

GUSTAVO REIS DE BRITO

**APLICAÇÃO DA MODELAGEM PREDITIVA DE DISTRIBUIÇÃO DE ESPÉCIES
COMO FERRAMENTA DE ESTUDO DA BIODIVERSIDADE**

ASSIS

2018

GUSTAVO REIS DE BRITO

**APLICAÇÃO DA MODELAGEM PREDITIVA DE DISTRIBUIÇÃO DE ESPÉCIES
COMO FERRAMENTA DE ESTUDO DA BIODIVERSIDADE**

Dissertação apresentada à Universidade Estadual Paulista (UNESP), Faculdade de Ciências e Letras, Assis, para a obtenção do título de Mestre em Biociências (Área de Conhecimento: Caracterização e Aplicação da Diversidade Biológica)

Orientador(a): Dr. Fernando Frei

Bolsista: CAPES

ASSIS

2018

Dados Internacionais de Catalogação na Publicação (CIP)
Biblioteca da F.C.L. – Assis – Unesp

B862a Brito, Gustavo Reis de
Aplicação da modelagem preditiva de distribuição de
espécies como ferramenta de estudo da biodiversidade /
Gustavo Reis de Brito. Assis, 2018.
121 p. : il.

Dissertação de Mestrado – Universidade Estadual Paulista
(UNESP), Faculdade de Ciências e Letras, Assis
Orientador: Dr. Fernando Frei

1. Modelagem preditiva. 2. Zoogeografia. 3. Conservação
biológica. 4. Ecologia. I. Título.

CDD 577

“O que o analfabeto em ecologia não percebe em relação a um ecossistema é que se trata de um sistema. [...] Um sistema tem ordem, uma correnteza que flui de um ponto a outro. Se algo represar a correnteza, a ordem desmoronará. [...] É por isso que a função mais elevada da ecologia é a compreensão das consequências. ”

(Frank Herbert, Duna)

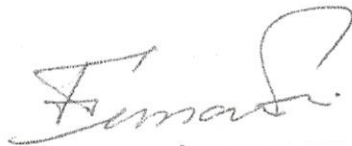
Gustavo Reis de Brito

**APLICAÇÃO DA MODELAGEM PREDITIVA DE
DISTRIBUIÇÃO DE ESPÉCIES COMO FERRAMENTA DE
ESTUDO DA BIODIVERSIDADE**

Dissertação apresentada à Universidade Estadual Paulista (UNESP), Faculdade de Ciências e Letras, Assis, para a obtenção do título de Mestrado Acadêmico em BIOCÊNCIAS (Área de Conhecimento: CARACTERIZAÇÃO E APLICAÇÃO DA DIVERSIDADE BIOLÓGICA)

Data da Aprovação: 22/02/2018

COMISSÃO EXAMINADORA



Presidente: Prof. Dr. Fernando Frei - UNESP/ASSIS



Membros: Prof. Dr. José Celso Rocha - UNESP/ASSIS

Profa. Dra. Maria Cecília Barbosa de Toledo - UNITAU/TAUBATÉ

AGRADECIMENTOS

Ao Dr. Fernando Frei, pela orientação ao longo de todos esses anos e pela confiança depositada no desenvolvimento desse projeto;

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES, pelo financiamento concedido sem o qual este projeto não seria possível;

À José Antonio e Daisa, meus pais, pelo inquestionável suporte e compreensão;

À Caroline Corrêa, pela amizade, companheirismo, motivação e suporte para prosseguir nessa jornada mesmo diante das dificuldades;

Ao Prof. Dr. Ciro César Zanini Branco, pelo apoio e contribuição às discussões deste projeto;

Ao Dr. Paulo Henrique Chaves Cordeiro do Museu Nacional do Rio de Janeiro, Universidade Federal do Rio de Janeiro, pela gentil contribuição com os dados utilizados para o desenvolvimento desse projeto;

À Dra. Solange Bongiovanni, pela amizade e apoio ao longo de todos esses anos, referência pessoal e profissional desde a graduação;

Aos demais mestres, amigos e colegas, o meu muito obrigado.

BRITO, G.R. **Aplicação da modelagem preditiva de distribuição de espécies como ferramenta de estudo da biodiversidade**. 2018. 121 p. Dissertação (Mestrado Acadêmico em Biociências). – Universidade Estadual Paulista (UNESP), Faculdade de Ciências e Letras, Assis, 2018.

RESUMO

Resumo - A Biologia da Conservação é uma ciência multidisciplinar surgia em meados dos anos 80 através da necessidade da junção de diferentes áreas do conhecimento frente às mudanças ambientais que afetam a biota como um todo. De maneira concomitante, o avanço das tecnologias permitiu a integração de áreas como a Ecologia com a computação, permitindo estudos que fossem capazes de gerar previsões não só atuais, mas futuras em relação às espécies e o ambiente em que estas estão inseridas. Conhecido como modelagem preditiva de distribuição de espécies, modelagem de nicho ecológico ou simplesmente modelagem preditiva, o processo de modelamento da relação entre espécies e ambiente se baseia em diferentes tipos de algoritmos computacionais visando atender não só a demanda por um conhecimento ecológico, mas atender principalmente estudos de conservação. O presente trabalho, demonstrou que a modelagem preditiva de distribuição de espécies é uma importante ferramenta aliada à Ecologia e à Biologia da Conservação e que, embora seja uma área em ascensão, ainda necessita de estudos quanto aos processos utilizados na produção dos modelos. Neste trabalho foi avaliada a interferência do tamanho da amostra no resultado final do modelo através da utilização de diferentes tamanhos amostrais para seis espécies de aves brasileiras, produzindo resultados que demonstram que o tamanho amostral é um dos principais pontos críticos para o processo de modelagem, requerendo atenção por parte do pesquisador para evitar modelos de baixa qualidade, ou ainda, que contenham informações que sub ou superestimam a real distribuição das espécies.

Palavras-chave: Modelagem preditiva. Distribuição de espécies. Conservação. Ecologia.

BRITO, G.R. **Application of predictive modeling of species distribution as a tool for the study of biodiversity**. 2018. 121 p. Dissertação (Mestrado em Biociências). – Universidade Estadual Paulista (UNESP), Faculdade de Ciências e Letras, Assis, 2018.

ABSTRACT

Abstract – Emerged in the mid-80s as a multidisciplinary science, the Conservation Biology was the result of the need to bring together different areas of knowledge in the face of environmental changes that affect the biota as a whole. At the same time, the advance of technologies permitted the integration of Ecology with Computation, allowing studies capable of generating not only current but future predictions regarding the species and the environment in which they are inserted. Known as species distribution modeling, ecological niche modeling, or simply, predictive modeling, the process of modeling the relationship between species and the environment is based on different types of computational algorithms, aimed at meeting not only the demand for ecological knowledge, but to attend the studies of conservation. The present work showed that the predictive modeling of species distribution is an important tool for Ecology and Conservation Biology and that although it is a growing area, it still needs studies on the process used in the production of the models. This study evaluated the interference of sample size in the final result of the model through the use of different sample sizes for six species of Brazilian birds, producing results that demonstrate that sample size is one of the main critical points for the modeling process, requiring attention on the part of the researcher to avoid low quality models or that contain information that under or overestimates the real distribution of the species.

Keywords: Predictive modeling. Species distribution. Conservation. Ecology.

Lista de Tabelas

Tabela 1 – Identificadores e descrição das variáveis bioclimáticas.....	39
Tabela 2 - Variáveis bioclimáticas utilizadas na modelagem das espécies estudadas.....	40
Tabela 3 - Valores de média e desvio padrão dos resultados de AUC.....	46
Tabela 4 - Valores médios de AUC após reamostragem.....	46
Tabela 5 - Valores de AUC obtidos na avaliação dos modelos construídos a partir dos tamanhos amostrais N_T , N_2 , N_3 e N_4 , espécie <i>Chiroxiphia caudata</i>	119
Tabela 6 - Valores de AUC obtidos na avaliação dos modelos construídos a partir dos tamanhos amostrais N_T , N_2 , N_3 e N_4 , espécie <i>Hemithraupis ruficapilla</i>	119
Tabela 7 - Valores de AUC obtidos na avaliação dos modelos construídos a partir dos tamanhos amostrais N_T , N_2 , N_3 e N_4 , espécie <i>Myiothlypis leucoblephara</i>	119
Tabela 8 - Valores de AUC obtidos na avaliação dos modelos construídos a partir dos tamanhos amostrais N_T , N_2 , N_3 e N_4 , espécie <i>Ramphocelus bresilius</i>	120
Tabela 9 - Valores de AUC obtidos na avaliação dos modelos construídos a partir dos tamanhos amostrais N_T , N_2 , N_3 e N_4 , espécie <i>Synallaxis ruficapilla</i>	120
Tabela 10 - Valores de AUC obtidos na avaliação dos modelos construídos a partir dos tamanhos amostrais N_T , N_2 , N_3 e N_4 , espécie <i>Tangara selodon</i>	120
Tabela 11 - Valores de AUC obtidos na avaliação dos modelos construídos a partir de amostras aleatórias, nos tamanhos amostrais N_2 , N_3 e N_4 para as espécies <i>Chiroxiphia caudata</i> , <i>Hemithraupis ruficapilla</i> e <i>Myiothlypis leucoblephara</i>	121

Tabela 12 - Valores de AUC obtidos na avaliação dos modelos construídos a partir de amostras aleatórias, nos tamanhos amostrais N_2 , N_3 e N_4 para as espécies *Ramphocelus bresilius*, *Synallaxis ruficapilla* e *Tangara seledon*.....121

Lista de Figuras

Figura 1 - Exibição do envelope climático para as variáveis ambientais Temperatura Média Anual e Precipitação Anual da espécie <i>Chiroxiphia caudata</i>	21
Figura 2 - Comparação entre os modelos de distribuição das espécies <i>Chiroxiphia caudata</i> , <i>Hemithraupis ruficapilla</i> e <i>Myiothlypis leucoblephara</i>	42
Figura 3 - Comparação entre os modelos de distribuição das espécies <i>Ramphocelus bresilius</i> , <i>Synallaxis ruficapilla</i> e <i>Tangara seledon</i>	43
Figura 4 - Percentual das áreas modeladas em comparação com os modelos originais.....	44
Figura 5 - Percentuais da área modelada em relação à variação da proporção de pseudo ausência por tamanho amostral.....	45
Figura 6 - Modelos de distribuição potencial de <i>Chiroxiphia caudata</i>	59
Figura 7 - Modelos de distribuição potencial de <i>Hemithraupis ruficapilla</i>	60
Figura 8 - Modelos de distribuição potencial de <i>Myiothlypis leucoblephara</i>	61
Figura 9 - Modelos de distribuição potencial de <i>Ramphocelus bresilius</i>	62
Figura 10 - Modelos de distribuição potencial de <i>Synallaxis ruficapilla</i>	63
Figura 11 - Modelos de distribuição potencial de <i>Tangara seledon</i>	64
Figura 12 – Mapas de distribuição potencial para 10 amostras aleatórias de 20 pontos, proporção de pseudo ausência 1x1, para a espécie <i>Chiroxiphia caudata</i>	65
Figura 13 – Mapas de distribuição potencial para 10 amostras aleatórias de 50 pontos, proporção de pseudo ausência 1x1, para a espécie <i>Chiroxiphia caudata</i>	66
Figura 14 – Mapas de distribuição potencial para 10 amostras aleatórias de 150 pontos, proporção de pseudo ausência 1x1, para a espécie <i>Chiroxiphia caudata</i>	67
Figura 15 – Mapas de distribuição potencial para 10 amostras aleatórias de 20 pontos, proporção de pseudo ausência 25x1, para a espécie <i>Chiroxiphia caudata</i>	68

Figura 16 – Mapas de distribuição potencial para 10 amostras aleatórias de 50 pontos, proporção de pseudo ausência 25x1, para a espécie <i>Chiroxiphia caudata</i>	69
Figura 17 – Mapas de distribuição potencial para 10 amostras aleatórias de 150 pontos, proporção de pseudo ausência 25x1, para a espécie <i>Chiroxiphia caudata</i>	70
Figura 18 – Mapas de distribuição potencial para 10 amostras aleatórias de 20 pontos, proporção de pseudo ausência 50x1, para a espécie <i>Chiroxiphia caudata</i>	71
Figura 19 – Mapas de distribuição potencial para 10 amostras aleatórias de 50 pontos, proporção de pseudo ausência 50x1, para a espécie <i>Chiroxiphia caudata</i>	72
Figura 20 – Mapas de distribuição potencial para 10 amostras aleatórias de 150 pontos, proporção de pseudo ausência 50x1, para a espécie <i>Chiroxiphia caudata</i>	73
Figura 21 – Mapas de distribuição potencial para 10 amostras aleatórias de 20 pontos, proporção de pseudo ausência 1x1, para a espécie <i>Hemithraupis ruficapilla</i>	74
Figura 22 – Mapas de distribuição potencial para 10 amostras aleatórias de 50 pontos, proporção de pseudo ausência 1x1, para a espécie <i>Hemithraupis ruficapilla</i>	75
Figura 23 – Mapas de distribuição potencial para 10 amostras aleatórias de 150 pontos, proporção de pseudo ausência 1x1, para a espécie <i>Hemithraupis ruficapilla</i>	76
Figura 24 – Mapas de distribuição potencial para 10 amostras aleatórias de 20 pontos, proporção de pseudo ausência 25x1, para a espécie <i>Hemithraupis ruficapilla</i>	77
Figura 25 – Mapas de distribuição potencial para 10 amostras aleatórias de 50 pontos, proporção de pseudo ausência 25x1, para a espécie <i>Hemithraupis ruficapilla</i>	78
Figura 26 – Mapas de distribuição potencial para 10 amostras aleatórias de 150 pontos, proporção de pseudo ausência 25x1, para a espécie <i>Hemithraupis ruficapilla</i>	79
Figura 27 – Mapas de distribuição potencial para 10 amostras aleatórias de 20 pontos, proporção de pseudo ausência 50x1, para a espécie <i>Hemithraupis ruficapilla</i>	80
Figura 28 – Mapas de distribuição potencial para 10 amostras aleatórias de 50 pontos, proporção de pseudo ausência 50x1, para a espécie <i>Hemithraupis ruficapilla</i>	81
Figura 29 – Mapas de distribuição potencial para 10 amostras aleatórias de 150 pontos, proporção de pseudo ausência 50x1, para a espécie <i>Hemithraupis ruficapilla</i>	82

Figura 30 – Mapas de distribuição potencial para 10 amostras aleatórias de 20 pontos, proporção de pseudo ausência 1x1, para a espécie <i>Myiothlypis leucoblephara</i>	83
Figura 31 – Mapas de distribuição potencial para 10 amostras aleatórias de 50 pontos, proporção de pseudo ausência 1x1, para a espécie <i>Myiothlypis leucoblephara</i>	84
Figura 32 – Mapas de distribuição potencial para 10 amostras aleatórias de 150 pontos, proporção de pseudo ausência 1x1, para a espécie <i>Myiothlypis leucoblephara</i>	85
Figura 33 – Mapas de distribuição potencial para 10 amostras aleatórias de 20 pontos, proporção de pseudo ausência 25x1, para a espécie <i>Myiothlypis leucoblephara</i>	86
Figura 34 – Mapas de distribuição potencial para 10 amostras aleatórias de 50 pontos, proporção de pseudo ausência 25x1, para a espécie <i>Myiothlypis leucoblephara</i>	87
Figura 35 – Mapas de distribuição potencial para 10 amostras aleatórias de 150 pontos, proporção de pseudo ausência 25x1, para a espécie <i>Myiothlypis leucoblephara</i>	88
Figura 36 – Mapas de distribuição potencial para 10 amostras aleatórias de 20 pontos, proporção de pseudo ausência 50x1, para a espécie <i>Myiothlypis leucoblephara</i>	89
Figura 37 – Mapas de distribuição potencial para 10 amostras aleatórias de 50 pontos, proporção de pseudo ausência 50x1, para a espécie <i>Myiothlypis leucoblephara</i>	90
Figura 38 – Mapas de distribuição potencial para 10 amostras aleatórias de 150 pontos, proporção de pseudo ausência 50x1, para a espécie <i>Myiothlypis leucoblephara</i>	91
Figura 39 – Mapas de distribuição potencial para 10 amostras aleatórias de 20 pontos, proporção de pseudo ausência 1x1, para a espécie <i>Ramphocelus bresilius</i>	92
Figura 40 – Mapas de distribuição potencial para 10 amostras aleatórias de 50 pontos, proporção de pseudo ausência 1x1, para a espécie <i>Ramphocelus bresilius</i>	93
Figura 41 – Mapas de distribuição potencial para 10 amostras aleatórias de 150 pontos, proporção de pseudo ausência 1x1, para a espécie <i>Ramphocelus bresilius</i>	94

Figura 42 – Mapas de distribuição potencial para 10 amostras aleatórias de 20 pontos, proporção de pseudo ausência 25x1, para a espécie <i>Ramphocelus bresilius</i>	95
Figura 43 – Mapas de distribuição potencial para 10 amostras aleatórias de 50 pontos, proporção de pseudo ausência 25x1, para a espécie <i>Ramphocelus bresilius</i>	96
Figura 44 – Mapas de distribuição potencial para 10 amostras aleatórias de 150 pontos, proporção de pseudo ausência 25x1, para a espécie <i>Ramphocelus bresilius</i>	97
Figura 45 – Mapas de distribuição potencial para 10 amostras aleatórias de 20 pontos, proporção de pseudo ausência 50x1, para a espécie <i>Ramphocelus bresilius</i>	98
Figura 46 – Mapas de distribuição potencial para 10 amostras aleatórias de 50 pontos, proporção de pseudo ausência 50x1, para a espécie <i>Ramphocelus bresilius</i>	99
Figura 47 – Mapas de distribuição potencial para 10 amostras aleatórias de 150 pontos, proporção de pseudo ausência 50x1, para a espécie <i>Ramphocelus bresilius</i>	100
Figura 48 – Mapas de distribuição potencial para 10 amostras aleatórias de 20 pontos, proporção de pseudo ausência 1x1, para a espécie <i>Synallaxis ruficapilla</i>	101
Figura 49 – Mapas de distribuição potencial para 10 amostras aleatórias de 50 pontos, proporção de pseudo ausência 1x1, para a espécie <i>Synallaxis ruficapilla</i>	102
Figura 50 – Mapas de distribuição potencial para 10 amostras aleatórias de 150 pontos, proporção de pseudo ausência 1x1, para a espécie <i>Synallaxis ruficapilla</i>	103
Figura 51 – Mapas de distribuição potencial para 10 amostras aleatórias de 20 pontos, proporção de pseudo ausência 25x1, para a espécie <i>Synallaxis ruficapilla</i>	104
Figura 52 – Mapas de distribuição potencial para 10 amostras aleatórias de 50 pontos, proporção de pseudo ausência 25x1, para a espécie <i>Synallaxis ruficapilla</i>	105
Figura 53 – Mapas de distribuição potencial para 10 amostras aleatórias de 150 pontos, proporção de pseudo ausência 25x1, para a espécie <i>Synallaxis ruficapilla</i>	106
Figura 54 – Mapas de distribuição potencial para 10 amostras aleatórias de 20 pontos, proporção de pseudo ausência 50x1, para a espécie <i>Synallaxis ruficapilla</i>	107

Figura 55 – Mapas de distribuição potencial para 10 amostras aleatórias de 50 pontos, proporção de pseudo ausência 50x1, para a espécie <i>Synallaxis ruficapilla</i>	108
Figura 56 – Mapas de distribuição potencial para 10 amostras aleatórias de 150 pontos, proporção de pseudo ausência 50x1, para a espécie <i>Synallaxis ruficapilla</i>	109
Figura 57 – Mapas de distribuição potencial para 10 amostras aleatórias de 20 pontos, proporção de pseudo ausência 1x1, para a espécie <i>Tangara seledon</i>	110
Figura 58 – Mapas de distribuição potencial para 10 amostras aleatórias de 50 pontos, proporção de pseudo ausência 1x1, para a espécie <i>Tangara seledon</i>	111
Figura 59 – Mapas de distribuição potencial para 10 amostras aleatórias de 150 pontos, proporção de pseudo ausência 1x1, para a espécie <i>Tangara seledon</i>	112
Figura 60 – Mapas de distribuição potencial para 10 amostras aleatórias de 20 pontos, proporção de pseudo ausência 25x1, para a espécie <i>Tangara seledon</i>	113
Figura 61 – Mapas de distribuição potencial para 10 amostras aleatórias de 50 pontos, proporção de pseudo ausência 25x1, para a espécie <i>Tangara seledon</i>	114
Figura 62 – Mapas de distribuição potencial para 10 amostras aleatórias de 150 pontos, proporção de pseudo ausência 25x1, para a espécie <i>Tangara seledon</i>	115
Figura 63 – Mapas de distribuição potencial para 10 amostras aleatórias de 20 pontos, proporção de pseudo ausência 50x1, para a espécie <i>Tangara seledon</i>	116
Figura 64 – Mapas de distribuição potencial para 10 amostras aleatórias de 50 pontos, proporção de pseudo ausência 50x1, para a espécie <i>Tangara seledon</i>	117
Figura 65 – Mapas de distribuição potencial para 10 amostras aleatórias de 150 pontos, proporção de pseudo ausência 50x1, para a espécie <i>Tangara seledon</i>	118

SUMÁRIO

Lista de Tabelas	9
Lista de Figuras	Erro! Indicador não definido.
INTRODUÇÃO GERAL	16
JUSTIFICATIVA	19
OBJETIVO.....	19
MÉTODOS	19
Escolha do algoritmo.....	19
Biologia das espécies	22
Chiroxiphia caudata	22
Hemithraupis ruficapilla.....	22
Myiothlypis leucoblephara.....	23
Ramphocelus bresilius.....	23
Synallaxis ruficapilla.....	24
Tangara seledon	24
REFERÊNCIAS BIBLIOGRÁFICAS	25
ARTIGO.....	31
INTRODUÇÃO	34
MÉTODOS.....	37
RESULTADOS.....	41
DISCUSSÃO	47
CONCLUSÃO	51
REFERÊNCIAS.....	52
CONCLUSÃO GERAL	58
APÊNDICE A – Mapas de distribuição potencial das espécies <i>Chiroxiphia caudata</i> , <i>Hemithraupis ruficapilla</i> , <i>Myiothlypis leucoblephara</i> , <i>Ramphocelus bresilius</i> , <i>Synallaxis ruficapilla</i> e <i>Tangara seledon</i>	59
APÊNDICE B – Mapas de distribuição potencial das espécies <i>Chiroxiphia caudata</i> , <i>Hemithraupis ruficapilla</i> , <i>Myiothlypis leucoblephara</i> , <i>Ramphocelus bresilius</i> , <i>Synallaxis ruficapilla</i> e <i>Tangara seledon</i> após processo de reamostragem	65
APÊNDICE C – Tabelas de valores AUC obtidos durante os processos de modelagem em diferentes tamanhos amostrais.....	119
APÊNDICE D – Tabelas de valores AUC obtidos durante os processos de modelagem para os conjuntos N2, N3 e N4 após reamostragem, nas proporções de pseudo ausência de 1x1, 25x1 e 50x1	121

INTRODUÇÃO GERAL

Surgida como ciência multidisciplinar em meados dos anos de 1980, a Biologia da Conservação visa auxiliar nos esforços de manejo e preservação da fauna e flora diante do crítico cenário pelo qual passa a biodiversidade (Perez-Sweeney *et al.*, 2003).

A integração dos estudos de biodiversidade com a computação deu origem aos primeiros modelos preditivos de distribuição de espécies e, posteriormente, softwares que incluíam diferentes algoritmos computacionais, resultando em ferramentas mais robustas aplicáveis à projetos voltados à conservação da biodiversidade (Giannini *et al.* 2011).

O processamento digital de dados, aliado ao conhecimento biológico, permitiu o avanço considerável no que tange à modelagem preditiva de distribuição de espécies, haja vista que a complexidade das interações entre os seres vivos e a escala que essas interações atingem requer de qualquer estudo de biodiversidade a execução de diferentes tarefas visando a coleta, documentação, análise e armazenamento de informações espaço-temporais das espécies (Michener *et al.*, 2012).

Em linhas gerais um modelo preditivo engloba os dados de ocorrência (presença e ausência) de uma espécie em um local e as variáveis ambientais da referida região. Integrando esses dois grupos de informações através de um algoritmo, o resultado é um conjunto de dados que permite visualizar a distribuição da espécie em relação às condições ambientais da região (Anderson *et al.* 2003).

De acordo com Guisan e Thuiller (2005), o interesse na computação para estudar a influência das variáveis ambientais nas distribuições de espécies se concentrou principalmente nas décadas de 1970, 1980 e 1990.

Os trabalhos pioneiros de Austin (1971), Nix *et al.* (1977), Ferrier (1984) e os livros de Verner *et al.* (1986) e Margules e Austin (1991) são referências no âmbito dos estudos de modelagem de distribuição de espécies através do uso da informática como ferramenta nos estudos de distribuição de espécies aplicados à projetos de conservação.

Por outro lado, a dificuldade em se definir o conceito de nicho exemplifica o quanto as interações e escalas existentes dentro de um sistema ecológico tornam a construção do conhecimento um campo onde a multidisciplinaridade torna-se imprescindível (Elith e Leathwick, 2009).

Conforme Giannini *et al.* (2011), Joseph Grinnel discutiu o nicho de várias espécies de aves e mamíferos em seus estudos (Grinnell 1904, 1917, 1924) apresentando-o como um atributo espacial; posteriormente Elton (1927) esboçaria nicho como sendo o papel de uma dada espécie dentro de uma cadeia alimentar.

A definição mais conhecida de nicho surgiu na década de 1940 através dos trabalhos de George Evelyn Hutchinson, conceituando nicho como “uma região de um hiper-volume n -dimensional”. De acordo com Hutchinson (1944), definido um número (n) de fatores limitantes à sobrevivência do organismo, a variação dos limites de tolerância a estes fatores constituiria um hiper-volume com n dimensões.

Uma visão estatisticamente mais rigorosa foi proposta por Chase e Leibold (2003), onde, através das definições conceituais de Grinnell (1924), Elton (1927) e Hutchinson (1957), nicho corresponderia às “condições ambientais que permitem à espécie suprir suas necessidades mínimas, de tal forma que a taxa de natalidade de uma população local seja igual ou maior à sua taxa de mortalidade, junto com o conjunto de efeitos *per capita* da espécie nessas condições ambientais”.

É devido à complexidade na conceituação de nicho e às diferentes visões no que tange quais fatores devem ser considerados para definir tal conceito que a metodologia de estudo das dinâmicas populacionais através de modelos computacionais seja tratada ora como modelagem de nicho ecológico (Austin 1992; Soberón e Peterson 2005), ora como modelagem de distribuição geográfica (Austin, 2002).

Brown e Lomolino (2006) apresentam a ideia de que a distribuição geográfica de uma espécie pode ser considerada como um resultado de sua história evolutiva e ecológica, resultado este influenciado por fatores de diferentes escalas e intensidades que atuam durante todo o histórico de desenvolvimento da espécie.

Como o foco de uma modelagem preditiva de espécies é determinar sua distribuição potencial e os fatores que possam ser considerados essenciais à sua existência em determinado local, considerações como a de Brown e Lomolino (2006) são pontos fundamentais para se definir como um modelo preditivo de espécies deve ser chamado.

Ainda, de acordo com Jiménez-Valverde *et al.* (2008b), seria necessário evitar a utilização do termo “modelagem de nicho” uma vez que, além de ser um termo complexo, o mesmo implica considerar efeitos não só dos fatores bióticos, mas abióticos também, complementando a ideia apresentada por Brown e Lomolino (2006).

Discorrendo sobre as técnicas de modelagens de distribuição, Lobo *et al.* (2008) apresenta que “estes modelos utilizam técnicas cada vez mais sofisticadas, mas muitas vezes falta um exame detalhado da qualidade dos dados de ocorrência biológicas em que se baseiam”.

Essa reflexão demonstra o quanto a compreensão da interação entre os seres vivos e o ambiente ainda é um dos grandes desafios da Ecologia moderna e, considerando

as abordagens mais recentes, optou-se por chamar a metodologia desenvolvida no presente trabalho de Modelagem Preditiva de Distribuição de Espécies.

JUSTIFICATIVA

Tendo em vista que, para realizar uma modelagem preditiva de distribuição de espécies faz-se necessário o acesso a uma base de dados multivariada/multidisciplinar, o modo como essa base foi criada torna-se um fator a ser considerado na hora de acessar os dados.

A padronização do esforço amostral é fundamental para a construção da base de dados, uma vez que o resultado da coleta de informações em campo pode gerar bases com dados insuficientes, pouco aproveitáveis para os processos de modelagem.

OBJETIVO

O presente trabalho teve por objetivo comparar os modelos preditivos para seis espécies de aves brasileiras para diferentes tamanhos amostrais através do algoritmo BIOCLIM.

MÉTODOS

Escolha do algoritmo

Desenvolvido por Nix (1986) na década de 1980, o BIOCLIM continua sendo um algoritmo bastante utilizado em trabalhos que envolvam a modelagem preditiva de distribuição de espécies. Embora tenha sido aplicado às modelagens de plantas e animais australianos, o algoritmo também se provou funcional em estudos realizados na África e América do Norte, resultando em boas modelagens preditivas de distribuição (Doran & Olsen, 2001). Trabalhos como os de Lindenmayer et al. (1991), Lindenmayer et al. (1996), Campbell et al. (1999), Sumner e Dickman (1998) e Jackson e Claridge

(1999) levaram o BIOCLIM a se tornar um importante e reconhecida ferramenta no campo dos estudos de conservação.

É graças a sua aplicabilidade nos mais diversos estudos que o BIOCLIM se mantém como um dos algoritmos mais utilizados em modelagens preditivas de distribuição de espécies, mesmo diante de algoritmos mais robustos e modernos como o MAXENT. São exemplos de aplicações atuais os trabalhos de Gabriel, Bennett e Muhammad (2015) com espécies de coníferas para reflorestamento na Nigéria, as modelagens de Braga (2015) para 20 espécies de aves da família *Tyrannidae* no continente Americano no que tange as áreas de reprodução e a interferência de mudanças climáticas sobre o processo migratório, as modelagens e o estudo do efeito amostral por Miranda (2012) utilizando espécies de *Vellozia*, endêmicas da Cadeia do Espinhaço, MG.

O modo de operação do algoritmo envolve a modelagem de quatro tipos de áreas de distribuição potencial com base no conceito de envelopes climáticos: áreas completamente fora do intervalo percentual 0-100 do envelope para uma ou mais variáveis climáticas levam o código "0". Áreas dentro do intervalo percentual 5-95 são marcadas com o código "3", enquanto que aquelas que estejam fora do referido intervalo, mas dentro do percentual 2.5-97.5 são identificadas com o código "2". O código "1" é definido para áreas que estão fora do intervalo 2.5-97.5, mas dentro do intervalo percentual 0-100 para todas as variáveis climáticas do modelo (Hijmans et al., 2012).

Modelos construídos através do algoritmo BIOCLIM, especificamente, combinam pontos de ocorrência da espécie dentro de uma área com dados climáticos correspondentes à mesma área (Fig 1). O processamento dos dados climáticos e dos dados observacionais permite estimar a amplitude de variação ambiental que é adequado

para a espécie em estudo (Guisan & Thuiller, 2005; Franklin, 2010; Hartley et al., 2010).

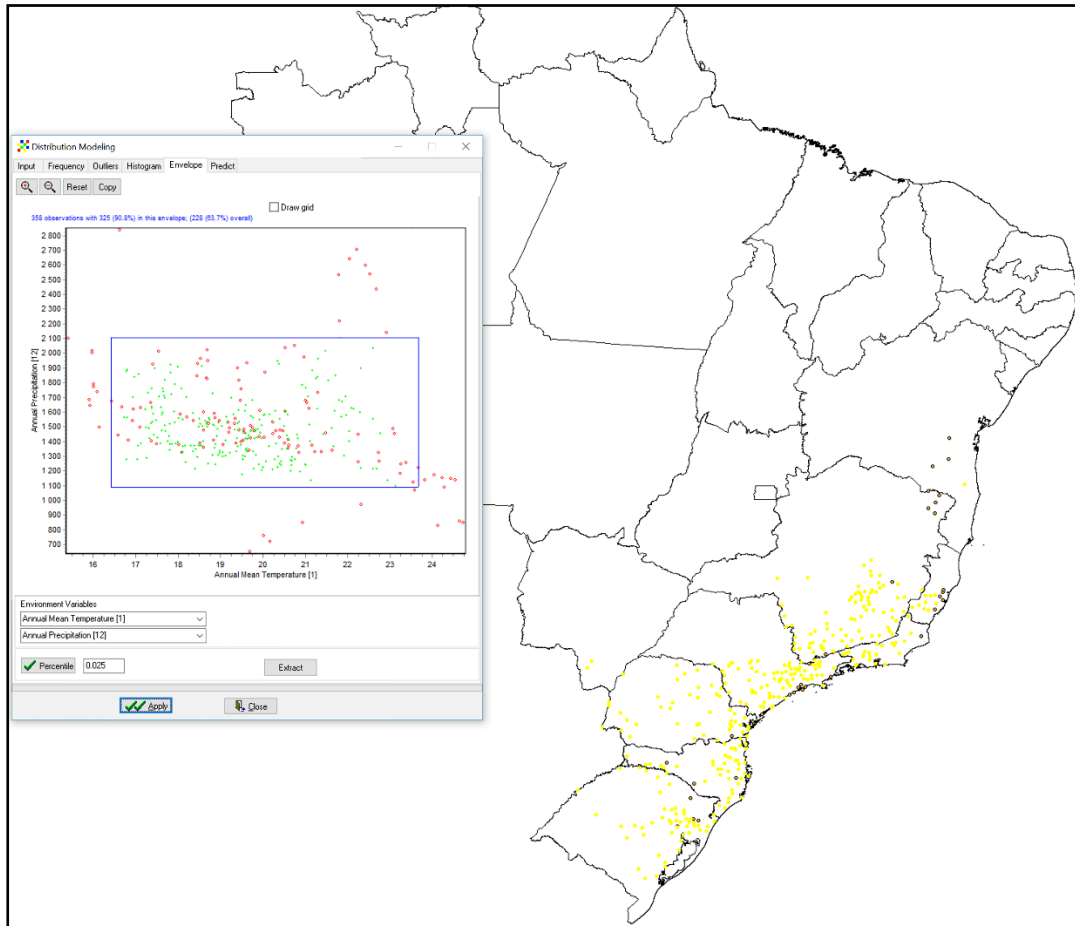


FIG. 1. Exibição do envelope climático para as variáveis ambientais Temperatura Média Anual e Precipitação Anual da espécie *Chiroxiphia caudata*. Os pontos coloridos em amarelo são aqueles que se enquadram no envelope climático definido.

Deste modo é possível estimar a atual distribuição de uma espécie e gerar previsões de novas áreas em potencial, ou seja, aquelas que possuam características ambientais iguais ou semelhantes à área original da espécie. Esse tipo de modelagem é elaborado com base no mecanismo de nicho grineliiano, delimitando o conjunto de condições ambientais adequadas para a espécie, dando origem a um envelope bioclimático (Soberón, 2007).

O resultado final da modelagem é um mapa com as distribuições potenciais da espécie baseado em uma escala de cor onde as cores exibidas pelo modelo correspondem ao agrupamento dos pontos de observação ao longo da área modelada com base no conceito de envelope climático utilizado pelo algoritmo.

Biologia das espécies

Chiroxiphia caudata

O Tangará, também chamado de Tangará-dançarino ou dançador, é uma ave passeriforme da família Pipridae. Medindo cerca de 13 centímetros de comprimento, é uma espécie com dimorfismo sexual marcante: os machos possuem uma plumagem azul-celeste, cauda preta e uma coroa vermelho-brilhante no alto da cabeça, enquanto as fêmeas são verde-escuras e com cauda mais longa que a dos machos; indivíduos jovens são verde-oliva diferindo das fêmeas apenas devido à coroa vermelha no alto da cabeça.

É uma espécie de hábitos alimentares onívoros, alimentando-se de bagas e pequenos artrópodes, apreciando também a fruta-do-sabiá e magnólia-amarela. Em termos reprodutivos, as fêmeas do Tangará possuem território próprio ao redor do ninho e costumam depositar dois ovos, incubando-os durante 18 dias; o sistema de reprodução é poligínico, com vários machos se agregando em locais tradicionais para exposições cooperativas, respeitando uma hierarquia linear que pode persistir por anos (WikiAves, 2008).

Hemithraupis ruficapilla

Conhecida popularmente como Saíra-ferrugem, é uma espécie endêmica do Brasil, com 14 centímetros de comprimento e peso aproximado de 13 gramas. É possível notar um grau de dimorfismo sexual através da coloração da plumagem, com fêmeas

quase totalmente esverdeadas em contraste à complexidade dos machos que apresentam a cabeça na cor ferrugem, lados do pescoço amarelados e peito marrom-claro.

Comum em bordas de florestas, bosques e capoeiras, apresentando comportamentos de forrageamento em bandos mistos ou pequenos grupos; se alimenta tanto de frutos como de insetos, principalmente larvas encontradas nos troncos das árvores, local onde costuma construir seu ninho (WikiAves, 2008).

Myiothlypis leucoblephara

Ave passeriforme da família Parulidae, medindo cerca de 14,5 centímetros e com peso entre 14 e 21 gramas, o *M.leucoblephara* é conhecido popularmente como Pula-pula-assobiador. Apresenta coloração acinzentada e asas e cauda em tons amarelados-oliva, sendo distinguidos entre um ou outro indivíduo através de variações sutilmente discretas na estrutura do canto.

Possui hábitos característicos de sub-bosques das serras altas cobertas por mata, sendo frequente também em taquarais. É insetívoro e procura seu alimento de preferência no solo, onde também constrói seu ninho sob troncos caídos ou na base de arvores (WikiAves, 2008).

Ramphocelus bresilius

Espécie endêmica do Brasil e símbolo da Mata Atlântica, o popularmente conhecido Tiê-sangue é uma ave passeriforme da família Thraupidae, pesando cerca de 31 gramas e medindo 19 centímetros de comprimento. Apresenta dimorfismo sexual, sendo a plumagem da fêmea de cor parda e marrom-avermelhada em contraste com o vermelho-vivo exibido pelos machos.

A reprodução da espécie se dá na primavera e no verão, com ninhos construídos na forma de cesto; as posturas ocorrem de duas a três vezes por temporada, compreendendo um número de dois ou três ovos, sendo estes incubados apenas pela fêmea.

São frugívoros, tendo predileção pelos frutos da embaúba e grande apreço pelos frutos da fruta-de-sabiá ou marianeira (*Acnistus arborescens*), podendo ainda se alimentar de insetos e vermes. Costuma viver aos pares e frequentar comedouros, podendo ser encontrado em capoeiras baixas, bordas de florestas, restingas e plantações, ou ainda em parques e praças de cidades (WikiAves, 2008).

Synallaxis ruficapilla

Também chamado de Pichororé, é uma ave passeriforme da família Furnariidae, medindo entre 13 e 17 centímetros e pesando entre 12 e 16 gramas. Possui uma coroa inteiramente castanho avermelhada brilhante, com garganta e peito em tons acinzentados, facilitando a identificação de um indivíduo imaturo graças as diferenças nas cores exibidas pela plumagem.

Seus hábitos alimentares incluem insetos e suas larvas, aranhas, opiliões e outros artrópodes e ainda moluscos; vive nas brenhas internas e na orla da mata, podendo pular no solo em algumas ocasiões. Seus hábitos reprodutivos envolvem a construção de ninhos através de um amontoado denso e comprido de gravetos, com entrada superior e, quase sempre, “decorado” com peles de cobras e lagartos (WikiAves, 2008).

Tangara seledon

Conhecido popularmente como Saíra-sete-cores, o Tangara seledon é uma ave passeriforme da família Thraupidae, medindo cerca de 13,5 centímetros e pesando cerca de 18 gramas. Os machos apresentam plumagem complexa e brilhante, diferindo das

fêmeas apenas na intensidade das cores exibidas. É uma ave monogâmica, com postura de dois a quatro ovos por ciclo reprodutivo, incubados pela fêmea.

É frugívoro, apreciando frutos não só silvestres, mas também goiaba, mamão, ameixa e caju; pode alimentar-se de insetos também. Seus hábitos de forrageamento facilitam sua observação pois costuma se juntar em pequenos grupos que buscam alimento ao longo das bordas da floresta podendo ainda forragear em áreas antropizadas (WikiAves, 2008).

REFERÊNCIAS BIBLIOGRÁFICAS

ANDERSON, R.P.; LEW, D.; PETERSON, A.T. Evaluating predictive models of species distributions: criteria for selecting optimal models. *Ecol. Model.*, v. 162, p. 211–232, 2003.

AUSTIN, M.P. Role of regression analysis in plant ecology. *Proc. Ecol. Soc. Aust.*, v. 6, p. 63–75, 1971.

AUSTIN, M.P. Modelling the environmental niche of plants: implications for plant community response to elevated CO₂ levels. *Australian Journal of Botany*, v. 40, p. 615–30, 1992.

AUSTIN, M.P. Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecological Modelling*, v. 157, p. 101–18, 2002.

BRAGA, M. R. Mudanças climáticas e migração de aves (Tyrannidae) nas Américas. 2015. 1 recurso online (xiv, 43p). Dissertação (mestrado) – Universidade Estadual de Campinas, Instituto de Biologia, Campinas, SP. Disponível em: <http://www.biblioteca-digital.unicamp.br/document/?code=000953033>. Acesso em: 13 jun. 2016.

BROWN, J. H. & LOMOLINO, M.V. 2006. Biogeografia. 2nd ed. Ribeirão Preto: Funpec, 2006. 692p

CAMPBELL, B.D.; MITCHELL, N.D.; FIELD, T.R. Climate profiles of temperature C3 and subtropical C4 species in New Zealand pastures. New Zealand Journal of Agricultural Research, v. 42, p. 223-233, 1999.

CHASE, J.M. & LEIBOLD, M.A. Ecological niches: linking classical and contemporary approaches. Chicago: The University of Chicago Press, 2003.

DORAN, B., & OLSEN, P. Customizing BIOCLIM to investigate spatial and temporal variations in highly mobile species. In: Proceedings of the 6th International Conference on GeoComputation, 24-26 September, 2001, University of Queensland, Brisbane, Australia.

ELITH, J. & LEATHWICK, J. R. Species distribution models: ecological explanation and prediction across space and time. Annual Review of Ecology, Evolution and Systematics, v. 40, p. 677-97, 2009.

ELTON, C. Animal ecology. London: Sidwig and Jackson, 1927. 296p.

FERRIER, S. The status of the Rufous Scrub-Bird *Atrichornis rufescens*: habitat, geographical variation and abundance. PhD Thesis, Armidale, Australia, University of New England, 1984.

FRANKLIN, J. Moving beyond static species distribution models in support of conservation biogeography. Divers. Distrib., v. 16, p. 321–330, 2010.

