Biodiversity Informatics, 11, 2016, 1-11

COMPLETENESS OF DIGITAL ACCESSIBLE KNOWLEDGE OF THE PLANTS OF GHANA

ALEX ASASE[1] AND A. TOWNSEND PETERSON[2]
[1]*Department of Botany, University of Ghana, P. O. Box LG 55 Legon, Ghana*
[2]*Biodiversity Institute, University of Kansas, Lawrence, KS 66045, USA*

Abstract.—Providing comprehensive, informative, primary, research-grade biodiversity information represents an important focus of biodiversity informatics initiatives. Recent efforts within Ghana have digitized >90% of primary biodiversity data records associated with specimen sheets in Ghanaian herbaria; additional herbarium data are available from other institutions via biodiversity informatics initiatives such as the Global Biodiversity Information Facility. However, data on the plants of Ghana have not as yet been integrated and assessed to establish how complete site inventories are, so that appropriate levels of confidence can be applied. In this study, we assessed inventory completeness and identified gaps in current Digital Accessible Knowledge (DAK) of the plants of Ghana, to prioritize areas for future surveys and inventories. We evaluated the completeness of inventories at ½° spatial resolution using statistics that summarize inventory completeness, and characterized gaps in coverage in terms of geographic distance and climatic difference from well-documented sites across the country. The southwestern and southeastern parts of the country held many well-known grid cells; the largest spatial gaps were found in central and northern parts of the country. Climatic difference showed contrasting patterns, with a dramatic gap in coverage in central-northern Ghana. This study provides a detailed case study of how to prioritize for new botanical surveys and inventories based on existing DAK.

Key words.—biodiversity informatics, primary data, inventory completeness, data gaps, Ghana, flora, botanical surveys

Biodiversity informatics may be defined as the application of information technologies to the management, algorithmic exploration, analysis, and interpretation of primary data regarding life, with a particular focus at the species level of organization (Soberón and Peterson, 2004). It is a rather new field, with the earliest citation of the term only 17 years ago (Schalk, 1998). The most important biodiversity information centers on primary data, such as records of occurrences of species (particularly when vouchered by specimens), although many secondary sources (e.g., atlases, species accounts, distribution maps) exist as well (Costello et al., 2013). Such data have accumulated over centuries, but only relatively recently have they been converted into digital formats (Guralnick et al., 2007) and shared openly via data portals (Graham et al., 2004).

Major activities in biodiversity informatics are currently data-centered, focused in three areas: (i) data extraction and capture, (2) data compilation and serving, and (3) data display and visualization (Peterson et al., 2010). The past two decades have seen advances and improvements in information technology (e.g. large-capacity electronic storage media, Internet and data portals, distributed database technology); development of efficient data digitization workflows; changes in policies of owners of primary biodiversity data (e.g., see large-scale initiatives toward digitization of specimen data in Naturalis Biodiversity Centre, Leiden, and Muséum National d'Histoire Naturelle (MNHN), Paris), as well as establishment of global and regional biodiversity information initiatives (e.g., Global Biodiversity Information Facility, GBIF; Atlas of Living Australia, ALA). These initiatives have contributed to massive accumulation and serving of primary biodiversity data records via the Internet: e.g., GBIF currently serves >648M primary data records on its data portal (accessed 13 May 2016). However, this forward progress could be threatened if the data do not prove sufficiently useful to biodiversity researchers, managers, and decision makers (Peterson et al., 2010).

Primary biodiversity data have myriad applications, providing an information base that is crucial to addressing challenges of sustainable development and decision-making about natural resources and environments (Chapman, 2005; Sousa-Baena et al., 2013). Digital Accessible Knowledge (DAK) regarding biodiversity comprises primary data records that are in digital format,

