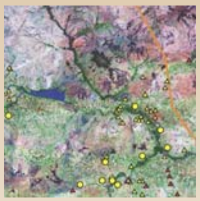
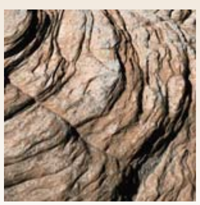
A Probabilistic Approach for Estimating the Spatial Extent of Pesticide Agricultural Use Sites and Potential Co-occurrence with Listed Species for Use in Ecological Risk Assessments





Presented by:

Katie Budreski[1], Michael Winchell[1], Lauren Padilla[1],
JiSu Bang[2], Richard Brain[2]

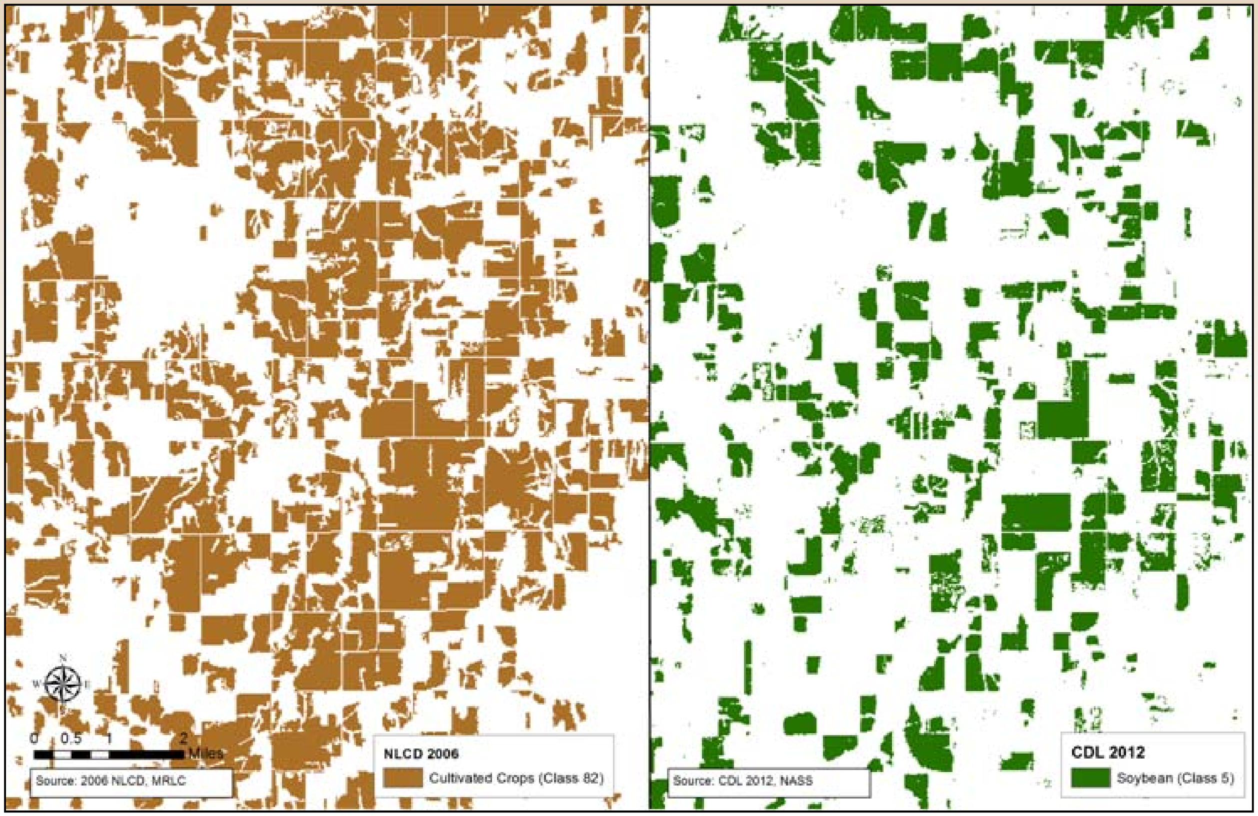
1. Stone Environmental, Inc. 2. Syngenta Crop Protection

EMPM, April 28th, 2015

Background

- An important component of endangered species assessments (ESAs) is the definition of crop footprints that represent potential sites for pesticide applications based on labeled uses.

- Historically, crop footprints have often been based on generalized land use datasets, (i.e., NLCD), without information concerning specific crops or historic use.



Background

- The Endangered Species Act and the National Academy of Sciences (NAS) recommends using best available data in ESAs and have listed the Cropland Data Layer (CDL, NASS) as a source of best available crop/land cover information (NAS 2013).
- The NAS report also recommends incorporating probabilistic approaches within ESAs to address uncertainty.
- The current approach proposed by EPA and the Services for crop footprint development uses multiple years of best available land cover data (CDL) to account for crop rotation and uncertainty, resulting in a deterministic crop footprint.