GABRIEL, S.; BENNETT, C.; BA, A.M.; HENRY, S. Modeling the Suitability Index of Selected Conifers on Mambilla Plateau Taraba State, Nigeria: Implication on Planted Forest. *International Journal of Agroforestry Remote Sensing and GIS*, v. 1, n. 1-9, 2015.

GIANNINI, Tereza C. et al. Desafios atuais da modelagem preditiva de distribuição de espécies. *Rio de Janeiro: Rodriguésia*, v. 63, n. 3, p. 733-749, Sept. 2011.

GRINNELL, J. The origin and distribution of the chestnut-backed chickadee. *Auk*, v. 21, p. 375-377, 1904.

GRINNELL, J. The niche-relationship of the California thrasher. *Auk*, v. 34, p. 427-433, 1917.

GRINNELL, J. Geography and evolution. *Ecology*, v. 5, p. 225-229, 1924.

GUISAN, A.; THUILLER, W. Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, v. 8, p. 993–1009, 2005.

HIJMANS, R.J. L.; GUARINOM, P. DIVA-GIS Version 7.5 Manual. Mathur, p. 77, 2012.

HUTCHINSON, G.E. Limnological studies in Connecticut. Part 7. A critical examination of the supposed relationship between phytoplankton periodicity and chemical changes in lake waters. *Ecology*, v. 35, p. 3-26, 1944.

HUTCHINSON, G.E. Concluding remarks. *Cold Spring Harbour Symposium on Quantitative Biology*, v. 22, p. 415-427, 1957.

JACKSON, S.M. & A. CLARIDGE. Climatic modelling of the distribution of the mahogany glider (*Petaurus gracilis*), and the squirrel glider (*P. norfolcensis*). *Collingwood: Australian Journal of Zoology*, v. 24, p. 47-57, 1999.

JIMÉNEZ-VALVERDE, A.; LOBO, J.M. & HORTAL, J. Not as good as they seem: the importance of concepts in species distribution modeling. *Diversity and Distributions*, v. 14, p. 885-890, 2008b.

LINDENMAYER, D.B.; NIX, H.A.; MCMAHON, J.P.; HUTCHINSON, M.F. & TANTON, M.T. The conservation of Leadbeater's Possum, *Gymnobelideus leadbeateri* (McCoy) – A case study of the use of bioclimatic modelling. *Journal of Biogeography*, v. 18, p. 371–383, 1991.

LINDENMAYER, D.B.; MACKEY, B. & NIX, H.A. Climatic analyses of the distribution of four commercially important wood production eucalypt trees from southeastern Australia. *Australian Forestry*, v. 59, p. 11–26, 1996.

LOBO, J.M. Database records as a surrogate for sampling effort provide higher species richness estimations. *Biodiversity and Conservation*, v. 17, p. 873-881, 2008.

MARGULES, C.R. & AUSTIN, M.P. *Nature Conservation: Cost Effective Biological Survey and Data Analysis*. CSIRO, Canberra, Australia, 1991.

MICHENER, W. K. et al. Participatory design of DataONE—Enabling cyberinfrastructure for the biological and environmental sciences. *Ecological Informatics*, v. 11, p. 5–15, 2012.

MIRANDA, M. D. Modelos de distribuição de espécies de *Vellozia* (Velloziaceae) endêmicas da cadeia do espinhaço e o efeito amostral sobre os mapas preditivos. 2012. 92 f. Dissertação (Mestrado em Biologia Vegetal). Instituto de Ciências Biológicas da Universidade Federal de Minas Gerais, Belo Horizonte, 2012.

NIX, H.A. A biogeographic analysis of Australian elapid snakes. Atlas of elapid snakes of Australia: Australian flora and fauna series 7 (ed. by R. Longmore), p. 4–15. Bureau of Flora and Fauna, Canberra, 1986.

NIX, H.; MCMAHON, J. & MACKENZIE, D. Potential areas of production and the future of pigeon pea and other grain legumes in Australia. In: The potential for pigeon pea in Australia. Proceedings of Pigeon Pea (*Cajanus cajan* (L.) Millsp.) Field Day (eds Wallis, E.S. & Whiteman, P.C.). University of Queensland, Queensland, Australia, p. 5/1–5/12, 1977.

PEREZ-SWEENEY, B. M.; RODRIGUES, F.P. and MELNICK, D.J. Metodologias moleculares utilizadas em genética da conservação. In CULLEN Jr., L., RUDRAN, R., and VALLADARES-PÁDUA, C. (eds.). Manual Brasileiro em Biologia da Conservação. Curitiba, Editora da Universidade Federal do Paraná e Fundação O Boticário de Proteção à Natureza, p. 343-380, 2003.

SOBERÓN, J. & PETERSON, A.T. Interpretation of models of fundamental ecological niches and species distributional areas. *Biodiversity Informatics*, v. 2, p. 1-10, 2005.

SOBERÓN, J. Grinnellian and Eltonian niches and geographic distributions of species. *Ecology Letters*, Oxford, v. 10, p. 1115-1123, 2007.

SUMNER, J. M. and DICKMAN, C. R. Distribution and identity of species in the *Antechinus stuartii*–*A. flavipes* group (Marsupialia: Dasyuridae) in south-eastern Australia. *Australian Journal of Zoology*, v. 46, p. 27–41, 1998.

VERNER, J.; MORRISON, M.L. & RALPH, C.J. *Wildlife 2000: Modelling Habitat Relationships of Terrestrial Vertebrates*. Madison, Wisconsin: University of Wisconsin Press. 1986.

WikiAves (2008) [*tangará (Chiroxiphia caudata)*]. WikiAves, a Enciclopédia das Aves do Brasil. Disponível em: <http://www.wikiaves.com.br/tangara>. Acesso em: 12/2/2016.

WikiAves (2008) [*saíra-ferrugem (Hemithraupis ruficapilla)*]. WikiAves, a Enciclopédia das Aves do Brasil. Disponível em: <http://www.wikiaves.com.br/saira-ferrugem>. Acesso em: 12/2/2016.

WikiAves (2008) [*pula-pula-assobiador (Myiothlypis leucoblephara)*]. WikiAves, a Enciclopédia das Aves do Brasil. Disponível em: <http://www.wikiaves.com.br/pula-pula-assobiador>. Acesso em: 12/2/2016.

WikiAves (2008) [*tiê-sangue (Ramphocelus bresilius)*]. WikiAves, a Enciclopédia das Aves do Brasil. Disponível em: <http://www.wikiaves.com.br/tie-sangue>. Acesso em: 12/2/2016.

WikiAves (2008) [*pichororé (Synallaxis ruficapilla)*]. WikiAves, a Enciclopédia das Aves do Brasil. Disponível em: <http://www.wikiaves.com.br/pichorore>. Acesso em: 12/2/2016.

WikiAves (2008) [*saíra-sete-cores (Tangara seledon)*]. WikiAves, a Enciclopédia das Aves do Brasil. Disponível em: <http://www.wikiaves.com.br/saira-sete-cores>. Acesso em: 12/2/2016.

ARTIGO

ESTUDO DO EFEITO DA VARIAÇÃO DO TAMANHO AMOSTRAL EM MODELAGENS PREDITIVAS

DE DISTRIBUIÇÃO DE ESPÉCIES BASEADAS EM UM ALGORITMO DE ENVELOPE CLIMÁTICO

Gustavo Reis de Brito

Abstract. - Species distribution modeling emerged as an important tool for Ecology and Conservation Biology studies. Among them, one of the most used are the predictive distribution models based on the concept of climatic envelope. Although the application of predictive modeling has been on the rise since the earliest studies in the 1960s, the discussion about the sample effect in the modeling process still lacks further studies. Based on this, the present work sought to analyze the influence of sample size on modeling performance through four distinct sample sizes for six Brazilian bird species using BIOCLIM as the computational algorithm. The obtained results indicate that the reduction of the sample size influences the performance of the algorithm, reducing the modeled area and underestimating the potential distribution of the studied species. It was also verified that the ecological characteristics of the species influence the modeling process, exhibiting areas of greater potential near the places where most points of observation are concentrated. Therefore, studies aimed at predictive modeling of species distribution should consider the sample size and the ecological knowledge of the species as critical factors for the modeling process.

Keywords: BIOCLIM, Climate envelope, Distribution modeling

**Study of the effect of sample size variation in predictive modelling of species distribution
based on climatic envelope algorithm**

Resumo. A modelagem preditiva de distribuição de espécies surgiu como ferramenta importante para os estudos de Ecologia e Biologia da Conservação. Um dos tipos mais aplicados de modelagem são baseados no conceito de envelope climático. Muito embora a aplicação da modelagem preditiva esteja em ascensão desde os primeiros estudos na década de 1960, a discussão sobre o efeito amostral no processo de modelagem ainda carece de mais estudos. Baseado nisso, o presente trabalho procurou analisar a influência do tamanho amostral na performance da modelagem através de quatro tamanhos amostrais distintos para seis espécies de aves brasileiras, utilizando como algoritmo computacional o BIOCLIM. Os resultados obtidos apontam que a redução do tamanho amostral influencia a performance do algoritmo, reduzindo a área modelada e subestimando a distribuição potencial das espécies estudadas. Foi verificado também que as características ecológicas das espécies influenciam o processo de modelagem, exibindo áreas de maior potencial próximas dos locais onde se concentram a maioria dos pontos de observação. Logo, trabalhos voltados à modelagem preditiva de distribuição de espécies devem considerar o tamanho amostral e o conhecimento ecológico das espécies como fatores críticos para o processo de modelagem.

Palavras-chave: BIOCLIM, Envelope climático, Modelagem de distribuição

INTRODUÇÃO

O desenvolvimento dos Modelos de Distribuição de Espécies (Species Distribution Models - SDMs, em inglês) é resultado da integração do conhecimento biológico com as diversas tecnologias (Giannini et al., 2011), unindo áreas como Ecologia e Computação na criação de modelos que permitam ao pesquisador estudar a distribuição real e potencial de espécies (Anderson et al., 2003).

Existem diferentes classificações para um modelo de distribuição de espécies, sendo duas as mais utilizadas: Modelos Lineares Generalizados (Generalised Linear Models - GLM) e Modelos Aditivos Generalizados (Generalised Additive Models - GAM) (Austin, 2007).

Modelos GLM são uma forma derivada dos modelos lineares clássicos (McCullagh and Nelder, 1989), dos quais a regressão logística apresenta-se como um dos métodos mais aplicados na modelagem de distribuição de espécies (Guisan e Zimmermann, 2000; Rushton et al., 2004). Em contra-partida, modelos GAM são apresentados como uma forma “aprimorada” dos modelos lineares generalizados por permitirem identificar e caracterizar efeitos não lineares da regressão (Hastie e Tibshirani, 1990).

A construção de um modelo preditivo linear ou aditivo requer a utilização de variáveis ambientais em conjunto com informações georreferenciadas referentes à presença e, em casos específicos, a ausência da espécie no local a ser modelado. Além da classificação de modelos GLM ou GAM, estes se subdividem em modelos que consideram apenas a presença da espécie (*presence only*) e aqueles que necessitam também da ausência (*presence-absence*) (Dettmers e Bart, 1999).

O total de pontos de presença da espécie é resultado do esforço amostral do pesquisador na coleta de dados em campo, em bases de dados virtuais ou em museus. A discussão sobre

qual o tamanho amostral “ideal” necessita de novos estudos justamente por esse tema ser dependente de inúmeros fatores como a raridade e endemismo da espécie (Kunin & Gaston, 1993), a dificuldade da amostragem na área habitada pela espécie ou ainda a indisponibilidade dos dados em formato eletrônico (Graham et al., 2004).

O efeito da falta de consenso entre quais tamanhos amostrais a utilizar em uma modelagem preditiva de distribuição de espécies é evidenciado na diversidade de tamanhos amostrais encontrados na literatura. Cassemiro, Gouveia e Diniz-Filho (2012) modelaram a distribuição de *Rhinella granulosa* com 71 pontos de presença, enquanto Stein (2015) utilizou 218 pontos de presença ao modelar a mesma espécie.

Outros trabalhos, por sua vez, visam discutir a influência de diferentes tamanhos amostrais no processo de modelagem sem de fato discutir diretamente quais os impactos dessa falta de consenso para a modelagem como um todo, como é o caso de Wisz et al. (2008) que gerou predições para 46 espécies usando três tamanhos amostrais distintos (100, 30 e 10 pontos de presença, respectivamente) e Bean, Stafford e Brashares (2012) que, por sua vez, utilizaram conjuntos amostrais contendo 5, 10, 100 e 1000 pontos de presença.

Além da diversidade de tamanhos amostrais, é comum a muitos trabalhos não apresentar a quantidade de pontos de presença utilizados, como visto em Gregorini et al. (2007), Siqueira e Durigan (2007), Giovanelli et al. (2008), Álvares (2009) e Nabout et al. (2009).

Por outro lado, amostrar a ausência da espécie pode ser uma tarefa difícil e/ou enviesada de acordo com a espécie e a região estudada (Dennis e Thomas, 2000; Reddy e Dávalos, 2003; Schulman et al., 2007). Alguns tipos de modelagem utilizam a amostragem de *background*, o que permite a criação de um conjunto de dados de pontos de pseudo ausência (Philips et al., 2009).

Diferente da utilização de dados de ausência, modelos que utilizam pseudo ausência são também classificados como *presence only*, justamente pela necessidade da simulação de pontos de ausência da espécie em estudo. (Tsoar et al., 2007).

As discussões sobre o total de pontos de pseudo ausência a se utilizar na modelagem de uma espécie aparecem com pouca frequência na literatura, gerando trabalhos com os mais variados tamanhos amostrais de pseudo ausência como é o caso de Engler, Guisan e Rechsteiner (2004) que utilizam de 45 a 77 pontos de pseudo ausência em suas modelagens. Wisz e Guisan (2009) simulam 10.000 pontos, e Zaniwski, Lehmann e Overton (2002) usam até 19.875 pontos de pseudo ausência.

Stockwell e Peterson (2002) e Hernandez, Graham, Master e Albert (2006) apresentam em seus trabalhos discussões a respeito da influência do tamanho amostral utilizado na modelagem preditiva de distribuição de uma espécie, demonstrando que é necessário definir um conjunto de dados baseado não só na performance do algoritmo que irá modelar a espécie, mas o conhecimento da biologia da mesma.

Baseado nas discussões sobre a influência do tamanho amostral na qualidade do modelo final, o presente trabalho teve por objetivo avaliar modelos de distribuição para 6 espécies de aves brasileiras, utilizando o algoritmo BIOCLIM, através da variação dos tamanhos amostrais durante a construção dos modelos para cada uma das espécies, partindo do tamanho amostral original.

MÉTODOS

Seleção do algoritmo

Embora existam diferentes algoritmos voltados às práticas de modelagem de distribuição de espécies, como é o caso do MaxEnt, GARP e Domain, a escolha do algoritmo BIOCLIM para o estudo da influência do tamanho amostral sobre modelos preditivos de distribuição de espécies foi baseada em fatores como o pioneirismo do algoritmo e a facilidade na compreensão dos processos de modelagem.

Quatro tipos de áreas de distribuição potencial são modeladas com base envelopes climáticos definidos pelo algoritmo: áreas completamente fora do intervalo percentual 0-100 do envelope para uma ou mais variáveis climáticas levam o código "0". Áreas dentro do intervalo percentual 5-95 são marcadas com o código "3", enquanto que aquelas que estejam fora do referido intervalo, mas dentro do percentual 2.5-97.5 são identificadas com o código "2". O código "1" é definido para áreas que estão fora do intervalo 2.5-97.5, mas dentro do intervalo percentual 0-100 para todas as variáveis climáticas do modelo (Hijmans et al., 2012).

As cores utilizadas no modelo final correspondem às distribuições potenciais da espécie ao longo da área modelada com base no conceito de envelope climático utilizado pelo algoritmo.

Seleção das espécies e definição dos tamanhos amostrais

Seis espécies de aves brasileiras foram utilizadas nos trabalhos de modelagem preditiva. Os tamanhos amostrais originais variaram de 250 a 600 pontos de presença devidamente georreferenciados e catalogados, oriundos do Museu Nacional – UFRJ, Rio de Janeiro, RJ.

Partindo do tamanho amostral original de cada espécie (N_T), três novas amostragens foram realizadas através do software InStat, de forma aleatória e sem reposição, para a obtenção de três novos conjuntos amostrais contendo $20(N_4)$, $50(N_3)$ e $150(N_2)$ pontos de presença.

Ao final de todas as simulações durante a primeira etapa do trabalho foram obtidos 240 modelos preditivos de distribuição de espécies, sendo quarenta modelos por espécie baseado no total de conjuntos amostrais (quatro tamanhos amostrais por espécie) e no total de pontos de pseudo ausência utilizado na modelagem de cada espécie.

A segunda etapa do presente trabalho compreendeu a obtenção de 10 amostras aleatórias nos tamanhos N_2 , N_3 e N_4 para cada uma das espécies a partir do tamanho N_T , totalizando 30 modelos por espécie, 180 modelos no geral. Para cada amostra foi produzido três modelos diferentes com base na proporção de pontos de pseudo ausência para cada ponto de presença da amostra, a saber: 1x1, 25x1 e 50x1.

Biologia das espécies

Baseado em informações da base de dados Wiki Aves (<http://www.wikiaves.com.br>), o presente trabalho utilizou seis espécies de aves brasileiras para testar a influência do tamanho amostral na modelagem de distribuição de cada espécie.

Das seis espécies escolhidas, duas são endêmicas do Brasil, sendo encontradas no bioma de Mata Atlântica (*Ramphocelus bresilius*) e em florestas de terrenos baixos subtropicais e tropicais úmidas, florestas de montanha subtropicais e tropicais (*Hemithraupis ruficapilla*), respectivamente.

Synallaxis ruficapilla, *Tangara seledon*, *Myiothlypis leucoblephara* e *Chiroxiphia caudata*, por sua vez, se apresentam como espécies de distribuição mais generalista em relação aos pontos coletados, com ampla distribuição pelos estados do Sul, Sudeste, Centro-Oeste até partes do Nordeste.

Seleção das variáveis ambientais

O conjunto de 19 variáveis bioclimáticas (Tabela 1) obtido na base de dados do Worldclim (www.worldclim.org) é derivado dos valores mensais de temperatura e precipitação. Representando tendências anuais, sazonalidade e fatores extremos e/ou limitantes, são as principais variáveis ambientais utilizadas para modelar a distribuição de espécies.

Tabela 1. Identificadores e descrição das variáveis bioclimáticas

Identificador	Variável bioclimática
BIO1	Annual Mean Temperature
BIO2	Mean Diurnal Range (Mean of monthly (max temp - min temp))
BIO3	Isothermality (BIO2/BIO7) (* 100)
BIO4	Temperature Seasonality (standard deviation *100)
BIO5	Max Temperature of Warmest Month
BIO6	Min Temperature of Coldest Month
BIO7	Temperature Annual Range (BIO5-BIO6)
BIO8	Mean Temperature of Wettest Quarter
BIO9	Mean Temperature of Driest Quarter
BIO10	Mean Temperature of Warmest Quarter
BIO11	Mean Temperature of Coldest Quarter
BIO12	Annual Precipitation
BIO13	Precipitation of Wettest Month
BIO14	Precipitation of Driest Month
BIO15	Precipitation Seasonality (Coefficient of Variation)
BIO16	Precipitation of Wettest Quarter
BIO17	Precipitation of Driest Quarter
BIO18	Precipitation of Warmest Quarter
BIO19	Precipitation of Coldest Quarter

Utilizando o software NCSS foi realizada a análise de multicolineariedade visando excluir do conjunto de variáveis bioclimáticas aquelas que fossem correlacionadas, reduzindo assim o número de variáveis utilizado na modelagem de cada espécie (Tabela 2).

Tabela 2. Variáveis bioclimáticas utilizadas na modelagem das espécies estudadas

Espécie	Variáveis utilizadas
<i>Chiroxiphia caudata</i>	BIO3, BIO7, BIO8, BIO9, BIO12, BIO13, BIO18
<i>Hemithraupis ruficapilla</i>	BIO3, BIO5, BIO7, BIO8, BIO13, BIO18, BIO19
<i>Myiothlypis leucoblephara</i>	BIO3, BIO5, BIO6, BIO7, BIO8, BIO9, BIO12, BIO13
<i>Ramphocelus bresilius</i>	BIO2, BIO4, BIO5, BIO13, BIO14, BIO19
<i>Synallaxis ruficapilla</i>	BIO3, BIO7, BIO8, BIO9, BIO12, BIO13, BIO18
<i>Tangara seledon</i>	BIO2, BIO3, BIO8, BIO9, BIO13, BIO19

Geração de pontos de pseudo ausência

A primeira etapa dos processos de modelagem utilizou dez conjuntos variados de pontos de pseudo ausência, propostos como padrão para as modelagens das espécies, tendo o menor dos conjuntos 5 pontos de pseudo ausência e o maior, 50 pontos de ausência para cada ponto de presença. Por sua vez, os processos de modelagens realizados durante a segunda etapa se basearam em três proporções de pseudo ausência, sendo estas 1x1, 25x1 e 50x1, respectivamente.

Obtenção da curva ROC, valores de AUC e mensuração da área modelada

Cada modelo foi avaliado através da técnica do *Receiver Operator Curve*, a partir da qual foram obtidos os valores AUC pelo mesmo software utilizado para as modelagens (DIVA-GIS). Posteriormente, para cada modelo foi feito o cálculo da área potencial, através do software de edição de imagens Image J, onde foi realizada a contagem de partículas (pixels) após a transformação dos modelos RGB em uma escala de preto e branco.

Os valores de área obtidos foram utilizados para comparar os diversos modelos preditivos, gerados a partir de diferentes tamanhos amostrais e com diferentes quantidades de pontos de pseudo ausência, com o modelo base de distribuição da espécie.

RESULTADOS

Modelos. Os modelos produzidos nas duas fases do trabalho apresentam resultados semelhantes no que tange à alteração da área modelada. Existem alterações significativas nas áreas modeladas conforme se varia o tamanho amostral, ficando evidente quando feitas as comparações com o modelo original de distribuição espécie (Figs 2 e 3). Pode-se observar que os modelos das menores amostras são geralmente restritos aos pontos, o que denota menores potenciais de distribuição (Fig. 3, B2).

As áreas do modelo tendem a acompanhar a distribuição das espécies pelo ambiente. Os modelos A1, B1, C1 (Fig 2) e B1 (Fig 3), baseados em espécies de distribuição mais generalista, apresentam áreas e cores melhores distribuídas acompanhando os pontos de observação das espécies. Por outro lado, os modelos A1 e C1 (Fig 3) possuem áreas de distribuição com tendência mais restritivas, estando as cores de maior distribuição potencial (e.g. vermelho, laranja e amarelo) nas áreas onde há maior concentração dos pontos de observação.

Essa característica tende a se manter em todos os modelos, não sofrendo total interferência do tamanho amostral da espécie, indicando que os modelos são influenciados não só pela amostra, mas por características ecológicas da espécie em estudo e pelo processo de modelagem utilizado pelo algoritmo escolhido.

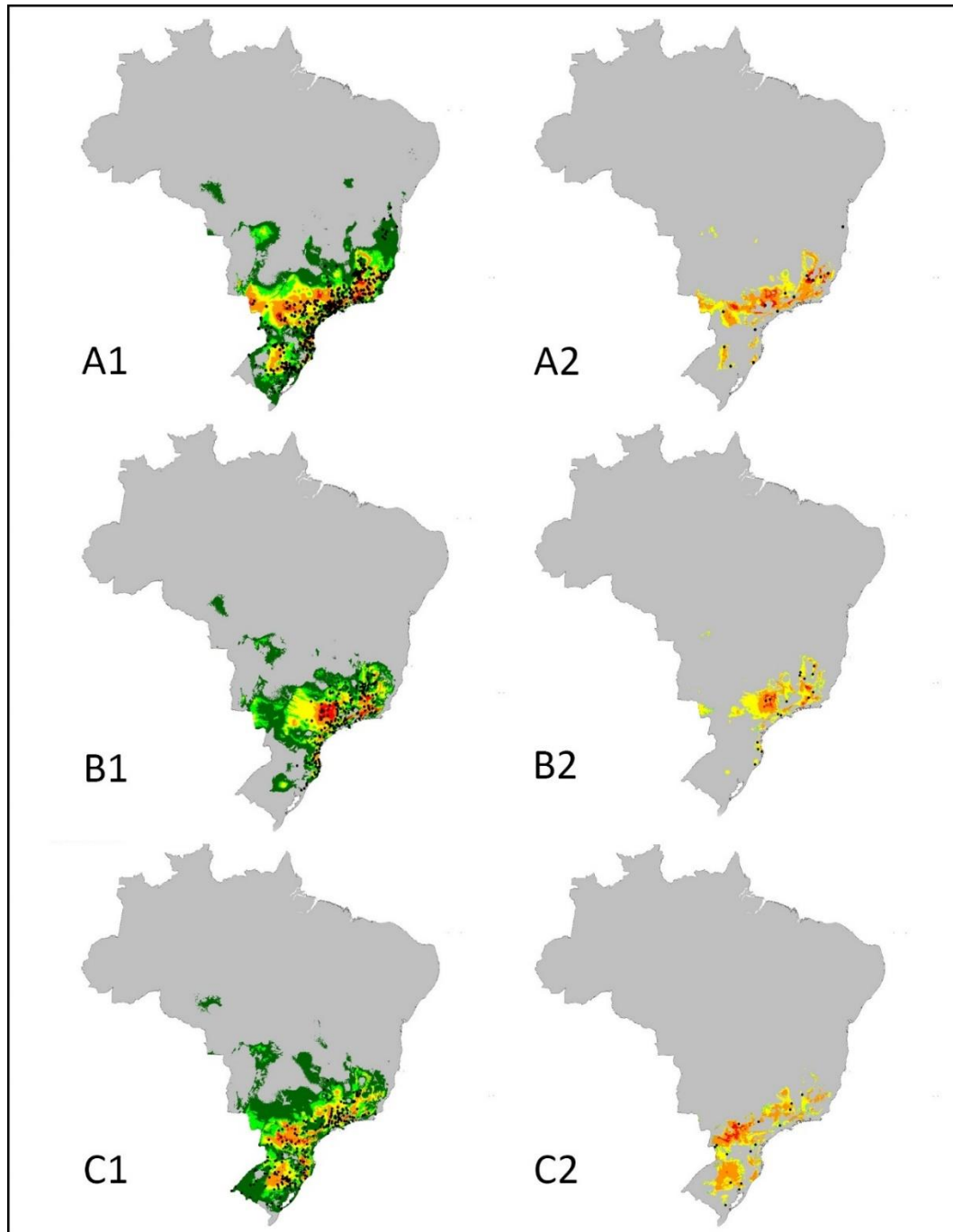


FIG. 2. Comparação entre o modelo de distribuição original e o modelo criado a partir de 20 pontos amostrais para as espécies *Chiroxiphia caudata* (A1 e A2 – 600 e 20 pontos, respectivamente), *Hemithraupis ruficapilla* (B1 e B2 – 250 e 20 pontos) e *Myiothlypis leucoblephara* (C1 e C2 – 300 e 20 pontos).

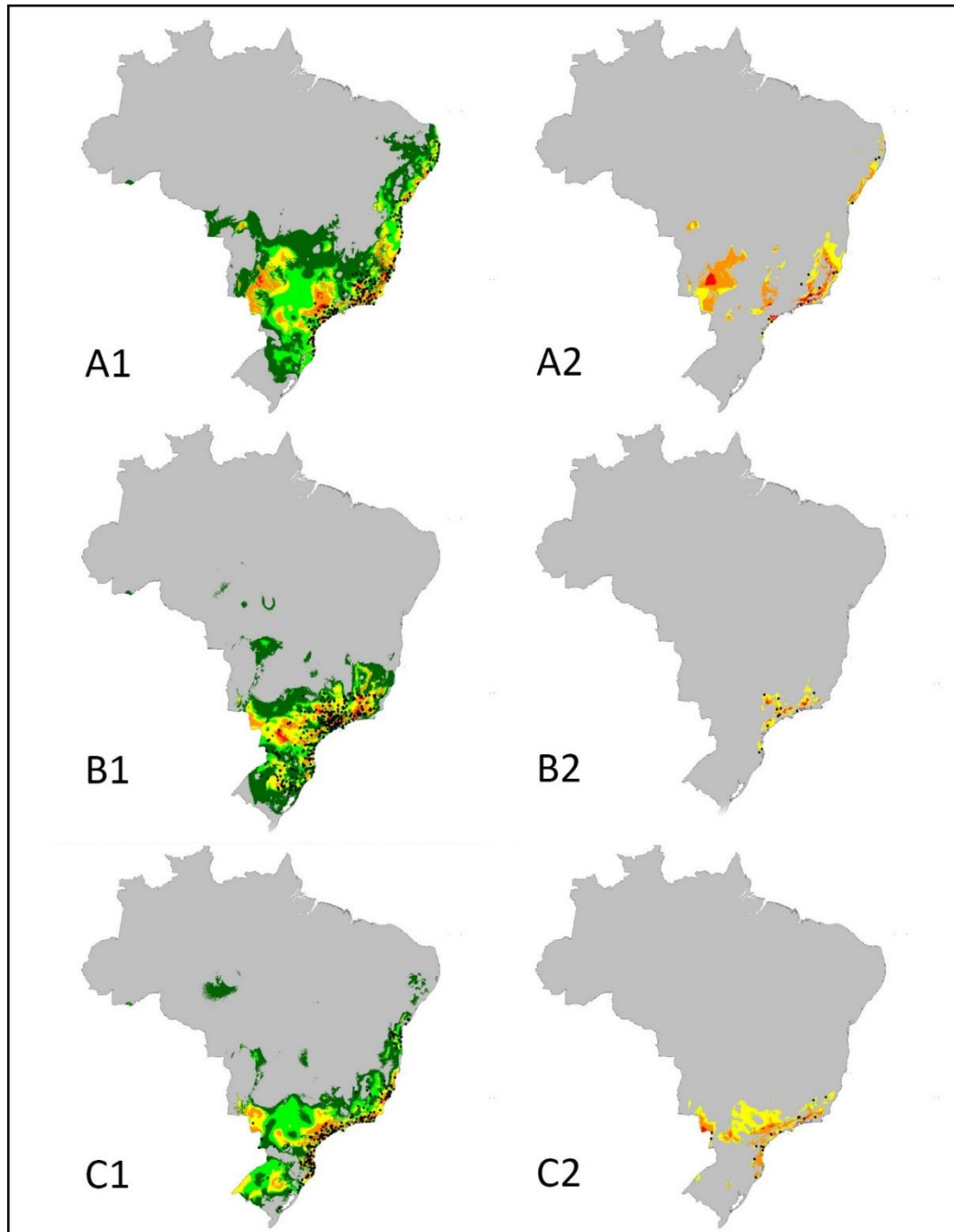


FIG. 3. Comparação entre o modelo de distribuição original e o modelo criado a partir de 20 pontos amostrais para as espécies *Ramphocelus bresilius* (A1 e A2 – 300 e 20 pontos), *Synalaxis ruficapilla* (B1 e B2 – 400 e 20 pontos) e *Tangara seledon* (C1 e C2 - 300 e 20 pontos).

Varição da área modelada. Conforme a Figura 4, a divisão do tamanho amostral original de cada espécie em $N_2 = 150$, $N_3 = 50$ e $N_4 = 20$ produz modelos com variabilidades acentuadas de área a partir de $N_2 = 150$, quando comparadas à área original de cada espécie. A maior variabilidade de área modelada é encontrada no tamanho amostral $N_3 = 50$, se mantendo quase constante conforme a redução amostral se aproxima de $N_4 = 20$. Os valores obtidos na modelagem utilizando o menor tamanho amostral, além de apresentarem a menor variabilidade quando comparados entre si, também apresentam a maior redução de área modelada em comparação ao tamanho amostral original, o que, para o algoritmo em estudo, pode indicar um padrão de modelos subestimados para a distribuição potencial de espécies.

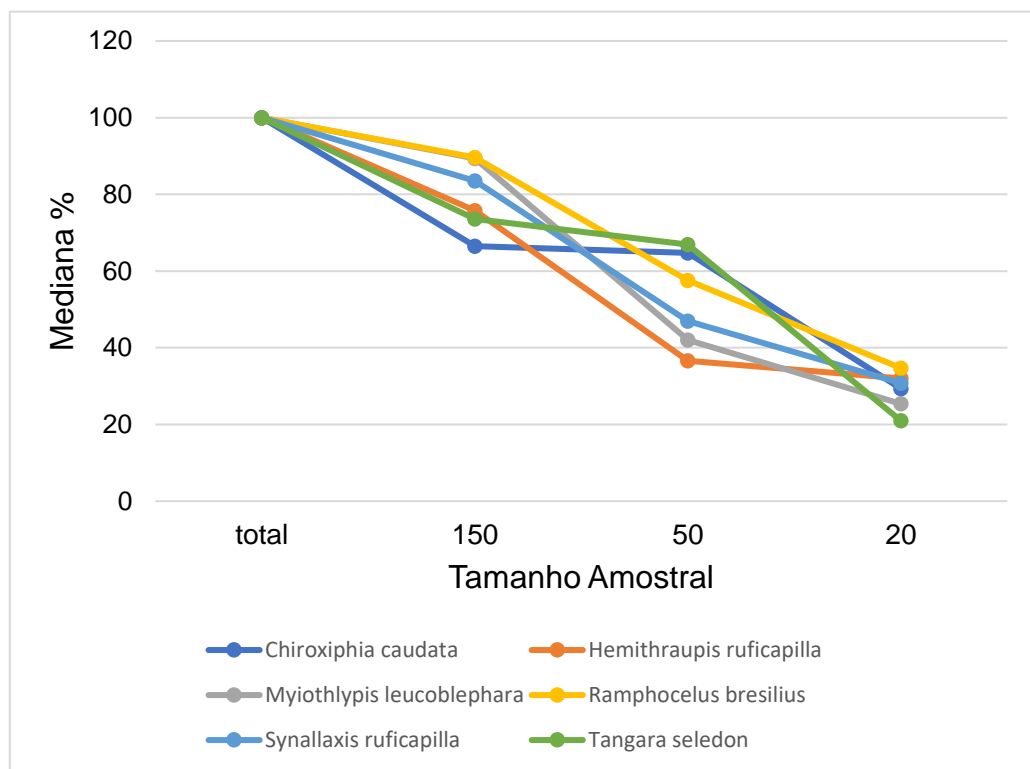


FIG. 4. Percentual da área modelada.

O mesmo resultado foi encontrado durante a segunda etapa do trabalho (Figura 5), onde para a mesma espécie foram utilizadas 10 amostras aleatórias obtidas a partir do tamanho

original N_T , demonstrando que a **redução** do tamanho das amostras e amostras menores tendem a **reduzir** a área modelada observada, apresentando perdas de até 80% das informações contidas nos modelos criados a partir dos tamanhos amostrais N_T .

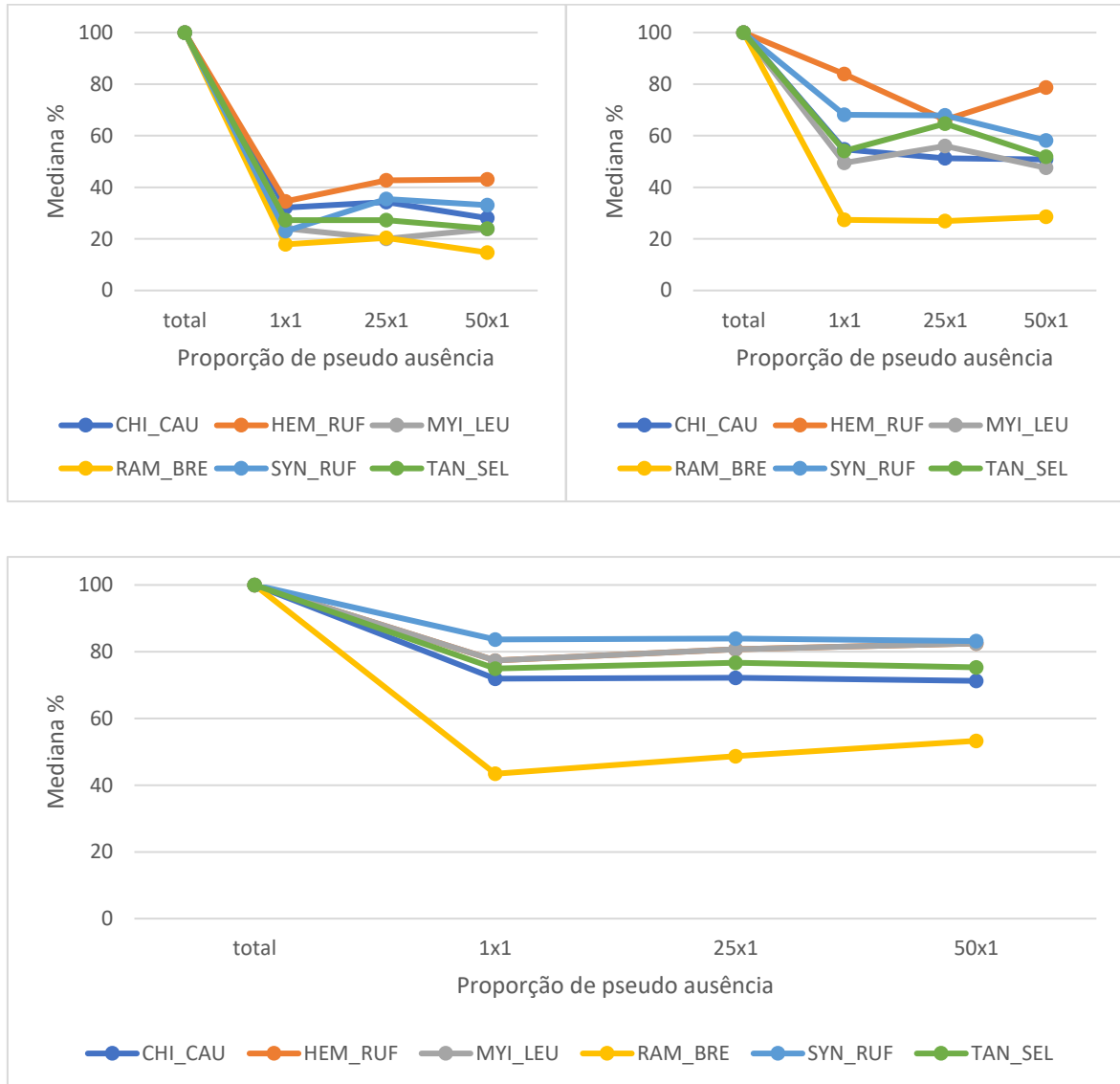


FIG. 5. Percentuais da área modelada em relação à variação da proporção de pseudo ausência por tamanho amostral: 20 pontos (no alto, à esq.), 50 pontos (no alto, à dir.) e 150 pontos (embaixo)

Valores de AUC. A obtenção dos valores de AUC para cada modelo nas duas etapas do trabalho demonstram que existe pouca variabilidade entre os três primeiros tamanhos amostrais

(N_T , N_2 e N_3) quando feitas as comparações intra e inter-espécies, mesmo após a reamostragem de 10 conjuntos amostrais para cada espécie na segunda etapa do trabalho (Tabela 4).

Em contrapartida, tamanhos amostrais menores (e.g. $N \leq 20$) apresentam as maiores variabilidades em comparação à tamanhos amostrais maiores. Os valores de AUC obtidos em modelos gerados a partir de amostras menores tendem a ser insatisfatórios, apresentando pouca confiabilidade nas áreas de distribuição potencial das espécies preditas pelos modelos (Tabela 3).

Tabela 3. Valores de média e desvio padrão dos resultados de AUC.

	AMOSTRA TOTAL		150		50		20	
	média	Desvio	média	Desvio	média	Desvio	média	Desvio
CHI_CAU	0,93	0,01	0,89	0,02	0,83	0,06	0,69	0,11
HEM_RUF	0,92	0,02	0,90	0,04	0,81	0,05	0,77	0,10
MYI_LEU	0,91	0,01	0,90	0,03	0,82	0,07	0,62	0,13
RAM_BRE	0,90	0,02	0,88	0,03	0,82	0,09	0,77	0,11
SYN_RUF	0,94	0,01	0,90	0,03	0,81	0,07	0,74	0,15
TAN_SEL	0,93	0,01	0,92	0,04	0,89	0,08	0,68	0,09

CHI_CAU *Chiroxiphia caudata*; **HEM_RUF** *Hemithraupis ruficapilla*; **MYI_LEU** *Myiothlypis leucoblephara*; **RAM_BRE** *Ramphocelus bresilius*; **SYN_RUF** *Synallaxis ruficapilla*; **TAN_SEL** *Tangara seledon*

Tabela 4. Valores médios de AUC obtidos após reamostragem

	20			50			150		
	1x1	25x1	50x1	1x1	25x1	50x1	1x1	25x1	50x1
CHI_CAU	0,712	0,722	0,772	0,875	0,809	0,777	0,896	0,891	0,845
HEM_RUF	0,598	0,698	0,685	0,867	0,786	0,847	0,897	0,876	0,827
MYI_LEU	0,690	0,598	0,666	0,854	0,852	0,786	0,888	0,885	0,823
RAM_BRE	0,667	0,585	0,670	0,743	0,738	0,755	0,791	0,789	0,843
SYN_RUF	0,594	0,766	0,778	0,894	0,879	0,808	0,902	0,911	0,844
TAN_SEL	0,733	0,724	0,699	0,815	0,842	0,809	0,754	0,906	0,865

É importante ressaltar que, embora tamanhos amostrais como N_3 ($n=50$) apresentem valores de AUC considerados satisfatórios ($AUC \geq 0,80$), quando feitas as comparações entre esses valores e os tamanhos de área modelada, torna-se evidente que o AUC pode mascarar a perda

de informações dos modelos, o que pode levar à conclusões equivocadas quanto a confiabilidade dos modelos.

DISCUSSÃO

Embora seja uma área em ascensão dentro da Ecologia, a modelagem da distribuição de espécies enfrenta algumas dificuldades no seu desenvolvimento. Um problema recorrente é o fato de que muitos dos dados utilizados para modelar espécies podem compor tamanhos amostrais insuficientes ou ainda serem resultado de coletas de dados temporal e espacialmente tendenciosas (Hernandez et al., 2006).

Os trabalhos de Wisz et al. (2008) e Bean, Stafford e Brashares (2012) discutem que, embora alguns estudos tenham demonstrado que os modelos criados a partir de pequenas amostras podem ser precisos e acurados, muitos dos testes realizados com pequenas amostras assumem o conhecimento pleno da distribuição completa de uma espécie para calcular estimativas de precisão.

Os resultados apresentados nas Figuras 2 e 3 indicam que a área modelada é reduzida quando comparamos as amostras originais em relação às menores amostras, e as maiores distribuições potenciais, indicadas pelas cores quentes (vermelho, laranja e amarelo), são limitadas aos pontos amostrais na maior parte das espécies e, portanto, a alteração da distribuição potencial da espécie evidencia a perda de informações importantes do modelo com a variação do tamanho amostral. Resultados semelhantes são encontradas na literatura, como visto em Araujo et al. (2005), Guisan & Thuiller (2005), Elith et al. (2006), Hernandez et al. (2006) e Wisz et al. (2008).

