

An Objective GIS Screening Tool for Rating the Suitability of Land for Septic-based Development

John A. Welhan

Carol Moore

Staff Report S-12-1
August 2012

Idaho Geological Survey
Morrill Hall, Third Floor
University of Idaho
Moscow, Idaho
83844-3014

An Objective GIS Screening Tool for Rating the Suitability of Land for Septic-based Development

John A. Welhan

Carol Moore

*Staff reports present timely information for
public distribution. This publication may
not conform to the agency's standards.*

Staff Report S-12-1
August 2012

Idaho Geological Survey
Morrill Hall, Third Floor
University of Idaho
Moscow, Idaho
83844-3014

CONTENTS

Summary.....	1
1. Introduction.....	2
1.1. Statement of the problem.....	2
1.2. Terminology.....	3
2. Approach and Methodology.....	4
2.1. Geographic study area.....	4
2.2. Types of information utilized.....	4
2.3. Class ranking.....	8
2.4. Data preparation.....	8
2.4.1. Topographic slope.....	8
2.4.2. Soil drainage.....	8
2.4.3. Soil permeability.....	8
2.4.4. Septic density.....	10
2.4.5. Depth to ground water.....	11
2.4.6. Overburden thickness.....	11
2.4.7. Hydrogeologic influences.....	11
2.4.8. Water-quality (nitrate).....	14
2.5. Fatal conditions.....	15
2.6. Statistical calibration.....	15
2.7. Sensitivity analysis.....	16
3. Results and Discussion.....	16
3.1. Calibrated information layers.....	16
3.2. Relative layer weights.....	16
3.3. Calibrated septic suitability.....	21
3.4. Sensitivity analysis and final classification.....	22
3.4.1. Using calibrated information.....	22
3.4.1.1. Sensitivity analysis protocol.....	22
3.4.1.2. Setting an objective suitability threshold.....	23
3.4.2. Using uncalibrated information.....	25
3.5. Comparison of results.....	29
4. Conclusions and Recommendations.....	33
5. Acknowledgments.....	36
6. References Cited.....	36
Appendices	
A. Examples of DRASTIC-like Assessments in Montana and Wyoming.....	38
B. GIS Information Layers Compiled for This Study.....	43
C. Water-Quality Data for Lower Portneuf River Valley.....	45
D. Worked Example of the Statistical Calibration Process.....	48
E. ArcToolbox Script for Permutation / Sensitivity Analysis.....	67
F. An Objective Classification Procedure Based on a Sensitivity Analysis.....	74

TABLES

Table 1. Principal hydrogeological factors in a conventional DRASTIC assessment.....	3
Table 2. Information layers used in the assessment of septic suitability in the LPRV.....	7
Table 3. Final ordinal class ranges used to evaluate septic suitability in the LPRV.....	17
Table 4. Layer weights assigned in this study.....	18
Table 5. Ways of converting the nominal classes in the Hydrogeologic Influences layer.....	27
Table 6. Ordinal class ranks assigned to uncalibrated layers.....	28
Table 7. Layer weights assigned to uncalibrated information.....	29
Table 8. Comparison of nitrate-N concentrations in classified areas.....	33

ILLUSTRATIONS

Figure 1. Geographic extent of the study area.....	5
Figure 2. Topographic slope classes across the study area.....	9
Figure 3. Soil drainage, according to NRCS SSURGO.....	9
Figure 4. Relative soil permeability, estimated from SSURGO grain-size data.....	10
Figure 5. Septic density computed from a Gaussian kernel density function.....	12
Figure 6. Depth to water grouped in seven classes from <30 to 100 feet below surface.....	12
Figure 7. Overburden thickness as classified from surficial geologic maps.....	13
Figure 8. Geologic and hydrologic factors relevant to nitrate contamination in the LPRV.....	13
Figure 9. Availability of nitrate-N concentration data across the study area.....	14
Figure 10. Comparison of uncalibrated and calibrated depth-to-ground water rasters.....	18
Figure 11. Comparison of uncalibrated and calibrated septic density classifications.....	19
Figure 12. Uncalibrated and calibrated hydrogeologic influences.....	19
Figure 13. Uncalibrated and calibrated (soil drainage classes.....	20
Figure 14. Uncalibrated and calibrated soil permeability classes.....	20
Figure 15. Arbitrarily classified slope map and a binary mask created from it.....	21
Figure 16. Septic suitability ratings for the study area.....	22
Figure 17. Cell-wise means and standard deviations of the suitability rating map.....	24
Figure 18. Fraction of the study area classified as unsuitable.....	24
Figure 19. Areas with coefficients of variation at or above the 75th percentile.....	26
Figure 20. Final septic suitability map.....	26
Figure 21. A comparison of septic suitability ratings.....	30
Figure 22. Areas classified as Suitable, with nitrate-N concentrations overlain.....	31
Figure 23. Areas classified as Suitable and Uncertain.....	31
Figure 24. Enlarged view of southern aquifer showing nitrate-N concentrations.....	32
Figure 25. Availability of nitrate-N data in areas that were classified.....	32

An Objective GIS Screening Tool for Rating the Suitability of Land for Septic-based Development

John A. Welhan¹

Carol Moore²

SUMMARY

A rating method to evaluate the suitability of land for septic sewage disposal was developed using a DRASTIC-like formalism and tested with data from the lower Portneuf River valley (LPRV) in northern Bannock County, Idaho. The method uses readily available information, is easy to apply and is entirely objective, relying on statistical analysis rather than subjective judgment. A single Hydrogeologic Influences layer replaces three of the original DRASTIC layers (aquifer media, conductivity, and unsaturated zone impact) that are some of the most difficult to quantify in data-poor situations.

Statistical analysis of correlations between ground water nitrate concentrations and the information content in each layer overcomes one of the most serious deficiencies of the original DRASTIC method, and a statistically-based classification method eliminates subjectivity when designating areas that are suitable and unsuitable. The classification method relies on sensitivity analysis to identify areas in which the predicted ratings are least certain, information which is then used to optimize the delineation of three classes: areas that are suitable for septic-based development, areas that are unsuitable, and areas in which classification is least certain and in which more information will be required. The entire process was freed of subjective judgment to enhance technical defensibility and adoption in the planning and regulatory community.

¹ Idaho Geological Survey, Pocatello Satellite Office

² Idaho State University, Department of Geosciences

1. INTRODUCTION

This report encapsulates the findings of a study funded by IDEQ that was intended to: (i) estimate the suitability of land in northern Bannock County for septic sewage disposal; (ii) develop an objective and technically defensible method for doing so; (iii) that could be adapted and applied in any location in Idaho using readily available data.

