Tree species consistent co-occurrence in seasonal tropical forests: an approach through association rules analysis
Cléber R. Souza¹*, Vinicius A. Maia¹, Natália de Aguiar-Campos¹, Camila L. Farrapo¹, Rubens M. Santos¹
¹Forest Sciences Department, Federal University of Lavras, Lavras, Minas Gerais, Brazil.

Abstract

Aim of study: Assessing the existence of consistent co-occurrence between tree species that characterize seasonal tropical forests, using a novel data mining methodology; and evaluating the taxonomic and functional similarities between associated species.

Area of study: forty-four seasonal forest sites with permanent plots (40.2 ha of total sample) located in Southeast Brazil, from which we obtained species occurrences.

Materials and methods: we applied association rules analysis (ARA) to the dataset of species occurrence in sites considering the criteria of support equal to or greater than 0.63 and confidence equal to or greater than 0.8 to obtain the first set of associations rules between pairs of species. This set was then submitted to Fisher’s criteria exact p-value less than 0.05, lift equal to or greater than 1.1 and coverage equal to or greater than 0.63. We considered these criteria to be able to select non-random and consistent occurring associations.

Main results: We obtained a final result of 238 rules for semideciduous forest and 11 rules for deciduous forests, composed of species characteristic of vegetation types. Co-occurrences are formed mainly by non-confamilial species, which have similar functional characteristics (potential size and wood density). There is a difference in the importance of co-occurrence between forest types, which tends to be less in deciduous forests.

Research highlights: The results point out the feasibility of applying ARA to ecological datasets as a tool for detecting ecological patterns of coexistence between species and the ecosystems functioning.

Keywords: data mining; coexistence; semideciduous forests; deciduous forests; biotic interaction.

Introduction

Tropical forests are the most biodiverse ecosystems in the world, where hundreds of species coexist in the same space (Wright, 2002; Barlow et al., 2018). Species coexisting in these ecosystems use the same local set of resources and interact in positive or negative ways (e.g., facilitation, competition); the multidimensional nature of these interactions has significant consequences to ecosystem function due to driving the full community use of resources and their final patterns of important ecosystems attributes such as carbon stock and uptake and biodiversity (Wright, 2002; Hart & Marshall, 2013; Barraclough, 2015; Schmid et al., 2020). Understanding the underlying mechanisms of species interactions in tropical forests has been a central issue in ecology, especially considering its fundamental role in maintaining biodiversity (Chesson, 2000; Wrigth, 2002; Hart et al., 2017).

In general, species that coexist have different ecological requirements or different ecological niches (Wright, 2002; Amarasekare et al., 2004; Barraclough, 2015; Kraft et al., 2015). When their niches overlap, species may face short-term consequences in the local community scale, such as competitive exclusion; and long-term consequences in evolutionary or biome scales, such as niche differentiation and habitat partitioning (Wright, 2002; Barraclough, 2015; Kraft et al., 2015; Chen et al., 2020). In addition, the mechanisms of species coexistence are also influenced by resource availability, which may strongly control the importance of interactions in
community assembly and define its most important mechanisms (Holmgren & Scheffer, 2010; Cadotte & Tucker, 2017). For example, in harsh environments process such as facilitation may have relatively greater importance than competition (Wright, 2002; Carrion et al., 2017). In addition, under restrictive conditions, species adopt a series of physiological mechanisms to assist in their survival by increasing their resistance to stressful environmental conditions and enhancing their ability to obtain and use resources efficiently (Cadotte & Tucker, 2017; van der Sande et al., 2017). These mechanisms can be observed in broad ecological strategies, such as resprouting, scleromorphic traits or deciduousness, or in the level of adaptation of internal structures and associated processes (e.g., variations in diameter and length of vessels, stomatal conductance) (Zeppel et al., 2015; Pausas et al., 2016; Jimenez-Rodriguez et al., 2018; He et al., 2019).

Species coexistence also affects forest ecosystem function, stability, resilience and important ecosystem services such as carbon uptake and carbon stocks (Barraclough, 2015; van der Sande et al., 2017; van der Plas, 2019). High species diversity usually promotes high productivity and ecosystem stability because the niches of high numbers of coexisting species tend to be complementary rather than overlapping (Tilman, 1999; van der Sande et al., 2017). Coexisting species with different requirements are able to exploit available resources more efficiently and, consequently, the overall ecosystem achieves higher productivity (Chesson, 2000; van der Sande et al., 2017).

