Utilizing GIS and Land Management Data to Assess Onsite Wastewater Facilities in Small Communities

Alan T. Brewer, M.S., R.E.H.S

Loudoun County, Virginia

About Loudoun County, VA

- Population : 272,000
- County Seat : Leesburg
- Land Area: 517 Square Miles
- Water and Sewer Utilities
  - Eastern Loudoun: Public
  - Western Loudoun : Private

First Objective

Create a risk assessment tool, based on existing data, for jurisdictions to utilize to predict future needs from failing and/or nonexistent onsite sewage dispersal systems.

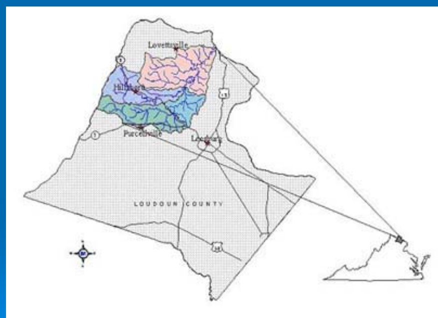
Traditional Community Wastewater Assessments

- Resource Intensive
  - Record Searches
  - Community Surveys
  - Site Evaluations
- Poor Community Participation
- Point in Time Assessment of Site Conditions

Second Objective

- Create a GIS map of the Catoctin Creek watershed in Loudoun County, Virginia, that illustrates high risk areas for potential pollution of surface waters due to inadequate onsite sewage systems.

Catoctin Creek Watershed



Catoctin Creek Watershed

- Approximately 59,000 Acres
- 6254 Parcels of land
- Impaired by Fecal Coliform
  - Livestock
  - Wildlife
  - Human

Onsite Sewage System Failures

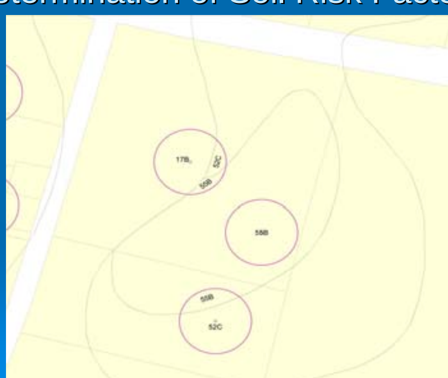
- Poor Soil and Site Conditions
- Small Lot Sizes
- Old Sewage Systems

* Each parcel in the watershed can be assigned a risk factor in each one of these categories

Soil and Site Conditions

- Risk factor assigned based on mapped soil potential for sewage system installation (Loudoun County's *Interpretive Guide to the Use of Soils*)
  - Good = 1
  - Fair = 2
  - Poor = 3
  - Very Poor = 4

Determination of Soil Risk Factor



Sewage System Age

- Onsite sewage systems in the watershed were divided into quartiles by age and assigned another risk factor.
  - 1st quartile =1 (less than 7 years old)
  - 2nd quartile =2 (7 to 17 years old)
  - 3rd quartile = 3 (18 to 31 years old)
  - 4th quartile = 4 (greater than 31 years old)

Parcel Size

- Onsite sewage systems in the watershed were divided into quartiles by parcel size and assigned another risk factor.
  - 1st quartile = 1 (greater than 10 acres)
  - 2nd quartile = 2 (3.33 to 10 acres)
  - 3rd quartile = 3 (.68 to 3.32 acres)
  - 4th quartile = 4 (less than .68 acres)

Default Assigned Values

- Vacant parcels assigned the lowest risk factor in all categories.
- Parcels with dwellings but no onsite sewage system mapped were assigned the highest risk factor in all categories.

Statistics

- The likelihood that soil conditions, sewage system age, and parcel size are all equally significant in causing system failure is extremely remote.
- Logistic regression performed to determine the significance of each category.

Logistic Regression

- Quantifies the association between a risk factor (i.e. soil) and an event (i.e. drainfield failure).
- Finds an equation that predicts a dependent variable (must be binary) from one or more independent variables.

Assumptions of Logistic Regression

- Cases are representative of a larger population.
- Each case is selected independently of others.
- Independent variables are independent (test for correlation).

Study Design Considerations

- There should be 5 to 10 cases for each independent variable (example: 3 variables (soil, age, size) times 5 events; sample size must be 15.
- Do not choose variables without a clear hypothesis.
- Utilize existing data where possible.
- Absolutely no more than 20 independent variables.