A Figura 4 apresenta as medianas para as seis espécies modeladas neste trabalho em contraste com os quatro tamanhos amostrais. É possível notar que, conforme o tamanho amostral

se reduz (e.g. N_T para N_2 , N_2 para N_3), os valores de mediana se tornam mais variáveis quando feitas as comparações entre espécies. Ao atingir o menor tamanho amostral (N_4), os modelos produzem menores valores de área e tendem a exibir menor variabilidade quando comparado entre as espécies, restringindo as áreas potenciais para a espécie, o que corrobora o padrão encontrado nas modelagens com a variação do tamanho amostral.

A obtenção dos valores de AUC a partir da curva ROC apresentados na Tabela 3 permitiram avaliar a interferência do tamanho amostral na performance do modelo através da medição de sua precisão e acurácia. Analisando as informações contidas na referida tabela, é possível notar que existem pequenas variabilidades nos valores de ROC quando comparados modelos de espécies diferentes com tamanhos amostrais iguais (N_T , N_2 e N_3). Por outro lado, um tamanho amostral reduzido (N_4) pode produzir valores de AUC abaixo de 0,7, podendo estes ser considerados insatisfatórios (Margotto, 2010), dados os critérios do pesquisador, para atestar a confiabilidade de um modelo e maior variabilidade quando comparados entre as diferentes espécies modeladas.

Para Araujo et al. (2005) o grau de acurácia de um modelo de distribuição de espécies deve ser levado em conta de acordo com a finalidade do modelo. O autor ainda discute que diferentes predições podem ser obtidas dependendo da região, tamanho da amostra e do algoritmo utilizado nas modelagens. Hernandez et al. (2006) acrescenta que os modelos baseados em conjuntos de dados reduzidos dependem de informações que muitas vezes não são comuns a todas as espécies, como distribuição geográfica e tolerância ambiental restritas.

Os resultados de AUC encontrados no presente trabalho permitiram observar que, embora seja uma medida padrão de validação de um modelo preditivo, os valores de AUC podem ser influenciados por fatores como tamanho amostral e pela quantidade de pontos de ausência

e/ou pseudo ausência utilizados na modelagem. Essa constatação surge após a comparação entre modelos com perdas de áreas variando de 40% a até 80% da área original total e seus respectivos valores de AUC que, muitas das vezes, poderiam ser classificados como aceitáveis devido serem iguais ou superiores a 0,80.

A utilização dos valores de AUC para validar os modelos produzidos neste trabalho mostrou-se como uma medida dúbia e influenciável ao permitir a validação de um modelo produzido a partir de um tamanho amostral reduzido como confiável na ocasião em que este perdeu 80% das informações originais contidas em um modelo criado a partir de um tamanho amostral maior.

Lobo et al. (2008) retrata 5 razões pelas quais o valor de AUC não deveria ser utilizado como medida de validação e, destas, a quinta razão corrobora os resultados verificados neste trabalho: a extensão total da área modelada influencia muito a taxa de ausências bem previstas e os escores da AUC.

Além da influência do algoritmo e do tamanho amostral na diminuição da área final modelada, os modelos gerados a partir do tamanho amostral N_4 ($N = 20$) também sofreram influência das características ecológicas das espécies modeladas, conforme discutido por Araujo e Williams (2000), Stockwell e Peterson (2002), Thuiller 2003 e Segurado e Araujo 2004.

Características ecológicas das espécies modeladas tendem a afetar o potencial de precisão do modelo devido ao fato de que espécies em espaços geográficos e ambientais generalizados geralmente são mais difíceis de modelar do que as espécies com distribuições espaciais compactas. Desta forma pode-se observar que as espécies *Chiroxiphia caudata*, *Hemithraupis ruficapilla*, *Myiothlypis leucoblephara* e *Synallaxis ruficapilla*, por apresentarem um padrão de dispersão maior permitem ao algoritmo uma maior abrangência de envelopes climáticos e,

portanto, resultam em áreas de predição melhor distribuídas em conformidade com os pontos de amostragem de cada espécie mesmo quando utilizados tamanhos amostrais reduzidos (Figs 2 e 3).

Em contrapartida, as espécies *Ramphocelus bresilius* e *Tangara seledon* apresentam modelos mais restritos quando comparados aos modelos de espécies amplamente distribuídas. Isso pode ser explicado por características espécie-específicas como endemismo (*Ramphocelus bresilius*) e limitação da distribuição da espécie atrelada à características ambientais (*Tangara seledon*).

Muito embora seja evidente a influência do tamanho amostral e das características ecológicas da espécie na performance do modelo, é possível notar que o modelo apresenta, ainda, pequenas manchas de áreas potenciais muito distantes dos pontos de observação e da área central modelada (Figs 2 e 3). Essas áreas podem ser irrelevantes para o modelo devido possuírem um tamanho muito reduzido e estarem muito distantes da área onde a espécie ocorre.

Conforme Stockwell e Peterson (2002), a adaptação ecológica local por subpopulações é mais provável resposta do por que características ecológicas também influenciam a modelagem. Em espécies amplamente distribuídas, diferentes preferências de habitat em partes discretas do alcance da espécie surgiriam e isso afetaria o modelo como um todo. Na modelagem climática, cada subpopulação teria uma faixa climática distinta em que ocorre e, portanto, quando a espécie é modelada como um todo em toda a sua área geográfica, a faixa climática total engloba condições climáticas não adequadas para a ocupação, superestimando assim a amplitude climática das espécies.

CONCLUSÃO

O presente trabalho, ao testar a modelagem de seis espécies brasileiras com diferentes tamanhos amostrais através do algoritmo BIOCLIM obteve resultados e conclusões semelhantes aos de Cumming (2000), Pearce & Ferrier (2000), Stockwell & Peterson (2002), Kadmon *et al.* (2003), Reese *et al.* (2005), Hernandez *et al.* (2006) e Wisz *et al.* (2008) atestando que o pesquisador deve atentar para a qualidade dos dados com que trabalha, especialmente no que tange às características ecológicas da espécie, a finalidade do modelo, o tamanho da área modelada, o tamanho amostral e a origem dos dados utilizados no processo de modelagem. Deve-se ter cuidado com o uso mecanicista dos diversos *softwares* de modelagem.

Os valores de AUC devem ser analisados com cautela, uma vez que os dados obtidos no presente trabalho se mostraram passíveis de influência não só do tamanho amostral, mas do total de pontos de pseudo ausência e do tamanho da área modelada, corroborado pelas críticas e análises feitas por Lobo *et al.* (2008).

A alteração nas distribuições potenciais e a diminuição dos valores de área obtidos com a redução dos tamanhos amostrais utilizados no presente trabalho demonstram que os modelos finais gerados a partir de tamanhos amostrais menores talvez não poderiam ser utilizados para algumas finalidades (e.g. estudos de conservação) devido à baixa qualidade exibida pelos mesmos, em conformidade com a afirmação de Araujo *et al.* (2005) de que a acurácia da modelagem dependerá de sua finalidade.

Trabalhos como os de Guisan & Thuiller (2005) e Elith *et al.* (2006) demonstram que, além de considerar o tamanho amostral a se utilizar, o tipo de algoritmo utilizado terá impacto significativo na modelagem das espécies, haja vista que a resposta do algoritmo diante de um número reduzido de informações irá depender da abordagem tomada pelo mesmo.

Diante dos resultados obtidos no presente trabalho, deve-se evitar o uso de amostras menores do que 50 pontos de presença em modelagens com o algoritmo BIOCLIM devido à influência do conjunto amostral sobre o processo de modelagem e na qualidade do modelo final.

Se espera que as discussões e conclusões obtidas no presente artigo venham a instigar outros trabalhos de modelagem preditiva de distribuição de espécies e lançar sobre estes as reflexões quanto à influência da qualidade dos dados utilizados no processo final de modelagem e na tomada de decisões embasadas em modelos de distribuição.

REFERÊNCIAS

Álvares, G. F. R. (2009). Taxonomia, Distribuição Geográfica Potencial e Conservação das Espécies de *Phyllomedusa* do Grupo *hypochondrialis*. Dissertação de Mestrado em Engenharia Florestal, Publicação PGEFL – 116/09, Departamento de Engenharia Florestal, Universidade de Brasília, DF, 95p.

Anderson RP, Lew D, Peterson AT (2003) Evaluating predictive models of species' distributions: criteria for selecting optimal models. *Ecol Model* 162:211–232

Araújo, M. B., Pearson, R. G., Thuiller, W., & Erhard, M. (2005). Validation of species–climate impact models under climate change. *Global Change Biology*, 11, 1504–1513.

Araújo, M.B. & Williams, P.H. (2000) Selecting areas for species persistence using occurrence data. *Biological Conservation*, 96, 331–345.

Austin, M. Species distribution models and ecological theory: A critical assessment and some possible new approaches. *Ecological Modelling*, v. 200, n. 1-2, p. 1-19, jan. 2007.

Bean WT, Stafford R, Brashares JS. 2012. The effects of small sample size and sample bias on threshold selection and accuracy assessment of species distribution models. *Ecography* 35: 250-258.

Casemiro, F. A. S., Gouveia, S. F. & Diniz-Filho, J. A. F. 2012. Distribuição de *Rhinella granulosa*: integrando envelopes bioclimáticos e respostas ecofisiológicas. *Revista da Biologia*, 8: 38- 44.

Cumming, G.S. (2000) Using between-model comparisons to fine-tune linear models of species ranges. *Journal of Biogeography*, 27, 441–455.

Dennis, R., and C. Thomas. 2000. Bias in butterfly distribution maps: the influence of hot spots and recorder's home range. *Journal of Insect Conservation* 4:73–77.

Dettmers, R. e Bart, J. A GIS modeling method applied to predicting forest songbird habitat. *Ecological Applications*, 9, 152–163, 1999.

Elith, J., Graham, C.H., Anderson, R.P., Dudik, M., Ferrier, S., Guisan, A., Hijmans, R.J., Huettmann, F., Leathwick, J.R., Lehmann, A., Li, J., Lohmann, L.G., Loiselle, B.A., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J.M., Peterson, A.T., Phillips, S.J., Richardson, K., Scachetti-Pereira, R., Schapire, R.E., Soberon, J., Williams, S., Wisz, M.S., Zimmermann, N.E., 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 29, 129–151.

Engler, R., Guisan, A., & Rechsteiner, L. (2004). An improved approach for predicting the distribution of rare and endangered species from occurrence and pseudo-absence data. *Journal of Applied Ecology*, 41, 263–274.

Giannini, T. C. et al. Desafios atuais da modelagem preditiva de distribuição de espécies. *Rodriguésia*, Rio de Janeiro, v. 63, n. 3, p. 733-749, set. 2011.

GIOVANELLI, J.G.R.; HADDAD, C.F.B.; ALEXANDRINO, J. Predicting the potential distribution of the alien invasive American bullfrog (*Lithobates catesbeianus*) in Brazil. *Biological Invasions*, v.10, p. 585–590, 2008.

GRAHAM, C.H.; FERRIER, S.; HUETTMAN, F.; MORITZ, C.; PETERSON, A.T. New developments in museum-based informatics and applications in biodiversity analysis. *Trends in Ecology and Evolution*, v. 19, p. 497-503, 2004.

Gregorini MZ, Rodolfo AM, Cândido FJr, Tôrres NM 2007. Modelagem de distribuição geográfica do tamanduá-bandeira (*Myrmecophaga tridactyla*) e sua ocorrência em unidades de conservação no estado do Paraná. In: VIII Congresso de Ecologia do Brasil, Caxambu – MG.

Guisan, A. & Thuiller, W. (2005). Predicting species distribution: offering more than simple habitat models. *Ecol. Lett.*, **8**, 993–1009.

Guisan, A., Zimmermann, N.E. Predictive habitat distribution models in ecology. *Ecol. Model.* 135, 147–186, 2000.

Hastie, T., Tibshirani, R. *Generalised Additive Models*. Chapman and Hall, London, 1990.

Hernandez, P.A., Graham, C.H., Master, L.L., Albert, D.L., 2006. The effect of sample size and species characteristics on performance of different species distribution modeling methods. *Ecography* 29, 773–785.

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. *International Journal of Climatology* 25: 1965-1978

Hijmans, R.J., L. Guarinon, P. Mathur. 2012. *DIVA-GIS Version 7.5 Manual*. 77 p.

Kadmon, R., Farber, O. & Danin, A. (2003) A systematic analysis of factors affecting the performance of climatic envelope models. *Ecological Applications*, 13, 853–867.

Kunin, W.E. & Gaston, K.J. (1993) The biology of rarity – Patterns, causes and consequences. *Trends in Ecology and Evolution*, 8, 298–301.

Margotto, P. R. (2010). Curva ROC como fazer e interpretar no SPSS. Disponível em: http://www.paulomargotto.com.br/documentos/Curva_ROC_SPSS.pdf

McCullagh, P., Nelder, J.A. *Generalized Linear Models*, second ed. Chapman and Hall, London, 1989.

Nabout, João & Terribile, Levi & Bini, Luis & Diniz-Filho, Jose Alexandre. (2009). Phylogenetic autocorrelation and heritability of geographic range size, shape and position of fiddler crabs, genus *Uca* (Crustacea, Decapoda). *Journal of Zoological Systematics and Evolutionary Research*. 48. 102 - 108. 10.1111/j.1439-0469.2009.00531.x.

Pearce, J. and Ferrier, S. Evaluating the predictive performance of habitat models developed using logistic regression. - *Ecol. Model.* 133: 225-245, 2000.

Phillips, S. J. et al. Sample selection bias and presence-only distribution models: implications for background and pseudoabsence data. *Ecol. Appl.* 19: 1811-1819, 2009.

Reddy, S., and L. M. Da'valos. 2003. Geographical sampling bias and its implications for conservation priorities in Africa. *Journal of Biogeography* 30:1719–1727.

Reese, G.C., Wilson, K.R., Hoeting, J.A. & Flather, C.H. (2005) Factors affecting species distribution predictions: a simulation modeling experiment. *Ecological Applications*, 15, 554– 564.

Rushton, S.P., Ormerod, S.J., Kerby, G. New paradigms for modelling species distributions? *J. Appl. Ecol.* 41, 193–200, 2004.

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

Segurado, P & Araújo, Miguel. (2004). An evaluation of methods for modelling species distributions. *Journal of Biogeography*. 31. 1555-1568. 10.1111/j.1365-2699.2004.01076.x. SI-QUEIRA, Marinez Ferreira de; DURIGAN, Giselda. Modelagem da distribuição geográfica de espécies lenhosas de cerrado no Estado de São Paulo. *Rev. bras. Bot.*, São Paulo , v. 30, n. 2, p. 233-243, June 2007.

STEIN, Mariana Garcez. Biogeografia e conservação dos anfíbios da Caatinga. 2015. 60 f., il. Dissertação (Mestrado em Ecologia)—Universidade de Brasília, Brasília, 2015.

Stockwell, D.R., Peterson, A.T., 2002. Effects of sample size on accuracy of species distribution models. *Ecol. Model.* 148, 1–13.

Thuiller W (2003) BIOMOD: optimising predictions of species distributions and projecting potential future shifts under global change. *Global Change Biology*, 9, 1353–1362

Tsoar, A., Allouche, O., Steinitz, O., Rotem, D. and Kadmon, R. A comparative evaluation of presence-only methods for modelling species distribution. *Diversity and Distributions*, 13: 397–405, 2007. doi:10.1111/j.1472-464

WikiAves (2008) [*tangará (Chiroxiphia caudata)*]. WikiAves, a Enciclopédia das Aves do Brasil. Disponível em: <http://www.wikiaves.com.br/tangara>. Acesso em: 12/2/2016.

WikiAves (2008) [*sáira-ferrugem (Hemithraupis ruficapilla)*]. WikiAves, a Enciclopédia das Aves do Brasil. Disponível em: <http://www.wikiaves.com.br/saira-ferrugem>. Acesso em: 12/2/2016.

WikiAves (2008) [*pula-pula-assobiador (Myiothlypis leucoblephara)*]. WikiAves, a Enciclopédia das Aves do Brasil. Disponível em: <http://www.wikiaves.com.br/pula-pula-assobiador>. Acesso em: 12/2/2016.

WikiAves (2008) [*tiê-sangue (Ramphocelus bresilius)*]. WikiAves, a Enciclopédia das Aves do Brasil. Disponível em: <http://www.wikiaves.com.br/tie-sangue>. Acesso em: 12/2/2016.

WikiAves (2008) [*pichororé (Synallaxis ruficapilla)*]. WikiAves, a Enciclopédia das Aves do Brasil. Disponível em: <http://www.wikiaves.com.br/pichorore>. Acesso em: 12/2/2016.

WikiAves (2008) [*sáira-sete-cores (Tangara seledon)*]. WikiAves, a Enciclopédia das Aves do Brasil. Disponível em: <http://www.wikiaves.com.br/saira-sete-cores>. Acesso em: 12/2/2016. Wisz, M.S. & Guisan, A. (2009) Do pseudo-absence selection strategies influence species distribution models and their predictions? An information-theoretic approach based on simulated data. *BMC Ecology*, 9, 8.

Wisz, M.S., Hijmans, R.J., Li, J., Peterson, A.T., Graham, C.H., Guisan, A., 2008. Effects of sample size on the performance of species distribution models. *Divers. Distrib.* 14, 763–773.

Zaniewski, A.E., Lehmann, A. & Overton, J.M. (2002) Predicting species spatial distributions using presence-only data: a case study of native New Zealand ferns. *Ecological Modelling*, 157, 261–280.

CONCLUSÃO GERAL

A modelagem preditiva de distribuição de espécies se apresenta como importante ferramenta no estudo das espécies, permitindo não só o avanço do conhecimento ecológico, mas o delineamento de estudos e projetos voltados à conservação das espécies.

Entretanto, como toda ferramenta, os princípios de modelagem da distribuição de espécies podem se tornar mecânicos e gerar estudos que se resumem ao uso de um algoritmo e/ou software visando apenas o resultado final, sem uma compreensão satisfatória do processamento dos dados pelo algoritmo ou ainda, sem o conhecimento biológico necessário para inferir algo sobre a espécie estudada.

Como visto no presente trabalho, a escolha do algoritmo, características ecológicas das espécies e a qualidade da base de dados são informações que requerem total atenção por parte do pesquisador, haja vista que, um modelo preditivo tem por sua prerrogativa auxiliar não só na compreensão geográfica da espécie, mas em sua distribuição temporal.

Os processos de modelagem da distribuição de uma espécie, são, portanto, passos importantes para a Ecologia e para a Biologia da Conservação, necessitando de bases sólidas para que se produzam resultados com precisão suficiente para a tomada de decisões.

APÊNDICE A – Mapas de distribuição potencial das espécies *Chiroxiphia caudata*, *Hemithraupis ruficapilla*, *Myiothlypis leucoblephara*, *Ramphocelus bresilius*, *Synallaxis ruficapilla* e *Tangara seledon*

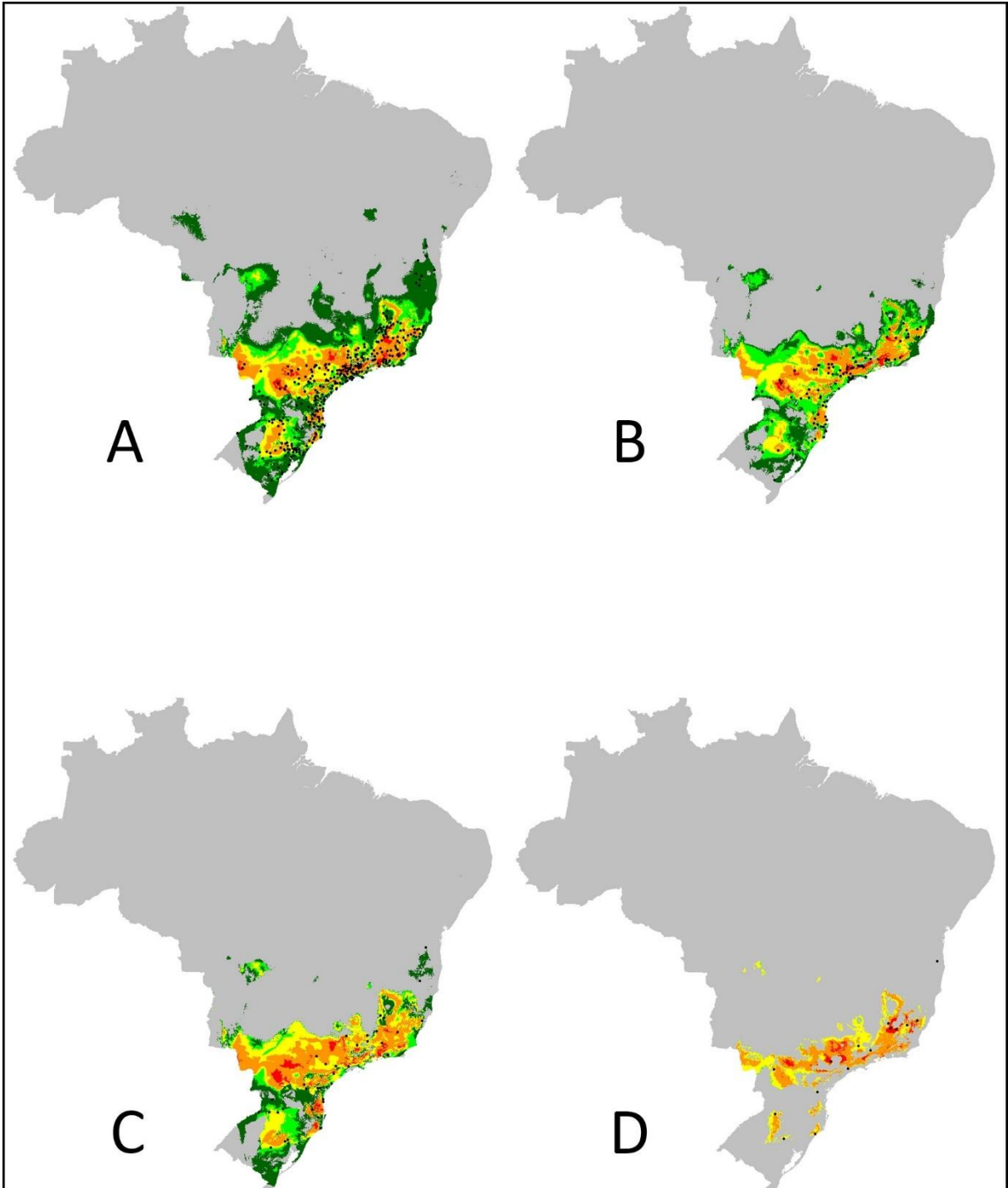


FIG 6. Modelos de distribuição potencial de *Chiroxiphia caudata* nos tamanhos amostrais A ($N_T = 600$), B ($N_2 = 150$), C ($N_3 = 50$) e D ($N_4 = 20$).

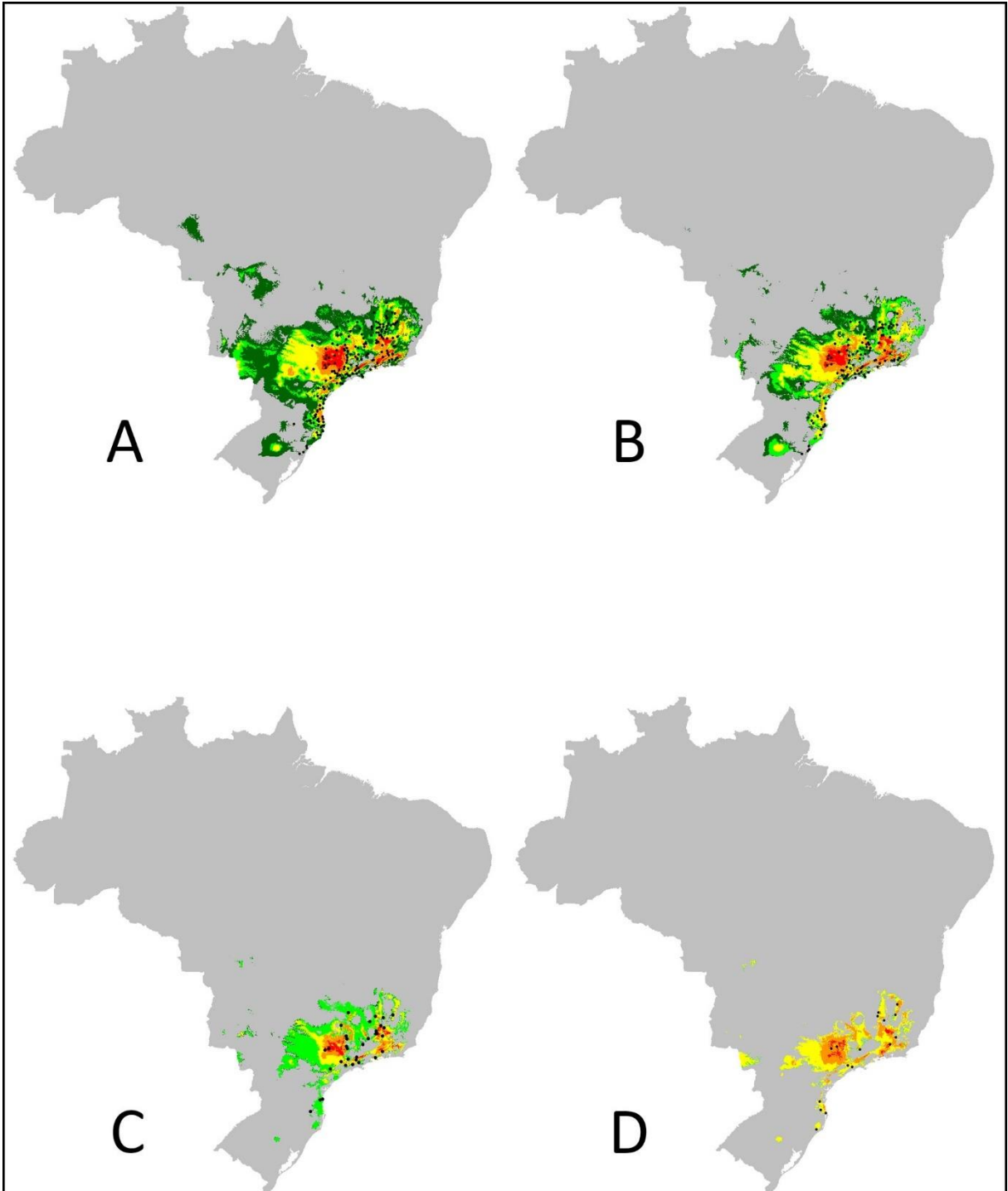


FIG 7. Modelos de distribuição potencial de *Hemithraupis ruficapilla* nos tamanhos amostrais A ($N_T = 250$), B ($N_2 = 150$), C ($N_3 = 50$) e D ($N_4 = 20$).

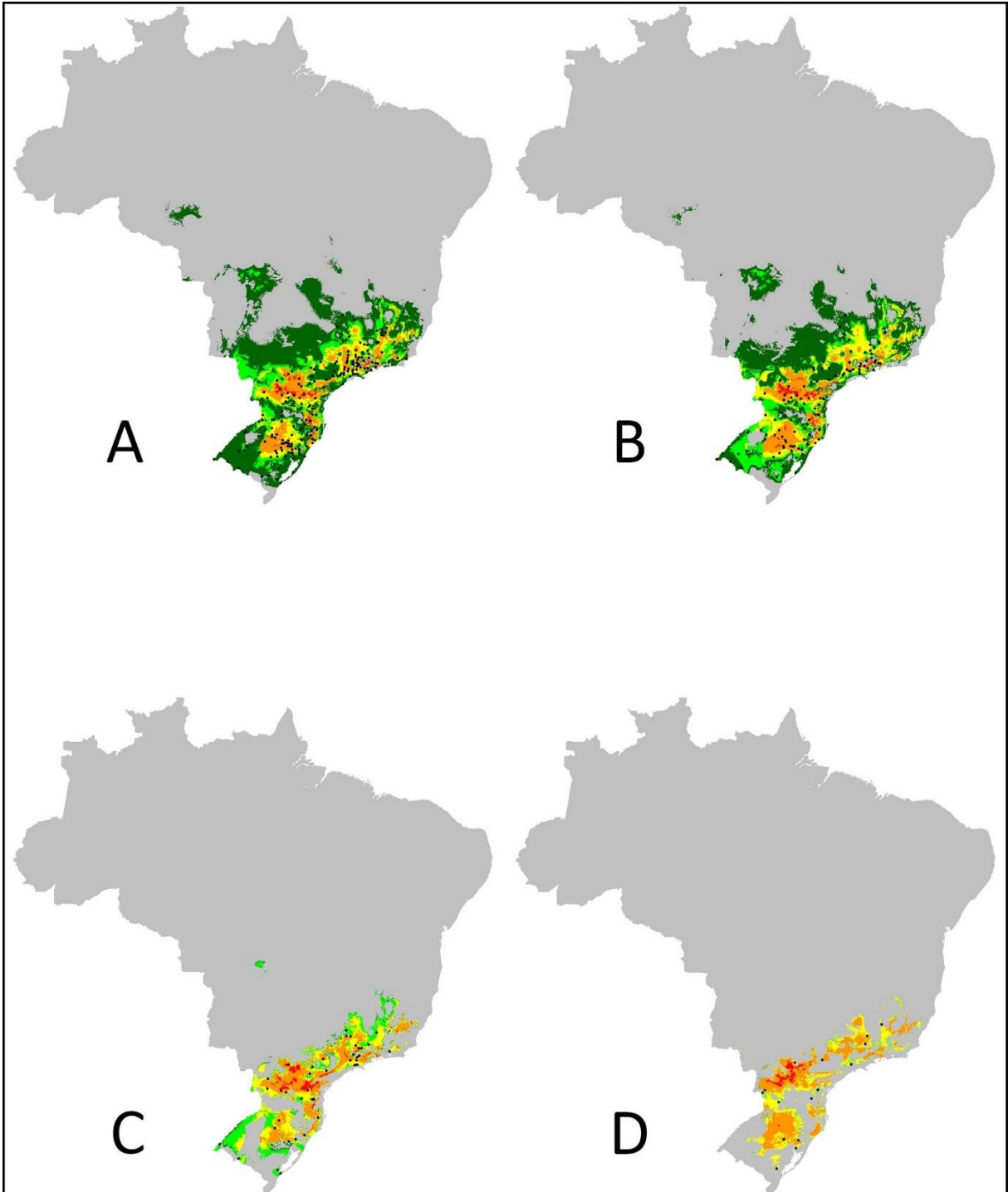


FIG 8. Modelos de distribuição potencial de *Myiothlypis leucoblephara* nos tamanhos amostrais A ($N_T = 300$), B ($N_2 = 150$), C ($N_3 = 50$) e D ($N_4 = 20$).

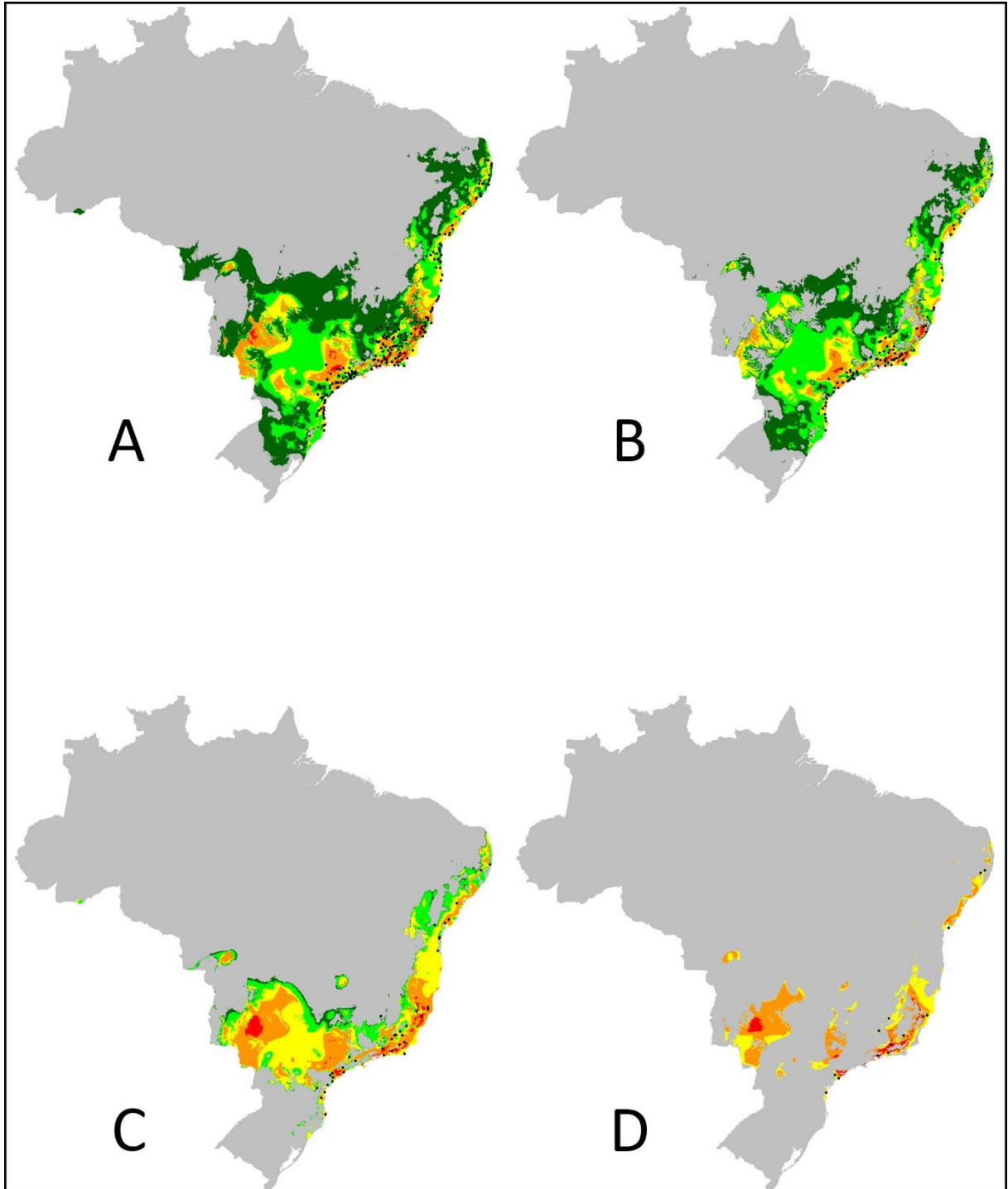


FIG 9. Modelos de distribuição potencial de *Ramphocelus bresilius* nos tamanhos amostrais A ($N_T = 300$), B ($N_2 = 150$), C ($N_3 = 50$) e D ($N_4 = 20$).

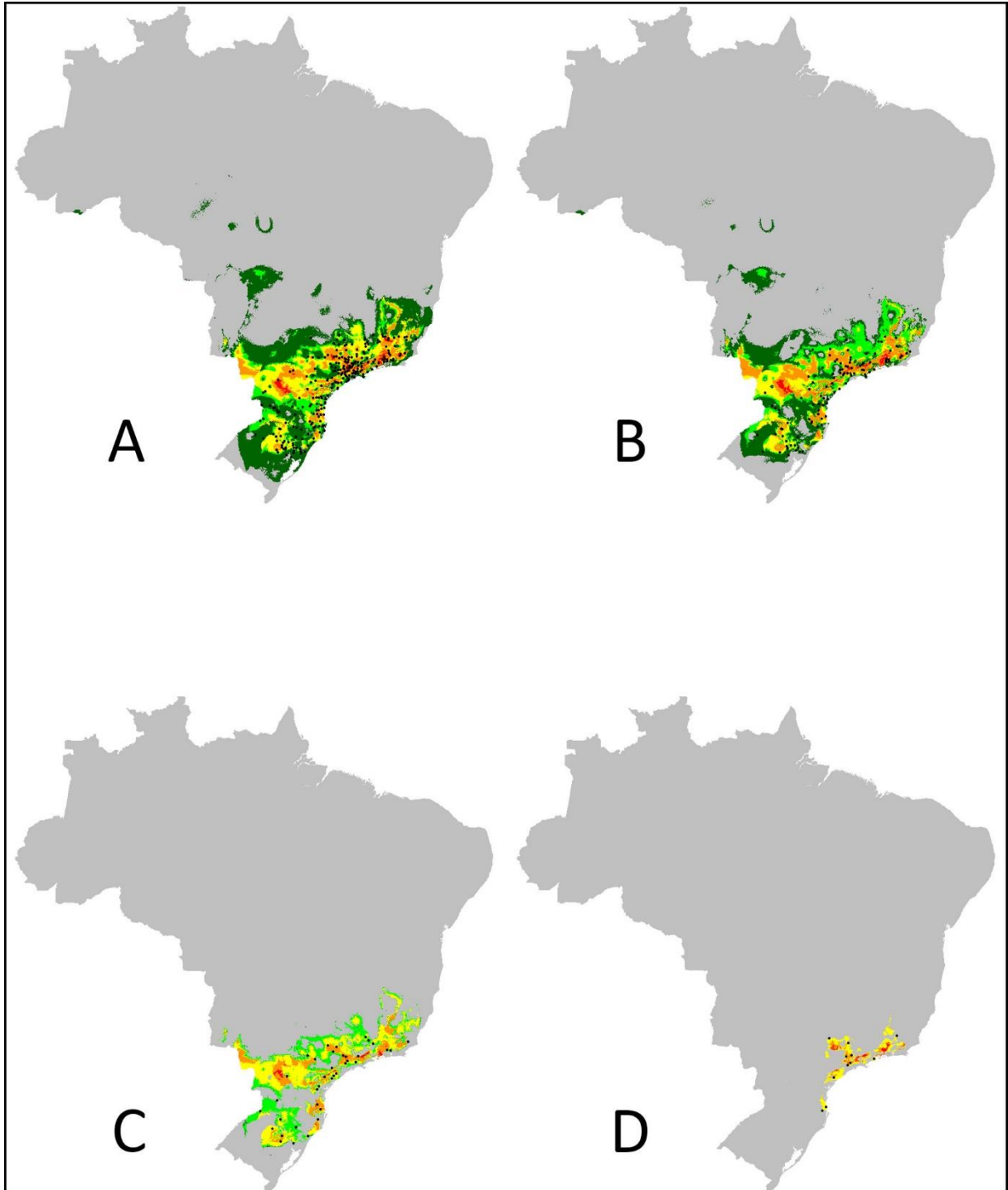


FIG 10. Modelos de distribuição potencial de *Synallaxis ruficapilla* nos tamanhos amostrais A ($N_T = 400$), B ($N_2 = 150$), C ($N_3 = 50$) e D ($N_4 = 20$).

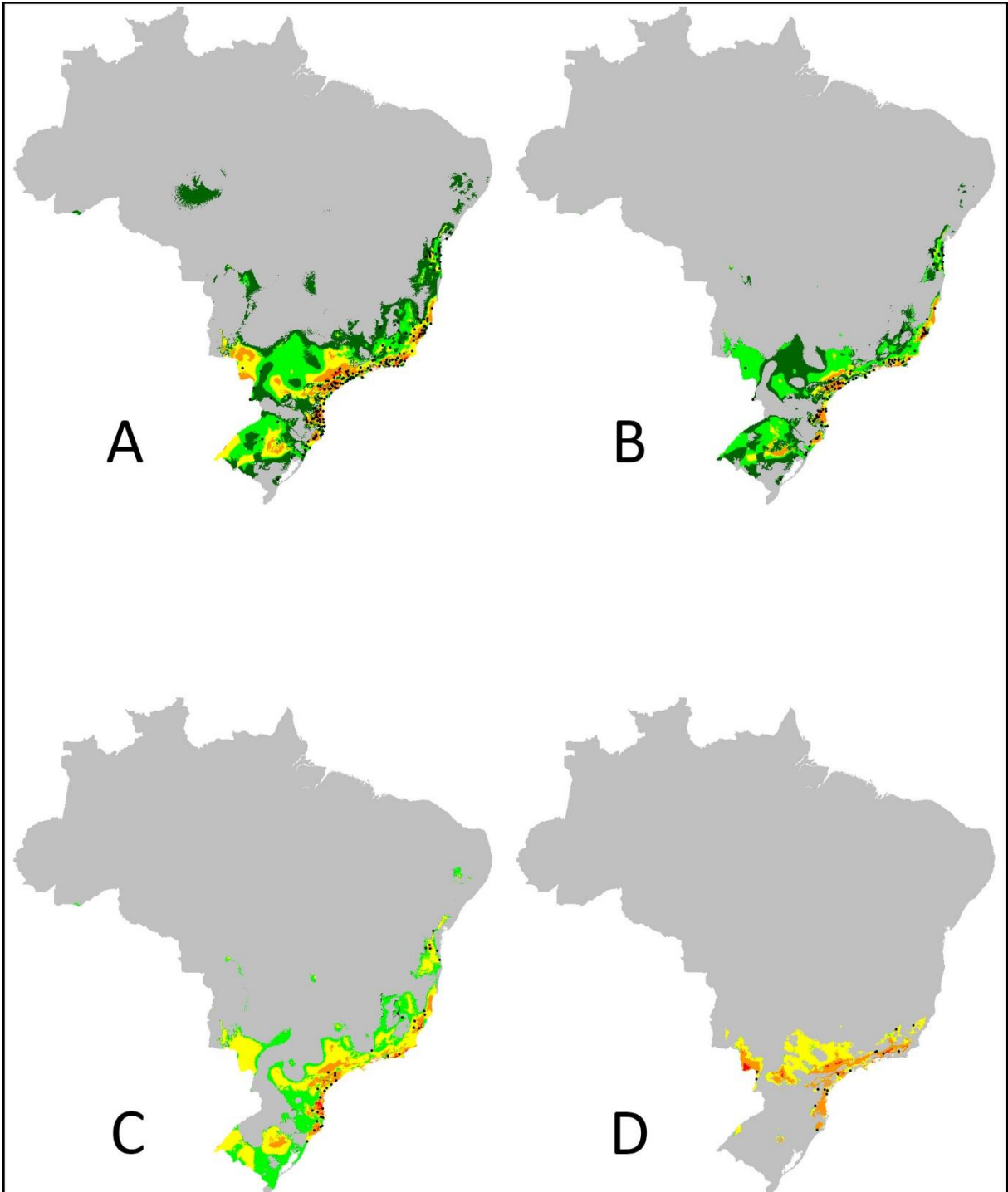


FIG 11. Modelos de distribuição potencial de *Tangara seledon* nos tamanhos amostrais A ($N_T = 300$), B ($N_2 = 150$), C ($N_3 = 50$) e D ($N_4 = 20$).

