

Citizen science data on urban forageable plants: a case study in Brazil

Filipi Miranda Soares^{1,2,*}, Luís Ferreira Pires², Maria Carolina Garcia³, Lidio Coradin⁴, Natalia Pirani Ghilardi-Lopes⁵, Rubens Rangel Silva⁶, Aline Martins de Carvalho⁷, Anand Gavai⁸, Yamine Bouzembrak⁹, Benildes Coura Moreira dos Santos Maculan^{10,11}, Sheina Koffler¹², Uiara Bandineli Montedo¹, Debora Pignatari Drucker¹³, Raquel Santiago¹⁴, Maria Clara Peres de Carvalho¹⁵, Ana Carolina da Silva Lima¹¹, Hillary Dandara Elias Gabriel¹, Stephanie Gabriele Mendonça de França¹, Karoline Reis de Almeida¹, Bárbara Junqueira dos Santos¹⁶ and Antonio Mauro Saraiva¹

- 1 Escola Politécnica, Universidade de São Paulo, São Paulo, SP, 05508-010, Brazil
- 2 Faculty of Electrical Engineering, Mathematics and Computer Science, University of Twente, Enschede, 7522 NB, Netherlands
- 3 Programa de Pós Graduação em Arquitetura, Urbanismo e Design, Centro Universitário Belas Artes de São Paulo, São Paulo, SP, 04018-010, Brazil
- 4 Plants for the Future Project, Brasília, DF, 70772-090, Brazil
- 5 Centro de Ciências Naturais e Humanas, Universidade Federal do ABC, São Bernardo do Campo, SP, 09606-045, Brazil
- 6 Centro Universitário Una, Belo Horizonte, MG, 30160-011, Brazil
- 7 Departamento de Nutrição, Faculdade de Saúde Pública, Universidade de São Paulo, São Paulo, SP, 01246-904, Brazil
- 8 Faculty of Behavioural, Management and Social Sciences (BMS), Industrial Engineering & Business Information Systems (IEBIS), University of Twente, Enschede, 7522 NB, Netherlands
- 9 Information Technology Group, Wageningen University and Research, Wageningen, 6700 HB, Netherlands
- 10 Programa de Pós-Graduação em Gestão & Organização do Conhecimento, Universidade Federal de Minas Gerais, Belo Horizonte, MG, 31270-901, Brazil
- 11 Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, MG, 31270-901, Brazil
- 12 Instituto de Estudos Avançados, Universidade de São Paulo, São Paulo, SP, 05508-060, Brazil
- 13 Embrapa Agricultura Digital, Campinas, SP, 13083-886, Brazil
- 14 Faculdade de Nutrição, Universidade Federal de Goiás, Goiânia, GO, 74605-080, Brazil
- 15 Ciências e Humanidades, Universidade de São Paulo, Escola de Artes, São Paulo, SP, 03828-000, Brazil
- 16 Instituto de Pesquisas Energéticas e Nucleares, Universidade de São Paulo, São Paulo, SP, 05508-000, Brazil

Submitted: 28 September 2023
Accepted: 08 February 2024
Published: 21 February 2024

* Corresponding author. E-mail: filipisoares@usp.br; f.mirandasoares@utwente.nl

Published by GigaScience Press.

Preprint submitted at <https://doi.org/10.1101/2024.01.22.575882>

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Gigabyte, 2024, 1–16

ABSTRACT

This paper presents two key data sets derived from the *Pomar Urbano* project. The first data set is a comprehensive catalog of edible fruit-bearing plant species, native or introduced to Brazil. The second data set, sourced from the iNaturalist platform, tracks the distribution and monitoring of these plants within urban landscapes across Brazil. The study includes data from the capitals of all 27 federative units of Brazil, focusing on the ten cities that contributed the most observations as of August 2023. The research emphasizes the significance of citizen science in urban biodiversity monitoring and its potential to contribute to various fields, including food and nutrition, creative industry, study of plant phenology, and machine learning applications.

We expect the data sets presented in this paper to serve as resources for further studies in urban foraging, food security, cultural ecosystem services, and environmental sustainability.

Subjects Animal and Plant Sciences, Botany, Ecology

DATA DESCRIPTION

This paper introduces two data sets related to Brazilian fruit-bearing plants. Data Set 1 presents a taxonomic list of these plant species, both native and introduced, which includes comprehensive scientific names of plants with human-edible fruits. This data set contains Scientific names (at species and family levels), Vernacular names in Portuguese, Establishment means (native or introduced), Higher geography (geographic distribution of the species in Brazil), and the last update of a given species name within the data set. A more detailed data set documentation is available on Zenodo [1].

Data Set 2, hosted on iNaturalist, compiles observations from the capitals of all 27 Brazilian federative units and specifically focuses on the species listed in Data Set 1. A backup of this data set was obtained from iNaturalist on August 22nd, 2023. The dataset features 47 columns, capturing details such as observation time, location, license, and taxonomic identification [2]. It provides an extensive taxonomic breakdown, covering classifications from kingdom and phylum down to species, subspecies, and variety (in some cases). Notably, iNaturalist allows the export of data sets annotated with Darwin Core metadata [3], ensuring data interoperability with numerous applications that accommodate data standardized with the Darwin Core standard, such as the Global Biodiversity Information Facility (GBIF).

Interoperability with GBIF is vital for several reasons. It improves the visibility and accessibility of biodiversity data worldwide, making it easier for researchers and policymakers to access this information [4]. GBIF is considered “the most extensive resource that provides access to open, integrated data on species occurrence” [5, p. 1]. This is especially important for supporting global collaborations in conservation and biodiversity research [4–6]. GBIF also ensures that data reporting is standardized and of high quality, which is crucial for reliable biodiversity studies [4, 5]. Finally, it enables more effective comparisons and cross-referencing of biodiversity data from different regions and species, enriching biodiversity research on species distribution [5].

CONTEXT

Urban food forests

According to the United Nations Environment Programme [7], Brazil boasts the world’s richest biodiversity, accounting for 15 to 20% of the global biodiversity [7, 8]. This vast biodiversity translates into a diverse array of foods available to Brazilians. Yet, as Gomes [9] underscores, this rich biodiversity is not adequately reflected in the diets of various social groups in the country. In fact, a mere 1.3% of Brazilians actively consume biodiverse foods. Geographic location, ethnicity, age, food insecurity, gender, and educational level influence these consumption patterns [9].

The unique combination of Brazil’s tropical climate and diverse ecosystems creates favorable conditions for cultivating a wide variety of fruits, many of which are native

species. These fruits are not only grown in agricultural settings but also are widely used in urban afforestation and contribute to the aesthetics of parks, squares, and sidewalks in numerous cities.

Fruit-bearing plants in urban areas, also known as urban food forests [10], may offer multiple benefits, such as enhancing food availability, particularly in areas with limited access to fresh fruit, thereby addressing food scarcity issues and improving nutritional security [11, 12]. Furthermore, they promote human health and well-being by fostering a connection with nature, providing opportunities for physical activity, and improving air quality through carbon sequestration and pollution mitigation [13].

Urban food forests also contribute to preserving and promoting biodiversity, ensuring the conservation of native fruit varieties, supporting local ecosystems, providing shelter, food, and nesting sites, and improving environmental conditions [11, 14, 15].