A number of different methodologies have been used to assess ground-water vulnerability to contamination, all of which have strengths and weaknesses (Focazio and others, 2002). Physical process models and even the simplest nutrient-pathogen spreadsheet models (Howarth and others, 2002; Wicherski, 2006) require site-specific data and pertain only to that area for which they were constructed. Index-based methods such as the popular DRASTIC model (Aller and others, 1987) are simple to implement in a geographic information system (GIS) and provide a screening tool that is applicable over a wide geographic area. However, index-based methods have important limitations. First, each "layer" of information utilized in the ranking must be classified (often arbitrarily); the resulting classes are assigned numerical ranks that indicate how each informs the overall vulnerability assessment; finally, weighting factors are applied to each layer, representing their relative importance in the overall vulnerability assessment. The weighted sum of the layers represents the relative vulnerability rating for a particular contaminant, as assessed over the geographic area in which the information layers were defined:

$$\text{Rating} = w_D D + w_R R + w_A A + w_S S + w_T T + w_I I + w_C C$$

where capital letters represent the various layers in the DRASTIC model and the w_i represent each layer's weight. **Table 1** summarizes the layers and recommended layer weights that comprise a conventional DRASTIC assessment. Each letter of the DRASTIC acronym represents one of seven hydrogeologic factors that affect the mobility of a contaminant from the land surface to the water table and/or its accumulation in an aquifer.

Index-based ranking methods, by their nature, provide only a relative measure of vulnerability. A DRASTIC assessment cannot "replace site specific investigations or preclude the consideration of [other factors] which may be important" (Aller and others, 1987). In practice, an index-based rating is a useful screening tool that should be part of a multifaceted decision-making framework that utilizes multiple tools (Focazio and others, 2002).

1.1 STATEMENT OF THE PROBLEM

The goal was to develop a simple, robust modification of the DRASTIC index-based method and rank the relative suitability of land for septic-based development in an urban / suburban setting. This study focused on the lower Portneuf River valley (LPRV), but the methodology was designed to be applicable in any regulatory or planning jurisdiction.

The methodology must be technically sound so that ground-water management policy based on it will be technically defensible. However, it is often the case that all of the information required for a full DRASTIC assessment is unavailable. In particular, information on factors such as aquifer media (A), the unsaturated zone (I) or hydraulic conductivity (C) may be hard to come by. Planners and regulators may be unaware of such data; it may not exist; or the

Table 1. Principal hydrogeological factors affecting contaminant mobility in a conventional DRASTIC assessment.

<u>Hydrogeologic Factor</u>	<u>Physical Relevance</u>	<u>Suggested Weight*</u>
(D) depth to water	vertical distance between the aquifer and the surface promotes attenuation	5
(R) effective recharge	recharge from rain/snow, canals, etc. accelerates vertical movement of contaminants from the surface	4
(A) aquifer characteristics	an aquifer's vulnerability to contamination is greater if it is shallow, unconfined, and/or comprised of permeable materials	3
(S) soil characteristics	vertical migration of contaminants is greatest through permeable or thin soil	2
(T) topography (slope)	infiltration and hence recharge is greater on flat ground and slopes less than ca. 5 %	1
(I) unsaturated zone impact	an aquifer's vulnerability to contamination is greater if it is overlain by permeable rather than impermeable material	5
(C) hydraulic conductivity	contamination spreads more readily in more permeable aquifers	3

* Table 2 in Aller and others (1987)

jurisdiction may lack the technical resources to compile and convert it to a format that can be utilized. This study emphasized the use of GIS information that is commonly available to planners and regulators or that could be readily converted to GIS format. The most important physical factors pertinent to septic contamination were prioritized, and subjectivity arising from arbitrary classification schemes was eliminated from the analysis process.

To preserve as much of the expert knowledge framework that was encapsulated within the DRASTIC model, the original inter-layer weights assigned by Aller and others (1987) were used in this study, adjusted for differences in layer classifications.

1.2. TERMINOLOGY

In this report, the terms "layer" and "information layer" refer to GIS raster maps that are synonymous with Aller and others' (1987) "mappable factors" or physical characteristics that affect ground water vulnerability. The categories within each layer are termed "classes" and the ordinal values assigned to each class are termed "class ranks" or simply "ranks" (Aller and others called these "factor ranges" and "factor ratings," respectively). The result of combining class

ranks from multiple layers is to create a map of what Aller and others called DRASTIC "index" values, herein termed "septic suitability ratings."

2. APPROACH AND METHODOLOGY

GIS technology is widely utilized in resource management and is growing rapidly in the field of urban and rural planning. To promote adoption of the technology by planners, this report provides guidance on developing GIS data sets for any geographic area, using the LPRV as an example. All GIS data sets and manipulations described in this report are specific to ESRI's ArcMap software.

2.1. GEOGRAPHIC STUDY AREA

This study was funded by the Idaho Department of Environmental Quality (IDEQ) and was very much purpose-driven. Specific constraints were imposed on the study at its inception: (i) the need for a defensible planning tool; (ii) a methodology that would be simple enough to adopt by city and county planners; and (iii) the results would be useful in ongoing efforts by IDEQ, the Idaho Geological Survey (IGS) and other state and federal agencies to develop a ground-water protection overlay for Bannock County as part of its comprehensive plan.

The study area was selected in consultation with IDEQ's Pocatello Regional Office on the basis of data availability, immediate need, and relevant geographic extent. The study area shown in **Figure 1** includes a large part of the LPRV watershed, including the cities of Pocatello and Chubbuck. The valley has almost no irrigated agriculture and the overwhelming source of ground water nitrate contamination has been shown to be septic leachate (Meehan, 2005).

2.2. TYPES OF INFORMATION UTILIZED

Of the original layers incorporated in DRASTIC, some are more relevant than others in the assessment of septic suitability. First, soil properties and the unsaturated zone are important in determining septic leachate mobility and how effectively it is attenuated, filtered and transformed prior to reaching the water table. Fractured rock provides very little attenuation and filtering so its presence at the surface (or its depth below surface) is a critical measure of septic suitability.

As has been observed in almost all ground water vulnerability studies, depth to ground water and effective rate of local recharge to the water table are very important indicators of pollution vulnerability. In southeast Idaho, where seasonal evapotranspiration exceeds annual precipitation by a factor of 2:1 or more, recharge from precipitation is negligible at all but the highest elevations. Artificial recharge sources like canals can be important localized sources of artificial recharge, but canals are nonexistent in all but the northernmost LPRV.