APÊNDICE B – Mapas de distribuição potencial das espécies *Chiroxiphia caudata*, *Hemithraupis ruficapilla*, *Myiothlypis leucoblephara*, *Ramphocelus bresilius*, *Synallaxis ruficapilla* e *Tangara seledon* após processo de reamostragem

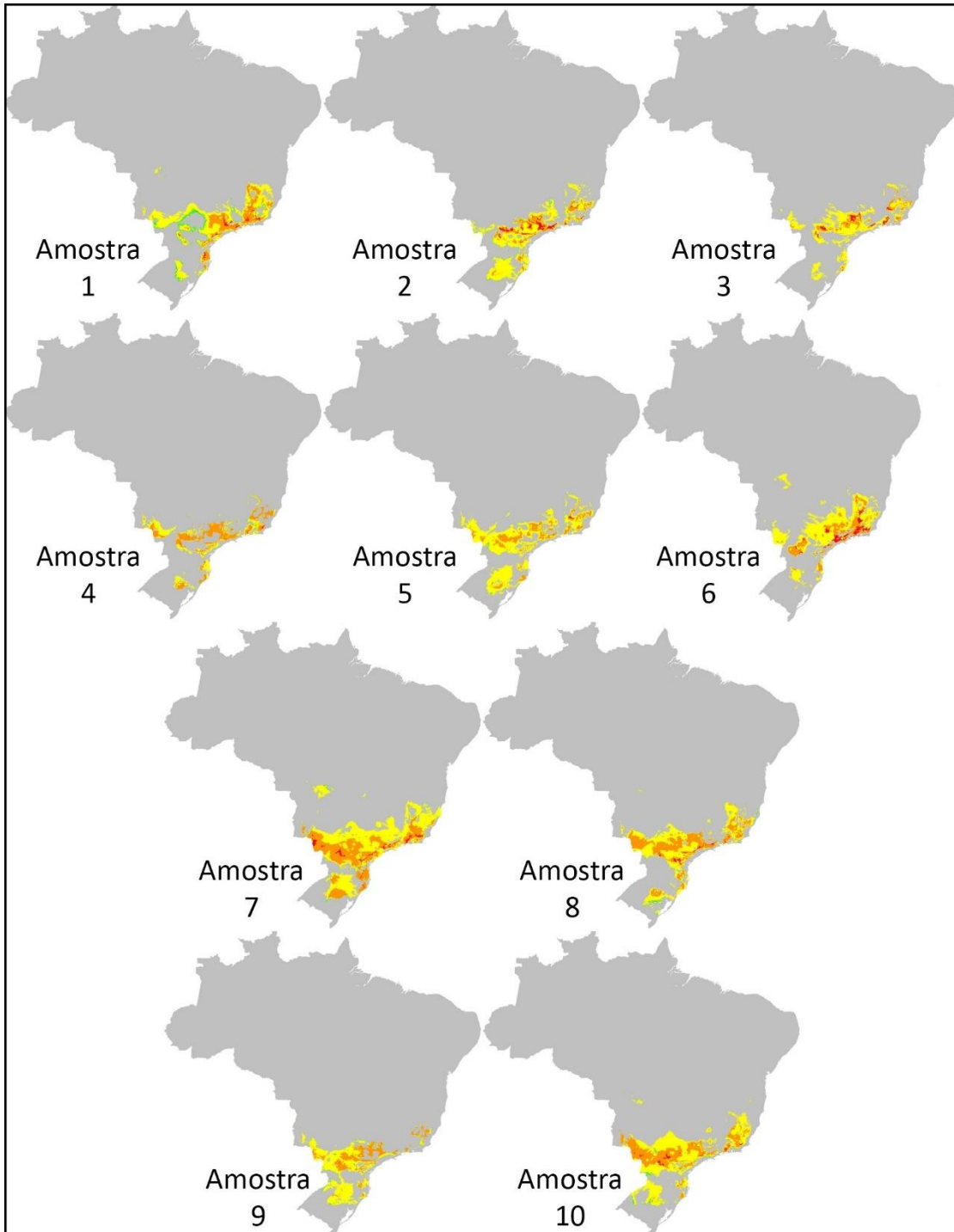


FIG 12. Modelagem da distribuição potencial de *Chiroxiphia caudata* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 20 pontos de presença, na proporção de pseudo ausência 1x1

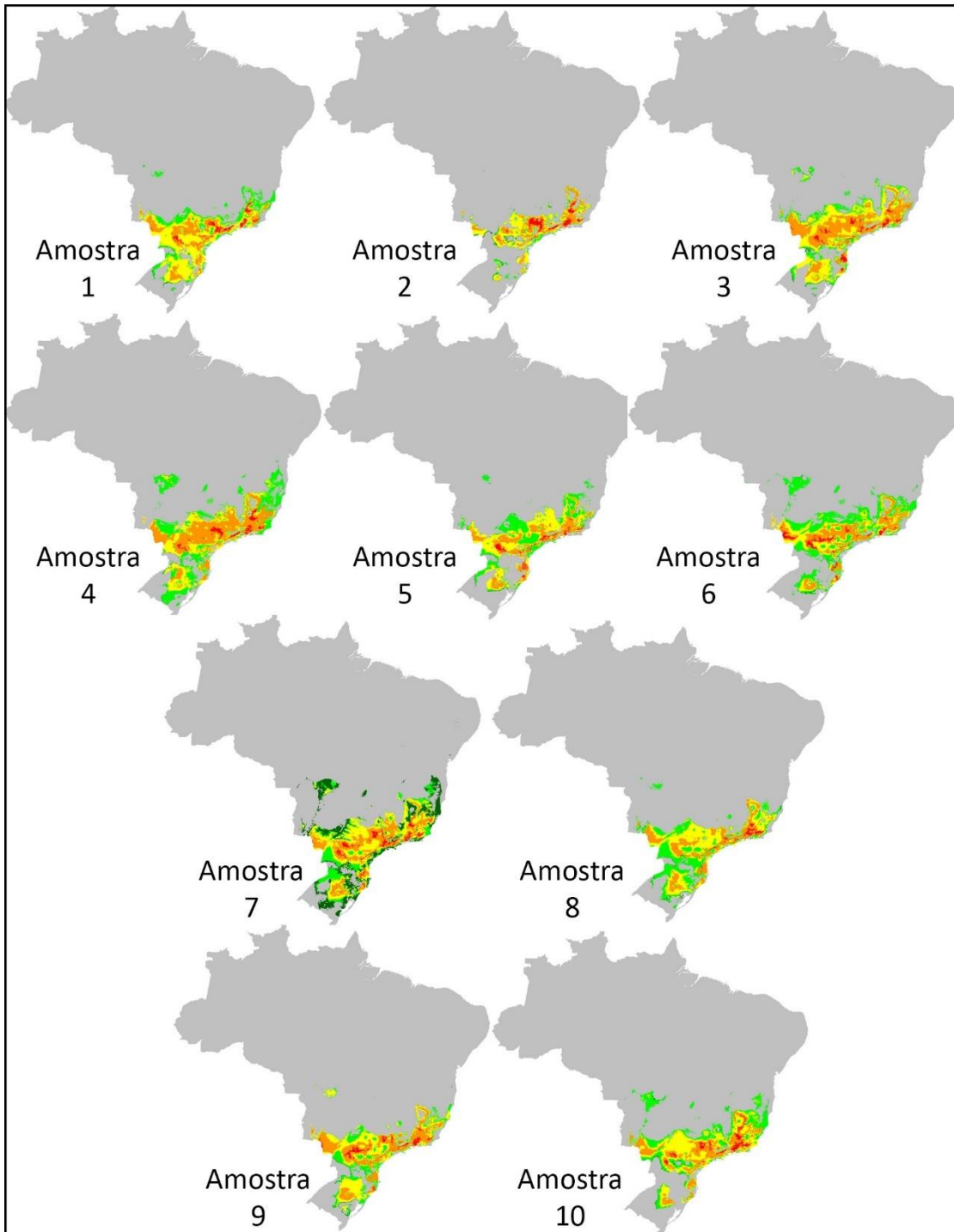


FIG 13. Modelagem da distribuição potencial de *Chiroxiphia caudata* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 50 pontos de presença, na proporção de pseudo ausência 1x1

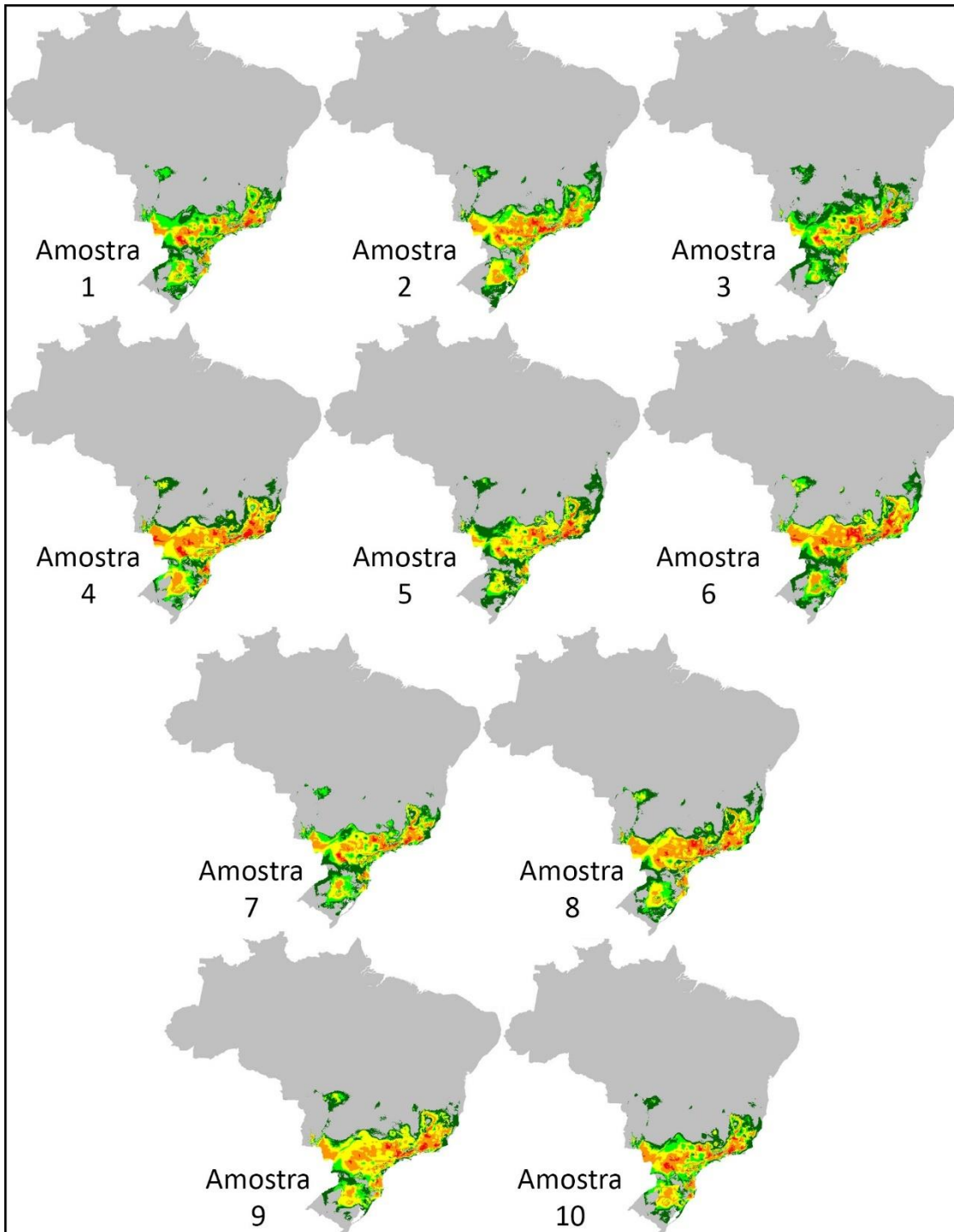


FIG 14. Modelagem da distribuição potencial de *Chiroxiphia caudata* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 150 pontos de presença, na proporção de pseudo ausência 1x1

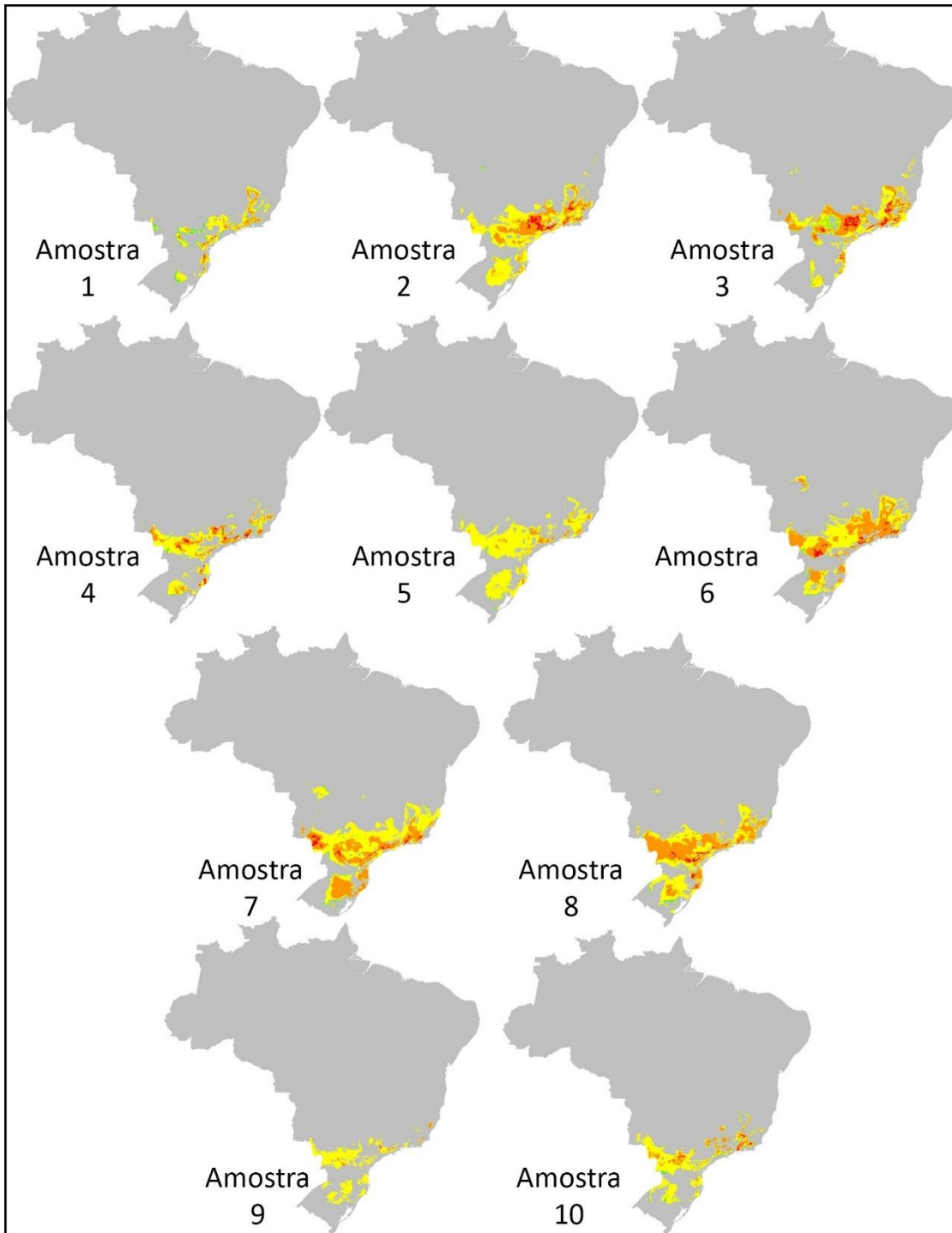


FIG 15. Modelagem da distribuição potencial de *Chiroxiphia caudata* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 20 pontos de presença, na proporção de pseudo ausência 25x1

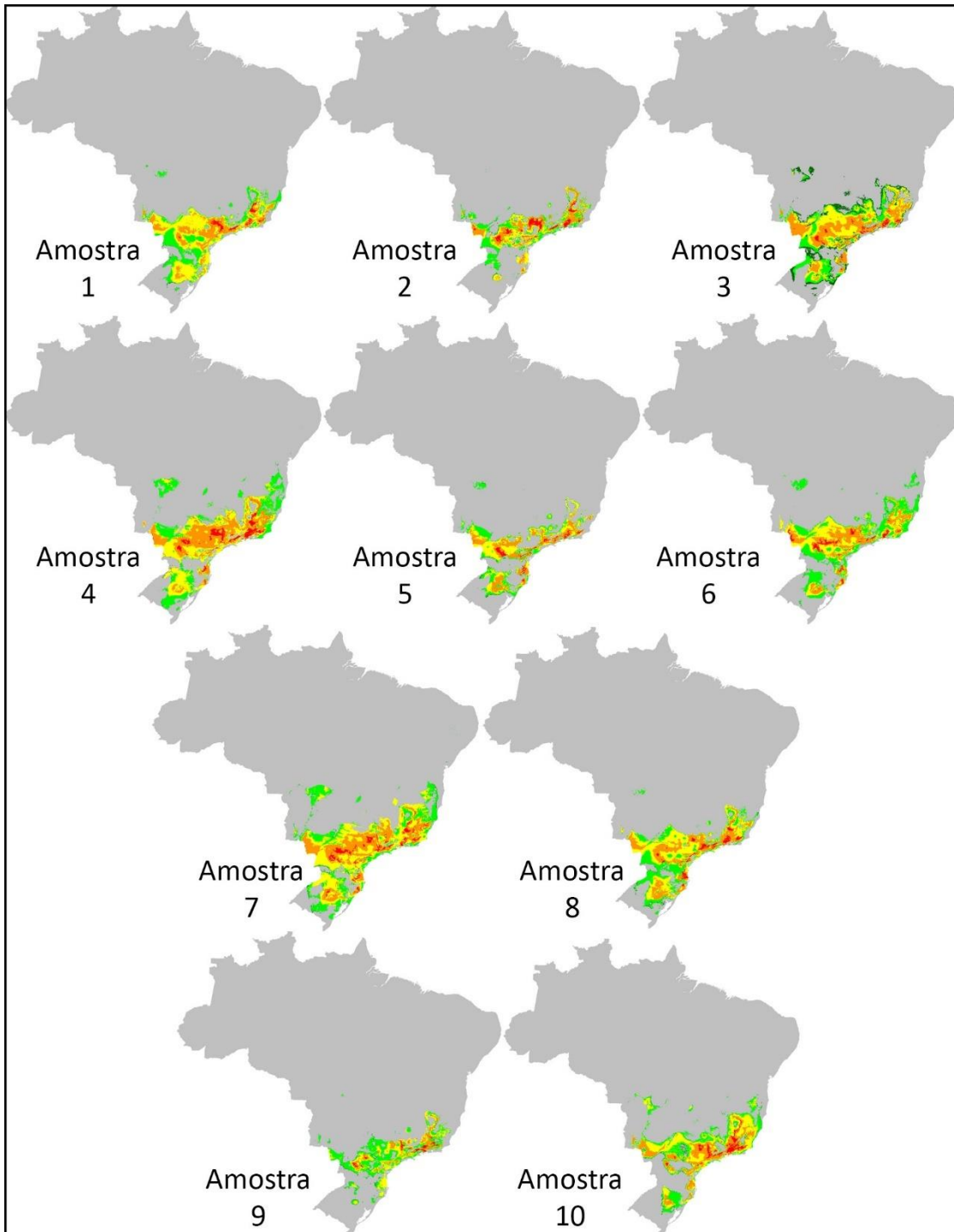


FIG 16. Modelagem da distribuição potencial de *Chiroxiphia caudata* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 50 pontos de presença, na proporção de pseudo ausência 25x1

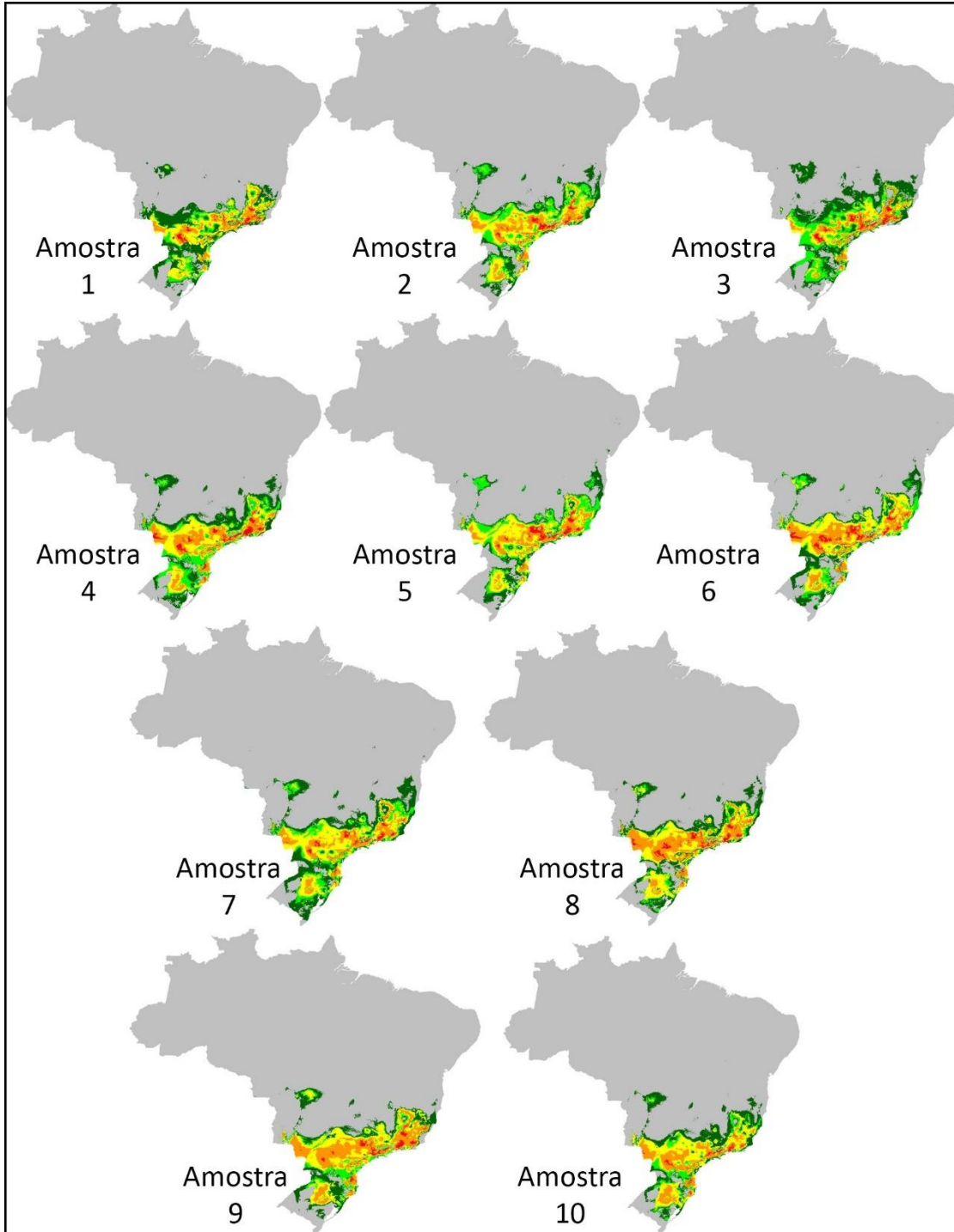


FIG 17. Modelagem da distribuição potencial de *Chiroxiphia caudata* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 150 pontos de presença, na proporção de pseudo ausência 25x1

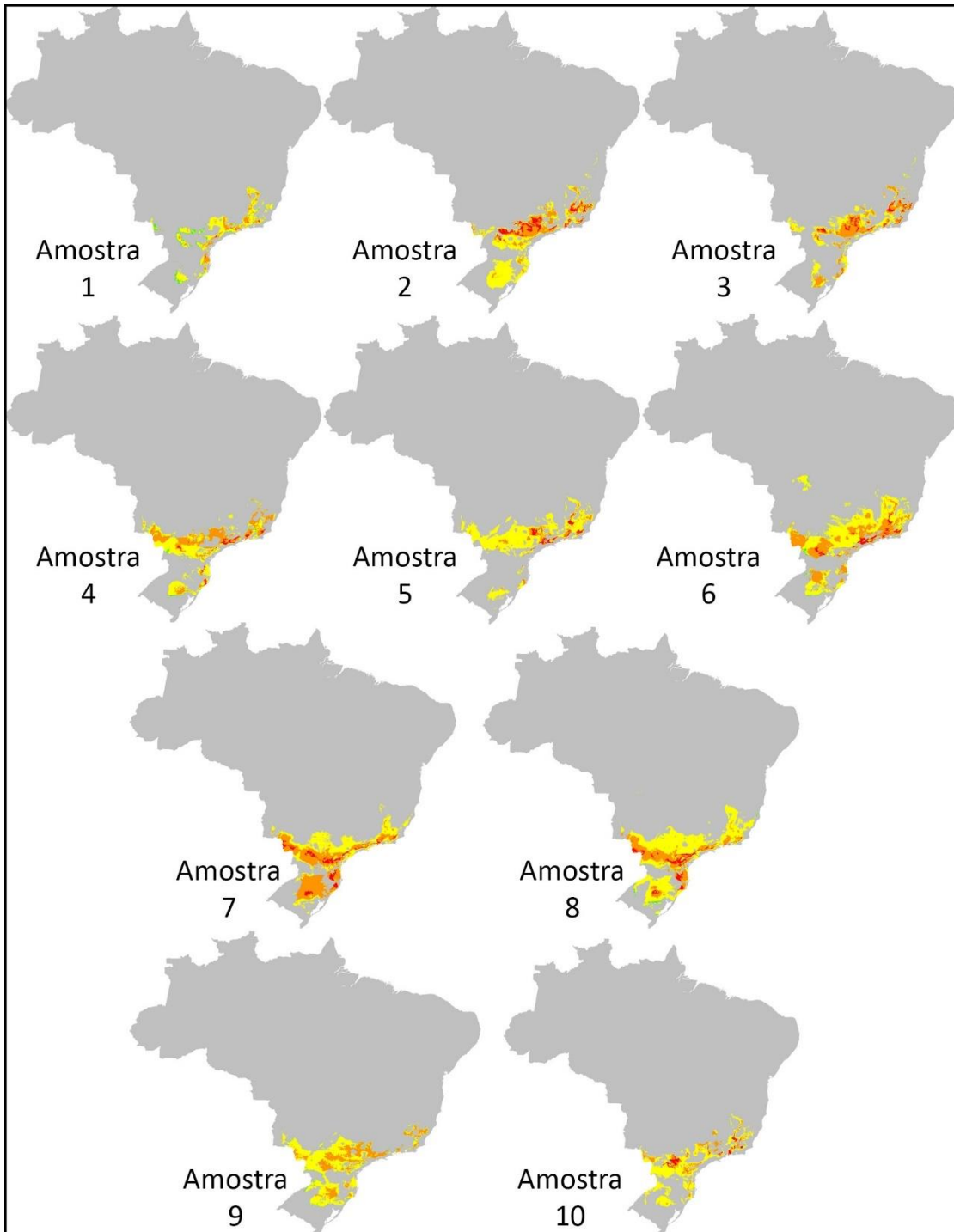


FIG 18. Modelagem da distribuição potencial de *Chiroxiphia caudata* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 20 pontos de presença, na proporção de pseudo ausência 50x1

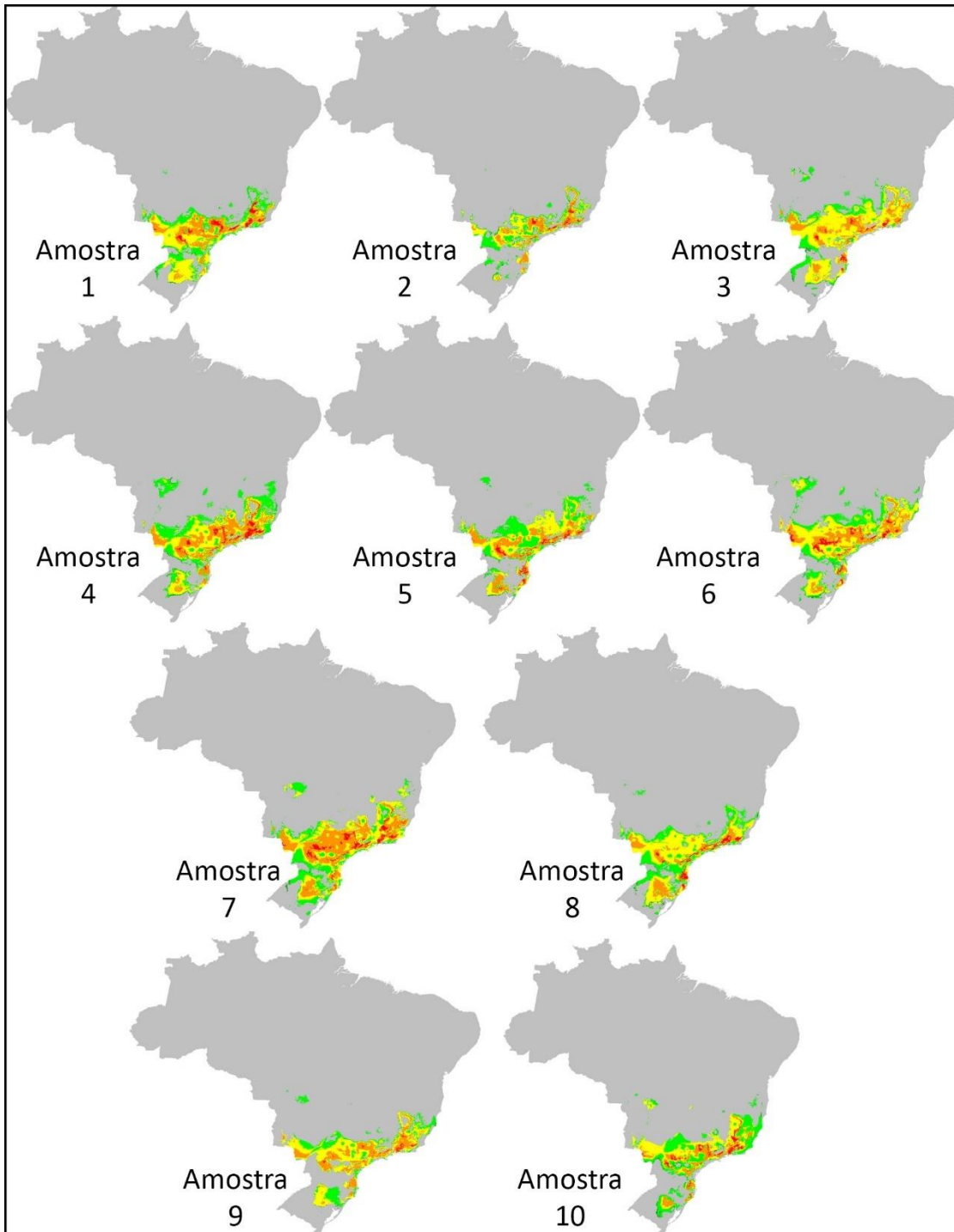


FIG 19. Modelagem da distribuição potencial de *Chiroxiphia caudata* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 50 pontos de presença, na proporção de pseudo ausência 50x1

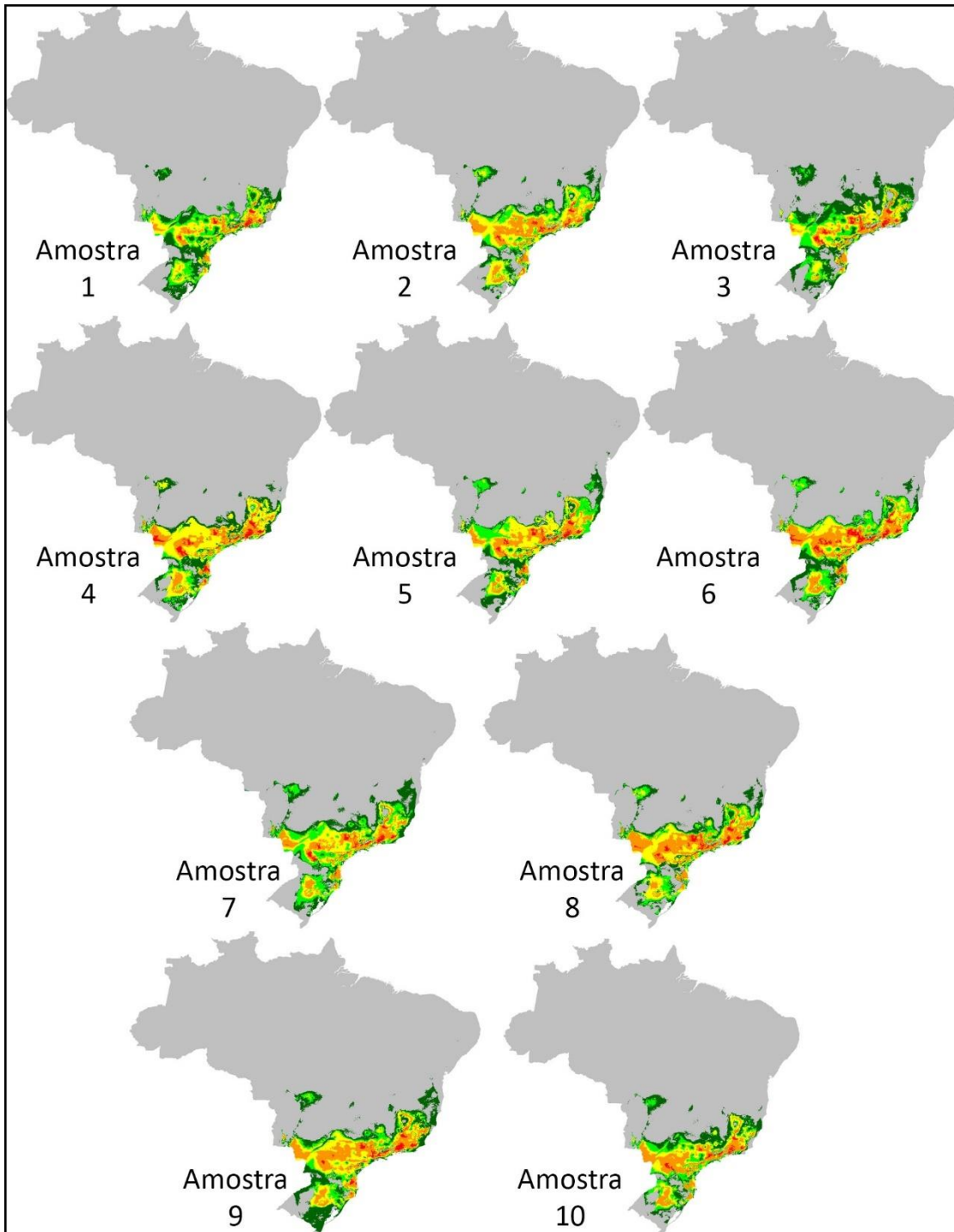


FIG 20. Modelagem da distribuição potencial de *Chiroxiphia caudata* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 150 pontos de presença, na proporção de pseudo ausência 50x1

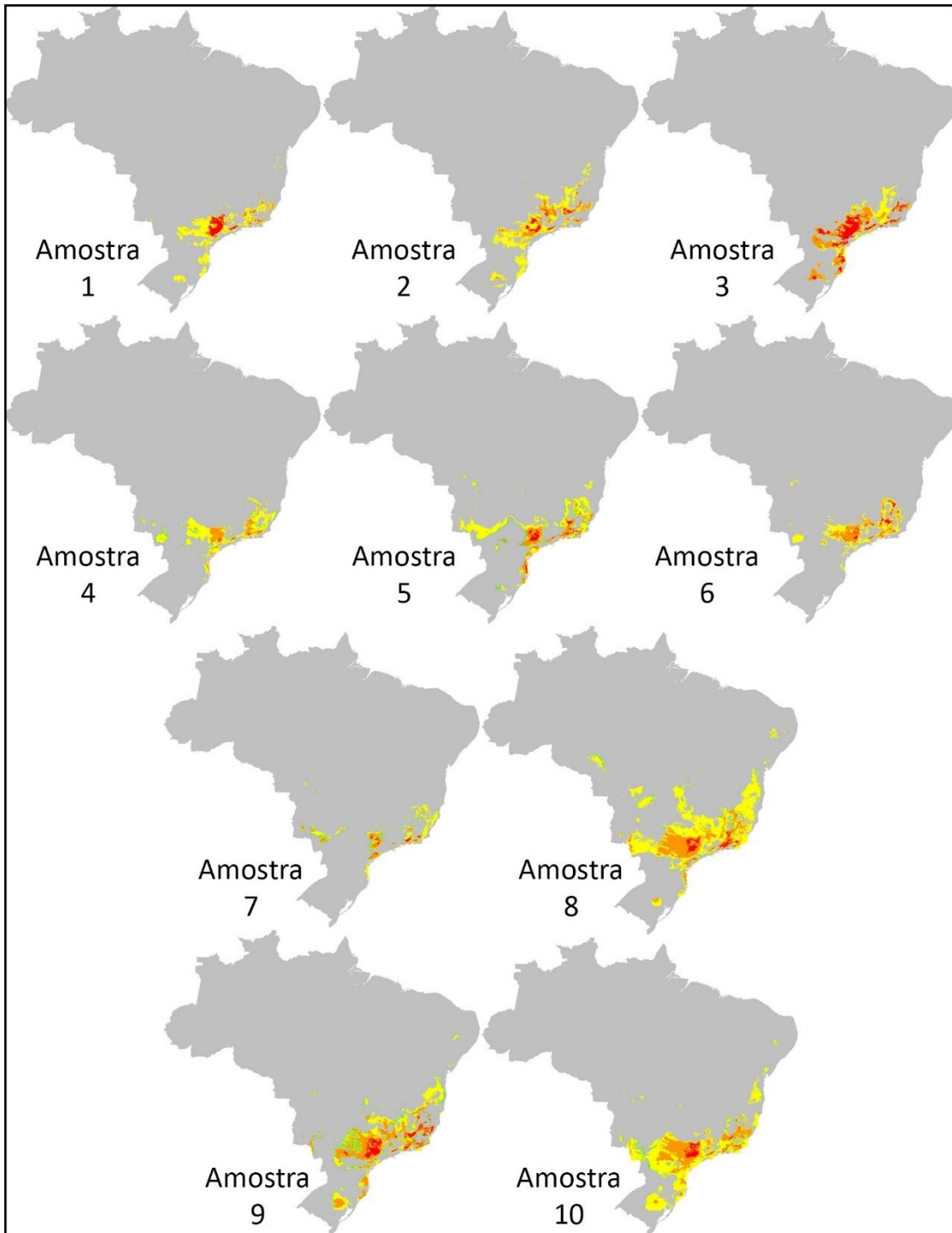


FIG 12. Modelagem da distribuição potencial de *Hemithraupis ruficapilla* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 20 pontos de presença, na proporção de pseudo ausência 1x1

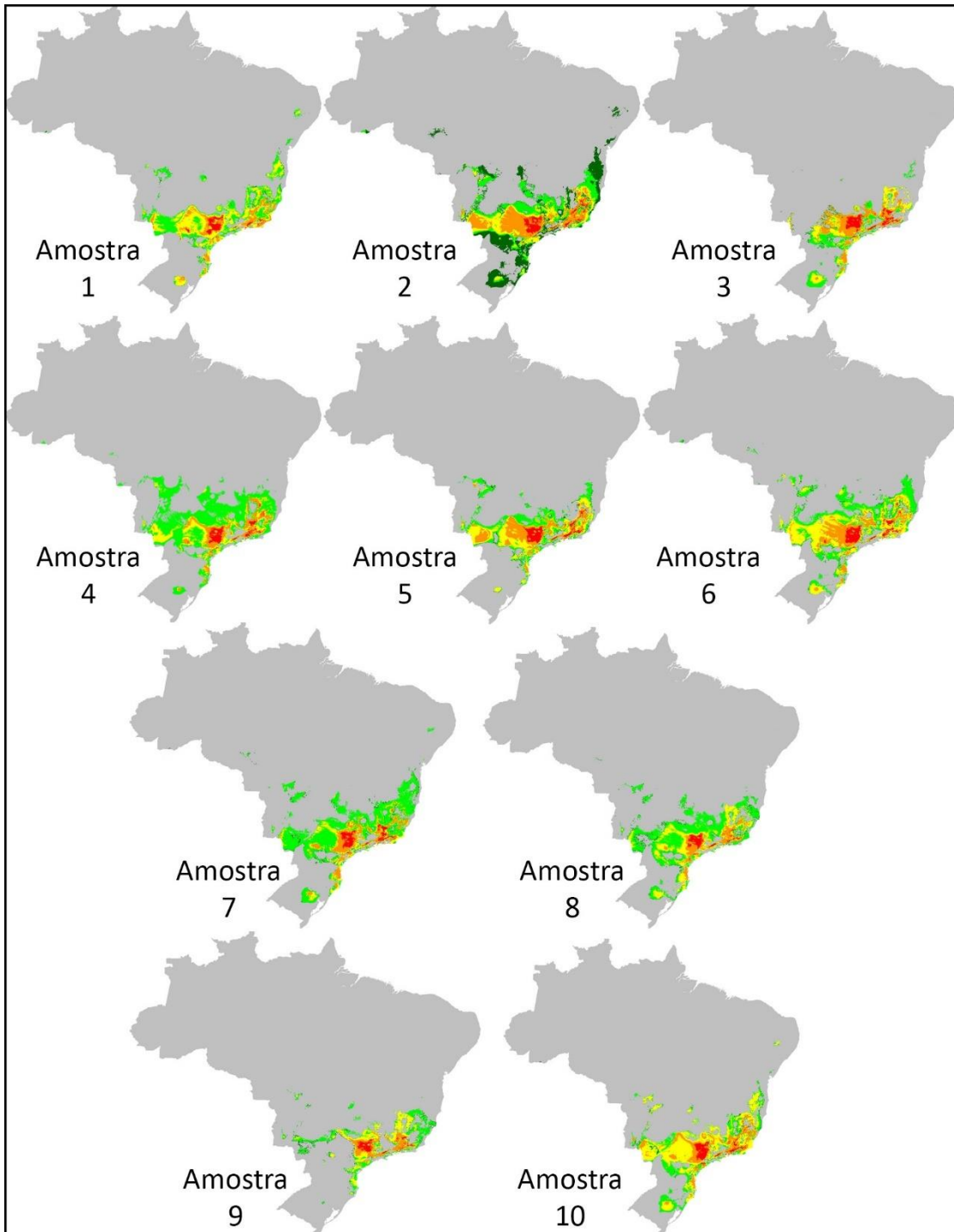


FIG 13. Modelagem da distribuição potencial de *Hemithraupis ruficapilla* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 50 pontos de presença, na proporção de pseudo ausência 1x1

