



# Role of Data Science and Machine Learning in Managing Plant Invasion Under Climate Change: A Case Study with ILORA Database in India

Amiya Ranjan Bhowmick  
Institute of Chemical Technology, Mumbai  
ICT Mumbai

Joint ICTP-IAEA Advanced School/Workshop  
On Machine Learning in Citizen Science  
27 February - 3 March 2023



# Agenda



## About the Aliens

Who they are  
Why they are of concern



## Indian scenario

Current status  
Research and policy needs



## Looking inside ILORA

Inspirations  
Species and variables



## Analytics

Data Analytical perspectives  
Case studies



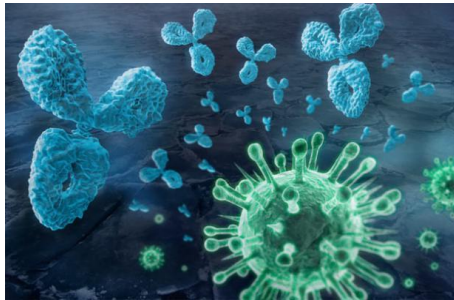
## Closing

Summary  
Questions and Answers



# Alien species

A being that came from a different place



# Alien species

What's in a name?...



- Alien
- Non-native
- Exotic
- Non-indigenous
- Introduced
- Naturalized
- Invasive
- Weed
- Noxious

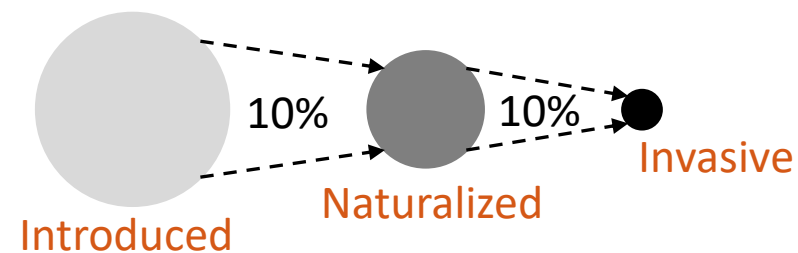




# Alien species

## What's in a name?...

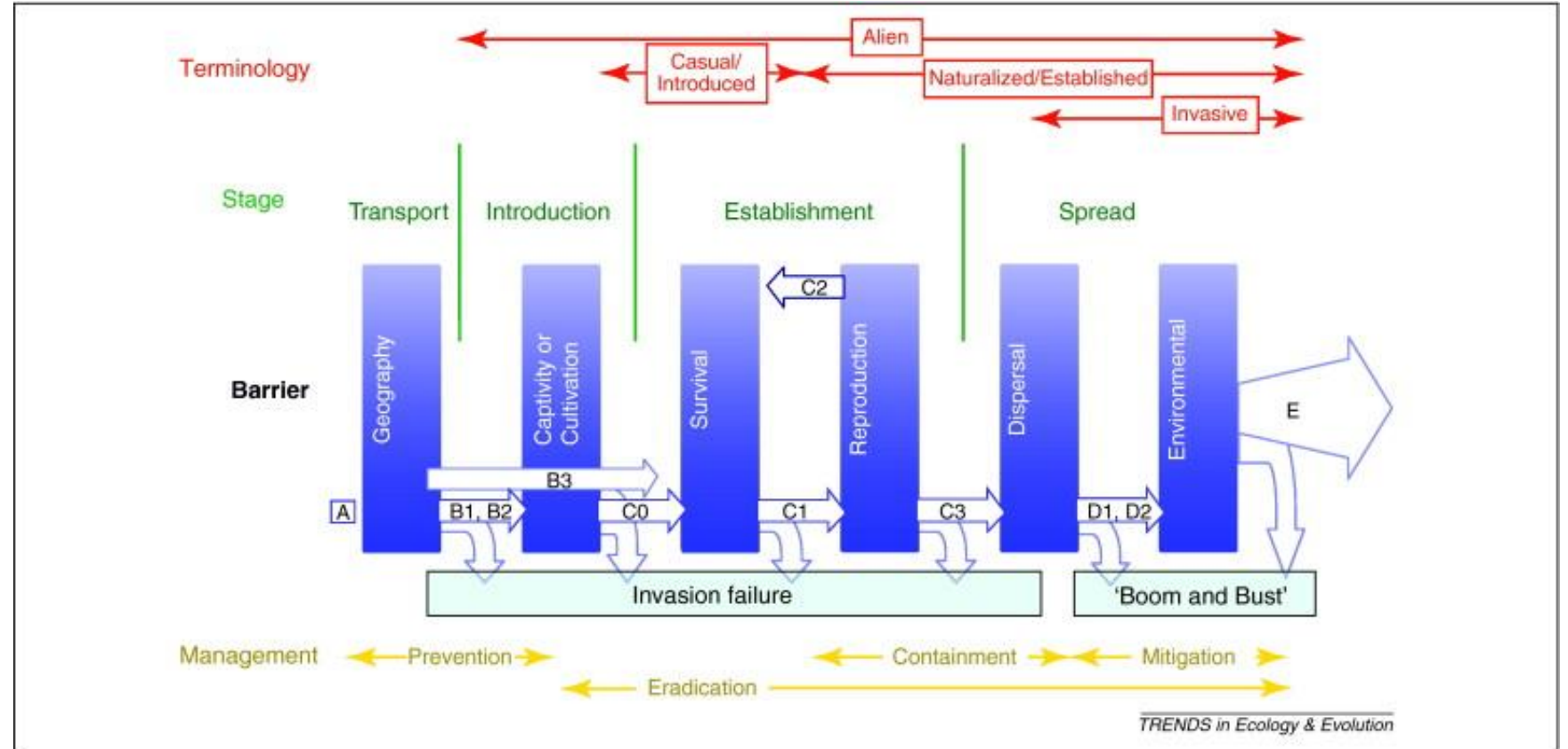
- Alien → Species just enters a geographic space, i.e., crosses the transport barrier, and is being cultivated or in captivity
- Non-native
- Exotic
- Non-indigenous
- **Introduced** → Species is in the wild, outside human control, i.e., crosses cultivation/captivity barrier
- **Naturalized** → Species is surviving in the wild, without human assistance, i.e., crosses survival barrier
- **Invasive** → Species is reproducing on its own and starts dispersing to new environment, i.e., crosses reproduction, dispersal and environmental barriers
- Weed
- Noxious



# Alien species

What's in a name?...

- Alien
- Non-native
- Exotic
- Non-indigenous
- Introduced
- Naturalized
- Invasive
- Weed
- Noxious



Blackburn et al. *Trends Ecol. Evol.* 2011; 26: 333-339

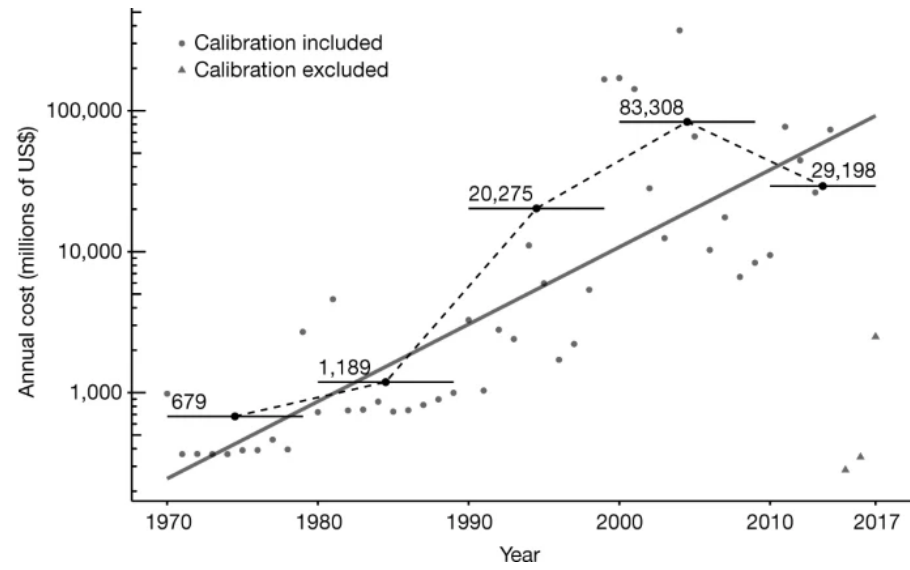
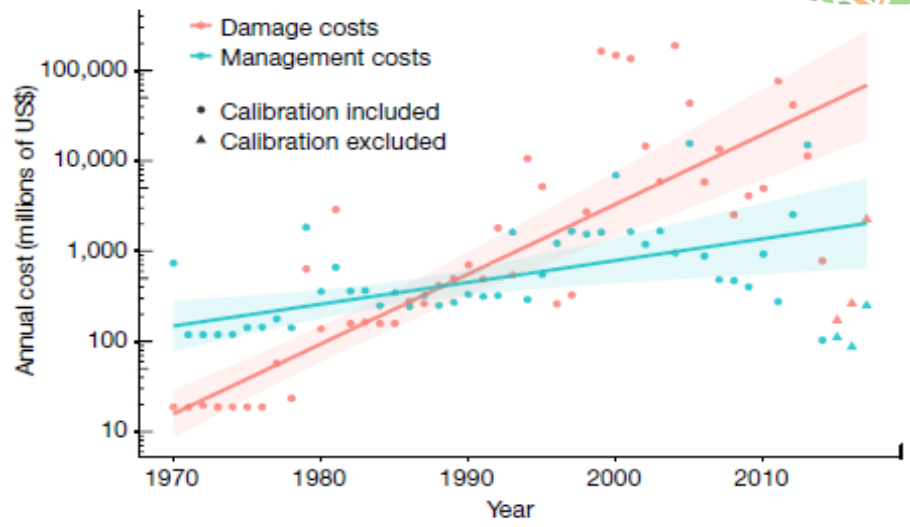
# Alien species

Why we should care



Photo: Revive&Restore

- Number of alien species increasing
- Predicted increase in invasion levels in emerging economies
- High and rising economic cost



Diagne et al. *Nature* 2021; 592: 571–576

# Global perspective

## Economic issue?...

- Alien
  - Non-native
  - Exotic
  - Non-indigenous
  - Introduced
  - Naturalized
  - Invasive
  - Weed
  - Noxious
- In the 2001–02 financial year, the combined estimated cost (**economic losses and control**) of invasive species was **\$9.8 billion, rising to \$13.6 billion** in the 11–12 financial year (**Hoffmann and Broadhurst, 2016, NeoBiota**).
  - Approximately **\$726 million of grants funded through the Commonwealth of Australia** was spent on invasive species management and research between 1996 to 2013.
  - In 01–02, total **national expenditure** on invasive species was \$2.31 billion, rising to \$3.77 billion in 11–12.
  - Of the approximately 2700 exotic plants species now established within **Australia, 429 species** have been declared noxious or are under some form of **legislative control (NRMCMC 2006)** with considerably more subject to eradication and control measures such as plant species listed on the National Environmental Alert List.

