Advanced Methods for Agricultural and Agro-environmental Monitoring

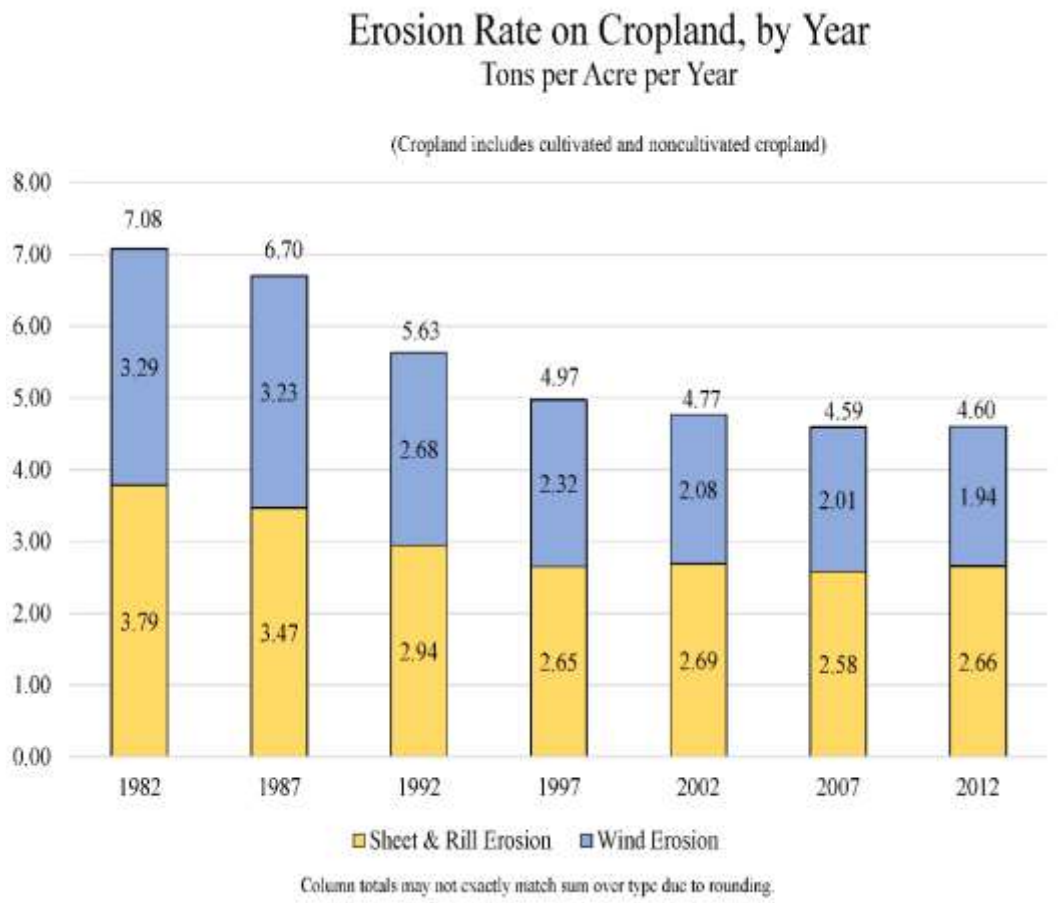
Emily Berg, Zhengyuan Zhu, Sarah Nusser, and Wayne Fuller

Outline

1. Introduction to the National Resources Inventory
2. Hierarchical Bayesian models for NRI county estimates
3. A compromise approach to spatially stratified sampling for the Conservation Effects Assessment Project

NRI Background -- Objectives

- Monitors natural resources and agriculture on nonfederal US land
- Inventory years
  - 1982, 1987, 1992, 1997, 2000-2012
- Land cover/use
  - Corn, soybeans, urban
- Erosion
  - Water, wind
- Change over time



NRI Background – Sample Designs

- Area frame
  - Sampling unit = segment
  - 3 points per segment
- Foundation sample (1982-1997)
  - ~300,000 segments observed every 5 years
    - ~800,000 points
- Annual samples (2000-present)
  - Core panel ~40,000 segments observed every year
  - Rotation panels ~30,000 segments observed less frequently
  - Core and rotation are stratified samples of foundation

