

**UNIVERSIDADE FEDERAL DE ALAGOAS
INSTITUTO DE CIÊNCIAS BIOLÓGICAS E DA SAÚDE
Programa de Pós-Graduação em Diversidade Biológica e
Conservação nos Trópicos**

BRUNO UMBELINO DA SILVA SANTOS

**Mapeando a perda de diversidade florística na
Amazônia brasileira em função do desmatamento
histórico e futuro**

**MACEIÓ - ALAGOAS
Fevereiro/2018**

BRUNO UMBELINO DA SILVA SANTOS

**Mapeando a perda de diversidade florística na
Amazônia brasileira em função do
desmatamento histórico e futuro**

Dissertação apresentada ao Programa de Pós-Graduação em Diversidade Biológica e Conservação nos Trópicos, Instituto de Ciências Biológicas e da Saúde da Universidade Federal de Alagoas, como requisito para obtenção do título de Mestre em CIÊNCIAS BIOLÓGICAS, área de concentração em Conservação da Biodiversidade Tropical.

MACEIÓ - ALAGOAS
Fevereiro/2018

**Catalogação na fonte
Universidade Federal de Alagoas
Biblioteca Central**

Bibliotecário Responsável: Janis Christine Angelina Cavalcante

S237m Santos, Bruno Umbelino da Silva.

Mapeando a perda de diversidade florística na Amazônia brasileira em função do desmatamento histórico e futuro. / Bruno Umbelino da Silva Santos. – 2018.

54f. : il. color.

Orientadora: Ana Cláudia Mendes Malhado.

Coorientadores: Juliana Stropp; Ricardo Correia.

Dissertação (Mestrado em Diversidade e Ecologia de Organismos Tropicais) – Universidade Federal de Alagoas. Instituto de Ciências Biológicas e da Saúde, 2018.

Bibliografia: f. 49-52.

1. Diversidade florística. 2. Desmatamento – Amazônia - Brasil.

3. Biodiversidade. 4. Plantas com flores – Conservação de espécie.

I. Título.

CDU: 574

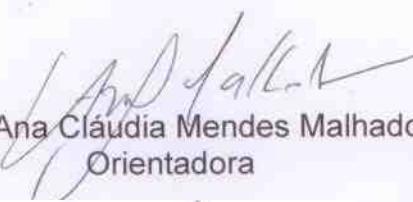
Folha de aprovação

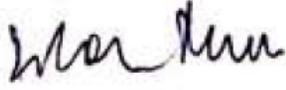
Bruno Umbelino da Silva Santos

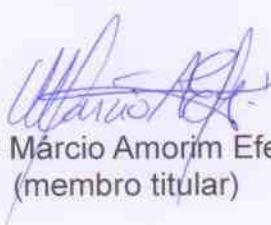
Mapeando a perda do conhecimento biológico na Amazônia em função do desmatamento histórico e futuro

Dissertação apresentada ao Programa de Pós-Graduação em Diversidade Biológica e Conservação nos Trópicos, Instituto de Ciências Biológicas e da Saúde, Universidade Federal de Alagoas, como requisito para obtenção do título de Mestre em CIÊNCIAS BIOLÓGICAS, área de concentração em Diversidade e ecologia de organismos tropicais.

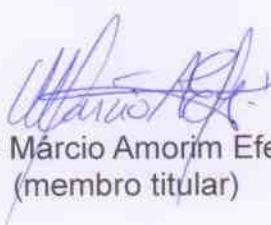
Dissertação aprovada em 27 de fevereiro de 2018.


Profa. Dra. Ana Cláudia Mendes Malhado/UFAL
Orientadora


Profa. Dra. Débora Pignatari Drucker/EMBRAPA-SP
(membro titular)


Maíra Benchimol de Souza

Profa. Dra. Maíra Benchimol de Souza/UESC
(membro titular)


Prof. Dr. Márcio Amorim Efe/UFAL
(membro titular)

MACEIÓ - AL
Fevereiro / 2018

O presente trabalho de dissertação foi realizado no Laboratório de Conservação do Século XXI - (LACOS 21) do Instituto de Ciências Biológicas e da Saúde da Universidade Federal de Alagoas, e teve as seguintes fontes finanziadoras:

Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Ministério da Ciência e Tecnologia, Brasil.

Global Biodiversity Information Facility – GBIF através do programa Jovens Pesquisadores 2016.

Orientação: Dr^a Ana Cláudia Mendes Malhado

Co-orientação: Dr^a Juliana Stropp & Dr Ricardo Correia

A ciência colaborativa.

AGRADECIMENTOS

Ao universo pelas oportunidades de ser alguém melhor em constante processo de evolução e aprendizado

A Ana Malhado, Ricardo Correia e Juliana Stropp pelo aprendizado e amadurecimento nesses dois anos de orientação e oportunidades oferecidas ao longo desta caminhada.

Ao João Campos Silva pela paciência e por te me mostrado que fazer ciência é algo que tem que ser prazeroso.

Ao Richard pelos livros publicados que me forneceram base conceitual.

A Thainá Lessa e Jhonatan Guedes pelas contribuições em correções textuais.

A Janisson Santos, Thainá Lessa e Felipe Alexandre no apoio aos mapas.

A Davi Teles e Nicolli Albuquerque no apoio ao R.

A Garrett Tusler pelo apoio no inglês.

Aos membros da qualificação Marcos Vital, Nídia Fabré e Rafael Loyola e os membros da banca de defesa Marcio Efe, Maíra Benchimol e Debora Drucker pelas contribuições de melhorias para a pesquisa.

A Deividy Lima, sempre presente nos momentos difíceis, apoio e incentivo.

A minha família de amigos Val Barbosa, Gabi Harumi, Betânia Avelar, Maria Flor, Valentim Barbosa, Ingrid Mendes, Debora Gomes, Fabrício Salvioni e Maiara Carvalho que me deram apoio de longe vibrando boas energias da Amazônia.

Aos meus familiares de sangue

A todos os colegas do Laboratório de Conservação do Século XXI da UFAL e do Herbário Rondoniensi da Universidade Federal de Rondônia

Aos docentes do PPG-DiBiCT, pelos esforços em melhorar as pesquisas em biodiversidade.

RESUMO

A informação digital sobre a ocorrência de espécies no tempo e no espaço é essencial para fins científicos e de conservação, mas geralmente é incompleta e tende a se degradar com o tempo. Em regiões como a Amazônia, onde grandes áreas já foram desmatadas, o conhecimento existente da biodiversidade pode não representar as condições atuais. Nesse trabalho, temos como objetivos: (1) Caracterizar quão completos e precisos são os dados de plantas terrestres para Amazônia brasileira disponíveis em plataformas de acesso livre; (2) Investigar quando e onde os registros botânicos ocorreram ao longo do tempo; (3) Avaliar o quanto que o conhecimento sobre a diversidade florística foi degradado pelo desmatamento histórico e futuro. Nós compilamos os registros de ocorrências de espécies da classe Magnoliopsida para a região da Amazônia Brasileira a partir de três bancos de dados de acesso livre. Calculamos a completude dos registros temporal e espacial disponíveis na região e como sua distribuição é afetada pelo desmatamento, passado e futuro. Do total de 608.451 registros pertencentes à Amazônia brasileira, aproximadamente 38% dos registros permaneceram com informações completas, com base nos critérios estabelecidos neste estudo. Todas as unidades amostrais com registros somaram 3.227, 10% foram amostras florísticas contendo pelo menos dois registros. Após a influência do desmatamento histórico (até 2016), perdemos 49 unidades/áreas amostradas com valor de completude $\geq 0,5$, em áreas no noroeste e no centro da Amazônia brasileira. No cenário de desmatamento futuro (2050), não teremos unidades amostrais bem representadas na Amazônia brasileira, de acordo com os dados utilizados nesse estudo. As unidades amostrais sem dados digitalmente disponíveis correspondem a 42% do total, e aproximadamente 12% foram áreas com diversidade florística perdida no atual cenário de desmatamento. Precisamos saber mais sobre a diversidade florística nas áreas não amostradas ou com baixo número de coletas, como as terras indígenas e áreas que poderão ser desmatadas. A comunidade científica deve avaliar a qualidade e a completude dos dados disponíveis para as ações de mitigação de impactos a biodiversidade.

Palavras-chave

Amazônia, qualidade dos dados, desmatamento, plantas com flores, Global Biodiversity Information Facility, SpeciesLink, SiBBr, viés temporal, dados de ocorrência de espécies.

ABSTRACT

Digital information on the occurrence of species in time and space is essential for scientific and conservation purposes, but is generally incomplete and tends to degrade over time. In regions such as the Amazon, where large areas have already been deforested, existing knowledge may not represent current conditions. Here, we explore: (1) To characterize how complete and accurate are the plant data for Brazilian Amazonia available on free access platforms; (2) Investigate when and where botanical records occurred over time; (3) Evaluate how much floristic diversity has been degraded by historical and future deforestation. We compiled records of species occurrences of the Magnoliopsida class for the Brazilian Amazon region from three free access databases. We calculate the completeness of the temporal and spatial records available in the region and how their distribution is affected by deforestation, past and future. Of the total of 608,451 records belonging to the Brazilian Amazon, approximately 38% of the records remained with complete information, based on the criteria established in this study. All sample units with records totaled 3,227, 10% were floristic samples containing at least two records. After the influence of historical deforestation (until 2016), we lost 49 sampled areas / areas with a value of completeness ≥ 0.5 , in areas in the northwest and central Brazilian Amazon. In the future deforestation scenario (2050), we will not have representative sample units in the Brazilian Amazon, according to the data used in this study. Sample areas without knowledge with the digitally available data correspond to 42% of the total, and approximately 12% were areas with floristic diversity lost in the current deforestation scenario. We need to know more about floristic diversity in non-sampled or low collection areas such as indigenous lands and areas that will be lost. The local and regional scientific community, together with the international community, should assess how comprehensive and accurate the available data are for mitigating biodiversity impacts.

Keywords

Amazon, data quality, deforestation, flowering plants, Global Biodiversity Information Facility, SpeciesLink, SiBBr, temporal bias, species occurrence data.

Lista de figuras

Figure 1 Spatial evolution of species occurrence data through the years on Brazilian Amazon between 1656 and 2016. For the period of 1656 until 1976 we had 60.348 records valid; 1656 until 1986 about 120.339; 1656 until 1996 had 162.976; 1656 until 2006 were 188.395 and from 1656 until 2016 we had 228.548 records valid.....	23
Figure 2 Distribution of sampled units according to completeness values ($C_i \geq 0,5$ and $C_i \geq 0,7$) (a) and (c) are according to $C_i \geq 0,5$ near to highways and protected areas and indigenous land. (b) and (d) is according well sampled units ($C_i \geq 0,7$).....	25
Figure 3 Sampled units and well sampled units lost by historical deforestation. (a) and (c) is about remain sampled units (blue points) near to highways and in protected areas and indigenous land. (b) and (d) is about well sampled units already lost by deforestation.....	26
Figure 4 Lost sampled units lost by future deforestation. (a) is about remain samples units near to highways and (b) about near or inside of protected areas and indigenous land.	27
Figure 5 Areas lost by historical and future deforestation.....	29

Sumário

1 APRESENTAÇÃO	1
2 REVISÃO DA LITERATURA	2
2.1 LACUNAS DO CONHECIMENTO SOBRE A BIODIVERSIDADE	2
2.2 SISTEMAS DE INFORMAÇÃO	4
2.3 CONHECIMENTO FLORÍSTICO DA AMAZÔNIA	6
2.4 DESMATAMENTO	8
REFERÊNCIAS	11
3 JUSTIFICATIVA	15
4 MAPPING FLORISTIC KNOWLEDGE LOSS IN THE BRAZILIAN AMAZON DUE TO HISTORICAL AND FUTURE DEFORESTATION	16
4.1 INTRODUCTION	16
4.2 METHODS	18
4.2.1 Data acquisition	18
4.2.2 Inventory completeness	19
4.2.3 Deforestation data and species-occurrence losses	20
4.3 RESULTS	21
4.3.1 Initial characterization	21
4.3.2 Knowledge losses	21
4.3.3 Spatial distribution of inventory completeness	24
4.3.4 Knowledge and data loss	28
4.4 DISCUSSION	29
Acknowledgments	32
REFERENCES	34
5 CONCLUSÃO GERAL	38

1 APRESENTAÇÃO

O presente trabalho está organizado em dois capítulos principais. O Capítulo I refere-se à revisão da literatura, e é subdividido em quatro partes que ajudarão o leitor a compreender o tema da pesquisa: a primeira aborda as lacunas de conhecimento, a segunda sobre a importância dos sistemas de informação de dados de biodiversidade; a terceira sobre a biodiversidade amazônica e finalmente a quarta e última parte, que aborda o desmatamento na região da Amazônia brasileira, enfatizando a perda de biodiversidade. O Capítulo II é apresentado em formato de artigo científico, no formato da revista “Global Ecology and Biogeography”, intitulado como “*Mapping biodiversity knowledge loss in the Amazon due to historical and future deforestation*”.

