Workshop on Visualisation in Environmental Sciences (EnvirVis) (2018)
R. Bujack and K. Rink and D. Zeckzer (Editors)

TreeeX: Exploring the Diversity of Tree Species

S. Jänicke

Image and Signal Processing Group, Institute of Computer Science, Leipzig University, Leipzig, Germany

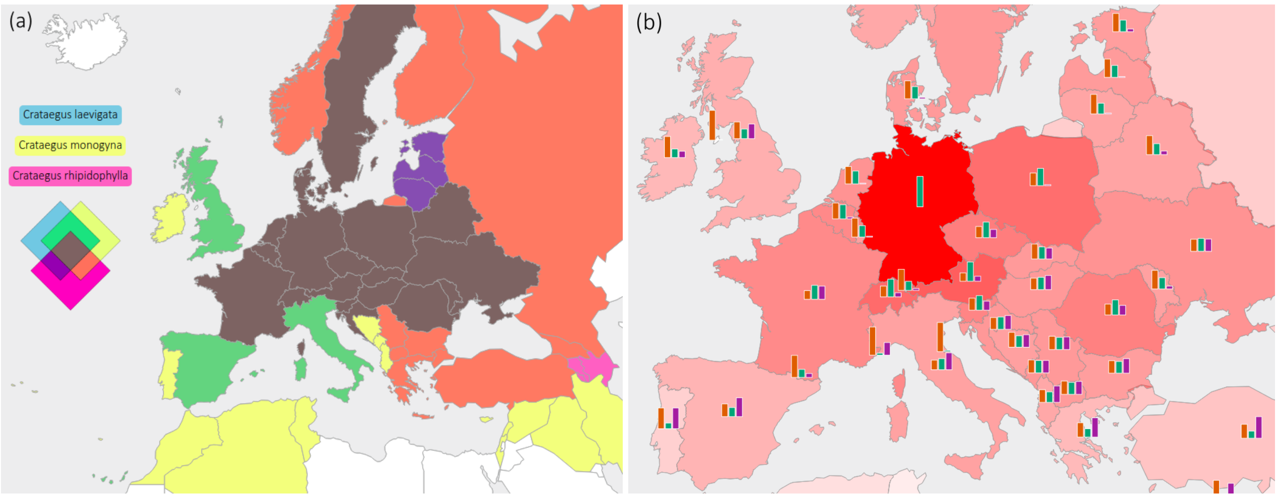


Figure 1: *Two TreeeX choropleth maps focusing on Germany. Figure 1(a) contrasts the countries where the three hawthorn species indigenous in Germany grow, and Figure 1(b) illustrates the biodiversity of tree species in Germany in comparison to other European countries.*

Abstract

The GlobalTreeSearch database provides a mapping between 60,065 tree species known to science and the countries where these trees grow. TreeeX is a visual exploration system that supports multifaceted analyses of the GlobalTreeSearch data. Focusing on the entire earth or a group of countries, investigating research questions on biodiversity are visually supported by interactive choropleth maps that color countries according to frequency, diversity, or uniqueness of prevalent tree species. Focusing on a single country, similarities and differences to other countries can be analyzed in detail. Several examples outline the system's capability of delivering insights concerning the geographical diversity of tree species.

1. Introduction

The importance of trees for the ecological system has been discussed in a multitude of publications. As trees are at the base of the trophic pyramid [KS84], they frame the habitat of other species that depend on prevalent tree species. Taking the human as example [FPZ*14], we need trees as resources for food, for building material, for medicine, and we often choose woodlands for recreational purposes. Also, the significant role of trees for stabilizing the Earth's climate system, thus, the necessity of preserving forests and woodlands is a well-known fact [PvdST*15].

In order to assess, monitor, and manage tree species diversity on a global, regional, and/or national level, researchers of Botanic Gardens Conservation International (BGCI) combined over 500 published sources on regional, taxonomic, and country-specific tree growth [BROS17] and made the resultant geographical mapping of tree species available in the form of the GlobalTreeSearch online database [BGC18]. Users of the database are provided with a textual search interface that allows for searching tree species and/or tree genera for a given country. The search result is a textual extract of the database listing family, taxon name and corresponding author of the tree species that match the given query; Figure 2 shows an example output after searching for hawthorn species in Germany. Though lists for individual queries can be downloaded in a tabular format, a comparative analysis of the geographical diversity of tree growth is not supported.