Questions related to the maintenance and management of species coexistence in tropical forests have been addressed with different approaches, such as theoretical and statistical models and empirical data (Chesson, 2000; Wright, 2002; Hart et al., 2017; Chen et al., 2020; Schmid et al., 2020). The emergence of new methodologies of machine learning and statistical analysis may potentially contribute to our understanding of important ecological patterns: coexistence between species, their relation to environment and also in identifying indicators species of the ecosystems. An example is provided by the data-mining technique association rule analysis, or ARA, that is an important market baskets analysis (Silverstein et al., 1998; Rossi et al., 2014). ARA is widely used in online product sales through different algorithms for allowing efficient identification of associations between elements in extensive datasets (Agrawal et al., 1993; Ferrarini & Tomaselli, 2010; Zumel et al., 2019). However, despite its potential, its use in scientific contexts is still scarce and, to our knowledge, even more so in ecological studies. The few examples of use of ARA in ecological contexts are Leote et al. (2020), that used to identify indicator species of arthropods; and Ferrarini & Tomaselli (2010) and Rossi et al., (2014) that used ARA to identify relations between vegetation and environmental attributes in studies at landscape level.

Here, we used ARA to assess patterns of species coexistence using as study case the tropical seasonal forest (deciduous and semideciduous), that may be broadly characterized by enduring high temperatures and the alternation between a rainy and a dry season every year, that imply deciduousness of most of the species under strong water stress (DRYFLOR, 2016). Most studies on tropical species coexistence have mainly focused on rainforests, despite of less the significant contribution of seasonal forests to tropical biodiversity and ecosystem service provision (Sunderland et al., 2015; DRYFLOR, 2016). We thus applied ARA on the data of 44 sites of tropical seasonal forests in order to (1) identify consistent associations between species (i.e., patterns of species coexistence) that may characterize these forest types, (2) evaluate functional and taxonomic similarities between consistently associated species. In addition, we discuss the feasibility of applying ARA into ecological research as a tool to identify patterns of species coexistence, structure, function and diversity of communities.

**Material and methods**

**Data set**

Here we used a dataset of 44 sites of seasonal tropical forests in central-eastern Brazil (Fig. 1; Table S1 [suppl.]), with 22 semideciduous forest sites and 22 deciduous forests sites. Semideciduous forests are associated with the Atlantic Forest domain and between 50 and 70% of the forest canopy loses its leaves in the dry season (Neves et al., 2017). These sites are under a Köppen Cwa climate (subtropical with dry winters and rainy summers), with average annual rainfall between 1400 and 1500 mm and average monthly temperature between 19 and 20 °C. Deciduous forests are affiliated with the Caatinga domain (a nucleus of seasonally dry tropical forests - SDTF) in which more than 70% of the species lose their leaves in the rainy season (Pennington et al., 2009), and that is located in the southern Caatinga or in islands of fertile soil in Cerrado domain (Fig. 1). These sites are located under a Koppen Aw/As climate transition (tropical dry winter), with average annual rainfall between 750 and 1000 mm and average monthly temperature between 22 to 24.6 °C (Maia et al., 2020).

Each site was sampled with a varied number of permanent sampling units (5 to 128 per site; total of 1077) with varying dimensions (400 m² in most cases, but a few with 225 and 300 m²) depending on local terrain features and fragment size. Total sampled area was 40.2 ha: 23.68 ha in semideciduous forests and 16.52 ha in deciduous forests. Within each sampling unit, we measured and identified to the species level all trees with a diameter at breast height (DBH; 1.30 m above the ground) ≥ 5 cm.
Plant identification followed APG IV (APG, 2016) and name standardization followed REFLORA (Flora do Brasil, 2020), using the flora package (Carvalho, 2016) implemented in the environment R v. 3.6.1 (2020). Forest inventory data are stored in the ForestPlots.net system (https://www.forestplots.net; codes present in Table S1 [suppl.]) and are available upon request. These data were then used to generate two matrices of species occurrence on the sites, one for each vegetation type (semideciduous and deciduous), which were used in the analyses described below. The final matrices had 664 species occurring in 22 semideciduous forests and 433 species occurring in 22 deciduous forests.

