4-2014

Biodiversity Informatics – Big Data for Biodiversity Conservation and Ecological Forecasting

Shaily Menon
Grand Valley State University, menons@gvsu.edu

Biodiversity Informatics

Big Data for Biodiversity Conservation and Ecological Forecasting

Shaily Menon
Biology Department and CLAS Deans Office

Grand Valley State University

BIODIVERSITY INFORMATICS

Definition

Information about biological diversity

Creating, integrating, analyzing, and understanding information

Applications

Ecological niche modeling, ecological forecasting

Primary Data

Historical and current species occurrence points

Big Data Source

GBIF GBIF.ORG
Free and open access to biodiversity data

Global Biodiversity Information Facility

Biodiversity Informatics
(data about biodiversity)

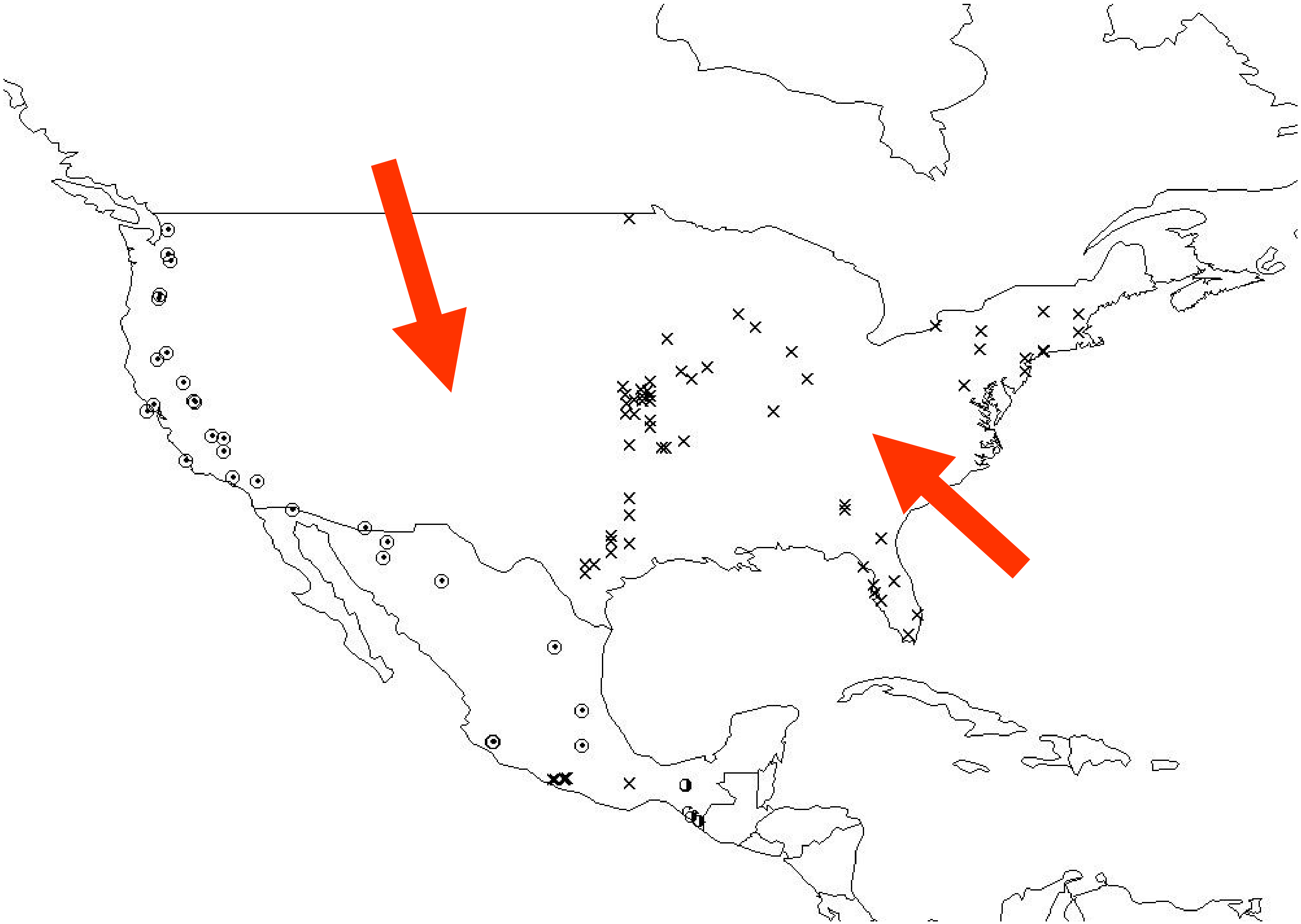
↓

Ecological Niche Modeling
(predict species distributions from data)

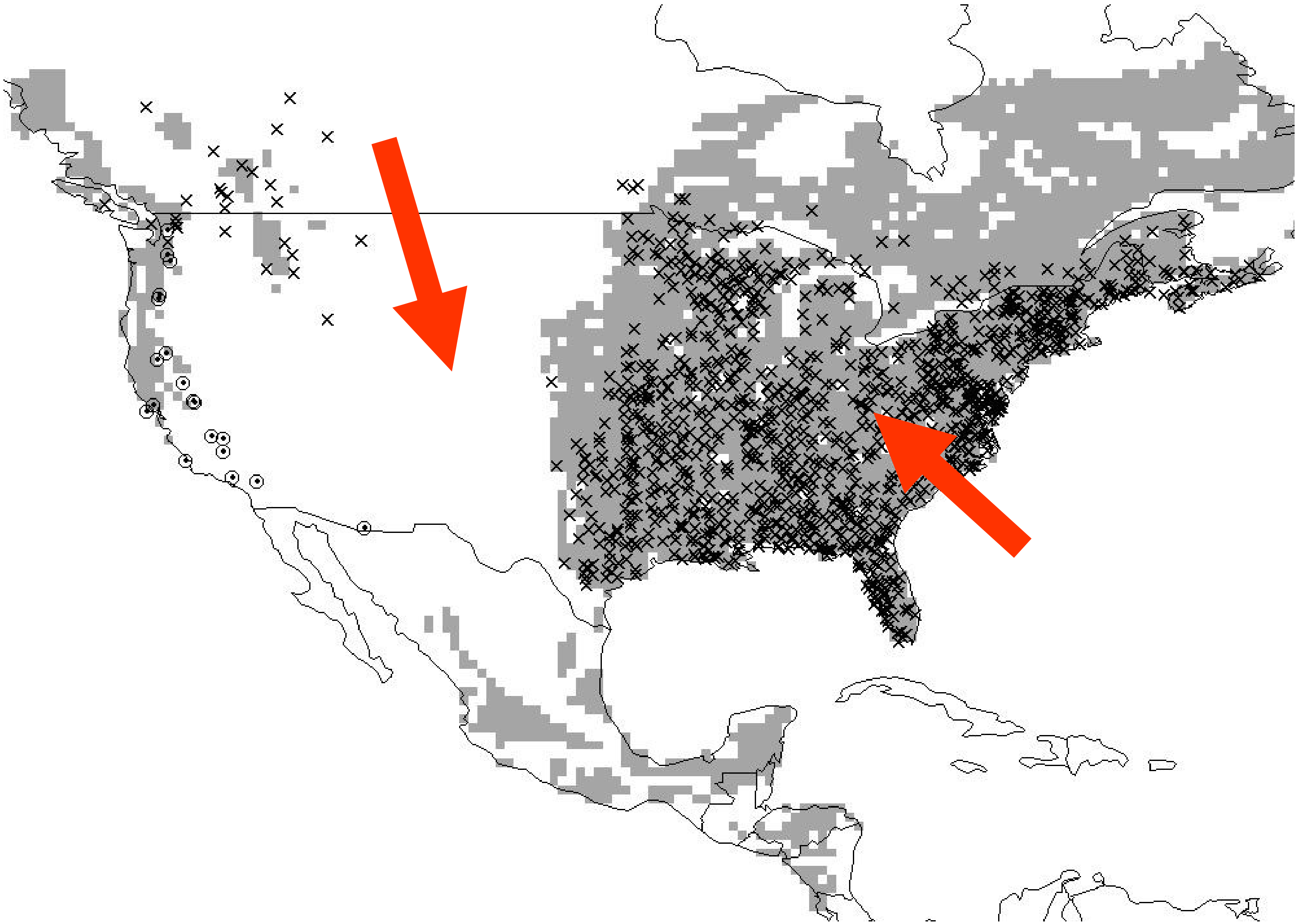
↓

Ecological Forecasting
(forecast effects of change)

