



openModeller

new computational tools for ecological niche modelling



CRIA, INPE, Poli-USP

IEA-USP 2008



Why openModeller?

- Proof of concept: specimen occurrence data as infrastructure for research and policy making
- First prototype developed as part of the speciesLink project (2001-2005)
 - Open source, C++, available at sourceforge
 - Attracted international collaboration: University of Kansas and University of Reading



Why openModeller?

- Modeling process is usually quite **complex and time consuming**
- requires **great expertise** in a number of software and tools (different formats, different software, ...)
- framework **capable of dealing with different projections, coordinate systems, and formats**
- By using such a framework, specialists would be able to **concentrate more on the analysis** than on preparation of data

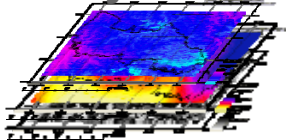


Thematic project funded by Fapesp (2006 – 2009)

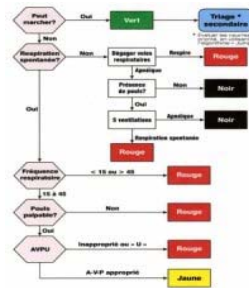
Occurrence
data



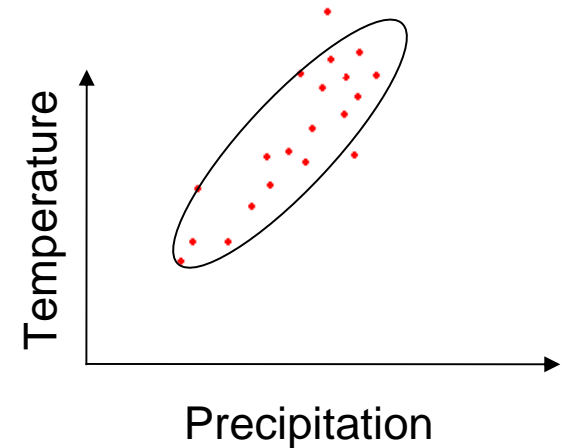
Environmental
layers



algorithm



Distribution model
(environmental space)

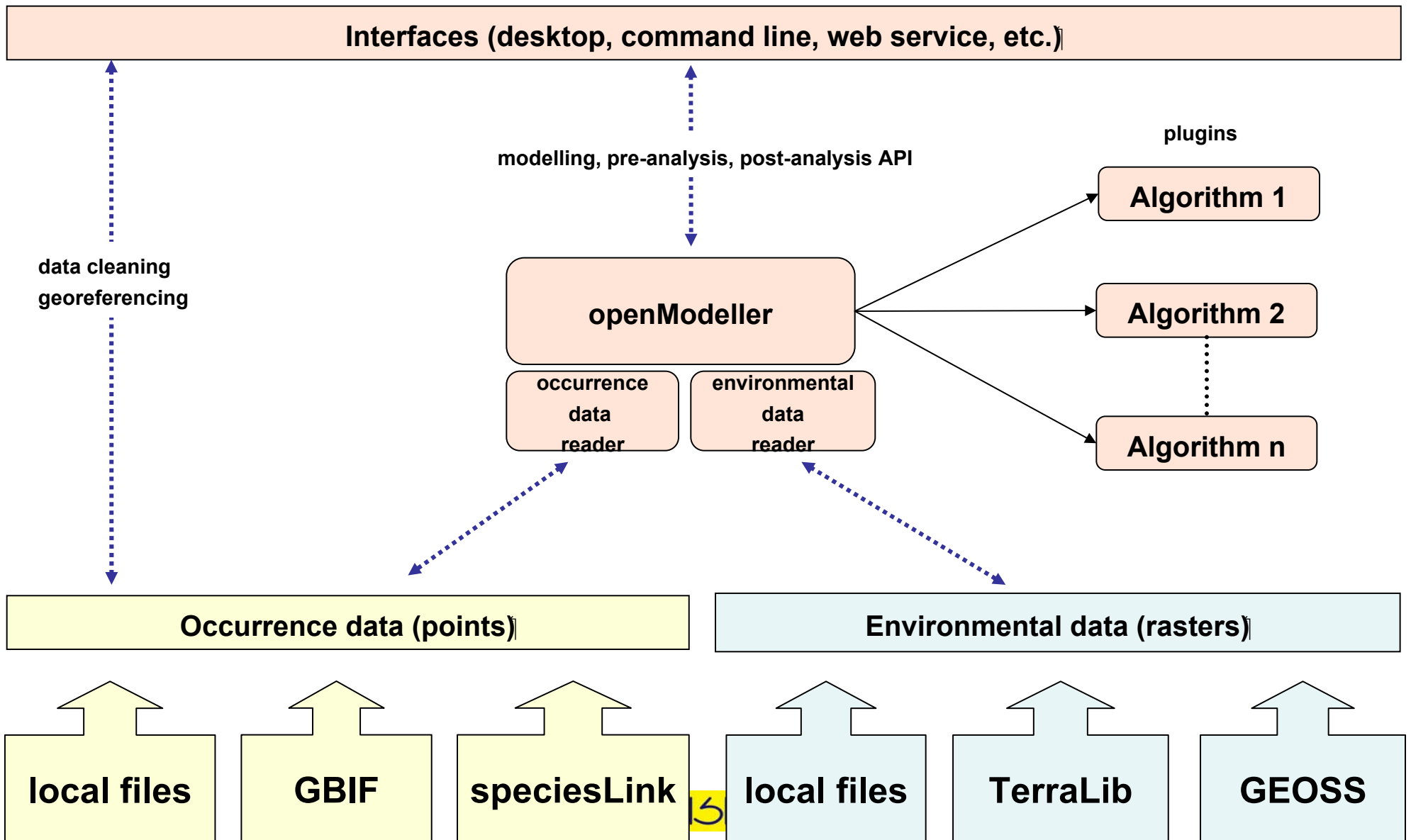


openModeller basic functions

- Work with multiple algorithms.
- Allow different interfaces on top of the same library (desktop, command line, web)
- Support different input formats (for both points and rasters).
- Platform independent (runs on GNU/Linux, Mac OSX and Windows).