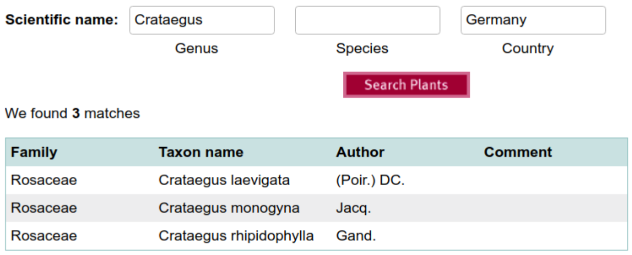


Figure 2: *Searching for hawthorns (Crataegus) in Germany using the GlobalTreeSearch online database*

TreeeX was designed in order to provide multifaceted visual access to the GlobalTreeSearch data. As the geographical mapping is given on the country-level, TreeeX supports analyses on the biodiversity of tree species with interactive choropleth maps that illustrate different scenarios. Next to comparing the distribution of different tree species, the system allows for exploring aspects on species richness and uniqueness. In addition, similarities and differences of a selected country compared to other countries can be analyzed.

2. Related Work

Geovisualizations as means of illustrating aspects on biodiversity have been frequently used; exhaustive overviews of related techniques from thematic cartography and geovisualization are provided by Slocum et al. [SMKH09] and by Andrienko et al. [AA05].

Usually, thematic maps are based on a qualitative data set for a small geographical area, be it a national park or a whole country. Debinski et al. [DKJ99] use a geographical information system (GIS) to categorize habitats, and to determine relationships between remotely sensed habitat categorizations and species distribution patterns in the Greater Yellowstone Ecosystem. Thematic maps have also been used for a qualitative study of vegetation in Madagascar to draw implications and recommendations for the conservation of biodiversity [DPM98]. Similarly, Madden [Mad04] uses various heat maps to assess vegetation patterns in Great Smoky Mountains National Park. Setturu et al. [SR16] use categorial maps to visualize landscape dynamics in National Parks of Central Western Ghats. Categorial pixel-based dot maps, where each tree species or genera in the study receives a certain color, can be used to visualize detailed information on the regional distribution of tree species [RFN*08, HHP*16]. Predominance tag maps can be used to overlay such maps with textual information on locally predominant tree species [RCSJ18].

Geovisualizations are further central to analyzing time-dependent tree cover changes. Carnaval et al. [CM08] use heat maps to model the spatial range of the Brazilian Atlantic forest under three climatic scenarios (current climate, 6000 and 21,000 years ago) in order to predict patterns of current biodiversity. Based on analysis of nearly 30,000 Landsat images, Kim et al. [KSN*14] illustrate worldwide forest-cover change from 1990 to 2000 with heat maps. Heat maps are also used by Hansen et al. [HPM*13] to visualize tree cover, forest loss, and forest gain, and by Allen et al. [ABM15] for the analysis of worldwide locations of substantial tree mortality and forest die-off from hotter drought in the Anthropocene.

Linked views that interlink geospatial and temporal information are also valuable for exploring biodiversity data. Jänicke et al. [JS14] use GeoTemCo [JHSS12, JHS13] for the geospatial-temporal mapping and biodiversity analyses of the benthic invertebrate fauna. CommonGIS, a similar system, is used to combine forest ecosystem models with exploratory data visualization for the analysis of long-term simulation results [CKM*05].

Image processing techniques are used for generating information about the diversity of trees in forests. While Clark et al. [CRC05] use high spectral and spatial resolution imagery for the automated species-level classification of individual tree crowns in a tropical rain forest, Simons et al. [SHTA14] propose a point-based rendering method to support the remote sensing of forests. Musasabi [KKO*15] aims at simulating the growth and changes of forests depending on the species of the tree and the land conditions.

3. TreeeX System

TreeeX is composed of an interactive choropleth map that is colored according to three different analysis modes: species analysis, global diversity, and country of interest. A screenshot of the system is given in Figure 3.

