Using Federal Environmental Data for Exposure Assessment in Epidemiologic Studies of Cancer

Mary H. Ward, Ph.D.

wardm@mail.nih.gov

Occupational and Environmental Epidemiology Branch
Division of Cancer Epidemiology and Genetics

NIH NATIONAL CANCER INSTITUTE

Geospatial Statistics, Tools, Data, Practices, Opportunities and Challenges in the Federal Agencies
Geospatial Interest Group of the Federal Committee on Statistical Methodology

October 16, 2015

Linkages to environmental data – NCI studies of cancer etiology

- U.S. Census data (e.g. income, education, housing, population density)

- Land use data (e.g. roads/traffic, agricultural fields [pesticides])

- Environmental monitoring data:
  - Drinking water contaminants (public water supplies)
  - Modeled levels in private wells
  - Industrial emissions (Toxic Release Inventory, Dioxin Emission databases)
  - Air pollutants (e.g. $PM_{2.5}$)
  - Ultraviolet light

Using Geographic Information Systems for exposure assessment

- **Mapping the study population, exposure sources, and exposure assessment (fate and transport modeling)**

Nuckols JR et al., *Environ Health Perspect;* 2004

- **Issues/Examples:**
  - **Accuracy of locations (geocoding)**
  - **Agricultural pesticides and emissions from animal feeding operations**
  - **Residential mobility**
  - **Drinking water contaminants**
    - **Public water supply monitoring data**
    - **Private wells – modeled estimates**

Locate the study population (and exposure sources)

- Global Positioning System (GPS)
  - High accuracy (m)
  - Easily added to home visit
  - Field studies can track individual activity patterns (e.g. commuting)
  - Not always feasible especially for large cohort studies

- Geocode current and past addresses

NCI-SEER NHL case-control study

- Aim: To identify potential environmental causes of NHL
- Study design:
  - 1321 cases (diagnosed 1998-2000), 1057 controls
  - SEER cancer registries: Detroit, Seattle, Los Angeles County, State of Iowa
  - Age at diagnosis: 20 – 74
  - Home visit, questionnaire:
    - Residential History
    - House dust samples
    - Blood samples
- Exposure Period: lifetime history (1923 - 2000)

Study area and GPS locations of current homes in the NCI-SEER NHL case-control study

Is geocoding accurate enough?

- We compared geocoded addresses to GPS measurements at home interview
- Calculated the positional accuracy for residences located inside and outside town boundaries (rural) in Iowa

Positional accuracy (m) by location of home (rural, in town)

	Commercial firm		Geocoding in ArcGIS	
	Town (n=159)	Rural (n=56)	Town (n=152)	Rural (n=53)
Median	50	212	56	88
(IQR)	(28-83)	(89-747)	(36-92)	(44-254)
p-value	<0.0001		0.0013	

- Rural addresses had greater errors
- Town size did not affect the positional accuracy for ‘urban’ addresses by either method

Ward MH et al, *Epidemiology; 2005*

Agricultural Health Study

- Prospective cohort of pesticide applicators and their spouses in IA and NC in 1993-1996; follow-up interviews in 2000-3 & 2007-10

- ~90,000 farmer/pesticide applicators and spouses

- >150,000 addresses including past homes from short residential history

- ~66% of homes are rural

- Compared rooftop locations (digital orthophotos), E911 locations, and geocoded addresses

Distance between the E911 address location (blue) and the roof top location (red)