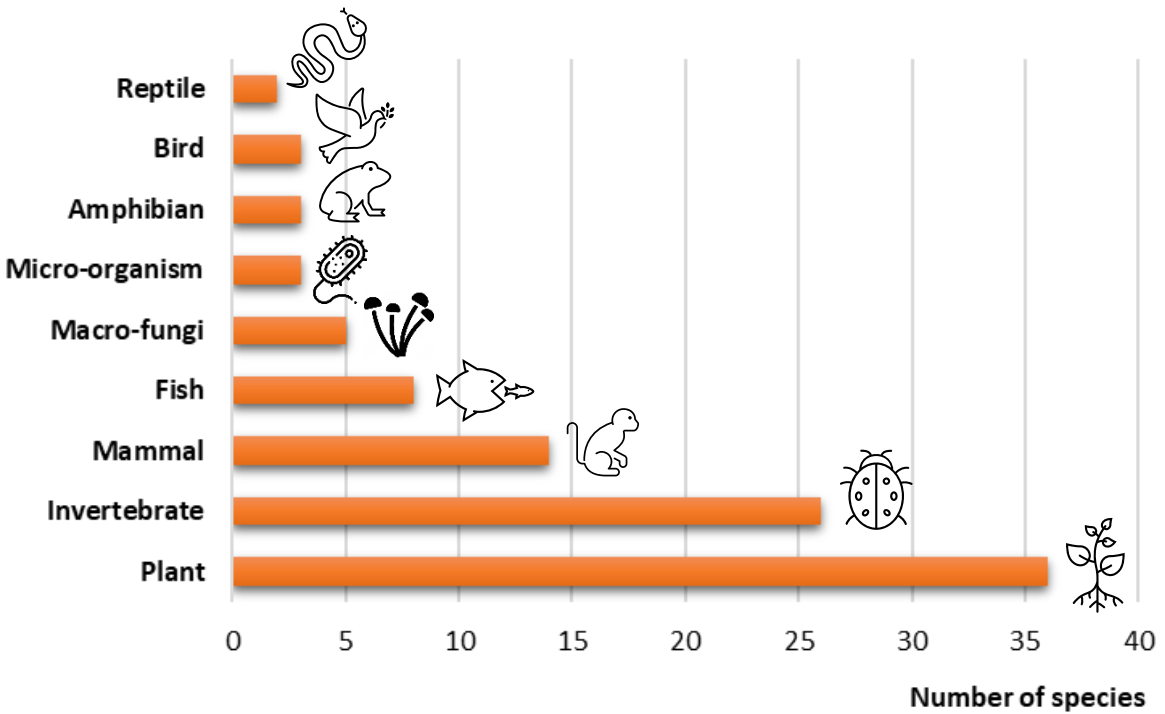
**Source:** <http://www.environment.gov.au/biodiversity/invasive/weeds/weeds/lists/alert.html>



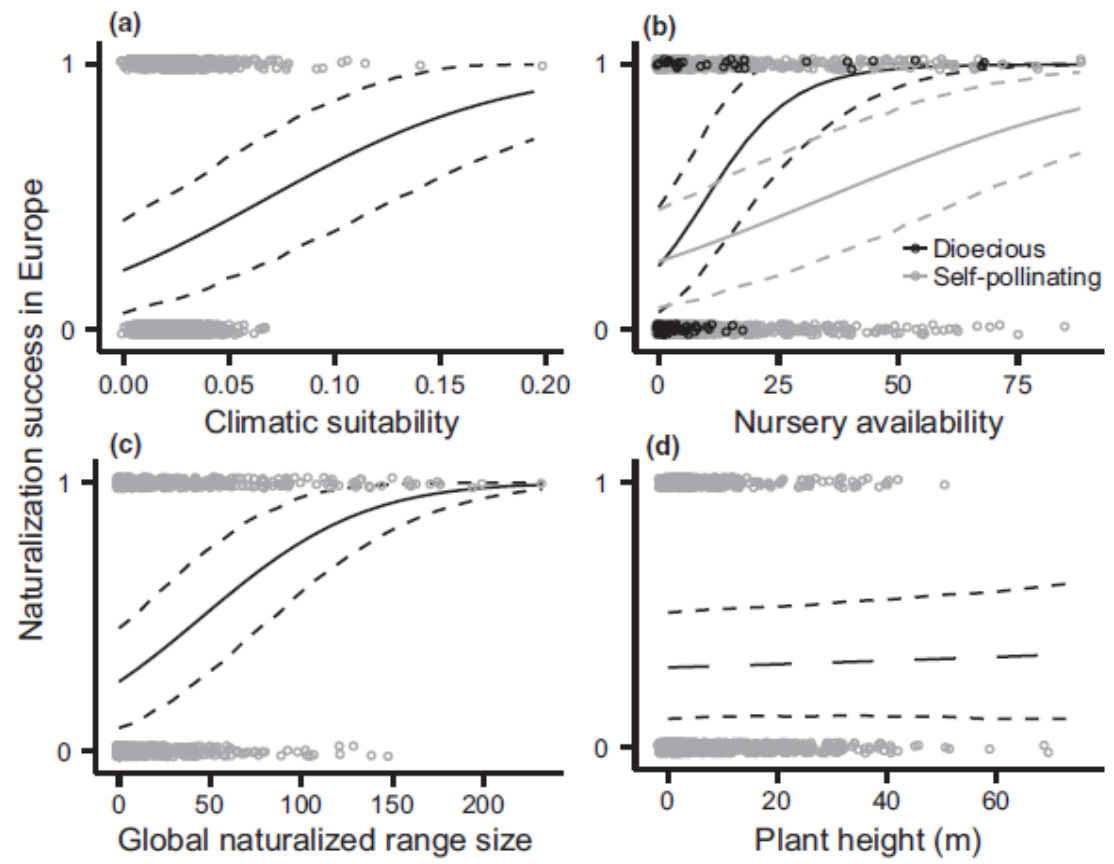
Process Initiated

# Other countries started

Invasive plants are the major shareholder



Lowe et al. 2000, 100 OF THE WORLD'S WORST INVASIVE ALIEN SPECIES, The Invasive Species Specialist Group (ISSG) a specialist group of the Species Survival Commission (SSC) of the World Conservation Union (IUCN)



Haeuser et al. J Appl Ecol. 2018;1–10.

- Identified species with **high future naturalization risk** and species with **high projected increases in naturalization potential** in Europe under **climate change**.
- This species list allows for prioritization of monitoring and **regulation of ornamental plants** to mitigate the invasion debt.

# Alien plants in India

Some popular ones: Its everywhere!

*Parthenium hysterophorus*,  
Herb/Shrub



*Eichhornia crassipes*,  
Herb Aquatic



*Lantana camara*,  
Shrub



*Acacia auriculiformis*,  
Tree



*Mikania micrantha*,  
Vine

# Citizen Science (Indian Context)

## Google Search: Citizen Science in India

- Citizen Science for **Biodiversity**
  - <https://citsci-india.org/>
  - Total number of projects: 31
- 
- Project on Invasive plants: **1**
  - Machine Learning: **Only for experts**

CitSci India

[Home](#)
[About](#)
[Public Sessions](#)
[CitSci India 2022](#)
[Working Groups](#)
[Past Events](#)
[CitSci Projects](#)

Citizen Science for Biodiversity



### Mapping Invasive Alien Plants

Species that are introduced from elsewhere are known as alien or exotic species. Some introduced species become invasive, with negative impacts on native biodiversity, ecosystem services and human wellbeing. In India we lack systematic information about invasive species distributions that could enable prioritising species (or habitats) to manage. The aim of this project is to map invasive plant distributions and to create greater public awareness about invasive alien species and their potential ecological and socioeconomic impacts.

**Year of Initiation:** 2021

**Project Coordinator:** Reshnu Raj R S

**E-mail Address:** [reshnu.raj@atree.org](mailto:reshnu.raj@atree.org)

**Role in the project:** Project Coordinator

**Institutional Affiliation:** Ashoka Trust for Research in Ecology and the Environment (ATREE)

**Taxa:** Herbs, Grasses, Shrubs, Vines, Trees

**Status:** Active



Image credits: Ankila Hiremath

**Website:** <http://miap.atree.org>

**Tags:** [monitoring](#), [mapping](#), [invasive alien species](#)



# Citizen Science (Indian Context)

Google Search: Citizen Science in India

- Center for Citizen Science
- <http://citizenscience.in/>
- SATARK Landslide Prediction System
- Microbial analysis of river Ganga
- Microbial analysis of 'Dust Storm' event
- Asteroid occultation observation
- Projects: Meghdoot, Shayadri
- Project on Invasive plants: 0
- Machine Learning: **Only for experts**



**CENTER FOR CITIZEN SCIENCE**  
Understanding Nature for Better Future

HOME ABOUT CCS महाराष्ट्र खगोल संमेलन KHAGOL VISHWA RESEARCH PROJECTS PROJECT MEGHDOOT PROJECT SAHYADRI  
CCS IN MEDIA CONTACT US

## About CCS

**Centre for Citizen Science (CCS)**

Center for Citizen Science- CCS is Pune based voluntary organisation. Main objective of CCS is to promote Citizen Science in India and to solve various social, environmental issues with the help Science - Technology. CCS has been working as a bridge between society and research institutes

Search form

**RECENT POSTS**

२६ ते २८ जानेवारी दरम्यान पुण्यात महाराष्ट्र खगोल संमेलन  
JANUARY 6, 2023

सरस्वती नदीच्या शोधदात  
AUGUST 15, 2020

# Indian perspective

## Economic issue?...



*Prosopis juliflora*



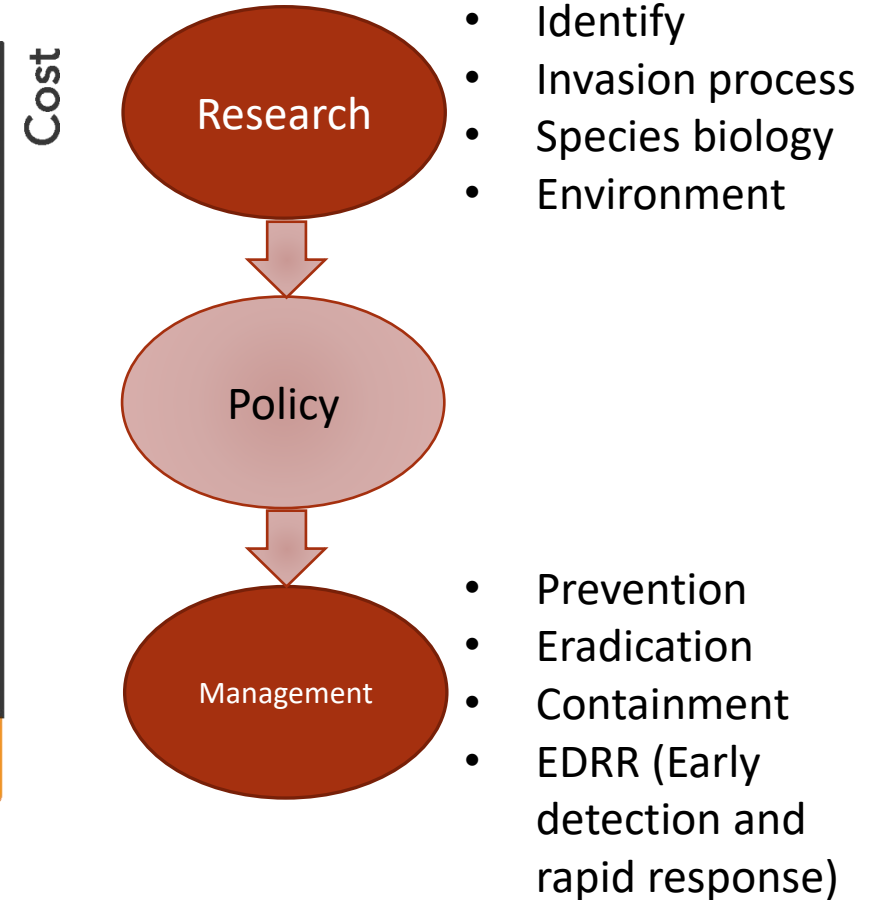
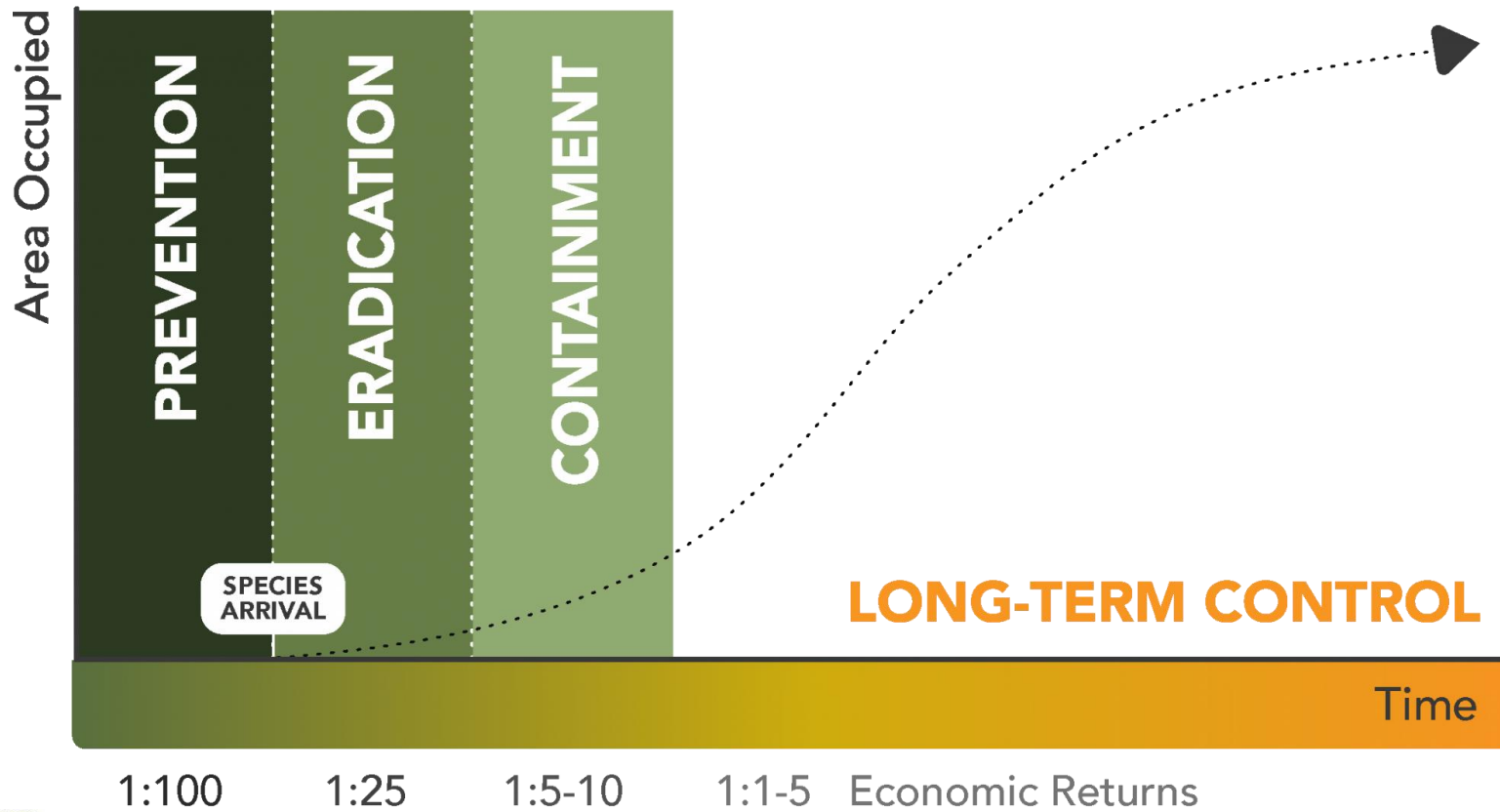
*Lantana camara*

