Integrating Invasive Species Data: Solutions for Data Collection, Management, and Dissemination

Alycia W. Crall
Graduate Research Assistant
University of Wisconsin-Madison

Tom Stohlgren, Catherine Jarnevich, Tracy Davern, Geneva Chong (USGS), Jim Graham, Greg Newman, Paul Evangelista, Sara Simonson, David Barnett, Mohammed Kalkhan, Rick Shory, Hilary Drucker, Sunil Kumar, Nate Ament (CSU/NREL)
**The Problem**

- **Invasive species**
  - Impact biodiversity, ecosystem function, human health, and the economy
  - $120 billion spent annually in the US
  - 46% of species on threatened and endangered list are listed because of invasive species

- **Issues of Data Synergy**
  - Large data gaps (spatial and temporal)
  - Crosier and Stohlgren 2004 (Weed Technology)
  - Crall et al. 2006 (Frontiers in Ecology and the Environment)
    - 319 databases
    - 43% not online
Why not share data?

- **Sensitive Data**
  - Threatened and endangered species
  - Private property
  - Biological control agents
  - Species of cultural significance
  - Data not yet published

- **Inability to Share**
  - Data on paper, not digital
  - No website to provide data online
  - No database management system in place
Major partners

- Colorado State University-NREL
- United States Geological Survey-USGS
- National Aeronautics and Space Administration-NASA
Implementation of sensitive data security

- Different levels of access to website pages and functionality based on user’s login
  - User Levels: Guest, User, Tester, Instigator, Expert
- Data sharing agreement required
- Uploaded data assigned to project
  - Project Roles: Contributor, Reviewer, Authority, Manager
  - Can assign data sensitivity during upload
  - Given four “data fuzzing” choices: 7.5 min, 30 min, ½°, 1°
  - Exact locations can be used in general analyses

---

Gather Data

- Field Methods
  - Description
  - Equipment needed
  - How to setup
  - References

- Field Tools
  - Project Envelope
    - use with ArcGIS
    - survey full environmental range of species
  - EcoNab
    - field data collection program that runs on a PDA with Palm OS
  - GODMSurveyor
Field Tools

The Past

+ Manual entry

The Future

+ Automatic upload
Browse Data

- **By Organization**
  - university, non-profit, etc.

- **By Location**
  - politically defined locales such as states, counties, national parks, wildlife refuges

- **By Species**
  - scientific name, common name, NRCS code

- **By Project**
  - data contributed by project members and managed by project managers

- **By Map**
Example Map: *Tamarix* sp. locations
Example Map – Points, lines, polygons
Contribute Data

- **Survey Data**
  - tab-delimited text file or shapefiles
  - single species or many species
  - many locations or single political boundary
  - captures organism location, date, attributes, auxiliary data, treatment data

- **New Sightings**
  - enter single point sighting

- **Data Standards**
  - set of standards users can follow to upload disparate data in a standard format
### Contribute Data: Survey Data

**Organism Information Definition for 'textfile_test.txt'**

**Next Step:** Select a data structure below by clicking on a tab.

#### Dominance:

<table>
<thead>
<tr>
<th>Organism</th>
<th>Dominant</th>
<th>Secondary</th>
<th>Tertiary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tamarix</td>
<td>12</td>
<td>5</td>
<td>15</td>
</tr>
<tr>
<td>Tamarix</td>
<td>8</td>
<td>4</td>
<td>9</td>
</tr>
<tr>
<td>Tamarix</td>
<td>7</td>
<td>7</td>
<td>12</td>
</tr>
<tr>
<td>Euphorbia esula</td>
<td>34</td>
<td>2</td>
<td>13</td>
</tr>
</tbody>
</table>

**Specified Species:** Each column has information on a different species but of the same attribute. Example: Percent cover of various tree species.

<table>
<thead>
<tr>
<th>Organism</th>
<th>Cover (%)</th>
<th>Height (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tamarix ramosissima</td>
<td>80</td>
<td>1</td>
</tr>
<tr>
<td>Elaeagnus angustifolia</td>
<td>50</td>
<td>3</td>
</tr>
<tr>
<td>Tamarix</td>
<td>22</td>
<td>2</td>
</tr>
<tr>
<td>Euphorbia esula</td>
<td>12</td>
<td>2</td>
</tr>
</tbody>
</table>

#### Organisms by Row:

<table>
<thead>
<tr>
<th>Organism</th>
<th>Cover (%)</th>
<th>Height (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tamarix ramosissima</td>
<td>80</td>
<td>1</td>
</tr>
<tr>
<td>Elaeagnus angustifolia</td>
<td>50</td>
<td>3</td>
</tr>
<tr>
<td>Tamarix</td>
<td>22</td>
<td>2</td>
</tr>
<tr>
<td>Euphorbia esula</td>
<td>12</td>
<td>2</td>
</tr>
</tbody>
</table>

