Modeling Risk of Japanese Encephalitis in Korea

Penny Masuoka, Terry A. Klein, Heung C. Kim, David Claborn, Nicole Achee, Richard Andre, Judith Chamberlin, Kevin Taylor, Susan George, Jennifer Small, Assaf Anyamba, Michael Sardelis, Andrew Au, and John Grieco

Uniformed Services University of the Health Sciences
US Army 18th Medical Command
University of Maryland Baltimore County
National Center for Medical Intelligence
University of Maryland University College

Funding:
DoD Global Emerging Infections Surveillance and Response System (GEIS)

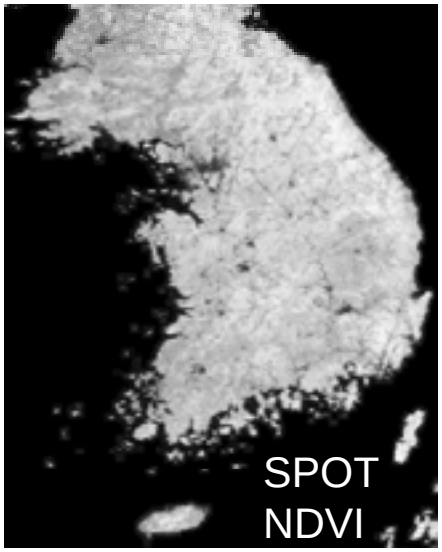
Background on Japanese Encephalitis

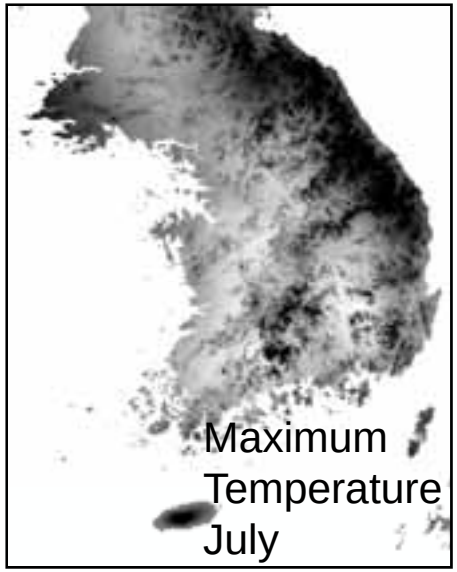
- A primary vector of the virus is the mosquito *Culex tritaeniorhynchus*.
- Over 35,000 cases are reported each year (10,000 deaths).
- Vaccine is used in Korean civilian population. Immunity wanes in older populations.
- US military personnel, not vaccinated [adverse reaction, cost ($100/series), no clinical cases].
- In temperate countries, peak mosquito populations appear in August with most clinical cases recognized in September.
- Prevention of JE depends in part on early identification of mosquitoes.
- Mosquito populations are affected by factors such as elevation, precipitation, temperature, and land cover (including rice).

Culex tritaeniorhynchus Ovipositing

Objective of this project

- Model potential risk areas for JE using land cover, elevation, NDVI, climate data and mosquito occurrence data

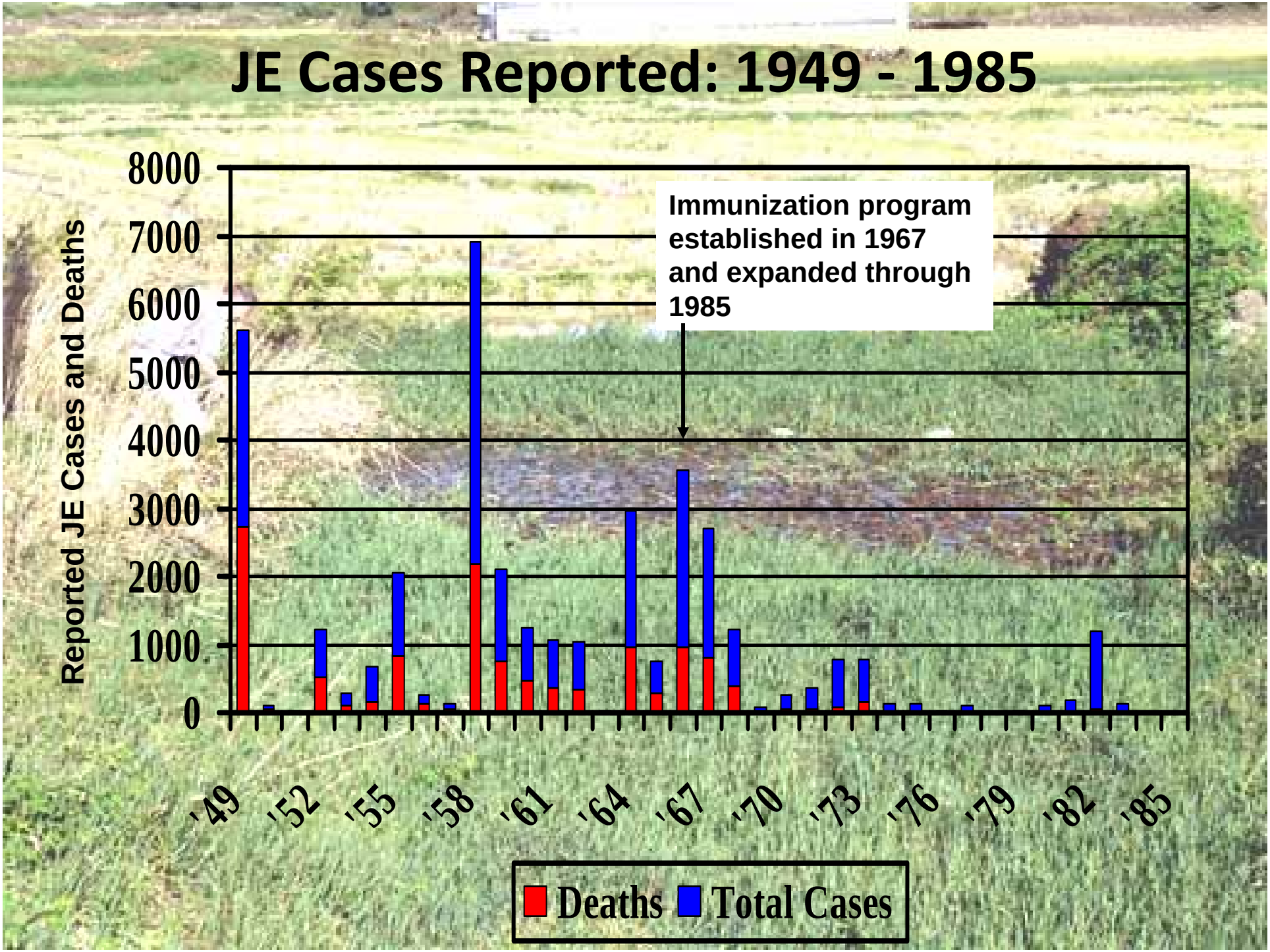




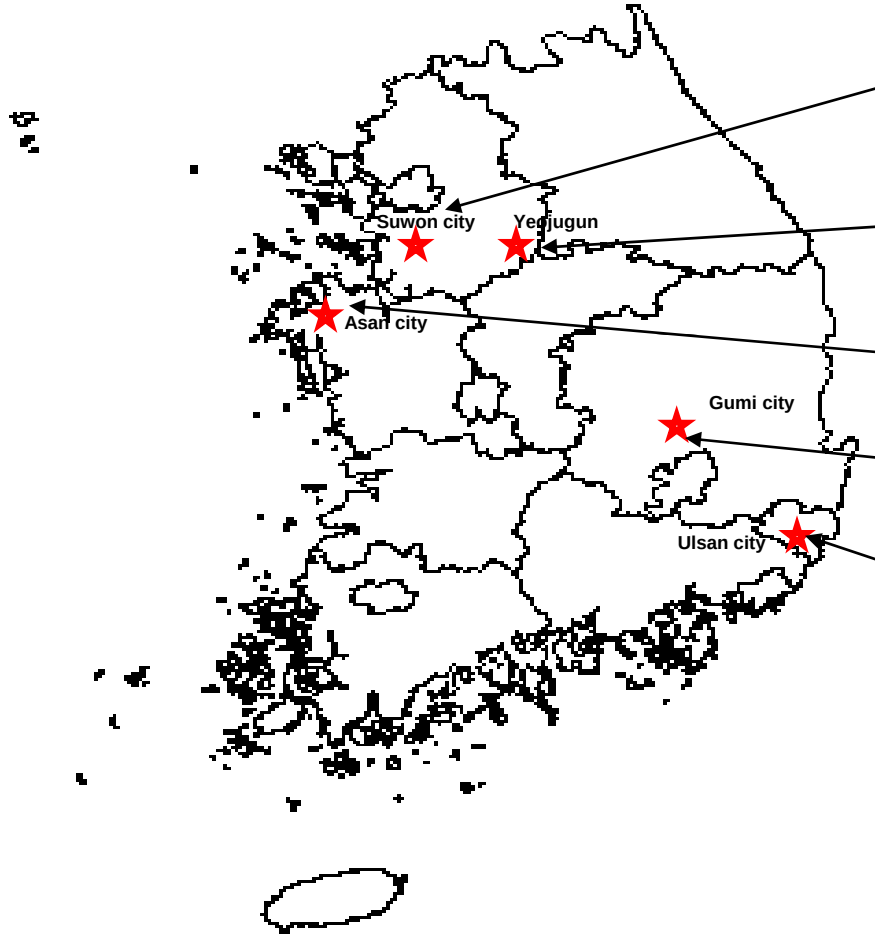
Geographical Distribution of Japanese Encephalitis Virus

(from CDC, Traveler's Health: Yellow book – Health information for international travel, 2005-2006)