Motivation: Probabilistic Crop Footprint for use in ESAs

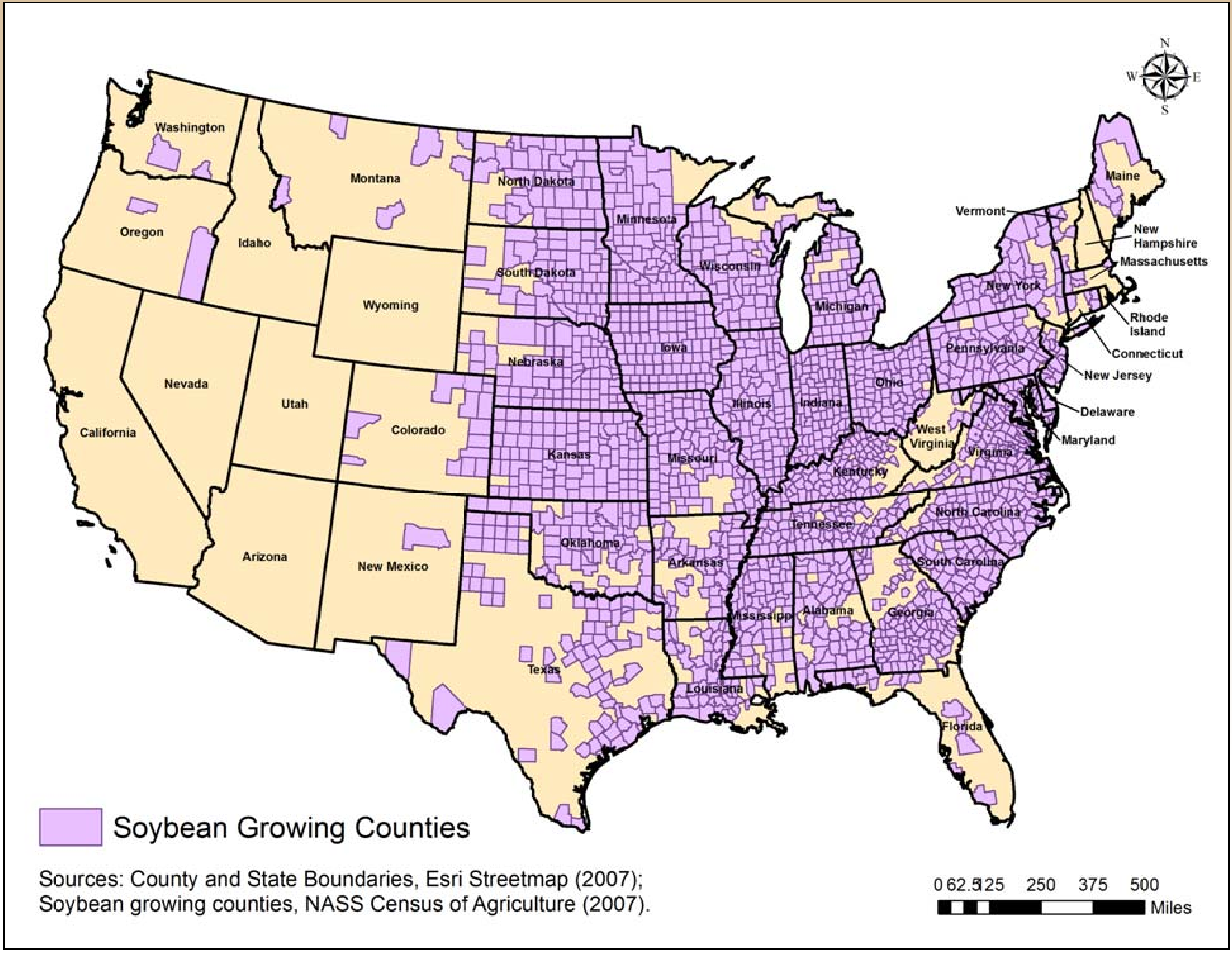
- Previous crop footprint methods based on generalized land use datasets (e.g., NLCD) are overly conservative in the representation of crop footprints for individual crops.
- Proposed crop footprint methods are unable to quantify the likelihood of co-occurrence of potential pesticide use sites and habitat areas, which may be advantageous at Step 2 and Step 3 of the proposed ESA process.
- Proposed crop footprint methods do not account for all uncertainty information available with the land cover datasets.

Objectives

- Develop a methodology that uses publically available, high resolution geospatial datasets to create refined crop footprints representing the probability of crop presence and thus potential pesticide use in any given year.
  - Account for misclassification of crop classes in the **Cropland Data Layer (CDL)** by incorporating accuracy assessment information by state, year, and crop, and include 5 years of data.
  - Adjust CDL misclassification probability based on accuracy assessment information from the **National Land Cover Dataset (NLCD)** by land cover class,
  - Scale crop probabilities at the state level by comparing against **NASS surveys** of reported planted acres by crop

Example Case Study

- Study Area
  - The continental US was evaluated to test the proposed approach.
  - Soybean was used as the example target crop for pesticide applications



Probabilistic Crop Footprint Methodology: 5-Year CDL, Bayesian Probability Method

Base-layer Development:

Step 1: CDL Footprint By Year (Target Crop with Original Class Preserved)

Step 2: Overlay CDL Footprint for Target Crop by Year with NLCD (All Original Classes Preserved)

Commission Adjustment:

Step 3: Calculate Bayesian probability based on CDL accuracy by year, state, and class; and NLCD accuracy by NLCD class.

Omission Adjustment:

Step 4: Omission Error Adjustment based on 2007 Ag Census or NASS Quick Stats.

Final Yearly Crop Footprint:

Step 5: Merge Commission Adjusted Raster with Omission Adjusted Raster

Combine Yearly Crop Footprints:

Step 6: Average 5 Years of Probabilistic Crop Footprint Rasters

Final Crop Footprint Check:

Step 7: For Original CDL Crop Pixels with a Final Probability of 0, assign 0.0001

Probabilistic Crop Footprint Methodology
5-Year CDL, Bayesian Probability Method

Commission Adjustment: Adjusting for errors associated with pixels that were incorrectly classified as soybean.

Legend:

Non-Soy	All pixels not classified as soybean
Soy (T)	Classified as soybean and is soybean on the ground
Soy (F)	Classified as soybean, but is not soybean on the ground

Non-Soy	Non-Soy	Non-Soy	Non-Soy	Non-Soy
Soy (T)	Soy (T)	Non-Soy	Non-Soy	Non-Soy
Soy (T)	Soy (T)	Non-Soy	Soy (F)	Soy (F)

Omission Adjustment: Adjusting for errors associated with pixels that are soybean but incorrectly classified as another crop or land use.

