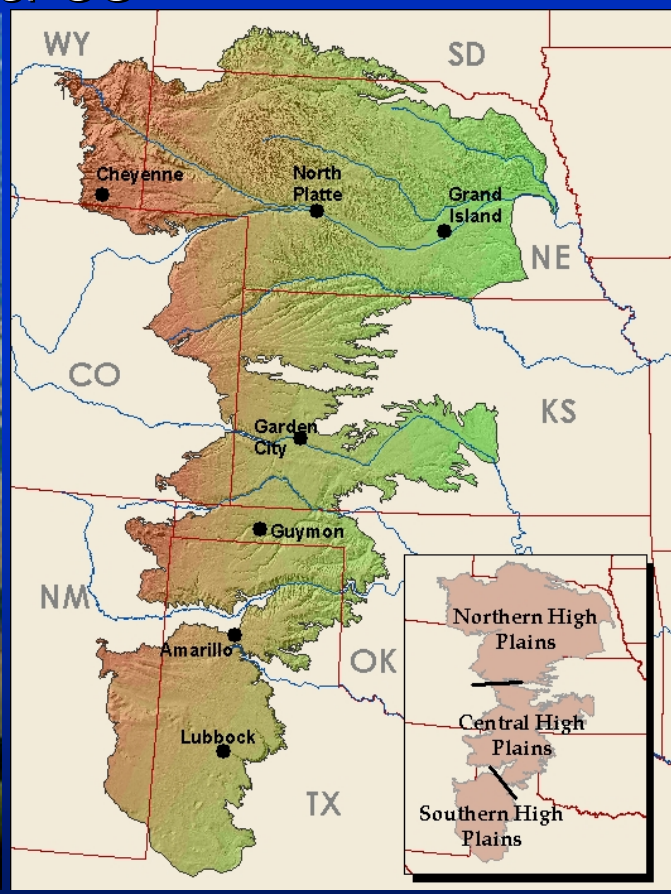
Using Logistic Regression to Assess Regional Ground-Water Vulnerability: High Plains Aquifer

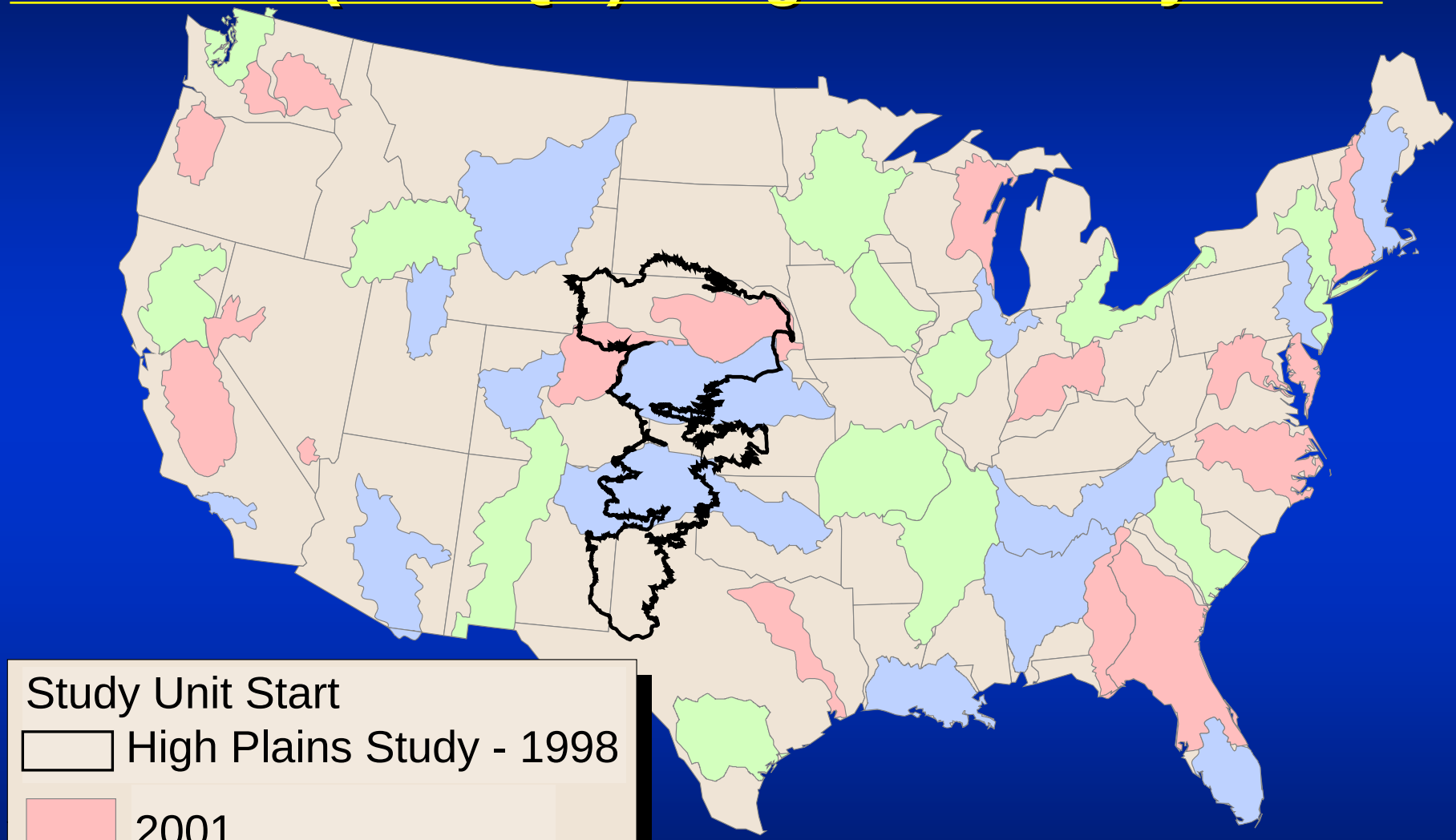
Jason Gurdak

U.S. Geological Survey, Denver CO



USGS science for a changing world

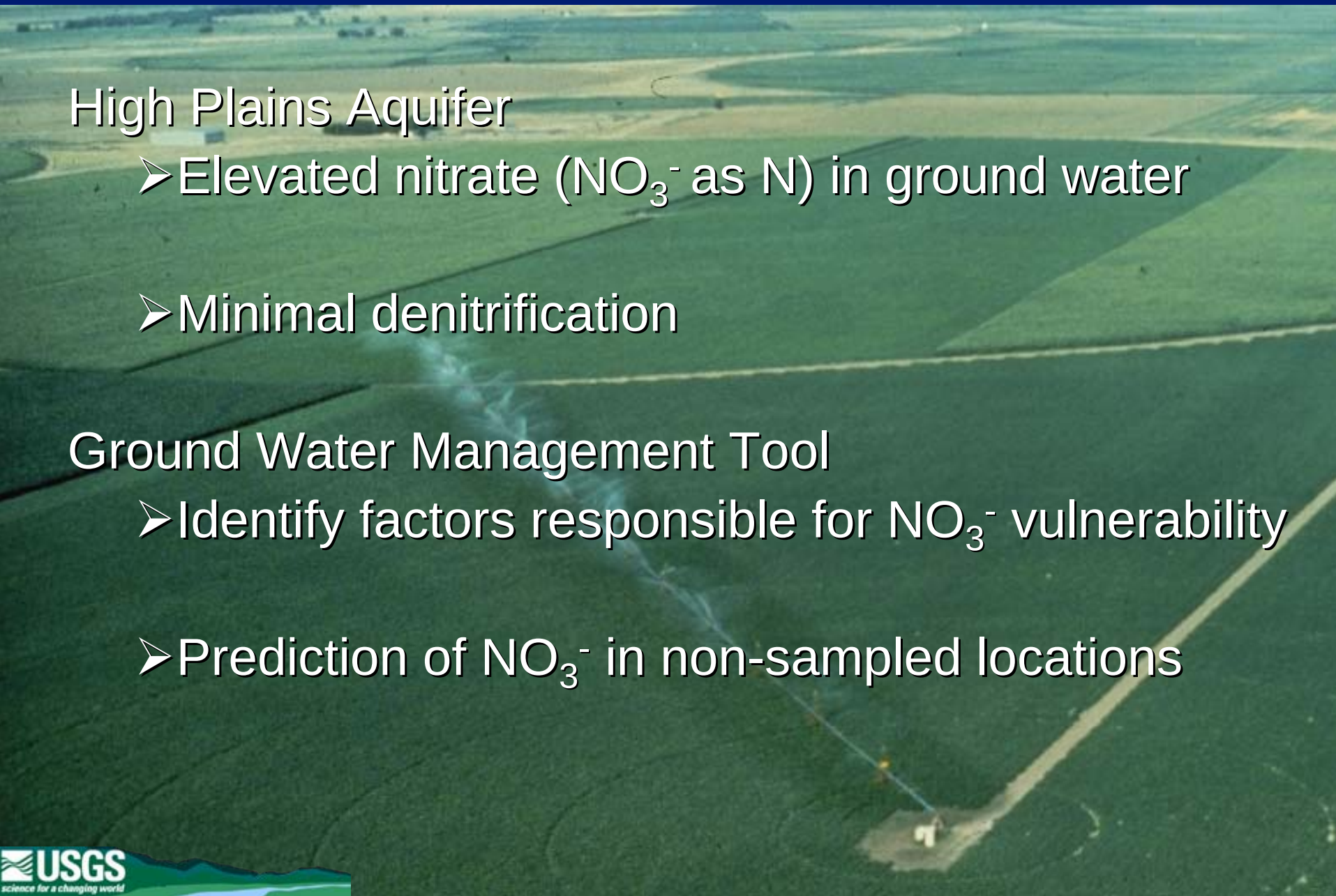
USGS National Water Quality Assessment (NAWQA) Program – Study Units