2 REVISÃO DA LITERATURA

2.1 LACUNAS DO CONHECIMENTO SOBRE A BIODIVERSIDADE

Os registros de ocorrência de espécies são a base para muitos trabalhos de biogeografia, ecologia, evolução e conservação (Hortal *et al.* 2015). No entanto, apesar de tal conhecimento sobre a biodiversidade ser amplamente descrito e o detalhamento do registro ser fundamental para qualquer tipo de avaliação, nem sempre os dados de ocorrência de espécies são precisos (Ladle & Whittaker 2014). Podemos definir como lacunas de conhecimento em biodiversidade a assimetria entre o conhecimento realizado e o conhecimento completo em uma determinada escala de tempo. Hortal *et al.* (2015) sugerem que as deficiências de dados da biodiversidade podem ser amplamente agrupadas em sete categorias principais, as quais correspondem às linhas de conhecimentos acerca da sistemática, biogeografia, biologia populacional, evolução, ecologia funcional, tolerâncias abióticas e interações ecológicas.

A lacuna de conhecimento sobre a distribuição de espécies é descrita como déficit Wallaceano (Lomolino 2004; Ladle & Whittaker 2014, p. 113; Hortal *et al.* 2015). Alguns autores consideram também a dinâmica geográfica das forças de extinção nessa lacuna, especialmente a dinâmica geográfica das civilizações humanas (Lomolino 2004). Esta escassez de informação se dá pela deficiência de trabalhos de inventários de espécies, sobretudo em países em desenvolvimento onde faltam recursos necessários para execução de inventários de campo e trabalhos de identificação. Por exemplo, o déficit Wallaceano é particularmente elevado nas regiões remotas e inacessíveis, como as florestas do sudoeste da Amazônia e na bacia do Congo (Hortal *et al.* 2015). Estima-se que aproximadamente 40% da Amazônia nunca foi pesquisada e que para algumas espécies de plantas a distribuição geográfica não apresenta ocorrência precisa nesta região (Schulman *et al.* 2007).

O grande problema quanto às estimativas de riquezas de espécies é explicado pelo déficit Lineano (Lomolino 2004; Hortal *et al.*, 2015) referindo-se à quantidade de espécies formalmente descritas e o número real de espécies. Este déficit é alto para os organismos com menor tamanho corporal e

amplitude do nicho ecológico, visto que esse padrão se mantém dentro de grupos taxonômicos (Riddle *et al.*, 2011). Entretanto, a magnitude desse déficit ainda é desconhecida, pois o número de espécies formalmente descrito passa por mudanças em virtude de novas descrições e revisões (Baselga *et al.*, 2010, May 2010). Outro ponto a ser considerado é a sensibilidade do número de espécies estimado pelos métodos de estimativas e os valores dos parâmetros (May 2010).

Mesmo com a variedade de métodos de estimativas para levantamento de espécies, os dados são passíveis de degradação temporal, e destinar um uso correto para os mesmos é de extrema importância para o uso no presente e no futuro (Tessarolo *et al.*, 2017). Essas perdas temporais podem ser causadas por dinamismo dos sistemas naturais, como extinções locais, imigração, invasões biológicas (Tessarolo *et al.*, 2017); em função das mudanças de categorias da biodiversidade de relações taxonômicas adotada pelos cientistas (Ladle & Hortal 2013); e por mudanças na distribuição e composição das espécies (Dornelas *et al.*, 2013). Tais tendências temporais têm o potencial de dificultar a descrição de variações temporais na biodiversidade e, assim, limitar a nossa capacidade de modelar os efeitos dos impactos nas mudanças globais e avaliar os resultados dos modelos dinâmicos de distribuição de espécies (Tessarolo *et al.*, 2017).

No entanto, enquanto a influência do tempo na degradação dos dados é bem conhecida em outros campos científicos e incorporada com sucesso em protocolos para avaliar a qualidade dos dados (Kennedy *et al.*, 2014). A qualidade potencial dos dados de biodiversidade não teve a mesma atenção ao longo do tempo, possivelmente por haver vários conjuntos de dados existentes na atualidade (Tessarolo *et al.*, 2017). Identificar estas lacunas de informação e os fatores que limitam a disseminação de informação sobre biodiversidade é prioridade, tanto em níveis políticos quanto científicos (Meyer *et al.*, 2015). Este ponto é de particular interesse para a conservação, uma vez que a escassez de dados relevantes sobre os números de espécies e distribuição dificultam a precisão das ações conservacionistas (Canhos *et al.*, 2015).

2.2 SISTEMAS DE INFORMAÇÃO

A disponibilidade de registros de espécies e ocorrência geográfica em plataformas na internet é de grande importância para a existência de uma rede de colaboradores de informação, sejam instituições governamentais (i.e. universidades e institutos federais) ou não (i.e. ONGs). Dentro deste ambiente, dados e conhecimentos compartilhados são mais propensos a serem utilizados, não só para o desenvolvimento científico, mas também nas políticas ambientais (Canhos *et al.* 2015).

Os sistemas de informações podem ter vários níveis de complexidade, desde um simples sistema de arquivamento de documentos até sistemas bem complexos que além de compilar também apresentam ferramentas para prover diferentes informações para os usuários. Com o advento da bioinformática, novas ferramentas e técnicas para armazenar, processar e comunicar dados biológicos permitiram a criação de grandes bancos de dados sobre biodiversidade com dados descritivos que podem incluir registros de comportamento, ocorrência, ecologia, morfologia, fisiologia, hábitos, ciclo de vida, evolução entre outros (Meyer *et al.* 2015). O termo bioinformática, que antes era usado apenas nas áreas genômicas, se expandiu com outro foco para as demais áreas científicas como biogeografia, macroecologia e conservação. No Brasil, foi observado um aumento de 74% de descrição de novas espécies vegetais, a partir da criação do Herbário Virtual pelo Ministério da Ciência, Tecnologia e Inovação (MCTI) que conta com investimento de recursos humanos na área de taxonomia (Canhos *et al.* 2015). Entretanto não se sabe se esse aumento se deu pela visibilidade de trabalhos realizados antes da criação da plataforma ou pela real descrição de novas espécies.

Tais sistemas de informação facilitam a acessibilidade de dados sobre a biodiversidade em todo o mundo promovida por museus, jardins botânicos, universidades e demais instituições que possuem coleções biológicas (Sunderland 2013). Esta facilidade logo teve apoio da comunidade de pesquisadores para fornecer plataformas de acesso livre dos dados de biodiversidade. No entanto, nem todos os grupos de pesquisas disponibilizam

dados nessas plataformas, ou permitem que outros pesquisadores utilizem os dados.

Mesmo que os dados disponibilizados sejam de fácil acesso, devem-se ter critérios quanto à utilização dos dados incompletos para uso específico (Veiga *et al.*, 2017). Por exemplo, avaliações inapropriadas quanto ao uso dos dados incompletos e a má gestão dos mesmos, comprometem estudos de modelagem das distribuições de espécies (Veiga *et al.*, 2017), e a quantidade de dados disponíveis pode não ser apropriada para determinadas abordagens (Beck *et al.*, 2014). Um exemplo claro disso é a precisão dos cálculos para estimar a biodiversidade local ou regional. Em função da necessidade de dados biogeográficos serem robustos para o desenvolvimento da biogeografia da conservação, esforços para mitigar os déficits Lineano e Wallaceano impulsionaram diversas iniciativas para superar esses desafios (Ladle & Whittaker 2014). Dentre elas podemos destacar “O catálogo da vida” (<http://www.catalogueoflife.org/>), que providencia informações sobre sinônimos das espécies e pode auxiliar na identificação do nome mais correto e aceito atualmente. Os registros de ocorrência de espécies também têm recebido atenção a nível global, como a plataforma Global Biodiversity Information Facility (GBIF: <https://www.gbif.org/>). No Brasil ocorrem iniciativas semelhantes como o SpeciesLink (<http://www.splink.org.br/>) criado em 2002 e SiBBr (<http://www.sibbr.gov.br/>) mais recente, criado em 2010.

Mesmo que essas iniciativas globais e nacionais forneçam dados importantes para o desenvolvimento de pesquisas, ainda é necessário desenvolver ferramentas mais robustas para aplicações específicas, como a conservação, manejo e uso sustentável da biodiversidade (Ladle & Whittaker 2014). A disponibilização desses não anula a importância do levantamento do conhecimento em campo e em melhorias para determinações taxonômicas e filogenéticas, bem como nas áreas da ecologia e autoecologia (Ladle & Whittaker 2014). Além disso, é importante que os gerenciadores das plataformas de acesso livre reforcem os critérios de qualidade dos metadados a serem disponibilizados, assim como os pesquisadores devem ter a preocupação de filtrar os dados incompletos a fim de evitar ruídos e tendências errôneas (Hopkins *et al.*, 2007; Hortal *et al.*, 2007). Coordenadas geográficas,

nome científico e ano de coleta são exemplos de informações essenciais em dados sobre biodiversidade e a sua presença indica a qualidade dos metadados.

2.3 CONHECIMENTO FLORÍSTICO DA AMAZÔNIA

O Brasil é considerado o país com maior riqueza de espécies em escala global, abrangendo cerca de 20% da diversidade de angiospermas, possuindo a flora mais rica e maior endemismo de plantas vasculares (56%) entre os países da região neotropical (Sobral & Stehmann, 2009; Forzza *et al.*, 2012). Mesmo não existindo um consenso quanto ao número de espécies de plantas, estudos estimaram que cerca de 40 a 50 mil espécies são encontradas em todo território brasileiro (Shepherd 2003; Lewinsohn & Prado 2005) e só para a Amazônia estima-se que a região possua cerca de 40 mil espécies vasculares de plantas, das quais aproximadamente 30 mil são endêmicas (Mittermeier *et al.*, 2003). Embora possua importância global, regional e local, o conhecimento gerado na Amazônia ao longo do tempo ainda é escasso considerando seus limites territoriais (Hopkins *et al.*, 2007; Malhado *et al.*, 2013).

A Amazônia se estende por nove países: Bolívia, Brasil, Colômbia, Equador, Guiana, Guiana Francesa, Peru, Suriname e Venezuela (WWF 2017), com aproximadamente 6,74 milhões km². No Brasil, é o maior bioma (60,1%) brasileiro cobrindo uma área de aproximadamente 4.198.551 km² se dividindo em nove estados: Acre, Amapá, Amazonas, Pará, Rondônia, Roraima, Mato Grosso, Maranhão e Tocantins. Segundo a WWF (2017), já se tem classificado cientificamente 427 espécies de mamíferos, 1.294 de aves, 378 de répteis, 427 de anfíbios e 3.000 de peixes. Quanto aos artrópodes, são conhecidas 1.800 espécies de borboletas na Amazônia, 3.000 espécies de formigas, 2.500 a 3.000 espécies de abelhas na Amazônia (MUSEO EMILIO GOELDI, 2018).