JE Cases Reported: 1949 - 1985



Distribution of JE Cases, 2007



46 yr-old male: Taxi driver
(Confirmed 17 Sep 07)

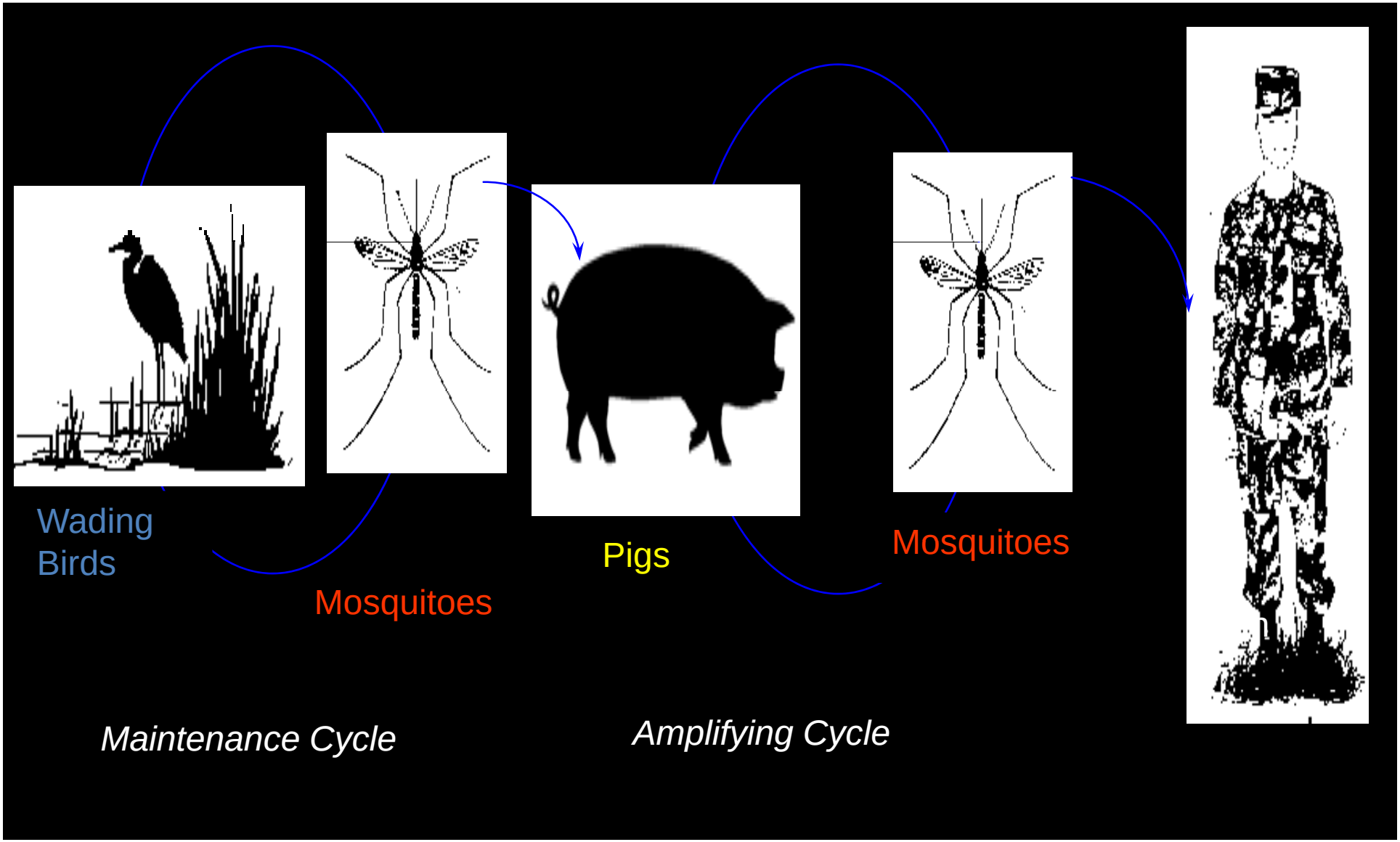
85 yr-old male: Student
(Confirmed 29 Oct 07)

51 yr-old female: Restaurant employee.
(Confirmed 6 Oct 07)

31 yr-old male: Temporary driver
(Confirmed 19 Sep 07)

52 yr-old male: Aquaculture
(Confirmed 28 Sep 07)

Transmission Cycles



Geographical & Environmental Factors

Adult Mosquito Collection Locations

Mosquitoes, by Species, Positive for Japanese Encephalitis, 2000 – 2004

Year	Mosquito Species	Number Tested	Positive Pools
2000	*Cx. tritaeniorhynchus*	4,281	14
	Ae. vexans nipponii	4,331	0
2002	*Cx. tritaeniorhynchus*	6,425	2
	Ae. vexans nipponii	7,530	0
	Ae. albopictus	4	0
2003	*Cx. tritaeniorhynchus*	0	0
	Ae. vexans nipponii	3,636	0
	Cx. pipiens	163	0
	Oc. koreicus	4	0
2004	*Cx. tritaeniorhynchus*	16,396	8
	Cx. bitaeniorhynchus	214	0
	Ae. vexans nipponii	4,633	0
	Cx. pipiens	1,116	0
	Ar. subalbatus	41	0
	Mn. unifromis	23	0

Environmental Data

- Land Cover Data
  - Boston University land cover from MODIS data
  - USGS land cover from AVHRR
- SPOT Vegetation NDVI
- WorldClim Data
  - 50 year averages of temperature and precipitation by month
- Shuttle Radar Topography Mission (SRTM) Elevation Data

Modeling species distribution Using Maxent

A number of programs have been written to model species distribution based on known locations and environmental factors:

Desktop GARP
OpenModeler
Maxent
others

We used Maxent:

- one of the better programs (Elith et al. 2006)
- good for small numbers (Hernandez et al. 2006)
- no absence data are used

Maxent

- Program that uses the maximum-entropy approach for species habitat modeling

- Input:
  - Known locations of a species in a comma-delimited file
    - Species name x-point y-point
  - Environmental layers (climate, elevation, land cover) in an ASCII grid format (export from ArcGIS)

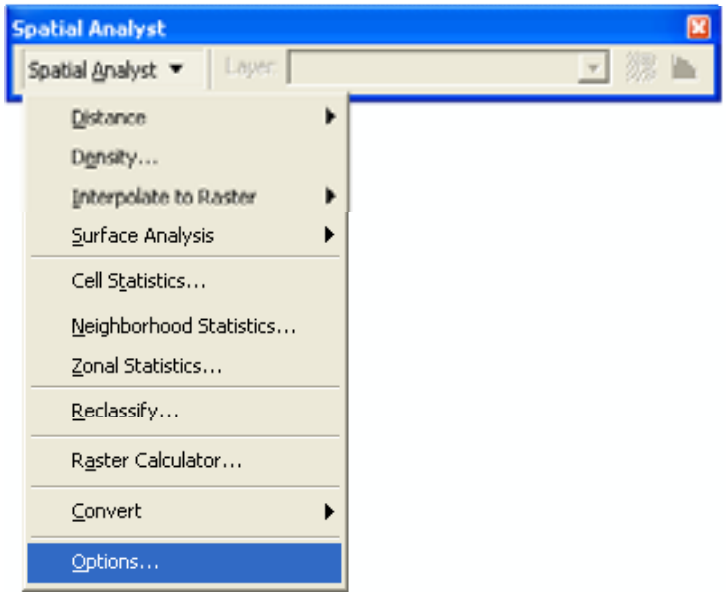
- Output:
  - Probability map of species potential distribution

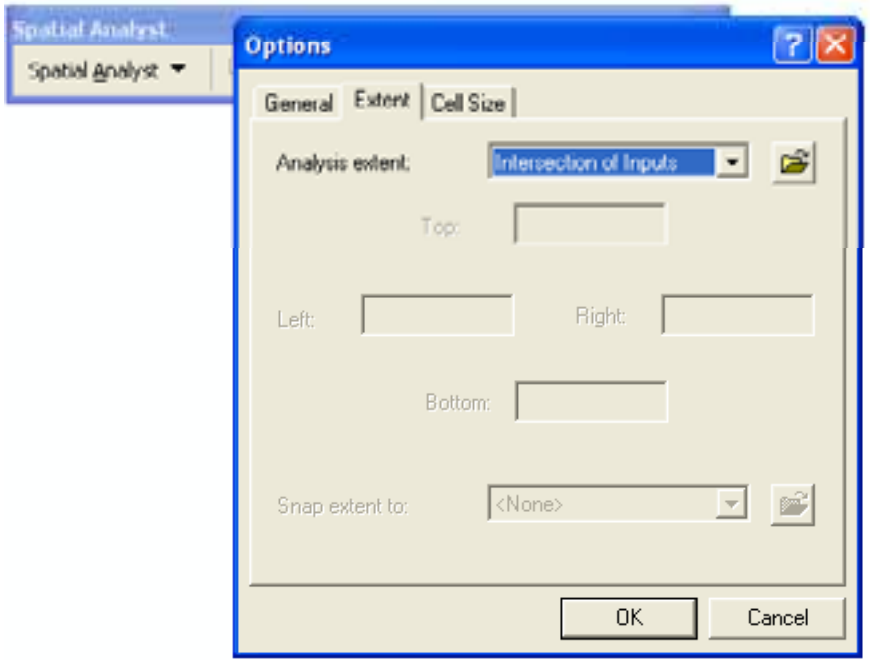
Modeling species distribution

In Maxent:

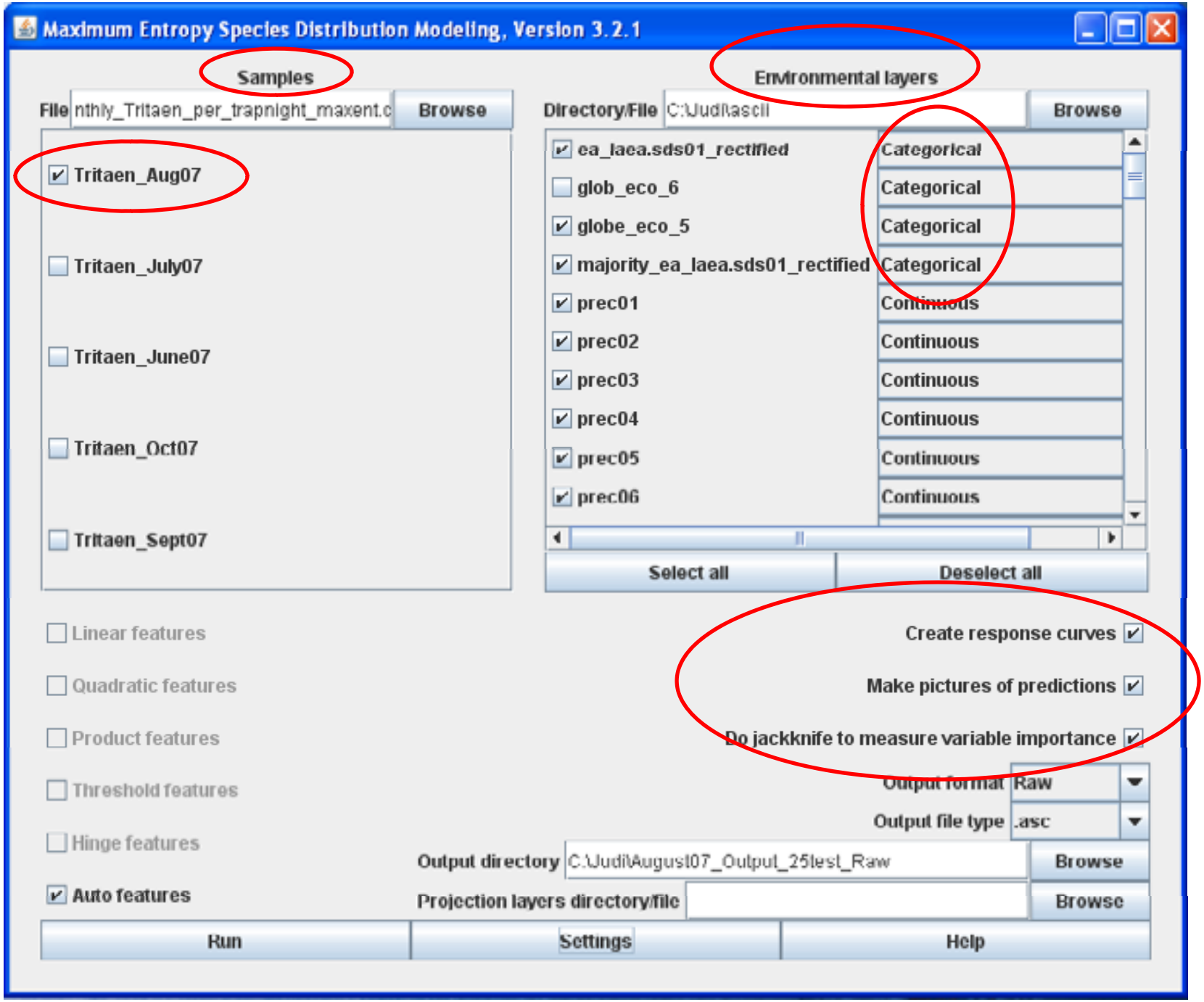
- All layers must have same projection.
- All raster layers must have same extent and cell size.

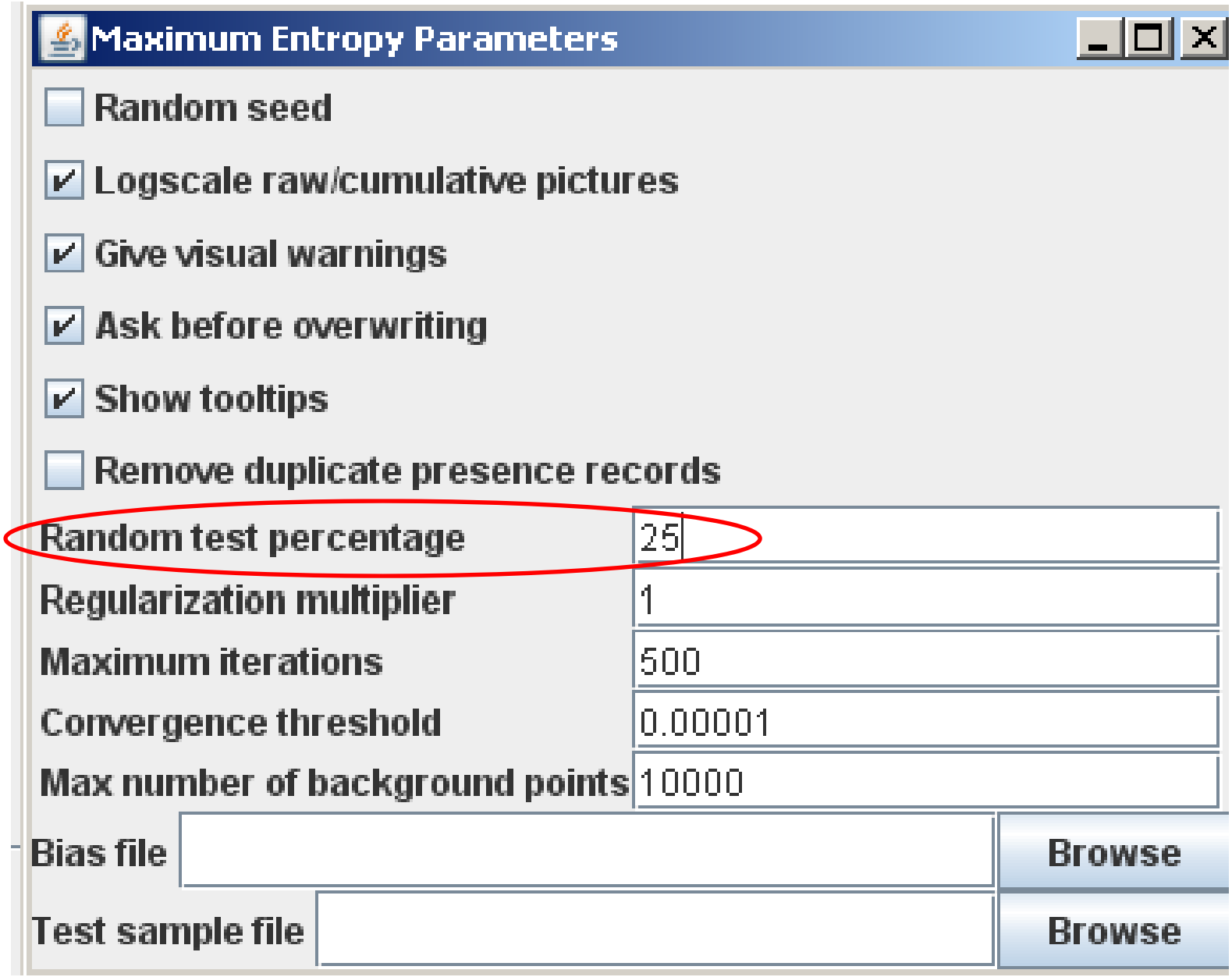
Use ArcGIS raster calculator to get same extent and cell size as one selected layer.