As explored by Hurley *et al.* [16], these green spaces contribute significantly to other non-food-related applications. Hurley and Emery [17] delved into the intricate relationship between urban forests and their utility for food, medicine, craft, and other purposes, focusing on woody species. They highlighted the material connections and diverse uses in artistic and cultural contexts of these green spaces, underscoring their multifaceted importance in urban settings [17].

By acknowledging the potential of edible green infrastructures, urban areas can mitigate the adverse impacts of urbanization while simultaneously enhancing the sustainability and resilience of cities [11, 12].

Use of citizen science for urban biodiversity monitoring

Citizen Science (CS) is a collaborative endeavor between scientists and the public that involves the active participation of members of society in scientific research. Participating in a CS initiative to collect data on fruit-bearing plants locally can be motivated by various reasons, including curiosity and people's desire to engage in these initiatives.

CS platforms offer a conducive environment for citizen scientists to contribute to science and learn. iNaturalist, for instance, serves as an online social network that supports a community of users that exchange biodiversity information, fostering collaborative learning about the natural world [18]. It also offers both a mobile and a web-based platform that allows users to upload photos to be analyzed by Artificial Intelligence, which provides organism identification based on these photos and other variables, such as the location of the observation.

The wealth of data available on iNaturalist can be leveraged for many research initiatives that center around monitoring urban biodiversity. Some notable studies have utilized iNaturalist to track and understand the presence of alien species. For example, in the Tyumen Region, Russia, researchers [19] compared data from the entire region with the Tyumen Urban Area to assess the distribution of four alien plant species. Additionally, in Kyiv city, the expansion of the invasive Balkan slug (*Tandonia kusceri*) in urban landscapes was studied [20].

Insects (class Insecta) have also been a focal point of research in urban environments using iNaturalist. Studies have explored various aspects of insect biodiversity, such as the species richness of butterflies (Lepidoptera: Rhopalocera) in Los Angeles [21], the compilation of a species checklist of hawkmoths (Lepidoptera) in Baguio City, Philippines [22], and the monitoring of Eastern Carpenter Bees (*Xylocopa virginica*) in urban areas of

the United States and Canada to detect wing pigment loss [23]. CS projects using iNaturalist have also targeted other animal groups, including mammals, birds, amphibians, and reptiles. For instance, studies monitored the presence of large mammals in urban centers in North America during the COVID-19 pandemic [24]. Urban red foxes and coyotes were monitored using iNaturalist observations in North America [25, 26]. Bird sightings in the city of Chicago, Illinois, were documented through iNaturalist [27]. Moreover, research characterized amphibians' and reptile species' richness in ecological recreational parks in Mérida, Yucatán, Mexico [28], and in the Edith L. Moore Nature Sanctuary in Houston, Texas [29].

Several studies with broader objectives have also leveraged iNaturalist data to understand ecological interactions in cities [30–32], monitor urban biodiversity [33, 34], explore human use of urban green areas in Europe [35], and examine insect management practices for urban farming [36].

This range of research studies demonstrates iNaturalist's versatility and potential as a tool for exploring urban biodiversity and ecological dynamics. It also indicates the platform's potential application in creating open-source data bases on forageable plants, as explored by [37].

METHODS

Data set 1

Version 3 of the list of native and exotic fruit-bearing plant species cultivated in Brazil was built upon two main sources based on specialists' recommendations: the book *Brazilian Fruits and Cultivated Exotics* [38] and the book series *Plants for the Future*. The latter includes volumes specifically dedicated to species of economic value from four regions of Brazil, namely the South [39], Midwest [40], Northeast [41], and North [8]. A botanist carefully reviewed this list in the following steps:

- (i) Scientific Name Verification. The primary task involved checking each scientific name to ensure the species name's existence, correctness, and validity. The review focused on determining if the species' names were accepted or synonyms at the current time. Key resources for this verification included the Flora and Fungi of Brazil catalog [42], [Tropicos.org](https://www.tropicos.org) by the Missouri Botanical Garden [43], the International Plant Names Index [44], Kew Botanic Gardens Plants of the World Online [45], and the New York Botanical Garden Vascular Plant Types Catalog [46]. It is important to note that scientific names are dynamic; a species considered valid today might become a synonym tomorrow, and vice versa. Thus, we will continuously update this list.
- (ii) Origin Check. A thorough check was conducted for each species to determine whether it was native or introduced. This involved extensive online research and consultation of relevant literature, primarily focusing on the *Plants for the Future* book series [8, 39–41].
- (iii) Two new columns were added to Version 3 of the list: Botanical Family and Geographical Distribution. Information for these columns was primarily sourced from the Flora and Fungi of Brazil catalog and [Tropicos.org](https://www.tropicos.org).
- (iv) Handling Synonyms. Many species included in Version 2 of the list from the reference books were identified as synonyms according to the consulted online catalogs. In such cases, the currently valid scientific names replaced the synonyms.
- (v) Removed species. In Version 3 of the list, some plant species from Version 2 were excluded due to their lack of importance as food sources.

(vi) Inclusion of Additional Species. Some fruit-bearing species listed in the *Plants for the Future* books [8, 39–41] that were missing in Version 2 were included in Version 3.

Data set 2

iNaturalist provides tools for managing observations of interest, known as Projects [47]. We used two types of projects:

- Collection Projects function as a saved database search with enhanced features for data visualization, such as a distinctive banner, icon, URL, and a journal that enables communication with project followers [47]. When creating a Collection project, specific requirements can be set, such as “taxa, place(s), users, dates, and quality grade”. Each time the project’s page is accessed, “iNaturalist will perform a quick search and display all observations that match the project’s requirements” [47].
- Umbrella Projects are designed to compile, compare, or promote multiple collection projects [47].

The criteria for the *Pomar Urbano* collection projects were established as follows: user-registered observations must be about fruit-bearing plant species listed in the taxonomic compilation of Data Set 1. Furthermore, each observation should originate from one of the capitals of Brazil’s 27 federative units, such as São Paulo, Rio de Janeiro, or Brasília. To compile all this data, the umbrella project has been set up to aggregate data from all 27 corresponding collection projects [48].

DATA SAMPLE

Version 3 of Data Set 1 counts 429 fruit-bearing species. Notably, 398 of these species have been observed by iNaturalist users in various locations across Brazil as of 22nd August 2023. *Pomar Urbano* has amassed a collection of over 10,943 observations throughout Brazil, referred to in this paper as Data Set 2. This data set was generated both from opportunistic records on iNaturalist, without the use of active engagement strategies by the project coordinators, and biological surveying (bioblitz).

Figure 1 shows the top 10 cities with the highest number of observations. This ranking spans all five macro-regions of Brazil. Brasília (Midwest region) leads this ranking, followed by São Paulo and Rio de Janeiro (both in the Southeast region).

Figure 2 depicts some observations from Brasília. Among the ten most frequently observed species, seven are native to Brazil, namely *Caryocar brasiliense*, *Eugenia dysenterica*, *Psidium guajava*, *Eugenia uniflora*, *Bixa orellana*, *Hymenaea courbaril*, and *Anacardium humile*. The species with the highest number of observations is *Caryocar brasiliense*, which is endemic to the Brazilian Cerrado biome, encompassing the entire territory of Brasília [49]. This particular species holds significant cultural importance in Midwest Brazilian Cuisine according to [50], potentially accounting for its widespread presence and dissemination throughout Brasília.