Legend:

Soy	Classified as soybean
Non-Soy (T)	Not classified as soybean and are not soybean on the ground
Non-Soy (F)	Not classified as soybean, but IS soybean on the ground

Non-Soy (T)	Non-Soy (T)	Non-Soy (T)	Non-Soy (T)	Non-Soy (T)
Soy	Soy	Non-Soy (F)	Non-Soy (T)	Non-Soy (T)
Soy	Soy	Non-Soy (F)	Soy	Soy

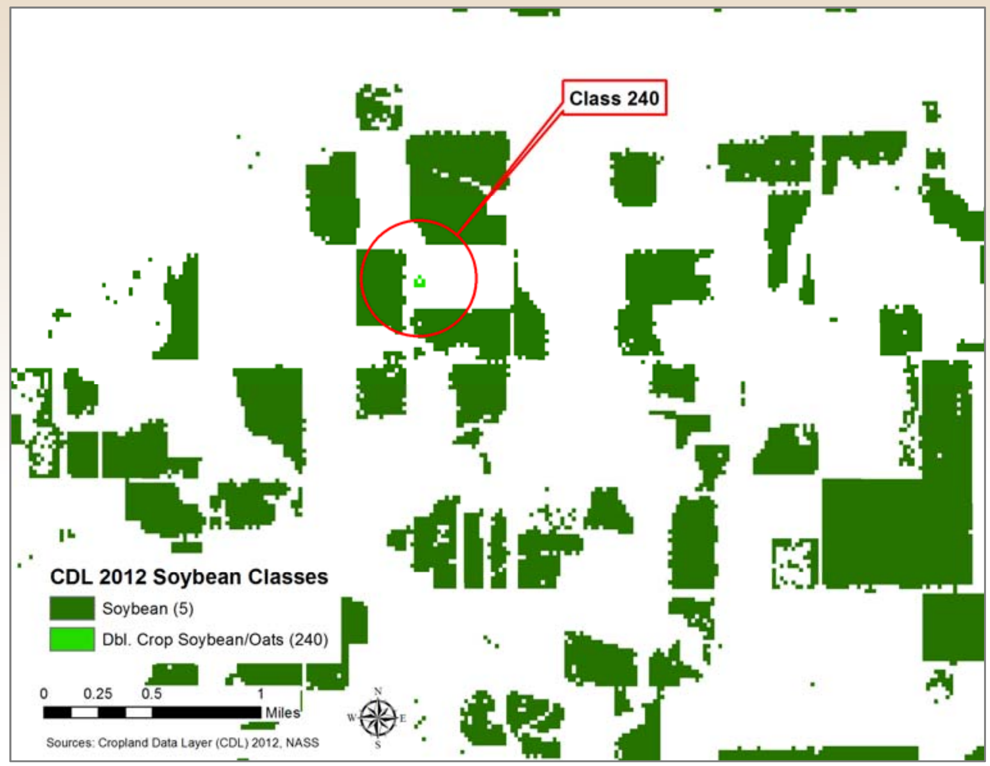
8

Probabilistic Crop Footprint Methodology
5-Year CDL, Bayesian Probability Method

- Step 1: Base-layer Development
  - CDL Footprint By Year (Target Crop with Original Class Preserved)

Single Year CDL

Single Year CDL: Soybean Classes Extracted



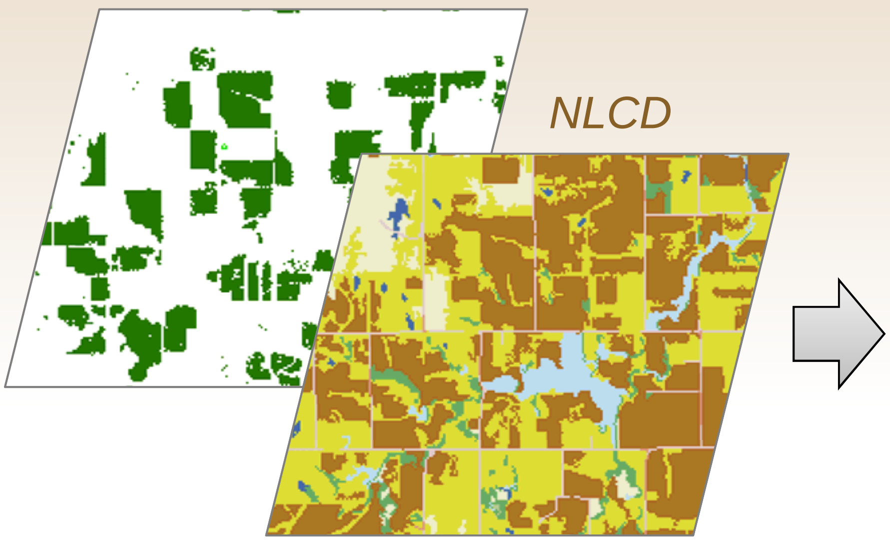
Example: 2012 CDL for Anderson County, Kansas