Maxent Main Screen

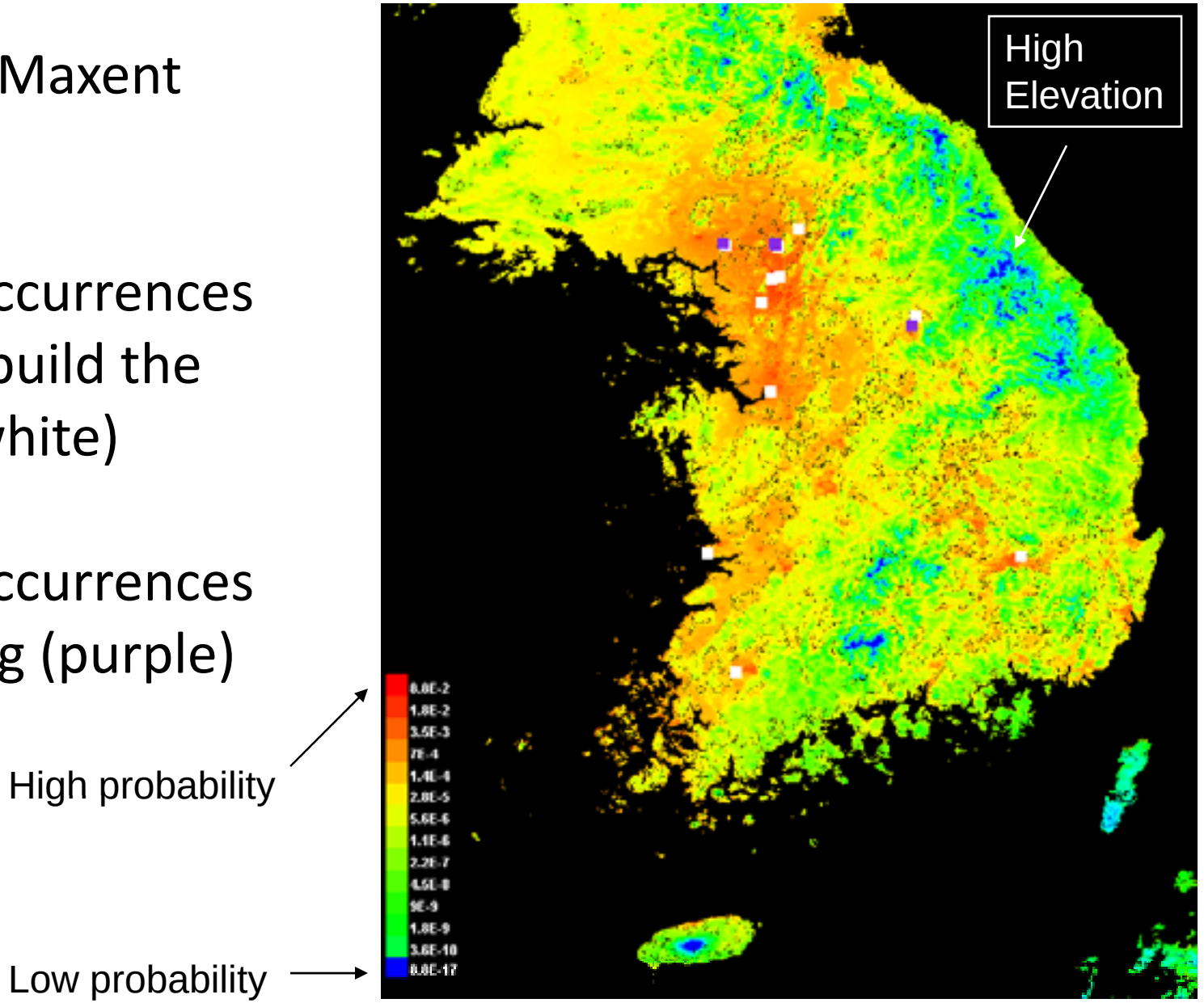



Maximum Entropy Parameters
Random seed
Logscale raw/cumulative pictures
Give visual warnings
Ask before overwriting
Show tooltips
Remove duplicate presence records
Random test percentage
25
Regularization multiplier
1
Maximum iterations
500
Convergence threshold
0.00001
Max number of background points
10000
Bias file
Browse
Test sample file
Browse

Example Maxent Output

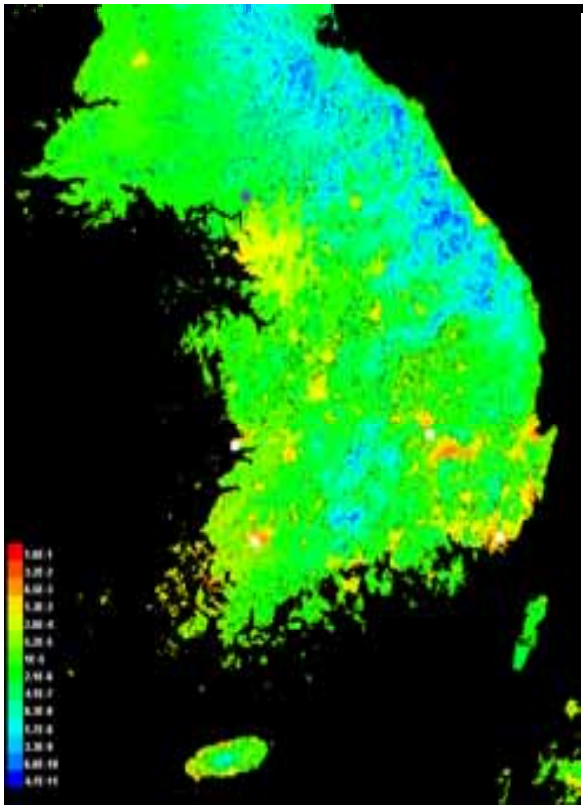
75% of occurrences using to build the model (white)

25% of occurrences for testing (purple)

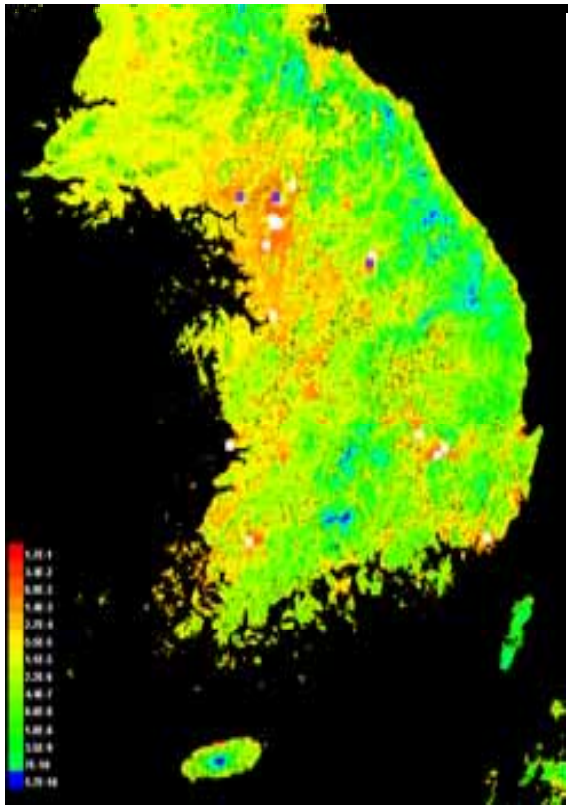


August 2007

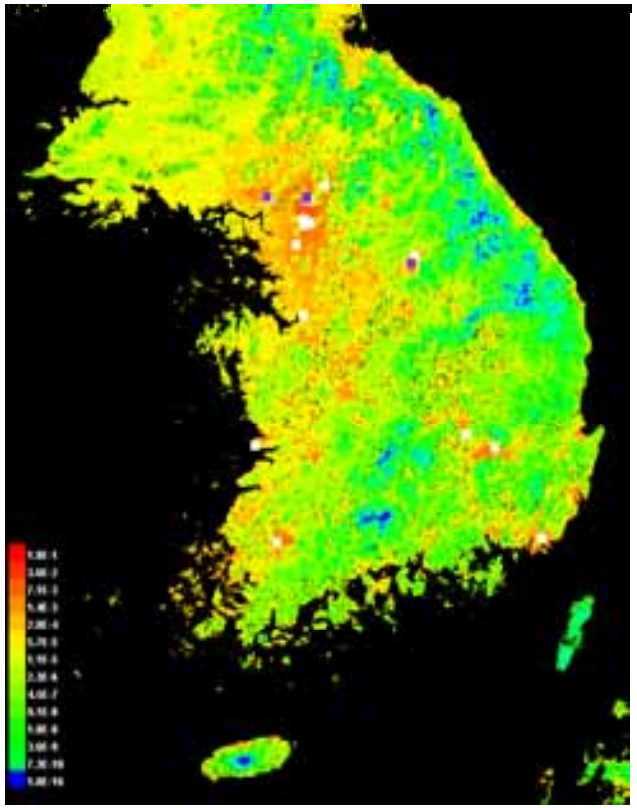
Predicted Distribution of *Cx. tritaeniorhynchus* 2005



