

Uso de modelagem de nicho ecológico em conservação de espécies de Passifloraceae no Estado de São Paulo

Renato De Giovanni

Trabalho final apresentado ao Programa de Mestrado Profissional em Conservação da Biodiversidade e Desenvolvimento Sustentável do Instituto de Pesquisas Ecológicas como parte dos requisitos para a obtenção do título de Mestre em Ecologia.

Orientadora: Dr^a. Flávia de Souza Rocha

Co-orientadora: Dr^a. Marinez Ferreira de Siqueira

Co-orientador: Dr. Luís Carlos Bernacci

Nazaré Paulista – SP

2011

Agradecimentos

À minha esposa, Rosely, que acompanhou como ninguém todo meu esforço, soube compreender as dificuldades e me apoiou durante todo o trabalho. É justamente nos momentos mais difíceis que o verdadeiro amor se prova!

A toda minha família, que sempre me incentivou e com quem não pude estar tão presente quanto gostaria nos últimos anos. Espero que tenham compreendido.

Aos meus orientadores, Flávia, Marinez e Luís, de quem recebi inúmeros conselhos, críticas e ideias que foram fundamentais para construção deste trabalho.

Aos colegas Flávia Santos Pinto e A. Townsend Peterson por todos os comentários e sugestões.

A todos da escola – professores, alunos e funcionários – fonte preciosa de convívio e aprendizado. Em particular à Ana, Natividad e Rafael, amigos que tive a oportunidade de conhecer melhor. Bons momentos passamos juntos!

À ESCAS, cuja bolsa proporcionou enorme ajuda para realização do curso.

Ao CRIA, Centro de Referência em Informação Ambiental, que ofereceu infra-estrutura computacional para geração dos modelos e que acreditou no potencial deste trabalho.

Aos colegas do IAC, Instituto Agrônomo de Campinas, onde tive a oportunidade única de conhecer de perto um herbário e participar de suas atividades.

Índice

Resumo.....	6
Introdução Geral.....	7
Objetivo	8
Objetivos específicos.....	9
Referências	9
PARTE I – Towards a comprehensive workflow for generating ecological niche models with herbarium data: a case study with Passifloraceae in Brazil	12
Abstract	12
Introduction.....	12
Material and methods	15
Species names.....	16
Environmental data	17
Specimen data	18
Modelling strategy	21
Results.....	24
Discussion	30
Conclusions	34
Acknowledgments.....	35
References	36
Annex I	43
Annex II	44
PARTE II – Mapas de distribuição potencial e indicadores de conservação para espécies de Passifloraceae que ocorrem no Estado de São Paulo	45
Introdução.....	45
Métodos.....	45
Resultados.....	48
<i>Passiflora actinia</i> Hook.	49
<i>Passiflora alata</i> Curtis	50

<i>Passiflora amethystina</i> J.C.Mikan.....	52
<i>Passiflora caerulea</i> L.	55
<i>Passiflora campanulata</i> Mast.....	57
<i>Passiflora capsularis</i> L.....	58
<i>Passiflora catharinensis</i> Sacco.....	60
<i>Passiflora cervii</i> M.A.M.Azevedo	61
<i>Passiflora cincinnata</i> Mast.....	62
<i>Passiflora clathrata</i> Mast.....	64
<i>Passiflora deidamioides</i> Harms	65
<i>Passiflora edulis</i> Sims.....	66
<i>Passiflora eichleriana</i> Mast.....	69
<i>Passiflora elegans</i> Mast.....	70
<i>Passiflora foetida</i> L.....	71
<i>Passiflora glaucescens</i> Killip	73
<i>Passiflora haematostigma</i> Mart. ex Mast.....	74
<i>Passiflora ischnoclada</i> Harms.....	76
<i>Passiflora jilekii</i> Wawra	77
<i>Passiflora lepidota</i> Mast.....	79
<i>Passiflora loefgrenii</i> Vitta	80
<i>Passiflora malacophylla</i> Mast.....	81
<i>Passiflora marginata</i> Mast.....	82
<i>Passiflora mendoncae</i> Harms.....	83
<i>Passiflora miersii</i> Mast.....	84
<i>Passiflora misera</i> Kunth.....	86
<i>Passiflora morifolia</i> Mast.....	88
<i>Passiflora mucronata</i> Lam.....	89
<i>Passiflora organensis</i> Gardner	91
<i>Passiflora ovalis</i> Vell. ex M.Roem.....	93

<i>Passiflora pardifolia</i> Mast.....	95
<i>Passiflora pentagona</i> Mast.	96
<i>Passiflora pohlii</i> Mast.....	98
<i>Passiflora racemosa</i> Brot.....	100
<i>Passiflora rhamnifolia</i> Mast.....	101
<i>Passiflora setulosa</i> Killip	103
<i>Passiflora sidifolia</i> M.Roem.	104
<i>Passiflora suberosa</i> L.	106
<i>Passiflora tenuifila</i> Killip	108
<i>Passiflora transversalis</i> M.A.M.Azevedo.....	109
<i>Passiflora tricuspis</i> Mast.	110
<i>Passiflora truncata</i> Regel.....	111
<i>Passiflora urnifolia</i> Rusby	112
<i>Passiflora vellozii</i> Gardner	113
<i>Passiflora villosa</i> Vell.	114
Discussão	116
Conclusões e Recomendações	117
Referências	118
Anexo I – Espécies com distribuição pouco conhecida em São Paulo.....	120
Anexo II – Espécies com distribuição potencialmente restrita em São Paulo.....	120
Anexo III – Espécies sem registro recente de coleta em São Paulo.....	121

Resumo

Modelos de nicho ecológico vêm sendo largamente utilizados em conservação. Sua importância tem se acentuado em função da atual taxa de perda de habitat, da escassez de dados completos sobre a distribuição de espécies e da iminência de possíveis alterações climáticas de maior intensidade. Várias ferramentas de modelagem encontram-se disponíveis, bem como o acesso a dados primários de biodiversidade tem crescido muito ao longo dos últimos anos. Tal fato decorre em grande parte do surgimento de redes especializadas no compartilhamento destes dados via Internet. Apesar disso, modelos de nicho ecológico ainda têm sido essencialmente produzidos de forma isolada por pesquisadores utilizando métodos que não permitem a geração de tais modelos em larga escala, ou seja, para um grande número de espécies. Este trabalho tem como objetivo principal definir um procedimento capaz de automatizar grande parte do processo de geração de modelos de nicho ecológico, sem com isso prescindir da intervenção humana em etapas onde é importante analisar, rever e fazer eventuais correções. O procedimento proposto prevê mecanismos de retroalimentação buscando a geração contínua de melhores modelos ao apontar lacunas de conhecimento e ao ser capaz de incorporar novos dados ao longo do tempo. Serviços Web providos por autoridades taxonômicas, servidores de modelagem e redes de acesso a dados de biodiversidade integram o procedimento proposto. Diferentes estratégias de modelagem foram utilizadas dependendo do número de pontos disponíveis para cada espécie após a aplicação de diversos filtros de qualidade. Todas as estratégias tomaram como base a utilização de variáveis ambientais em alta resolução (1 km²) e dados de ocorrência de espécies provenientes de herbários. Os modelos de nicho ecológico foram treinados utilizando apenas registros que indicassem possuir suficiente precisão na coordenada geográfica, e em seguida foram testados com um novo método utilizando os registros que foram descartados por não possuir precisão suficiente. Todo o procedimento foi testado com as 135 espécies da família Passifloraceae que ocorrem no Brasil. O produto final deste trabalho ficou dividido em duas etapas: Na primeira delas o procedimento de modelagem é descrito e seus resultados são apresentados em inglês sob a forma de um artigo científico a ser submetido a um periódico internacional. Na segunda etapa, os modelos que foram gerados para espécies que ocorrem no Estado de São Paulo são apresentados sob a forma de mapas de distribuição potencial, perfis climáticos e indicadores de conservação conforme a disponibilidade de dados para cada espécie. O procedimento aqui proposto servirá como base para um novo sistema de informação integrado ao Herbário Virtual da Flora e dos Fungos, tendo seu escopo ampliado para outras famílias botânicas.

Introdução Geral

Conhecer a distribuição e as necessidades ambientais de cada espécie é tarefa fundamental para direcionar ações de conservação. Uma das ferramentas que têm sido mais utilizadas para identificar características ambientais favoráveis à sobrevivência de uma espécie bem como visualizar o resultado no espaço geográfico sob diferentes cenários ambientais é a modelagem de nicho ecológico (Sillero 2011) – também conhecida por modelagem de distribuição potencial.

Diversas técnicas de modelagem de nicho ecológico têm se mostrado importantes no estudo da biogeografia de espécies, encontrando ampla utilização em áreas como ecologia, biologia da conservação e biologia evolutiva (Elith *et al.* 2006). Fatores como o grau de degradação do meio ambiente, a escassez de dados sobre a ocorrência de determinadas espécies e a iminência de alterações climáticas tornam ainda mais oportuna e urgente a utilização de tais técnicas.

Em particular, vários exemplos demonstram a utilidade de modelos de nicho ecológico na definição de estratégias para conservação, visto que podem auxiliar na escolha de áreas prioritárias (Ferrier *et al.* 2002; Nelson & Boots 2008), indicar áreas para reintrodução de espécies (Pearce & Lindenmayer 1998), direcionar a busca por novas populações (Bourg *et al.* 2005; Guisan *et al.* 2006), prevenir a invasão de espécies exóticas em hábitat de espécies ameaçadas (Higgins *et al.* 1999; Peterson 2003), avaliar o impacto de mudanças climáticas (Araújo & New 2007; Hannah *et al.* 2007), avaliar o risco de exposição a doenças (Peterson *et al.* 2006; Peterson *et al.* 2007), direcionar ações de recomposição de áreas degradadas (Siqueira 2005) entre outras aplicações.

O conceito de nicho ecológico foi introduzido por Grinnell (1917) que o definiu como sendo o conjunto de condições ecológicas em que uma espécie consegue manter populações viáveis sem necessidade de fluxo migratório. Entretanto, no contexto da modelagem, frequentemente essa noção se restringe às condições abióticas, aproximando-se mais do conceito de nicho fundamental (Hutchinson 1957). É importante observar que a distribuição real de uma espécie frequentemente não corresponde à distribuição potencial produzida por um modelo. Populações estáveis só podem ser encontradas em regiões nas quais a espécie teve acesso desde a sua região de origem, seja por vias naturais ou por fatores antrópicos (Soberón & Peterson 2005). Por outro lado, modelos de nicho ecológico já demonstraram elevada capacidade preditiva, mesmo em prazos muito longos, de fenômenos biogeográficos. Alguns estudos apontam que os nichos modelados tendem a se conservar por períodos evolutivos e ecológicos (Peterson *et al.* 1999).

Apesar da disponibilidade de várias técnicas e ferramentas de modelagem, seus resultados ainda permanecem em grande parte dispersos e inacessíveis, sendo normalmente produzidos em pequena escala por pesquisadores e projetos. A geração de modelos em larga escala para fins de conservação requer a definição de procedimentos que possam ser automatizados e replicados em diferentes situações. Historicamente poucas iniciativas abordaram de alguma maneira a geração de modelos de nicho ecológico em larga escala, sendo possível mencionar Lifemapper (Stockwell *et al.* 2008), AquaMaps (Kaschner *et al.* 2008), Biodiversity World (Pahwa *et al.* 2006) e SEEK (Pennington *et al.* 2007). A primeira delas foi uma das pioneiras, tendo gerado milhares de mapas de distribuição potencial que, no entanto, aparentemente foram todos descartados e nunca encontraram utilização,

possivelmente por não levar em conta aspectos de qualidade. A segunda, com foco em organismos marinhos, também gerou milhares de mapas de distribuição potencial, todos atualmente disponíveis na Internet. Por fim, as duas últimas tiveram um escopo mais amplo e voltado à pesquisa acadêmica, abordando em caráter experimental a questão da geração de modelos de nicho ecológico através de novas tecnologias computacionais para aplicações de longa duração e automação de fluxos de trabalho.

A geração de modelos de nicho ecológico em larga escala depende também de amplo acesso a dados de biodiversidade, estes disponíveis principalmente através de coleções biológicas e bancos de dados de observação. O potencial do uso de dados de coleções biológicas em modelagem de nicho ecológico tem sido já há algum tempo divulgado (Graham *et al.* 2004), apresentando vantagens importantes em relação a dados de observação, uma vez que cada registro encontra respaldo físico em espécimes que podem ser consultados e analisados a qualquer momento. Tal vantagem também permite que seja feita uma constante atualização taxonômica dos registros ao longo do tempo. Entretanto, o uso efetivo de tais dados encontra uma série de desafios (Soberón & Peterson 2004), incluindo: 1) baixa quantidade de dados para muitas espécies; 2) viés geográfico e taxonômico das coletas realizadas ao longo da história; 3) presença de erros e imprecisões nos dados; 4) ausência de informações geográficas em muitas coletas. Além disso, ao direcionarmos o uso de modelos gerados com tais dados para tomada de decisão, novos desafios são encontrados, pois existem várias técnicas de modelagem e até o momento nenhuma delas provou ser superior em todos os casos (Segurado & Araújo 2004; Elith *et al.* 2006). Estratégias efetivas para geração de modelos de nicho ecológico com base em dados de coleções biológicas devem levar em conta todos esses fatores, sendo o presente trabalho uma tentativa neste sentido.

O procedimento de modelagem proposto neste trabalho foi testado com dados de herbários para todas as espécies da família Passifloraceae que ocorrem no Brasil. Em seguida, os resultados foram apresentados em maior detalhe para espécies que ocorrem no Estado de São Paulo. A família Passifloraceae é constituída por mais de 500 espécies ao redor do mundo, principalmente trepadeiras herbáceas ou lenhosas, sendo a maioria delas nativa da região Neotropical (Killip 1938). Estima-se que 135 espécies sejam nativas do Brasil (Cervi *et al.* 2010) e que pelo menos 38 ocorram no Estado de São Paulo (Bernacci *et al.* 2003). Algumas espécies possuem importância econômica, seja por seu caráter ornamental, medicinal ou como fonte de alimento. Outras são pouco conhecidas, e algumas são consideradas ameaçadas, figurando tanto na Lista Vermelha de Espécies da Flora do Brasil (MMA 2008) quanto na lista de espécies ameaçadas do Estado de São Paulo (Mamede *et al.* 2007). A diversidade aliada à importância econômica e ecológica de suas espécies torna esta família particularmente interessante como objeto de estudo para geração de modelos de distribuição potencial com fins de conservação.

Objetivo

Definir um procedimento para gerar modelos de nicho ecológico em larga escala (para muitas espécies) a partir de dados de herbários, utilizando como estudo de caso espécies da família Passifloraceae que ocorrem no Brasil e explorando os resultados em maior detalhe para espécies que ocorrem no Estado de São Paulo.

Objetivos específicos

- Criar um procedimento para geração de modelos de nicho ecológico em larga escala para espécies da flora que possa futuramente ser incorporado a um sistema de informação.
- Possibilitar intervenção humana em etapas que não são passíveis de automatização, permitindo que dados e resultados possam ser analisados, revisados e selecionados por especialistas.
- Prover mecanismos de controle de qualidade tanto nos dados de entrada quanto nos resultados da modelagem, incluindo filtros que facilitem a seleção de dados para geração de modelos em alta resolução e fornecendo parâmetros que ajudem a decidir sobre a adequação dos resultados ao uso em conservação.
- Direcionar o procedimento para uso de dados cujo acesso seja público, permitindo análise e revisão dos mesmos a qualquer momento, além de dar credibilidade e transparência ao processo.
- Incluir estratégias para lidar com poucos pontos de ocorrência de espécies.
- Prover mecanismos capazes de orientar ações que contribuam para obtenção de novos dados, permitindo a geração de modelos para espécies antes com dados insuficientes, bem como a geração de melhores modelos, criando um ciclo contínuo em busca de resultados mais precisos e com melhor qualidade.
- Incluir meios que facilitem a tomada de decisão com base nos resultados.
- Testar o procedimento com espécies de Passifloraceae que ocorrem no Brasil.
- Detalhar os resultados para espécies que ocorrem no estado de São Paulo, gerando perfis climáticos e indicadores de conservação quando houver dados suficientes.

Referências

Araújo, M.B. & New, M. (2007) Ensemble forecasting of species distributions. *Trends in Ecology & Evolution*, 22 (1): 42-47.

Bernacci, L.C. (2003) Passifloraceae. In Wanderley, M.G.L.; Shepherd, G.J.; Giulietti, A.M. & Melhem, T.S. (coord.) *Flora Fanerogâmica do Estado de São Paulo*. vol. 3: 247-248. RiMa/FAPESP, São Paulo.

Bourg, N. A.; McShea, W. J. & Gill, D. E. (2005) Putting a CART before the search: Successful habitat prediction for a rare forest herb. *Ecology* 86: 2793-2804.

Cervi, A.C., Milward-de-Azevedo, M.A., Bernacci, L.C. (2010) *Passifloraceae in Lista de Espécies da Flora do Brasil*. Jardim Botânico do Rio de Janeiro (<http://floradobrasil.jbrj.gov.br/2010/FB000182>).

Elith, J.; Graham, C.H.; Anderson, R.P.; Dudík, M.; Ferrier, S.; Guisan, A.; Hijmans, R.J.; Huettmann F.; Leathwick, J.R.; Lehmann, A.L.; Li, J.; Lohman, L.G.; Loiselle, B.A.; Manion, G.; Moritz, C.; Nakamura, M.; Nakazawa, Y.; Overton, J.Mc.C.; Peterson, A.T.; Phillips, S.J.; Richardson, K.; Scachetti-Pereira, R.; Schapire, R.E.; Soberón, J.;

- Williams, S.; Wisz, M.S. & Zimmermann, N.E. (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 29:129–151.
- Ferrier, S.; Watson G.; Pearce, J. & Drielsma M. (2002) Extended statistical approaches to modelling spatial pattern in biodiversity: the north-east New South Walles experience. I. Species-level modelling. *Biodiversity and Conservation* 11: 2275-2307.
- Graham, C.H.; Ferrier, S.; Huettman, F.; Moritz, C. & Peterson, A.T. (2004) New developments in museum-based informatics and applications in biodiversity analysis. *Trends in Ecology and Evolution*, 19(9).
- Grinnell, J. (1917) Field tests of theories concerning distributional control. *The American Naturalist*. 51: 115-128.
- Guisan, A., Lehmann, A.; Ferrier, S.; Austin, M.; Overton, J. MC. C.; Aspinall, R. & Hastie, T. (2006) Making better biogeographical predictions of species' distributions. *Journal of Applied Ecology* 43: 386-392.
- Hannah, L.; Midgley, G.; Andelman, S.; Araújo, M.; Hughes, G.; Martinez-Meyer, E.; Pearson, R. & Williams, P. (2007) Protected area needs in a changing climate. *Frontiers in Ecology and the Environment* 5: 131-138.
- Higgins, S. I.; Richardson, D. M.; Cowling, R. M. & Trinder-Smith, T. H. (1999) Predicting the landscape-scale distribution of alien plants and their threat to plant diversity. *Conservation Biology* 13: 303-313.
- Hutchinson, G.E. (1957) Concluding remarks. *Cold Spring Harbor Symposia on Quantitative Biology*, 22(2): 415–427.
- Kaschner, K.; Ready, J.S.; Agbayani, E.; Rius, J.; Kesner-Reyes, K.; Eastwood, P.D.; South, A.B.; Kullander, S.O.; Rees, T.; Close, C.H.; Watson, R.; Pauly, D. & Froese, R. (2008) AquaMaps: Predicted range maps for aquatic species. Publicação eletrônica, www.aquamaps.org, Version 08/2010.
- Killip, E.P. (1938) The American species of Passifloraceae. *Field Museum of Natural History, Botanical Series*, 49: 1-613.
- Mamede, M.C.H.; Souza, V.C.; Prado, J.; Barros, F.; Wanderley, M.G.L. & Rando, J.G. (2007) Livro Vermelho das Espécies Vegetais Ameaçadas do Estado de SP. Instituto de Botânica, São Paulo, 165p.
- MMA (2008) Instrução Normativa MMA nº 06, 23 de setembro de 2008.
- Nelson, T. A. & Boots, B. (2008) Detecting spatial hot spots in landscape ecology. *Ecography* 31: 556-566.
- Pahwa, J.S.; Brewer, P.; Sutton, T.; Yesson, C.; Burgess, M.; Xu, X.; Jones, A.C.; White, R.J.; Gray, W.A.; Fiddian, N.J.; Bisby, F.A.; Culham, A.; Caithness, N.; Scoble, M.; Williams, P. & Bhagwat, S. (2006) Biodiversity World: A Problem-Solving Environment for Analysing Biodiversity Patterns, em: Proc. 6th IEEE International Symposium on Cluster Computing and the Grid (CCGRID 2006), Singapura.

- Pearce, J. & Lindenmayer, D. (1998) Bioclimatic analysis to enhance reintroduction biology of the endangered helmeted honeyeater (*Linchenostomus melanops cassidix*) in southeastern Australia. *Restoration Ecology* 6: 238-243.
- Pennington, D.D.; Higgins, D.; Peterson, A.T.; Jones, M.B.; Ludäscher, B. & Bowers, S. (2007) Ecological Niche Modelling Using the Kepler Workflow System, in: Taylor, I.J., Deelman, E., Gannon, D.B., Shields, M. (Eds.), *Workflows for e-Science*, 91-108, Springer, Berlin.
- Peterson, A. T. (2003) Predicting the geography of species' invasions via Ecological Niche Modeling. *Quarterly Review of Biology* 78(4): 419-433.
- Peterson, A.T.; Soberón, J. & Sanchez-Cordero, V. (1999) Conservatism of ecological niches in evolutionary time. *Science* 285: 1265-1267.
- Peterson, A. T.; Lash, R. R.; Carroll, D. S. & Johnson, K. M. (2006) Geographic potential for outbreaks of Marburg hemorrhagic fever. *American Journal of Tropical Medicine & Hygiene* 75: 9-15.
- Peterson, A. T.; Benz, B. W. & Papeş, M. (2007) Highly pathogenic H5N1 avian influenza: Entry pathways into North America via bird migration. *PLoS ONE* 2: e261.
- Segurado, P. & Araújo, M.B. (2004) An evaluation of methods for modelling species distributions. *Journal of Biogeography*, 31: 1555–1568.
- Sillero, N. (2011) What does ecological modelling model? A proposed classification of ecological niche models based on their underlying methods. *Ecological Modelling*, 222(8): 1343-1346.
- Siqueira, M. F. (2005) Uso de Modelagem de nicho fundamental na avaliação do padrão de distribuição geográfica de espécies vegetais. Tese de Doutorado. Ciências da Engenharia Ambiental, Escola de Engenharia de São Carlos, USP.
- Soberón, J. & Peterson, A.T. (2004) Biodiversity informatics: managing and applying primary biodiversity data. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 359: 689-98.
- Soberón, J. & Peterson, A.T. (2005) Interpretation of models of fundamental ecological niches and species' distributional areas. *Biodiversity Informatics*, 2, 1–10.
- Stockwell, D.R.B, Beach, J.H., Stewart, A., Vorontsov, G., Vieglais, D. & Scachetti-Pereira, R. (2008) The use of the GARP genetic algorithm and internet grid computing in the Lifemapper world atlas of species biodiversity. *Ecological Modelling*, 195 (1-2): 139-145.

PARTE I – Towards a comprehensive workflow for generating ecological niche models with herbarium data: a case study with Passifloraceae in Brazil

Abstract

Ecological niche models have been widely used in conservation. Its importance is accentuated by the current rate of habitat loss, the paucity of occurrence data for certain species and the imminence of possibly more intense climatic changes. Many modelling tools already exist, and its main input – biodiversity data – is becoming increasingly available through online networks. Nevertheless, ecological niche models have been mostly produced in isolation by individual researchers in ways that do not scale well for a large numbers of species. The work described here is an attempt to define a workflow to automate as much as possible the whole modelling process, but also enabling human intervention in certain tasks to analyze, review and make specific changes. The process makes use of feedback loops to generate better models over the time by indicating data gaps and continually checking when new data becomes available. Online services and data sources, such as taxonomic authority files, modelling servers and biodiversity data networks are an integral part of the workflow. Different modelling strategies were used depending on the remaining number of unique selected locality records for each species after applying several data quality filters. All modelling strategies were based on high resolution environmental layers (1 km²) and presence-only species occurrence data from herbaria. Models were trained with records having enough locality precision and tested with a new procedure based on records that were discarded due to positional uncertainty. A pilot experiment for the workflow was carried out for the 135 species of Passifloraceae that are known to occur in Brazil. Results showed that 13% of the species had enough locality data (at least 20 points) to produce models combining five niche modelling techniques, from where environmental profiles and conservation parameters can be extracted. Other 16% had between 5 and 19 localities to be used with exploratory modelling, and 36% of the species had between 1 and 4 localities, in which case environmental dissimilarity models were generated to guide new surveys. Finally, 35% of the species had no suitable localities, providing important information to prioritize further field research and digitization efforts.

Introduction

Recent studies are confirming that the current rate of biodiversity loss has possibly reached critical levels (Barnosky *et al.* 2011; Rockström *et al.* 2009). The advent of global warming and its potential synergistic interactions with other threats (Stork 2009) contributes to a scenario that requires urgent conservation measures. In this context, priority must clearly be given to making the most effective use of the available data, as well as using existing predictive tools to guide conservation actions, overcome data gaps and provide the necessary support for decision making.

Ecological niche modelling and biodiversity data from natural history collections have emerged as a promising alternative for the task (Graham *et al.* 2004; Soberón & Peterson 2004). This combination is gradually gaining momentum as novel modelling methods are researched (Elith *et al.* 2006; Zimmermann *et al.* 2010), incorporated into existing tools (Muñoz *et al.* 2010; Guo & Liu 2010; Thuiller *et al.* 2009; Diniz-Filho *et al.* 2009), and more data becomes available in digital format.

Natural history collections are estimated to contain between 1.2 and 2.1 billion specimens collected from nature during the last centuries (Ariño 2010). These specimens are the foundation of all our knowledge about biodiversity on Earth, also providing a unique baseline data in a period when human impact on nature increased dramatically (Chapman & Busby 1994). Access to this data is being improved as natural history collections digitize their holdings and share records with biodiversity data networks. By documenting where and when each specimen was collected, natural history collections are an important source of data for most ecological niche modelling methods. The most common one is known as the correlative approach (Soberón & Peterson 2005), which relates sample localities from the species' range with spatially explicit environmental data so that specific algorithms can build mathematical representations of the species' environmental requirements. Thus, results can be seen as an approximation of the species' niche, originally defined by Grinnell (1917) as being the range of ecological conditions within which a species can maintain populations. Such models can then be used to understand and predict species occurrences under different scenarios. For this reason, and for its broad-spectrum utility (see Elith *et al.* 2006 for a comprehensive list), ecological niche modelling has been extensively used in conservation planning and biodiversity management. We actually reached a point where, even with relatively small sample sizes and all sorts of biases that can be found in data coming from biological collections (Yesson *et al.* 2007; Tobler *et al.* 2007; Kadmon *et al.* 2004), ecological niche modelling can still be used successfully in many cases (Wisz *et al.* 2008; Guisan *et al.* 2007; Elith *et al.* 2006).

Despite this fact, ecological niche models have been mostly produced by individuals in ways that do not scale well for a large number of species (Pahwa *et al.* 2006). Additionally, species occurrence records that are used to generate models frequently come from privately held spreadsheets or databases, making it difficult for other people to inspect the data and reproduce the experiments. Finally, results from the modelling process also do not tend to be made publicly available in a digital format to promote their usage. To date, at least two initiatives tried to address these issues: Lifemapper (Stockwell *et al.* 2008) and AquaMaps (Kaschner *et al.* 2008). Inspired by the SETI@home project (Sullivan *et al.* 2007), Lifemapper generated models for more than 100000 taxa through screensavers installed on desktop computers (Stockwell *et al.* 2008) and is currently rebuilding its repository of models using a centralized Cluster. AquaMaps also created an approach to generate a large number of models, currently making available on the Web 11500 distribution maps for species of fishes, marine mammals and invertebrates.

Two other initiatives, the Science Environment for Ecological Knowledge project (Michener *et al.* 2005) and the Biodiversity World project (Pahwa *et al.* 2006), explored the idea of creating scientific workflows to enable a wide range of biodiversity research. Although they had a broad scope, both included specific niche modelling case studies taking advantage of Grid computing (Pahwa *et al.* 2006; Pennington *et al.* 2007).

All these initiatives led to important advances, some of them by developing the notion of and building large databases of species distribution models, others by promoting the usage of Web Services and high-throughput computing resources, as well as creating experimental workflows to produce such models. However, there are many practical issues that need to be carefully addressed when ecological niche models are generated with data from natural history collections (Soberón & Peterson 2004; Chapman 2005; Guralnick *et al.* 2007; Heap & Culham 2010), including errors and uncertainties (both taxonomic and geographic), small sample sizes, and model quality control – especially when results will be made available for general use.

Many species have very few records available from online networks, either due to limitations and biases in the historical collecting effort, or simply to the fact that most data is still not digitized. Moreover, most records available from biodiversity data networks do not contain any coordinates (Wieczorek *et al.* 2004). In the opposite situation, species with many occurrence records can have their distribution modelled by several algorithms – each one with particular features and showing distinct performance under different conditions (Elith & Graham 2009; Lorena *et al.* 2008; Guisan *et al.* 2007). Many comparisons were made between existing algorithms, so far with no individual algorithm proving to be superior in all cases (Segurado & Araújo 2004; Elith *et al.* 2006; Lorena *et al.* 2011). This fact emphasizes the importance of including specific strategies in a generic ecological niche modelling workflow to deal with different sample sizes and multiple algorithms.

Here we propose a comprehensive workflow to generate ecological niche models based on herbarium data trying to take into account these issues and at the same time producing a single model per species to facilitate its usage. Feedback loops and partial results can be used to report errors and inconsistencies to data providers, guide further research, and therefore allow better models to be created over time. The whole strategy depends on a series of factors including knowledge about the ecology and taxonomy of the species, knowledge about the geography and history of the area of interest, and also access to all relevant data and tools. These requirements are hard to be met on a global basis, and even in a more regional or local context this is only possible after significant effort from a lot of people and institutions.

Many important initiatives that have been developed in Brazil over the last years are now culminating in a unique opportunity to meet the necessary conditions for this kind of research, such as the official checklist of the Brazilian Flora that was recently released (Forzza *et al.* 2010), and the *speciesLink* network (Canhos *et al.* 2004) that has reached 4 million records from biological collections and is steadily growing at half million records per year. If we consider only plant specimens from Brazilian collections, *speciesLink* currently provides access to ~2.3 million records, which corresponds to approximately 38% of all holdings stored in Brazilian herbaria (assumed be around 6 million specimens according to Egler & Santos 2006). This proportion is considerably higher than previous digitization estimates of 5-10% for biological collections worldwide (Beaman & Conn 2003). Additionally, ~41% of these plant records deposited in Brazilian herbaria have coordinates – another proportion far greater than previous global estimates of 1% (Guralnick *et al.* 2006).

The workflow was experimentally tested in a first pass with all species of Passifloraceae that occur in Brazil. This family, well-known for the passion-flowers, is mainly comprised by herbaceous or woody vines, although there are also trees and shrubs among its species (Killip 1938). More than 500 species of Passifloraceae exist around the world, but they are mostly found in the Neotropics (Bernacci *et al.* 2003). There is an additional challenge for ecological niche modelling regarding this family since it contains many cultivated species, including species of commercial importance such as *Passiflora edulis* Sims and *Passiflora alata* Curtis. Together, these two species accounted for an annual production of 713515 ton in 2009 only in Brazil (IBGE 2011). This fact requires additional procedures to distinguish between specimens collected from cultivated individuals and those collected from the wild. In Brazil, there are 135 native species of Passifloraceae (Cervi *et al.* 2010), some of them are wide spread across the country, others are specific to certain regions and biomes, and still there are others with a very constrained distribution, making this family of particular interest for this kind of study. Additionally, five species are listed as threatened by the latest red list

of the Brazilian Flora (MMA 2008): *Passiflora hatschbachii* Cervi, *Passiflora imbeana* Sacco, *Passiflora ischnoclada* Harms, *Passiflora margaritae* Sacco and *Passiflora saccoi* Cervi.

Material and methods

The workflow presented here (figure 1) was designed to interact with external resources (remote data providers or tools) through Web Service interfaces, i.e. standardized ways for programs to communicate with each other over the Web. This way, the most critical parts of the process can be automated, allowing ecological niche models to be generated for a large number of species. All steps were implemented in the Python programming language. Flow control, however, was mostly performed manually, since the main idea was to produce a proof of concept for the whole modelling strategy. All data needed by the workflow and generated by it was stored into a local database during the process.

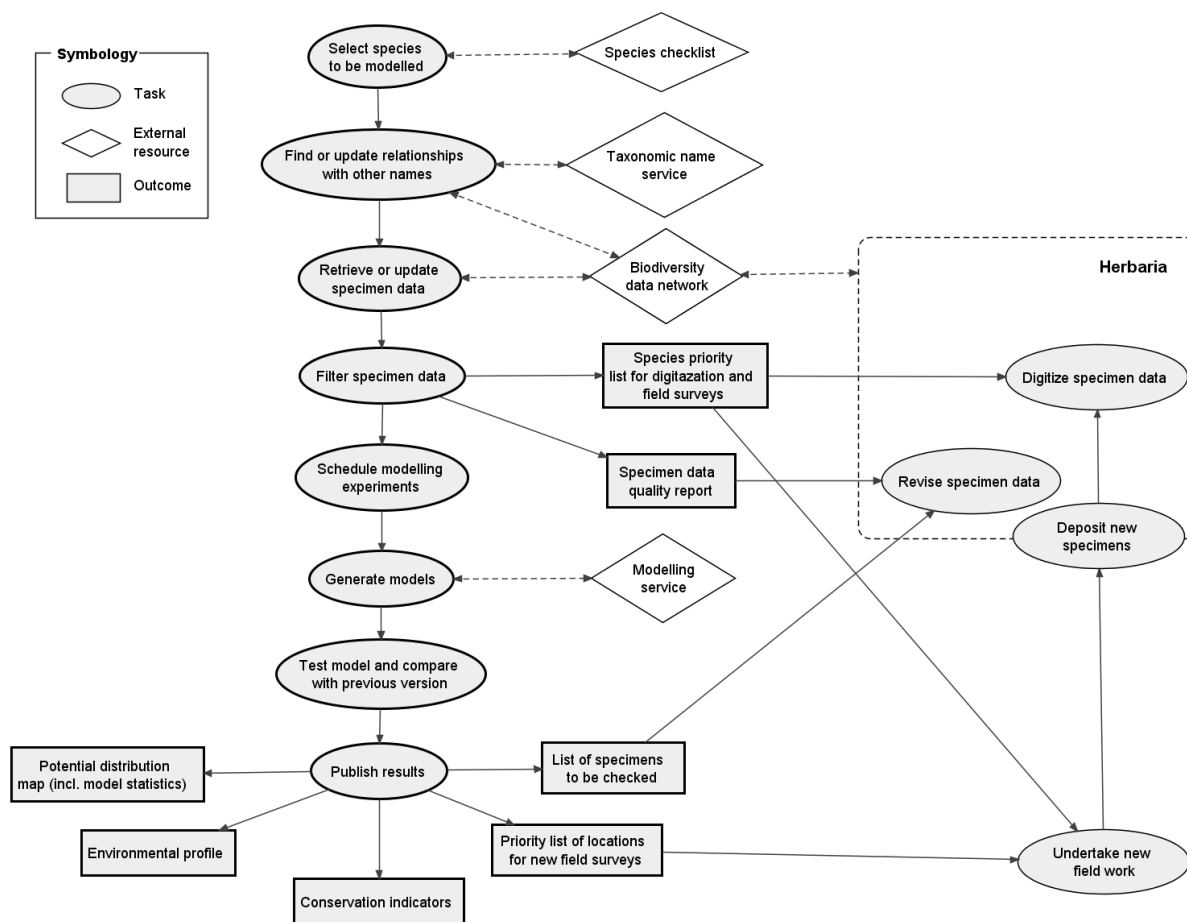


Figure 1: Overview of the workflow used to generate the ecological niche models. Tasks on the left side (darker border) were performed as part of this work. Tasks on the right side can be guided by outcomes from the workflow and may provide new data to improve results in subsequent iterations.

The process starts by selecting one or more species from the official checklist of the Brazilian Flora. Since this is the only external resource that currently still lacks a Web Service, all accepted names for species of Passifloraceae that occur in Brazil were initially imported from the checklist into the local database.

Three other external resources were used during the workflow: 1) A taxonomic name service provided relationships between accepted names and other names that can still be found in materials deposited in herbaria. This was achieved by interacting with the

Tropicos Web Services (Tropicos 2011). 2) Species occurrence records were retrieved from the *speciesLink* network through its public Web Service based on the TAPIR protocol (Giovanni *et al.* 2010) and on the Darwin Core data standard (Wieczorek *et al.* 2009). More than 30 institutions that are listed in the end of the article provided data to this work through *speciesLink*. 3) Most models were generated, tested and projected remotely by interacting with a modelling server hosted at the Reference Center on Environmental Information (CRIA). This was performed through the openModeller Web Service API (Muñoz *et al.* 2011).

Although it should be possible to carry out all these steps automatically, each task was followed by expert review allowing specific corrections and adjustments to be made, such as adding more related names or tagging specimen records for revision. Many outcomes from the workflow, not just ecological niche models but other intermediary results presented later, are expected to influence field work, research and digitization efforts, which may in turn gather new data that can be used in subsequent workflow iterations. A single iteration was performed as part of this work: from species selection to the first set of ecological niche models and all derived outcomes.

Species names

Specimen identifications in biological collections are subject to many nomenclatural and taxonomic variations, posing additional challenges when using data from different sources (Soberón & Peterson 2004). Despite the known limitations of using scientific names as identifiers for biological taxa (Kennedy *et al.* 2005), which also includes risks for species distribution modelling (Franz *et al.* 2008), scientific names are still the main way of accessing primary biodiversity data (Chapman 2005). This condition will likely persist for considerable time, since either creating a full prospect of all taxon concept relationships or revising and updating all specimen identifications is a colossal task. Meanwhile, taxonomic and nomenclatural databases can be used to increase the number of records per species by including unambiguous synonyms and known misspellings during data retrieval, as well as excluding homonyms and varieties potentially referring to different species.

These tasks were mostly performed by interacting with the Tropicos Web Services, but always followed by expert review. Only one species from the checklist was not found in the Tropicos database (*Passiflora boticarioana* Cervi) and seven other species had to be manually associated with Tropicos name entries due to different authorship or use of orthographical variants. The following types of names were taken into account to improve specimen data selection:

- Incompatible homonyms: Names published with the same binomial of an accepted name but referring to a different species. For example, *Passiflora rubra* L. is currently an accepted name, but *Passiflora rubra* Vell. is now seen as *Passiflora organensis* Gardner. In such cases, specimen records from the species being modelled were later flagged for exclusion from the modelling process if no author name was provided or if the author name did not match the expected value. Five incompatible homonyms were detected (annex I).
- Incompatible varieties: Varieties that share the same binomial with an accepted name but are now considered a new species after taxonomic revision. For example, *Passiflora amethystina* J.C.Mikan is an accepted name, but *Passiflora amethystina* var. *bolosii* Cervi is now seen as *Passiflora loefgrenii* Vitta. Specimen records named with incompatible varieties were not used in the

modelling process of the species with the same binomial. However, they could be used to model the new species which now corresponds to that variety (in case the new species was in the list of species to be modelled). Fourteen incompatible varieties were detected (annex I).

- Unambiguous synonyms: Different binomials that can be used to search for specimen records of the species being modelled provided they do not have any incompatible homonyms. Many synonyms came from the checklist of the Brazilian Flora and from literature (e.g. Killip 1938; Cervi 1997; Bernacci *et al.* 2003). In total, 23 synonyms were registered (annex I).

Name misspellings were also registered as alternative names, this time by interacting with the specimen web service. This was possible because *speciesLink* offers a mechanism to retrieve all distinct specific epithets given one or more genera (in our case, *Ancistrothyrsus* Harms, *Dilkea* Mast., *Mitostemma* Mast. and *Passiflora* L.). This was achieved by means of TAPIR inventory requests (Giovanni *et al.* 2010). Inventory responses can be browsed to find possible misspellings followed by confirmation in the Tropicos name service that they do not correspond to a published name. In total, 45 misspellings and orthographical variants were identified and registered.

Environmental data

The following criteria guided the selection of environmental layers to be used in the modelling process:

- Ecological relevance of predictor variables, giving particular emphasis on climate, since climatic variables seem to be the main drivers controlling the distribution of most plant species (Woodward 1987).
- Availability of high resolution georeferenced raster data (~1 km), allowing use case scenarios at local scales for the generated models even at the later risk of not finding enough occurrence records with compatible precision.
- Variables having continuous ranges, since most algorithms cannot deal with categorical data.
- Preference to rasters with global coverage, in case models need to be projected in other geographical regions in the future.

These conditions lead to an initial selection of twenty two layers: Nineteen bioclimatic variables from WorldClim (Hijmans *et al.* 2005); altitude, also from WorldClim; and potential evapotranspiration and aridity index, both from the Global High-Resolution Soil-Water Balance dataset (Zomer *et al.* 2007). WorldClim bioclimatic variables were preferred over the monthly averages since they better represent seasonality, annual trends and extreme climatic conditions (Hijmans *et al.* 2005). Altitude was chosen as an indirect measure of other variables that can influence plant distribution, such as incident solar radiation and barometric pressure (Gale 2004). Unfortunately no continuous soil data was found for the whole study area.

This initial data set was further refined to avoid redundancy and to minimize the risk of model over fitting – especially in situations when only a small number of species occurrence points are available. This was achieved by performing a correlation analysis after randomly sampling the environmental conditions across the study area ($n = 10000$) and calculating a correlation matrix with Pearson coefficients. The final set of

layers was selected by successively removing one of the layers involved in each high correlation detected ($r > 0.75$, including visual inspection). Layers from simpler (non-composite) variables that could better explain plant distribution were favoured. By using smaller environment dimensions with variables that are easier to understand we also expect to facilitate model interpretation in the future.

The final set of layers used in the modelling process included 8 variables: mean diurnal temperature range (BIO2), maximum temperature of warmest month (BIO5), minimum temperature of coldest month (BIO6), precipitation of wettest quarter (BIO16), precipitation of driest quarter (BIO17), precipitation of warmest quarter (BIO18), precipitation of coldest quarter (BIO19) and altitude (ALT).

Specimen data

Specimen data was retrieved from *speciesLink* using the genus and the specific epithet from the accepted name and from alternative names, with proper care regarding incompatible homonyms and varieties as explained earlier. Since models would be generated for species, all specimens identified as accepted varieties or old varieties compatible with the species were considered valid records for the species.

Each record retrieved was processed to estimate values not directly provided by most biological collections and to associate free textual content with standard values. Country interpretation was performed by testing exact matches to different strings representing the country. State and county were compared against names and abbreviations from a standard gazetteer (IBGE 2003) using the Python *diffliB* module, which measures the similarity of two character sequences with a match ratio value between 0 (no match) and 1 (full match). Best ratios above 0.83 for names were accepted, allowing partial matches caused by typos or misspellings. The same gazetteer also provided georeferenced shapes for each county, allowing later verification whether the coordinates of each record were within the boundaries of the specified county.

Records collected from cultivated individuals were automatically detected by searching for a few different sequences of characters (such as "cultiv") in the observations about the record, but still allowing experts to review the automatic selection later.

Since high positional uncertainties can degrade model performance (Guisan *et al.* 2007; Graham *et al.* 2008; Fernandez *et al.* 2009) – especially when high resolution environmental layers are used – the following sources of positional uncertainty were calculated: coordinate precision, datum error and whether retrospective georeferencing by county was used. Precision errors were calculated according to Wiczorek *et al.* (2004) for all records with coordinates. In this case, verbatim coordinate values were always preferred as the input, since decimal longitude and latitude values derived from another format usually do not reflect the original precision. Datum errors were also calculated since none of the collections seemed to store this information. Three options for horizontal datum were considered: WGS84, SAD69 and "Córrego Alegre". Datum errors would be negligible in the study area if the alternatives were restricted to WGS84 and SAD69. However, when "Córrego Alegre" is included, errors can be greater than 100 meters for points in the south or northwest of the country (figure 2). The overall location uncertainty was calculated as the precision error added to the datum error. In the few occasions when a coordinate uncertainty was already specified by the data provider, this value was used instead of the calculated error.

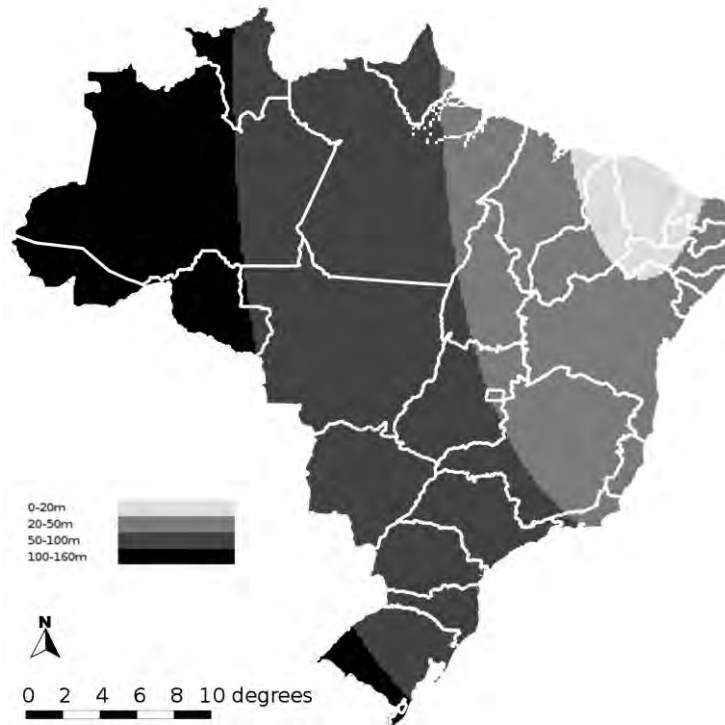


Figure 2: Uncertainty ranges (in meters) when no datum is indicated for coordinates in Brazil. Values were calculated as the maximum distance between points with the same coordinates but a different datum (Wieczorek *et al.* 2004), including WGS84, SAD69 and “Córrego Alegre”.

When each point has an associated uncertainty and is represented by a circle following the point-radius method described by Wieczorek *et al.* (2004), the circle’s position could ultimately be checked in relation to the raster pixels when retrieving environmental data. A precise match between the circle and a given pixel can only be guaranteed if the entire circle is within the pixel boundaries. Otherwise, if the circle intersects with multiple pixels, different strategies could be used to better estimate the corresponding environmental value. Regardless the strategy, as the uncertainty increases in relation to the raster resolution the final result may indicate a totally inaccurate environmental condition for the species, which means that it is always important to define a maximum acceptable uncertainty. In this work, environmental conditions were retrieved based only on the provided coordinates for the species (to our knowledge, none of the existing ecological niche modelling software packages can handle position uncertainty). We chose 500 m (half the environmental data resolution at the equator) as the maximum uncertainty for two reasons: 1) Any uncertainty above zero and not greater than the pixel size have an associated risk that the real position is actually related to a neighbour pixel – even points measured with high precision GPS. In most cases this is probably acceptable, unless the point is in the edge of the ecological niche and the pixel indicates unsuitable conditions for the species. 2) Any choice of a maximum uncertainty value below the pixel size becomes a matter of choosing the minimum acceptable probability of the real point being in the same pixel of the provided coordinates (figure 3). The probability function can be approximated by an exponential decay: $P(u) = \exp(-1.1 \cdot u)$, giving, for example, probabilities of 0.89, 0.72 and 0.51 for uncertainties of 100 m, 300 m and 600 m, respectively. Points with an uncertainty value close to 500 m will have an approximate probability of 0.6 to be correctly associated with the given pixel.

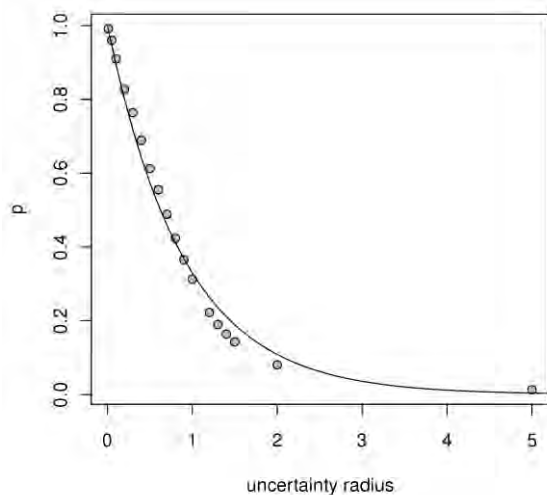


Figure 3: When a measured point with a given uncertainty radius falls within a 1x1 pixel, this graph shows the probability of the “real” location being inside the same pixel. Results were obtained through simulation by generating 1000 random points inside the pixel for each radius to be tested and then taking the average intersection area between the circle (defined by the point and the corresponding radius) and the pixel. Dots indicate the average probability from the simulation and the line shows the result of a non-linear regression (exponential decay) to approximate the probability function.

Retrospective georeferencing by county was detected by calculating the distance between the record coordinates and the county coordinates provided by a standard gazetteer (IBGE 2003). In Brazil, positional uncertainties at the level of counties have an average magnitude of 35 km ($\sigma = 45.3$), possibly reaching more than 500 km in some counties from the Amazon region.

The following criteria were used to automatically discard records for model training:

- Records without coordinates.
- Records collected or observed outside the study area.
- Observation or living collection records (as indicated by the content of the Darwin Core term "basisOfRecord"). Only specimen records were used.
- Records identified as being collected from cultivated individuals.
- Records with a location uncertainty greater than 500 m.
- Records with coordinates outside the indicated county (when the shortest distance to the county border was greater than 2 km plus the location uncertainty).
- Records that were likely retrospective georeferenced at the county level (when coordinates were less than 50 m from the standard coordinates for the county).
- Records without an author or with an unexpected author in the scientific name (only when the identified species has incompatible homonyms).

Additionally, records with the same collector name and collector number were considered duplicates, and in this case only one record was selected based on its location quality (prioritizing records with no data conflicts and with the smallest uncertainty). If at least one of the duplicates had a different identification, then all

duplicates were excluded from the modelling process and tagged with an identification conflict to be reported back to the respective collections.

Finally, a reference raster with the same resolution of the environmental layers was used to get the associated pixel for each selected record and guarantee that only one record per pixel was used for each species. When more than one record shared the same pixel, the record with best location quality was preferred.

All records that were only discarded for having high positional uncertainty were selected for model testing. This includes: records with location uncertainty greater than 500 m, records that were identified as being retrospectively georeferenced at the county level, and records without coordinates but with an associated county name that could be interpreted. In the first case, the area to be tested was approximated by a circle centred in the provided coordinates. In the two other cases, the area to be tested was the county polygon. This initial selection was further refined by excluding:

- Records whose polygon included one of the training points.
- Records whose polygon entirely included another test record, therefore favouring test records with better precision. Note that duplicate test records were also excluded here.

One important aspect is that no changes were made to the original data coming from biological collections. All procedures tried to estimate additional data based on what is provided and detect inconsistencies to select which records would be used in each part of the modelling process. Additionally, species whose records did not pass any of these filters were included in a list so that they can be prioritized in digitization programs and field work (annex II). The list was ordered by the corresponding number of records available from *speciesLink*, the number of distinct counties where they were collected (as a proxy to their geographical range) and the year of the last collecting event.

Modelling strategy

The modelling strategy used by the workflow builds on the notion that model ensembles (Thuiller 2004; Araújo & New 2007) can produce better results than individual models generated by single algorithms (Marmion *et al.* 2009) and that each algorithm requires a minimum number of presence points to reach an acceptable level of stability and produce more accurate results (Mateo *et al.* 2010). Additionally, merging results from different algorithms into a single consensus model addresses a usability issue by not exposing users to a wide range of models shaped by the idiosyncrasies of each algorithm.

Considering that natural history collections are restricted to species occurrence records and that the usage of absence data in ecological niche modelling is strongly subject to pitfalls (Peterson *et al.* 2008; Hirzel *et al.* 2002) all algorithms used by the workflow follow the correlative approach (Soberón & Peterson 2005) and require presence-only data – except for algorithms that are known to produce accurate results with totally random pseudo-absences, such as GARP (Stockwell & Peters 1999). The strategy can be seen as an incremental model ensemble approach, where each algorithm is only used if the species has a minimum number of presence points (figure 4) and, when more than one algorithm is used, a single consensus model is created. As the number of presences increases, more algorithms are used and more robust is expected to be the result. A final post-processing step was also carried out to avoid possible over predictions by some of the algorithms involved.

Maxent (Phillips *et al.* 2006) was used with a minimum number of 5 presence points (Hernandez *et al.* 2006). GARP with the Best Subsets procedure (Anderson *et al.* 2003) (GARP BS) was used with at least 10 presence points (Stockwell & Peterson 2002). The other algorithms – ENFA (Hirzel *et al.* 2002), one-class Support Vector Machines (Schölkopf *et al.* 2001) (SVM) and Mahalanobis Distance (Farber & Kadmon 2003) (MD) – were all used with a minimum number of 20 points, following a more liberal approach than existing recommendations for these algorithms, which are: three times the number of environmental predictors for ENFA (Hirzel *et al.* 2011), meaning 24 points in this case, 40 points for SVMs (Drake *et al.* 2006) and 50 points for MD (Farber & Kadmon 2003). This decision was made since the effect of model aggregation and the final post-processing step can compensate an eventual poorer performance of individual algorithms.