The city of São Paulo presents a distinct scenario (see Figure 3). Among the most observed species in the city, only four of them (*Eugenia uniflora*, *Syagrus romanzoffiana*, *Plinia cauliflora*, and *Psidium guajava*) are native to Brazil. This highlights a significant difference in species composition compared to Brasília.

Furthermore, there were notable variations when considering the species that garnered the highest number of observations in each city. Only *Psidium guajava*, *Eugenia uniflora*,

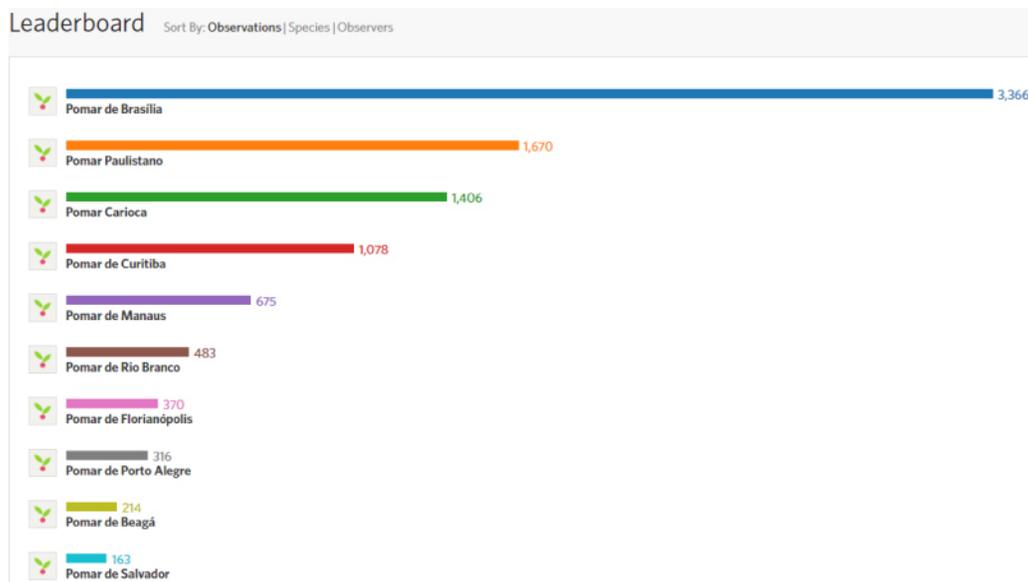


Figure 1. Ten Brazilian cities with the highest number of observations of fruit-bearing species.
 Note: Screenshot from the “Pomar Urbano” project on iNaturalist, based on observations collected as of August 2023. This image represents the data available at that time and may not reflect the most recent updates or additions. To view the most up-to-date data and additional contributions, visit the project online at: <https://www.inaturalist.org/projects/pomar-urbano>.

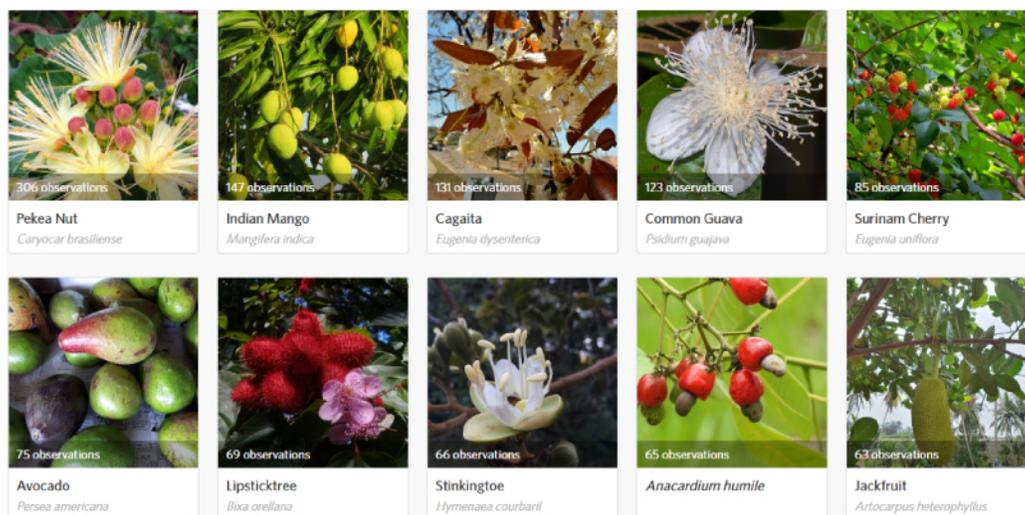


Figure 2. Ten most observed species in Brasília.
 Note: Screenshot from the “Pomar de Brasília” project on iNaturalist, based on observations collected as of August 2023. This image represents the data available at that time and may not reflect the most recent updates or additions. To view the most up-to-date data and additional contributions, visit the project online at: <https://www.inaturalist.org/projects/pomar-de-brasilia>.

and *Persea americana* emerged within the most observed species in both Brasília and São Paulo, since these cities are situated in regions characterized by distinct biomes, differences in species diversity between these locations were expected. The contrasting

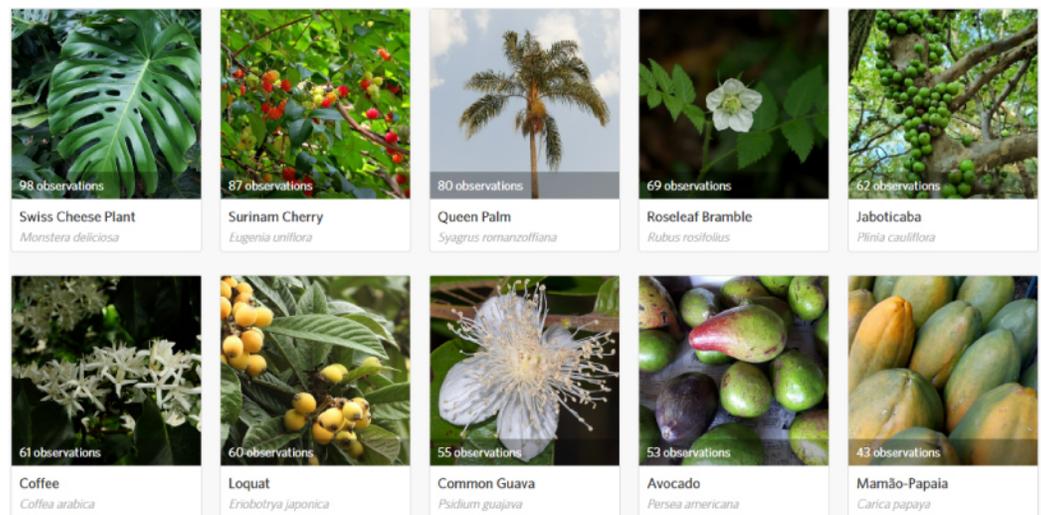


Figure 3. Ten most observed species in São Paulo.

Note 1: Screenshot from the “Pomar Paulistano” project on iNaturalist, based on observations collected as of August 2023. This image represents the data available at that time and may not reflect the most recent updates or additions. To view the most up-to-date data and additional contributions, visit the project online at: <https://www.inaturalist.org/projects/pomar-paulistano>. Note 2: Recent taxonomic advancements have led to the synonymization of *Plinia cauliflora* with *Plinia peruviana*, with the latter now being the preferred nomenclature, as detailed in the Flora and Funga of Brazil [42]. However, at the time of this article’s publication, this taxonomic update had not been reflected on iNaturalist, resulting in the figure retaining the previous name, *Plinia cauliflora*.

biomes contribute to the variation in the composition and distribution of fruit-bearing species, highlighting the significance of considering regional ecological factors when analyzing biodiversity patterns.