Outline

- Ground-Water Vulnerability?
- Conceptual Model –
  High Plains Aquifer
- Methods
  - Ground water flow and particle tracking simulations
  - Logistic regression
- Vulnerability to Nitrate
- Uncertainty Estimates
  - Latin hypercube sampling simulations

Why Assess Ground-Water Vulnerability



High Plains Aquifer

- Elevated nitrate (NO_3^- as N) in ground water
- Minimal denitrification

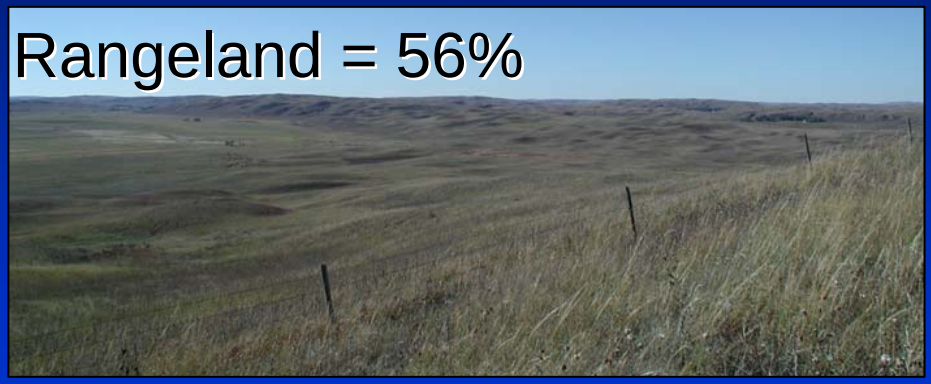
Ground Water Management Tool

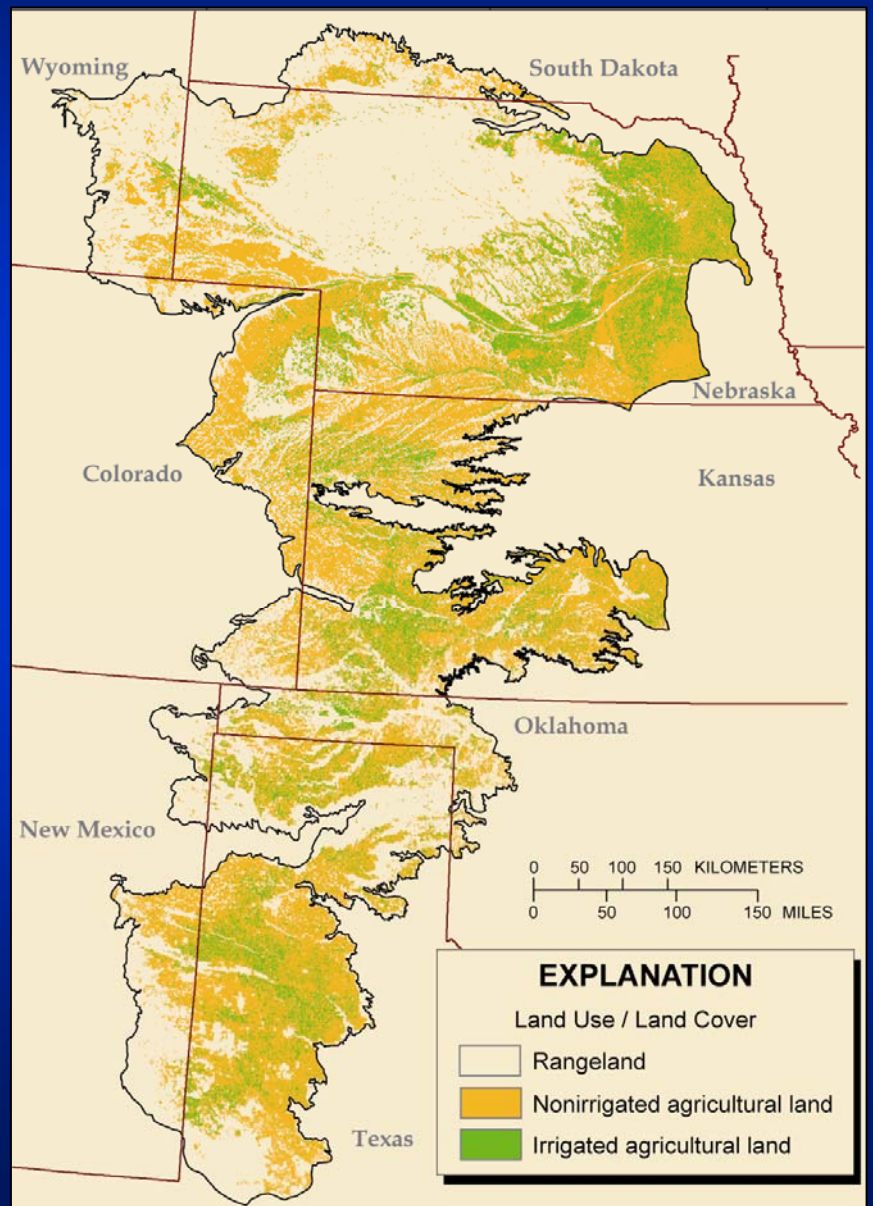
- Identify factors responsible for NO_3^- vulnerability
- Prediction of NO_3^- in non-sampled locations

