	Eigenvalues		
	1	2	3
Variables	2053.154	88.997	7.318
R_1	-0.557759	0.285425	-0.097421
R_2	-0.200217	-0.220554	0.424748
R_3	-0.486617	0.217634	-0.192129
R_4	-0.187064	-0.392865	0.639681
R_5	-0.485939	0.111360	-0.031971
R_6	-0.303610	-0.077944	0.259211
R_7	-0.220711	-0.806043	-0.543857
Cumulative (%)	95.3	99.3	99.7

The results of data projection into two and three eigenvectors chosen in eigenvalues descending order are presented in Figures 1 and 2.

In order to calculate clusters of regions that minimizes the loss of information about the nests sizes, we have applied the iterative algorithm described in the previous section for different values of Information Bottleneck (IB) parameters t and β. From a random initial probability distribution $P(t)$, the conditional probability values $p(t|x)$ were calculate by the iteration process, which proves to be stable, with up to 15 iterations to converge. Initial condition testing was performed for each set of parameters t and β in order to determine solutions that minimized the functional $L[p(t|x)]$. The cluster tree diagram obtained for different values of parameter t as $\beta \to \infty$ is shown in Figure 3.

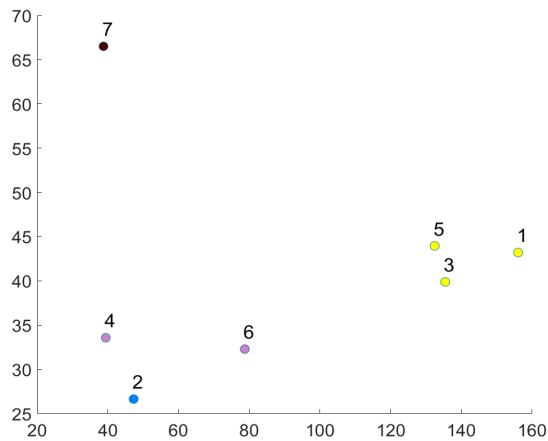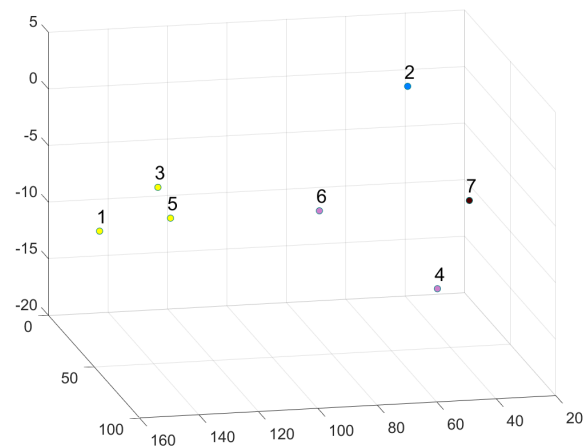**Figure 1**. Data from nests size of Table 1 projected into two eigenvectors with higher eigenvalues.