In Rio de Janeiro (see Figure 4), the abundance of native fruit-bearing species is also evident. Among the ten most observed species, a noteworthy seven of them are native to Brazil, including *Eugenia uniflora*, *Syagrus romanzoffiana*, *Lecythis pisonis*, *Myrciaria glazioviana*, *Pachira aquatica*, *Opuntia monacantha*, and *Eugenia brasiliensis*.

Both *Eugenia uniflora* and *Syagrus romanzoffiana* emerge as prominent species within the top ten observations across all three cities. The remarkable prevalence of *Eugenia uniflora* can be attributed to its adaptability to diverse climate and soil conditions [51]. This adaptability, coupled with the increasing economic interest in the berries produced by this plant, has led to its cultivation in various regions worldwide, not just in Brazil [51].

The significant presence of *Syagrus romanzoffiana* in the three cities can be attributed to its adaptability to the biomes prevalent in these regions, including the Atlantic Forest and Cerrado [52]. This species’ fruit pulp and nut can be consumed by humans and also holds great value for various groups of animals, indicating its ecological importance in providing food resources [52, 53]. In addition, native communities believe that this palm species is a sacred plant, attributed with cosmogonic and cosmological spiritual significance [53]. Furthermore, recent studies have shed light on the medicinal potential of *Syagrus romanzoffiana*, demonstrating its effectiveness in treating infections and chronic diseases [54]. Additionally, the seeds of this plant have been found to possess a high concentration of oil, making them a promising candidate for biodiesel production [55].

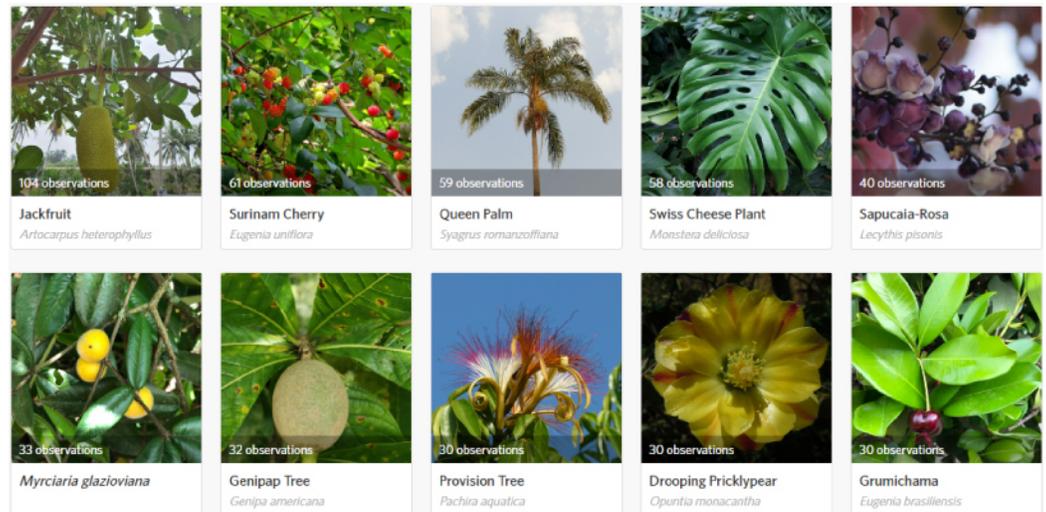


Figure 4. Ten most observed species in Rio de Janeiro.

Note: Screenshot from the “*Pomar Carioca*” project on iNaturalist, based on observations collected as of August 2023. This image represents the data available at that time and may not reflect the most recent updates or additions. To view the most up-to-date data and additional contributions, visit the project online at: <https://www.inaturalist.org/projects/pomar-carioca>.

DATA VALIDATION AND QUALITY CONTROL

Species identification on iNaturalist leverages a Computer Vision Model (CVM) to process observation images and suggest potential species names [56–58]. Although the CVM lacks data on all 2 million species registered worldwide, it successfully identifies about 85% of observations [57, 58]. For species not recognized by the CVM, it recommends higher taxonomic classifications like genus or family [58].

A key strength of iNaturalist is its community of specialists, known as identifiers, who significantly enhance the platform’s species identification accuracy [56]. These identifiers also contribute to the CVM by providing quality data and validating observations’ identifications [56]. Their expertise is particularly valuable in distinguishing between similar species, thereby refining the CVM and improving its suggestions over time [56, 58].

The ongoing contributions of these identifiers are vital for the CVM’s continuous improvement. By providing a high-quality dataset and expanding the model’s species coverage, they significantly boost its overall performance. For more insights on the impact of identifiers on iNaturalist data quality, refer to [56].

iNaturalist enforces data quality criteria, classifying records into various levels. The highest is “Research Grade” (RG), characterized by specific criteria including having a valid date, location, photo or sound, being of a non-captive or non-cultivated organism, and community consensus on the organism identification at the species level or lower [59]. GBIF indexes RG observations, making them accessible for download [60].

In the *Pomar Urbano* project, prioritizing RG data from iNaturalist would be logical due to the necessity for accurate plant species identification, especially when health risks from non-edible fruits are concerned. However, a significant challenge arises since the project focuses on forageable plants, many of which are cultivated and thus are not eligible for RG status on iNaturalist.

This categorization is particularly complex in urban settings where the line between cultivated and wild plants can blur. For instance, a plant initially cultivated by humans might become naturalized in an urban environment, raising questions about its categorization. This ambiguity is evident in discussions within the iNaturalist community, where there is an ongoing debate about how to handle observations of plants that might have been planted initially but have become self-sustaining and naturalized in an urban context. For further insight into the complexities surrounding the classification of plants as “cultivated” in urban contexts, refer to the iNaturalist forum discussion [61].

Given these nuances, relying exclusively on RG data in the *Pomar Urbano* project could be limiting. The project’s focus on urban environments means many valuable observations might be of cultivated plants. Therefore, while RG observations are preferred for their verified accuracy, the project would benefit from a more inclusive approach, considering data also from non-RG observations. Still, we believe this data set could also be relevant to GBIF and other biodiversity data platforms, considering the potential interest in data that sheds light on the distribution and characteristics of economically important plant species within urban ecosystems.

REUSE POTENTIAL

The data sets presented in this paper provide a foundational platform for understanding the diversity of fruit-bearing plants in select Brazilian cities. We assert that this contributes to various open research questions in the existing literature on urban foraging and ecosystem services in urban environments, including, but not limited to:

- Examining how the current species composition in cities forms pre-existing landscapes that are both edible and useful, as explored in the works of [16, 17, 62];
- Developing open data sources focused on Urban Food Forests, as highlighted by [37];
- Investigating the role of urban forests in supporting cultural ecosystem services, as discussed by [63];
- Analyzing urban foraging as a means to enhance food security [64–69];
- Facilitating the sharing of ecological knowledge, as mentioned in [70];
- Contributing to social-ecological resilience, as studied by [71];
- Aiding in reconnecting urban residents with nature and biodiversity, as described by [72–74].