USGS
science for a changing world

Conceptual Model – Spatial Variability

Land Use / Land Cover



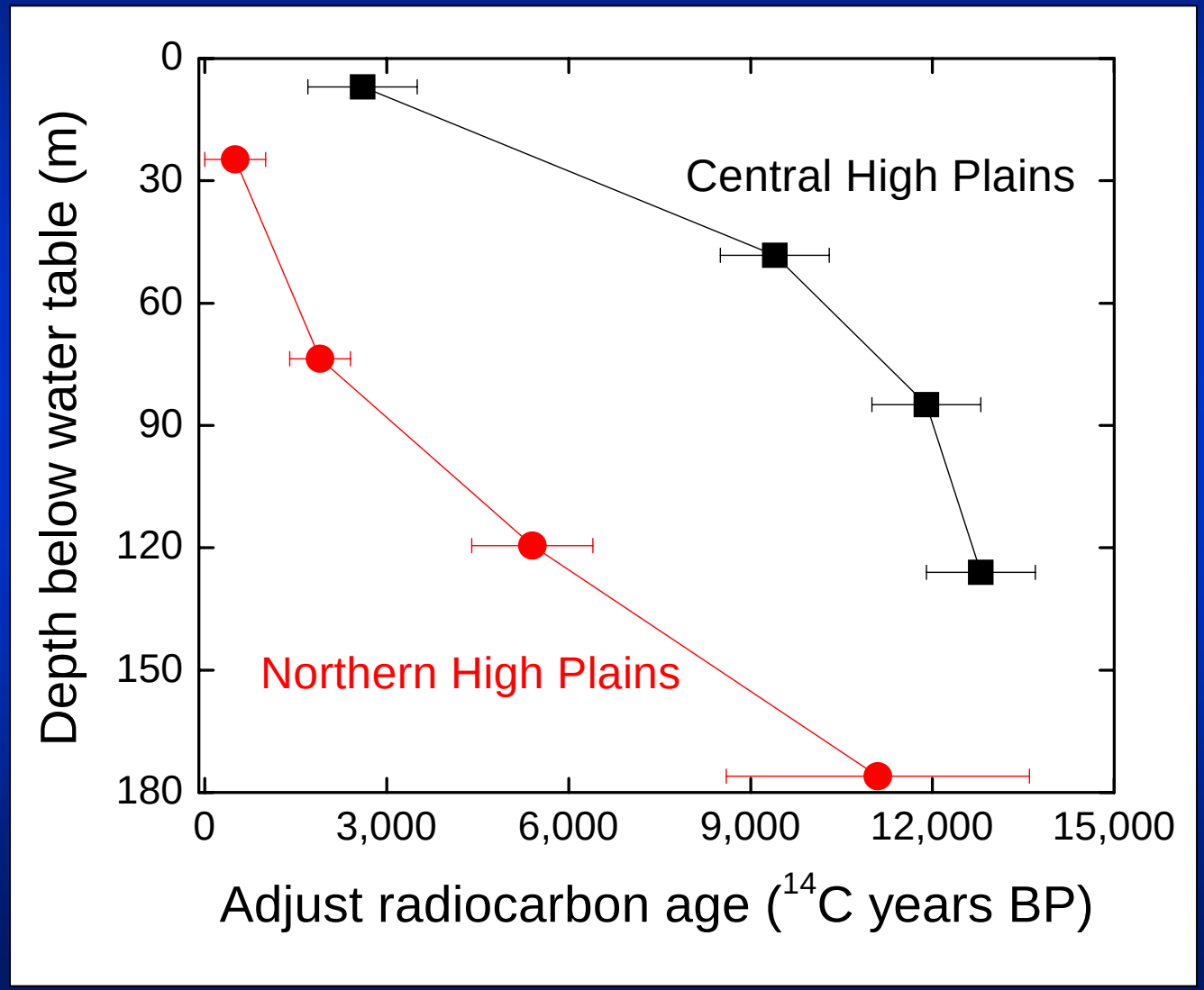


USGS science for a changing world

Conceptual Model – Temporal Variability

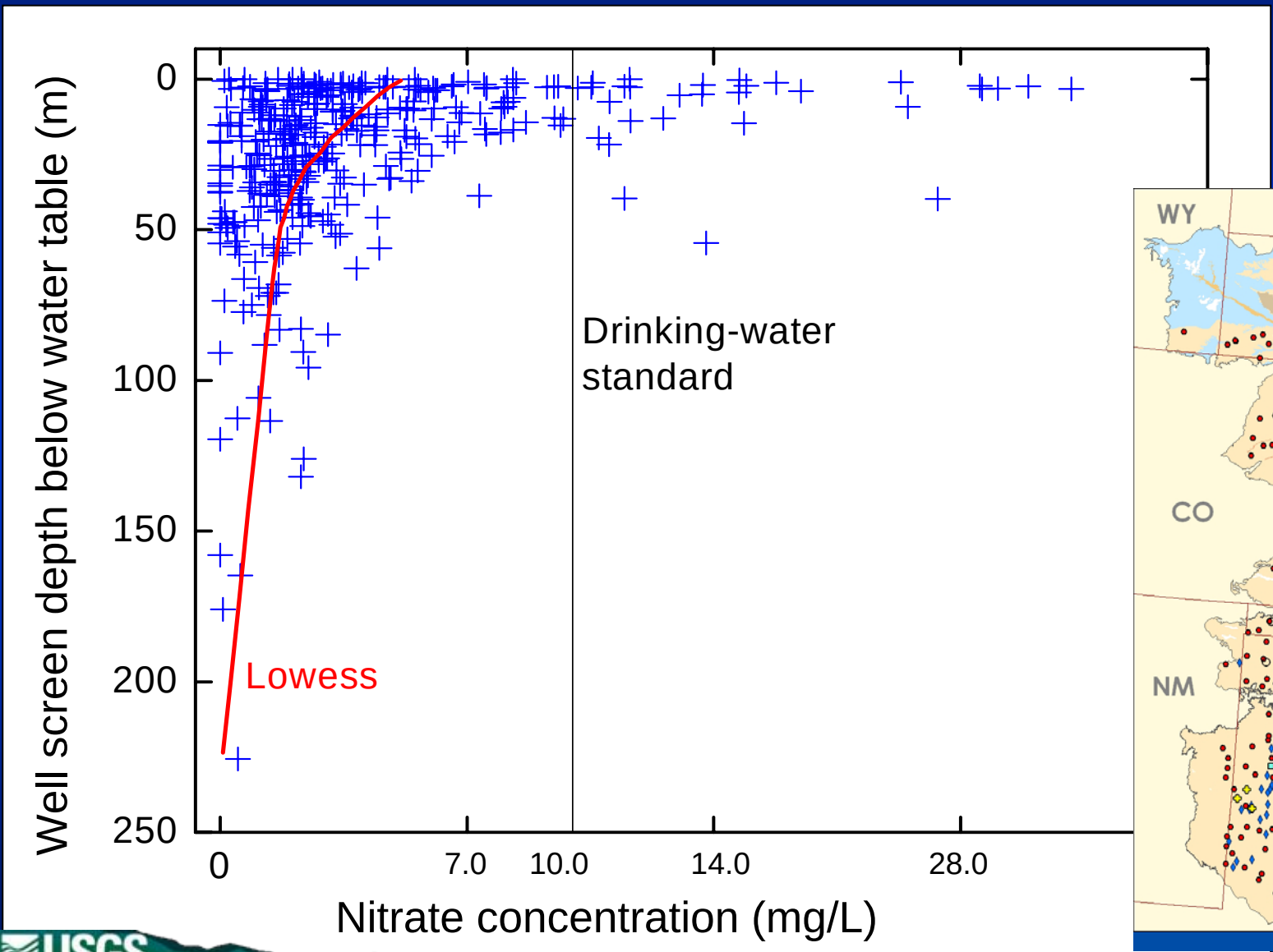
Paleorecharge nitrate:
1.9 to 3.5 mg/L

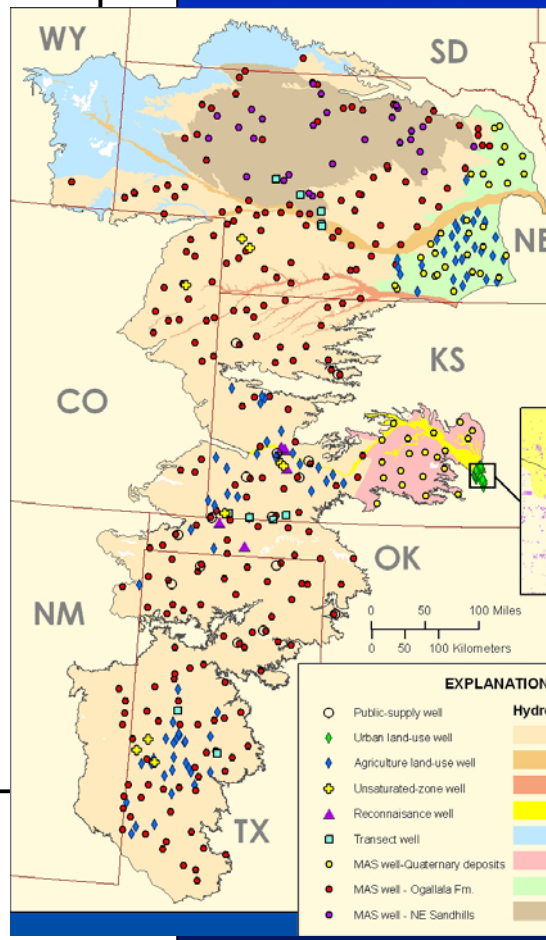
Vertical Gradients in Ground Water Age



from - McMahon and others (2004)

Spatial and Temporal Variability – Nitrate Concentrations in Ground Water





USGS science for a changing world

Objectives – Vulnerability Assessment

1. Identify factors responsible for vulnerability to NO_3^-

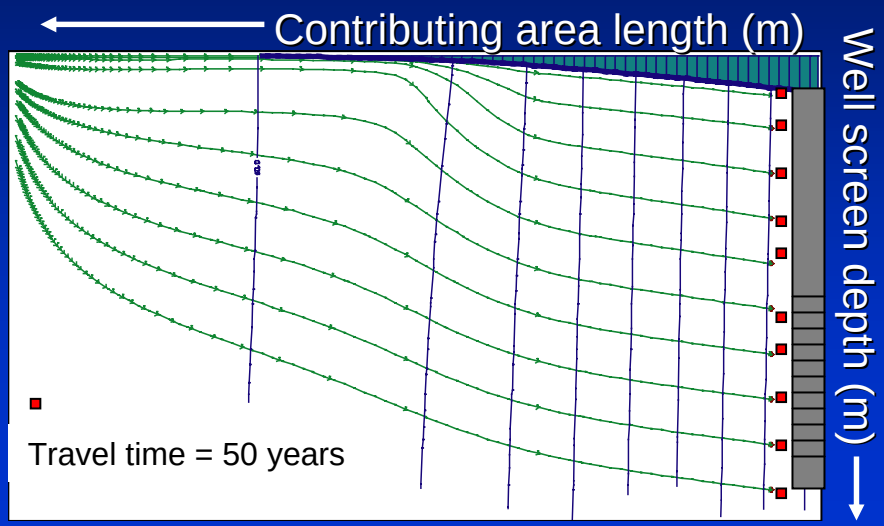
2. Develop empirical model and map that predicts the probability of detecting concentrations of NO_3^- (as N) > 4 mg L^{-1} in recently recharged ground water (< 50 years) of the High Plains aquifer.