Figure 2. Data from nests size of Table 1 projected into three eigenvectors with higher eigenvalues.
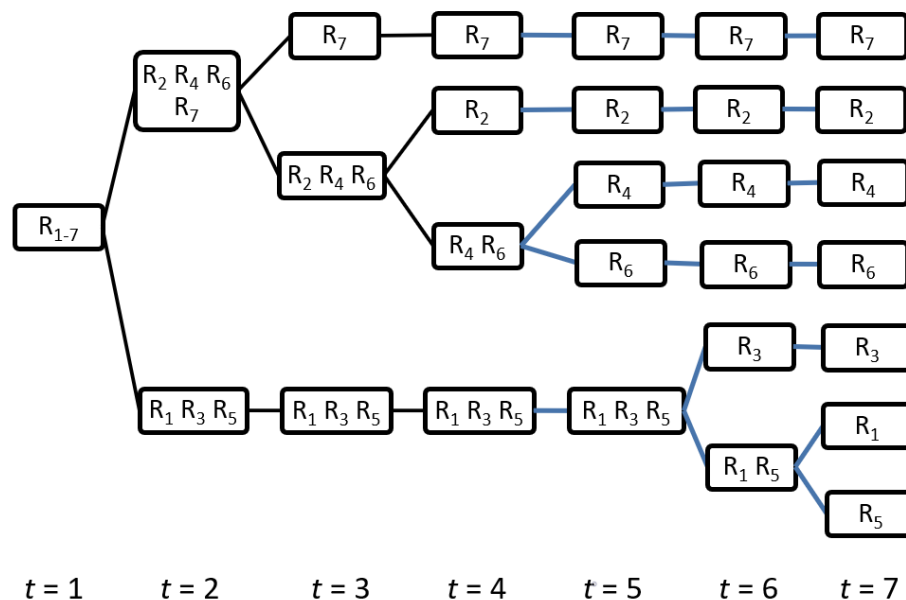

Figure 3. Regions clusters (t) formed using Information Bottleneck method associated with the occurrence of nests areas (X).

Regions clusters (Y) formed using IB method associated with occurrence of nests areas (X) determined from conditional probability $p(t|x)$, are also obtained as a result from the iterative calculations. Results are shown in Figure 4 for the three main clusters.

Observation of the projections shown on Figure 1 and 2, obtained with PCA, suggests three clusters formed by a separation of regions $\{R_1, R_3, R_5\}$, $\{R_2, R_4, R_6\}$ and Region 7, isolated from the other sets. Although three clusters are observed in the PCA results (Figure 1), it may be of interest to identify different clusters within the data.
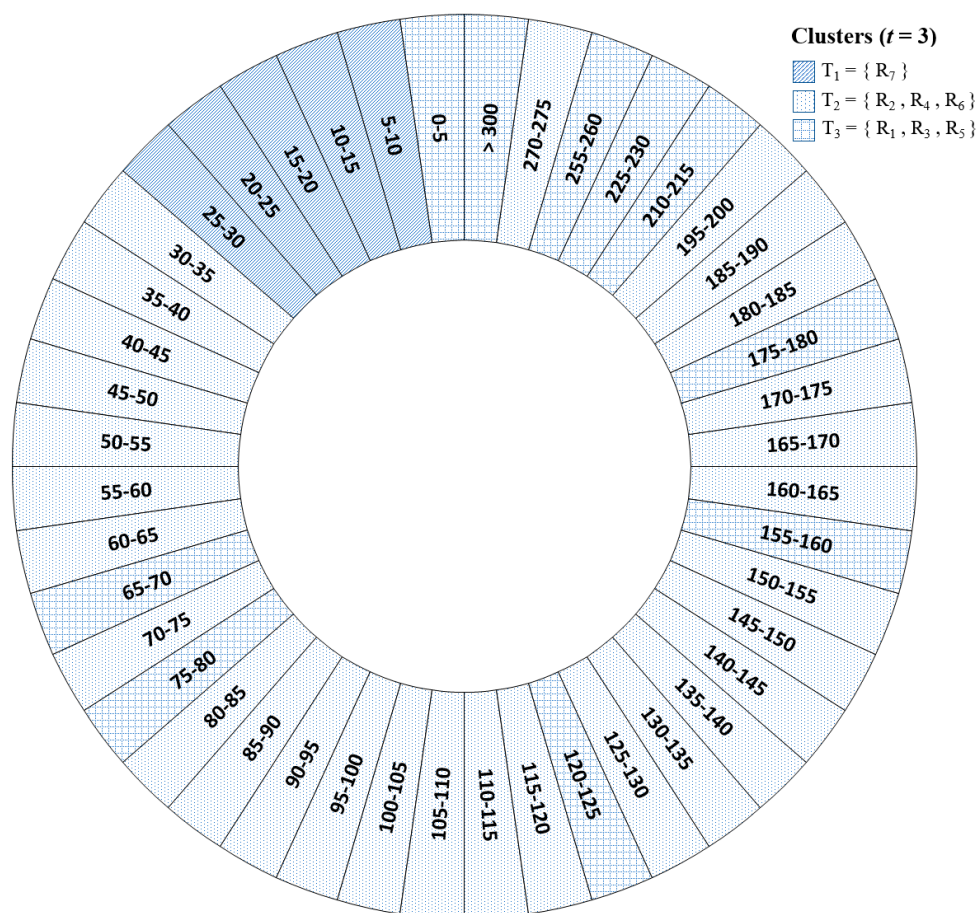
Figure 4. Ants nests occurence in regions *R* from Table 1 associated to each cluster generated by Information Bottleneck with *t*=3.