- In the Western Ghats, for example, **lantana (*Lantana camara*)**, a Central and South America shrub that was introduced to India in **1809** as a **garden ornamental**, has spread extensively.
- In Karnataka's Biligiri Rangaswamy Temple Tiger Reserve, Soliga farmers have suffered from reduced abundances of (Non-timber forest products) NTFPs, such as amla, which they harvested for supplementary income. Lantana affects regeneration of other forest plants that wild herbivores depended upon.
- **Prosopis (*Prosopis juliflora*)**, variously known as **vilayati keekar, gando bawar and seemai karuvel** in different parts of the country. Also from South America, this thorny tree was introduced to India around **1870** for its **extreme drought tolerance** and as a **source of fuelwood**.
- Rapid spread in **Rajasthan, Gujarat, Maharashtra, Tamil Nadu, Andhra Pradesh and Karnataka**. In parts of Tamil Nadu, the invasion has entirely transformed agricultural lands and grazing commons, forcing people to seek new livelihood opportunities.

Source: The Wire: <https://thewire.in/environment/invasive-species-prosopis-lantana>

# Data-driven action plan

When and how we should care

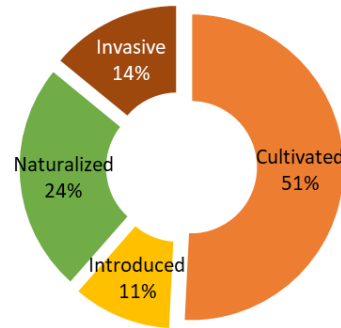




# Alien plants in India

Multiple checklists, little coordination

- Alien
- Non-native
- Exotic
- Non-indigenous
- Introduced
- Naturalized
- Invasive
- Weed
- Noxious



**1599 alien species**  
*Alien Flora of India*  
 Khuroo et al. 2012



**Invasive plant species:**  
 54 terrestrial, 8 aquatic  
*National Biodiversity Authority, 2018*

SpringerLink

**471 naturalized plants**

2018

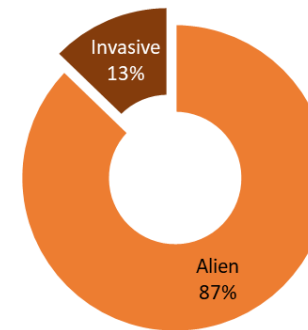
Alien Floras and Faunas 1 | Published: 13 December 2017

Naturalized alien flora of the Indian states: biogeographic patterns, taxonomic structure and drivers of species richness

Inderjit, Jan Pergl, Mark van Kleunen, Martin Hejda, Cherukuri Raghavendra Babu, Sudipto Majumdar, Paramjit Singh, Surendra Pratap Singh, Sugali Salamma, Boyina Ravi Prasad Rao & Petr Pyšek

*Biological Invasions* 20, 1625–1638 (2018) | [Cite this article](#)

1279 Accesses | 16 Citations | [Metrics](#)

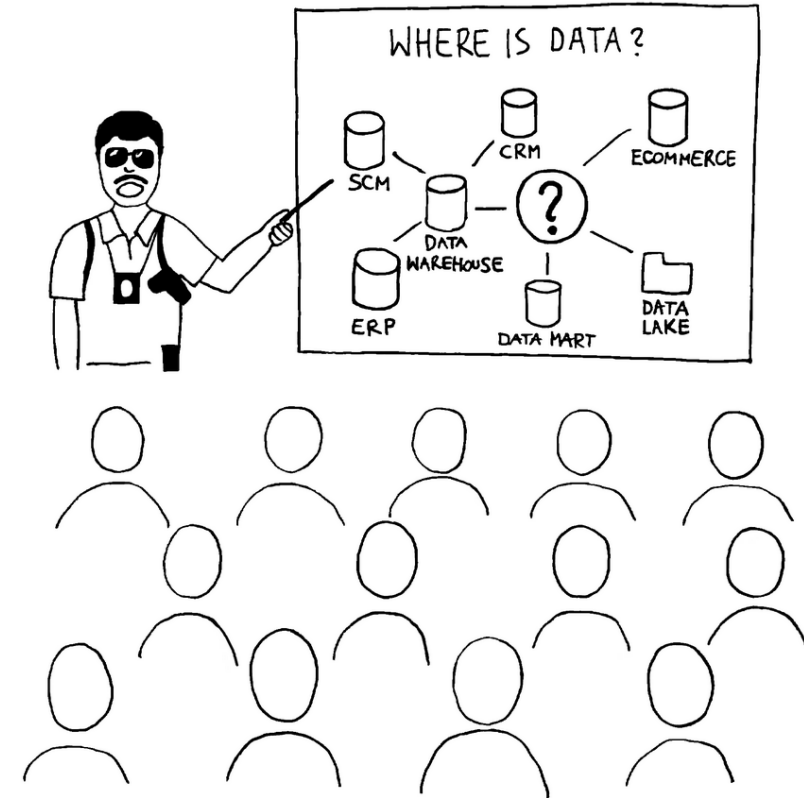


**2082 alien species**  
*Global Register of Introduced and Invasive species, 2020*

# Alien plants in India

## Where is the data?

- Scattered across multiple databases
- Intensive and informed query-directed search
- Tedious access to data
- Low resolution at national and regional scales
- Lack of comprehensive database



YOU'LL WORK IN TEAMS OF TWO,  
EACH LOOKING INTO 540 TABLES AND FILES.

# ILORA

## Inspirations

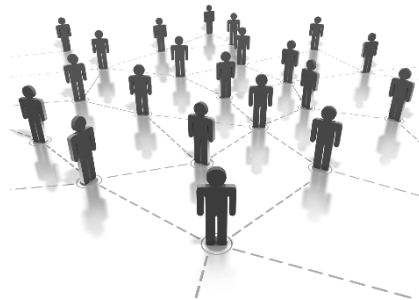
- A database having information of variables responsible for invasion process for INDIAN alien plant species