Null Hypothesis

- The independent variable (soil, age or parcel size) does not cause sewage dispersal system failure in the Catoctin Creek Watershed.
- Compared to a significance level of .05.

Logistic Regression

- Results of the most recent 100 onsite system evaluations in the watershed were examined (28 failing, 72 functioning)
- Each of the 100 parcels assigned a risk factor for the age, site conditions, parcel size
- Parcels assigned dependent variable value of 1 if failing and 0 if functioning
- Logistic Regression Analysis Performed

Results

Logistic Regression (Size, Age and Soil)

	B	S.E.	Df	P	Exp(B)
SIZE	.948	.328	1	.004	2.580
AGE	.824	.268	1	.002	2.280
SOIL	-.323	.261	1	.216	.724
Const	-4.5811	.097	1	.000	.010

* Reject null hypothesis for parcel size and system age.

Predicted Risk For Each Parcel

- Multiple regression formula:

$$Y = a + (b_1)(X_1) + (b_2)(X_2) \ldots (b_k)(X_k)$$

Where:

- Y = the predicted risk factor for the parcel
- a = regression constant
- b = the regression coefficient
- X = the independent variable risk factor

*Only include significant independent variables

Predicted Probability For Onsite Sewage System Failure

Logit Transformation Performed (only two possible values (failing or not failing) and not a linear relationship).

$$p = \frac{e^Y}{1 + e^Y}$$

Where:

p = transformed predicted risk factor or predicted probability

e = the inverse natural log

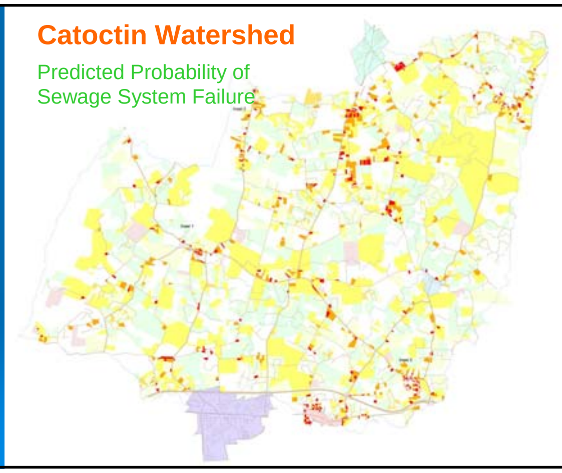
$Y = a + (b_1)(X_1) + (b_2)(X_2) \ldots (b_k)(X_k)$

Predicted Probability

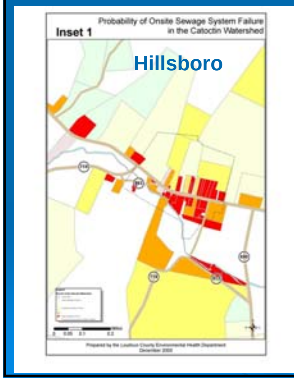
- Each parcel in the watershed is assigned its predicted probability value.
- Values ranged from .04 to .83 with values closer to 1 having systems more likely to fail.

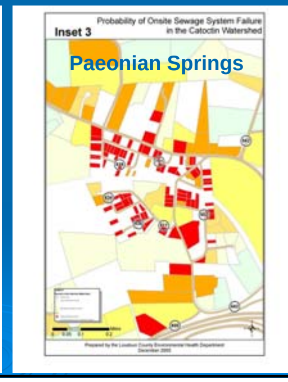
Catoctin Watershed

Predicted Probability of Sewage System Failure

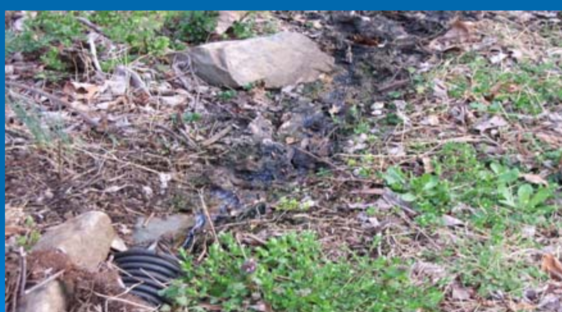


Predicted Probability of Sewage System Failure That Can Be Mapped





Straight Pipe, Hillsboro



System Failure, Mountain Road

No System, Mountain Road

QUESTIONS?