Orthophoto is 2 feet (0.6 m) resolution

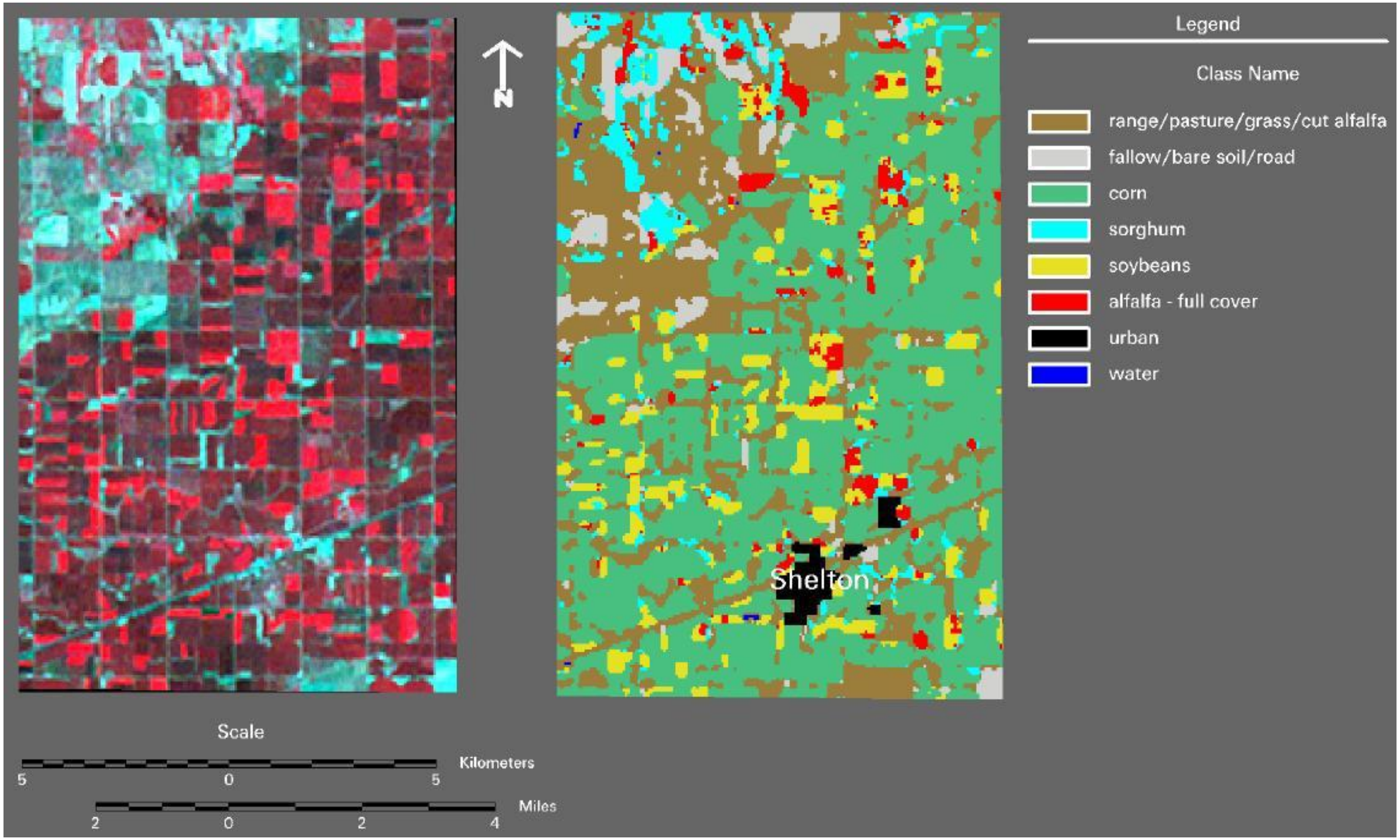
Positional error by geo-location method

Comparison	Median (meters)	Interquartile range (m)
Overall: Address-match vs. Rooftop	90	47-215
E911 vs. Rooftop	39	22-61
Rural: Address-match vs. Rooftop	147	78-353
E911 vs. Rooftop	51	39-83
In-town: Address-match vs. Rooftop	45	27-68
E911 vs. Rooftop	19	17-23

Jones RR et al., *Int J Health Geographics*; 2014

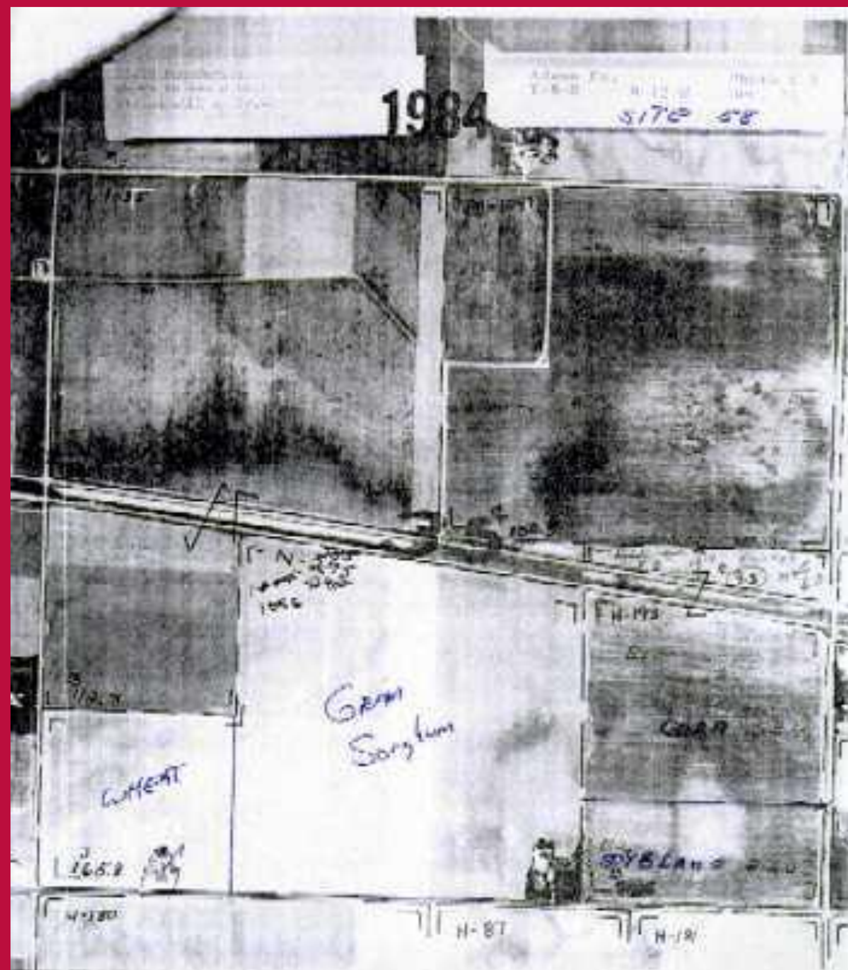
Estimating residential exposure to agricultural pesticides – NCI-SEER NHL study

Historical NASA Landsat imagery can be used to create crop maps



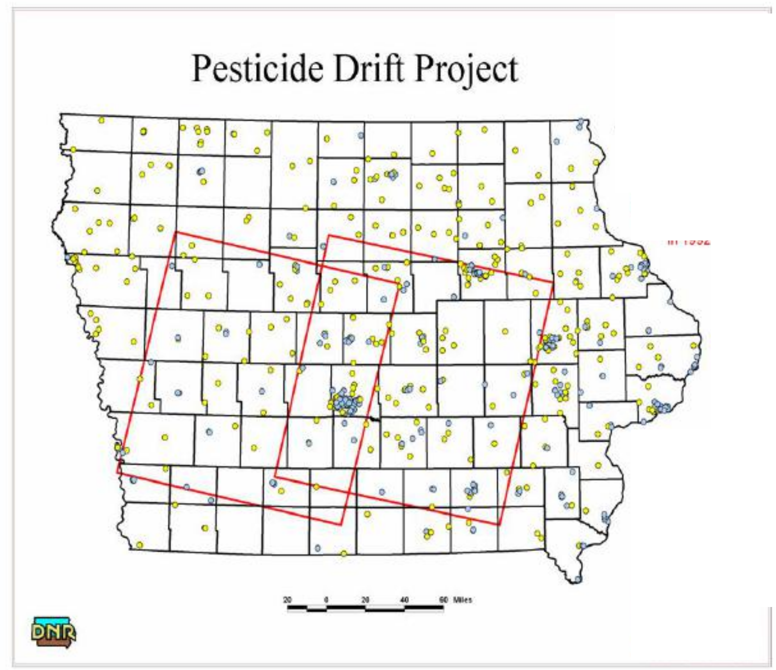
Ward MH et al., *Environ Health Perspect*; 2000