In the context of the *Pomar Urbano* project, data reuse includes:

Food and nutrition

To engage and inspire individuals to explore the culinary potential of locally available fruit, we partnered up with the food laboratory *Sustentarea* at the University of São Paulo. In this facility, we document food recipes showcasing the most frequently encountered fruit in São Paulo, Brazil. Additionally, we have conducted parallel experiments in Brasília. These recipes will be shared across social media platforms to encourage people to prepare easy, affordable, tasty dishes using these regional ingredients. Figure 5 shows the results of some of these experiments.

As our initiative progresses, we envision expanding our recipe collection to include fruit from other cities. The book *Biodiversidade Brasileira: sabores e aromas* (Brazilian Biodiversity: flavors and aromas) [75] is also being used as reference for preparing Brazilian recipes using fruit.

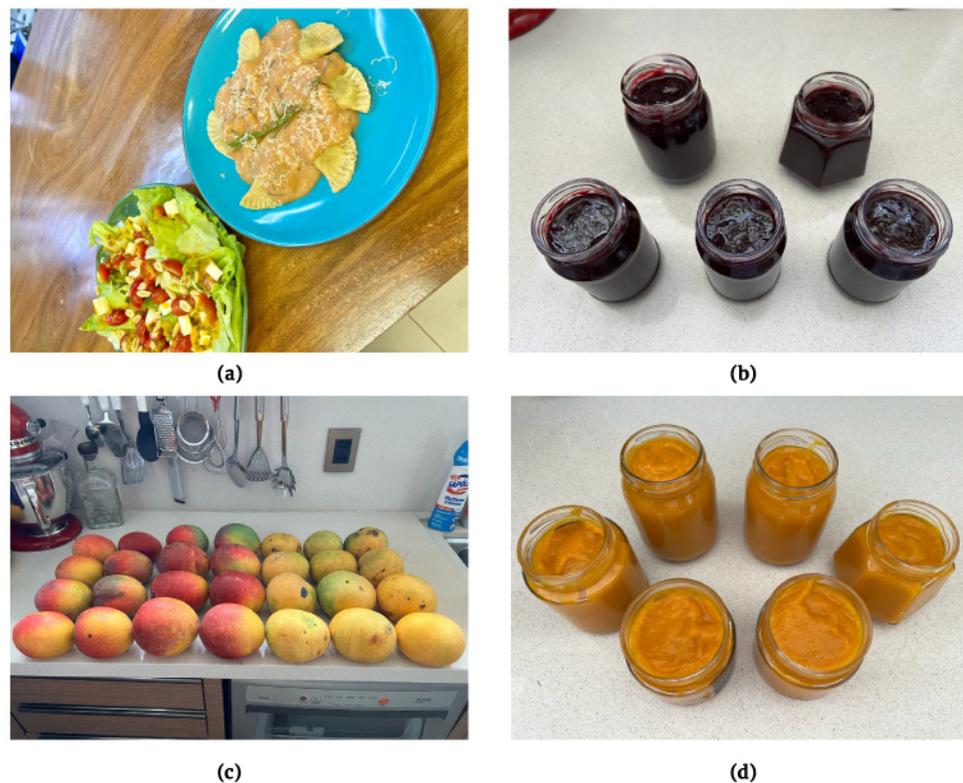


Figure 5. Recipes prepared for *Pomar Urbano*.

(a) Green salad with passion fruit (*Passiflora sp.*) sauce and a taioba (*Xanthosoma taioba*) tortelli with guava (*Psidium guajava*) sauce. Recipe production: Ana Maria Bertolini, Gabriela Rigote, Natalie Marinho Dantas. Photo: Gabriela Rigote [48]. (b) Jaboticaba (*Plinia peruviana*) jam, prepared in December 2023 in Brasília, Brazil. Recipe preparation and Photo: Lidio Coradin and Vera Coradin. (c) Freshly foraged mangoes, Brasília, Brazil, December 2023. This photograph shows ripe mangoes (*Mangifera indica*), handpicked during a foraging expedition. Photo: Lidio Coradin. (d) Mango chutney prepared by Lidio Coradin and Vera Coradin with foraged mango in Brasília, Brazil, December 2023. Photo: Lidio Coradin.

Creative industry

The creative industry, also known as creative economy, encompasses a diverse range of sectors dedicated to generating and commercializing creativity, art, and original content. This industry spans advertising, architecture, arts and crafts, design (including fashion), film, music, performing arts, publishing, and software (like video games) [76]. It thrives on innovation and caters to cultural and economic demands, often blurring the lines between art and commerce. The creative industry may benefit from our data in many ways.

We are identifying Brazilian digital influencers on prominent platforms such as TikTok and Instagram, focusing on those who delve into topics like sustainable diets, vegetarianism, science, nature, and related areas. We aim to collaborate with these influencers, seeking their endorsement to introduce our initiative to their followers. By leveraging data from our project, these influencers can generate content while we gain the advantage of increased user participation in our data collection efforts.

Beyond content creation, information on fruit-bearing plants can inspire the development of diverse products. Figure 6 showcases products crafted by Brazilian artists inspired by observations recorded on *Pomar Urbano*.



Figure 6. Product designs inspired by Pomar Urbano.

(a) A porcelain cup featuring a print inspired by the Night Blooming Cactus flower (*Epiphyllum oxypetalum*), observed during a research survey in São Paulo, Brazil. The design was created by Fashion Design students Kelly Cristina Soares Barbieri, Larissa Galdino de Souza Costa, and Karollina Brandão Araújo Cosso at Centro Universitário Belas Artes de São Paulo, supervised by Maria Carolina Garcia [48]. (b) Print for a tablecloth inspired by the guava tree flower (*Psidium guajava*), using the traditional technique of Brazilian Chita. Created by Luciana Mendonça, a student of Interior Design at the Centro Universitário Belas Artes de São Paulo under the supervision of Maria Carolina Garcia [48]. (c) Digital illustrations showcasing a diverse array of Brazilian plants (from the up left corner to the right down corner: *Annona mucosa*, *Mauritia flexuosa*, *Campomanesia phaea*, *Theobroma grandiflorum*, *Campomanesia xanthocarpa*, and *Plinia peruviana*). Designed for the Pomar Urbano project based on multiple photos retrieved from [iNaturalist.org](https://www.inaturalist.org). Illustrations: Bruna Stefani Perin and Fernanda Beatriz Fernandez Correa. Supervision: Rubens Rangel Silva (Ánima Educação).

Understanding plant phenology

Numerous reference sources offer information on the timing of fruiting and flowering for various species in Brazil. Notably, the work conducted by Coradin, Camillo, and Vieira [8] stands out, presenting an extensive table that details the aforementioned plant phenology aspects, including their respective peaks, for species found in the North of Brazil. Drawing from these existing literature resources, we anticipate harnessing the provided data to support our future analyses concerning the fruiting seasons of each species.

Our primary objective is to compare the data collected by citizen scientists with the findings presented by reference sources. Through this comparative analysis, we aim to ascertain whether the information gathered through CS initiatives yields similar outcomes to those achieved by professional scientists. This investigation aims to assess the accuracy and reliability of CS data collection within the specific framework of our project.

By systematically examining and contrasting these results, we can shed light on the efficacy and potential of CS as a means of data collection for such projects. This comparison will serve as a crucial step towards enhancing our understanding of seasonal fruit patterns while also showcasing the contributions that citizen scientists can make in advancing scientific knowledge. Furthermore, this data can potentially highlight temporal shifts in the reproductive seasonality of the monitored plant species in the long term.