*Open access*



*Dynamic*



*Networking*

## Indian ALien FLOra InfoRMAtion database

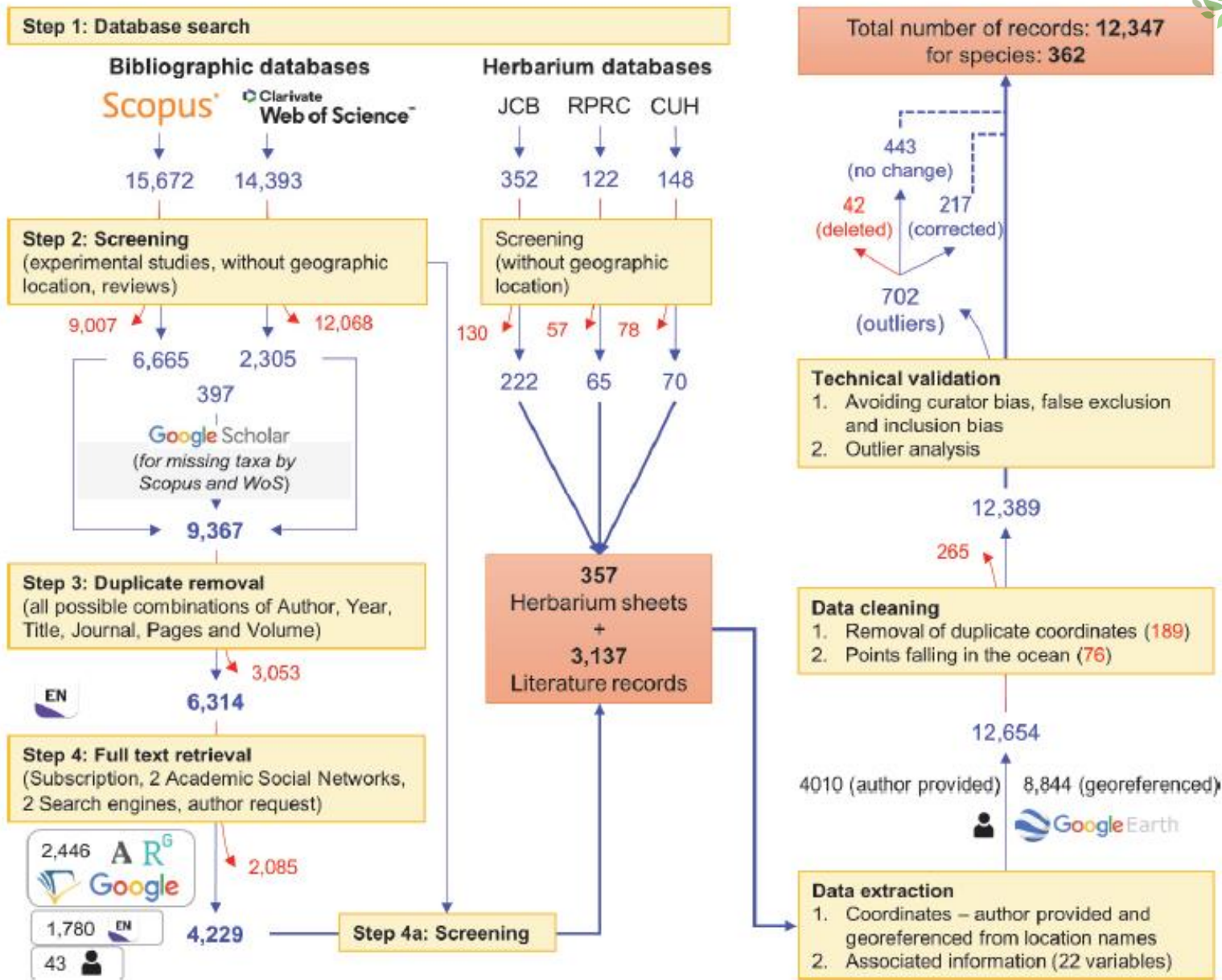




# ILORA

## The process

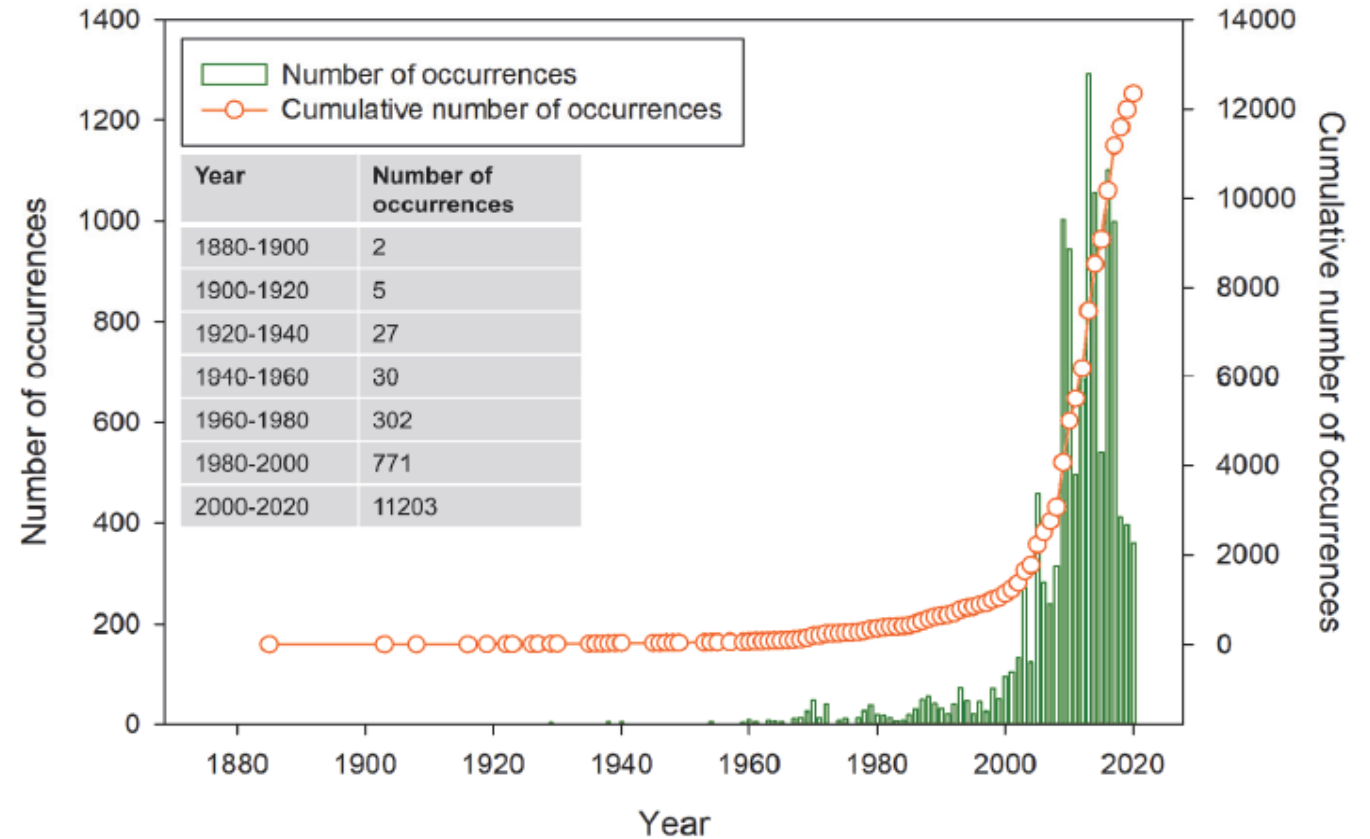
- Automated data extraction process using codes written in Python (Thanks to Interns and PhD students)
- Manual verification of each record



# ILORA

## Overview

- Number of species: **1747**
- Number of variables: **14**
  - Species categorization
  - General information
  - Introduction history
  - Biogeography
  - Uses
  - Occurrence and distribution
  - Climate
- Number of sources: **22**
- Standardized following international stand

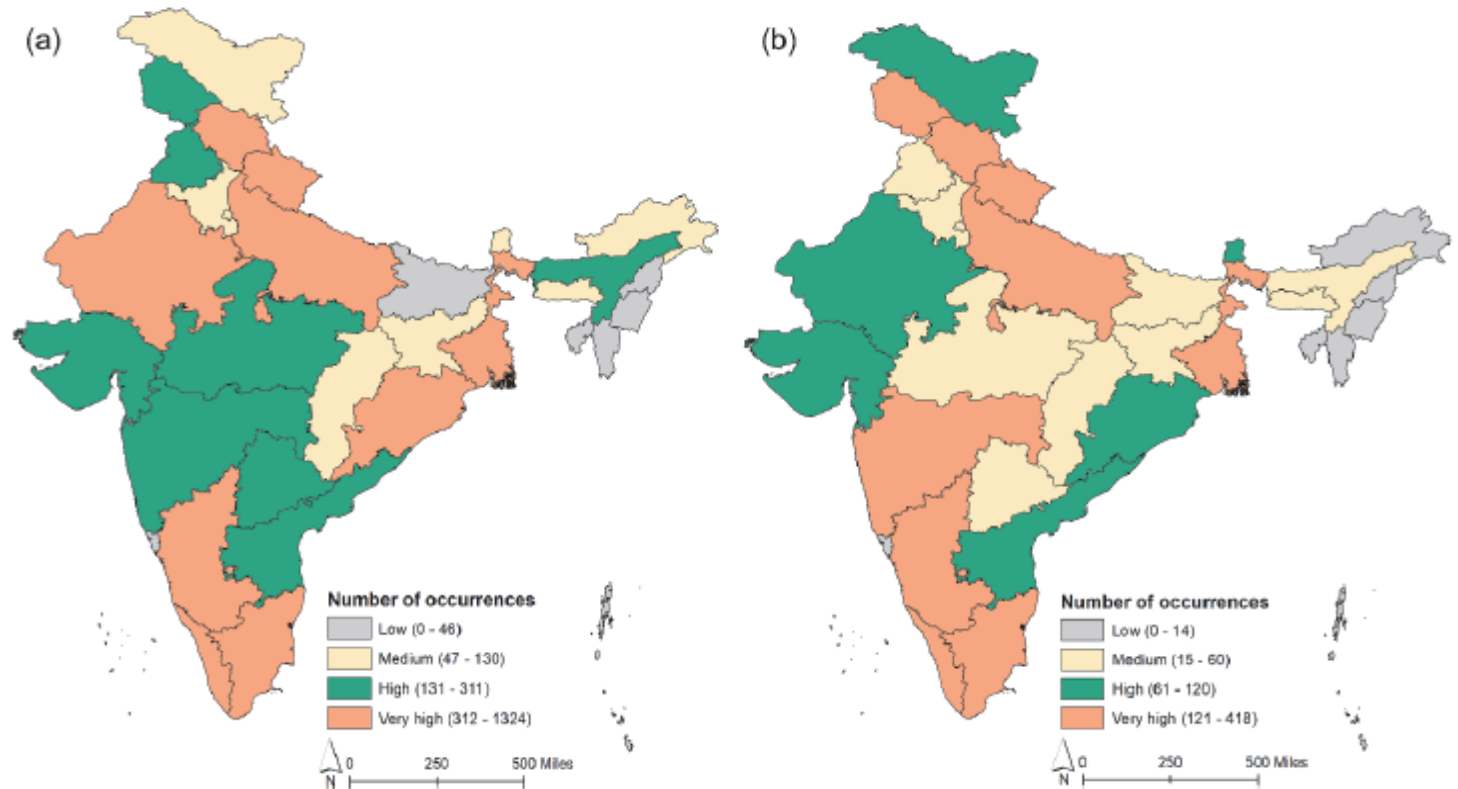


*Prajapati et al. 2022, Ecology 103 (11): e3794.*

# ILORA

## Overview

- State-wise distribution of occurrence data of (a) invasive and b) naturalized alien species in India.



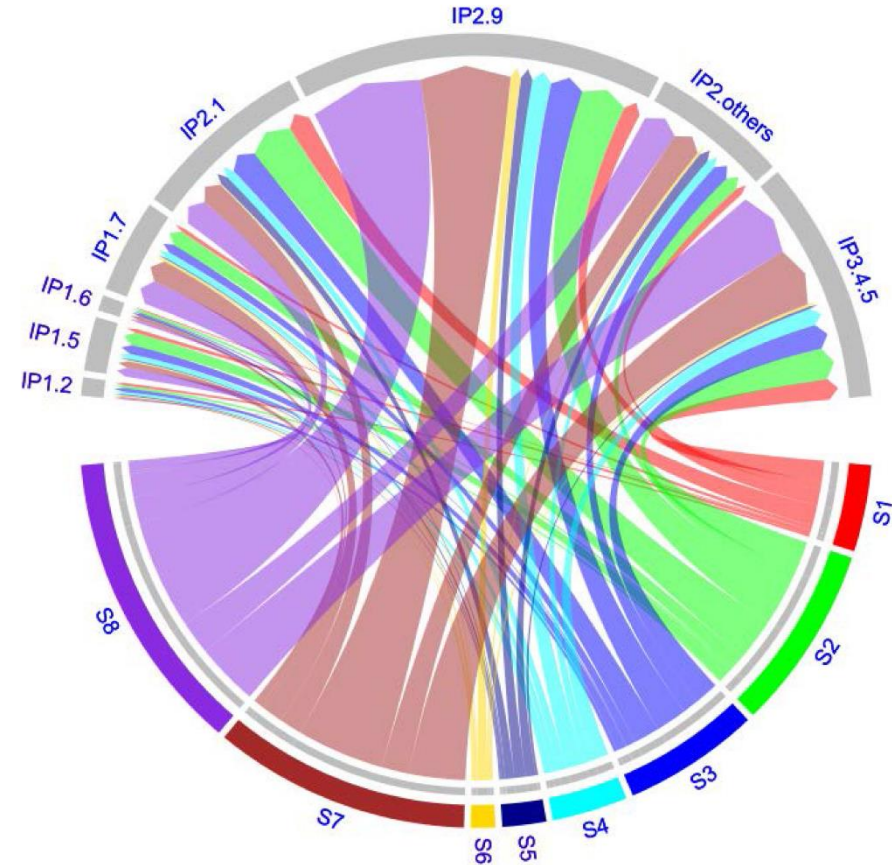


# ILORA

## How and When - Introduction history

- Introduction pathways (**how** the species came) - categorical
- First record date (year **when** the species was first recorded) - numeric
- Earlier and more pathways of entry often facilitate successful naturalization and invasion

IP: Introduction Pathways –  
**1.2:** Erosion control; **1.5:** Landscape improvement; **1.6:** Conservation; **1.7:** Release in nature for use; **2.1:** Agriculture; **2.9:** Ornamental; **2.others:** Botanical garden + Aquarium species + Forestry + Horticulture + Food + Other escape; **3.4.5:** Contaminant (Nursery, Food, Plant, Seed, Habitat materials), Stowaway (Packing, Ballast water), Interconnected waterways

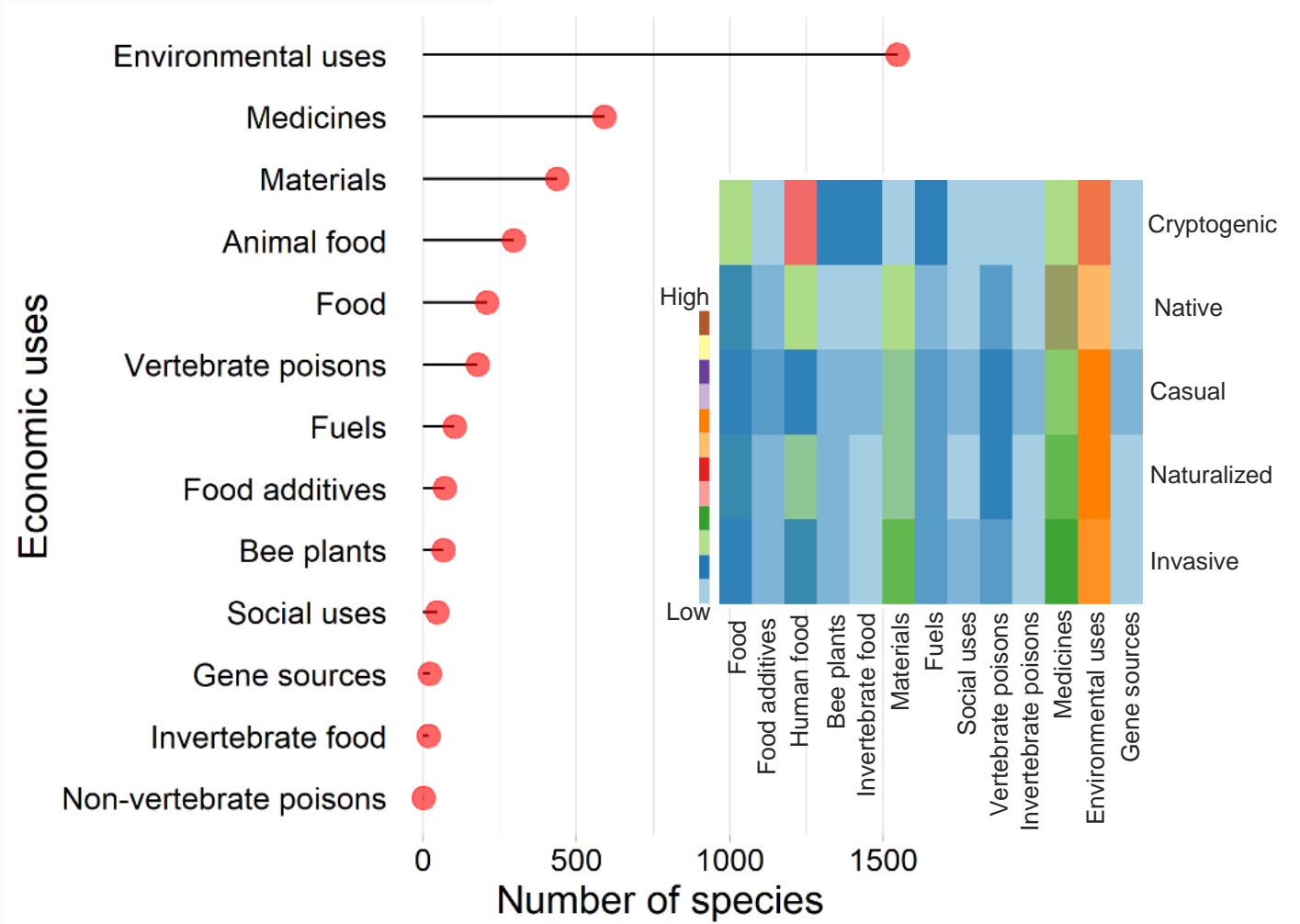


S1: Europe; S2: Africa; S3: Asia-Temperate; S4: Asia-Tropical;  
 S5: Australasia; S6: Pacific; S7: Northern America; S8: Southern America