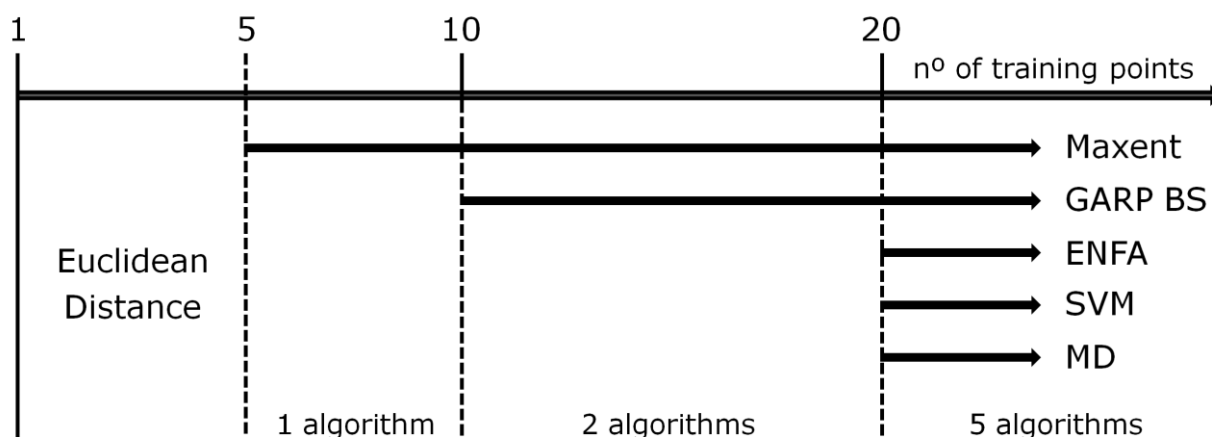


Figure 4: Algorithms used for each range of training points. Final models were generated by combining the individual models from each algorithm into a consensus model after converting them to a binary response based on the LPT (lowest presence threshold) above zero, and then discarding predictions when there was not enough agreement between the algorithms (minimum level of agreement based on simple majority).

There are different approaches to generate consensus models and some of them have already been compared, resulting in the recommendation to use weighted averages or mean values across the individual predictions (Marmion *et al.* 2009). Since there is still some debate about methods used to measure algorithm performance (Lobo *et al.* 2008; Peterson *et al.* 2008) and also to start with a simpler strategy we decided not to assign weights to each individual model. The approach used here is similar to the mean values technique, except that each individual model was first converted into a binary output by applying a threshold before getting merged. There are also different methods to calculate thresholds and an existing comparison suggests either the use of prevalence (also known as lowest presence threshold – LPT) or average probability/suitability (AVG) as two simple and efficient approaches (Liu *et al.* 2005). Since AVG thresholds tend to be high, producing quite restrictive predicted ranges, we used LPT with the additional constraint of getting the lowest presence threshold above zero to handle possible omissions that could result in the whole study area predicted as suitable for the species. An interesting property when consensus models are generated this way is that prediction values indicate the number of algorithms agreeing about suitable conditions. Potential distribution maps can then be post-processed to avoid possible over predictions, showing results only when there is a minimum level of agreement between the algorithms. The minimum level of agreement adopted here was based on simple majority, meaning 1 when only Maxent was used, 2 when Maxent and GARP BS were used, and 3 when all 5 algorithms were used. This corresponds to

final post-processing thresholds of 1.0, 1.0 and 0.6, respectively (predictions below these values were converted to zero).

Most models were remotely generated with the openModeller framework (Muñoz *et al.* 2011) and then post-processed locally to generate the final consensus, with the only exception of Maxent, whose models were all generated locally using the original Maxent software (Phillips *et al.* 2006). The following parameters were used for each algorithm: Maxent with auto-features (maximum iterations = 500, convergence threshold = 0.00001), GARP BS with 100 runs (training proportion = 50%, soft omission, models under omission threshold = 20, commission threshold = 50%, convergence limit = 0.01, population size = 50, resamples = 2500), SVM with radial basis function ($\gamma = 1/\text{number of variables}$, $\nu = 0.05$), ENFA with the broken stick method for discarding components, and MD with no distance limit to the central distribution point, forcing distances to be expressed as chi-square probabilities. The GARP version used here is known in the literature as OM-GARP (Elith *et al.* 2006). The same 10000 points used for the environmental data correlation analysis served as background points for Maxent and ENFA, and as pseudo-absence points for GARP BS.

External tests were performed using a new approach based on locality data with positional uncertainty higher than the maximum value used to select training points (500 m). External omission errors and model significance were calculated with this method. All tests were based on binary model responses (presence/absence) using the same thresholds associated with the minimum level of agreement. When binary models are tested with point data, significance can be calculated by means of one-tailed binomial tests with the probability of success being the proportion of the area predicted by the model in relation to the study area (Anderson *et al.* 2002). However, since we are dealing with low precision data represented by polygons, the basic premise for binomial tests to have equal probabilities across all events is not fulfilled. Clearly, the larger the area of a polygon and the wider its range, the bigger is the probability of success (in our case, success is interpreted as any kind of intersection of the polygon with the projected model). When each event to be tested has a different probability of success, the Poisson-binomial test can be used. This is an analogous situation to calculating significance for the leave-one-out procedure (Pearson *et al.* 2007).

Although there is now an efficient way to calculate Poisson-binomial tests (Fernandez & Williams 2010), even for a large number of events, individual probabilities for each polygon had to be approximated with a computing-intensive procedure. As mentioned earlier, records with low precision locality data were already represented by a circle. The remaining localities with high positional uncertainty, i.e., those associated with county polygons, were also approximated here by a minimum bounding circle. With this simplification, the probability of a projected model to intersect with any circle to be tested can be seen as the probability of the centre of the circle falling inside the predicted model area buffered by the circle radius. In other words, individual probabilities were approximated by the corresponding buffered area ratios of the projected model. An important detail here is that, after buffering the model area, the result has to be cropped by the study area since it may have extrapolated its limits. It should also be noted that this approach overestimates p-values, since both county polygons and low precision localities are approximated by a larger geometry.

Species with less than 5 points were not left out of the process. An environmental dissimilarity technique was used to indicate potential areas where the species may occur. Such models were generated with the Environmental Distance algorithm available in openModeller using the Euclidean metric to measure distances in relation

to the centroid of all available presences in the environmental space. An arbitrary limit for the maximum accepted normalized distance was set to 0.1. These models were not converted into binary outputs or even tested – they were generated to serve as preliminary models mainly to guide new fieldwork that may lead to more points and more robust models in the future (Siqueira *et al.* 2009). Models generated in the next two ranges of points (5-19) can theoretically produce more accurate results, but were in principle intended for exploratory modelling following existing recommendations (see Wisz *et al.* 2008). On the other hand, species with at least 20 points were expected to produce robust models potentially useful in conservation.

Finally, environmental profiles were calculated for all species with at least 20 points, showing for each variable how much the projected area covers the environmental conditions of the whole study area and indicating the minimum, maximum and average values based on the training points (figure 11). We envisage this output as being useful not only to better understand the environmental preferences for each species, but also to detect possible omissions, over predictions and outliers.

Results

A total number of 11531 records were retrieved from *speciesLink*, from where 545 records (5%) could only be found by means of synonyms or misspellings. Only 7 records were automatically flagged as being identified with incompatible homonyms. In total, 11210 records (97%) were collected in Brazil and 5487 (48%) had coordinates (figure 5). This number also includes records from one data provider whose coordinates had to be imported from an external file into the database, since all their records were originally marked as having coordinates withheld.



Figure 5: Original points from all species without applying any filters. As expected, most points are concentrated on the east, where human population densities are higher and fieldwork is more frequent.

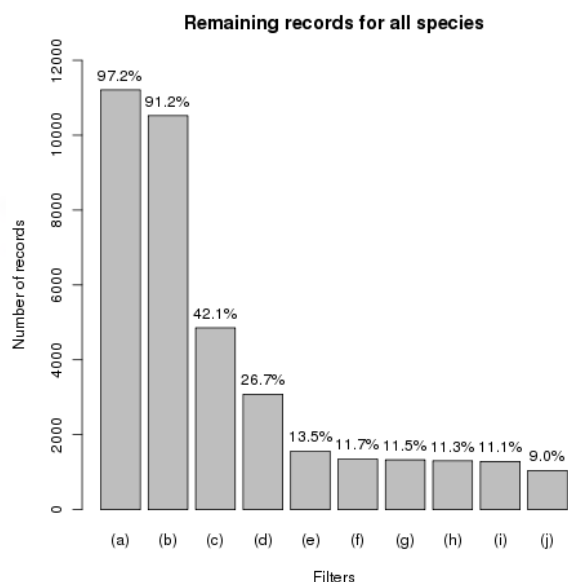


Figure 6: Remaining occurrence records for model training after successively applying filters to all records from all species (n = 11531): (a) Gathered in Brazil, (b) From the wild, (c) With coordinates, (d) Coordinate uncertainty < 500 m, (e) Distance from county reference coordinates > 50m, (f) Removing georeferencing conflicts, (g) Removing identification conflicts and taxonomic issues, (h) Only specimen records, (i) Eliminating duplicates, (j) Unique pixels per species.

Records with coordinates actually added up to 49% of all records collected or observed in Brazil. However, after removing those apparently georeferenced at the county-level, this proportion dropped to 30%. The average coordinate uncertainty error (discarding records georeferenced by county) was 1.66 km ($\sigma = 9.4$), while the biggest uncertainty below the precision threshold adopted here was 324 m. The final selection of training points resulted in 1041 records (9.0%) (figure 6). On the other hand, 2966 records were selected for external tests – almost three times the number of training points.

During the selection process, 16 records were manually flagged as having suspect identification, 7 as having suspect location and 114 as being likely collected from cultivated individuals. These records were included in a list of specimens to be checked and constitute another outcome from the workflow. Only records that were automatically selected for model training or model tests were inspected, therefore manual inspection resulted in 4% of the originally selected records being removed from the modelling process.

Five species (*P. balbis* Feuillet, *P. cryptopetala* Hoehne, *P. margaritae* Sacco, *P. rufa* Feuillet & J.M.MacDougal and *P. reitzii* Sacco) had no records at all in *speciesLink*. Added to the other 42 species whose records did not pass the data quality filters, this group was included in one of the outputs from the workflow: a list of species that should be prioritized in digitization programs and field surveys (annex II). On the opposite side, only two species (*P. cincinnata* Mast. and *P. foetida* L.) had more than 100 selected records (table 1). Discarding the 5 species with no records from *speciesLink*, the average number of records selected per species for model training was 8.

Table 1: Overall results for each range of training points.

No. of training points	No. of species	Average intrinsic omission rate	Average extrinsic omission rate	No. of significant models (p<0.001)
0	47 (35%)	-	-	-
1 to 4	48 (36%)	-	-	-
5 to 9	12 (9%)	0	0.18	8
10 to 19	10 (7%)	0.051	0.32	7
20 or more	18 (13%)	0.019	0.14	17

In most cases, the number of testing localities was much greater than the number of training points (table 2) and *P. mucugeana* Lam. was the only species with no localities available for tests. Almost all models generated with at least 20 points performed significantly better than random ($p < 0.01$), even for species with wider distribution such as *P. coccinea* Aubl. and *P. cincinnata* Mast., and a well-known generalist species such as *P. foetida* L. (table 2). The only exception was *P. misera* Kunth, which is also widely distributed. With only one exception, internal omission rates were below 0.05, and just a few external omission rates were above 0.2, with *P. mucronata* Lam. reaching the highest external omission rate of 0.44. Proportional predicted areas varied from 0.01 to 0.72, reflecting the different distribution patterns. Projected models were also reviewed by an expert and compared with the known distribution of the species in the official checklist of the Brazilian Flora (Forzza *et al.* 2010), showing many similarities (figure 7). This comparison was only made as an additional assessment with no further implications, although it could result in relevant observations that could be attached to each model when making them available to the public. It should also be noted that, when such models are projected into the landscape, the corresponding maps represent

the potential distribution of the species, not the real one. Therefore, comparisons with the real distribution should be made with care – especially regarding over predictions.

In average, models for species with 10-19 points presented higher internal and external omission rates (table 1), but still almost all models were significant (table 2). In many cases projected models showed similarities with the expected distribution (figure 8).

Table 2: Model results ordered by the number of training points. Intrinsic omission indicates the rate of training points not predicted by the model. Extrinsic omission indicates the rate of testing locations not predicted by the model. Predicted area indicates the proportion of the model area in relation to the study area. Median buffered area provides an indication of the external test precision (a value equal to the predicted area would mean that all testing localities were points, while greater values mean a higher number of testing localities with lower precision). P-values are the result of one-tailed Poisson-binomial tests on model prediction for the testing locations.

Species	Training locations	Testing locations	Intrinsic omission rate	Extrinsic omission rate	Predicted area	Median buffered area	P
<i>Passiflora cincinnata</i>	121	184	0.025	0.054	0.415	0.526	< 0.01
<i>Passiflora foetida</i>	107	252	0	0.036	0.726	0.803	< 0.01
<i>Passiflora galbana</i>	56	95	0	0.16	0.091	0.136	< 0.01
<i>Passiflora edmundoi</i>	55	35	0	0.11	0.09	0.146	< 0.01
<i>Passiflora alata</i>	39	130	0.026	0.023	0.33	0.437	< 0.01
<i>Passiflora setacea</i>	39	53	0	0.17	0.166	0.180	< 0.01
<i>Passiflora luetzelburgii</i>	34	10	0.029	0	0.063	0.067	< 0.01
<i>Passiflora amethystina</i>	33	159	0	0.031	0.252	0.361	< 0.01
<i>Passiflora ovalis</i>	31	36	0	0.028	0.026	0.053	< 0.01
<i>Passiflora edulis</i>	29	162	0.034	0.093	0.161	0.254	< 0.01
<i>Passiflora coccinea</i>	28	80	0	0.1	0.529	0.541	< 0.01
<i>Passiflora misera</i>	25	100	0	0.23	0.583	0.736	0.18
<i>Passiflora suberosa</i>	25	107	0	0.22	0.169	0.266	< 0.01
<i>Passiflora recurva</i>	24	30	0.042	0.033	0.047	0.053	< 0.01
<i>Passiflora mucronata</i>	23	62	0.087	0.44	0.01	0.014	< 0.01
<i>Passiflora capsularis</i>	22	139	0	0.16	0.103	0.176	< 0.01
<i>Passiflora jilekii</i>	21	62	0	0.16	0.093	0.165	< 0.01
<i>Passiflora organensis</i>	21	91	0	0.22	0.062	0.140	< 0.01
<i>Passiflora kermesina</i>	18	57	0.056	0.47	0.057	0.12	< 0.01
<i>Passiflora speciosa</i>	18	41	0	0.12	0.086	0.178	< 0.01
<i>Passiflora nitida</i>	17	39	0	0.46	0.331	0.372	0.24
<i>Passiflora pohlii</i>	16	63	0.062	0.24	0.173	0.384	< 0.01
<i>Passiflora haematostigma</i>	15	74	0.2	0.19	0.173	0.389	< 0.01
<i>Passiflora rhamnifolia</i>	11	27	0	0.26	0.077	0.09	< 0.01
<i>Passiflora trintae</i>	11	9	0	0.11	0.029	0.079	< 0.01
<i>Passiflora watsoniana</i>	11	15	0.091	0.33	0.024	0.055	< 0.01
<i>Passiflora auriculata</i>	10	26	0.1	0.54	0.357	0.384	0.44
<i>Passiflora laurifolia</i>	10	6	0	0.5	0.143	0.158	0.08
<i>Passiflora caerulea</i>	9	59	0	0.2	0.059	0.09	< 0.01
<i>Passiflora glandulosa</i>	9	39	0	0.13	0.325	0.334	< 0.01
<i>Passiflora sidifolia</i>	8	24	0	0.083	0.102	0.174	< 0.01
<i>Passiflora mansoi</i>	6	30	0	0.1	0.657	0.709	0.01
<i>Passiflora miersii</i>	6	44	0	0.068	0.347	0.468	< 0.01
<i>Passiflora mucugeana</i>	6	0	0	-	0.003	-	-
<i>Passiflora villosa</i>	6	40	0	0.1	0.126	0.299	< 0.01
<i>Mitostemma brevifilis</i>	5	10	0	0.2	0.323	0.472	0.03
<i>Passiflora acuminata</i>	5	10	0	0.1	0.400	0.433	< 0.01
<i>Passiflora pentagona</i>	5	22	0	0.5	0.003	0.013	< 0.01
<i>Passiflora spinosa</i>	5	5	0	0.2	0.194	0.201	< 0.01
<i>Passiflora vespertilio</i>	5	24	0	0.33	0.483	0.486	0.29

Models with 5-9 training points had zero internal omission (table 2) since they were generated with a single algorithm that produces predictions above zero across the entire study area and was tested against the lowest presence threshold. The average extrinsic omission rate was lower than for the previous range of points. Most models were still significant, although in some cases there was clear over prediction. This can be seen from the proportional predicted area for *P. mansoi* (Mart.) Mast., which covered more than half the Brazilian territory when the species mostly occurs in the central part of Brazil. In some cases, projected models also showed omitted areas when compared with the expected distribution, as it can be seen for *P. caerulea* L. (figure 9). In this specific example, however, it may be possible that the projected model is closer to the real distribution, since this species is largely cultivated as ornamental all over the country.

All localities associated with omission errors (16 related to internal tests and 367 to external tests) constitute another output from the workflow, indicating places where new samples collected with better precision could improve the next models of the species' ecological niche in future workflow iterations. Species with less than 5 points were not tested, but had their models generated also to help guiding new field work (figure 10).

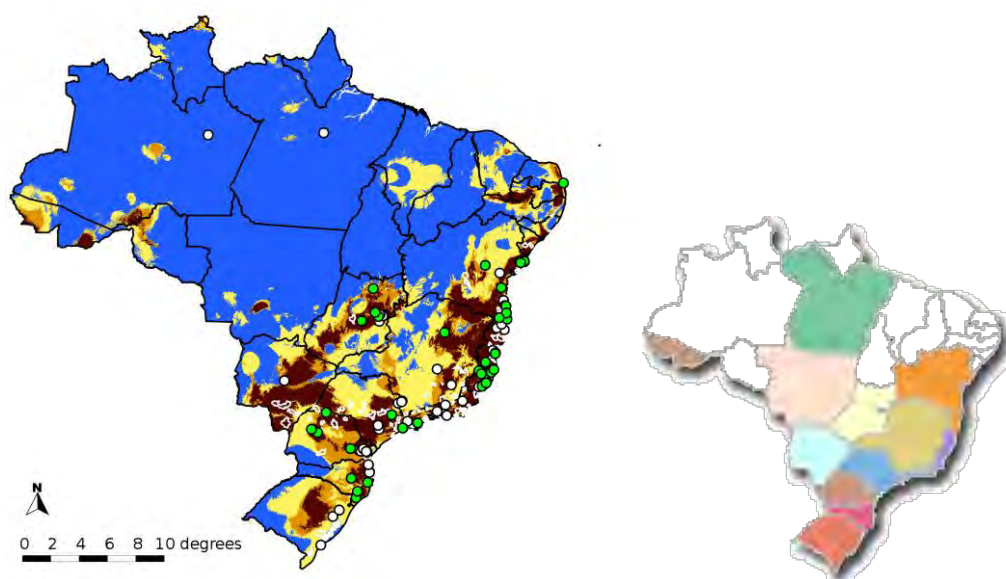


Figure 7: Projected model for *P. alata* Curtis combining ENFA, GARP BS, MD, Maxent and one-class SVM (left side). Dark brown indicates agreement between the five algorithms, dark orange four algorithms, yellow three algorithms and blue no prediction. Green dots are training points while white dots are testing records with low precision and white polygons are testing records representing county boundaries. On the right side is the species distribution by state according to expert knowledge (Cervi *et al.* 2010).

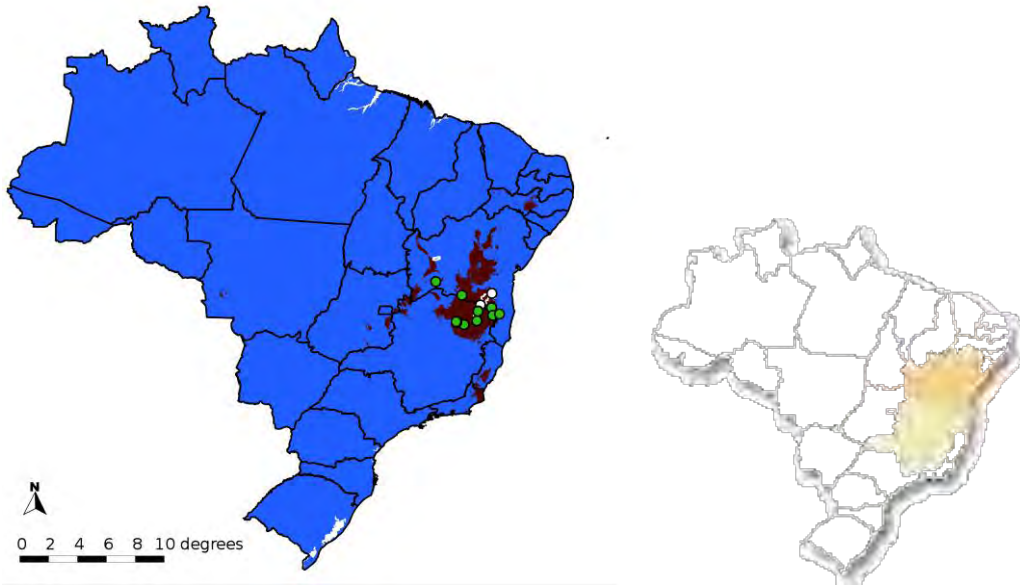


Figure 8: Projected model for *P. trintaе* Sacco combining GARP BS and Maxent (left side). Dark brown indicates agreement between the two algorithms and blue no prediction. Green dots are training points while white dots are testing records with low precision and white polygons are testing records representing county boundaries. On the right is the species distribution by state according to expert knowledge (Cervi *et al.* 2010).

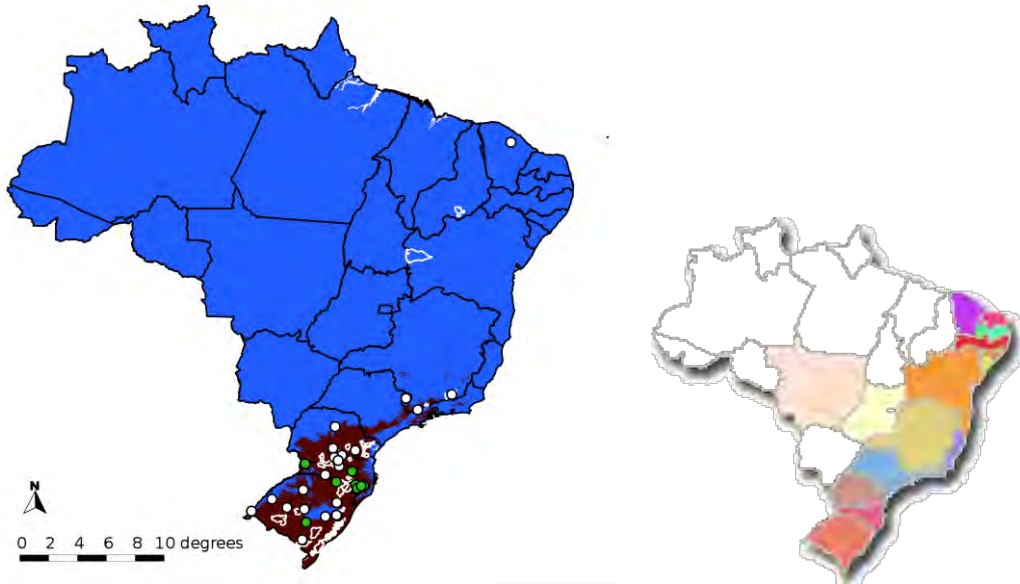


Figure 9: Projected model for *P. caerulea* L. using only Maxent and converting the result into a binary response using LPT (left side). Dark brown indicates suitable areas and blue no prediction. Green dots are training points while white dots are testing records with low precision and white polygons are testing records representing county boundaries. On the right is the species distribution by state according to expert knowledge (Cervi *et al.* 2010).

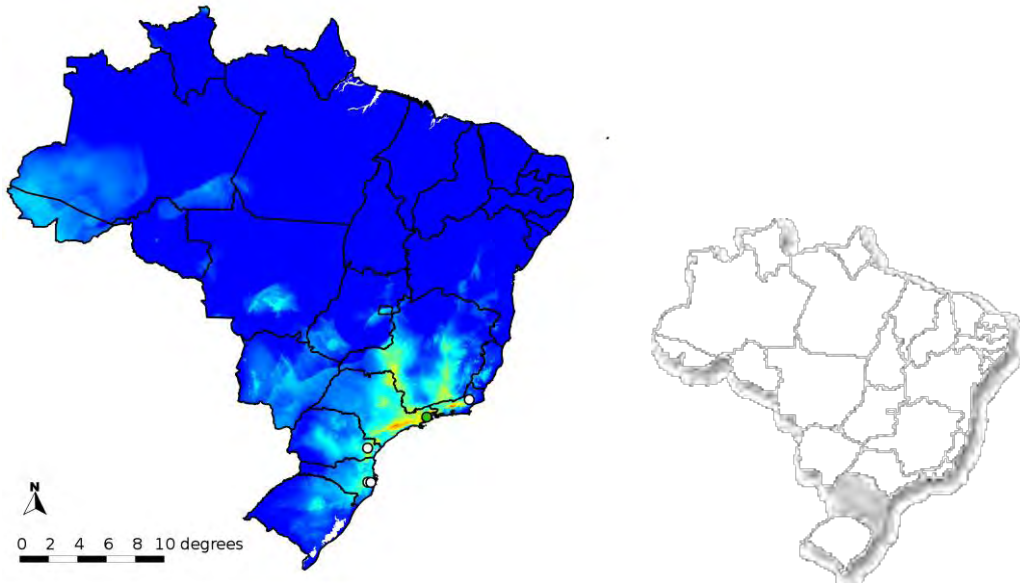


Figure 10: Projected model for *P. catharinensis* Sacco using Euclidean distance (left side). Red indicates high environmental similarity and dark blue indicates no prediction, Green dots are training points while white dots are testing records with low precision. On the right is the species distribution by state according to expert knowledge (Cervi *et al.* 2010).

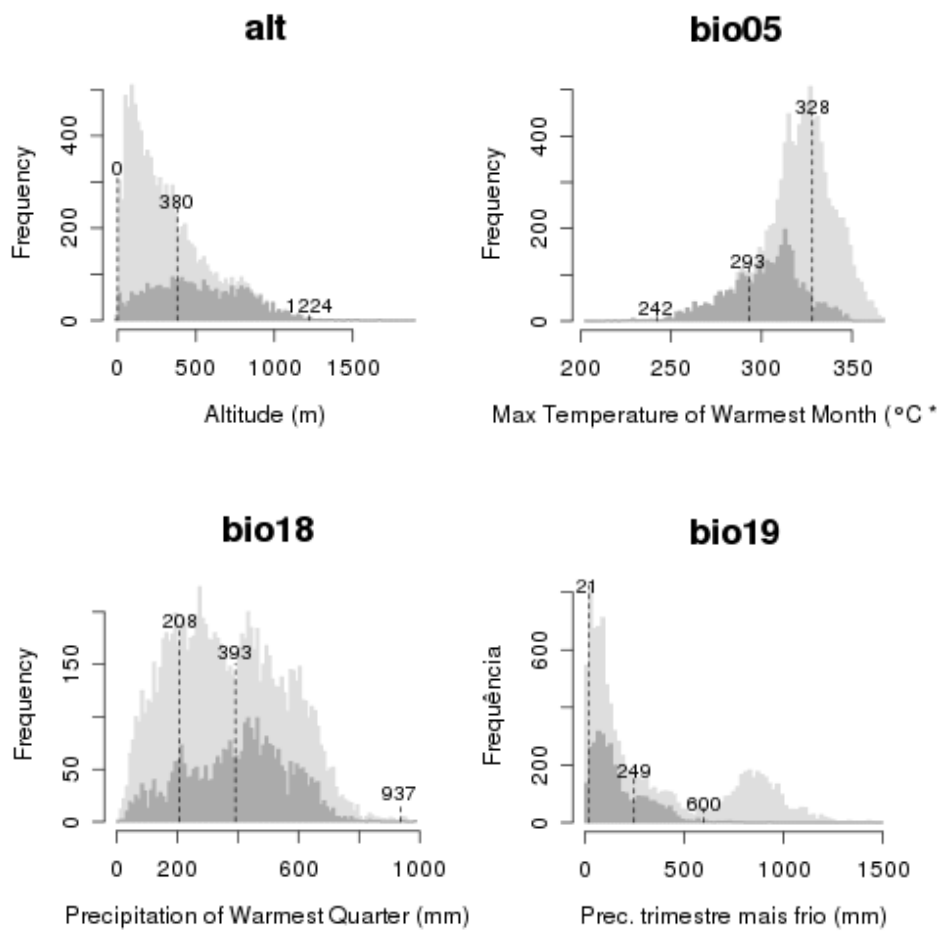


Figure 11: Environmental profile for *P. alata* Curtis using only four variables. Light grey shows the histogram for the whole study area and dark grey shows the histogram for the projected model area. Minimum, maximum and average values based on the training points are also displayed.

Discussion

The expressive proportion of records discarded for model training corroborates the importance of data quality filters when using data from biological collections – especially when models are generated with only a few points. In this circumstance, model stability can be an issue, since each point tends to exert greater influence in the final result (Wisniewski *et al.* 2008). Interestingly, the use of more restrictive filters did not result in a smaller number of records comparing with a recent study for tropical plants. After analyzing data from biodiversity data networks, Feeley & Silman (2011) also found an average of ~8 usable spatially unique points for ecological niche modelling per species in South America for the same resolution. Similarities can also be observed in the number of species per range of points: 33%, 9.4%, 2.7% and 0.7% of the species had 5, 20, 50, 100 or more points respectively (Feeley & Silman 2011), while here we found 31.5%, 13.8%, 3% and 1.5% (percentages calculated without species with no records in *speciesLink*). However, there are many methodological differences between the two studies, so this comparison can only serve as an initial reference. It may be the case that species of Passifloraceae could have been more frequently collected than most species from other families (Passifloraceae is currently ranked as the 36th most collected botanical family in *speciesLink*). For a better comparison and before making generalisations it would be necessary to extend this work to other families.

The use of synonyms and misspellings to retrieve data resulted in a relatively low number of additional records. However, for some species this approach made the difference between having an initial model or not, or having a better model built with more algorithms. For example, the single point used in the environmental dissimilarity model of *P. urnifolia* Rusby came from an orthographic variant. Four of the five points used to model *P. pentagona* Mast. came from synonyms, as well as 30 out of 31 points used to model *P. ovalis* Vell. ex M.Roem.

Among the species with more than 20 points, there were not just common ones with wide distribution. Species such as *P. ovalis* Vell. ex M.Roem. and *P. mucronata* Lam. inhabit a restricted range near the coast in the Atlantic Forest and may be of conservation concern, while *P. luetzelburgii* Harms was only known from its type specimen until recently (Vitta & Bernacci 2004). Together with *P. galbana* Mast. and *P. edmundoi* Sacco, the distribution of *P. luetzelburgii* Harms is mainly delimited by one of the most threatened biomes of the country: the Caatinga (Brazilian semi-arid).

Even being fairly restrictive, the set of filters used by the workflow still do not encompass other known data cleaning techniques, such as checking for a possible conflict between the provided altitude and the corresponding value from a Digital Elevation Model at the given coordinates (Chapman 2005), looking for inconsistencies in collectors' itineraries (Peterson *et al.* 2004) and verifying the species kingdom (Heap & Culham 2010). This last one is of particular importance for workflows with a broader taxonomic coverage, since there can be homonyms across different kingdoms.

The process of selecting records for model training resulted in another outcome that can be of special importance to conservation: the list of species that did not pass all data quality filters (annex II). This can help prioritizing digitization programs and field surveys. Most species that are rare or have a narrow distribution will likely be in this list, as can be inferred from the work of Du Puy *et al.* (2002), who found that 82% of the species that were of conservation concern had less than 10 specimen records after assessing the conservation status of all woody endemic legumes of Madagascar.

As expected, the automatic filters could not detect problems in all situations. For this reason, it was important to provide ways for human intervention during the process, allowing an expert to indicate possible misidentifications, geographic errors and whether specimens were likely collected from cultivated areas. By keeping these annotations in an independent database, the workflow could proceed based on initial assumptions that can be verified later and reported back to the original collections.

Even with all automatic filters and human interventions, it is possible that points apparently having enough precision do not correspond to the original locality. This may be caused by digitization errors or by retrospectively georeferencing based on another nearby location, such as a farm or a village. If this is not detected, algorithms may be trained with inaccurate environmental conditions, potentially degrading model performance. However, all the following conditions must be satisfied for this to happen: 1) Such records must pass all data quality filters (note that digitization errors can be a source of geographical conflicts and records may also be indirectly discarded by other filters); 2) Points must be associated with an environmental condition clearly unsuitable for the species being modelled (note that positional inaccuracy may not be a problem in regions where environmental conditions are essentially the same across wide areas, as it happens in many parts of the Amazon region); 3) Records must also pass human inspection; 4) In some cases, it may be necessary for more than one inaccurate record to pass all previous conditions in order to degrade model performance, since algorithms such as GARP, SVM and MD can omit training points in the final result.

The approach used to select training points is not just consistent with the aim of generating models in high resolution, but it also followed a previous indication that quality (meaning spatial resolution and positional accuracy of input data) seems to be more important than the number of training points (Engler *et al.* 2004). Additionally, despite discarding many records due to positional uncertainty, these were all used to perform external tests with a new approach. Even records without coordinates were used, as long as the county could be interpreted. This approach proved to be useful for almost all species that were tested, except one that did not have any testing records available. In such cases, it should be possible to use existing techniques based on training points (n-fold cross validation or leave-one-out tests). These techniques may also be the only alternative if the available testing localities have a comparable or even larger size than the predicted distribution. Species like *P. ischnoclada* Harms and *P. hatschbachii* Cervi will likely be in this situation in the future since they are only known to occur in a small area – possibly being micro-endemic. For now, both species did not have enough points to generate a testable model.

An important aspect regarding the new model validation is that when a predicted area partially intersected with a polygon being tested, this was considered a success (i.e. if the specimen was collected somewhere in a certain polygon and the model predicts suitable conditions for the species in at least part of that polygon, this was considered a successful prediction). In such cases, it may well be possible that the original precise location of the collecting event was not inside the intersection, and in theory it is also possible that the intersection actually contains unsuitable conditions for the species. Given the level of uncertainty involved, any conclusions about the correctness of model predictions inside the polygon being tested should be regarded as arbitrary. Only in extreme situations – if the entire polygon is inside the predicted area or if the entire polygon is outside the predicted area – could we draw a precise conclusion. Under partial intersections, the only way to check if model predictions are correct in a finer scale is with additional field work, at the risk of not finding the species anymore due to habitat loss or other reasons. Hence, model significance should be viewed as the

probability of attaining each success at the level of uncertainty of the corresponding locality being tested.

Test results provided a good indication that, perhaps with only one or two exceptions, models that were generated with at least 20 points can be potentially useful for conservation. Additionally, although being in a condition originally intended for exploratory modelling, some of the species with less than 20 points also produced models with low omission errors and high significance. In cases where omission rates are considered too high, each locality associated with an internal or external omission can be used to guide new field work. As more records become available, new models with a more complete sampling of the species distribution can be generated in further workflow iterations. Another approach to overcome high omission errors could be to gradually decrease individual model thresholds (or the final cut off threshold in the consensus model) until an acceptable external omission rate is achieved.

Although the workflow presented here tries to be comprehensive enough to handle data scarcity and to generate more reliable results as each species reaches a certain data-availability condition, some simplifications were made which can be the object of future improvements. Using a single study area for all species when each may have quite distinct distribution patterns is something that can both influence the generated models and the corresponding tests. Giovanelli *et al.* (2010) show that even keeping the same input data, algorithms may still produce different results depending on the extent of the study area, and that algorithms such as GARP are more vulnerable to this effect. In this case, broader study areas may result in broader predicted areas, especially for species with narrower distributions. On the other hand, the approach used here for species with at least 10 training points will only produce over predictions if the majority of algorithms over predict the same areas.

The influence of the study area in model tests is well-known and affects all sorts of model validation that depend on the proportion of predicted areas (Lobo *et al.* 2008). In our case, model significance is directly influenced, and two cases illustrate this fact. The model for *P. mucronata* Lam. produced one of the smallest predicted areas (table 2) because the species is only found near the coast. Although the external omission rate was high, model significance was also high ($p < 0.01$) since each success when testing a locality has a very low associated probability. On the opposite situation, generalist species such as *P. misera* Kunth produce larger predicted areas with a high probability of success for each locality to be tested, resulting in poorer significance – which does not necessarily mean a bad model. Additionally, the new test based on lower precision data introduces another factor that must be taken into account. If most localities tested have very low precision, the overall test becomes weaker, which also degrades significance. The extent of this effect can be estimated by looking at the individual probabilities of each locality to be tested (i.e., the proportional predicted buffered areas). A high number of localities with very low precision will increase the median buffered area proportions (table 2). This may also explain the poor significance for the model of *P. misera* Kunth, since the median buffered area is considerably higher than the original predicted area. For this reason, model statistics need to be carefully interpreted before making decisions.

Still regarding the influence of study areas, although there are recommendations to build models using datasets with 50% of prevalence, particularly if optimal niche thresholds are being sought (Liu *et al.* 2005), a recent study indicates that this area should ideally be the area that has been accessible to the species via dispersal during the past (Barve *et al.* 2011). Estimating this area, however, is currently a complex

process since it depends on knowledge about when and where each species originated, dispersal characteristics, geography of the area of interest and the corresponding environmental changes over the relevant time period (Barve *et al.* 2011). Very few studies exist about the biogeographical history of Passifloraceae. Muschner (2005) estimated that *Passiflora*, the most representative genus of the family, started to diversify around 42 million years ago. This suggests that most species considered in this work are probably very old with a complex biogeographical history during which some of them managed to colonize vast regions of the country, while others were confined to very specific areas. On the one hand, the almost continental dimensions of Brazil likely contain sufficiently diverse environmental conditions allowing efficient sampling of the species' niches. This is reinforced by the fact that 88 species (65%) are considered endemic to the country (Cervi *et al.* 2010). On the other hand, such dimensions may overestimate the accessible area for species with narrower distributions. This is an important subject for further investigation which may lead to more efficient methods to determine different study areas for each species, preferably in a way that could be automated based on existing data.

Another decision was to use a single set of layers for all species. It may be possible that for certain species some of the layers that were discarded in the correlation analysis could provide additional information to algorithms, since the magnitude of each detected correlation can be different in different parts of the study area. This is also related to the fact that different species may have different climatic factors limiting their distribution. One possible improvement in this respect could be to perform a pre-analysis on the available layers for each species, potentially selecting a different set of layers in each situation.

The choice of using a single study area with such dimensions and a single set of environmental layers also brings additional constraints. One known limitation regarding environmental layers is that no soil data was used since it was not available for the whole study area. This is of special concern for species whose distribution could be mainly constrained by edaphic conditions. It is likely that continuous soil data in the desired resolution could be available for specific parts of the country that may sufficiently cover the entire distribution of certain species. In any case, new environmental layers can be used in the workflow as more data becomes available in the future.

Finally, all models were generated with the same parameter values for each algorithm, in most cases using default settings. Although default parameter values tend to get consolidated over the time as each algorithm is used under different conditions, there can be some margin to improve model performance by tuning parameters (see Guo *et al.* 2005; Phillips & Dudík 2008 for examples). Therefore, in some cases the workflow could possibly generate better results with an additional step for tuning specific parameters.

As a last remark, restricting the scope to specimen data did not lead to a significant impact on the final number of selected records. The reason is that currently only one data provider connected to *speciesLink* contains plant observation data. If a similar workflow is used for other taxonomic groups with other biodiversity data networks, this can make a big difference. For example, bird watchers around the world already provide many millions of records through online networks (Devictor *et al.* 2010). On the other hand, the choice of using only biological collections' data adds more transparency and reliability to the process, since each record used to generate the ecological niche models is backed by a specimen that can be examined and re-identified whenever

necessary. This choice also offers advantages in the long-term, since specimen curation tends to keep records up-to-date with the current taxonomy (Graham *et al.* 2004).

Conclusions

Data from natural history collections is certainly prone to all sorts of errors and uncertainties, deserving special care to be used. At the same time, revising and fixing specimen data is a major task that can be greatly optimized for specific purposes – in this case for understanding and predicting species' distributions. By applying several restrictive filters on herbarium data, attention can be focused in a much smaller number of records with a greater potential to be used in ecological niche modelling.

Considering previous estimates that natural history collections have a low number of records digitally available, a small proportion of georeferenced records and an even smaller proportion of records that are suitable for ecological niche modelling, the use case presented here can be a sign that this scenario may be changing. Even including quite restrictive filters it is already possible to generate ecological niche models in high resolution with potential use in conservation for many species – not just widespread and well-known species. A new test using lower precision data was also presented, offering a new opportunity to measure model quality through external tests using more records from natural history collections that would otherwise be discarded for high-resolution experiments.

Effective workflows can be designed to carry out these tasks, and recent advances in biodiversity informatics allow most part of the process to be automated. Nevertheless, human intervention is still important and can also be facilitated by existing technologies, as can be seen from the process to build the checklist of the Brazilian Flora – a collaborative work that involved more than 400 taxonomists during less than 2 years to create a reference list of 40982 species through a Web application (Forzza *et al.* 2010). Similar efforts can certainly be carried out to produce additional results, such as participating in the generation of ecological niche models for many of those species.

The need of expert review and some of the functionalities of the data filters are also related to the fact that many collections' management software do not offer the possibility to indicate more detailed positional data, such as coordinate uncertainty and horizontal datum for GPS measurements, as well as to indicate more details about the collecting event in a structured way (such as whether specimens were collected from the wild or not). Existing standards for data exchange such as Darwin Core already include specific terms for all these purposes, but they can only be used if collectors include the corresponding data when depositing new materials and if the original data providers allow this information to be stored in a standard format.

The workflow presented here does not claim or intend to produce better ecological niche models than those produced individually by researchers who frequently have access to more data about the distribution of the species and who work with a wider range of tools for more specific goals. Instead, it intends to provide a platform for collaboration capable of generating models with enough quality for a large number of species based on data that is publicly available.

The approach proposed here not only includes specific strategies to ensure data quality control (both input and output), but also to deal with low number of training points, use as much data as possible to test models, facilitate usage by decision makers and provide feedback to the original data providers. Lack of data can be regarded as an

opportunity to prioritize digitization and research, and data quality issues as an opportunity to improve our knowledge. The proactive nature of the workflow presented here may also offer opportunities for funding agencies, policy makers and researchers to better direct resources for conservation. Although the workflow is still not implemented as a full-featured application and the use case involves a single family, the concept is intended to serve as the basis for a future system dedicated to more families of the Brazilian flora and it may also provide useful references and insights to similar initiatives.

Acknowledgments

We would like to thank Flávia Santos Pinto and A. Townsend Peterson for all their comments and suggestions on the manuscript, and the following collections that contributed with data to this work:

Banco de DNA do Jardim Botânico do Rio de Janeiro
Carpoteca do Instituto Nacional de Pesquisas da Amazônia
Carpoteca da Universidade Federal de Pernambuco
Coleção de Fanerógamas do Herbário do Estado "Maria Eneyda P. Kaufmann Fidalgo"
Herbário do Instituto Nacional de Pesquisas da Amazônia
Herbário do Museu de Biologia Professor Mello Leitão
Herbário Central da Universidade Federal do Espírito Santo
Herbário da Escola Superior de Agricultura Luiz de Queiroz
Herbário Dárdano de Andrade Lima
Herbário da Universidade de Brasília
Herbário da Universidade Estadual de Campinas
Herbário da Universidade Estadual de Feira de Santana
Herbário da Universidade Estadual de Londrina
Herbário da Universidade Federal de Sergipe
Herbário da Universidade Federal do Acre
Herbário da Universidade Tecnológica Federal do Paraná Campus Campo Mourão
Herbário de São José do Rio Preto
Herbário Dimitri Sucre Benjamin
Herbário do Departamento de Botânica da Universidade Federal do Paraná
Herbário do Instituto Agronômico de Campinas
Herbário do Instituto do Meio Ambiente do Estado de Alagoas
Herbário Dom Bento Pickel
Herbário do Museu Botânico Municipal de Curitiba
Herbário do Parque da Ciência Newton Freire Maia
Herbário do Trópico Semiárido
Herbário Dr. Roberto Miguel Klein
Herbário Lauro Pires Xavier
Herbário Prisco Bezerra
Herbário Sérgio Tavares
Herbário da Universidade Estadual de Maringá
Herbário da Universidade Federal do Rio Grande do Norte
Herbário Vale do São Francisco
Herbier Virtuel Saint-Hilaire
Missouri Botanical Garden
Smithsonian Department of Botany
The New York Botanical Garden

References

- Anderson R.P., Gómez-Laverde, M., Peterson, A.T., 2002. Geographical distributions of spiny pocket mice in South America: insights from predictive models. *Global Ecology & Biogeography*, 11 (2): 131-141.
- Anderson R.P., Lew, D., Peterson, A.T., 2003. Evaluating predictive models of species' distributions: criteria for selecting optimal models. *Ecological Modelling*, 162: 211–232.
- Araújo, M.B., New, M., 2007. Ensemble forecasting of species distributions. *Trends in Ecology & Evolution*, 22 (1): 42-47. doi:10.1016/j.tree.2006.09.010.
- Ariño, A.H., 2010. Approaches to estimating the universe of natural history collections data. *Biodiversity Informatics*, 7 (2): 81-92.
- Barnosky, A.D., Matzke, N., Tomiya, S., Wogan, G.O.U., Swartz, B., Quental, T.B., Marshall, C., McGuire, J.L., Lindsey, E.L., Maguire, K.C., Mersey, B., Ferrer, E.A., 2011. Has the Earth's sixth mass extinction already arrived? *Nature*, 471: 51-57. doi:10.1038/nature09678.
- Barve, N., Barve, V., Jiménez-Valverde, A., Lira-Noriega, A., Maher, S.P., Peterson, A.T., Soberón, J., Villalobos, F., 2011. The crucial role of the accessible area in ecological niche modelling and species distribution modelling. *Ecological Modelling*, 222: 1810-1819. doi: 10.1016/j.ecolmodel.2011.02.011.
- Beaman, R.S., Conn, B.J., 2003. Automated geoparsing and georeferencing of Malesian collection locality data. *Telopea*, 10 (1): 43-52.
- Bernacci, L.C., 2003. Passifloraceae, in: Wanderley, M.G.L., Shepherd, G.J., Giulietti, A.M., Melhem, T.S. (Coord.) *Flora Fanerogâmica do Estado de São Paulo*, 3: 247-248. RiMa/FAPESP, São Paulo.
- Canhos, V.P., Souza, S., Giovanni, R., Canhos, D.A.L., 2004. Global Biodiversity Informatics: Setting the scene for a "New World" of ecological modelling. *Biodiversity Informatics*, 1: 1-13.
- Cervi, A.C., 1997. Passifloraceae do Brasil: estudo do gênero *Passiflora* L., subgênero *Passiflora*. *Fontqueria*, 45: 1-92.
- Cervi, A.C., Milward-de-Azevedo, M.A., Bernacci, L.C., 2010. *Passifloraceae in Lista de Espécies da Flora do Brasil*. Jardim Botânico do Rio de Janeiro (<http://floradobrasil.jbrj.gov.br/2010/FB000182>).
- Chapman, A.D., 2005. Principles and methods of data cleaning – Primary species and species occurrence data, version 1.0. Report for the Global Biodiversity Information Facility, Copenhagen, 72 pp.
- Chapman, A.D., Busby, J.R., 1994. Linking plant species information to continental biodiversity inventory, climate and environmental monitoring, in: Miller, R.I. (Ed.) *Mapping the Diversity of Nature*, 177-195, London: Chapman and Hall.

Devictor, V., Whittaker, R.J., Beltrame, C., 2010. Beyond scarcity: citizen science programmes as useful tools for conservation biogeography. *Diversity and Distributions* 16 (3): 354–362.

Diniz-Filho, J.A.F., Bini, L.M., Rangel, T.F., Loyola, R.D., Hof, C., Nogués-Bravo, D., Araújo, M.B., 2009. Partitioning and mapping uncertainties in ensembles of forecasts of species turnover under climate change. *Ecography*, 32: 897-906. doi:10.1111/j.1600-0587.2009.06196.x.

Drake, J.M., Randin, C., Guisan, A., 2006. Modelling ecological niches with support vector machines. *Journal of Applied Ecology*, 43: 424-432. doi:10.1111/j.1365-2664.2006.01141.x.

Du Puy, D.J., Labat, J.-N., Rabevohitra, R., Villiers, J.-F., Bosser, J., Moat, J., 2002. *The Leguminosae of Madagascar*. Kew, UK: Royal Botanic Gardens.

Egler, I. & Santos, M.M. (Coord.), 2006. Diretrizes e estratégias para a modernização de coleções biológicas brasileiras e a consolidação de sistemas integrados de informação sobre biodiversidade. MCT/CGEE. Brasília.

Elith, J., Graham, C.H., 2009. Do they? How do they? WHY do they differ? On finding reasons for differing performances of species distribution models. *Ecography* 32 (1): 66–77. doi:10.1111/j.1600-0587.2008.05505.x.

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, G., Nakamura, M., Nakazawa, Y., Overton, J.McC., Peterson, A.T., Phillips, S.J., Richardson, K., Scachetti-Pereira, R., Schapire, R.E., Soberón, 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.

Farber, O., Kadmon, R., 2003. Assessment of alternative approaches for bioclimatic modeling with special emphasis on the Mahalanobis distance. *Ecological Modelling*, 160: 115–130.

Feeley, K.J., Silman, M.R., 2011. The data void in modelling current and future distributions of tropical species. *Global Change Biology*, 17: 626–630. doi:10.1111/j.1365-2486.2010.02239.x.

Fernandez, M., Blum, S., Reiche, S., Guo, Q., Holzman, B., Hamilton, H., 2009. Locality uncertainty and the differential performance of four common niche-based modelling techniques. *Biodiversity Informatics*, 6: 36-52.

Fernandez, M., Williams, S., 2010. Closed-form expression for the Poisson-Binomial probability density function. *IEEE Transactions on Aerospace and Electronic Systems*, 46 (2): 803-817. doi: 10.1109/TAES.2010.5461658.

Forzza, R.C., Leitman, P.M., Costa, A.F., Carvalho Jr., A.A., Peixoto, A.L., Walter, B.M.T., Bicudo, C., Zappi, D., Costa, D.P., Lleras, E., Martinelli, G., Lima, H.C., Prado, J., Stehmann, J.R., Baumgratz, J.F.A., Pirani, J.R., Sylvestre, L., Maia, L.C., Lohmann,

L.G., Queiroz, L.P., Silveira, M., Coelho, M.N., Mamede, M.C., Bastos, M.N.C., Morim, M.P., Barbosa, M.R., Menezes, M., Hopkins, M., Secco, R., Cavalcanti, T.B., Souza, V.C., 2010. Introdução, in: Lista de Espécies da Flora do Brasil. Jardim Botânico do Rio de Janeiro.

Franz, N.M., Peet, R.K., Weakley, A.S., 2008. On the use of taxonomic concepts in support of biodiversity research and taxonomy, in: Wheeler, Q.D. (Ed.) The New Taxonomy. Systematics Association Special Volume (74): 63–86, Boca Raton, FL: Taylor & Francis.

Gale, J., 2004. Plants and Altitude – Revisited. *Annals of Botany*, 94 (2), 199.

Giovanelli, J.G.R., Siqueira, M.F., Haddad, C.F.B., Alexandrino, J., 2010. Modeling a spatially restricted distribution in the Neotropics: How the size of calibration area affects the performance of five presence-only methods. *Ecological Modelling*, 221: 215–224. doi: 10.1016/j.ecolmodel.2009.10.009.

Giovanni, R., Copp, C., Döring, M., Güntsch, A., Vieglais, D., Hobern, D., Torre, J., Wieczorek, J., Gales, R., Hyam, R., Blum, S., Perry, S., 2010. TAPIR – TDWG Access Protocol for Information Retrieval, Protocol Specification, Version 1.0. http://www.tdwg.org/dav/subgroups/tapir/1.0/docs/tdwg_tapir_specification_2010-05-05.htm (accessed 2011-01-13)

Graham, C.H., Elith, J., Hijmans, R., Guisan, A., Peterson, A.T., Loiselle, B., 2008. The influence of spatial errors in species occurrence data used in distribution models. *Journal of Applied Ecology*, 45: 239-247.

Graham, C.H., Ferrier, S., Huettman, F., Moritz, C., Peterson, A.T., 2004. New developments in museum-based informatics and applications in biodiversity analysis. *Trends in Ecology and Evolution*, 19 (9).

Grinnell, J., 1917. Field tests of theories concerning distributional control. *American Naturalist*, 51: 115-128.

Guisan, A., Zimmermann, N.E., Elith, J., Graham, C.H., Phillips, S.J., Peterson, A.T., 2007. What matters for predicting the occurrences of trees: techniques, data, or species' characteristics? *Ecological Monographs*, 77 (4): 615–630.

Guo, Q., Kelly, M., Graham, C.H., 2005. Support vector machines for predicting distribution of Sudden Oak Death in California. *Ecological Modelling*, 182: 75-90. doi:10.1016/j.ecolmodel.2004.07.012.

Guo, Q., Liu, Y., 2010. ModEco: an integrated software package for ecological niche modelling. *Ecography*, 33: 637-642. doi:10.1111/j.1600-0587.2010.06416.x.

Guralnick, R.P., Hill, A.W., Lane, M., 2007. Towards a collaborative, global infrastructure for biodiversity assessment. *Ecology Letters*, 10 (8): 663–672. doi:10.1111/j.1461-0248.2007.01063.x.

Guralnick, R.P., Wieczorek, J., Beaman, R., Hijmans, R.J. & the BioGeomancer Working Group. 2006. BioGeomancer: Automated georeferencing to map the world's biodiversity data. *PLoS Biology*, 4 (11): e381. doi:10.1371/journal.pbio.0040381.