3.1. Species Analysis

In order to give a first impression of the most frequent tree species in the database, a tag cloud is shown with words scaled accordingly. Especially for novice users, the tag cloud aids as a starting point due to their intuitiveness, and their widespread usage to display summaries of textual data [VW08]. When hovering a tag, the number of countries where the chosen tree species grows is shown, and on mouse click, the corresponding countries are highlighted in a certain color; tree species can also be directly requested with a keyword search. The interface enables comparing the distributions of different tree species visually. When multiple species are prevalent in a country, the different colors are overlaid using a subtractive color mixing scheme. As the number of colors that can be easily distinguished by humans is limited by 12 [War04], though extendable, comparing a maximum of three tree species is suggested, which results in a color map consisting of seven distinctive colors in the CMYK color model as illustrated in Figure 4.

species	C	Y	M	K
A	50	0	0	0
B	0	50	0	0
C	0	0	50	0
$A \wedge B$	50	50	0	0
$A \wedge C$	50	0	50	0
$B \wedge C$	0	50	50	0
$A \wedge B \wedge C$	50	50	50	0

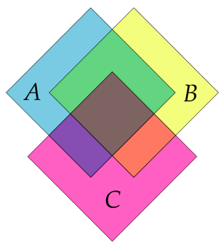


Figure 4: *CYMK color mixing for the species A, B and C*

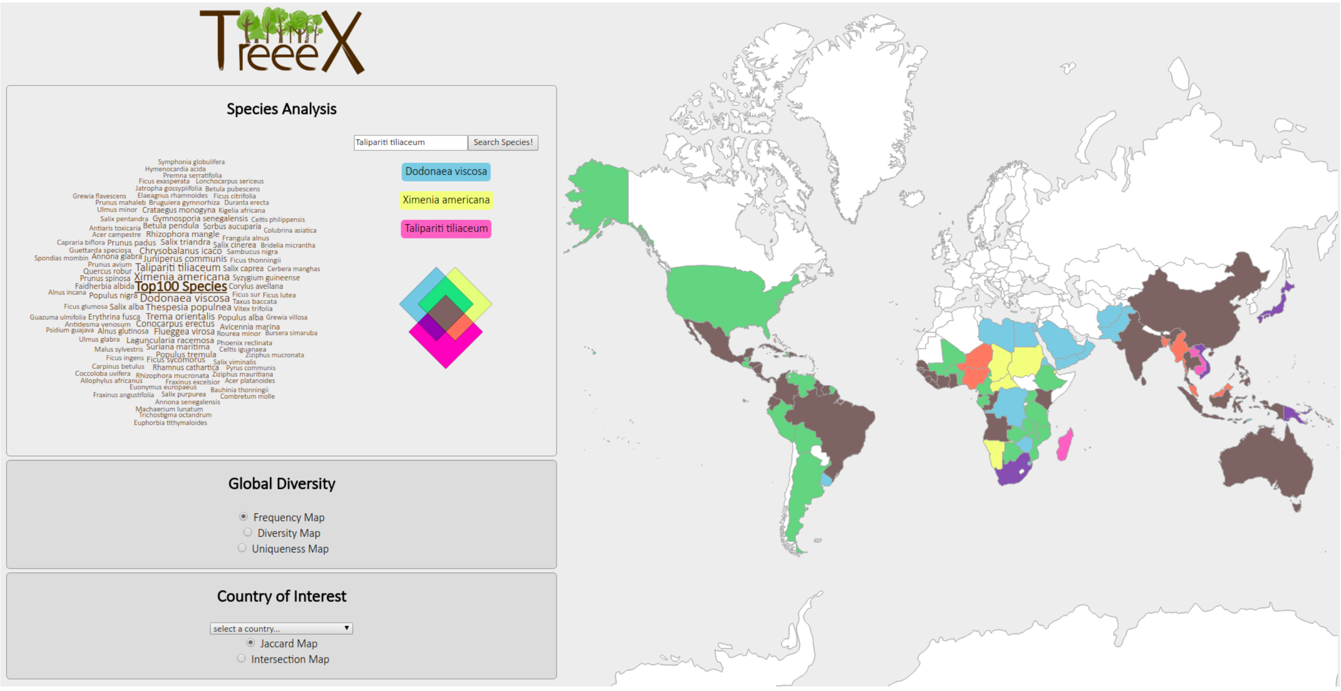


Figure 3: *Screenshot of TreeeX: The countries are colored according to the distributions of the three most widespread tree species Dodonaea viscosa (blue), Ximenia americana (yellow) and Talipariti tiliaceum (pink).*