# ILORA

## Uses

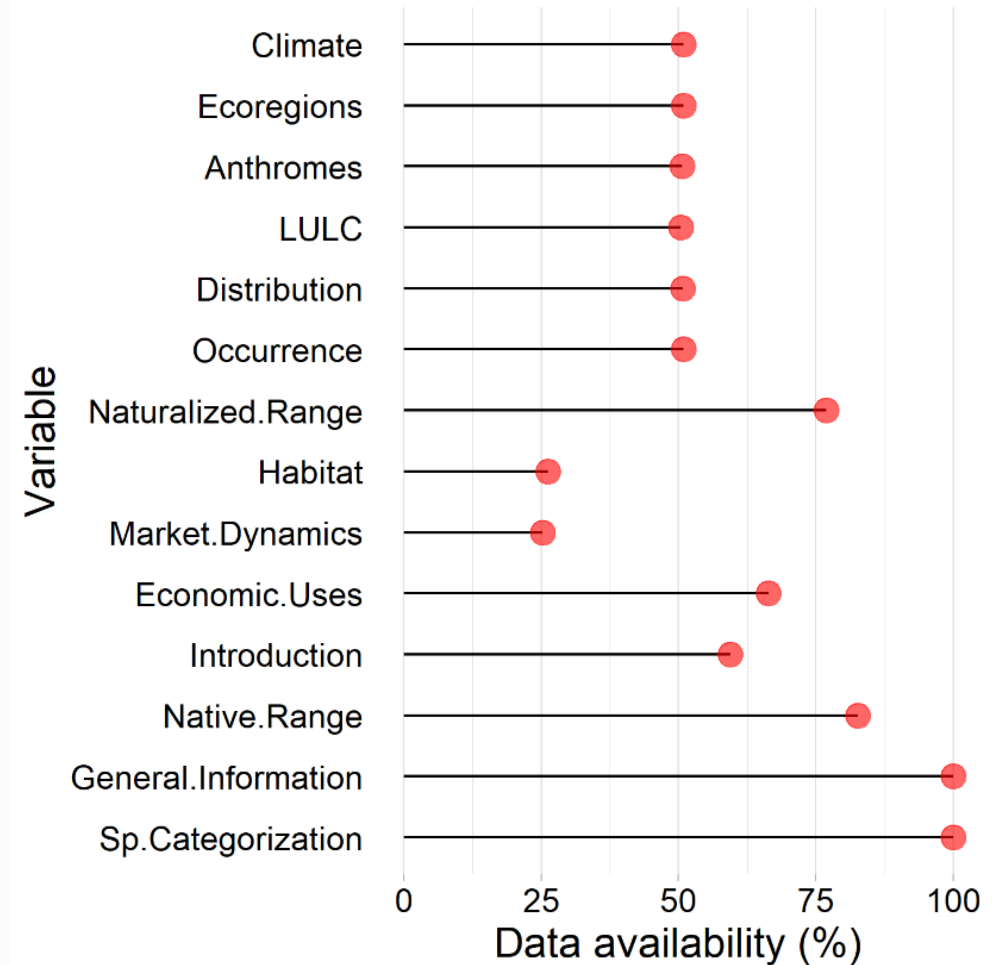
- Categories (13) and subcategories (50) of uses - categorical
- More number of uses facilitate alien species naturalization and invasion



# ILORA

## Summary

- Number of species: **1747**
- Number of variables: **14**
  - Species categorization
  - General information
  - Introduction history
  - Biogeography
  - Uses
  - Occurrence and distribution
  - Climate
- Number of sources: **22**
- Standardized following international standard



# How we use ILORA

## The database itself



Home About Data Submit data Projects Citation Contact

ILORA

- Robust **data validation**
- Variables arranged in **CSV** files
- Published as **open access**
- Dedicated **website**



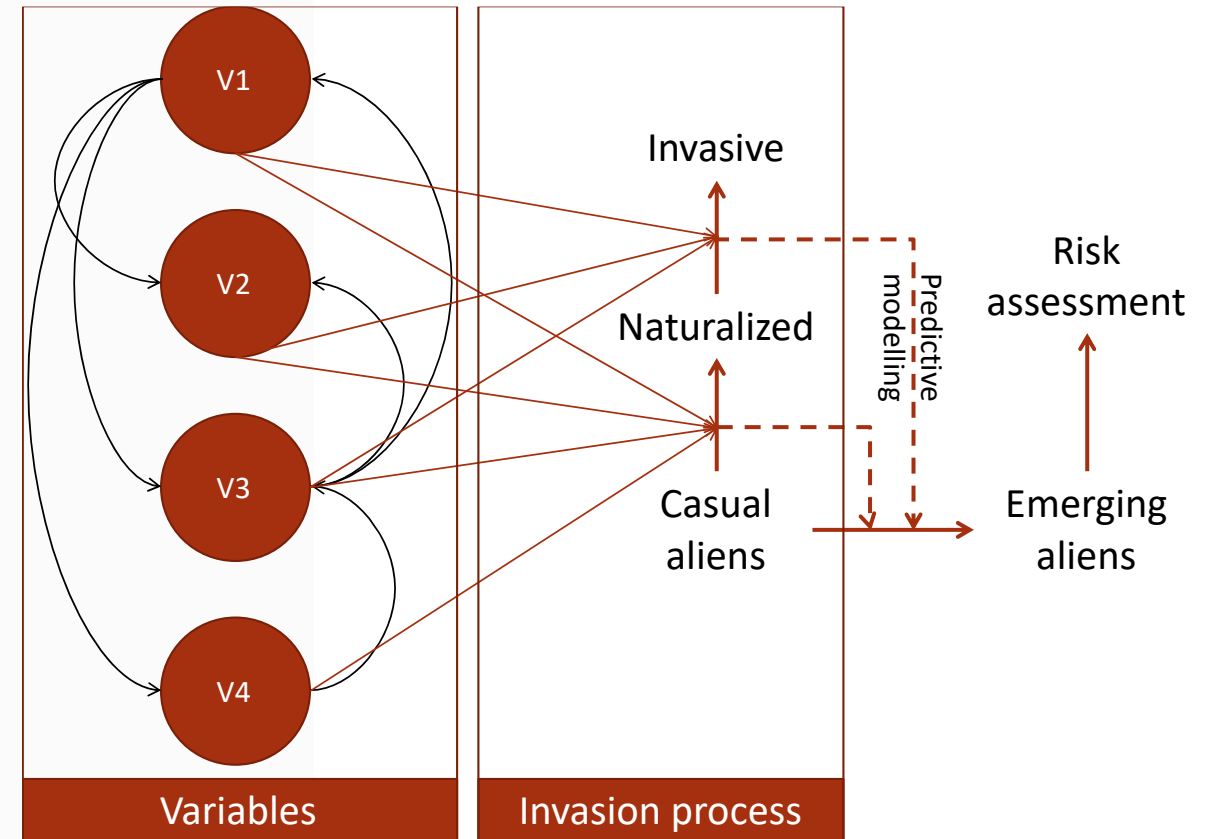
<https://ilora2020.wixsite.com/ilora2020>



# How we use ILORA

## Identifying emerging invaders

- Considered the invasive, naturalized and casual aliens
- Based on the relationship between variables and invasion process
- Identify the aliens which can become invasives
- Risk assessment
- Proactive management

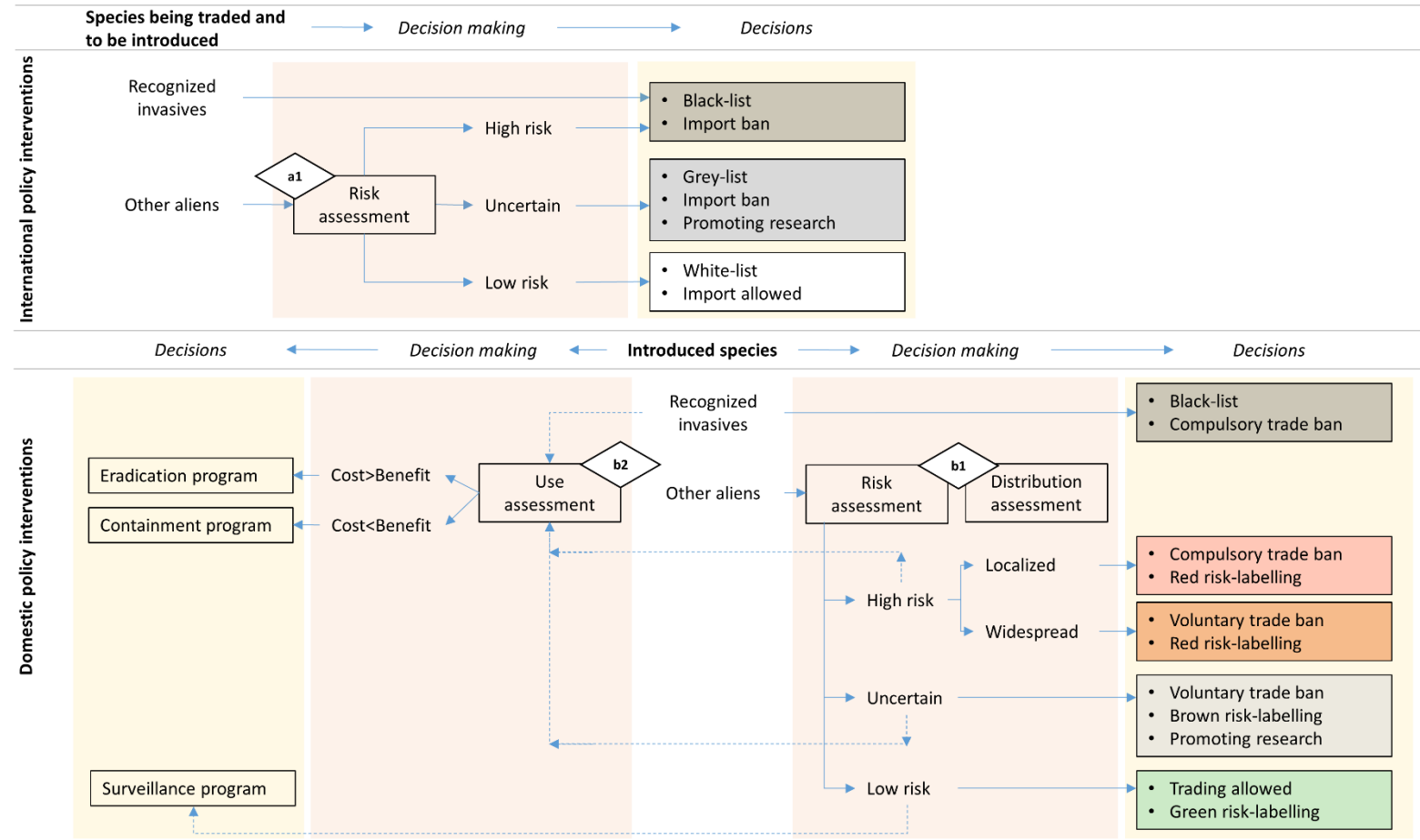


Banerjee et al. *J. Environ. Manage.* 2021; 294:113054

# Policy framework

An action plan specific to India

- Policy interventions proposed for regulating trade and managing invasive alien plant species in India (a) international and (b) domestic



