Week 4: Introduction to Logistic Regression, Species Distribution and Habitat Selection

Lecture: Tuesday, 06 April 2021 and Thursday, 08 April 2021
8:30 am – 10:20 pm EDT

A Road Map for the Week

Tuesday

- Introduction to Logistic Regression
    - When Appropriate to Use (Presence/Absence Data)
    - Assessing Model Assumptions
- Creating a Spatial Database
    - Getting Raster Data into R
    - Manipulating Data Layers
    - Extract Remotely Sensed Data at Occurrence Locations

Thursday

- Full Example using Data on Addax Occurrence from Niger
    - Summarize Dataset
    - Apply Logistic Regression to Model Occurrence Probability
    - Graph Response Curves and Interpret Coefficients
    - Create a Predictive Surface

Instructors:
Ramiro Crego (CregoRD@si.edu)
Jared Stabach (StabachJ@si.edu)

Week 4: Introduction to Logistic Regression, Species Distribution and Habitat Selection

Lecture: Tuesday, 06 April 2021 and Thursday, 08 April 2021
8:30 am – 10:20 pm EDT

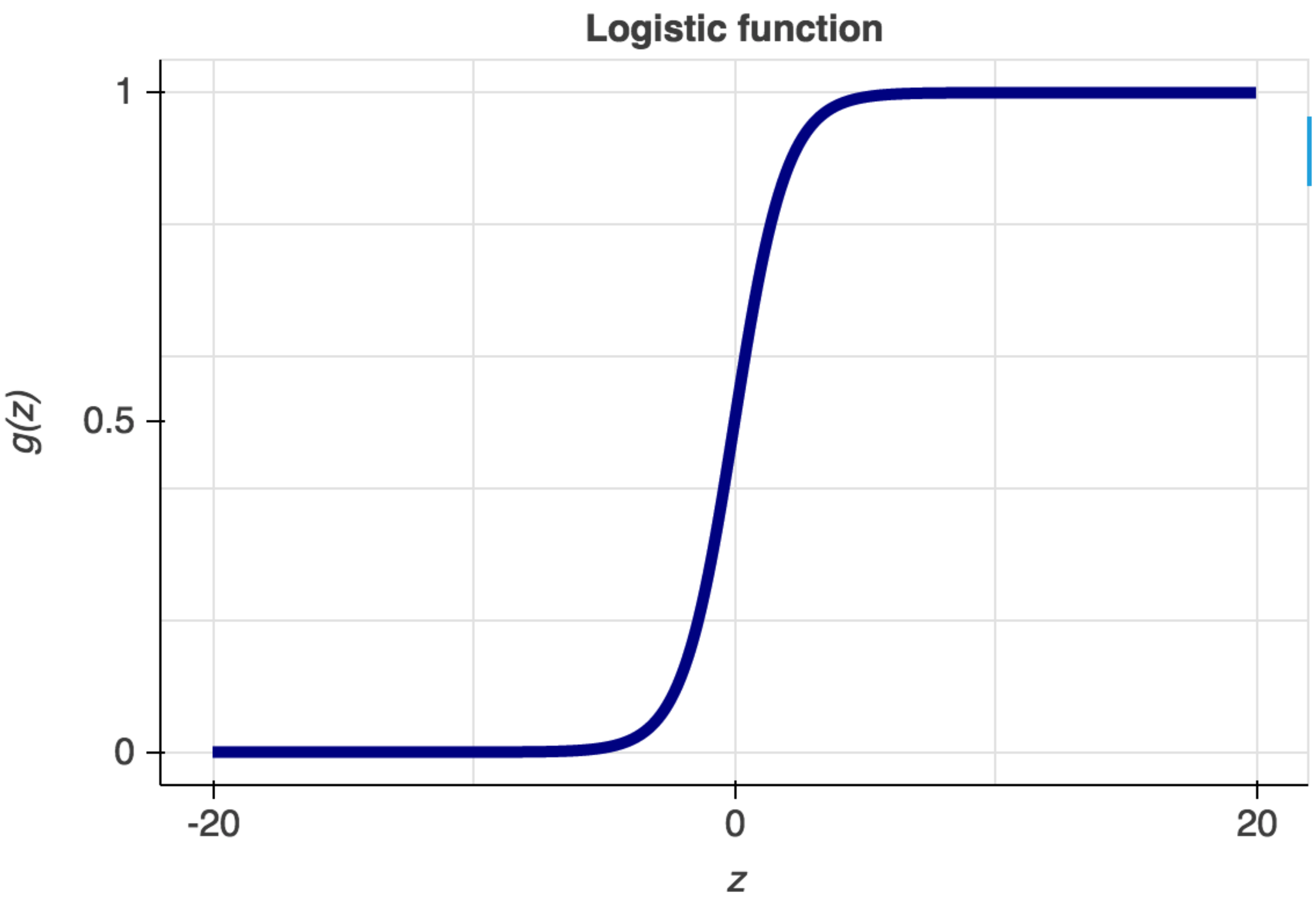
<u>Learning Objectives</u>

- Understand the Basics of Logistic Regression
- Access Available Spatial Data in R
- Become Comfortable with Manipulating Spatial Data
- Understand How to Fit and Interpret Model Results
- Graph Variable Response Curves
- Predict a Surface for Visual Interpretation