Machine learning applications

The safety and authenticity of food can be safeguarded through data derived from fruit-bearing plants, serving as a foundational reference for genuine fruit attributes. Machine Learning (ML) algorithms can then be employed to pinpoint irregularities in fruit quality or composition, thereby enabling the detection of deceptive practices and potential food safety hazards [77]. Furthermore, within urban settings, the local production of fruits and vegetables can be seamlessly integrated into the urban supply chain to enhance supply chain management. ML algorithms can optimize transportation logistics and minimize food wastage, leading to a more sustainable and efficient urban food system. Moreover, in the face of environmental challenges, ML offers a powerful tool to analyze long-term data from urban fruit crops. This data-driven approach can provide valuable insights into understanding the impact of climate change and urbanization on plant growth and biodiversity. Such knowledge can be pivotal in guiding urban planning decisions and conservation initiatives.

In conclusion, when coupled with ML capabilities, the wealth of data produced by fruit-bearing plants in urban environments can significantly impact issues related to food authenticity, safety, and environmental concerns. These applications enhance the quality of urban life and contribute to global endeavors to build sustainable and resilient urban ecosystems.

For additional insights into citizen science for urban foraging, refer to the related commentary paper in GigaScience [78].

DATA AVAILABILITY

The datasets are available on Zenodo: Data Set [1] and Data Set [2]. Supporting data is also available via the GigaScience database, GigaDB [79].

LIST OF ABBREVIATIONS

AI: Artificial Intelligence; CS: Citizen Science; CVM: Computer Vision Model; RG: Research Grade.

DECLARATIONS

Competing interests

The authors declare that they have no competing interest regarding the publication of this work. There are no financial, personal, or professional relationships that could be perceived as potentially biasing the content presented in this manuscript.

Authors' contributions

Study Conceptualization: FMS, MCG, AMS, LFP, BCMSM. Data curation: FMS, LC, MCPC, ACSL, SGME, HDEG, BJS. Formal Analysis: FMS, LC, LFP, RRS, NPGL. Funding acquisition: AMS, UBM, BCMSM, MCG, RRS. Investigation: FMS, LFP, MCG, LC, NPGL, RRS, AMC, BCMSM, SK, UBM, DPD, RS, AMS. Methodology: FMS, NPGL, SK, UBM. Project administration: FMS, AMS. Supervision: AMS, LFP, BCMSM, UBM, AMC, MCG, RRS. Validation: LC. Data Visualization: FMS, RRS. Writing – original draft: FMS, LFP, LC, NPGL, AG, YB, SK. Writing – review & editing: All authors made significant contributions to the review and editing of this manuscript.

Funding

FMS thanks the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) (Process number: 21/15125-0, and 22/08385-8). BCMSM thanks the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (Process number: 303650/2019-2). AMS, NPGL, SK, and FMS thank FAPESP (Process number: 2018/14994-1). SK thanks FAPESP (Process number: 2019/26760-8).

Acknowledgements

We express our gratitude to the more than 2,600 citizen scientists who have made valuable contributions through their observations.

REFERENCES

- 1 Soares FM, Coradin L, Lima ACS et al. Taxonomic list of Brazilian fruit-bearing plants for human use (Version 3.0). Zenodo, November 2023; <https://doi.org/10.5281/zenodo.10212850>.
- 2 Soares FM, Pires LF, Garcia MC et al. Fruit-bearing plant species observations in Brazilian cities (Version 2.0). Zenodo, August 2023; <https://doi.org/10.5281/zenodo.8273049>.
- 3 Wieczorek J, Bloom D, Guralnick R et al. Darwin core: an evolving community-developed biodiversity data standard. *PLoS One*, 2012; 7(1): e29715. doi:10.1371/journal.pone.0029715.
- 4 Yesson C, Brewer PW, Sutton T et al. How global is the global biodiversity information facility? *PLoS One*, 2007; 2(11): e1124. doi:10.1371/journal.pone.0001124.
- 5 Ivanova NV, Shashkov MP. The possibilities of GBIF data use in ecological research. *Russ. J. Ecol.*, 2021; 52(1): 1–8. doi:10.1134/S1067413621010069.
- 6 Lane MA, Edwards JL. The global biodiversity information facility (GBIF). In: Curry GB, Humphries CJ (eds), *Biodiversity Databases: Techniques, Politics, and Applications*. Boca Raton, London, New York: CRC Press, 2007. doi:10.1201/9781439832547.
- 7 United Nations Environment Programme. Megadiverse Brazil giving biodiversity an online boost. <https://www.unep.org/news-and-stories/story/megadiverse-brazil-giving-biodiversity-online-boost#:~:text=Brazil%5C%20is%5C%20at%5C%20the%5C%20top,more%5C%20than%5C%204%5C%2000%5C%20plant%5C%20species> (visited on 05/26/2023).
- 8 Coradin L, Camillo J, Vieira ICG (eds) *Especies nativas da flora brasileira de valor econômico atual ou potencial: plantas para o futuro – Região Norte*. Ministério do Meio Ambiente, 2022.
- 9 Gomes SM, Chaves VM, de Carvalho AM et al. Biodiversity is overlooked in the diets of different social groups in Brazil. *Sci. Rep.*, 2023; 13(1): 7509. doi:10.1038/s41598-023-34543-8.
- 10 Clark KH, Nicholas KA. Introducing urban food forestry: a multifunctional approach to increase food security and provide ecosystem services. *Landsc. Ecol.*, 2013; 28(9): 1649–1669. doi:10.1007/s10980-013-9903-z.
- 11 Food and Agriculture Organization of the United Nations. *Unasylva 250: Forests and Sustainable Cities*. FAO Unasylva, 2018. <https://www.fao.org/documents/card/es/c/I8707EN>.
- 12 Salbitano F, Borelli S, Conigliaro M et al. Guidelines on urban and peri-urban forestry. In: FAO Forestry Paper. vol. 178, Food and Agriculture Organization of the United Nations (FAO), 2016. <https://www.fao.org/documents/card/es/c/e068e0d9-0c97-41c7-a856-05556a1bd10b>.