Gaps in Knowledge of Biodiversity



Predictive Modeling
Allows inferring into gaps in knowledge






Examples of Applications

- Species discovery
- Predicting spread of invasive and disease species
- Forecasting effects of climate change

Predicting distributions of known and unknown reptile species in Madagascar

Christopher J. Raxworthy[1], Enrique Martínez-Meyer[2], Ned Horning[1], Ronald A. Nussbaum[3], Gregory E. Schneider[3], Miguel A. Ortega-Huerta[2] & A. Townsend Peterson[4]

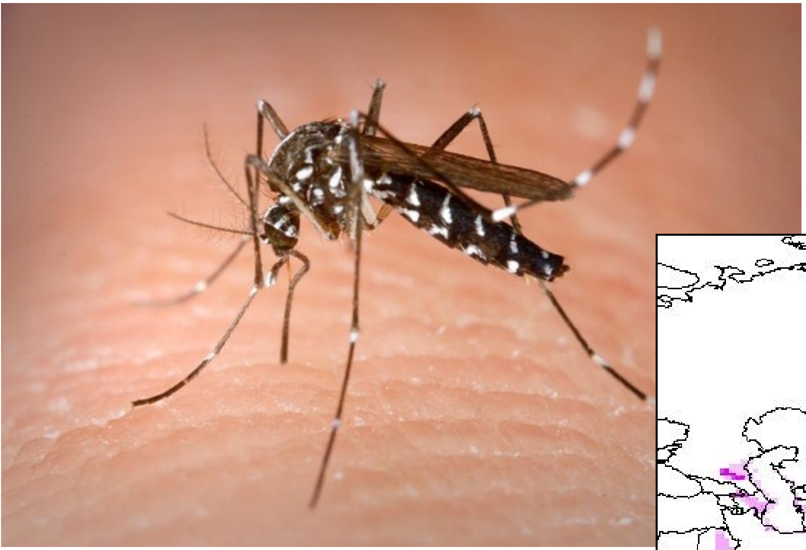
[1]*American Museum of Natural History, Central Park West at 79th Street, New York, New York 10024-5192, USA*
[2]*Instituto de Biología, Universidad Nacional Autónoma de México, Ciudad Universitaria, Mexico City 04510, Mexico*
[3]*Museum of Zoology, University of Michigan, Ann Arbor, Michigan 48109-1079, USA*
[4]*Natural History Museum & Biodiversity Research Center, The University of Kansas, Lawrence, Kansas 66045-2454, USA*

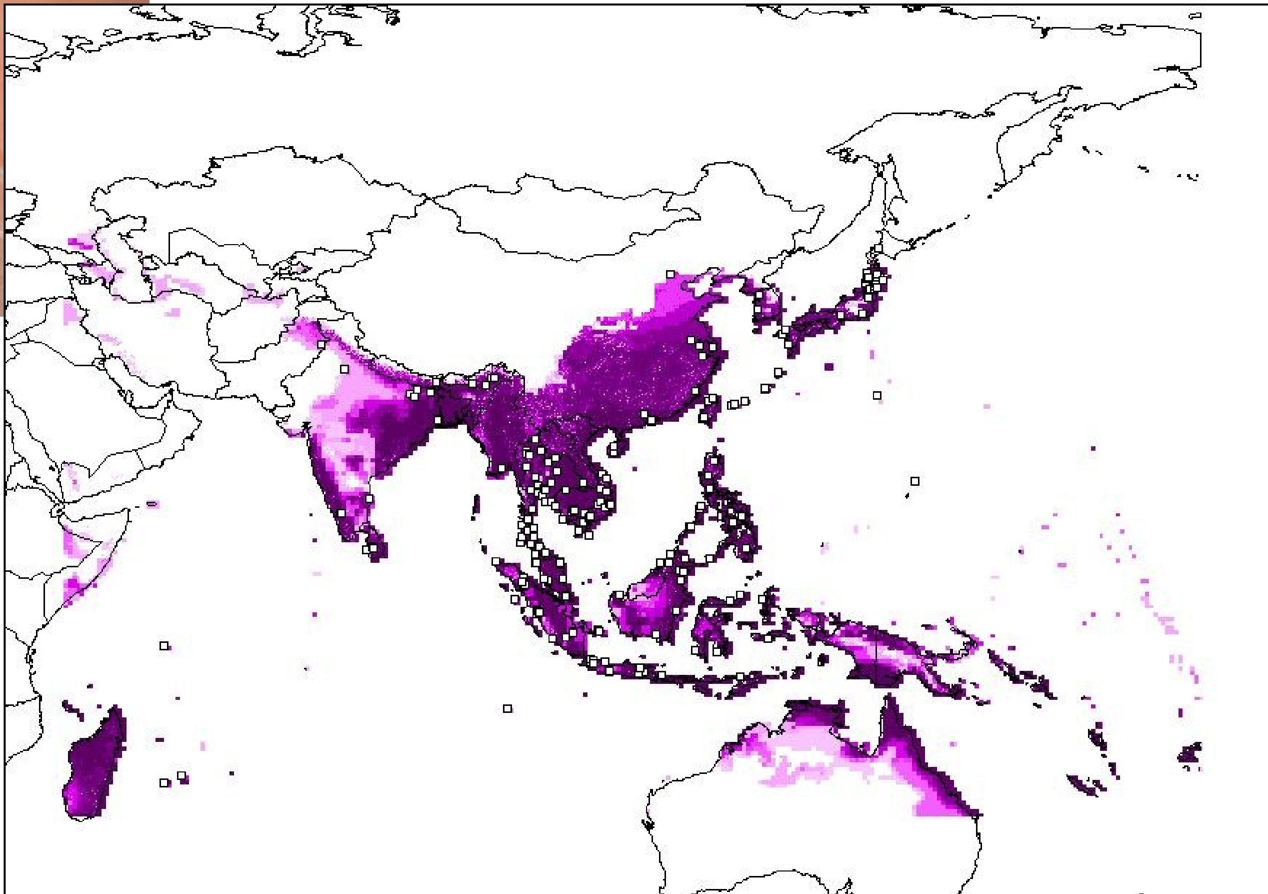
Despite the importance of tropical biodiversity[1], informative species distributional data are seldom available for biogeographical study or setting conservation priorities[2,3]. Modelling ecological niche distributions of species offers a potential solution[4–7]; however, the utility of old locality data from museums, and of more recent remotely sensed satellite data, remains poorly explored, especially for rapidly changing tropical landscapes. Using 29 modern data sets of environmental land coverage and 621 chameleon occurrence localities from Madagascar (historical and recent), here we demonstrate a significant ability of our niche models in predicting species distribution. At 11 recently inventoried sites, highest predictive success (85.1%) was obtained for models based only on modern occurrence data (74.7% and 82.8% predictive success, respectively, for pre-1978 and all data combined). Notably, these models also identified three intersecting areas of over-prediction that recently yielded seven chameleon species new to science. We conclude that ecological niche modelling using recent locality records and readily available environmental coverage data provides informative biogeographical data for poorly known tropical landscapes, and offers innovative potential for the discovery of unknown distributional areas and unknown species.