July 2005



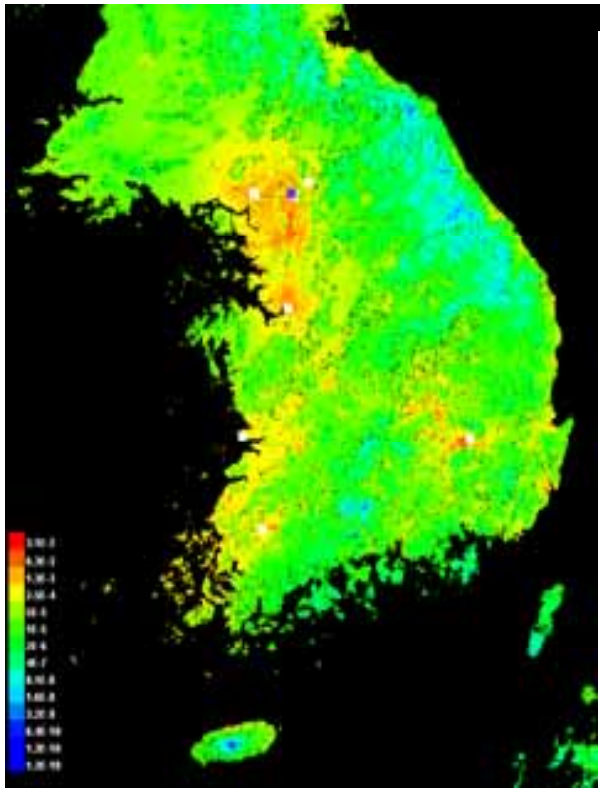
August 2005



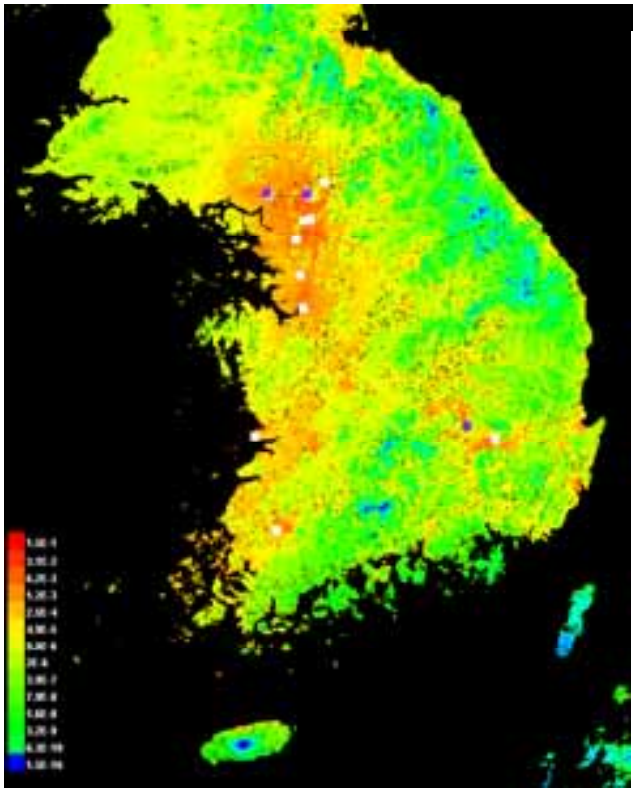
September 2005

Predicted Distribution of *Cx. tritaeniorhynchus* 2006

Not enough mosquitoes to run Maxent for July



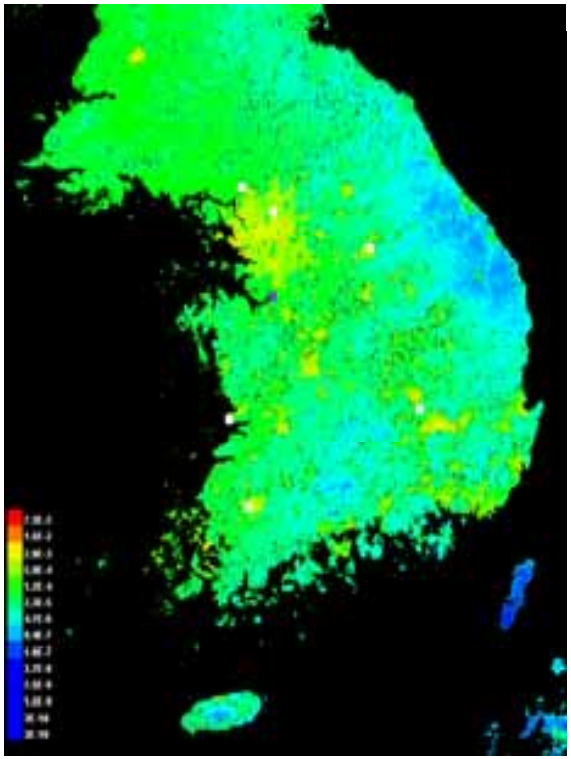
August 2006



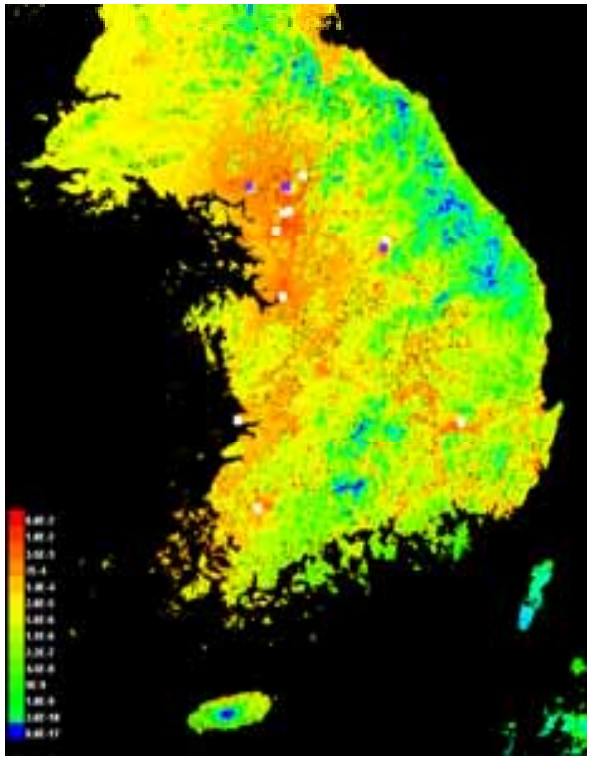
September 2006

July 2006

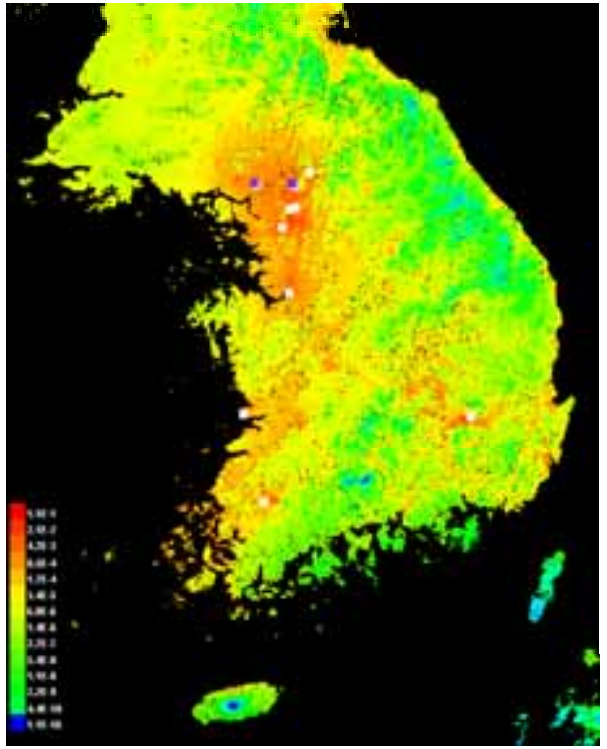
Predicted Distribution of *Cx. tritaeniorhynchus* 2007