- 13 **Lwasa S, Mugagga F, Wahab B et al.** Urban and peri-urban agriculture and forestry: Transcending poverty alleviation to climate change mitigation and adaptation. *Urban Clim.*, 2014; 7: 92–106. doi:10.1016/j.uclim.2013.10.007.
- 14 **Batista DAM, Maciel EA, Machado DL et al.** Urban greening dynamics in a brazil central city as a subsidy for public policies. *Urban Ecosyst.*, 2023; 26: 845–856. doi:10.1007/s11252-023-01344-w.
- 15 **Freitas SR, Tambosi LR, Ghilardi-Lopes NP et al.** Spatial and temporal variation of potential resource availability provided by street trees in southeastern Brazil. *Urban Ecosyst.*, 2020; 23(5): 1051–1062. doi:10.1007/s11252-020-00974-8.
- 16 **Hurley PT, Emery MR, McLain R et al.** Whose urban forest? The political ecology of foraging urban nontimber forest products. In: Isenhour C, McDonogh G, Checker M (eds), Sustainability in the Global City. 1st ed, Cambridge University Press, 2015; pp. 187–212, doi:10.1017/CBO9781139923316.014.
- 17 **Hurley PT, Emery MR.** Locating provisioning ecosystem services in urban forests: Forageable woody species in New York City, USA. *Landsc. Urban Plan.*, 2018; 170: 266–275. doi:10.1016/j.landurbplan.2017.09.025.
- 18 **iNaturalist.** [inaturalist]: What is it. <https://www.inaturalist.org/pages/what+is+it> (visited on 06/05/2023).
- 19 **Khapugin AA, Kuzmin IV, Ivanova LA.** Distribution of four alien plants in Tyumen Region (Western Siberia): contribution of citizen science and expert data. *Wulfenia*, 2021; 28: 151–160. doi:10.5281/zenodo.6326883.
- 20 **Balashov I, Markova A.** Expansion of the invasive balkan slug *Tandonia Kusceri* (stylommatophora: Milacidae): A new frontier in northern ukraine and other new records. *Folia Malacolo.*, 2023; 31(1): 24–31. doi:10.12657/folmal.031.004.
- 21 **Prudic KL, Oliver JC, Brown BV et al.** Comparisons of citizen science data-gathering approaches to evaluate urban butterfly diversity. *Insects*, 2018; 9(4): 186. doi:10.3390/insects9040186.
- 22 **Mape N, Alisto L, Kitching I.** Hawkmoths of baguio city, philippines: a preliminary checklist derived from photo-based observation records available on iNaturalist and philippine lepidoptera internet platforms. *Philipp. J. Sci.*, 2022; 151(4): 1447–1453. doi:10.56899/151.04.12.
- 23 **Fogel NS, Thompson MF, Muniz PA et al.** The prevalence and manifestation of wing de-melanization in the eastern carpenter bee (Hymenoptera, Apidae *Xylocopa virginica*) associated with urban areas. *J. Kans. Entomol. Soc.*, 2023; 94(4): 235–266. doi:10.2317/0022-8567-94.4.235.
- 24 **Vardi R, Berger-Tal O, Roll U.** iNaturalist insights illuminate COVID-19 effects on large mammals in urban centers. *Biol. Conserv.*, 2021; 254: 108953. doi:10.1016/j.biocon.2021.108953.
- 25 **Mueller MA, Drake D, Allen ML.** Using citizen science to inform urban canid management. *Landsc. Urban Plan.*, 2019; 189: 362–371. doi:10.1016/j.landurbplan.2019.04.023.
- 26 **Drake D, Dubay S, Allen ML.** Evaluating human–coyote encounters in an urban landscape using citizen science. *J. Urban Ecol.*, 2021; 7(1): juaa032. doi:10.1093/jue/juaa032.
- 27 **Lopez B, Minor E, Crooks A.** Insights into human-wildlife interactions in cities from bird sightings recorded online. *Landsc. Urban Plan.*, 2020; 196: 103742. doi:10.1016/j.landurbplan.2019.103742.
- 28 **Cervera PEN.** Anfibios Y reptiles en parques recreativos ecológicos de la ciudad de mérida, yucatán, méxico. *Revista Latinoamericana de Herpetología*, 2021; 4(1): 82–94. doi:10.22201/fe.25942158e.2021.1.196.
- 29 **Jones D, Foshee B, Fitzgerald L.** A herpetological survey of edith l. Moore nature sanctuary. *Check List*, 2021; 17(1): 27–38. doi:10.15560/17.1.27.
- 30 **Herrera DJ, Cove MV.** Camera trap serendipity and citizen science point to broader effects of urban heat islands on food webs. *Food Webs*, 2020; 25: e00176. doi:10.1016/j.fooweb.2020.e00176.
- 31 **Marín-Gómez OH, Rodríguez Flores C, del Coro Arizmendi M.** Assessing ecological interactions in urban areas using citizen science data: Insights from hummingbird–plant meta-networks in a tropical megacity. *Urban For. Urban Green.*, 2022; 74: 127658. doi:10.1016/j.ufug.2022.127658.
- 32 **Putman BJ, Williams R, Li E et al.** The power of community science to quantify ecological interactions in cities. *Sci. Rep.*, 2021; 11(1): 3069. doi:10.1038/s41598-021-82491-y.
- 33 **Callaghan CT, Ozeroff I, Hitchcock C et al.** Capitalizing on opportunistic citizen science data to monitor urban biodiversity: A multi-taxa framework. *Biol. Conserv.*, 2020; 251: 108753. doi:10.1016/j.biocon.2020.108753.

- 34 **Li E, Parker SS, Pauly GB et al.** An urban biodiversity assessment framework that combines an urban habitat classification scheme and citizen science data. *Front. Ecol. Evol.*, 2019; 7: 277. doi:10.3389/fevo.2019.00277.
- 35 **Cambria VE, Campagnaro T, Trentanovi G et al.** Citizen science data to measure human use of green areas and forests in european cities. *Forests*, 2021; 12(6): 779. doi:10.3390/f12060779.
- 36 **Prudic KL, Wilson JK, Toshack MC et al.** Creating the urban farmer's almanac with citizen science data. *Insects*, 2019; 10(9): 294. doi:10.3390/insects10090294.
- 37 **Stark PB, Miller D, Carlson TJ et al.** Open-source food: nutrition, toxicology, and availability of wild edible greens in the east bay. *PLoS One*, 2019; 14(1): e0202450. doi:10.1371/journal.pone.0202450.
- 38 **Lorenzi H, Bacher L, Lacerda M et al.** Frutas brasileiras e exóticas cultivadas. Instituto Plantarum, 2006.
- 39 Coradin L, Siminski A, Reis A (eds) Espécies nativas da flora brasileira de valor econômico atual ou potencial: plantas para o futuro – Região Sul. Ministério do Meio Ambiente, 2011.
- 40 Vieira RF, Camillo J, Coradin L (eds) Espécies nativas da flora brasileira de valor econômico atual ou potencial: plantas para o futuro – Região Centro-Oeste. Ministério do Meio Ambiente, 2016.
- 41 Coradin L, Camillo J, Pareyn FGC (eds) Espécies nativas da flora brasileira de valor econômico atual ou potencial: plantas para o futuro – Região Nordeste. Ministério do Meio Ambiente, 2018.
- 42 **do Rio de Janeiro JB.** Flora e funga do brasil. 2024; <http://floradobrasil.jbrj.gov.br/> (Accessed 31 January 2024).
- 43 **Missouri Botanical Garden.** Tropicos.org. 2024; <https://tropicos.org> (Accessed 31 January 2024).
- 44 **The Royal Botanic Gardens, Kew and Harvard University Herbaria & Libraries and Australian National Herbarium.** International plant names index. 2023; <http://www.ipni.org> (Accessed 25 July 2023).
- 45 **Royal Botanic Gardens, Kew.** Plants of the world online. 2024; <http://www.plantsoftheworldonline.org/> (Accessed 31 January 2024).
- 46 **New York Botanical Garden.** Vascular plant types catalog. 2024; <https://sciweb.nybg.org/science2/hcol/vasc/index.asp.html> (Accessed 31 January 2024).
- 47 **iNaturalist.** Managing projects. <https://www.inaturalist.org/pages/managing-projects> (visited on 06/05/2023).
- 48 **Soares FM, Pires LF, Garcia MC et al.** Optimizing the monitoring of urban fruit-bearing flora with citizen science: An overview of the pomar urbano initiative. *Biodivers. Inf. Sci. Stand.*, 2023; 7: e112009. doi:10.3897/biss.7.112009.
- 49 **Governo do Distrito Federal.** Geografia. <https://www.df.gov.br/333/> (visited on 05/27/2023).
- 50 **Antunes MJM, dos Santos VVO, de Paula VA.** Culinária do Centro-Oeste: o pequi como iguaria natural. Etec Profa. Anna de Oliveira Ferraz, 2020. <http://ric.cps.sp.gov.br/handle/123456789/7160>.
- 51 **Bezerra JEF, de Lira Junior JS, da Silva Junior JF.** *Eugenia uniflora*: pitanga. In: Coradin L, Camillo J, Pareyn FGC (eds), Espécies nativas da flora brasileira de valor econômico atual ou potencial: plantas para o futuro: região Nordeste. Ministério do Meio Ambiente, 2018; pp. 155–169.
- 52 **Carvalho PER.** Jerivá: *Syagrus romanzoffiana*. In: Espécies Arbóreas Brasileiras. Embrapa, 2006; pp. 309–315.
- 53 **Jorge N, da Silva AC, Veronezi CM.** *Syagrus romanzoffiana*. In: de Lima FF, Lescano CH, de Oliveira IP (eds), Fruits of the Brazilian Cerrado: Composition and Functional Benefits. Springer International Publishing, 2021; pp. 141–159, doi:10.1007/978-3-030-62949-6_9.
- 54 **de Lacerda Coriolano D, Alves MHME, Cavalcanti IMF.** Biological potential of products obtained from palm trees of the genus syagrus. *Evid. Based Complementary Altern. Med.*, 2021; 2021: 1–11. doi:10.1155/2021/5580126.
- 55 **Moreira MAC, Payret Arrúa ME, Antunes AC et al.** Characterization of *Syagrus romanzoffiana* oil aiming at biodiesel production. *Ind. Crops Prod.*, 2013; 48: 57–60. doi:10.1016/j.indcrop.2013.04.006.
- 56 **Campbell CJ, Barve V, Belitz MW et al.** Identifying the identifiers: how inaturalist facilitates collaborative, research-relevant data generation and why it matters for biodiversity science. *BioScience*, 2023; 73(7): 533–541. doi:10.1093/biosci/biad051.
- 57 **Ueda K-I.** An overview of computer vision in iNaturalist. *Biodivers. Inf. Sci. Stand.*, 2020; 4: e59133. doi:10.3897/biss.4.59133.

