**Supplementary Material**

**Limited and socially unequal contributions of rain gardens to native biodiversity in a tropical megacity**

Jonathan Almeida¹,2, Ana Novoa²,3\* and Jean Paul Metzger¹

¹ Ecology Department, Institute of Biosciences, University of São Paulo, S. of Matão, 321 - Butantã, São Paulo - SP, 05508-090, Brazil

² Experimental Station of Arid Zones, Spanish National Research Council (CSIC), Av. Sacramento, s/n, 04120 Almería, Spain

3 Institute of Botany, Department of Invasion Ecology, The Czech Academy of Sciences, CZ19 252 43 Průhonice, Czech Republic

*\*Corresponding author: novoa.perez.ana@gmail.com*

1. **Data collection**

**1.1 Sampling Design**

*1.1.1 Vegetation sampling*

After taxonomic identification, we investigated the biogeographical origin of the species using the databases of Flora do Brazil and Plants of the World. To ensure the standardization of scientific names, we followed the nomenclature of Plants of the World Online. In this study, we considered as native species those with a natural distribution in the state of São Paulo, considering both the Atlantic Forest and Cerrado biomes, which encompass the study area. The classification of exotic and invasive species was based on the Global Register of Introduced and Invasive Species (GRIIS) (Ziller et al. 2020), the Global Naturalized Alien Flora (GloNAF) (van Kleunen et al. 2019).

Finally, we classified species as ornamental or ruderal based on their use according to the Flora do Brasil. We consulted the biodiversity inventory of the municipality of São Paulo and the lists of species cultivated in municipal nurseries, particularly for the assessment of ornamental plants (Secretaria do Verde e do Meio Ambiente, 2016; Divisão de Produção e Herbário Municipal – DPHM, 2024), and specialized websites for ornamental species (Jardineiro.net; Plantei). For ruderal plants, we consulted the list from Embrapa for ruderal classification (Moreira & Bragança 2011).

*1.1.2 Management assessment*

To assess the level of management in the studied areas, we took photographic records and applied a scoring system based on four criteria. The management level was defined as the sum of these scores, ranging from 0 (unmanaged) to 4 (highly managed) (Figure 1):

1. Vegetation delimitation: A score of 1 was assigned to sites with clearly defined vegetation boundaries (e.g., flower beds) and 0 to those without defined boundaries (e.g., vacant lots or squares where vegetation extended onto sidewalks).
2. Presence of ornamental species: A score of 1 was assigned for presence and 0 for absence.
3. Presence of ruderal species: A score of 0 was assigned for presence and 1 for absence, as these species are not intentionally cultivated.
4. Evidence of recent management: Areas showing signs of pruning or landscaping received a score of 1, while unmanaged areas with spontaneous vegetation were scored 0.

Interface gráfica do usuário, Aplicativo

O conteúdo gerado por IA pode estar incorreto.

Figure 1. Examples of rain gardens and adjacent areas with different levels of management. No rain gardens were recorded without spatial delimitation or ornamental plants; therefore, there are no examples with a management level below 2. In this figure, for rain gardens, a score of 2 corresponds to the presence of ruderal species and no evidence of recent management; a score of 3 indicates the presence of ruderal species with evidence of recent management; and a score of 4 represents the absence of ruderal species with evidence of recent management .For adjacent areas, a score of 0 corresponds to plots with no delimitation, the presence of ruderal species, the absence of ornamental species, and no recent management. A score of 1 refers to delimited vegetation with the presence of ruderal species, the absence of ornamental species, and no recent management. A score of 2 indicates delimited vegetation with the presence of both ruderal and ornamental species, but no recent management. A score of 3 represents delimited vegetation with both ruderal and ornamental species and evidence of recent management. A score of 4 corresponds to delimited vegetation with the absence of ruderal species, the presence of ornamental species, and evidence of recent management.