Discussions

Applying Information Bottleneck (IB) is possible to improve the subjective graphical of principal component analysis (PCA) results by providing a precise approach to group the regions for a given number of clusters (Slonim & Tishby, 2000; Zaidi et al., 2020). For small β values, we observed maximum compression at which just one cluster was formed regardless of the cluster´s t parameter choice. For the interval $3.5 < \beta < 4.8$ two clusters were formed with the region sets T_1: $\{R_2, R_4, R_6, R_7\}$ and T_2: $\{R_1, R_3, R_5\}$. No other cluster was formed for β parameter within this interval. Such behavior is obtained from $p(y|t)$ conditional probability which are the same for different values of t. This means that for a fixed β increasing t does not lead to new cluster structures. Above $\beta = 4.8$ three clusters were obtained,

T_1: $\{R_7\}$, T_2: $\{R_2, R_4, R_6\}$ and T_3: $\{R_1, R_3, R_5\}$, which are the maximum partition up to $\beta \approx 9.2$. With $\beta > 9.2$ the $\{R_2, R_4, R_6\}$ cluster splits into $\{R_4, R_6\}$ and $\{R_2\}$.

Within PCA´s 2D data projection, the separation described is not clear, but within the 3D figure (Figure 2), it may be observed that R_4 and R_6 are geometrically closer to each other than to R_2. It is interesting to point out this result as an example in which PCA dimensionality reduction can lead to a loss of information with respect to clustering. Information Bottleneck method analysis, on the other hand, does not requires prior dimensionality reduction in order to compress data into clusters (Tishby et al., 2000). This method also allows gradual data compression when parameter β is increased, as may be observed in a cluster tree diagram as shown in Figure 3. A misleading interpretation could be performed by administrators, when neighboring regions, for example

regions 5 and 6 (Andrade et al., 2014) are considered as belonging to the same cluster.

Using IB with PCA on the data is a consistent approach to group the regions for a given number of clusters. For two groups, the IB method showed that regions similar with respect to infestation levels are 2, 4, 6, and 7 in a group, 1, 3 and 5 in another. As parameter cluster t grows, new clusters emerge. The set $\{R_2, R_4, R_6\}$ is dismembered, while there is a significant tendency to preserve $\{R_1, R_3, R_5\}$, which are kept together for up to 5, as seen in Figure 3. Region 7 is shown with distinct characteristics from the others for every parameter choice above $t = 2$.

The significant damage caused by leaf-cutting ants to pinus and eucalyptus plantations, as observed by Zanetti (2014), has led Brazilian foresters to invest in monitoring systems to reduce costs and environmental impacts. The quantification of population density (Sousa-Souto, 2013) of the pests is one of the aspects of an integrated pest management that has been developed in Brazil to keep the pest population below the economic injury level (Zanetti et al., 2014). Based on information theory, the method presented in this article can add new information to help optimize resource management because it identifies and groups similar regions with respect to nest size distributions (Croft & Hoyt, 1983).

The occurrence of the leaf-cutting ant nests can be related to several environmental factors such as sunlight exposure, humidity, altitude and food availability (Doncaster, 1981; Kaspari et al., 2000). To a lesser degree, dispersion methods and competitive strategies may be considered (Ricklefs & Schluter, 1994; Tanner & Keller, 2012).