Artificial recharge due to septic drain field operation is a source of localized recharge that has been overlooked in many DRASTIC assessments. In particular, the cumulative impact of

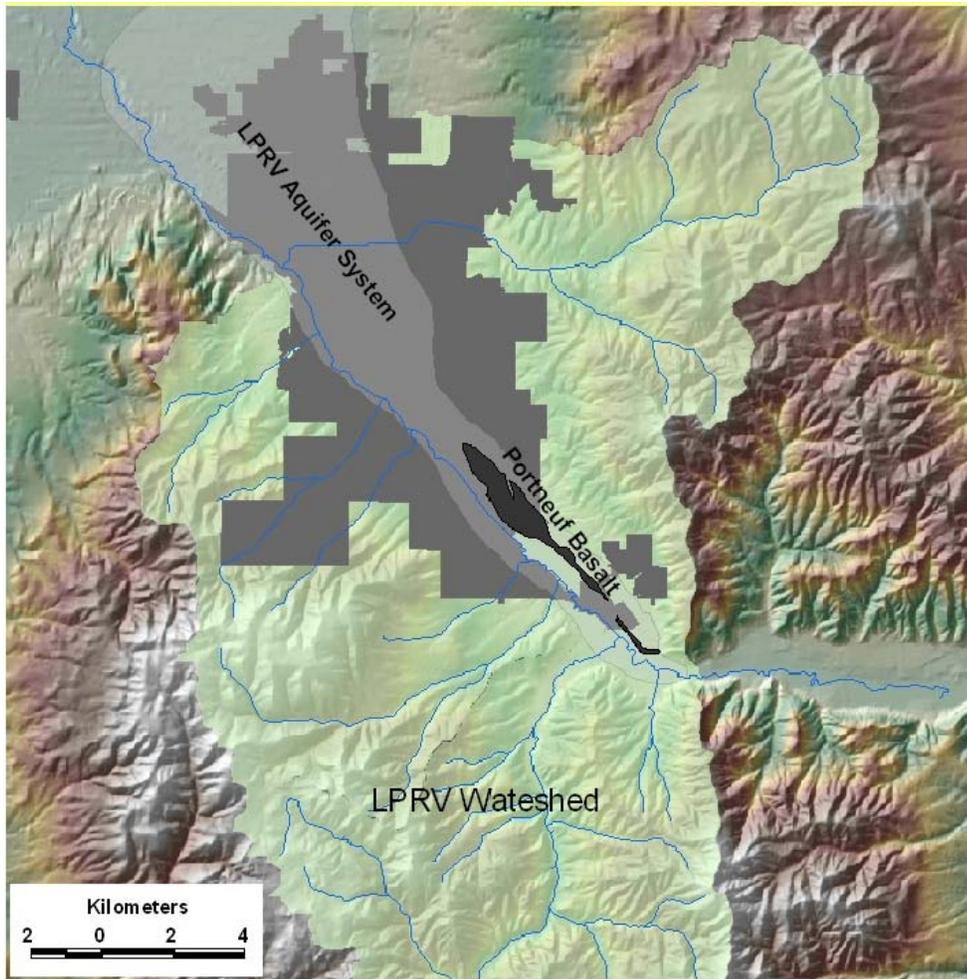


Figure 1. Geographic extent of the study area, showing lower Portneuf River valley watershed (light green), LPRV aquifer system (light grey), and the combined municipal boundary of Pocatello and Chuibuck (dark grey).

multiple sources of septic recharge needs to be considered (Meehan, 2005). Therefore, the spatial density of existing septic drain fields is very important factor that directly affects septic suitability.

The cumulative nature of septic impacts on ground-water quality also suggests that aquifer characteristics are an important consideration, although not necessarily in the manner that Aller and others (1987) devised. Factors such as the degree of confinement and the rate at which ground water moves through an aquifer and dilutes contaminants may be more important to consider than its hydraulic conductivity or the material that comprises the aquifer.

Based on the above considerations, as well as the fact that relevant subsurface data is lacking in many areas of Idaho, it is suggested that the following factors comprise a modified DRASTIC assessment of septic suitability:

- (i) average depth to ground water (derived from IDWR's on-line database);

- (ii) septic drain field locations (estimated from county land parcel maps);
- (iii) soil characteristics like relative permeability, drainage, or others (derived from NRCS's SSURGO database); and
- (iv) areas of hydrogeologically relevant characteristics that can affect contaminant mobility to, and accumulation within, an aquifer such as:
 - thin soil, shallow or exposed fractured bedrock
 - confined vs. unconfined aquifer conditions
 - extremely slow ground water movement, ineffective dilution of contaminants
 - unusual aquifer conditions (noxious odors, poor water quality, iron staining)

Typically, the above information can be gleaned from surface geologic maps, well drillers and drilling reports, and local residents' observational knowledge.

Topographic slope was not considered to be an essential layer in rating septic suitability because, up to a point, slopes are engineered to accommodate drain fields. Instead, slope information was considered as one example of how various "fatal" criteria could be incorporated in a septic suitability rating: for example, areas known to be affected by historic contamination that originates at the surface (e.g., from old CAFO facilities, abandoned septic systems, and other unusual cases of polluted ground water). All such information should be considered "anecdotally relevant" to the presence of conditions that may be unsuitable for septic placement.

To develop a methodology that can be widely applied, three criteria were considered in selecting information: the data should be (i) readily accessible; (ii) already in or easily converted to GIS format; and (iii) relevant to the contaminant vulnerability issue being assessed (septic suitability in this case). **Table 2** provides a summary of the information that was utilized in this study and the ranks that were assigned to the categories within each layer.

The layers summarized in Table 2 represent three types of information: numerical data (e.g., percent slope; septic systems per square kilometer); ranked categorical (or "ordinal") relationships (e.g., low-medium-high permeability or depth to water); and unranked categorical information known as nominal data³. Examples of the latter include soil types, geologic units, land use classes and zoning classifications. Nominal information is essential in all types of decision-making. However, because the classes cannot be ranked in any meaningful order, the inclusion of nominal data in an index-based model is a highly subjective undertaking, entirely predicated on the order in which the classes are arranged and the class ranks that are assigned.

The "Hydrogeologic Influences" layer shown in Table 2 is an example of nominal information that, in the LPRV, was arbitrarily classified into seven classes (see Section 2.3.7):

- 0 areas with very low ground-water flow rates (1 class);
- 1 confined aquifer conditions (1 class);
- 2 presence or absence of fractured rock (1 class);
- 3 depth of overburden (3 classes); and
- 4 historically contaminated areas (1 class).

The only way such information can be objectively incorporated in a DRASTIC-like model is to transform its categories via a statistical calibration process, such as Rupert (2001) did to incorporate land use as an information layer. This is a type of data transformation that produces an objective relative ranking (ordinal classification) of nominal classes. Land use was not considered to be an appropriate criterion in this study because it can change over time and

³categorical information that lacks a natural ordering of its classes

because it may not be objective or relevant in many urban / suburban situations. For example, different types of urban and suburban land use can have very different impacts on ground-water quality depending on whether particular areas are sewered or not, on whether land use zoning has recently changed, or what type of residential irrigation practices and fertilizer usage prevail. In this study, the spatial density of existing septic systems was deemed a far more objective and relevant land-use criterion.