*1.1.3 Environmental, socio-economic and management variables*

Table 1. Variables initially considered to assess the influence of environmental, socioeconomic, and management characteristics on plant diversity in rain gardens and adjacent areas. In the 'Selected' column, 'yes' indicates variables included in the final models.

| **Variables** | **Selected** |
| --- | --- |
| **Environmental** | |
| Vegetation volume (height x area) in sampled buffer | No |
| Total vegetation proportion in sampled buffer | No |
| Vegetation proportion of arboreal low coverage | No |
| Vegetation proportion of arboreal medium coverage | No |
| **Vegetation proportion of herbaceous and shrubby coverage** | **Yes** |
| Water proportion in the sampled buffer | No |
| Shortest distance from the rain garden to the nearest Atlantic Forest fragment defined by conservation plan of São Paulo State | No |
| Area of the nearest Atlantic Forest fragment | No |
| Shortest distance from the rain garden to the nearest environmental park | No |
| **Shortest distance from the rain garden to the nearest Environmental protection area** | **Yes** |
| APA area in square meters | No |
| **Shortest distance from the rain garden to the nearest watercourse** | **Yes** |
| Watercourse perimeter in square meters | No |
| Watercourse area in square kilometers | No |
| Building volume in buffer of 200 m² radius surrounding the garden (sampled buffer) | No |
| **Building proportion in the sampled buffer** | **Yes** |
| **Shortest distance from the rain garden to the nearest large road (arterial, expressway or highway)** | **Yes** |
| **Management** | |
| **Rain garden area** | **Yes** |
| **Rain garden management level** | **Yes** |
| **Adjacent area management level** | **Yes** |
| **Socioeconomic** | |
| Municipal Developing Human Index per district (considering income, education and longevity information) | No |
| **Municipal Developing Human Index focusing on incoming per district – income per capita** | **Yes** |

**2. Data Analysis**

*2.1 Diversity Index combining cover of stoloniferous species and abundance (non-stoloniferous species) data*

Follow below the formula and description used to calculate the diversity indices (alpha and beta) combining cover and abundance data:

Where:

* : Weight of species with abundance data – number of species with abundance data
* : Weight of species with cover data – number of species with cover data
* ​: Diversity index calculated using abundance data
* ​: Diversity index calculated using cover data

*2.2 Statistical families of each response variable*

To identify the most appropriate probability distribution for each response variable, we tested candidate distributions based on theoretical statistical assumptions using the fitdist function from the fitdistrplus package (Delignette-Muller, 2015).

Table 2. Statistical distribution families of response variables, corresponding R package family names, and packages used for analysis.

| **Variables** | **Family (package family code)** | **Package** |
| --- | --- | --- |
| General richness inside | Poisson (poisson) | lme4 |
| General richness outside |
| General simpson index inside | Normal (gaussian) |
| General simpson index outside |
| General shannon index inside |
| General shannon index outside |
| Native shannon index outside |
| Exotic shannon index inside |
| Exotic shannon index outside |
| Exotic simpson index outside |
| General Bray-Curtis index | Beta (BE) | gamlss |
| Exotic simpson index inside |
| Native simpson index outside |
| Native simpson index inside | Beta Inflated (BEINF) |
| Exotic Bray-Curtis index |
| Native Bray-Curtis index |
| Native shannon index inside | Zero Adjusted Gamma (ZAGA) |
| Exotic richness inside | Zero Inflated Negative Binomial (NBI) |
| Exotic richness outside |
| Native richness outside |
| Native richness inside |

*2.3 Exploratory models to select most frequent explanatory variables and models diagnosis*

Given the numerous explanatory variables with potential effects on the diversity of rain gardens, we conducted an initial selection to identify the most relevant variables (Table 1). We built linear models considering diversity indices (richness, Simpson, Bray-Curtis) as response variables and tested combinations of up to three environmental, socioeconomic, and management variables as predictors. To avoid collinearity, we restricted combinations to variables with correlations below 0.6. We selected the most frequently occurring variables in the initial models. A cutoff criterion was applied to exclude variables whose occurrence frequency in the models was less than half the frequency of any other variable. The application of this criterion resulted in the selection of nine explanatory variables, corresponding to the set of variables most frequently included in the models for each response variable. (Table 1).