Probabilistic Crop Footprint Methodology
5-Year CDL, Bayesian Probability Method

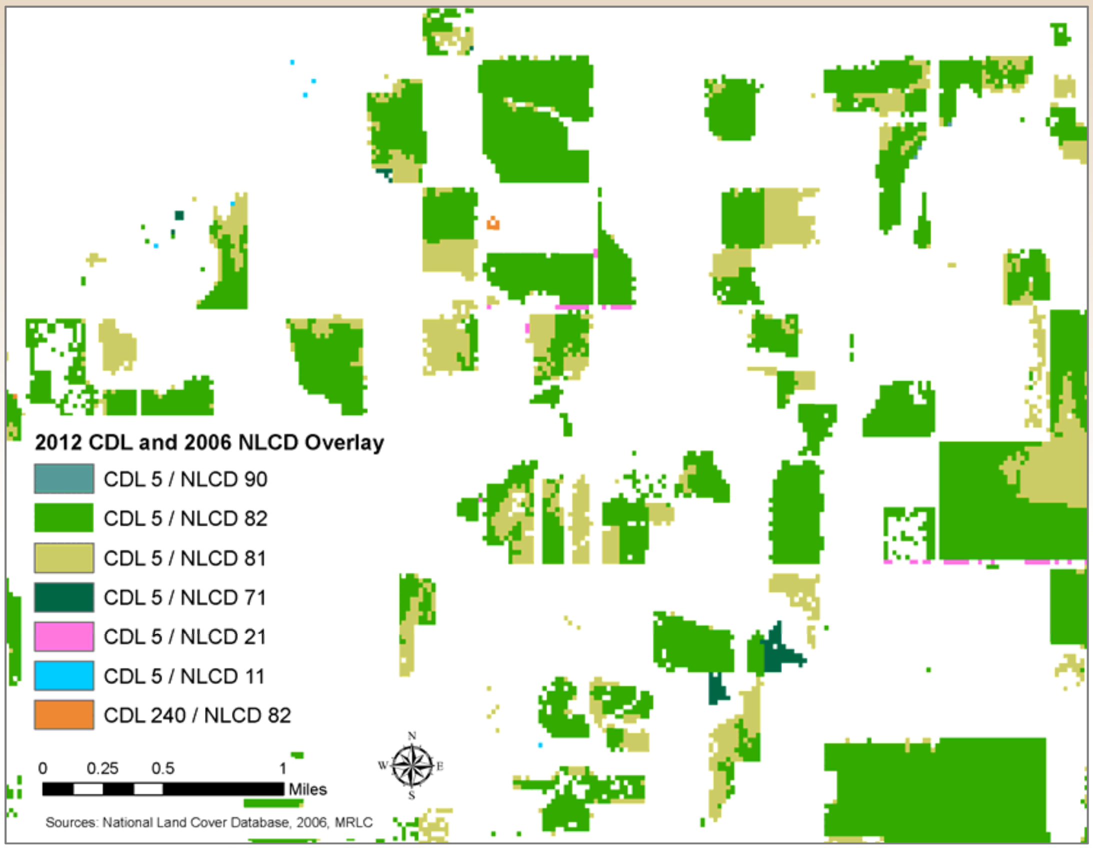
- Step 2: Base-layer Development
  - Overlay CDL Footprint for Target Crop by Year with NLCD (All Original Classes Preserved)

Single Year CDL: Soybean Classes Extracted

NLCD



Single Year CDL and NLCD Overlay



Example: 2012 CDL and 2006 NLCD for Anderson County, Kansas

Probabilistic Crop Footprint Methodology
5-Year CDL, Bayesian Probability Method

- Step 3: Commission Adjustment
  - Calculate Bayesian probability based on CDL accuracy by year, state, and class; and NLCD accuracy by NLCD class.
  - Bayes theorem is used to determine the posterior probability a pixel is soybean assuming a prior probability from CDL user's accuracy and conditional probabilities from NLCD overlap. Posterior probabilities may both increase or decrease depending on the NLCD overlap.

Results:

-Pixel with CDL 'soy' AND NLCD 'cultivated crops' → higher probability of being soy

-Pixel with CDL 'soy' AND NLCD other class → lower probability of being soy

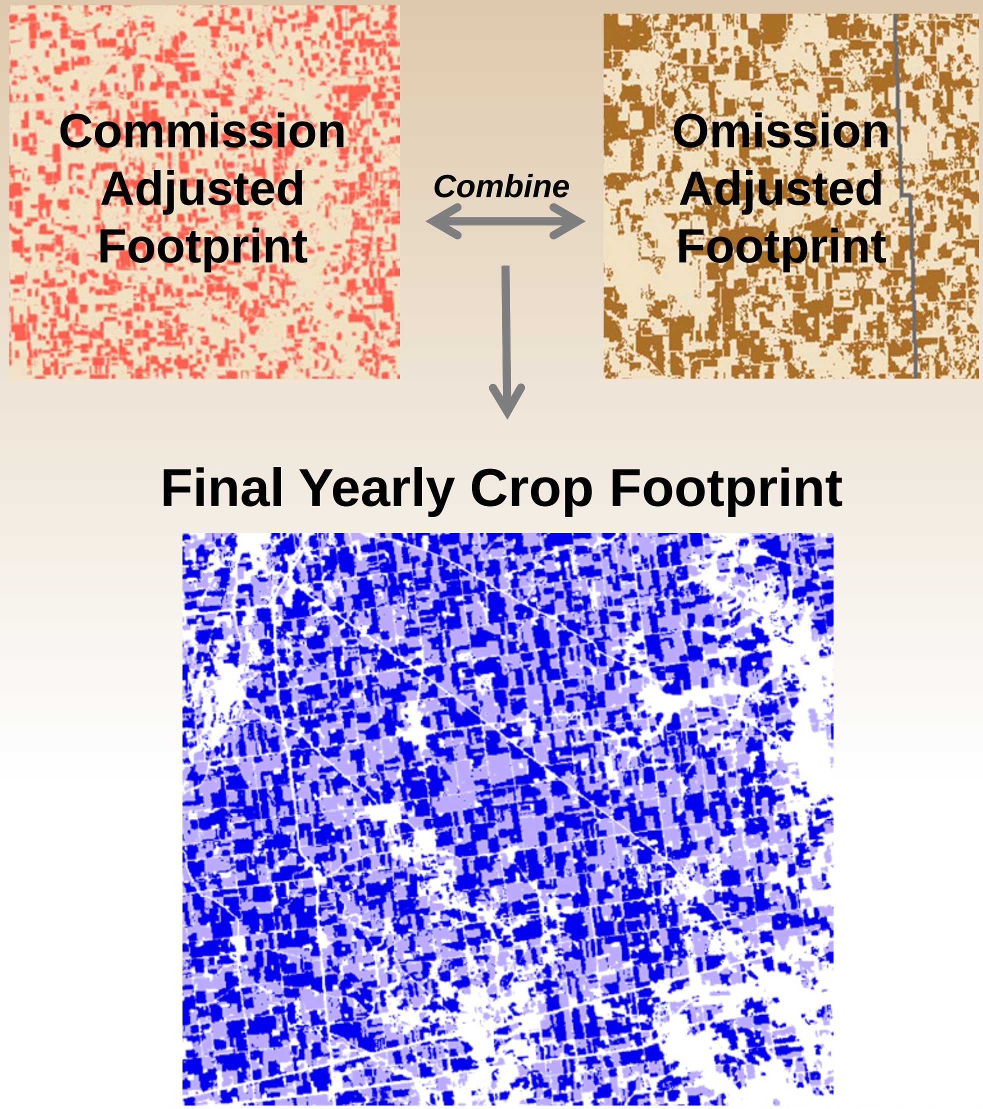
Probabilistic Crop Footprint Methodology
5-Year CDL, Bayesian Probability Method