Farm Service Agency has historical aerial photographs with locations of crops



Estimating agricultural pesticides: NCI-SEER NHL case-control study

- **Satellite imagery used for crop maps**
  - **1998-2000**
- **Dust samples**
- **14 agricultural herbicides**



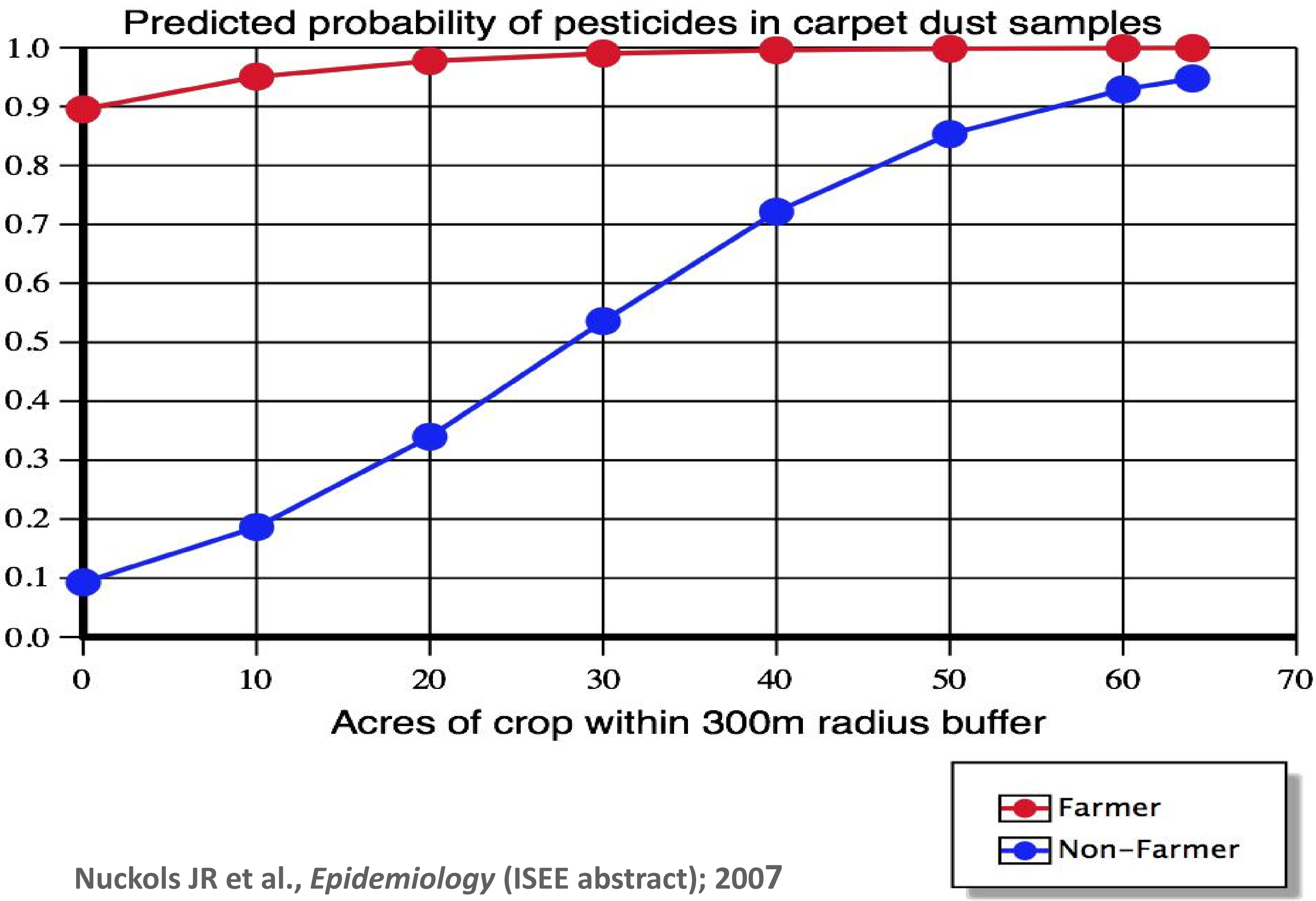
Ward MH et al., *Environ Health Perspect*; 2006

Iowa population with agricultural crops within various buffer distances of residence

Buffer (meters)	All residences	Within towns
100	29%	18%
250	44%	38%
500	60%	52%

Ward MH et al., *Environ Health Perspect*; 2006

Probability of detecting agricultural herbicides in homes of farmers (red) and nonfarmers (blue)