Sudden Oak Death
(*Phytophthora ramorum*)

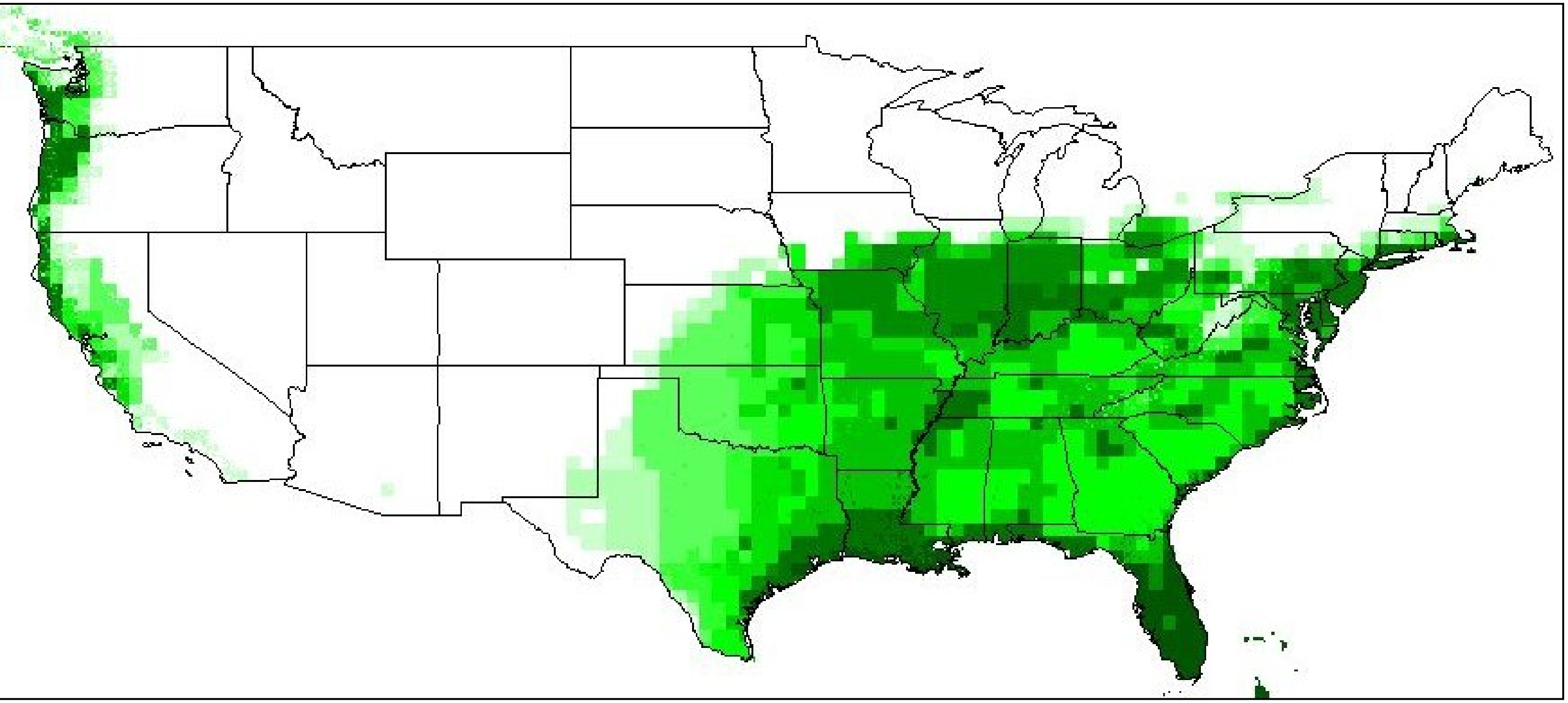
Aedes albopictus

Tiger mosquito – epidemiologically important vector for transmission of many viral pathogens including Yellow fever virus and dengue fever





Aedes albopictus in the USA



Predicted niche model

Current spread


8 January 2004
International weekly journal of science
nature
£10.00
www.nature.com/nature
Feeling the heat
Biodiversity losses due to global warming

Climate Change Forecasting

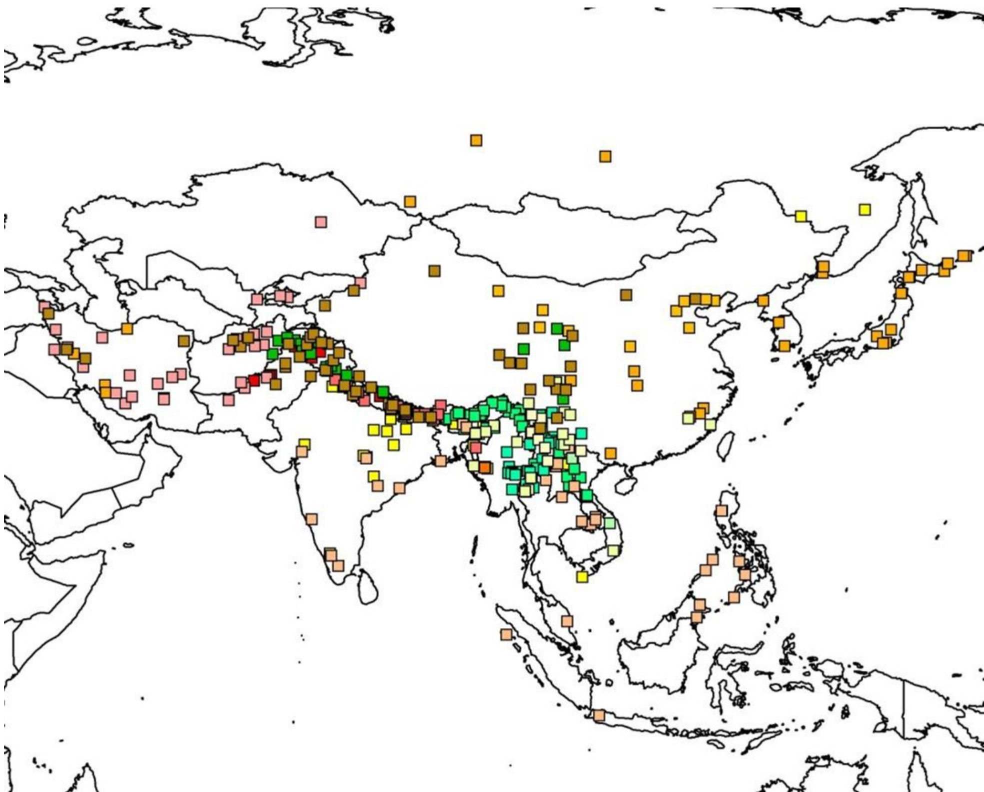
The Wilson Journal of Ornithology 120(4):692–699, 2008

PRELIMINARY ANALYSIS OF THE ECOLOGY AND GEOGRAPHY OF THE ASIAN NUTHATCHES (AVES: SITTIDAE)

SHAILY MENON,[1,5] ZAFAR-UL ISLAM,[2,4] JORGE SOBERÓN,[3] AND A. TOWNSEND PETERSON[3]