Table 2. Information layers and arbitrarily assigned class ranks that were used in the assessment of septic suitability in the LPRV

<u>Information Layer</u>	<u>Assigned Ranks¹</u>	<u>Sources of Information</u>
Overburden depth	1 - 3	Idaho Geological Survey (IGS) 1:24000 geologic maps http://www.idahogeology.org/
Depth to water	0 - 6	Idaho Dept. of Water Resources (IDWR) well drillers' database http://www.idwr.idaho.gov/
Septic recharge	1 - 3	Bannock County parcel database and municipal sewer system
Soil drainage	0 - 6	NRCS SSURGO soils database http://soils.usda.gov/
Soil permeability	1 - 3	
Hydrogeologic influences	1 - 7 ²	Compiled from various sources (see Section 2.4.7)
Nitrate water quality	n.a.	Compiled from various sources (see Section 2.4.9)
Slope	1 - 7	USGS digital elevation models http://www.geocommunity.com/

¹ values assigned to the categories within each layer; see Section 2.4

² the order of nominal class ranks is completely arbitrary; see text, and Section 2.4.7

2.3. CLASS RANKING

The DRASTIC method is amenable to either risk or suitability assessment because risk and suitability are complimentary measures. Thus, a DRASTIC index rating can represent either relative risk or relative suitability depending on the sense in which ranks are assigned to the data categories within a layer. In this study, the concept of septic suitability was deemed simpler to understand and convey to users than septic risk. Therefore, all layers were assigned class ranks that range from low to high, where the lowest rank represents the least suitable condition (highest contamination risk) and the highest rank, the most suitable condition (lowest contamination risk). For example, depth to ground water in the LPRV ranges from 19 to more than 100 feet, a range that was classified into seven arbitrary classes with the deepest class assigned a rank of 0 and the shallowest, a rank of 6.

2.4. DATA PREPARATION

Considerable effort was expended on acquiring appropriate information, vetting it for consistency and accuracy, and converting it to a GIS-compatible format. Data were represented in ESRI raster format and projected into Idaho Transverse Mercator NAD 1983 coordinates; all rasters were created as floating type with a cell size of approximately 100 x 100 meters. All GIS data developed in this study are summarized in **Appendix C**.

2.4.1. Topographic slope

Slope was calculated from a 30-meter digital elevation model (DEM) and arbitrarily classified into seven categories. **Figure 2** shows the slope map that was produced in this manner.

2.4.2. Soil drainage

Soil drainage information was derived from the NRCS SSURGO database via a spatial join of geographic location information ("soilmu_a_ID711.shp") and tabular data on soil properties ("muaggat.dbf"). The resulting shapefile was converted into a raster whose attribute table was modified to include a field with drainage classes ranked from 0 to 6 (most to least well drained, respectively, as classified in the SSURGO database). These drainage classes were used as a proxy for the soils' ability to attenuate septic drain field leachate. **Figure 3** summarizes this layer's geographic coverage and the spatial variability of its classes.

2.4.3. Soil permeability

The soil permeability classes in the study area were inadequate for rating septic suitability because only two classes exist, with more than 90 percent of the study area in one class. To portray more meaningful geographic variations, soil grain size data in the SSURGO database⁴ were extracted and converted to four rasters showing each grain size's spatial variability. ArcMap's Raster Calculator was used to identify raster cells having grain sizes that bracket the sieve size that passes the finest 10% soil fraction. This d_{10} grain size was found to lie between a grain size of 0 mm (D_1) and 0.074 mm (D_2). Raster cells where the 0.074 mm grain

⁴weight percent of material passing sieve sizes 4, 10, 40 and 200 (4.7, 1.65, 0.42, 0.074 mm)

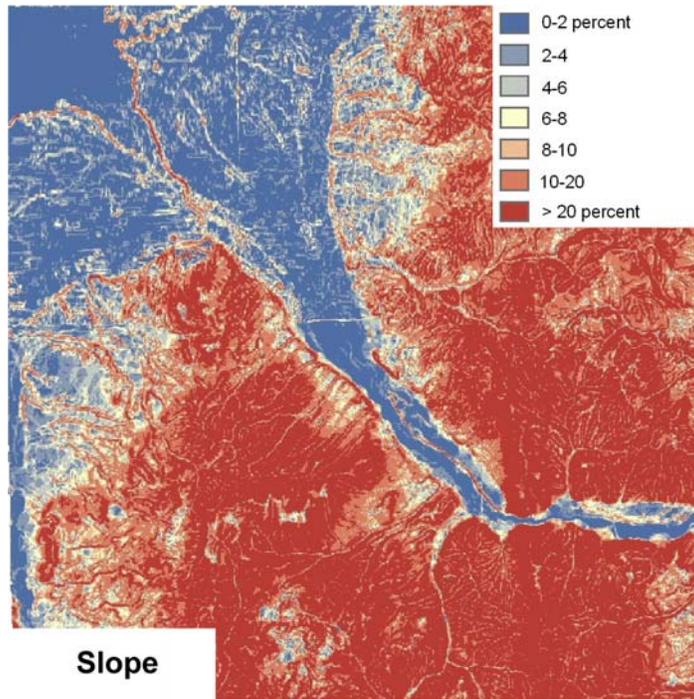


Figure 2. Topographic slope classes, in percent, across the study area.

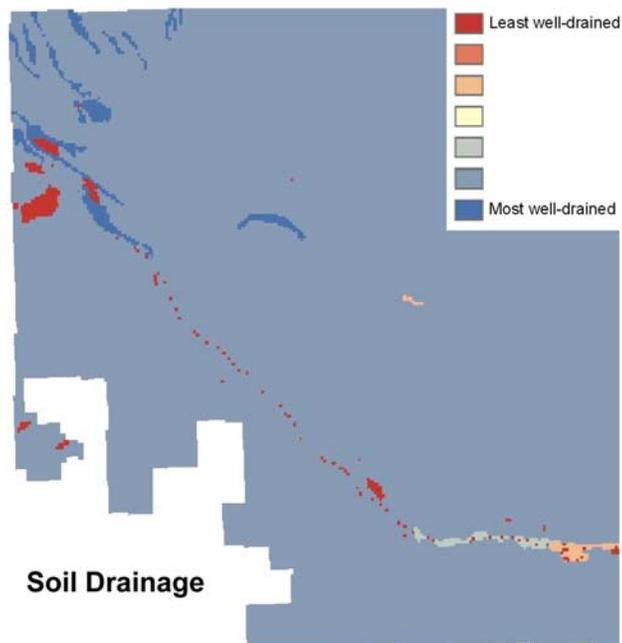


Figure 3. Soil drainage, according to NRCS SSURGO classified from least to most well-drained.

size exceeds 10 percent were identified, and slopes for the local grain size curves at each location were estimated using the relationship

$$\sigma = (P_2 - P_1) / (D_2 - D_1)$$

where P_2 is the 0.074 mm raster value, $P_1 = 0$, $D_2 = 0.074$ mm, and $D_1 = 0$ mm. The corresponding d_{10} value was then estimated with the expression $d_{10} = [(10 - P_1) + (D_1 * \sigma)] / \sigma$ and a relative permeability, K_H , was estimated from the d_{10} value using a modified Hazen formula ($K_H = d_{10}^2$; Fetter, 1994). The resulting range of relative K_H values (0.00006 - 0.0006) was then graphed and classified into three groups, the highest 20 percent of which was assigned an ordinal rank of 1 (least suitable); the lowest 20 percent, a rank of 2; and the intermediate range, a rank of 3 (most suitable). **Figure 4** summarizes the raster map and the spatial distribution of its classes.