Instructors:
Ramiro Crego (CregoRD@si.edu)
Jared Stabach (StabachJ@si.edu)

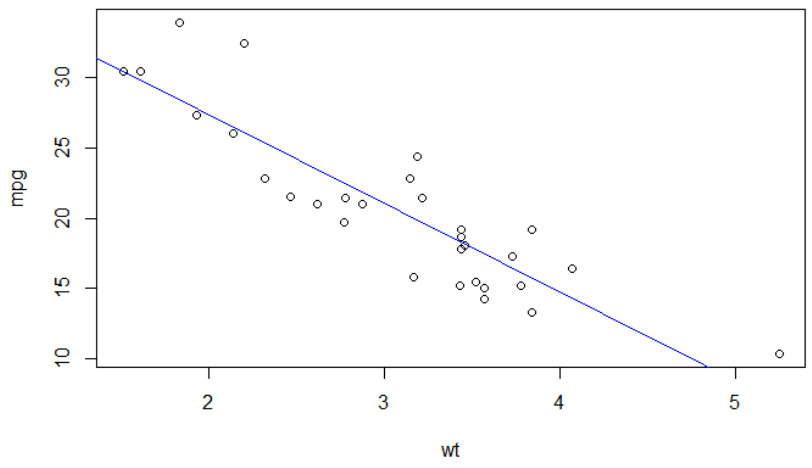
Smithsonian Institution

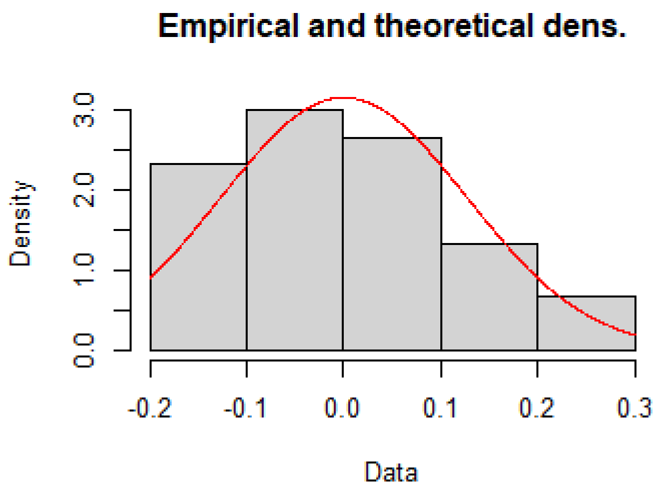
Some basics about Generalized Linear Models



General Linear Models

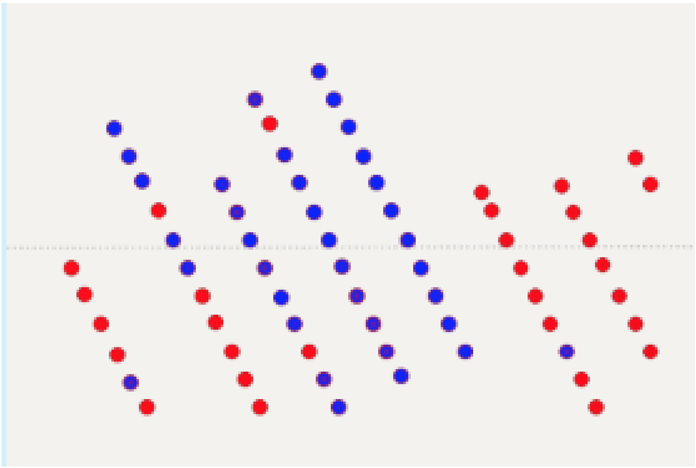
- **General linear models** (regression, ANOVA, etc.) are restricted by the degree to which the residuals conform to normality. The residuals follow a normal distribution.

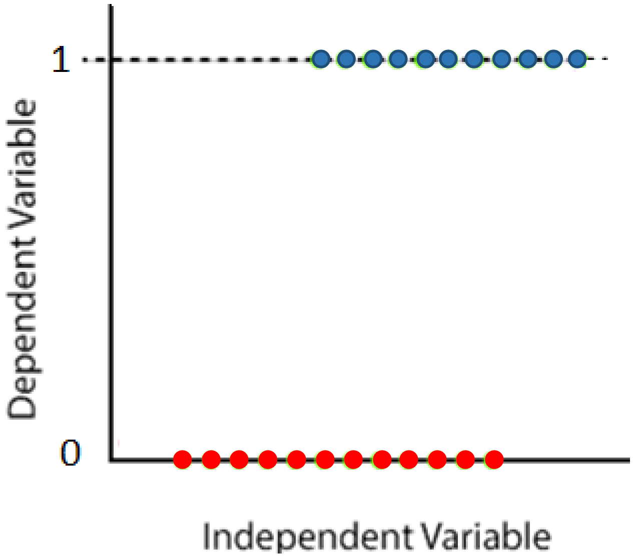




Generalized Linear Models (GLM)

- A primary goal of basic ecological and applied conservation research is to understand how species are distributed across space and through time. Generally, animal locations or abundance data do not follow a normal distribution. As a result, other type of model structures are needed.