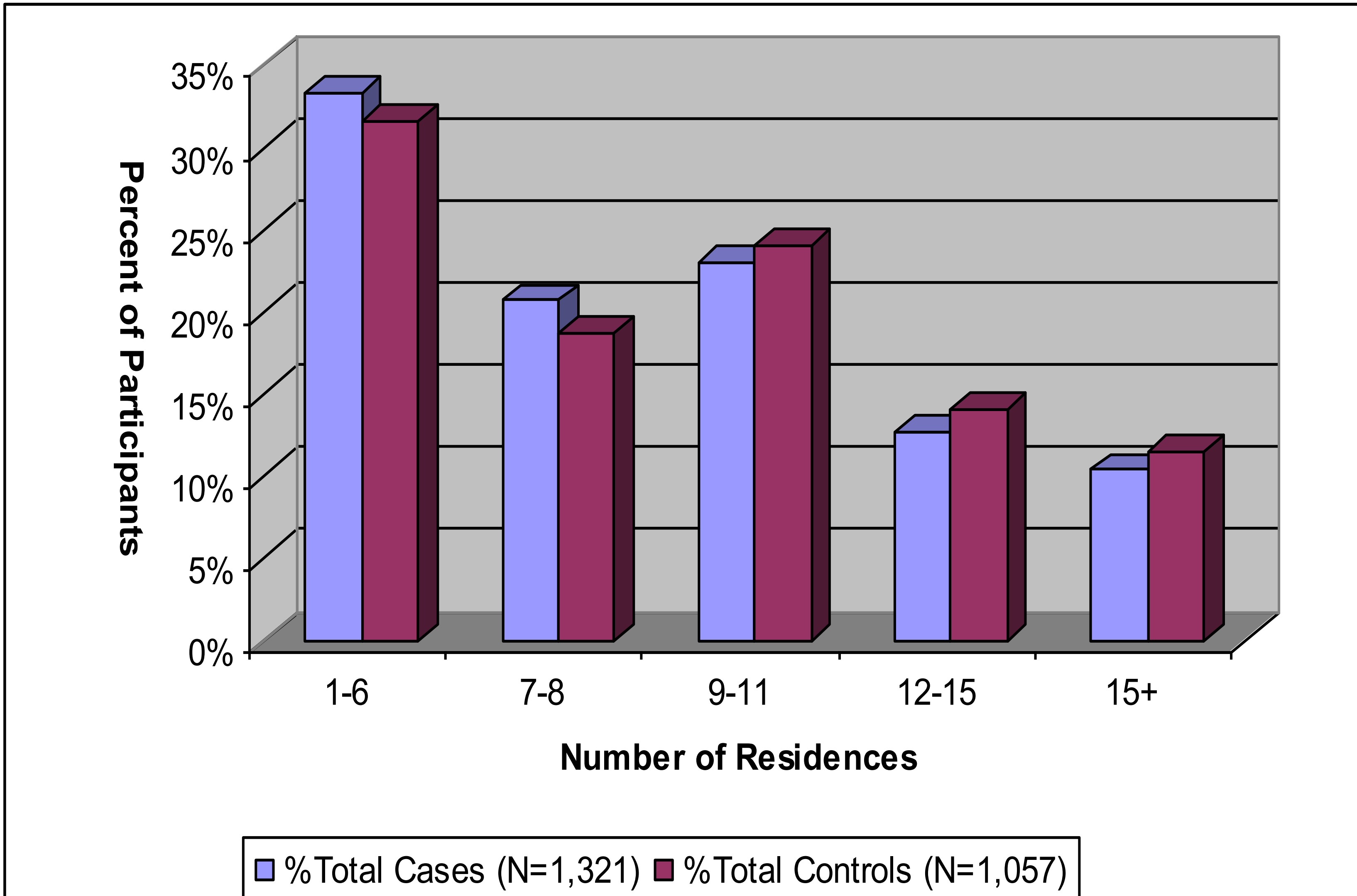


Nuckols JR et al., *Epidemiology* (ISEE abstract); 2007

NCI-SEER NHL Study: Current and past homes

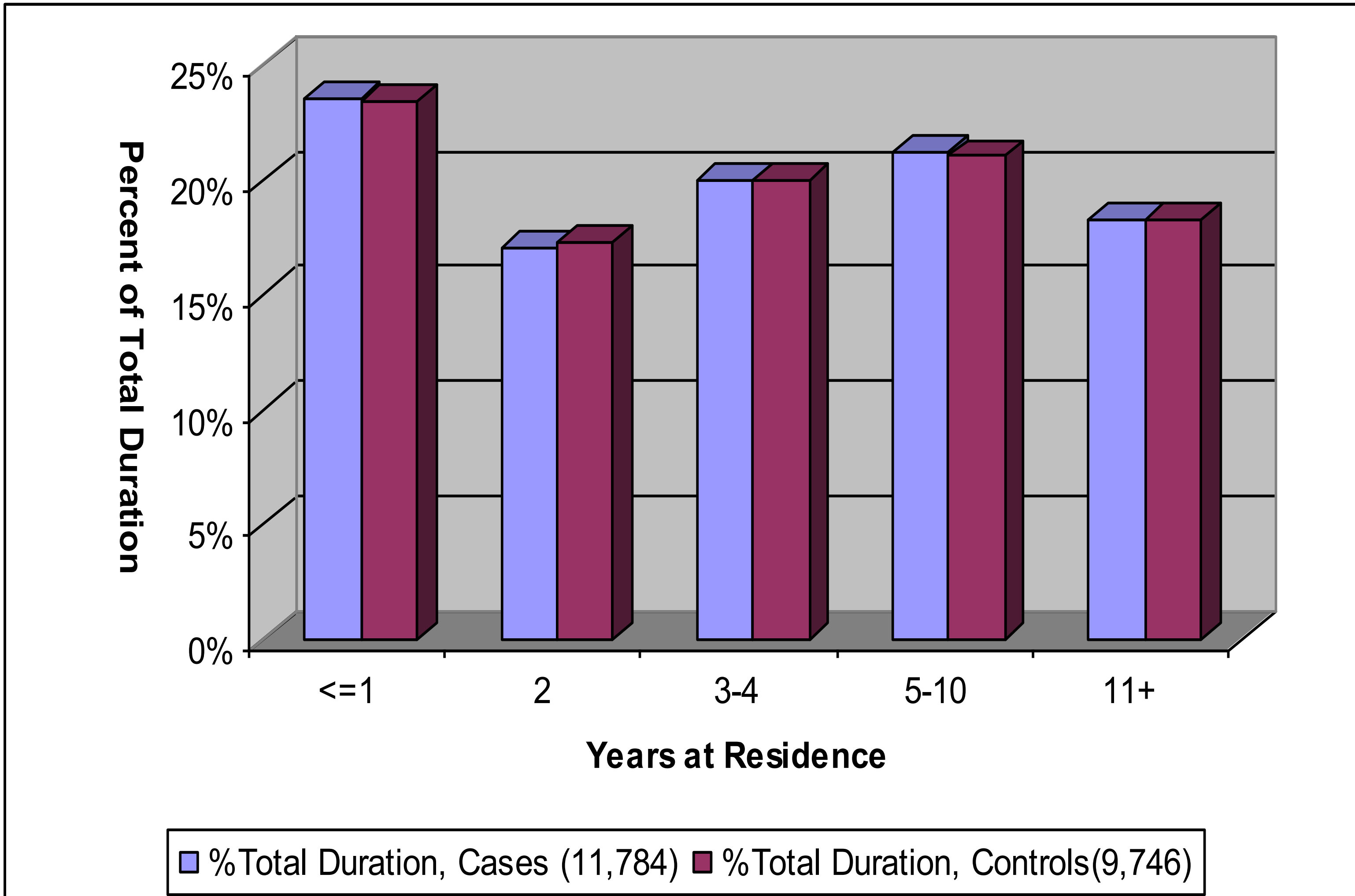
Nuckols JR et al., *Epidemiology* (ISEE abstract); 2009