3. Quantify uncertainty associated with these vulnerability predictions.

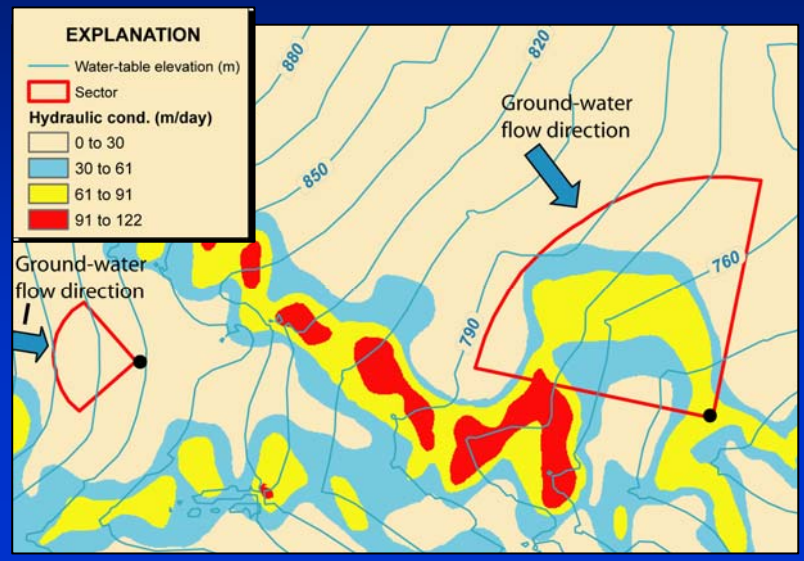
USGS
science for a changing world

Methods

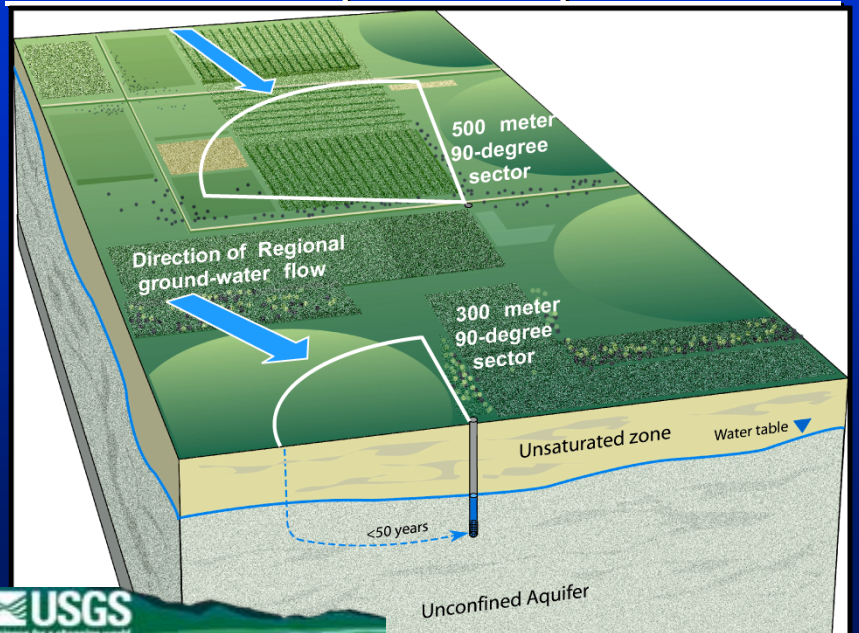
1. Particle Tracking Simulations



2. Define Contributing Area



3. Extract Explanatory Variables



4. Independent Validation

336 wells selected that intercept recently (<50 years) recharged ground water.

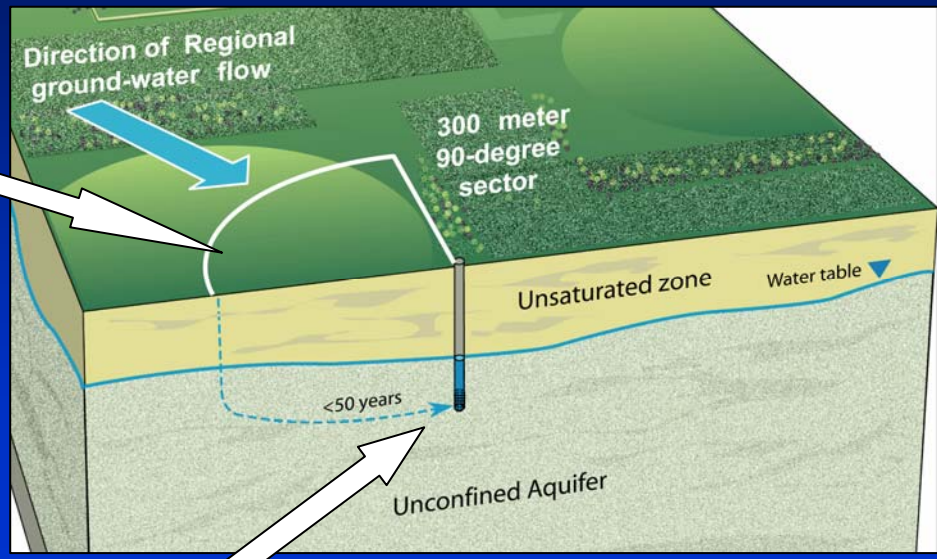
- Ground-water age – Tritium (^3H) method
  86% of wells analyzed for ^3H indicate that ground water was recharged during the last 50 years.

USGS

Methods – Logistic Regression Modeling

Explanatory Variables:

- Land Use
  - Irrigated agricultural land
  - Non-irrigated agricultural land
  - Rangeland
- Nitrogen application
  - Commercial fertilizer
  - Manure
  - Soil residuals
  - Atmospheric deposition
- Soils (clay, organic matter, permeability, thickness, etc)
- Unsaturated zone lithology
- Depth to water
- Saturated thickness
- Number of irrigation wells
- Number of playa basins



Response Variable:

- Nitrate (NO_3^- as N) concentrations greater than "background".
  - Background concentration = 4 mg/L (paleorecharge water nitrate concentrations: 1.9 to 3.5 mg/L)

Logistic Regression Model:

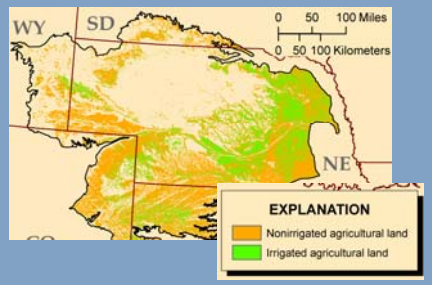
- Predicts the probability of detecting nitrate > 4 mg/L in recently (< 50 years) recharged ground water.

USGS science for a changing world

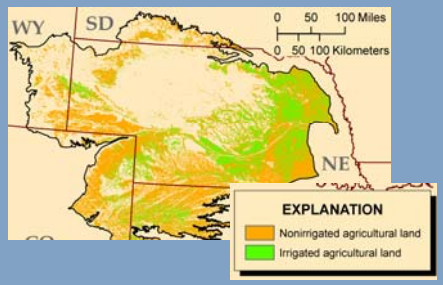
Results – Logistic Regression Models

Northern High Plains (NHP) Model:

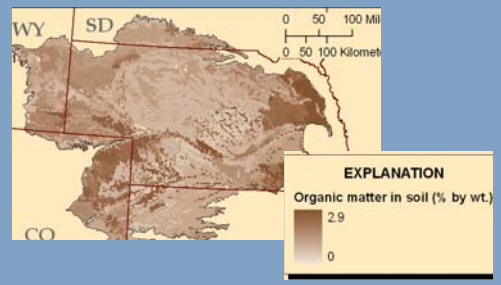
Non-Irrigated Ag (1.566)

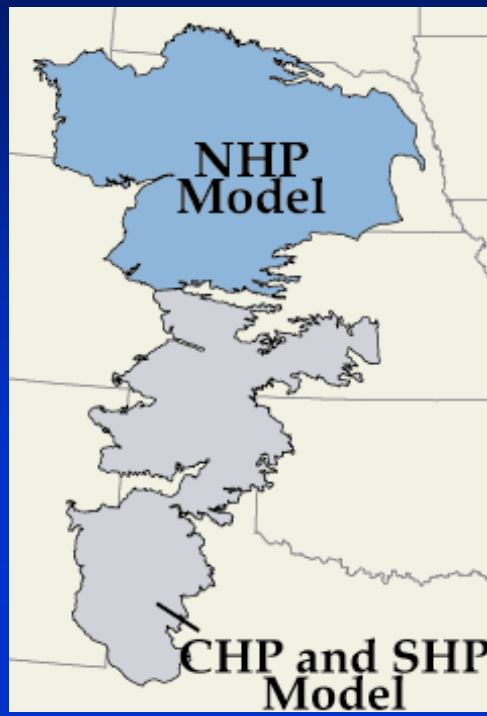


Irrigated Ag (1.133)



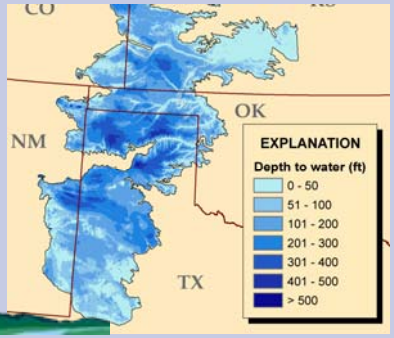
Soil Organic Matter (-0.741)



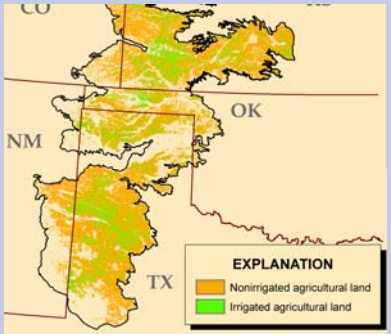


Central and Southern High Plains (CHP and SHP) Model:

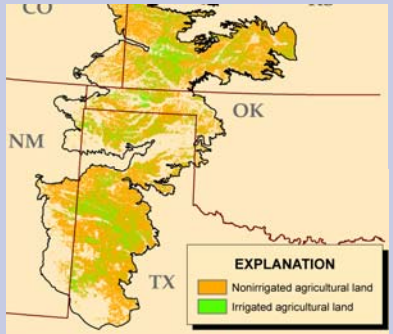
UZ Thickness (-1.621)



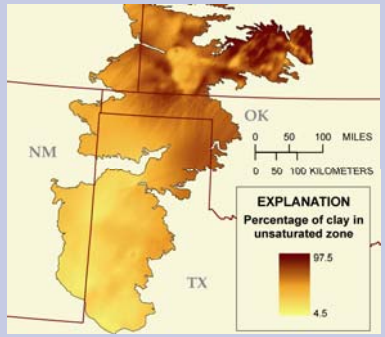
Non-Irrigated Ag (0.909)



Irrigated Ag (0.710)



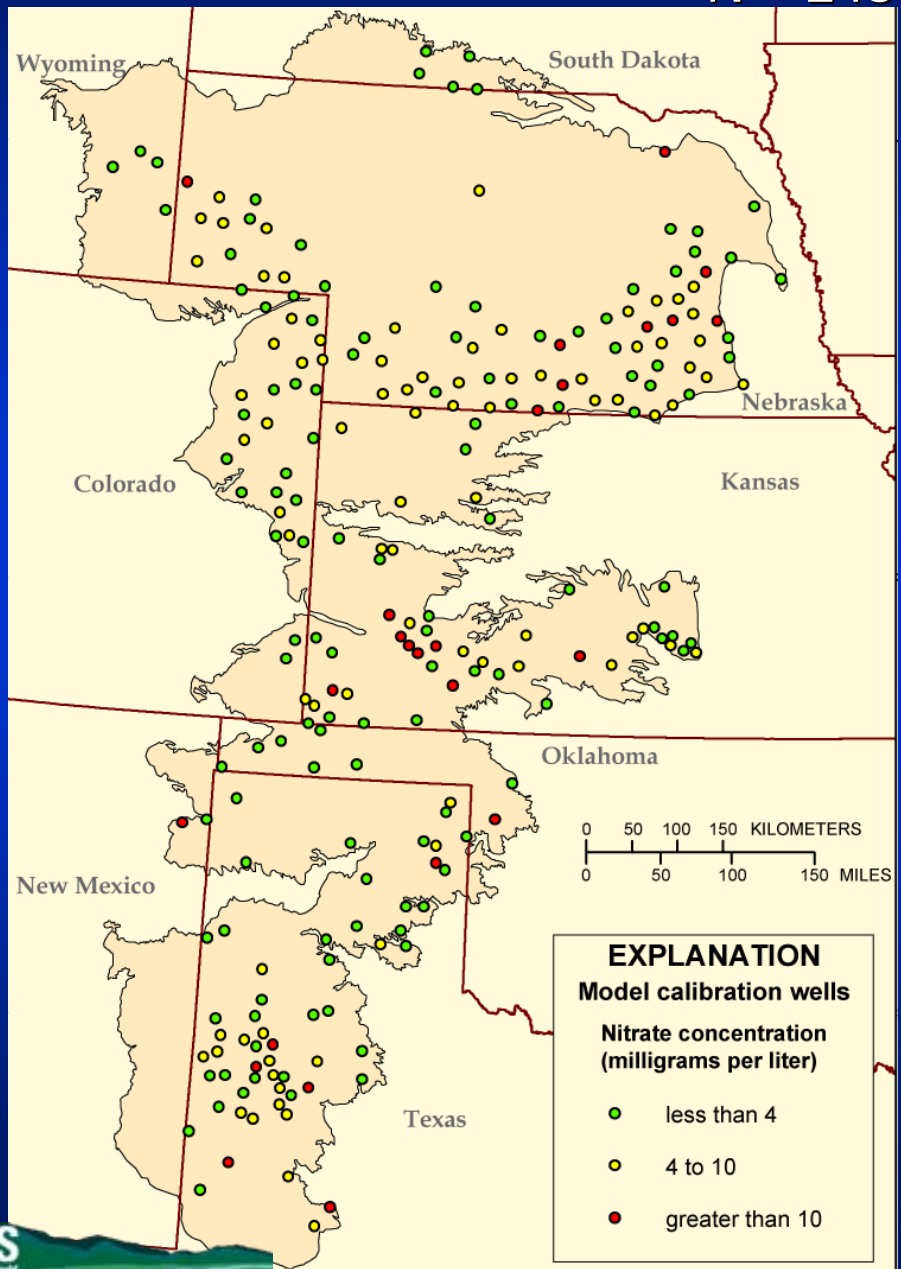
Clay in UZ (%) (-0.678)