July 2007

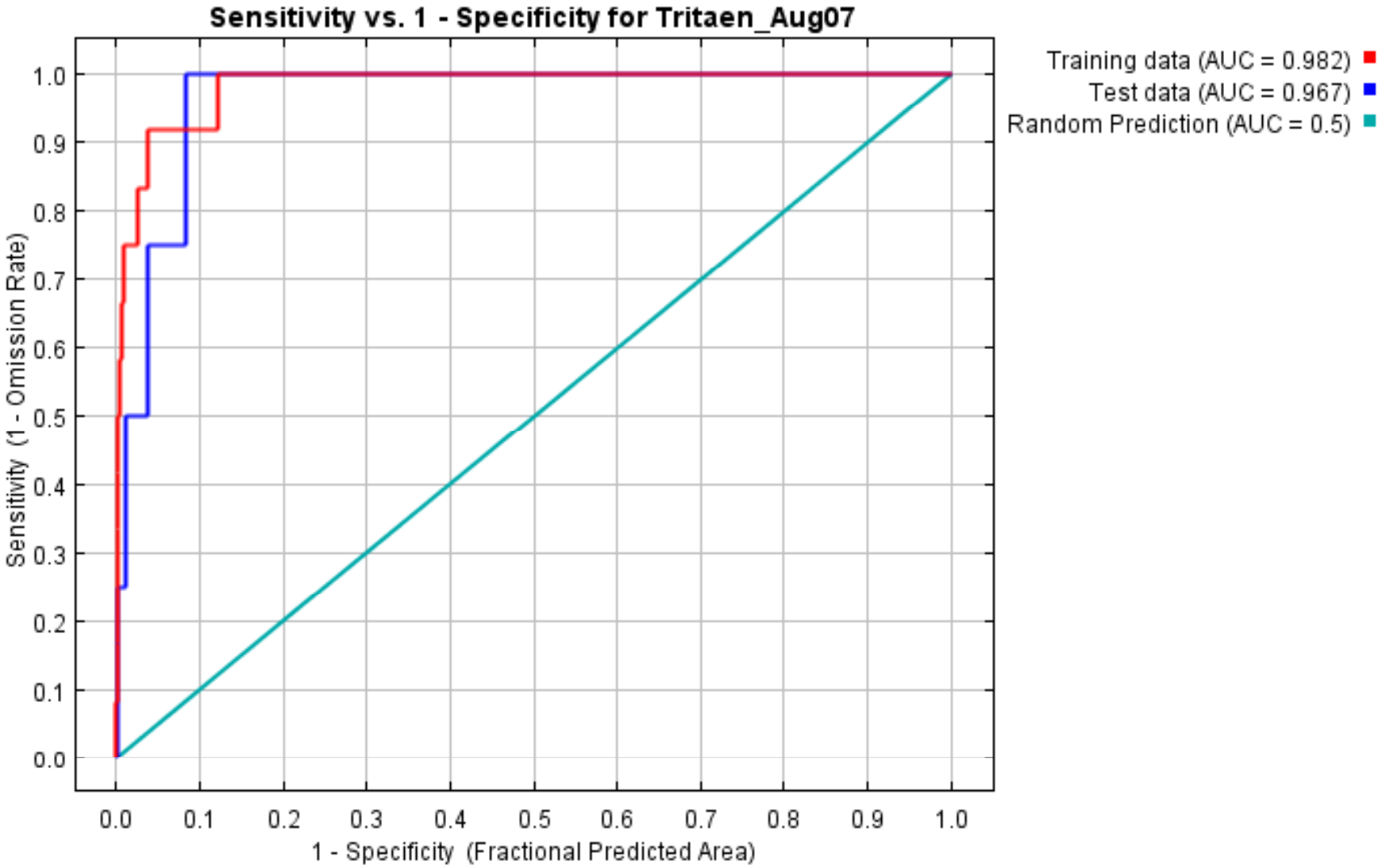


August 2007



September 2007

Assessment of the model reliability



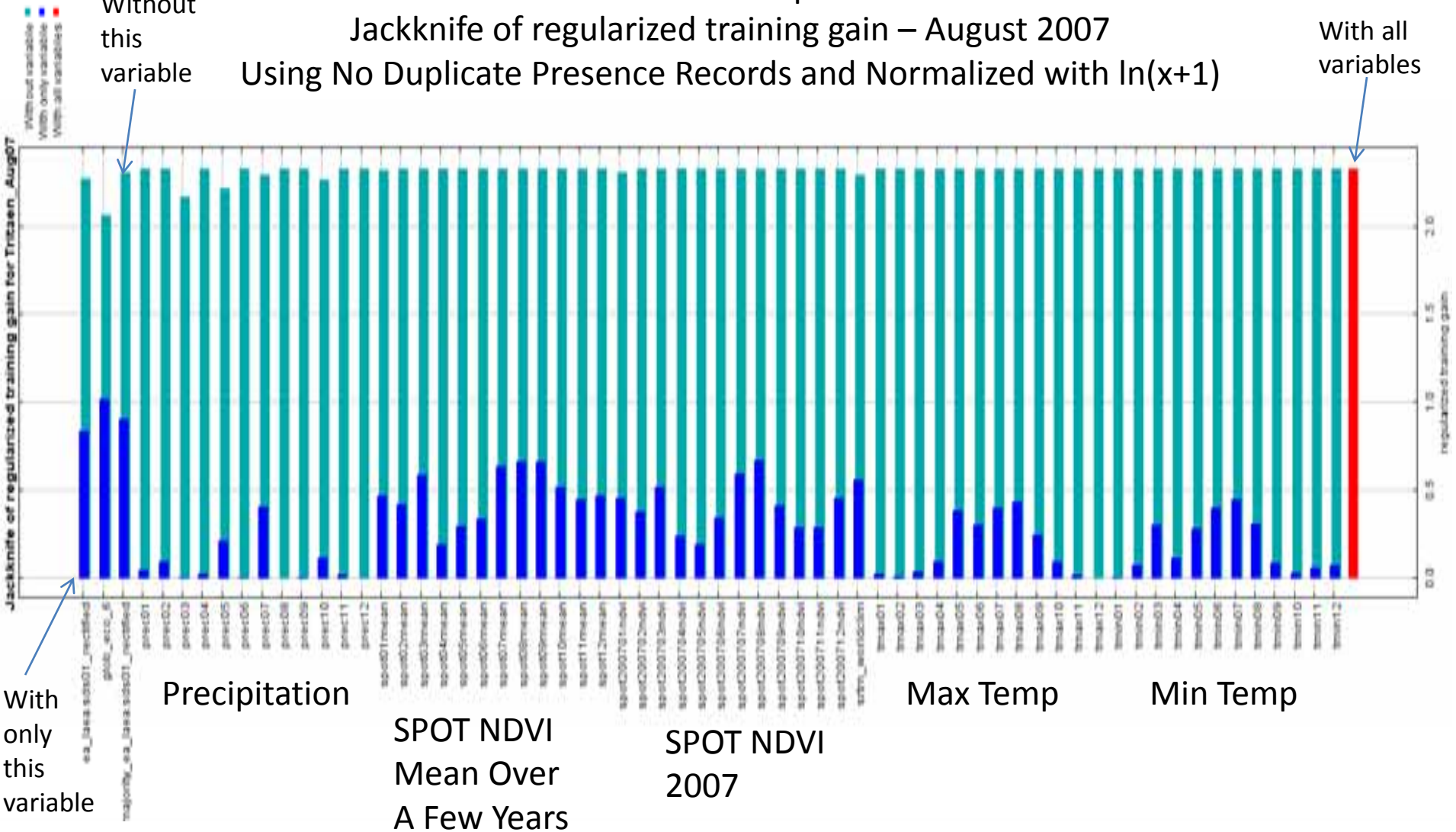
Maxent Output

Jackknife of regularized training gain – August 2007

Using No Duplicate Presence Records and Normalized with ln(x+1)

Without this variable

With all variables



With only this variable

Precipitation

SPOT NDVI Mean Over A Few Years

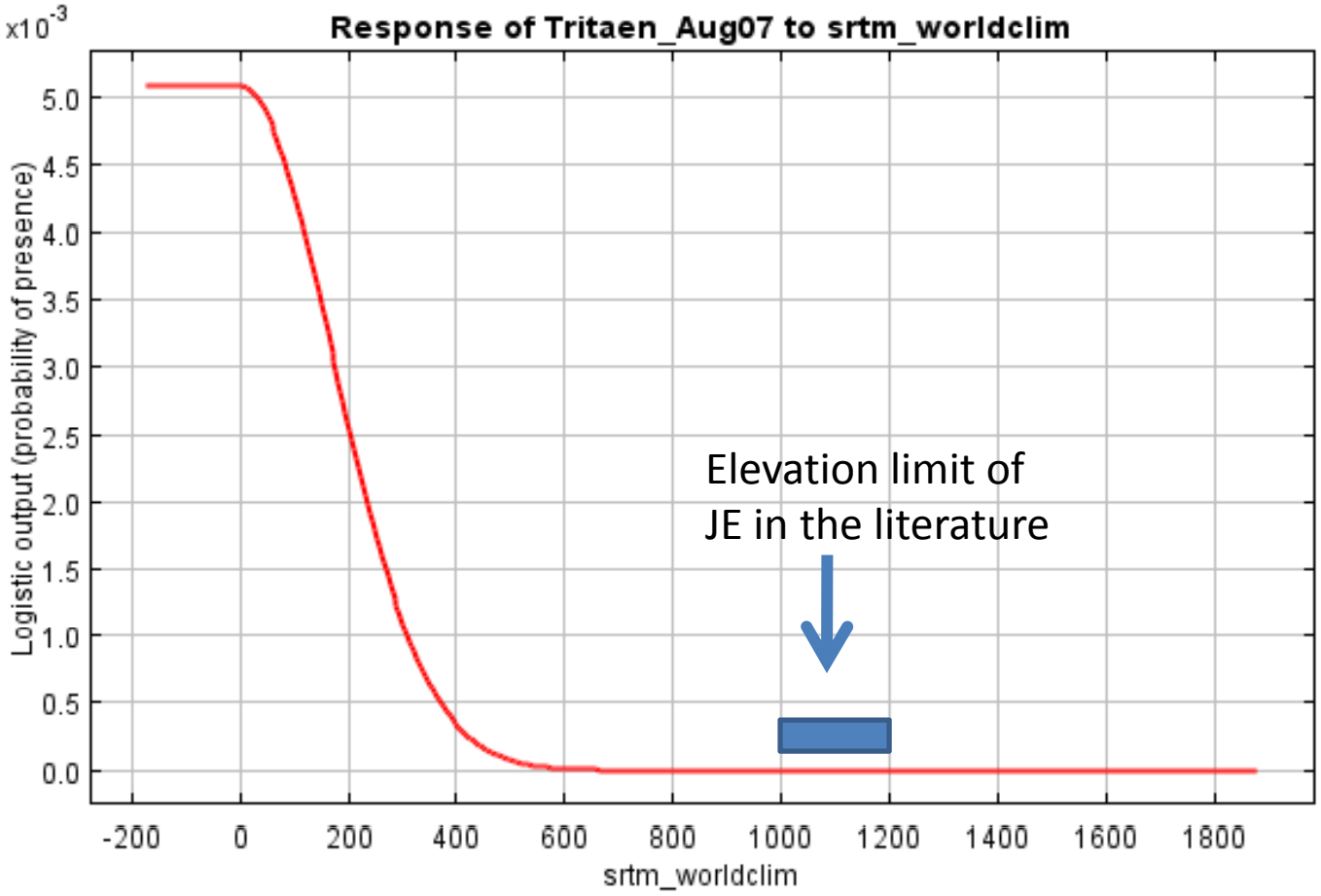
SPOT NDVI 2007

Max Temp

Min Temp

Land cover

Graph of individual variable shows the response of the species to the variable



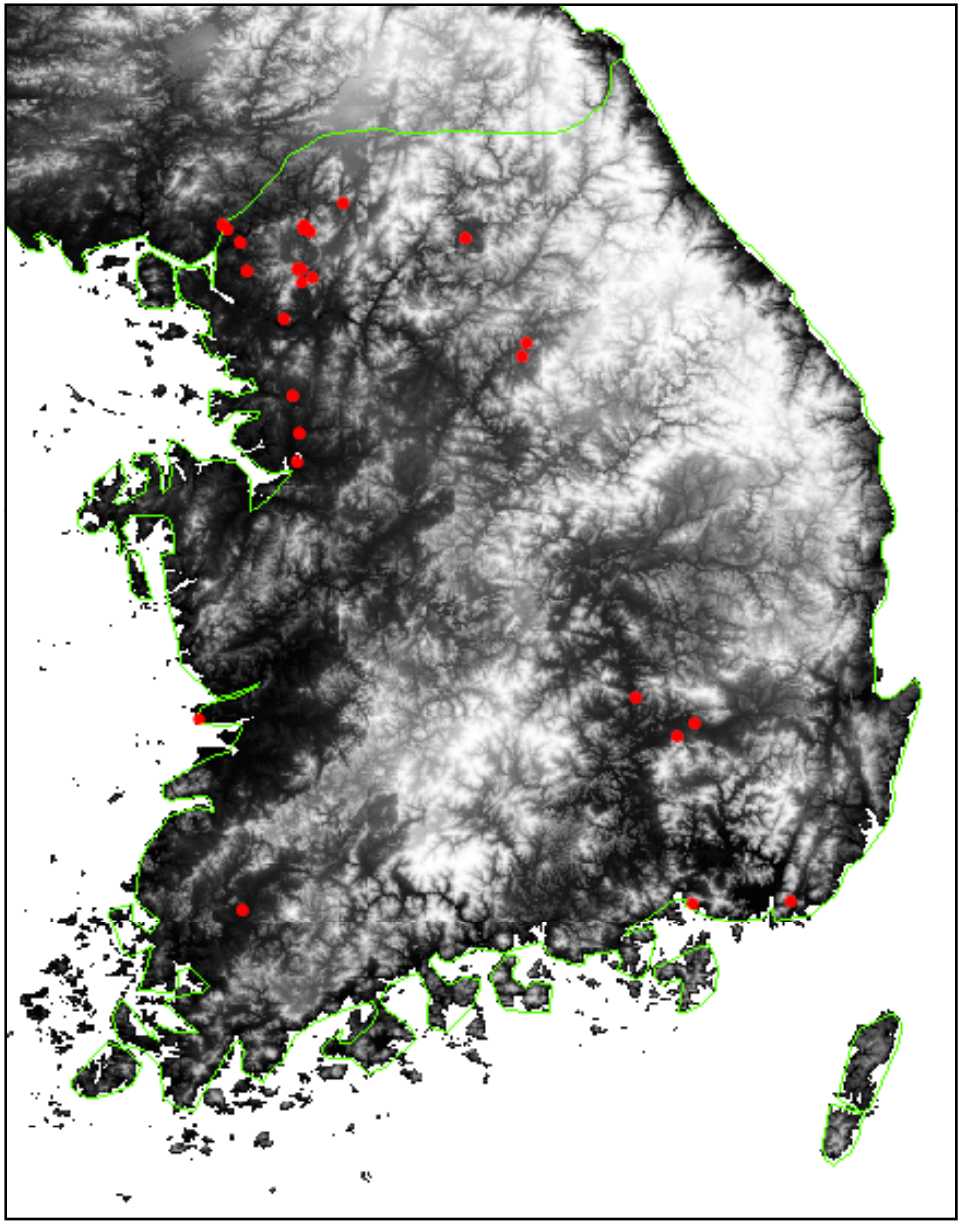
Elevation (meters)

Elevations at each camp derived from SRTM digital elevation data

Actual values of camps range from 8 to 310 meters.