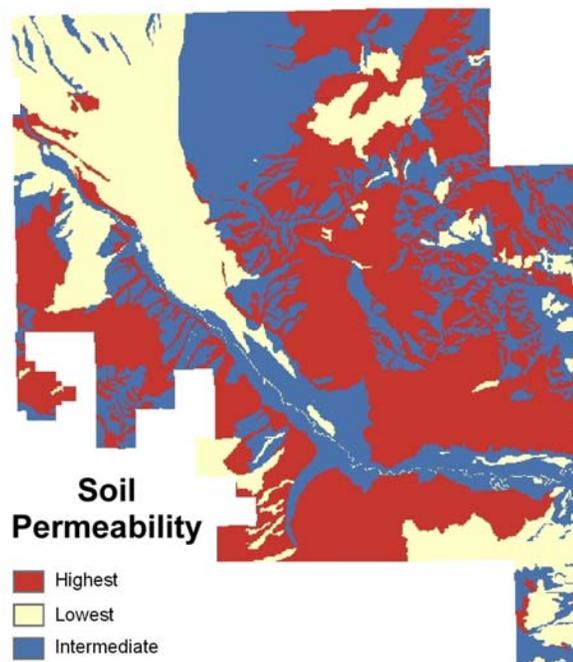


Figure 4. Relative soil permeability, estimated from SSURGO grain-size data.

2.4.4. Septic density

In 2002, IDEQ created a map of north Bannock County's septic-developed lands based on Bannock County's land parcel database (M. Byrd, written comm., 2002). The process that was developed could be applied to any county's parcel database: An addressed parcel more than 200 feet from a municipal sewer line was assumed to be serviced by a private septic system; the centroid of that land parcel represents that drain field's approximate location. The map portrays more than 3000 septic drain field locations and reflects the geographic extent of septic-based development in the LPRV as of 2002. It is still considered a good approximation of septic usage because septic-based development in Bannock County declined sharply from an average annual

growth rate of more than 50 permits per year prior to 2002 to less than a dozen per year since (S. Ernst, pers. comm., 2010). The septic location map provides an objective measure of land use that is directly relevant to nitrate contamination from septic sources. It provides a relative measure of the intensity of localized artificial recharge that transports nitrate to the water table.

ArcToolbox's kernel density function (500 meter sampling radius) was used to convert the shapefile of septic locations to a spatial density raster (number of septic systems per square kilometer). The raster was classified into three classes: no septics; fewer than ten times the mean spatial density; and up to the maximum spatial density. Ordinal ranks were assigned to these classes, with 1 corresponding to the highest septic density (least suitable for future septic development) and 3 corresponding to the lowest density (most suitable). **Figure 5** summarizes the layer's geographic coverage and its classes.

2.4.5. Depth to ground water

Recent work by the IGS (Welhan, in prep.) on the eastern Snake River Plain has shown that depth-to-water information in IDWR's on-line database of permitted water wells ('wells.shp' data file, <http://www.idwr.idaho.gov/ftp/gisdata/>) can be used to estimate regional depth-to-water maps relatively accurately. A map produced in 2002 for a previous project in the LPRV was used in this study. Depths to the seasonal high water table vary from 19 to more than 100 feet below surface and were arbitrarily classified into seven classes ranked from 0 to 6 (<30, 30-42.5, 42.5-53.5, 53.5-65, 65-75, 75-85, >85 ft). **Figure 6** depicts the resulting raster and its classes.

2.4.6. Overburden thickness

Basic geologic information for use in Section 2.4.8 was obtained from the IGS. GIS versions of three geologic maps were spatially joined and classified into three ordinal categories:

- 0 depth to bedrock of zero (where bedrock outcrops at the surface);
- 1 depth to bedrock of between 0 and 100 feet (where unconsolidated materials were mapped at the surface everywhere except over the valley aquifer); and
- 2 depth to bedrock in excess of 100 feet (beneath the valley aquifer).

Justification for the latter two classes was obtained from an examination of drillers' logs. The three classes were assigned ordinal ranks of 1 (no overburden, lowest septic suitability) to 3 (thickest overburden, most suitable). **Figure 7** shows the resulting raster map and its classes.

2.4.7. Hydrogeologic influences

In order to maintain conformity with the original DRASTIC approach and to simplify the process of creating information layers in situations where subsurface information is sparse, three of the original DRASTIC layers were combined into a single layer of hydrogeologically relevant information. As discussed in Section 2.2, areas that correspond to different overburden thicknesses, confined aquifer conditions, slow ground-water movement, and known historic contamination were represented on a single map and segregated into seven arbitrary classes. Relevant hydrogeologic information was gleaned from prior work (Welhan and others, 1996; Welhan, 2006), anecdotal information from residents, examination of well log information, and hydrogeologic inference. To simplify the synthesis of such information and make the DRASTIC method easier to apply in areas of limited data availability, this layer's classes are descriptive (nominal) rather than quantitative (ordinal). **Figure 8** shows the resulting raster. Because the classes represent a nominal information scale, their order is arbitrary.

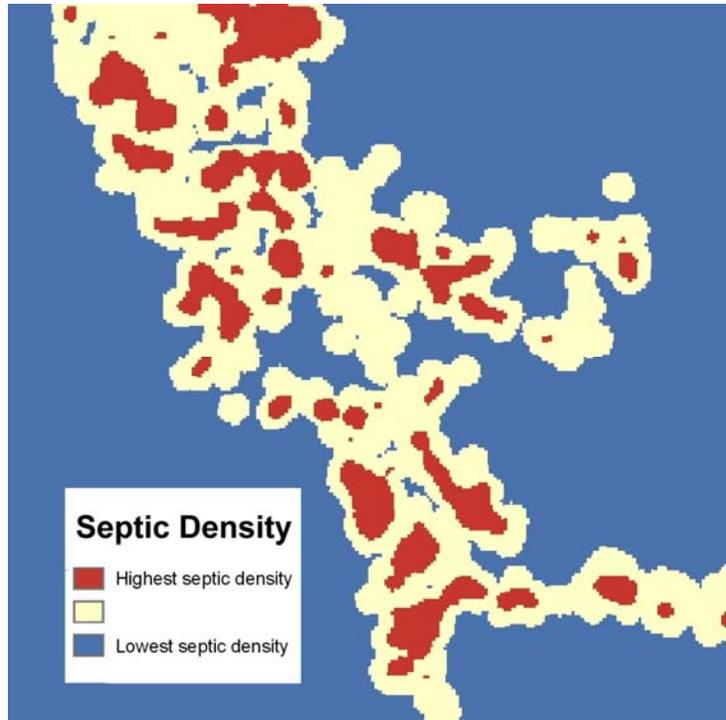


Figure 5. Septic density computed from a Gaussian kernel density function.