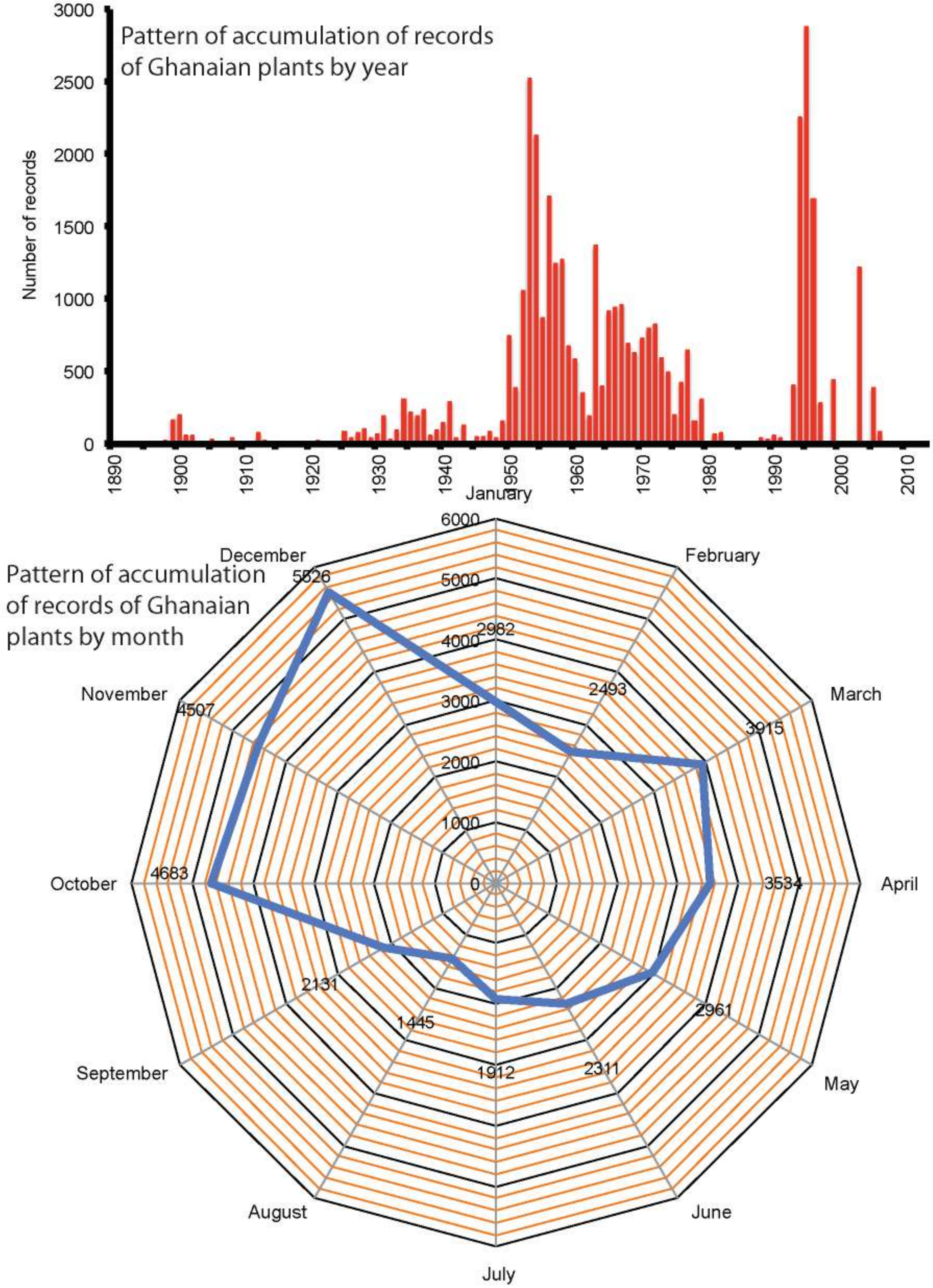

Figure 1. Graphs showing accumulation of records of Ghanaian plants through time (a) years, and (b) during the year.

accessible globally without cost, and integrated with the broader university of such data (Sousa-Baena et al., 2013). Some exciting examples of uses of DAK exist, including for prioritizing areas for conservation, assessing geographic potential for species invasions, and understanding ecological and evolutionary processes (e.g., Mora et al., 2008; Nakamura & Soberón, 2008).

In Ghana, significant efforts have been invested in digitization of and providing access to primary biodiversity data on the plants of the country. Although the National Biodiversity Strategy for Ghana estimated 3227 plant species (2974 indigenous and 252 introduced) in Ghana (Ministry of Environment and Science, 2002), no consensus exists on how many plant species occur in the country. Still, DAK on the plants (including fungi and algae that are traditionally studied in the field of botany) of Ghana is relatively large, based on >90% of primary biodiversity data records derived from specimens in Ghanaian herbaria, plus data from other institutions served through biodiversity informatics initiatives such as GBIF. Data on plants of Ghana have not been integrated and assessed to establish how complete are site inventories across the country, so that appropriate levels of confidence can be applied; these gaps in knowledge affect directly the fitness-for-use of the data (Otegui et al., 2013). As a consequence, this study undertook detailed assessment of DAK on plants of Ghana to identify and highlight gaps in knowledge.

METHODS

Data were obtained from two major sources: (1) a BRAHMS database on plants of Ghana that includes data associated with plant specimens in the collections of the University of Ghana, Resource Support Management Centre (RSMC) of Ghana Forestry Commission, Aburi Botanic Gardens, and Centre for Scientific Research into Plant Medicine (CSRPM), Ghana [1] and (2) records of plants collected from Ghana downloaded from the GBIF data portal (accessed 11 January 2015). The database on plants of Ghana included a total of 53,509 records captured from Ghanaian herbarium specimen sheets, including 10,765 records of Ghanaian plants from University of Wageningen (WAG). The GBIF data contributed a further 9673 records after cleaning (see below).

Data were cleaned via an iterative series of inspections and visualizations designed to detect and document inconsistencies. First, we created lists of unique names in each dataset in Microsoft Excel, and inspected them for repeated versions of the same taxonomic concepts: misspellings, name variants, different versions of authority information, etc. Such repeated name variants were flagged, checked via independent sources, and corrected to produce single scientific names that correctly referred to single taxa. Second, we checked for geographic coordinates that fell outside of the country, but that were referred to it. Next, within the country, we checked for consistency between textual description of regions (equivalent to provinces or states in other countries) and geographic coordinates. In each case, where possible, we corrected the data record; where no clear correction was possible, we discarded data, recording data losses at each step in the cleaning process. Lastly, we discarded data records for which information on year, month, or day of collection was lacking; we created a unique 'stamp' of time as year_month_day.

We aggregated point-based occurrence data to ½° spatial resolution grid across the country. This choice of spatial resolution was the product of an analysis balancing benefits of aggregating data (e.g., larger sample sizes) versus disadvantages (e.g., loss of spatial resolution across larger areas). The procedure consists of examining the relative change in area-adjusted variance of the data with increasing grid-cell size, and selecting the finest resolution at which the trend of the slope of the overall variance *versus* area curve changed most (Ariño et al., in prep.); it is similar to the concept of selecting the largest sample size beyond which no significant increase in diversity is expected (Ariño et al. 2008).

In this study, a ½° spatial resolution offered the best balance between spatial resolution and inventory completeness, and was consistent with the spatial resolution used in a previous analysis of Brazilian plants (Sousa-Baena et al. 2013) and wild palms of Benin (Idohou et al. 2015). We produced aggregation grid shapefiles in the Vector Grid module of QGIS, version 2.4, added the coarse-resolution grid identification codes to each occurrence datum, and aggregated occurrence data into coarse-resolution aggregation squares. In Excel, we explored relations between species identity, time, and aggregation grid-square. We calculated (1) the total number of records available from each grid

[1] http://herbaria.plants.ox.ac.uk/bol/.