3.2. Country of Interest

Focusing on a single country c_i, each country c_j of the choropleth map is colored in red with a saturation level according to the similarity of tree species $T(c_i)$ and $T(c_j)$ of c_i and c_j. Two coloring modes are possible. The Jaccard index defined as

$$J(c_i, c_j) = \frac{|T(c_i) \cap T(c_j)|}{|T(c_i) \cup T(c_j)|}$$

considers the cardinalities of both sets of tree species, thus, providing a general view on the similarity of the biodiversities of c_i and c_j. On the other hand, when the user likes to see the countries sharing most of c_i's tree species, thus, providing similar habitat conditions, the cardinality of $T(c_j)$ is disregarded and only the overlap size compared to the richness of tree species of c_i is considered as

$$I(c_i, c_j) = \frac{|T(c_i) \cap T(c_j)|}{|T(c_i)|}.$$

For exploration purposes, each country is attached with a bar chart reflecting these proportions. The center bar colored green illustrates the amount of shared species, the orange bar on the left the amount of tree species only prevalent in c_i, and the purple bar on the right the amount of tree species only prevalent in c_j. When hovering the bar chart, a popup window is shown that provides several information, and it can be expanded for a more detailed investigation of individual tree species. An example focusing on Germany is shown in Figure 1(b).

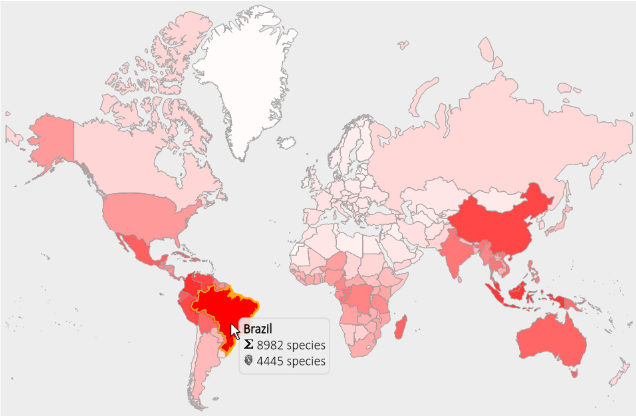
3.3. Global Analysis

Various research questions on the biodiversity of tree growth can be investigated in this mode for a given set of countries. The choropleth map can be colored according to three different aspects of global diversity:

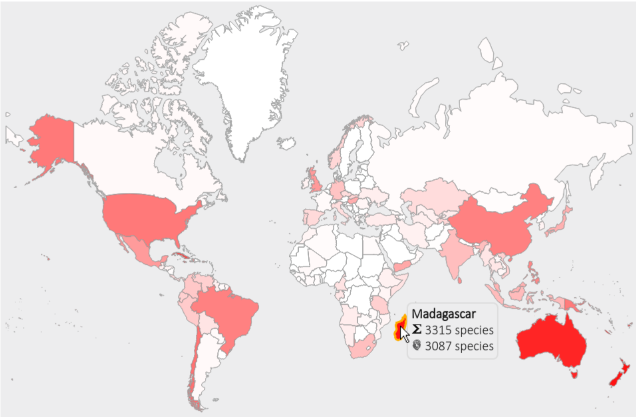
Frequency. Countries are colored in red with a saturation level according to the total of different tree species divided by the total of the country with most tree species. Currently, the database holds Brazil with 8,982 tree species as country with the highest biodiversity.

Uniqueness. Countries are colored in red with a saturation level according to the uniqueness of tree species, that is the number of tree species unique in a country divided by the country's total of tree species. Currently, Madagascar is listed as the country with the most unique biodiversity as 3,087 out of 3,315 tree species only grow in Madagascar.

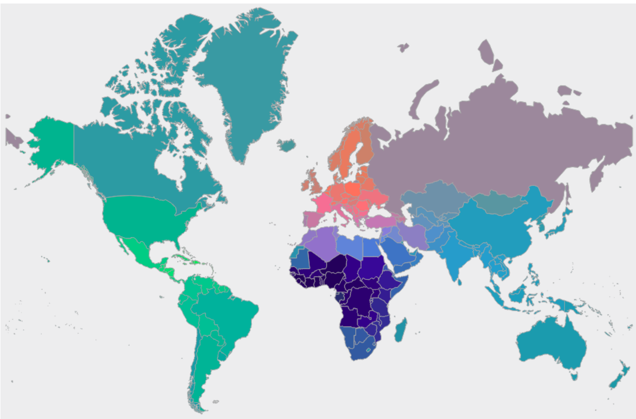
Diversity. The Country-of-Interest functionality gives a comparative view on the similarities of tree species of a chosen country compared to all other countries. In order to generate a global view on the changing biodiversity among the given set of countries, this mode takes the similarities of all country tuples at once into account. Therefore, a $n \times n$ distance matrix $D = [d_{ij}]$ for n countries is calculated. d_{ij} denotes the dissimilarity of tree species between the countries c_i and c_j, which is defined as $d_{ij} = 1 - J(c_i, c_j)$. Then, a classical multidimensional scaling (MDS) algorithm [BG05] is executed that determines a position $p_i = (x_i, y_i, z_i)$ in the 3D space $S = [0,1]_x \times [0,1]_y \times [0,1]_z$ for each country c_i while preserving distances between countries as well as possible. S is then scaled to