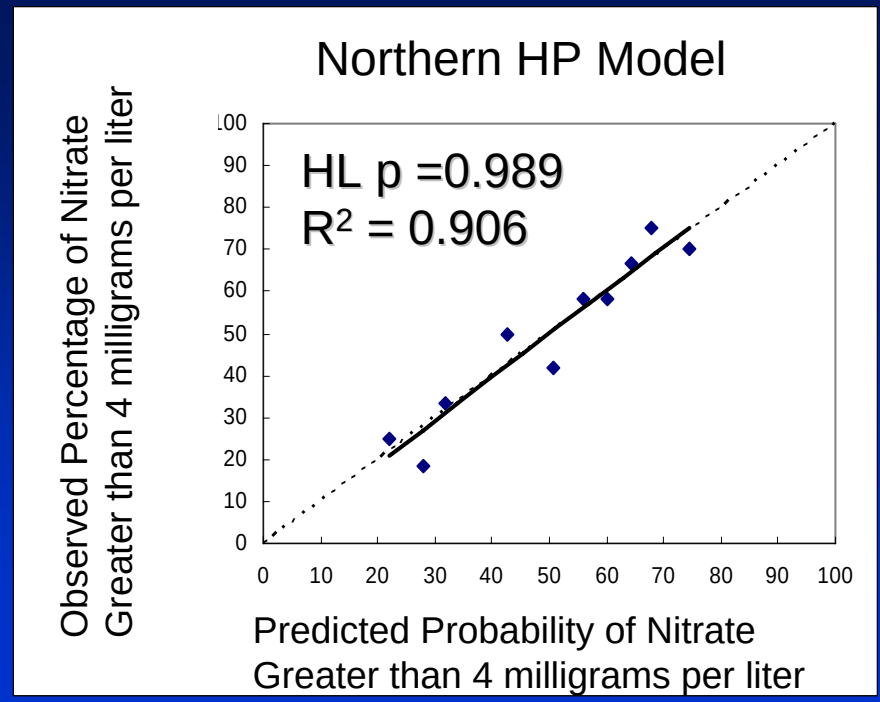


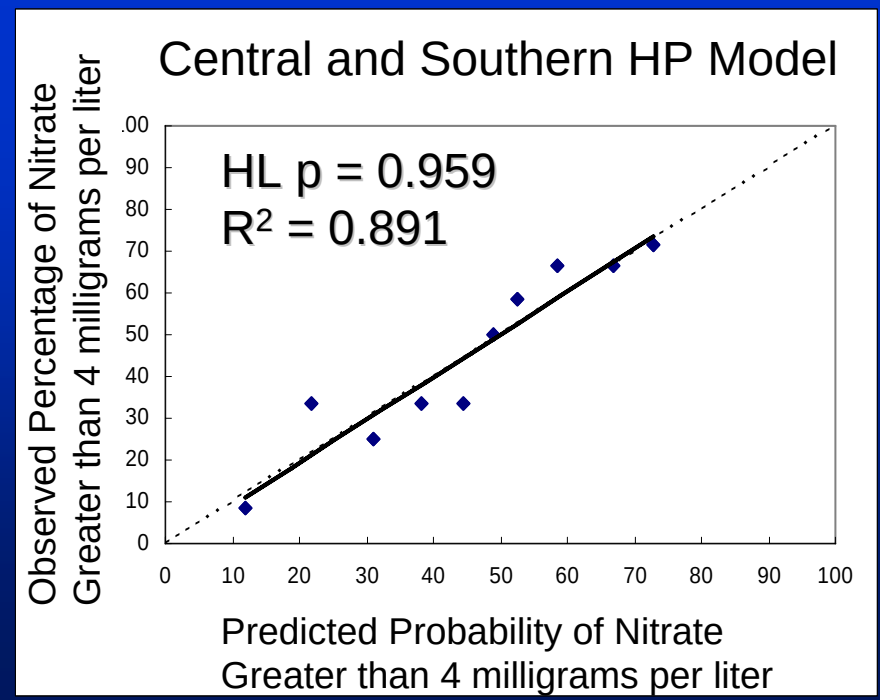
USGS

Model Calibration

N = 243





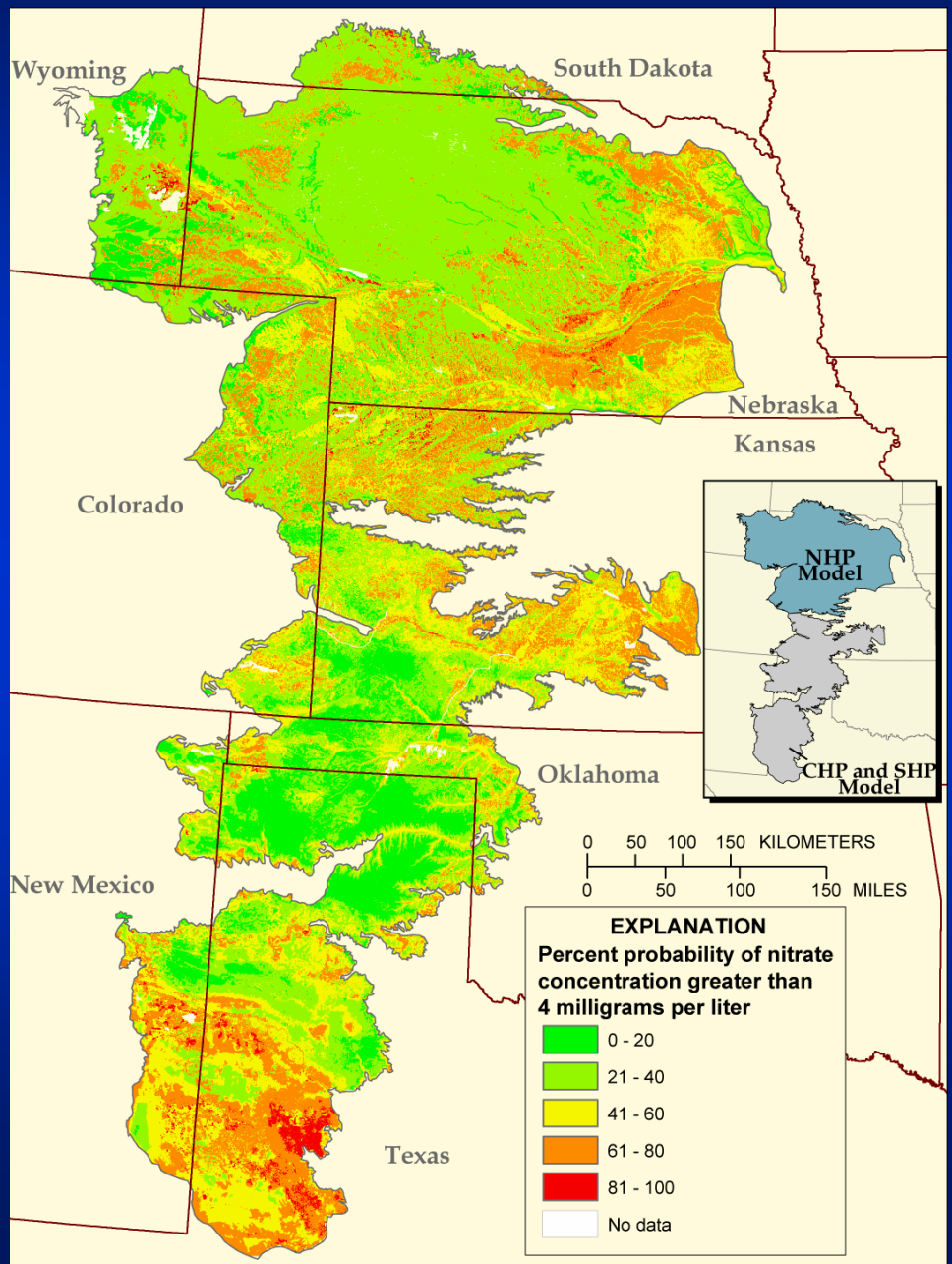


USGS

Vulnerability Map

Predicted probability of detecting $NO_3 > 4$ mg L^{-1} in recently recharged (<50 years) ground water.

Predicted Probability		Percentage of Study Area (%)
0 - 20	:	11.9
21 - 40	:	41.3
41 - 60	:	25.6
61 - 80	:	19.6
81 - 100	:	1.3

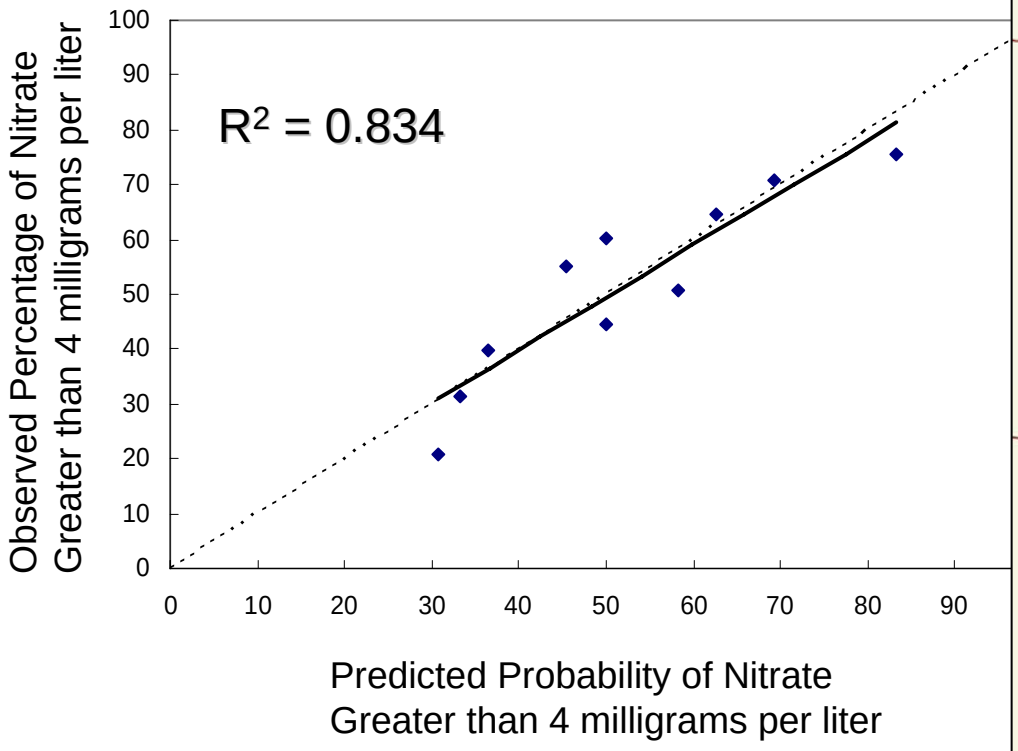


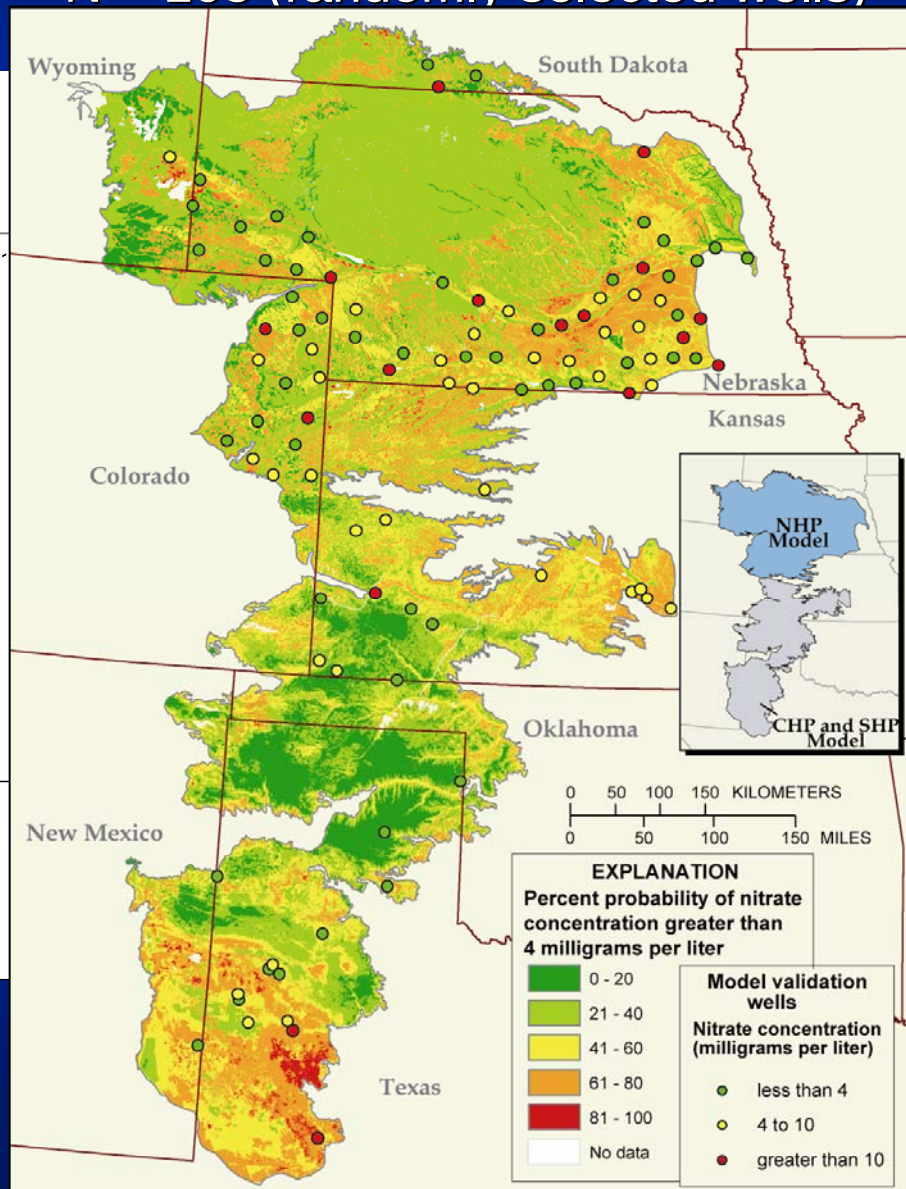
USGS science for a changing world

Model Validation

N = 105 (randomly selected wells)