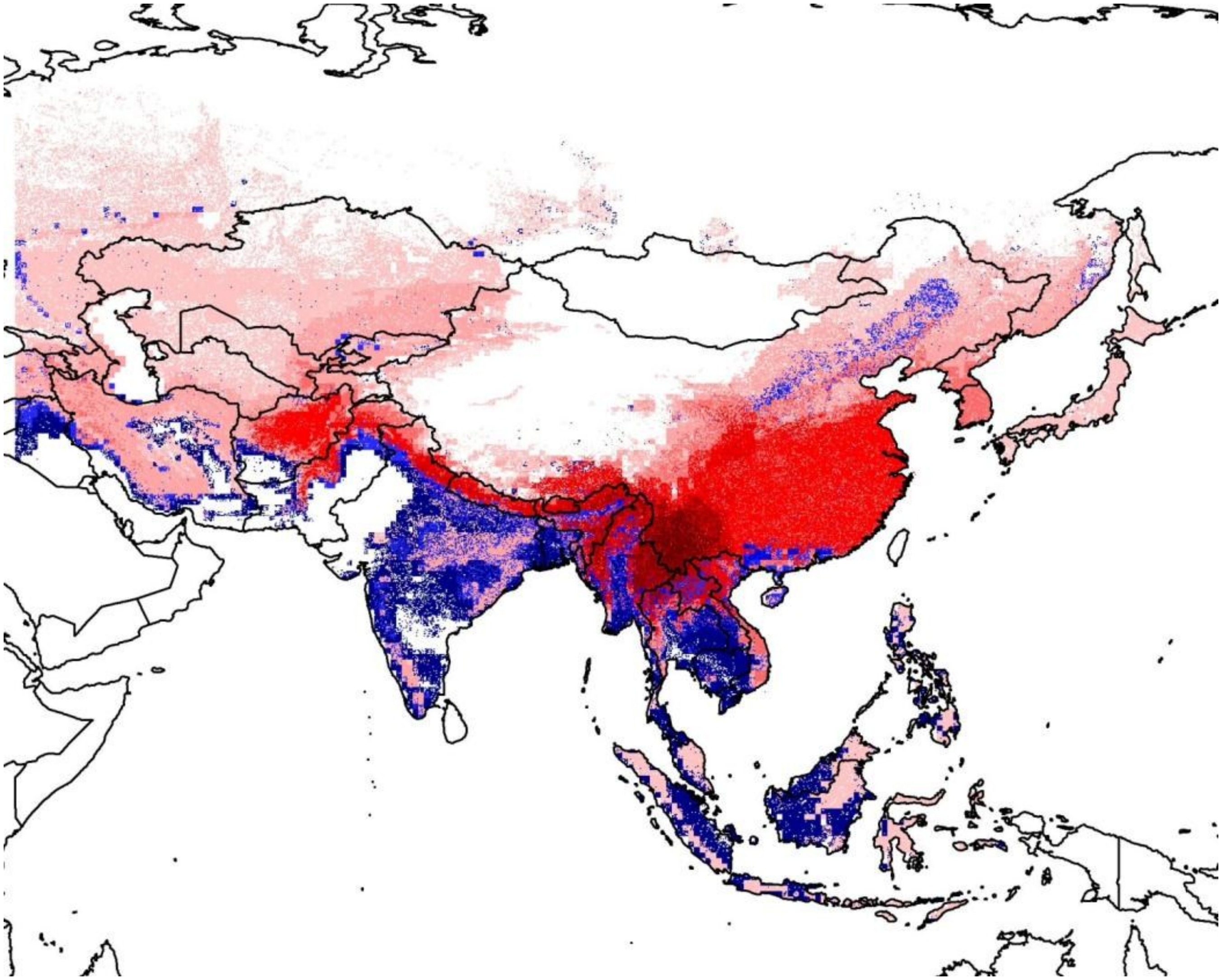
ABSTRACT.—We explored distributions of Asian nuthatch species in ecological and geographic space using ecological niche modeling based on occurrence data associated with specimens and observations. Nuthatches represent a well-defined clade occurring throughout the Northern Hemisphere, but are most diverse in southern Asia where 15 of the 24 species occur and where the lineage is believed to have evolved. Species richness was focused in a narrow east–west band corresponding to the forested parts of the Himalayas with a maximum number of nine species predicted present in these foci. The distributional predictions have a mid-elevation focus with highest species diversity between 1,000 and 2,000 m. Niche breadth and volume were positively related, but accumulation of distributional area (niche volume) decreased with additional environmental combinations (niche breadth). The extent of potential range filling, a measure of distributional disequilibrium, was connected with montane habit (R^2 = 0.422) indicating that montane situations limit the distributional potential of species.
Received 13 September 2007. Accepted 1 February 2008.

Asian nuthatches - occurrence points



Climate Change Forecasting

Eurasian nuthatches - Current diversity and predicted areas of high species loss (in blue)