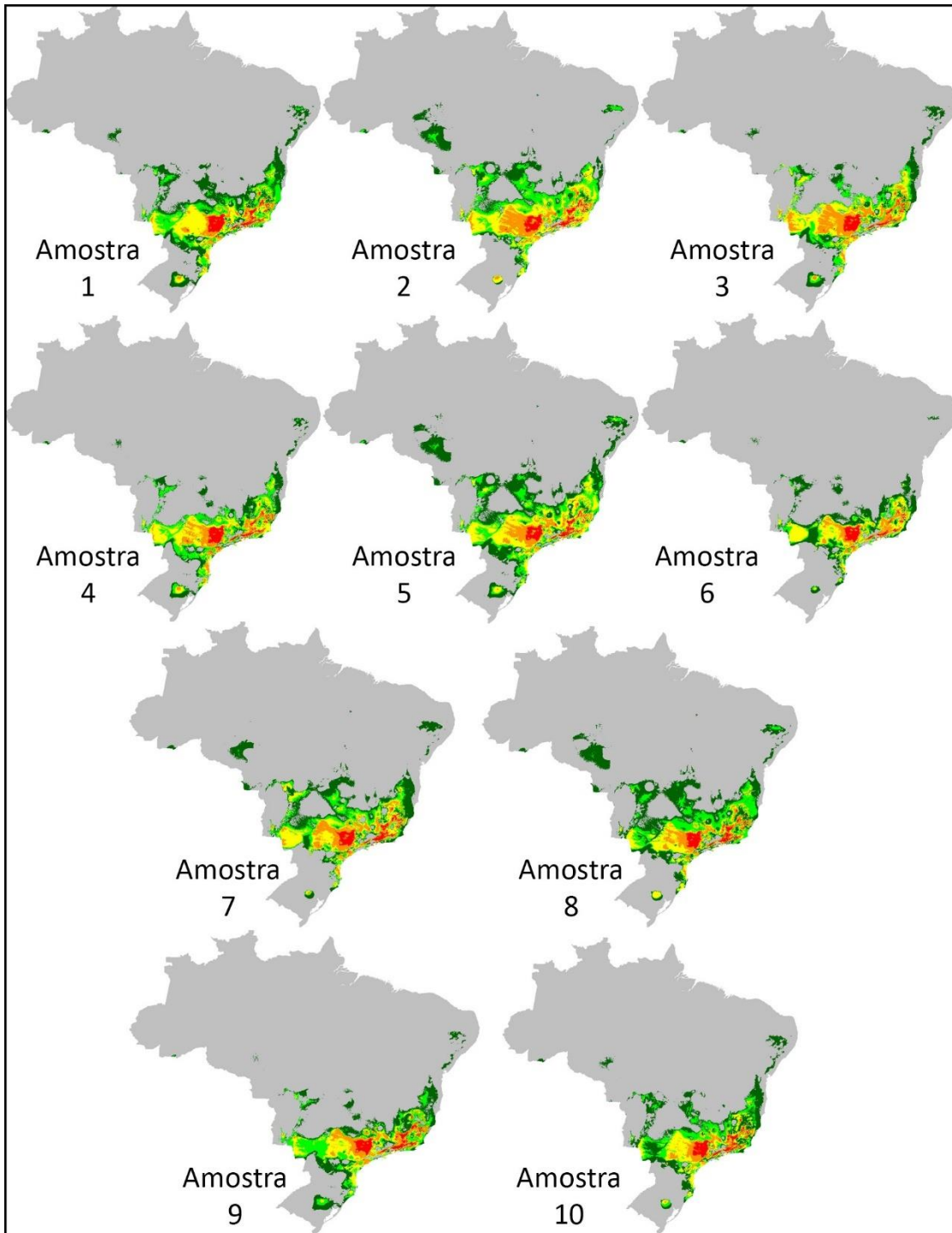


FIG 14. Modelagem da distribuição potencial de *Hemithraupis ruficapilla* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 150 pontos de presença, na proporção de pseudo ausência 1x1

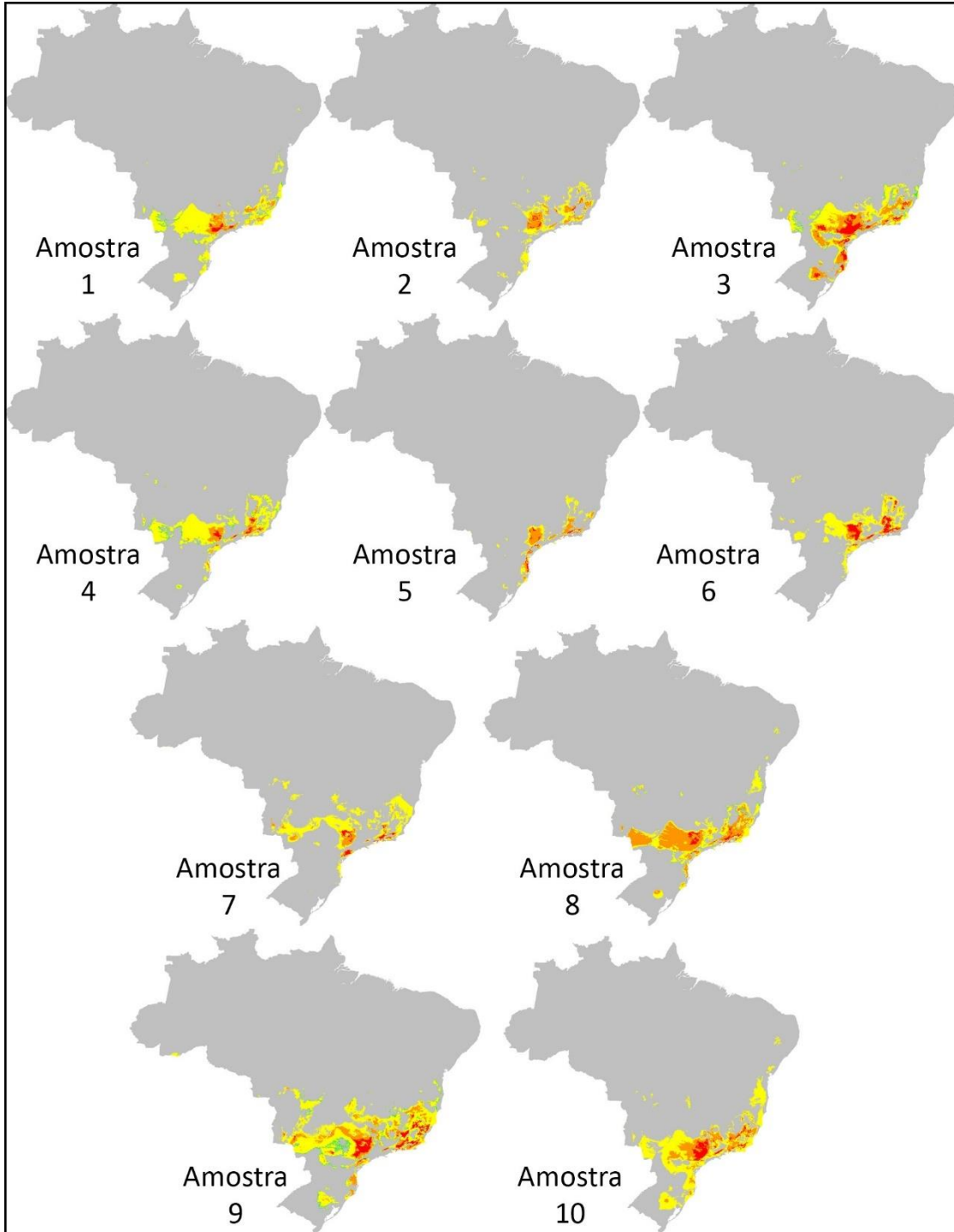


FIG 15. Modelagem da distribuição potencial de *Hemithraupis ruficapilla* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 20 pontos de presença, na proporção de pseudo ausência 25x1

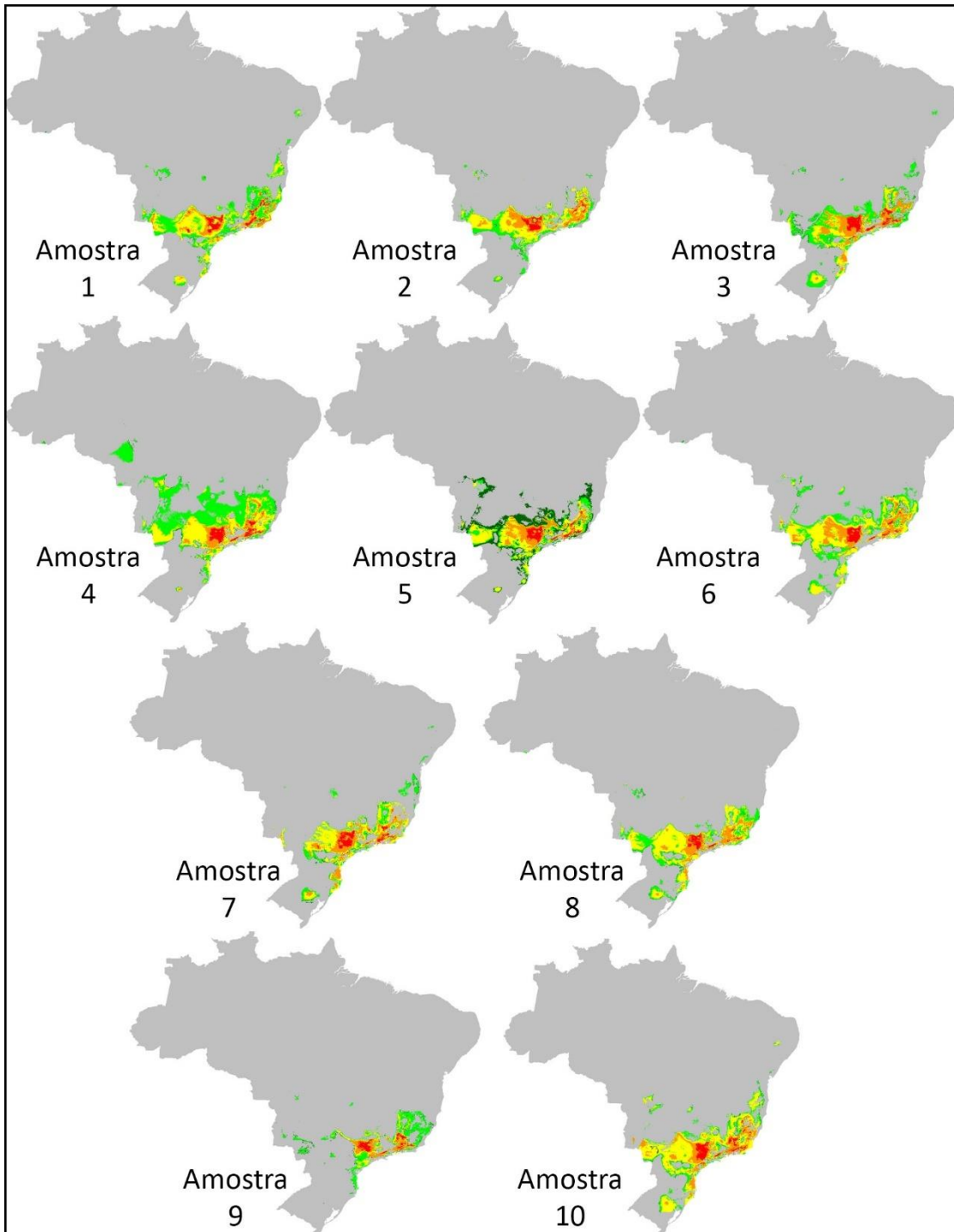


FIG 16. Modelagem da distribuição potencial de *Hemithraupis ruficapilla* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 50 pontos de presença, na proporção de pseudo ausência 25x1

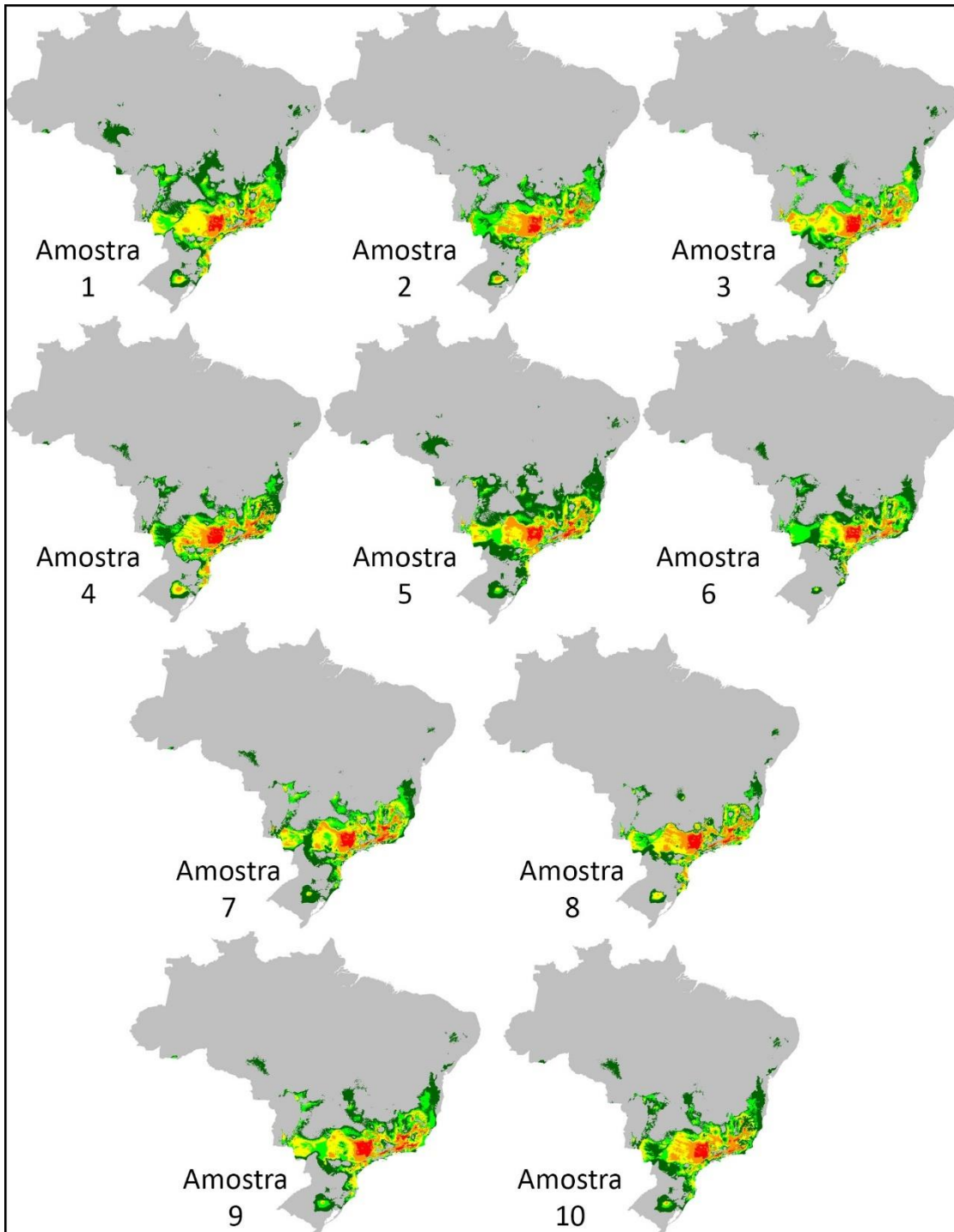


FIG 17. Modelagem da distribuição potencial de *Hemithraupis ruficapilla* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 150 pontos de presença, na proporção de pseudo ausência 25x1

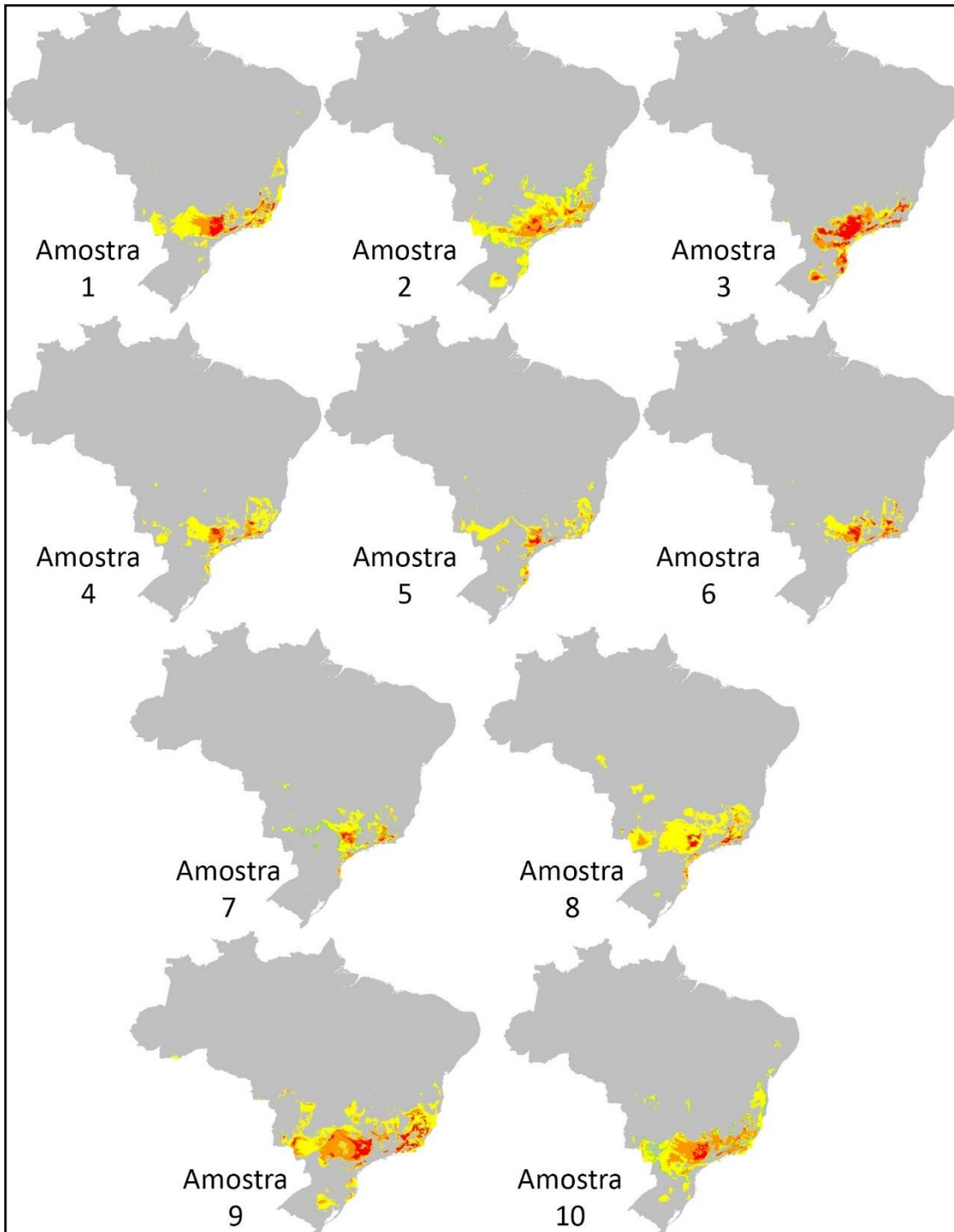


FIG 18. Modelagem da distribuição potencial de *Hemithraupis ruficapilla* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 20 pontos de presença, na proporção de pseudo ausência 50x1

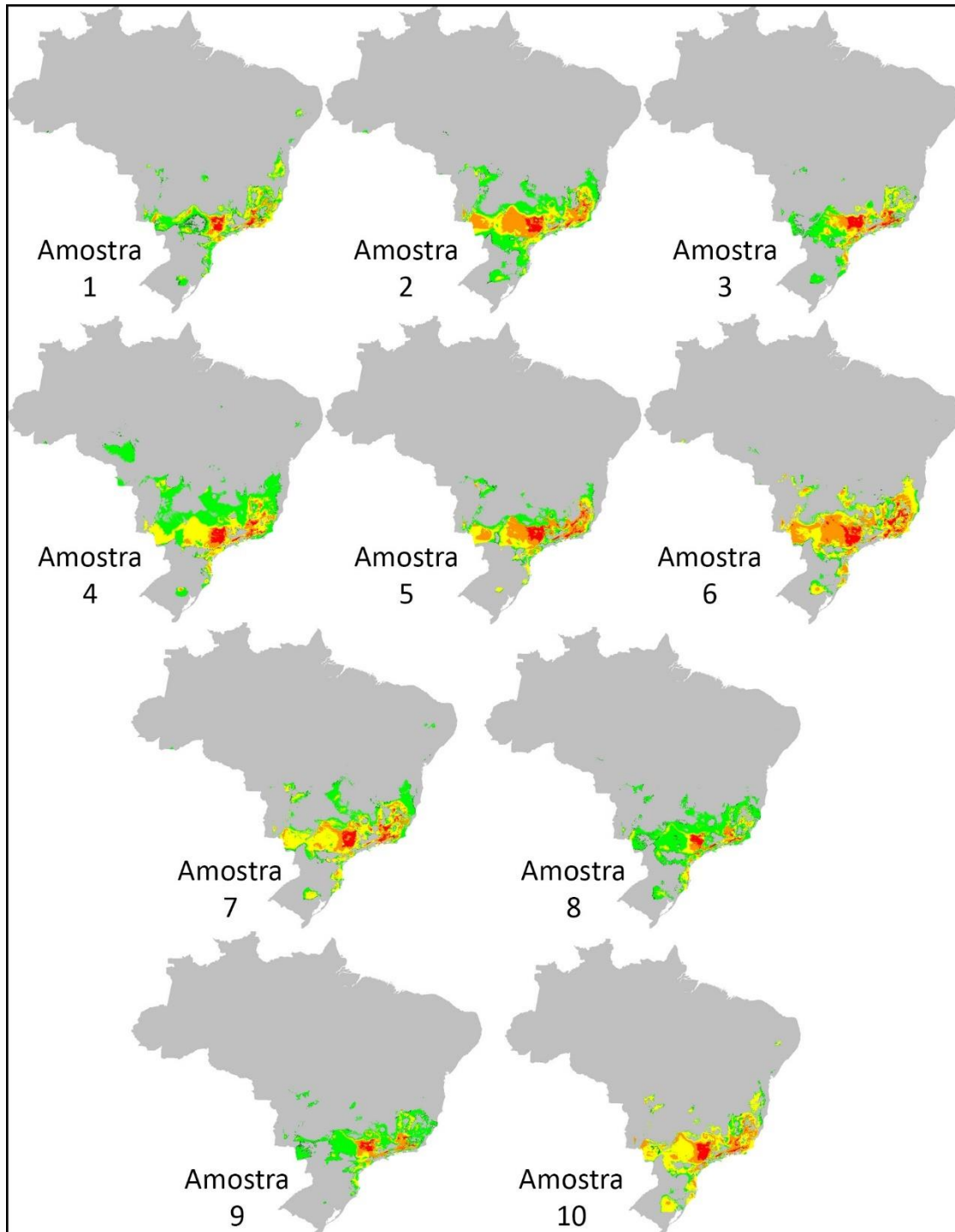


FIG 19. Modelagem da distribuição potencial de *Hemithraupis ruficapilla* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 50 pontos de presença, na proporção de pseudo ausência 50x1

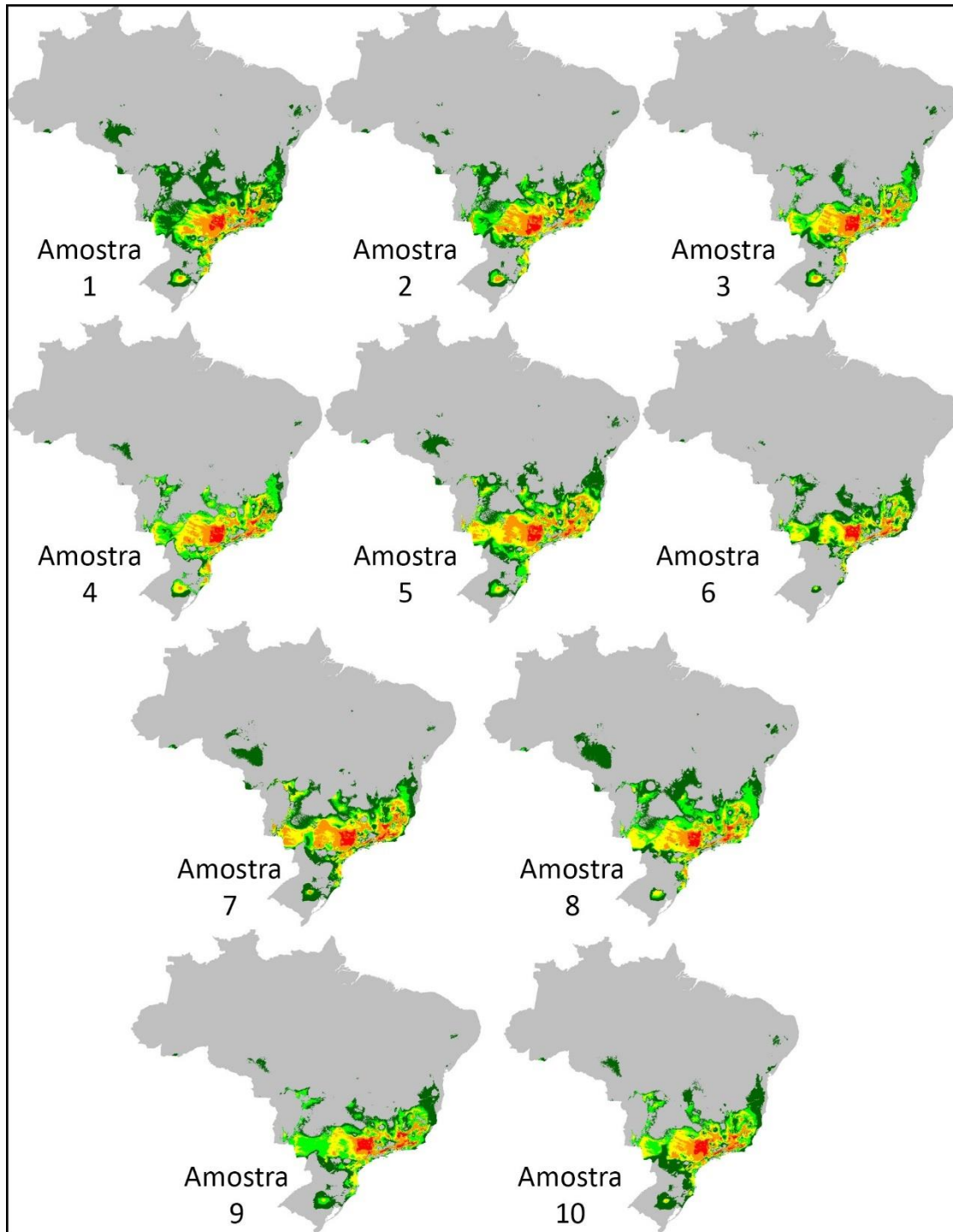


FIG 20. Modelagem da distribuição potencial de *Hemithraupis ruficapilla* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 150 pontos de presença, na proporção de pseudo ausência 50x1

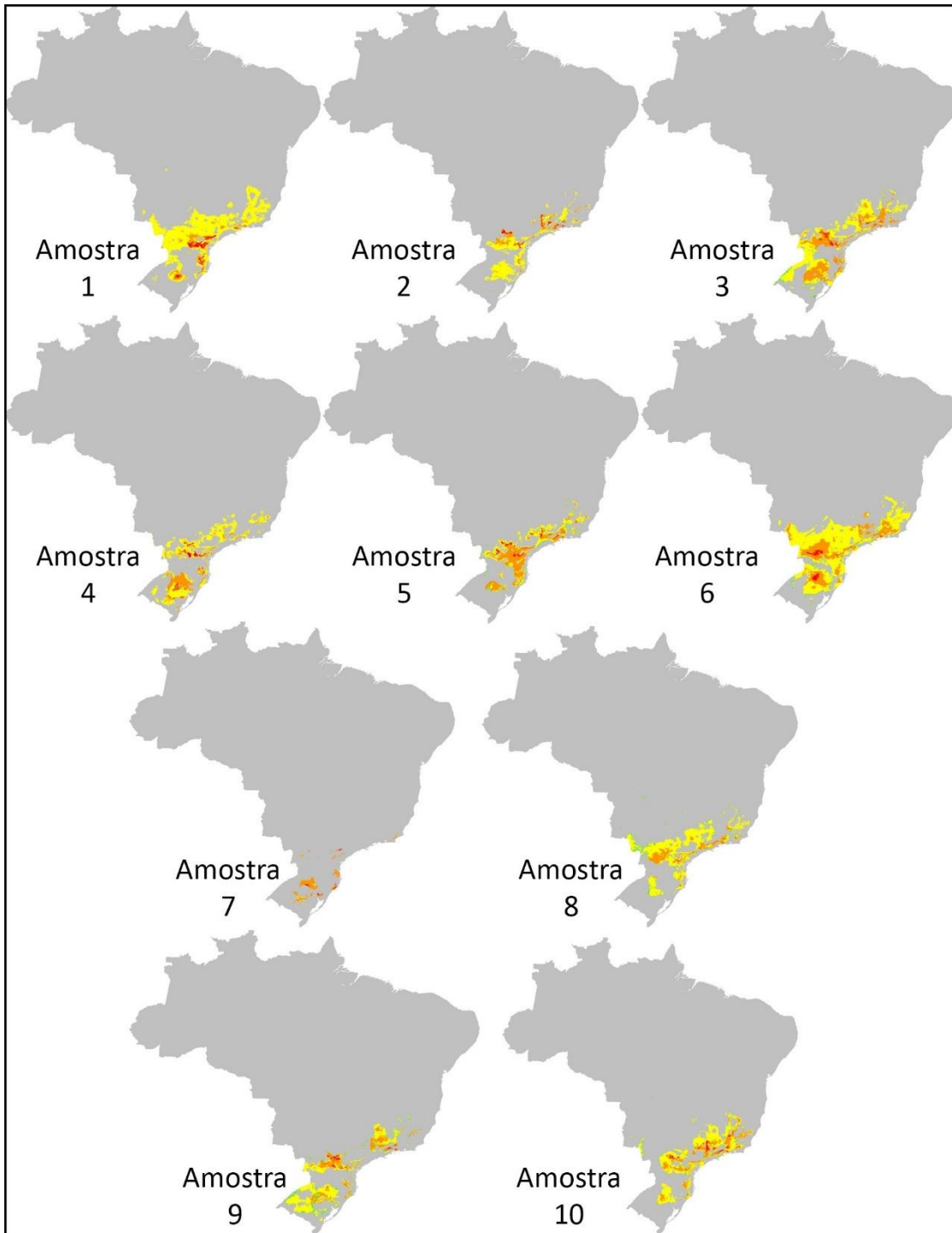


FIG 12. Modelagem da distribuição potencial de *Myiothlypis leucoblephara* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 20 pontos de presença, na proporção de pseudo ausência 1x1

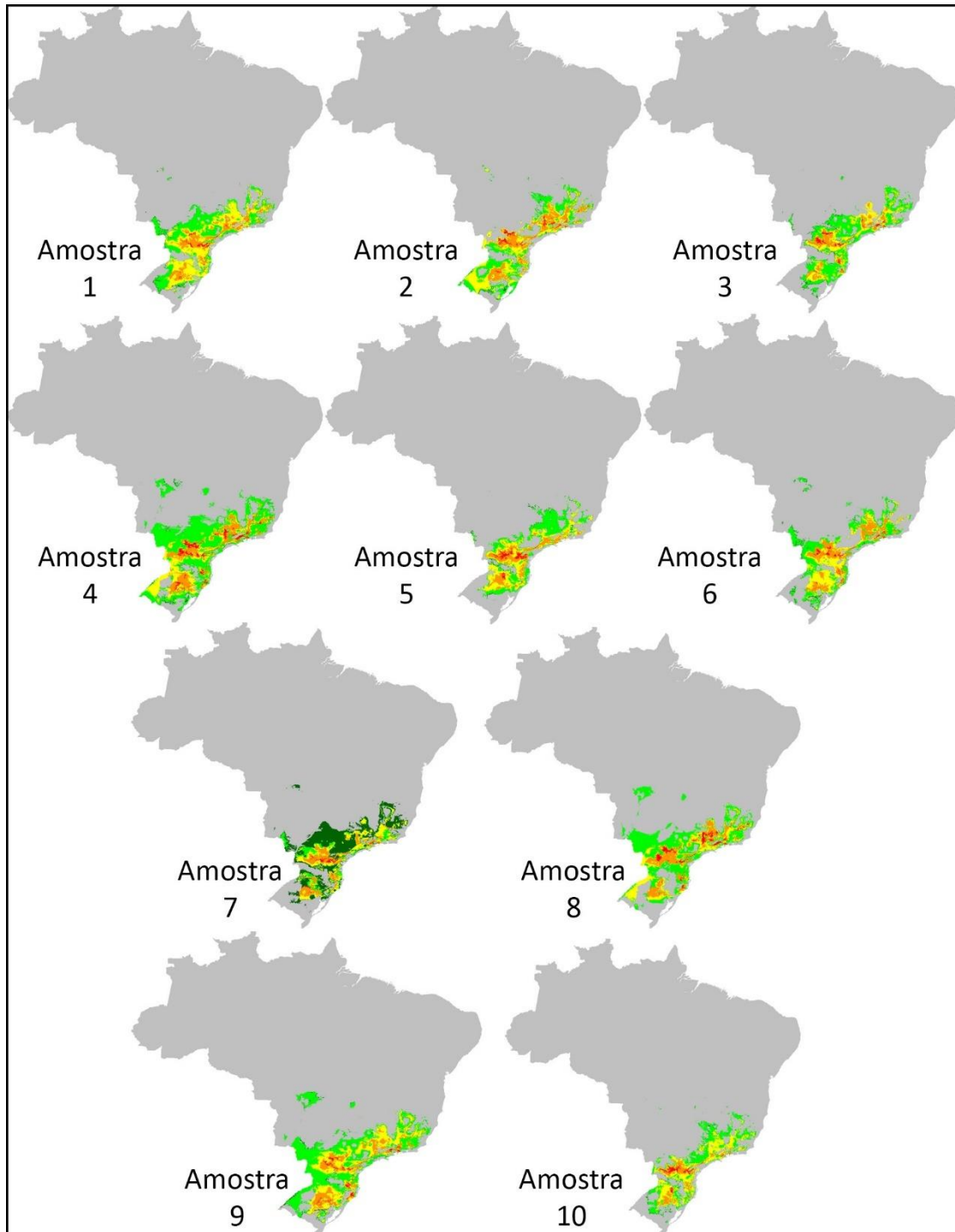


FIG 13. Modelagem da distribuição potencial de *Myiothlypis leucoblephara* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 50 pontos de presença, na proporção de pseudo ausência 1x1

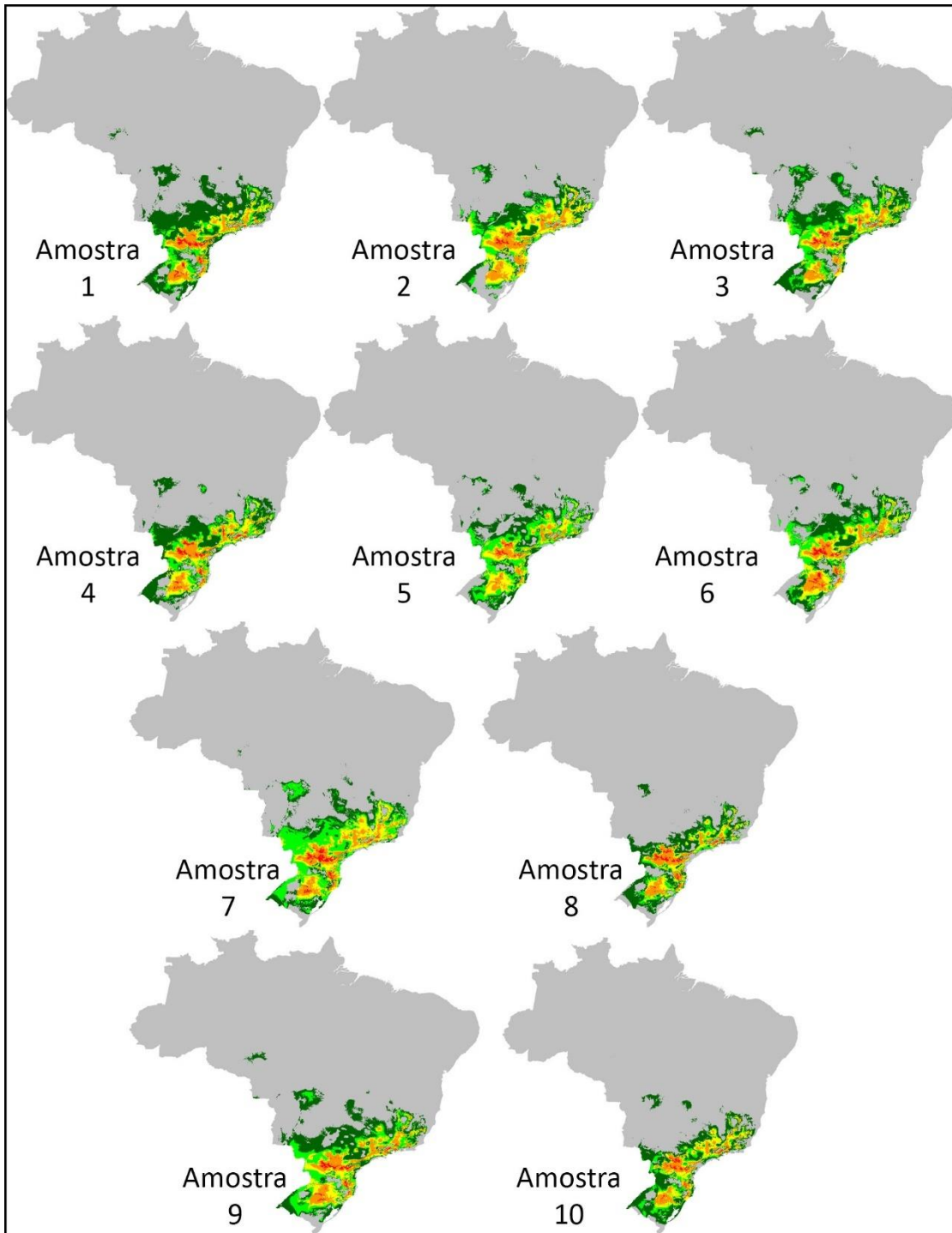


FIG 14. Modelagem da distribuição potencial de *Myiothlypis leucoblephara* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 150 pontos de presença, na proporção de pseudo ausência 1x1

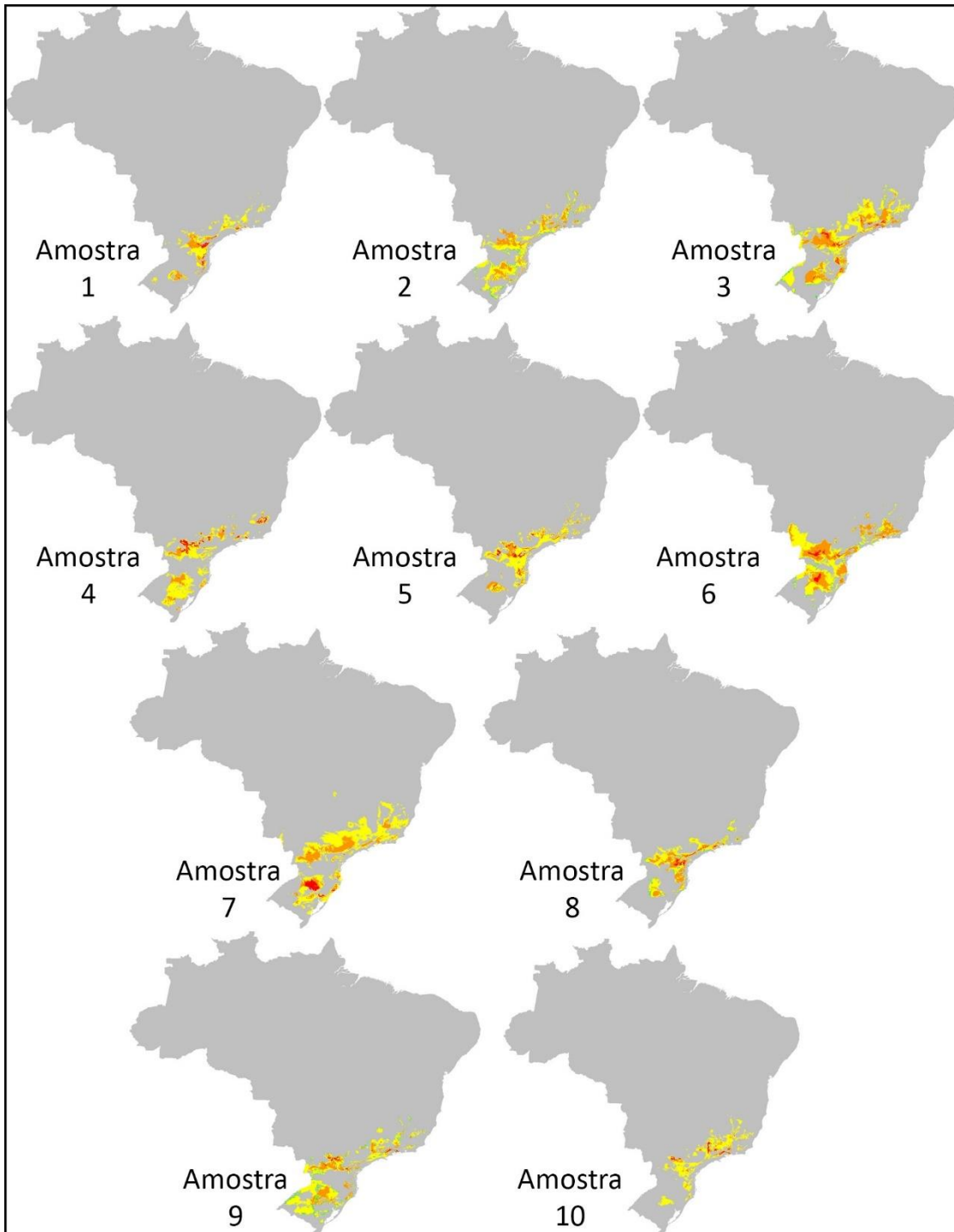


FIG 15. Modelagem da distribuição potencial de *Myiothlypis leucoblephara* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 20 pontos de presença, na proporção de pseudo ausência 25x1