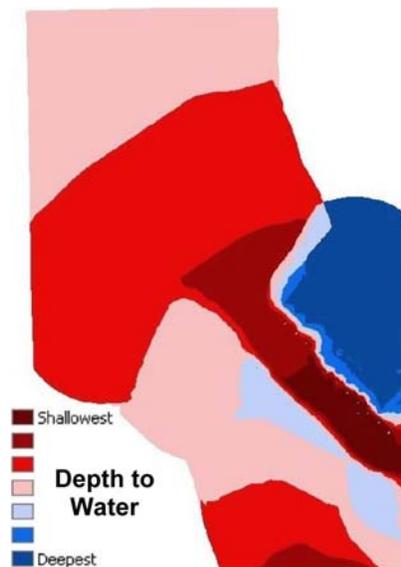


Figure 6. Depth to water grouped in seven classes from <30 to 100 feet below surface. The spatial extent of this layer is much smaller than other information layers due to the limited spatial extent of water well data in the LPRV.

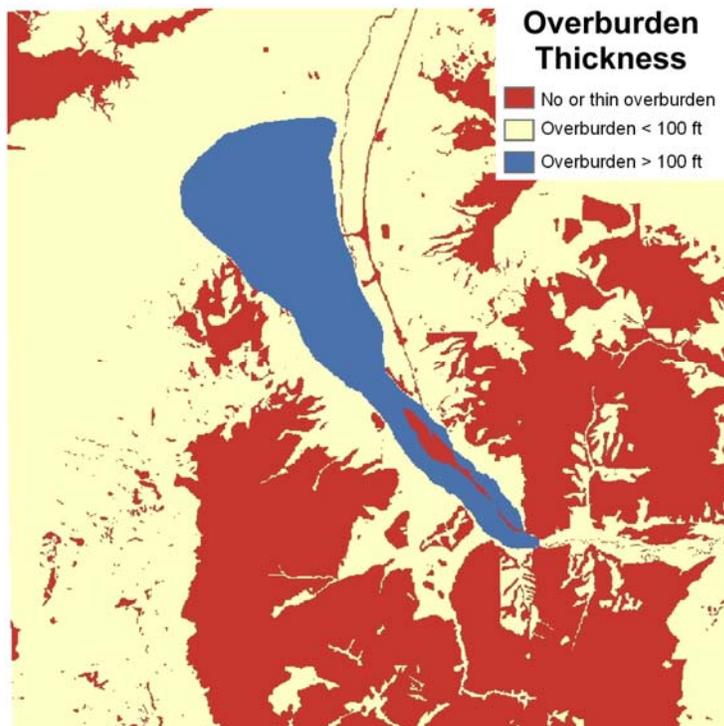


Figure 7. Overburden thickness as classified from surficial geologic maps.

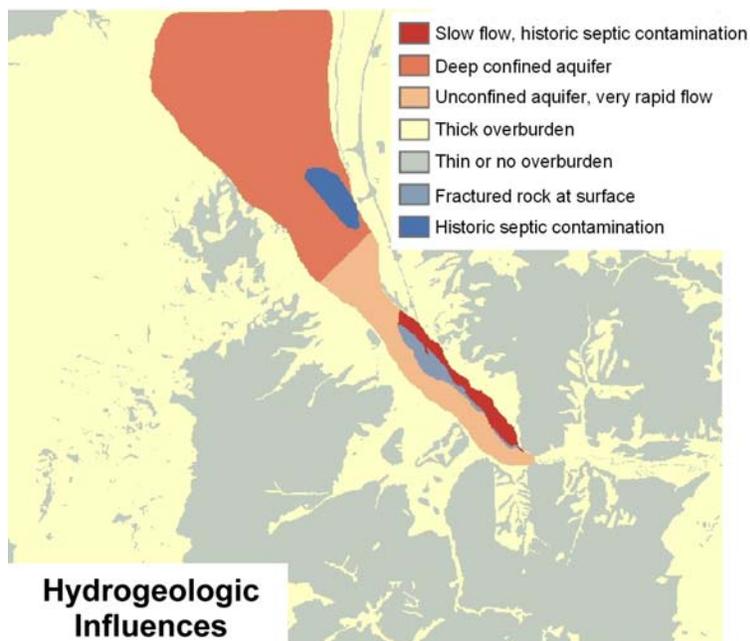


Figure 8. Geologic and hydrologic factors relevant to nitrate contamination in the LPRV. Areas so mapped can represent a mixture of quantitative and descriptive information or only descriptive information. Depending on the type of information that is available.

2.4.8. Water-quality (nitrate)

Water quality data were compiled from five sources: (i) IDEQ regulatory sampling programs; (ii) the USGS / IDWR statewide monitoring well network; (iii) City of Pocatello and Chubbuck municipal well sampling; (iv) water-quality sampling conducted for Bannock County; and (v) past ISU graduate thesis research. All available data up to and including 2005 were checked for consistency. Outliers were flagged or deleted as necessary and the information was compiled into a single attribute table with geographic coordinates, data sources, sampling, and editing comments. Because nitrate levels have remained fairly constant over time and information in many areas is sparse, the nitrate data were not segregated by time but considered in the aggregate. **Figure 9** summarizes the spatial availability of nitrate data and its variability. The attribute table is described in **Appendix C**.