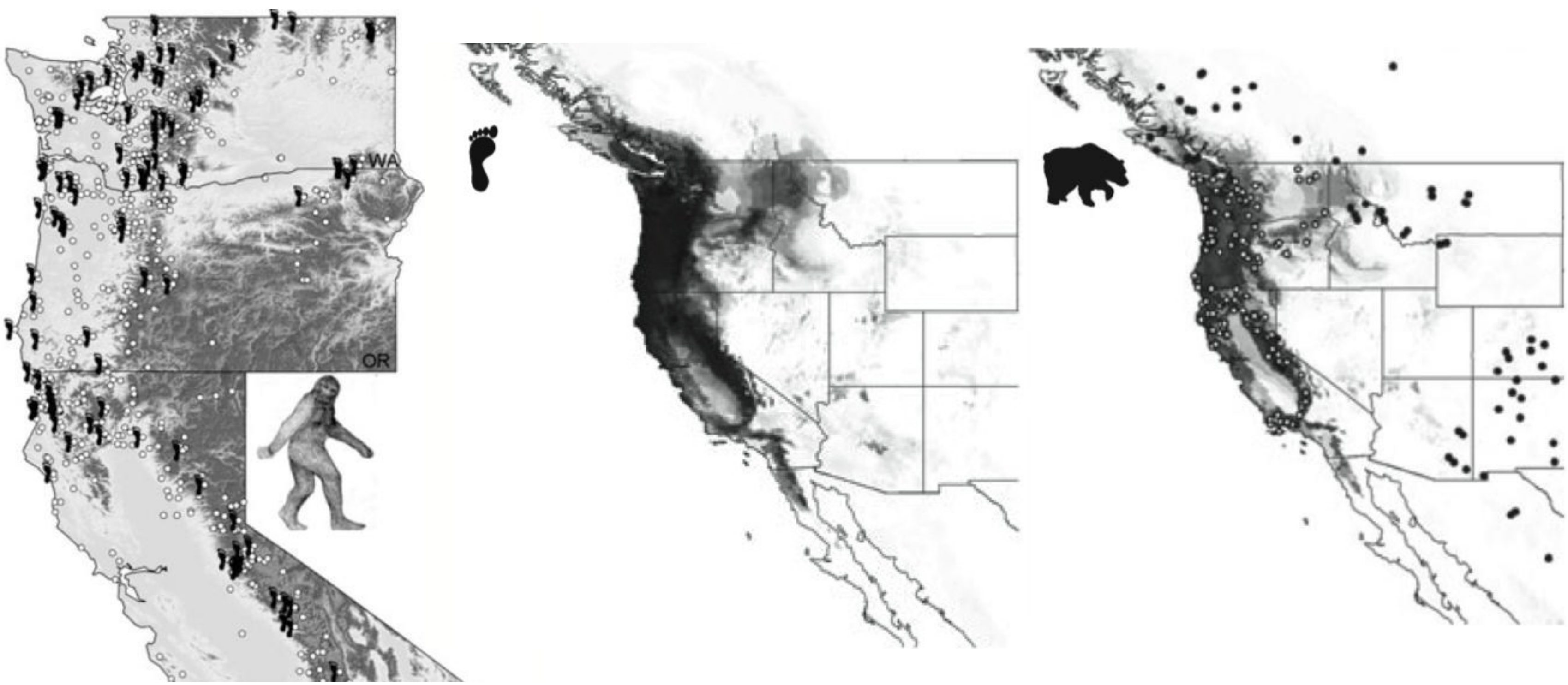


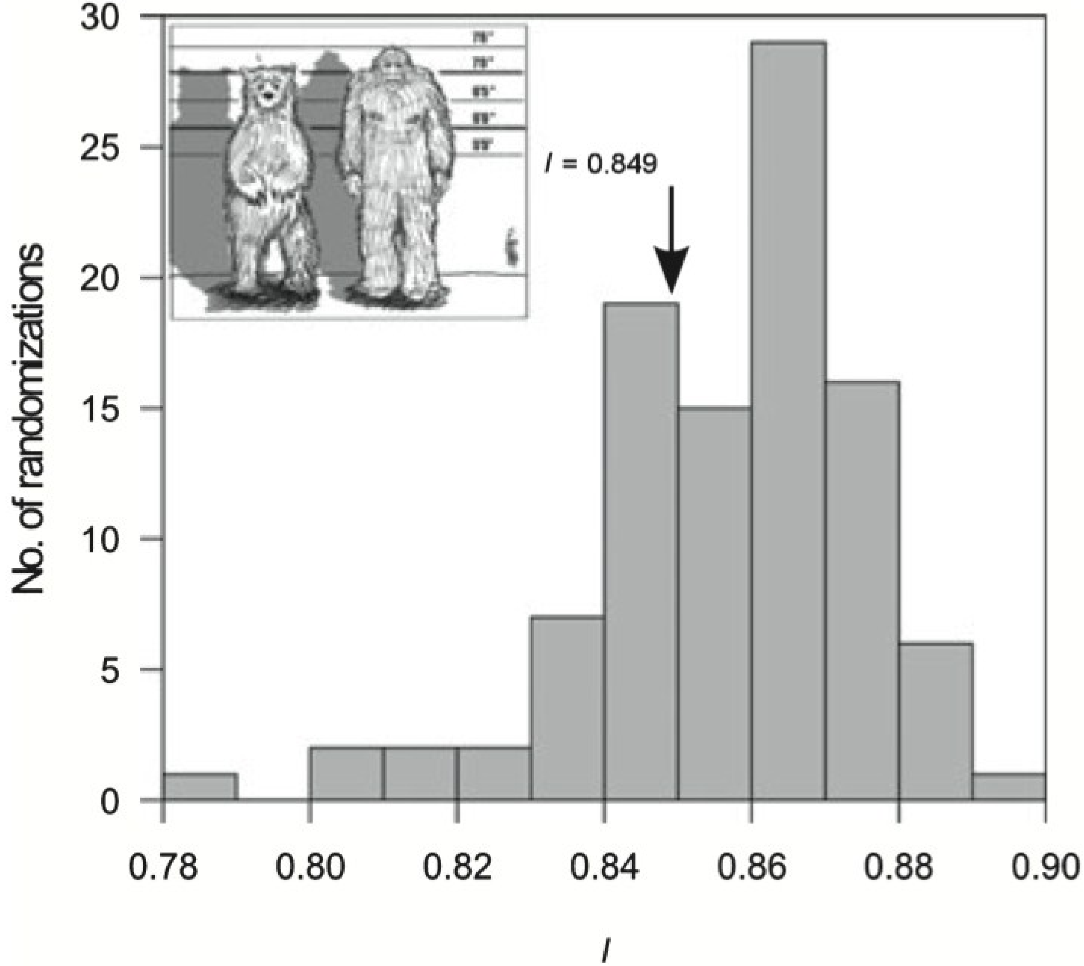
Journal of Biogeography (*J. Biogeogr.*) (2009) **36**, 1623–1627

GUEST EDITORIAL

Predicting the distribution of Sasquatch in western North America: anything goes with ecological niche modelling

J. D. Lozier[1]*, P. Aniello[2] and M. J. Hickerson[3]





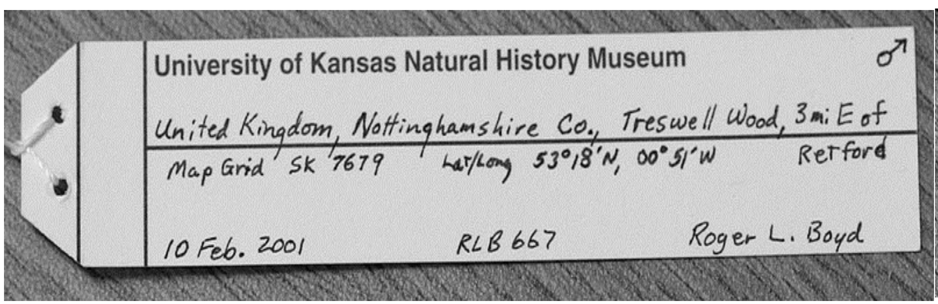
The observed value of I = 0.849 indeed indicates a high degree of overlap, and falls well within the null distribution generated from maxent runs for 100 randomizations of Bigfoot and black bear coordinates (Fig. 3; P < observed = 0.32). Thus, the two 'species' do not demonstrate significant niche differentiation with respect to the selected bioclimatic variables. Although it is possible that Sasquatch and *Ursus americanus* share such remarkably similar bioclimatic requirements, we nonetheless suspect that many Bigfoot sightings are, in fact, of black bears.

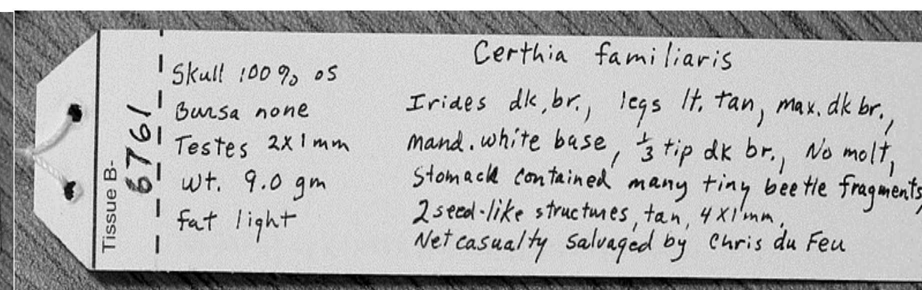
Sources of occurrence data?


Secondary Information
Secondary information
Primary data

Historical Data

Natural history museums
Literature (monographs, papers)