NHL Number of Residences



Nuckols JR et al., *Epidemiology* (ISEE abstract); 2009

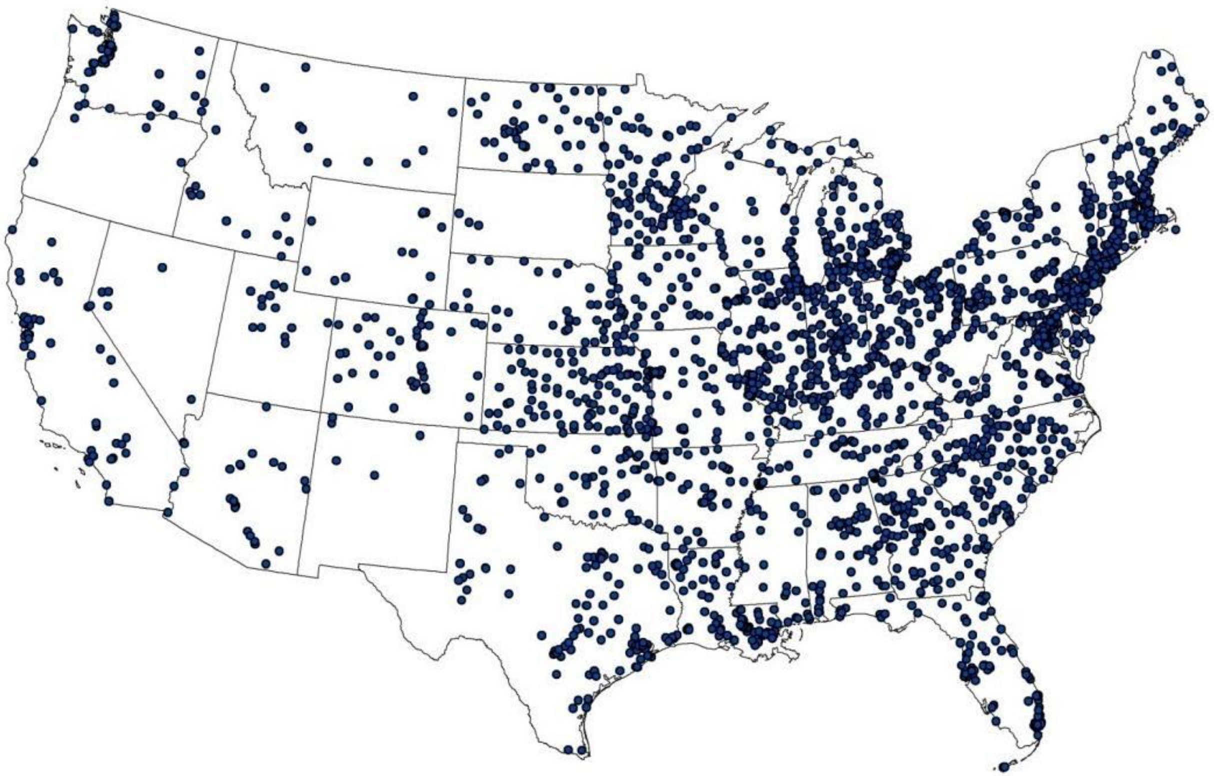
NHL Duration at each Residence



Nuckols JR et al., *Epidemiology* (ISEE abstract)*;* 2009

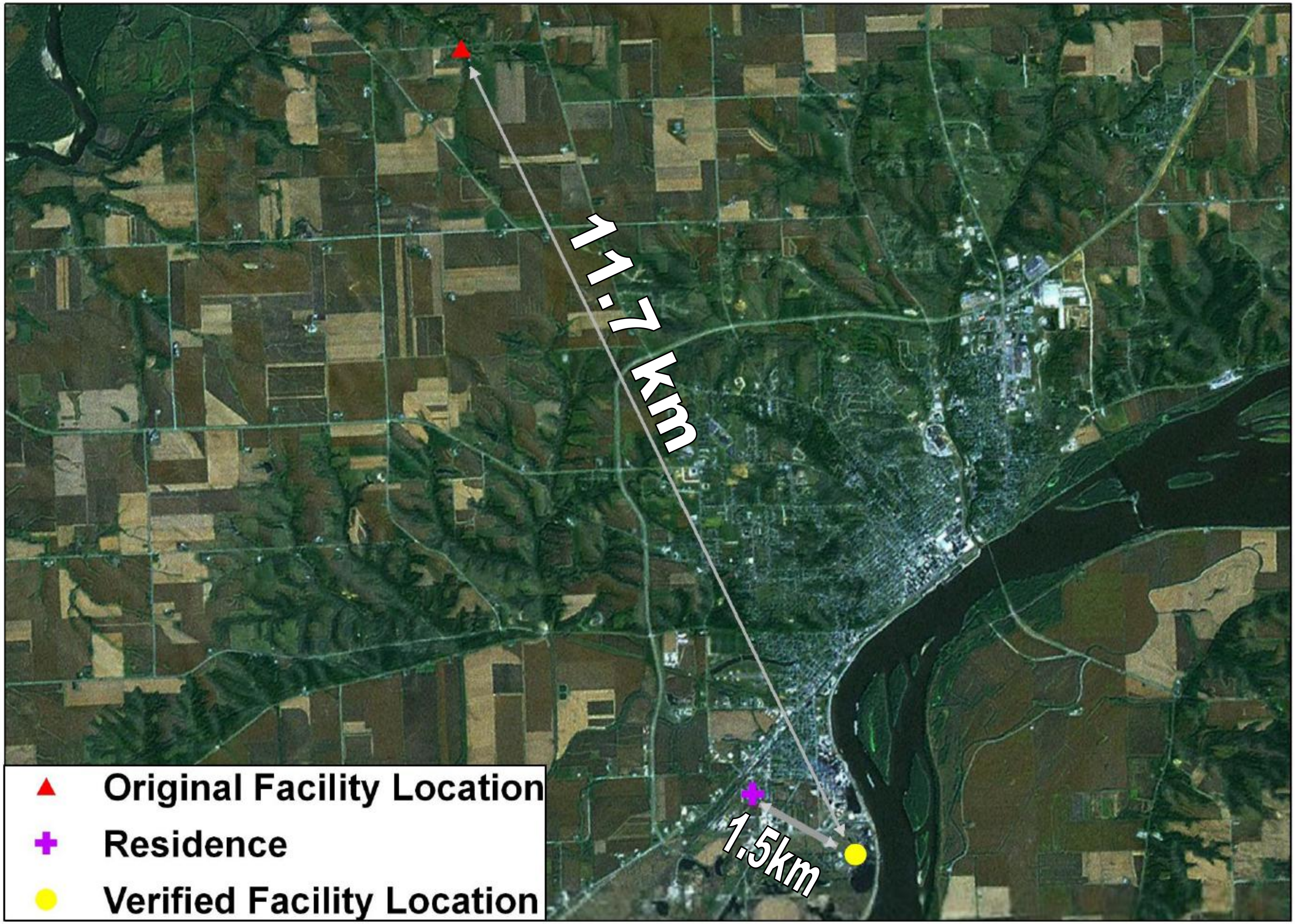
DIOXIN EMITTING FACILITIES

EPA INVENTORY, 1987 and 1995



EPA: An inventory of sources and environmental releases of dioxin-like compounds in the United States for the years 1987, 1995, and 2000. U.S. Environmental Protection Agency; 2006.

Original and Verified Facility and Residence




1,500 meters

Distance between EPA Facility Geocode and Verified Location (Km)

	Cement Kilns	Coal Power Plants	Hazard Waste Incinerator	Medical Waste	Municip. Waste	Sewage Sludge incinerator