<https://ilora2020.wixsite.com/ilora2020C>

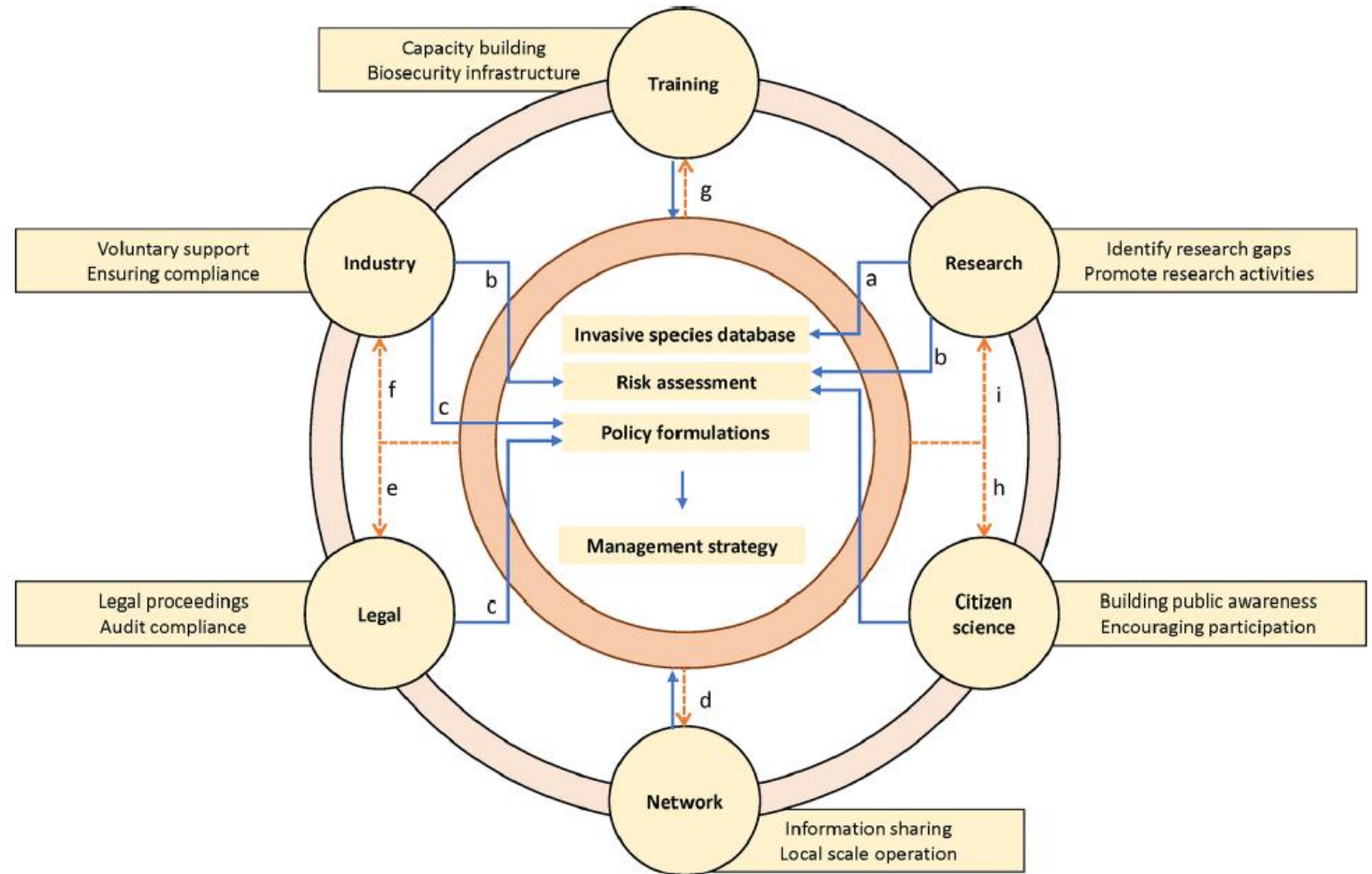
Banerjee et al. *Environmental Science and Policy*. 124, 64-72, 2021



# Multidisciplinary approach

An action plan specific to India

- A decentralized system tasked with coordination between different agencies involved in the trading of invasive alien plant species in India

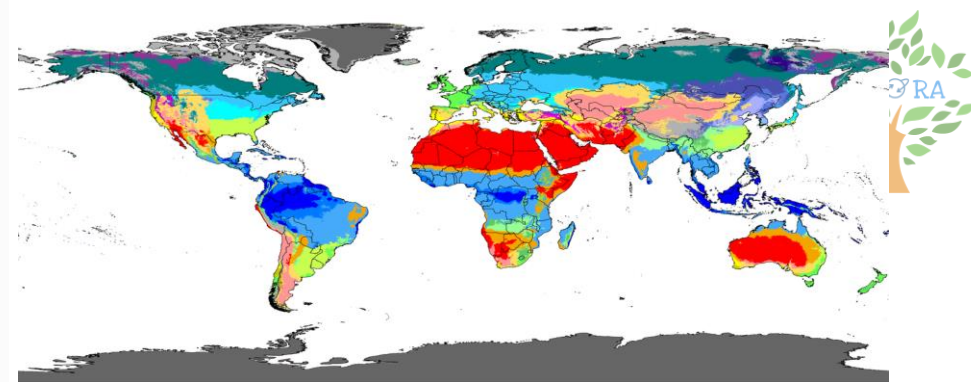


<https://ilora2020.wixsite.com/ilora2020>

Banerjee et al. *Environmental Science and Policy*. 124, 64-72, 2021

# Emerging invaders

## Data processing and statistical modelling



### Species selection

- ILORA
- 220 IAPS
- Occurrence (www.gbif.org)
- Occurrence Records 6355150

### Data Cleaning

- Removing missing coordinates
- Duplicates removal
- Removing points falling in ocean
- Removing geographic bias

### Delimiting geographic range

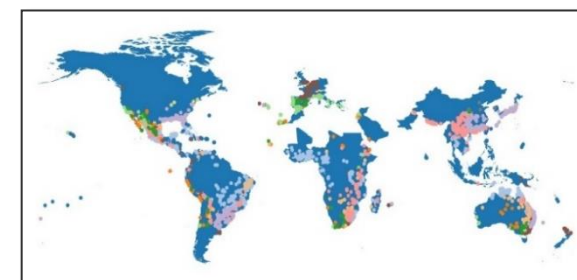
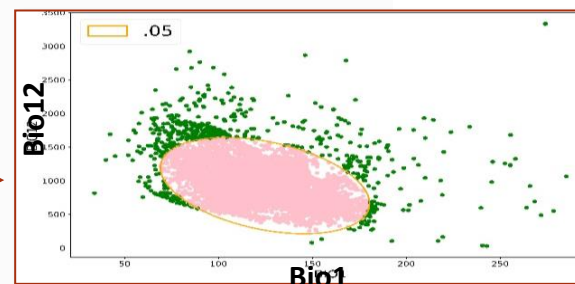
- Extract temperature (Bio1)
- Extract precipitation (Bio12)
- 95% ellipsoid in Bio1 and Bio12
- Occurrence within confidence ellipsoid

### Background selection

- Köppen-Geiger climate Class [www.gloh2o.org/koppen](http://www.gloh2o.org/koppen)
- Climate data extraction [www.worldclim.org](http://www.worldclim.org)
- Future climate data (RCP 4.5 and RCP 6.0)
- Union of K-G classes for occurrence records

A		B		C	
Af	BWh	Csa	Cwa	Cfa	
Am	BWk	Csb	Cwb	Cfb	
Aw	BSh	Csc	Cwc	Cfc	
As	BSk				

D			E
Dsa	Dwa	Dfa	ET
Dsb	Dwb	Dfb	
Dsc	Dwc	Dfc	EF
Dsd	Dwd	Dfd	



Basic Introduction: <https://sites.google.com/site/amiyaiitb/research/ecological-niche-modelling>

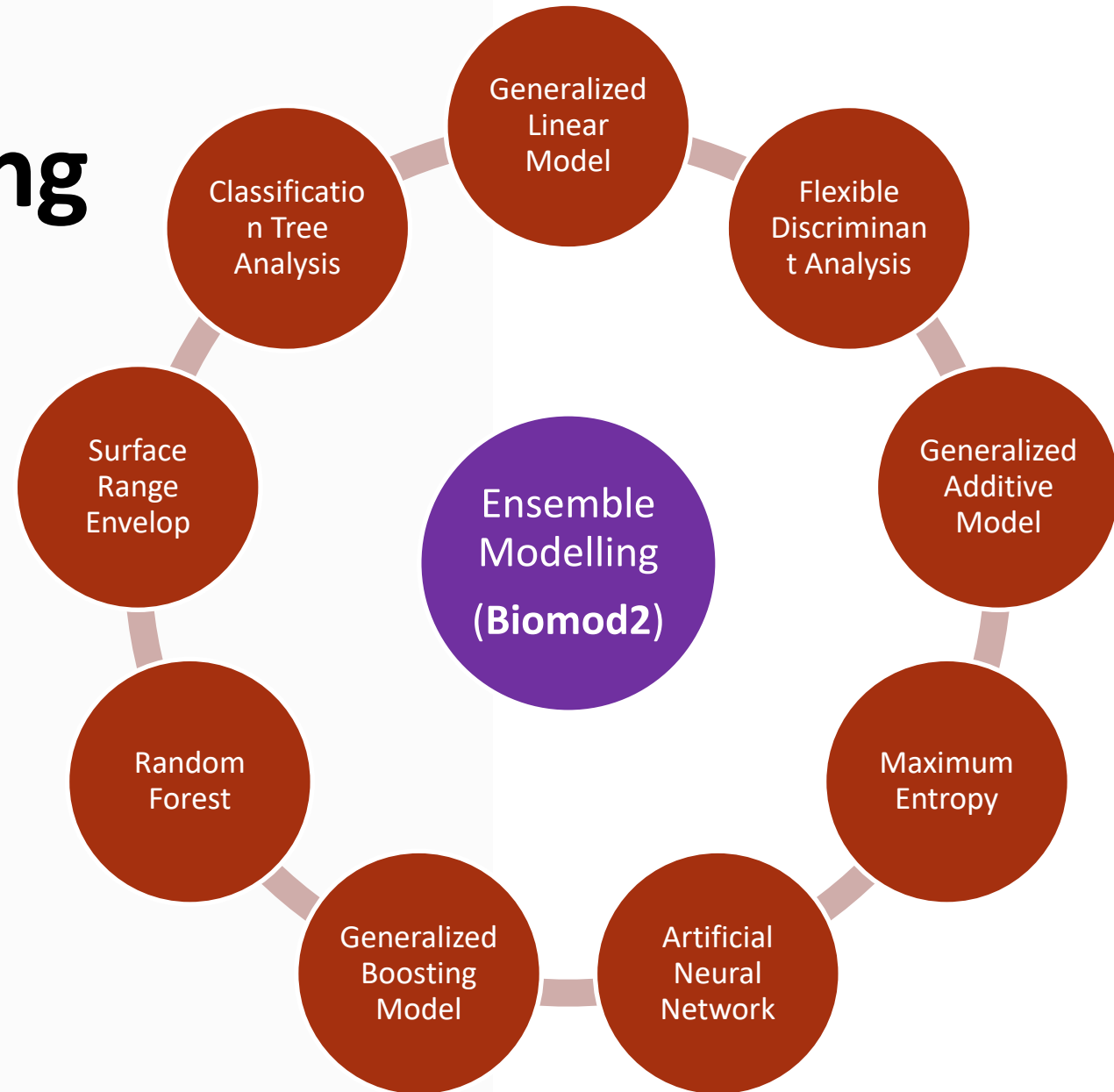




# Ensemble modelling

## biomod2: Ensemble Platform for Species Distribution Modeling

- Generalized Linear Model ([Searle and McCulloch, 2001](#))
- Flexible Discriminant Analysis ([Hastie et al., 2009](#))
- Generalized Additive Model ([Guisan et al., 2002](#))
- Maximum Entropy ([Favretti, 2017](#))
- Artificial Neural Network ([Zhang, 2010](#))
- Generalized Boosting Model ([Einziger et al., 2019](#))
- Random Forest ([Nordhausen, 2014](#))
- Surface Range Envelop ([Hannah, 2012](#))
- Classification Tree Analysis ([Breiman et al., 2017](#))