(a) *Color coding according to the total of tree species*



(b) *Color coding according to the uniqueness of tree species*



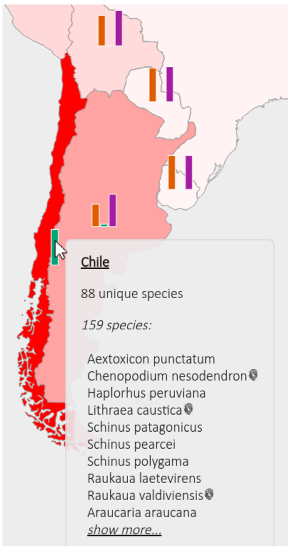
(c) *Color coding according to MDS of countries' tree species similarities*

Figure 5: *Choropleth maps illustrating aspects of global diversity*

the RGB cube, so that the color for a country is determined by

$$R_i = \frac{\frac{r_{max}-r_x}{2} + x_i}{r_{max}} \cdot 255$$

$$G_i = \frac{\frac{r_{max}-r_y}{2} + y_i}{r_{max}} \cdot 255$$



(a) *Browsing tree species in Chile*

(b) *Pepper tree distribution areas*

Figure 6: *Pepper tree species in Chile*

$$B_i = \frac{\frac{r_{max}-r_z}{2} + z_i}{r_{max}} \cdot 255.$$

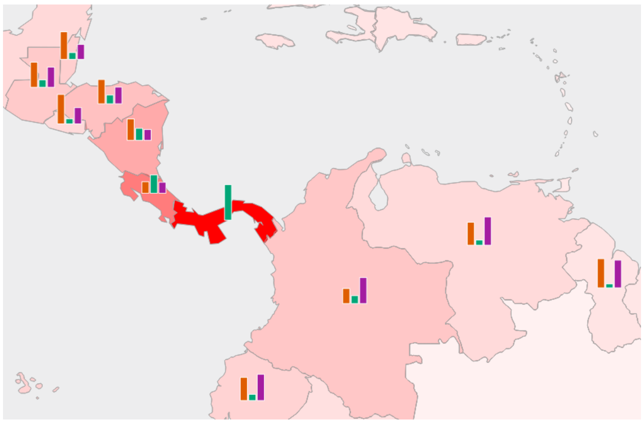
r_x defines the spanned range of all x-values, thus, determined as $r_x = x_{max} - x_{min}$. r_y and r_z are defined equally, and $r_{max} = \max\{r_x, r_y, r_z\}$.

4. Working with TreeeX

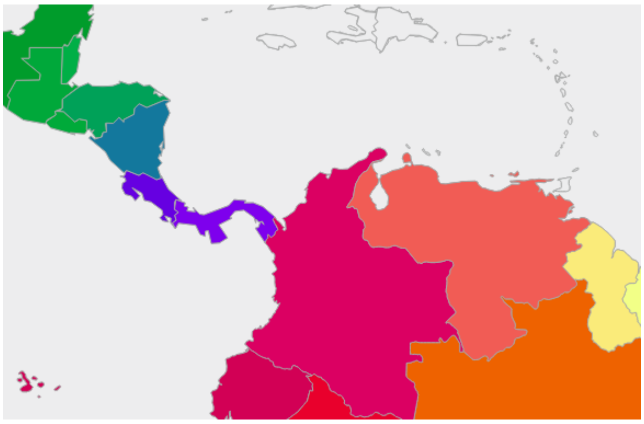
In order to illustrate the value of the proposed system to investigate different research questions on the biodiversity of tree species, this section outlines typical workflows with TreeeX.