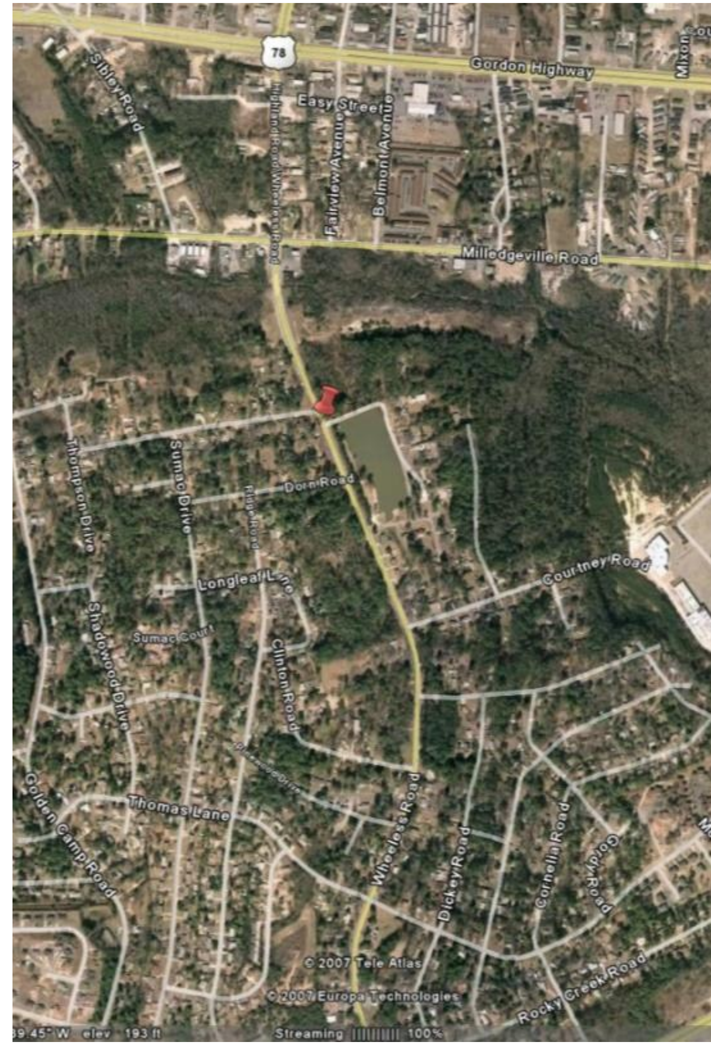




Georeferencing

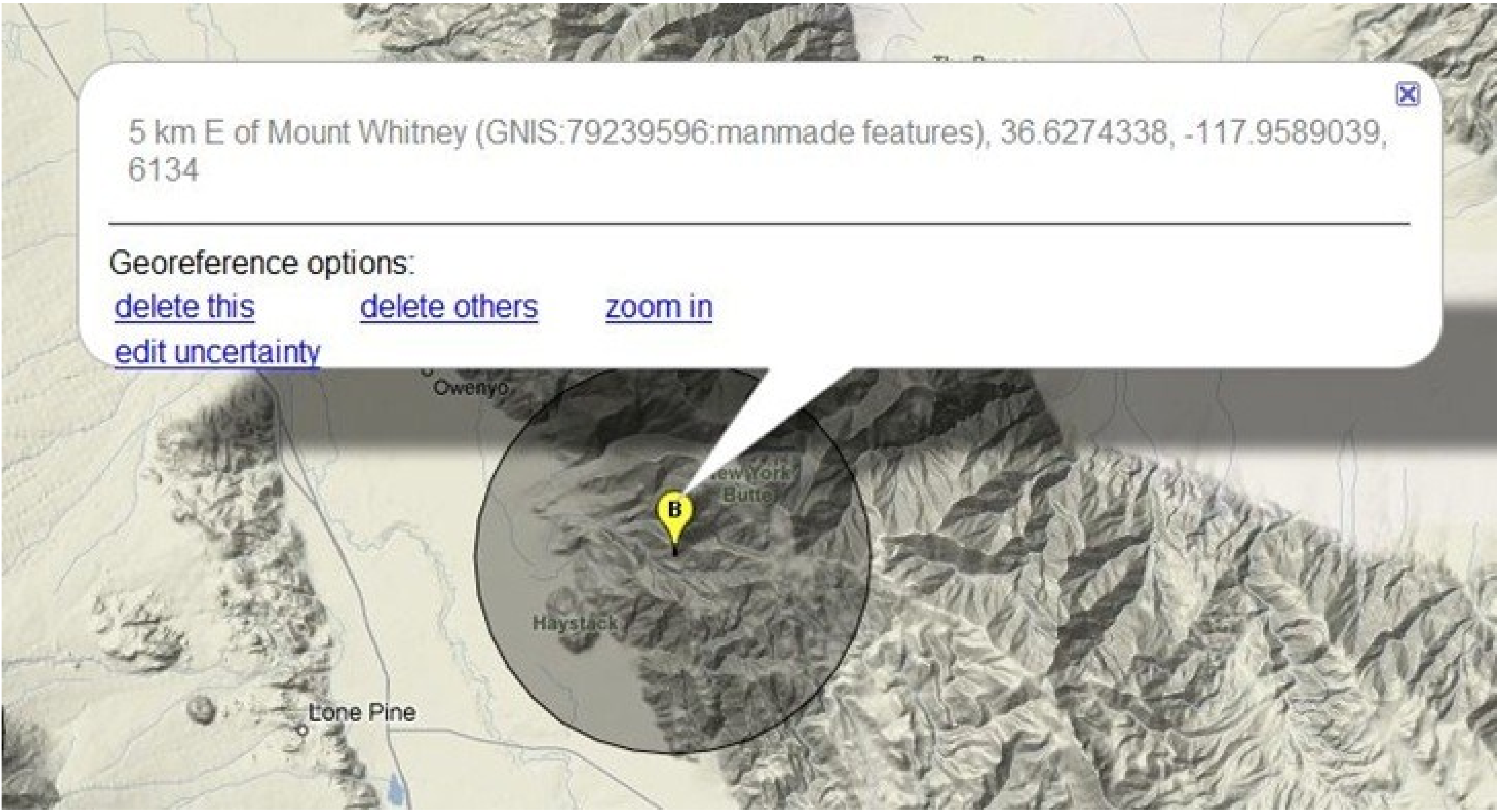
Assigning geographic coordinates to a locality description

USA, Georgia, Richmond Co., inside city limits of Augusta, Wheeless Rd., 0.5 mi S of Gordon Hwy