#### Single Species:

<table>
<thead>
<tr>
<th>X</th>
<th>Y</th>
<th>% Cover</th>
</tr>
</thead>
<tbody>
<tr>
<td>-104.5</td>
<td>40.56</td>
<td>15</td>
</tr>
<tr>
<td>104.36</td>
<td>40.38</td>
<td>9</td>
</tr>
<tr>
<td>-105.94</td>
<td>41.36</td>
<td>12</td>
</tr>
<tr>
<td>103.35</td>
<td>39.98</td>
<td>13</td>
</tr>
</tbody>
</table>

**Describe the columns in your file**

- Column Has: **Organism**
  - Contains: Organism Name
  - Format: Scientific Name
  - Dominance: Primary (1st)
  - Secondary (2nd)

- Column Has: **Treatment**
  - Contains: Control Agent
  - Value Format: Custom Code

- Column Has: **Attribute**
  - Contains: Number
  - Species column: DOMINANT
  - Attribute types are: In next pull down menu

**DOMINANT**

<table>
<thead>
<tr>
<th>Tamarix</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elaeagnus angustifolia</td>
</tr>
<tr>
<td>Tamarix</td>
</tr>
<tr>
<td>Euphorbia esula</td>
</tr>
</tbody>
</table>

**BIOCONTROL**

<table>
<thead>
<tr>
<th>Diothabda elongaia</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aphiodea lacertosa</td>
<td>500</td>
</tr>
<tr>
<td>Aphiodea nigriscul</td>
<td>500</td>
</tr>
</tbody>
</table>

**NUMBER**

<table>
<thead>
<tr>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
</tr>
<tr>
<td>500</td>
</tr>
<tr>
<td>500</td>
</tr>
</tbody>
</table>
Contribute Data: New Sightings

Add a new sighting

Ensure proper project
Your current working project is NISS - General User
Change?

Please select a species
The species for this sighting: Browse for a species
Presence/Absence:
○ Present ○ Absent

Describe location information

Projection Type:  ○ Latitude / Longitude ○ UTM

X (Longitude or UTM Easting):
○ Western Hemisphere ○ Eastern

Y (Latitude or UTM Northing):
○ Northern Hemisphere ○ Southern

UTM Zone:
− Select a Zone −

Datum:
○ WGS_84

Area Name (optional):
NewSighting for NISS - General User.

Enter the date of the observation

Date:
○ October ○ 30th ○ 2006

Submit
Contribute Data: Add Point Via Map
1) GISIN exchange protocol

2) Your database export to spreadsheet, survey addition

3) Survey addition from shapefiles or spreadsheets

4) Field form data entry on the web

5) Direct from GPS or PDA

6) Enter single coordinates or click on map

Multiple Ways to Contribute Data

...and there can be MANY contributors to ONE project
Analyze Data

- **Spreadsheets**
  - manage and create spreadsheets of data from the database
  - columns include project name, visit date, scientific name, common name, present, percent cover, height, location data, raster layers
  - can filter by species, area, start and end date, project

- **Georasters**
  - Upload georasters
  - Download georasters

- **Statistics**
  - Descriptive statistics (mean, min, max, etc.)
  - Simple and multiple linear regression
  - Regression tree
  - Simple and multiple logistic regression
### Analyze Data: Spreadsheets

**Many of the features on this page are under construction!**

<table>
<thead>
<tr>
<th>Spreadsheets</th>
<th>Options</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alewife Test Data Set - MOnsoon test12</td>
<td>New</td>
</tr>
<tr>
<td>western test proj 3 data empty spreadsheet?</td>
<td>Delete</td>
</tr>
<tr>
<td>GSENM tamarisk- test data- do not delete Testing for defect</td>
<td>View Data</td>
</tr>
<tr>
<td>Testing for defect 2 California Tamarisk Quads</td>
<td>Update Data</td>
</tr>
<tr>
<td>Colorado Dalmation Toadflax Tamarisk from Needles Quad</td>
<td>Download Data</td>
</tr>
<tr>
<td></td>
<td>Analyze Data</td>
</tr>
</tbody>
</table>
Analyze Data: Multiple Logistic Regression

Analysis Settings

Select analysis settings

<table>
<thead>
<tr>
<th>Analysis Type:</th>
<th>Multiple Logistic Regression</th>
</tr>
</thead>
<tbody>
<tr>
<td>Response Variable:</td>
<td>Present</td>
</tr>
<tr>
<td># of Predictors:</td>
<td>6</td>
</tr>
<tr>
<td>Predictor # 1:</td>
<td>Average annual precipitation,</td>
</tr>
<tr>
<td>Predictor # 2:</td>
<td>Average annual temperature,</td>
</tr>
<tr>
<td>Predictor # 3:</td>
<td>Elevation</td>
</tr>
<tr>
<td>Predictor # 4:</td>
<td>Distance to Water</td>
</tr>
<tr>
<td>Predictor # 5:</td>
<td>MODIS EVI three year mean,</td>
</tr>
<tr>
<td>Predictor # 6:</td>
<td>MODIS EVI three year range,</td>
</tr>
</tbody>
</table>