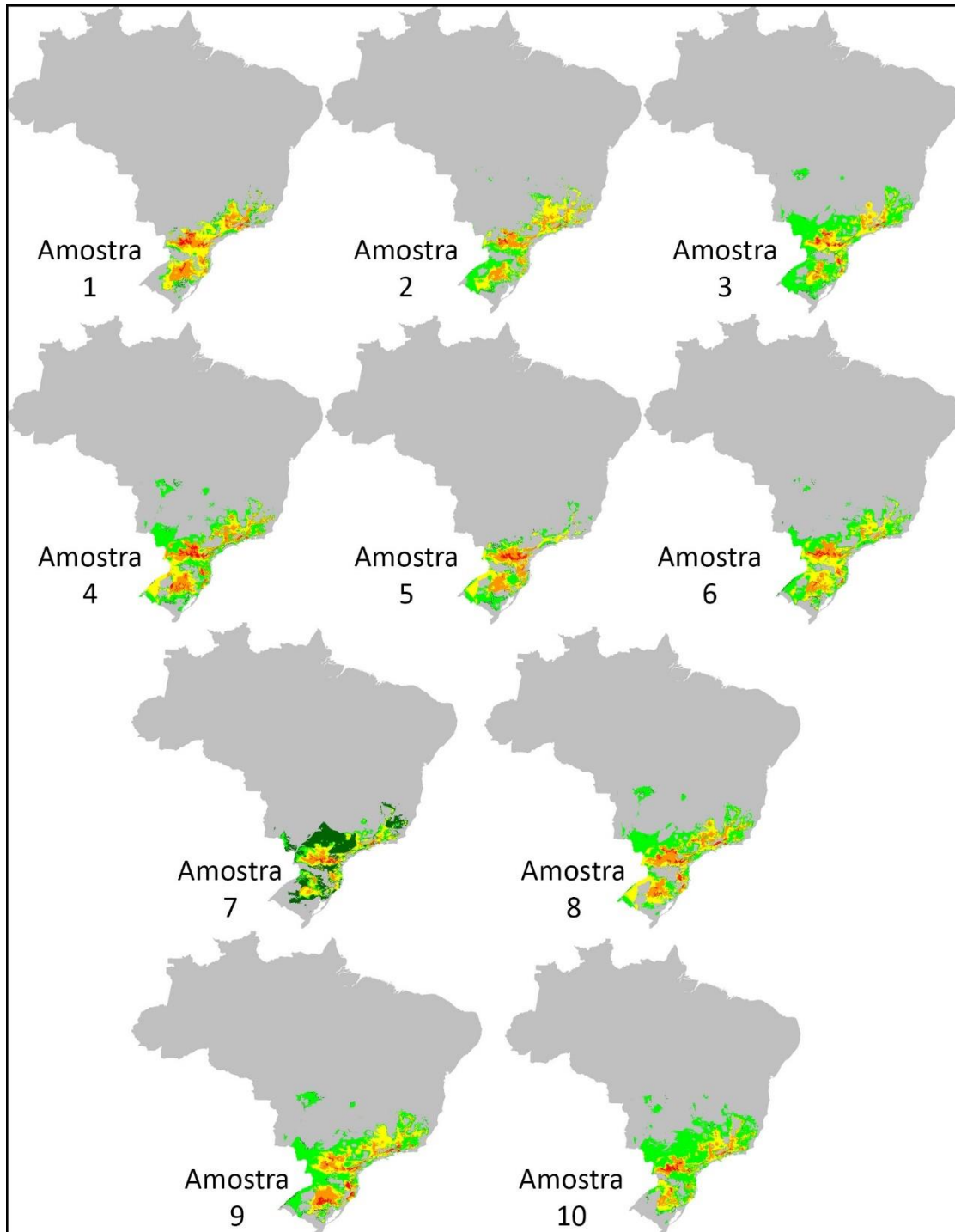


FIG 16. Modelagem da distribuição potencial de *Myiothlypis leucoblephara* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 50 pontos de presença, na proporção de pseudo ausência 25x1

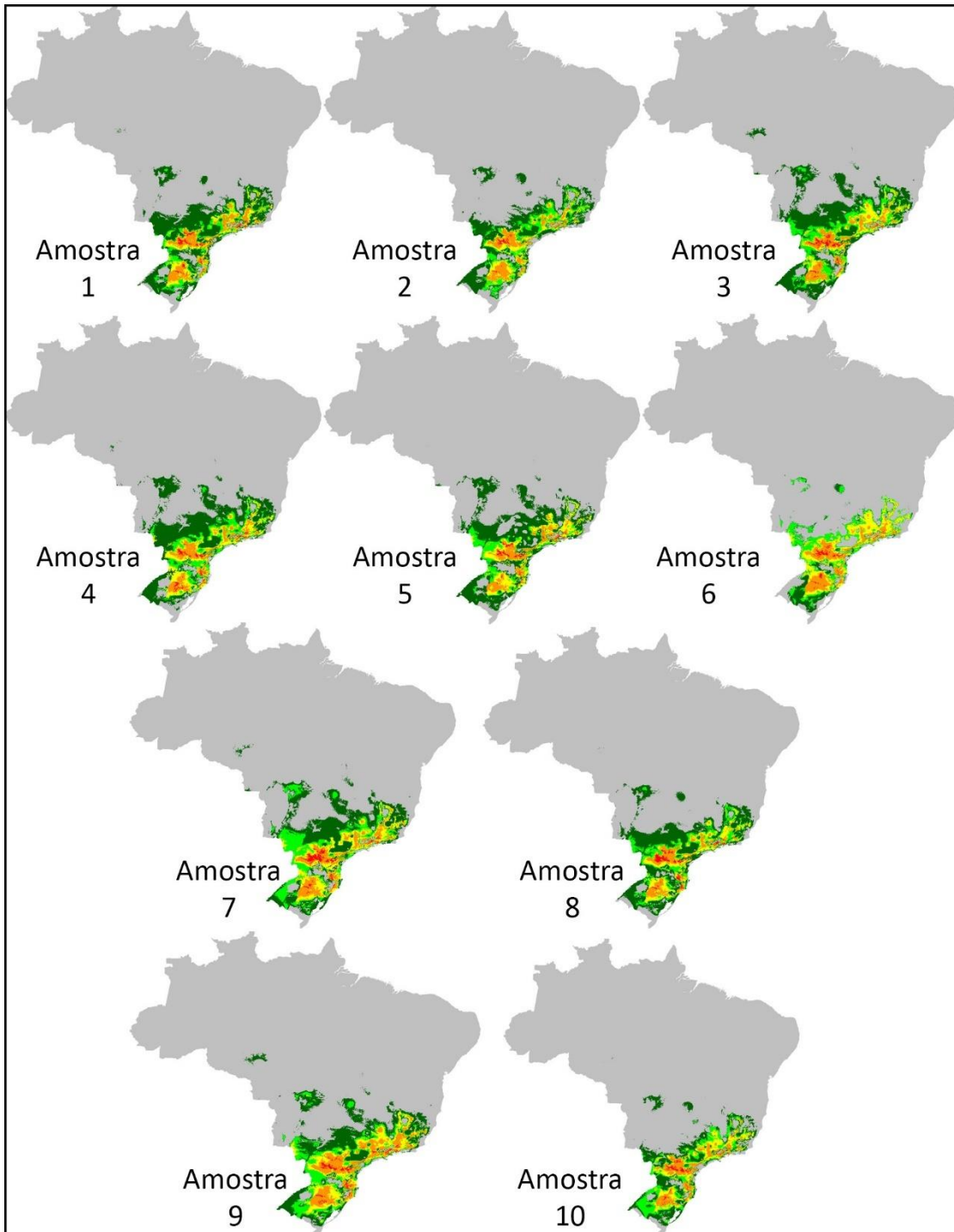


FIG 17. Modelagem da distribuição potencial de *Myiothlypis leucoblephara* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 150 pontos de presença, na proporção de pseudo ausência 25x1

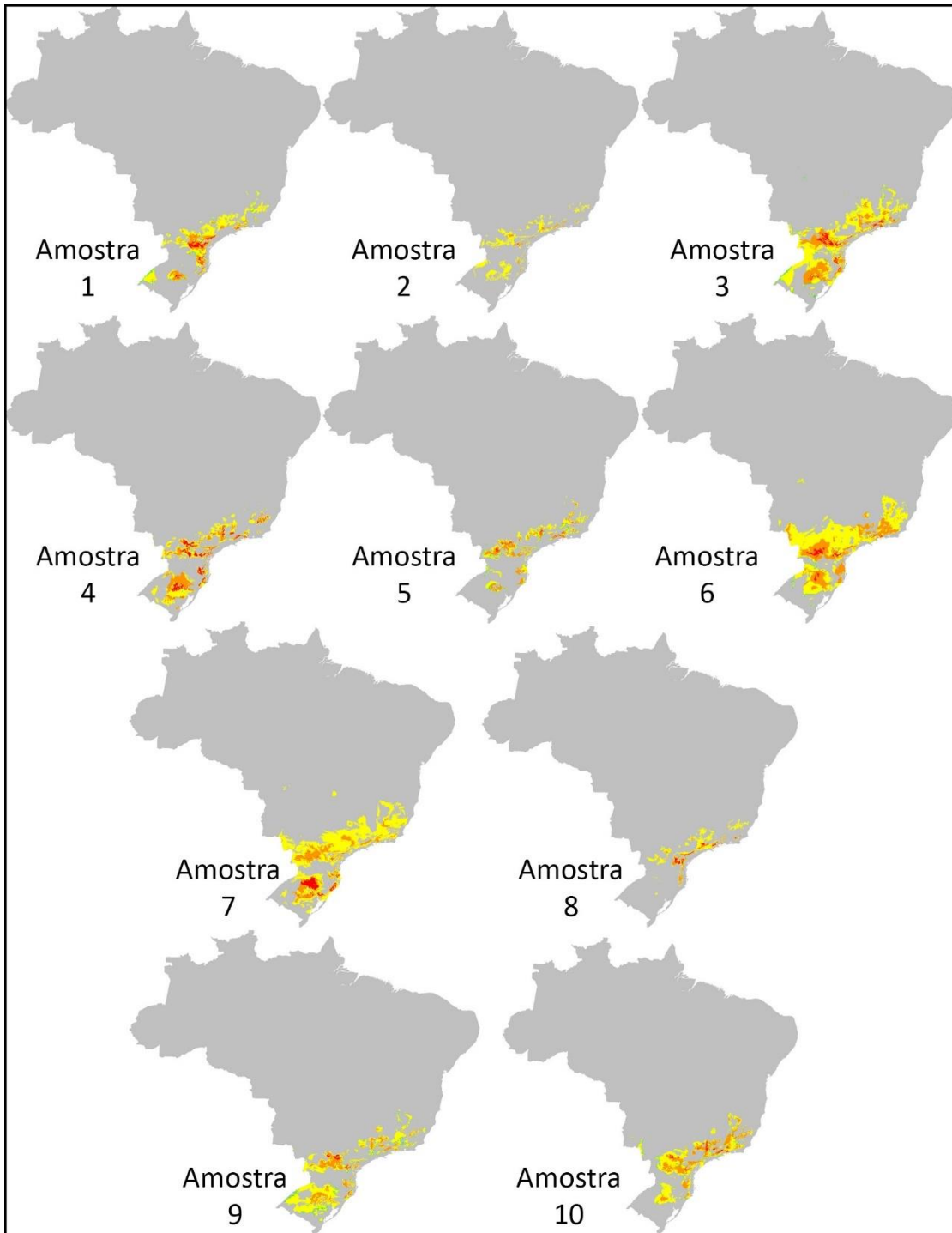


FIG 18. Modelagem da distribuição potencial de *Myiothlypis leucoblephara* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 20 pontos de presença, na proporção de pseudo ausência 50x1

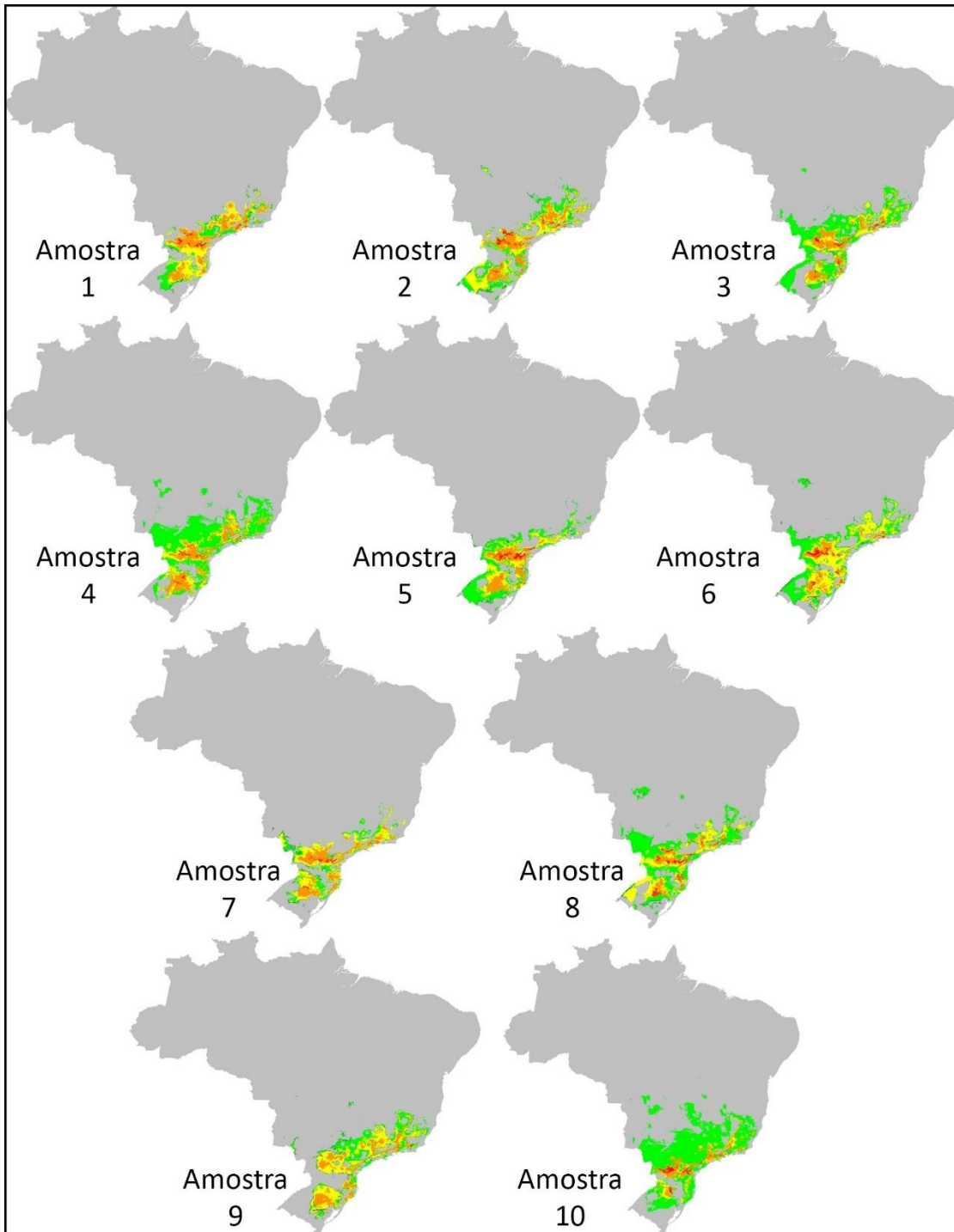


FIG 19. Modelagem da distribuição potencial de *Myiothlypis leucoblephara* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 50 pontos de presença, na proporção de pseudo ausência 50x1

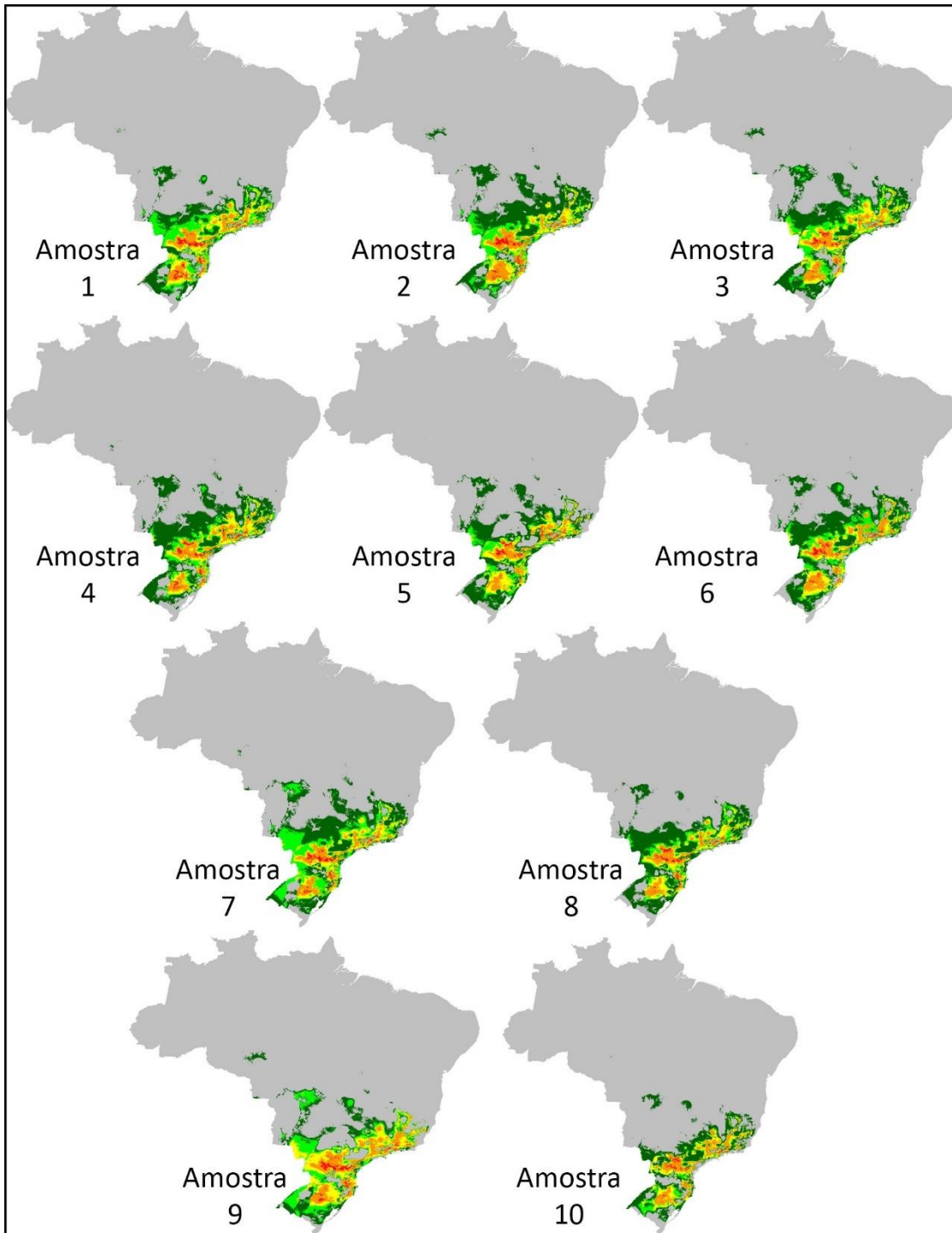


FIG 20. Modelagem da distribuição potencial de *Myiothlypis leucoblephara* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 150 pontos de presença, na proporção de pseudo ausência 50x1

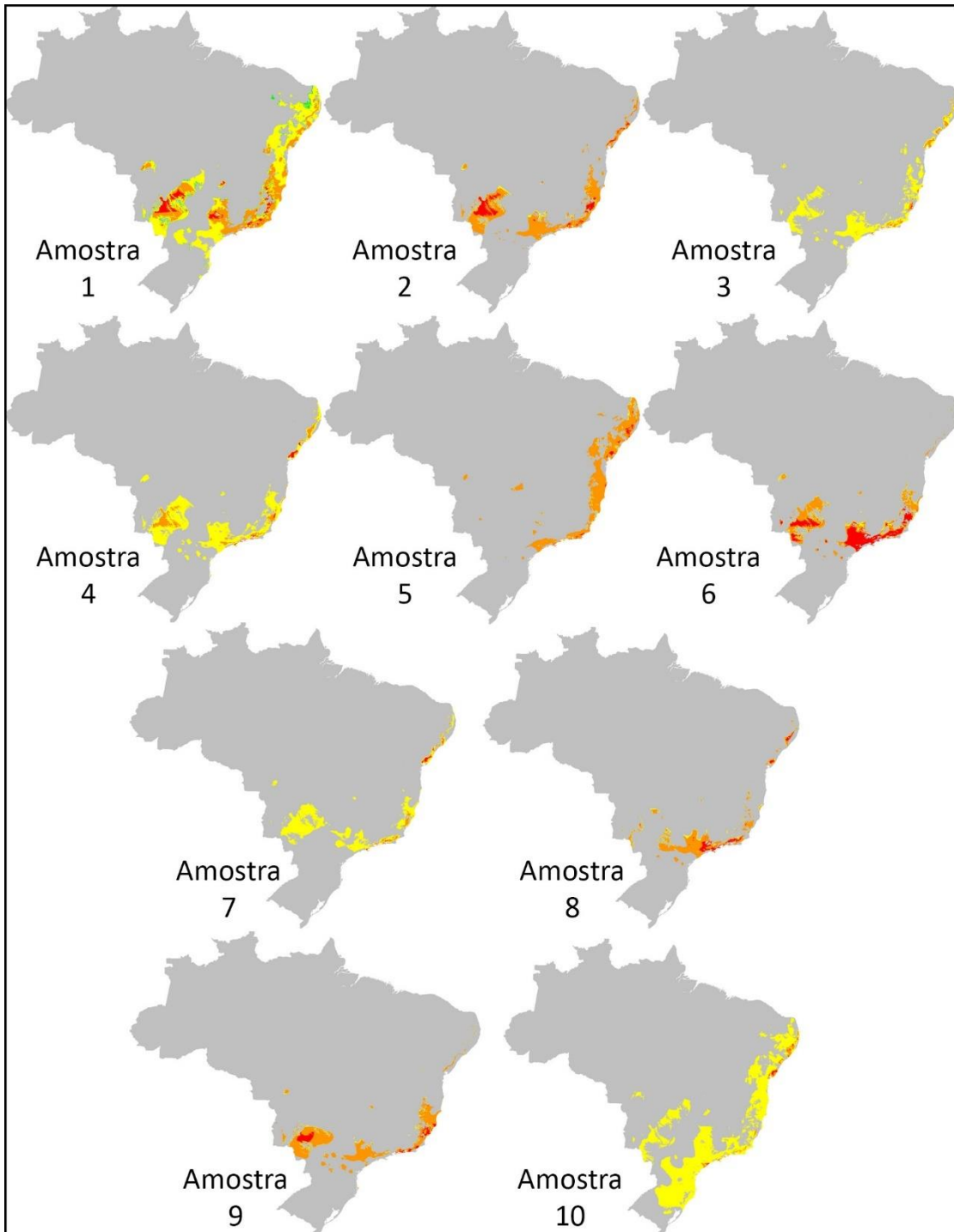


FIG 12. Modelagem da distribuição potencial de *Ramphocelus bresilius* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 20 pontos de presença, na proporção de pseudo ausência 1x1

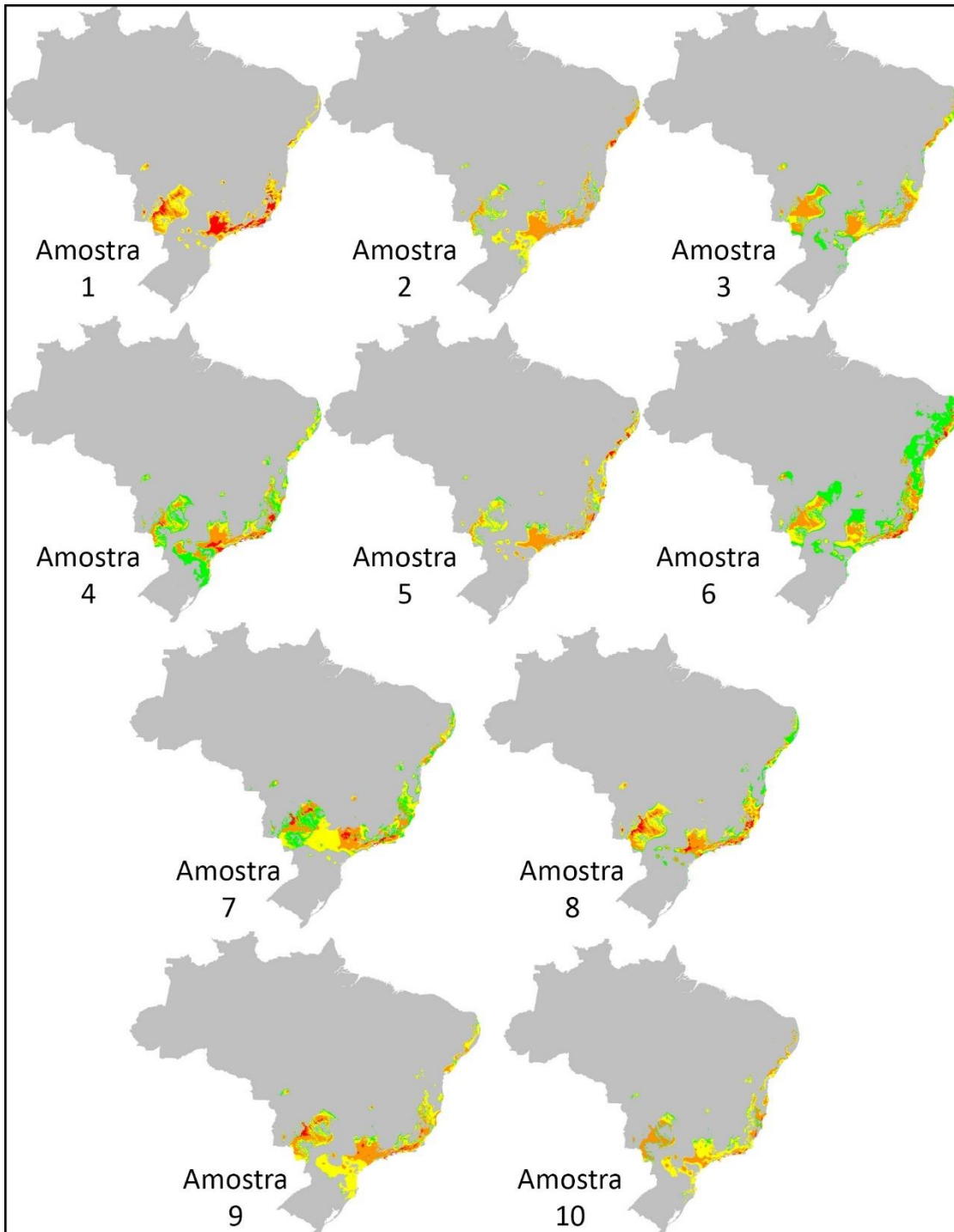


FIG 13. Modelagem da distribuição potencial de *Ramphocelus bresilius* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 50 pontos de presença, na proporção de pseudo ausência 1x1

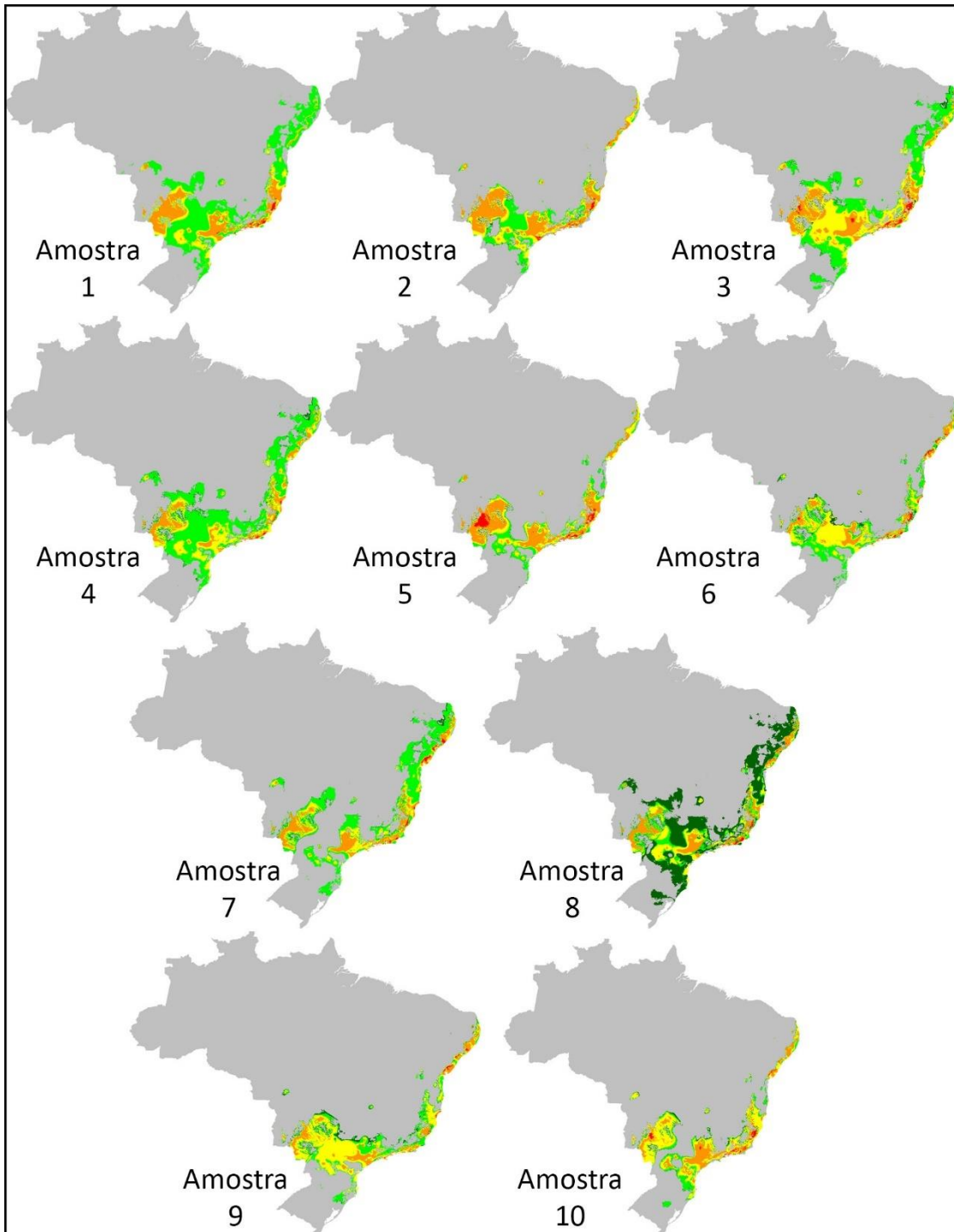


FIG 14. Modelagem da distribuição potencial de *Ramphocelus bresilius* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 150 pontos de presença, na proporção de pseudo ausência 1x1

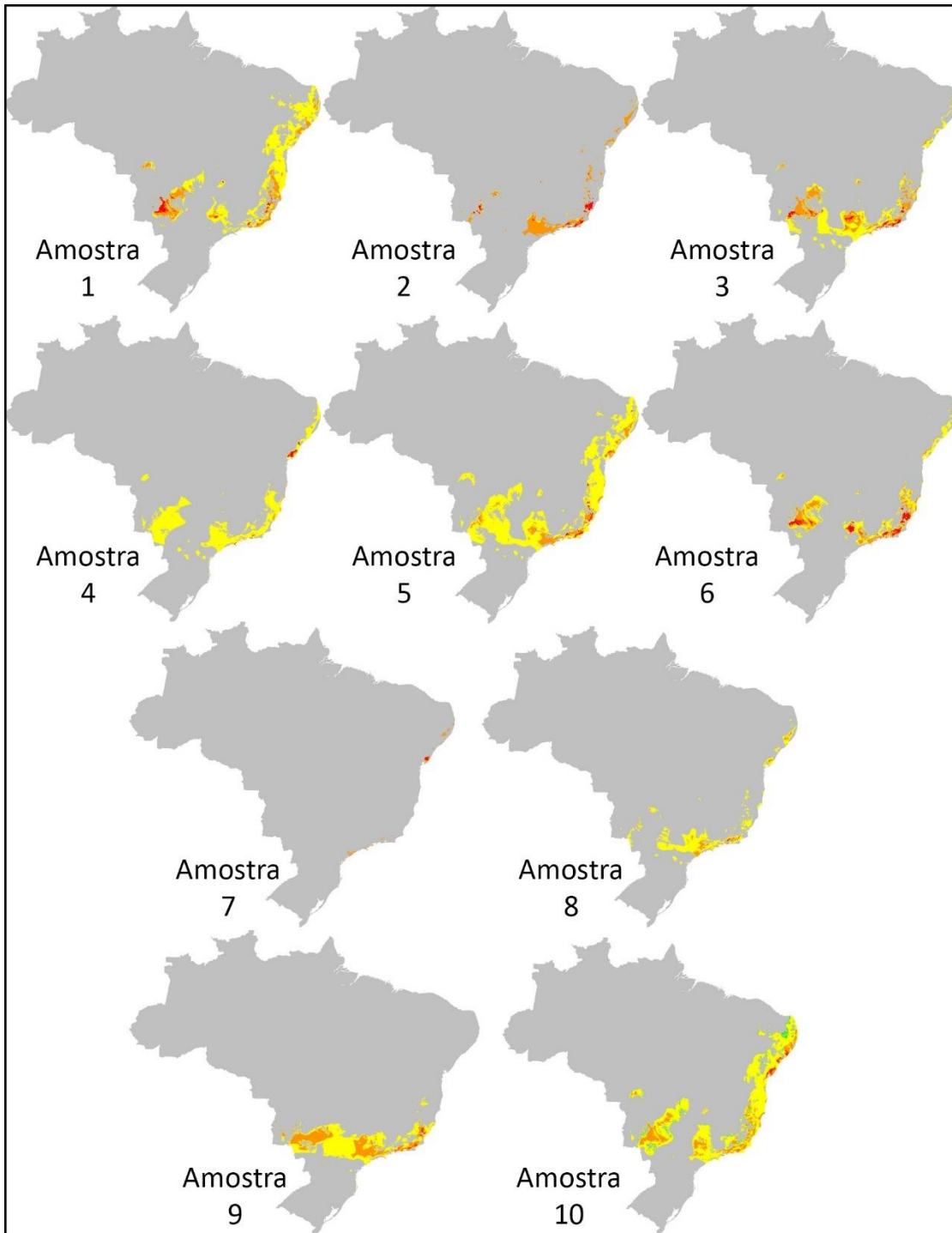


FIG 15. Modelagem da distribuição potencial de *Ramphocelus bresilius* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 20 pontos de presença, na proporção de pseudo ausência 25x1

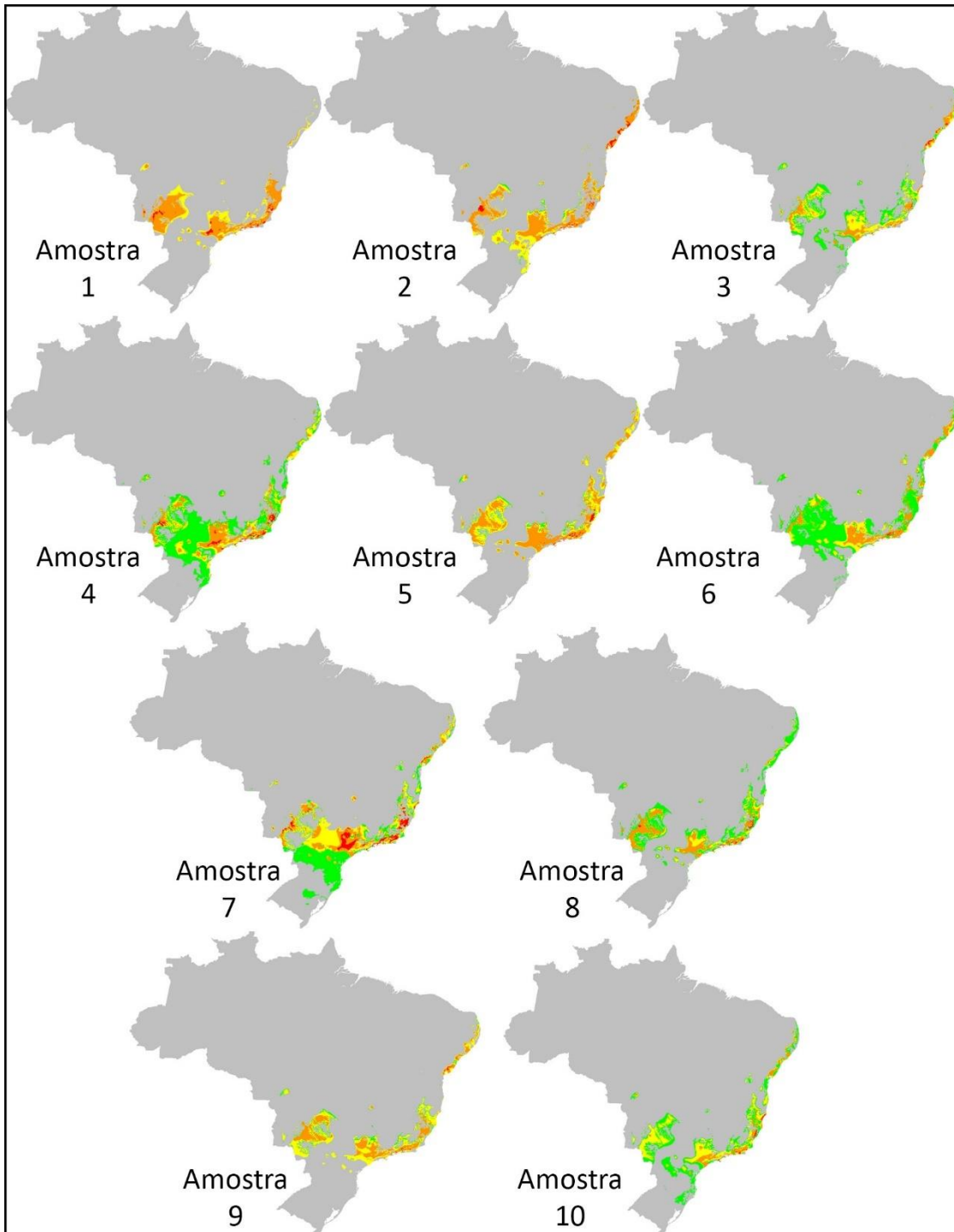


FIG 16. Modelagem da distribuição potencial de *Ramphocelus bresilius* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 50 pontos de presença, na proporção de pseudo ausência 25x1

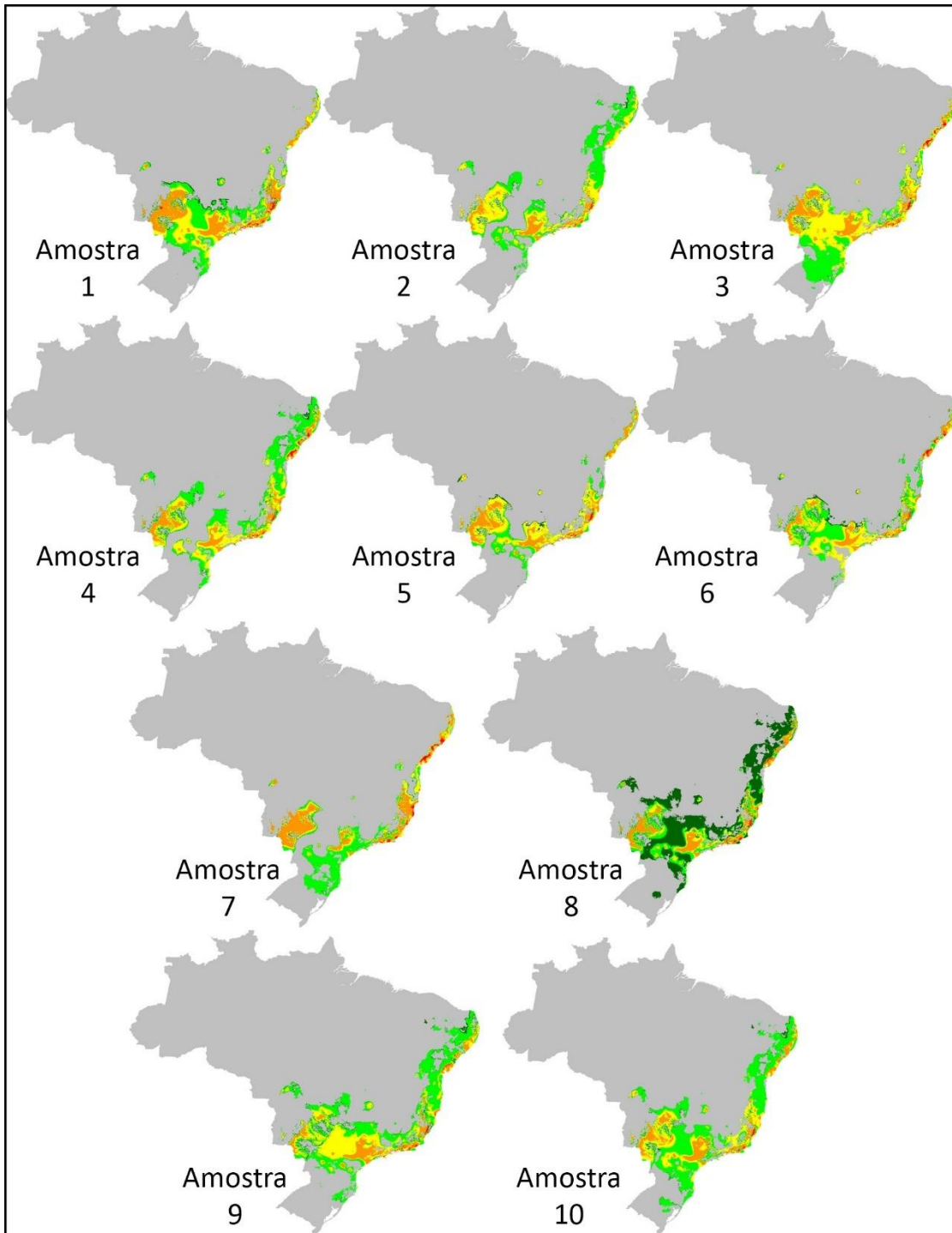


FIG 17. Modelagem da distribuição potencial de *Ramphocelus bresilius* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 150 pontos de presença, na proporção de pseudo ausência 25x1