openModeller architecture



Reference Center on Environmental Information



Brazilian not-for-profit, non-governmental organization that aims to contribute towards a more sustainable use of Brazil's biodiversity through dissemination of high quality information.



www.cria.org.br

The *speciesLink* network (splink.cria.org.br)

english

o projeto



species Link



UEPG

USP

speciesLink é um sistema distribuído de Informação que integra em tempo real, dados primários de coleções científicas. O sistema foi desenvolvido graças ao apoio das instituições: FAPESP, GBIF, JRS Foundation e CRIA.

novidades

159 coleções e sub-coleções
2,909,295 registros on-line
1,299,714 georeferenciados
260,132 nomes diferentes de espécies
22 oct 2008 - 03:05

indicadores

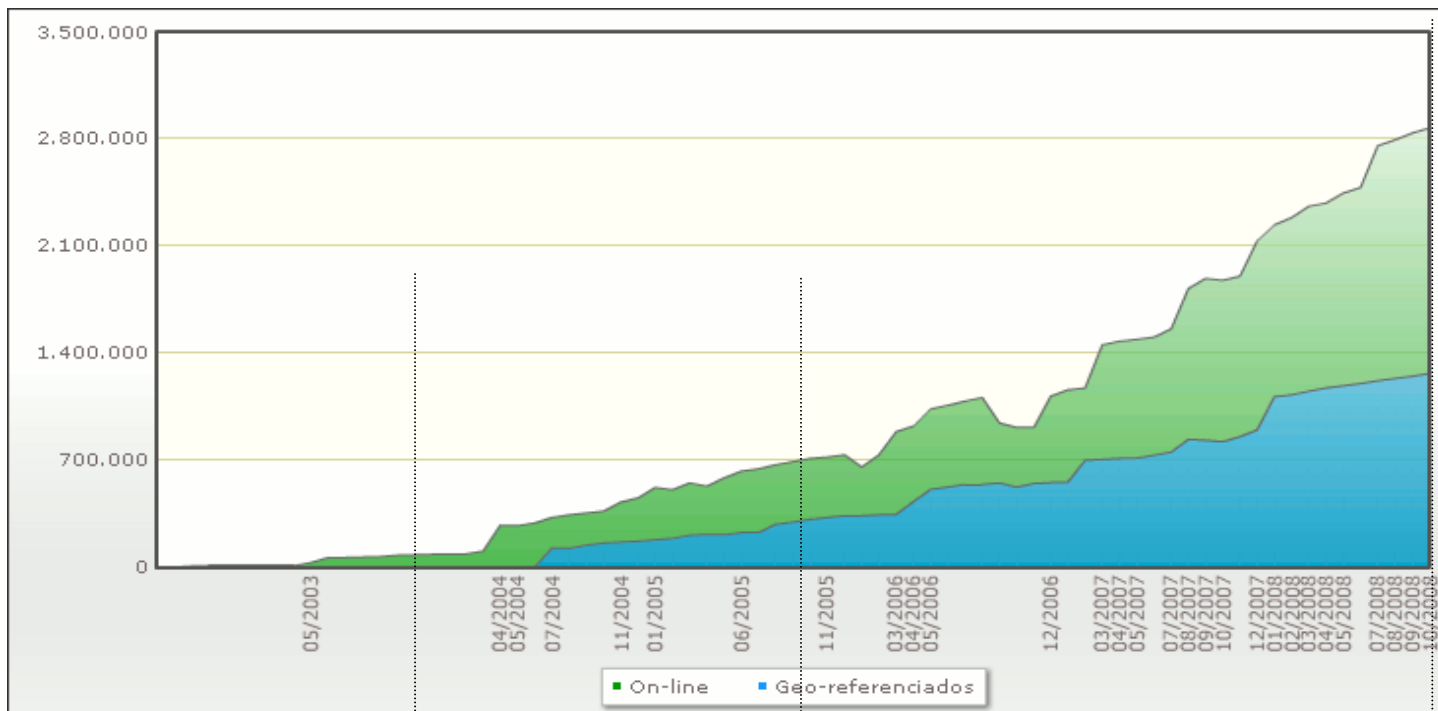


dados e ferramentas



POLI USP

speciesLink evolution



45% georef

FAPESP Nov 2005

~700k records

~40 collections

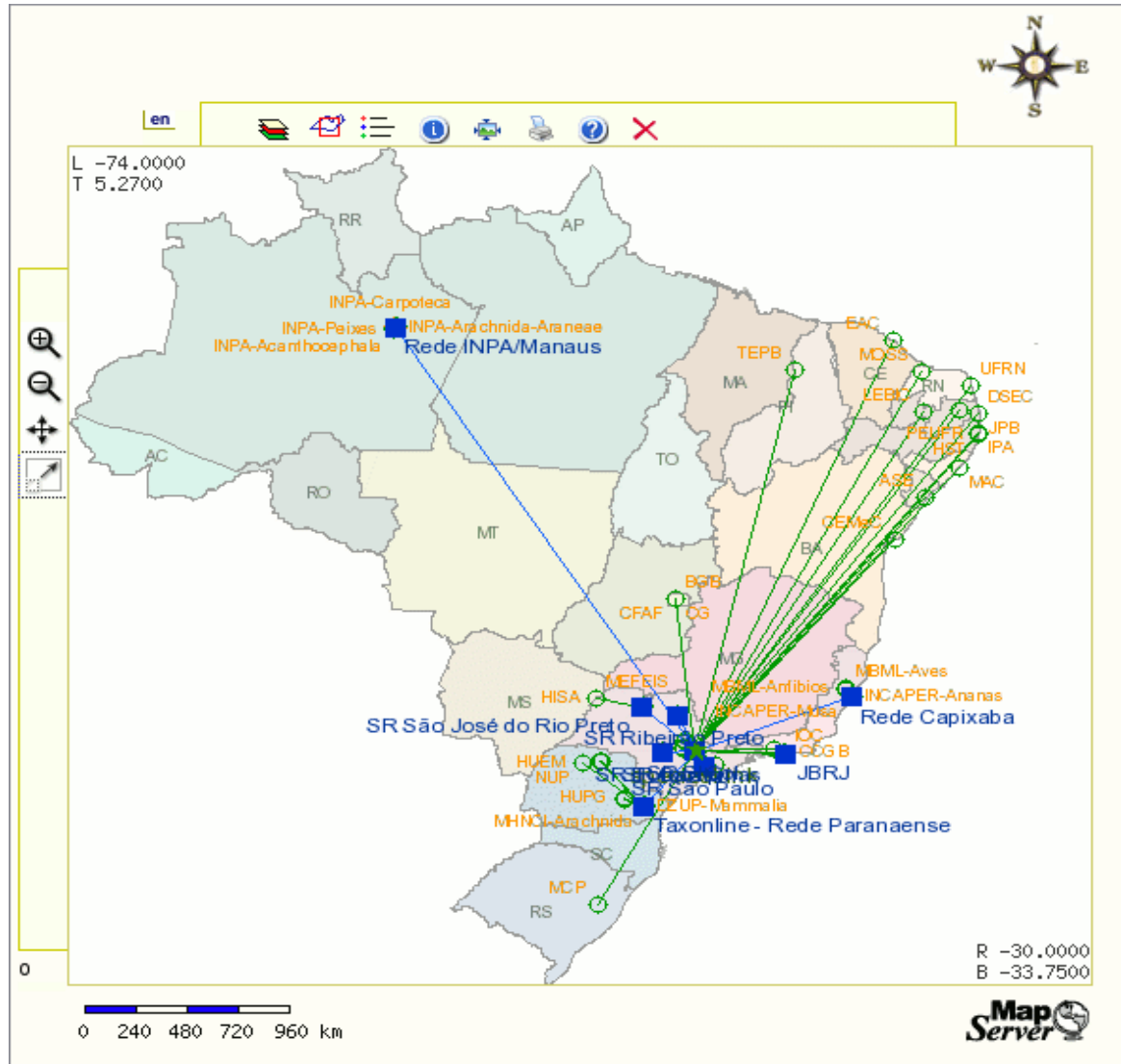
Out 2008

~3M records

~ 160 collections and subcollections



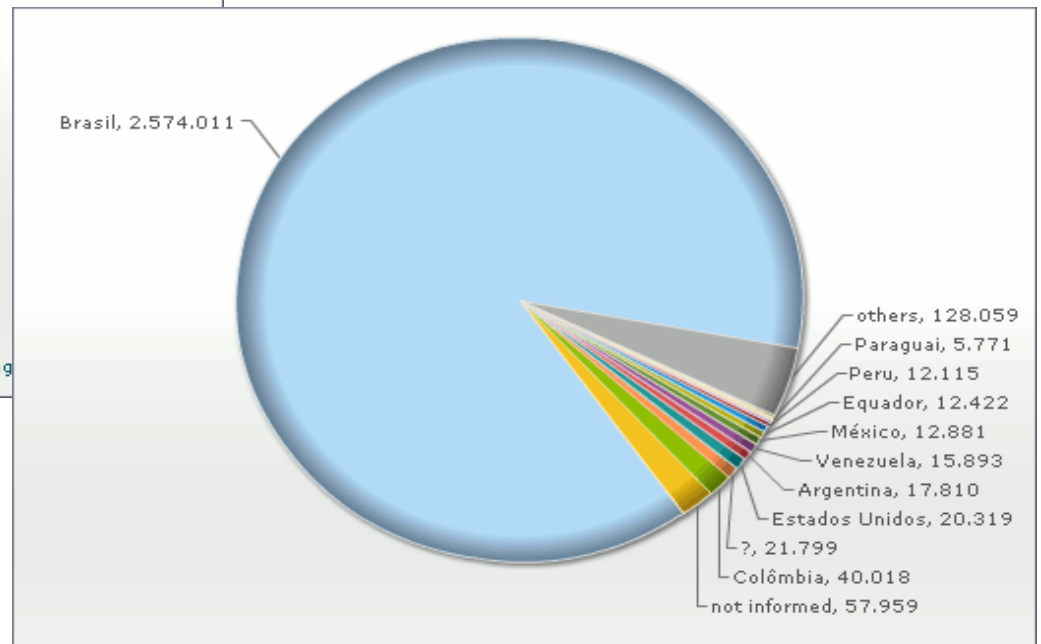
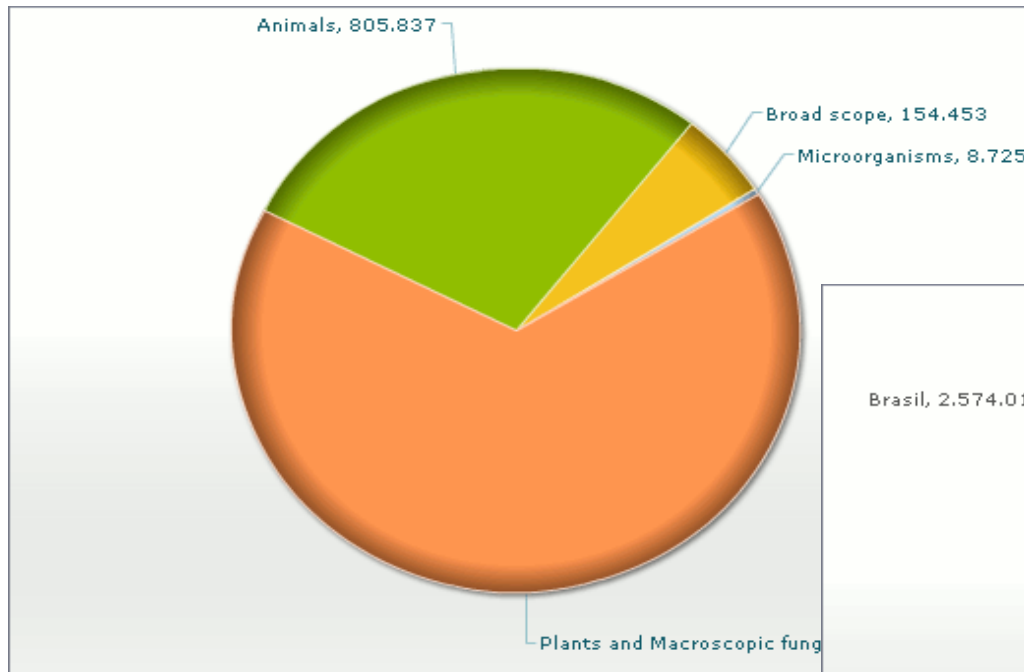
speciesLink regional servers