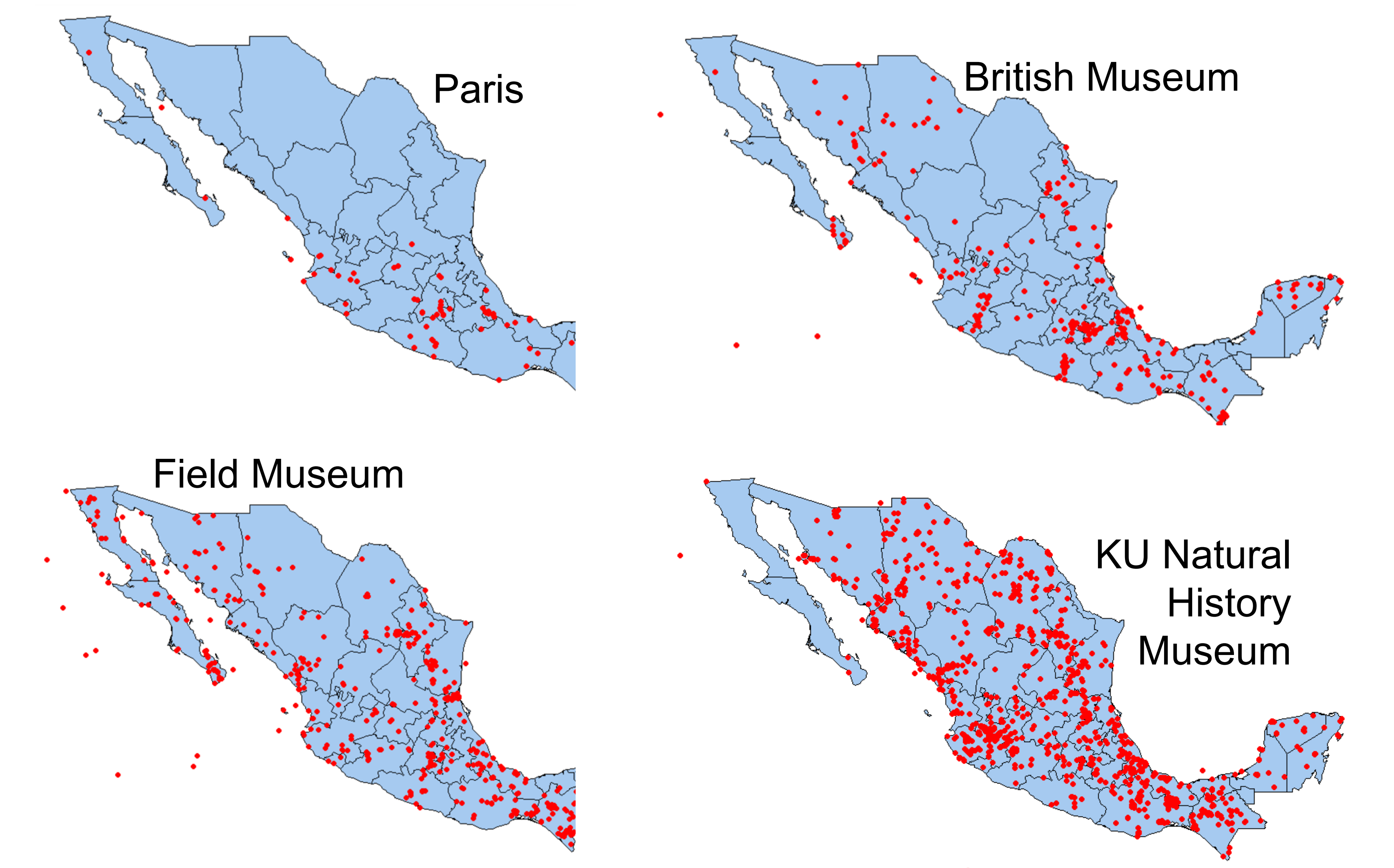


Locality georeferenced to Latitude: 33.444642 Longitude:-82.045296

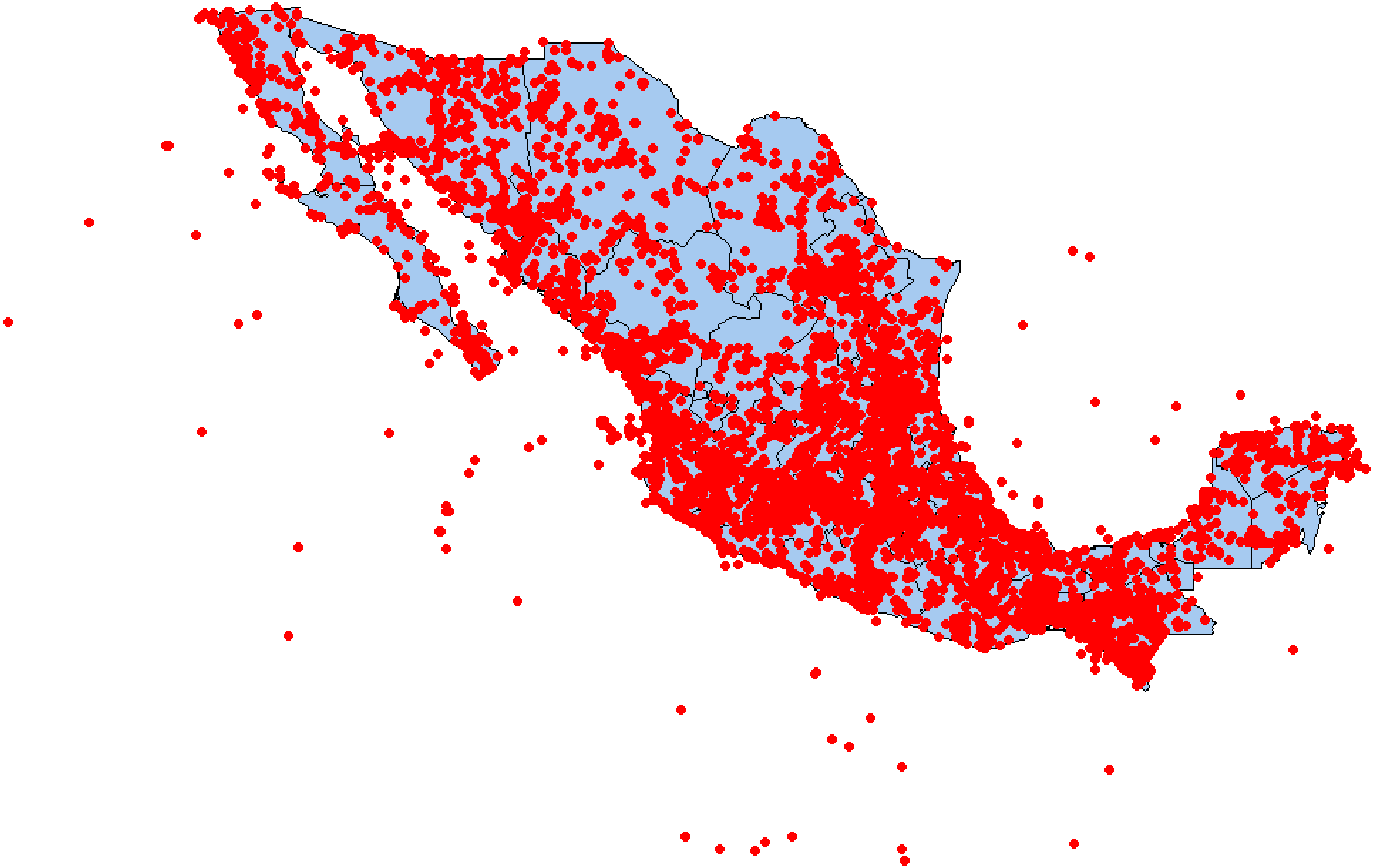
BioGeomancer

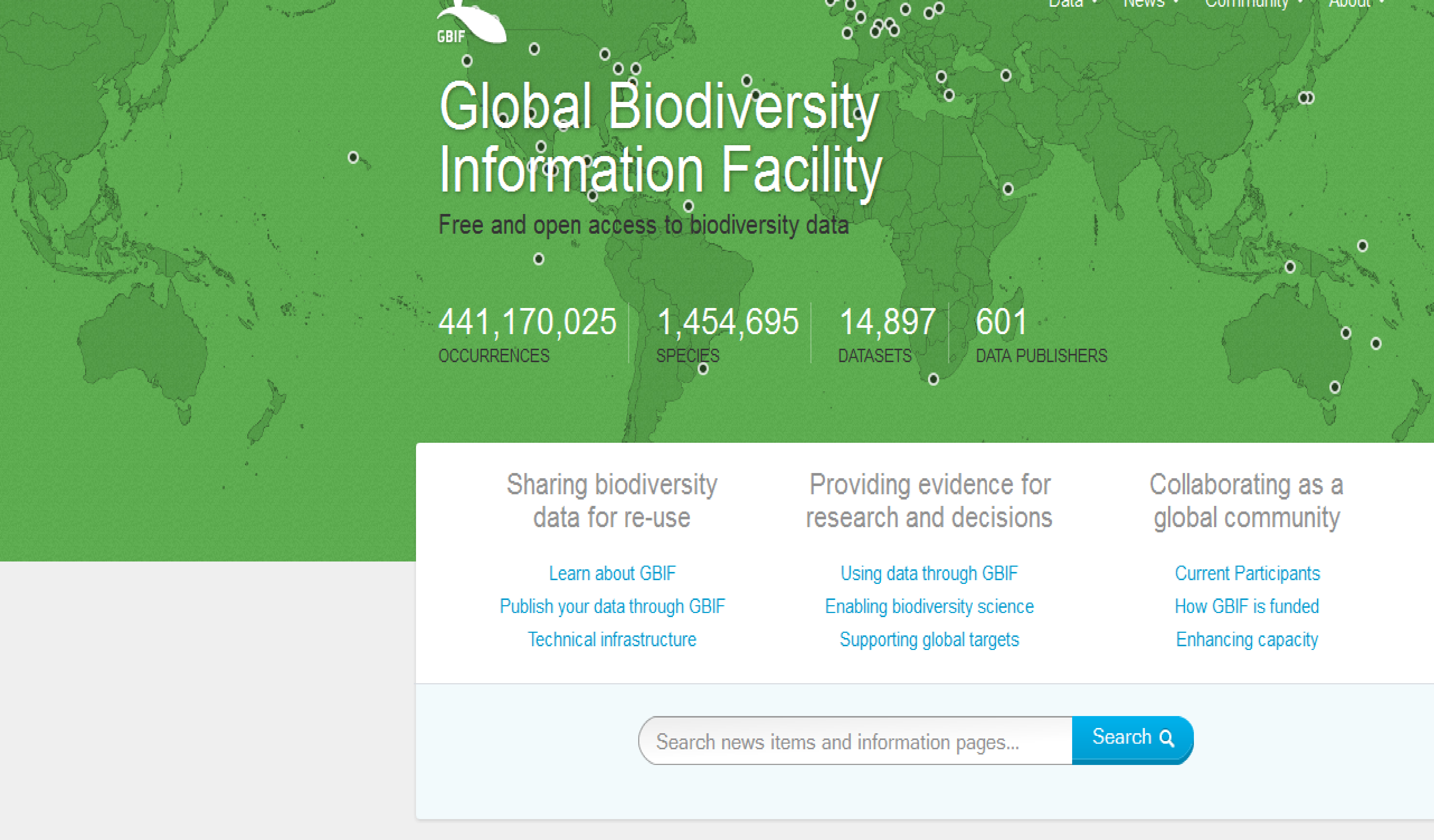


Example - Mexican Bird Collections



“World Museum” Mexican Birds



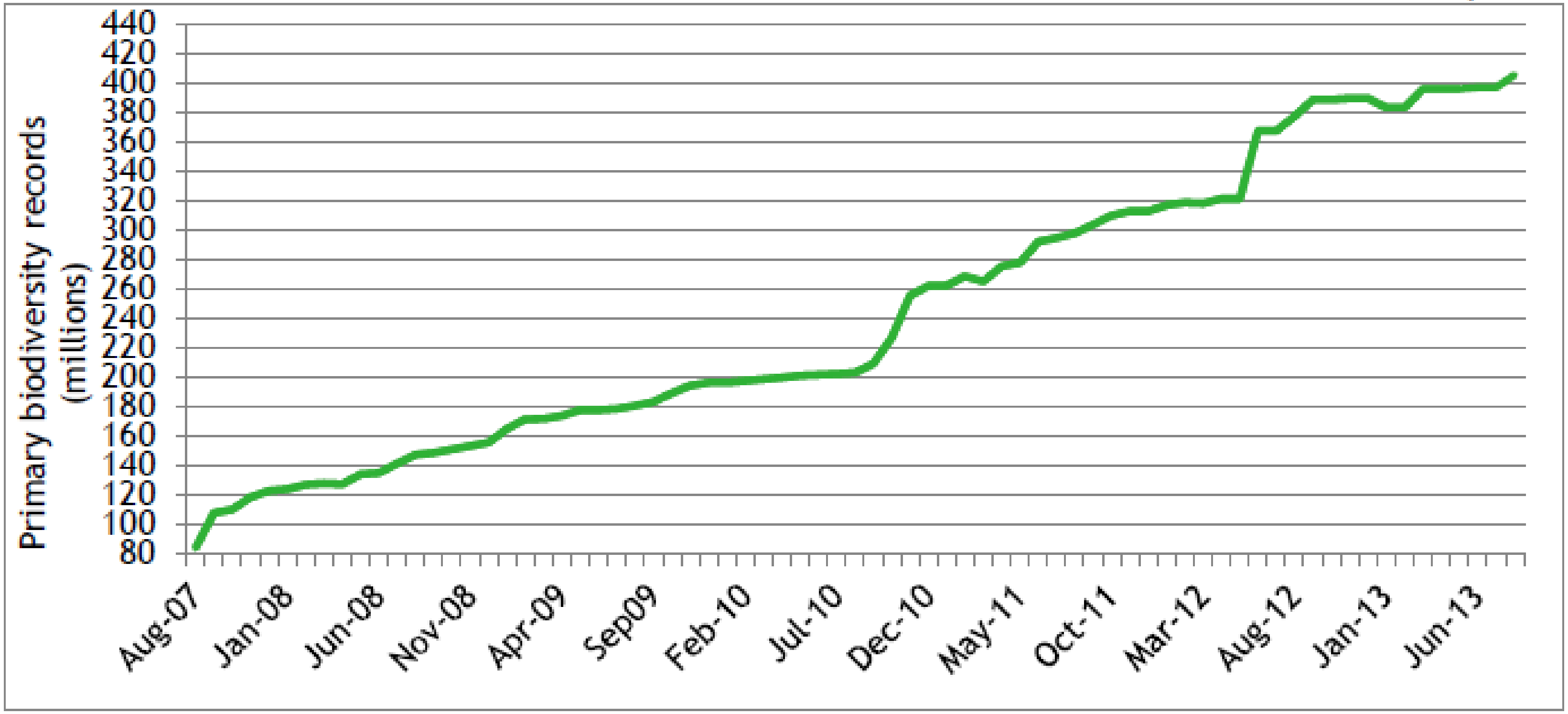

GBIF
Data
News
Community
About
Global Biodiversity Information Facility
Free and open access to biodiversity data
441,170,025
OCCURRENCES
1,454,695
SPECIES
14,897
DATASETS
601
DATA PUBLISHERS
Sharing biodiversity data for re-use
Learn about GBIF
Publish your data through GBIF
Technical infrastructure
Providing evidence for research and decisions
Using data through GBIF
Enabling biodiversity science
Supporting global targets
Collaborating as a global community
Current Participants
How GBIF is funded
Enhancing capacity
Search news items and information pages...
Search

The GBIF data portal – a window on biodiversity

- The GBIF data portal offers a single online access point to over **400 million biodiversity records** from more than **10,000 datasets** published by more than **500 institutions**, ranging from museum specimens collected from the earliest days of natural history exploration, to current observations by 'citizen scientists' and monitoring from research expeditions.

- Since its launch in 2007, the volume of data made accessible through the GBIF data portal and associated web services has continued to grow.

Growth in data records served by GBIF

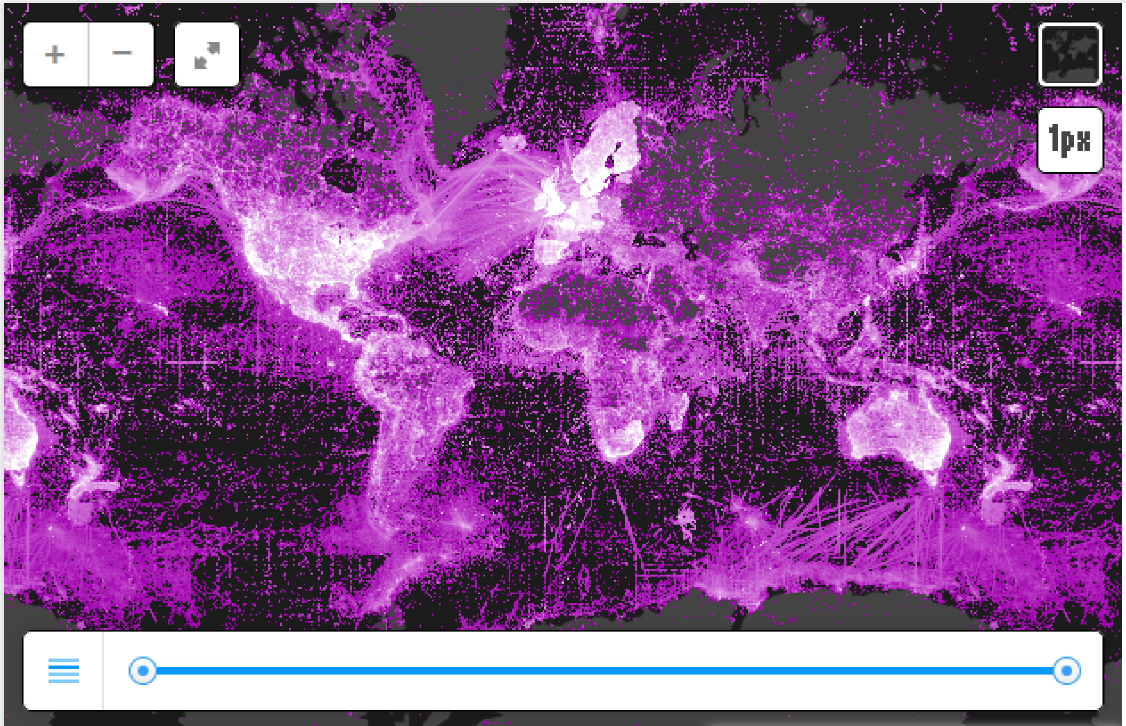


Explore 441,170,026 occurrences

Occurrence records document evidence of a named organism in nature. Through this portal, you can search, view and download records that are published through the GBIF network.

441,170,026
occurrences records

377,065,516
georeferenced records



Georeferenced data

VIEW RECORDS

All records | In viewable area

ABOUT

This map shows the density of all 377,065,516 georeferenced occurrence records published through the GBIF network.

To explore the records, zoom into the map or click on the links above and add further filters to customize search results.

GBIF NEWS

Amateur divers share species data through GBIF

Species observations from thousands of scuba divers all over the world are now freely accessible via the GBIF portal.