Continue

Back to spreadsheets
Analyze Data: Multiple Logistic Regression

### Analysis Results

**Response Variable:** Present

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model Intercept:</td>
<td>17.28823</td>
<td>0</td>
</tr>
<tr>
<td>Average annual precipitation,</td>
<td>0.07562426</td>
<td>0</td>
</tr>
<tr>
<td>Average annual temperature,</td>
<td>-0.1597557</td>
<td>0.0901</td>
</tr>
<tr>
<td>MODIS three year composite of EVI, Range</td>
<td>-0.0010652232</td>
<td>0</td>
</tr>
<tr>
<td>MODIS three year composite of EVI, Mean</td>
<td>0.001473194</td>
<td>0</td>
</tr>
<tr>
<td>Elevation,</td>
<td>-0.01004329</td>
<td>0</td>
</tr>
<tr>
<td>Distance to Water,</td>
<td>0.0001222648</td>
<td>0</td>
</tr>
</tbody>
</table>

**Null Deviance:** 2787.353

**Residual Deviance:** 1012.422

**Deviance explained:** 0.637

**AIC (Akaike's Information Criterion):** 1026.422

**AICc (AIC corrected for small sample size):** 1026.473

**Graphs of residuals versus fitted**

[Link to Predicted Surface]

[Back to spreadsheets]
Analyze Data: Predictive Surface

<table>
<thead>
<tr>
<th>Surface</th>
<th>Legend</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image-url" alt="Image of surface prediction" /></td>
<td><img src="legend-url" alt="Legend image" /></td>
</tr>
</tbody>
</table>

When you click "Save", your raster will be saved to your GeoRaster folder where you can then add it to our map application to view it along with your original data used to create the prediction surface.
Analyze Data: Map Display
Download Data

allows you to download organism locations as a textfile, csv file, or shapefile

| Input Field Data | Automatically add satellite data |

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>-99.969</td>
<td>38.856</td>
<td>-99.969</td>
<td>38.853</td>
<td>1.892</td>
<td>56.266</td>
<td>11.888</td>
<td>4.2353</td>
<td>2090.8</td>
<td>610.69</td>
<td>172.73</td>
</tr>
<tr>
<td>-100.07</td>
<td>38.914</td>
<td>-100.07</td>
<td>38.814</td>
<td>2.1</td>
<td>56.238</td>
<td>11.905</td>
<td>4.2223</td>
<td>1917.7</td>
<td>805.95</td>
<td>164.39</td>
</tr>
<tr>
<td>-99.930</td>
<td>38.714</td>
<td>-99.930</td>
<td>38.714</td>
<td>1.597</td>
<td>56.405</td>
<td>11.906</td>
<td>4.3495</td>
<td>2162.1</td>
<td>1037.4</td>
<td>194.96</td>
</tr>
<tr>
<td>-99.915</td>
<td>38.811</td>
<td>-99.915</td>
<td>38.811</td>
<td>1.4</td>
<td>56.311</td>
<td>11.974</td>
<td>4.3166</td>
<td>2070.4</td>
<td>825.44</td>
<td>181.42</td>
</tr>
<tr>
<td>-100.05</td>
<td>38.921</td>
<td>-100.05</td>
<td>38.921</td>
<td>0.4</td>
<td>59.294</td>
<td>11.597</td>
<td>4.0366</td>
<td>2044.1</td>
<td>974.59</td>
<td>188.92</td>
</tr>
<tr>
<td>-99.833</td>
<td>38.957</td>
<td>-99.833</td>
<td>38.857</td>
<td>2.1</td>
<td>57.733</td>
<td>11.844</td>
<td>4.2797</td>
<td>2140.7</td>
<td>670.49</td>
<td>176.06</td>
</tr>
</tbody>
</table>
Geared towards citizen scientist organizations in need of a data management system

New Features
- Online tutorials
- Monitoring protocols
- QA/QC protocols
- Customizable data entry forms
- Cheap, easy to use digital field tools
Additional New Feature

- Early Warning System
  - automatically sends emails out to users to warn them of a new non-native location on adjacent lands
Our Role in the Midwest

- Free Web-Based Data Management System
  - Caveats
    - Short term grants primarily focused on new features
    - Limited customer support staff
  - Joint Grant Opportunities
    - More cost effective to pool existing resources rather than reinventing wheel across US
    - Provide support staff and hardware for larger user base
Acknowledgements:

Tom Stohlgren, Catherine Jarnevich, Tracy Davern, Geneva Chong (USGS), Jim Graham, Greg Newman, Paul Evangelista, Dave Barnett, Rick Shory, Suni Kumar, Sara Simonson, Nate Ament, and Mohammed Kalkhan (CSU), with help from . . . John Kartesz (BONAP), Bruce Peterjohn, Pam Fuller (USGS), Curt Flather (USFS), John Schnase, Jeff Morisette, Ed Sheffner, Woody Turner (NASA) and many others!