Diversos fatores podem determinar a distribuição e composição da biodiversidade na Amazônia. As condições edafoclimáticas, com períodos secos e chuvosos exercem maiores influências na biodiversidade na escala regional, enquanto que em menor escala, as interações bióticas e abióticas influenciam com maior força sobre as espécies (ter Steege *et al.*, 2003;

Malhado *et al.*, 2013). O padrão populacional encontrado para a maioria das espécies arbóreas amazônicas é raro, sendo muitas sensíveis a qualquer alteração de habitat (Terborgh *et al.*, 1990; Thiollay, 1994). Nesse contexto, as espécies consideradas raras são aquelas que aparecem em baixa densidade nos métodos de amostragem, uma vez que o tamanho da área amostral pode vir a influenciar na quantidade de espécies localmente raras.

Os inventários florísticos de longa duração são extremamente importantes para aumentar o conhecimento sobre os padrões de diversidade e distribuição da flora amazônica. De acordo com estimativas do levantamento de parcelas permanentes (ter Steege *et al.*, 2013), foram coletadas 4.962 espécies (com nome validado) em 1.170 parcelas, e os autores estimaram que existam 16 mil espécies de plantas arbóreas na Amazônia. Um número maior de espécies foi encontrado por ter Steege *et al.*, (2016) ao investigarem os registros de ocorrências de espécies depositados em plataformas online, GBIF e SpeciesLink respectivamente. Os resultados obtidos por eles foram 11.187 espécies com nomes válidos pertencentes a 530.025 coleções entre os anos de 1707 e 2015. De acordo com o mesmo estudo, apenas 489 espécies possuíam ocorrência confiável para Amazônia do total de espécies verificadas (11.676).

Entretanto, os dados apresentados no estudo de ter Steege *et al.*, (2016) consideram as espécies não amazônicas e cultivadas para região (Cardoso *et al.*, 2017), superestimando assim os valores de riquezas. Para Cardoso *et al.*, (2017) aproximadamente 4.670 dos 11.676 dos nomes totais listados não estão corretos pelas seguintes situações: espécies comprovadamente não amazônicas, espécies listadas mais de uma vez com sinônimos e variantes de ortografia, listagem de espécies não específicas para árvores, inclusão de espécies do Velho Mundo devido a erros de banco de dados e a inclusão de espécies cultivadas não-amazônicas. Com isso a lista de 11.676 espécies de ter Steege *et al.*, (2016) cairia para 7.006 espécies arbóreas.

Mesmo que haja inconsistências sobre o número real de espécies de plantas para a região, o viés amostral de coleta é bem documentado na literatura. Estudos realizados por Nelson *et al.* (1990) mostraram uma forte correlação entre as áreas consideradas “centros de riquezas e endemismo” com áreas de intensa amostragem. Padrão semelhante foi observado por

Oliveira *et al.*, (2016) ao avaliar a distribuição de aves, aranhas e angiospermas para os biomas brasileiros, em que a média de distância entre os registros e acessibilidade foi entre 1-5 km.

A falta de dados e/ou dados incompletos ou enviesados podem interferir nos resultados de estudos que não procuram minimizar essas influências, prejudicando no conhecimento sobre a biodiversidade e no conhecimento biogeográfico (Malhado *et al.*, 2013). Iniciativas de redes de pesquisas na Amazônia, como Rainfor (<http://www.rainfor.org/pt>), Amazon Tree Diversity Network – ATDN (<http://atdn.myspecies.info/>) e o próprio Serviço Florestal Brasileiro (<http://www.florestal.gov.br/>) têm demonstrado grandes incentivos no avanço do conhecimento florístico nas regiões onde realizam monitoramentos e avaliações da ameaça da biodiversidade.

2.4 DESMATAMENTO

Os maiores causadores de perda de biodiversidade na Amazônia são o desmatamento, degradação (queimadas, caça predatória, exploração madeireira sem o manejo rotativo de áreas, entre outros) e as mudanças climáticas (Fearnside, 1999) e também é onde tem as maiores taxas acumuladas de desmatamento, aproximadamente 80% (Soares-Filho *et al.*, 2006). Nas extensões sul e sudoeste dos limites da floresta que fazem divisa com o bioma Cerrado são onde se encontram as maiores perdas de habitat, região esta conhecida como “arco do desmatamento” (Morton *et al.*, 2006). A abertura de estradas não pavimentadas facilita a criação de novas frentes ilegais de extração florestal por madeireiros (Soares-Filho *et al.*, 2006), assim como a ocupação ilícita em glebas públicas federais para expansão agropecuária, ou seja, grilagem em terras públicas (MMA 2018).

A desaceleração da perda florestal se dá por diversos fatores, dentre os quais podemos destacar as políticas públicas contra o desmatamento, a participação de ONGs e a influência das práticas econômicas locais (Nepstad *et al.*, 2014). Monitoramento via imagens de satélite vem contribuindo para conter o desmatamento, como o projeto PRODES desenvolvido pelo Instituto

de Pesquisas Espaciais (INPE). Este monitoramento usa como referência o desmatamento por corte raso em toda a extensão da Amazônia Legal e produz, desde 1988, relatórios anuais das taxas de desmatamento dentro e fora de unidades de conservação e terras indígenas na região, que são usadas pelos tomadores de decisões para o estabelecimento de políticas públicas (INPE 2018).

As estimativas geradas a partir dos dados do monitoramento PRODES são consideradas confiáveis pelos cientistas tanto no âmbito nacional quanto internacional (INPE 2018). O PRODES utiliza imagens provenientes de satélites da classe LANDSAT (20 a 30 metros de resolução espacial) que busca minimizar os problemas de cobertura de nuvens (INPE 2018). Entretanto, comparações entre os sistemas PRODES e Global Forest Change (GFC) mostraram diferenças significativas quanto à precisão do monitoramento de desmatamento (Richards *et al.*, 2017). Os mesmos autores destacam que, desde 2008 o sistema PRODES se tornou menos representativo nas estimativas para mitigação de Gases de Efeito Estufa (GEE) oriunda do desmatamento que poderia ser evitado. Os mesmos autores ainda apontam as melhorias no sistema PRODES quanto às detecções de perda de floresta não-antropogênica, florestas secundárias, florestas nas regiões mais secas da bacia e degradação florestal em pequenas clareiras. Mesmo com as falhas no monitoramento em algumas regiões, o PRODES se mostrou uma ferramenta útil para combater o desmatamento. De 1988 a 2017, 428.398 km² de áreas foram perdidas pelo desmatamento acumulado (INPE - <http://www.obt.inpe.br/prodes/dashboard/prodes-rates.html>).

Quanto às políticas públicas, o governo brasileiro aprovou em 2004 o Plano de Ação para Prevenção e Controle de Desmatamento na Amazônia Legal (PPCDAM), dividido em quatro fases: o ordenamento fundiário e territorial, monitoramento e controle ambiental, fomento das atividades produtivas sustentáveis da região e instrumentos econômicos e normativos (MMA 2018). Após a aprovação do plano, a taxa anual de 27.772 km² em 2004 passou para 7.989 km² em 2016, com redução de 70% em uma escala de 10 anos. Entretanto esses números não devem ser comemorados, visto que entre

os anos de 2015 e 2016, a taxa de desmatamento aumentou em 1,782 km² (MMA 2018).

Outras iniciativas também foram tomadas pelo governo brasileiro no âmbito de políticas públicas, como o desenvolvimento do Sistema Nacional de Unidades de Conservação, no ano de 2000; expansão do Programa de Redes de Áreas Protegidas da Amazônia (ARPA) em unidades de conservação com suporte financeiro do Fundo Brasileiro para a Biodiversidade (Funbio) (Nepstad *et al.*, 2014); o programa de Redução de Emissões provenientes de Desmatamento e Degradação Florestal (REDD+), que tem como objetivo mitigar os feitos das mudanças climáticas por meio da eliminação do desmatamento ilegal, da conservação e da recuperação dos ecossistemas florestais e do desenvolvimento sustentável econômico de baixo carbono (MMA 2018); Cadastro Ambiental Rural (CAR) auxiliando no planejamento de imóveis rurais e em recuperação de áreas degradadas, bem como na formação de corredores ecológicos e a conservação dos demais recursos naturais (MMA 2018); Bolsa Verde que beneficiam famílias em situação de pobreza em áreas para a conservação ambiental, respeitando as regras de utilização dos recursos (MMA 2018). O grande desafio do Brasil é saber aproveitar os recursos naturais sem destruir massivamente as áreas florestais. A valorização da permanência das áreas florestais parece não ter destaque ou importância na economia, seja regional ou local (Becker 2005).

REFERÊNCIAS

- Baselga, A., Lobo, J. M., Hortal, J., Jiménez-Valverde, A., & Gómez, J. F. (2010). Assessing alpha and beta taxonomy in eupelmid wasps: determinants of the probability of describing good species and synonyms. *Journal of Zoological Systematics and Evolutionary Research*, 48(1), 40-49.
- Beck, J., Böller, M., Erhardt, A., & Schwanghart, W. (2014). Spatial bias in the GBIF database and its effect on modeling species' geographic distributions. *Ecological Informatics*, 19, 10-15.
- Becker, B. K. (2005). Geopolítica da amazônia. *Estudos avançados*, 19(53), 71-86.
- Brown JH, Lomolino MV. (1998). Biogeography. Sunderland, MA: Sinauer. 2nd ed
- Busby, J. R. (1979). Australian Biotaxonomic information system. Introduction and data interchange standards. Canberra: Australian Biological Resources Study, Department of Science and the Environment, 25.
- Canhos DAL, Sousa-Baena MS, de Souza S, Maia LC, Stehmann JR, Canhos VP, et al. (2015) The Importance of Biodiversity E-infrastructures for Megadiverse Countries. *PLoS Biol* 13(7): e1002204. doi:10.1371/journal.pbio.1002204.
- Dos Santos, J. G., Malhado, A. C., Ladle, R. J., Correia, R. A., & Costa, M. H. (2015). Geographic trends and information deficits in Amazonian conservation research. *Biodiversity and conservation*, 24(11), 2853-2863.
- Dornelas, M., Magurran, A. E., Buckland, S. T., Chao, A., Chazdon, R. L., Colwell, R. K., ... & McGill, B. (2013). Quantifying temporal change in biodiversity: challenges and opportunities. In *Proc. R. Soc. B* (Vol. 280, No. 1750, p. 20121931). The Royal Society.
- Cardoso P, Erwin TL, Borges PA, New TR. (2011). The seven impediments in invertebrate conservation and how to overcome them. *Biol. Conserv.* 144:2647–55
- Cardoso, D., Särkinen, T., Alexander, S., Amorim, A. M., Bitrich, V., Celis, M., ... & Goldenberg, R. (2017). Amazon plant diversity revealed by a taxonomically verified species list. *Proceedings of the National Academy of Sciences*, 114(40), 10695-10700.
- Diniz-Filho JAF, Loyola RD, Raia P, Mooers AO, Bini LM. (2013). Darwinian shortfalls in biodiversity conservation. *Trends Ecol. Evol.* 28:689–95
- Edwards, J. L., Lane, M. A., & Nielsen, E. S. (2000). Interoperability of biodiversity databases: biodiversity information on every desktop. *Science*, 289(5488), 2312-2314.
- Fearnside, P. M. (1999). Biodiversity as an environmental service in Brazil's Amazonian forests: risks, value and conservation. *Environmental conservation*, 26(4), 305-321.
- Forzza, R.C., Baumgratz, J.F.A., Bicudo, C.E.M. et al. (2012) New Brazilian floristic list highlights conservation challenges. *BioScience*, 62,39–45.