- Step 4: Omission Adjustment
  - Omission Error Adjustment based on 2007 Ag Census or NASS Quick Stats.
  - Assumption: Survey data from AgCensus and NASS Quick Stats represents the “true” crop acreage.
  - In all omission adjustment cases, no pixel is added or removed from the overall 5-Year CDL Crop Footprint.
  - Pixel probabilities are scaled to meet AgCensus/QuickStats acreage

Probabilistic Crop Footprint Methodology
5-Year CDL, Bayesian Probability Method

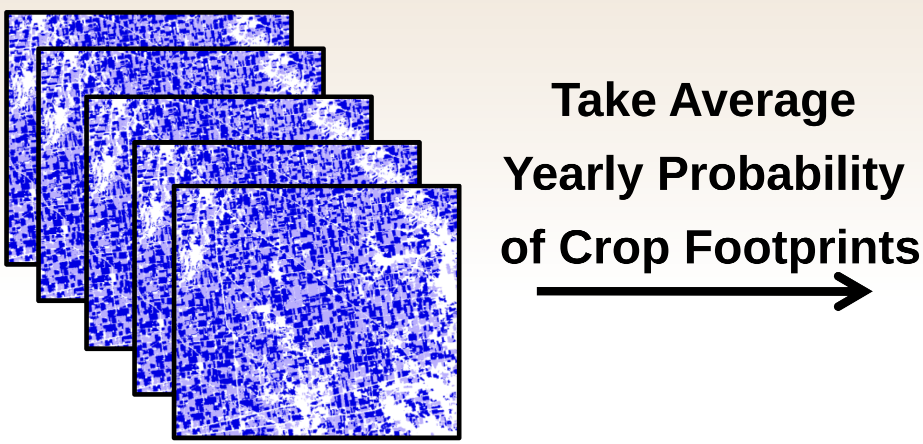
- Step 5: Final Yearly Crop Footprint
  - Combine Commission Adjusted Footprint with Omission Adjusted Footprint

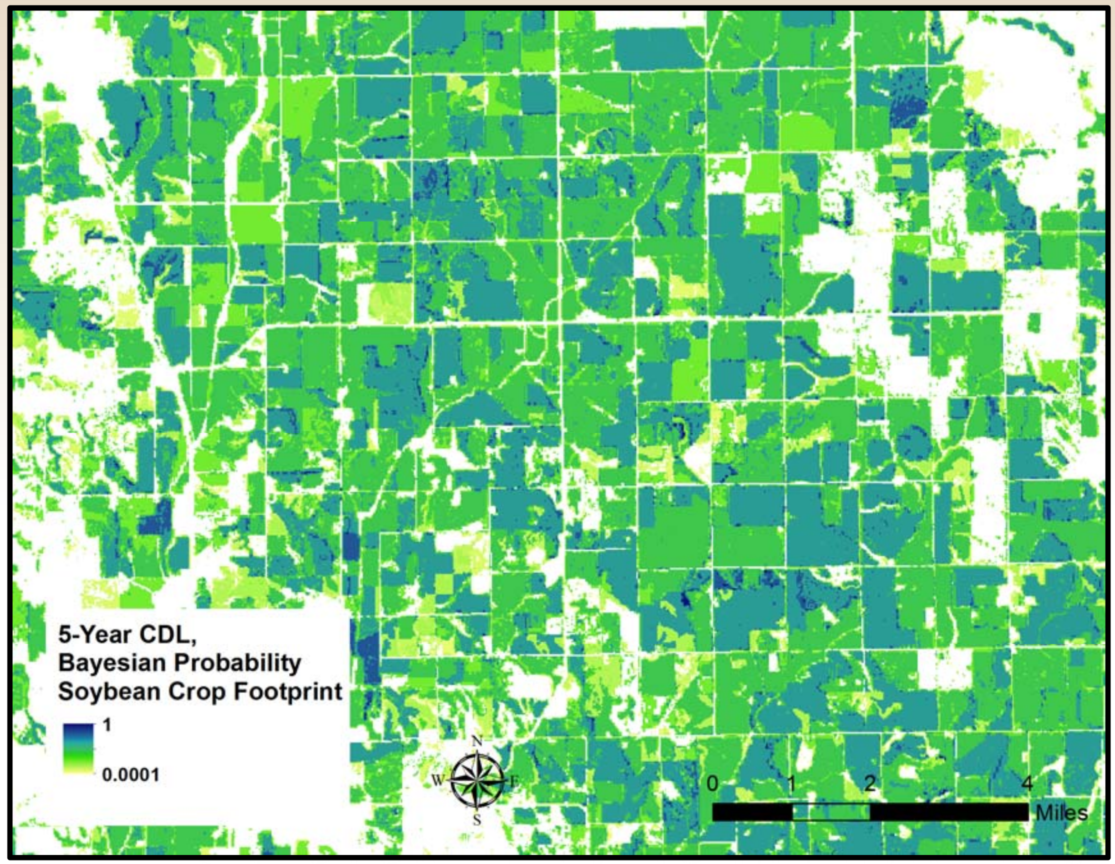
EXAMPLE



Probabilistic Crop Footprint Methodology
5-Year CDL, Bayesian Probability Method

- Step 6: Combine Yearly Crop Footprints
  - Average 5 Years of Probabilistic Crop Footprints
- Step 7: Final Crop Footprint Check
  - For Original CDL Crop Pixels with a Final Probability of 0, assign 0.0001