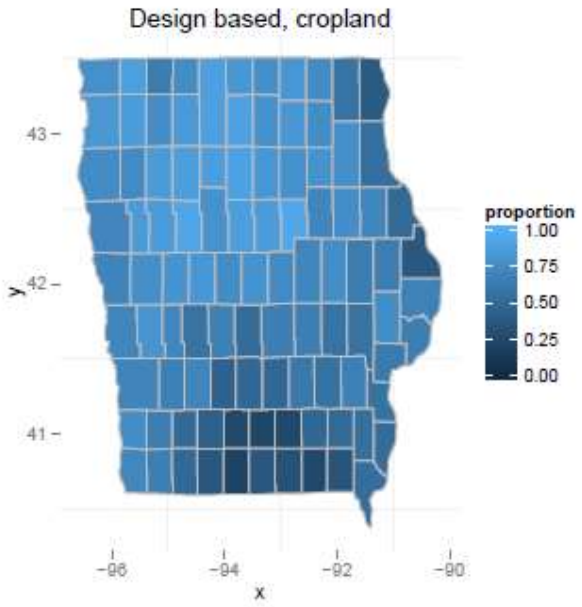
NRI Background – Data Collection

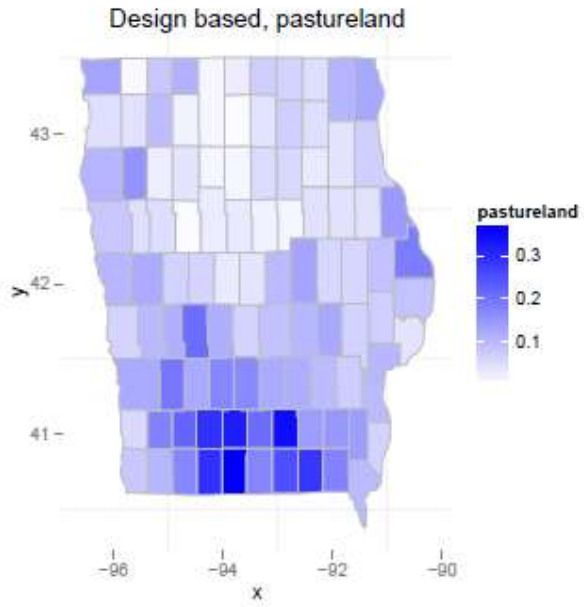
- Data collection
  - Interpretation of aerial photographs of sampled segments
  - “Local data” – administrative information for certain kinds of points
    - Ex: cropland, wetlands, soils
  - Custom software

NRI: County Estimates

- Published estimation domains
  - State, region, nation
- Estimates are of interest for counties
  - Proportion of each county classified in different cover/use categories
  - Current estimates for counties can be judged unreliable
- Model-based small area estimation
  - Incorporate external auxiliary information
  - Borrow information from neighboring domains

NRI County Level Estimation: Spatial Structure





	Cultivated Crop	Pasture
Geary's C	0.21	0.35
P-value	<0.001	<0.001

NRI County Level Estimation: Hierarchical Bayesian Model

- Model for NRI estimators: Generalized Dirichlet
  - Unequal sampling variances
  - Sum-to one restriction
  - Multivariate relationships preserved
- Model for true proportions: logistic-normal
  - Incorporates covariates obtained from satellite imagery
  - Conditionally autoregressive spatial structure
- Bayesian inference
  - Gibbs sampling