Hortal, J., Lobo, J. M., & JIMÉNEZ-VALVERDE, A. L. B. E. R. T. O. (2007). Limitations of biodiversity databases: case study on seed-plant diversity in Tenerife, Canary Islands. *Conservation Biology*, 21(3), 853-863.

Hortal, J., de Bello, F., Diniz-Filho, J. A. F., Lewinsohn, T. M., Lobo, J. M., & Ladle, R. J. (2015). Seven shortfalls that beset large-scale knowledge of biodiversity. *Annual Review of Ecology, Evolution, and Systematics*, 46, 523-549.

Ladle, R. J. & Whittaker, R.J. (2014). Biogeografia e Preservação Ambiental. Ed ANDREI.

Ladle, R. J., & Jepson, P. (2008). Toward a biocultural theory of avoided extinction. *Conservation Letters*, 1(3), 111-118.

Lewinsohn, T. M., & Prado, P. I. (2005). How many species are there in Brazil?. *Conservation Biology*, 19(3), 619-624.

Ladle, R., & Hortal, J. (2013). Mapping species distributions: living with uncertainty. *Frontiers of Biogeography*, 5(1).

Lomolino MV. (2004). Conservation biogeography. In *Frontiers of Biogeography: New Directions in the Geography of Nature*, ed. MV Lomolino, LR Heaney, pp. 293–96. Sunderland, MA: Sinauer.

Malhado, A. C. M. et al. (2013). The ecological biogeography of Amazonia. *Frontiers of Biogeography*, v.5, n.2.

Malhi, Y., Roberts, J. T., Betts, R. A., Killeen, T. J., Li, W., & Nobre, C. A. (2008). Climate change, deforestation, and the fate of the Amazon. *science*, 319(5860), 169-172.

May, R. M. (2010). Tropical arthropod species, more or less?. *Science*, 329(5987), 41-42.

Meyer C, Kreft H, Guralnick R, Jetz W (2015) Global priorities for an effective information basis of biodiversity distributions. *Nat Commun* doi:10.1038/ncomms9221

Mittermeier, R. A. et al. (2003). Wilderness and biodiversity conservation. *Proceedings of the National Academy of Sciences of the United States of America*, v.100, n.18, p.10309–10313. doi:10.1073/pnas.1732458100

MMA: <http://www.mma.gov.br/biodiversidade/biodiversidade-global/impactos>. Acesso em 11/07/2017

Morton, D. C., DeFries, R. S., Shimabukuro, Y. E., Anderson, L. O., Arai, E., del Bon Espírito-Santo, F., ... & Morisette, J. (2006). Cropland expansion changes deforestation dynamics in the southern Brazilian Amazon. *Proceedings of the National Academy of Sciences*, 103(39), 14637-14641.

Nelson, B. W., Ferreira, C. A., da Silva, M. F., & Kawasaki, M. L. (1990). Endemism centres, refugia and botanical collection density in Brazilian Amazonia. *Nature*, 345(6277), 714.

- Nepstad, D. C., de Carvalho, C. R., Davidson, E. A., Jipp, P. H., Lefebvre, P. A., Negreiros, G. H., ... & Vieira, S. (1994). The role of deep roots in the hydrological and carbon cycles of Amazonian forests and pastures. *Nature*, 372(6507), 666.
- Richards, P., Arima, E., VanWey, L., Cohn, A., & Bhattarai, N. (2017). Are Brazil's deforesters avoiding detection?. *Conservation letters*, 10(4), 470-476.
- Riddle, B. R., Ladle, R. J., Lourie, S. A., & Whittaker, R. J. (2011). Basic biogeography: estimating biodiversity and mapping nature. *Conservation biogeography*, 45-92.
- Schulman, L., Toivonen, T. and Ruokolainen, K. (2007), Analysing botanical collecting effort in Amazonia and correcting for it in species range estimation. *Journal of Biogeography*, 34: 1388–1399. doi:10.1111/j.1365-2699.2007.01716.x
- Shepherd, G.J. (2003) Terrestrial plants. Evaluation of the state of knowledge on biological diversity in Brazil: executive summary (ed. by T.M. Lewinsohn), pp. 53–55. Programa Nacional da Diversidade Biológica, Ministério do Meio Ambiente, Brasília.
- Soares-Filho, B. S., Nepstad, D. C., Curran, L. M., Cerqueira, G. C., Garcia, R. A., Ramos, C. A., ... & Schlesinger, P. (2006). Modelling conservation in the Amazon basin. *Nature*, 440(7083), 520.
- Soberón, J., & Nakamura, M. (2009). Niches and distributional areas: concepts, methods, and assumptions. *Proceedings of the National Academy of Sciences*, 106(Supplement 2), 19644-19650.
- Sobral, M. & Stehmann, J.R. (2009) An analysis of new angiosperm species discoveries in Brazil (1990–2006). *Taxon*, 58, 227–232.
- Sunderland, M. E. (2013). Computerizing natural history collections. *Endeavour*, 37(3), 150-161.
- Ter Steege, H., Pitman, N. C., Sabatier, D., Baraloto, C., Salomão, R. P., Guevara, J. E., ... Monteagudo, A. (2013). Hyperdominance in the Amazonian tree flora. *Science*, 342(6156), 1243092.
- Ter Steege, H., Pitman, N., Sabatier, D., Castellanos, H., Van Der Hout, P., Daly, D. C., ... & Duivenvoorden, J. (2003). A spatial model of tree α -diversity and tree density for the Amazon. *Biodiversity & Conservation*, 12(11), 2255-2277.
- Terborgh, J., Robinson, S. K., Parker, T. A., Munn, C. A., & Pierpont, N. (1990). Structure and organization of an Amazonian forest bird community. *Ecological Monographs*, 60(2), 213-238.
- Thiollay, J. M. (1994). Structure, density and rarity in an Amazonian rainforest bird community. *Journal of Tropical Ecology*, 10(4), 449-481.
- Veiga AK, Saraiva AM, Chapman AD, Morris PJ, Gendreau C, Schigel D, et al. (2017) A conceptual framework for quality assessment and management of biodiversity data. *PLoS ONE* 12(6): e0178731. <https://doi.org/10.1371/journal.pone.0178731>

Whittaker, R. J., Araújo, M. B., Jepson, P., Ladle, R. J., Watson, J. E., & Willis, K. J. (2005). Conservation biogeography: assessment and prospect. *Diversity and distributions*, 11(1), 3-23.

WWF, 2017
http://www.wwf.org.br/natureza_brasileira/areas_prioritarias/amazonia1/bioma_amazonia/index.cfm

Yang, W., Ma, K., & Kreft, H. (2013). Geographical sampling bias in a large distributional database and its effects on species richness–environment models. *Journal of Biogeography*, 40(8), 1415-1426.

Yang, W., Ma, K., & Kreft, H. (2014). Environmental and socio-economic factors shaping the geography of floristic collections in China. *Global Ecology and Biogeography*, 23(11), 1284-1292.

MUSEO EMILIO GOELDI, 2018. http://marte.museu-goeldi.br/marcioayres/index.php?option=com_content&view=article&id=10&Itemid=11

3 JUSTIFICATIVA

No presente trabalho, abordamos conceitos da biogeografia da conservação e macroecologia para compreender e quantificar o acúmulo e a perda da biodiversidade de plantas na Amazônia brasileira. Especificamente, avaliamos 360 anos de amostragem de plantas, buscando compreender o quão completo são os dados de ocorrência de espécies disponibilizados em plataformas de acesso livre. Abordamos também o método de estimativa de biodiversidade proposto por Sousa-Baena *et al.*, (2014) que aborda o quanto de espécies esperadas ocorrem no local com base nos dados de ocorrência, associando essa informação com a principal problemática de perda de biodiversidade na Amazônia, o desmatamento. Para isto, buscamos dados históricos de desmatamento a partir do monitoramento realizado pelo INPE e estimamos os dados futuros através de modelagens de cenários de mudanças de habitat proposto por Soares-Filho *et al.*, (2005) para quantificar o quanto de informação de biodiversidade florística foi degradada pelo tempo e pelas mudanças de habitat.

4 MAPPING FLORISTIC KNOWLEDGE LOSS IN THE BRAZILIAN AMAZON DUE TO HISTORICAL AND FUTURE DEFORESTATION

Bruno Umbelino^{1*}, Juliana Stropp¹, Ricardo Correia^{1,2,3}, João Vitor Campos-Silva¹, Richard J. Ladle^{1,2} e Ana C. M. Malhado¹

¹Institute of Biological Sciences and Health, Federal University of Alagoas, Campus A. C. Simões, Av. Lourival Melo Mota, s/n Tabuleiro dos Martins, Maceió, AL, Brazil

²School of Geography and the Environment, University of Oxford, South Parks Road, Oxford OX1 3QY, United Kingdom

³DBIO & CESAM-Centre for Environmental and Marine Studies, University of Aveiro, Aveiro, Portugal

*Correspondence: Bruno Umbelino. E-mail: brunosilvasantos27@gmail.com

4.1 INTRODUCTION

Biodiversity knowledge play a central role in the effectiveness of conservation programs because information on species distributions, particularly for endangered species in highly impacted areas is essential to subsidize frontline conservation actions, which can mitigate the impacts caused by anthropogenic disturbances. Species distribution databases need to ensure the quality (complete data and age) to guarantee the usefulness and reliability of these studies (Rocchini et al., 2011). Among the applied outputs we highlight: (1) prioritizing investments for the recovery of species; (2) directing the establishment of protected areas; (3) limit the exploitation of species and their habitats and (4) report the status of the species in the environment (Possingham et al., 2002). Lack of knowledge about species occurrence, generally described in the literature as the Wallacean shortfall (Lomolino, 2004; Ladle & Whittaker, 2014), is one of the major biodiversity knowledge shortfalls (Hortal et al., 2015). Another important knowledge shortfall involves species that have not yet been described, the co-called Linnean shortfall. This

shortfall includes species that have never been recorded, or those that already been collected, but remain undescribed, as a consequence of the lack of funding and capacity in global taxonomy (Hortal et al., 2015).

The Amazon basin hosts nearly half of the world's remaining tropical forests (Hansen et al., 2013) and is one of the most biodiverse regions worldwide (Loyola et al., 2014). Due its colossal extent, many of its areas are remote and practically inaccessible, therefore it is no surprise that the biodiversity knowledge of the region is still incomplete and deficient. Most of the existing knowledge available about Amazonian biodiversity is clearly concentrated around large cities, close to research centers and along large rivers and roads (Nelson et al., 1990; Schulman et al., Hopkins, 2007; Dos Santos et al., 2015, Oliveira et al., 2016). As a result, it is estimated that approximately 40% of the Amazon has ever been researched and for some many plant species, the known geographic distribution does not accurately represent occurrence in the region (Schulman et al., 2007). Another reason for concern is that oftentimes, the areas that do not have detailed botanical documentation coincide with areas of intense habitat destruction, and will thus never be scientifically documented (Sousa-Baena et al., 2014). To know the situation about degradation of plants diversity is important because is the first organism to suffer due to deforestation.

Although there is disagreement about effectiveness about the use of species occurrence among researchers (Sousa-Baena et al., 2014; Yang et al., 2014; Engemann et al., 2015), the data represent a strong tools to be used as basis in the areas of knowledge such as biogeography, ecology, evolution and conservation (Hortal et al., 2015), temporal trends and the general quality of this data receives less attention (Boakes et al., 2010). Species occurrence data are usually characterized by three main attributes (Rondinini et al., 2006): species name, geographic location and collection / observation dates. Most research assessing the existing biodiversity knowledge have been approaching the spatial dimension, identifying areas without knowledge and how the localization errors can affect the conclusions. Temporal degradation of biodiversity data quality is inevitable due to ecological dynamics (e.g., local extinctions, immigration, biological invasions) (Tessarolo et al., 2017). This process can be

gradual, due to taxonomic or environmental change, or more abrupt as is the case of large habitat changes due to deforestation for example (Ladle & Hortal, 2013).