A possible application of the procedure is spatial–temporal investigations of the dissemination of ants around the globe by identifying correlations among different regions (Dillier & Wehner, 2004). It also may be applied to other types of crops and pest infestations, such as coffee leaf miners (Scalon et al., 2011), *Alabama argillacea* (cotton caterpillar) (Fernandes et al., 2003) and larvae of corn (Farias et al., 2001). Those are some examples of species whose spatial distribution have been investigated to determine control strategies, the optimization of sampling techniques, and also the determination of economic damage.

Conclusions

This work shows the application of Information Bottleneck (IB) method combined with principal component analysis (PCA) to group nests of leaf-cutter ants according to their occurrence distribution in each area investigated. The analysis showed similarity among regions regarding to the occurrence of nests. While IB allows a hierarchical cluster analysis, PCA furnishes a visual aid by reducing sample variable dimension.

The application of the IB directly related the clusters of regions and nests, ensuring by means of information theory that these clusters were identified with the least loss of information. In turn, the results of the PCA, in addition to confirming the results of the IB, show that this method can be used as an instrument for visual analysis of the proximity and distancing of each region based on nest distributions.

However, it is important to note that for this visualization to be performed with the PCA, it is necessary to reduce the random variable related to the nest sizes to three dimensions. The loss of information can lead to interpretation errors, which is not the case with the IB results. Thus, we conclude that the application of PCA and IB together can be useful for analyzing data acquired on leaf-cutting ant infestations. For future work, the results can be used to cross-check data related to altitude, humidity, temperature and soil types to discover the parameters that define the similarities identified.

Acknowledgements

References

Abraf. Brazilian Association of Producers of Planted Forests. **Statistical yearbook 2013**. Brasília, DF, 2013.

Andrade, J. M. A. et al. Avaliação do método estatístico de análise da componente principal (PCA) para níveis de infestação de formigas cortadeiras em plantações de eucalipto. **Revista Científica Eletrônica Estácio Uniseb**, v. 3, p. 181-199, 2014.

Barbato, D. M. & De Groote, J. J. Extracting relevant structures from self-determination theory questionnaires via information bottleneck method. **arXiv preprint**, arXiv:1806.01345, 2018.

Begon, M. et al. **Ecology**: individuals, populations and communities. 3rd ed. Oxford: Black-Well Science, 1996. https://doi.org/10.2307/2960512.

Bollazzi, M. et al. Efficiency and soil contamination during underground application of insecticides: control of leaf-cutting ants with thermal foggers. **Journal of Pest Science**, v. 87, n. 1, p. 181-189, 2014. https://doi.org/10.1007/s10340-013-0525-7.

Buczkowski, G. et al. Polyacrylamide hydrogels: an effective tool for delivering liquid baits to pest ants (Hymenoptera: formicidae). **Journal of Economic Entomology**, v. 107, n. 2, p. 748-757, 2014. https://doi.org/10.1603/ec13508

Caldeira, M. A. et al. Distribuição espacial de sauveiros (Hymenoptera: Formicidae) em eucaliptais, **Cerne**, v. 11, n. 1, p. 34-39, 2005.

Choate, B. & Francis, A. D. The influence of insecticides and vegetation in structuring formica mound ant communities (Hymenoptera: formicidae) in Maine Lowbush Blueberry. **Journal of Economic Entomology**, v. 106, n. 2, p. 716-726, 2013. https://doi.org/10.1603/ec12273.

Croft, B. A. & Hoyt, S. C. **Integrated management of insect pests of pome and stone fruits**. New York: Wiley Interscience, 1983. https://doi.org/10.1016/0261-2194(85)90045-6.

Dillier, F. X. & Wehner, R. Spatio-temporal patterns of colony distribution in monodomous and polydomous species of North African desert ants, genus Cataglyphis. **Insectes Sociaux**, v. 51, n. 2, p. 186-196, 2004. https://doi.org/10.1007/s00040-003-0722-0.