Combined Northern Model and
Central and Southern Model





USGS science for a changing world

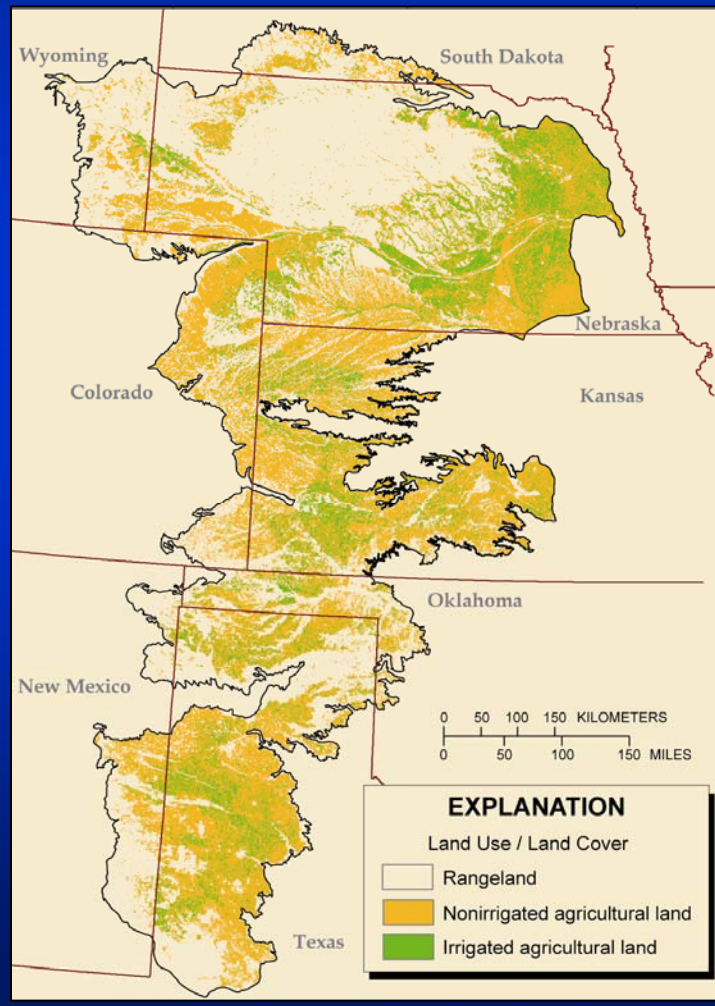
Uncertainty of Vulnerability Predictions

Two Sources of Error:

1. Estimated coefficients in the Logistic Regression model, due to lack of spatial coverage of monitoring wells.
2. GIS explanatory variables, due to accuracy, precision, and processing of GIS data.