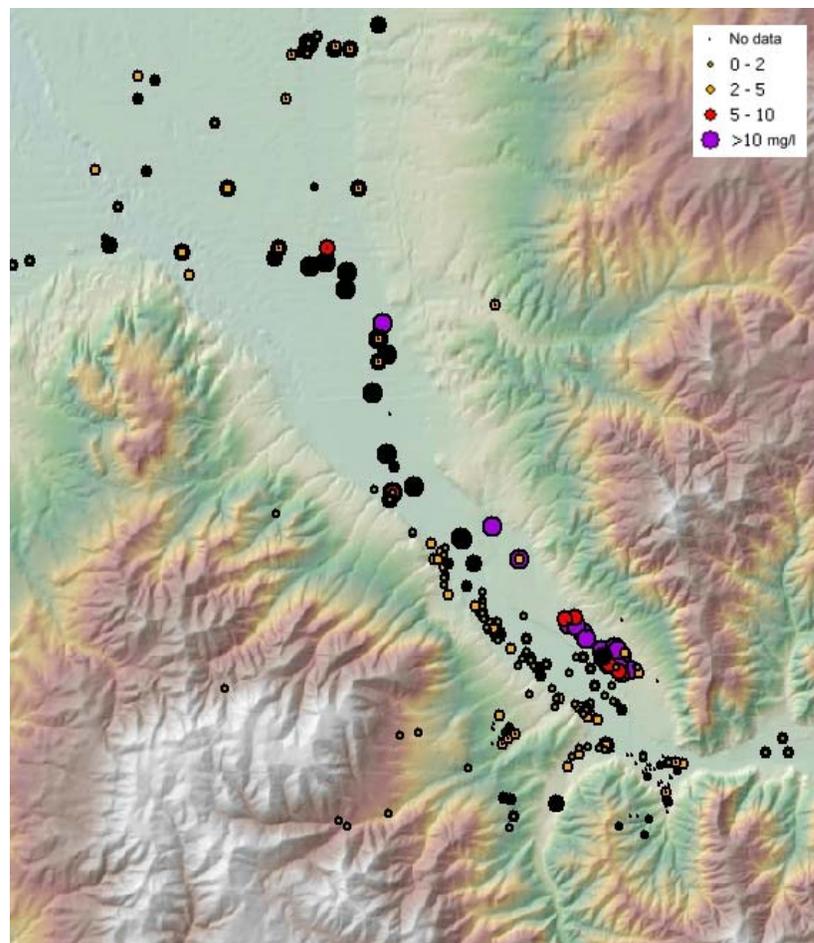


Figure 9. Availability of nitrate-N concentration data (mg/l) across the study area.

2.5. FATAL CONDITIONS

The DRASTIC model uses a weighted sum to represent the relative influence of each information layer on ground water vulnerability. The relative risk or suitability rating, R , is obtained by multiplying each DRASTIC layer (D, R, \dots, C) with an appropriate weight, w_i :

$$R = w_D D + w_R R + \dots + w_C C \quad \text{Eq'n (1)}$$

From a planning perspective, it might be desirable to include nonphysical considerations in Equation 1 (such as economic factors or management boundaries) as well as physical criteria not directly related to nitrate mobility but which nonetheless inform on septic siting suitability (e.g., steep slopes, sensitive or protected lands). To do so, however, risks conflating technically defensible decision criteria with other factors whose justification may be political or regulatory.

A more transparent (and defensible) approach would be to additively combine only physical factors in a DRASTIC-like rating and to incorporate other factors via separate binary masks that reflect "fatal" siting conditions. For example, areas having slopes greater than 20 percent might be excluded from consideration regardless of physical suitability because of building codes, development ordinances, or other policies that restrict development on steep slopes. Such a condition could be easily incorporated into a septic suitability rating by classifying the slope raster into a binary mask, S , that contains only two classes, $[0, 1]$, where '0' represents areas with slopes >20 percent (unsuitable for development). Several binary masks, M_i , could be combined with the DRASTIC rating in a single raster via a multiplicative operation:

$$R = (M_1 * M_2 * M_3) * (w_D D + w_R R + \dots + w_C C) \quad \text{Eq'n (2)}$$

In this example, areas with slopes steeper than 20 percent receive a rating of zero regardless of their septic suitability. Other fatal criteria could be incorporated in this manner including restricted wellhead protection zones, flood plains, sensitive recharge areas, and protected lands.

2.6. STATISTICAL CALIBRATION

The purpose of statistical calibration is to objectively assign ordinal ranks to a data layer by correlating them with ground water contamination (nitrate concentrations), including the objective ordering of nominal classes so they can be assigned ordinal ranks. The process involves these steps (**Appendix D** provides a detailed worked example):

- 1) segregate the nitrate data into n groups corresponding to the sampling locations that fall within the geographic boundaries of an information layer's n classes;
- 2) perform a Wilcoxon signed-rank statistical test⁵ on all unique pairs of grouped nitrate data to compare their medians at a specified confidence level;
- 3) if the medians of any pair of groups are statistically indistinguishable, then their nitrate data are regrouped into a single data set;
- 4) repeat steps (2) and (3) until the medians of all data groupings are statistically different;
- 5) assign a rank of 1 to the class(es) having the highest median nitrate and incrementally higher ranks to classes with progressively higher concentrations;
- 6) repeat steps (1) through (5) for all information layers.

⁵available in any basic statistical software package such as Minitab, S-Plus, or SAS

2.7. SENSITIVITY ANALYSIS

Besides using statistical calibration to rank classes within a layer, Rupert (2001) used it to adjust layer weights. This study did not do so because the process introduces more a level of subjectivity that is contrary to the goals of this project. Instead, a simpler procedure was devised: Layer weights were randomly varied to determine how much the septic suitability ratings varied in response; from these results, specific geographic areas were identified where the suitability ratings have the greatest uncertainty (are most sensitive to the choice of layer weights).

To automate this process, an ArcMap script was created to generate a large number of random permutations for any specified weighting scheme. Cell-wise statistics (means and standard deviations) calculated from the ensemble of permutations were used to create a map of the Coefficient of Variation ($CV = \text{standard deviation} / \text{mean}$). Suitability ratings that are spatially associated with the highest CV values indicate areas in which the predicted ratings are the least certain. The script is documented in **Appendix E** and included on CD-ROM together with instructions on how to configure, run and interpret its output.

3. RESULTS AND DISCUSSION

3.1. CALIBRATED INFORMATION LAYERS

Table 3 summarizes the number of class ranks in each information layer following the statistical calibration process on the original, arbitrarily assigned classes. **Figures 10 to 14** provide a side-by-side comparison of the calibrated and uncalibrated rasters and indicate where ordinal and nominal classes were combined and/or re-ranked. Note that regardless of how class ranks were originally assigned (Table 2), statistical calibration standardizes the assigned ranks to a common minimum value ('1') that increases with increasing septic suitability.

Figure 15 shows a binary slope mask, classified at a 20 percent slope threshold, that will be used as an example to demonstrate how policy-based information could be incorporated into the rating analysis.

3.2. RELATIVE LAYER WEIGHTS

The relative weights assigned to individual layers in the DRASTIC rating scheme can have a "drastic" effect on the suitability rating: Different relative weights will produce wildly different suitability ratings. Layer weights in this study were chosen to conform to those recommended by Aller and others (1987) in order to preserve the technical rationale and expert judgment that went into DRASTIC's relative weight assignments. In order to compensate for differences in the numbers of classes within each layer, DRASTIC weights were adjusted so that each layer's contribution to the suitability rating would be comparable to the corresponding layer's relative contribution in the original DRASTIC model.

The adjusted weight was calculated by constraining the average relative contribution of each layer to the overall rating to be the same as the corresponding layer's average relative

Table 3. Class ranks assigned to lower Portneuf River valley information layers following statistical calibration are all ordinal (1 = least suitable).

<u>Information Layers</u>	<u>Range of calibrated class ranks</u>
<i>Vulnerability Criteria</i>	
Depth to water	1 - 2
Septic recharge	1 - 3
Soil drainage	1 - 5
Soil permeability	1 - 3
Hydrogeologic influences	1 - 3
<i>Fatal Policy Criteria</i> (example)	
Slope	1 - 4

contribution in a DRASTIC rating. The midpoint of a layer's range of class ranks was assumed to represent its average relative contribution to the overall rating. The product of each layer's adjusted weight and the midpoint of the class ranks in Table 3 was then set equal to the product of the corresponding DRASTIC layer weight and midpoint rank to solve for the adjusted weight.