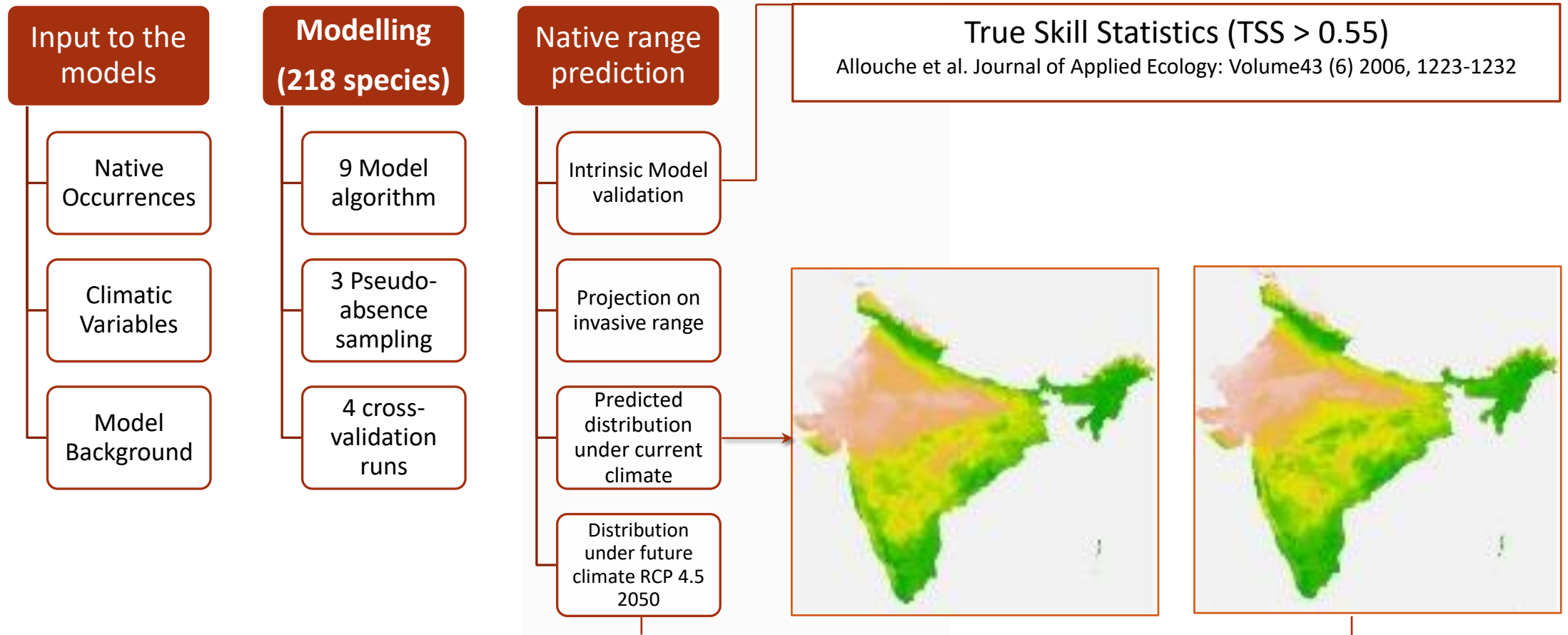


# Selection of predictors

## Identifying invasion hotspots

- First four bioclimatic variables from the first two principal components with highest loadings in PC1 & PC2 using PCA package in R. ([Guisan et al., 2017](#))
- Variables with a correlation value less than 0.7 using stats package ([Braunisch et al., 2013](#))
- Principal components developed globally using bioclimatic variables using the **kuenm** package in R and used first 5 principal components as predictors ([Cobos et al., 2019](#))
- Variation Inflation Factor with a threshold as 5 using **regclass** package in R ([Mpakairi et al., 2017](#); [Rodríguez-Rey et al., 2019](#))
- Based on a pilot run on 10 randomly selected species, the performance of the ensemble model approach has been evaluated using all the above four variable selection strategies.
- No significant difference was obtained for different variable selection predictors.

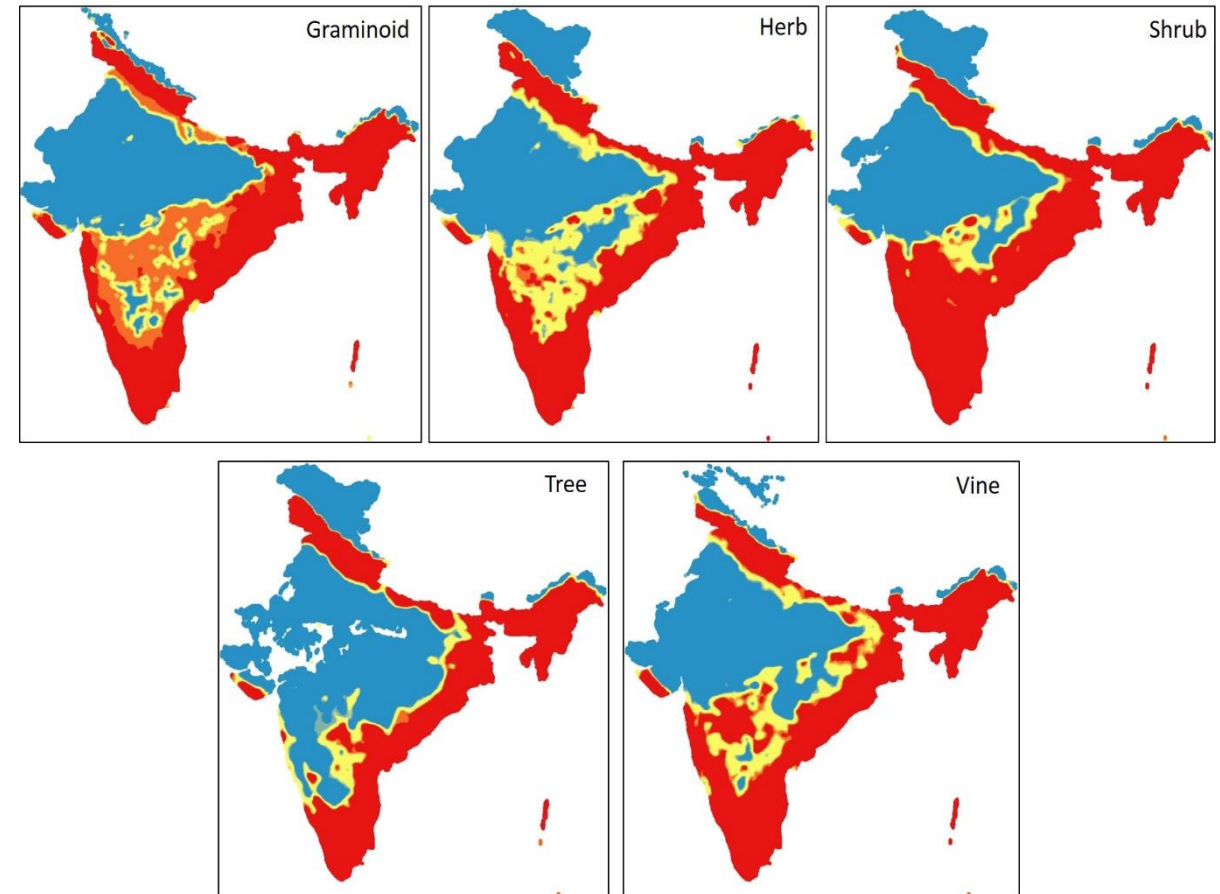
# Ensemble predictions



# How we use ILORA

## Identifying invasion hotspots

- Considered the invasive aliens (218)
- Ensemble species distribution modelling framework
- Identify climatically suitable areas
- Both under current and future climates
- Management implications
- The analysis has been carried out using QGIS plugin







# FAQs - ILORA

## Time to wrap up

- From where can I read ILORA publications?

While ILORA is open access, some of our papers are not. Check the abstracts in [ResearchGate](#) and send us request for full texts. We will respond as soon as (humanly) possible.

Received: 3 July 2021 | Accepted: 20 September 2021

DOI: 10.1002/2688-8319.12105

DATA ARTICLE



## ILORA: A database of alien vascular flora of India

Vidushi Pant<sup>1</sup> | Chinmay Patwardhan<sup>2</sup> | Kshitij Patil<sup>2</sup> | Amiya Ranjan Bhowmick<sup>2</sup> |  
Abhishek Mukherjee<sup>3</sup> | Achyut Kumar Banerjee<sup>4</sup>

Environmental Science and Policy 124 (2021) 64–72

Contents lists available at ScienceDirect

**Environmental Science and Policy**

journal homepage: [www.elsevier.com/locate/envsci](http://www.elsevier.com/locate/envsci)




An integrated policy framework and plan of action to prevent and control plant invasions in India

Achyut Kumar Banerjee<sup>a,\*</sup>, Anzar Ahmad Khuroo<sup>b</sup>, Katharina Dehnen-Schmutz<sup>c</sup>, Vidushi Pant<sup>d</sup>, Chinmay Patwardhan<sup>e</sup>, Amiya Ranjan Bhowmick<sup>e</sup>, Abhishek Mukherjee<sup>f</sup>

Journal of Environmental Management 294 (2021) 113054

Contents lists available at ScienceDirect

**Journal of Environmental Management**

journal homepage: [www.elsevier.com/locate/jenvman](http://www.elsevier.com/locate/jenvman)




Different factors influence naturalization and invasion processes – A case study of Indian alien flora provides management insights

Achyut Kumar Banerjee<sup>a,\*</sup>, Jyoti Prajapati<sup>b</sup>, Amiya Ranjan Bhowmick<sup>b</sup>, Yelin Huang<sup>a</sup>, Abhishek Mukherjee<sup>c</sup>

Received: 21 February 2022 | Revised: 20 May 2022 | Accepted: 24 May 2022

DOI: 10.1002/ecy.3794

DATA PAPER



## An occurrence data set for invasive and naturalized alien plants in India

Jyoti Prajapati<sup>1,2</sup> | Abhijit Singh<sup>2</sup> | Kshitij Patil<sup>1</sup> |  
Amiya Ranjan Bhowmick<sup>1</sup> | Abhishek Mukherjee<sup>2</sup> | Yelin Huang<sup>3</sup> |  
Achyut Kumar Banerjee<sup>3</sup>



# How you can use ILORA

Become a user and contributor

- Easy and **open access** to all data
- Query-based **data retrieval** system
- **Multidisciplinary applications**
- Provision for **data submission**
- Data **validation**
- Associated with **source information**

The screenshot displays the ILORA web interface. On the left is a dark sidebar with the ILORA logo and navigation links: Home, Data, Visualisation, and Release. Below these is a section titled 'How do you want to look at the data?' with a 'Species Search' dropdown menu. Underneath, it says 'Choose Species or clear with backspace and start typing' and shows 'Mikania micrantha Kunth' selected in the dropdown. A green 'Download' button is at the bottom of the sidebar.

The main content area shows a search result for 'Mikania micrantha Kunth'. At the top, it says 'Show 20 entries' and has a search box. Below this is a table of information:


Variable	Mikania micrantha Kunth
Taxonomic information (Class, Order, Family)	Magnoliopsida
Invasion status	Invasive
General information (Common name, Vernacular name, Growth habit, Duration, Group)	Climbing Hempweed, Uri hingchabi (mni); Lahare Banamaaraa (npi); Japanhlo (lus); Japani-lota (asm), Herb;Vine, Perennial, Dicot
Native range (TDWG Level 2 Names)	Mexico, Central America, Caribbean, Northern South America, Western South America, Brazil, Southern South America
Introduction pathway(s)	Release in nature for use;Landscape improvement in the wild
First record date	1884
Economic uses	Forage, Traditional Medicine, Soil Improver, Landscape Improver, Green Manure, Ornamental Use
Sold by nurseries	No
Habitat	Terrestrial Managed, Terrestrial Natural
Naturalized range (TDWG Level 2 names) [TDWG Level 4 names in ILORA_8_NaturalizedRange.csv]	Southwestern Pacific, Australia, China, South-Central Pacific, Malaysia, Indian Subcontinent, Western Indian Ocean, Indo-China, Eastern Asia, Papuasias

# How you can use ILORA


Become a user and contributor

- Standard for data submission (Version 1.1)
- [Data submission guidelines](#)

Please wait, the data submission form may take some time to load



Data formatting guidelines



### Data submission\_ILORA v1.1

Portal for submitting occurrence related information of invasive and naturalized alien flora of India

**Name \***

First

Last

*Full name in capital letters*

**Affiliation**

**Submit data for \***

One species, one observation

One species, multiple observations

Multiple species

**Email \***

For one species, one observation - data can be submitted through this form.

For one species, multiple observations - users need to upload formatted data in this form.

For multiple species, multiple observations - users need to upload formatted data in this form.