25th	1.2	0.05	4.4	0.0	3.2	1.0
Median	4.0	0.12	11.5	0.0	7.8	4.6
75th	6.7	0.45	32.8	2.7	29.0	23.4
Mean	7.0	0.93	23.5	48.7	23.1	13.8

Pronk A et al., *Environ Health*; 2013

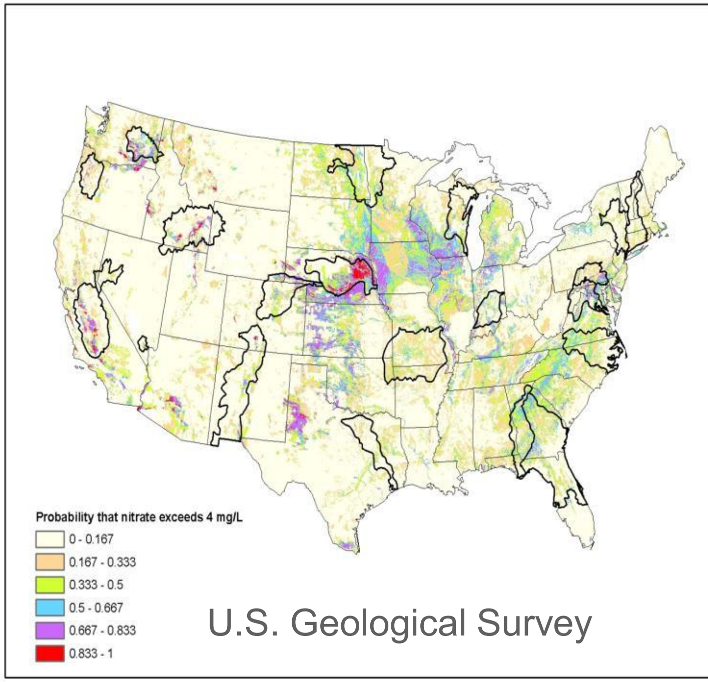
Nitrate in drinking water: Sources and exposures

- **Nitrogen fertilizers, animal and human waste**

- **Maximum contaminant level: 10 mg/L as NO_3-N**

- **Highest exposures:**
  private well users

- **Measurements are sparse**



Drinking water nitrate and cancer in two cohorts in agricultural areas

- Iowa Women's Health Study (IWHS):
  - 42,000 postmenopausal women
  - Drinking water source and duration in 1989
  - 73% use public supplies, 25% private wells
  - Many surface water supplies