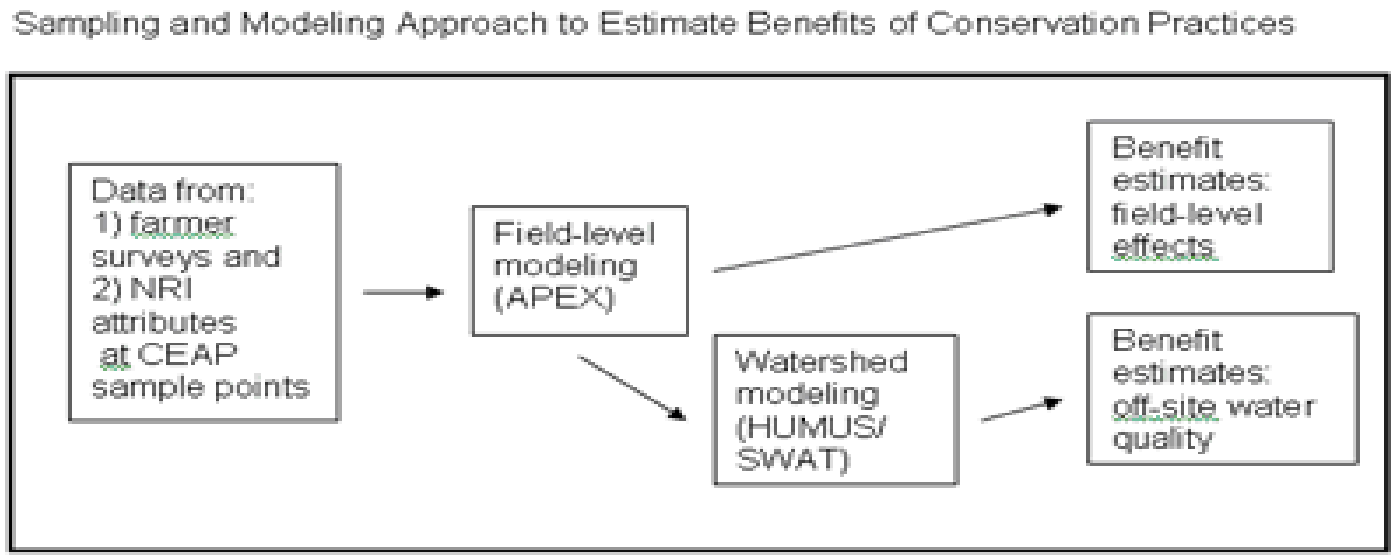
NRI County Level Estimation: Variance Comparison

- (Posterior variance)/(Est. variance of NRI)

	Cropland	Pasture	Remainder
Min	0.04	0.02	0.04
1st Quartile	0.12	0.05	0.17
Median	0.16	0.08	0.27
Mean	0.21	0.14	0.31
3rd Quartile	0.29	0.17	0.41
Max.	0.85	0.77	0.81

Conservation Effects Assessment Project

- Special interest in cropland points
  - Data collectors visit a subset of NRI points and collect more detailed information about crop managements and conservation practices



Conservation Effects Assessment Project

- Spatial spread desired for efficient sample designs
  - Points closer together are more similar than points farther apart
- Stratified sampling can improve spatial spread
- Information on the variability within a stratum is needed for variance estimation

Spatially Stratified Designs: Strengths and Weaknesses

- One per stratum – select one point from each stratum
  - Good spatial spread
  - No design-unbiased variance estimator – no estimate of within-stratum variance
- Two per stratum – select two points from each stratum
  - Possibility of clustering within a stratum
  - Variance estimation possible

Illustration for sample size n = 6						
One per stratum	**Stratum 1**	**Stratum 2**	**Stratum 3**	**Stratum 4**	**Stratum 5**	**Stratum 6**
Two per stratum	**Stratum 1**		**Stratum 2**		**Stratum 3**	

Spatially Stratified Designs

- Combination of one per stratum and two per stratum sampling
- Form n/2 pairs of strata
- Select 2 from a randomly selected stratum of a randomly selected pair

Illustration for sample size n = 6						
One per stratum	**Stratum 1**	**Stratum 2**	**Stratum 3**	**Stratum 4**	**Stratum 5**	**Stratum 6**
Compromise	**Pair 1 Stratum 1 1**	**Pair 1 Stratum 2 1**	**Pair 2 Stratum 3 1**	**Pair 2 Stratum 4 1**	**Pair 3**	
# to select					**Stratum 5 Select 2**	**Stratum 6 Select 0**

Compromise Designs for CEAP

- Extension – form K groups and apply the compromise procedure within each group
  - K = 2 is 2 per stratum
  - As K increases, design approaches 1 per stratum

Number to Select	Variance of Estimator	Variance of Variance Estimator X Number to select	Approx. Degrees of Freedom
2	1.50	5.00	54.0
3	1.33	4.00	53.3
4	1.25	5.17	36.3
5	1.20	6.40	25.4
10	1.10	16.29	7.5

  - Within variance = between variance = 1

Summary

- NRI and CEAP use diverse advanced statistical methods for both estimation and sample design

- Examples presented exploit spatial structure
  - Model based estimation for county estimates
  - Spatially stratified designs for CEAP

Thank You!

- Acknowledgements:
  - Xin Wang
  - Cooperative Agreement No. 68-3A75-4-122 between the USDA NaturalResources Conservation Service and the Center for Survey Statistics and Methodology at Iowa State University