**Association rule analysis**

We explored patterns of species coexistence in each vegetation type (semideciduous and deciduous forest) using association rule analysis (ARA, Agrawal et al., 1993; Silverstein et al., 1998; Ferrarini & Tomaselli 2010; Rossi et al. 2014; Leote et al., 2020). ARA is a data mining tool that identifies associations between categorical observations (in this case, species names) in large data sets. It proposes relationship rules in the template: “if species X, then species Y”. Thus, based on a set of events (often called “transactions”) and the presence of elements in these events, it is possible to relate the existing categories and associate their coexistence with a probability of occurrence and importance (Ferrarini & Tomaselli, 2010; Rossi et al., 2014; Zumel et al., 2019; Leote et al., 2020). In the case of our data, the events (or transactions) are the sites of collection and the set of elements are all tree species present in each one.

Each rule has two parts: the rule left side (RLS) is a unique category level of reference (e.g., one species) that has an occurrence frequency in seasonal forest sites, which is called support; and the rule right side (RRS) that is composed by one or more category levels and that is related to the RLS in a frequency called confidence. Confidence values are always based on the rule support value and inform how frequent is the rule in the dataset (Ferrarini & Tomaselli, 2010; Rossi et al., 2014; Zumel et al., 2019; Leote et al., 2020). Thus, if for a given rule the RLS species occurs in half of the sites (support of 0.5) and the whole rule has confidence of 1.0, the RRS species are associated with the RLS species in all its occurrences, that is, they coexist in half of the sites. The rules obtained present three other measures: lift, coverage and Fisher’s exact test p-value. Lift is a measure of the rule’s importance and is obtained through the relation between the observed rule confidence and the previously expected confidence, in which lift > 1 indicate the rule is more frequent than expected by initial constraints and lift equal to 1 indicates that rule’s species are independent from each other and that the rule is not real. That is, more the value of lift, greater are the chances of having the sp 2 if the site already has the sp1. In addition, in cases where the lift has values less than 1, there is a negative relationship between the species, in which the occurrence of the first can negatively affect the occurrence of the second. Coverage is a measure of the rule’s frequency that evaluate
the support of the species present in the RLS. Fisher’s exact test p-value is the rule’s significance obtained through Fisher’s exact test of contingency for small samples, and evaluate whether the rule is more significantly frequent than expected by chance (Hahsler, 2006). The full process of rules’ obtaining by applying constraints is summarized in Fig. 2.

Based on the occurrence data (presence/absence) of species in sites of each vegetation type (analysis were run separately for semideciduous and deciduous forests), we used the apriori algorithm (Borgelt & Kruse, 2002) from the arules package (Hahsler et al., 2020) in R v. 3.6.1 (2020) to obtain all pairwise association rules that met the following criteria: support equal to or greater than 0.63, which corresponds to the occurrence of the species in RLS in more than 60% of the sites (at least 14 sites of each vegetation type); and confidence of 0.8 for the rule. These criteria indicate that species on RRS must be related to the species on RLS in at least 80% of its occurrences. We focused on pairwise species associations aiming to reveal direct associations between them and analyze their similarities more closely. From the rules obtained, we selected those with significant Fisher’s exact p-value (<0.05), with coverage equal to or greater than 0.63 and lift greater than or equal to 1.1 (more than 10 % frequent than expected). By establishing these criteria, we were able to select strong combinations composed of species of wide occurrence, as well as to select associations of high importance that characterize the vegetation types.

Based on the selected rules, we explored the similarities between the coexisting species in each vegetation type. For that, we evaluated the compatibility between families of the coexisting species in selected rules of each vegetation type, as well as the similarities between them regarding two functional characteristics: potential size (DBH, cm), calculated as the 95th percentile of all DBH measurements of that species in the data set; and wood density (WD, g/cm³), extracted from a global database (Wood Density database; Zanne, 2009) that uses the average species WD value or the genus or family average when species-level data is unavailable. We chose these functional traits because they hint on species ecological behavior: wood density is associated with growth rate, mechanical resistance and hydraulic efficiency (Chave et al., 2009); while potential size is a proxy of resistance to canopy light and ability to reach advanced successional stages in forests (Falster & Westoby, 2005). With the functional data of the species in selected association rules (Table S2 [suppl.]), we thus quantified the difference between the functional traits of the species in each rule (left, RLS and right, RRS), evaluating the pattern of functional similarity between associated species. Graphics were made with the packages arulesViz (Hahsler, 2019) and ggplot (Wickham, 2016) for R program (R Core Team, 2020).