Heap, M.J., Culham, A., 2010. Automated Pre-processing Strategies for Species Occurrence Data Used in Biodiversity Modelling, in: Setchi, R., Jordanov, I., Howlett, R.J., Jain, L.C. (Eds.) Knowledge-Based and Intelligent Information and Engineering Systems: Lecture Notes in Computer Science, 6279, 517-526. SpringerLink. doi: 10.1007/978-3-642-15384-6.

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., Cameron, S.E., Parra, J.L., Jones, P.G., Jarvis, A., 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25: 1965–1978.

Hirzel, A.H., Hausser, J., Chessel, D., Perrin, N., 2002. Ecological-niche factor analysis: How to compute habitat-suitability maps without absence data? *Ecology*, 83 (7): 2027–2036.

Hirzel, A., Hausser, J., Perrin, N., 2011. Biomapper online documentation. Lausanne, Lab. for Conservation Biology. <http://www2.unil.ch/biomapper/faq.html> (accessed 2011-05-11)

IBGE, 2003. Base cartográfica integrada digital do Brasil ao milionésimo.

IBGE, Produção Agrícola Municipal, 2011. Sidra: <http://www.sidra.ibge.gov.br/bda/agric/default.asp?t=2> (accessed 2011-03-01)

Kadmon, R., Farber, O., Danin, A., 2004. Effect of roadside bias on the accuracy of predictive maps produced by bioclimatic models. *Ecological Applications*, 14 (2): 401–413.

Kaschner, K., Ready, J.S., Agbayani, E., Rius, J., Kesner-Reyes, K., Eastwood, P.D., South, A.B., Kullander, S.O., Rees, T., Close, C.H., Watson, R., Pauly, D., Froese, R., 2008. AquaMaps: Predicted range maps for aquatic species. World wide web electronic publication, www.aquamaps.org, Version 08/2010.

Kennedy, J., Kukla, R., Paterson, T., 2005. Scientific names are ambiguous as identifiers for biological taxa: Their context and definition are required for accurate data integration, in: Ludäscher, B., Raschid, L. (Eds.) *Data Integration in the Life Sciences: Proceedings of the Second International Workshop*, 3615, 80–95, San Diego, CA: SpringerLink. doi:10.1007/11530084_32.

Killip, E.P., 1938. The American species of Passifloraceae. *Field Museum of Natural History, Botanical Series*, 49: 1-613.

Liu, C., Berry, P.M., Dawson, T.P., Pearson, R.G., 2005. Selecting thresholds of occurrence in the prediction of species distributions. *Ecography*, 28 (3): 385-393.

Lobo, J. M., Jimenez-Valverde, A., Real, R., 2008. AUC: a misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography*, 17: 145–151.

Lorena, A.C., Jacintho, L.F.O., Siqueira, M.F., Giovanni, R., Lohmann, L.G., Carvalho, A.C.P.L.F., Yamamoto, M., 2011. Comparing Machine Learning Classifiers in Potential

Distribution Modelling. *Expert Systems with Applications*, 38 (5): 5268-5275. doi:0.1016/j.eswa.2010.10.031.

Lorena, A.C., Siqueira, M.F., Giovanni, R., Carvalho, A.C.P.L.F., Prati, R.C., 2008. Potential Distribution Modelling Using Machine Learning, in: Nguyen, N.T., Borzemski, L., Grzech, A., Ali M. (Eds.) *New Frontiers in Applied Artificial Intelligence: Lecture Notes in Computer Science*, 5027, 255-264, Wroclaw, Poland: SpringerLink. doi: 10.1007/978-3-540-69052-8.

Marmion, M., Parviainen, M., Luoto, M., Heikkinen, R.K., Thuiller, W., 2009. Evaluation of consensus methods in predictive species distribution modelling. *Diversity and Distributions*, 15 (1): 59–69. doi:10.1111/j.1472-4642.2008.00491.x

Mateo, R.G., Felicísimo, Á.M., Muñoz, J., 2010. Effects of the number of presences on reliability and stability of MARS species distribution models: the importance of regional niche variation and ecological heterogeneity. *Journal of Vegetation Science*, 21: 908–922. doi:10.1111/j.1654-1103.2010.01198.x

Michener, W., Beach, J., Bowers, S., Downey, L., Jones, M., Ludäscher, B., Pennington, D., Rajasekar, A., Romanello, S., Schildhauer, M., Vieglais, D., Zhang, J., 2005. SEEK: Data Integration and Workflow Solutions for Ecology, in: Ludäscher, B., Raschid, L. (Eds.) *Data Integration in the Life Sciences: Proceedings of the Second International Workshop*, 3615, 321–324, San Diego, CA: SpringerLink. doi:10.1007/11530084_32.

MMA, 2008. Instrução Normativa MMA nº 06, September 23, 2008.

Muñoz, M.E.S., Giovanni, R., Siqueira, M.F., Sutton, T., Brewer, P., Pereira, R.S., Canhos, D.A.L., Canhos, V.P., 2011. openModeller: a generic approach to species' potential distribution modelling. *Geoinformatica*, 15: 111-135.

Muschner, V.C., 2005. “Filogenia molecular, taxas evolutivas, tempo de divergência e herança organelar no gênero *Passiflora* L. (Passifloraceae)”, PhD thesis, Universidade Federal do Rio Grande do Sul, Porto Alegre.

Pahwa, J.S., Brewer, P., Sutton, T., Yesson, C., Burgess, M., Xu, X., Jones, A.C., White, R.J., Gray, W.A., Fiddian, N.J., Bisby, F.A., Culham, A., Caithness, N., Scoble, M., Williams, P., Bhagwat, S., 2006. Biodiversity World: A Problem-Solving Environment for Analysing Biodiversity Patterns, in: Proc. 6th IEEE International Symposium on Cluster Computing and the Grid (CCGRID 2006), Singapore.

Pearson, R.G., Raxworthy, C.J., Nakamura, M., Peterson, A.T., 2007. Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. *Journal of Biogeography*, 34: 102–117.

Pennington, D.D., Higgins, D., Peterson, A.T., Jones, M.B., Ludäscher, B., Bowers, S., 2007. Ecological Niche Modelling Using the Kepler Workflow System, in: Taylor, I.J., Deelman, E., Gannon, D.B., Shields, M. (Eds.), *Workflows for e-Science*, 91-108, Springer, Berlin.

Peterson, A.T., Navarro-Sigüenza, A.G., Pereira, R.S., 2004. Detecting errors in biodiversity data based on collectors' itineraries. *Bulletin of the British Ornithologists' Club*, 124: 143-151.

- Peterson, A.T., Papes, M., Soberón, J., 2008. Rethinking receiver operating characteristic analysis: applications in ecological niche modeling. *Ecological Modelling*, 213: 63-72.
- Phillips, S.J., Anderson, R.P., Schapire, R.E., 2006. Maximum entropy modelling of species geographic distributions. *Ecological Modelling*, 190: 231–259.
- Phillips, S.J., Dudík, M., 2008. Modelling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography*, 31 (2): 161–175.
- Rockström, J., Steffen, W., Noone, K., Persson, Å., Chapin III, F.S., Lambin, E.F., Lenton, T.M., Scheffer, M., Folke, C., Schellnhuber, H.J., Nykvist, B., Wit, C.A., Hughes, T., van der Leeuw, S., Rodhe, H., Sörlin, S., Snyder, P.K., Costanza, R., Svedin, U., Falkenmark, M., Karlberg, L., Corell, R.W., Fabry, V.J., Hansen, J., Walker, B., Liverman, D., Richardson, K., Crutzen, P., Foley, J.A., 2009. A safe operating space for humanity. *Nature* 461 (24).
- Schölkopf, B., Platt, J., Shawe-Taylor, J., Smola, A.J., Williamson, R.C., 2001. Estimating the support of a high-dimensional distribution. *Neural Computation*, 13: 1443-1471.
- Segurado, P., Araújo, M.B., 2004. An evaluation of methods for modelling species distributions. *Journal of Biogeography*, 31: 1555–1568.
- Siqueira, M.F., Durigan, G., De Marco Júnior, P., Peterson, A.T., 2009. Something from nothing: Using landscape similarity and ecological niche modelling to find rare or endangered species. *Journal for Nature Conservation*, 17 (1): 25-32.
- Soberón, J., Peterson, A.T., 2004. Biodiversity informatics: managing and applying primary biodiversity data. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 359: 689-98.
- Soberón, J., Peterson, A.T., 2005. Interpretation of models of fundamental ecological niches and species' distributional areas. *Biodiversity Informatics*, 2: 1-10.
- Stockwell, D.R.B., Beach, J.H., Stewart, A., Vorontsov, G., Vieglais, D., Scachetti-Pereira, R., 2008. The use of the GARP genetic algorithm and internet grid computing in the Lifemapper world atlas of species biodiversity. *Ecological Modelling*, 195 (1-2): 139-145.
- Stockwell, D.R.B., Peters, D., 1999. The GARP modelling system: problems and solutions to automated spatial prediction. *International Journal of Geographical Information Science*, 13: 143-158.
- Stockwell, D.R.B., Peterson A.T., 2002. Effects of sample size on accuracy of species distribution models. *Ecological Modelling*, 148: 1-13.
- Stork, N.E., Coddington, J.A., Colwell, R.K., Chazdon, R.L., Dick, C.W., Peres, C.A., Sloan, S., and Willis, K., 2009. Vulnerability and resilience of tropical forest species to land use change. *Conservation Biology*, 23 (6): 1438–1447. doi:10.1111/j.1523-1739.2009.01335.x.
- Sullivan, W.T., Werthimer, D., Bowyer, S., Cobb, J., Gedye, D., Anderson, D., 1997. A new major SETI project based on Project Serendip data and 100000 personal

computers, in: C.B. Cosmovici, S. Bowyer, and D. Werthimer (Eds.), IAU Colloq. 161: Astronomical and Biochemical Origins and the Search for Life in the Universe, Proc. of the Fifth Intl. Conf. on Bioastronomy, Editrice Compositori, Bologna, Italy.

Thuiller, W., 2004. Patterns and uncertainties of species' range shifts under climate change. *Global Change Biology*, 10: 2020–2027.

Thuiller, W., Lafourcade, B., Engler, R., Araújo, M.B., 2009. BIOMOD – a platform for ensemble forecasting of species distributions. *Ecography*, 32: 369–373. doi:10.1111/j.1600-0587.2008.05742.x.

Tobler, M., Honorio, E., Janovec, J., Reynel, C., 2007. Implications of collection patterns of botanical specimens on their usefulness for conservation planning: an example of two neotropical plant families (Moraceae and Myristicaceae) in Peru. *Biodiversity and Conservation*, 16 (3): 659–677. doi:10.1007/s10531-005-3373-9.

Tropicos, 2011. Botanical information system at the Missouri Botanical Garden: www.tropicos.org

Vitta, F.A., Bernacci, L.C., 2004. A new species of *Passiflora* in section *Tetrastylis* (Passifloraceae) and two overlooked species of *Passiflora* from Brazil. *Brittonia*, 56 (1): 89–95.

Wieczorek, J., Döring, M., Giovanni, R., Robertson, T., Vieglais, D., 2009. Darwin Core standard. <http://rs.tdwg.org/dwc/> (accessed 2011-01-26).

Wieczorek, J., Guo, Q., Hijmans, R.J., 2004. The point-radius method for georeferencing locality descriptions and calculating associated uncertainty. *International Journal of Geographical Information Science*, 18 (8): 745–767.

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. *Diversity and Distributions*, 14: 763–773. doi:10.1111/j.1472-4642.2008.00482.x.

Woodward, F.I., 1987. *Climate & Plant distribution*. Cambridge University Press, Cambridge.

Yesson, C., Brewer, P.W., Sutton, T., Caithness, N., Pahwa, J.S., Burgess, M., Gray, W.A., White, R.J., Jones, A.C., Bisby, F.A., Culham, A., 2007. How Global Is the Global Biodiversity Information Facility? *PLoS ONE*, 2 (11): e1124. doi:10.1371/journal.pone.0001124.

Zimmermann, N.E., Edwards Jr, T.C., Graham, C.H., Pearman, P.B., Svenning, J.C., 2010. New trends in species distribution modelling. *Ecography*, 33 (6): 985–989. doi:10.1111/j.1600-0587.2010.06953.x.

Zomer, R.J., Bossio, D.A., Trabucco, A., Yuanjie, L., Gupta, D.C., Singh, V.P., 2007. *Trees and Water: Smallholder Agroforestry on Irrigated Lands in Northern India*. Colombo, Sri Lanka: International Water Management Institute. IWMI Research Report, 122, pp 45.

Annex I

Incompatible homonyms (*):

Homonym	Currently accepted name
<i>Passiflora elegans</i> Triana & Planch.	<i>Passiflora quindensis</i> Killip.
<i>Passiflora rubra</i> Vell.	<i>Passiflora organensis</i> Gardner
<i>Passiflora vespertilio</i> Ker Gawl.	<i>Passiflora misera</i> Kunth
<i>Passiflora villosa</i> Dombey ex Triana & Planch.	<i>Passiflora menispermifolia</i> Kunth
<i>Passiflora villosa</i> Macfad.	<i>Passiflora suberosa</i> L.

(*) Names sharing the same binomial with an accepted name but referring to a different species.

Incompatible varieties (**):

Variety	Currently accepted name
<i>Dilkea johannesii</i> var. <i>parvifolia</i> Hoehne	<i>Dilkea acuminata</i> Mast.
<i>Passiflora amethystina</i> var. <i>bolosii</i> Cervi	<i>Passiflora loefgrenii</i> Vitta
<i>Passiflora capsularis</i> var. <i>geminiflora</i> Mast.	<i>Passiflora sexflora</i> Juss.
<i>Passiflora capsularis</i> var. <i>geminifolia</i> DC.	<i>Passiflora sexflora</i> Juss.
<i>Passiflora edulis</i> var. <i>kerii</i> (Spreng.) Mast.	<i>Passiflora incarnata</i> L.
<i>Passiflora foetida</i> var. <i>ciliata</i> (Aiton) Mast.	<i>Passiflora ciliata</i> Aiton
<i>Passiflora foetida</i> var. <i>hibiscifolia</i> (Lam.) Killip	<i>Passiflora ciliata</i> Aiton
<i>Passiflora foetida</i> var. <i>moritziana</i> (Planch.) Killip ex Pulle	<i>Passiflora moritziana</i> Planch.
<i>Passiflora foetida</i> var. <i>nicaraguensis</i> (Killip ex Standl.) Killip	<i>Passiflora ciliata</i> Aiton
<i>Passiflora foetida</i> var. <i>subintegra</i> Killip	<i>Passiflora ciliata</i> Aiton
<i>Passiflora foetida</i> var. <i>vitacea</i> Mast.	<i>Passiflora chrysophylla</i> Chodat
<i>Passiflora quadriglandulosa</i> var. <i>involucrata</i> (Mast.) Killip	<i>Passiflora involucrata</i> (Mast.) A.H.Gentry
<i>Passiflora vitifolia</i> var. <i>involucrata</i> Mast.	<i>Passiflora involucrata</i> (Mast.) A.H.Gentry
<i>Passiflora vitifolia</i> var. <i>minor</i> Mast.	<i>Passiflora quadriglandulosa</i> Rodschied

(**) Varieties sharing the same binomial with accepted names but referring to a different species.

Synonyms (***):

Accepted name	Synonyms
<i>Dilkea acuminata</i> Mast.	<i>Dilkea parviflora</i> Killip
<i>Dilkea johannesii</i> Barb.Rodr.	<i>Dilkea ulei</i> Harms
<i>Mitostemma brevifilis</i> Gontsch.	<i>Dilkea margaritae</i> Cervi
<i>Passiflora actinia</i> Hook.	<i>Passiflora paulensis</i> Killip
<i>Passiflora ambigua</i> Hemsl.	<i>Passiflora emiliae</i> Sacco
<i>Passiflora amethystina</i> J.C.Mikan	<i>Passiflora bangii</i> Mast.
	<i>Passiflora cornuta</i> Mast.
<i>Passiflora auriculata</i> Kunth	<i>Passiflora cayaponioides</i> Rusby
<i>Passiflora candida</i> (Poepp. & Endl.) Mast.	<i>Passiflora guedesii</i> Huber
<i>Passiflora costata</i> Mast.	<i>Passiflora inundata</i> Ducke
<i>Passiflora haematostigma</i> Mart. ex Mast.	<i>Passiflora platystyla</i> Mast.
<i>Passiflora jilekii</i> Wawra	<i>Passiflora canescens</i> Killip
	<i>Passiflora colorata</i> Mast.
	<i>Passiflora mediterranea</i> Vell.
<i>Passiflora morifolia</i> Mast.	<i>Passiflora erosa</i> Rusby
	<i>Passiflora warmingii</i> Mast.
<i>Passiflora oerstedii</i> Mast.	<i>Passiflora rojasii</i> Hassl. ex Harms
<i>Passiflora ovalis</i> Vell. ex M.Roem.	<i>Tetrastylis ovalis</i> (Vell. ex M. Roem.) Killip
<i>Passiflora pentagona</i> Mast.	<i>Passiflora alliacea</i> Barb. Rodr.
<i>Passiflora quadriglandulosa</i> Rodschied	<i>Passiflora vitifolia</i> var. <i>minor</i> Mast.
<i>Passiflora retipetala</i> Mast.	<i>Passiflora lonchophora</i> Harms
<i>Passiflora rubra</i> L.	<i>Passiflora cisnana</i> Harms
<i>Passiflora truncata</i> Regel	<i>Passiflora lawsoniana</i> Mast.

(***) This list does not include orthographical variants and misspellings.

Annex II

Priority list for digitization and field surveys (*)

Scientific name	Total number of records	Number of counties where the species was collected	Last collected
<i>Passiflora balbis</i> Feuillet	0	0	?
<i>Passiflora cryptopetala</i> Hoehne	0	0	?
<i>Passiflora margaritae</i> Sacco	0	0	?
<i>Passiflora reitzii</i> Sacco	0	0	?
<i>Passiflora rufa</i> Feuillet & J.M.MacDougal	0	0	?
<i>Passiflora chlorina</i> L.K. Escobar	1	1	?
<i>Passiflora spicata</i> Mast.	1	1	?
<i>Passiflora quadrifaria</i> Vanderpl.	1	1	1995
<i>Passiflora securiclata</i> Mast.	2	1	1889
<i>Passiflora castellanosi</i> Sacco	2	1	1967
<i>Passiflora nephrodes</i> Mast.	2	2	1962
<i>Passiflora cordistipula</i> Cervi	2	2	1977
<i>Passiflora longilobis</i> Hoehne	2	2	2003
<i>Passiflora hypoglauca</i> Harms	2	2	2004
<i>Passiflora leptopoda</i> Harms	3	1	1908
<i>Passiflora pedata</i> L.	3	1	1913
<i>Passiflora phaeocaula</i> Killip	3	2	2004
<i>Passiflora jiboiensis</i> M.A.M.Azevedo	4	1	?
<i>Passiflora ichthyura</i> Mast.	4	1	1839
<i>Passiflora pardifolia</i> Vanderpl.	4	2	2007
<i>Passiflora glaucescens</i> Killip	4	4	2006
<i>Passiflora sclerophylla</i> Harms	6	1	1994
<i>Passiflora involucrata</i> (Mast.) A.H.Gentry	6	3	1976
<i>Ancistrothyrsus tessmannii</i> Harms	6	3	2007
<i>Passiflora trifasciata</i> Lem.	7	1	2006
<i>Passiflora retipetala</i> Mast.	7	2	1932
<i>Passiflora ernesti</i> Harms	8	1	1929
<i>Passiflora menispermifolia</i> Kunth	8	2	2007
<i>Passiflora saxicola</i> Gontsch.	8	7	2008
<i>Passiflora filamentosa</i> Cav.	11	3	2008
<i>Dilkea acuminata</i> Mast.	11	5	1997
<i>Passiflora candollei</i> Triana & Planch.	11	7	2007
<i>Passiflora amalocarpa</i> Barb.Rodr.	12	1	1990
<i>Passiflora variolata</i> Poepp. & Endl.	13	4	1988
<i>Passiflora elegans</i> Mast.	14	4	2009
<i>Passiflora micropetala</i> Mart. ex Mast.	15	4	2008
<i>Passiflora farneyi</i> Pessoa & Cervi	15	5	2001
<i>Passiflora odontophylla</i> Harms ex Glaz.	15	5	2005
<i>Passiflora hatschbachii</i> Cervi	15	5	2007
<i>Passiflora gardneri</i> Mast.	17	6	2007
<i>Passiflora serrato-digitata</i> L.	18	5	2006
<i>Passiflora vitifolia</i> Kunth	18	6	2008
<i>Passiflora ceratocarpa</i> F.Silveira	21	5	2003
<i>Passiflora elliptica</i> Gardner	24	4	2004
<i>Passiflora malacophylla</i> Mast.	27	11	2009
<i>Passiflora cervii</i> M.A.M.Azevedo	35	4	2009
<i>Passiflora transversalis</i> M.A.M.Azevedo	66	10	2005

(*) All records from all species in this list did not pass the data quality tests for the modelling process.

PARTE II – Mapas de distribuição potencial e indicadores de conservação para espécies de Passifloraceae que ocorrem no Estado de São Paulo

Introdução

Neste capítulo os procedimentos de modelagem descritos no capítulo anterior foram utilizados com o objetivo de gerar mapas de distribuição potencial e indicadores de conservação para espécies de Passifloraceae que ocorrem no Estado de São Paulo (SP).

São Paulo é atualmente o estado mais urbanizado, industrializado e populoso do Brasil, sendo responsável por mais de 30% do PIB do país (IBGE 2007) e com uma população de quase 40 milhões de habitantes (IBGE 2010) distribuída num território de 248.209,426 km² (IBGE 2002). O desenvolvimento econômico foi acompanhado de grande impacto ambiental nos dois biomas que abrangem o estado: a Mata Atlântica e o Cerrado. Estima-se que a cobertura florestal que originalmente correspondia a mais de 80% do estado tenha sido reduzida a 3% até ano 2000 (Victor *et al.* 2005). Apesar da brutal redução e da existência de legislação específica para proteger os dois biomas, até hoje o desmatamento persiste a uma taxa de 1,1% no Cerrado e de 0,05% na Mata Atlântica, calculadas entre os anos de 2002 e 2008 e relativos à área legal do bioma (CSR/Ibama 2009; CSR/Ibama 2010). Em compensação, São Paulo possui um dos maiores percentuais do país de áreas protegidas por unidades de conservação (UCs) de proteção integral, cobrindo 3,5% de seu território (SMA 2009).

Dentre as 135 espécies de Passifloraceae que ocorrem no país (Cervi *et al.* 2010) 38 constam na Flora Fanerogâmica do Estado de São Paulo (Bernacci *et al.* 2003) sendo que outras quatro têm sua distribuição relacionada ao estado pela Lista de Espécies da Flora do Brasil (Cervi *et al.* 2010) e três, *P. catharinensis*, *P. pardifolia* e *P. rhamnifolia*, foram recentemente encontradas no estado, totalizando até o momento 45 espécies da família. Destas, cinco estão no Livro Vermelho das Espécies Vegetais Ameaçadas do Estado de São Paulo (Mamede *et al.* 2007) e uma está na Lista Vermelha de Espécies da Flora do Brasil (MMA 2008).

O objetivo deste trabalho foi o de obter indicadores de conservação a partir de registros de ocorrência e mapas de distribuição potencial para identificar espécies potencialmente raras ou com distribuição restrita que possam estar ameaçadas. Além de produzir subsídios para direcionar ações de conservação, trata-se de uma primeira tentativa de compreender melhor os diferentes padrões de distribuição e requisitos ambientais de cada espécie desta família.

Métodos

Todos os modelos foram gerados seguindo os mesmos procedimentos descritos no capítulo anterior, portanto utilizando registros de ocorrência provenientes de herbários (através da rede *speciesLink*), filtrando os resultados da seleção para evitar possíveis erros e imprecisões, e utilizando diferentes estratégias de modelagem conforme a quantidade final de pontos disponíveis para cada espécie. Os modelos foram treinados

com pontos de ocorrência coletados no Brasil, portanto amostrando de maneira mais abrangente as preferências ambientais de cada espécie. Por outro lado, os mapas de distribuição potencial foram posteriormente recortados com os limites do estado. Todos os modelos foram gerados em resolução de 30' utilizando as mesmas camadas ambientais selecionadas no capítulo anterior (tabela 1).

Tabela 1: Camadas ambientais utilizadas na modelagem.

Abrev.	Descrição	Unidade
bio02	Variação média de temperatura ao longo do dia	°C * 10
bio05	Temperatura máxima no mês mais quente do ano	°C * 10
bio06	Temperatura mínima no mês mais frio do ano	°C * 10
bio16	Precipitação no trimestre mais úmido do ano	mm
bio17	Precipitação no trimestre mais seco do ano	mm
bio18	Precipitação no trimestre mais quente do ano	mm
bio19	Precipitação no trimestre mais frio do ano	mm
alt	Altitude	m

Os resultados são apresentados por espécie, incluindo mapa de distribuição potencial (quando houve registros de ocorrência suficientes para a modelagem) e um conjunto de dados variável em função da estratégia de modelagem utilizada. Em todos os casos é apresentado o número total de pontos disponíveis através da rede *speciesLink*, servindo como um indicativo do quanto se conhece sobre a espécie. Em seguida é apresentado o número de pontos que foram utilizados na modelagem após filtragem e seleção e, destes, quantos foram coletados em SP.

Para espécies com mais de cinco pontos são exibidas as estatísticas do modelo, incluindo omissão interna tanto relativa a todos os pontos usados na construção do modelo quanto aos pontos cuja coleta foi realizada em SP. A omissão externa também foi calculada em relação a todos os polígonos de teste disponíveis e em relação àqueles referentes a coletas em SP. A significância foi calculada levando-se em consideração todos os polígonos de teste disponíveis de materiais coletados no Brasil. Não foi calculada significância para SP devido ao baixo número de registros por estado para a maioria das espécies.

Ainda para espécies com pelo menos cinco pontos foram calculadas em relação ao estado a área de distribuição potencial, a área de distribuição potencial dentro de remanescentes de vegetação e a área de distribuição potencial dentro de remanescentes de vegetação que se encontram dentro de unidades de conservação de proteção integral. O limite de corte aplicado ao modelo para separar áreas favoráveis à ocorrência da espécie foi o menor valor de predição do modelo entre os pontos de treino. Os mapas de remanescentes de vegetação utilizados referem-se ao ano de 2008 e foram provenientes do Projeto de Monitoramento do Desmatamento dos Biomas Brasileiros por Satélite (PMDBBS 2011) para a Mata Atlântica e para o Cerrado. Os mapas de unidades de conservação foram provenientes do “World

Database on Protected Areas” (IUCN & UNEP 2010). Foram excluídas as áreas com categorias IUCN IV, V e VI (áreas de manejo de espécie ou habitat, paisagens protegidas e áreas de manejo de recursos, respectivamente), além de áreas indígenas e áreas delimitadas por organizações internacionais, tal como reservas da biosfera da UNESCO. Assim restaram apenas unidades de conservação que se enquadram na categoria de proteção integral.

Todos os valores de áreas foram calculados após conversão dos mapas vetoriais tanto dos remanescentes de vegetação quanto das unidades de conservação para um formato matricial compatível com o mapa de distribuição da espécie (mesma escala). Os cálculos foram feitos de forma aproximada assumindo o valor de 0,8 km² por pixel em toda a área de estudo, resultando em erros menores que 1% para a área de distribuição potencial da espécie comparando-se com a forma mais precisa de cálculo de área através de mapas vetoriais. Considerando-se que a área de ocorrência real (no sentido de “área de ocupação”, e não “extensão de ocorrência”) tende a ser sempre menor que a área de ocorrência potencial, e quando a primeira é menor que 2.000 km² isso pode levar uma espécie a ser considerada vulnerável de acordo com os critérios da IUCN (IUCN 2001), áreas de ocorrência potencial abaixo de 10.000 km² foram utilizadas como critério para indicar espécies com distribuição potencialmente restrita. A área de ocorrência potencial foi considerada como a intersecção da área do modelo com as áreas de remanescentes de vegetação.

Foi também incluído para todas as espécies, quando disponível, o ano da última coleta e a respectiva referência ao material de herbário – neste caso considerando-se apenas coletas realizadas em SP em ambiente nativo, porém sujeitas a erros e imprecisões na localização e na identificação. Como em alguns casos a ausência de coleta por mais de 50 anos é tida como indicador de que a espécie possa estar extinta (Mamede *et al.* 2007) utilizou-se como critério a ausência de coleta há mais de 10 anos para listar espécies potencialmente ameaçadas. Além disso, são mencionados os municípios onde houve coleta no estado, sejam eles associados a pontos de treino precisos utilizados na modelagem (indicados em materiais sublinhados) ou a polígonos de teste com precisão insuficiente para a modelagem. Nos casos em que o respectivo ponto de treino ou polígono de teste foi omitido no resultado final do modelo, o material de herbário aparece em negrito, indicando locais onde poderiam ser realizadas novas coletas com maior precisão ou mesmo sugerindo a revisão do material no herbário (identificação e localização geográfica). Tais indicações não foram feitas para espécies com menos de cinco pontos disponíveis para modelagem, pois neste caso os modelos são mapas de dissimilaridade ambiental com o objetivo de orientar novas coletas, não tendo os mesmos sido testados como foram os demais modelos. Cabe observar que o material indicado no ano da última coleta pode algumas vezes não aparecer associado a nenhum município, seja porque o material apenas indica que a coleta foi realizada no estado, ou porque algum conflito nos dados impediu que o mesmo fosse utilizado na modelagem e nos testes.

Os pontos de treino e polígonos de teste também foram utilizados para indicar unidades de conservação onde a espécie foi ou pode ter sido coletada. Quando o ponto ou polígono ficou inteiramente contido dentro da área da UC, esta aparece

seguida do símbolo ●. Quando houve intersecção parcial de um polígono com a UC, foi utilizado o símbolo ◐. Em alguns casos é possível que haja indicação de UC sem que haja indicação de município pelo fato da espécie possuir registro de ocorrência num município pertencente a outro estado, porém adjacente à UC. Por outro lado, também é possível que não haja indicação de UC mesmo quando os registros de ocorrência estejam relacionados a municípios onde existem UCs. Neste caso os registros de ocorrência necessariamente possuem precisão melhor que a do município, não se sobrepondo a nenhuma UC.

Por fim, para espécies modeladas apenas com dissimilaridade ambiental são apresentados os dados ambientais (valores mínimo e máximo) referentes aos pontos utilizados no treino do modelo. Em espécies com pelo menos cinco pontos, foram calculados perfis ambientais mais completos. Estes foram feitos por variável ambiental, contendo um histograma da distribuição potencial da espécie no Brasil para a respectiva variável (cinza escuro) sobreposto ao histograma da variável ambiental em todo o Brasil (cinza claro) indicando também valores mínimo, médio e máximo relativos aos pontos de treino do modelo.

Resultados

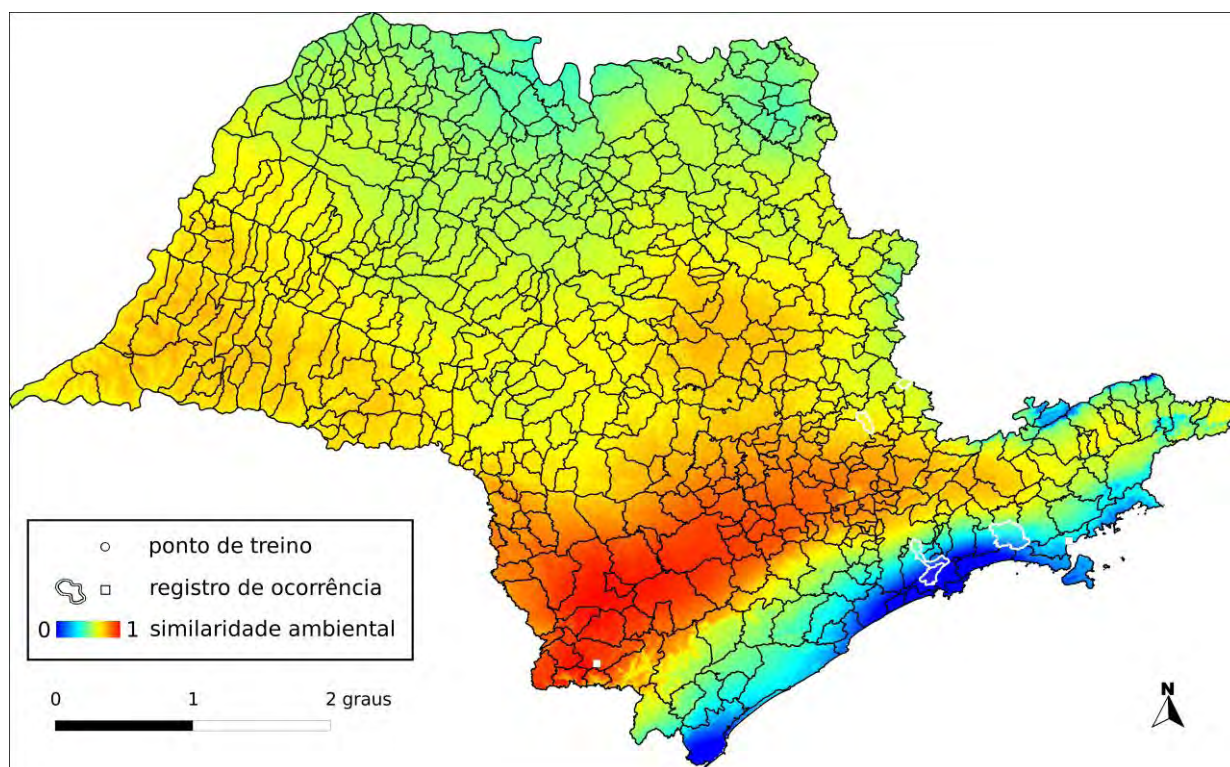
Dentre as 45 espécies incluídas no estudo, seis (*P. cervii*, *P. elegans*, *P. glaucescens*, *P. malacophylla*, *P. pardifolia* e *P. transversalis*) não tiveram mapa de distribuição potencial gerado por não possuírem nenhum ponto de ocorrência que tenha passado por todos os critérios de qualidade. Destas, três (*P. cervii*, *P. elegans* e *P. malacophylla*) além de *P. campanulata* e *P. racemosa* não possuem nenhum registro na rede *speciesLink* indicando que a coleta tenha sido realizada em SP, o que as coloca no topo da lista de espécies cuja distribuição no Estado é menos conhecida (anexo I).

Entre as espécies com pelo menos cinco pontos, duas (*P. mucronata* e *P. pentagona*) apresentaram área de ocorrência potencial menor que 10.000 km² em SP (anexo II), sugerindo atenção especial do ponto de vista de conservação devido a uma possível área de ocorrência real bastante restrita. Dezesseis espécies aparentemente não são coletadas há mais de 10 anos (anexo III), indicando também maior atenção quanto ao seu *status* de conservação.

Foram encontrados registros de ocorrência de cinco espécies cuja distribuição até o momento não inclui SP: *P. catharinensis* (IAC49355), *P. pardifolia* (UPCB60135), *P. quadrangularis* (MO282855), *P. rhamnifolia* (IAC44029) e *P. watsoniana* (MO100237700). Destas, apenas três foram incluídas neste trabalho, pois *P. quadrangularis* é tida como endêmica da região amazônica, onde seus frutos são inclusive apreciados para consumo, portanto o registro de ocorrência em SP deve ser investigado com maior detalhe. Quanto ao registro de *P. watsoniana*, este já era conhecido e também é bastante questionável segundo Bernacci *et al.* (2003).

Os resultados encontram-se organizados a seguir para cada espécie.

Passiflora actinia Hook.



Algoritmo: Distância Euclideana

Total de registros disponíveis na rede *speciesLink*: 167

Nº de pontos usados na modelagem: 2 (0 em SP)

Ano da última coleta em SP: 2004 (IAC44169)

Municípios com registros de ocorrência em SP: Cubatão (IAC33737), Pedreira (IAC36158), Salesópolis (SPSF11281), Santo André (IAC44169), Águas de Lindóia (HTSA1922)

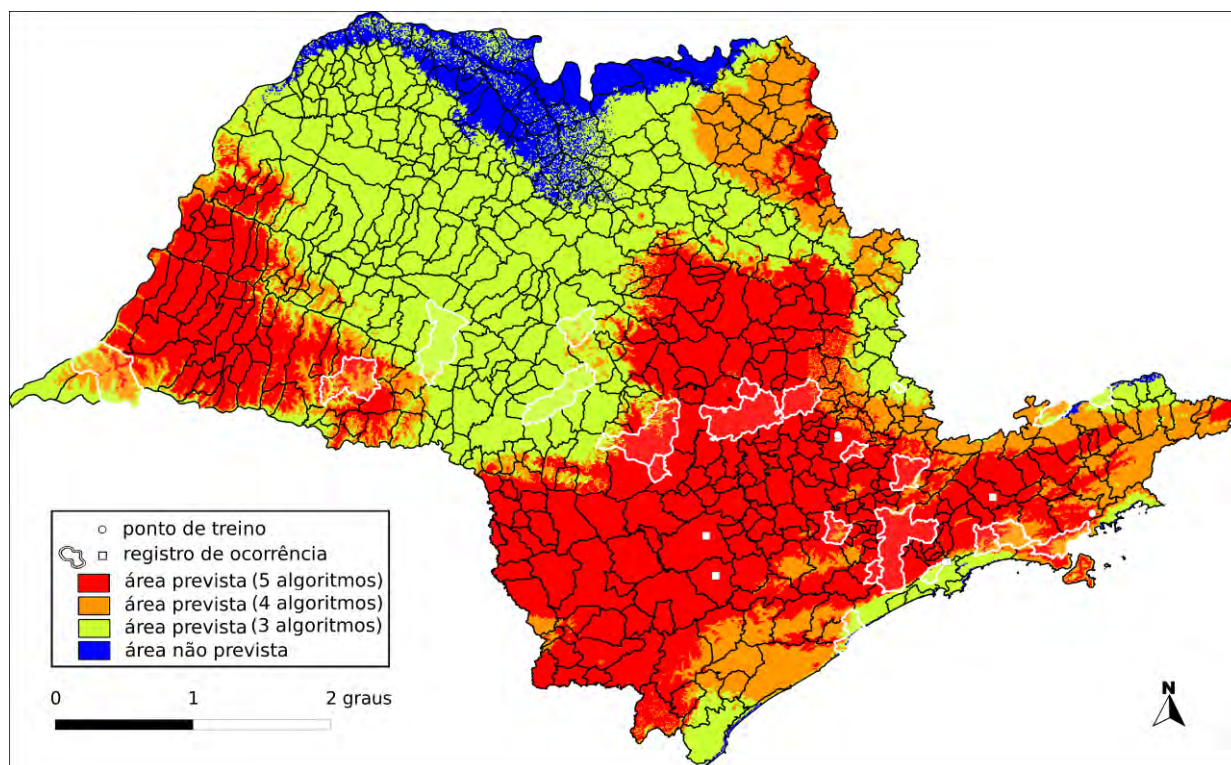
UCs¹: P.E. da Serra do Mar ●

Dados ambientais:

bio02 var. temp. dia (°C*10)	104 – 116	bio17 prec. tri. mais seco (mm)	103 – 268
bio05 temp. max. (°C*10)	272 – 283	bio18 prec. tri. mais quente (mm)	386 – 474
bio06 temp min. (°C*10)	79 – 124	bio19 prec. tri. mais frio (mm)	103 – 287
bio16 prec. tri. mais úmido (mm)	496 – 577	alt altitude (m)	849 – 889

¹ Registro de ocorrência inteiramente (●) ou parcialmente (◐) contido em UC.

Passiflora alata Curtis



Algoritmo: Consenso (ENFA, GARP BS, Maxent, MD e SVM)

Total de registros disponíveis na rede *speciesLink*: 425

Omissão interna (BR): 0,03 (n=39) **Omissão interna (SP):** 0,00 (n=3)

Omissão externa (BR): 0,02 (n=130) **Omissão externa (SP):** 0,00 (n=23)

Significância (BR): $p < 0,01$

Área do modelo (SP): 234.944 km² (94,7% da área do Estado)

Área em remanescentes (2008): 42.440 km² (18,1% da área do modelo)

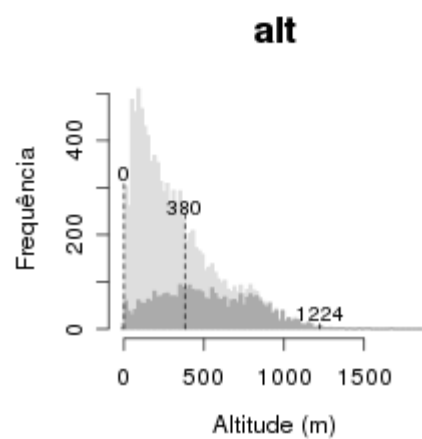
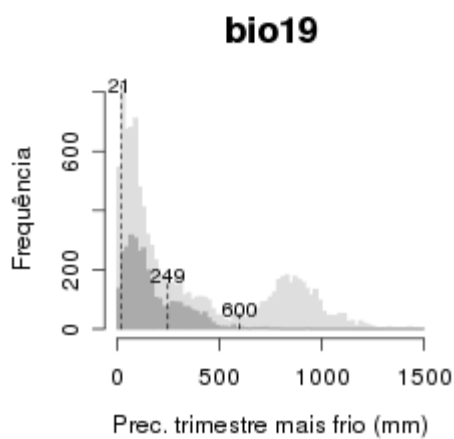
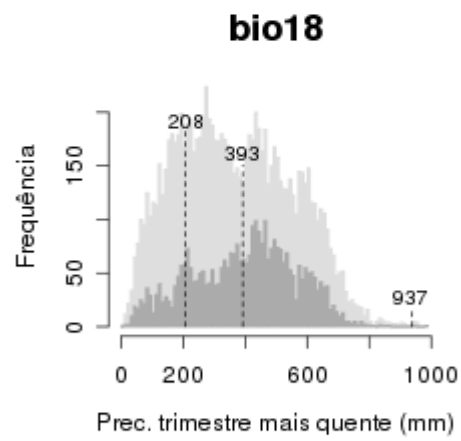
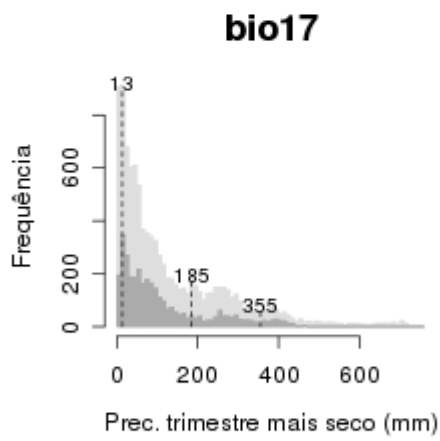
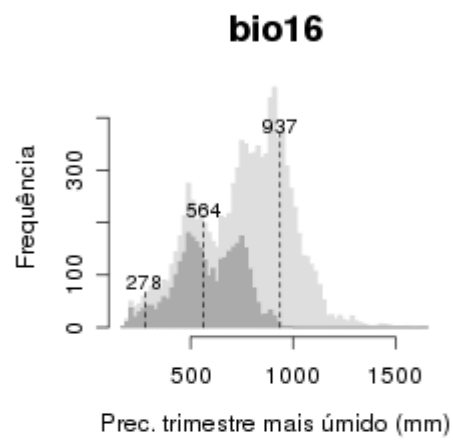
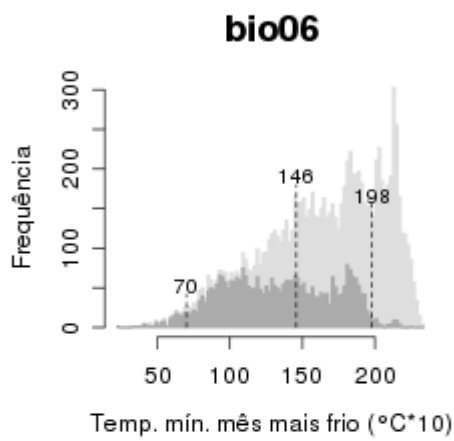
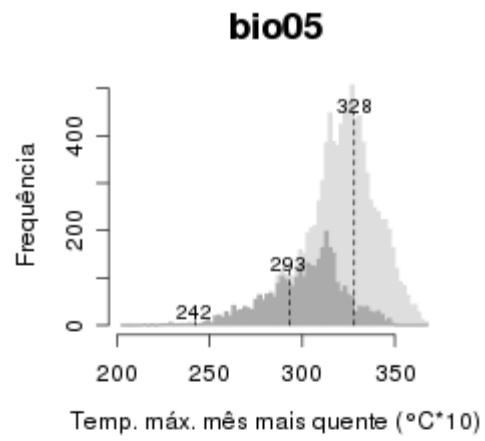
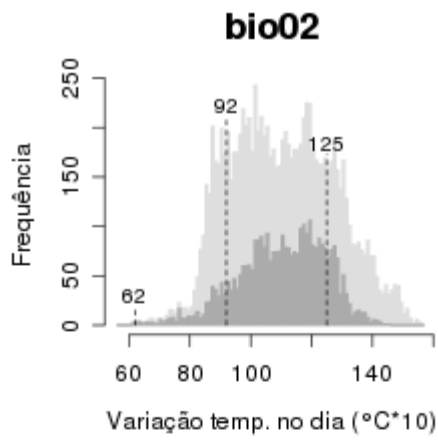
Área em remanescentes em UCs (2008): 7.901 km² (3,4% da área do modelo)

Ano da última coleta em SP: 2008 (IAC50283)

Municípios com registros de ocorrência em SP: Agudos (IAC35877), Arealva (IAC50283), Atibaia (UEC58769), Biritiba-Mirim (IAC42139), Botucatu (UPCB28145), Campinas (IAC37210), Campos do Jordão (UPCB19912), Caraguatatuba (SPSF28930), Charqueada (IAC40221), Cubatão (IAC32580), Embu-Guaçu (RB475374), Limeira (IAC32579), Lindóia (IAC24889), Marília (IAC7955), Natividade da Serra (IAC46146), Paraguaçu Paulista (SPSF17658), Pardinho (RB253078), Peruíbe (ESA11406), Piquete (IAC48704), Piracicaba (IAC32585), Salesópolis (IAC43034), Santo André (SP212619), São Paulo (IAC32582), São Roque (SPSF35474), Teodoro Sampaio (IAC32586), Valinhos (UEC12587)

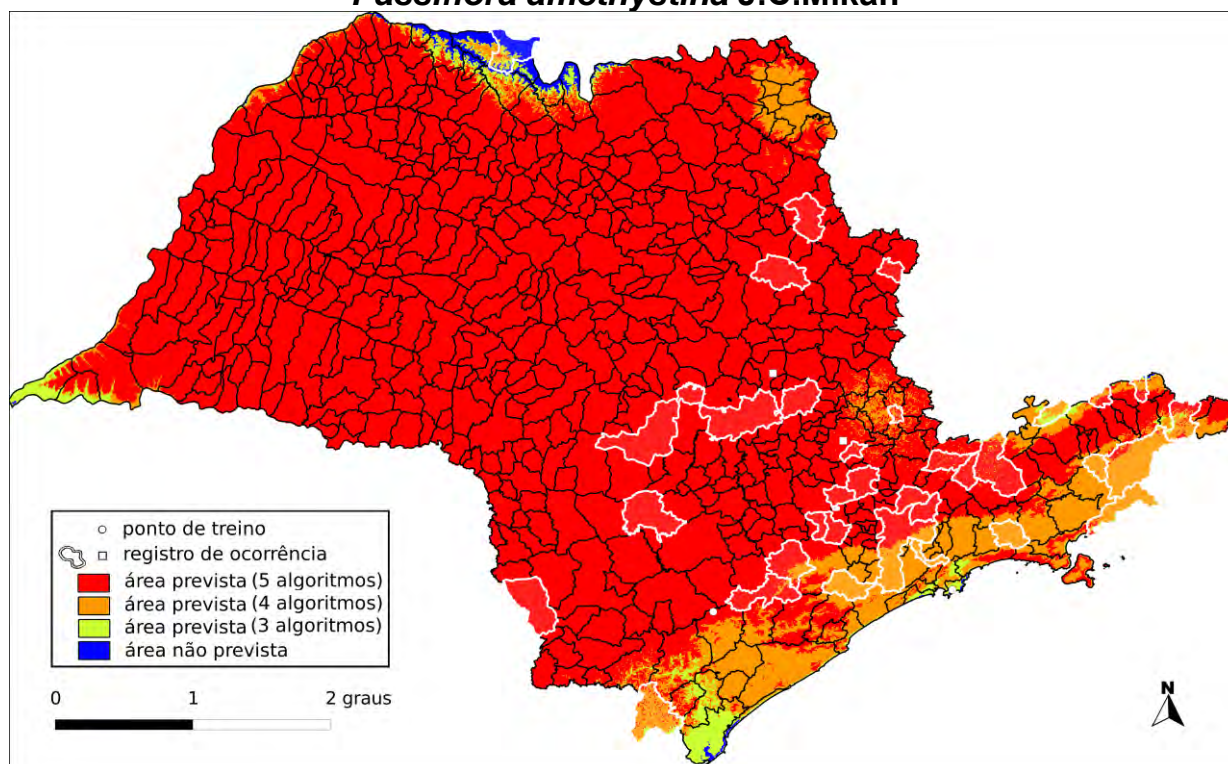
UCs²: P.E. de Campos do Jordão (●), P.E. do Jaraguá (●), P.E. Guarapiranga (●), E.E. Ibicatu (●), P.E. Alberto Löfgren (●), E.E. Mico Leão Preto (●), P.E. da Cantareira (●), Embu-Guaçu (●), P.N. da Serra da Bocaina (●), P.E. da Ara (●), E.E. Juréia-Itatins (●), E.E. de Valinhos (●), P.E. dos Mananciais de Campos do Jordão (●), P.E. do Morro do Diabo (●), P.E. da Serra do Mar (●), P.E. do Itinguçu (●)

² Registro de ocorrência inteiramente (●) ou parcialmente (◐) contido em UC.



Histogramas: Brasil (claro) e projeção do modelo no Brasil (escuro). Valores: mínimo, médio e máximo dos pontos de ocorrência.