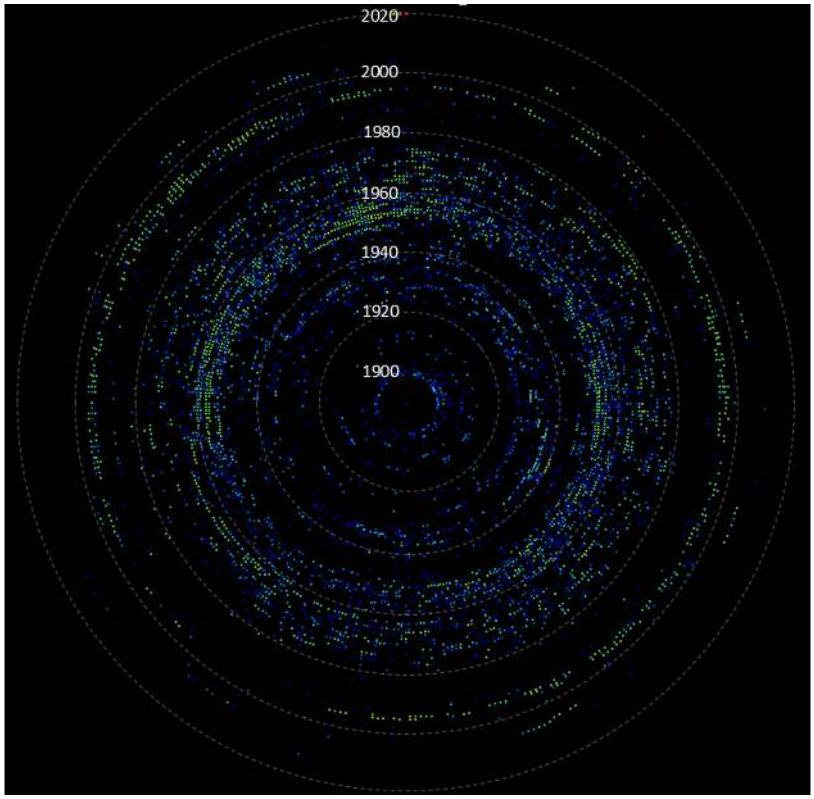

Figure 2. Chronhorogram showing temporal course of accumulation of records of Ghanaian plants. Radial dimension shows year of collection, and angle indicates day of the year. Color of dots: black = no records; a ramp of colors indicates numbers of records, from blue = few to red = many.

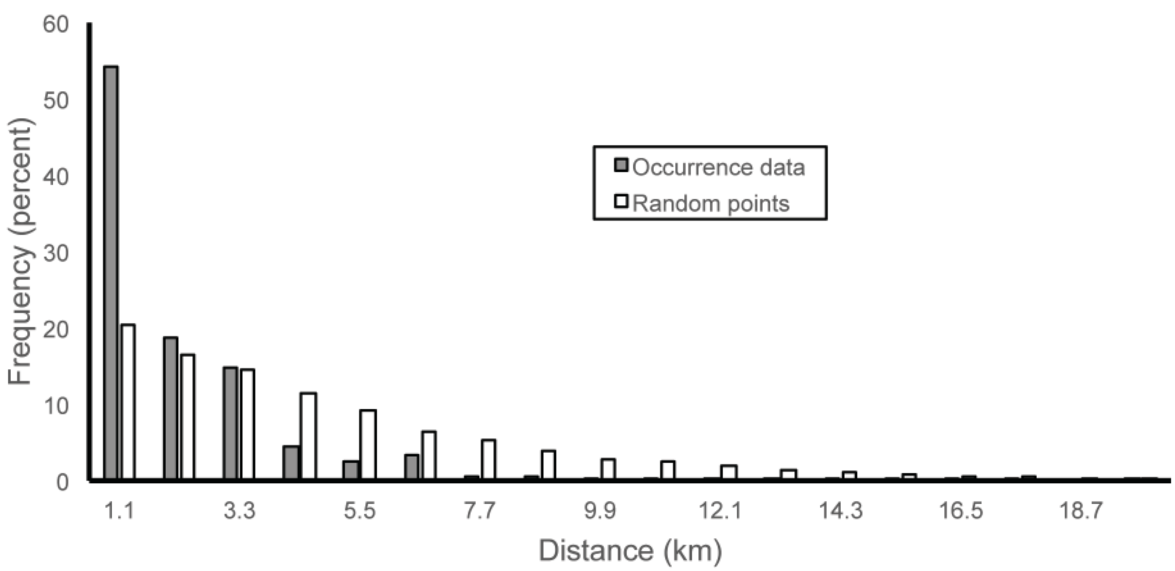

Figure 3. Summary of distances to nearest road for the available DAK for Ghanaian plants (gray bars) and for 5000 random points across the country. The concentration of the DAK along roads (i.e., within 1.1 km of roads) compared to random patterns across the country is clear.

square (termed N); (2) the total number of distinct species recorded from each grid square (S_{obs}); and (3) the number of species detected on one date only (a), and (4) the number of species detected on two dates only (b). Via equations provided by Chao (1987), we calculated the expected number of species (S_{exp}), as

$$S_{exp} = S_{obs} + \frac{a^2}{2b},$$

and inventory completeness (C) as $C = S_{obs} / S_{exp}$. We then explored plots of C *versus* N (Hortal et al. 2007) to assess appropriate and adequate definitions of relatively completely *versus* incompletely inventoried grid squares—we used criteria of $C \geq 0.4$ and $N \geq 1000$ as final definitions of well-inventoried areas.