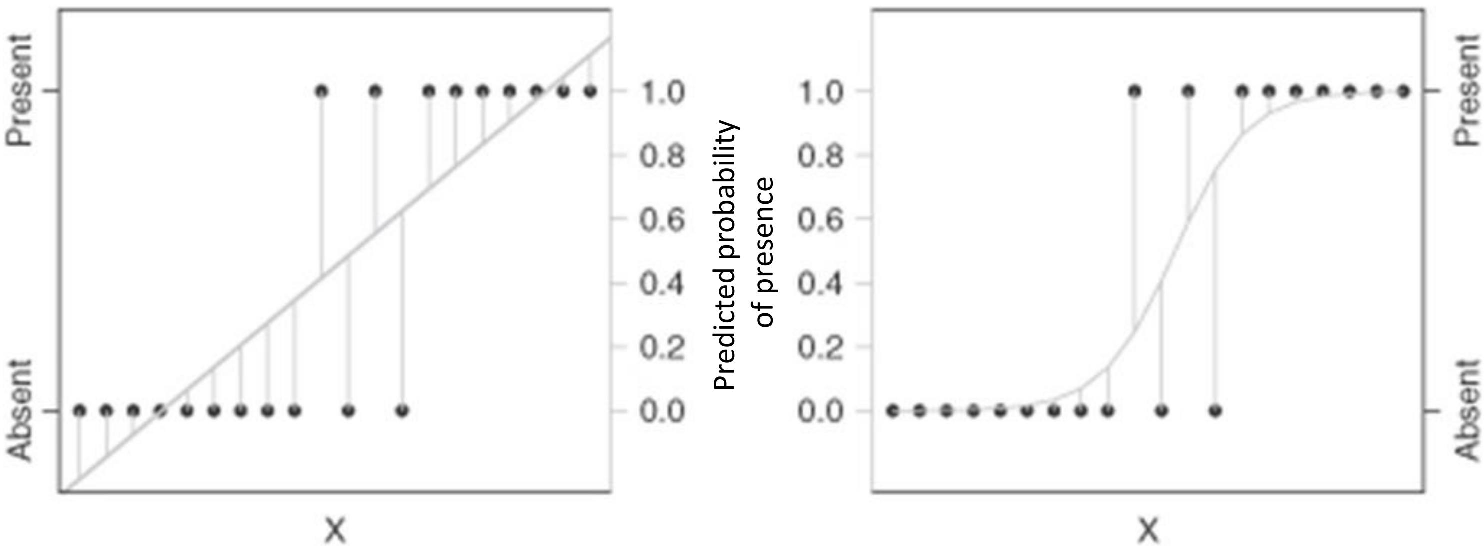




- Generalized linear models are used when the residuals are not normal, when there are non-linear relationships between dependent and independent variables, and/or when the variance in the dependent variable is not uniform across its range. E.g., **presence-absence data**.

Linear regression

GLM - Logistic regression



Predicted values can exceed
0 t o1 range

Prediction lies within
0 to 1 range

The GLM consists of three elements:

1. A probability distribution from the exponential family
    - In a logistic regression is a binomial distribution
2. A linear predictor η = Xβ
    - It is the quantity which incorporates the information about the independent variables ***β*** into the model.
3. A link function
    - It provides the relationship between the linear predictor and the expected value of the data
    - In a logistic regression we will use the **logit** link function