Such temporal biases have the potential to hamper descriptions of changes in biodiversity and thus limit our ability to model the effects of global changes (Tessarolo et al., 2017). However, while the influence of time on data degradation is well-known in other scientific fields and has been successfully incorporated into protocols to assess data quality (Kennedy et al., 2014), the temporal degradation of biodiversity data has been somewhat neglected. This aspect of biodiversity data remains poorly understood and managed, and many databases contain a high proportion of old records (Tessarolo et al., 2017). This is well illustrated by a recent study that analyzed African flowering plants from GBIF data, and the authors found that the majority of well surveyed cells contained old records (Stropp et al., 2016).

Here we explore the spatial and temporal quality of existing occurrence data of plant species in the Brazilian Amazon. Specifically, we address three questions: (1) How complete are species-occurrence data for flowering plants in Amazon rainforest available on open access platforms? (2) When and where did botanical sampling occur through time? (3) What is the size information have been lost already by historical deforestation and how many will be lost to the future deforestation? With these questions we will be able to identify less well-known areas and directions which can advance conservation science and practice.

4.2 METHODS

4.2.1 Data acquisition

We focused our analysis on plants of Class Magnoliopsida because they represent the most abundant class of all plants in the Amazon (Forzza et al., 2010), and were, therefore to considered as a proxy of plant biodiversity. Species occurrence data were downloaded from three freely accessible databases: Global Biodiversity Information Facility (<https://www.gbif.org/>), SpeciesLink (<http://www.splink.org.br/>) and Sistema de Informação sobre a Biodiversidade Brasileira (<http://www.sibbr.gov.br/>). All data were accessed on

December 1, 2017 and included only geo-referenced records of all available years (1656-2016) for the Brazilian Amazon region.

We then filtered our database to include only records that included complete and valid information on the three key knowledge dimensions: species names, geographic location and collection / observation dates (Rondinini et al., 2006). This process involved several steps: firstly, we eliminated specimens with coordinates identified outside of the Amazon boundaries, as defined by a shapefile obtained from the Brazilian Ministry of the Environmental (<http://mapas.mma.gov.br/i3geo>). Second, we evaluated the validity of names using the Taxonomic Name Resolution Service (TNRS) version 3.2 (Boyle et al., 2013), consulted during January 2018 and kept only specimens for which species names we obtained valid matches from TNRS. We considered as valid matches any species name that matched TNRS records with a score over 0.9. Third, we filtered out records containing missing information using the '*scrubr*' package (Chamberlain, 2016) for R software. For this step, we filtered out records missing information on latitude, longitude, collected date, taxonomy and removed duplicated records selecting unique combinations of species name, collection date and location by '*duplicated*' function of the same package. Finally, we also excluded all the species with non-amazonian and cultivated according to Cardoso et al. (2017).

4.2.2 Inventory completeness

We estimated the inventory completeness for Magnoliopsida class for the entire Brazilian Amazon considering sampling units (SUs) of 25 x 25 km². For each sampling unit, we calculated species accumulation curves considering the number of species collected from 1656 until 2016. We obtained estimates of the accumulation curve with "exact" and "random" procedures. When the predicted richness was estimated using the data generated by the methods "exact" and "random" at the same time, the mean of both richness values was used to calculate completeness (the percentage representing the observed number of species against the predicted one). We estimated inventory completeness based on Sousa-Baena et al. (2014) using 'KnowBR' R package (González & Lobo 2017). This method returns a completeness score (C_i) from zero to one,

with values close to one indicating a complete inventory (equation below). Each SU with a small number of records show an artificial values of C , because changes in a and b , causing estimates unstable in SUs with a small number of records (Sousa-Baena et al., 2014; Stropp et al. 2016).

$$C_i = \frac{S_{obs,i}}{S_{obs,i} + (a_i^2/2b_i)}$$

To define the edge of C_i values are stable or not, we assessed the relationship between C_i and number of records by each SU. After calculated for final cleaned data, we after data lost by historical and future deforestation. To do that, we saw how deforestation effects on our understanding about species abundance provided by inventories of biodiversity. We present estimates of inventory completeness for SUs with ≥ 100 specimens. To well-sampled unites we consider C_i value ≥ 0.7 .

4.2.3 Deforestation data and species-occurrence losses

We mapped both the complete and filtered records datasets by counting the number of records falling within $25 \times 25 \text{ km}^2$ grid cells defined within the entire Brazilian Amazon, for areas with and without forest. Grid cells were defined according to methodology employed by Instituto de Pesquisas Espaciais (INPE 2017) for deforestation monitoring. PRODES project developed by INPE monitoring uses as a reference deforestation by shallow cut across the length of the Amazon and produces, since 1988, annual reports of deforestation rates in and out of indigenous conservation and land units in the region. All the layers from PRODES project were used projecting by WGS 84/Pseudo-Mercator datum and 1:100.000 spatial scales according to MMA (2018). Temporal scale data of records, we divided by ten years according to INPE's monitoring data available. For initial database (before loses calculated) and final database. Thus, we analysed how much information has been lost up until the present day. For the future deforestation scenario, we use the modeling proposed by Soares-Filho et al. (2006) for the year 2050 only for the "*Business as usual*" scenario. This scenario assumes that recent deforestation trends will continue, highways currently scheduled for paving will be paved, compliance

with legislation requiring forest reserves on private land will remain low and new Protect Areas will not be created. And also, we quantified the areas, with and without knowledge, that have been and will be lost by deforestation. For this purpose, we calculated the initial percentage areas with and without information previous measurements and we took the difference of deforested areas. After each loss, we calculate again the percentage area of information and non-information. We used QGIS software version 2.14 (QGIS Development Team, 2018).

4.3 RESULTS

4.3.1 Initial characterization

The initial data contained 608.451 records for the Brazilian Amazon region. Of these, nine collections provided more than 10.000 records; 32 collections contributed between 1.000 and 10.000 records; 57 collections with more than 100 but less than 1.000 records; 60 with between 10 and 100 and 68 collection contributed less than 10 records. A large proportion of records (113.332) contained no information on data provenance. All those collections provided more data it is belonging to GBIF database, with three collections from Brazil (INPA, JBRJ and Cenargen) and six international, and together just this nine collection had approximately 59,5% (608.451). Among the three consulted databases, GBIF contained the most records (607.045), followed by SpeciesLink (1.034) and SiBBr (372). The total number of scientific collections contributing to the initial database was 226.

The taxonomic characterization of the data revealed records from 212 families, 2.379 genera and 45.999 species. Thirteen families had more than 10.000 records, approximately 65% of the total number of records. In contrast, 42 families had less than 10 records. The most abundant genera were *Manihot* and *Miconia* with more than 10.000 records or 21% from all data. On the other hand, 968 genera had less than ten records. Two species (*Manihot esculenta* and *Myrcia splendens*) had 19% and 12.524 species had records between 10 until 100, approximately 32%.

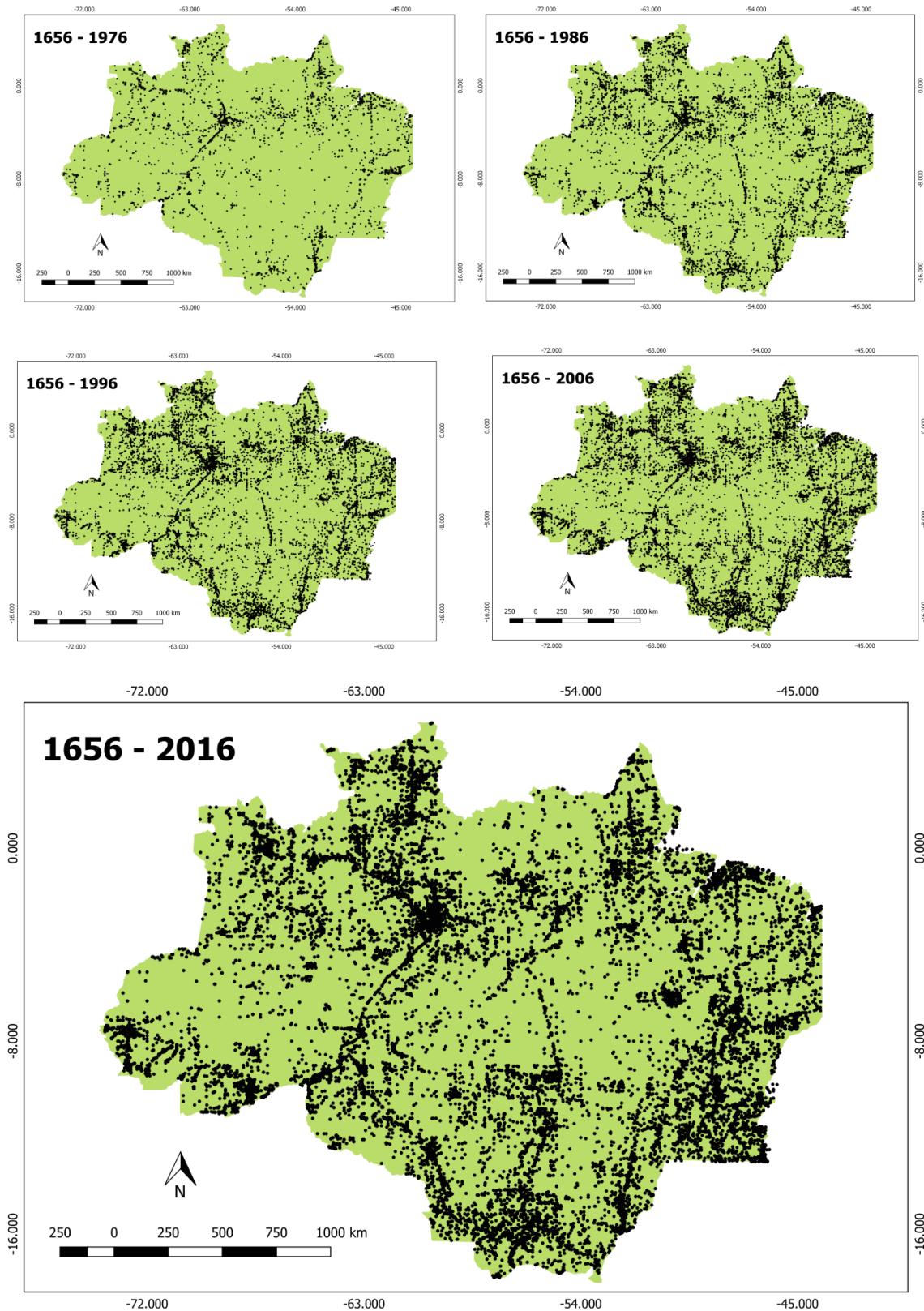
4.3.2 Knowledge losses

Of the total of 608.451 records belonging to the Brazilian Amazon, only ~38% (228.548) remained with complete information based on the criteria established in this study. The main problems in the data were: 158.203 duplicate records; 103.980 records missing collection date information; 100.199 without taxonomic information; 15.866 species non Amazonian and cultivated; 1.185 data identified until genera level, 408 records with coordinate problems and taxonomy with low identification confidence (under 0.9 match score) had the lowest losses with only 62 records. From three initial databases, just two remained with complete data GBIF (228.202) and SpeciesLink (346).

The total number of scientific collections contributing to the final database was 163. Of these, 10 collections provided more than 10.000 records; 48 collections contributed between 1.000 and 10.000 records; 78 collections with more than 100 but less than 1.000 records; 92 with between 10 and 100 and 98 collections remained with less than 10 records. All those collections provided more data kept belonging to GBIF database, with three collections from Brazil (INPA and JBRJ) and three international, and together just these five collections had approximately 66% (151.421).