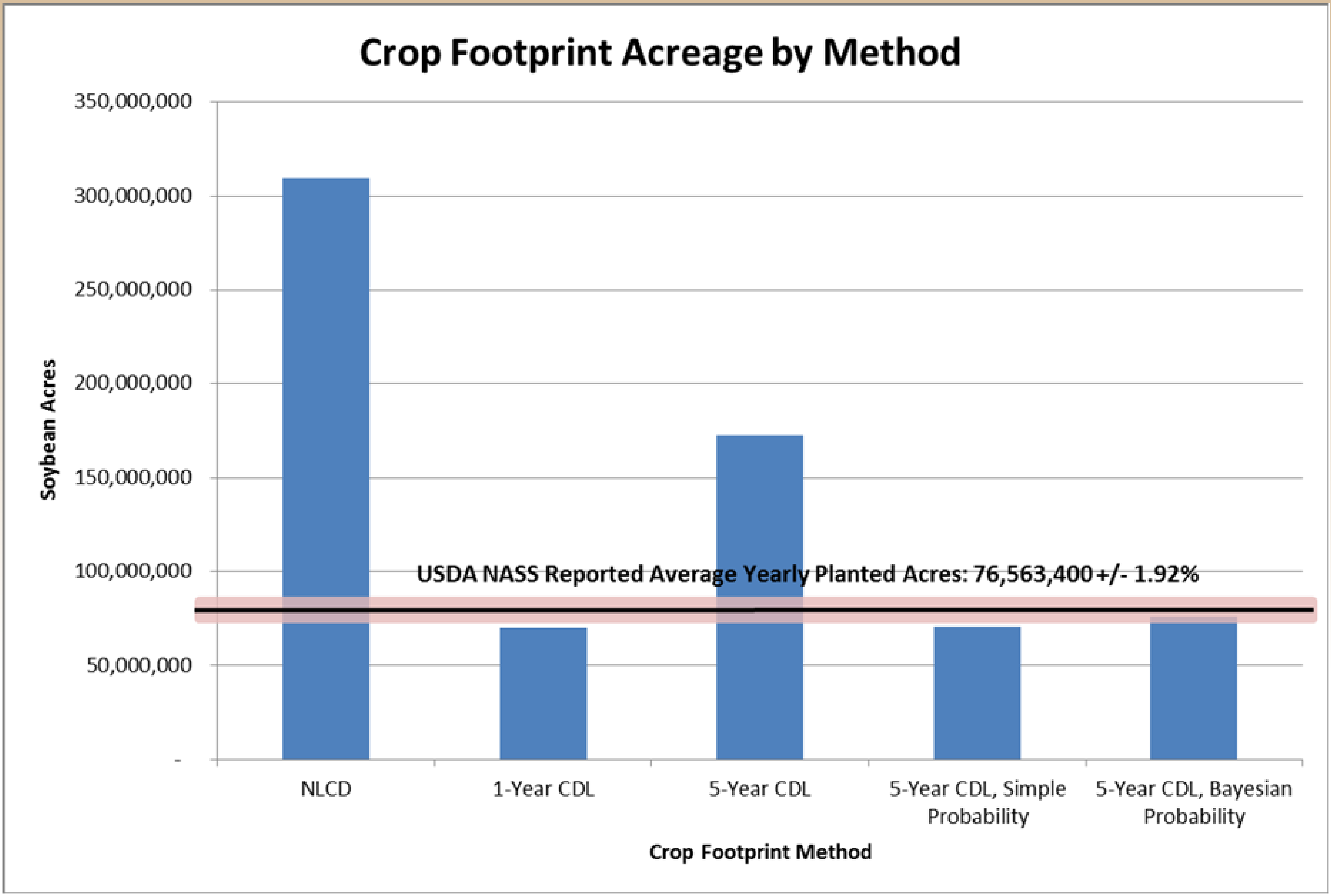




Comparison with Alternative Methods

- Compared Bayesian Probabilistic Method with 4 Alternate Methods
  - NLCD 2006, Cultivated Crop Class
  - 1-Year CDL, All Soybean Classes
  - 5-Year CDL, All Soybean Classes
  - 5-Year CDL, All Soybean Classes, Probability Based on # of Years Soybean is Present (e.g., soybean in 1 of 5 years would equal a probability of 0.20).

Results: Crop Footprint Comparison
National Crop *Footprint Acreage*



Results: Crop Footprint Comparison
Crop Footprint Extent

Example: Dallas County, Iowa

Results: Crop Footprint Comparison

Spurious Pixels

Example: Iron County, Missouri

Results: Crop Footprint Comparison
Crop Footprint Extent

- Crop footprint EXTENT is the same for all 5-year CDL methods (5-Year CDL, 5-Year CDL Simple Probability, and 5-Year CDL Bayes Probability)
  - **NO pixels removed for probabilistic methods**
- Spurious pixels are not removed, but assigned lower probabilities, due to overlap with non-agricultural NLCD classes.
- Final acreages are within the error bounds of known soybean acreages, based on NASS statistics.