Once we had established criteria for which grid squares could be considered as well-sampled, in QGIS, we linked the table with the grid square statistics (i.e., N, S_{obs}, S_{exp}, C) to the aggregation grid, and saved this file as a shapefile. Applying the criteria for 'well-sampled,' we created a shapefile of well-sampled grid squares, which we in turn converted to binary-valued (0 = not well-sampled, 1 = well-sampled) raster (geotiff) format using custom scripts in R (R Core Team 2014). This raster coverage was the basis for our identification of gaps, as follows.

We used the Proximity (Raster Distance) function in QGIS to summarize geographic distance across the country to the nearest well-sampled area. To create a parallel view of environmental difference from well-sampled areas, we plotted 5000 random points across the country, and used the Point Sampling Tool in QGIS to link each point to the geographic distance raster, and to raster coverages (2.5' spatial resolution) summarizing annual mean temperature and annual precipitation drawn from the WorldClim climate data archive (Hijmans et al. 2005).

We exported the attributes table associated with the random points, and analyzed further in Microsoft Excel. We first standardized the values of each environmental variable to the overall range of the variable as $(x_i - x_{min}) / (x_{max} - x_{min})$, where x_i is the particular observed value in question, thus rescaling the two variables on the same magnitude of overall variation. We then created a matrix of Euclidean distances in the two-dimensional climate space, relating all of the points with a geographic distance >0 (see above) to all of the points with geographic distance of zero. The latter represent points falling in well-sampled regions, whereas the former are scattered across the entire region; the points in well-sampled regions were assigned (by definition) environmental distances of zero. Finally, the environmental distances were imported into QGIS, and linked back to the random points shapefile.

RESULTS

The DAK of plants of Ghana consisted of a total of 38,400 cleaned and geo-referenced data records covering the period 1830-2012. Number of records per year ranged between 1 and 241, with an average of 46.2 data records per year; ~66% of the records were from the period 1950–1977 (Figure 1). Seasonal patterns in the records showed that most records were from the dry season (October to December), whereas the fewest records were from the rainy season (June to August; Figure 1). The chronological and seasonal pattern in data records can be visualized via the chronhorogram in Figure 2: we observed a rapid increase in numbers of data records between 1940 and 1980, and decreasing numbers of records thereafter.

The data showed an overwhelming tendency towards concentration of records in southern Ghana. Points of access (roads and rivers) were clearly visible in the spatial distribution of records, and indeed the DAK was significantly concentrated close to roads compared to random points (Figure 3). Total number of grid cells across Ghana for ½° resolution was 92; all held data and about 13% were classified as well-known sites. Plots of C against number of records per grid cell showed sample-size dependency in the range of 500-1000 records (Figure 4); hence, we defined well-sampled sites as those having ≥1000 records and $C \geq 0.4$, because stricter criteria would identify massive swaths of territory as not-well-sampled (see Sousa-Baena 2013).

Generally, well-known grid cells were concentrated in the southwestern and southeastern areas of the country, and the largest gaps were in the west-central and northeastern parts of the country (Figures 5 and 6). Climatic conditions are diverse in Ghana, with more homogenous climates in the south compared to those of the central and northern parts of the country; distinct climatic conditions exist in the northern and northwestern areas of the country.