Species Analysis. Researchers already stated that around 58% of all tree species are single country endemics [BROS17]. When browsing the tree species for a certain country, those unique species are marked with a fingerprint icon. Taking Chile with a total of 158 species as an example, unique and common species are already visible in the popup window as illustrated in Figure 6a. The three pepper tree species prevalent in Chile are not unique, which deserves a more detailed investigation of the distribution areas with the species analysis functionality. The resultant choropleth map marks Chile as the only country where all three species grow (see Figure 6b). A similar example is given in Figure 1(a) that shows the distribution areas of the three hawthorn species that grow in Germany. While all three species are endemic central Europe, *Crataegus monogyna* (common hawthorn) has the largest distribution area excluding the Baltic states.

Country of Interest. The role of Panama's biodiversity has been the subject of interest due to the global importance of the Panama Canal [CRIn*01]. Figure 7a illustrates a high similarity to the tree species prevalent in Costa Rica, and the farther away from Panama to the north or to the south the amount of overlap decreases. That Panama lies in the center of a continuum of a north-south biodiversity change is depicted in Figure 7b. Focusing on Germany 1(b)



(a) *Focusing on Panama*



(b) *Continuum of tree species diversity change with Panama in the center*

Figure 7: *The role of Panama*

shows that to the north, only few species are endemic that do not grow in Germany. On the other hand, the farther to the south, the more tree species occur that are not endemic in Germany.

Global Analysis. When comparing the diversities of tree species of all countries, thus, choosing the diversity view in the global analysis mode, yet regional similarities and geographical changes get visible. Looking at the entire earth (see Figure 5c), one can see several clusters of similarly colored countries sharing a certain amount of tree species. Expectedly, this is the case for countries belonging to the same continents, e.g., Europe and Africa. But also, countries of different continents receive similar shades illustrating a higher similarity of biodiversities. The diversity among these countries becomes visible on a smaller scale, thus, by assigning a larger color space. By first zooming to the Americas (see Figure 8a), and then, by zooming to South America, as shown in Figure 8b, five clusters of countries having similar biodiversities can be seen. Switching to the Country of Interest mode with a focus on Argentina confirms the diversity mapping with highest similarities to Paraguay and Uruguay that can be further investigated (see Figure 8c).

5. Conclusion

TreeeX is a system that supports the comparative geographical analysis of biodiversities on the country-level. For that purpose, TreeeX provides an interactive choropleth map that is colored according to the given research question. So, TreeeX enables a multifaceted visual access to a rich database fundamental for research on the biodiversity of tree species. Temporal information on biodiversity, which are not contained in the GlobalTreeSearch database, but pertinent for biodiversity research are currently not considered by TreeeX.

The TreeeX interface is designed the way that it can also be used for other kinds of data where geographical regions—countries, states, districts, etc.—have overlapping sets of entities. In order to foster future research, TreeeX is implemented as a user-configurable JavaScript library based on D3.js, and it will be made available as open-source soon.

References

[AA05] ANDRIENKO N., ANDRIENKO G.: *Exploratory Analysis of Spatial and Temporal Data: A Systematic Approach*. Springer, 2005. 2

[ABM15] ALLEN C. D., BRESHEARS D. D., MCDOWELL N. G.: On underestimation of global vulnerability to tree mortality and forest die-off from hotter drought in the Anthropocene. *Ecosphere 6*, 8 (2015), 1–55. art129. 2

[BG05] BORG I., GROENEN P.: *Modern Multidimensional Scaling: Theory and Applications*. Springer, 2005. 3

[BGC18] BGCI: GlobalTreeSearch online database, 2018. Botanic Gardens Conservation International. Richmond, U.K. Available at www.bgci.org. Accessed on 09/03/2018. 1

[BROS17] BEECH E., RIVERS M., OLDFIELD S., SMITH P. P.: GlobalTreeSearch: The first complete global database of tree species and country distributions. *Journal of Sustainable Forestry 36*, 5 (2017), 454–489. 1, 4

[CKM*05] CHERTOV O., KOMAROV A., MIKHAILOV A., ANDRIENKO G., ANDRIENKO N., GATALSKY P.: Geovisualization of Forest Simulation Modelling Results: A Case Study of Carbon Sequestration and Biodiversity. *Comput. Electron. Agric. 49*, 1 (Oct. 2005), 175–191. 2