- Agricultural Health Study (AHS):
  - 60% use private wells

Iowa Women's Health Study:
Public water supply exposure assessment

Median duration (years)

Public Water Supply Monitoring Data, 1955-1988

NO_3-N

Total Trihalomethanes (THMs)

City

N = 473

4

8

16

34

Exposure Metrics

Average NO_3-N and THMs

Years >1/2 Maximum Contaminant Level

82% used supply 16+ years

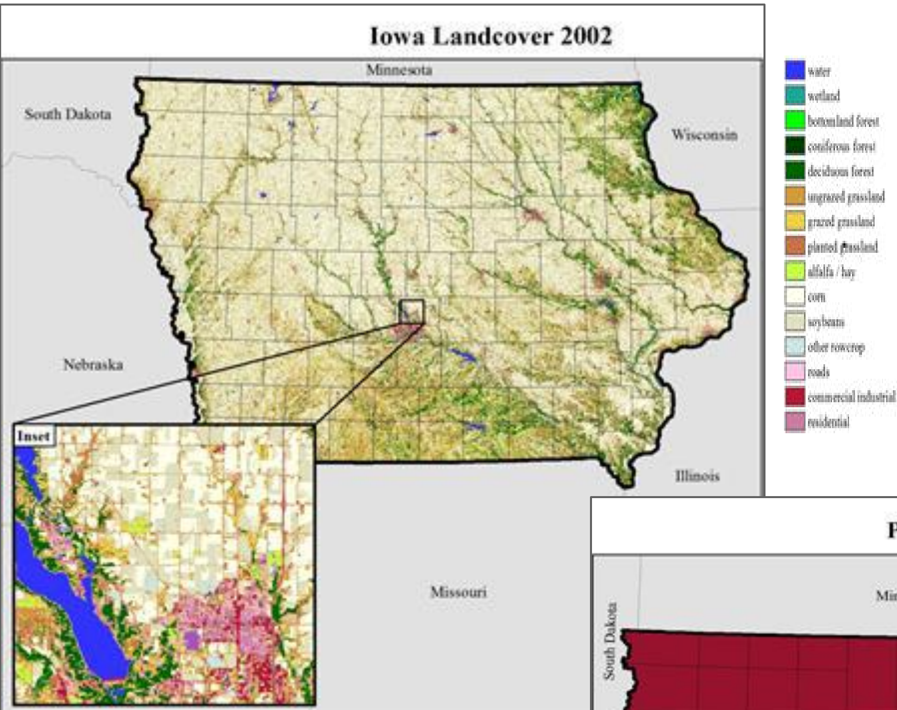
Ward MH et al., *Epidemiology*; 2010

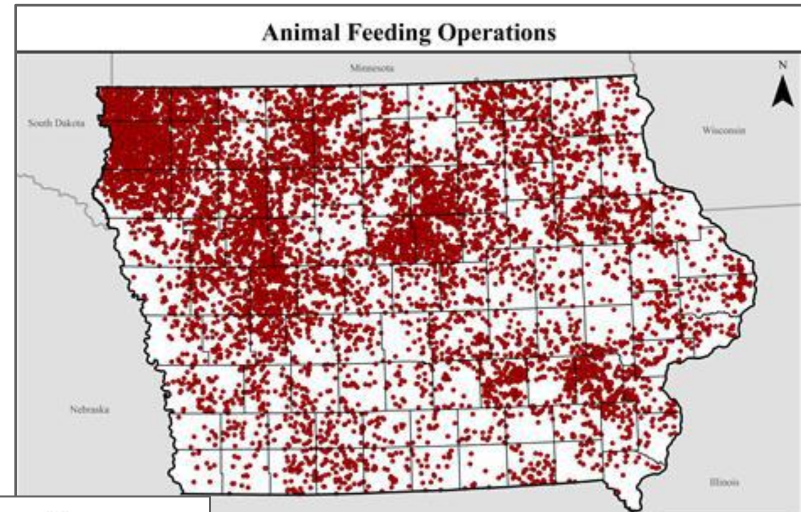
Inoue-Choi M et al., *Int J Cancer*; 2014

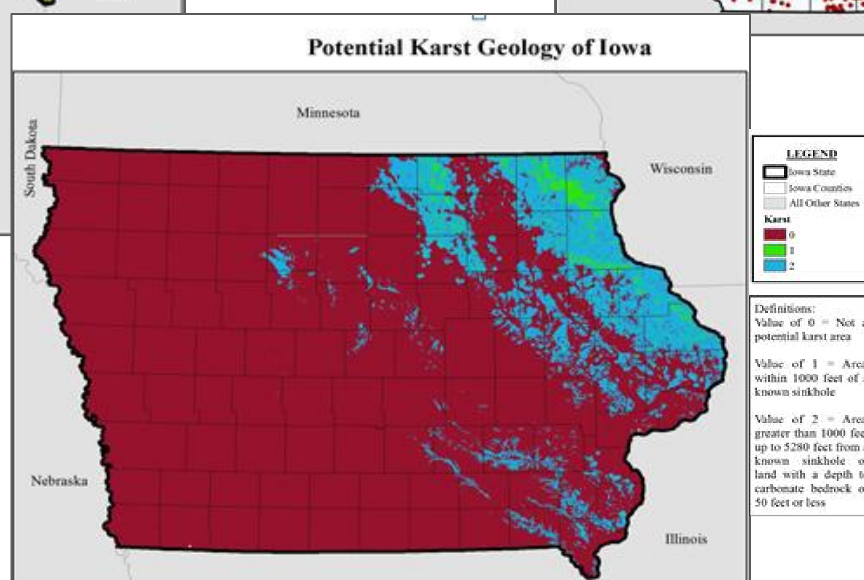
Jones RJ et al., *Environ Health Perspect*; in press