Results

Association rule analysis (ARA) with the established criteria resulted in 238 (out of 928 rules) significant and frequent pairwise species associations in semideciduous forests, and in 11 (out of 62) significant and frequent pairwise species associations in deciduous forests (Tables S3 and S4 [suppl.]). In the semideciduous forests, the pairwise associations are formed by a group of 33 species in 28 genera and 18 families. In the deciduous forests, the pairwise associations are formed by a group of 8 species in 8 genera and 5 families. In semideciduous forests, maximum support was 0.68 and maximum coverage was 0.68; maximum lift was 1.47 for the association between Dendropanax cuneatus (DC.) Decne. & Planch. (Araliaceae) and Annona dolabripetala Raddi (Annonaceae) (Table S3 [suppl.]; Fig 3 - a). In deciduous forests, maximum support was 0.68 and maximum coverage was 0.73; maximum lift was 1.21 for the association between Ptilochaeta bahiensis Turcz. (Malpighiaceae) and Cenostigma phluiosa (DC.) L.P. Queiroz (Fabaceae) (Table S4 [suppl.]).
Tree species consistent co-occurrence through association rules analysis

[Suppl.]; Fig 3 - b). In both semideciduous and deciduous forests, some species pairs were mutually related, with occurrence dependence found in both directions; i.e., the association was consistent when swapping the species in the left and right sides.

Most of the pairs are formed by taxonomically unrelated species, at least as far as the family level (Table S2 [Suppl.]). In the semideciduous forests, only 5.8% of all pairwise associations were between conspecific species. In the deciduous forests, Goniorrhachis marginata Taub. and Machaerium acutifolium Vogel. formed the only conspecific species pair (both belong to Fabaceae) (~ 9% of total; Table S2 [Suppl.]). In the rules selected for semideciduous forests, there are associations between species with different functional characteristics (Fig S1 [Suppl.]), but most associations occur between species with similar potential size (DBH) and wood density (Fig S1 - a [Suppl.]). In deciduous forests the trend seems to be the same, but the low number of rules selected prevents an accurate assessment (Fig S1 - b [Suppl.]). In general, association rules tend to be formed by taxonomically unrelated species that are similar but not functionally equal.

Discussion

Our results showed the existence of consistent pairwise associations between species in semideciduous and deciduous forests, which are characteristic of their communities. However, widespread associations of species of broad occurrence are more frequent and are associated with a greater species number in the semideciduous forest, in comparison with the deciduous forests. In addition, we found that these pairs are mainly formed by non-conspecific species (of different botanical families) with similar functional characteristics.

The widely occurring associations between species in these seasonal forests indicate that a set of broad species has their occurrence dependent on each other in communities, and that their associations are characteristic of these vegetation types and important to their structure and functioning. These co-occurrences are formed by species characteristic of these vegetation types, which have a wide occurrence in these forests conditioned by past evolutionary occupation processes (Oliveira-Filho & Fontes, 2000; Pennington et al., 2009; Santos et al., 2012; Moro et al., 2016; Neves et al., 2017). This result points out that there is a great chance of sampling the presented co-occurrences in tree community randomly sampled in these regions.

The greater number of species and coexistence rules observed in the semideciduous forests compared to the deciduous forests is related to the different environmental conditions between the two vegetation types. Because they are subjected to more restrictive ecological conditions such as high temperatures and a long dry season that explain the greater species deciduousness, often species in deciduous forests tend to occur in environments close
to their survival limit (Pennington et al., 2009; Allen et al., 2017). Thus, small variations in environmental conditions may generate variations in ecological filters and thus select a distinct set of species (Santos et al., 2012; Apgaua et al., 2015; Maia et al., 2020; Souza et al., 2020). In fact, these results reflected the most frequent species in our deciduous forest dataset. It is also important to highlight the different sampling intensity between vegetation types, that probably influenced the result due to a lower number of trees sampled, and consequently a lower chance to identify association. But we consider that the great difference (238 x 11 rules) is strong enough to indicate the trend of more association rules in semideciduous forests.