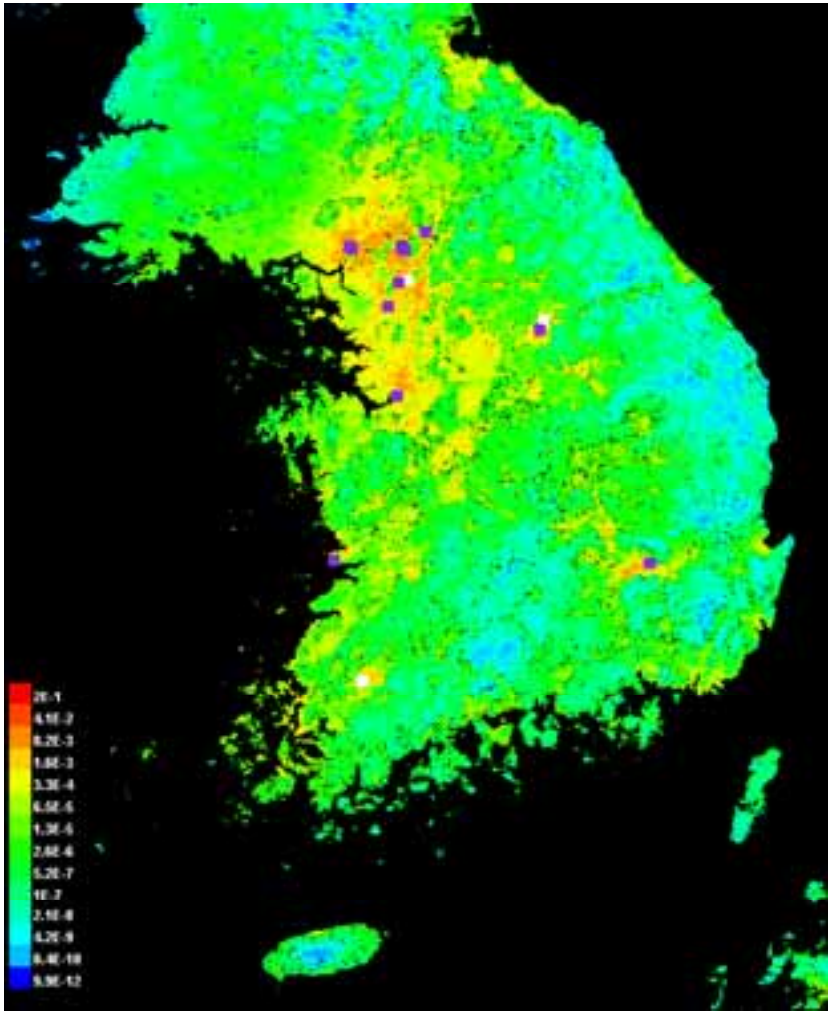
Entomologists think that this may represent full extent of mosquito habitat since no mosquitoes on steep slopes.

CAMP	Elevation
Bonifas	47
Carroll	104
Casey	57
Castle	181
Chinhae	251
Eagle	120
Essoyons	57
Giant	26
Hialeah	73
Hovey	148
Howze	25
Humphreys	8
Jackson	310
Kwangju	18
Liberty	45
Long	125
Mobile	60
Nimble	136
Osan	10
Page	72
Red Cloud	215
Stanley	58
Suwon	26
Taegu	38
Walker	67
Warrior	37
Yongsan	61



Other Issues in the Adult Mosquito Data

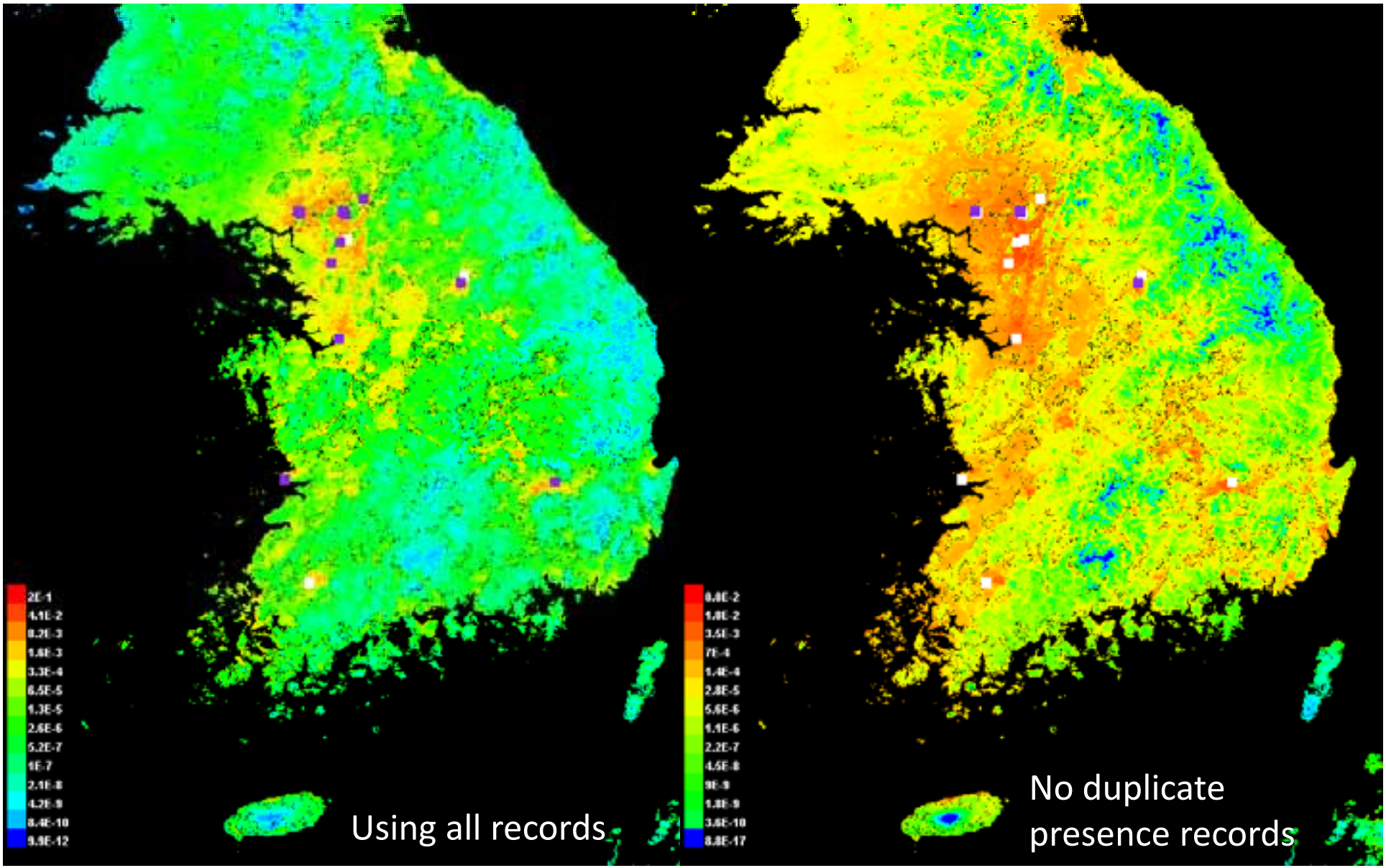
- Oversampling
  - Few collection locations (29 or less) and with high numbers of mosquitoes collected
  - Results in very high probabilities around bases and very low in other areas



Options to Remedy Oversampling

- Ignore it
  - Use every presence record
  - Get very high probabilities of mosquitoes around each camp
  - Not good
- Use ln(x+1) * 10 to normalize the data
  - Still get high probabilities around the camps even though the high values are somewhat suppressed
  - Better, but still not good.
- Using only one mosquito presence record at each site on each date
  - Appears to create best model