Passiflora amethystina J.C.Mikan



Algoritmo: Consenso (ENFA, GARP BS, Maxent, MD e SVM)

Total de registros disponíveis na rede *speciesLink*: 583

Omissão interna (BR): 0,00 (n=33)

Omissão interna (SP): 0,00 (n=2)

Omissão externa (BR): 0,03 (n=159)

Omissão externa (SP): 0,00 (n=35)

Significância (BR): $p < 0,01$

Área do modelo (SP): 246.212 km² (99,2% da área do Estado)

Área em remanescentes (2008): 43.327 km² (17,6% da área do modelo)

Área em remanescentes em UCs (2008): 7.940 km² (3,2% da área do modelo)

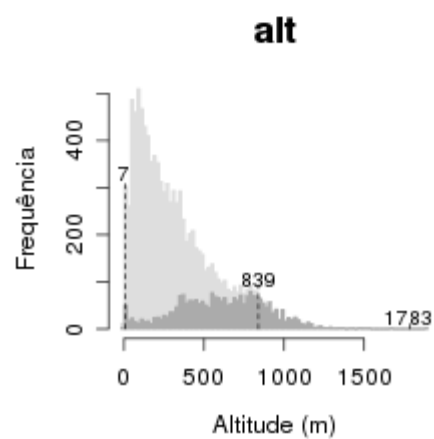
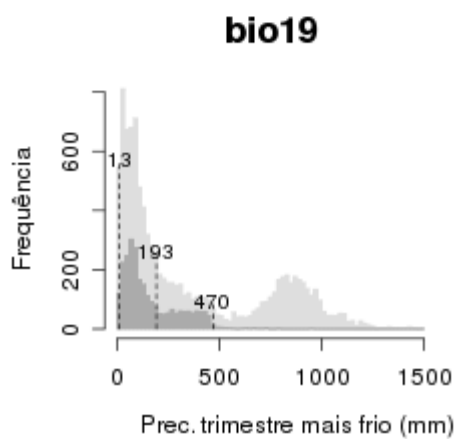
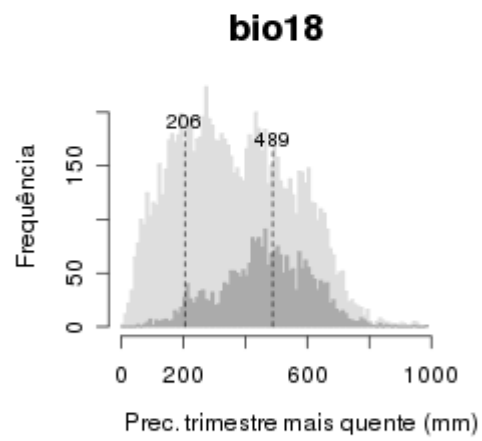
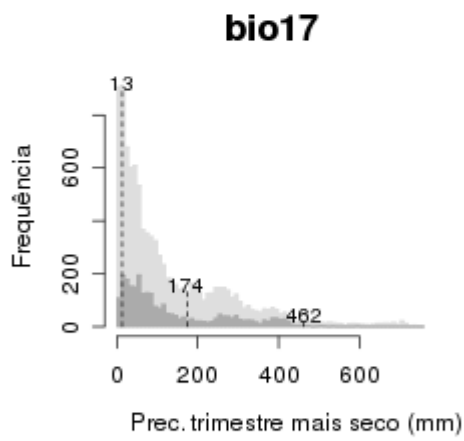
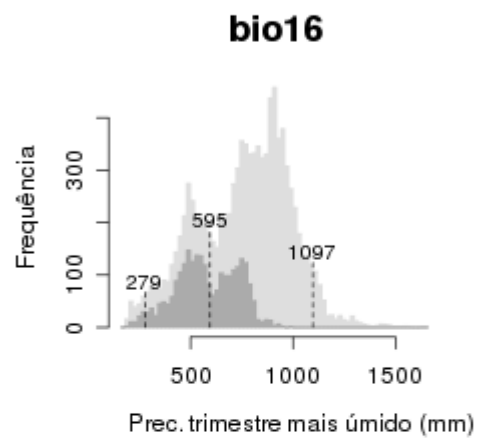
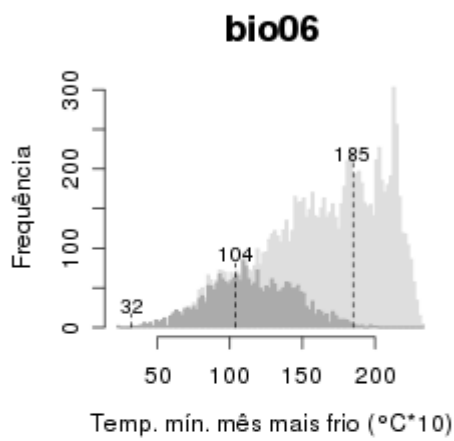
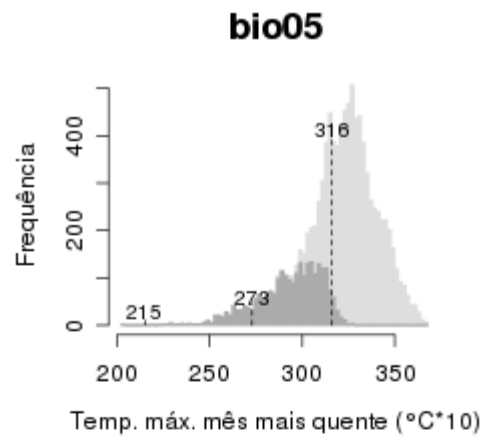
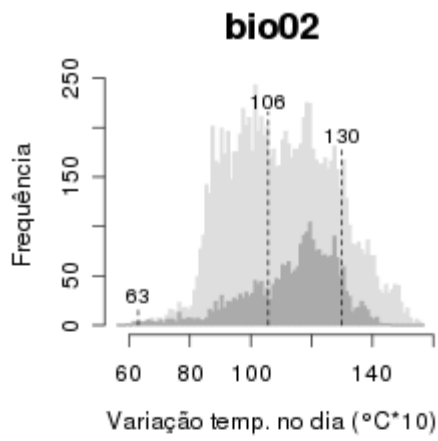
Ano da última coleta em SP: 2008 (IAC49357)

Municípios com registros de ocorrência em SP: Angatuba (IAC46225), Barra do Turvo (UPCB32754), Botucatu (IAC24590), Cabreúva (MBM151465), Cajuru (UEC60178), Campos do Jordão (RB215080), Cunha (IAC3621), Divinolândia (IAC21484), Guarujá (IAC49357), Guarulhos (SPSF44397), Igaratá (IAC33751), Itararé (UEC76145), Jundiá (UEC41158), Juquitiba (UEC75992), Lagoinha (RB23561), Lavrinhas (UEC88817), Limeira (IAC33749), Mairinque (IAC33755), Mairiporã (SPSF21263), Monte Alegre do Sul (IAC19323), Paulo de Faria (HSJRP10246), Piedade (IAC21929), Piquete (UEC83820), Piracaia (IAC24945), Piracicaba (ESA19919), Salesópolis (RB113962), Santa Maria da Serra (IAC21112), Santa Rita Do Passa Quatro (UEC12624), Santo André (IAC44168), Sete Barras (ESA79745), São José do Barreiro (RB214619), São José dos Campos (IAC45175), São Paulo (IAC4494), São Roque (IAC32345), Tapiraí (ESA22503), Ubatuba (IAC37819), Valinhos (IAC25375)

UCs³: E.E. de Angatuba (●), E.E. de Jataí (●), P.E. do Jaraguá (●), P.E. Alberto Löfgren (●), P.E. Carlos Botelho (●), E.E. de Tupinambás (●), P.E. do Jurupará (●), P.E. de Vassununga (●), E.E. de Paulo de Faria (●), E.E. Ibicatu (●), P.E. da Cantareira (●), P.E. da Serra do Mar (●)

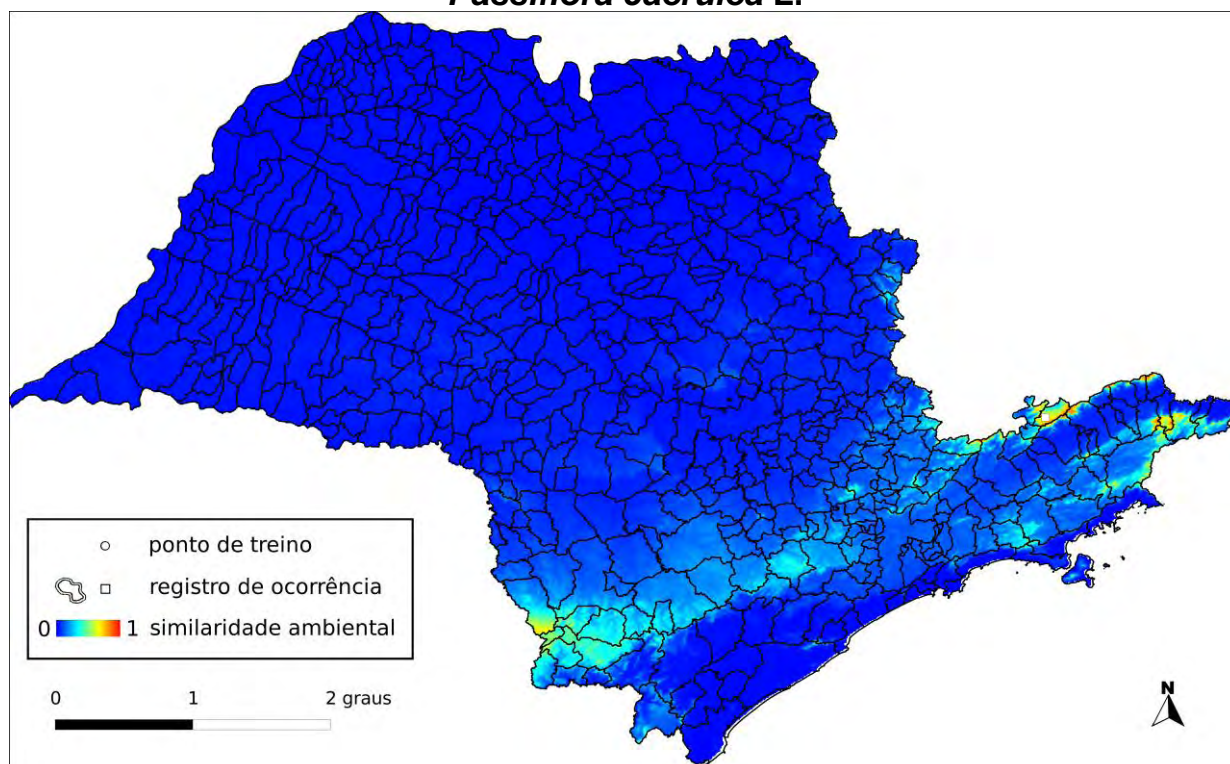
³ Registro de ocorrência inteiramente (●) ou parcialmente (◐) contido em UC.

●, P.N. da Serra da Bocaína ●, P.E. da Ilha Anchieta ●, P.E. de Guarapiranga ●, P.E. do Juquery ●, P.E. da Ara ●, E.E. de Valinhos ●, P.E. dos Mananciais de Campos do Jordão ●, P.E. de Campos do Jordão ●, P.E. de Jacupiranga ●



Histogramas: Brasil (claro) e projeção do modelo no Brasil (escuro). Valores: mínimo, médio e máximo dos pontos de ocorrência.

Passiflora caerulea L.



Algoritmo: Maxent

Total de registros disponíveis na rede *speciesLink*: 158

Omissão interna (BR): 0,00 (n=9) **Omissão interna (SP):** – (n=0)

Omissão externa (BR): 0,20 (n=59) **Omissão externa (SP):** – (n=0)

Significância (BR): $p < 0,01$

Área do modelo (SP): 30.010 km² (12,1% da área do Estado)

Área em remanescentes (2008): 12.692 km² (42,3% da área do modelo)

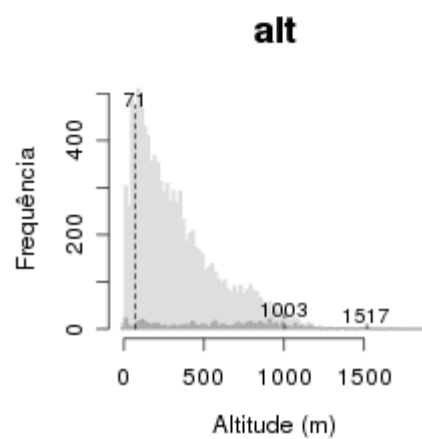
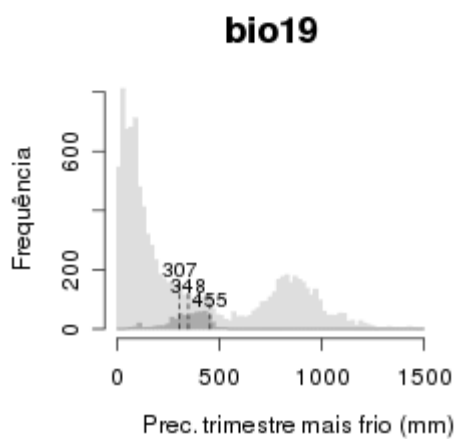
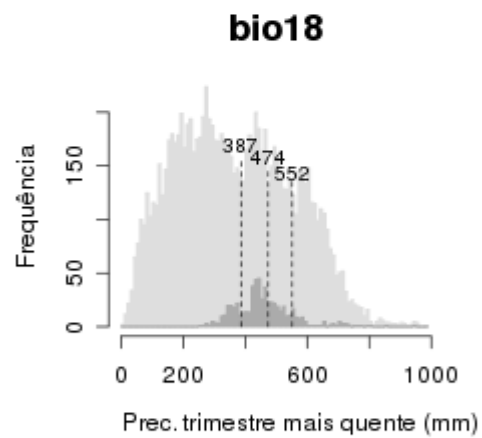
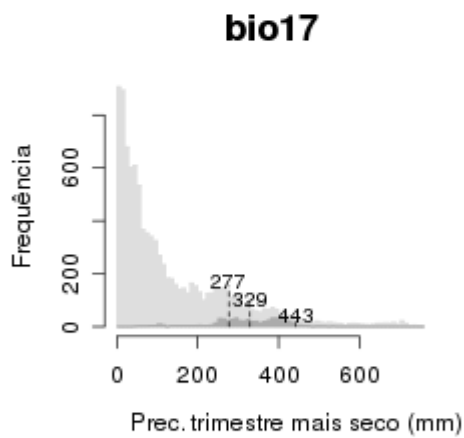
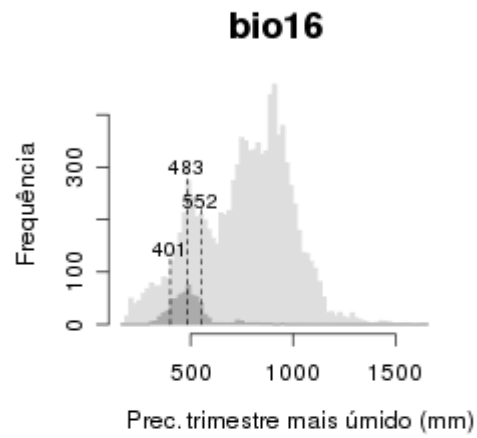
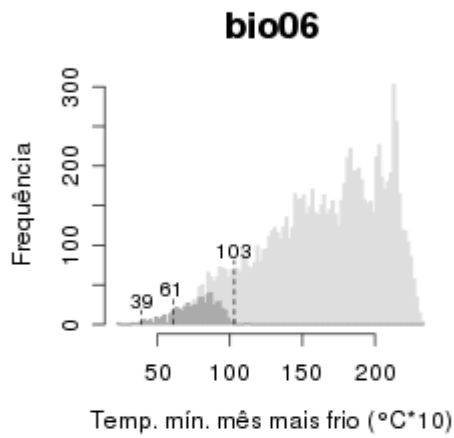
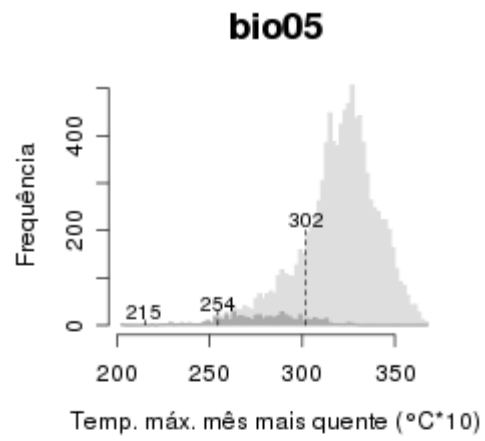
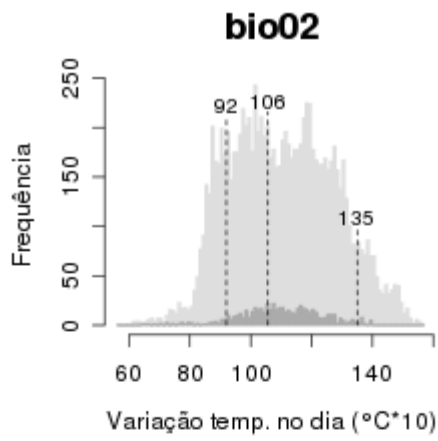
Área em remanescentes em UCs (2008): 2.389 km² (8% da área do modelo)

Ano da última coleta em SP: 1997 (IAC36161)

Municípios com registros de ocorrência em SP: sem informação

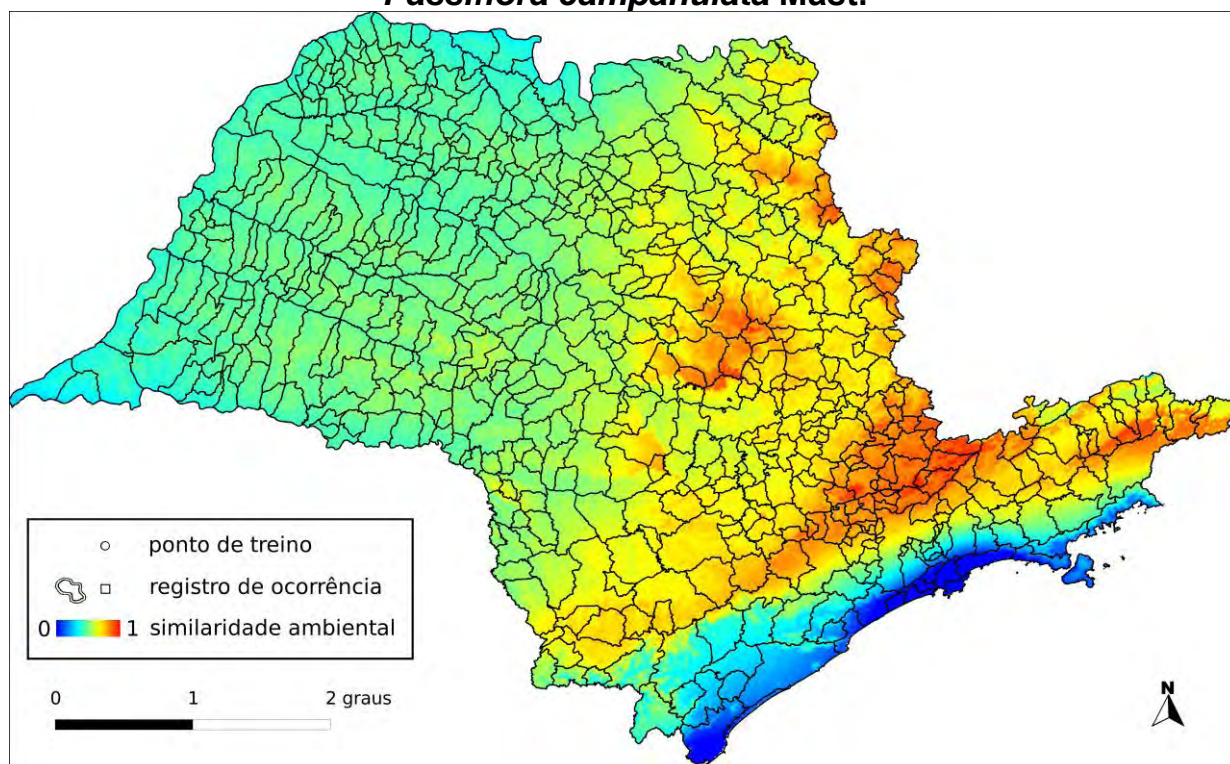
UCs⁴: P.E. de Jacupiranga ●

⁴ Registro de ocorrência inteiramente (●) ou parcialmente (◐) contido em UC.



Histogramas: Brasil (claro) e projeção do modelo no Brasil (escuro). Valores: mínimo, médio e máximo dos pontos de ocorrência.

Passiflora campanulata Mast.



Algoritmo: Distância Euclideana

Total de registros disponíveis na rede speciesLink: 68

Nº de pontos usados na modelagem: 3 (0 em SP)

Ano da última coleta em SP: 2008 (SPSF40872)

Municípios com registros de ocorrência em SP: sem informação

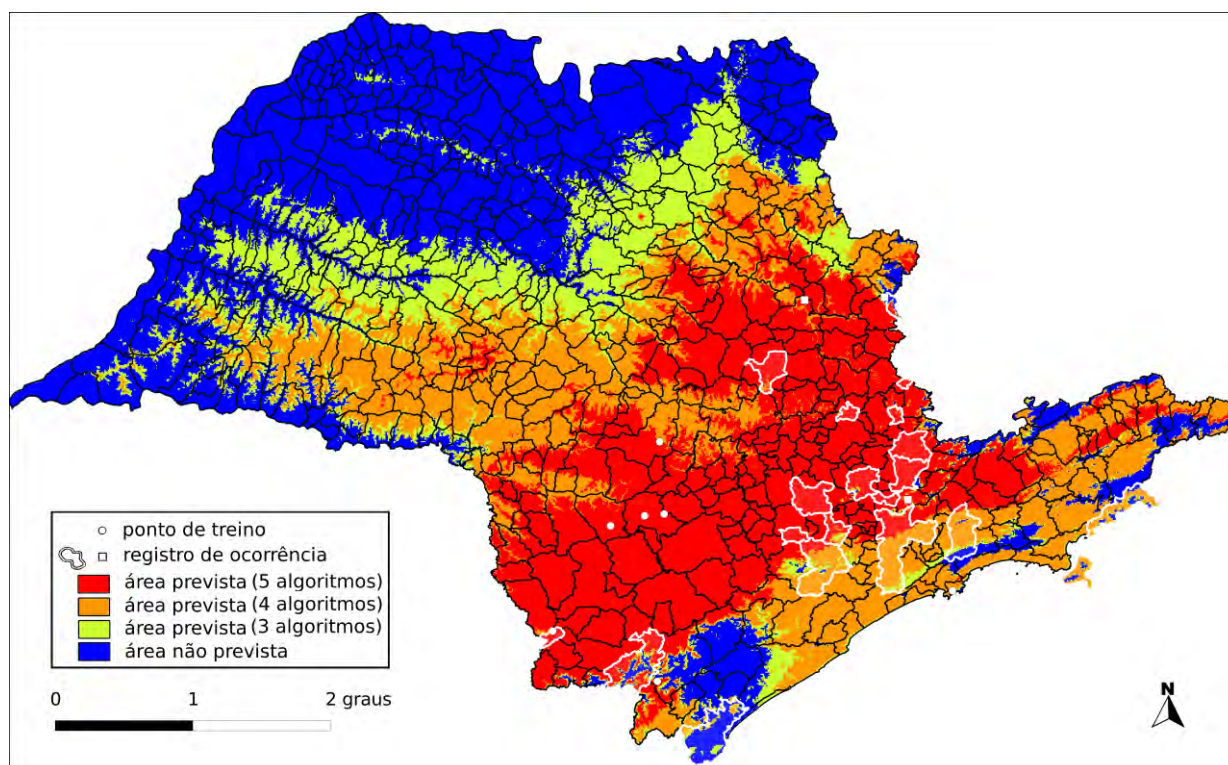
UCs⁵: P.E. da Serra do Mar ●, P.E. de Jacupiranga ●, P.N. da Serra da Bocaina ●

Dados ambientais:

bio02 var. temp. dia (°C*10)	94 – 125	bio17 prec. tri. mais seco (mm)	51 – 127
bio05 temp. max. (°C*10)	211 – 280	bio18 prec. tri. mais quente (mm)	351 – 745
bio06 temp min. (°C*10)	17 – 132	bio19 prec. tri. mais frio (mm)	70 – 127
bio16 prec. tri. mais úmido (mm)	543 – 811	alt altitude (m)	812 – 1859

⁵ Registro de ocorrência inteiramente (●) ou parcialmente (◐) contido em UC.

Passiflora capsularis L.



Algoritmo: Consenso (ENFA, GARP BS, Maxent, MD e SVM)

Total de registros disponíveis na rede *speciesLink*: 408

Omissão interna (BR): 0,00 (n=22)

Omissão interna (SP): 0,00 (n=7)

Omissão externa (BR): 0,16 (n=139)

Omissão externa (SP): 0,00 (n=20)

Significância (BR): $p < 0,01$

Área do modelo (SP): 165.977 km² (66,9% da área do Estado)

Área em remanescentes (2008): 32.326 km² (19,5% da área do modelo)

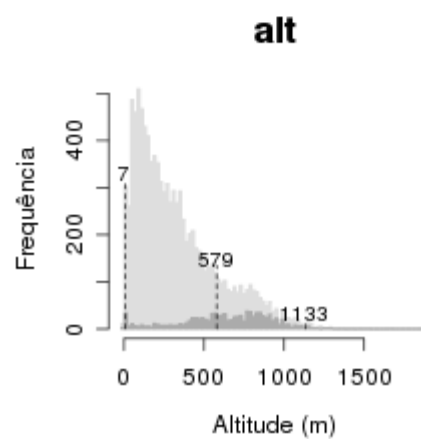
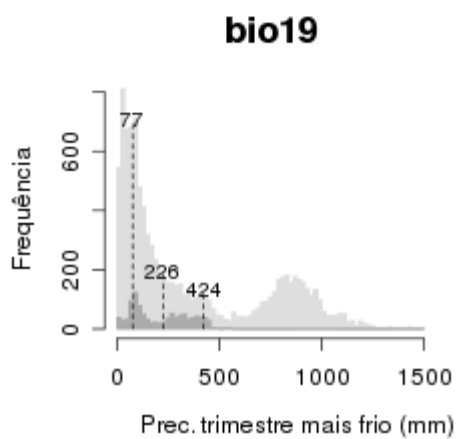
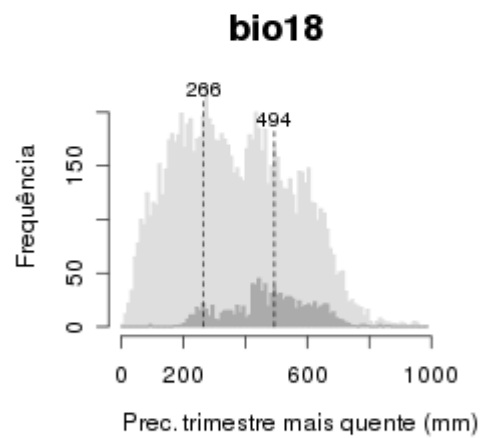
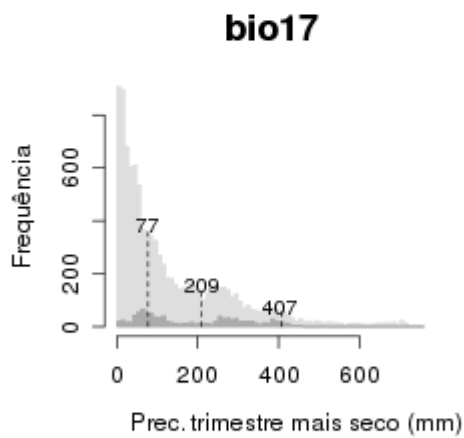
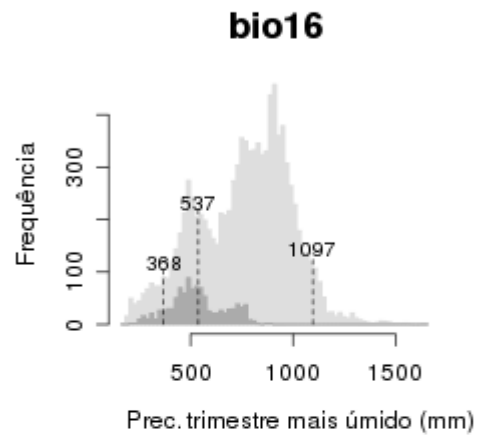
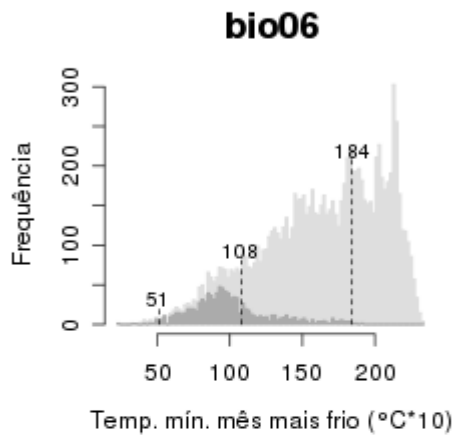
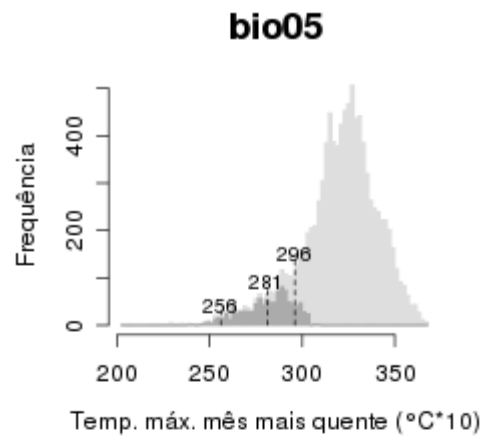
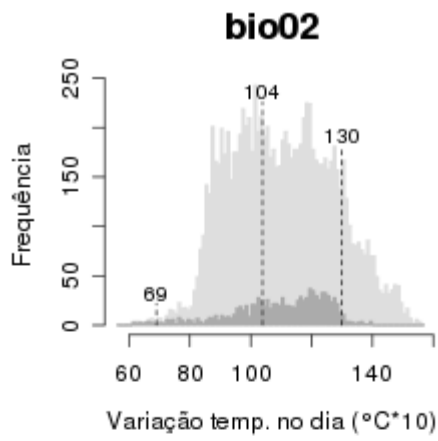
Área em remanescentes em UCs (2008): 6.386 km² (3,9% da área do modelo)

Ano da última coleta em SP: 2008 (SPSF39264)

Municípios com registros de ocorrência em SP: Angatuba (IAC46227, IAC51347), Atibaia (SP306540), Bom Sucesso de Itararé (ESA40683, IAC36731), Botucatu (UPCB40100), Bragança Paulista (IAC33766), Cananéia (SP197009), Eldorado (SPSF36466), Franco da Rocha (SPSF34493), Guarujá (IAC49356), Ibiúna (IAC41246), Ilhabela (SP82562), Iporanga (FUEL2081), Itu (IAC35964), Jaguariúna (RB312533), Jundiá (RB480947), Mogi das Cruzes (UPCB15743), Monte Alegre do Sul (IAC49547), Paranapanema (SPSF39264), Rio Claro (RB482642), São Paulo (IAC33767), São Roque (IAC7287), Ubatuba (IAC5458), Votorantim (RB283221), Águas da Prata (IAC42205), Águas de Lindóia (HTSA1745)

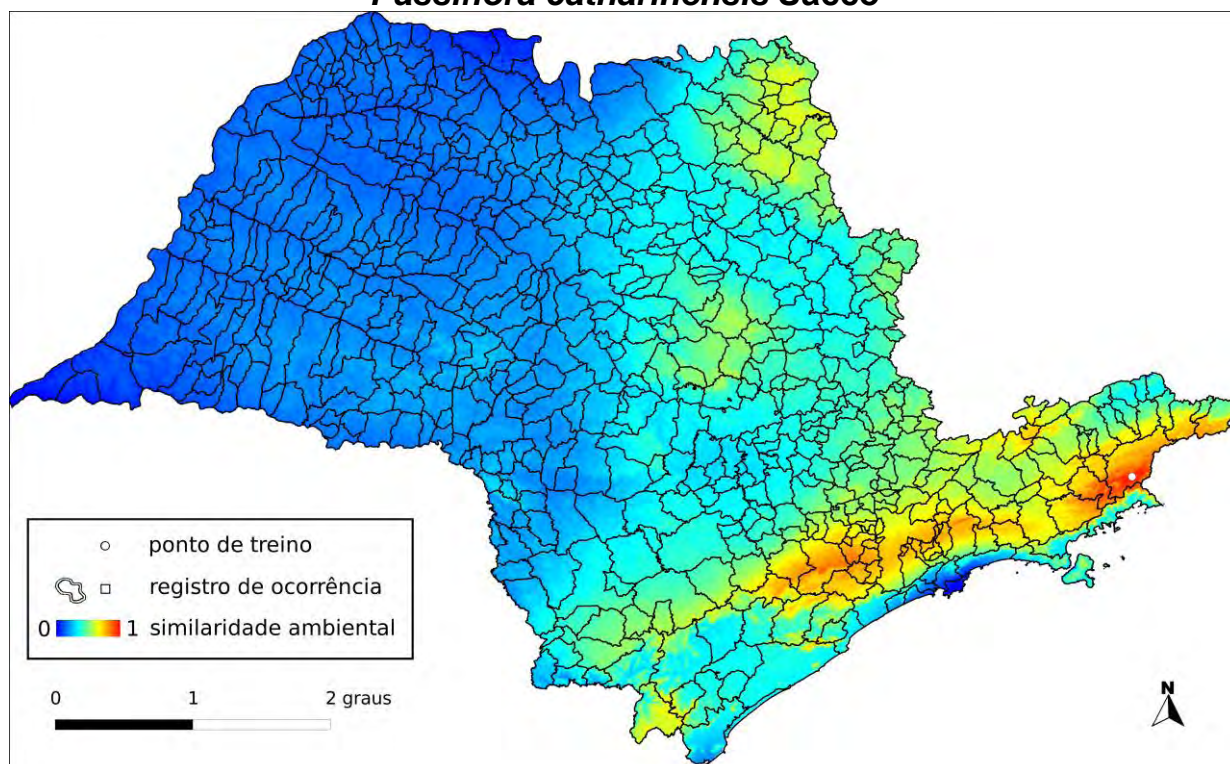
UCs⁶: P.E. de Ilhabela (●), P.E. do Juquery (●), P.E. de Jacupiranga (●), P.E. do Jaraguá (●), E.E. de Itapeti (●), P.E. Alberto Löfgren (●), E.E. de Xitué (●), P.E. Guarapiranga (●), E.E. dos Tupiniquins (●), P.E. do Alto do Ribeira (●), P.E. de Intervalos (●), P.E. da Cantareira (●), P.E. da Serra do Mar (●), P.N. da Serra da Bocaina (●), E.E. de Tupinambás (●), Ilha do Cardoso (●), P.E. da Ilha Anchieta (●), P.E. do Jurupará (●), E.E. de Paranapanema (●)

⁶ Registro de ocorrência inteiramente (●) ou parcialmente (●) contido em UC.



Histogramas: Brasil (claro) e projeção do modelo no Brasil (escuro). Valores: mínimo, médio e máximo dos pontos de ocorrência.

Passiflora catharinensis Sacco



Algoritmo: Distância Euclideana

Total de registros disponíveis na rede *speciesLink*: 8

Nº de pontos usados na modelagem: 1 (1 em SP)

Ano da última coleta em SP: 2008 (IAC49355)

Municípios com registros de ocorrência em SP: Cunha ([IAC49355](#))

UCs⁷: P.E. de Jacupiranga (●)

Dados ambientais:

bio02 var. temp. dia (°C*10)	95	bio17 prec. tri. mais seco (mm)	177
bio05 temp. max. (°C*10)	237	bio18 prec. tri. mais quente (mm)	785
bio06 temp min. (°C*10)	73	bio19 prec. tri. mais frio (mm)	177
bio16 prec. tri. mais úmido (mm)	818	alt altitude (m)	1185

⁷ Registro de ocorrência inteiramente (●) ou parcialmente (◐) contido em UC.

***Passiflora cervii* M.A.M.Azevedo**

Total de registros disponíveis na rede *speciesLink*: 35

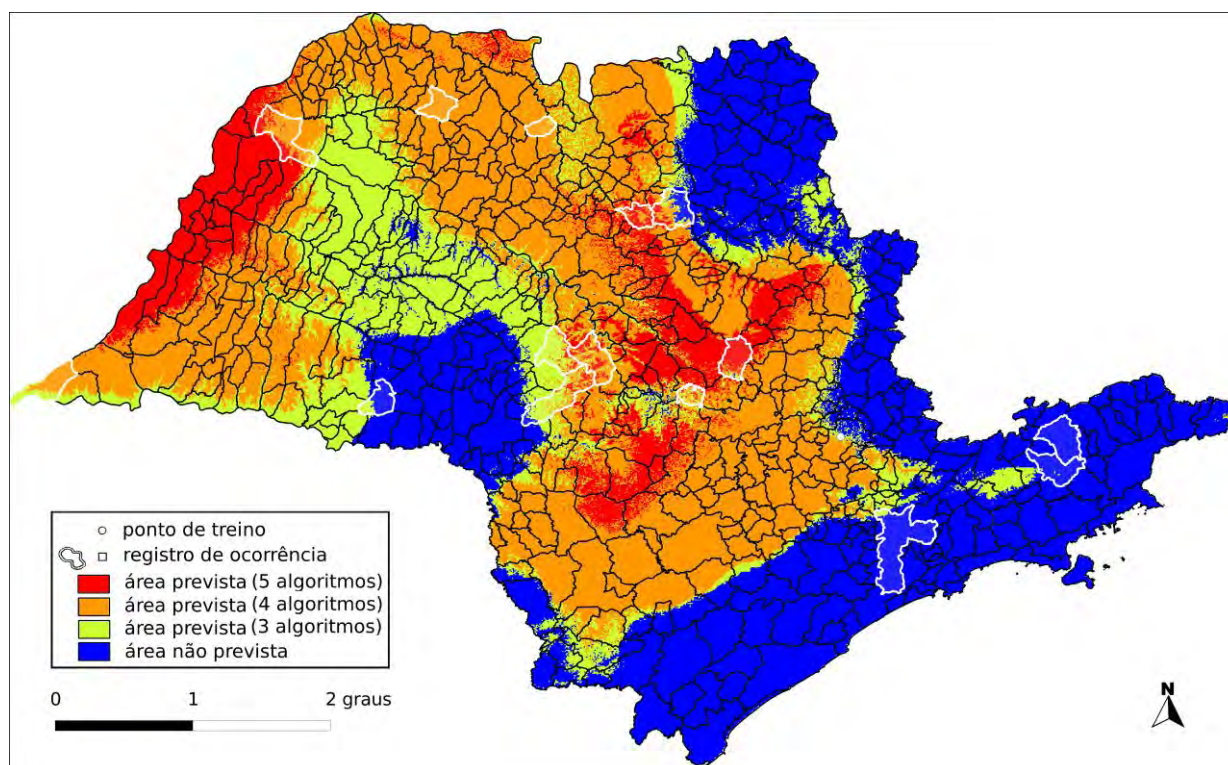
Nº de pontos usados na modelagem: 0

Ano da última coleta em SP: sem informação

Municípios com registros de ocorrência em SP: sem informação

UCs: sem informação

Passiflora cincinnata Mast.



Algoritmo: Consenso (ENFA, GARP BS, Maxent, MD e SVM)

Total de registros disponíveis na rede speciesLink: 714

Omissão interna (BR): 0,02 (n=121) **Omissão interna (SP):** 0,00 (n=1)

Omissão externa (BR): 0,05 (n=184) **Omissão externa (SP):** 0,07 (n=15)

Significância (BR): $p < 0,01$

Área do modelo (SP): 162.473 km² (65,5% da área do Estado)

Área em remanescentes (2008): 12.514 km² (7,7% da área do modelo)

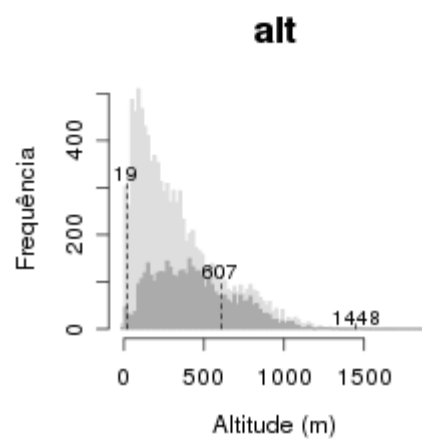
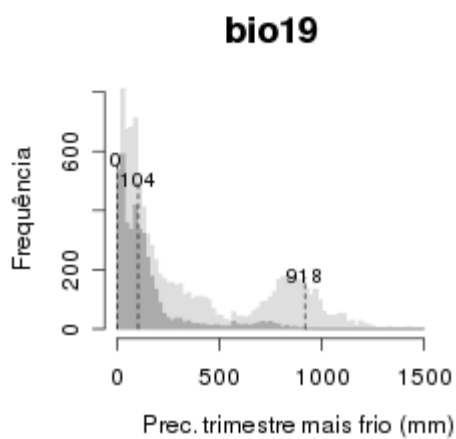
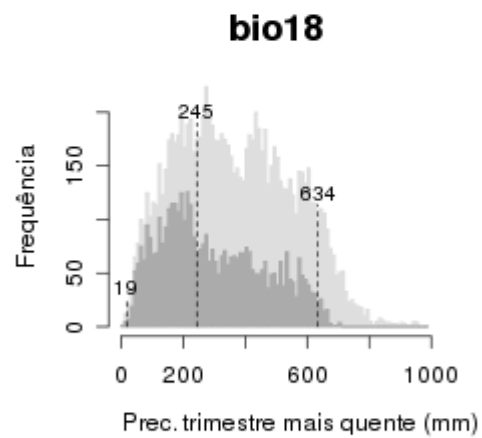
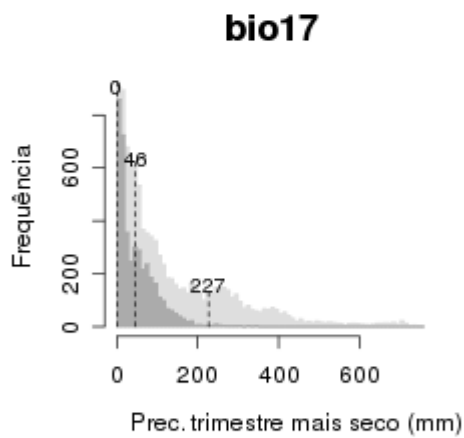
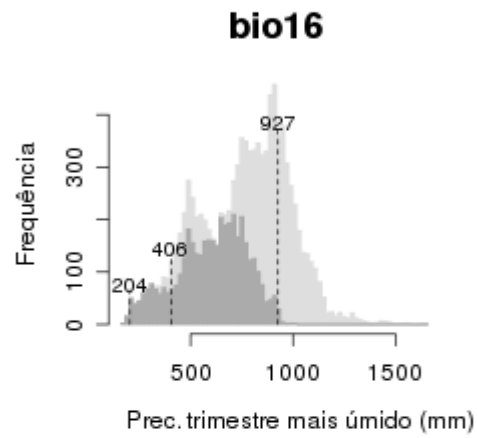
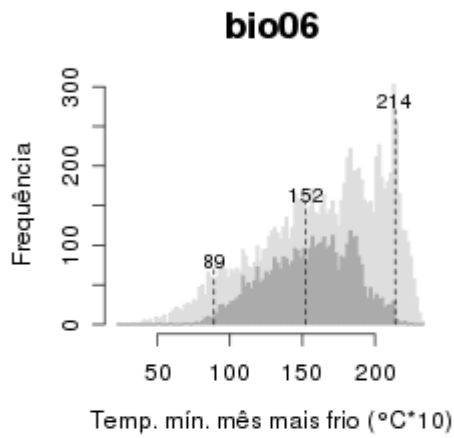
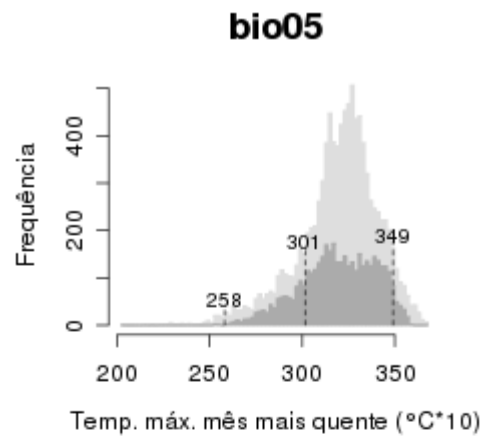
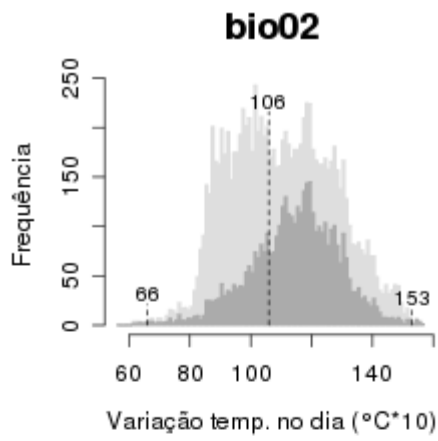
Área em remanescentes em UCs (2008): 595 km² (0,37% da área do modelo)

Ano da última coleta em SP: 2007 (IAC48098)

Municípios com registros de ocorrência em SP: Agudos (UPCB26092), Assis (SPSF35910), Bauru (IAC48062), Campinas (IAC33029), Itirapina (IAC37954), Jaboticabal (UPCB51096), Monte Alto (IAC31101), Onda Verde (SPSF18448), Pederneras (IAC21093), Pereira Barreto (IAC32735), Pindamonhangaba (**IAC19134**), Rosana (UPCB37424), Santa Maria da Serra (IAC20295), São Paulo (IAC24934), Taubaté (IAC32631), Votuporanga (IAC48098)

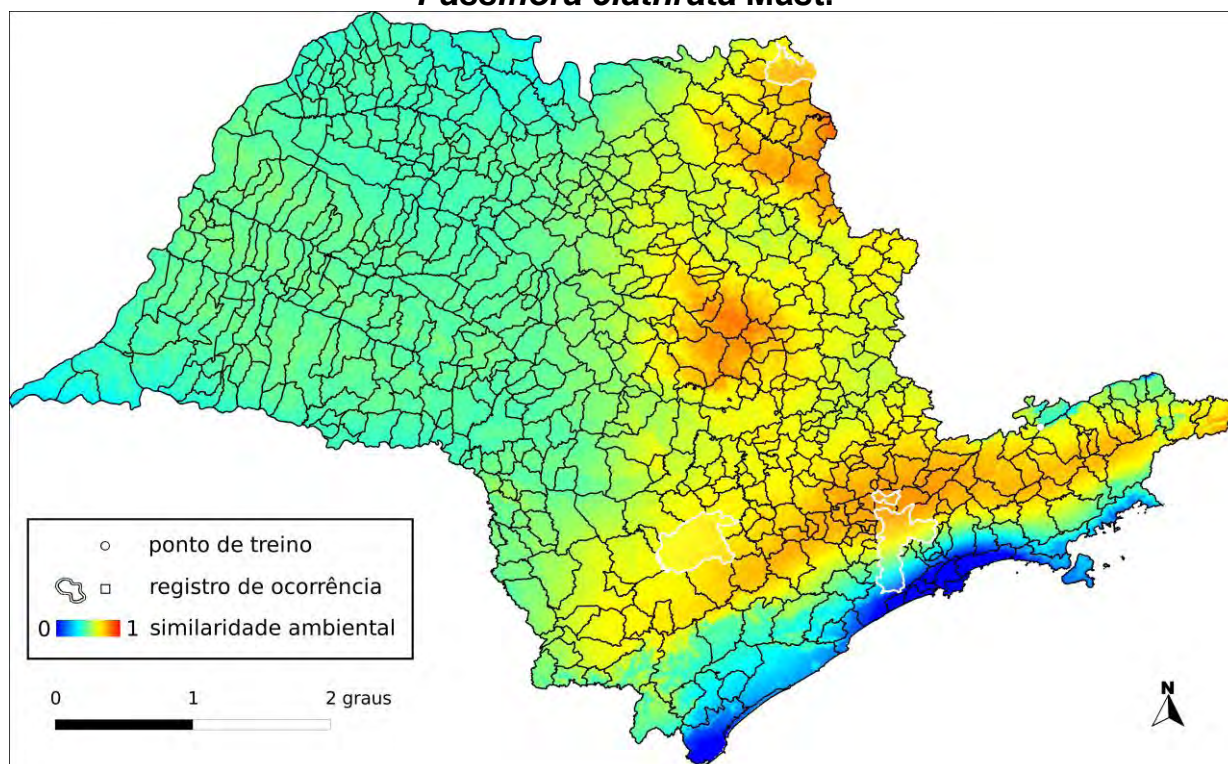
UCs⁸: P.E. de Campos do Jordão ●, E.E. Itirapina ●, P.E. do Jaraguá ●, P.E. Guarapiranga ●, E.E. de Assis ●, P.E. Alberto Löfgren ●, P.E. da Cantareira ●, P.E. da Serra do Mar ●, E.E. de Bauru ●, P.E. dos Mananciais de Campos do Jordão ●

⁸ Registro de ocorrência inteiramente (●) ou parcialmente (◐) contido em UC.



Histogramas: Brasil (claro) e projeção do modelo no Brasil (escuro). Valores: mínimo, médio e máximo dos pontos de ocorrência.

Passiflora clathrata Mast.



Algoritmo: Distância Euclideana

Total de registros disponíveis na rede speciesLink: 59

Nº de pontos usados na modelagem: 3 (1 em SP)

Ano da última coleta em SP: 1964 (UB5575)

Municípios com registros de ocorrência em SP: Campos do Jordão (UB5575), Franco da Rocha (SPSF30724), Itapetininga (RB69479), Pedregulho (SPSF23084), São Paulo (IAC32645)

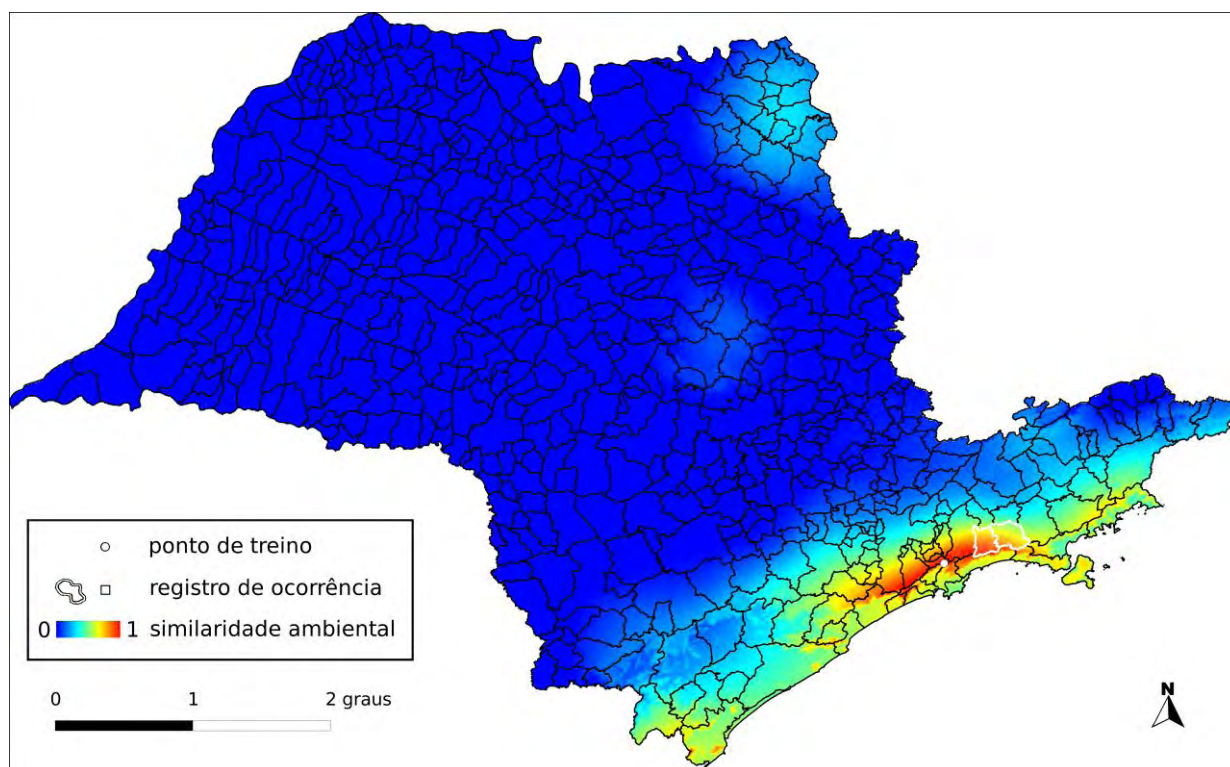
UCs⁹: P.E. do Juquery ●, P.E. do Jaraguá ●, P.E. Guarapiranga ●, F.E. Cesário Lange ●, P.E. da Cantareira ●, P.E. Alberto Löfgren ●, P.E. das Furnas do Bom Jesus ●, P.E. da Serra do Mar ●

Dados ambientais:

bio02 var. temp. dia (°C*10)	89 – 113	bio17 prec. tri. mais seco (mm)	16 – 114
bio05 temp. max. (°C*10)	217 – 297	bio18 prec. tri. mais quente (mm)	229 – 779
bio06 temp min. (°C*10)	42 – 164	bio19 prec. tri. mais frio (mm)	44 – 148
bio16 prec. tri. mais úmido (mm)	767 – 881	alt altitude (m)	967 – 1511

⁹ Registro de ocorrência inteiramente (●) ou parcialmente (◐) contido em UC.

Passiflora deidamioides Harms



Algoritmo: Distância Euclideana

Total de registros disponíveis na rede *speciesLink*: 26

Nº de pontos usados na modelagem: 1 (1 em SP)

Ano da última coleta em SP: 2004 (IAC44041)

Municípios com registros de ocorrência em SP: Biritiba-Mirim (IAC32607), Salesópolis (IAC46079), Santo André (SP197318)

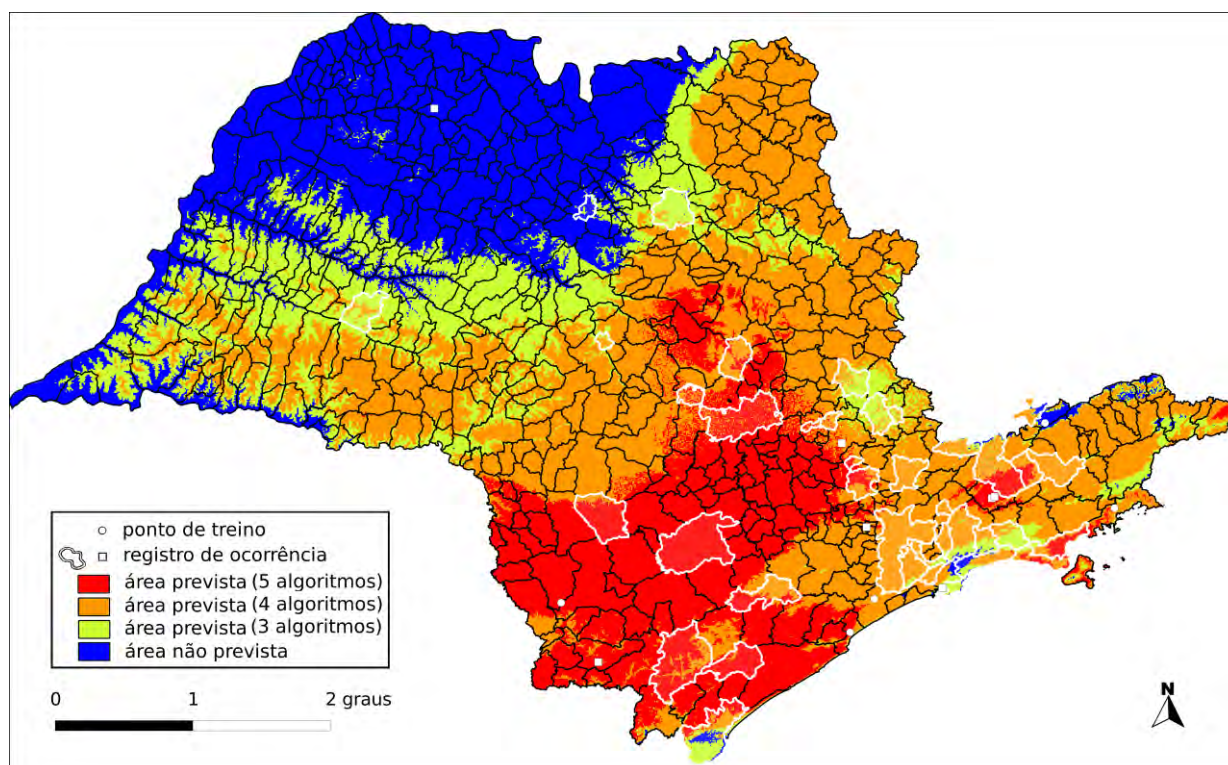
UCs¹⁰: P.E. da Serra do Mar ●, P.N. da Serra da Bocaina ○

Dados ambientais:

bio02 var. temp. dia (°C*10)	76	bio17 prec. tri. mais seco (mm)	302
bio05 temp. max. (°C*10)	242	bio18 prec. tri. mais quente (mm)	937
bio06 temp min. (°C*10)	98	bio19 prec. tri. mais frio (mm)	302
bio16 prec. tri. mais úmido (mm)	937	alt altitude (m)	669

¹⁰ Registro de ocorrência inteiramente (●) ou parcialmente (○) contido em UC.