+ coleções de países amazônicos

+ repatriação de dados NYBG, MOBOT, ...

speciesLink indicators



66% Plants

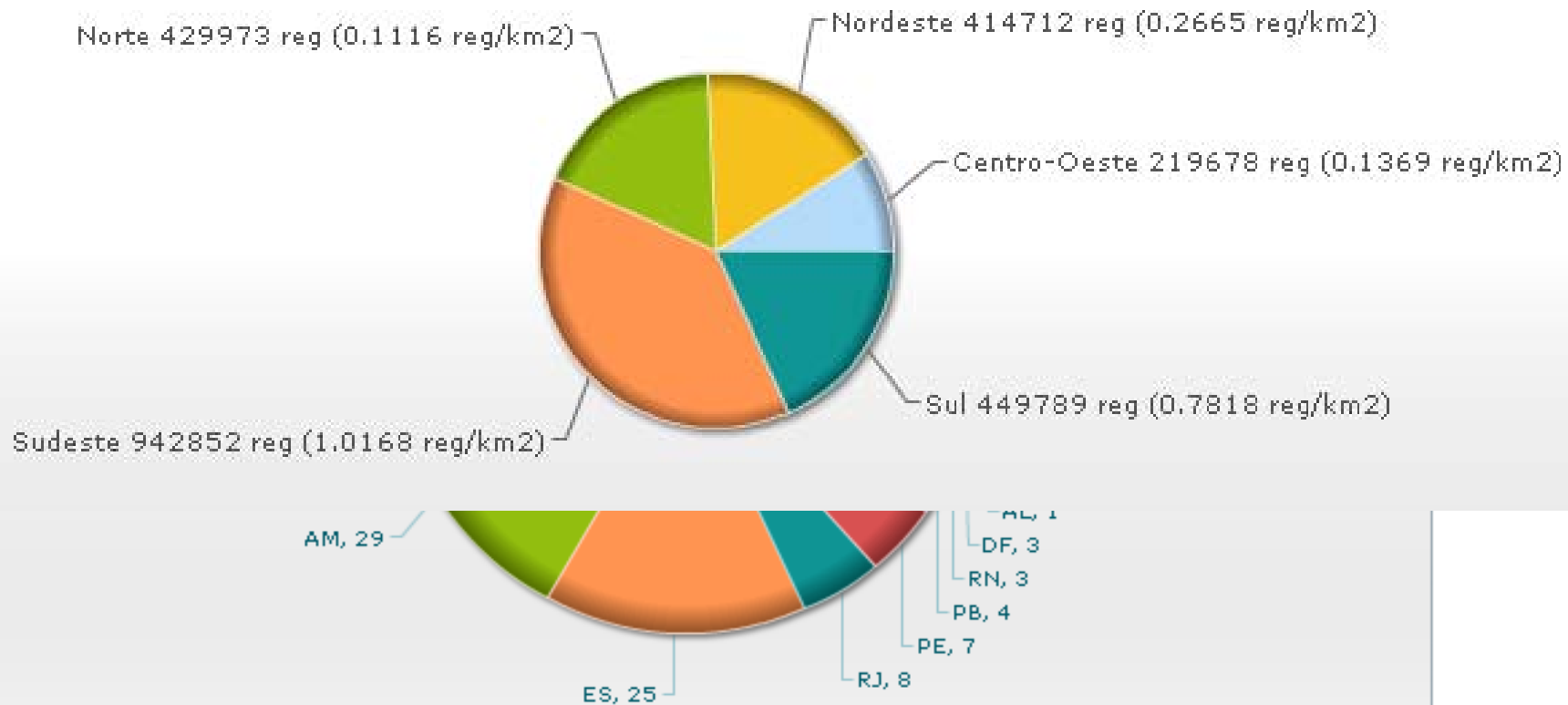
28% Animals

5% "broad scope"

1% microorganisms

85% from Brazilian collections

15% from foreign collections



Esse indicador mostra o número de coleções e sub-coleções conectadas a rede speciesLink por estado brasileiro.