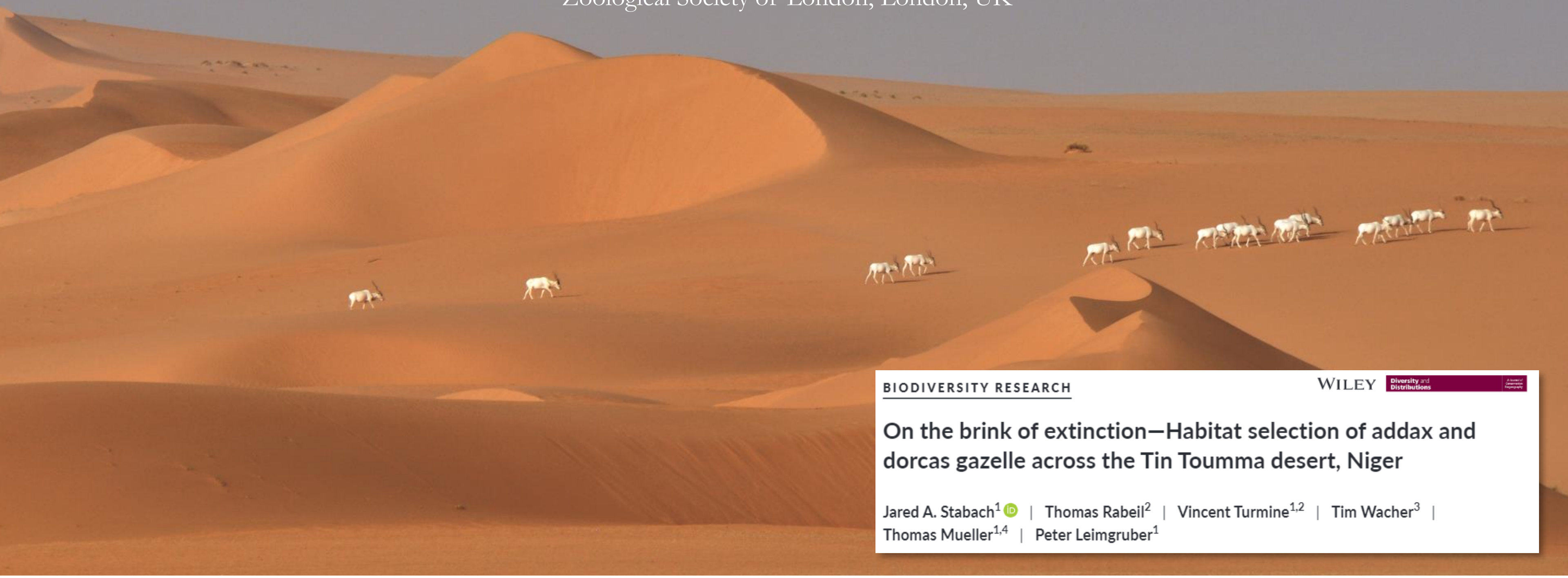
$$logit(p) = log(\frac{p}{1-p}) = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k.$$

$$p = \frac{exp(\beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k)}{1 + exp(\beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k)}.$$

Practical Example: On the Brink of Extinction
Habitat Selection of Addax Across the Tin Toumma desert, Niger

Jared A. Stabach[1], Thomas Rabeil[2], Vincent Turmine, Tim Wacher[3], Peter Leimgruber[1]

[1]Conservation Ecology Center, National Zoological Park, Smithsonian Conservation Biology Institute, Front Royal, USA
[2]Sahara Conservation Fund, L'Isle, CH
[3]Zoological Society of London, London, UK



BIODIVERSITY RESEARCH

WILEY Diversity and Distributions

On the brink of extinction—Habitat selection of addax and dorcas gazelle across the Tin Toumma desert, Niger

Jared A. Stabach[1] | Thomas Rabeil[2] | Vincent Turmine[1,2] | Tim Wacher[3] | Thomas Mueller[1,4] | Peter Leimgruber[1]

Smithsonian Conservation Biology Institute

SAHARA CONSERVATION FUND

ZSL LET'S WORK FOR WILDLIFE

Addax/Dorcas gazelle

Critically endangered
Formerly widespread across Sahelo-Saharan Africa
Highly adapted to hyper-arid conditions
Current population estimate: < 100 individuals

Main threats:
Overhunting

Vulnerable
Habitat generalist
Perceived affinity to Sahelian regions
Current population estimate: 35,000-40,000 individuals

Main threats:
Loss of habitat, disturbance, overhunting

Research Objectives

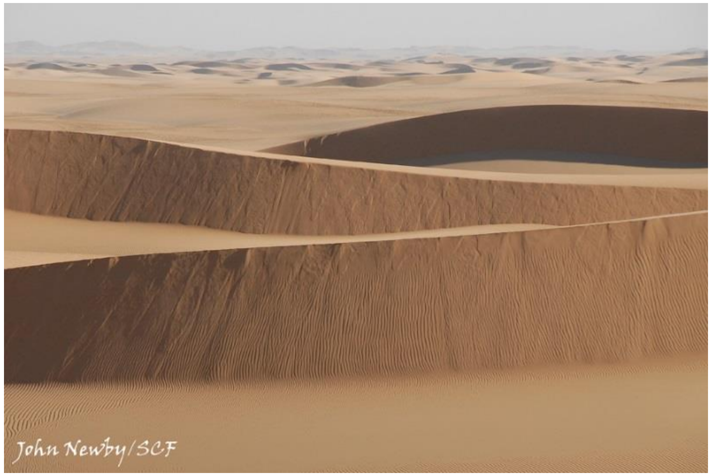
1) Identify the factors contributing to increased habitat suitability