The remaining 228.548 records are distributed by 197 families, 1.834 genera and 13.490 species. Records data over 1.000 on each taxonomic level, we had: ninety-eight families and eighty-two genera. About species level, we had 892 over 100 records (Supplementary Material). The earliest records in our final database was for the year 1656 (n=1) and the most recent was in 2016 (n=934). Over the years, there are 360 years with records for the Brazilian Amazon (Fig 1).

Figure 1 Spatial evolution of species occurrence data through the years on Brazilian Amazon between 1656 and 2016. For the period of 1656 until 1976 we had 60.348 records valid; 1656 until 1986 about 120.339; 1656 until 1996 had 162.976; 1656 until 2006 were 188.395 and from 1656 until 2016 we had 228.548 records valid.



4.3.3 Spatial distribution of inventory completeness

The total of SUs was 3.227 with 10% have been floristic sampling containing at least two records. Only 457 (14%) SUs contained at least 100 unique records. The estimates values of inventory completeness (C) ranged from 0.01 to 0.88. Only 79 (17%) SUs with more than 100 records and $C_i \geq 0,5$ (Fig 2a), many of these areas were located near highways. Furthermore, just 0,8% of the well-sampled SUs (4 SUs) remained (Fig2b).

After influence about historical deforestation (until 2016) we lost 49 SUs with $C_i \geq 0,5$ (Fig 3a), remained areas on northwest and central in Brazilian Amazon. For areas considered well sampled ($C_i \geq 0,7$) we observed that just two remained on the region located central and east region (Fig 3b). On the future deforestation (2050) we will not have well sampled areas in Brazilian Amazon, according to the current data (Fig 4a), losing more 37 SUs.

Figure 2 Distribution of sampled units according to completeness values ($C_i \geq 0,5$ and $C_i \geq 0,7$) (a) and (c) are according to $C_i \geq 0,5$ near to highways and protected areas and indigenous land. (b) and (d) is according well sampled units ($C_i \geq 0,7$).

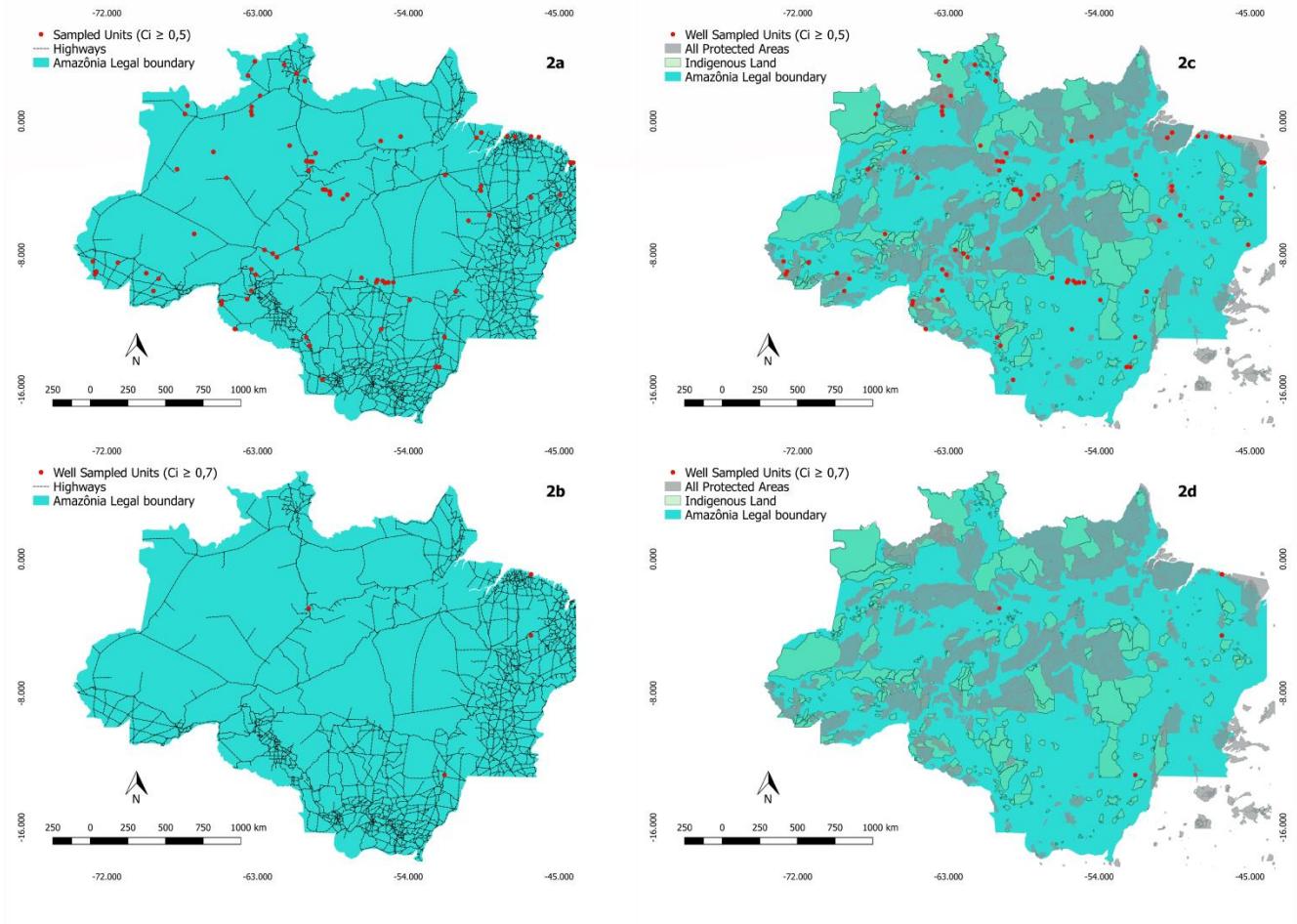


Figure 3 Sampled units and well sampled units lost by historical deforestation. (a) and (c) is about remain sampled units (blue points) near to highways and in protected areas and indigenous land. (b) and (d) is about well sampled units already lost by deforestation.

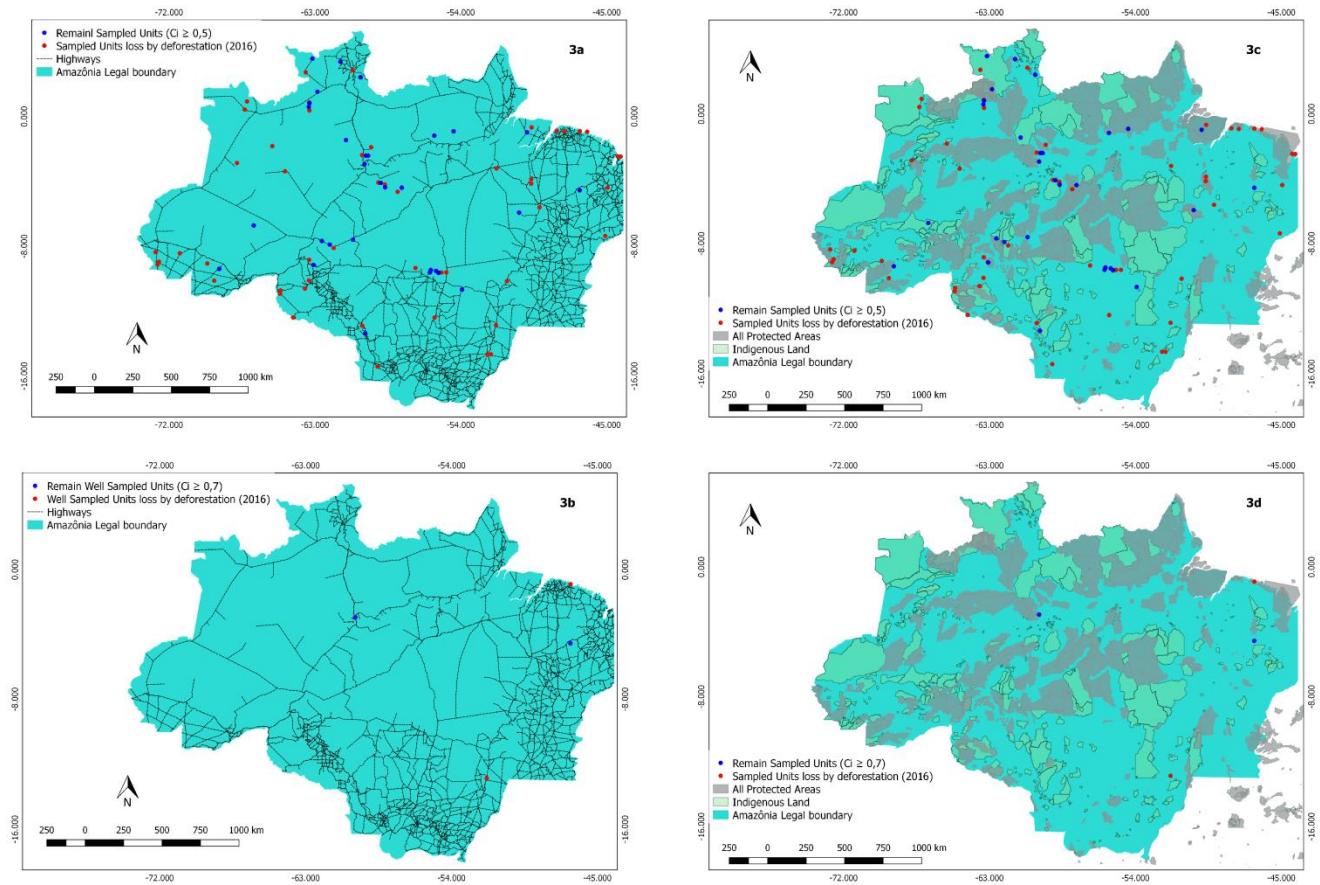
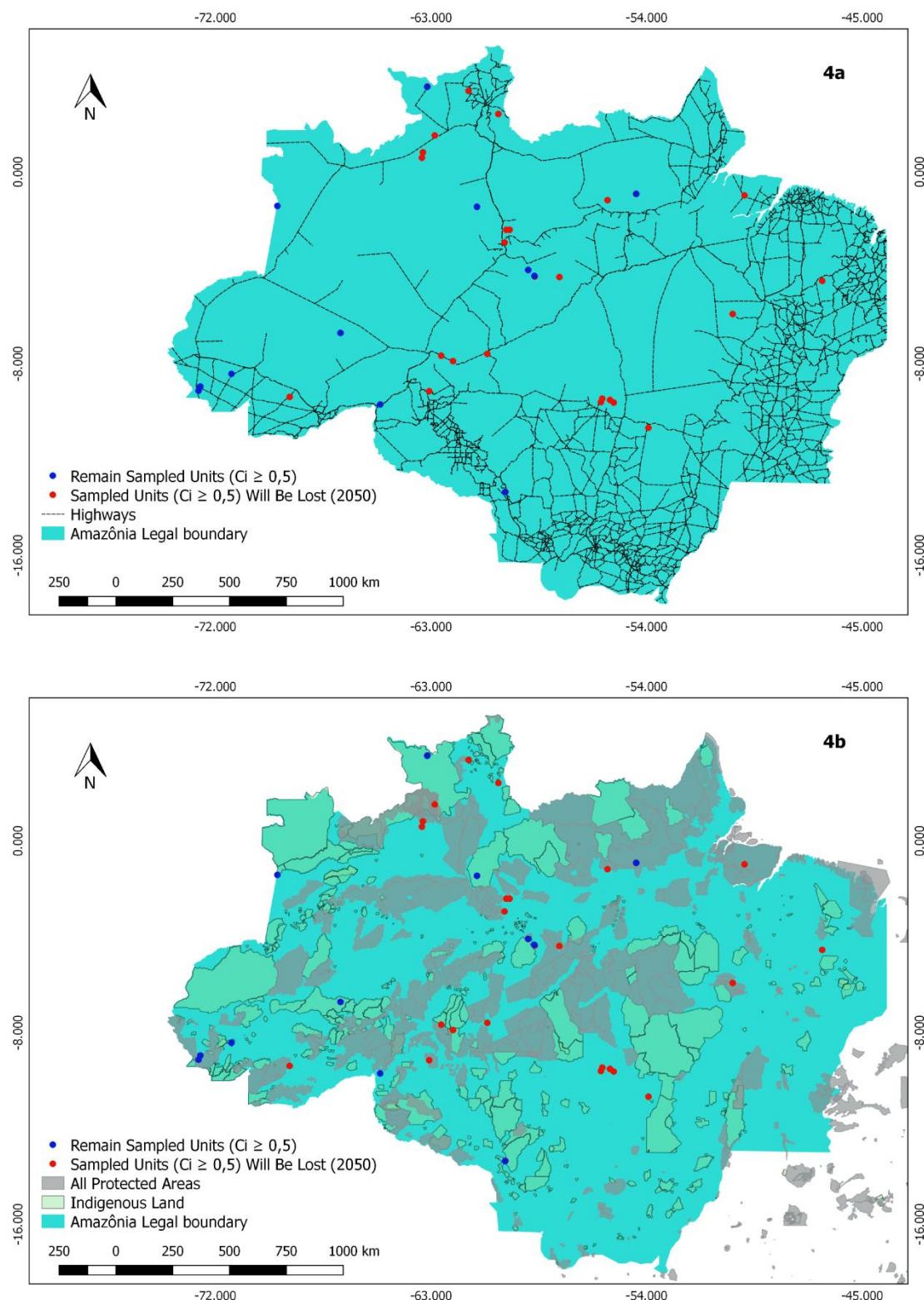