XML

ASE - Herbário da Universidade Federal de Sergipe

Universidade Federal de Sergipe

[Como citar e Condições de uso](#)

spLink	Instituição	Coleção	Catálogo	Nome científico	Forma de registro	Reino	Filo	Classe	Ordem	Família	Gênero	Espécies	Subespécie	Autor do nome
	UFS	ASE	5402	Tabebuia impetiginosa	S	Plantae	-	-	-	Bignoniaceae	Tabebuia	impetiginosa	-	(Mart. ex DC.)

**CPMA - Coleção de plantas medicinais e aromáticas**

Universidade Estadual de Campinas

[Como citar e Condições de uso](#)

spLink	Instituição	Coleção	Catálogo	Nome científico	Forma de registro	Reino	Filo	Classe	Ordem	Família	Gênero	Espécies	Subespécie	Autor do nome
	CPQBA	CPMA	1076	Tabebuia impetiginosa	S	Plantae	-	-	-	Bignoniaceae	Tabebuia	impetiginosa	-	(Mart. ex DC.)

EAC - Herbário Prisco Bezerra

Universidade Federal do Ceará

[Como citar e Condições de uso](#)

spLink	Instituição	Coleção	Catálogo	Nome científico	Forma de registro	Reino	Filo	Classe	Ordem	Família	Gênero	Espécies	Subespécie	Autor do nome
	UFC	EAC	41404	Tabebuia impetiginosa	-	Plantae	-	-	-	Bignoniaceae	Tabebuia	impetiginosa	-	(Mart. ex DC.)
	UFC	EAC	41405	Tabebuia impetiginosa	-	Plantae	-	-	-	Bignoniaceae	Tabebuia	impetiginosa	-	(Mart. ex DC.)

ESA - Herbário da Escola Superior de Agricultura Luiz de Queiroz

Escola Superior de Agricultura Luiz de Queiroz

[Como citar e Condições de uso](#)

spLink	Instituição	Coleção	Catálogo	Nome científico	Forma de registro	Reino	Filo	Classe	Ordem	Família	Gênero	Espécies	Subespécie	Autor do nome
	ESALQ	ESA	36432	Tabebuia impetiginosa	S	Plantae	Magnoliophyta	-	-	Bignoniaceae	Tabebuia	impetiginosa	-	(M.)
	ESALQ	ESA	36436	Tabebuia impetiginosa	S	Plantae	Magnoliophyta	-	-	Bignoniaceae	Tabebuia	impetiginosa	-	(M.)
	ESALQ	ESA	51663	Tabebuia impetiginosa	S	Plantae	Magnoliophyta	-	-	Bignoniaceae	Tabebuia	impetiginosa	-	-



Occurrence search

Please add filters and click Search to perform a search for occurrences records. Specify more filters to narrow your search to get more accurate results for the species, area or time period that is of interest.

Add search filter

Scientific name

Enter a scientific name and click Add Filter.

This filter will return any records that have a matching name given for the identification of the organism, regardless of how the organism is classified.

Your current search

Filters can be removed by clicking on the minus signs.

Classification includes Species:
Tabebuia impetiginosa

This search matches 646 occurrence records.

Actions

- View:** [Matching records as table](#) [Matching records on map](#)
- Specify:** [Data providers to be included in search](#) [Datasets to be included in search](#) [Countries to be included in search](#)
- Download:** [Spreadsheet of results](#) [Darwin core \(maximum 100,000\)](#) [Google Earth \(maximum 50,000\)](#) [Species in results](#)
- Create:** [New! Niche Model](#)

Sample results

Scientific Name	Dataset	Institution Code	Collection Code	Catalogue Number	Basis of Record	Date	Coordinates	Country	
<i>Tabebuia impetiginosa</i>	herbario	Museo Nacional de Costa Rica (MNCR)	Herbario Nacional de Costa Rica (CR)	51483	Specimen			Costa Rica	View
<i>Tabebuia impetiginosa</i>	herbario	Museo Nacional de Costa Rica (MNCR)	Herbario Nacional de Costa Rica (CR)	103192	Specimen			Costa Rica	View
<i>Tabebuia impetiginosa</i>	herbario	Museo Nacional de Costa Rica (MNCR)	Herbario Nacional de Costa Rica (CR)	109019	Specimen			Costa Rica	View
<i>Tabebuia impetiginosa</i>	SysTax	BAYRT	Okologisch Botanischer Garten Bayreuth	591273	Unknown				View
<i>Tabebuia</i>	Missouri Botanical	MO	Tropicos	479917	Specimen	05/01/1985		Brazil	View

Create Niche Model

This tool provides an integration between [openModeller](#) and the occurrence point data available within the portal. Its main purpose is to demonstrate the integration of GBIF-mediated occurrence data with other applications like modelling tools.

Please be aware that the significance of the modelling result varies greatly with the selection of meaningful parameters.



This integration takes the occurrence points provided by your search, submits these points and the layers selected to openModeller. openModeller then generates a probability distribution using the [Envelope Score Algorithm](#).