Passiflora edulis Sims



Algoritmo: Consenso (ENFA, GARP BS, Maxent, MD e SVM)

Total de registros disponíveis na rede *speciesLink*: 828

Omissão interna (BR): 0,03 (n=29) **Omissão interna (SP):** 0,20 (n=5)

Omissão externa (BR): 0,09 (n=162) **Omissão externa (SP):** 0,00 (n=39)

Significância (BR): $p < 0,01$

Área do modelo (SP): 189.712 km² (76,4% da área do Estado)

Área em remanescentes (2008): 39.421 km² (20,8% da área do modelo)

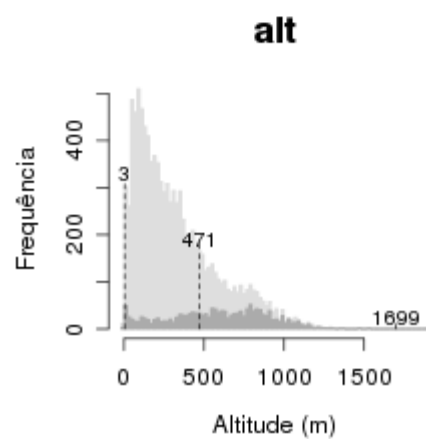
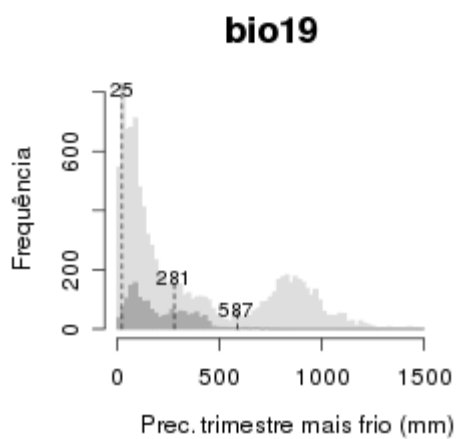
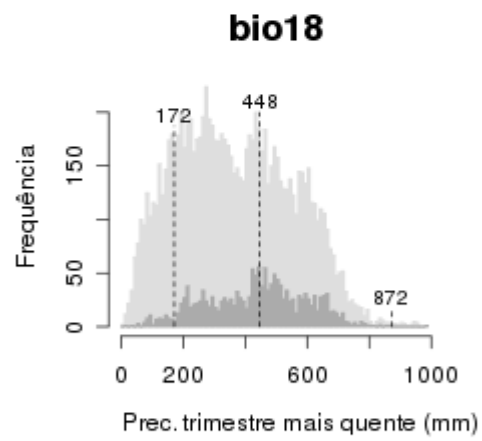
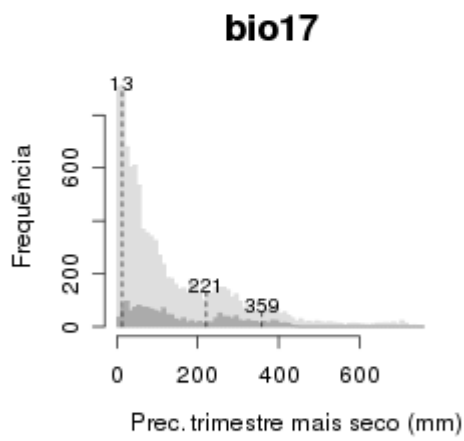
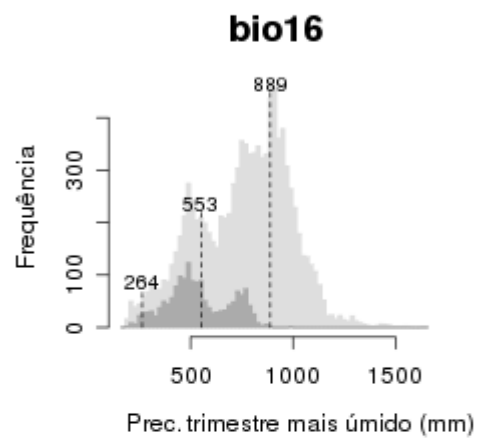
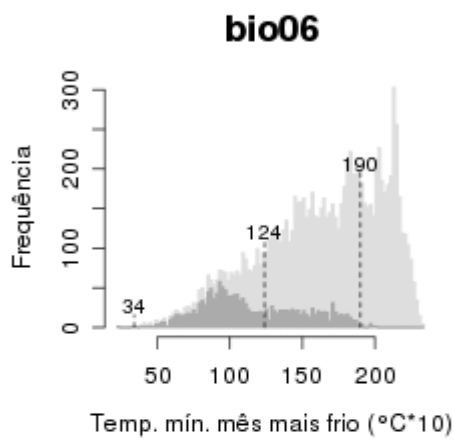
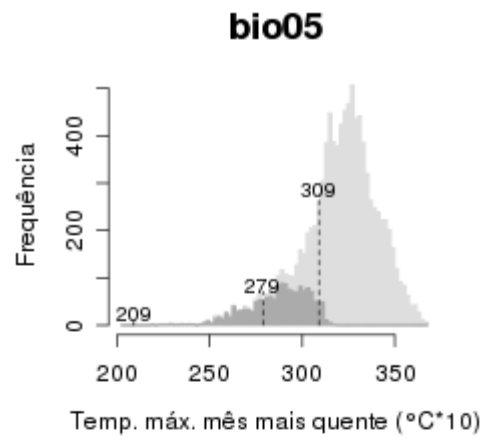
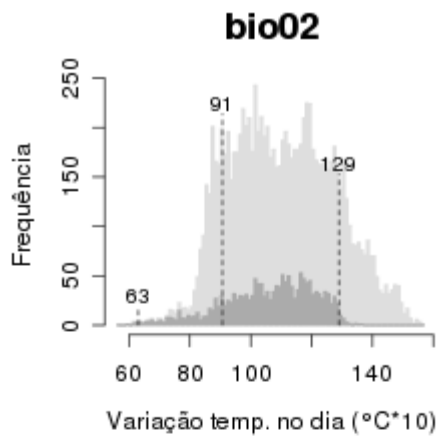
Área em remanescentes em UCs (2008): 7.442 km² (3,9% da área do modelo)

Ano da última coleta em SP: 2008 (SPSF40148)

Municípios com registros de ocorrência em SP: Amparo (ESA73151), Arujá (IAC32588), Atibaia (IAC21804), Bertioga (SP253176), Biritiba Mirim (SPSF09935), Boracéia (ESA678), Campos do Jordão (**SPSF37594**), Cananéia (IAC41248), Caraguatatuba (SPSF28814), Caçapava (ESA72863), Cubatão (MBM159785), Eldorado (ESA92633), Guarujá (IAC38097), Itanhaém (ESA75790), Itapetininga (IAC19532), Itapeva (SPSF40148), Itirapina (IAC32603), Jaboticabal (UPCB51101), Joanópolis (SP291438), Jundiaí (IAC37123), Mogi das Cruzes (UPCB15750), Mogi-Mirim (IAC37818), Monte Alegre do Sul (IAC38713), Nova Odessa (IAC32343), Paranapanema (ESA16796), Paulínia (UEC111509), Peruíbe (UPCB28156), Pindorama (IAC4750), Piracicaba (IAC40223), Registro (IAC32598), Salesópolis (IAC5839), Santa Maria da Serra (IAC21116), Santo André (IAC44354), Suzano (SP249574), São Bento do Sapucaí (SP291440), São Bernardo do Campo (UPCB15751), São José dos Campos (IAC4144), São Paulo (IAC32595), São Sebastião (ESA77469), Tapiraí (IAC32600), Taubaté (IAC2307), Tupã (FUEL4253), Ubatuba (IAC50565), Vinhedo (SPSF35459)

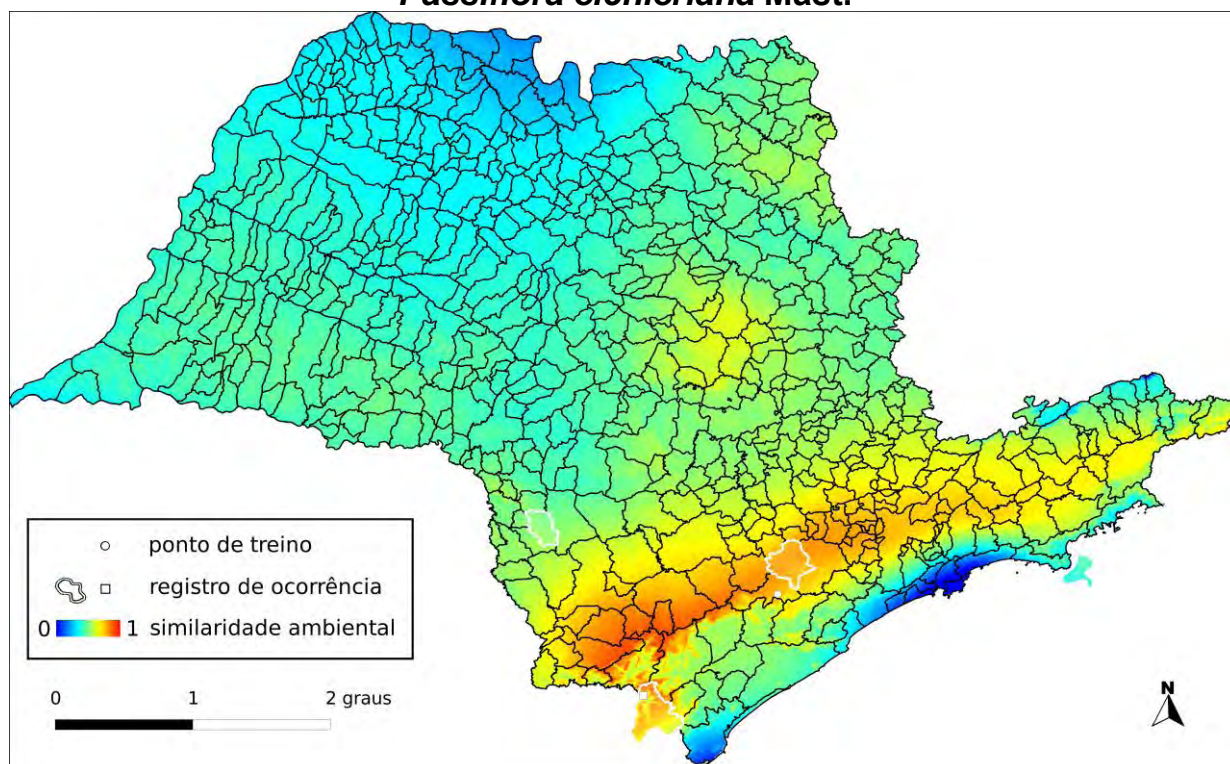
UCs¹¹: P.E. de Jacupiranga (●), E.E. Itirapina (●), P.E. Carlos Botelho (●), E.E. de Itapeti (●), E.E. Ibicatu (●), F.E. Cesário Lange (●), E.E. de Xitué (●), P.E. da Serra do Mar (●), P.E. Guarapiranga (●), E.E. dos Tupiniquins (●), P.E. de Intervalos (●), P.E. da Cantareira (●), P.E. Alberto Löfgren (●), P.N. da Serra da Bocaina (●), Ilha do Cardoso (●), P.E. do Jurupará (●), P.E. do Jaraguá (●), E.E. de Paranapanema (●)

¹¹ Registro de ocorrência inteiramente (●) ou parcialmente (●) contido em UC.



Histogramas: Brasil (claro) e projeção do modelo no Brasil (escuro). Valores: mínimo, médio e máximo dos pontos de ocorrência.

***Passiflora eichleriana* Mast.**



Algoritmo: Distância Euclideana

Total de registros disponíveis na rede *speciesLink*: 42

Nº de pontos usados na modelagem: 4 (1 em SP)

Ano da última coleta em SP: 1994 (IAC33743)

Municípios com registros de ocorrência em SP: Barra do Turvo (FUEL9610), Ilhabela (ESA21727), Piedade (IAC21429), Tapiraí (IAC33743), Taquarituba (IAC33772)

UCs¹²: P.E. de Ilhabela (●), P.E. do Jurupará (●), P.E. de Jacupiranga (●), E.E. de Tupinambás (●)

Dados ambientais:

bio02 var. temp. dia (°C*10)	90 – 115	bio17 prec. tri. mais seco (mm)	168 – 275
bio05 temp. max. (°C*10)	229 – 289	bio18 prec. tri. mais quente (mm)	466 – 687
bio06 temp min. (°C*10)	53 – 100	bio19 prec. tri. mais frio (mm)	168 – 314
bio16 prec. tri. mais úmido (mm)	466 – 724	alt altitude (m)	606 – 1409

¹² Registro de ocorrência inteiramente (●) ou parcialmente (◐) contido em UC.

***Passiflora elegans* Mast.**

Total de registros disponíveis na rede *speciesLink*: 14

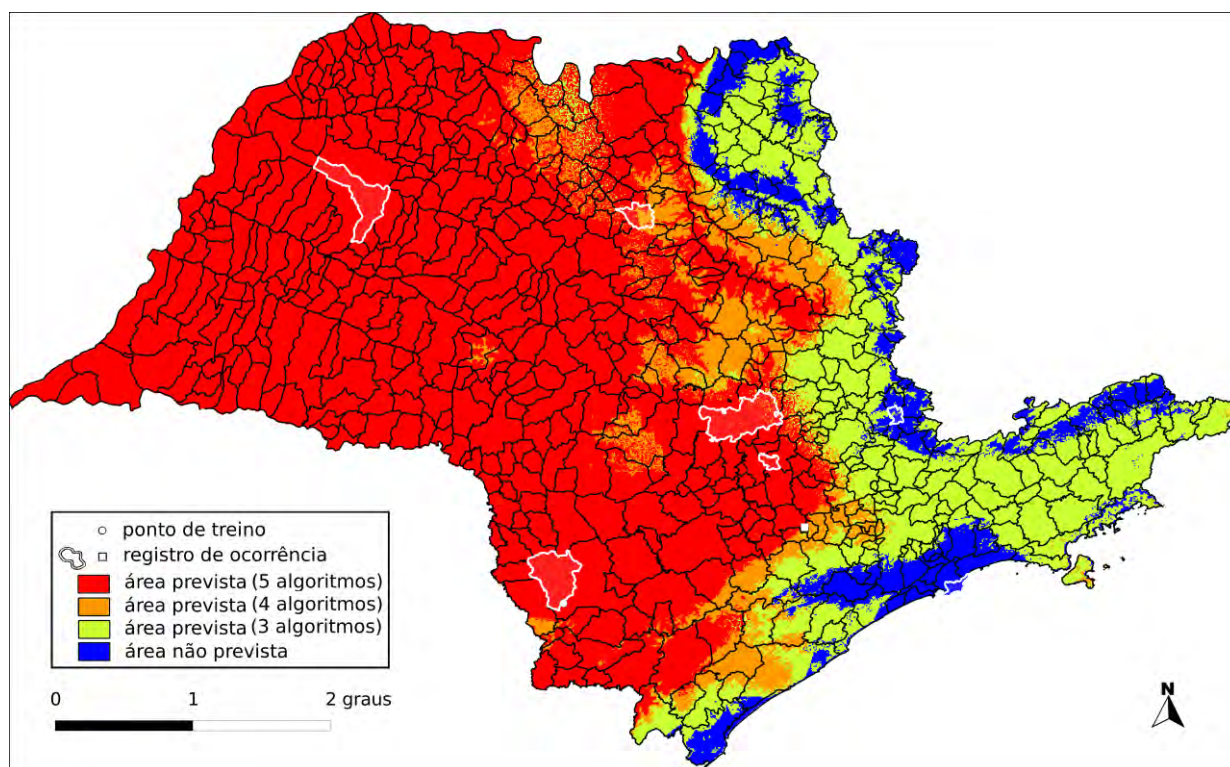
Nº de pontos usados na modelagem: 0

Ano da última coleta em SP: sem informação

Municípios com registros de ocorrência em SP: sem informação

UCs: sem informação

Passiflora foetida L.



Algoritmo: Consenso (ENFA, GARP BS, Maxent, MD e SVM)

Total de registros disponíveis na rede speciesLink: 955

Omissão interna (BR): 0,00 (n=107) **Omissão interna (SP):** 0,00 (n=1)

Omissão externa (BR): 0,04 (n=252) **Omissão externa (SP):** 0,14 (n=7)

Significância (BR): $p < 0,01$

Área do modelo (SP): 230.326 km² (92,8% da área do Estado)

Área em remanescentes (2008): 36.529 km² (15,9% da área do modelo)

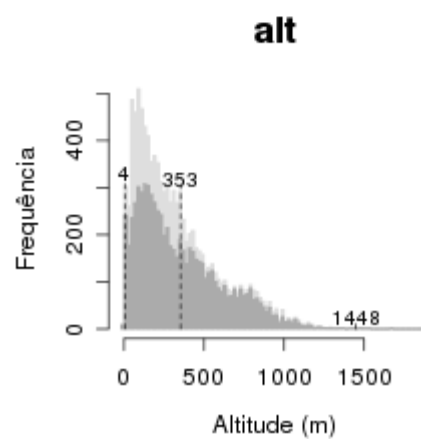
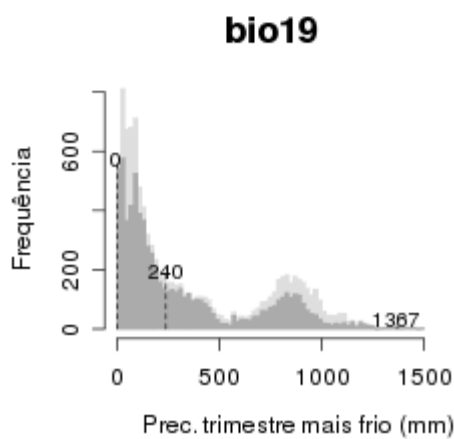
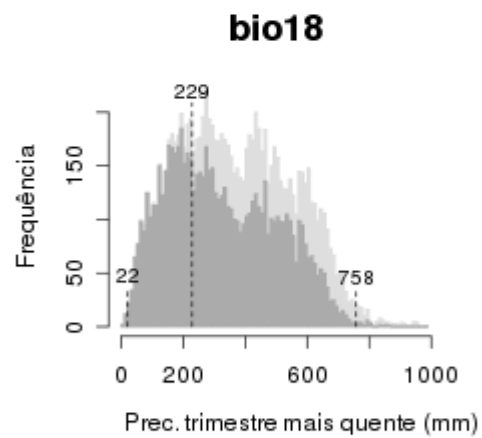
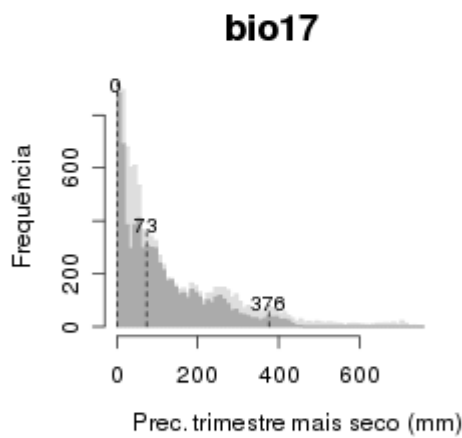
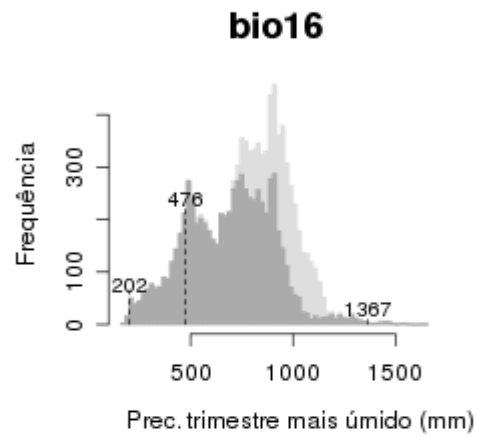
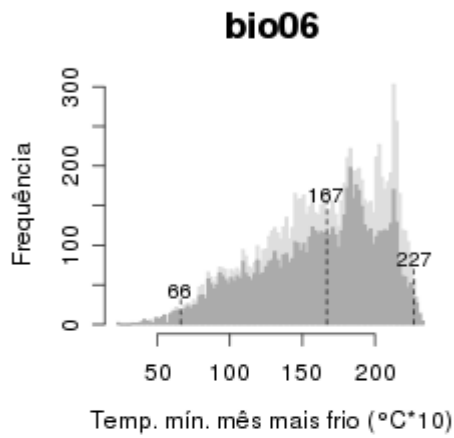
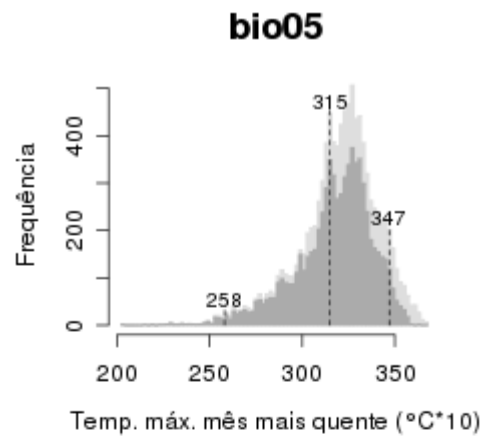
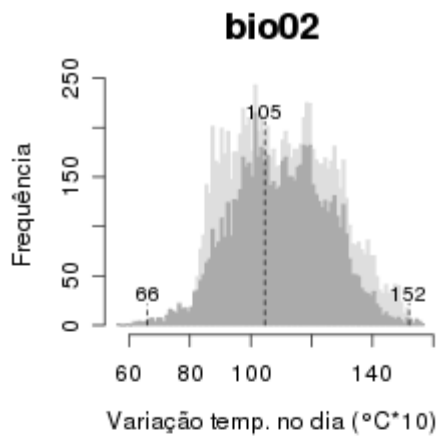
Área em remanescentes em UCs (2008): 6.079 km² (2,6% da área do modelo)

Ano da última coleta em SP: 2009 (IAC52791)

Municípios com registros de ocorrência em SP: Araçatuba (UPCB51417), Guarujá (IAC46722), Itaberá (IAC44078), Itapeva (SPSF41227), Monte Alegre do Sul (IAC52791), Monte Alto (IAC30006), Piracicaba (IAC32648), Rafard (IAC23604)

UCs¹³: E.E. de Itaberá ●, E.E. Ibicatu ●

¹³ Registro de ocorrência inteiramente (●) ou parcialmente (◐) contido em UC.



Histogramas: Brasil (claro) e projeção do modelo no Brasil (escuro). Valores: mínimo, médio e máximo dos pontos de ocorrência.

Passiflora glaucescens Killip

Total de registros disponíveis na rede *speciesLink*: 4

Nº de pontos usados na modelagem: 0

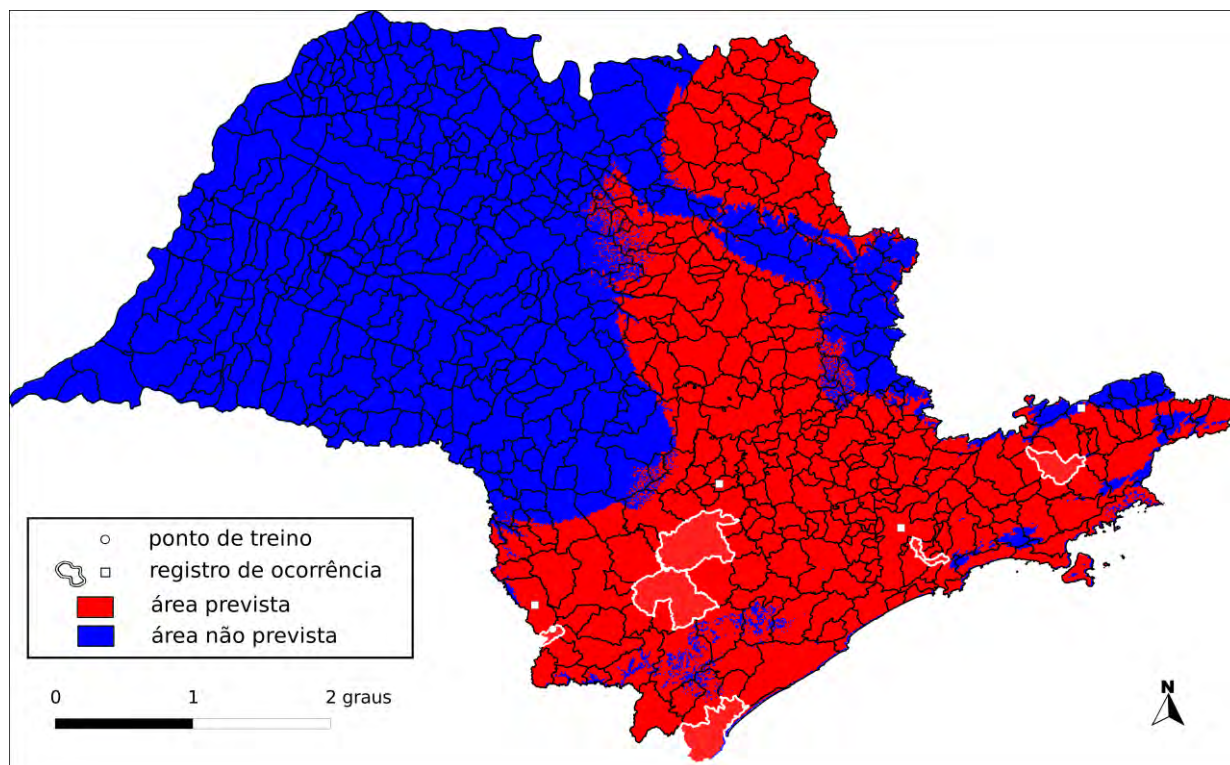
Ano da última coleta em SP: 1976 (RB195669)

Municípios com registros de ocorrência em SP: Itararé (RB195669), São Paulo (IAC37655)

UCs¹⁴: P.E. Alberto Löfgren ●, P.E. da Cantareira ●, P.E. do Jaraguá ●, P.E. Guarapiranga ●, P.E. da Serra do Mar ●

¹⁴ Registro de ocorrência inteiramente (●) ou parcialmente (◐) contido em UC.

Passiflora haematostigma Mart. ex Mast.



Algoritmo: Consenso (GARP BS e Maxent)

Total de registros disponíveis na rede speciesLink: 261

Omissão interna (BR): 0,20 (n=15) **Omissão interna (SP):** 0,00 (n=1)

Omissão externa (BR): 0,19 (n=74) **Omissão externa (SP):** 0,00 (n=6)

Significância (BR): $p < 0,01$

Área do modelo (SP): 111.234 km² (44,8% da área do Estado)

Área em remanescentes (2008): 31.571 km² (28,4% da área do modelo)

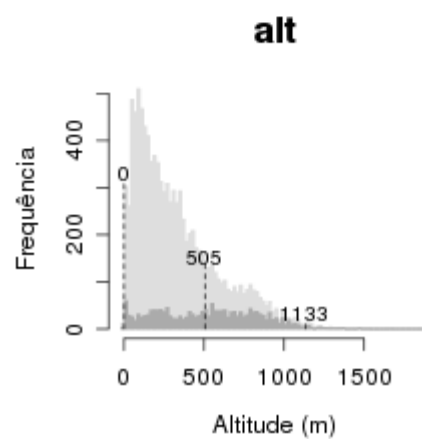
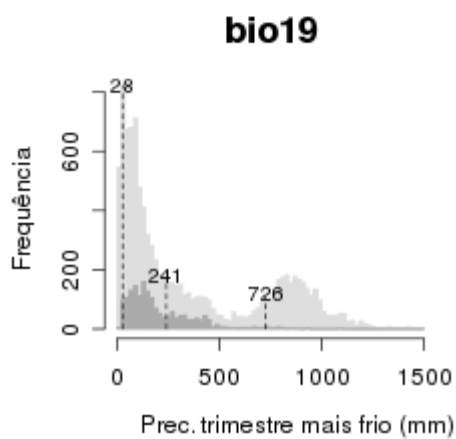
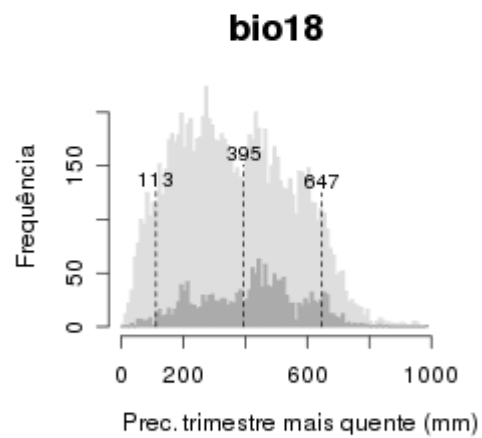
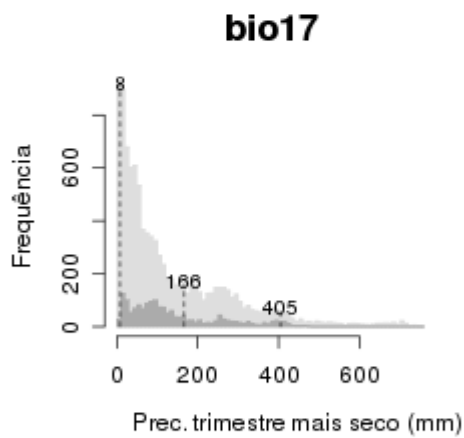
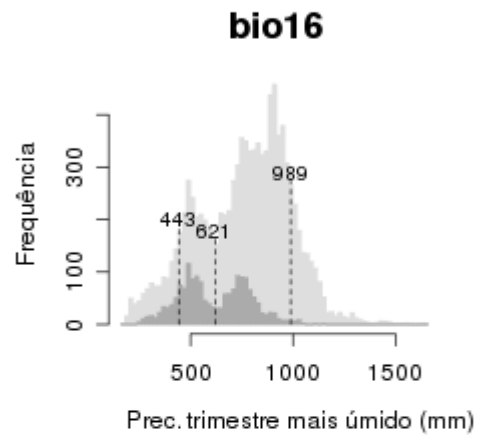
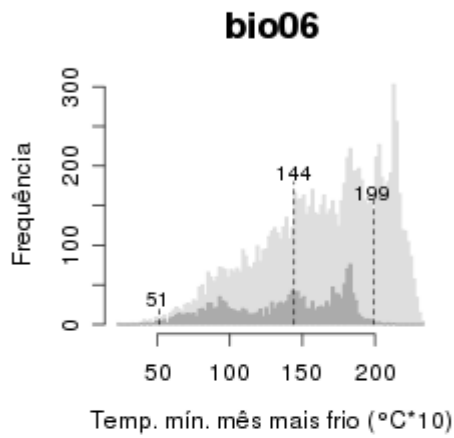
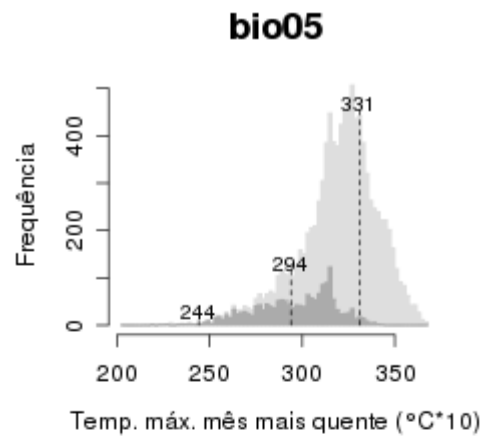
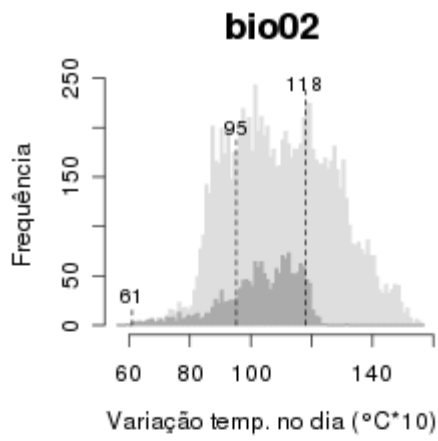
Área em remanescentes em UCs (2008): 6.665 km² (6% da área do modelo)

Ano da última coleta em SP: 2004 (IAC44040)

Municípios com registros de ocorrência em SP: Bom Sucesso de Itararé (ESA40580, IAC36730), Cananéia (UPCB18464), Capão Bonito (IAC31095), Itapetininga (MBM49358), Santo André (IAC44040), Taubaté (IAC22994)

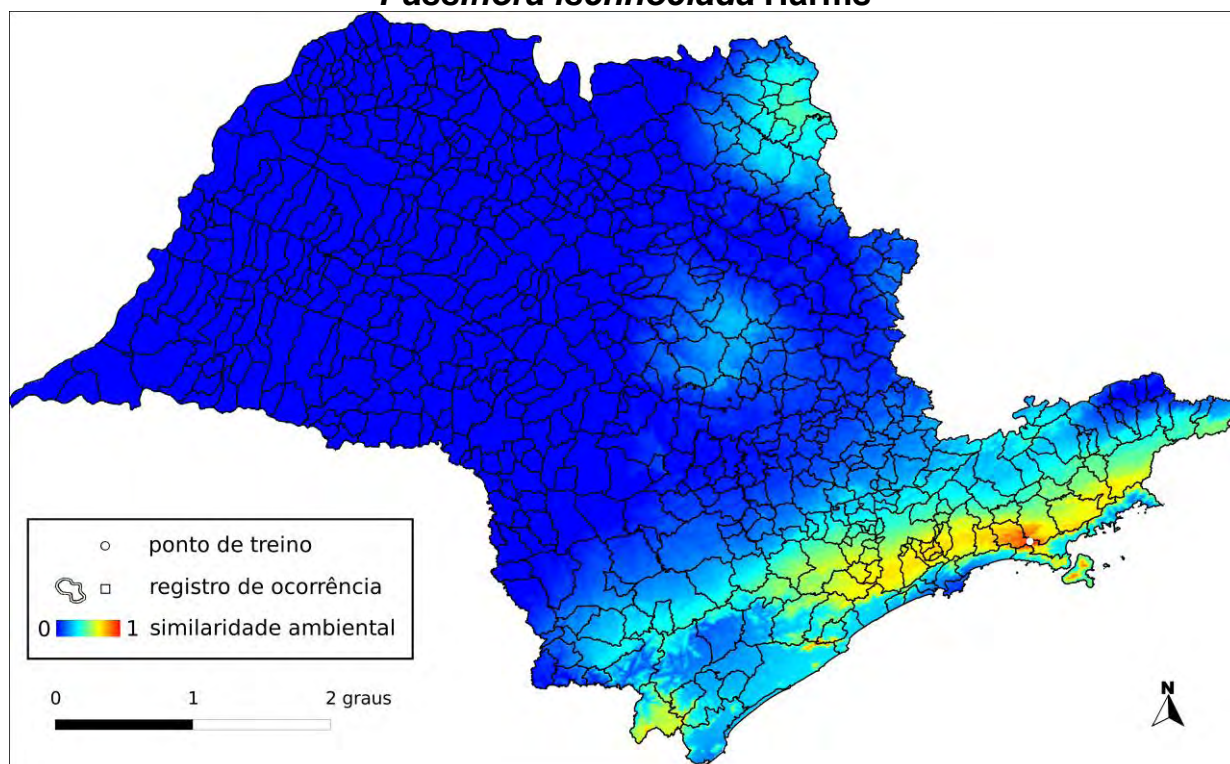
UCs¹⁵: P.E. de Campos do Jordão ●, P.E. de Jacupiranga ●, F.E. Cesário Lange ●, E.E. dos Tupiniquins ●, P.E. de Intervales ●, P.E. Carlos Botelho ●, Ilha do Cardoso ●, P.E. da Serra do Mar ●

¹⁵ Registro de ocorrência inteiramente (●) ou parcialmente (◐) contido em UC.



Histogramas: Brasil (claro) e projeção do modelo no Brasil (escuro). Valores: mínimo, médio e máximo dos pontos de ocorrência.

Passiflora ischnoclada Harms



Algoritmo: Distância Euclideana

Total de registros disponíveis na rede *speciesLink*: 6

Nº de pontos usados na modelagem: 1 (1 em SP)

Ano da última coleta em SP: 2003 (IAC44081)

Municípios com registros de ocorrência em SP: Paraibuna ([IAC44081](#))

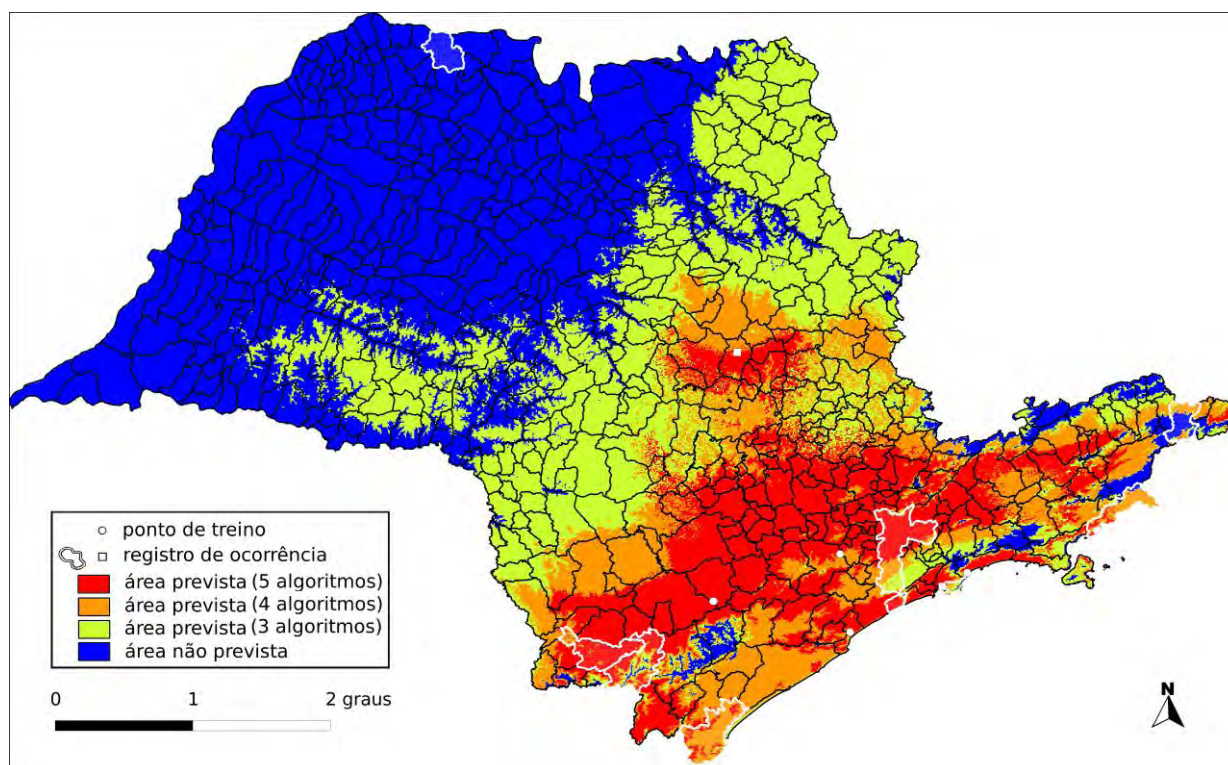
UCs¹⁶: P.E. da Serra do Mar ●

Dados ambientais:

bio02 var. temp. dia (°C*10)	76	bio17 prec. tri. mais seco (mm)	244
bio05 temp. max. (°C*10)	216	bio18 prec. tri. mais quente (mm)	759
bio06 temp min. (°C*10)	74	bio19 prec. tri. mais frio (mm)	244
bio16 prec. tri. mais úmido (mm)	776	alt altitude (m)	1243

¹⁶ Registro de ocorrência inteiramente (●) ou parcialmente (◐) contido em UC.

Passiflora jilekii Wawra



Algoritmo: Consenso (ENFA, GARP BS, Maxent, MD e SVM)

Total de registros disponíveis na rede speciesLink: 216

Omissão interna (BR): 0,00 (n=21) **Omissão interna (SP):** 0,00 (n=5)

Omissão externa (BR): 0,16 (n=62) **Omissão externa (SP):** 0,11 (n=9)

Significância (BR): $p < 0,01$

Área do modelo (SP): 141.764 km² (57,1% da área do Estado)

Área em remanescentes (2008): 34.184 km² (24,1% da área do modelo)

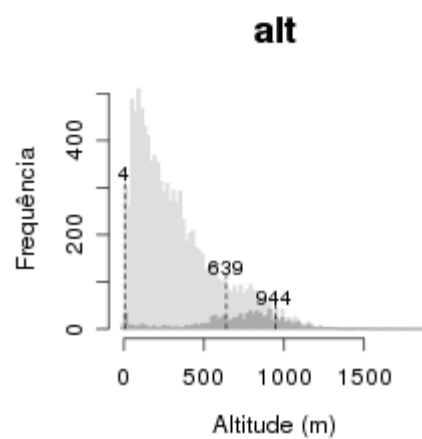
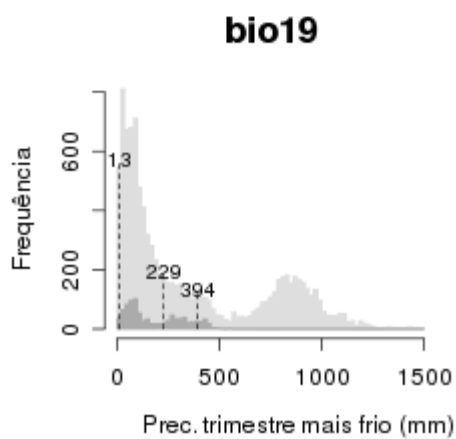
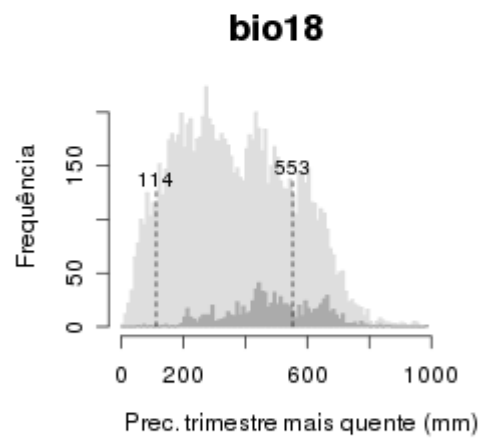
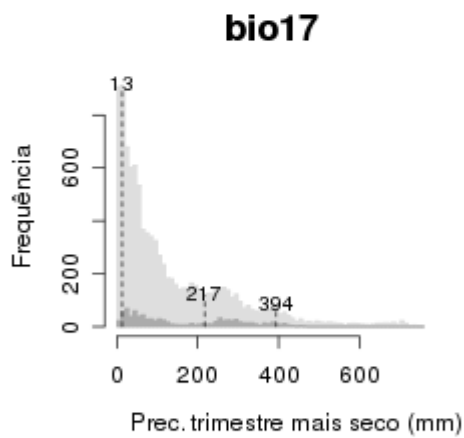
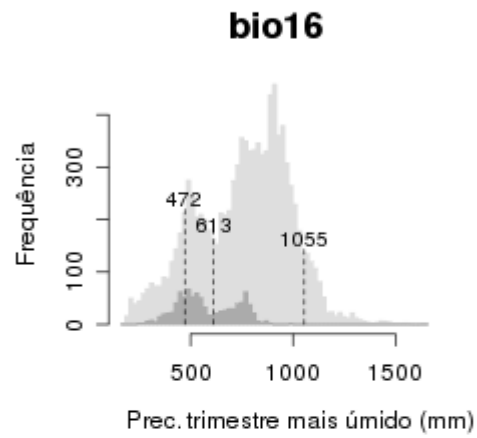
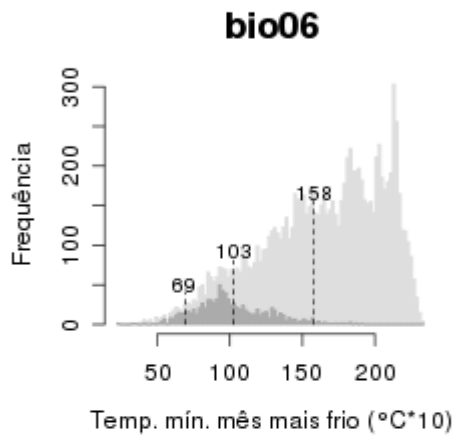
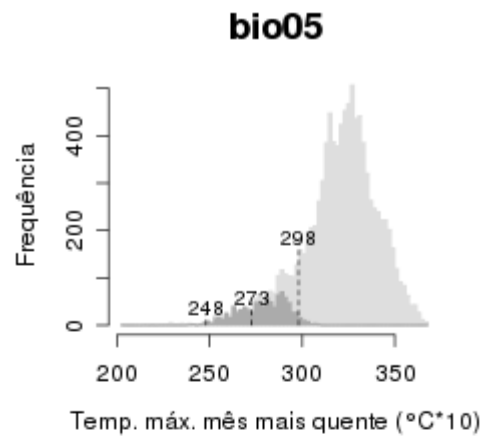
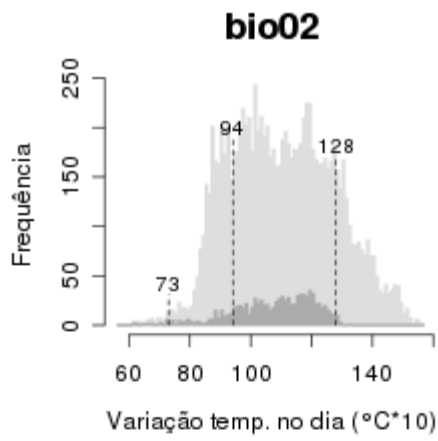
Área em remanescentes em UCs (2008): 6.741 km² (4,7% da área do modelo)

Ano da última coleta em SP: 2005 (IAC46723)

Municípios com registros de ocorrência em SP: Apiaí (IAC42204), Bertioga (UPCB39435), Cananéia (UPCB19995), Cardoso (**RB419964**), Guarujá (IAC46723), Ibiúna (IAC41023), Iporanga (RB115820), Mongaguá (UPCB20024), Peruíbe (UPCB28160), São José do Barreiro (RB482643), São Miguel Arcanjo (ESA79597,ESA79759), São Paulo (IAC33781), Ubatuba (IAC21113)

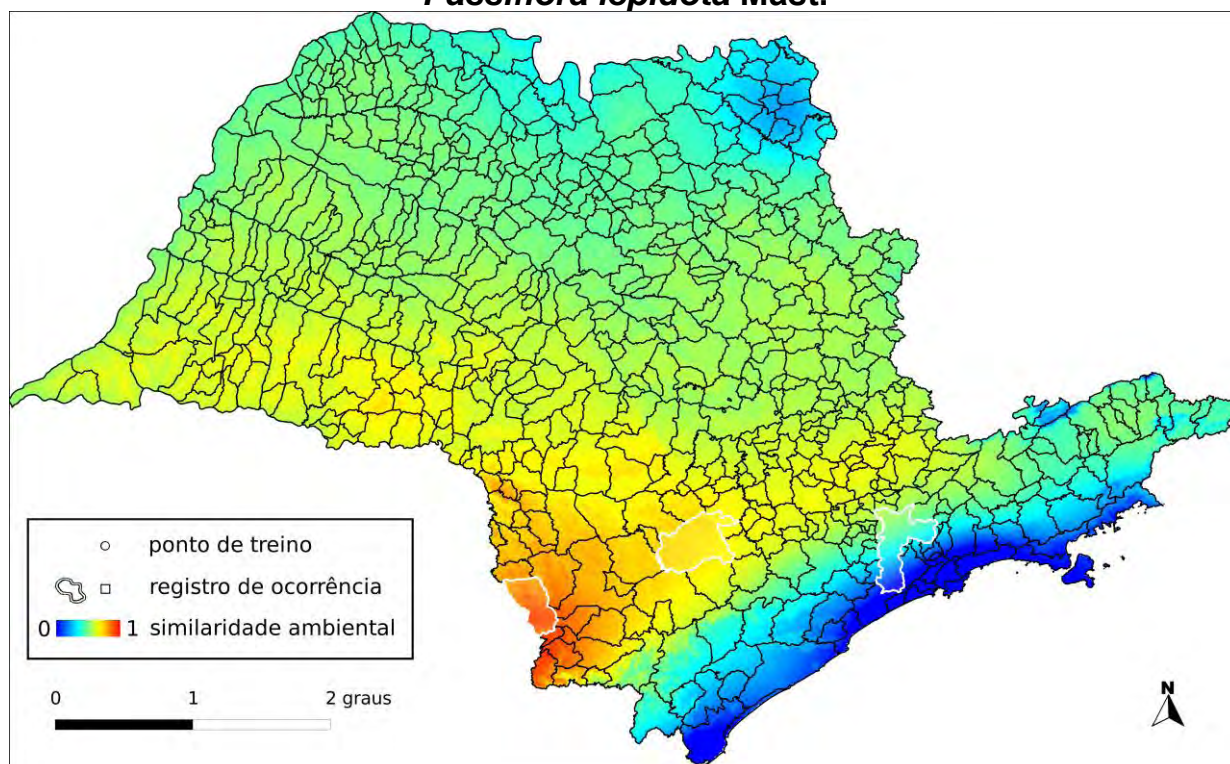
UCs¹⁷: P.E. de Jacupiranga ●, P.E. do Jaraguá ●, P.E. Guarapiranga ●, P.E. Alberto Löfgren ●, E.E. de Xitúé ●, E.E. dos Tupiniquins ●, P.E. do Alto do Ribeira ●, P.E. de Intervalos ●, P.E. da Cantareira ●, P.E. Carlos Botelho ●, P.N. da Serra da Bocaina ●, E.E. de Tupinambás ●, Ilha do Cardoso ●, P.E. da Ilha Anchieta ●, P.E. da Serra do Mar ●

¹⁷ Registro de ocorrência inteiramente (●) ou parcialmente (◐) contido em UC.



Histogramas: Brasil (claro) e projeção do modelo no Brasil (escuro). Valores: mínimo, médio e máximo dos pontos de ocorrência.

Passiflora lepidota Mast.



Algoritmo: Distância Euclideana

Total de registros disponíveis na rede speciesLink: 41

Nº de pontos usados na modelagem: 2 (0 em SP)

Ano da última coleta em SP: 1993 (IAC32647)

Municípios com registros de ocorrência em SP: Itapetininga (RB55460), Itararé (IAC32647), São Paulo (IAC32646)

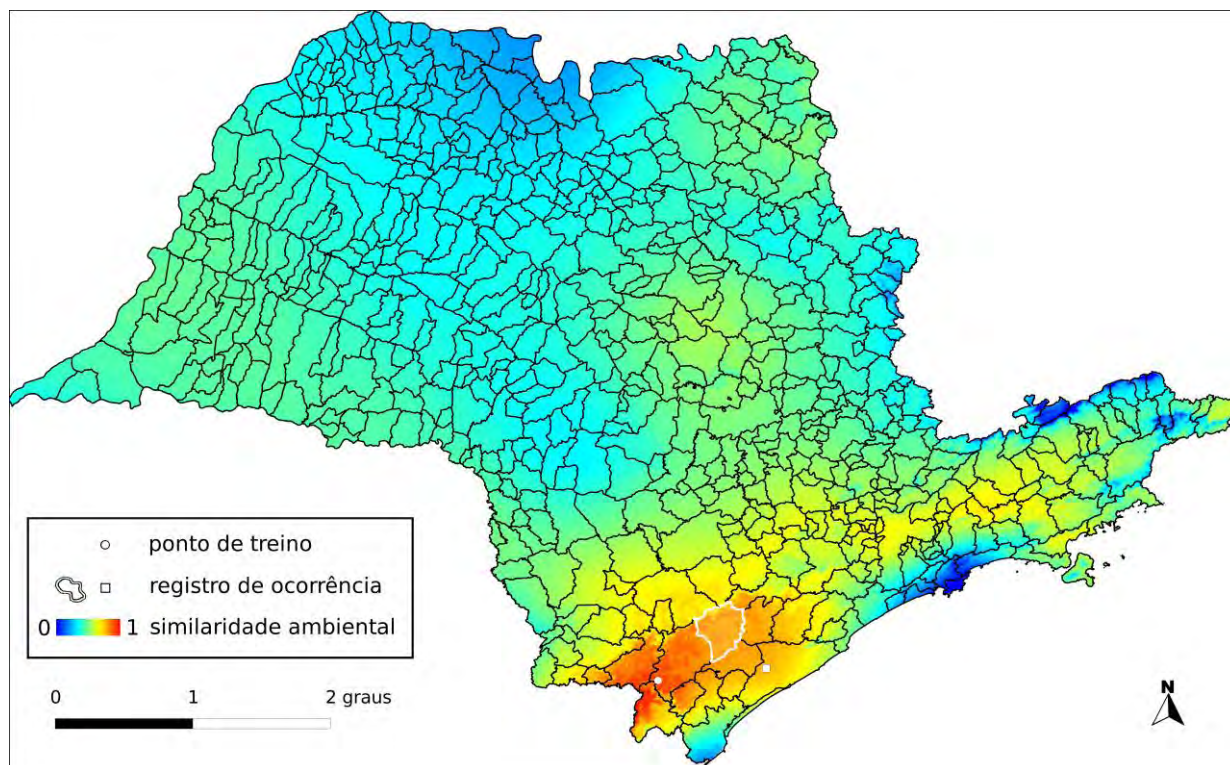
UCs¹⁸: P.E. do Jaraguá ●, P.E. Guarapiranga ●, F.E. Cesário Lange ●, P.E. da Cantareira ●, P.E. Alberto Löfgren ●, P.E. da Serra do Mar ●

Dados ambientais:

bio02 var. temp. dia (°C*10)	114 – 127	bio17 prec. tri. mais seco (mm)	202 – 271
bio05 temp. max. (°C*10)	263 – 275	bio18 prec. tri. mais quente (mm)	448 – 497
bio06 temp min. (°C*10)	53 – 70	bio19 prec. tri. mais frio (mm)	221 – 287
bio16 prec. tri. mais úmido (mm)	480 – 521	alt altitude (m)	813 – 964

¹⁸ Registro de ocorrência inteiramente (●) ou parcialmente (◐) contido em UC.