- Error Propagation
  - Latin Hypercube Sampling

Example:
80% Correct Identification

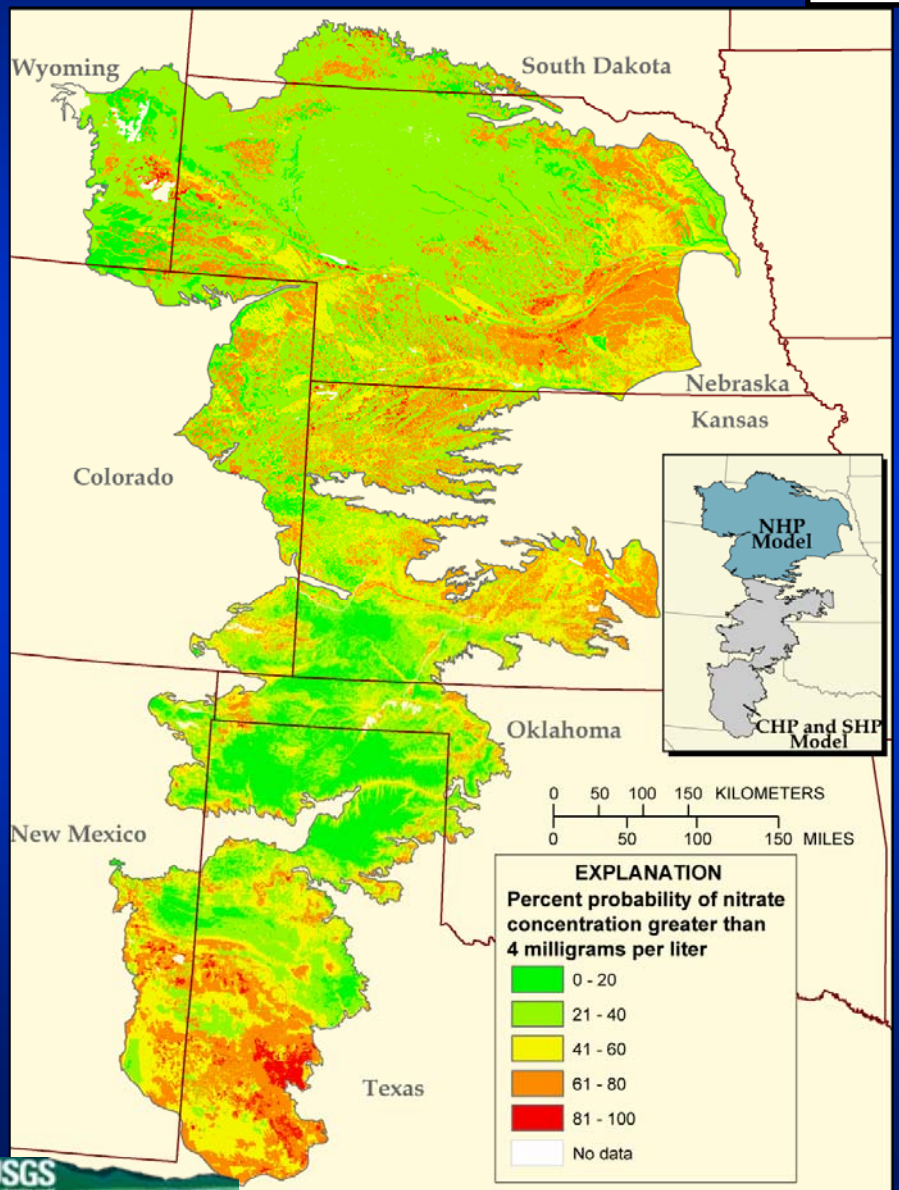


Qi and others, 2002

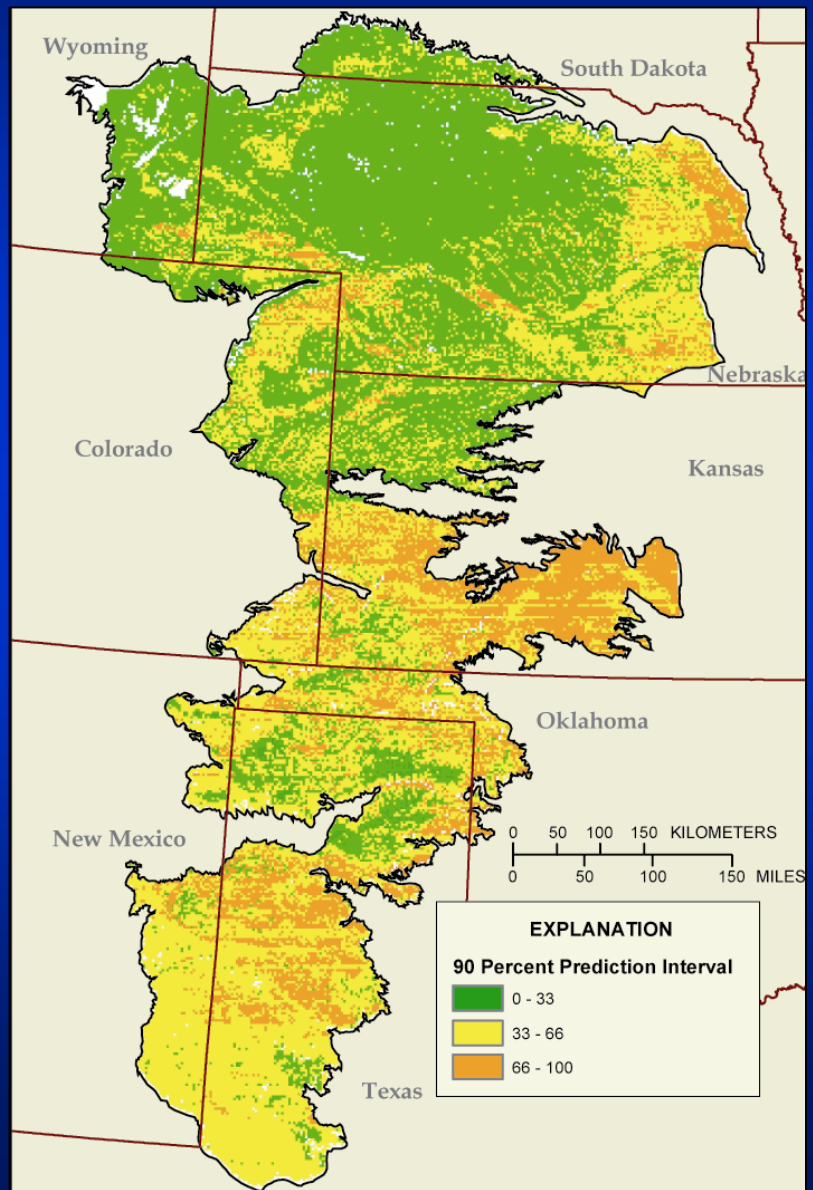
USGS

Uncertainty Estimation

Vulnerability Map



90 Percent Prediction Interval



USGS

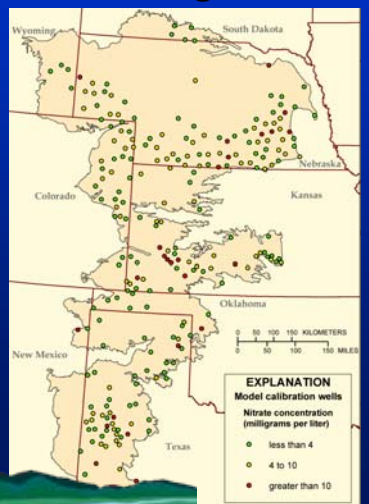
Conclusions

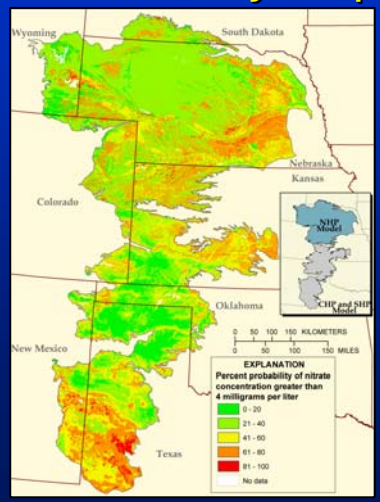
- Developing conceptual model(s) — important first step:
  - Define recently recharged (< 50 years) ground water
  - Delineate contributing areas using up-gradient 90-degree sectors in GIS

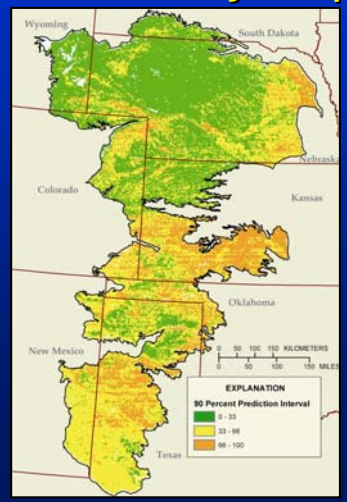
- Coupling logistic regression with GIS:
  - Effective method to extrapolation from monitoring network to aquifer.

- Vulnerability assessment — valuable ground-water management tool.
  - 20% of High Plains aquifer > 60% predicted probability of recently recharged ground water having nitrate > 4 mg/L (as N).

Monitoring Network → Vulnerability Map ↔ Uncertainty Map







USGS

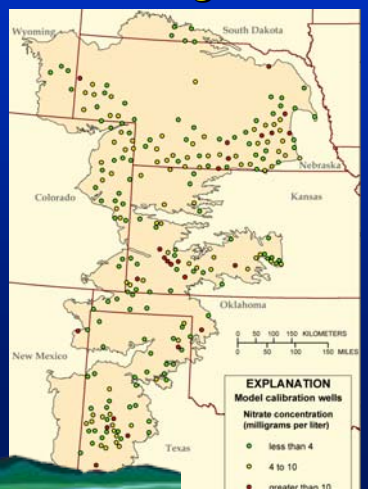
Conclusions – Continued

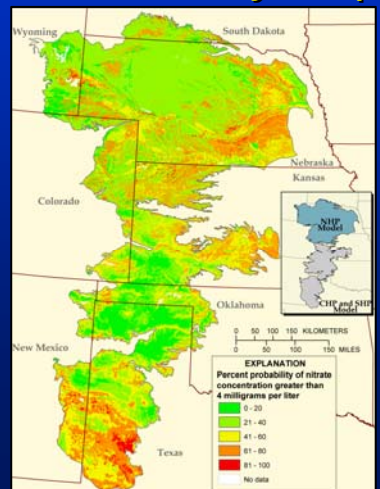
- Important factors — nitrate management in High Plains aquifer:
  1. Non-Irrigated agricultural land
  2. Irrigated agricultural land
  3. Clay content of the unsaturated zone
  4. Organic material of the soil
  5. Depth to water table (unsaturated zone thickness)

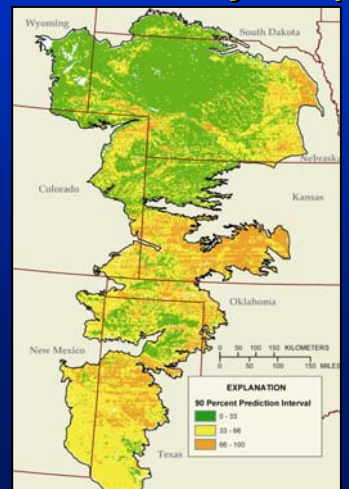
- Uncertainty analysis — practical tool for improved management.
  - Spatial estimates of model confidence across the aquifer.
  - Quantify how and where to improve future vulnerability models.
    - Additional monitoring wells and improved GIS data sets.

- Continuing need to quantify uncertainty associated with GIS data.

Monitoring Network → Vulnerability Map ↔ Uncertainty Map







USGS

http://co.water.usgs.gov/nawqa/hpgw/index.html

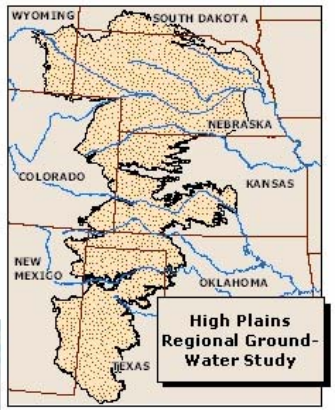
- Questions – jjgurdak@usgs.gov

USGS science for a changing world

National Water-Quality Assessment (NAWQA) Program

High Plains Regional Ground Water (HPGW) Study

As part of the the **NAtional Water-Quality Assessment (NAWQA) Program**, the USGS is evaluating ground-water quality in the **High Plains aquifer system**. The High Plains aquifer system underlies 175,000 square miles in parts of eight States (CO, KS, NE, NM, OK, SD, TX, and WY) (figure 1). Approximately 20 percent of the irrigated land in the United States is in the High Plains and about 30 percent of the ground water used for irrigation in the U.S. is pumped from the High Plains aquifer. **Irrigation withdrawals** in 2000 were **17 billion gallons per day**. In 2000, 1.9 million people were supplied by ground water from the High Plains aquifer with **total public-supply withdrawals of 315 million gallons per day**.



The quality of water in the High Plains aquifer generally is suitable for irrigation use but, in many places, the water does not meet U.S. Environmental Protection Agency drinking-water standards with respect to several dissolved constituents (dissolved solids/salinity, fluoride, chloride, and sulfate). Only sparsely scattered water-quality data (except Texas) are available for pesticides, volatile organic compounds, and trace metals in the High Plains aquifer system. Nutrient data are available, to a varying degree, across the aquifer.

Beginning in 1999 and continuing for a period of 6 years, the High Plains Regional Ground Water Study will intensively investigate the quality of ground-water resources within the study area. Investigations will begin in the Central High Plains and move to the Southern High Plains and Northern High Plains as the project progresses (figure 1). The first and most aerially extensive component of the intensive study phase is the "Occurrence and Distribution Assessment." The goal of this assessment is to characterize, in a nationally consistent manner, the broad-scale geographic variations of ground-water quality related to major contaminant sources and background conditions.

If you have any questions or comments related to the High Plains Regional Ground Water (HPGW) NAWQA, contact:
Breton Bruce, Project Manager
(bbruce@usgs.gov)

or write:
Breton Bruce, NAWQA Project Manager
U.S. Geological Survey
Box 25046, MS 415
Denver Federal Center
Denver, CO 80225

USGS