Please select environment to model within

▸ [Land](#)

▸ [Ocean](#)

Land Layers - provided by [Worldclim](#)



[Select all](#) | [Deselect all](#)

- | | |
|--|--|
| <input checked="" type="checkbox"/> Annual mean temperature | <input checked="" type="checkbox"/> Precipitation of wettest month |
| <input checked="" type="checkbox"/> Mean diurnal range | <input checked="" type="checkbox"/> Precipitation of driest month |
| <input checked="" type="checkbox"/> Isothermality | <input checked="" type="checkbox"/> Precipitation seasonality |
| <input checked="" type="checkbox"/> Temperature seasonality | <input checked="" type="checkbox"/> Precipitation of wettest quarter |
| <input checked="" type="checkbox"/> Maximum temperature of warmest month | <input checked="" type="checkbox"/> Precipitation of driest quarter |
| <input checked="" type="checkbox"/> Minimum temperature of coldest month | <input checked="" type="checkbox"/> Precipitation of warmest quarter |
| <input checked="" type="checkbox"/> Temperature annual range | <input checked="" type="checkbox"/> Precipitation of coldest quarter |
| <input checked="" type="checkbox"/> Mean temperature of wettest quarter | |
| <input checked="" type="checkbox"/> Mean temperature of driest quarter | |
| <input checked="" type="checkbox"/> Mean temperature of warmest quarter | |
| <input checked="" type="checkbox"/> Mean temperature of coldest quarter | |
| <input checked="" type="checkbox"/> Annual precipitation | |

Recently viewed

- [Occurrence search - Classification includes Species: Tabebuia impetiginosa](#)
- [Classification of Species: Tabebuia impetiginosa Standl.](#)
- [Classification](#)

Actions

- View:** [Matching records as table](#) [Matching records on map](#)
- Specify:** [Data providers to be included in search](#) [Datasets to be included in search](#) [Countries to be included in search](#)
- Download:** [Spreadsheet of results](#) [Darwin core \(maximum 100,000\)](#) [Google Earth \(maximum 50,000\)](#)
[Species in results](#)
- Create:** *New!* [Niche Model](#)

Your current search

Classification includes Species: *Tabebuia impetiginosa*

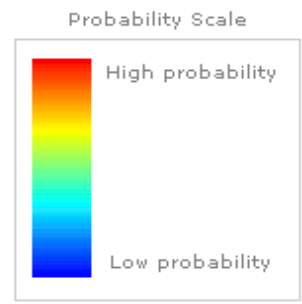
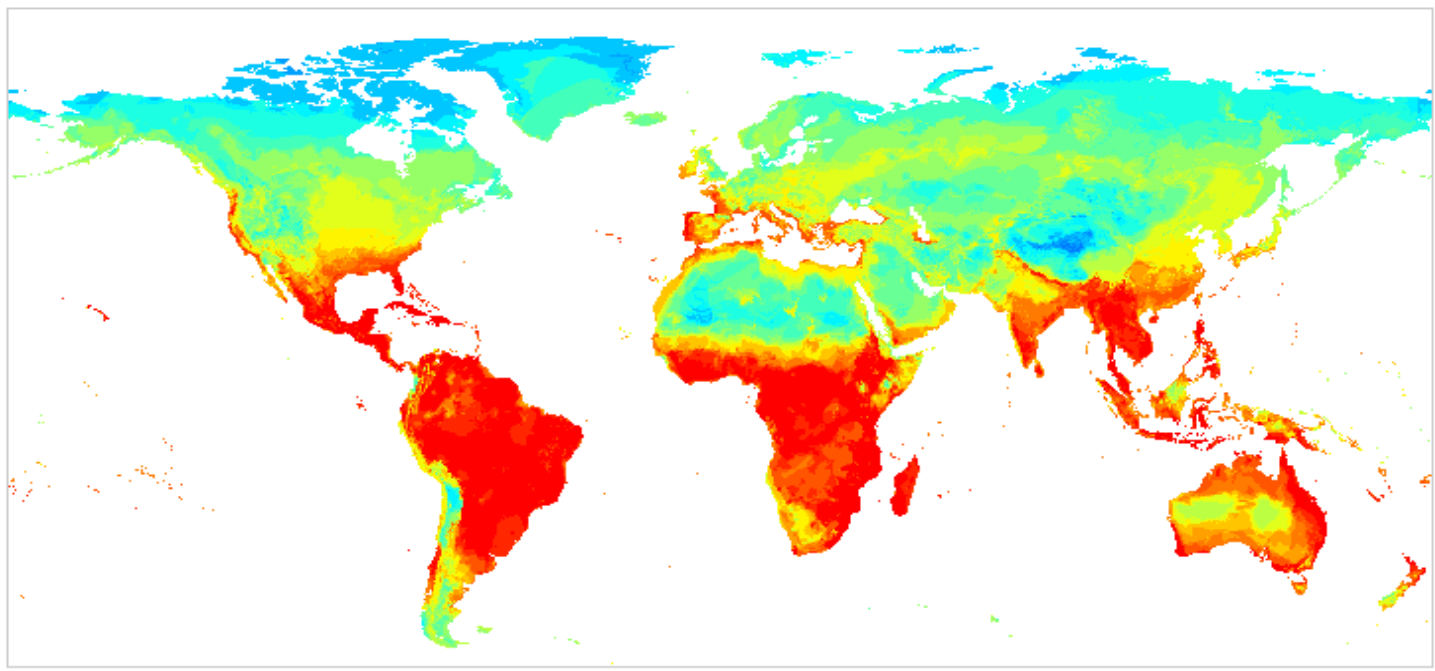
 [Change your current search](#)

Your Generated Niche Model

This tool provides an integration between openModeller and the occurrence point data available within the portal. Its main purpose is to demonstrate the integration of GBIF-mediated occurrence data with other applications like modelling tools.

Please be aware that the significance of the modelling result varies greatly with the selection of meaningful parameters.

This integration takes the occurrence points provided by your search, submits these points and the layers selected to openModeller. openModeller then generates a probability distribution using the **Envelope Score Algorithm**.