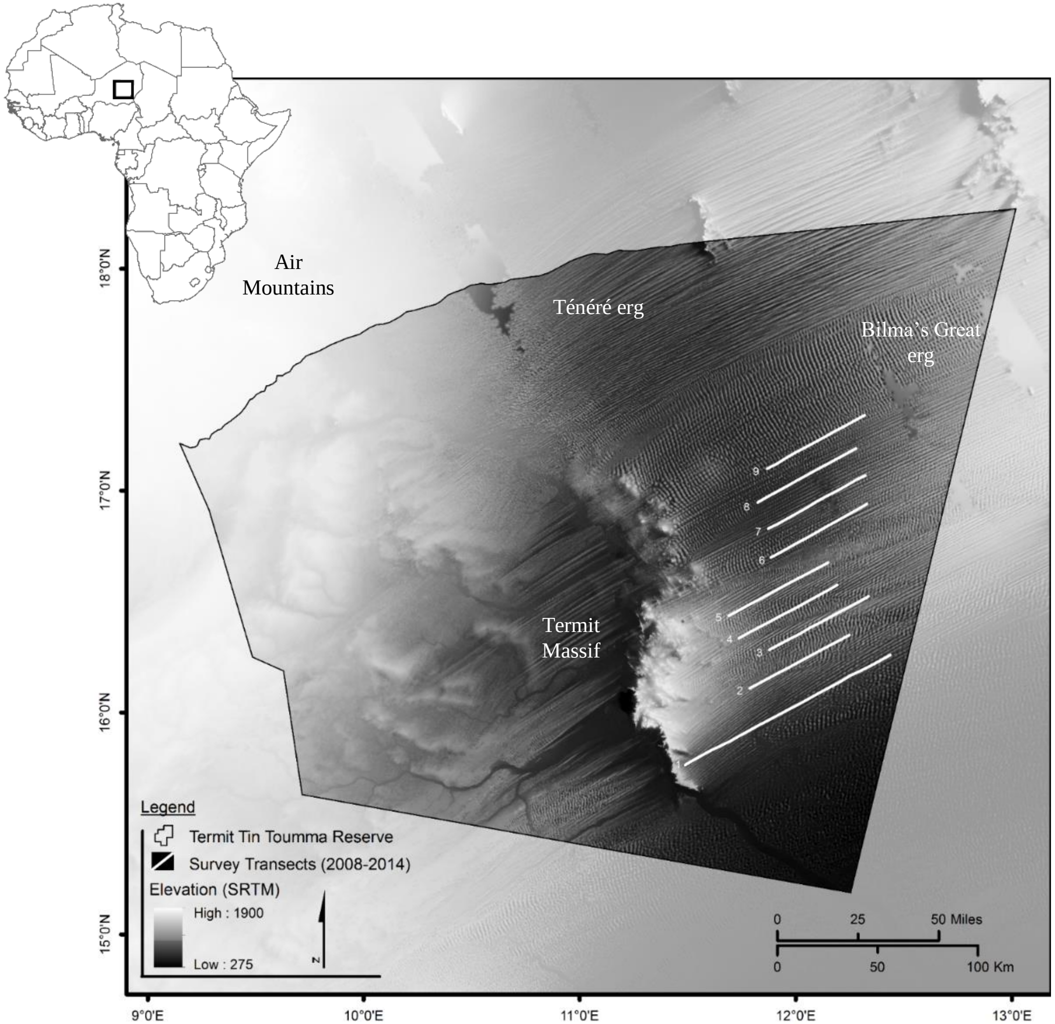
2) Examine shared resource use (A question for Homework)

3) Predict habitat suitability - guide future field surveys

Study Design

Line transect sampling (2008-2014)
Vehicle surveys
Plot locations – 5 km apart





Smithsonian Institution

Study Design

Line transect sampling (2008-2014)
Vehicle surveys
Plot locations – 5 km apart

Table 1. Occurrence and prevalence for Addax (*Addax nasomaculatus*) and Dorcas gazelle (*Gazella dorcas*) at sampling locations across the Tin Toumma desert, Niger.

		Transect	Plot	Occurrence (*n*)		Species Prevalence (%)		Site Prevalence[2] (%)	
Year	Season	ID[1]	Locations	Addax	D.gazelle	Addax	D.gazelle	One	Both
2008	Cold	1-5	66	22	28	33.3	42.4	68.2	7.6
2009	Dry	1-5	66	6	16	9.1	24.2	28.8	4.6
2009	Wet	1-5	66	8	19	12.1	28.8	39.4	1.5
2009	Cold	1-5	66	16	33	24.2	50.0	63.6	10.6
2010	Dry	1-5	66	4	15	6.1	22.7	24.2	4.6
2010	Dry	1-5	66	8	19	12.1	28.8	37.9	3.0
2010	Wet	1-5	66	3	13	4.6	19.7	24.2	0.0
2010	Cold	1-5, 8-9	88	14	29	15.9	33.0	44.3	4.6
2011	Dry	4-5, 8-9	44	7	4	15.9	9.1	15.9	9.1
2011	Wet	1, 4-5, 8-9	66	10	12	15.2	18.2	24.2	9.1
2012	Cold	4-9	66	8	9	12.1	13.6	21.2	4.6
2013	Cold	1-5	66	16	35	24.2	53.0	60.6	16.7
2014	Dry	4-9	66	4	9	6.1	13.6	16.7	3.0
2014	Wet	4-9	66	4	8	6.1	12.1	18.2	0.0
2014	Cold	2-6, 8	66	1	34	1.5	51.5	51.5	1.5
					Mean:	13.2	28.1	35.9	5.4