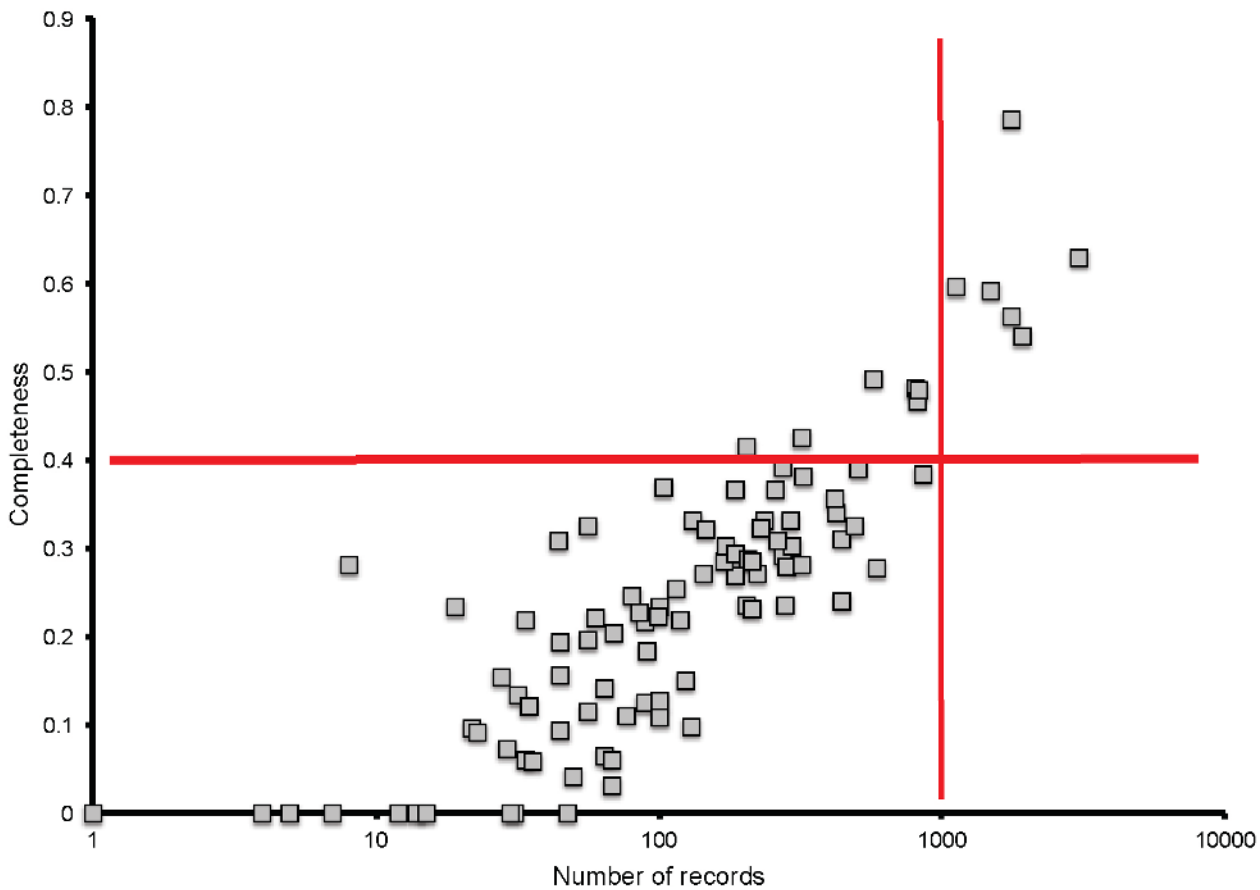

Figure 4. Plot of inventory completeness (C) against sample size (N) for grid cells across Ghana. Red lines indicate criteria for "well-known" with respect to each of the axes.

Climatic differences from well-known cells were most pronounced in a broad swath of the northern part of the country (Figure 6).

Combining these two views identified sites that are both geographically remote and environmentally different from well-known sites (Figure 6). Four areas fit these criteria: (1) northeastern Ghana, including the entire Upper West Region; (2) the north-central part of the Northern Region of the country around Tamale; (3) the west-central part of the country, including parts of the Brong-Ahafo Region around Bui National Park; and (4) the east-central part of Ghana, including parts of the northern Volta Region and the Brong-Ahafo Region, including Digya National Park.

DISCUSSION

Access to primary biodiversity data is critical to addressing challenges of sustainable development and decision-making (Sousa-Baena et al., 2013). Most primary biodiversity data, in excess of 6.5 x 10^8 data records, that have been shared openly are based on biological collections held in herbaria and museums, as well as observational data from citizen scientists. Our focus on DAK emphasizes data that are available to the broader scientific community for analysis and exploration (Sousa-Baena et al. 2013), such that biological collections for which associated data have not been digitized or that are digital but remain broadly unavailable are ignored (Costello et al., 2013). In contrast, information that is open and accessible has potential to impact science and conservation, as well as the care and curation of specimens (Sousa-Baena et al. 2013), such that digitization and sharing of primary biodiversity data is much to be encouraged (see Article 17, Convention on Biological Diversity).

Knowledge of inventory completeness is important to determine appropriate levels of confidence that can be applied to data-derived patterns of biodiversity across a region or a taxon (Soberón and Llorente, 1993; Colwell and Coddington, 1994; Gotelli and Colwell, 2001). Recent analyses have attempted to summarize the state of knowledge of plant diversity (Kier et al., 2005; Mutke and Barthlott, 2005), but few of these studies are based on primary biodiversity data (Soberón et al., 2000; Ariño et al., 2012; Sousa-Baena et al. 2013). Such studies have generally indicated high species richness at small numbers of well-sampled areas, and few sites that are well-known and comprehensively documented, but broad areas that remain poorly sampled. Particularly perplexing is when high species richness sites correspond closely to sites of high sampling intensity, as such situations suggest that "hotspots" in fact represent artifacts of incomplete sampling (Tobler et al., 2007; Ahrends et al., 2011).

In Ghana, few studies have evaluated the state of knowledge of distributions of plants; the few studies existing focused in the forest vegetation zone in the southern parts of the country, based on both herbarium records and observations from sampled plots, and characterized species' distribution patterns (Hall and Swaine, 1981; Hawthorne and Abu-Juam, 1995). Although the savanna vegetation zone covers about two-thirds of Ghana, no assessment has addressed knowledge of distributions of plants there. This study is also the first to be based on digital records that are openly accessible to the broader scientific community. Although DAK for Ghanaian plants should improve with time, both in quantity and quality, numbers of new collections have been decreasing steadily over recent years.

DAK completeness focuses on the consistency of information that is available, and offers a useful index about why a site has few or many species recorded (Soberón et al., 2000; Sousa-Baena et al. 2013). Here, we addressed questions about sites in Ghana where biodiversity knowledge is relatively reliable *versus* where information is incomplete. We found well-known sites principally in southeastern Ghana relatively close to the location of the Ghana Herbarium, where botanists and students have developed intensive collections. Southwestern Ghana is considered richest in terms of plant species in Ghana (Hawthorne and Abu-Juam, 1995); as a consequence, many botanists and indeed many large-scale projects have collected plants from the area. In northwestern Ghana, extensive plant collections have been undertaken around Mole National Park, such that that area is botanically well known. In this study, we identified knowledge gaps, and characterized them in terms of geographic distance and environmental difference from well-known sites, and see these sites as priority areas for botanical sampling. These gaps frequently result from no previous collecting visits to sites, but may in some cases reflect lack of digital access to collections that in truth exist (Sousa-Baena et al. 2013).