Doncaster, C. P. The spatial distribution of ant' nests on Ramsey Island, South Wales. **Journal of Animal Ecology**, Oxford, v. 50, p. 95-218, 1981. https://doi.org/10.2307/4040.

Farias, P. R. S. et al. Distribuição espacial da lagarta-do-cartucho, *Spodoptera frugiperda* (J. E. Smith) (Lepidoptera: noctuidae), na cultura do milho. **Neotropical Entomology**, v. 30, n. 4, p. 681-689, 2001. https://doi.org/10.1590/s1519-566x2001000400025.

Fernandes, M. G. et al. Distribuição espacial de Alabama argillacea (Hübner) (Lepidoptera: Noctuidae) em algodoeiro. **Neotropical Entomology**, v. 32, n. 1, p. 107-115, 2003. https://doi.org/10.1590/s1519-566x2003000100016.

Gao, M. Detecting spatial aggregation from distance sampling: a probability distribution model of nearest neighbor distance. **Ecological Research**, v. 28, p. 397–405, 2013. https://doi.org/10.1007/s11284-013-1029-x.

Hecht, R. M. et al. Speaker recognition by gaussian information bottleneck. **Interspeech**, p 1567–1570, 2009.

IBÁ. Indústria Brasileira de Árvores. **Relatório 2019 = Report 2019**. Brasília, DF, 2019. 80 p. Disponível em: https://www.iba.org/datafiles/publicacoes/relatorios/iba-relatorioanual2019.pdf. Acesso em: 19 nov. 2020.

Jolliffe, I. **Principal component analysis**. Berlin: Springer Berlin Heidelberg, 2011. p. 1094-1096. https://doi.org/10.1007/978-3-642-04898-2_455.

Kaspari, M et al. Energy, density, and constraints to species richness: ant assemblages along a productivity gradient. **The American Naturalist**, v. 155, p. 280-293, 2000. https://doi.org/10.1086/303313.

Laranjeiro, A. J. Integrated pest management at Aracruz Celulose. **Forest Ecological Management**, v. 65, n. 1, p. 45-52, 1994. https://doi.org/10.1016/0378-1127(94)90256-9.

Li, W. Mutual information functions versus correlation functions. **Journal of Statistical Physics**, v. 60, p. 823-837, 1990. https://doi.org:10.1007/BF01025996.

Morrison, L. W. et al. Potential global range expansion of the invasive fire ant, Solenopsis invicta. **Biological Invasions**, v. 6, n. 2, p. 183-191, 2004. https://doi.org/10.1023/b:binv.0000022135.96042.90.

Nickele, M. et al. Distribuição espacial de formigueiros de *Acromyrmex crassispinus* (Forel) (Hymenoptera: Formicidae) em plantios de *Pinus taeda*. **Neotropical Entomology**, v. 39, n. 6, p. 862-872, 2010. https://doi.org/10.1590/s1519-566x2010000600003.

Nickele, M. et al. Formigas cultivadoras de fungos: estado da arte e direcionamento para pesquisas futuras. **Pesquisa Florestal Brasileira**, v. 33, n. 73, p. 53-72, 2013. https://doi.org/10.4336/2013.pfb.33.73.403.

Oi, D. H. et al. Bait distribution among multiple colonies of Pharaoh ants (Hymenoptera: Formicidae). **Journal of Economical Entomology**, v. 93, n. 4, p. 1247-1255, 2000. https://doi.org/10.1603/0022-0493-93.4.1247.

Pérez, P. S. et al. Potential impact of the leaf-cutting ant Acromyrmex lobicornis on conifer plantations in northern Patagonia, Argentina. **Agricultural and Forest Entomology**, v. 13, n. 2, p. 191-196, 2011. https://doi.org/10.1111/j.1461-9563.2010.00515.x.