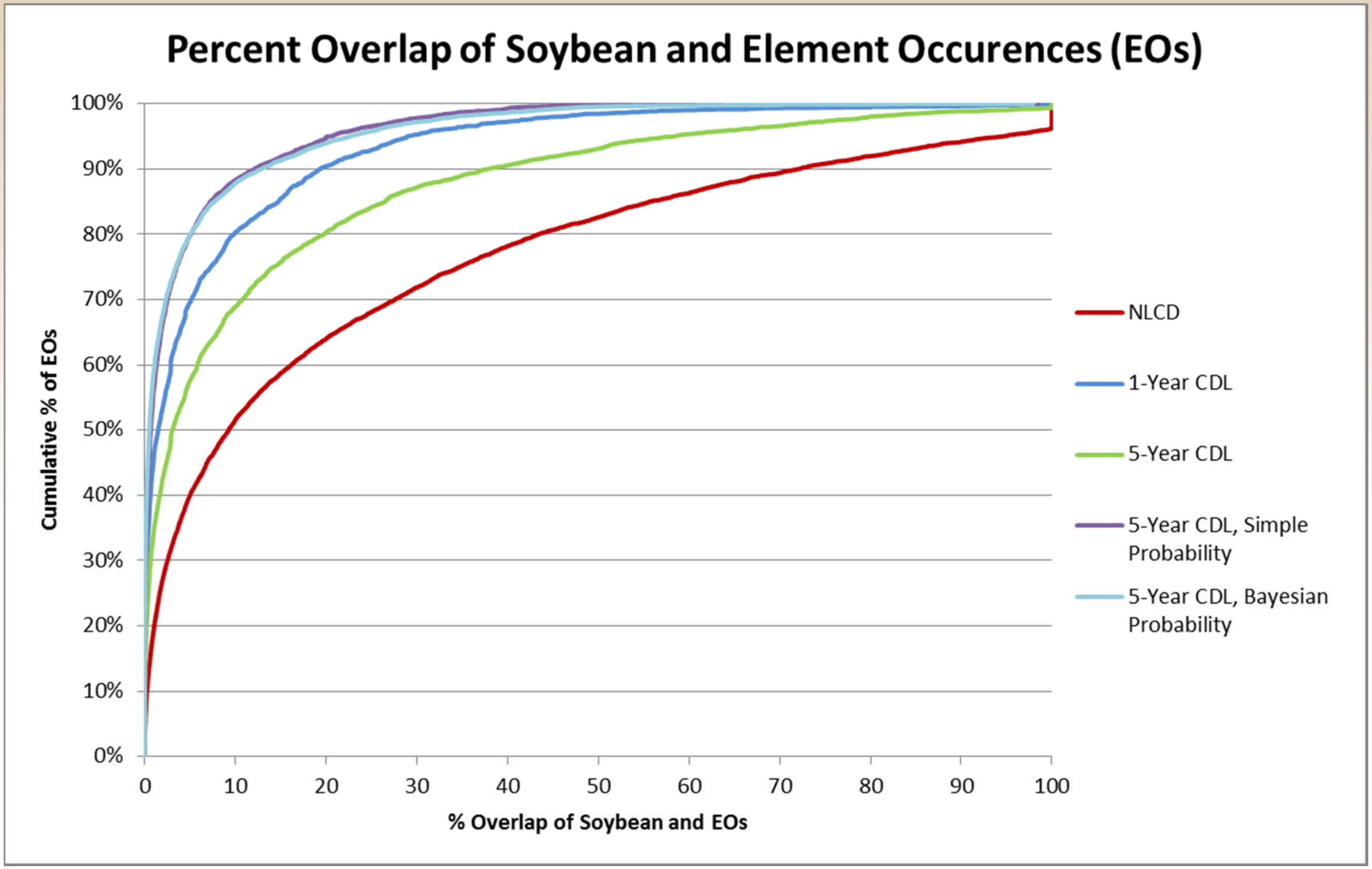
Results: Crop Footprint Comparison
Co-occurrence Analysis

- All 5-Year CDL methods result in the same number of species that have co-occurrence with soybean crop footprint
  - NO pixels removed for probabilistic methods

NLCD (Number of Species)	NLCD (% of All Species Evaluated)	1-Year CDL (Number Species)	1-Year CDL (% of All Species Evaluated)	5-Year CDL (Number Species)	5-Year CDL (% of All Species Evaluated)	5-Year CDL, Simple Probability (Number Species)	5-Year CDL, Simple Probability (% of All Species Evaluated)	5-Year CDL, Bayesian Probability (Number Species)	5-Year CDL, Bayesian Probability (% of All Species Evaluated)
511	48.99%	210	20.13%	276	26.46%	276	26.46%	276	26.46%