Most herbaria in Ghana now have their data

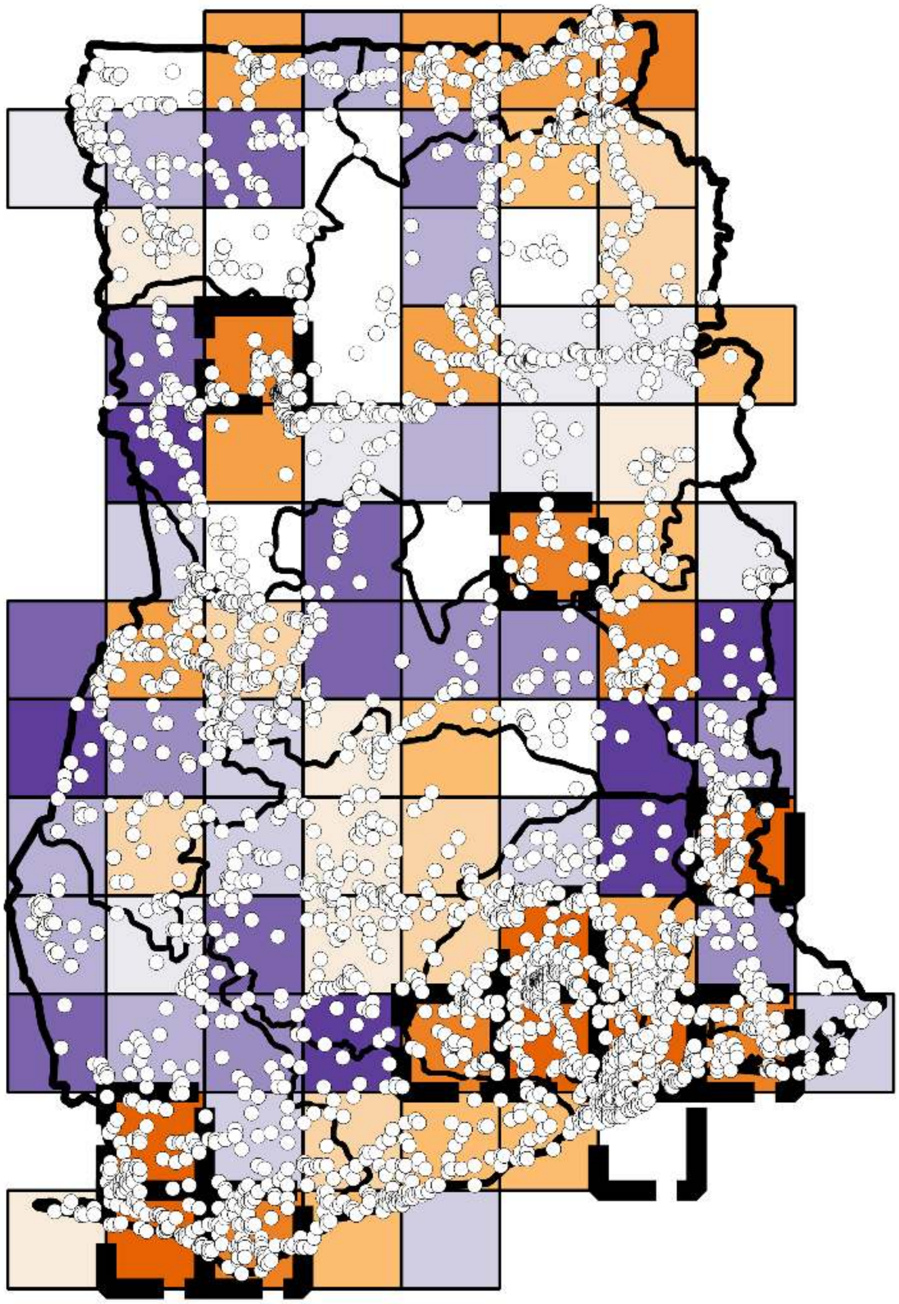

Figure 5. Geographic patterns of inventory completeness across Ghana based on ½° grid squares. Shading indicates inventory completeness (C), in a spectrum from violet (as low as 0) to red-brown (as high as 0.63). Thick dashed black outlines indicate those grid squares that fit the "well-known" criterion of $C \geq 0.4$ and ≥1000 records.