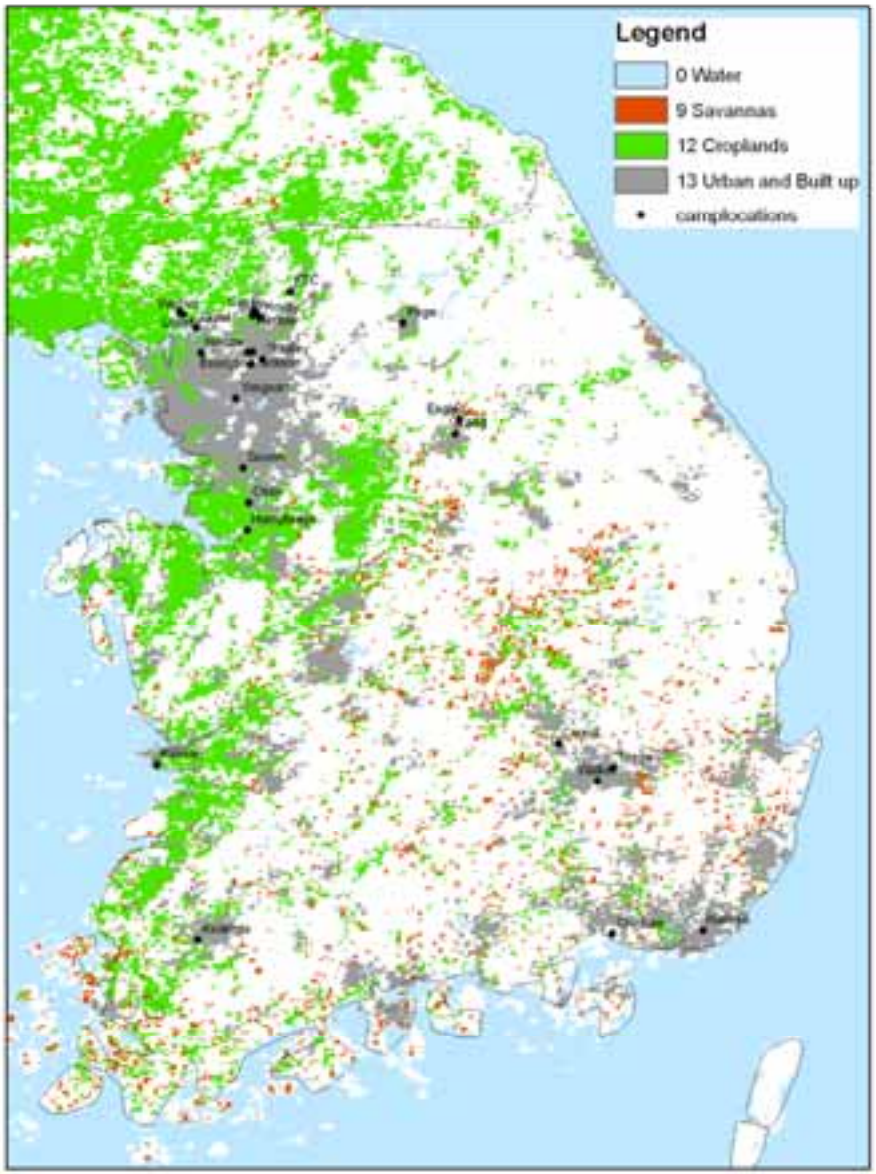
August 2007 Using ln(x+1) normalization



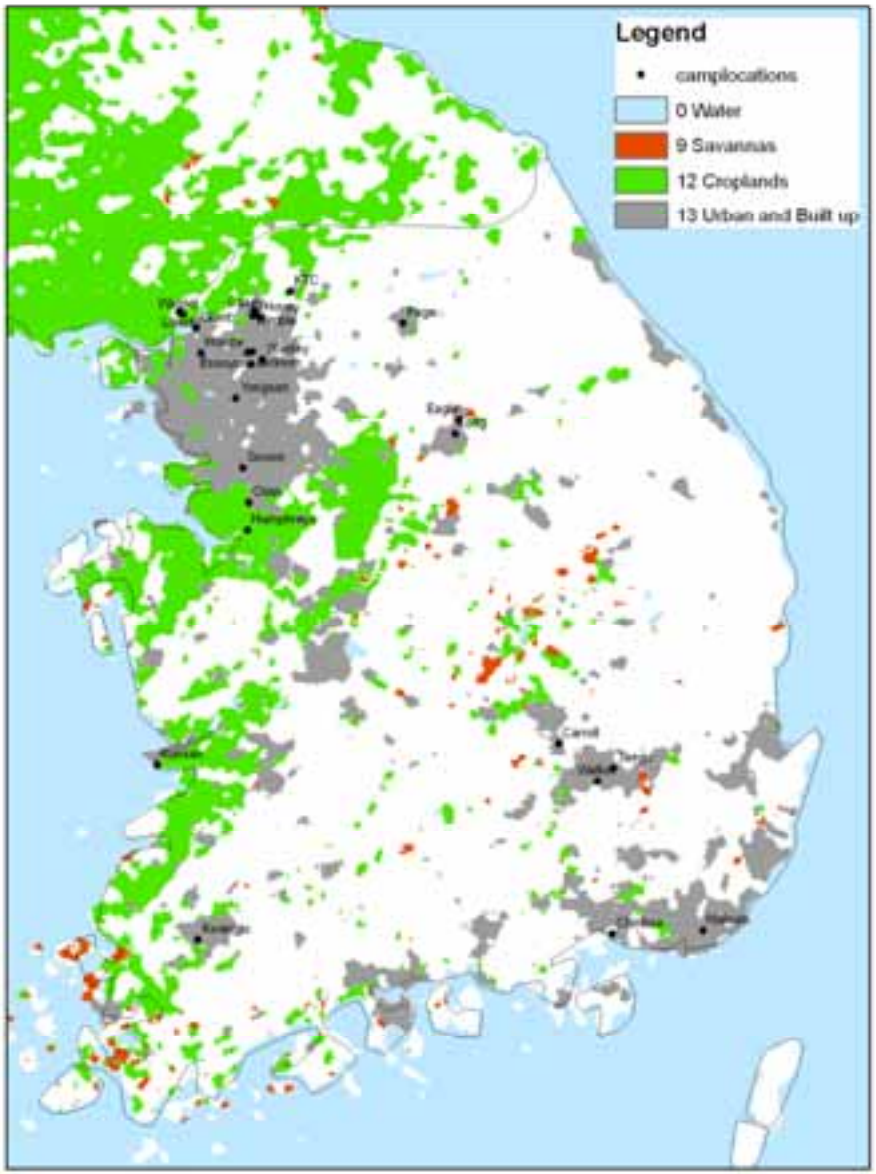
Other Issues in Adult Mosquito Data

- All collection sites at military bases (urban)
- Options to Remedy Urban Land Cover at Collection Sites:
  - More field collection sites in the future?
  - GIS ways to remedy?
  - Majority filter was used to determine the most frequent land cover class around each pixel in a 5 x 5 pixel area

Original Land Cover



Majority Filtered Land Cover

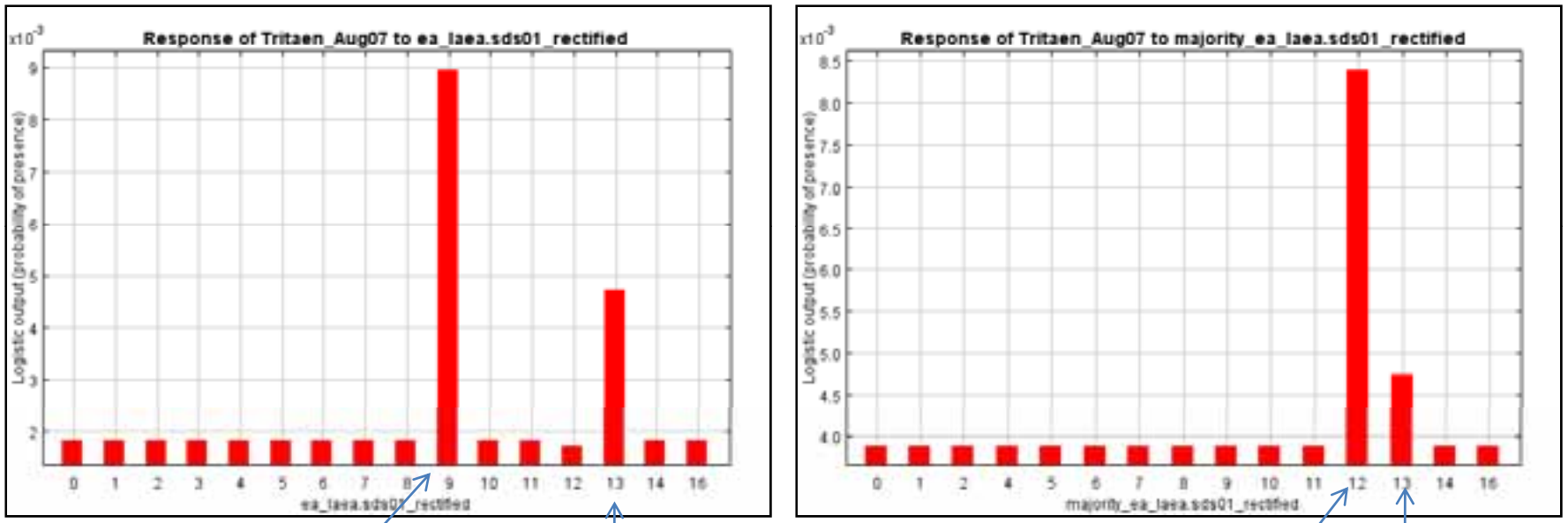


Response of Land Cover Classes to the Model

Original land cover layer

Land cover filtered with majority filter

Probability of presence



Savannas
(trees and grass)

Urban

Croplands

Urban

Histograms show how the logistic prediction changes as each environmental variable is varied, keeping all other environmental variables at their average sample value

Conclusions

- Models match what is already known about *Cx. tritaeniorhynchus*:
  - Peak probability in August/September
  - Lower probability in higher elevations
- Duplicate presence records for the same locations cause local high probabilities
  - Shows the need for more collection sites
- NDVI contributes to model (hopeful for future real-time predictions)
- Improvements needed:
  - More collection sites over full range of environmental variables
  - Methods to model the effect of nearby rice fields on mosquito populations on U.S. military bases.