Passiflora loefgrenii Vitta



Algoritmo: Distância Euclideana

Total de registros disponíveis na rede speciesLink: 19

Nº de pontos usados na modelagem: 2 (1 em SP)

Ano da última coleta em SP: 2003 (ESA84923)

Municípios com registros de ocorrência em SP: Eldorado ([IAC33788](#)), Sete Barras (ESA52140)

UCs¹⁹: Carlos Botelho (●), P.E. de Jacupiranga (●), P.N. da Serra da Bocaina (●), P.E. de Intervales (●)

Dados ambientais:

bio02 var. temp. dia (°C*10)	93 – 96	bio17 prec. tri. mais seco (mm)	211 – 280
bio05 temp. max. (°C*10)	284 – 297	bio18 prec. tri. mais quente (mm)	522 – 599
bio06 temp min. (°C*10)	97 – 113	bio19 prec. tri. mais frio (mm)	211 – 288
bio16 prec. tri. mais úmido (mm)	522 – 599	alt altitude (m)	30 – 510

¹⁹ Registro de ocorrência inteiramente (●) ou parcialmente (◐) contido em UC.

***Passiflora malacophylla* Mast.**

Total de registros disponíveis na rede *speciesLink*: 27

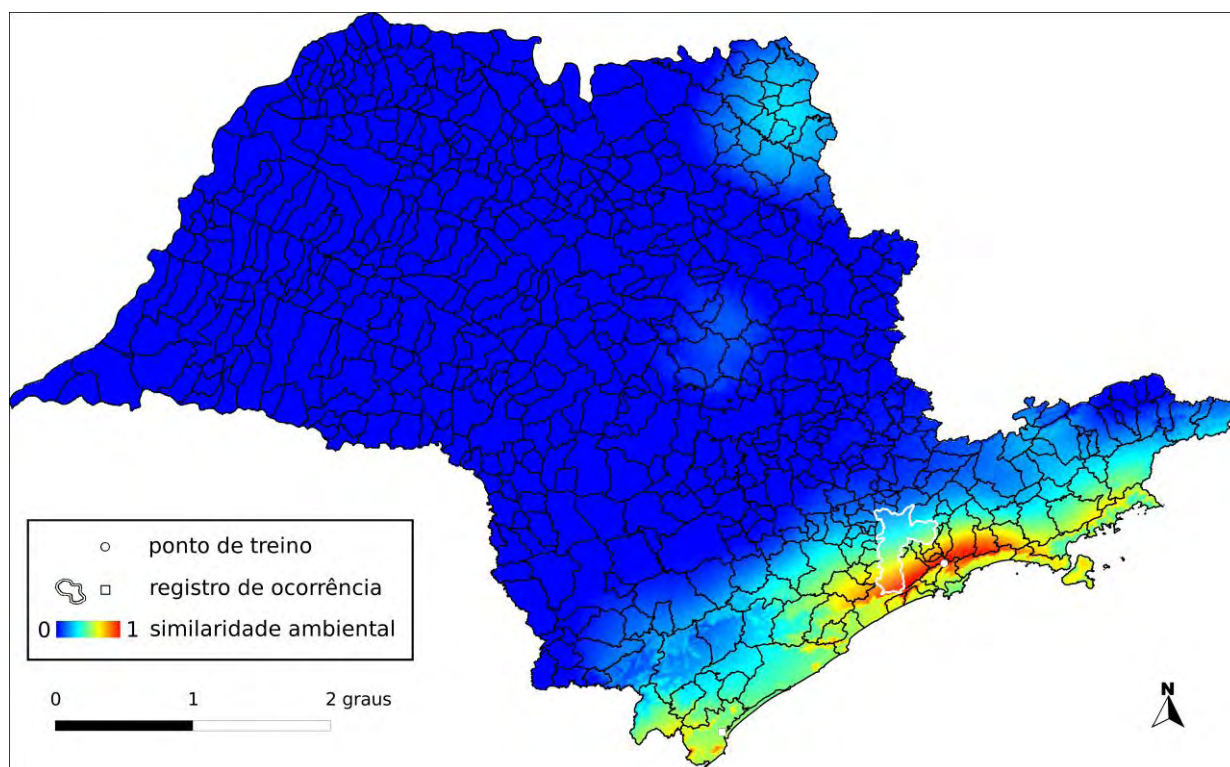
Nº de pontos usados na modelagem: 0

Ano da última coleta em SP: sem informação

Municípios com registros de ocorrência em SP: sem informação

UCs: sem informação

Passiflora marginata Mast.



Algoritmo: Distância Euclideana

Total de registros disponíveis na rede speciesLink: 11

Nº de pontos usados na modelagem: 1 (1 em SP)

Ano da última coleta em SP: 1984 (SP199400)

Municípios com registros de ocorrência em SP: Santo André (SP199400), São Paulo (IAC33791)

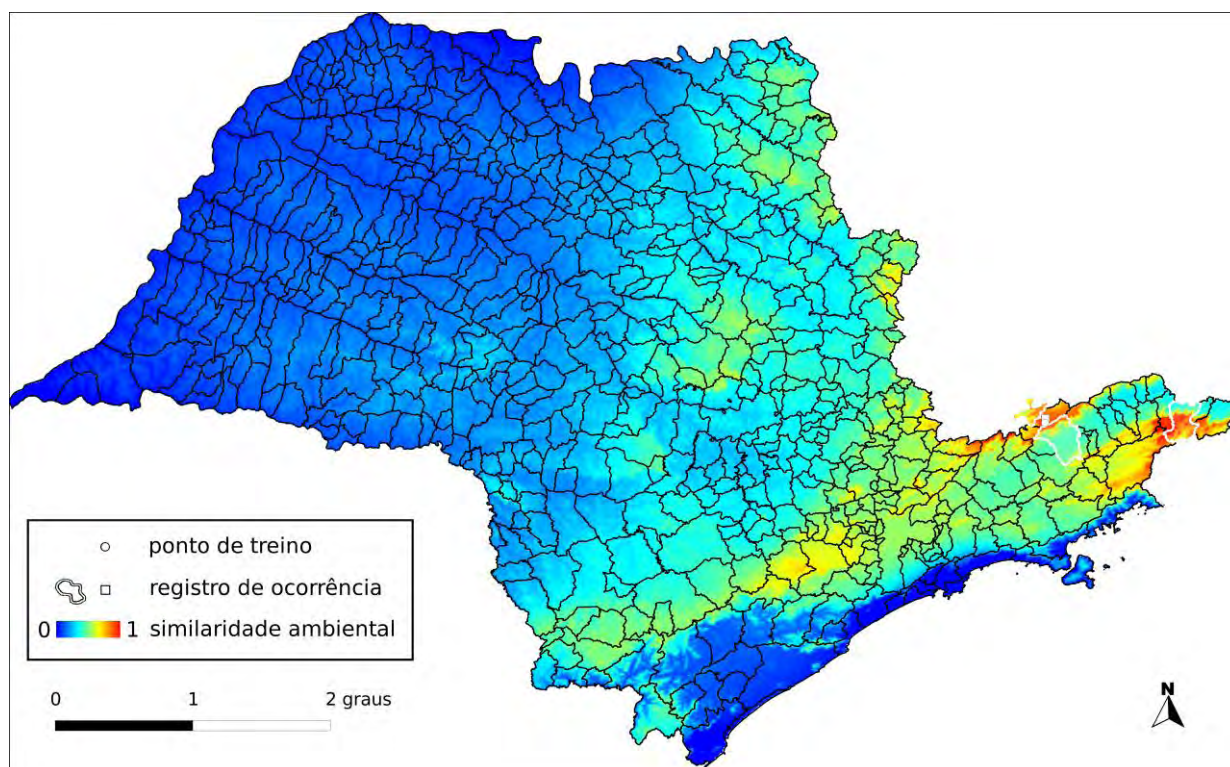
UCs²⁰: P.E. Alberto Löfgren (●), P.E. da Cantareira (●), P.E. do Jaraguá (●), P.E. Guarapiranga (●), P.E. da Serra do Mar (●)

Dados ambientais:

bio02 var. temp. dia (°C*10)	76	bio17 prec. tri. mais seco (mm)	302
bio05 temp. max. (°C*10)	242	bio18 prec. tri. mais quente (mm)	937
bio06 temp min. (°C*10)	98	bio19 prec. tri. mais frio (mm)	302
bio16 prec. tri. mais úmido (mm)	937	alt altitude (m)	669

²⁰ Registro de ocorrência inteiramente (●) ou parcialmente (●) contido em UC.

Passiflora mendoncae Harms



Algoritmo: Distância Euclideana

Total de registros disponíveis na rede *speciesLink*: 69

Nº de pontos usados na modelagem: 4 (1 em SP)

Ano da última coleta em SP: 2007 (SPSF37595)

Municípios com registros de ocorrência em SP: Campos do Jordão (SPSF37595), Pindamonhangaba (IAC42197), São Bento do Sapucaí (IAC49327), São José do Barreiro (IAC42743)

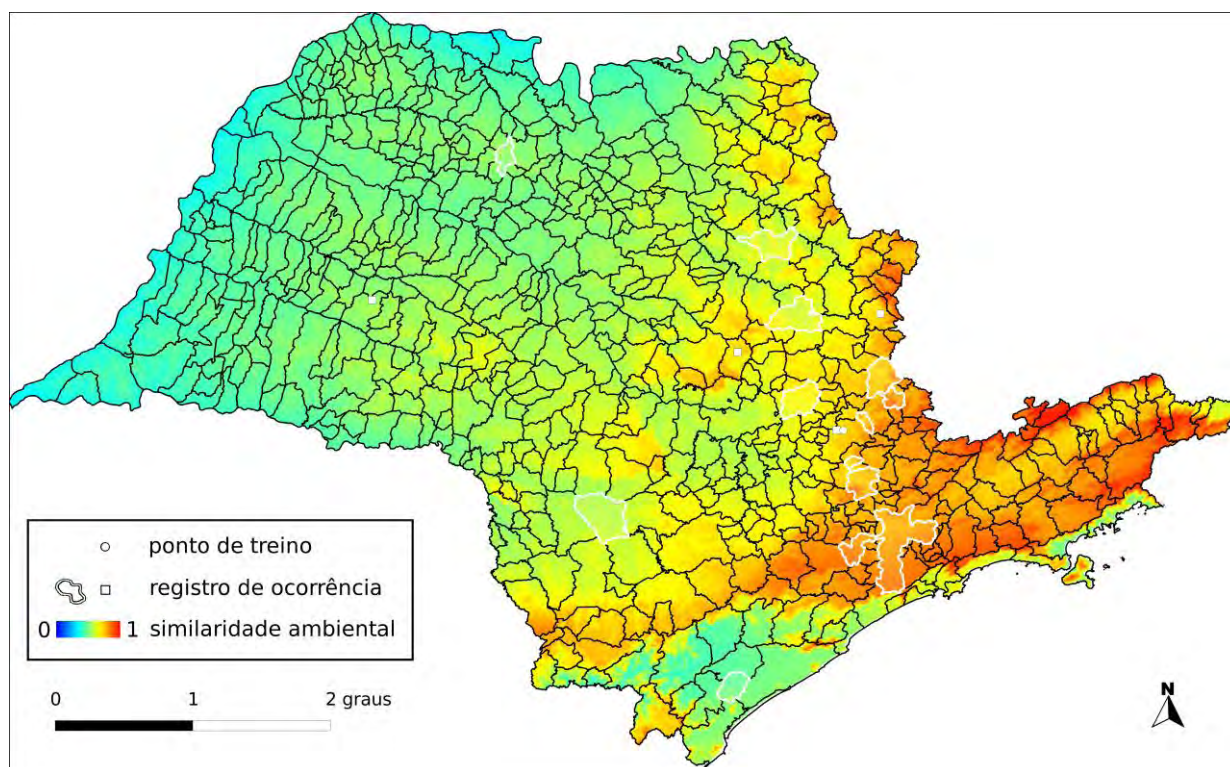
UCs²¹: P.E. de Campos do Jordão (●), P.N. da Serra da Bocaina (●), P.E. dos Mananciais de Campos do Jordão (●)

Dados ambientais:

bio02 var. temp. dia (°C*10)	88 – 112	bio17 prec. tri. mais seco (mm)	116 – 273
bio05 temp. max. (°C*10)	203 – 226	bio18 prec. tri. mais quente (mm)	663 – 795
bio06 temp min. (°C*10)	28 – 52	bio19 prec. tri. mais frio (mm)	116 – 296
bio16 prec. tri. mais úmido (mm)	669 – 860	alt altitude (m)	1387 – 1840

²¹ Registro de ocorrência inteiramente (●) ou parcialmente (◐) contido em UC.

Passiflora miersii Mast.



Algoritmo: Maxent

Total de registros disponíveis na rede *speciesLink*: 103

Omissão interna (BR): 0,00 (n=6) **Omissão interna (SP):** 0,00 (n=1)

Omissão externa (BR): 0,07 (n=44) **Omissão externa (SP):** 0,00 (n=13)

Significância (BR): $p < 0,01$

Área do modelo (SP): 248.000 km² (99,9% da área do Estado)

Área em remanescentes (2008): 43.412 km² (17,5% da área do modelo)

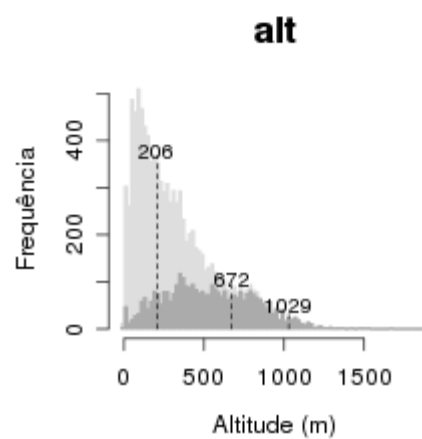
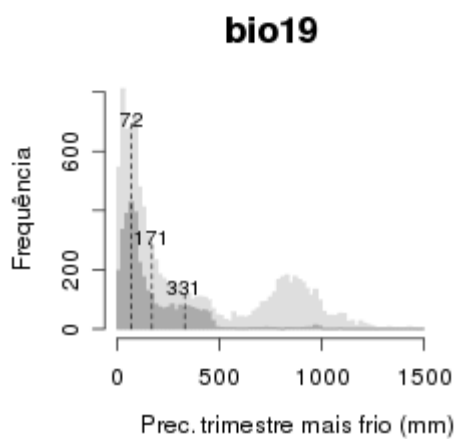
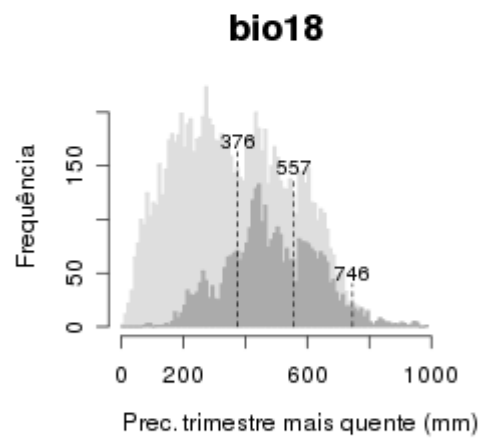
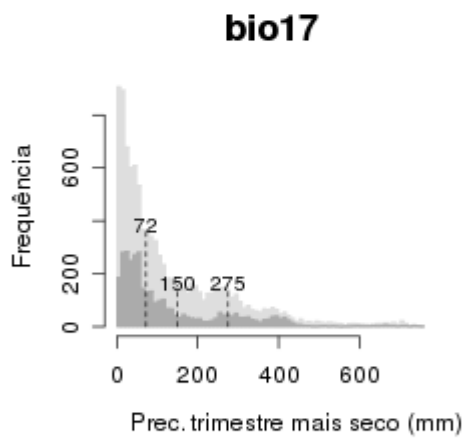
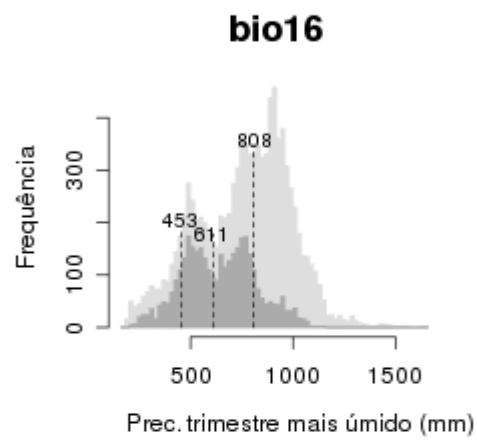
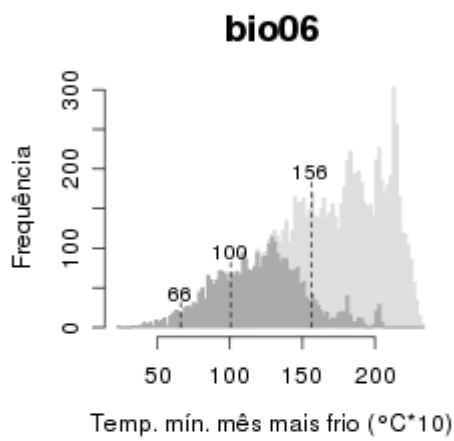
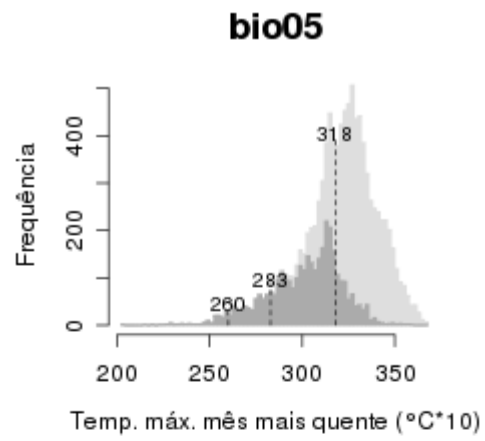
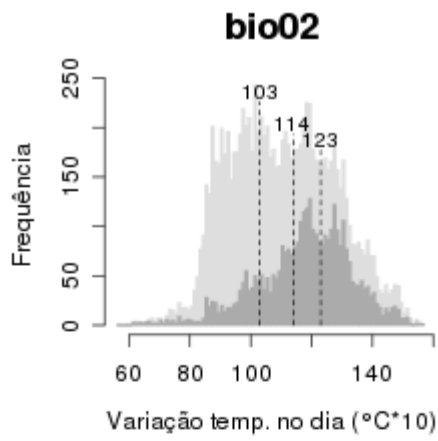
Área em remanescentes em UCs (2008): 7.920 km² (3,2% da área do modelo)

Ano da última coleta em SP: 2009 (IAC52119)

Municípios com registros de ocorrência em SP: Campinas (IAC46386), Cotia (IAC32635), Itapira (IAC52119), Jundiaí (MBM250672), Limeira (IAC32642), Mirassol (IAC35876), Paranapanema (SPSF39777), Parquera-Açu (IAC19071), Pedreira (IAC36162), Pirassununga (SPSF00762), Serra Negra (IAC32638), São Paulo (IAC32639), São Simão (IAC32643), Vinhedo (SPSF29874)

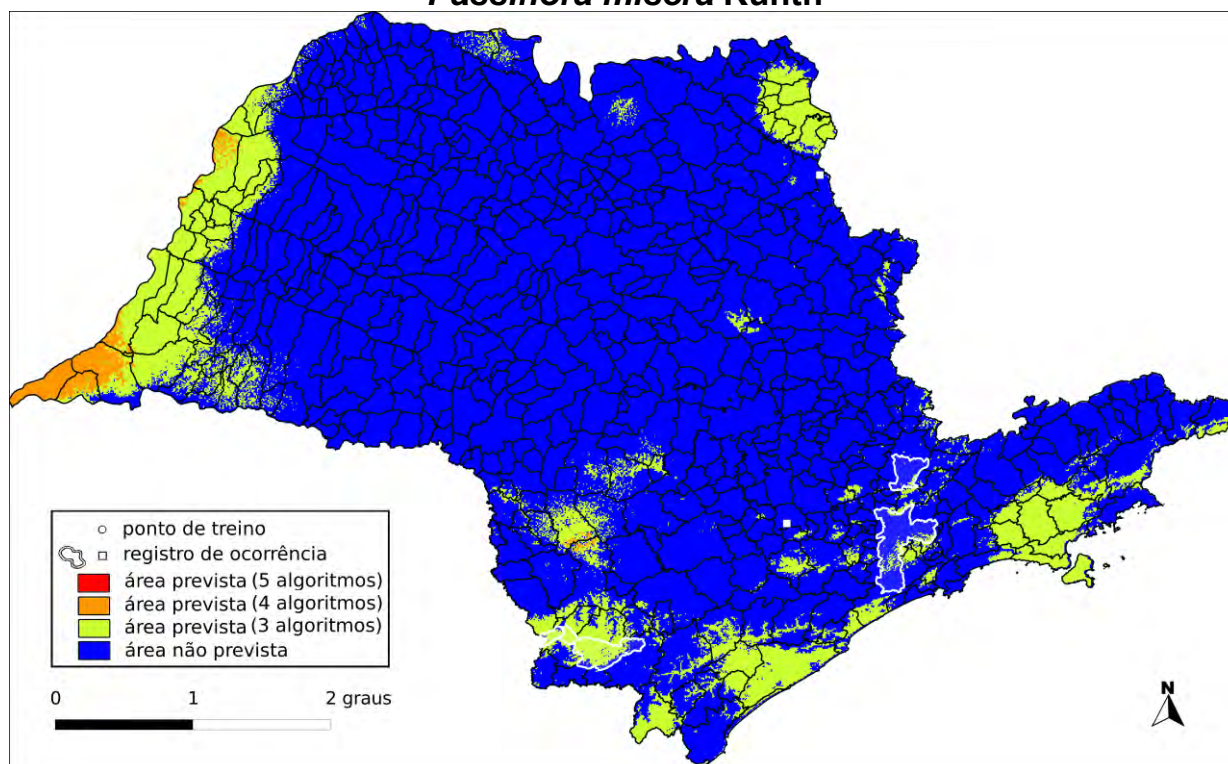
UCs²²: P.E. do Jaraguá ●, P.E. Guarapiranga ●, P.E. Alberto Löfgren ●, P.E. da Campina do Encantado ●, P.E. da Cantareira ●, P.E. da Serra do Mar ●, E.E. de Santa Maria ●, E.E. de Paranapanema ●

²² Registro de ocorrência inteiramente (●) ou parcialmente (◐) contido em UC.



Histogramas: Brasil (claro) e projeção do modelo no Brasil (escuro). Valores: mínimo, médio e máximo dos pontos de ocorrência.

Passiflora misera Kunth



Algoritmo: Consenso (ENFA, GARP BS, Maxent, MD e SVM)

Total de registros disponíveis na rede *speciesLink*: 353

Omissão interna (BR): 0,00 (n=25) **Omissão interna (SP):** – (n=0)

Omissão externa (BR): 0,23 (n=100) **Omissão externa (SP):** 0,00 (n=4)

Significância (BR): $p=0,18$

Área do modelo (SP): 36.183 km² (14,6% da área do Estado)

Área em remanescentes (2008): 10.214 km² (28,2% da área do modelo)

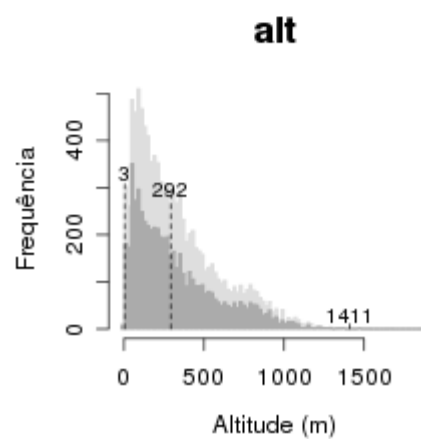
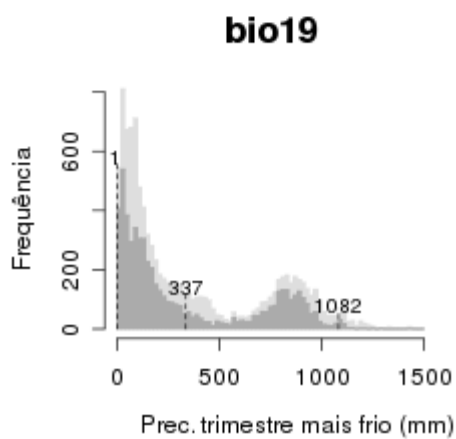
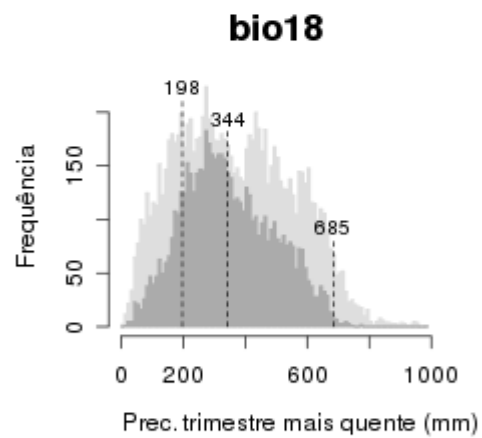
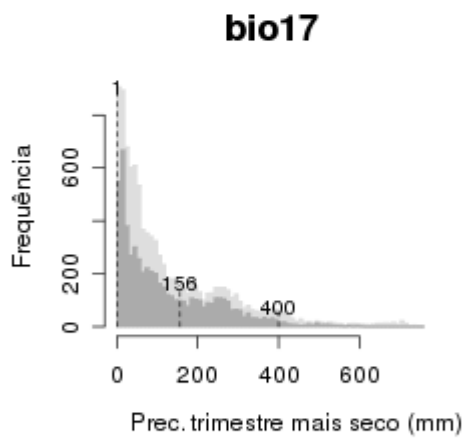
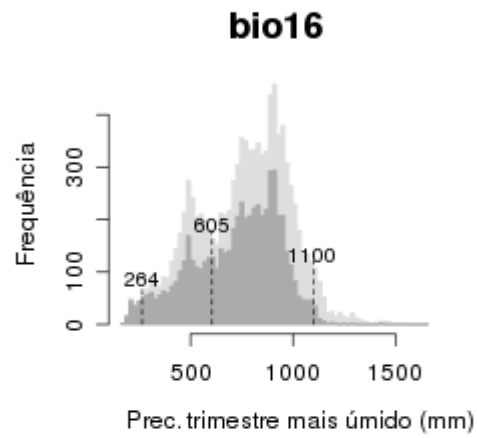
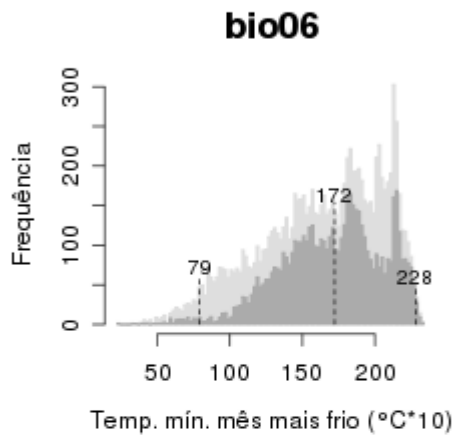
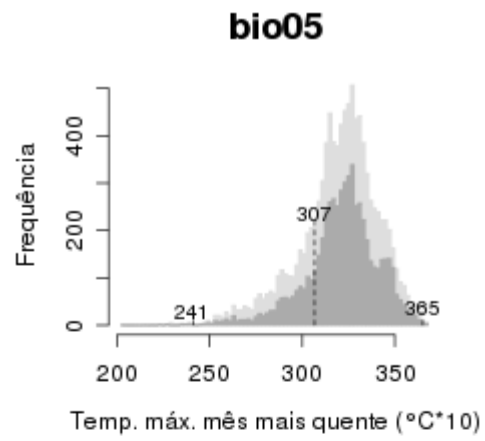
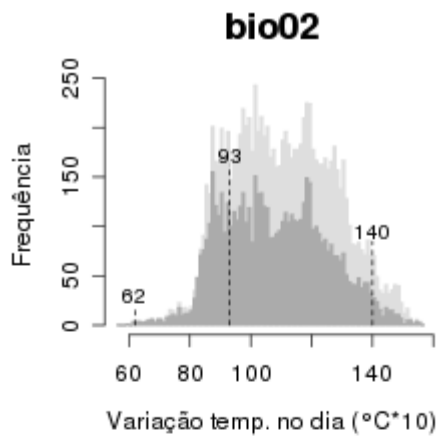
Área em remanescentes em UCs (2008): 2.738 km² (7,6% da área do modelo)

Ano da última coleta em SP: 1997 (IAC36728)

Municípios com registros de ocorrência em SP: Apiaí (ESA37723), Atibaia (MO100238528), Bom Sucesso de Itararé (IAC36728), São Paulo (IAC32628)

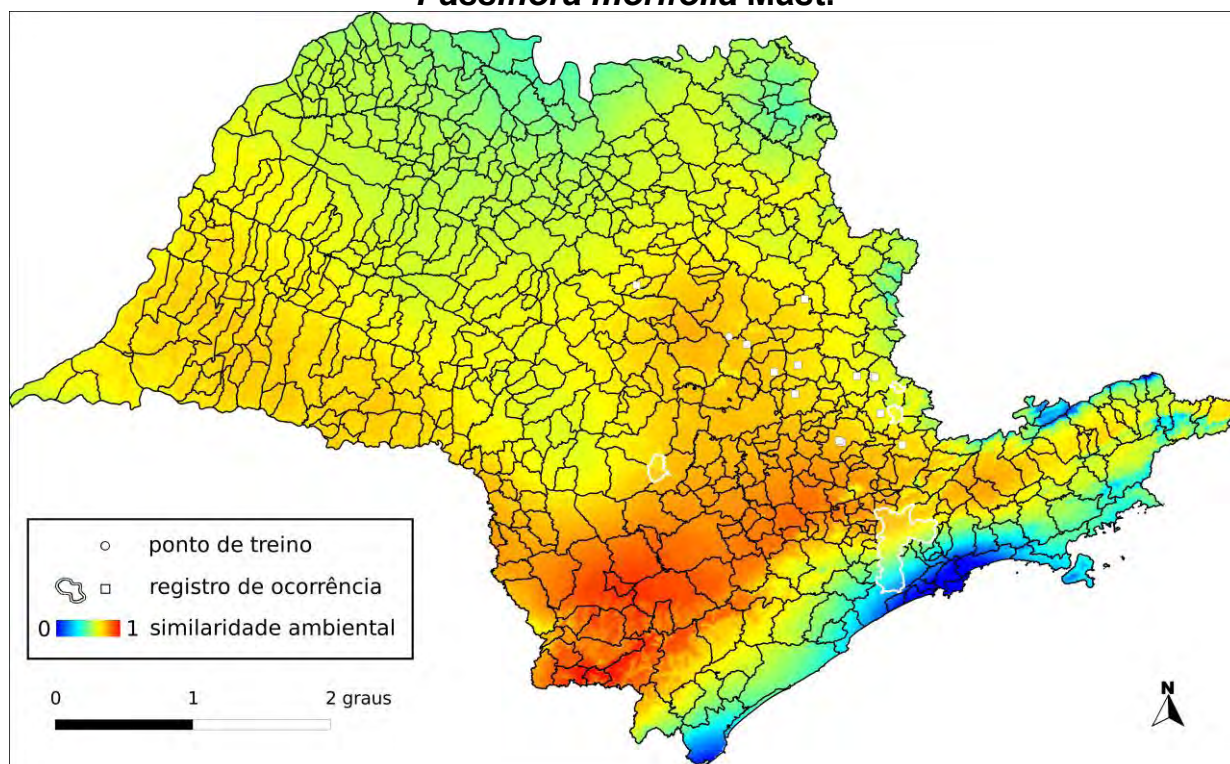
UCs²³: P.E. do Jaraguá ●, P.E. Guarapiranga ●, P.E. Alberto Löfgren ●, P.E. de Intervales ●, P.E. da Cantareira ●, P.E. da Serra do Mar ●, P.E. do Alto do Ribeira ●

²³ Registro de ocorrência inteiramente (●) ou parcialmente (◐) contido em UC.



Histogramas: Brasil (claro) e projeção do modelo no Brasil (escuro). Valores: mínimo, médio e máximo dos pontos de ocorrência.

Passiflora morifolia Mast.



Algoritmo: Distância Euclideana

Total de registros disponíveis na rede speciesLink: 99

Nº de pontos usados na modelagem: 4 (1 em SP)

Ano da última coleta em SP: 2007 (IAC49461)

Municípios com registros de ocorrência em SP: Lindóia (IAC29960), Monte Alegre do Sul (IAC39538), Pardinho (IAC49461), São Carlos (IAC32623), São Paulo (IAC32616)

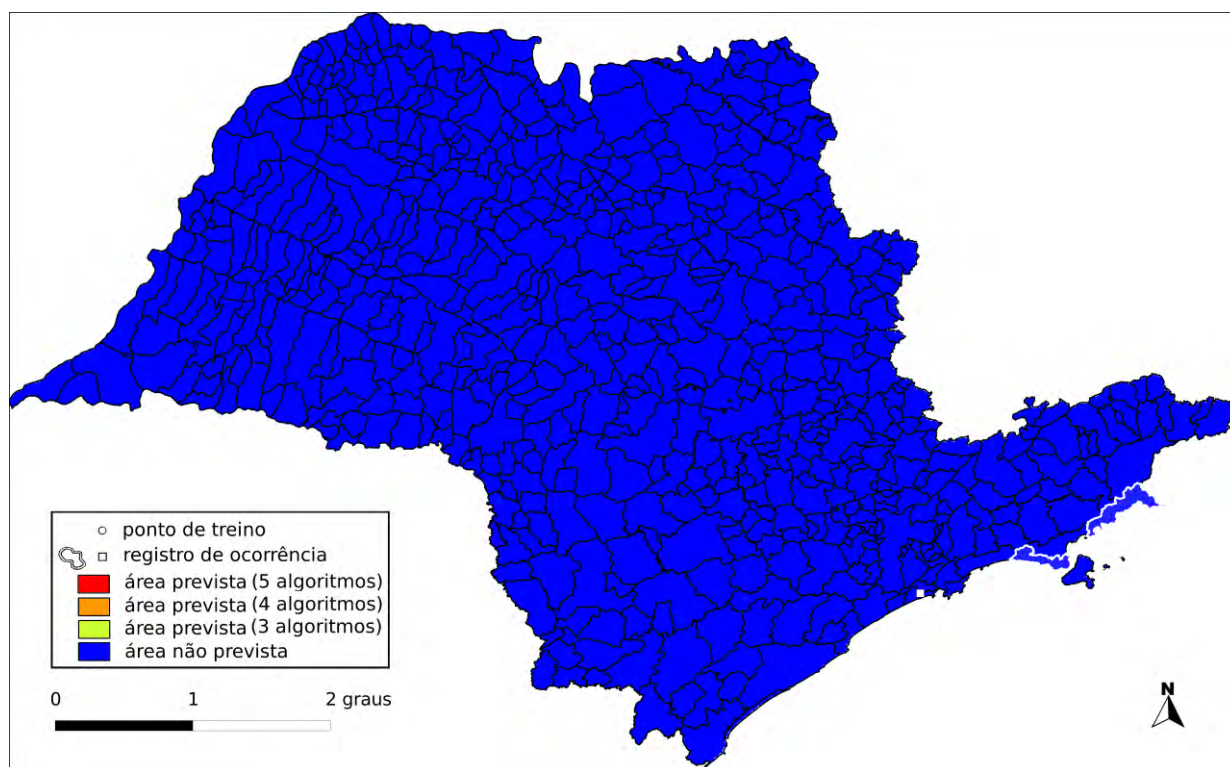
UCs²⁴: P.E. Alberto Löfgren ●, P.E. da Cantareira ●, P.E. do Jaraguá ●, P.E. Guarapiranga ●, P.E. da Serra do Mar ●

Dados ambientais:

bio02 var. temp. dia (°C*10)	86 – 124	bio17 prec. tri. mais seco (mm)	85 – 284
bio05 temp. max. (°C*10)	274 – 294	bio18 prec. tri. mais quente (mm)	470 – 608
bio06 temp min. (°C*10)	92 – 105	bio19 prec. tri. mais frio (mm)	108 – 324
bio16 prec. tri. mais úmido (mm)	504 – 675	alt altitude (m)	232 – 749

²⁴ Registro de ocorrência inteiramente (●) ou parcialmente (◐) contido em UC.

Passiflora mucronata Lam.



Algoritmo: Consenso (ENFA, GARP BS, Maxent, MD e SVM)

Total de registros disponíveis na rede *speciesLink*: 421

Omissão interna (BR): 0, (n=23)

Omissão interna (SP): – (n=0)

Omissão externa (BR): 0, (n=62)

Omissão externa (SP): 1,00 (n=2)

Significância (BR): $p < 0,01$

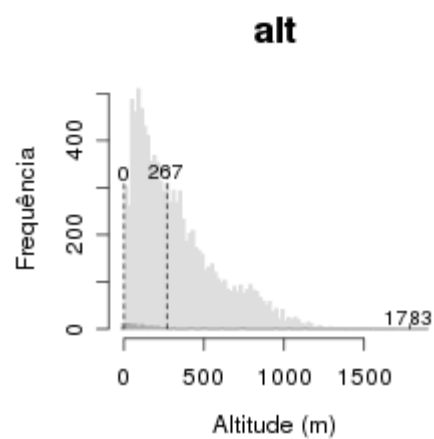
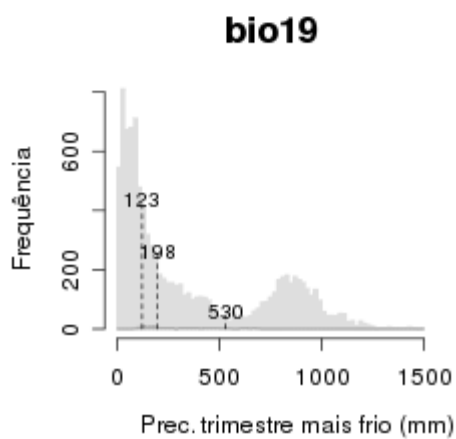
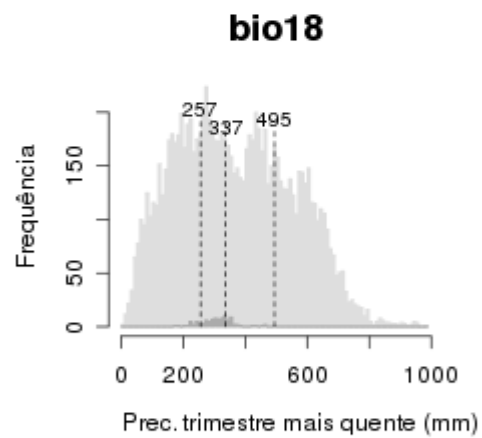
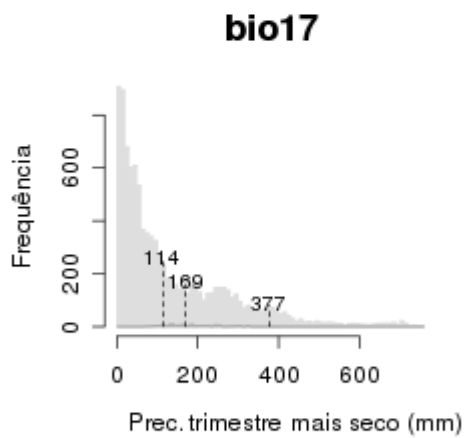
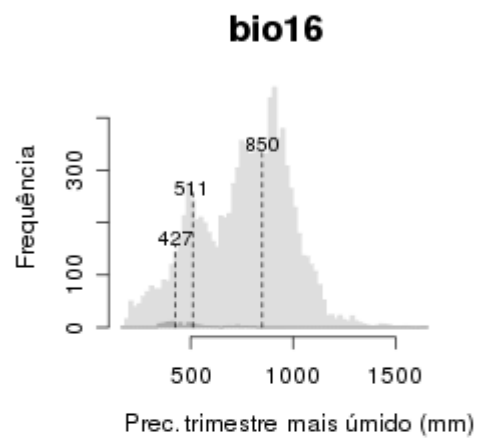
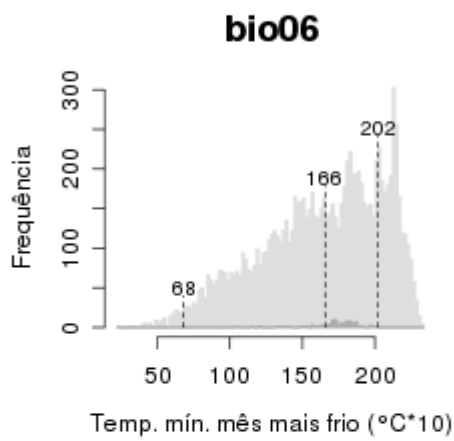
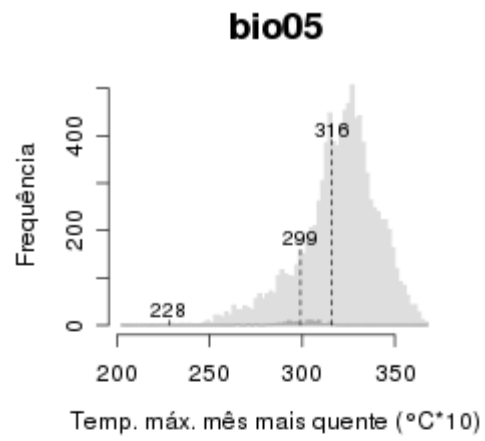
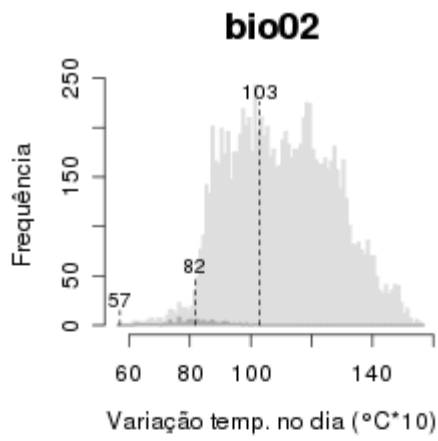
Área do modelo (SP): 0 km² (0% da área do Estado)

Ano da última coleta em SP: 2005 (IAC45532)

Municípios com registros de ocorrência em SP: São Sebastião (IAC36159), Ubatuba (RB325560)

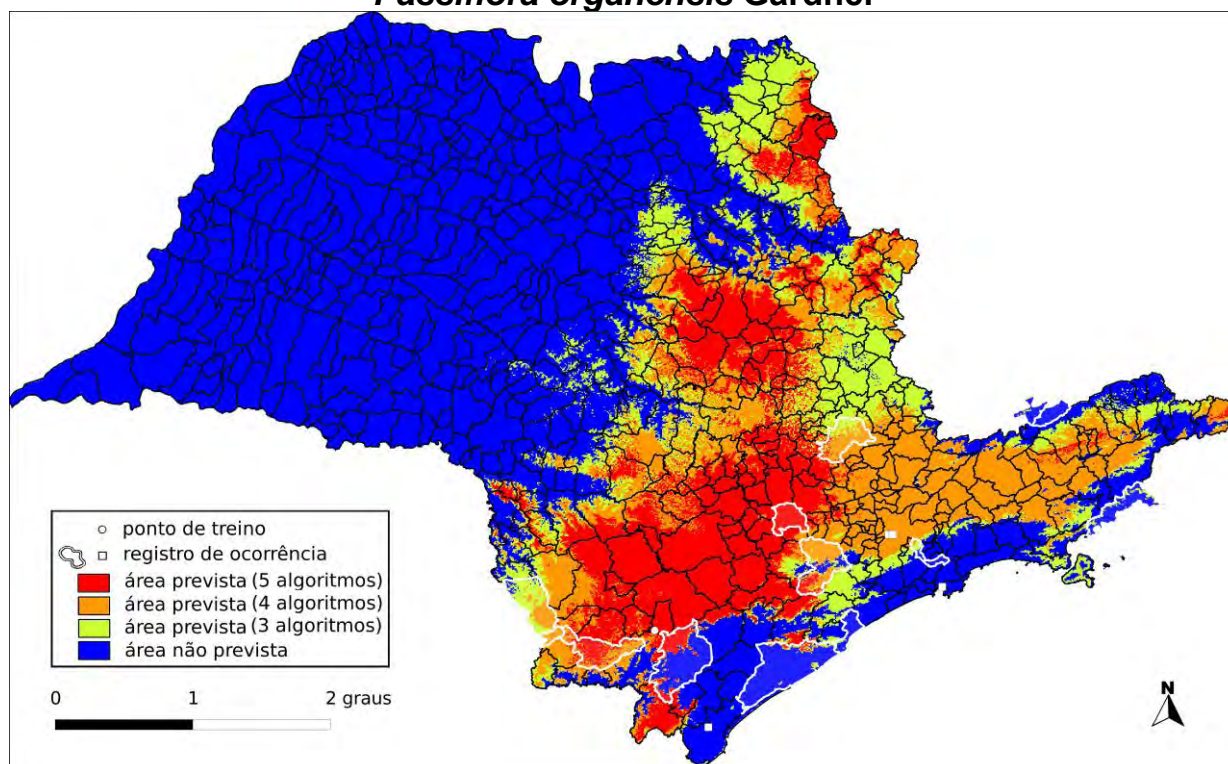
UCs²⁵: P.E. da Ilha Anchieta (●), P.E. da Serra do Mar (●), P.N. da Serra da Bocaina (●), E.E. de Tupinambás (●)

²⁵ Registro de ocorrência inteiramente (●) ou parcialmente (◐) contido em UC.



Histogramas: Brasil (claro) e projeção do modelo no Brasil (escuro). Valores: mínimo, médio e máximo dos pontos de ocorrência.

Passiflora organensis Gardner



Algoritmo: Consenso (ENFA, GARP BS, Maxent, MD e SVM)

Total de registros disponíveis na rede speciesLink: 278

Omissão interna (BR): 0,00 (n=21)

Omissão interna (SP): 0,00 (n=1)

Omissão externa (BR): 0,22 (n=91)

Omissão externa (SP): 0,08 (n=13)

Significância (BR): $p < 0,01$

Área do modelo (SP): 99.957 km² (40,3% da área do Estado)

Área em remanescentes (2008): 20.879 km² (20,9% da área do modelo)

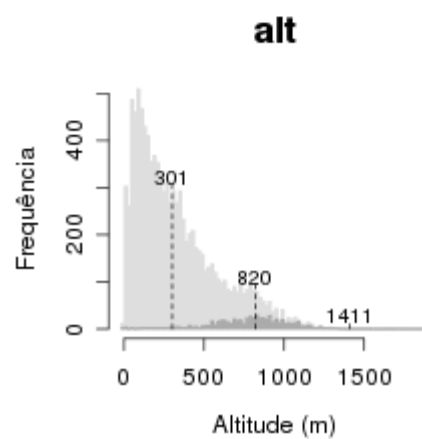
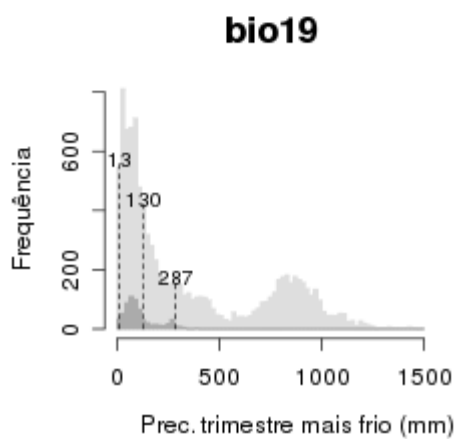
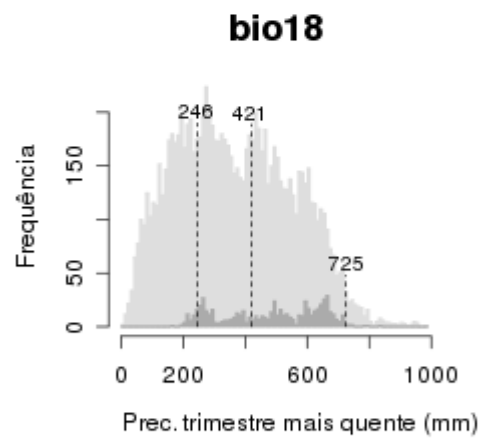
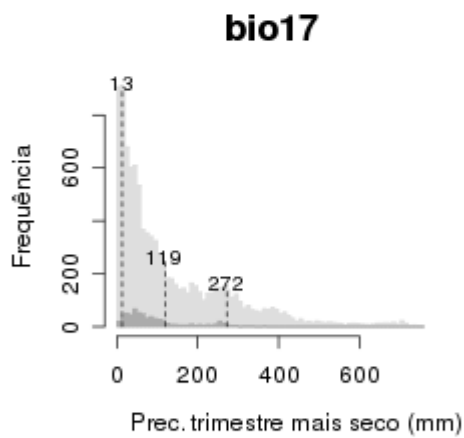
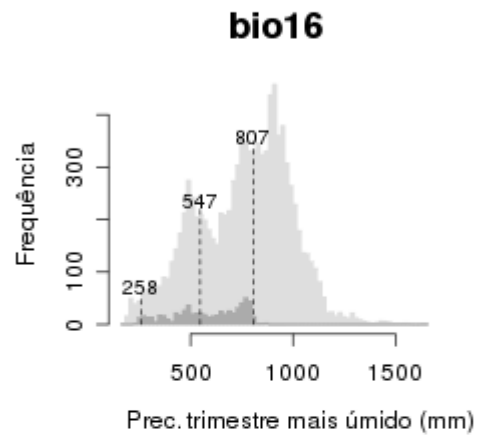
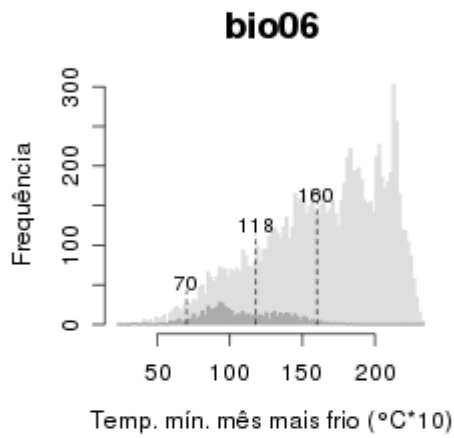
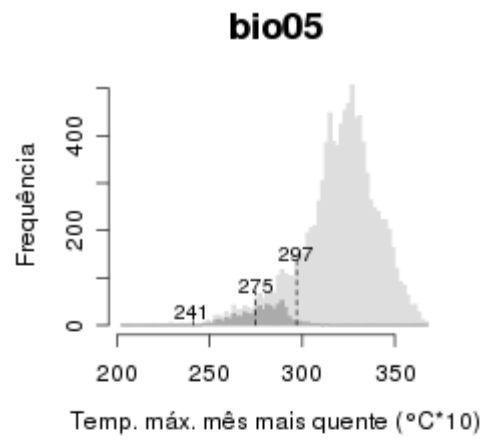
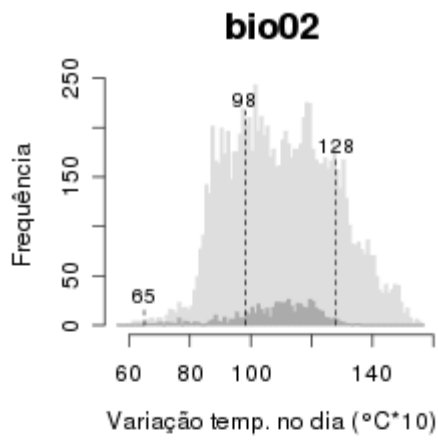
Área em remanescentes em UCs (2008): 3.750 km² (3,8% da área do modelo)

Ano da última coleta em SP: 2004 (IAC44061)

Municípios com registros de ocorrência em SP: Apiaí (IAC36421), Bom Sucesso de Itararé (IAC44061), Campinas (IAC2851), Campos de Jordão (**RB32585**), Eldorado (ESA92634), Ibiúna (IAC33808), Iguape (SP238315), Itararé (IAC33895), Peruíbe (SPSF25502), Ribeirão Grande (**ESA84924**), Santo André (IAC33804), Sorocaba (IAC19731), São Bento do Sapucaí (MO100246510), Ubatuba (IAC33811)

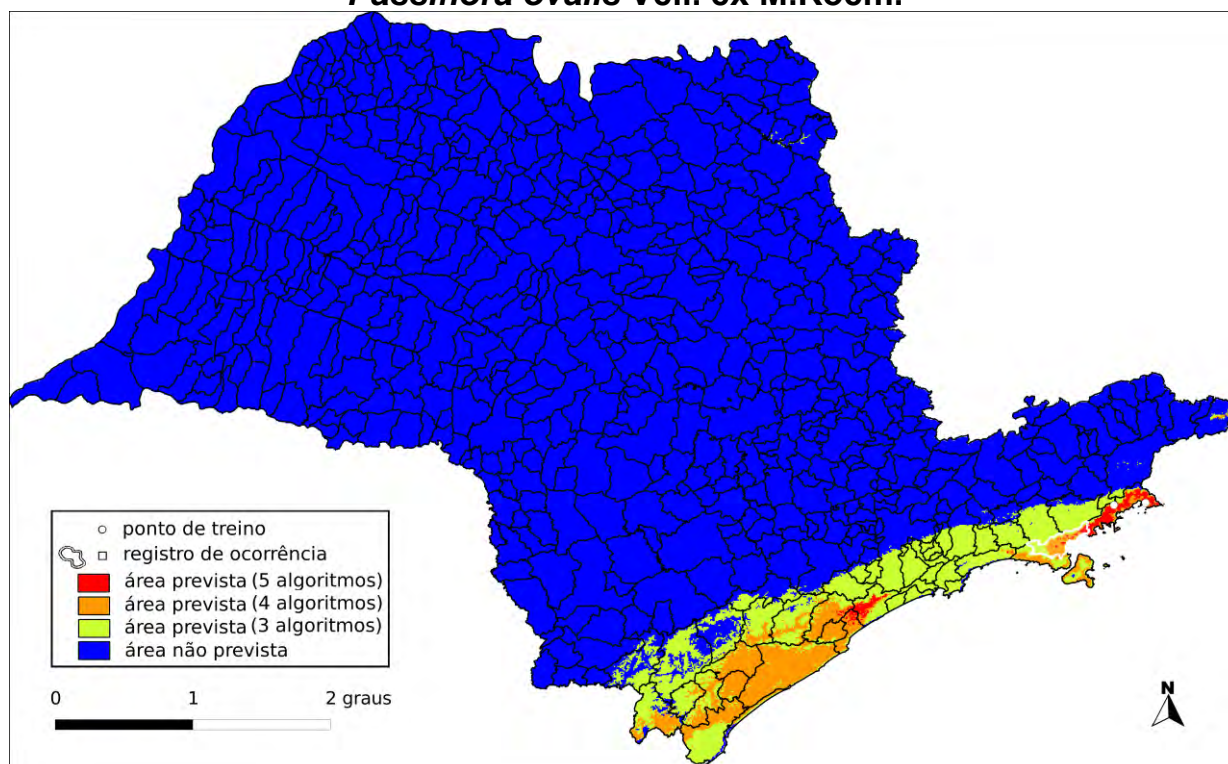
UCs²⁶: P.E. de Campos do Jordão (●), P.E. da Ara (●), P.E. da Serra do Mar (●), E.E. de Xitué (●), P.N. da Serra da Bocaina (●), P.E. de Intervalos (●), P.E. do Alto do Ribeira (●), E.E. Chaúas (●), E.E. de Tupinambás (●), P.E. do Prelado (●), E.E. Juréia-Itatins (●), P.E. da Campina do Encantado (●), P.E. da Ilha Anchieta (●), E.E. de Valinhos (●), P.E. do Jurupará (●), P.E. dos Mananciais de Campos do Jordão (●), P.E. de Jacupiranga (●), P.E. do Itinguçu (●)

²⁶ Registro de ocorrência inteiramente (●) ou parcialmente (●) contido em UC.