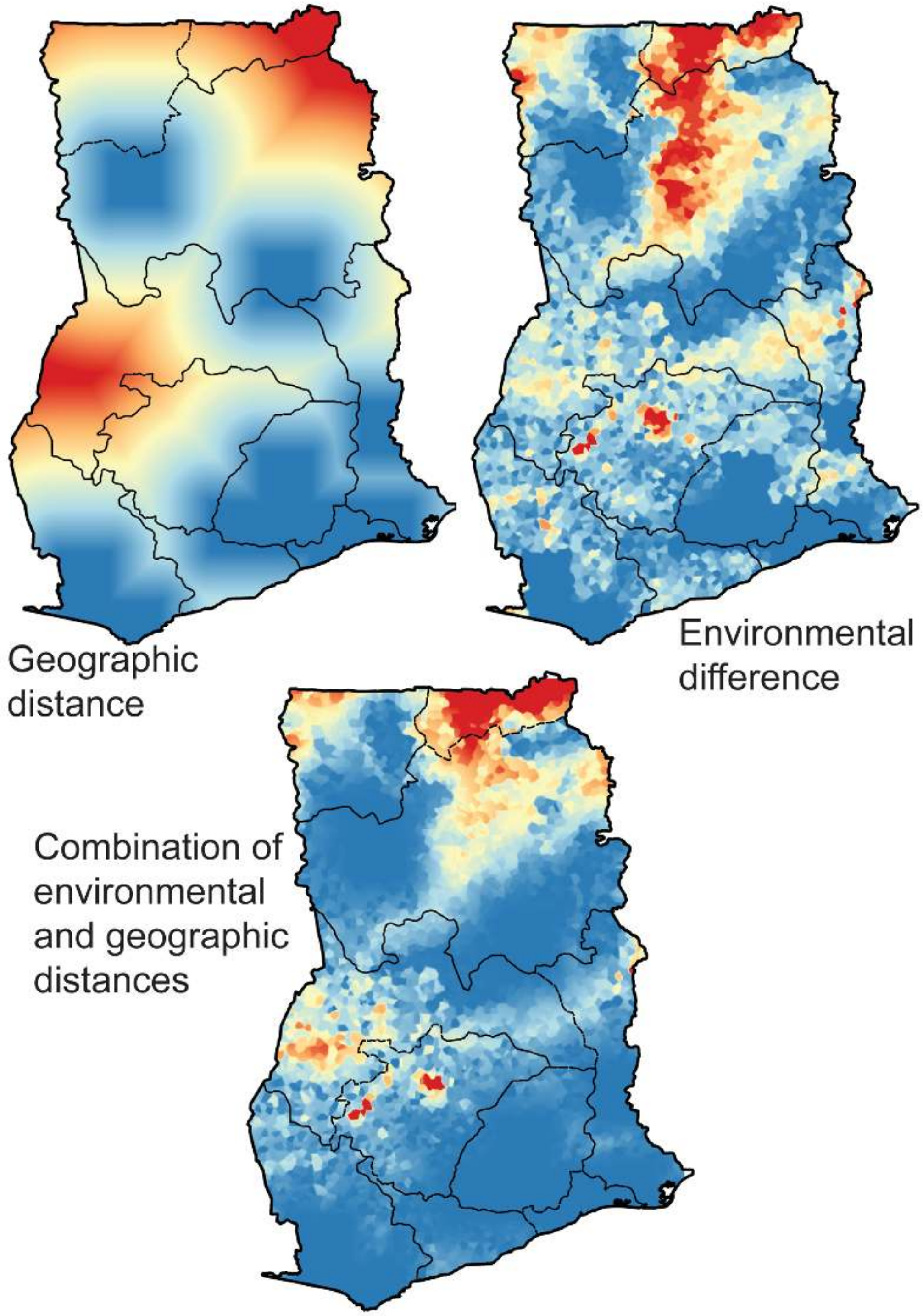

Figure 6. Visualization of geographic distance of ½° pixels across Ghana from well-known sites; climatic difference of ½° pixels across Ghana from well-known sites based on nearest-neighbor distance; and a combination of the two distances based on equal minimum-maximum scaling.

records in digital formats and openly available: ~80% of herbarium specimen sheets have been captured, although carpological and other collections remain to be digitized. In addition, many specimens of Ghanaian provenance are deposited in collections in other countries and remain to be digitized and shared. As such, the fastest way to improve DAK of Ghanaian plants is to fix data "leaks" among existing digital data records (Sousa-Baena et al. 2013). In particular, of the 53,509 records analyzed herein, 60 (0.1%) records had indeterminate names, 24,710 (46.2%) records had incomplete dates (missing year, month, or day), and 538 (1.0%) records lacked geographic coordinates, such that these records could not be included in our analyses. Another area of importance in terms of DAK improvement are the large amounts of botanical data associated with significant collections of Ghanaian plants held elsewhere in the world. Data repatriation from European and North American herbaria with large collections from Ghana is an important potential source of DAK of Ghanaian plants. Collaborative efforts linking West African botanists, and North American and European herbaria are now underway[2], and promise to develop and enable rich new DAK resources.

This study illustrates the importance of assessing completeness of DAK for prioritizing botanical surveys based on existing knowledge and benefits to be reaped from biodiversity data sharing and integration. Field sampling efforts should focus in areas identified herein as both environmentally different and geographically distant from well-known sites. Such information exists for only a few countries, such as Brazil (Sousa-Baena et al. 2013, Idohou et al. 2015, Koffi et al. 2015). This kind of information is important for strategic national policy (Soberón and Peterson, 2009), and is an important step towards meeting the Aichi Targets[3] and national commitments to the Clearing House Mechanism of the Convention on Biological Diversity[4].

Acknowledgments

We are most grateful to JRS Biodiversity Foundation for funding the first author in development of the database of the plants of Ghana. The authors are thankful to several student data capturists, particularly Cindy Asare, Gloria Oppong, Tony Asafo-Agyei, Amos Asase, and Joana Adofo. Preparation of this paper began at a workshop on National Biodiversity Diagnoses at Entebbe in Uganda supported by a JRS Biodiversity Foundation grant to the second author. We are grateful to Arturo Ariño, for developing the chronohorogram. We thank Arturo Ariño, Lindsay Campbell, and Kate Ingenloff for guidance and support in developing these analyses.

Literature Cited

Ahrends, A., N.D. Burgess, R.E. Gereau, R. Marchant, M.T. Bulling, J.C. Lovett, P.J. Platts, V.W. Kindemba, N. Owen, E. Fanning, and C. Rahbek. 2011. Funding begets biodiversity. Div. and Distrib. 17:191-200.

Ariño, A.H., C. Belascoáin, and R. Jordana. 2008. Optimal sampling for complexity in soil ecosystems. Pp. 220-230 in Minai A., and Y. Bar-Yam, Unifying themes in Complex Systems IV. Springer-Verlag, Berlin.

Ariño, A.H., J. Otegui, A. Villarroya, and A. Pérez de Zabalza. 2012. Primary biodiversity data records in the Pyrenees. Env. Eng. & Man. J. 11:1059-1075.

Chao, A.1987. Estimating the population size for capture-recapture data with unequal catchability. Biometrics 43:783-791.

Chapman, A.D. 2005. Uses of Primary Species-occurrence Data, version 1.0. Global Biodiversity Information Facility, Copenhagen.

Colwell, R.K., and J. A. Coddington. 1994. Estimating terrestrial biodiversity through extrapolation. Phil. Trans. Roy. Soc. B 335:101-118.