Environmental Variables

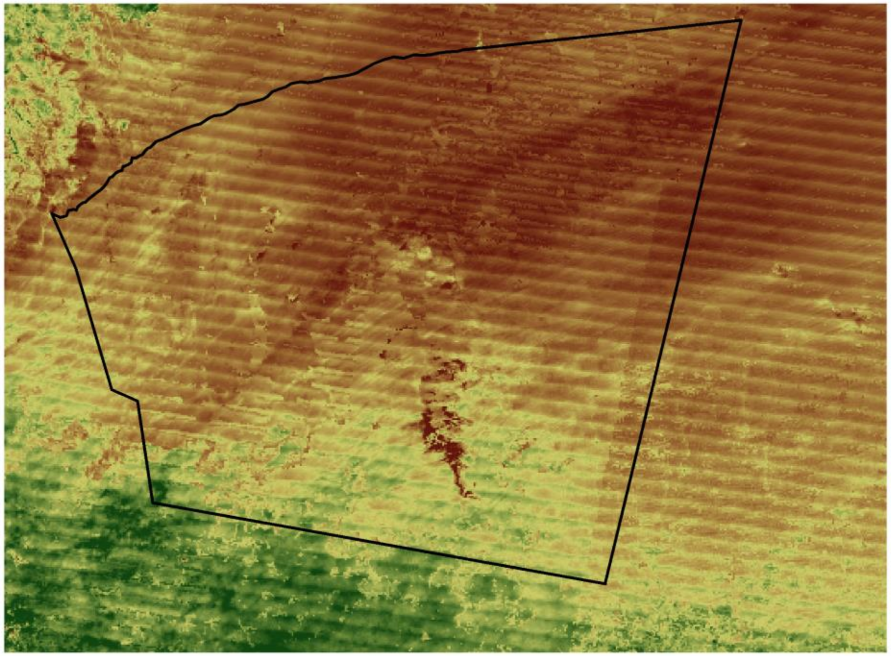
Remotely sensed variables

Surface Roughness

Difference in elevation between min/max values of a cell and its 8 surrounding neighbors

NDVI – Normalized Difference Vegetation Index

Vegetation greenness



2015 01 01

Field variables

Occurrence of Addax/Dorcas gazelle

Occurrence of Human disturbance

Occurrence of Perennials

Stipagrostis vulnerans

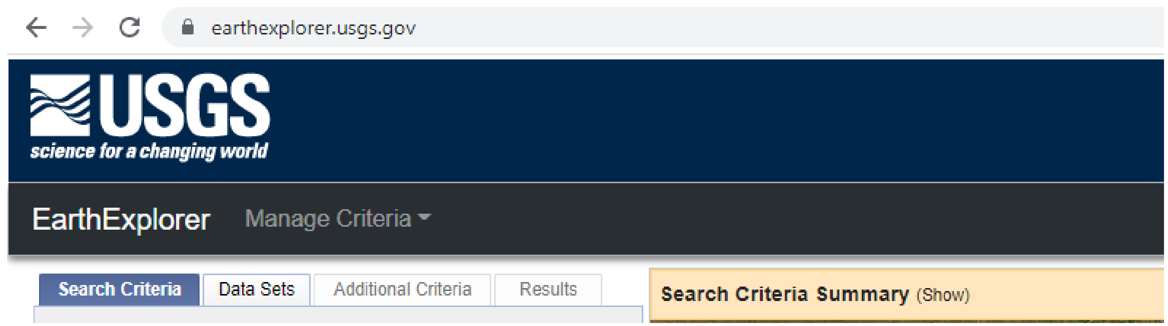
Stipagrostis acutiflora

Cornulaca monacontha

Survey Season (Cold/Dry/Wet)

Survey Year (2008-2014)

Environmental Variables



Where to get Data?
Many options

Remotely sensed variables

- Surface Roughness
  - Difference in elevation between min/max values of a cell and its 8 surrounding neighbors
- NDVI – Normalized Difference Vegetation Index
  - Vegetation greenness

Field variables

- Occurrence of Addax/Dorcas gazelle
- Occurrence of Human disturbance
- Occurrence of Perennials
  - *Stipagrostis vulnerans*
  - *Stipagrostis acutiflora*
  - *Cornulaca monacontha*
- Survey Season (Cold/Dry/Wet)
- Survey Year (2008-2014)