Good models depend on data
(quantity and quality),
tools, and knowledge



inpe

- Incluir 4 a 5 slides



Escola Politécnica da USP (Poli-USP)



- Dept. of Computing and Digital Systems Engineering
 - Agricultural Automation Laboratory (LAA)
 - Architecture, algorithms and software engineering
 - Architecture and High Performance Computing (LAHPC)
 - Cluster and grid computing
 - Adaptive Techniques Laboratory (LTA)
 - Adaptive algorithms
- Team:
 - 6 faculty members
 - 5 PhD, 1 MSc and 10 undergraduate students
 - Scholarships from FAPESP, CNPq, CAPES, FDTE



Contributions of Poli-USP to oM

- New architecture with focus on interoperability
- New algorithms: modeling and analysis
- High performance computing
- Software development process
 - Unit tests
 - Process documentation
- Case studies
 - Environmental services: pollinators



Architecture

- Facilitate integration with other applications
 - Make modeling available to other applications
 - Access data from distributed sources
- Examples
 - GBIF – Global Biodiversity Information Facility
 - IABIN – InterAmerican Biodiversity Information Network
- Paradigm: Service-oriented architecture
 - Break oM into a set of services
 - that can invoke (and be invoked by) other applications
 - Cluster: available as a high performance modeling service



Web interface to the service-oriented oM

open Modeller

login:
password:

Overview
New User
Using the System
Contact

Developers
Documents
Publications
Sources
Bugs
Downloads

Home

Species Occurrence Environmental Layers Probability Distribution

Welcome to the openModeller project home page!

openModeller aims to provide a flexible, user friendly, cross-platform environment where the entire process of conducting a fundamental niche modeling experiment can be carried out. The software includes facilities for reading species occurrence and environmental data, selection of environmental layers on which the model should be based, creating a fundamental niche model and projecting the model into an environmental scenario. A number of algorithms are provided as plugins, including GARP, Climate Space Model, Bioclimatic Envelopes, Support Vector Machines and others.

The project is currently being developed by the Centro de Referência em Informação Ambiental (CRIA), Escola Politécnica da USP (Poli), and Instituto Nacional de Pesquisas Espaciais (INPE) as an open-source initiative. It is funded by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), the Incofish project, and by individuals that have generously contributed their time. Previous collaborators include the BIOWorld project (University of Reading), the University of Kansas Natural History Museum & Biodiversity Research Center (KU), and other individual participants.

openModeller is a fundamental niche modelling library, providing a uniform method for modelling distribution patterns using a variety of modelling algorithms.

New Algorithms

- Importance of having more algorithms
 - Performance: computational and biological (functional)
- Modeling algorithms
 - GARP
 - AdaptGARP: based on adaptive techniques
 - P-GARP: parallel version for cluster execution
 - MaxEnt: maximizes entropy
 - Environmental Distance: Mahalanobis distance
 - Neural Networks: learn from examples
- Analysis algorithms
 - Jackknife: layers selection