Costello, M.J., W.K. Michener, M. Gahegan, Z.Q. Zhang, and P.E. Bourne. 2013. Biodiversity data should be published, cited, and peer reviewed. Trends Ecol. Evol. 28:454-461.

Gotelli, N.J., and R. K. Colwell. 2001. Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. Ecol. Lett. 4:379-391.

Graham, C.H., S. Ferrier, F. Huettman, C. Moritz, and A. T. Peterson. 2004. New developments in museum-based informatics and applications in biodiversity analysis. Trends Ecol. Evol. 19:497-503.

Guralnick, R.P., A.W. Hill, and M. Lane. 2007. Towards a collaborative, global infrastructure for biodiversity assessment. Ecol. Lett. 10:663-672.

Hall, J.B., and M. Swaine. 1981. Distribution and Ecology of Vascular Plants in Tropical Rainforest: Forest Vegetation in Ghana. W. Junk, The Hague.

Hawthorne, W. D., and M. Abu-Juam. 1995. Forest Protection in Ghana. International Union for the Conservation of Nature, Gland.

[2] http://jrsbiodiversity.org/grant/university-of-ghana-herbaria/.
[3] https://www.cbd.int/sp/targets.
[4] https://www.cbd.int/chm.

Hijmans, R.J., S.E. Cameron, J.L. Parra, P.G. Jones, and A. Jarvis. 2005. Very high resolution interpolated climate surfaces for global land areas. Int. J. Climatol. 25:1965-1978.

Hortal, J., J.M. Lobo, and A. Jimenez-Valverde. 2007. Limitations of biodiversity databases: Case study on seed-plant diversity in Tenerife, Canary Islands. Cons. Biol. 21:853-863.

Idohou R., A. H. Ariño, A. Assogbadjo, R. Glele Kakai, and B. Sinsin. 2015. Knowledge of diversity of wild palms (Arecaceae) in the Republic of Benin: Finding gaps in the national inventory by combining field and digital accessible knowledge. Biodiv. Inf. 10:45-55.

Kier, G., J. Mutke, E. Dinerstein, T.H. Ricketts, W. Kuper, H. Kreft, and W. Barthlott. 2005. Global patterns of plant diversity and floristic knowledge. J. Biogeogr. 32:1107-1116.

Koffi, K.J., A.F. Kouassi, C.Y. Adou Yao, A. Bakayoko, I.J. Ipou, and J. Bogaert. The present state of botanical investigations in Côte d'Ivoire. Biodiv. Inf. 10:56-64.

Ministry of Environment of Science. 2002. National Biodiversity Strategy of Ghana. Ministry of Environment and Science, Accra.

Mora,C., D.P. Tittensor, and R. A. Myers. 2008. The completeness of taxonomic inventories for describing the global diversity and distribution of marine fishes. Proc. Roy. Soc. B 275:149-155.

Mutke, J., and W. Barthlott. 2005. Patterns of vascular plant diversity at continental to global scales. Biologiske Skrifter 55:521-537.

Nakamura, M., and J. Soberón. 2008. Use of approximate inference in an index of completeness of biological inventories. Cons. Biol. 23:469-474.

Otegui, J., A.H. Ariño, M.A. Encinas, and F. Pando. 2013. Assessing the primary data hosted by the Spanish Node of the Global Biodiversity Informatics Facility (GBIF). PLoS ONE 8:e55144.

Pardo, I., M.P. Pata, D. Gómez, and M. B. García. 2013. A novel method to handle the effects of uneven sampling effort in biodiversity databases. PLoS ONE 8:e52786.

Peterson, A.T., A.G. Navarro-Sigüenza, and H. Benítez-Díaz. 1998. The need for continued scientific collecting: a geographic analysis of Mexican bird specimens. Ibis 140:288-294.

Peterson, A.T., S. Knapp, R. Guralnick, J. Soberón, and M. T. Holder. 2010. The big questions for biodiversity informatics. Syst. Biodiv. 8:159-168.

R Core Team. 2014. R: A Language and Environment for Statistical Computing; http://www.R-project.org/. R Foundation for Statistical Computing, Vienna.

Schalk, P.H. 1998. Management of marine natural resources through biodiversity informatics. Marine Pol. 22:269-280.

Soberón, J., and A.T. Peterson. 2004. Biodiversity informatics: managing and applying primary biodiversity data. Phil. Trans. Roy. Soc. B 359: 689-698.

Soberón, J. and A. T. Peterson. 2009. Monitoring biodiversity loss with primary species-occurrence data: Toward national-level indicators for the 2010 Target of Convention on Biological Diversity. Ambio 38:29-34.

Soberón, J., and J.E. Llorente. 1993. The use of species accumulation functions for the prediction of species richness. Cons. Biol. 7:480-488.

Soberón, J.M., J. B. Llorente and L. Oñate. 2000. The use of specimen-label databases for conservation purposes: An example using Mexican papilionid and pierid butterflies. Biodiv. & Conserv. 9:1441-1466.

Sousa-Baena, M.S., L.C. Garcia, and A.T. Peterson. 2013. Completeness of digital accessible knowledge of the plants of Brazil and priorities for survey and inventory. Div. Distrib. 20:369-381.

Sousa-Baena, M.S., L. C. Garcia, A. T. Peterson. 2014. Knowledge behind conservation status decisions: Data basis for 'Data Deficient' Brazilian plant species. Biol. Conserv. 173:80-89.

Tobler, M.W., E. Honorio, J. Janovec, and C. Reynel. 2007. Implications of collection patterns of botanical specimens on their usefulness for conservation planning: An example of two Neotropical plant families (Moraceae and Myristicaceae) in Peru. Biodiv. & Conserv. 16: 659-677.