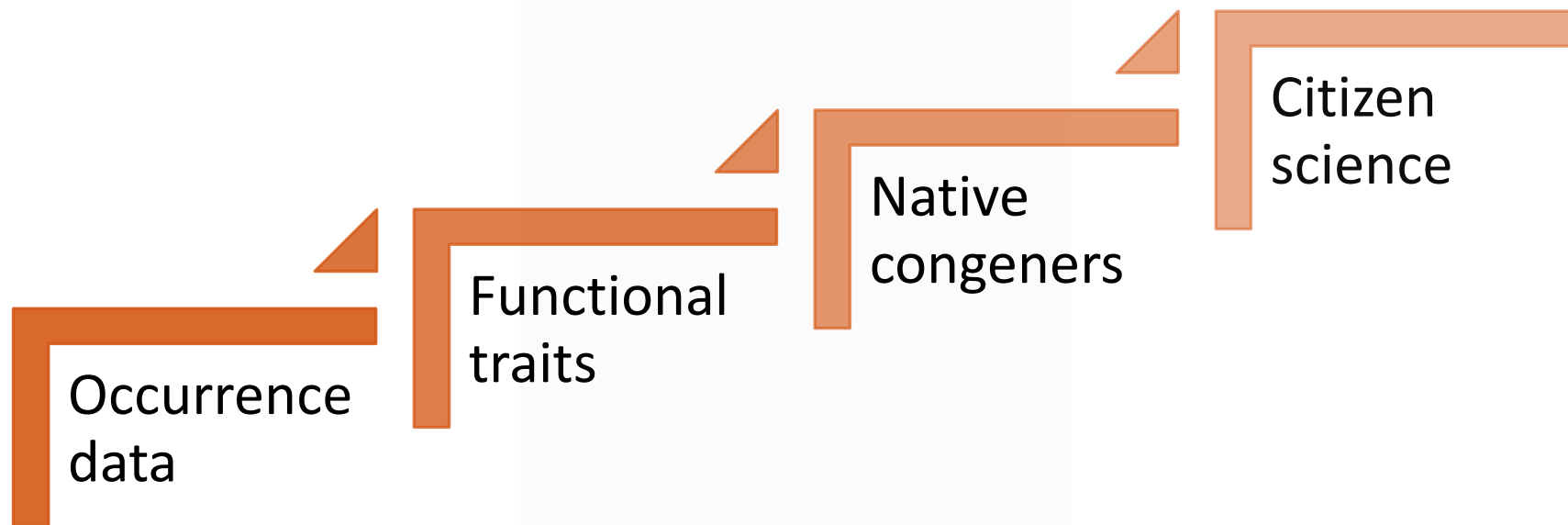
Data formatting guidelines are available [here](#).

search

The taskbar shows various application icons including a search bar, a folder, a blue folder, a mail icon, a Microsoft Excel icon, a calendar icon, a checkmark icon, a folder icon, a Google Chrome icon, a R logo icon, a P logo icon, a PDF icon, a weather icon showing 32°F, and system tray icons for network and volume.

# A long term vision

Be a part of this...



ILORA version 1.1

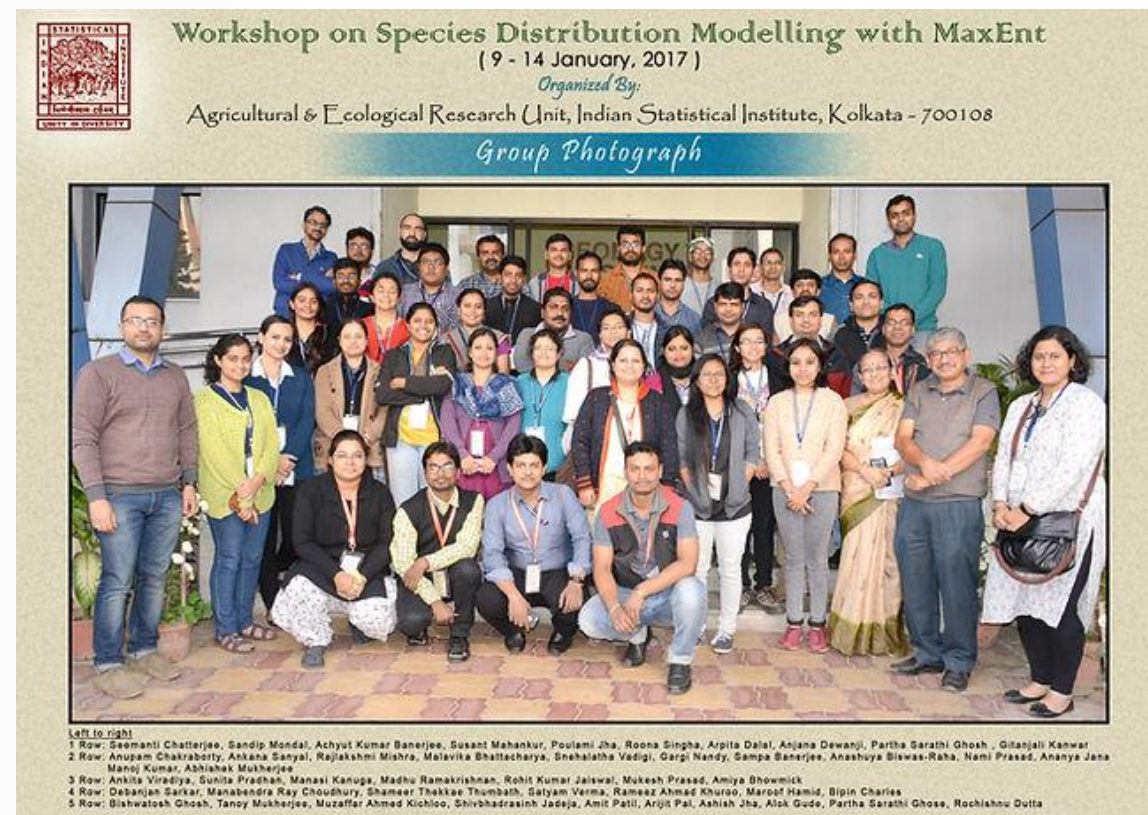
ILORA version 1.0



# Reaching ML Applications to Citizen Scientists

## Involve Citizens?...

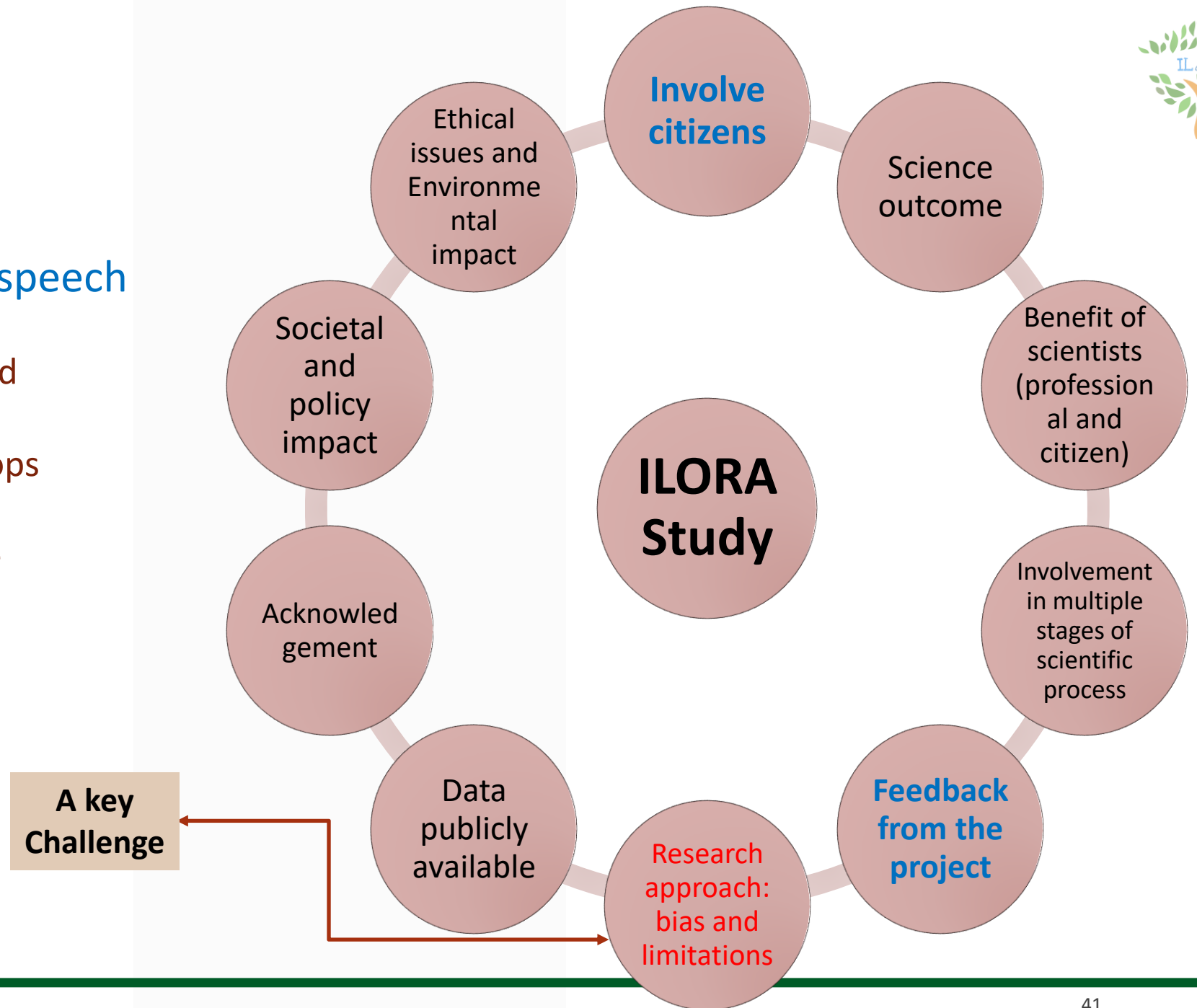
- Statistical Methods
  - Programming
  - Research Methodology
  - Biological Research
  - Ecological Data Science
  - Environmental research
  - NGOs
  - Students
  - Ph.D.
  - Postdocs
- Workshop on Species Distribution Modelling (2017, 2018)
  - Workshop on Statistical Methods for Interdisciplinary Researchers (Target group: Citizen Scientists, NGO) (10 workshops across India)



# 10 Principles

Thanks to Simona Carreto's speech

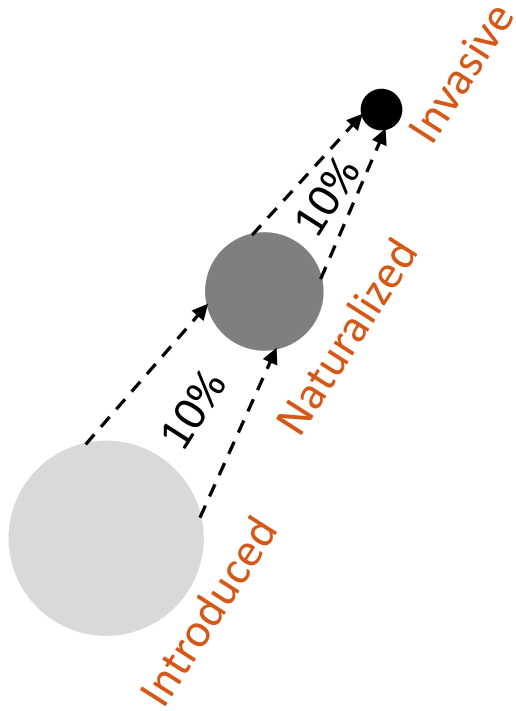
- Infusing machine learning and data science among applied researchers through workshops and conferences
- An action plan to convert the scientific findings into policy formulations.
- Not only data collection, infusing an objective way of doing science.



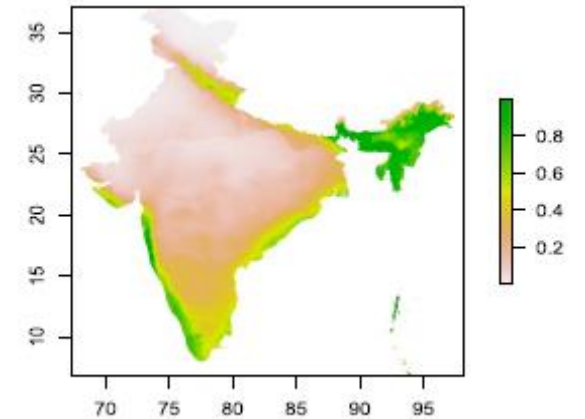
**A key Challenge**

# A collaborative framework

A long way to go...



- Estimate the risk associated with each naturalized species to be a successful invader
- **Identify their projected distributions under climate change**
- Develop an Application which compute the probability score for each species and may be installed in the borders
- It is not only a prediction problem, but also an inference problem.
- We would like to identify the important predictors so that best practices can be undertaken to minimize the risk.
- **Secure funding to develop the technology that include our research and integrate the policy/regulatory framework.**





# Acknowledgement

- Vidushi Pant, Sustainable India Finance Facility, Guntur, Andhra Pradesh
- Chinmay Patwardhan, Karlsruhe Institute of Technology, Germany
- Kshitij Anand Patil, Simon Fraser University, Canada
- Abhijit Singh, Indian Statistical Institute, Giridih



Achyut Kumar Banerjee  
Sun Yat-sen University, China



Abhishek Mukherjee  
Indian Statistical Institute, Giridih



Blessed to have wonderful and hardworking Ph.D. students  
Research: <https://sites.google.com/site/amiyaiitb/research>



*We are drowning in information, while starving for wisdom. The world henceforth will be run by synthesizers, people able to put together the right information at the right time, think critically about it, and make important choices wisely.*

# Thank you

QUESTIONS?

[ar.bhowmick@ictmumbai.edu.in](mailto:ar.bhowmick@ictmumbai.edu.in)

[amiyaiitb@gmail.com](mailto:amiyaiitb@gmail.com)

<https://sites.google.com/site/amiyaiitb/>



**Edward O. Wilson**

'Darwin's natural heir'

*Jun 10, 1929 - Dec 26, 2021*