Model Structure

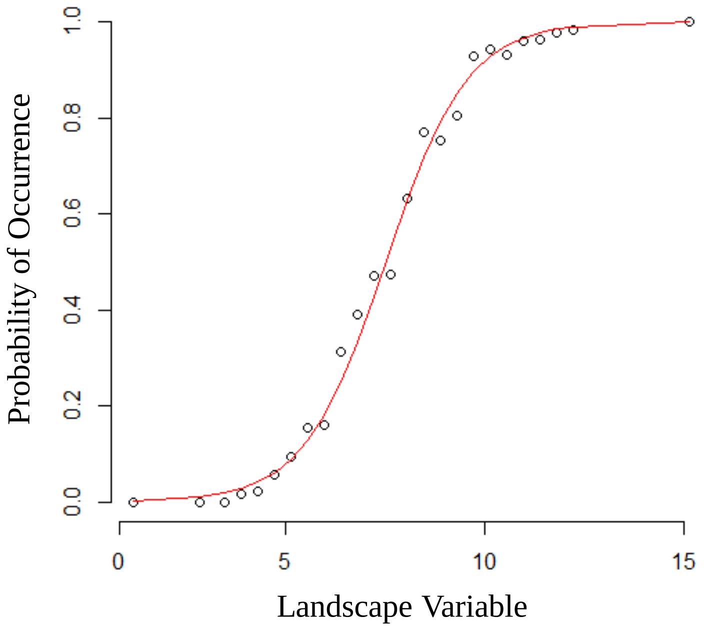
Generalized linear model (GLM) framework

Logistic regression

Addax models	K[a]	AIC_c	ΔAIC_c[b]
Rough + Rough^2 + NDVI + NDVI^2 + Humans + D.gazelle + *S.vulnerans* + *S.acutiflora* + *C.monacontha* + Season + Year	18	642.9	0.00
Reduced Model			
Rough + Rough^2 + NDVI + NDVI^2	5	744.6	101.7

[a]Number of estimable parameters.

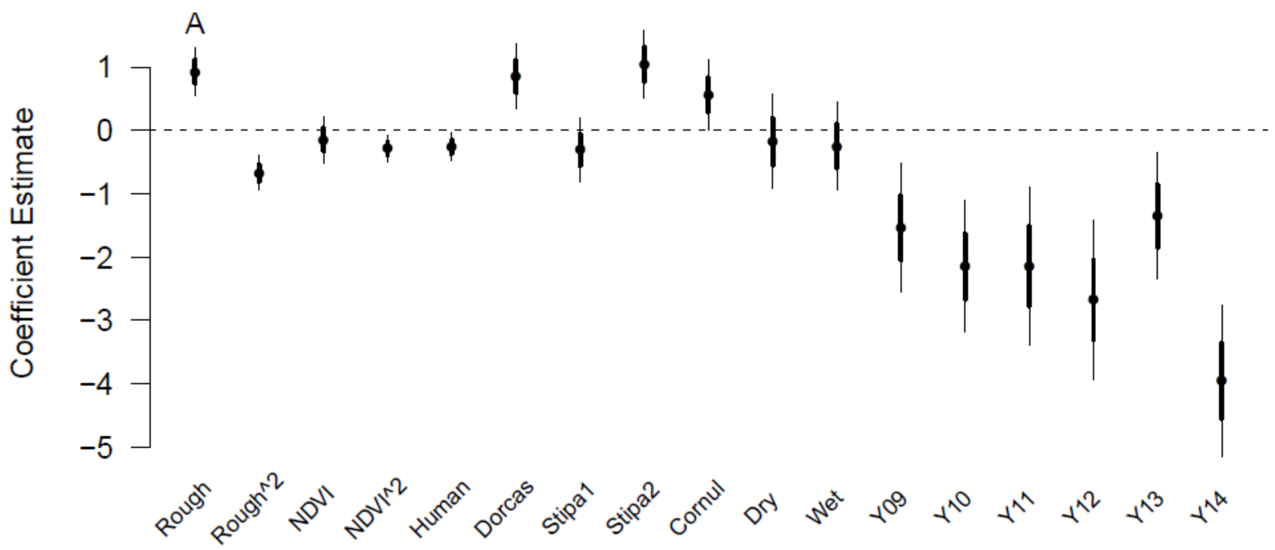
[b]Difference in value between Akaike's Information Criterion for small sample sizes (AIC_c) of the full and reduced variable model.