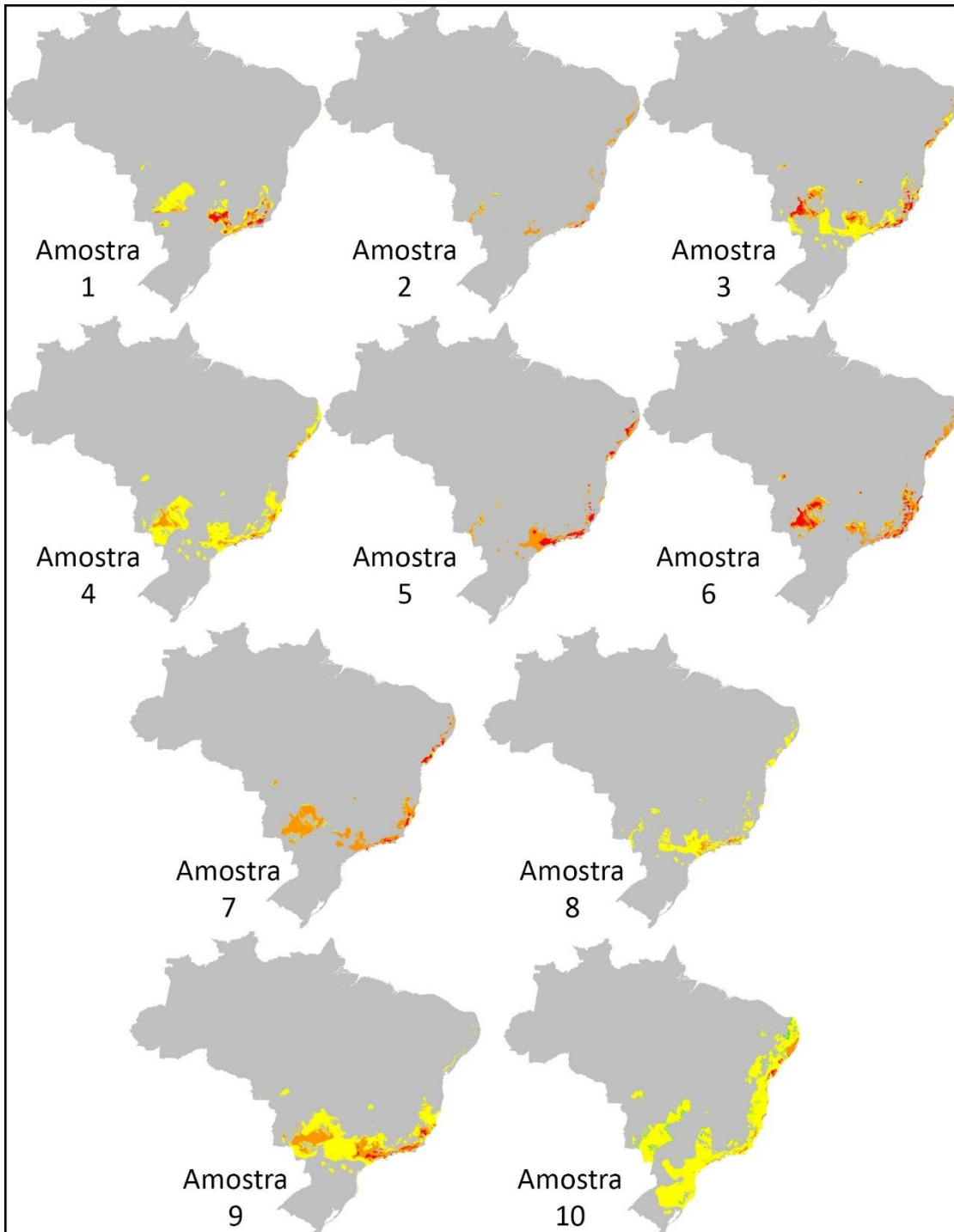


FIG 18. Modelagem da distribuição potencial de *Ramphocelus bresilius* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 20 pontos de presença, na proporção de pseudo ausência 50x1

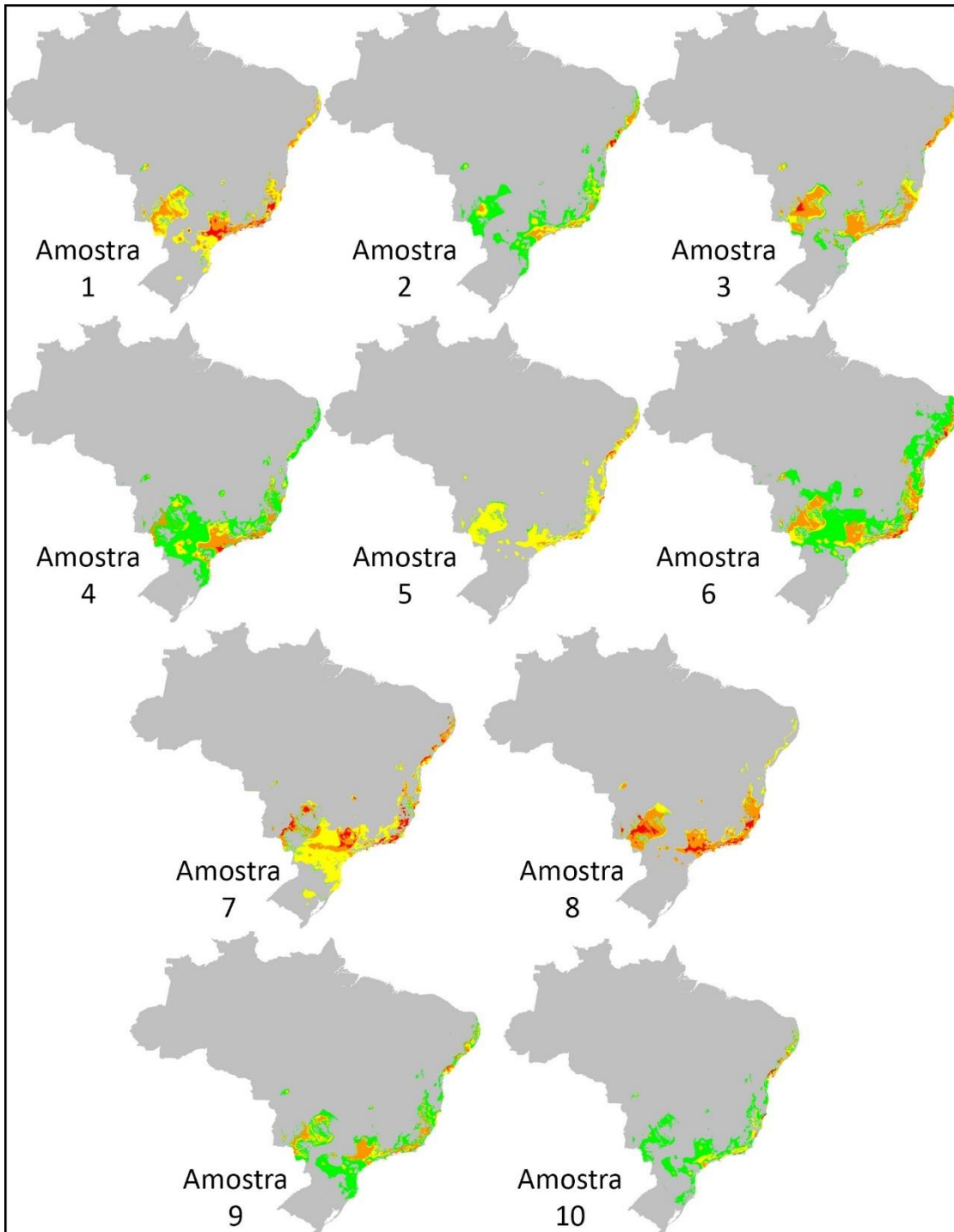


FIG 19. Modelagem da distribuição potencial de *Ramphocelus bresilius* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 50 pontos de presença, na proporção de pseudo ausência 50x1

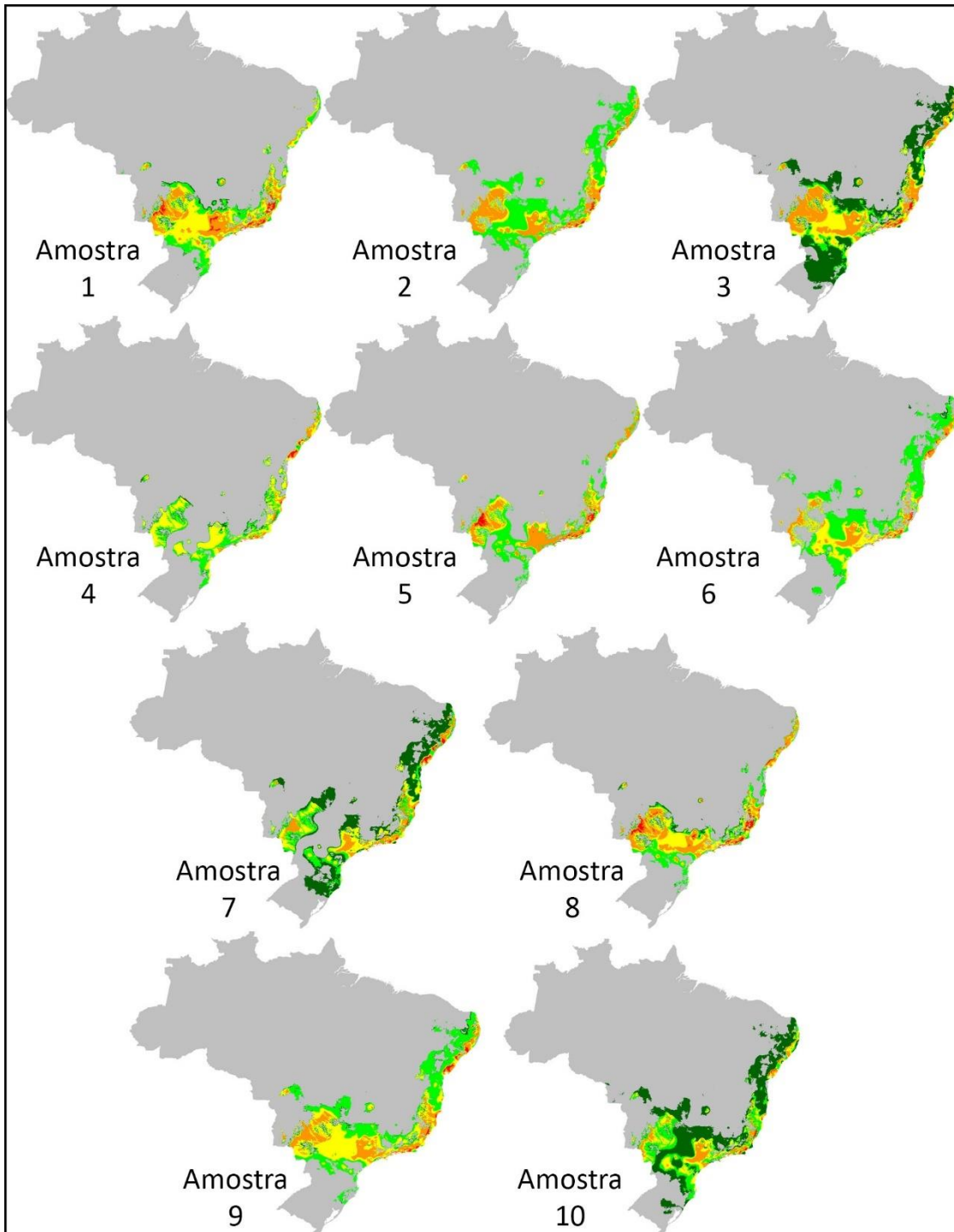


FIG 20. Modelagem da distribuição potencial de *Ramphocelus bresilius* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 150 pontos de presença, na proporção de pseudo ausência 50x1

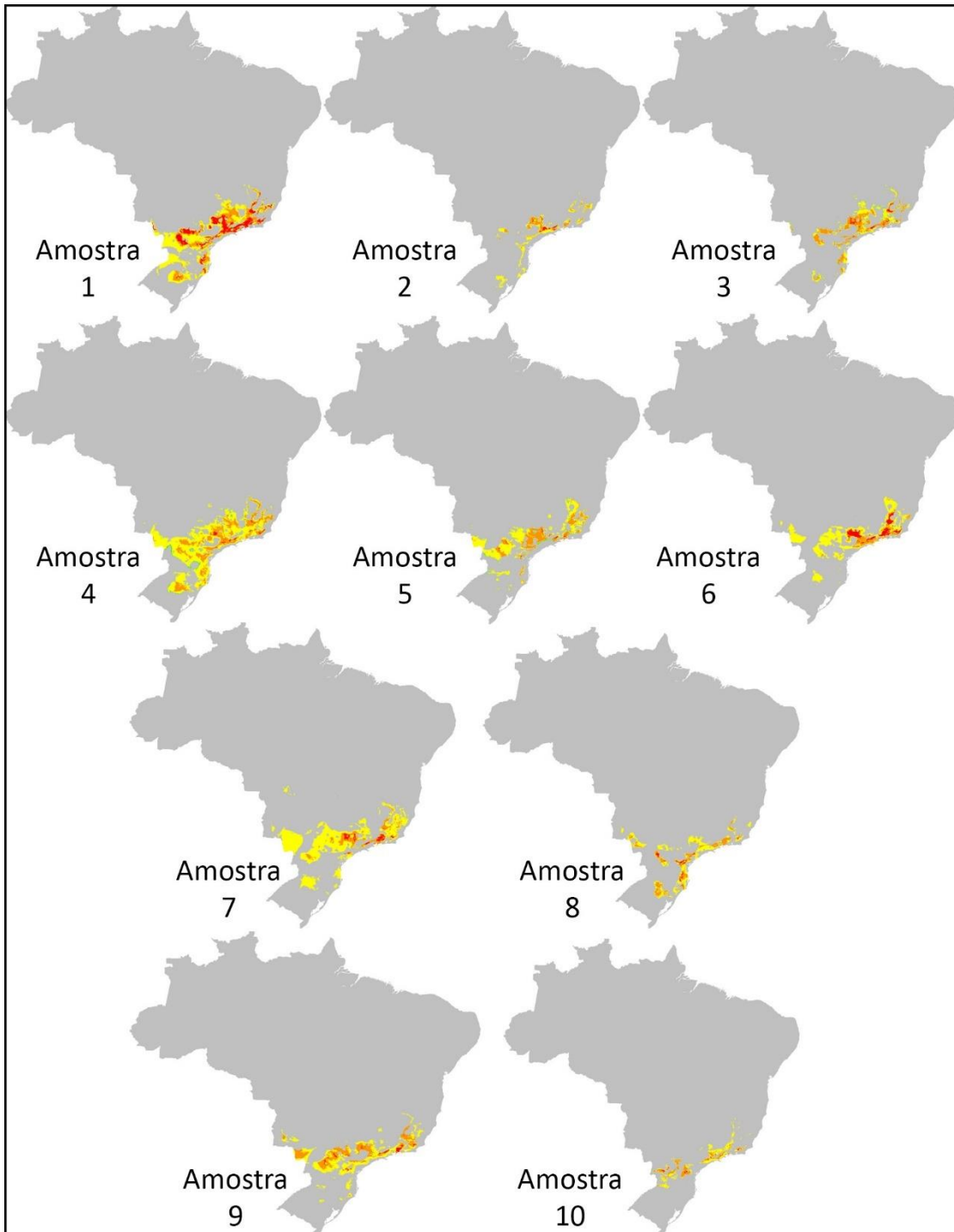


FIG 12. Modelagem da distribuição potencial de *Synallaxis ruficapilla* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 20 pontos de presença, na proporção de pseudo ausência 1x1

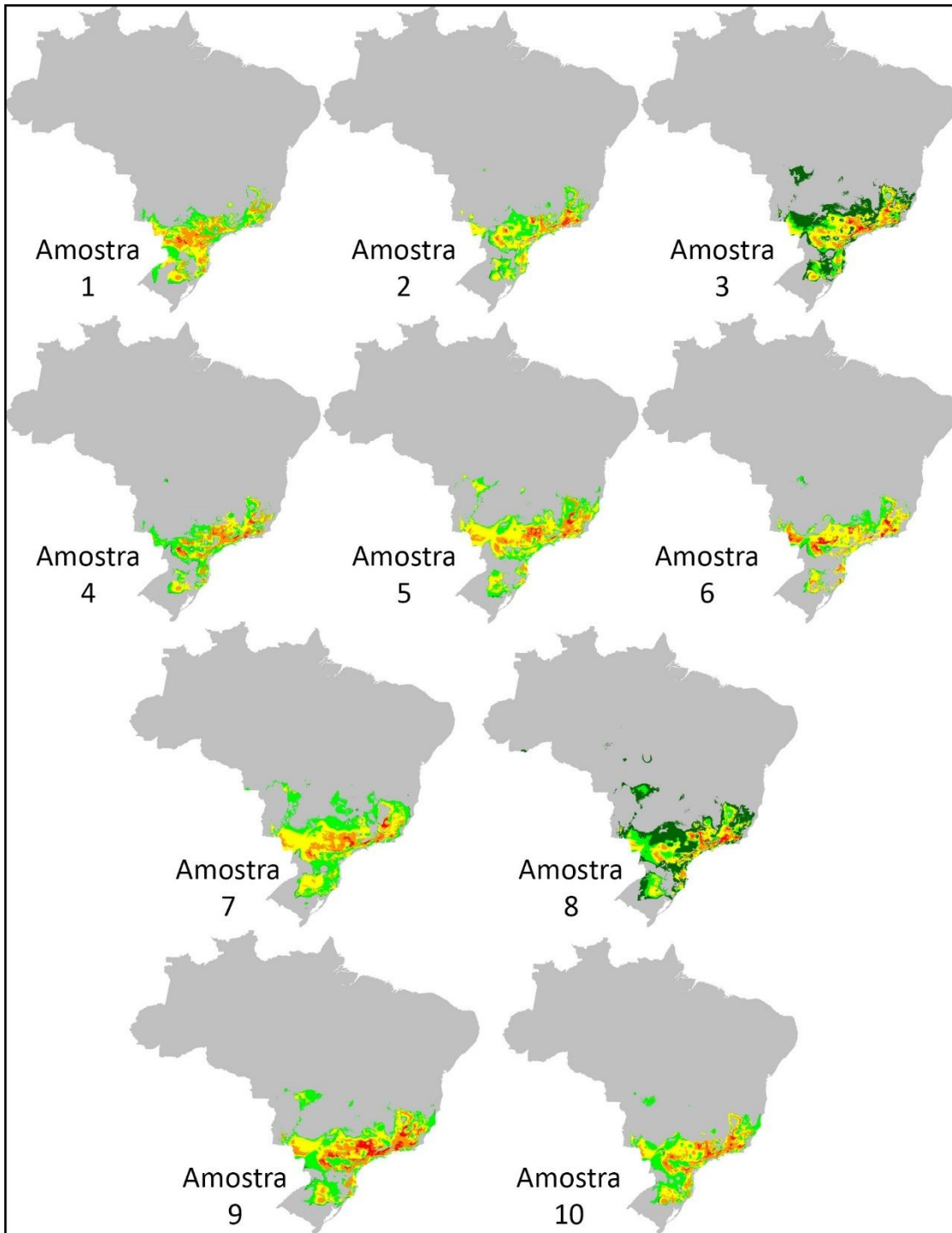


FIG 13. Modelagem da distribuição potencial de *Synallaxis ruficapilla* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 50 pontos de presença, na proporção de pseudo ausência 1x1

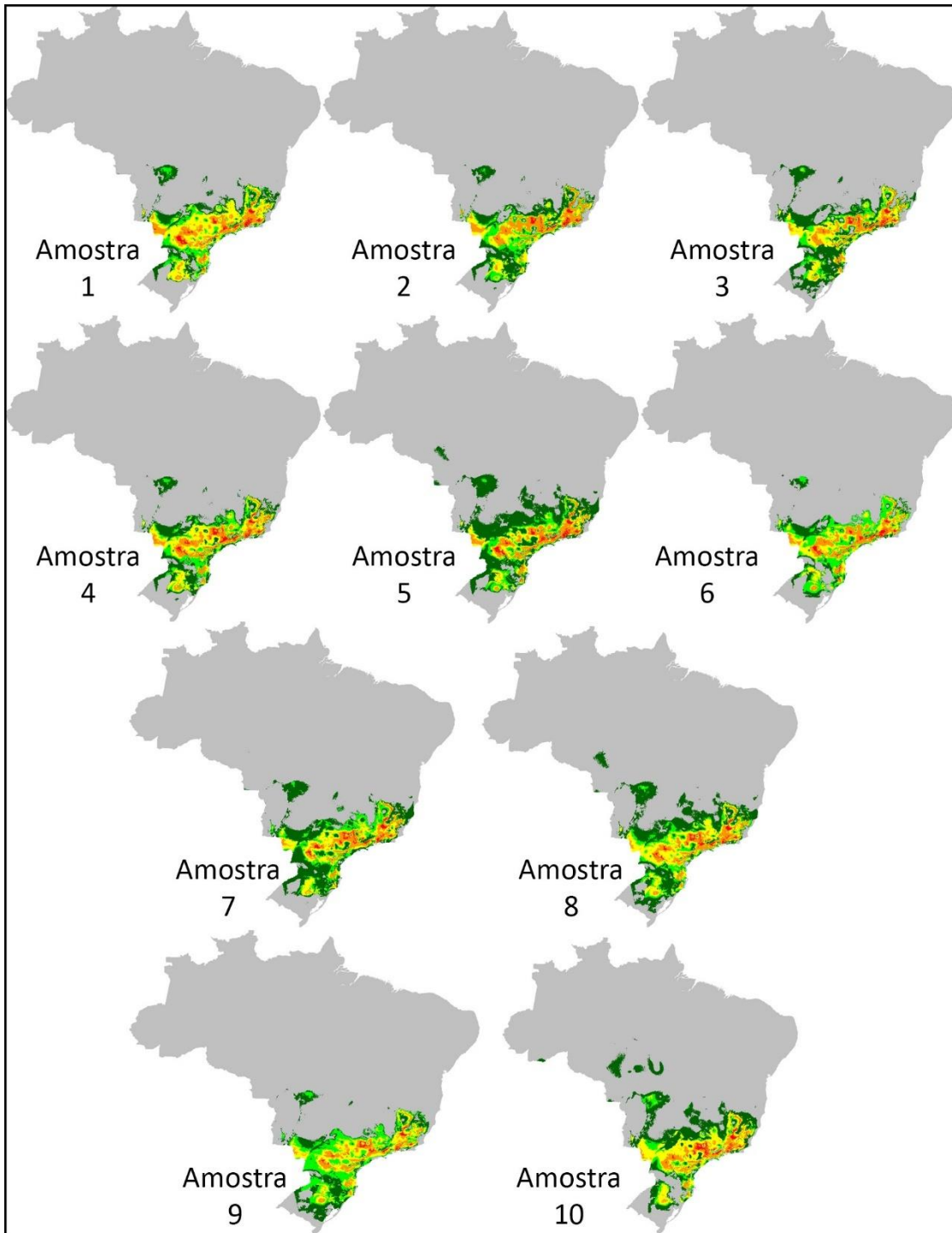


FIG 14. Modelagem da distribuição potencial de *Synallaxis ruficapilla* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 150 pontos de presença, na proporção de pseudo ausência 1x1

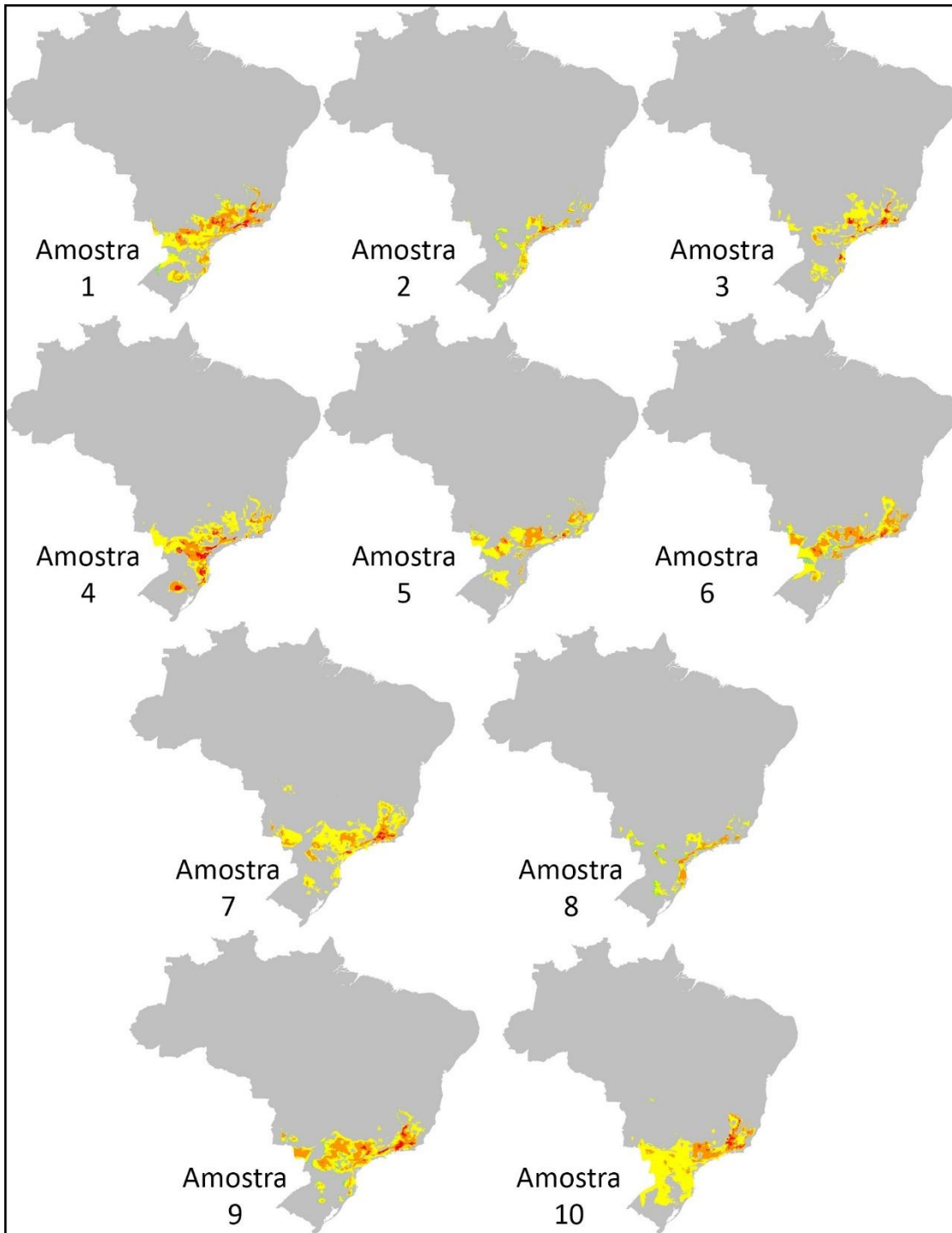


FIG 15. Modelagem da distribuição potencial de *Synallaxis ruficapilla* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 20 pontos de presença, na proporção de pseudo ausência 25x1

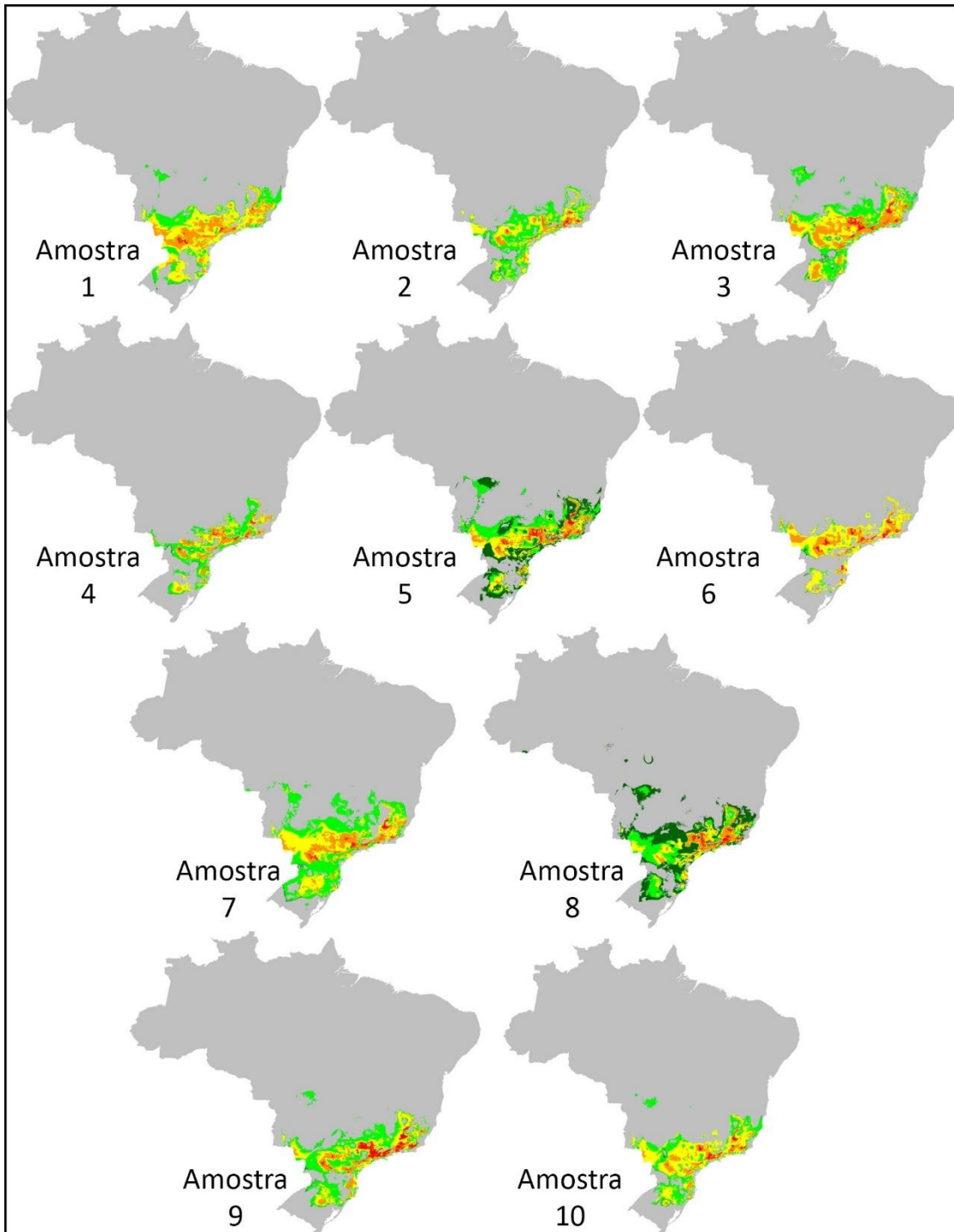


FIG 16. Modelagem da distribuição potencial de *Synallaxis ruficapilla* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 50 pontos de presença, na proporção de pseudo ausência 25x1

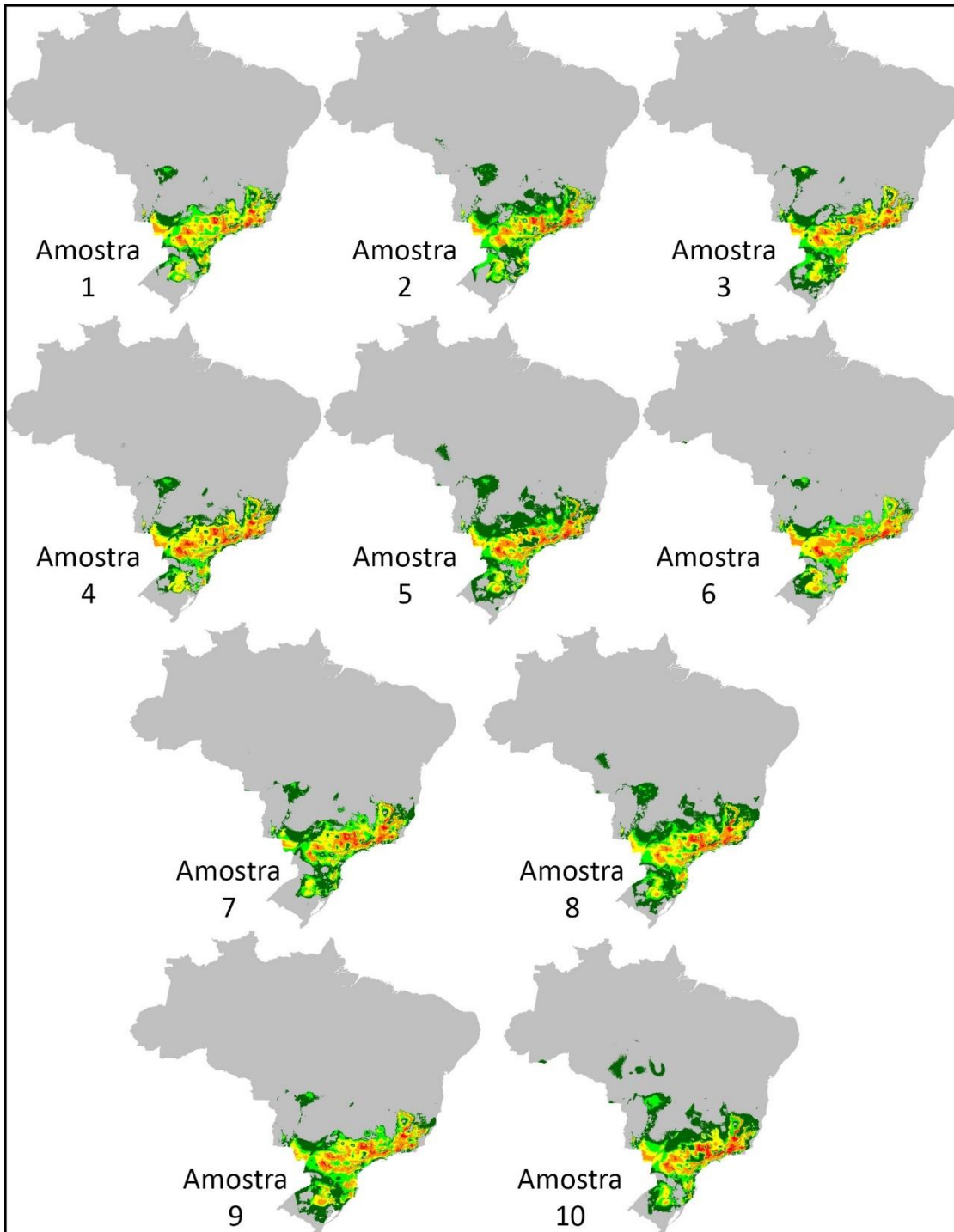


FIG 17. Modelagem da distribuição potencial de *Synallaxis ruficapilla* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 150 pontos de presença, na proporção de pseudo ausência 25x1

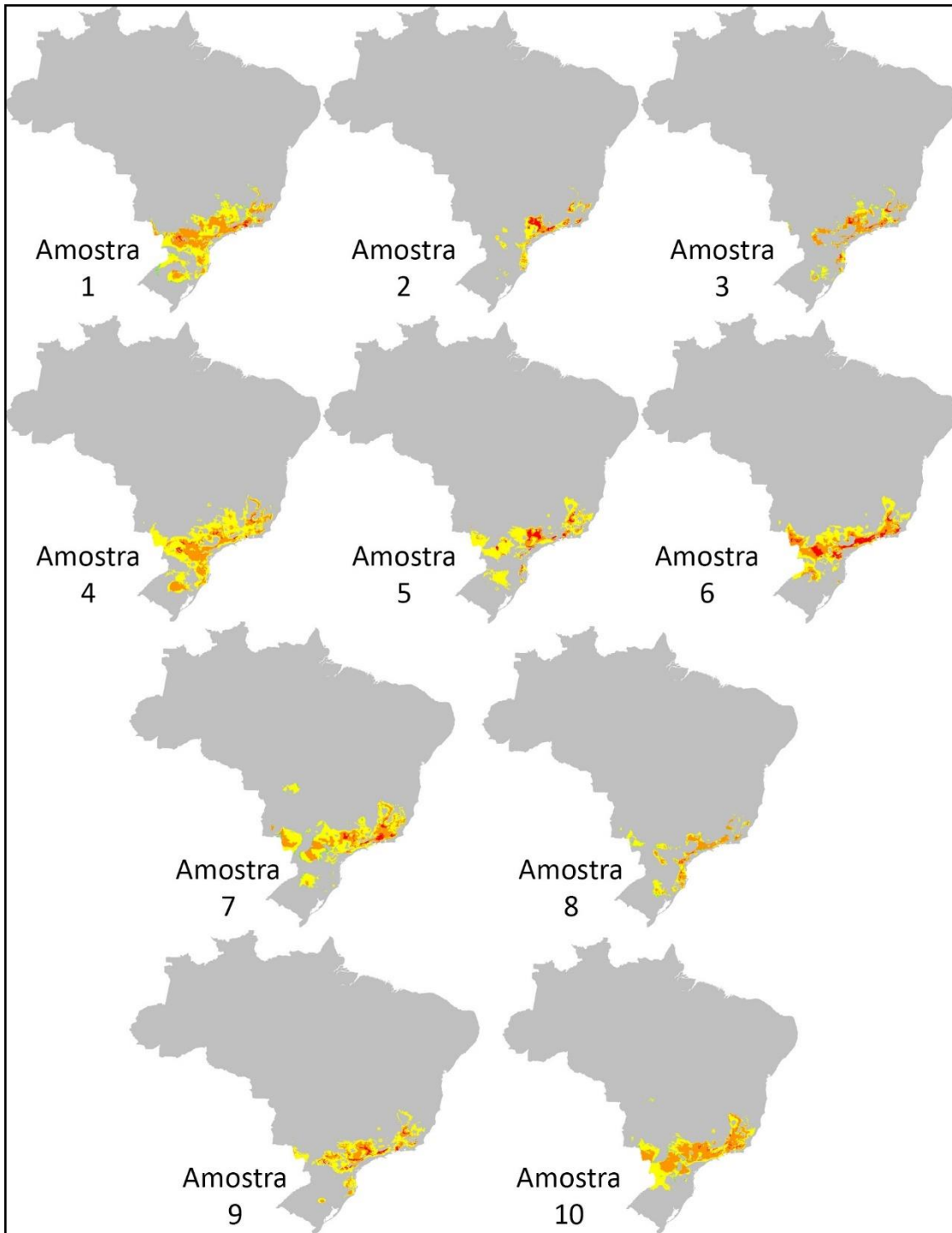


FIG 18. Modelagem da distribuição potencial de *Synallaxis ruficapilla* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 20 pontos de presença, na proporção de pseudo ausência 50x1

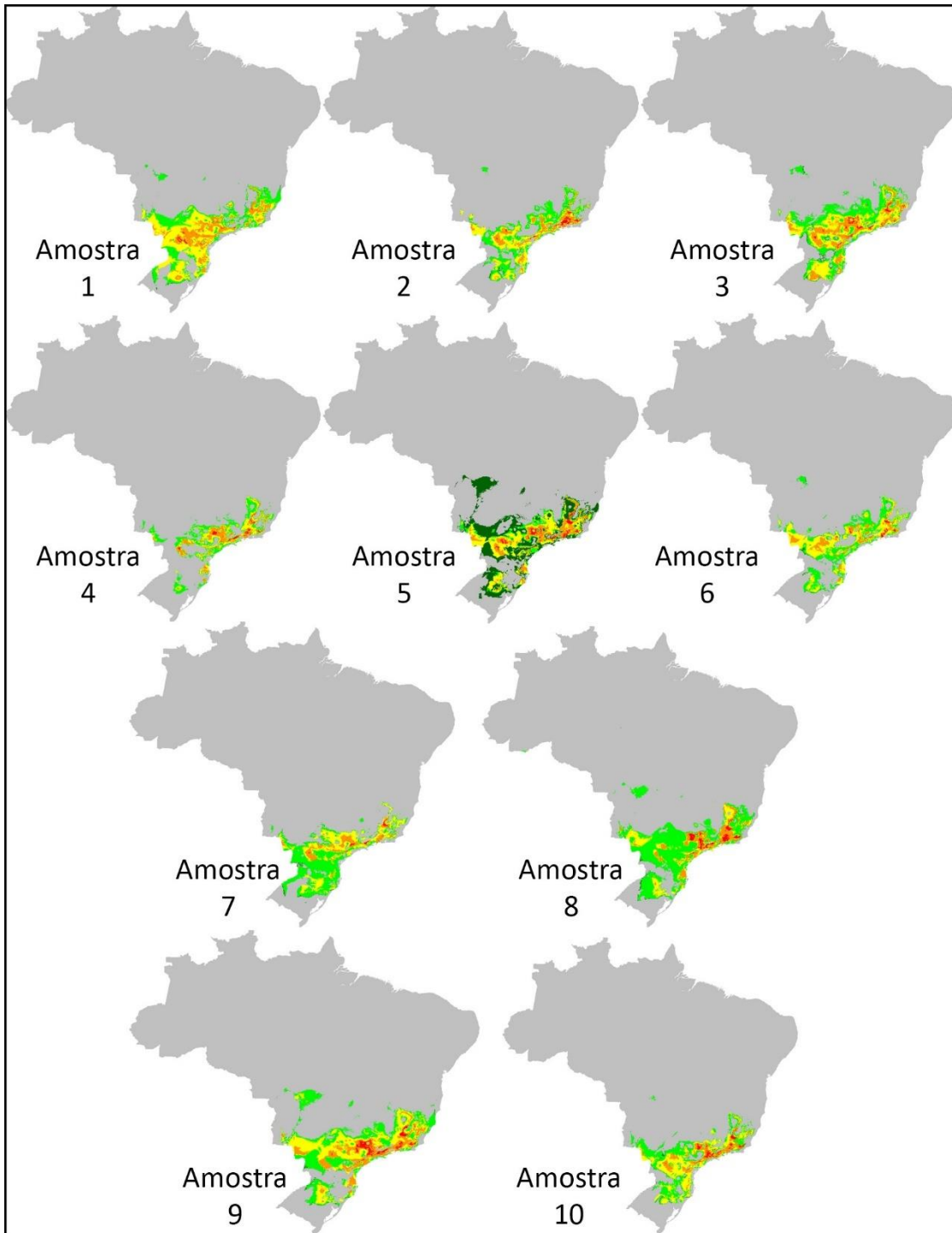


FIG 19. Modelagem da distribuição potencial de *Synallaxis ruficapilla* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 50 pontos de presença, na proporção de pseudo ausência 50x1

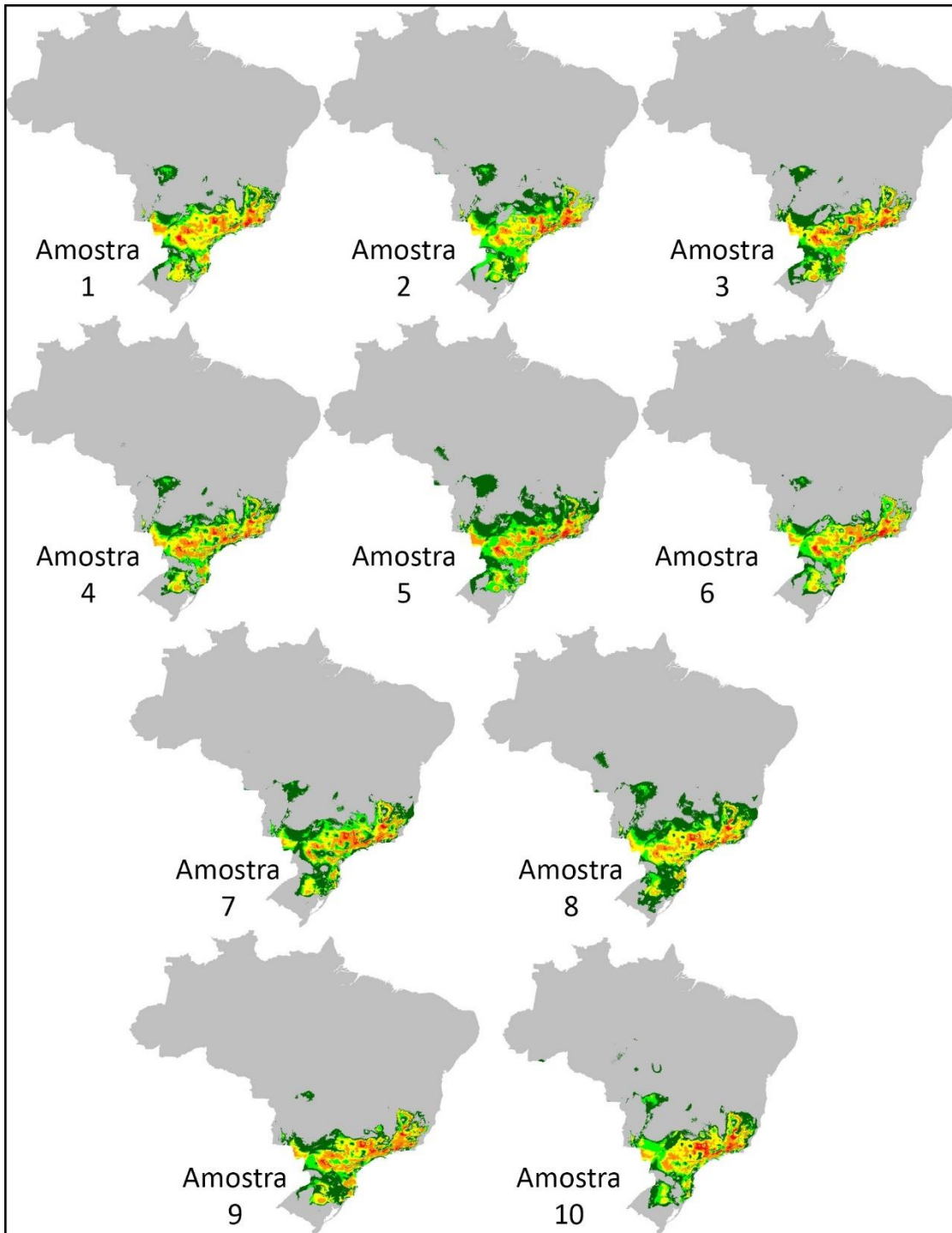


FIG 20. Modelagem da distribuição potencial de *Synallaxis ruficapilla* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 150 pontos de presença, na proporção de pseudo ausência 50x1