- 58 **iNaturalist**. iNaturalist computer vision explorations. 2022; https://www.inaturalist.org/pages/computer_vision_demo (visited on 01/02/2023).
- 59 **iNaturalist**. What is the data quality assessment and how do observations qualify to become “research grade”? <https://www.inaturalist.org/pages/help#quality> (visited on 05/27/2023).
- 60 **iNaturalist Contributors**. iNaturalist research-grade observations. 2023; <https://doi.org/10.15468/AB3S5X>.
- 61 **Sponsler D**. The category of “cultivated” is problematic for plants in urban landscapes. *iNatForum*. April 2019; <https://forum.inaturalist.org/t/the-category-of-cultivated-is-problematic-for-plants-in-urban-landscapes/2317> (Accessed 31 January 2024).
- 62 **Hurley PT, Becker S, Emery MR et al**. Estimating the alignment of tree species composition with foraging practice in philadelphia’s urban forest: Toward a rapid assessment of provisioning services. *Urban For. Urban Green.*, 2022; **68**: 127456. doi:10.1016/j.ufug.2021.127456.
- 63 **Plieninger T, Bieling C, Fagerholm N et al**. The role of cultural ecosystem services in landscape management and planning. *Curr. Opin. Environ. Sustain.*, 2015; **14**: 28–33. doi:10.1016/j.cosust.2015.02.006.
- 64 **Synk CM, Kim BF, Davis CA et al**. Gathering baltimore’s bounty: characterizing behaviors, motivations, and barriers of foragers in an urban ecosystem. *Urban For. Urban Green.*, 2017; **28**: 97–102. doi:10.1016/j.ufug.2017.10.007.
- 65 **Bunge A, Diemont SAW, Bunge JA et al**. Urban foraging for food security and sovereignty: quantifying edible forest yield in syracuse, new york using four common fruit- and nut-producing street tree species. *J. Urban Ecol.*, 2019; **5**(1): juy028. doi:10.1093/jue/juy028.
- 66 **Garekae H, Shackleton CM**. Urban foraging of wild plants in two medium-sized south african towns: People, perceptions and practices. *Urban For. Urban Green.*, 2020; **49**: 126581. doi:10.1016/j.ufug.2020.126581.
- 67 **Sardeshpande M, Shackleton C**. Fruits of the city: the nature, nurture and future of urban foraging. *People and Nat.*, 2023; **5**(1): 213–227. doi:10.1002/pan3.10428.
- 68 **Brito VV, Borelli S**. Urban food forestry and its role to increase food security: A brazilian overview and its potentialities. *Urban For. Urban Green.*, 2020; **56**: 126835. doi:10.1016/j.ufug.2020.126835.
- 69 **Garekae H, Shackleton CM**. Foraging wild food in urban spaces: the contribution of wild foods to urban dietary diversity in south africa. *Sustainability*, 2020; **12**(2): 678. doi:10.3390/su12020678.
- 70 **Landor-Yamagata J, Kowarik I, Fischer L**. Urban foraging in berlin: people, plants and practices within the metropolitan green infrastructure. *Sustainability*, 2018; **10**(6): 1873. doi:10.3390/su10061873.
- 71 **Sardeshpande M, Hurley PT, Mollee E et al**. How people foraging in urban greenspace can mobilize social–ecological resilience during covid-19 and beyond. *Front. Sustain. Cities*, 2021; **3**: 686254. doi:10.3389/frsc.2021.686254.
- 72 **Palliwoda J, Kowarik I, Von Der Lippe M**. Human-biodiversity interactions in urban parks: The species level matters. *Landsc. Urban Plan.*, 2017; **157**: 394–406. doi:10.1016/j.landurbplan.2016.09.003.
- 73 **Fischer LK, Kowarik I**. Connecting people to biodiversity in cities of tomorrow: Is urban foraging a powerful tool? *Ecol. Indic.*, 2020; **112**: 106087. doi:10.1016/j.ecolind.2020.106087.
- 74 **Schunko C, Brandner A**. Urban nature at the fingertips: investigating wild food foraging to enable nature interactions of urban dwellers. *Ambio*, 2022; **51**(5): 1168–1178. doi:10.1007/s13280-021-01648-1.
- 75 **Santiago RdAC, Coradin L (eds)** Biodiversidade brasileira: sabores e aromas. Ministério do Meio Ambiente, 2018.
- 76 **Urban 20**. Championing ‘Local’ Culture and Economy. Urban 20 India, 2023. https://www.u20india.org/Content/WhitePaper/6_CLCE_White%20Paper_1.pdf (Accessed 15 January 2024).
- 77 **Marvin HJP, Hoenderdaal W, Gavai AK et al**. Global media as an early warning tool for food fraud; an assessment of MedISys-FF. *Food Control*, 2022; **137**: 108961. doi:10.1016/j.foodcont.2022.108961.
- 78 **Soares FM, Pires LF, Garcia MC et al**. Leveraging citizen science for monitoring urban forageable plants. *GigaScience*, 2024; doi:10.1093/gigascience/giae007.
- 79 **Soares FM, Pires LF, Garcia MC et al**. Supporting data for “Citizen science data on urban forageable plants: a case study in Brazil”. *GigaScience Database*, 2024; <https://doi.org/10.5524/102499>.