Figure 4 Lost sampled units lost by future deforestation. (a) is about remain samples units near to highways and (b) about near or inside of protected areas and indigenous land.

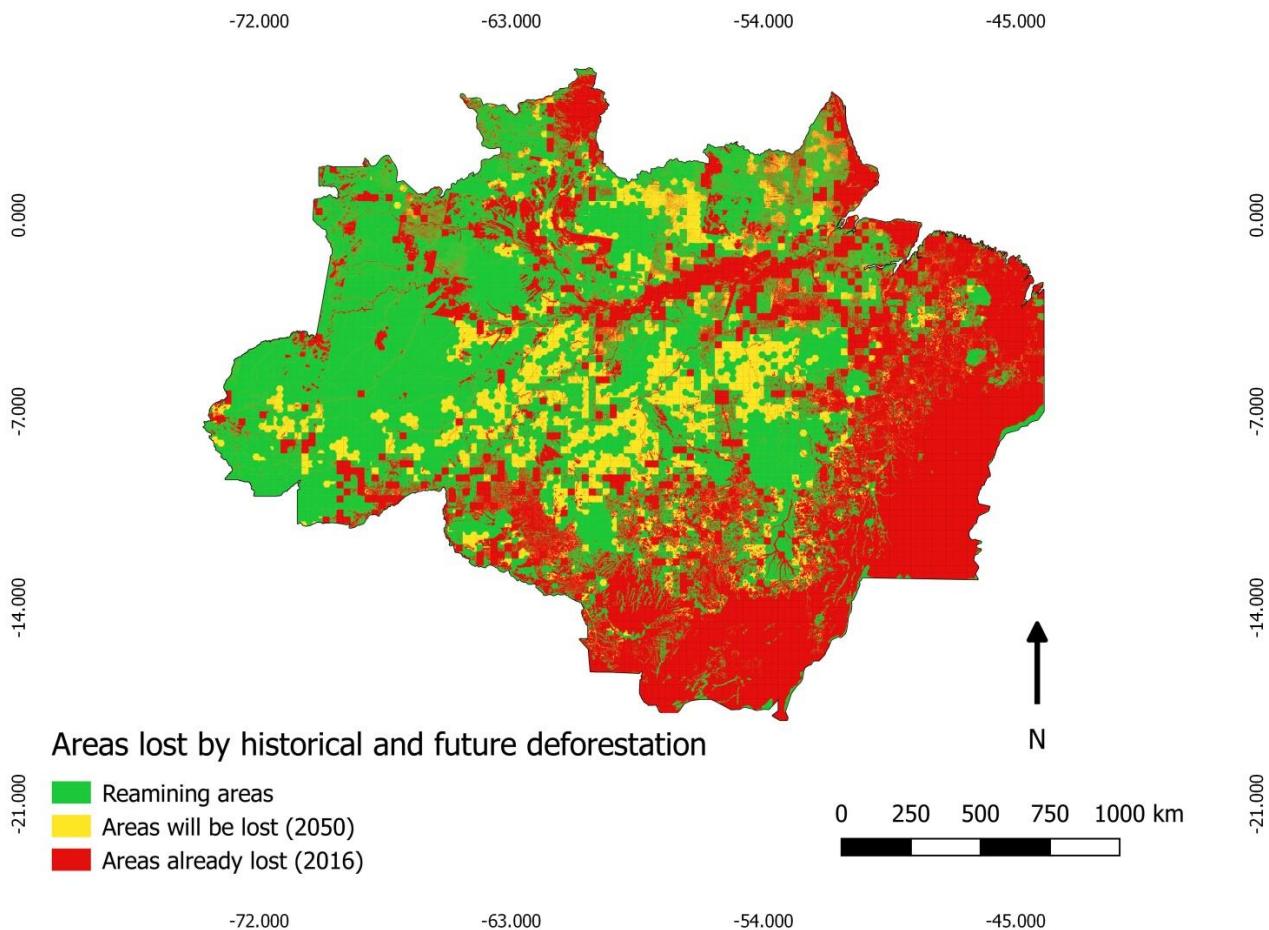


4.3.4 Knowledge and data loss

We also mapped the records throughout the protected areas system, including indigenous land. Around 35% of the records are located within the limits of protected areas and indigenous land. About 23% (52.763) are inside protected areas, including Sustainable Use (35.743) and Integral Protection (17.020) categories. Within Indigenous Land we observed 12% (27.535) of records (Supplementary material). About 65% of all records are out of protected areas, 13% (31.107) is belonging to “ghost data” (data already lost by historical deforestation). Most sampled areas are located inside protected areas and/or indigenous land (Fig 2c), but well-sampled areas were not. These well-sampled three they were out of protected areas and indigenous land, and just one there was in protected area on Pará state (Fig 2d). On the current deforestation, most sampled areas were lost including those inside of protected areas (Fig 3c). For those areas considered well sampled, kept to be out of protect areas and indigenous land (Fig 3d). To future deforestation, we probably will have sampled units on north and west regions (Fig 4b).

The areas without digital accessible knowledge correspond to 42% of the total. For these areas, about 7% correspond to areas that have already been lost by historical and current deforestation, knowledge is forever lost. If we consider the loss of areas by future projection, we will lose is about 3% of floristic diversity. The areas with digital accessible knowledge distributed in the Brazilian Amazon correspond to 58% of the totality. Approximately 12% were the areas lost in the current deforestation scenario. If we consider future losses, these areas of lost knowledge will be around 4% (Fig 5).

Figure 5 Areas lost by historical and future deforestation.



4.4 DISCUSSION

Our results show that the information about plant biodiversity in the Brazilian Amazon is impressively compromised. When we observe that 62% of records missed information, requiring the attention of the scientific community and decision makers, once these datasets can jeopardize the biodiversity dynamics understanding conservation practices and environmental policies. The major concern is specimen identification, which caused the greatest loss of data. Among the many implications for biodiversity surveys, the phenotypic variation (staining, behavior, and phenology) of the specimens aggravates the accuracy of identification (Ladle & Hortal, 2013). Two major problems noticed by Hortal et al. (2015) on these deficit of knowledge (Linnean) are the number of species formally described are constantly changing because of new descriptions, reviews and synonyms have not been resolved, and the difficulties in establishing a unique conception of species. Another difficulty is how the

number of species expected to be highly sensitive to the method adopted in the survey and to the values of the parameters used for the species estimates. Thus, most biological research that depends on species distribution information is likely to be based on incomplete maps, even when a variety of sampling methods are used (Colwell & Coddington, 1994). Failure to eliminate duplicate records brings another problem due to the false impression of abundance. Several techniques are proposed to increase spatial knowledge of species, such as potential modelling for the creation of priority areas for conservation. However, the uses of duplicate records for this purpose result in imprecise models about the actual potential occurrence of the species. In the face of increasing habitat reductions for human occupation, cattle and hydroelectric, biodiversity loss assessments based on methods of estimating species with complete and non-duplicate data result in better conservation plans for less wasted money (Whittaker et al., 2005). We emphasize that investment in conservation strategies based only on future distribution models can be delicate, since there is no guarantee that species will migrate to the geographic reach to the predicted sites (Lemes & Loyola, 2013) and whether they will be in condition to support of these new species weakened by the constant loss of habitats.

We observed a tendency of records for the most abundant groups, an expected pattern since much of the species in the Amazon present some pattern of rarity (Forzza et al., 2010). However, we do not have information about the sampling effort end survey method used in the field inventory. Beyond the information lost at the species level, we are losing them at the highest levels, such as genera and family. We recognize that these data are still relevant for studies that do not require identification until species level, such as the evaluation of gender or family distribution. But, undoubtedly the poor quality of Amazonian plant diversity can severely compromise our understanding about Amazonian biodiversity as a whole. Not excluding non-native species data from open digital herbariums, they compromise our understanding of how many species are likely to exist in tropical forests (Cardoso et al., 2017). Therefore, we emphasize that much of the floristic Amazonian biodiversity still needs to be revealed.

The areas that remained with good records during the evolution of deforestation were the same areas indicated by (Soares-Filho et al., 2005) as possible forest remnants. According to the same authors, the largest areas with biodiversity loss will be the ones closest to the roads, whether they are paved or not. This was observed in our study through biodiversity completeness analyses. For conservation plans to be effective, we must have solid data about species occurrence focusing in poorly-known and threatened areas (Sousa-Baena et al., 2014). Basically, there are few regions in Amazonas state within been sampled, while large areas remain unsampled.

Even if records with complete information are distributed across most of the entire length of the Brazilian Amazon, we observed that most areas have a low and biased number of samples. The difficulties of access to research areas in the Amazon region are already well documented in the literature (Nelson et al., 1990; Schulman et al., 2007). Thus, high levels of botanical knowledge are located close to cities (e.g. Manaus, Belém, Macapá, Porto Velho and Boa Vista) and some major rivers (e.g. Tocantins and Tapajós in Pará and Rio Negro and Madeira in the Amazonas) (Hopkins 2007). It can be expected that study sites are often located in such areas since doing so near the main access routes becomes more applicable in the logistic sense, especially if the work involves large amounts of equipment or requires frequent trips to the research sites (Dos Santos et al., 2015).

Gaps of knowledge correlate with those observed by investigating the distribution of scientific knowledge for the Amazon through published articles on conservation and biodiversity (Dos Santos et al. 2015). Our study shows that indigenous lands contain less data than protected areas. This pattern could be explained by the stringent legislation controlling access and activities in indigenous land making it difficult to implement research projects in those areas (Dos Santos et al. 2015). Opposite that, conservation units show conditions for researchers to make their research without barrier to access areas with potential to be studied (Laurance 2013; Dos Santos et al., 2015). Oliveira et al. (2016) studied the biases of biodiversity in areas easily accessible: he found the density of records for angiosperms, arthropods and vertebrates is concentrated less than 1 km away of major access routes in all Brazilian biomes. The same

authors observed that the proportions were similar for all the studied groups, and that the angiosperms presented more areas and records than the arthropods and vertebrates in all the biomes. These biases affect the understanding of spatial patterns of diversity and reduced efficiency of conservation actions (Grand et al., 2007).