Histogramas: Brasil (claro) e projeção do modelo no Brasil (escuro). Valores: mínimo, médio e máximo dos pontos de ocorrência.

Passiflora ovalis Vell. ex M.Roem.



Algoritmo: Consenso (ENFA, GARP BS, Maxent, MD e SVM)

Total de registros disponíveis na rede *speciesLink*: 184

Omissão interna (BR): 0,00 (n=31) **Omissão interna (SP):** 0,00 (n=3)

Omissão externa (BR): 0,03 (n=36) **Omissão externa (SP):** 0,00 (n=1)

Significância (BR): $p < 0,01$

Área do modelo (SP): 21.173 km² (8,5% da área do Estado)

Área em remanescentes (2008): 15.981 km² (75,5% da área do modelo)

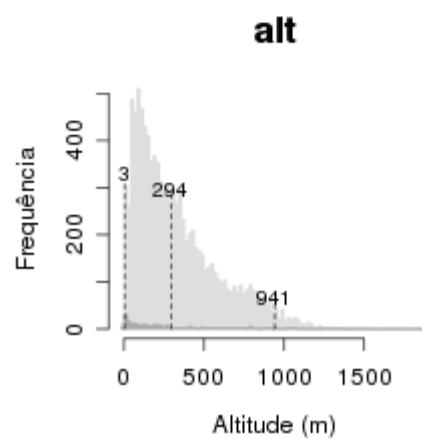
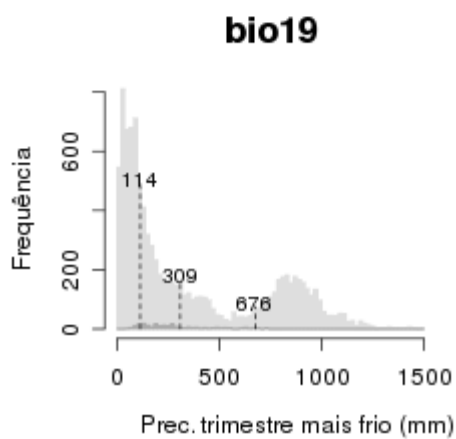
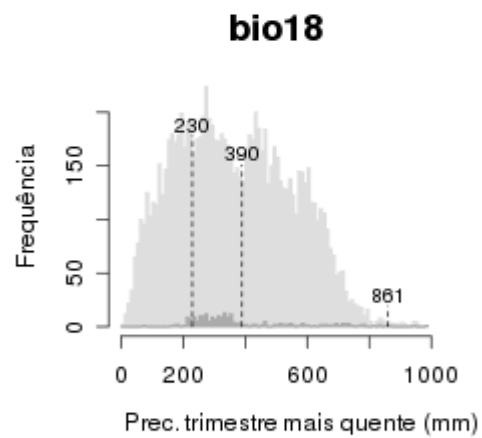
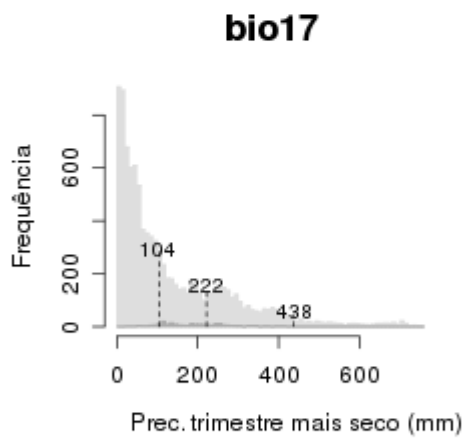
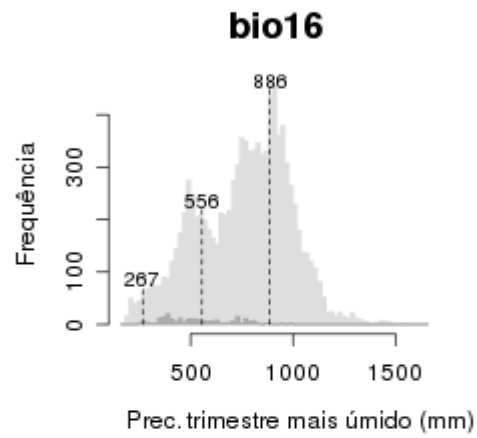
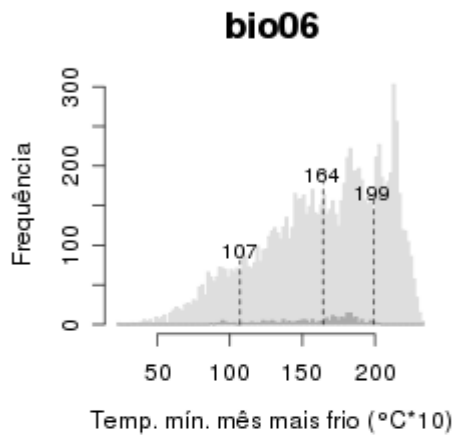
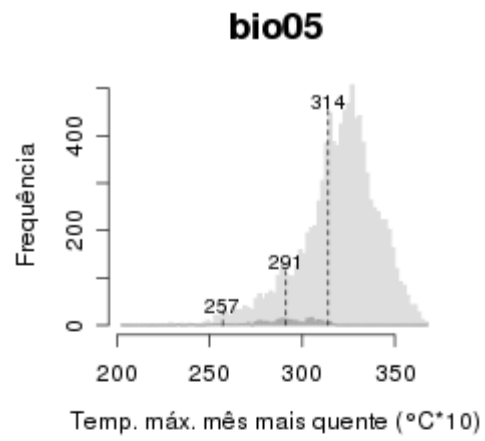
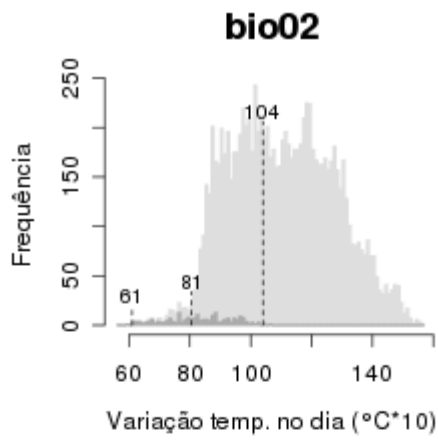
Área em remanescentes em UCs (2008): 5.686 km² (26,7% da área do modelo)

Ano da última coleta em SP: 2007 (IAC50609)

Municípios com registros de ocorrência em SP: Caraguatatuba (IAC37955), Ubatuba (IAC46738, IAC50609, HUEFS124242)

UCs²⁷: P.E. da Serra do Mar ●, P.N. da Serra da Bocaina ●

²⁷ Registro de ocorrência inteiramente (●) ou parcialmente (◐) contido em UC.



Histogramas: Brasil (claro) e projeção do modelo no Brasil (escuro). Valores: mínimo, médio e máximo dos pontos de ocorrência.

***Passiflora pardifolia* Mast.**

Total de registros disponíveis na rede *speciesLink*: 4

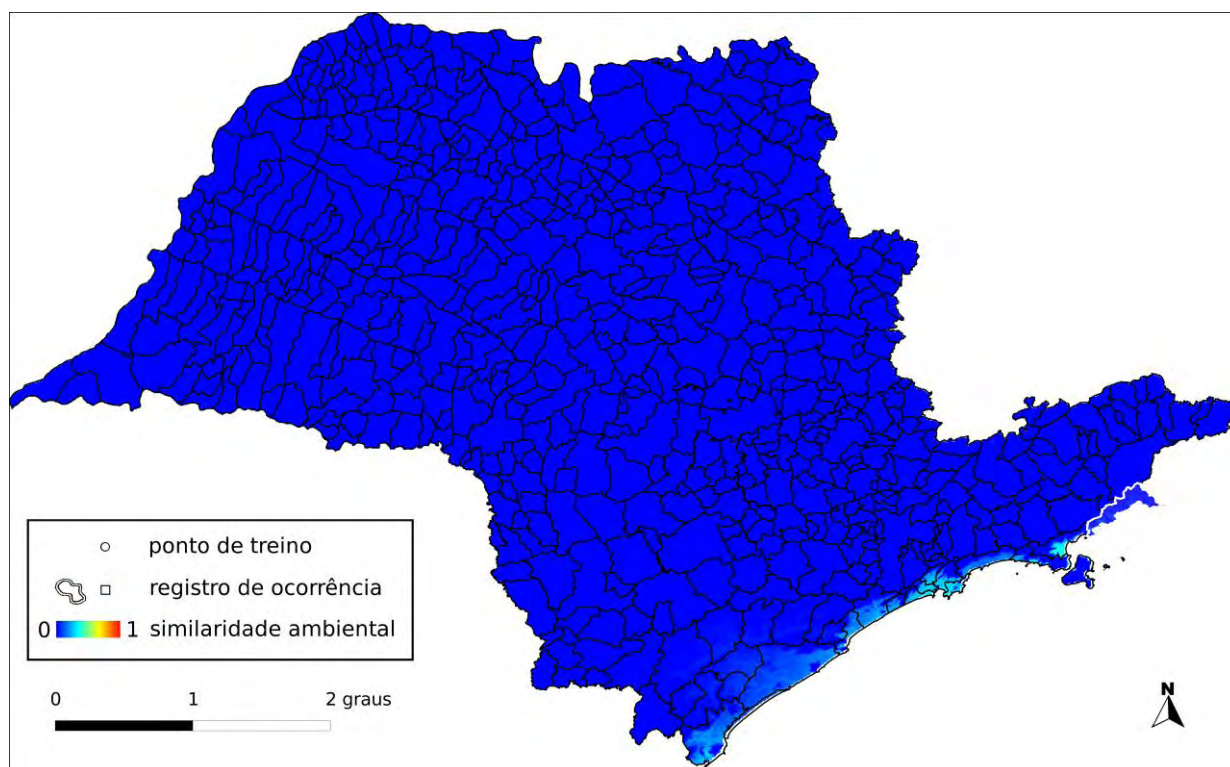
Nº de pontos usados na modelagem: 0

Ano da última coleta em SP: 2007

Municípios com registros de ocorrência em SP: Capão Bonito (UPCB60135)

UCs: sem informação

Passiflora pentagona Mast.



Algoritmo: Maxent

Total de registros disponíveis na rede speciesLink: 94

Omissão interna (BR): 0,00 (n=5) **Omissão interna (SP):** – (n=0)

Omissão externa (BR): 0,50 (n=22) **Omissão externa (SP):** 1,00 (n=1)

Significância (BR): $p < 0,01$

Área do modelo (SP): 53 km² (0,02% da área do Estado)

Área em remanescentes (2008): 18 km² (34,3% da área do modelo)

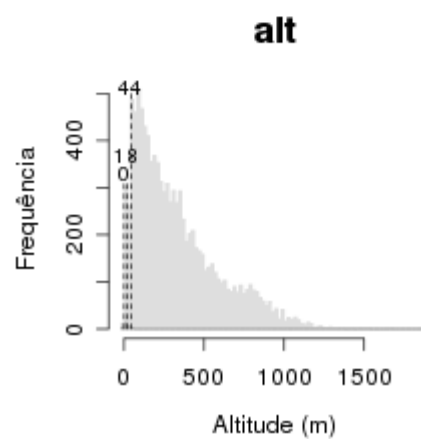
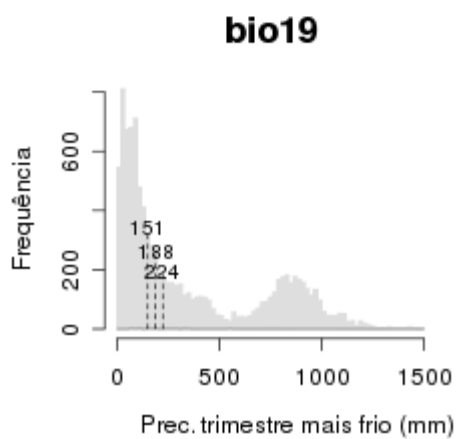
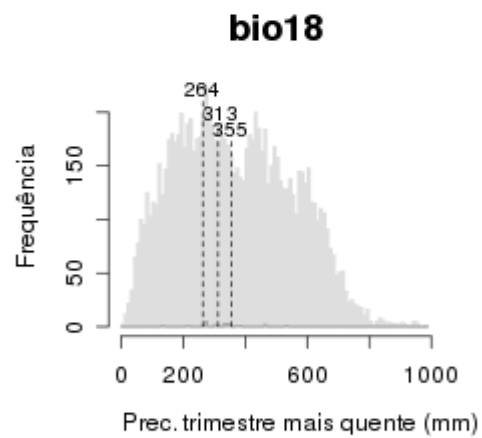
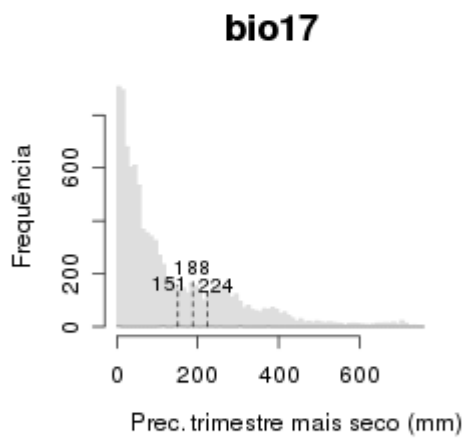
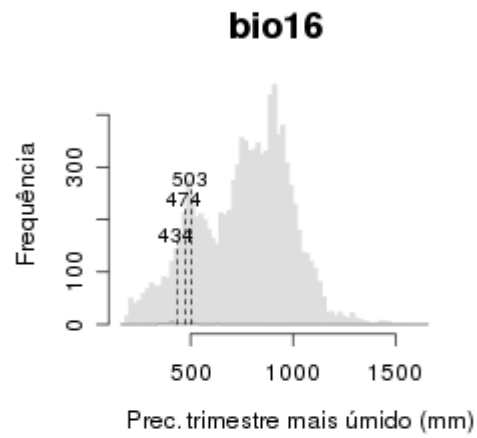
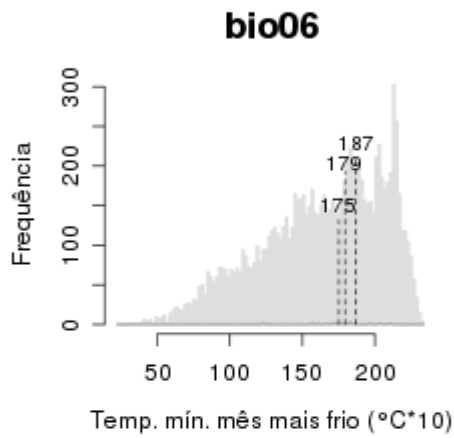
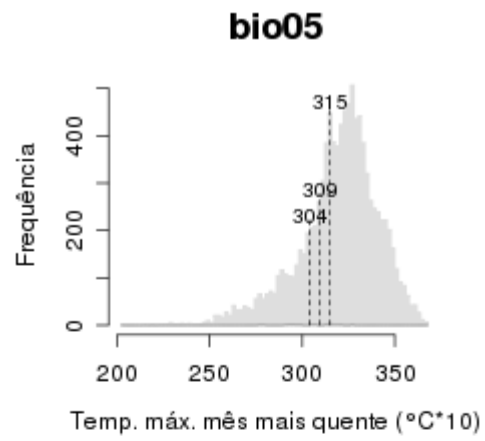
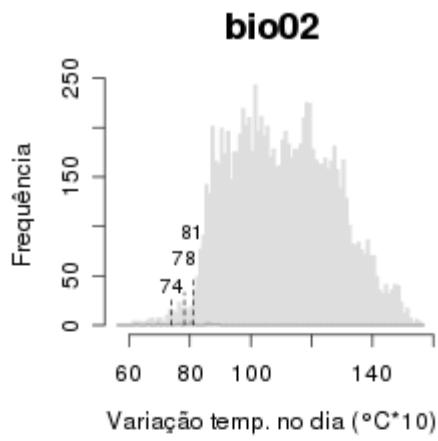
Área em remanescentes em UCs (2008): 3 km² (7,5% da área do modelo)

Ano da última coleta em SP: 2004 (IAC44352)

Municípios com registros de ocorrência em SP: Ubatuba (IAC44352)

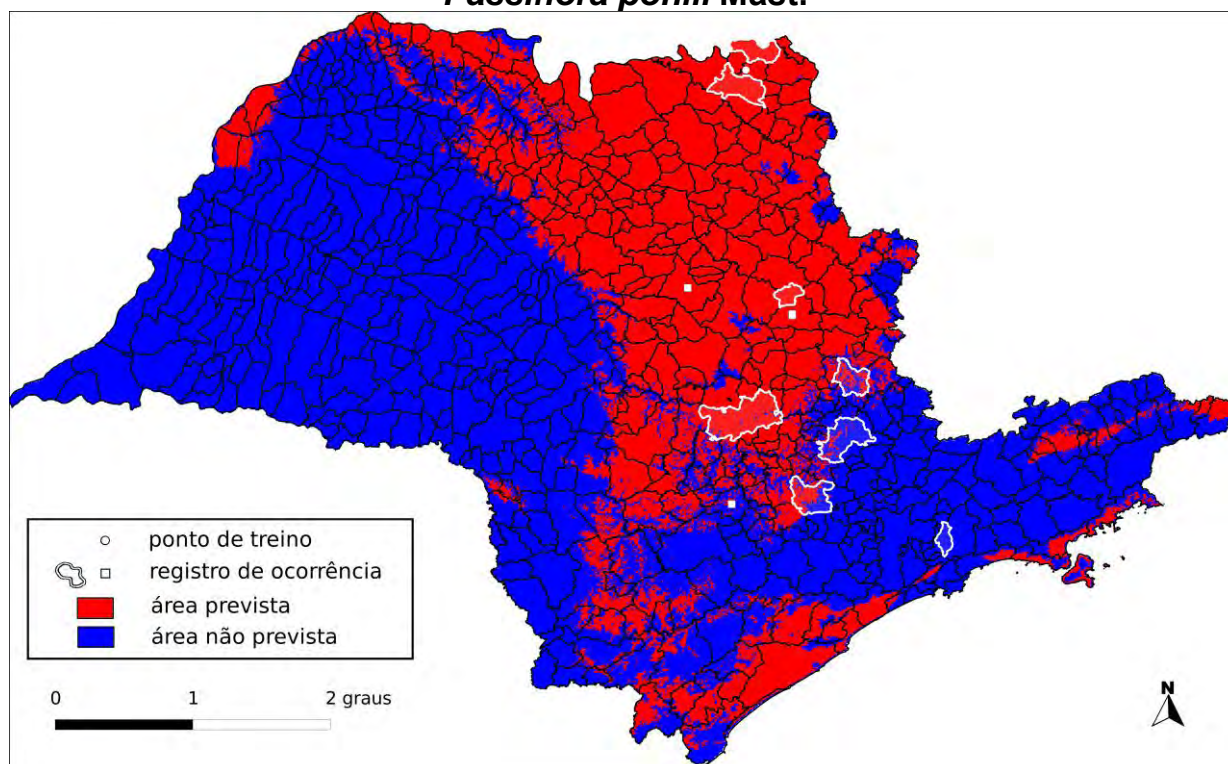
UCs²⁸: P.E. da Ilha Anchieta ●, P.E. da Serra do Mar ●, P.N. da Serra da Bocaina ●, E.E. de Tupinambás ●

²⁸ Registro de ocorrência inteiramente (●) ou parcialmente (◐) contido em UC.



Histogramas: Brasil (claro) e projeção do modelo no Brasil (escuro). Valores: mínimo, médio e máximo dos pontos de ocorrência.

Passiflora pohlii Mast.



Algoritmo: Consenso (GARP BS e Maxent)

Total de registros disponíveis na rede speciesLink: 173

Omissão interna (BR): 0,06 (n=16) **Omissão interna (SP):** 0,00 (n=1)

Omissão externa (BR): 0,24 (n=63) **Omissão externa (SP):** 0,12 (n=8)

Significância (BR): $p < 0,01$

Área do modelo (SP): 91.321 km² (36,8% da área do Estado)

Área em remanescentes (2008): 14.996 km² (16,4% da área do modelo)

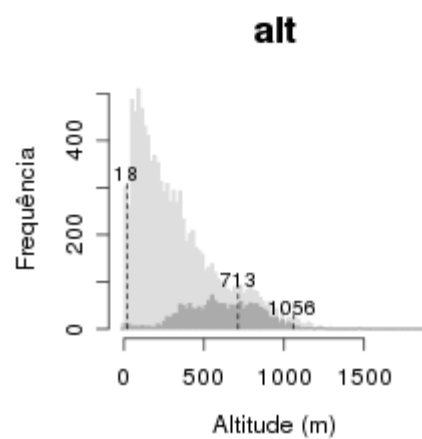
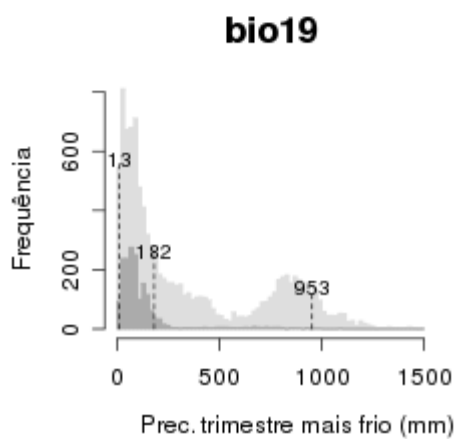
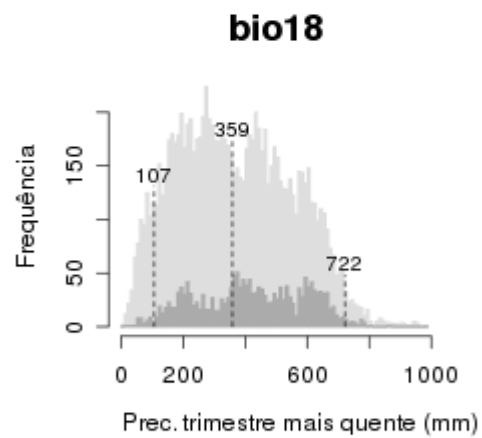
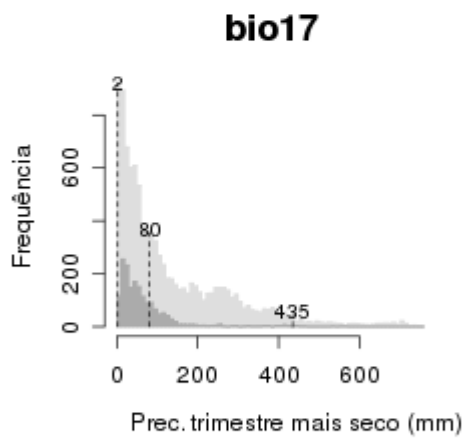
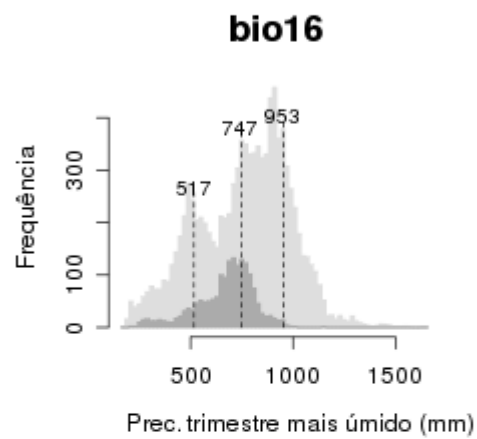
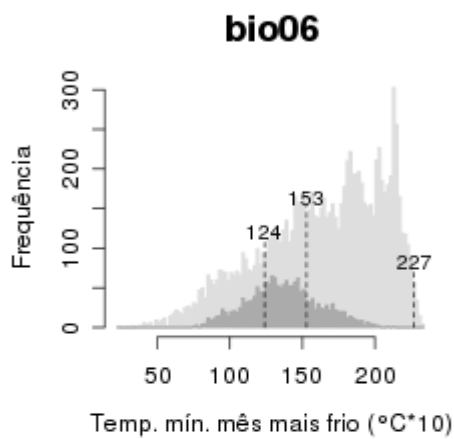
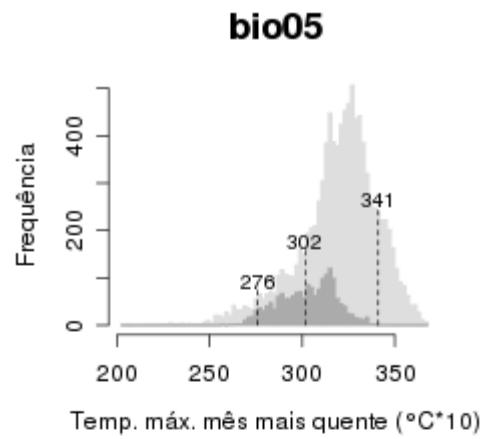
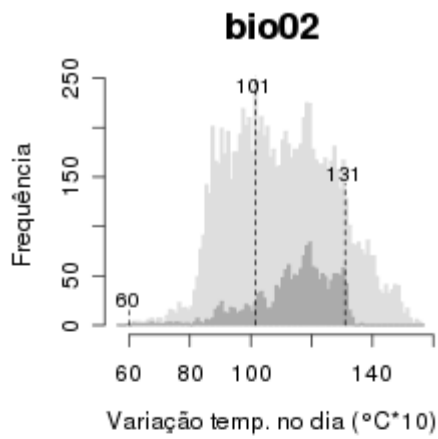
Área em remanescentes em UCs (2008): 2.620 km² (2,9% da área do modelo)

Ano da última coleta em SP: 1999 (IAC42203)

Municípios com registros de ocorrência em SP: Buritizal (IAC33818), Campinas (IAC5911), Igarapava (IAC33816), Itu (MO100238512), Ituverava (MO100254972), Mogi-Mirim (IAC42203), Piracicaba (IAC33817), Porto Ferreira (SPSF20847), Suzano (IAC33813)

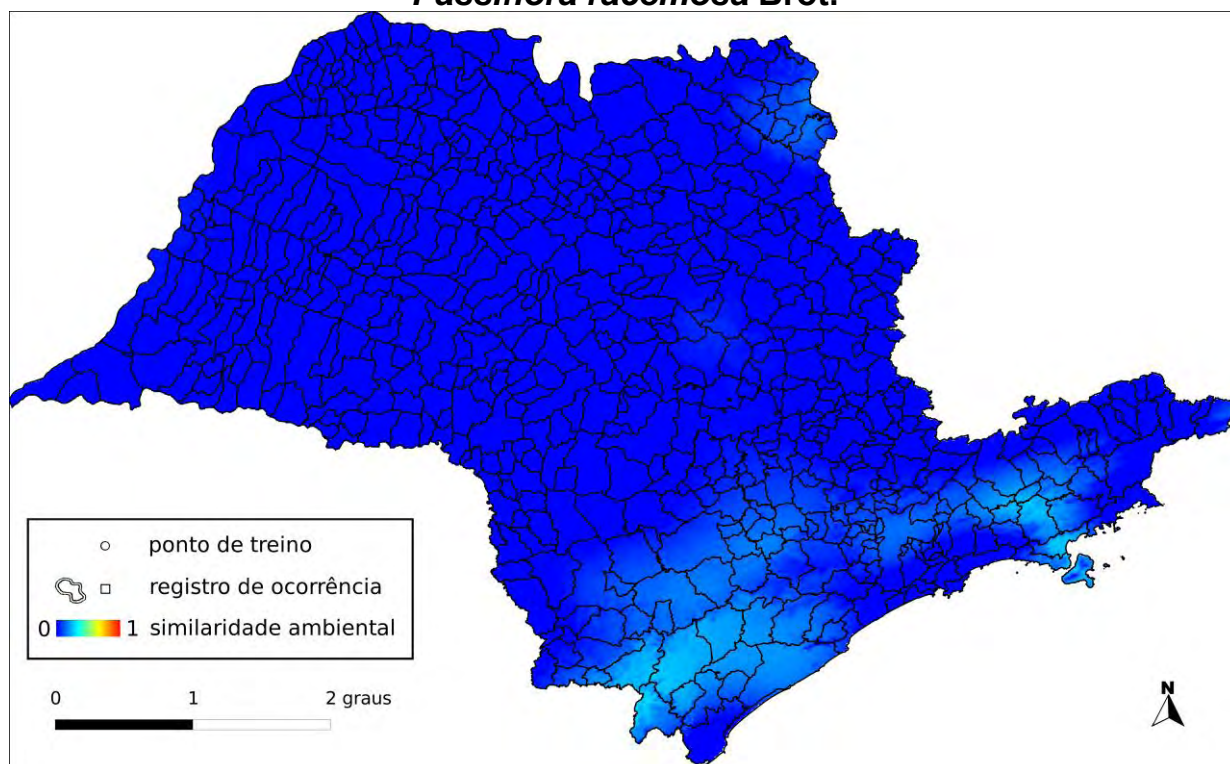
UCs²⁹: P.E. de Porto Ferreira ●, E.E. de Valinhos ●, P.E. da Ara ●, E.E. de Ibicatu ●

²⁹ Registro de ocorrência inteiramente (●) ou parcialmente (◐) contido em UC.



Histogramas: Brasil (claro) e projeção do modelo no Brasil (escuro). Valores: mínimo, médio e máximo dos pontos de ocorrência.

Passiflora racemosa Brot.



Algoritmo: Distância Euclideana

Total de registros disponíveis na rede speciesLink: 127

Nº de pontos usados na modelagem: 1 (0 em SP)

Ano da última coleta em SP: sem informação

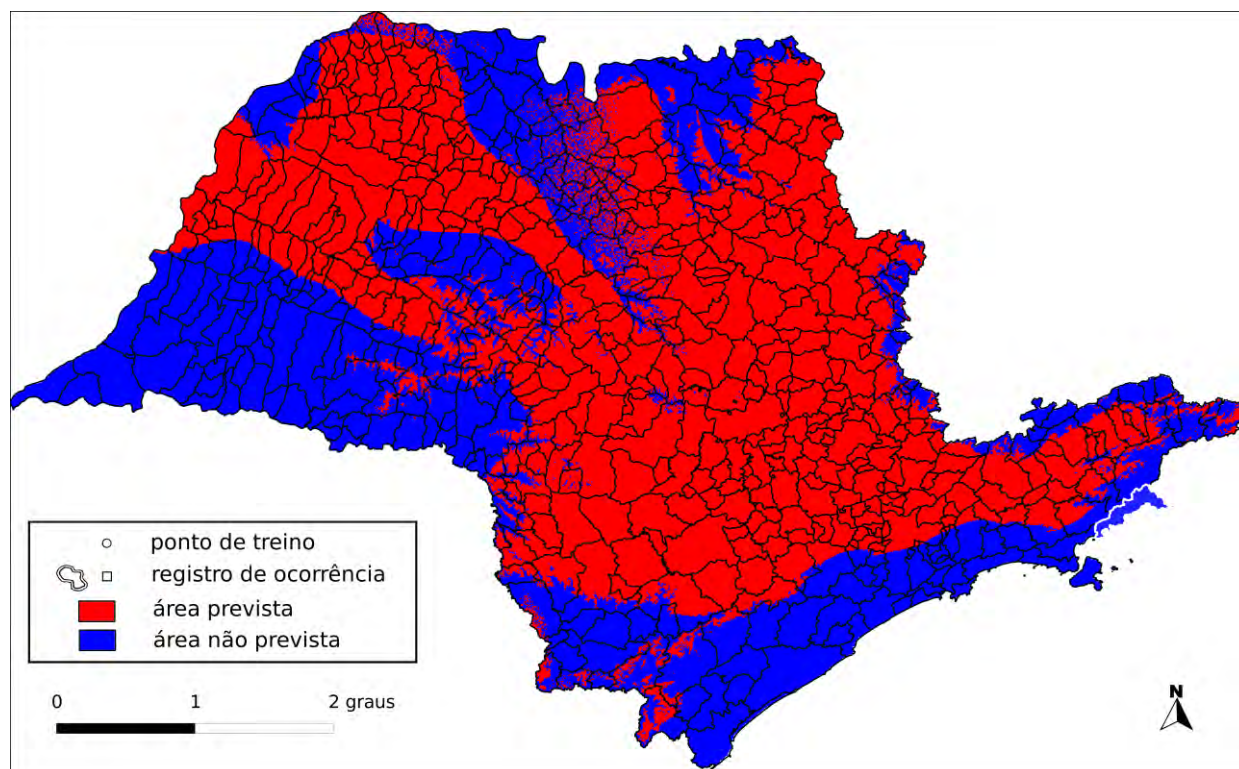
Municípios com registros de ocorrência em SP: sem informação

UCs: sem informação

Dados ambientais:

bio02 var. temp. dia (°C*10)	69	bio17 prec. tri. mais seco (mm)	118
bio05 temp. max. (°C*10)	288	bio18 prec. tri. mais quente (mm)	276
bio06 temp min. (°C*10)	167	bio19 prec. tri. mais frio (mm)	118
bio16 prec. tri. mais úmido (mm)	338	alt altitude (m)	136

Passiflora rhamnifolia Mast.



Algoritmo: Consenso (GARP BS e Maxent)

Total de registros disponíveis na rede *speciesLink*: 92

Omissão interna (BR): 0,00 (n=11) **Omissão interna (SP):** – (n=0)

Omissão externa (BR): 0,26 (n=27) **Omissão externa (SP):** 1,00 (n=1)

Significância (BR): $p < 0,01$

Área do modelo (SP): 148.712 km² (59,9% da área do Estado)

Área em remanescentes (2008): 16.947 km² (11,4% da área do modelo)

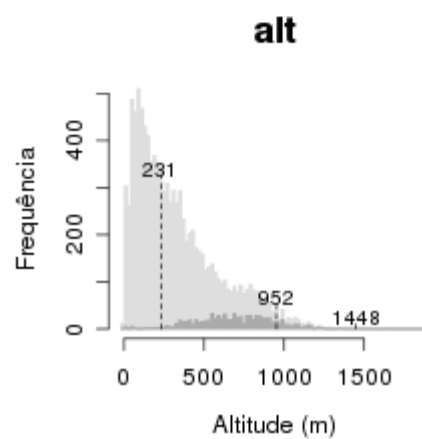
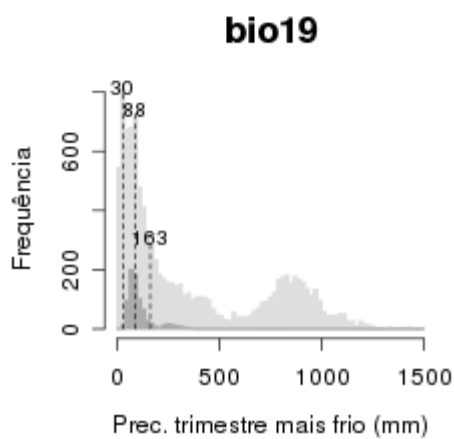
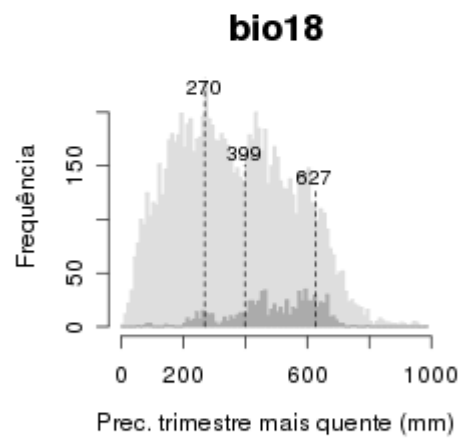
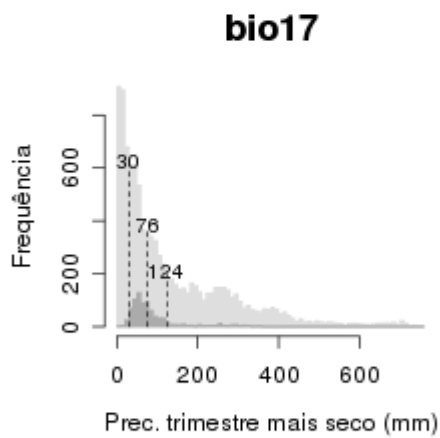
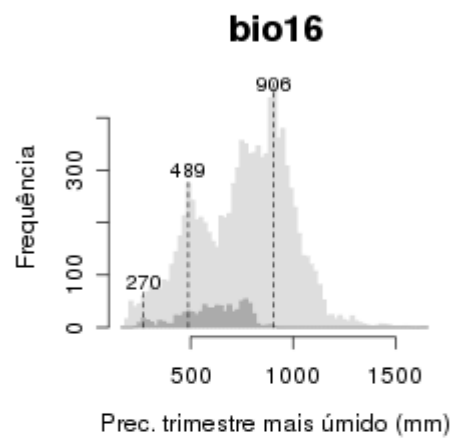
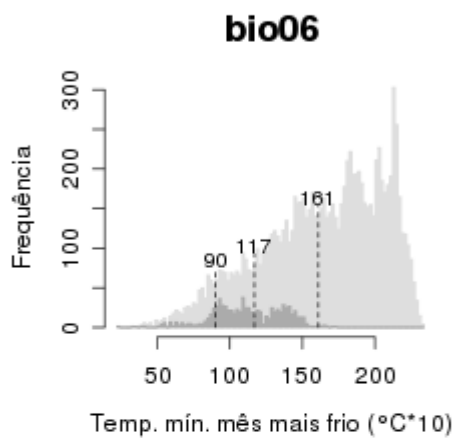
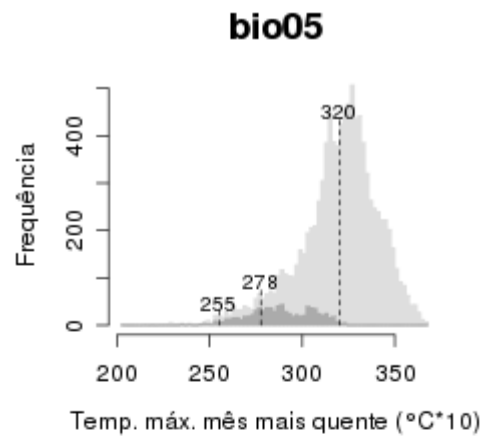
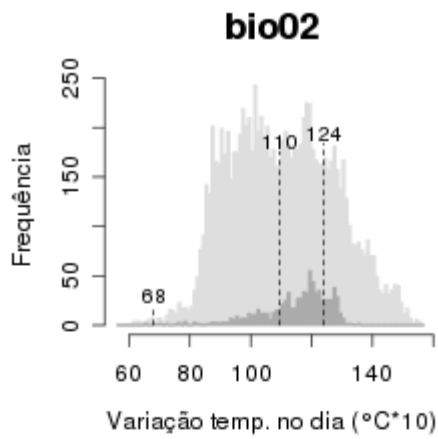
Área em remanescentes em UCs (2008): 1.293 km² (0,87% da área do modelo)

Ano da última coleta em SP: 2003 (IAC44029)

Municípios com registros de ocorrência em SP: Ubatuba (IAC44029)

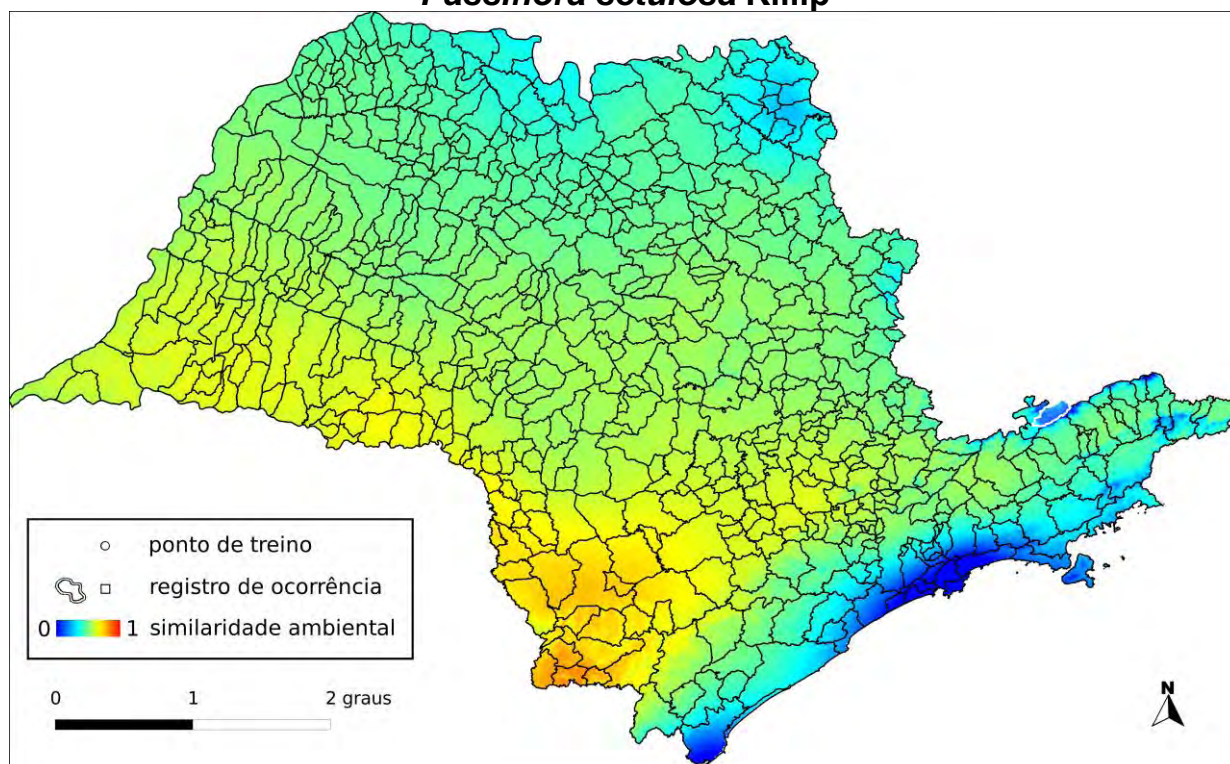
UCs³⁰: P.E. da Ilha Anchieta ●, P.E. da Serra do Mar ●, P.N. da Serra da Bocaina ●, E.E. de Tupinambás ●

³⁰ Registro de ocorrência inteiramente (●) ou parcialmente (◐) contido em UC.



Histogramas: Brasil (claro) e projeção do modelo no Brasil (escuro). Valores: mínimo, médio e máximo dos pontos de ocorrência.

Passiflora setulosa Killip



Algoritmo: Distância Euclideana

Total de registros disponíveis na rede speciesLink: 20

Nº de pontos usados na modelagem: 1 (0 em SP)

Ano da última coleta em SP: 1935 (IAC32574)

Municípios com registros de ocorrência em SP: Campos do Jordão (IAC32574)

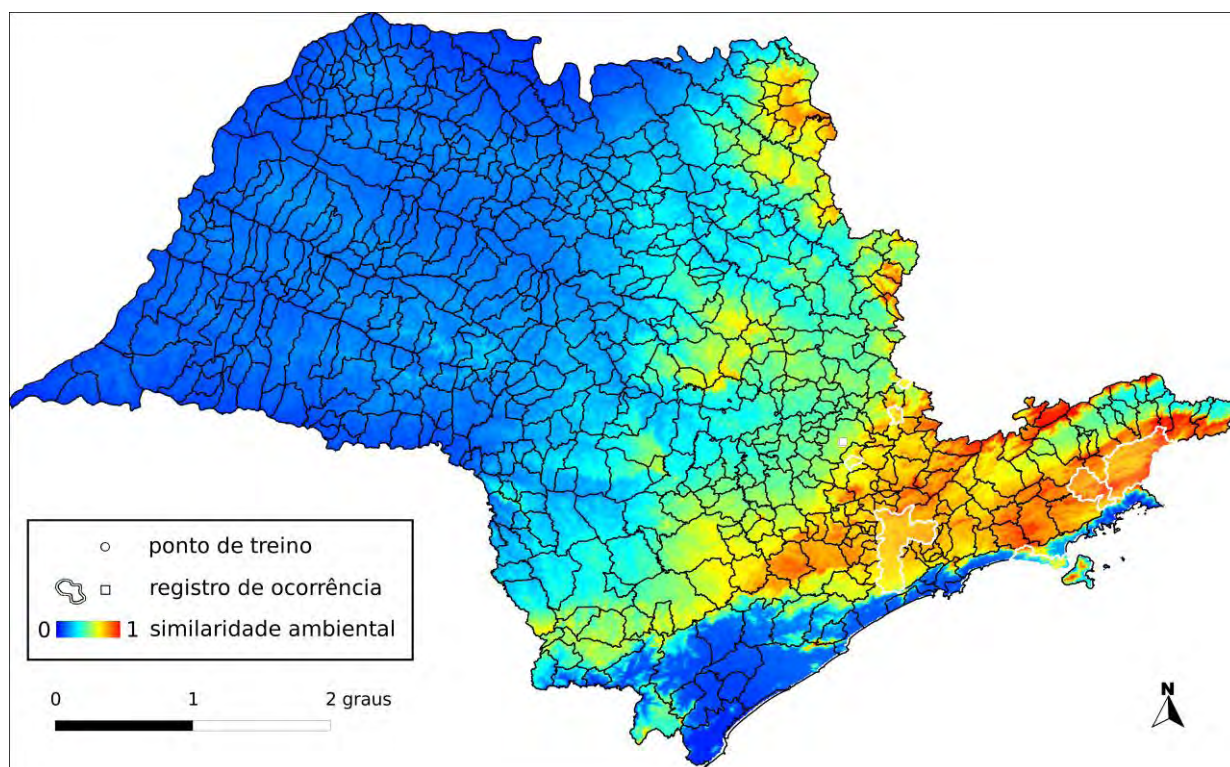
UCs³¹: P.E. de Campos do Jordão (●), P.E. dos Mananciais de Campos do Jordão (●)

Dados ambientais:

bio02 var. temp. dia (°C*10)	114	bio17 prec. tri. mais seco (mm)	276
bio05 temp. max. (°C*10)	287	bio18 prec. tri. mais quente (mm)	520
bio06 temp min. (°C*10)	100	bio19 prec. tri. mais frio (mm)	315
bio16 prec. tri. mais úmido (mm)	520	alt altitude (m)	621

³¹ Registro de ocorrência inteiramente (●) ou parcialmente (●) contido em UC.

Passiflora sidifolia M.Roem.



Algoritmo: Maxent

Total de registros disponíveis na rede speciesLink: 87

Omissão interna (BR): 0,00 (n=8) **Omissão interna (SP):** – (n=0)

Omissão externa (BR): 0,08 (n=24) **Omissão externa (SP):** 0,00 (n=7)

Significância (BR): $p < 0,01$

Área do modelo (SP): 96.118 km² (38,7% da área do Estado)

Área em remanescentes (2008): 23.892 km² (24,9% da área do modelo)

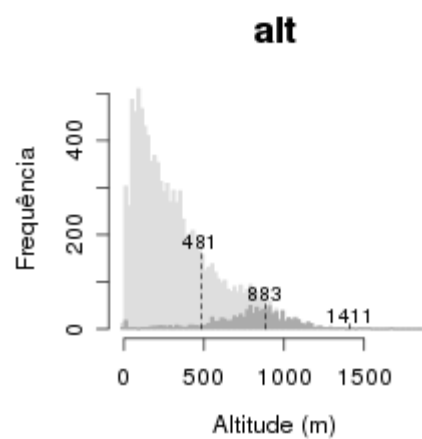
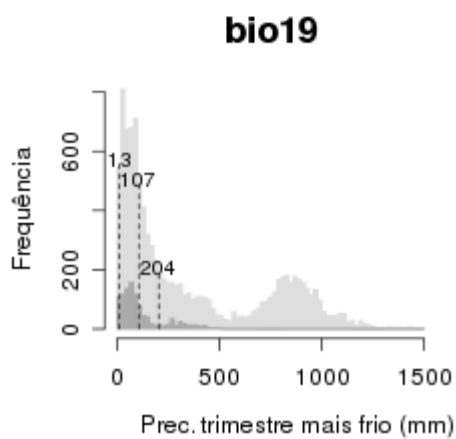
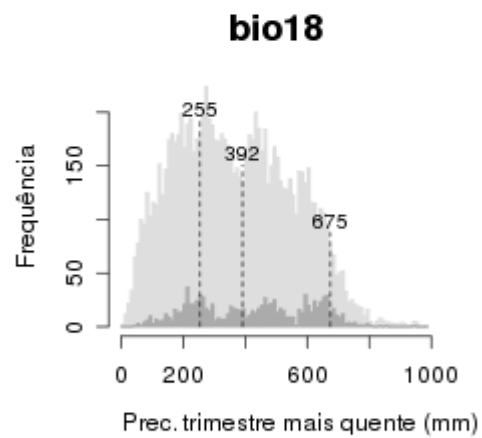
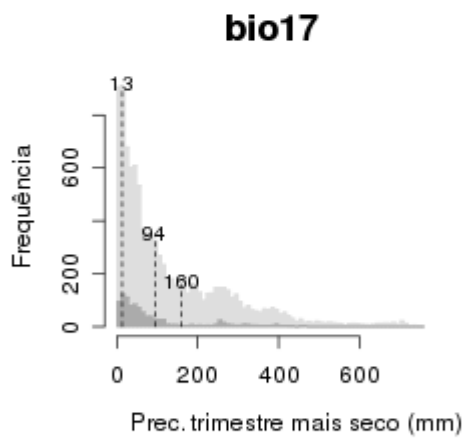
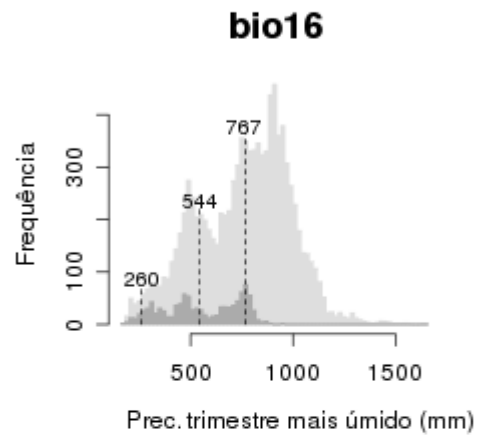
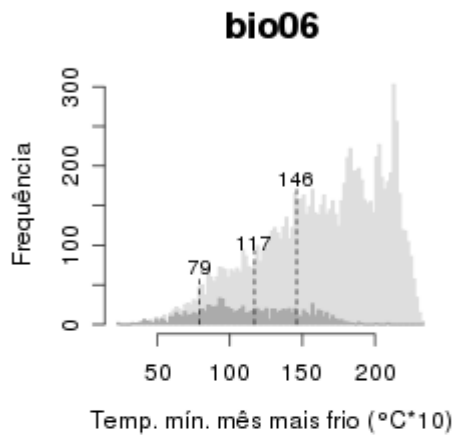
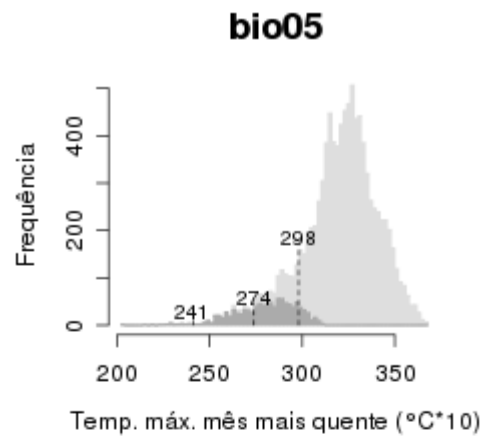
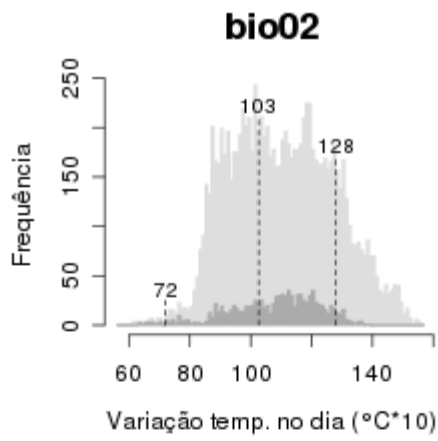
Área em remanescentes em UCs (2008): 4.728 km² (4,9% da área do modelo)

Ano da última coleta em SP: 2007 (IAC48230)

Municípios com registros de ocorrência em SP: Cunha (IAC40227), Monte Alegre do Sul (IAC28944), São Luiz do Paraitinga (UPCB51900), São Paulo (IAC37658), São Sebastião (IAC42201), Vinhedo (SPSF29940), Águas de Lindóia (IAC32344)

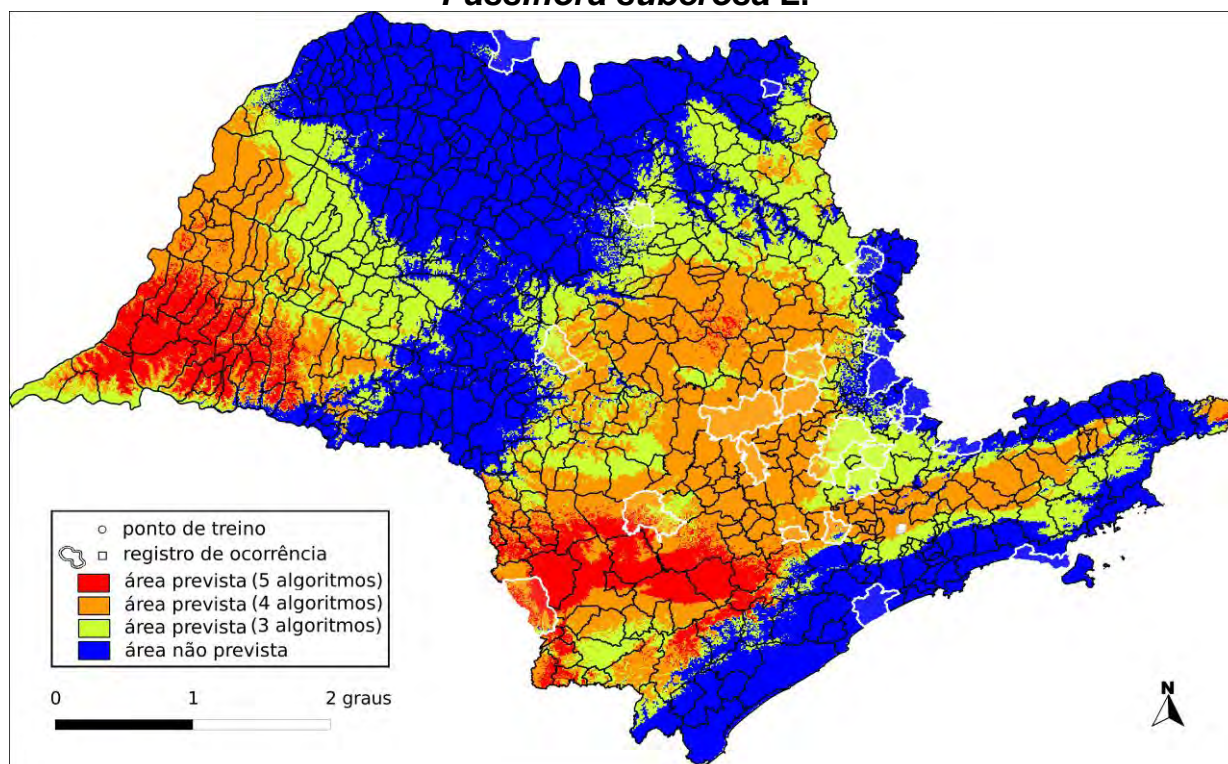
UCs³²: P.E. do Jaraguá ●, P.E. Guarapiranga ●, P.E. Alberto Löfgren ●, P.E. da Cantareira ●, P.E. da Serra do Mar ●, P.N. da Serra da Bocaina ●

³² Registro de ocorrência inteiramente (●) ou parcialmente (◐) contido em UC.