The citizen science platform Diveboard has published over 15,000 records from the 'electronic log books' submitted by its community of nearly 100,000 registered divers.

The dataset includes records of species occurrences from dives in all the world's oceans, as well as many inland water bodies.

GBIF NEWS

Providing instant access to data behind species discovery

Researchers and the public can now have immediate access to data underlying discovery of new species of life on Earth, under a new streamlined system linking taxonomic research with open data publication through GBIF and other networks.

GBIF NEWS

How plants weather the cold

A team of researchers has used data on nearly 30,000 species, shared by hundreds of institutions through GBIF, to cast new light on how plants colonized colder regions.

The study published in Nature journal assembled the largest evolutionary "timetree" to show the order in which flowering plants evolved strategies such as the shedding of leaves to move into areas with cold winters.

The international research team used more than 47 million occurrence records accessed via GBIF to determine the distributions of over 27,000 plant species. From these records, they were able to extract minimum temperatures from the Worldclim climate database, to flag which species are exposed to freezing across their ranges.

Featured GBIF data use

See all GBIF data use stories

Using models to inform conservation policies

Two studies, based on data from GBIF and other sources, define the distribution of the bearded wood partridge, and help inform conservation policy.

Designing marine protected areas off Mexico

Researchers look at methods to determine the ideal spacing between protected areas in the Gulf of California, ensuring connectivity based on the distances covered by larvae of fish species identified through GBIF.

Shifting niches and invasive species control

Researchers use data available through GBIF to investigate how species can shift their ecological niches in alien environments – complicating the prediction of invasion risks.

Acknowledgments

Town Peterson
Jorge Sobéron
Monica Papes

The University of Kansas
Natural History Museum

Xingong Li
Department of Geography
University of Kansas

BNHS India
Conserving Nature Since 1883

Zafar-ul Islam
Asad Rahmani
Bombay Natural History Society

National Science Foundation

NSF-ROA (Research Opportunity Award)

Grand Valley State University Faculty Research Grant