[CM08] CARNAVAL A. C., MORITZ C.: Historical climate modelling predicts patterns of current biodiversity in the Brazilian Atlantic forest. *Journal of Biogeography 35*, 7 (2008), 1187–1201. 2

[CRC05] CLARK M. L., ROBERTS D. A., CLARK D. B.: Hyperspectral discrimination of tropical rain forest tree species at leaf to crown scales. *Remote Sensing of Environment 96*, 3 (2005), 375 – 398. 2

[CRIn*01] CONDIT R., ROBINSON W. D., IBÁÑEZ R., AGUILAR S., SANJUR A., MARTÍNEZ R., STALLARD R. F., GARCÍA T., ANGEHR G. R., PETIT L., WRIGHT S. J., ROBINSON T. R., HECKADON S.: The Status of the Panama Canal Watershed and Its Biodiversity at the Beginning of the 21st CenturyLong-term ecological studies reveal a diverse flora and fauna near the Panama Canal, harbored within a corridor of forest stretching from the Caribbean to the Pacific, but deforestation, land degradation, erosion, and overhunting remain threats. *BioScience 51*, 5 (2001), 389–398. 4

[DKJ99] DEBINSKI D., KINDSCHER K., JAKUBAUSKAS M.: A remote sensing and GIS-based model of habitats and biodiversity in the Greater Yellowstone Ecosystem. *International Journal of Remote Sensing 20*, 17 (1999), 3281–3291. 2

[DPM98] DU PUY D. J., MOAT J.: Vegetation mapping and classification in Madagascar (using GIS): implications and recommendations for the conservation of biodiversity. *Chorology, taxonomy and ecology of the floras of Africa and Madagascar* (1998), 97–117. 2

(a) *Biodiversities in the Americas* **(b)** *Biodiversities in South America* **(c)** *Comparing Argentina and Uruguay*

Figure 8: *Comparing biodiversities on different zoom levels*

[FPZ*14] FITZJOHN R. G., PENNELL M. W., ZANNE A. E., STEVENS P. F., TANK D. C., CORNWELL W. K.: How much of the world is woody? *Journal of Ecology 102*, 5 (2014), 1266–1272. 1

[HHP*16] HEIDEN U., HOLZWARTH S., PINNEL N., REICHMUTH A., RACZKO E., HEURICH M., MÜLLER J., WANG Z., SKIDMORE A., ALI A., WANG T., DARVISHZADEH R., WEGMANN M.: Laboratory for Essential Biodiversity Variables (EBV) Concepts – The "Data Pool Initiative for the Bohemian Forest Ecosystem". In *Living Planet Symposium 2016* (2016). 2

[HPM*13] HANSEN M. C., POTAPOV P. V., MOORE R., HANCHER M., TURUBANOVA S. A., TYUKAVINA A., THAU D., STEHMAN S. V., GOETZ S. J., LOVELAND T. R., KOMMAREDDY A., EGOROV A., CHINI L., JUSTICE C. O., TOWNSHEND J. R. G.: High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science 342*, 6160 (2013), 850–853. 2

[JHS13] JÄNICKE S., HEINE C., SCHEUERMANN G.: GeoTemCo: Comparative Visualization of Geospatial-Temporal Data with Clutter Removal Based on Dynamic Delaunay Triangulations. In *Computer Vision, Imaging and Computer Graphics. Theory and Application*. Springer, 2013, pp. 160–175. 2

[JHSS12] JÄNICKE S., HEINE C., STOCKMANN R., SCHEUERMANN G.: Comparative Visualization of Geospatial-temporal Data. In *Proceedings of the International Conference on Computer Graphics Theory and Applications and International Conference on Information Visualization Theory and Applications (VISIGRAPP 2012)* (2012), pp. 613–625. 2

[JS14] JÄNICKE S., SCHEUERMANN G.: Utilizing GeoTemCo for Visualizing Environmental Data. In *Workshop on Visualisation in Environmental Sciences (EnvirVis)* (2014), Kolditz O., Rink K., Scheuermann G., (Eds.), The Eurographics Association. 2

[KKO*15] KATO T., KATO A., OKAMURA N., KANAI T., SUZUKI R., SHIRAI Y.: Musasabi: 2D/3D Intuitive and Detailed Visualization System for the Forest. In *ACM SIGGRAPH 2015 Posters* (New York, NY, USA, 2015), SIGGRAPH '15, ACM, pp. 79:1–79:1. 2