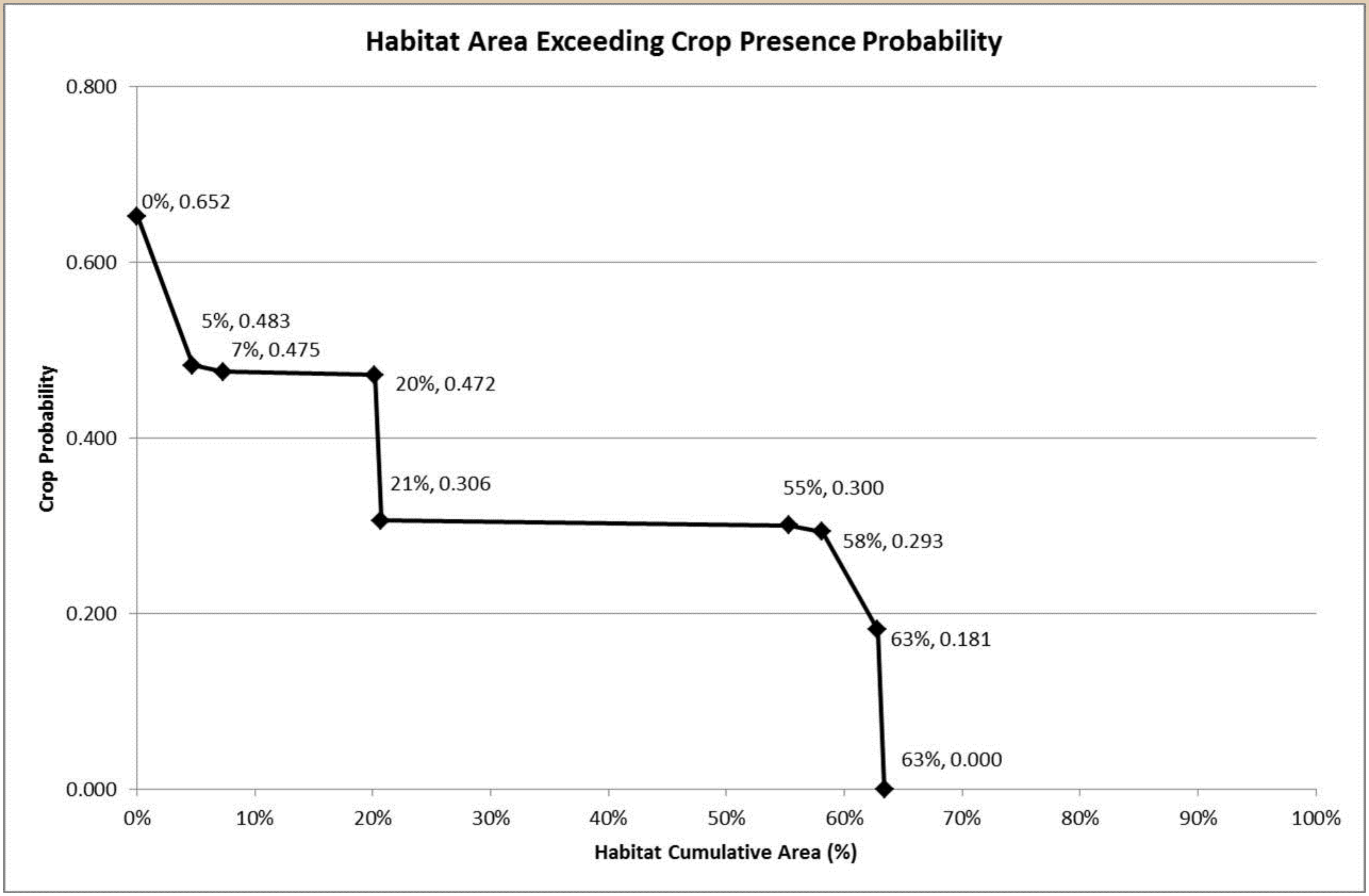
Results: Crop Footprint Comparison
Co-occurrence Analysis, Single EOs

- 5-Year CDL, Bayesian Probability: 59% of EOs have less than 1% overlap with soybean.
- 5-Year CDL: 34% of EOs have less than 1% overlap with soybean



Results: Crop Footprint Comparison
Co-occurrence Analysis, Single EOs

- 5-Year CDL, Bayesian Probability: 55% of the habitat area has a crop presence probability exceeding 0.300



Summary of Probabilistic Approach

- Use of best available, crop-level land cover data (Cropland Data Layer, USDA NASS)
- Use of multiple years of crop-level land cover data to account for crop rotation and uncertainty
- Probabilistic crop presence based on well established Bayesian approach and known uncertainty of land cover datasets
- The number of species that co-occur with the crop footprint and potential pesticide use is the same as all other 5-year CDL methods (can be used in Step 1 of proposed ESA method)
- However, the likelihood of co-occurrence of species habitat and potential pesticide use sites is better understood (can be used in Step 2 and Step 3 of proposed ESA approach)

Conclusions

- There are known and quantifiable uncertainties in land cover datasets that can be accounted for using probabilistic methods.
- The proposed method accounts for both *commission* and *omission* errors.
- Using yearly data can help understand and account for crop rotation and changes in land use over time.
- The probabilistic crop footprints represent the same amount of acreage as NASS reported acreages by state (+/- 95% CI)
- The probabilistic crop footprints help to reduce the influence of ‘spurious’ pixels without removing them from the analysis.
- The probabilistic crop footprint allows for conservative estimates of May Affect for ‘Step 1’ of the proposed ESA process, but allows for more detailed analysis and review of habitat level information for ‘Step 2’ and ‘Step 3’ of the proposed ESA process

References

Budreski, K., M. Winchell, L. Padilla, J. Bang, and R. Brain. A probabilistic approach for estimating the spatial extent of pesticide agricultural use sites and potential co-occurrence with listed species for use in ecological risk assessments. IEAM, *in review*.

Fry, J., G. Xian, S. Jin, J. Dewitz, C. Homer, L. Yang, C. Barnes, N. Herold, and J. Wickham, 2011. Completion of the 2006 National Land Cover Database for the Conterminous United States, PE&RS, Vol. 77(9):858-864.

USDA National Agricultural Statistics Service Cropland Data Layer. 2008-2012. Published crop-specific data layer [Online]. Available at http://nassgeodata.gmu.edu/CropScape/. USDA-NASS, Washington, DC.

USDA National Agricultural Statistics Service Quick Stats. 2008-2012. Published crop-specific statistics [Online]. Available at http://www.nass.usda.gov/Quick_Stats/. USDA-NASS, Washington, DC.

USDA National Agricultural Statistics Service 2007 Census of Agriculture. December 2009. Available at http://www.agcensus.usda.gov/Publications/2007/Full_Report/. USDA-NASS, Washington, DC.

Wickham, J.D., Stehman, S.V., Gass, L., Dewitz, J., Fry, J.A., and Wade, T.G. 2013. Accuracy assessment of NLCD 2006 land cover and impervious surface, Remote Sensing of Environment, Vol. 130, pp. 294-304.