The lower number of coexistences in deciduous forests can also be explained by the lower participation of interactions in community assembly in restrictive environments according to stress-gradient hypothesis (Holmgren & Scheffer 2010; Kraft et al. 2015; Cadotte & Tucker, 2017). In this perspective, associations thus would have a secondary role in the deciduous forest assembly and function and would be linked mainly to facilitation processes (Holmgren & Scheffer, 2010; Hart & Marshall, 2013; Cadotte & Tucker, 2017; Carrión et al. 2017). It is important to emphasize that adopting broader criteria in the measures of association rule analysis (more permissive values) may allow the achievement of a greater number of association rules, but without considering the assumption of high frequency in the communities.

The result found that most of the significant coexistences are composed of non-confamilial and functionally similar species (especially in semideciduous species where the number of relationships is greater) suggests the existence of niche adaptation processes for the occupation of the same environments. Species of different families with similar ecological requirements may thus have undergone niche differentiation processes after the occupation of habitats, as a way of persistence and avoiding competitive exclusion due to niche overlap (Wright, 2002; Barraclough, 2015; Cadotte & Tucker 2017; Chen et al. 2020). In addition, the coexistence between these species may have been consolidated after processes of modification of their ecological patterns, such as decrease of representativity in abundance and/or biomass (Holmgren & Scheffer, 2010; Hart & Marshall, 2013; Kraft et al., 2015; Hart et al., 2017; Chen et al., 2020). Thus, the species coexistence seasonal tropical forest would be associated with changes in ecological requirements and persistence strategies by species, consequently impacting ecosystem functioning (Wright, 2002; Hart & Marshall, 2013; Kraft et al., 2015).

Here we demonstrate the feasibility of applying the novel association rule analysis (ARA) methodology to ecological studies of tree communities in seasonal tropical forests, but that can be extended to other biological groups. As demonstrated here, the methodology is promising to identify coexistence relationships (both positive and negative) in large ecological data sets, which can also be used to identify relationships with environmental drivers synthesized into categorical variables, although in this case other robust analysis approaches may be more appropriate (Ferrarin & Tomasselli, 2010; Rossi et al., 2014; Leote et al., 2020). The use of ARA approach can also be used to identify indicators species in ecosystems, being an alternative to traditional approaches in situations of extensive datasets (Leote et al., 2020). In vegetation studies we call attention to the need of sampling intensity and for the also need of inclusion criterium standardization, since species richness and composition are strongly influenced by them. Our results indicate that association rules analysis is an interesting alternative to the traditional analysis of ecological data and that it should be incorporated into future studies, joint to other novel data mining and data science tools (e.g., machine learning, neural networks, natural language processing and artificial intelligence) broadly used in non-scientific contexts (Hahsler, 2006; Zumel et al., 2019). All of these approaches are already in consolidated use with positive results in organizations, so that they can also assist in the resolution of relevant ecological issues, in addition to allowing a revisiting of consolidated patterns, as they offer new perspectives in relation to the data. Its use must also be associated with the incorporation of other objects of ecological studies, in order to contribute to the understanding of biodiversity patterns and ecosystem functioning.

Acknowledgments

To Federal University of Lavras (UFLA), State of Minas Gerais Research Foundation (FAPEMIG), National Council for Scientific and Technological Development (CNPq) and to Coordination for the Improvement of Higher Education Personnel (CAPES) for all the support.

References

Agrawall R, Imielinski T, Swami A, 1993. Mining association rules between sets of items in large databases. In: Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data; Buneman P, Jadodia S (eds), pp. 207-216. ACM Press, Washington, USA. https://doi.org/10.1145/170036.170072

Allen K, Dupuy JM, Gei MG, Hulshof C, Medvigy D, Pizano C, Salgado-Begret B, Smith CM, Trierweiler A, Van Bloem S, 2017. Will seasonally dry tropical forests be sensitive or resistant to future changes in rainfall regimes? Environ Res Lett 12(2): 023001. https://doi.org/10.1088/1748-9326/aa5968