Pitt, J. P. et al. Predicting Argentine ant spread over the heterogeneous landscape using a spatially explicit stochastic model. **Ecological Applications**, v. 19, n. 5, p. 1176-1186, 2009. https://doi.org/10.1890/08-1777.1.

Reis, A. et al. Mechanized systematic application of granulated ants baits in eucalyptus plantations in maintenance phase. **Cerne**, v. 21, n. 3, p. 423-428, 2016. http://dx.doi.org/10.1590/01047760201521031331.

Ricklefs, R. E. & Schluter, D. **Species diversity in ecological communities**: historical and geographical perspectives, 2nd ed. Chicago: University of Chicago Press, 1994. https://doi.org/10.1007/bf02803887.

Scalon, J. D. et al. Spatial distribution of the coffee-leaf-miner (Leucoptera coffeella) in an organic coffee (*Coffea arabica* L.) field in formation, **Coffee Science**, v. 6, n. 3, p. 226-232, 2011.

Slonim, N. & Tishby, N. Document clustering using word clusters via the information bottleneck method. Annual international ACM SIGIR conference on Research and development in information retrieval, 23, Athens, 2000. **Proceedings** [...]. New York: ACM, 2000. p. 208-215,https://doi.org/10.1145/345508.345578.

Slonim, N. & Tishby, N. **Document clustering using word clusters via the information bottleneck method.** Proceedings of the 23rd annual international ACM SIGIR conference on Research and development in information retrieval. ACM, New York, p. 208-215, 2000. https://doi.org/10.1145/345508.345578.

Souza, A. et al. Economic damage level for leaf-cutting ants in function of the productivity index of eucalyptus plantations in an atlantic forest region. **Neotropical Entomology**, v. 40, n. 4, p. 483-488, 2011.

Sousa-Souto, L. et al. Spatial distribution of *Acromyrmex balzani* (Emery) (Hymenoptera: Formicidae: Attini) nests using two sampling nethods. **Sociobiology**, v. 60, n. 2, p. 162-168, 2013. https://doi.org/10.13102/sociobiology.v60i2.162-168.

Strouse, D. J. & Schwab, D. J. The information bottleneck and geometric clustering. **Neural Computation**, v. 31, n. 3, p. 596-612, 2019. https://doi.org/10.1162/neco_a_01136.

Tanner, C. J. & Keller, L. Nest distribution varies with dispersal method and familiarity-mediated aggression for two sympatric ants. **Animal Behaviour**, v. 84, n. 5, p. 1151-1158, 2012. https://doi.org/10.1016/j.anbehav.2012.08.018.

Taylor, L. R. Assessing and interpreting the spatial distributions of insect populations. **Annual Review of Entomology**, v. 29, p. 321-357, 1984. https://doi.org/10.1146/annurev.en.29.010184.001541.

Tishby, N. et al. **The information bottleneck method**. v. 1, 24 abr, 2000. https://arXiv:preprint physics/0004057.

Tishby, N. & Zaslavsky, N. Deep learning and the information bottleneck principle. In 2015 IEEE Information Theory Workshop, 2015, Jerusalem. [**Proceedings**...]. [S.l.]: IEEE, 2015. p. 1-5. https://doi.org/10.1109/ITW.2015.7133169.

Vilela, E. F. Status of leaf-cutting ant control in forest plantations in Brazil. **Fire Ants and Leaf-Cutting Ants**, p. 399–408, Mar. 2019. https://doi.org/10.1201/9780429038266-33.

Zanetti, R. et al. An overview of integrated management of leaf-cutting ants (Hymenoptera: Formicidae) in Brazilian forest plantations. **Forests**, v. 5, n. 3, p. 439-454, 2014. https://doi.org/10.3390/f5030439.

Zaidi, A. et al. On the information bottleneck problems: models, connections, applications and information theoretic views. **Entropy**, v. 5, n. 3, p. 151, 2020. https://doi.org/10.3390/e22020151.