Results

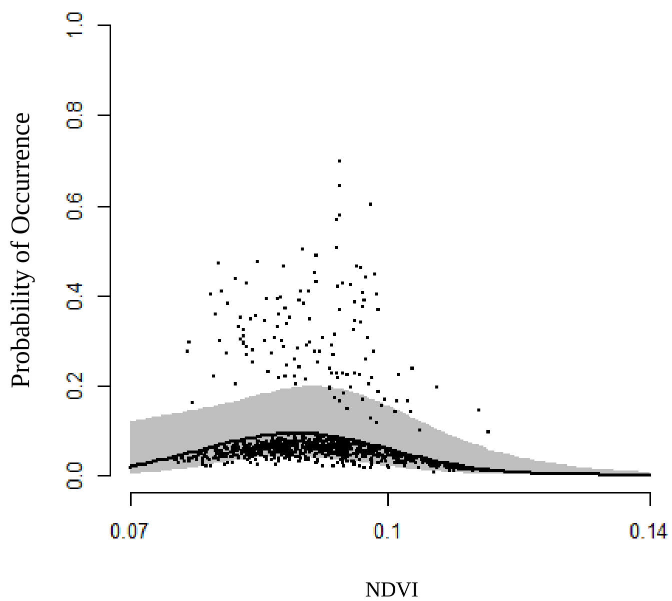
Evaluate Parameter Responses

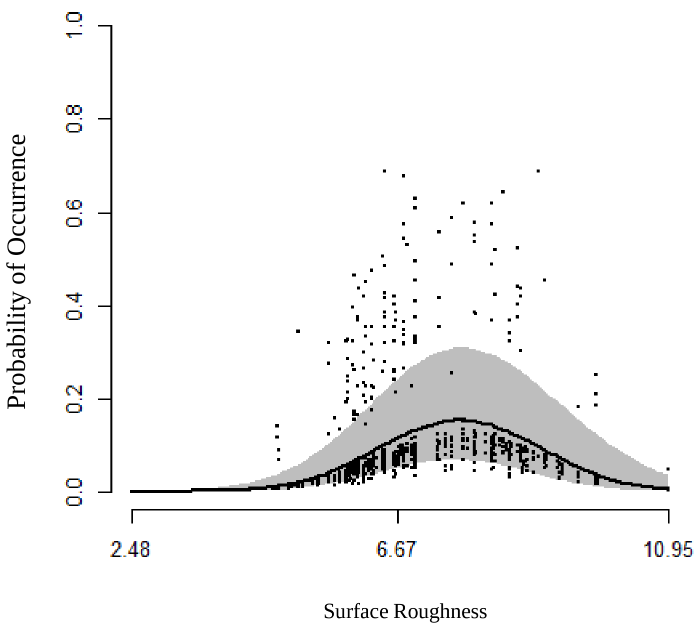
Addax



Results

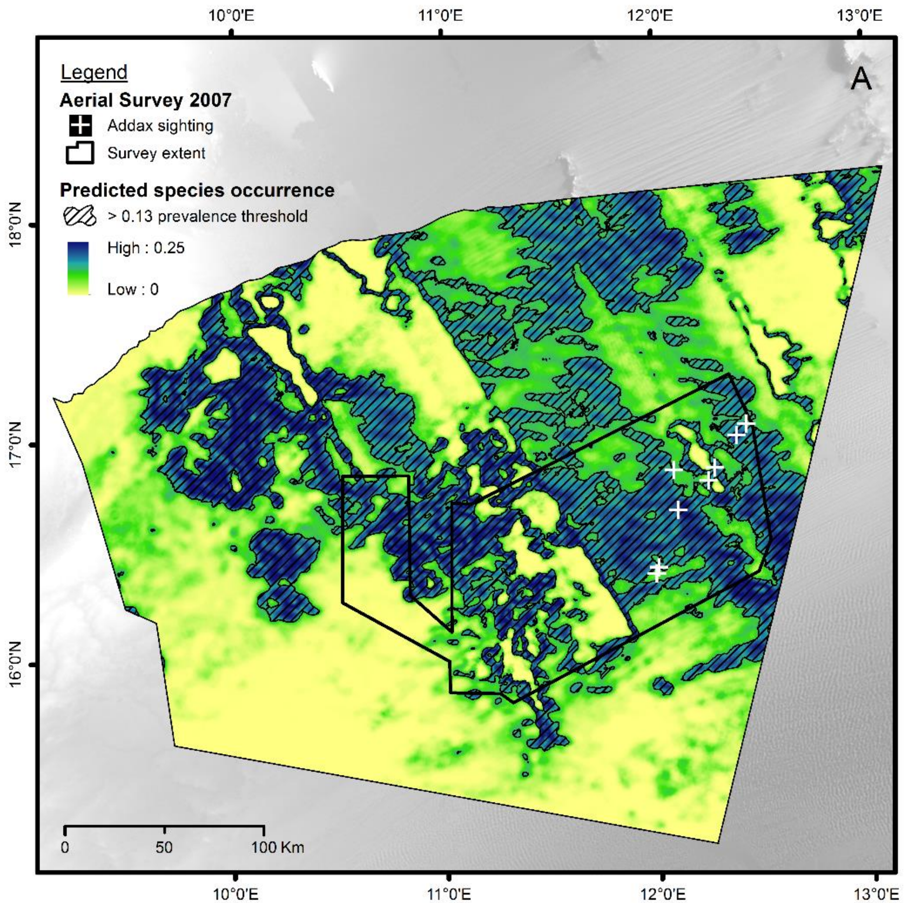
Evaluate Parameter Responses





Results

Predictive Surface



Summary

1) Addax - Year had the strongest effect in species occurrence model
2) Remote sensing data alone can help guide future field surveys
3) Urgent conservation action is required if this species is to persist into the future.

Acknowledgements

Funding and Support:

- Sahara Conservation Fund
- French Facility for Global Environment
- European Union
- Ministry of Environment – Niger
- Directorate General of Water & Forestry – Niger
- Department of Wildlife, Hunting and Protected Areas - Niger

The many local field assistants and staff members that made this research possible

Smithsonian Conservation Biology Institute

SAHARA CONSERVATION FUND

ZSL LET'S WORK FOR WILDLIFE