Amarasekare P, Hoopes MF, Mouquet N, Holyoak M, 2004 Mechanisms of coexistence in competitive me-
Tree species consistent co-occurrence through association rules analysis

Ferrarinì A, Tomaselli M, 2010. A new approach to the analysis of adjacencies: Potentials for landscape insights. Ecol Model 221 (16): 1889-1896. https://doi.org/10.1016/j.ecolmodel.2010.04.020

Flora do Brasil 2020 em construção, 2020. Jardim Botânico do Rio de Janeiro. Available at: http://floradobrasil.jbrj.gov.br

Hahsler M, 2006. A model-based frequency constraint for mining associations from transaction data. Data Min Knowl Discov 13(2): 137-166. https://doi.org/10.1007/s10618-005-0026-2

Hahsler M, 2019. arulesViz: Visualizing Association Rules and Frequent Itemsets. R package version 1.3-3. https://CRAN.R-project.org/package=arulesViz

Hahsler M, Buchta C, Gruen B, Hornik K, 2020. arules: Mining Association Rules and Frequent Itemsets. R package version 1.6-6, https://CRAN.R-project.org/package=arules

Hart S, Marshall DJ, 2013. Environmental stress, facilitation, competition, and coexistence. Ecology 94: 2719-2731. https://doi.org/10.1890/12-0804.1

Hart SP, Usinowicz J, Levine JM, 2017. The spatial scales of species coexistence. Nat. Ecol. Evol. 1: 1066-1073. https://doi.org/10.1038/s41559-017-0230-7

He T, Lamont BB, Pausas JG, 2019. Fire as a key driver of Earth's biodiversity. Biol Rev Camb Philos Soc 94: 1983-2010. https://doi.org/10.1111/brv.12544

Holmgren M, Scheffer M, 2010. Strong facilitation in mild environments: the stress gradient hypothesis revisited. J Ecol 98(6): 1269-1275. https://doi.org/10.1111/j.1365-2745.2010.01709.x

Kraft NJ, Adler PB, Godoy O, James EC, Fuller S, Levine JM, 2015. Community assembly, coexistence and the environmental filtering metaphor. Funct Ecol 29: 592-599. https://doi.org/10.1111/1365-2435.12345

Jimenez-Rodriguez DL, Alvarez-Añorve MY, Pineda-Cortes M, Flores-Puerto JL, Benítez-Malvido J, Oyama K, Avila-Cabaddilla LD, 2018. Structural and functional traits predict short term response of tropical dry forests to a high intensity hurricane. For Ecol Manage 426: 101-114. https://doi.org/10.1007/j.foforeco.2018.04.009

Leote P, Cajaiba RL, Cabral JA, Bresecovit AD, Santos M, 2020. Are data-mining techniques useful for selecting ecological indicators in biodiverse regions? Bridges between market basket analysis and indicator value analysis from a case study in the neotropics. Ecol Indic 109: 105833. https://doi.org/10.1016/j.ecolind.2019.105833

Maia VA, Souza CR, Aguiar-Campos N, Fagundes NCA, Santos ABM, Paula GGP, Santos PF, Silva WB, Menino GCO, Santos RM, 2020. Interactions between climate and soil shape tree community assembly and above-ground woody biomass of tropical dry forests. For Ecol Manage 474: 118348. https://doi.org/10.1016/j.foreco.2020.118348
Moro MF, Lughadha EM, Araújo FS, Martins FR, 2016. A phytogeographical metaanalysis of the semiarid Caatinga domain in Brazil. Bot Rev 82(2): 91-148. https://doi.org/10.1007/s12229-016-9164-z

Neves DM, Dexter KG, Pennington RT, Valente ASM, Bueno ML, Eisenlohr PV, Fontes MAL, Miranda PL, Moreira SN, Rezende VL, et al., 2017. Dissecting a biodiversity hotspot: the importance of environmentally marginal habitats in the Atlantic Forest Domain of South America. Divers Distrib 23(8): 898-909. https://doi.org/10.1111/ddi.12581

Oliveira-Filho AT, Fontes MAL, 2000. Patterns of floristic differentiation among Atlantic Forests in southeastern Brazil and the influence of climate. Biotropica 32: 793-810. https://doi.org/10.1111/j.1744-7429.2000.tb00619.x