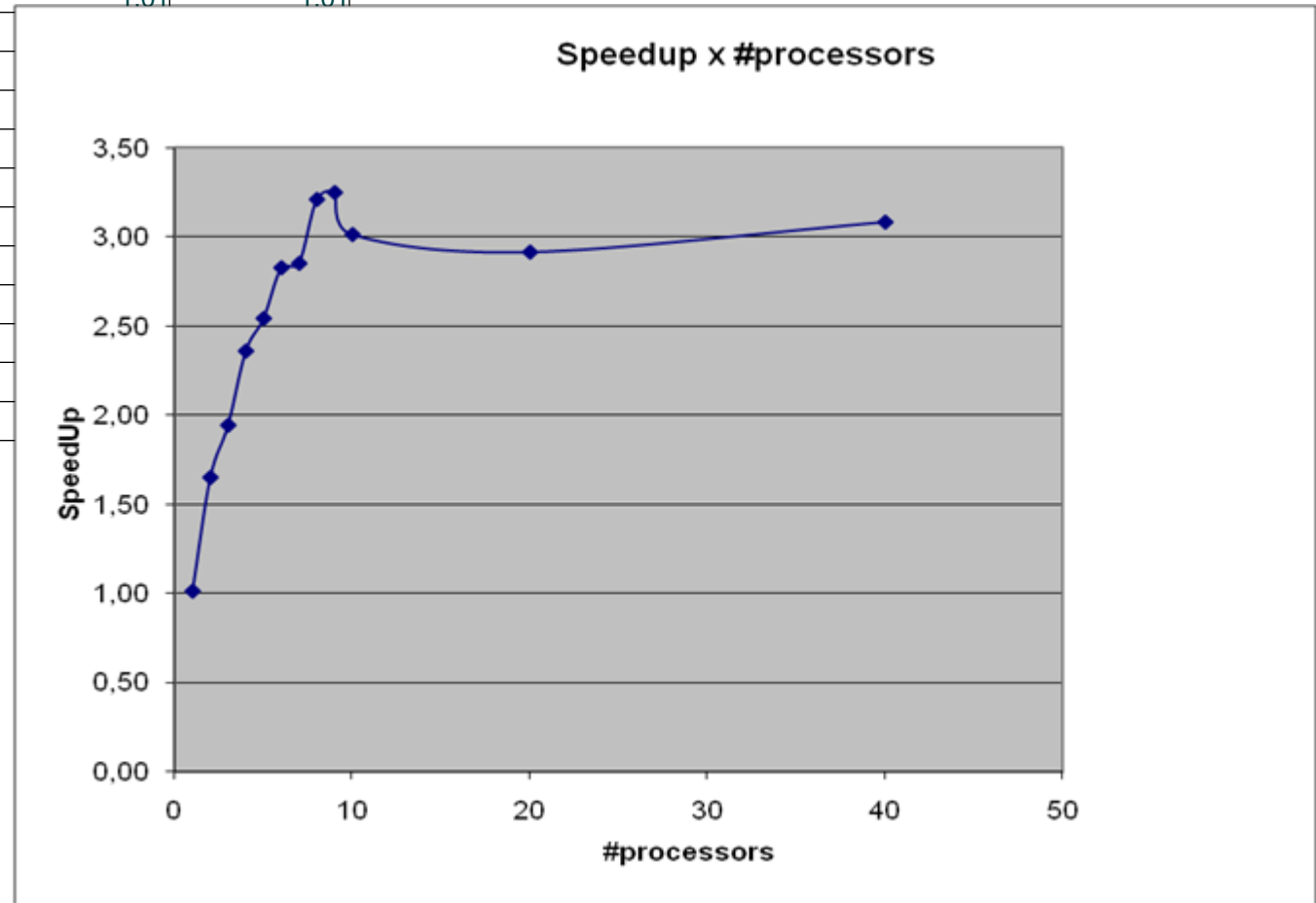


Performance improvement

T(GARP)	T(PGARP)	np	speedup	efficiency
0,68817	0,6817	1	1.01	1.01
0,68817	0,4176	2		
0,68817	0,3545	3		
0,68817	0,2919	4		
0,68817	0,2708	5		
0,68817	0,2436	6		
0,68817	0,2414	7		
0,68817	0,2144	8		
0,68817	0,2119	9		
0,68817	0,2285	10		
0,68817	0,2362	20		
0,68817	0,2233	40		

Obs.:

O *speedup* representa o número de vezes que um algoritmo paralelo é mais rápido do que um seqüência, considerando o melhor algoritmo seqüencial disponível.



High performance computing

- Why high performance computing?
 - Modeling and projection are computer intensive
 - Up to many days to complete on a desktop computer!
- Computer cluster infrastructure
 - SGI 88 cores: 11 nodes, 2 quad Intel Xeon 2.0 GHz and 8 GB ram memory;
 - Hosted at USP-CCE Datacenter
 - Dedicated to oM
- Parallel version of oM's algorithms
 - P-GARP, P-bestSubsets,
- Cluster power available on oM Desktop



Case study: ecosystem services – pollination

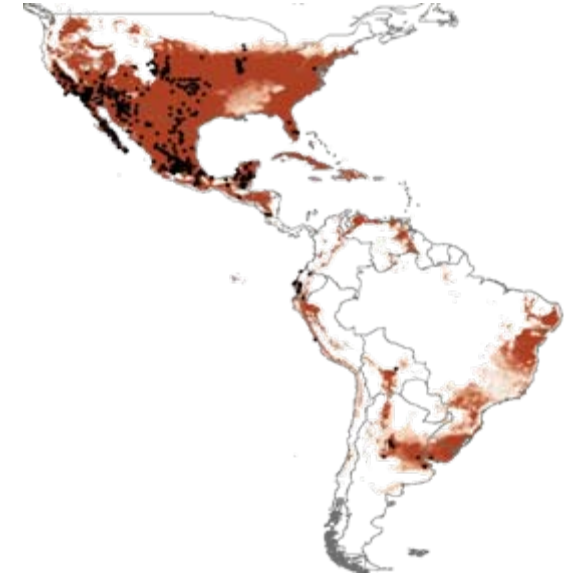
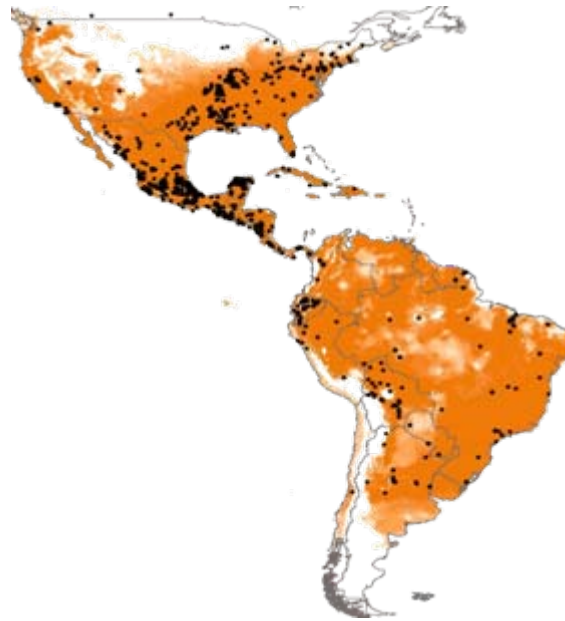
Peponapis spp x *Cucurbita* spp

- Tereza C.GIANNINI (Thesis project)
- *Peponapis* (Eucerini, Apidae)
 - bees specialized on flowers of the genus *Cucurbita*. Neotropical distribution with diversity center in Mexico.
- *Cucurbita* spp
 - pumpkins: 13 wild and 5 cultivated species. Diversity center in Mexico.
- Distributions patterns?



Case study

- Distribution of pollinators x commercial x native pumpkin



Peponapis
spp



P. pruinosa

Cultivated
pumpkins



C. pepo pepo

wild pumpkins



C. sp.



Future plans & Team (Poli-USP)

•Plans

- Web Interface to service-architecture, cluster and grid computing,
- more algorithms: parallel versions of Jackknife, MaxEnt, ANN
- Human resources on modelling: users and developers
- Case studies

•Team

- Profs: AM Saraiva, PLP Correa, LM Sato, JJ Neto, ET Midorikawa, RLA Rocha, E Murakami (UDESC)
- Graduate Students: FS Santana, F.Rodrigues, ESC Rodrigues, RL Stange, SL Stanzani, TC Giannini.
- Undergrad. students: Danilo, Mariana, Albert, Alex, Igor, Jefferson...



Thank you

<http://openmodeller.sf.net>