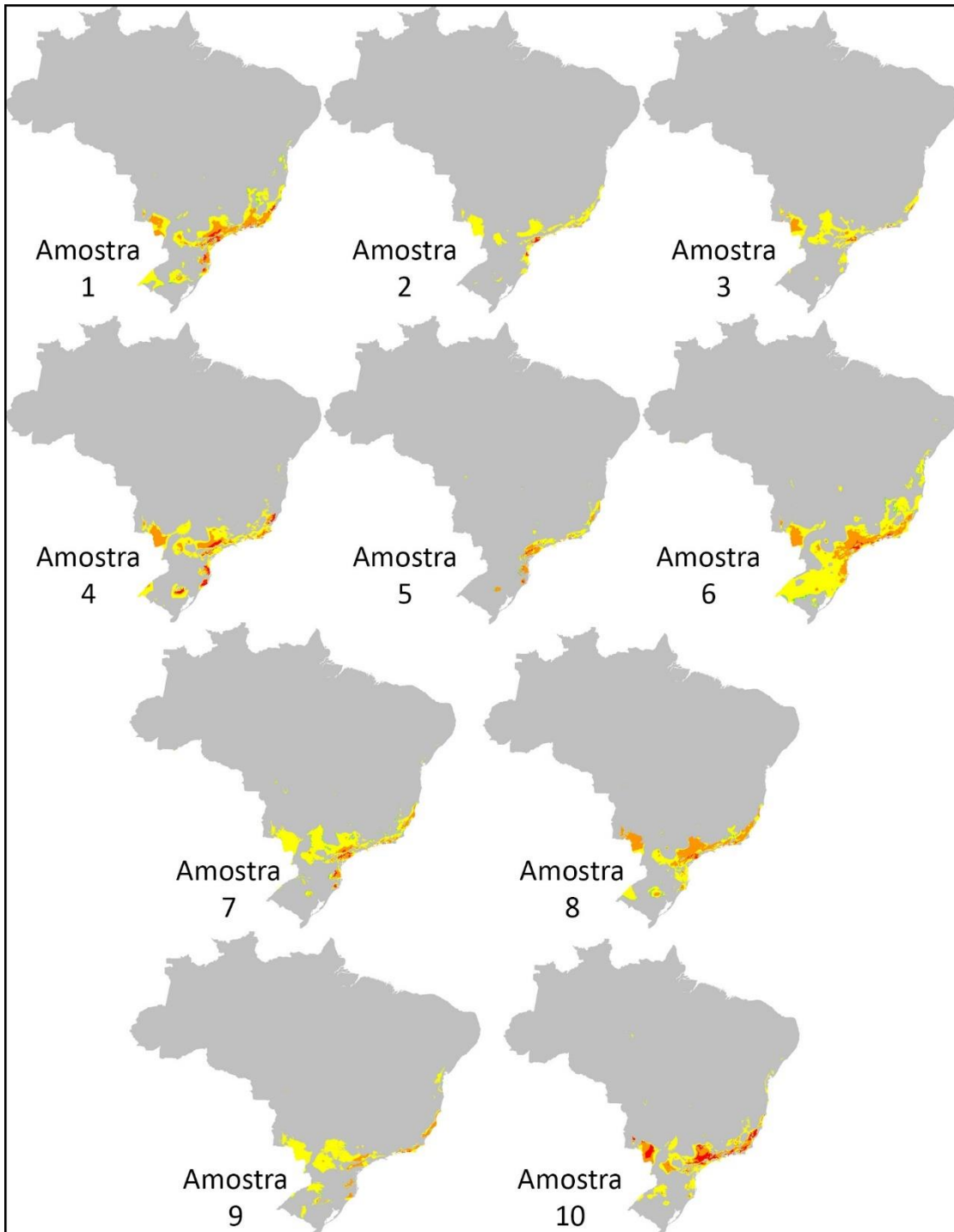


FIG 12. Modelagem da distribuição potencial de *Tangara seledon* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 20 pontos de presença, na proporção de pseudo ausência 1x1

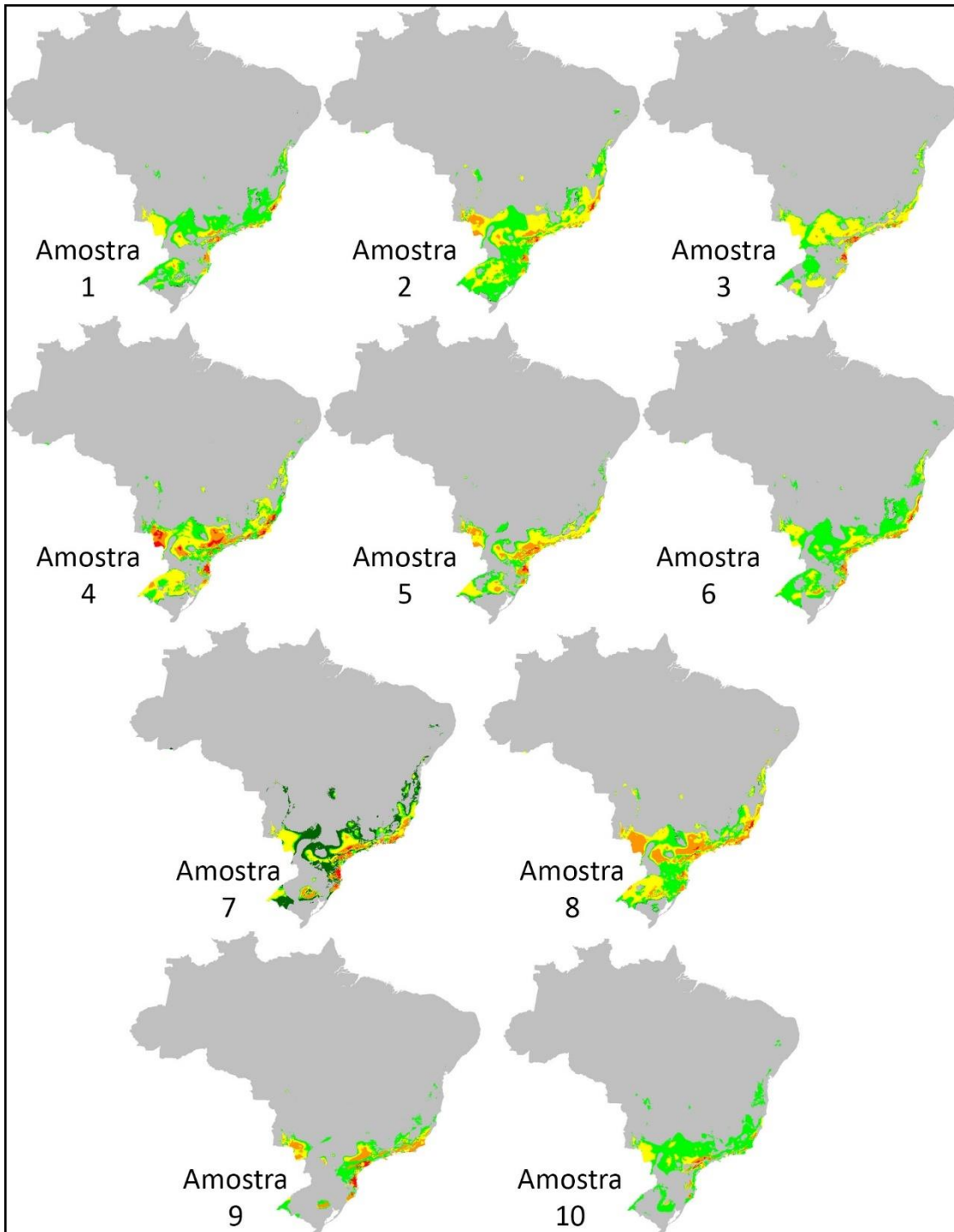


FIG 13. Modelagem da distribuição potencial de *Tangara seledon* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 50 pontos de presença, na proporção de pseudo ausência 1x1

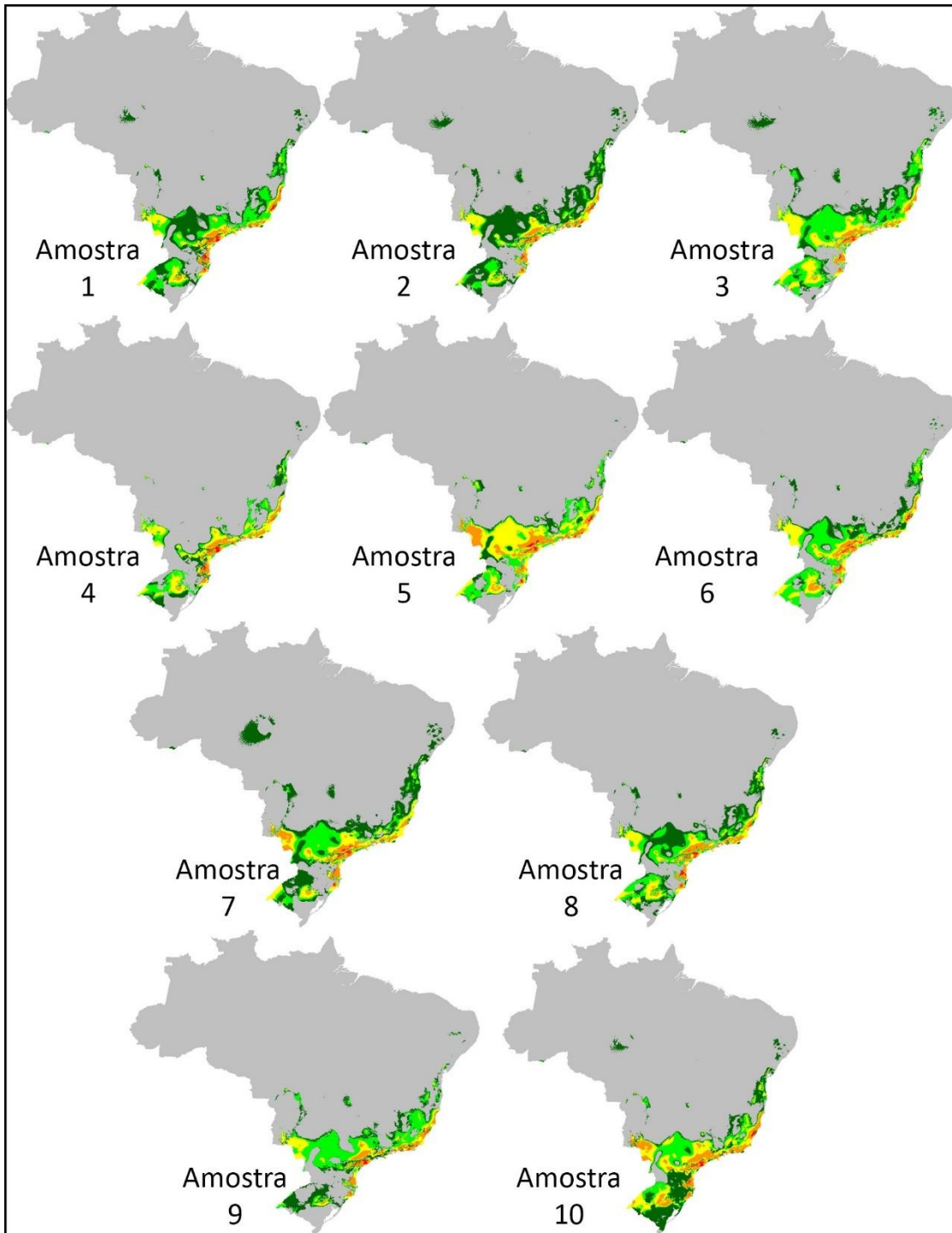


FIG 14. Modelagem da distribuição potencial de *Tangara seledon* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 150 pontos de presença, na proporção de pseudo ausência 1x1

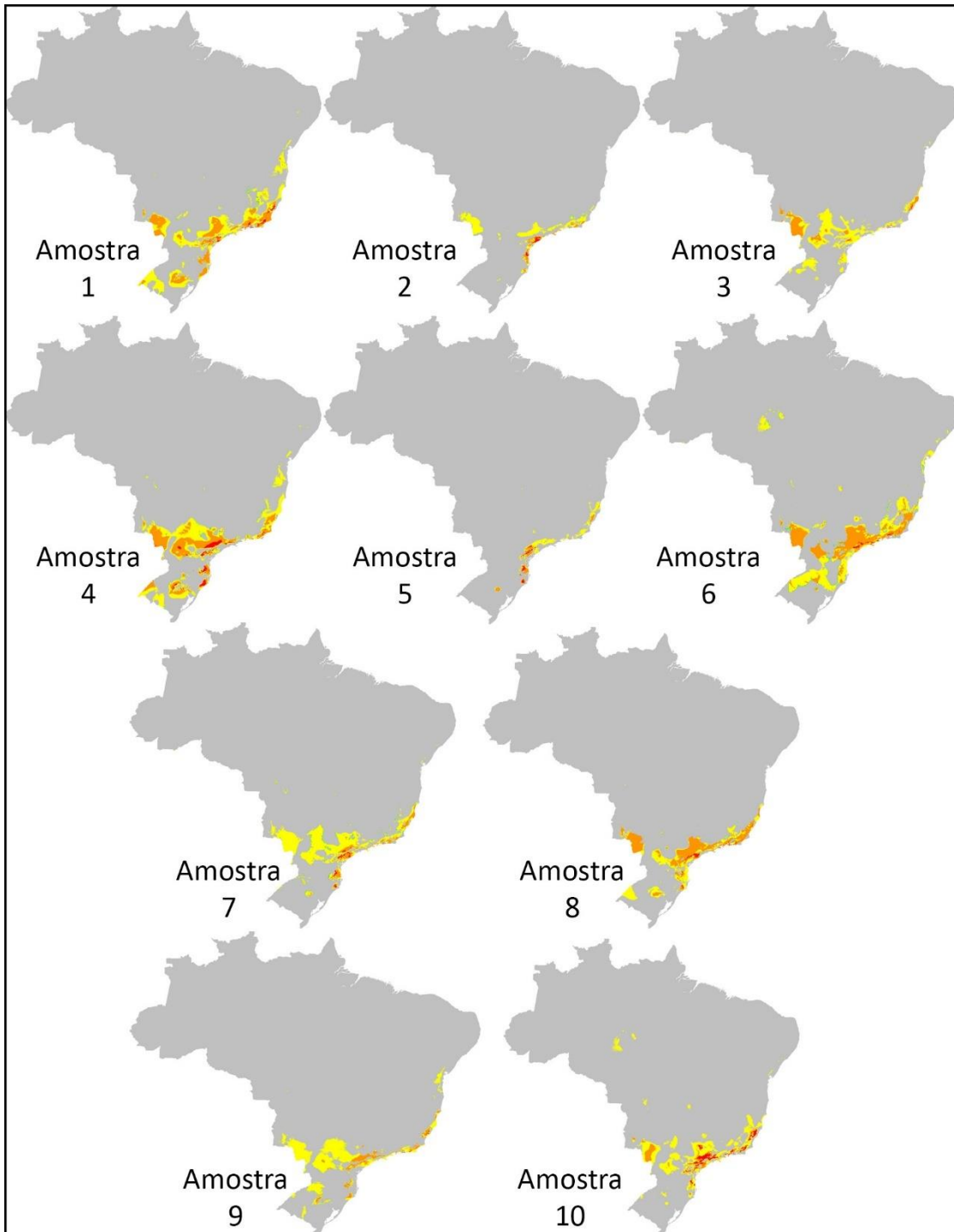


FIG 15. Modelagem da distribuição potencial de *Tangara seledon* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 20 pontos de presença, na proporção de pseudo ausência 25x1

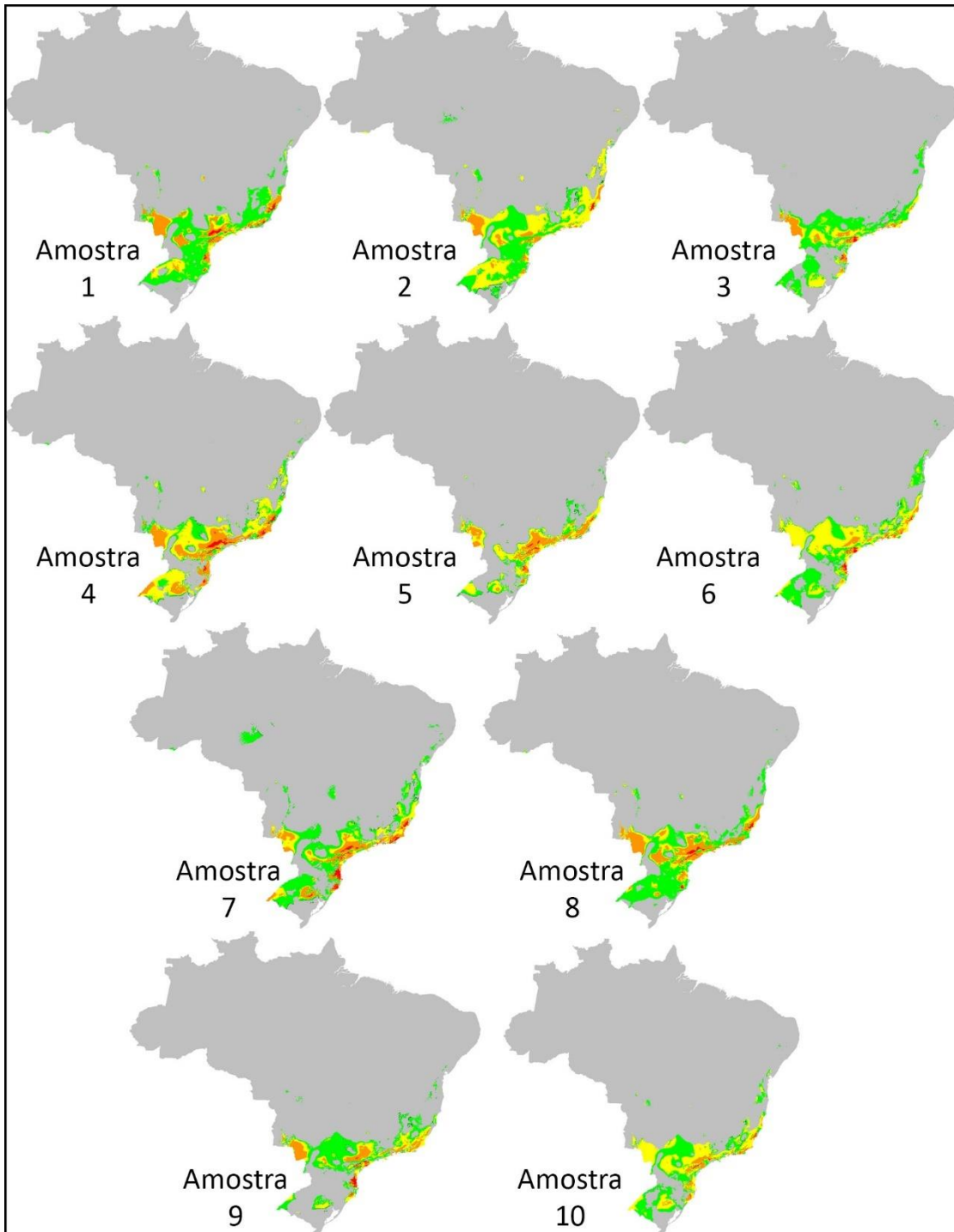


FIG 16. Modelagem da distribuição potencial de *Tangara seledon* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 50 pontos de presença, na proporção de pseudo ausência 25x1

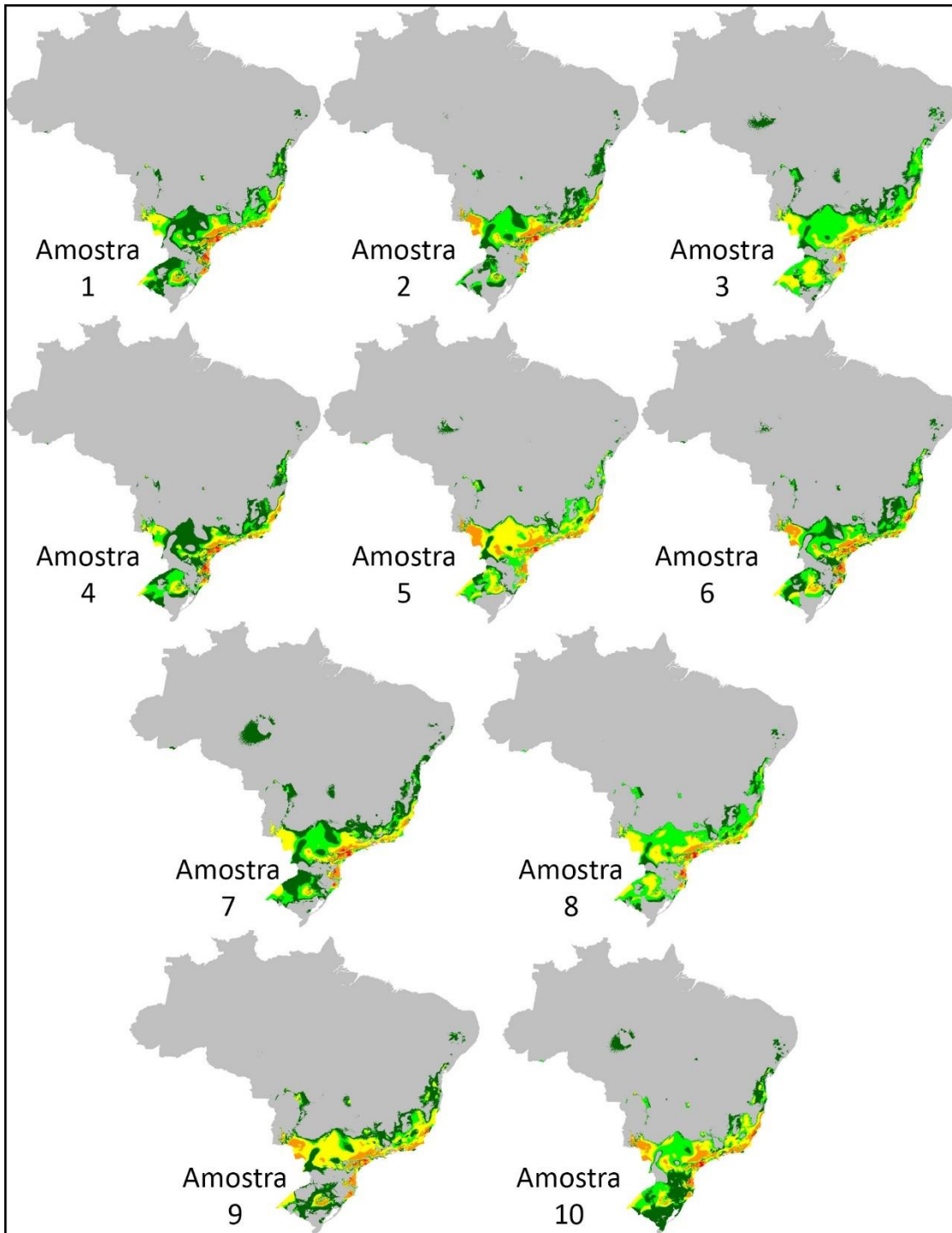


FIG 17. Modelagem da distribuição potencial de *Tangara seledon* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 150 pontos de presença, na proporção de pseudo ausência 25x1

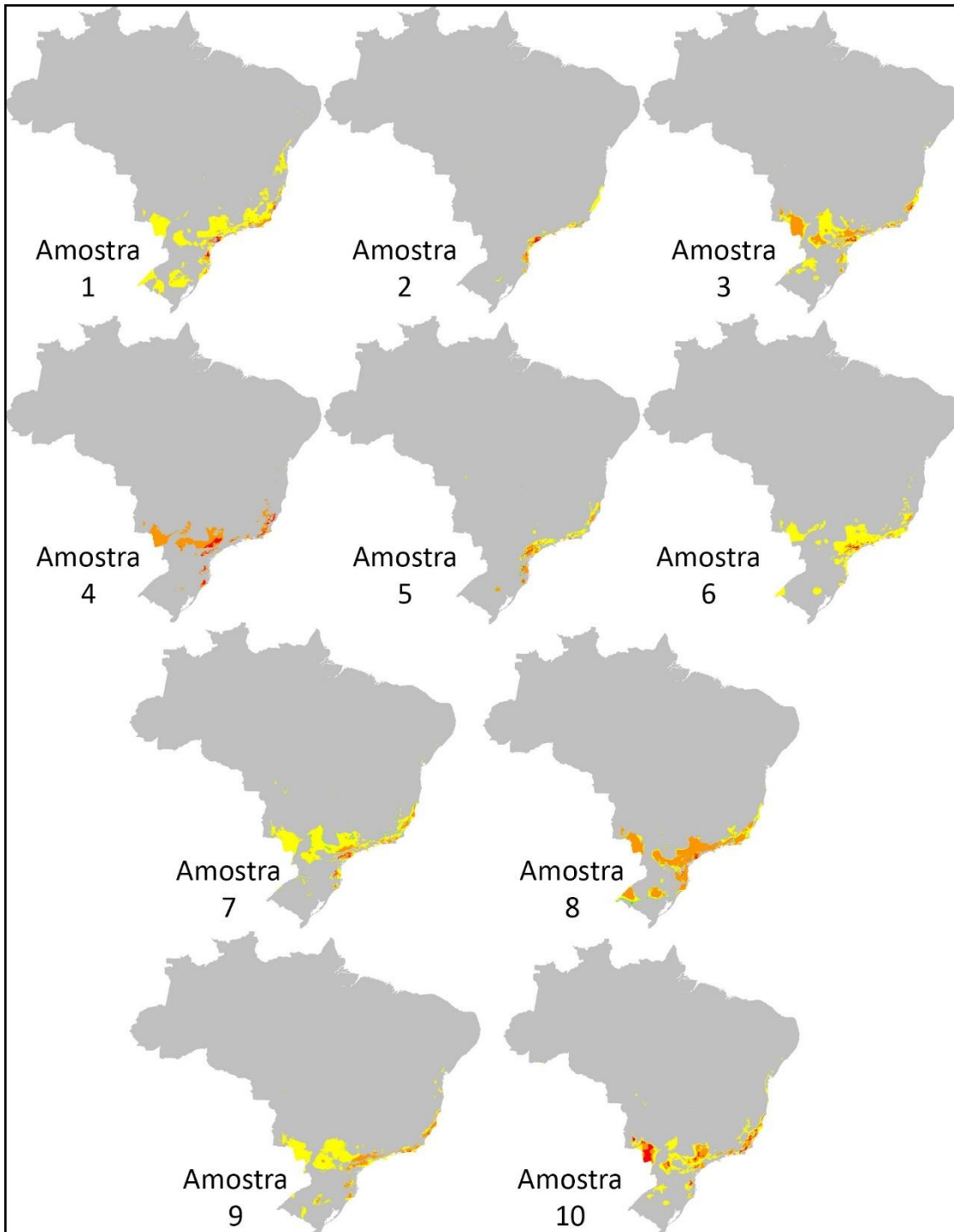


FIG 18. Modelagem da distribuição potencial de *Tangara seledon* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 20 pontos de presença, na proporção de pseudo ausência 50x1

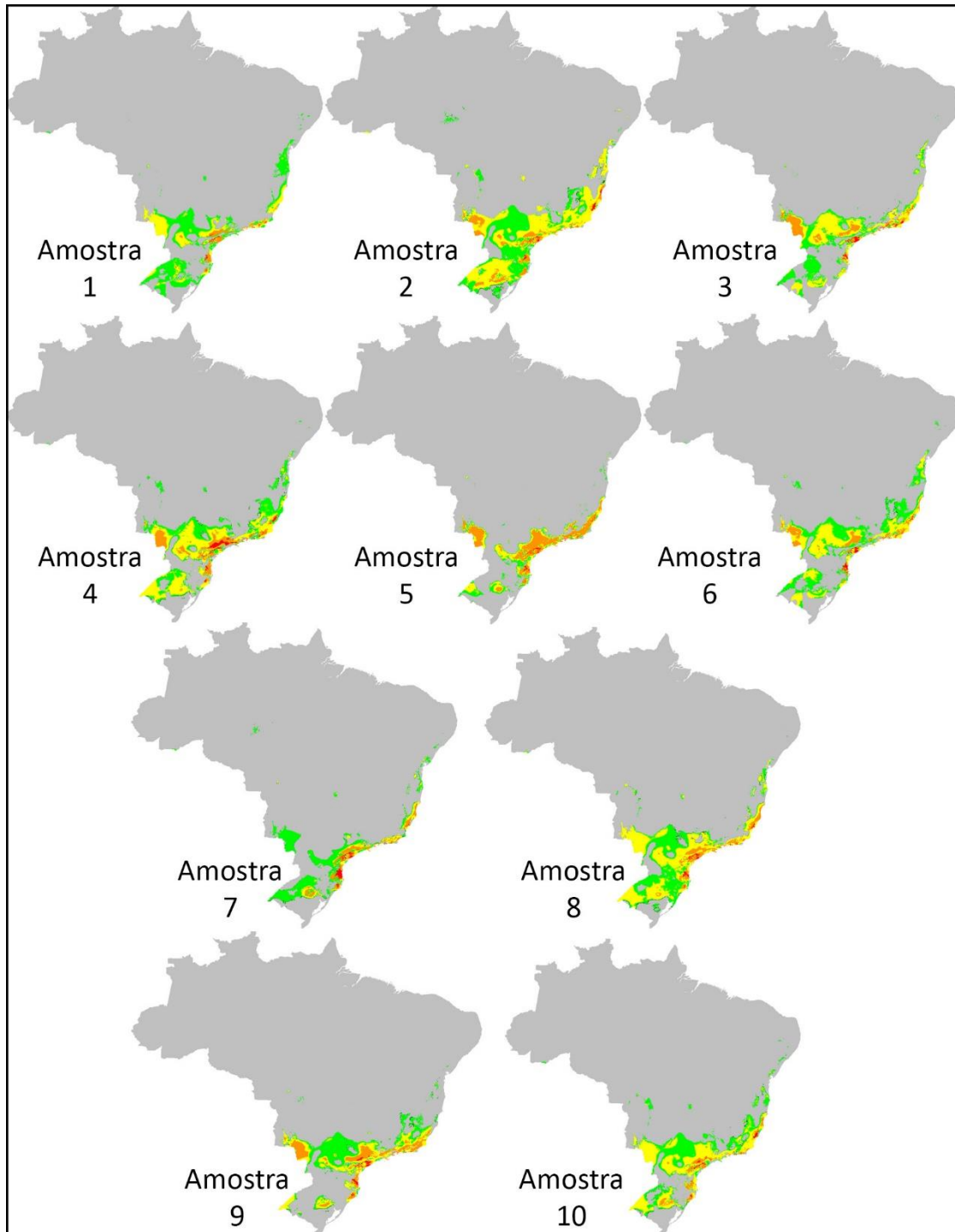


FIG 19. Modelagem da distribuição potencial de *Tangara seledon* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 50 pontos de presença, na proporção de pseudo ausência 50x1

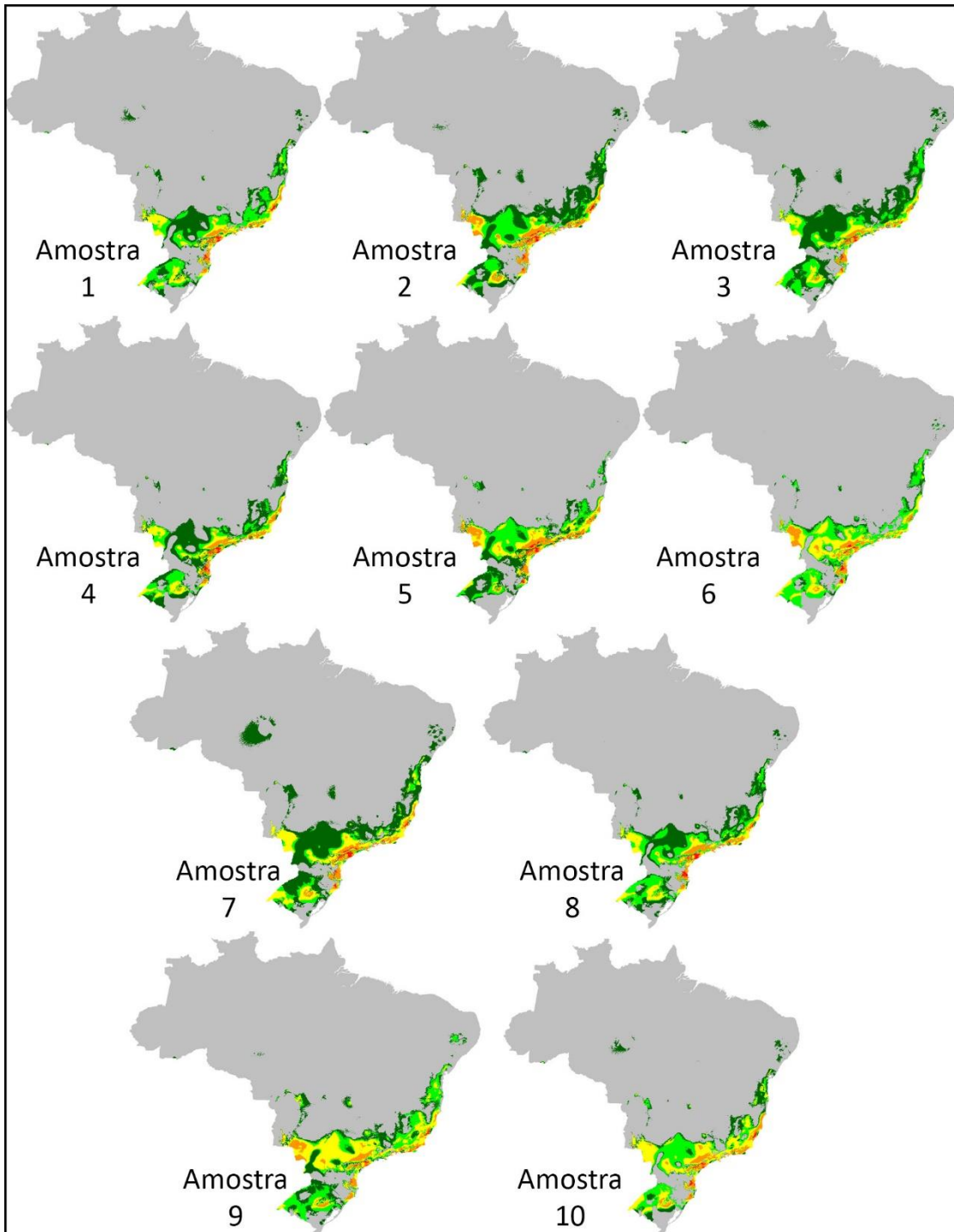


FIG 20. Modelagem da distribuição potencial de *Tangara seledon* a partir da reamostragem de pontos do conjunto N_T em 10 amostras aleatórias contendo 150 pontos de presença, na proporção de pseudo ausência 50x1

APÊNDICE C – Tabelas de valores AUC obtidos durante os processos de modelagem em diferentes tamanhos amostrais

Tabela 5 – Valores de AUC obtidos na avaliação dos modelos construídos a partir dos tamanhos amostrais N_T , N_2 , N_3 e N_4 , espécie *Chiroxiphia caudata*

	600	150	50	20
5	0,931	0,904	0,809	0,633
10	0,934	0,927	0,883	0,960
15	0,932	0,853	0,807	0,607
20	0,933	0,871	0,834	0,724
25	0,932	0,912	0,806	0,599
30	0,933	0,869	0,872	0,727
35	0,921	0,882	0,940	0,726
40	0,926	0,901	0,714	0,635
45	0,921	0,913	0,784	0,713
50	0,902	0,905	0,900	0,574

Tabela 6 – Valores de AUC obtidos na avaliação dos modelos construídos a partir dos tamanhos amostrais N_T , N_2 , N_3 e N_4 , espécie *Hemithraupis ruficapilla*

	250	150	50	20
5	0,931	0,966	0,719	0,795
10	0,886	0,942	0,770	0,839
15	0,910	0,930	0,836	0,978
20	0,944	0,829	0,792	0,651
25	0,901	0,911	0,826	0,662
30	0,927	0,944	0,771	0,776
35	0,935	0,901	0,865	0,788
40	0,901	0,871	0,887	0,851
45	0,937	0,908	0,806	0,794
50	0,897	0,847	0,821	0,611

Tabela 7 – Valores de AUC obtidos na avaliação dos modelos construídos a partir dos tamanhos amostrais N_T , N_2 , N_3 e N_4 , espécie *Myiothlypis leucoblephara*

	300	150	50	20
5	0,913	0,883	0,775	0,465
10	0,919	0,851	0,778	0,589
15	0,927	0,920	0,966	0,626
20	0,909	0,911	0,813	0,590
25	0,916	0,932	0,781	0,787
30	0,890	0,842	0,830	0,471
35	0,911	0,903	0,917	0,482
40	0,891	0,899	0,688	0,714

45	0,918	0,936	0,855	0,625
50	0,931	0,927	0,820	0,875

Tabela 8 – Valores de AUC obtidos na avaliação dos modelos construídos a partir dos tamanhos amostrais N_T , N_2 , N_3 e N_4 , espécie *Ramphocelus bresilius*

	300	150	50	20
5	0,882	0,875	0,754	0,870
10	0,921	0,878	0,765	0,743
15	0,876	0,897	0,945	0,725
20	0,852	0,890	0,937	0,769
25	0,902	0,868	0,682	0,718
30	0,899	0,914	0,867	0,896
35	0,884	0,923	0,885	0,716
40	0,913	0,827	0,765	0,977
45	0,913	0,884	0,724	0,725
50	0,928	0,853	0,835	0,562

Tabela 9 – Valores de AUC obtidos na avaliação dos modelos construídos a partir dos tamanhos amostrais N_T , N_2 , N_3 e N_4 , espécie *Synallaxis ruficapilla*

	400	150	50	20
5	0,940	0,886	0,831	0,669
10	0,941	0,889	0,898	0,664
15	0,948	0,920	0,718	0,988
20	0,924	0,926	0,753	0,611
25	0,955	0,841	0,846	0,893
30	0,928	0,913	0,719	0,940
35	0,926	0,944	0,874	0,732
40	0,946	0,866	0,911	0,610
45	0,937	0,914	0,838	0,807
50	0,936	0,879	0,761	0,497

Tabela 10 – Valores de AUC obtidos na avaliação dos modelos construídos a partir dos tamanhos amostrais N_T , N_2 , N_3 e N_4 , espécie *Tangara seledon*

	300	150	50	20
5	0,946	0,945	0,973	0,788
10	0,921	0,917	0,922	0,639
15	0,947	0,942	0,928	0,576
20	0,906	0,902	0,867	0,608
25	0,941	0,900	0,756	0,847
30	0,911	0,947	0,967	0,698
35	0,939	0,936	0,743	0,691
40	0,933	0,971	0,934	0,652

45	0,944	0,946	0,904	0,567
50	0,932	0,839	0,857	0,732

APÊNDICE D – Tabelas de valores AUC obtidos durante os processos de modelagem para os conjuntos N2, N3 e N4 após reamostragem, nas proporções de pseudo ausência de 1x1, 25x1 e 50x1

Tabela 11 – Valores de AUC obtidos na avaliação dos modelos construídos a partir de amostras aleatórias, nos tamanhos amostrais N₂, N₃ e N₄ para as espécies *Chiroxiphia caudata*, *Hemithraupis ruficapilla* e *Myiothlypis leucoblephara*

PA	TA	CHI_CAU		HEM_RUF		MYI_LEU	
		Média	DP	Média	DP	Média	DP
1:1	20	0,712	0,17	0,598	0,10	0,690	0,14
	50	0,875	0,06	0,867	0,05	0,854	0,06
	150	0,896	0,03	0,897	0,02	0,888	0,02
25:1	20	0,722	0,13	0,698	0,14	0,598	0,11
	50	0,809	0,09	0,786	0,06	0,852	0,07
	150	0,891	0,04	0,876	0,04	0,885	0,03
50:1	20	0,772	0,10	0,685	0,15	0,666	0,17
	50	0,777	0,04	0,847	0,05	0,786	0,05
	150	0,845	0,08	0,827	0,06	0,823	0,03

TA – Tamanho Amostral PA – Pseudo ausência DP – Desvio-padrão

CHI_CAU - *Chiroxiphia caudata* HEM_RUF – *Hemithraupis ruficapilla* MYI_LEU – *Myiothlypis leucoblephara*

Tabela 12 – Valores de AUC obtidos na avaliação dos modelos construídos a partir de amostras aleatórias, nos tamanhos amostrais N₂, N₃ e N₄ para as espécies *Ramphocelus bresilius*, *Synallaxis ruficapilla* e *Tangara seledon*

PA	TA	RAM_BRE		SYN_RUF		TAN_SEL	
		Média	DP	Média	DP	Média	DP
1:1	20	0,67	0,12	0,59	0,13	0,73	0,15
	50	0,74	0,07	0,89	0,05	0,82	0,12
	150	0,79	0,09	0,90	0,03	0,75	0,05
25:1	20	0,59	0,08	0,77	0,11	0,72	0,12
	50	0,74	0,09	0,88	0,05	0,84	0,05
	150	0,79	0,06	0,91	0,03	0,91	0,04
50:1	20	0,65	0,16	0,77	0,16	0,71	0,14
	50	0,76	0,09	0,83	0,06	0,80	0,05
	150	0,83	0,06	0,84	0,06	0,85	0,07

TA – Tamanho Amostral PA – Pseudo ausência DP – Desvio-padrão

RAM_BRE – *Ramphocelus bresilius* SYN_RUF – *Synallaxis ruficapilla* TAN_SEL – *Tangara seledon*