[KS84] KENNEDY C. E. J., SOUTHWOOD T. R. E.: The Number of Species of Insects Associated with British Trees: A Re-Analysis. *Journal of Animal Ecology 53*, 2 (1984), 455–478. 1

[KSN*14] KIM D.-H., SEXTON J. O., NOOJIPADY P., HUANG C., ANAND A., CHANNAN S., FENG M., TOWNSHEND J. R.: Global, landsat-based forest-cover change from 1990 to 2000. *Remote Sensing of Environment 155* (2014), 178 – 193. 2

[Mad04] MADDEN M.: Vegetation modeling, analysis and visualization in US National Parks. *International Archives of Photogrammetry and Remote Sensing 35* (2004), 1287–1293. 2

[PvdST*15] POORTER L., VAN DER SANDE M. T., THOMPSON J., ARETS E. J. M. M., ALARCÓN A., ÁLVAREZ-SÁNCHEZ J., ASCARRUNZ N., BALVANERA P., BARAJAS-GUZMÁN G., BOIT A., BONGERS F., CARVALHO F. A., CASANOVES F., CORNEJO-TENORIO G., COSTA F. R. C., DE CASTILHO C. V., DUIVENVOORDEN J. F., DUTRIEUX L. P., ENQUIST B. J., FERNÁNDEZ-MÉNDEZ F., FINEGAN B., GORMLEY L. H. L., HEALEY J. R., HOOSBEEK M. R., IBARRA-MANRÍQUEZ G., JUNQUEIRA A. B., LEVIS C., LICONA J. C., LISBOA L. S., MAGNUSSON W. E., MARTÍNEZ-RAMOS M., MARTÍNEZ-YRIZAR A., MARTORANO L. G., MASKELL L. C., MAZZEI L., MEAVE J. A., MORA F., MUÑOZ R., NYTCH C., PANSONATO M. P., PARR T. W., PAZ H., PÉREZ-GARCÍA E. A., RENTERÍA L. Y., RODRÍGUEZ-VELAZQUEZ J., ROZENDAAL D. M. A., RUSCHEL A. R., SAKSCHEWSKI B., SALGADO-NEGRET B., SCHIETTI J., SIMÕES M., SINCLAIR F. L., SOUZA P. F., SOUZA F. C., STROPP J., TER STEEGE H., SWENSON N. G., THONICKE K., TOLEDO M., URIARTE M., VAN DER HOUT P., WALKER P., ZAMORA N., PEÑA CLAROS M.: Diversity enhances carbon storage in tropical forests. *Global Ecology and Biogeography 24*, 11 (2015), 1314–1328. 1

[RCSJ18] RECKZIEGEL M., CHEEMA M. F., SCHEUERMANN G., JÄNICKE S.: Predominance Tag Maps. *IEEE Transactions on Visualization and Computer Graphics* (2018). 2

[RFN*08] RUEFENACHT B., FINCO M., NELSON M., CZAPLEWSKI R., HELMER E., BLACKARD J., HOLDEN G., LISTER A., SALAJANU D., WEYERMANN D., WINTERBERGER K.: Conterminous U.S. and Alaska Forest Type Mapping Using Forest Inventory and Analysis Data.

Photogrammetric Engineering & Remote Sensing 74, 11 (2008), 1379–1388. 2

[SHTA14] SIMONS L., HE S., TITTMAN P., AMENTA N.: Point-based rendering of forest LiDAR. In *Workshop on Visualisation in Environmental Sciences (EnvirVis), The Eurographics Association* (2014), pp. 19–23. 2

[SMKH09] SLOCUM T. A., MCMASTER R. B., KESSLER F. C., HOWARD H. H.: *Thematic Cartography and Geovisualization*, 3rd, international ed. Prentice Hall Series in Geographic Information Science. Prentice Hall, 2009. 2

[SR16] SETTURU B., RAMACHANDRA T.: Visualization of Landscape Dynamics in National Parks of Central Western Ghats. In *proceedings of 10thBiennial Lake Conference–2016* (2016), pp. 28–31. 2

[VW08] VIEGAS F., WATTENBERG M.: TIMELINES: Tag Clouds and the Case for Vernacular Visualization. *interactions 15*, 4 (July 2008), 49–52. 2

[War04] WARE C.: *Information Visualization: Perception for Design*, 3rd ed. Morgan Kaufmann, 2004. 2