Deforestation is a major cause of loss of biological knowledge, and while an optimistic scenario of governance could reverse this trend, reconciling economic interest with biodiversity conservation is not easy (Soares-Filho et al., 2005), and it does not seem to be a priority of the current governments (Campos-Silva et al., 2015). We show that the major extensions of areas with knowledge were more than areas without knowledge. These losses are important to know to redirecting searches those areas that will be lost. The Wallacean shortfall hinders our understanding about the patterns and ecological distributions of species, affecting dramatically the estimates of the status of threats adopted by IUCN (Hortal et al., 2015), since it indicates where taxa and regions are priorities for conservation (Riddle et al., 2011). For those areas with knowledge lost, we can do future scenarios of how long these data of species occurrence will be valid, with the rate of how many species data we are losing by converting habitat.

Brazilian society is facing a dramatic political crisis, but the government should change its political will, increasing their effort on conservation of Amazonia. The strengthening of the governance includes the increase of funding and human resources increase the efforts in low sampled areas, increase the understanding of biodiversity in remote places and within indigenous land, respecting their cosmology and sovereignty within the area, and the development of smart platforms of data management. The international community also should evaluate the quality of information available to researchers and decision makers, and develop new technologies to improve the system of data acquisition.

Acknowledgments

This work was funded Brazilian National Council for Scientific and Technological Development CNPq and Global Biodiversity Information Facility through Young Researcher award (2016). We thank Marcos Vital, Nídia Fabré,

Rafael Loyola, Marcio Efe, Maíra Benchimol and Debora Drucker for contributions on the study.

REFERENCES

- Boakes, E. H., McGowan, P. J., Fuller, R. A., Chang-qing, D., Clark, N. E., O'Connor, K., Mace, G. M. (2010). Distorted views of biodiversity: spatial and temporal bias in species occurrence data. *PLoS biology*, 8(6), e1000385.
- Boyle, B., Hopkins, N., Lu, Z., Garay, J. A. R., Mozzherin, D., Rees, T., ... Lowry, S. (2013). The taxonomic name resolution service: an online tool for automated standardization of plant names. *BMC bioinformatics*, 14(1), 16.
- Colwell, R. K., & Coddington, J. A. (1994). Estimating terrestrial biodiversity through extrapolation. *Philosophical Transactions: Biological Sciences*, 101-118.
- Campos-Silva, João Vitor, Sinomar Ferreira da Fonseca Junior, and Carlos Augusto da Silva Peres. "Policy reversals do not bode well for conservation in Brazilian Amazonia." *Natureza et Conservacao* 13.2 (2015): 193-195.
- Dos Santos, J. G., Malhado, A. C., Ladle, R. J., Correia, R. A., Costa, M. H. (2015). Geographic trends and information deficits in Amazonian conservation research. *Biodiversity and conservation*, 24(11), 2853-2863.
- Engemann, K., Enquist, B. J., Sandel, B., Boyle, B., Jørgensen, P. M., Morueta-Holme, N., ... Svenning, J. C. (2015). Limited sampling hampers "big data" estimation of species richness in a tropical biodiversity hotspot. *Ecology and evolution*, 5(3), 807-820.
- Forzza, R. C., Baumgratz, J. F. A., Costa, A., Hopkins, M., Leitman, P. M., Lohmann, L. G., ... Pirani, J. R. (2010). As angiospermas do Brasil. 78-88.
- Grand, J., Cummings, M. P., Rebelo, T. G., Ricketts, T. H., Neel, M. C. (2007). Biased data reduce efficiency and effectiveness of conservation reserve networks. *Ecology Letters*, 10(5), 364-374.
- Hopkins, M. J. (2007). Modelling the known and unknown plant biodiversity of the Amazon Basin. *Journal of Biogeography*, 34(8), 1400-1411.
- Hortal, J., de Bello, F., Diniz-Filho, J. A. F., Lewinsohn, T. M., Lobo, J. M., & Ladle, R. J. (2015). Seven shortfalls that beset large-scale knowledge of biodiversity. *Annual Review of Ecology, Evolution, and Systematics*, 46, 523-549.

INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS - INPE. Coordenação Geral de Observação da Terra - OBT. Projeto PRODES: monitoramento da Floresta Amazônica Brasileira por Satélite. Disponível em: <<http://www.obt.inpe.br/prodes/>> acessado em agosto de 2017.

Kennedy, R. E., Andréfouët, S., Cohen, W. B., Gómez, C., Griffiths, P., Hais, M., ... Zhu, Z. (2014). Bringing an ecological view of change to Landsatbased remote sensing. *Frontiers in Ecology and the Environment*, 12, 339–346.

Ladle, R. J. & Whittaker, R.J. (2014). Biogeografia e Preservação Ambiental. Ed ANDREI

Ladle, R., & Hortal, J. (2013). Mapping species distributions: living with uncertainty. *Frontiers of Biogeography*, 5(1).

Lemes, P., & Loyola, R. D. (2013). Accommodating species climate-forced dispersal and uncertainties in spatial conservation planning. *PloS one*, 8(1), e54323.

Lomolino MV. (2004). Conservation biogeography. In *Frontiers of Biogeography: New Directions in the Geography of Nature*, ed. MV Lomolino, LR Heaney, pp. 293–96. Sunderland, MA: Sinauer.

Loyola, R., Machado, N., Vila Nova, D., Martins, E., & Martinelli, G. (2014). *Áreas prioritárias para conservação e uso sustentável da flora brasileira ameaçada de extinção*.

Nelson, B. W., Ferreira, C. A., da Silva, M. F., Kawasaki, M. L. (1990). Endemism centres, refugia and botanical collection density in Brazilian Amazonia. *Nature*, 345(6277), 714-716.

Oliveira, U., Paglia, A. P., Brescovit, A. D., Carvalho, C. J., Silva, D. P., Rezende, D. T., ... Ascher, J. S. (2016). The strong influence of collection bias on biodiversity knowledge shortfalls of Brazilian terrestrial biodiversity. *Diversity and Distributions*, 22(12), 1232-1244.

Possingham, H. P., Andelman, S. J., Burgman, M. A., Medellín, R. A., Master, L. L., Keith, D. A. (2002). Limits to the use of threatened species lists. *Trends in ecology & evolution*, 17(11), 503-507.

QGIS Development Team, (2017). QGIS Geographic Information System. Open Source Geospatial Foundation Project. <http://qgis.osgeo.org>

Riddle, B. R., Ladle, R. J., Lourie, S. A., & Whittaker, R. J. (2011). Basic biogeography: estimating biodiversity and mapping nature. *Conservation biogeography*, 45-92.

Rocchini, D., Hortal, J., Lengyel, S., Lobo, J. M., Jimenez-Valverde, A., Ricotta, C., ... Chiarucci, A. (2011). Accounting for uncertainty when mapping species distributions: the need for maps of ignorance. *Progress in Physical Geography*, 35(2), 211-226.

Rondinini, C., Wilson, K. A., Boitani, L., Grantham, H., & Possingham, H. P. (2006). Tradeoffs of different types of species occurrence data for use in systematic conservation planning. *Ecology letters*, 9(10), 1136-1145.

Schulman, L., Toivonen, T., & Ruokolainen, K. (2007). Analysing botanical collecting effort in Amazonia and correcting for it in species range estimation. *Journal of Biogeography*, 34(8), 1388-1399.

Soares-Filho, B., Moutinho, P., Nepstad, D., Anderson, A., Rodrigues, H., Garcia, R., ... Silvestrini, R. (2010). Role of Brazilian Amazon protected areas in climate change mitigation. *Proceedings of the National Academy of Sciences*, 107(24), 10821-10826.

Sousa-Baena, M. S., Garcia, L. C., Peterson, A. T. (2014). Completeness of digital accessible knowledge of the plants of Brazil and priorities for survey and inventory. *Diversity and Distributions*, 20(4), 369-381

Stropp, J., Ladle, R. J., Malhado, M., Ana, C., Hortal, J., Gaffuri, J., ... & Mayaux, P. (2016). Mapping ignorance: 300 years of collecting flowering plants in Africa. *Global Ecology and Biogeography*, 25(9), 1085-1096.

Ter Steege, H., Pitman, N. C., Sabatier, D., Baraloto, C., Salomão, R. P., Guevara, J. E., ... Monteagudo, A. (2013). Hyperdominance in the Amazonian tree flora. *Science*, 342(6156), 1243092.

Ter Steege, H., Vaessen, R. W., Cárdenas-López, D., Sabatier, D., Antonelli, A., De Oliveira, S. M., ... Salomão, R. P. (2016). The discovery of the Amazonian tree flora with an updated checklist of all known tree taxa. *Scientific reports*, 6, 29549.

Tessarolo, G., Ladle, R., Rangel, T., & Hortal, J. (2017). Temporal degradation of data limits biodiversity research. *Ecology and evolution*, 7(17), 6863-6870.

Yang, W., Ma, K., Kreft, H. (2014). Environmental and socio- economic factors shaping the geography of floristic collections in China. *Global ecology and biogeography*, 23(11), 1284-1292.

Cardoso, D., Särkinen, T., Alexander, S., Amorim, A. M., Bitrich, V., Celis, M., ... & Goldenberg, R. (2017). Amazon plant diversity revealed by a taxonomically verified species list. *Proceedings of the National Academy of Sciences*, 114(40), 10695-10700.

Laurance WF (2013) Does research help to safeguard protected areas? Trends Ecol Evol 28:261–266

5 CONCLUSÃO GERAL

A perda de biodiversidade é observada em todo o planeta Terra em função das dinâmicas ambientais, sejam elas pelas condições naturais quanto às causadas pelas ações humanas. Os efeitos desta última são mais expressivos e preocupantes, constituindo a era geológica mais devastadora com altas taxas de extinções de espécies em curto período de tempo. Na região dos trópicos esses efeitos se tornam mais agravantes em função de não possuir conhecimento total das espécies existentes.

Ao passo que vamos conhecendo a biodiversidade, é percebido que as dificuldades de acesso às áreas implicam no cenário ideal de amostragem. Priorizar as áreas não inventariadas ou pouco inventariadas se torna cada vez mais essenciais para conhecer a biodiversidade antes do processo de fragmentação. Os dados de ocorrência de espécies se tornam ótimas ferramentas para conhecer os padrões de distribuição de espécies em regiões onde as espécies já não ocorrem atualmente, seja nas escalas local, regional ou global. Entretanto, ter consciência de que nem todos os dados de ocorrência de espécies possuem informações completas sobre a data de coleta, local (coordenadas) e identificação, torna-se desafiador a utilização dos mesmos para avaliações da biodiversidade e ações de mitigação ambiental. Os critérios para limpeza dos dados é de extrema relevância para conhecermos a qualidade de informação que esta sendo adquirida por parte dos pesquisadores, visto que os recursos financeiros destinados às pesquisas são de origem pública e escassa.

O compartilhamento de dados sobre a biodiversidade por parte dos grupos de pesquisas para comunidade científica ou civil, é de total importância para conscientização do estado de ameaça das espécies, formar redes de pesquisas e facilitar o acesso remoto. As trocas de experiências entre grupos de pesquisas de diversos lugares dá oportunidade de ações conservacionistas mais efetivas. Além disso, protocolos de levantamentos de flora e fauna não padronizados é uma realidade comum em todo o planeta, entretanto torna-los

padrão entre as linhas de pesquisas e automatiza-los seria uma forma de minimizar os erros quanto às dinâmicas ecológicas das espécies.

Conhecer e saber utilizar a biodiversidade sem comprometer os serviços ecossistêmicos é de responsabilidade de uma única espécie: *Homo sapiens sapiens*.