Histogramas: Brasil (claro) e projeção do modelo no Brasil (escuro). Valores: mínimo, médio e máximo dos pontos de ocorrência.

Passiflora suberosa L.



Algoritmo: Consenso (ENFA, GARP BS, Maxent, MD e SVM)

Total de registros disponíveis na rede speciesLink: 351

Omissão interna (BR): 0,00 (n=25) **Omissão interna (SP):** – (n=0)

Omissão externa (BR): 0,22 (n=107) **Omissão externa (SP):** 0,08 (n=25)

Significância (BR): $p < 0,01$

Área do modelo (SP): 149.470 km² (60,2% da área do Estado)

Área em remanescentes (2008): 21.289 km² (14,2% da área do modelo)

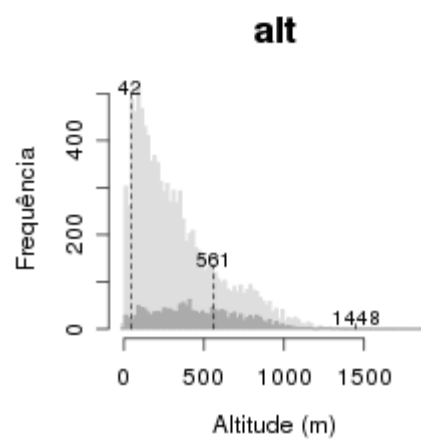
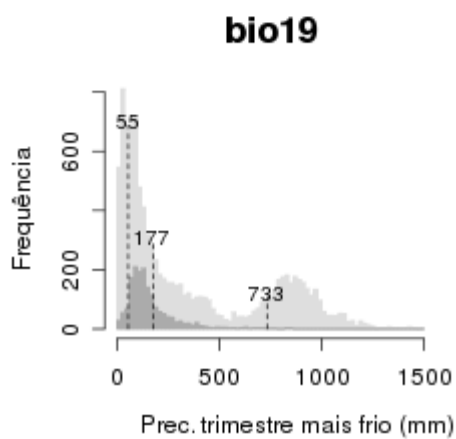
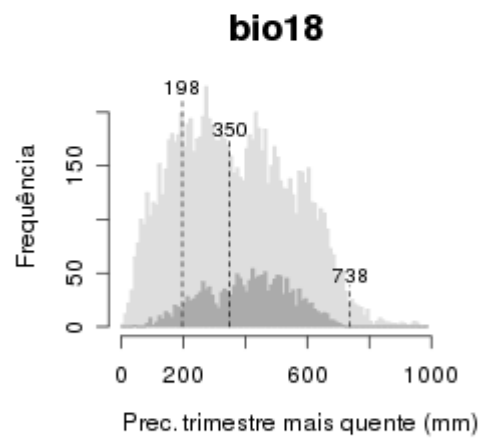
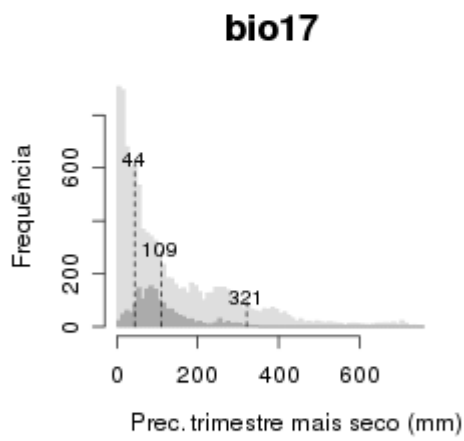
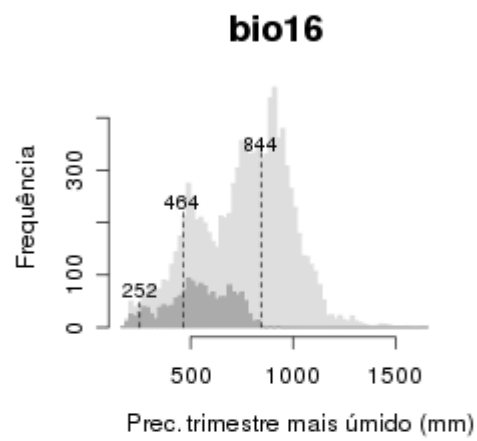
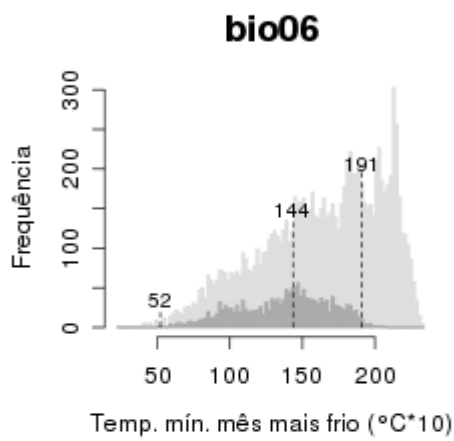
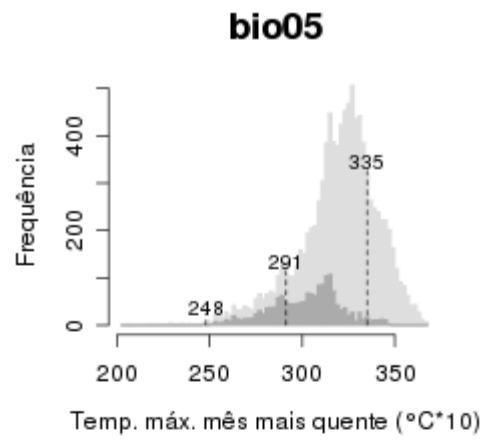
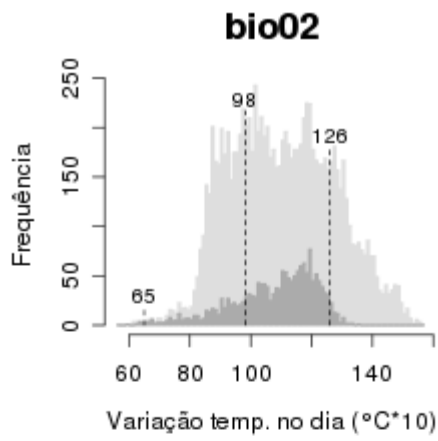
Área em remanescentes em UCs (2008): 2.411 km² (1,6% da área do modelo)

Ano da última coleta em SP: 2009 (IAC52101)

Municípios com registros de ocorrência em SP: Angatuba (IAC26528), Araras (SP114203), Bauru (IAC32615), Campinas (IAC2891), Espírito Santo do Pinhal (SP78805), Indaiatuba (SP119768), Itanhaém (**RB482678**), Itapira (IAC32609), Itararé (SP101664), Itatiba (IAC52101), Jeriquara (SP115757), Joanópolis (SP108720), Jundiá (IAC35554), Limeira (IAC32613), Monte Alegre do Sul (SP78813), Monte Alto (IAC39317), Paulo de Faria (HSJRP29345), Pinhalzinho (RB482702), Piracicaba (IAC4371), Socorro (SP356841), São José do Rio Pardo (IAC33888), São Roque (SP291447), São Sebastião (**IAC36160**), Tietê (IAC30453), Votorantim (RB200581)

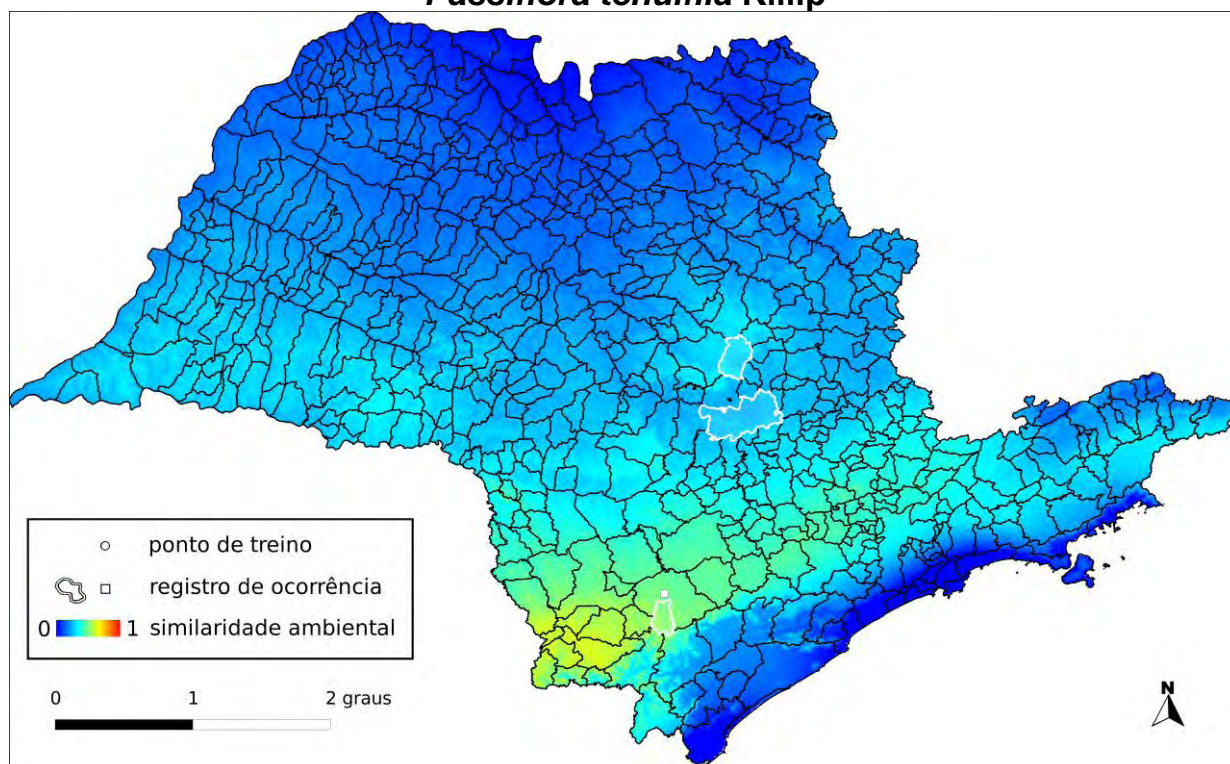
UCs³³: E.E. de Angatuba ●, E.E. de Paulo de Faria ●, E.E. Ibicatu ●, P.E. da Serra do Mar ●, E.E. Mogi Guaçu ●, P.N. da Serra da Bocaina ●, E.E. de Bauru ●, P.E. da Ara ●, E.E. de Valinhos ●

³³ Registro de ocorrência inteiramente (●) ou parcialmente (◐) contido em UC.



Histogramas: Brasil (claro) e projeção do modelo no Brasil (escuro). Valores: mínimo, médio e máximo dos pontos de ocorrência.

Passiflora tenuifila Killip



Algoritmo: Distância Euclideana

Total de registros disponíveis na rede speciesLink: 52

Nº de pontos usados na modelagem: 1 (0 em SP)

Ano da última coleta em SP: 2005 (IAC46787)

Municípios com registros de ocorrência em SP: Itirapina (HSJRP00646), Piracicaba (IAC46787), Ribeirão Grande (UPCB31864)

UCs³⁴: E.E. de Xitué ●, E.E. Itirapina ●, P.E. de Intervalos ●, E.E. Ibicatu ●

Dados ambientais:

bio02 var. temp. dia (°C*10)	109	bio17 prec. tri. mais seco (mm)	351
bio05 temp. max. (°C*10)	244	bio18 prec. tri. mais quente (mm)	440
bio06 temp min. (°C*10)	44	bio19 prec. tri. mais frio (mm)	370
bio16 prec. tri. mais úmido (mm)	447	alt altitude (m)	1090

³⁴ Registro de ocorrência inteiramente (●) ou parcialmente (◐) contido em UC.

***Passiflora transversalis* M.A.M.Azevedo**

Total de registros disponíveis na rede *speciesLink*: 66

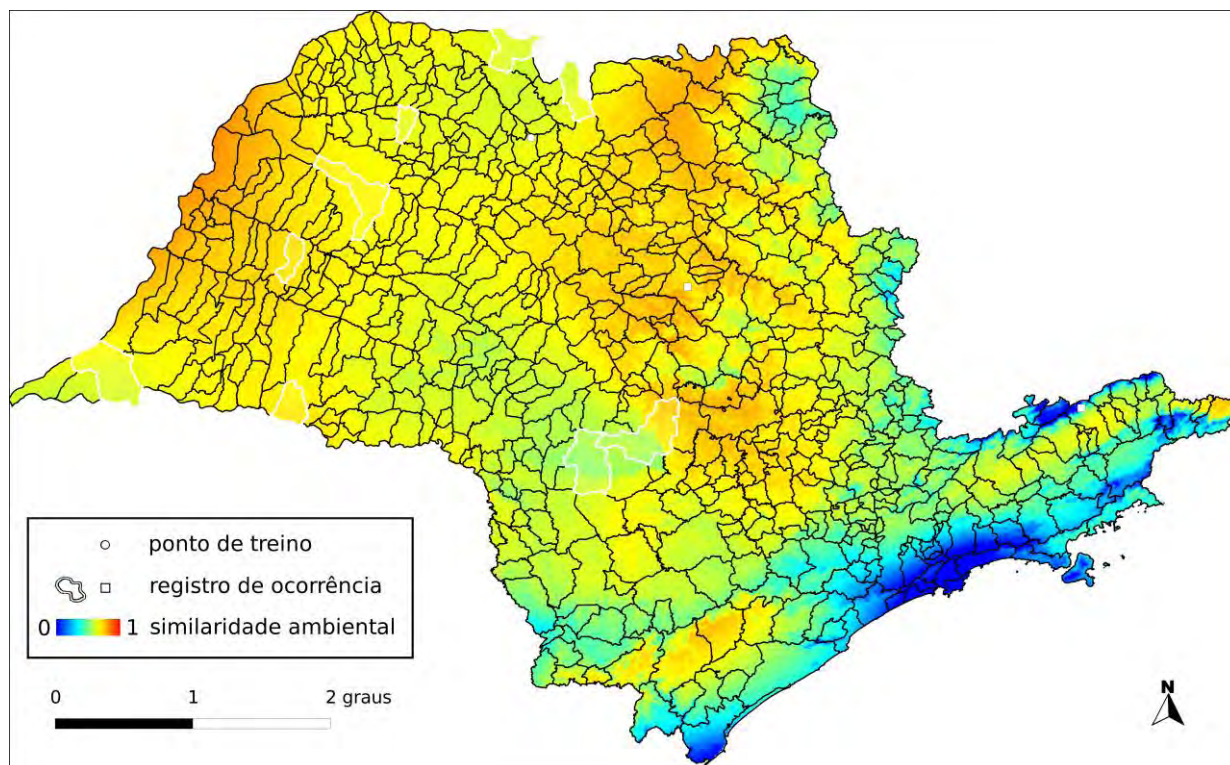
Nº de pontos usados na modelagem: 0

Ano da última coleta em SP: 1939 (IAC32626)

Municípios com registros de ocorrência em SP: Guarulhos (IAC32626), São Paulo (IAC32627)

UCs: sem informação

Passiflora tricuspis Mast.



Algoritmo: Distância Euclideana

Total de registros disponíveis na rede speciesLink: 153

Nº de pontos usados na modelagem: 3 (1 em SP)

Ano da última coleta em SP: 2009 (IAC51445)

Municípios com registros de ocorrência em SP: Adamantina (IAC48053), Araçatuba (UPCB51415), Avaré (MO100238559), Botucatu (IAC32629), Guaraci (SPSF21332), Iepê (FUEL4167), Magda (IAC29576), Paulo de Faria (HSJRP10729), São José do Rio Preto (IAC51445), Teodoro Sampaio (IAC32630)

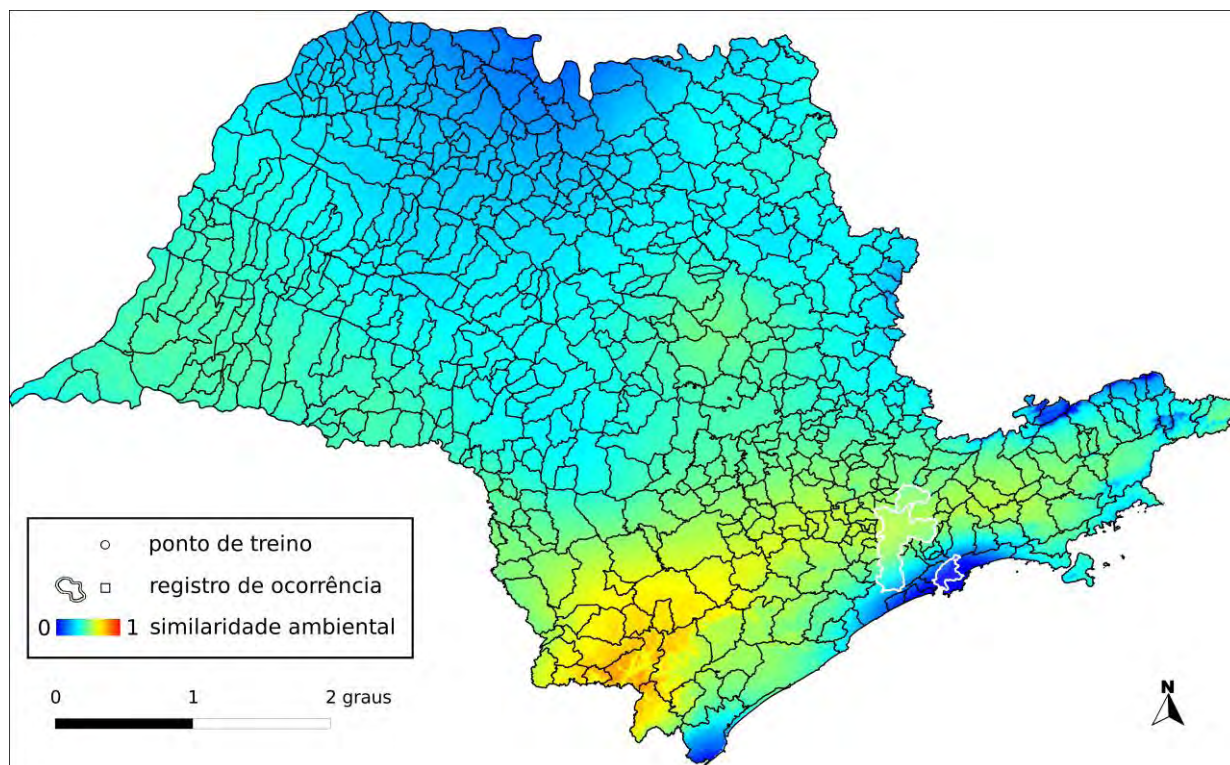
UCs³⁵: E.E. de Paulo de Faria (●), P.E. do Morro do Diabo (●), E.E. Mico Leão Preto (●)

Dados ambientais:

bio02 var. temp. dia (°C*10)	101 – 129	bio17 prec. tri. mais seco (mm)	60 – 148
bio05 temp. max. (°C*10)	306 – 332	bio18 prec. tri. mais quente (mm)	491 – 598
bio06 temp min. (°C*10)	118 – 183	bio19 prec. tri. mais frio (mm)	87 – 187
bio16 prec. tri. mais úmido (mm)	561 – 867	alt altitude (m)	101 – 502

³⁵ Registro de ocorrência inteiramente (●) ou parcialmente (◐) contido em UC.

Passiflora truncata Regel



Algoritmo: Distância Euclideana

Total de registros disponíveis na rede speciesLink: 53

Nº de pontos usados na modelagem: 2 (0 em SP)

Ano da última coleta em SP: 2000 (IAC44371)

Municípios com registros de ocorrência em SP: Mairiporã (IAC44371), Santos (RB361207), São Paulo (IAC45184)

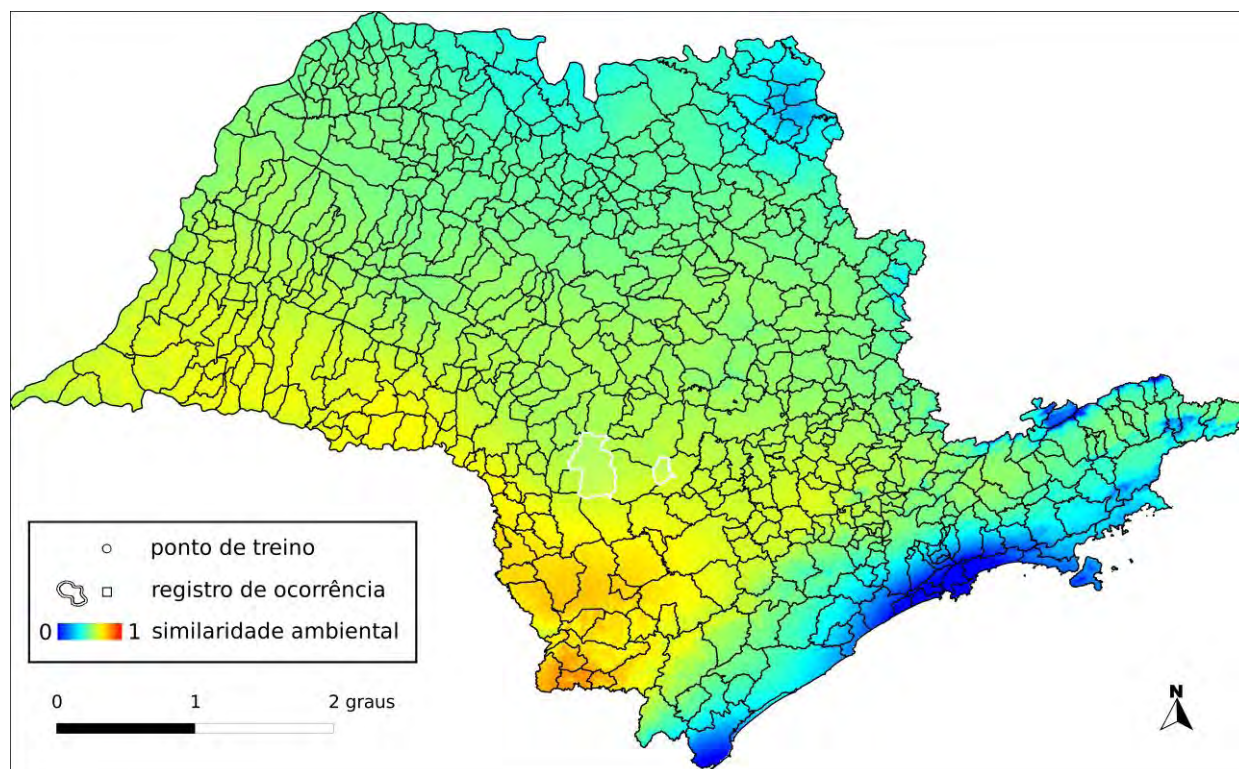
UCs³⁶: P.E. do Juquery (●), P.E. de Jacupiranga (●), P.E. do Jaraguá (●), P.E. Guarapiranga (●), P.E. Alberto Löfgren (●), P.E. da Cantareira (●), P.E. da Serra do Mar (●)

Dados ambientais:

bio02 var. temp. dia (°C*10)	94 – 102	bio17 prec. tri. mais seco (mm)	283 – 291
bio05 temp. max. (°C*10)	274 – 279	bio18 prec. tri. mais quente (mm)	469 – 511
bio06 temp min. (°C*10)	81 – 97	bio19 prec. tri. mais frio (mm)	305 – 317
bio16 prec. tri. mais úmido (mm)	469 – 511	alt altitude (m)	476 – 647

³⁶ Registro de ocorrência inteiramente (●) ou parcialmente (◐) contido em UC.

Passiflora urnifolia Rusby



Algoritmo: Distância Euclideana

Total de registros disponíveis na rede *speciesLink*: 38

Nº de pontos usados na modelagem: 1 (0 em SP)

Ano da última coleta em SP: 2007 (IAC49460)

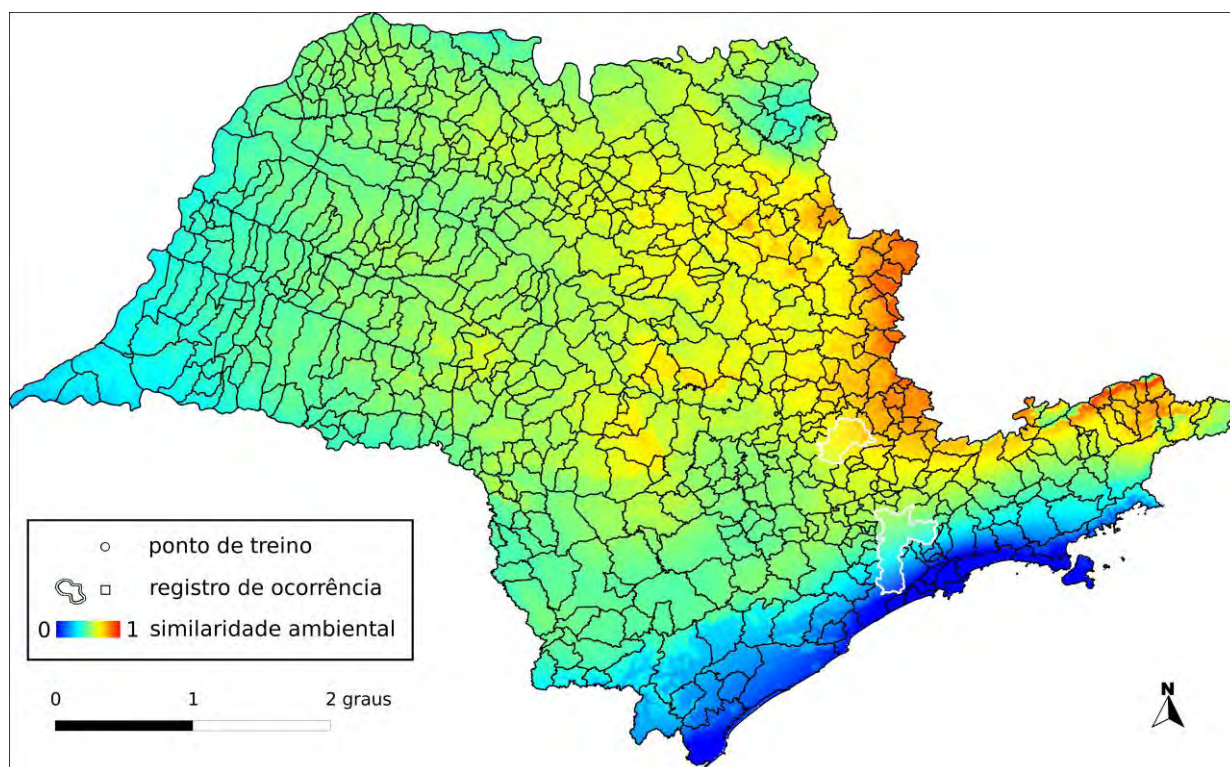
Municípios com registros de ocorrência em SP: Avaré (MBM154742), Pardinho (IAC49460)

UCs: sem informação

Dados ambientais:

bio02 var. temp. dia (°C*10)	115	bio17 prec. tri. mais seco (mm)	276
bio05 temp. max. (°C*10)	289	bio18 prec. tri. mais quente (mm)	520
bio06 temp min. (°C*10)	100	bio19 prec. tri. mais frio (mm)	315
bio16 prec. tri. mais úmido (mm)	520	alt altitude (m)	611

Passiflora vellozii Gardner



Algoritmo: Distância Euclideana

Total de registros disponíveis na rede *speciesLink*: 38

Nº de pontos usados na modelagem: 1 (0 em SP)

Ano da última coleta em SP: 1995 (IAC32656)

Municípios com registros de ocorrência em SP: Campinas (SP35581), São Paulo (SP290807)

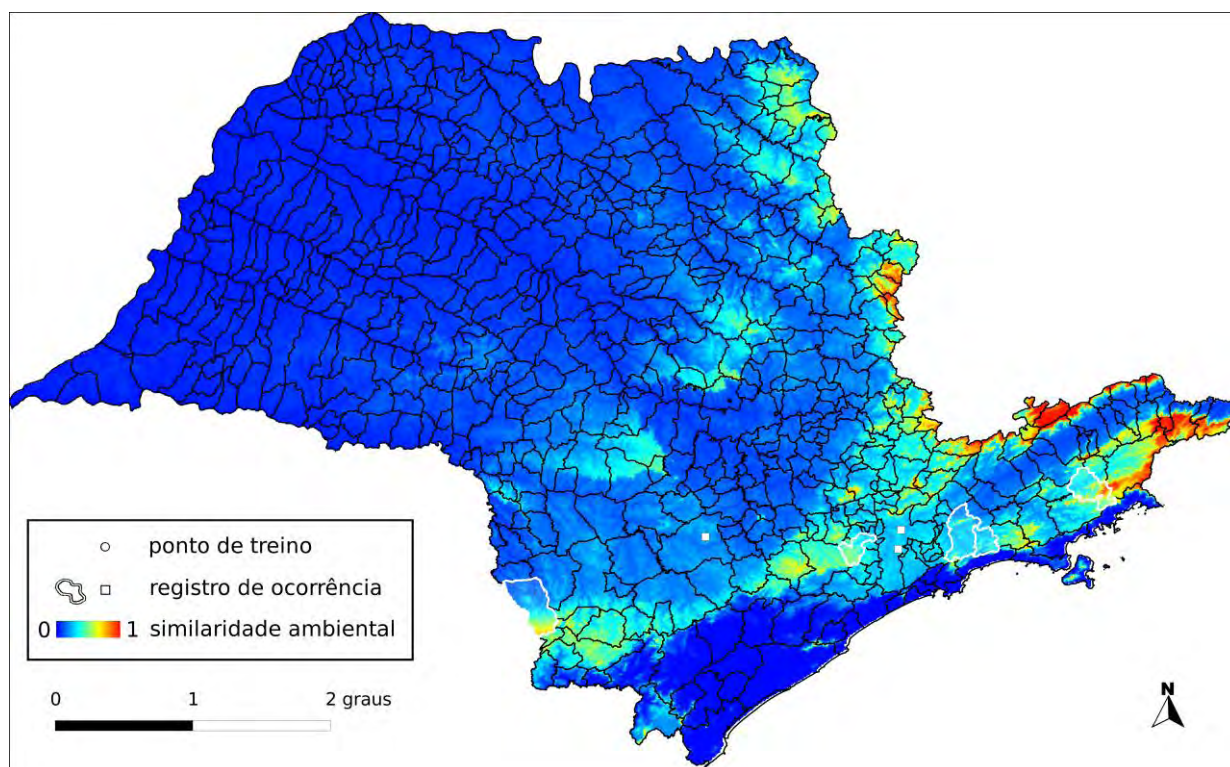
UCs³⁷: P.E. do Jaraguá ●, P.E. Guarapiranga ●, P.E. Alberto Löfgren ●, P.E. da Cantareira ●, P.E. da Serra do Mar ●, P.N. da Serra da Bocaina ●, P.E. da Ara ●, E.E. de Valinhos ●

Dados ambientais:

bio02 var. temp. dia (°C*10)	123	bio17 prec. tri. mais seco (mm)	72
bio05 temp. max. (°C*10)	260	bio18 prec. tri. mais quente (mm)	746
bio06 temp min. (°C*10)	66	bio19 prec. tri. mais frio (mm)	72
bio16 prec. tri. mais úmido (mm)	808	alt altitude (m)	1029

³⁷ Registro de ocorrência inteiramente (●) ou parcialmente (◐) contido em UC.

Passiflora villosa Vell.



Algoritmo: Maxent

Total de registros disponíveis na rede speciesLink: 105

Omissão interna (BR): 0,00 (n=6) **Omissão interna (SP):** – (n=0)

Omissão externa (BR): 0,10 (n=40) **Omissão externa (SP):** 0,00 (n=5)

Significância (BR): $p < 0,01$

Área do modelo (SP): 64.040 km² (25,8% da área do Estado)

Área em remanescentes (2008): 18.594 km² (29% da área do modelo)

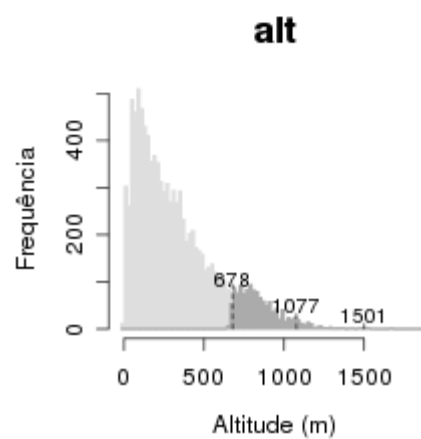
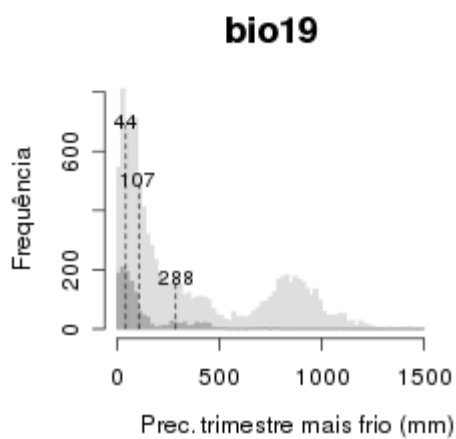
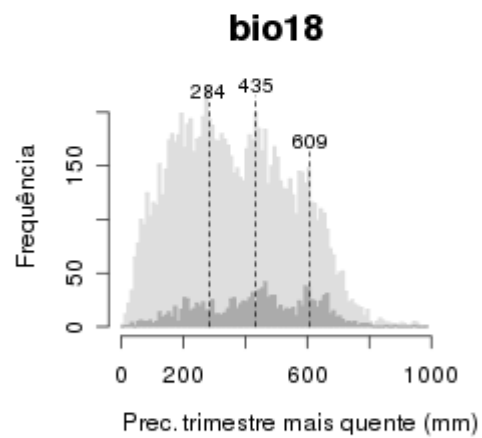
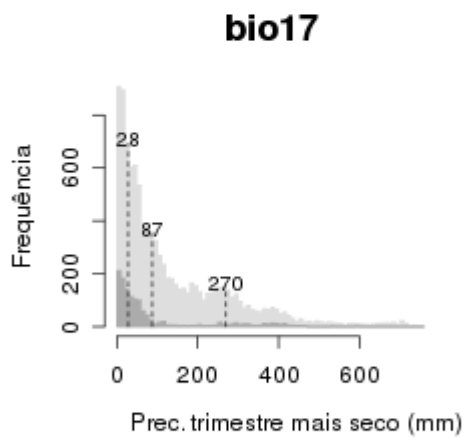
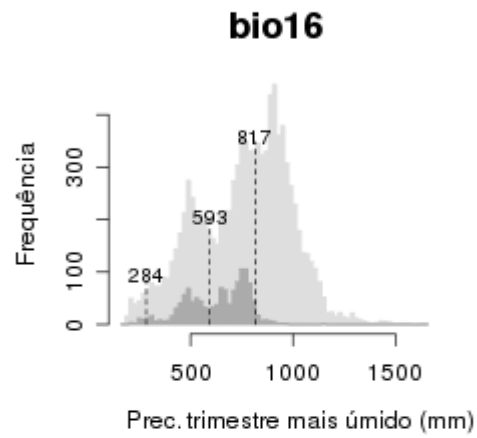
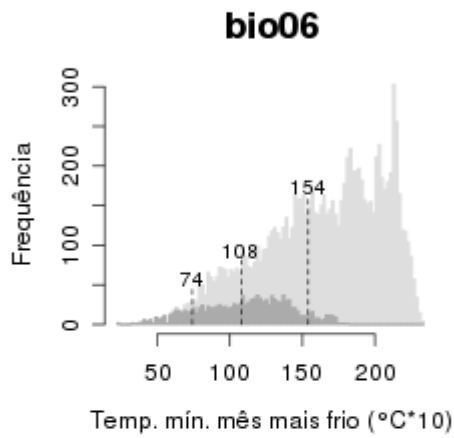
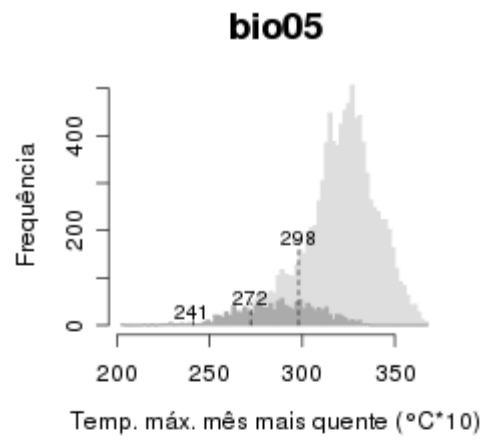
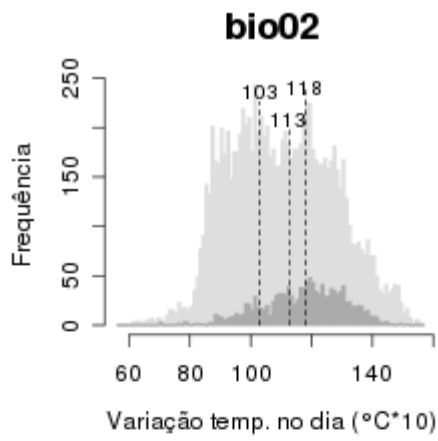
Área em remanescentes em UCs (2008): 3.179 km² (5% da área do modelo)

Ano da última coleta em SP: 2002 (IAC42137)

Municípios com registros de ocorrência em SP: Biritiba-Mirim (IAC42137), Cotia (RB45467), Itararé (MBM208804), Mogi das Cruzes (UPCB15744), São Luís do Paraitinga (IAC21114)

UCs³⁸: P.E. da Serra do Mar ●, P.E. de Jacupiranga ●, E.E. de Itapeti ●

³⁸ Registro de ocorrência inteiramente (●) ou parcialmente (◐) contido em UC.



Histogramas: Brasil (claro) e projeção do modelo no Brasil (escuro). Valores: mínimo, médio e máximo dos pontos de ocorrência.

Discussão

A abordagem utilizada buscou aproveitar ao máximo os dados de herbário existentes sobre cada espécie, incluindo diferentes estratégias de modelagem com fins de conservação. Entretanto, o bom uso dos resultados requer entendimento sobre as limitações e potencialidades de cada estratégia. Em casos onde o número de pontos de ocorrência é muito baixo, modelos de dissimilaridade ambiental oferecem uma primeira aproximação quanto ao potencial de ocorrência da espécie. Quando aplicados a espécies com distribuição notadamente restrita ou pouco conhecida, podem ser uma importante ferramenta para ampliar o conhecimento sobre os requisitos ambientais e a distribuição da espécie (Siqueira *et al.* 2009). Este pode ser o caso de *P. catharinensis*, recentemente coletada pela primeira vez no Estado, e de *P. ischnoclada*, possivelmente microendêmica (Bernacci *et al.* 2003). Porém, quando aplicados a espécies de ampla distribuição que por algum motivo não possuem pontos de ocorrência suficientes para gerar modelos mais robustos, o mapa gerado pode não oferecer indicação precisa quanto ao potencial de distribuição da espécie. Mesmo em tais casos, é importante observar que o mapa pode ao menos evidenciar locais onde a amostragem da espécie seria estratégica do ponto de vista ambiental para conseguir gerar modelos mais robustos no futuro. Este parece ser o caso de *P. actinia*, que possui mais de cem registros disponíveis na rede *speciesLink* e cujos dois únicos pontos usados na construção do modelo não foram suficientes para indicar áreas próximas ao litoral norte de SP, onde sabe-se que existem registros de ocorrência da espécie, como propícias a sua ocorrência.

Para espécies com menos de cinco pontos, além de terem sido gerados modelos de dissimilaridade ambiental, os dados de herbários foram capazes de fornecer indicações importantes para priorizar ações de conservação, uma vez que espécies raras e com distribuição bastante restrita tendem a ter poucos dados. É interessante observar que mesmo adotando critérios simples como os que foram utilizados aqui para selecionar espécies cuja distribuição é menos conhecida em SP (anexo I) foi possível obter uma boa aproximação sobre quais espécies demandam maior atenção para conservação. Entre as 15 espécies selecionadas, estão todas as cinco espécies da família que constam no Livro Vermelho das Espécies Vegetais Ameaçadas do Estado de São Paulo (Mamede *et al.* 2007) bem como a única espécie dessa família que ocorre em São Paulo e que consta na Lista Vermelha de Espécies da Flora do Brasil (MMA 2008). Uma possível exceção está relacionada com *P. rhamnifolia*, cuja distribuição em SP provavelmente é bastante restrita (litoral norte do estado), mas que não foi incluída em nenhum dos anexos visto que possui coleta recente, teve seu modelo gerado com mais de 5 pontos, e a área de ocorrência potencial foi bastante grande.

As técnicas de modelagem utilizadas em espécies com mais de quatro pontos permitiram calcular novos dados, além de fornecer uma melhor aproximação quanto à distribuição da espécie. Mesmo assim, é importante mencionar que modelos gerados com baixo número de pontos devem ser vistos com ressalvas e utilizados de forma exploratória (Wisn *et al.* 2008). Além das espécies modeladas com técnicas de dissimilaridade ambiental, outros oito modelos foram gerados com número

considerado ainda baixo de pontos (menos de 20), sendo que em um dos casos a área de ocorrência potencial foi extremamente baixa (*P. pentagona*, 18km²), sugerindo maior cuidado para essa espécie.

Entre as espécies com mais de 20 pontos, chama a atenção o caso de *P. mucronata*, cujo mapa de distribuição potencial não incluiu nenhuma área em SP. Tal fato pode ser explicado levando-se em conta que o litoral norte de SP é justamente considerado como um dos limites da distribuição da espécie (Bernacci *et al.* 2003) – região com características ambientais possivelmente distintas das demais áreas de ocorrência, não tendo sido ela incluída entre os pontos de treino do modelo.

À exceção de *P. mucronata* e *P. pentagona*, os erros de omissão para registros em SP foram em geral baixos e quase todos os modelos foram significativos ($p < 0,01$). O único modelo que não apresentou boa significância foi o de *P. misera*. Neste caso, embora o mapa de distribuição potencial tenha sido bastante restrito para o estado, o modelo incluiu todos os municípios com registro de ocorrência.

O fato de cinco espécies não possuírem nenhum registro de coleta em SP mostra que esforços de digitação ainda podem contribuir com dados importantes. Como exemplo, em Bernacci *et al.* (2003) há referência a três materiais de *P. elegans* e um de *P. malacophylla* coletados em SP que ainda não estão disponíveis na rede *speciesLink*.

Outro dado relevante nos resultados quanto à conservação das espécies é que praticamente todas que apresentaram registros de ocorrência em SP ou já foram coletadas em Unidades de Conservação ou potencialmente ocorrem em alguma Unidade de Conservação próxima aos locais onde houve coleta. As únicas exceções foram *P. cervii*, *P. elegans*, *P. malacophylla*, *P. pardifolia*, *P. racemosa*, *P. transversalis* e *P. urnifolia*.

Conclusões e Recomendações

Os resultados deste trabalho indicam que é possível ajudar a direcionar ações de conservação utilizando dados de herbários disponíveis através da Internet e técnicas de modelagem de nicho ecológico. Tal fato torna-se particularmente importante na medida em que limitações de recursos humanos e financeiros constantemente obrigam o estabelecimento de prioridades na definição de tais ações. Uma possível estratégia seria empreender esforços de digitação, pesquisa e trabalho de campo de forma a:

- Priorizar espécies pouco conhecidas ou com distribuição muito restrita que possam estar mais vulneráveis à extinção;
- Garantir que todas as espécies tenham sido coletadas há menos de 10 anos;
- Garantir que todas as espécies tenham ao menos uma coleta confirmada dentro de Unidades de Conservação;

- Alcançar um número mínimo de pontos de ocorrência com precisão suficiente para gerar modelos de nicho ecológico em alta resolução com qualidade suficiente para cálculo de indicadores de conservação.

Seguindo tal abordagem, podem ser detectadas situações que requeiram a elaboração de planos de ação específicos para determinadas espécies em função do grau de ameaça ao qual estejam submetidas. *P. ischnoclada*, por exemplo, pode ter sua distribuição de fato bastante restrita devido ao clima. Neste caso, modelos de nicho ecológico poderiam ajudar a confirmar ou não tal hipótese, ajudando a direcionar ações mais específicas de conservação principalmente levando-se em conta o possível impacto de mudanças climáticas. Em tais situações, os perfis ambientais podem se mostrar particularmente importantes em planos de conservação *ex situ*, indicando condições ideais de cultivo.

Futuramente, riscos de extinção por perda de habitat também poderão ser calculados com base na área de ocorrência potencial da espécie, desde que mapas de remanescentes de vegetação relativos a outros anos estejam acessíveis. Neste caso, um interessante tema a ser pesquisado seria como aproximar a área de ocorrência real a partir da área de ocorrência potencial.

A abordagem aqui apresentada poderia facilmente se estender a outras famílias botânicas, prevendo atividades de monitoramento ao longo do tempo. Na medida em que mais dados forem obtidos, modelos mais robustos poderão ser gerados, ajudando a compreender melhor a distribuição e os requisitos ambientais de cada espécie.

Referências

Bernacci, L.C. (2003) Passifloraceae. Em: Wanderley, M.G.L.; Shepherd, G.J.; Giulietti, A.M. & Melhem, T.S. (coord.) Flora Fanerogâmica do Estado de São Paulo. vol. 3: 247-248. RiMa/FAPESP, São Paulo.

Cervi, A.C., Milward-de-Azevedo, M.A., Bernacci, L.C. (2010) *Passifloraceae* in Lista de Espécies da Flora do Brasil. Jardim Botânico do Rio de Janeiro (<http://floradobrasil.jbrj.gov.br/2010/FB000182>).

CSR/Ibama (2009) Relatório Técnico de Monitoramento do Desmatamento no Bioma Cerrado, 2002 a 2008 – Dados Revisados. Novembro de 2009. http://siscom.ibama.gov.br/monitorabiomas/cerrado/Relatorio%20tecnico_Monitoramento%20Desmate_Bioma%20Cerrado_CSR_REV.pdf (acesso em 2011-05-27)

CSR/Ibama (2010) Relatório Técnico de Monitoramento do Desmatamento no Bioma Mata Atlântica, 2002 a 2008. 1 de Novembro de 2010. http://siscom.ibama.gov.br/monitorabiomas/mataatlantica/RELATORIO_PMDBBS_MATA_ATLANTICA_2002-2008.pdf (acesso em 2011-05-27)

IBGE (2002) Área Territorial Oficial. Resolução N°5. 10 de outubro de 2002.

IBGE (2007) Contas Regionais do Brasil 2002-2005. 26 de novembro de 2007. http://www.ibge.gov.br/home/presidencia/noticias/noticia_visualiza.php?id_noticia=1039&id_pagina=1 (acesso em 2011-05-29)

IBGE (2010) Dados do Censo 2010. Diário Oficial da União. 4 de novembro de 2010.

IUCN (2001) IUCN Red List Categories and Criteria version 3.1. <http://www.iucnredlist.org/technical-documents/categories-and-criteria/2001-categories-criteria> (acesso em 2011-04-26)

IUCN & UNEP (2010) The World Database on Protected Areas (WDPA). UNEP-WCMC. Cambridge, UK. <http://www.protectedplanet.net> (acesso em 2011-04-26)

Mamede, M.C.H.; Souza, V.C.; Prado, J.; Barros, F.; Wanderley, M.G.L. & Rando, J.G. (2007) Livro Vermelho das Espécies Vegetais Ameaçadas do Estado de SP. Instituto de Botânica, São Paulo, 165p.

MMA (2008) Instrução Normativa MMA nº 06, 23 de setembro de 2008.

PMDBBS (2011) Programa de Monitoramento do Desmatamento dos Biomas Brasileiros por Satélite, Secretaria de Biodiversidade e Florestas do Ministério do Meio Ambiente. <http://siscom.ibama.gov.br/monitorabiomas/> (acesso em 2011-05-27).

Siqueira, M.F., Durigan, G., De Marco Júnior, P. & Peterson, A.T. (2009) Something from nothing: Using landscape similarity and ecological niche modelling to find rare or endangered species. *Journal for Nature Conservation*, 17 (1): 25-32.

SMA (2009) Unidades de conservação da natureza. Secretaria do Meio Ambiente do Estado de São Paulo, Fundação Florestal, Luiz Roberto Numa de Oliveira (coord.), São Paulo, 104p.

Victor, M.A.M., Cavalli, A.C., Guillaumon, J.R. & Filho, R.S. (2005) Cem Anos de Devastação: revisitada 30 anos depois. Ministério do Meio Ambiente, Secretaria de Biodiversidade e Florestas, Brasília, 72p.

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. *Diversity and Distributions*, 14: 763–773. doi:10.1111/j.1472-4642.2008.00482.x.

Anexo I – Espécies com distribuição pouco conhecida em São Paulo

Critério: espécies com menos de 5 pontos usados na modelagem e com registro de ocorrência em menos de 3 municípios no Estado.

Espécie	Nº de municípios em que há registro de ocorrência	Nº registros usados no modelo (SP) *	Nº registros usados no modelo (BR) **
<i>Passiflora cervii</i>	0	0	0
<i>Passiflora elegans</i>	0	0	0
<i>Passiflora malacophylla</i>	0	0	0
<i>Passiflora racemosa</i>	0	0	1
<i>Passiflora campanulata</i>	0	0	3
<i>Passiflora pardifolia</i>	1	0	0
<i>Passiflora setulosa</i>	1	0	1
<i>Passiflora ischnoclada</i>	1	1	1
<i>Passiflora catharinensis</i>	1	1	1
<i>Passiflora glaucescens</i>	2	0	0
<i>Passiflora transversalis</i>	2	0	0
<i>Passiflora urnifolia</i>	2	0	1
<i>Passiflora vellozii</i>	2	0	1
<i>Passiflora marginata</i>	2	1	1
<i>Passiflora loefgrenii</i>	2	1	2

(*) coletas realizadas no Estado de São Paulo.

(**) coletas realizadas no Brasil.

Anexo II – Espécies com distribuição potencialmente restrita em São Paulo

Critério: espécies com pelo menos 5 pontos usados na modelagem e com área de distribuição potencial menor que 10.000 km² em remanescentes de vegetação no Estado.

Espécie	Área de ocorrência potencial em SP (km ²)	Nº registros usados no modelo (BR)
<i>Passiflora mucronata</i>	0	23
<i>Passiflora pentagona</i>	18	5

Anexo III – Espécies sem registro recente de coleta em São Paulo

Critério: ano da última coleta anterior a 2001.

Espécie	Ano da última coleta (referência <i>speciesLink</i>)	Material mais recente examinado na Flora Fanerogâmica de SP	Nº registros usados no modelo (BR)
<i>Passiflora cervii</i>	–	–	0
<i>Passiflora elegans</i>	–	1999	0
<i>Passiflora malacophylla</i>	–	1892	0
<i>Passiflora racemosa</i>	–	1952	1
<i>Passiflora setulosa</i>	1935	1951	1
<i>Passiflora transversalis</i>	1939	–	0
<i>Passiflora clathrata</i>	1964	1999	3
<i>Passiflora glaucescens</i>	1976	–	0
<i>Passiflora lepidota</i>	1993	1993	2
<i>Passiflora marginata</i>	1994	1915	1
<i>Passiflora eichleriana</i>	1994	1994	4
<i>Passiflora vellozii</i>	1995	1995	1
<i>Passiflora caerulea</i>	1997	1997	9
<i>Passiflora misera</i>	1997	1985	25
<i>Passiflora pohlii</i>	1999	1995	16
<i>Passiflora truncata</i>	2000	1988	2