GIS-based model of nitrate in private wells

- **~34,000 nitrate measurements** (1980-2000s)
- **Evaluated >150 variables (e.g., land use, animal feeding operations, geology, soils)**







Random forest model performed best

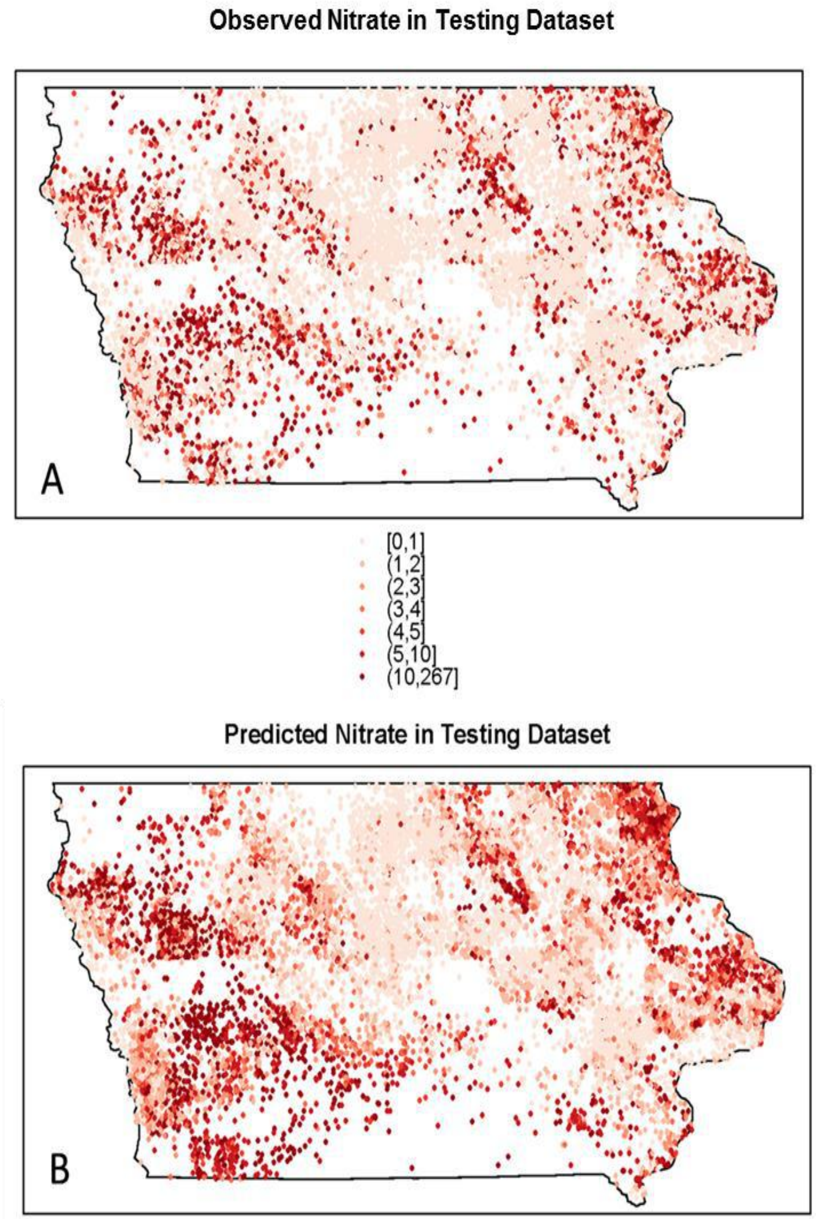
-66 variables explained 77% of variation in training dataset:

- Well depth
- Geologic features – karst geology, sinkholes
- Slope, elevation
- Animal feeding operations
- Agricultural land (1990)
- Precipitation
- Soil characteristics at well screen
- Year

Wheeler DC/Nolan BT et al., *STOTEN* 2015

Sensitivity and Specificity (5 mg/L)

		NO$_3$-N Observed	
		≥5 mg/L	<5 mg/L
NO$_3$-N Predicted	≥5 mg/L	**Sensitivity** 2615 **(67%)**	2598 (14%)
	<5 mg/L	1280 (33%)	**Specificity** 15660 **(86%)**



Wheeler DC/Nolan BT et al, *STOTEN*; 2015

GIS for exposure assessment - summary

- Assess exposures not easily obtained by questionnaires or biomonitoring
- Need to assess and quantify uncertainty in spatial data
- Validation of GIS-based exposure metrics for assessing individual exposure
- Include activity patterns to refine exposure assessment

Nuckols JR et al, *Environ Health Perspect;* 2004

Ward MH and Wartenberg D, *Am J Epidemiol*; 2006

GIS for improving exposure assessment in cancer studies

- **Will be increasingly useful in future - data availability, improved technology for satellite imagery, geocoding**

- **Requires an interdisciplinary approach:**
  - **Geographers, environmental engineers, chemists, environmental scientists, epidemiologists**

Nuckols JR et al, *Environ Health Perspect;* 2004

Ward MH and Wartenberg D, *Am J Epidemiol*; 2006

Acknowledgements/Co-authors

Jim Cerhan, Mayo College of Medicine

Anneclaire De Roos, Drexel U.

Jim Giglierano, Iowa Geologic Survey

Susan Maxwell, Colorado State, USGS

Tom Nolan, USGS

Jay Nuckols, Colorado State U.

Philip Riggs, Colorado State U.

Kim Robien, George Washington U.

Gerry Rushton, U Iowa

Peter Weyer, U Iowa

David Wheeler, Virginia Commonwealth U.

NCI:

Joanne Colt

Maki Inoue-Choi

Rena Jones

Jay Lubin

Anjoeka Pronk

Westat:

Matt Airola

Abigail Flory

DEPARTMENT OF HEALTH & HUMAN SERVICES • USA

NIH NATIONAL CANCER INSTITUTE

www.cancer.gov www.cancer.gov/espanol