Pausas JG, Pratt RB, Keeley JE, Jacobsen AL, Ramirez AR, Vilagrosa A, Paula S, Kaneakua-Pia IN, Davis SD, 2016. Towards understanding resprouting at the global scale. New Phytol 209: 945-954. https://doi.org/10.1111/nph.13644

Pennington RT, Lavin M, Oliveira-Filho AT, 2009. Woody plant diversity, evolution, and ecology in the tropics: perspectives from seasonally dry tropical forests. Annu Rev Ecol Evol Syst 40: 437-457. https://doi.org/10.1146/annurev.ecolsys.110308.120327

R Core Team, 2020. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna. ISBN 3-900051-07-0. URL https://www.r-project.org/

Rossi G, Ferrarini A, Dowgiarlo G, Carton A, Gentili R, Tomaselli M, 2014. Detecting complex relations among vegetation, soil and geomorphology. An in-depth method applied to a case study in the Apennines (Italy). Ecol Comp 17: 87-98. https://doi.org/10.1016/j.ecocom.2013.11.002

Santos RM, Oliveira-Filho AT, Eisenlohr PV, Queiroz LP, Cardoso DBOS, Rodal MJN, 2012. Identity and relationships of the Arboreal Caatinga among other floristic units of seasonally dry tropical forests (SDTFs) of north-eastern and Central Brazil. Ecol Evol 2: 409-428. https://doi.org/10.1002/ece3.91

Schmid JS, Taubert F, Wiegand T, Sun I, Huth A, 2020. Network science applied to forest megaplots: tropical tree species coexist in small-world networks. Sci Rep 10: 1-10. https://doi.org/10.1038/s41598-020-70052-8

Silverstein C, Brin S, Motwani R, 1998. Beyond market baskets: generalizing association rules to dependence rules. Data Min. Knowl. Disc. 2: 39-68. https://doi.org/10.1023/A:1009713703947

Souza C, Morel JD, Santos ABM, Silva WB, Maia VA, Coelho PA, Rezende VL, Santos RM, 2020. Small-scale edaphic heterogeneity as a floristic-structural complexity driver in Seasonally Dry Tropical Forests tree communities. J For Res (Harbin) 31: 2347-2357. https://doi.org/10.1007/s11676-019-01013-9

Sunderland T, Apgaaua D, Baldauf C, Blackie R, Colfer C, Cunningham AB, Dexter K, Djoudi H, Gautier D, Gumbo D, et al., 2015. Global dry forests: a prologue. Int. For. Rev 17: 1-9. https://doi.org/10.1505/146554815815834813

Tilman D, 1999. The ecological consequences of changes in biodiversity: a search for general principles. Ecology 80: 1455-1474. https://doi.org/10.1890/0012-9658(1999)080[1455:TECOEI]2.0.CO;2

van der Plas F, 2019. Biodiversity and ecosystem functioning in naturally assembled communities. Biol Rev Camb Philos Soc 94: 1220-1245. https://doi.org/10.101111/nph.12499

van der Sande, MT, Poorter, L, Kooistra, L, Balvanera, P, Thonicke, K, Thompson, J, Arets, EJMM, Alainz, NG, Jones, L, Mora, F, et al., 2017. Biodiversity in species, traits, and structure determines carbon stocks and uptake in tropical forests. Biotropica 49: 593-603. https://doi.org/10.1111/btp.12453

Wickham H, 2016. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York. ISBN 978-3-319-24277-4.

Wright SJ, 2002. Plant diversity in tropical forests: a review of mechanisms of species coexistence. Oecologia 130: 1-14. https://doi.org/10.1007/s004420100809

Zanne AE, Lopez-Gonzalez G, Coomes DA, Ilic J, Jansen S, Lewis SL, Miles, RB, Swenson, NG, Wiemann, MC, Chase, J, 2009. Global wood density database. Dryad Digital repository.

Zeppel MJB, Harrison SP, Adams HD, Kelley DI, Li G, Tissue DT, Palmer A, McDowell NG, 2015. Drought and resprouting plants. New Phytol 206: 583-589. https://doi.org/10.1007/nph.13205

Zumel N, Mount J, Porzak J, 2019. Practical data science with R. Shelter Island, NY: Manning, 2019.