We evaluated model fit by examining residual diagnostics using the simulateResiduals function from the DHARMa package (Hartig, 2024) for models fitted with the lme4 package. Since DHARMa does not fully support models fitted with gamlss, we alternatively assessed the normality of residuals using the Shapiro–Wilk test (shapiro.test function from the stats package; R Core Team, 2023) and tested for heteroskedasticity using the Breusch–Pagan test (bptest function from the lmtest package; Zeileis & Hothorn, 2002). We also used the qqnorm, qqline, and qqplot functions from the stats package to visually inspect the residual distribution.

**3. References**

Delignette-Muller ML, Dutang C (2015). “fitdistrplus: An R Package for Fitting Distributions.” Journal of Statistical Software, 64(4), 1–34. doi:10.18637/jss.v064.i04.

Hartig F (2024). DHARMa: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models. R package version 0.4.7, <https://CRAN.R-project.org/package=DHARMa>.

Jardineiro.net. (s.d.). *Tudo sobre jardinagem*. Acesso em 02 de abril de 2025. <https://www.jardineiro.net>

Moreira, H. J. C., & Bragança, H. B. N. (2011). Manual de identificação de plantas infestantes: hortifrúti. São Paulo: FMC Agricultural Products.

Plantei. (s.d.). *Dicas de jardinagem e cultivo*. Acesso em 02 de abril de 2025. <https://www.plantei.com.br>

Prefeitura de São Paulo. (2016). Inventário da Biodiversidade do Município de São Paulo. Secretaria do Verde e do Meio Ambiente.

Prefeitura de São Paulo. (2024). VIVEIRO ARTHUR ETZEL - DPHM 2. Secretaria do Verde e do Meio Ambiente. Divisão de Produção e Herbário Municipal.

Prefeitura de São Paulo. (2024). VIVEIRO MANEQUINHO LOPES - DPHM 1. Secretaria do Verde e do Meio Ambiente. Divisão de Produção e Herbário Municipal.

R Core Team (2023). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.

van Kleunen M, Pyšek P, Dawson W, Essl F, Kreft H, Pergl J, Weigelt P, Stein A, Dullinger S, König C, Lenzner B, Maurel N, Moser D, Seebens H, Kartesz J, Nishino M, Aleksanyan A, Ansong M, Antonova LA, Barcelona JF, Breckle SW, Brundu G, Cabezas FJ, Cárdenas D, Cárdenas-Toro J, Castaño N, Chacón E, Chatelain C, Conn B, de Sá Dechoum M, Dufour-Dror JM, Ebel AL, Figueiredo E, Fragman-Sapir O, Fuentes N, Groom QJ, Henderson L, Inderjit, Jogan N, Krestov P, Kupriyanov A, Masciadri S, Meerman J, Morozova O, Nickrent D, Nowak A, Patzelt A, Pelser PB, Shu W sheng, Thomas J, Uludag A, Velayos M, Verkhosina A, Villaseñor JL, Weber E, Wieringa JJ, Yazlık A, Zeddam A, Zykova E, Winter M (2019) The Global Naturalized Alien Flora (GloNAF) database. Ecology 100: 1–2. <https://doi.org/10.1002/ecy.2542>

Zeileis A and Hothorn T (2002). Diagnostic Checking in Regression Relationships. R News 2(3), 7-10. URL https://CRAN.R-project.org/doc/Rnews/

Ziller S, Zenni R, Souza Bastos L, Possato Rossi V, Wong L J, Pagad S (2020). Global Register of Introduced and Invasive Species - Brazil. Version 1.5. Invasive Species Specialist Group ISSG. Checklist dataset https://doi.org/10.15468/i0avrm accessed via GBIF.org on 2025-03-05.