For example, from Table 3 the range of class ranks for Septic Recharge is [1, 3]; the midpoint is 2. The midpoint of DRASTIC's class rank for the same layer is 6⁶ and the layer weight is 4 (Table 5; Aller and others, 1987). Therefore, the adjusted layer weight assigned to Septic Recharge is:

$$4 \times 6 / 2 = 12.$$

Table 4 summarizes the layers and weights that were used to rate septic suitability in the LPRV. Because the Hydrogeologic Influences layer represents a combination of three DRASTIC layers, the weighted sum of the corresponding DRASTIC weights and class midpoints was used to calculate an adjusted layer weight in the following manner:

$$\text{Hydro-Influences Weight} = [A \cdot w_A + C \cdot w_C + I \cdot w_I] / 2 \quad \text{Eq'n (3)}$$

where A, C, and I are the midpoints of the ranges of class ranks for DRASTIC's Aquifer Media, Conductivity and Unsaturated zone layers, respectively; the w_i are their DRASTIC weights; and the factor of 2 is the midpoint of the Hydrogeologic Influences class ranks in Table 3.

⁶DRASTIC's midpoint rank is that of the middle class (or average of two middle classes)

Table 4. Layer weights assigned in this study so as to preserve, as much as possible, the relative weight assignments of Aller and others' (1987) while objectivity accounting for differences in the number and ranks of calibrated layer classes.

<u>DRASTIC layer and weight</u>		<u>Equivalent calibrated layer and weight in this study¹</u>	
Depth to water	5	Depth to water	16.7
Net recharge	4	Septic recharge	12
Aquifer media and conductivity	3	Hydrogeologic influences	31.5
Unsaturated zone	5		
Soil media	2	Soil drainage	2.4 ²
		Soil permeability	2.4
Topography	1	Slope	n.a. (binary mask)

¹ see text for how adjusted weights were assigned based on class ranks

² if only soil drainage is considered, weight = 4; if only permeability, weight = 6

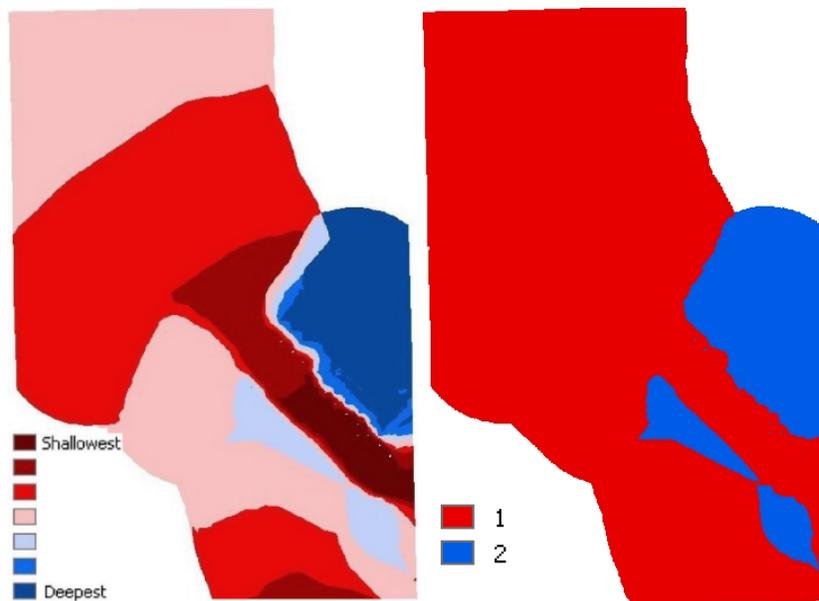


Figure 10. Comparison of uncalibrated (left) and calibrated (right) depth-to-ground water rasters. Note that in the statistically calibrated information layer on the right, only two categories are statistically justified (1 = least suitable; 2 = most suitable).

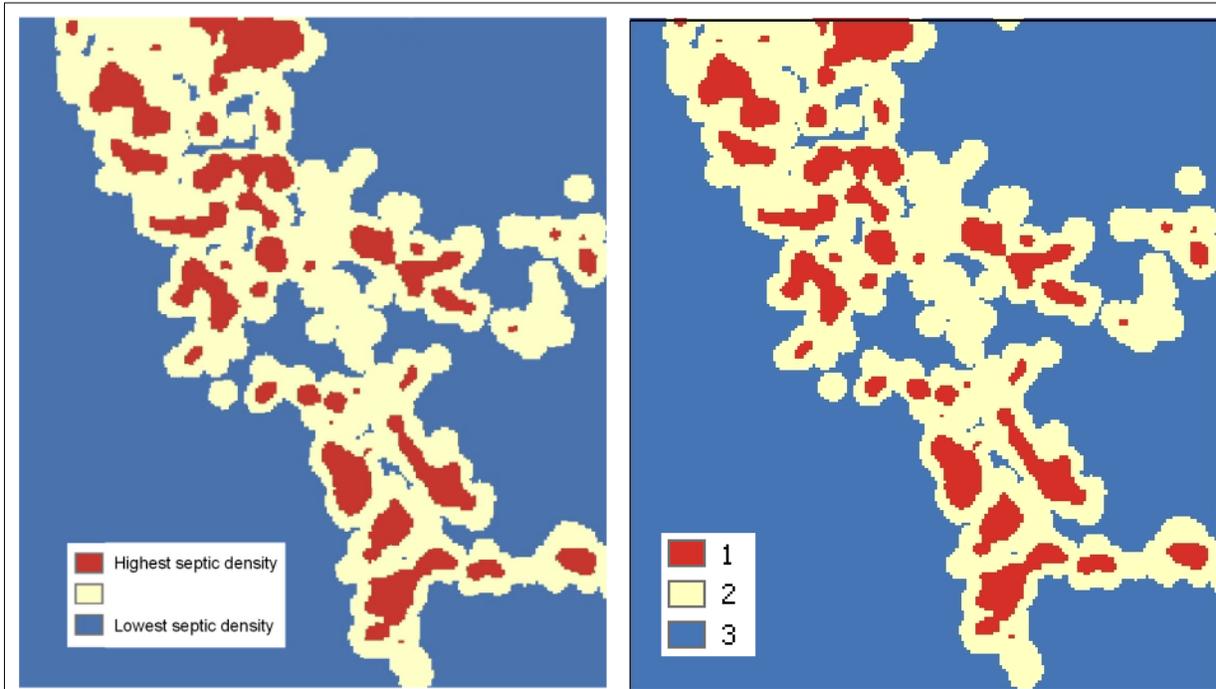


Figure 11. Comparison of uncalibrated (left) and calibrated (right) septic density classifications, showing that all three arbitrarily defined classes are statistically justified (1 = least suitable; 3 = most suitable).

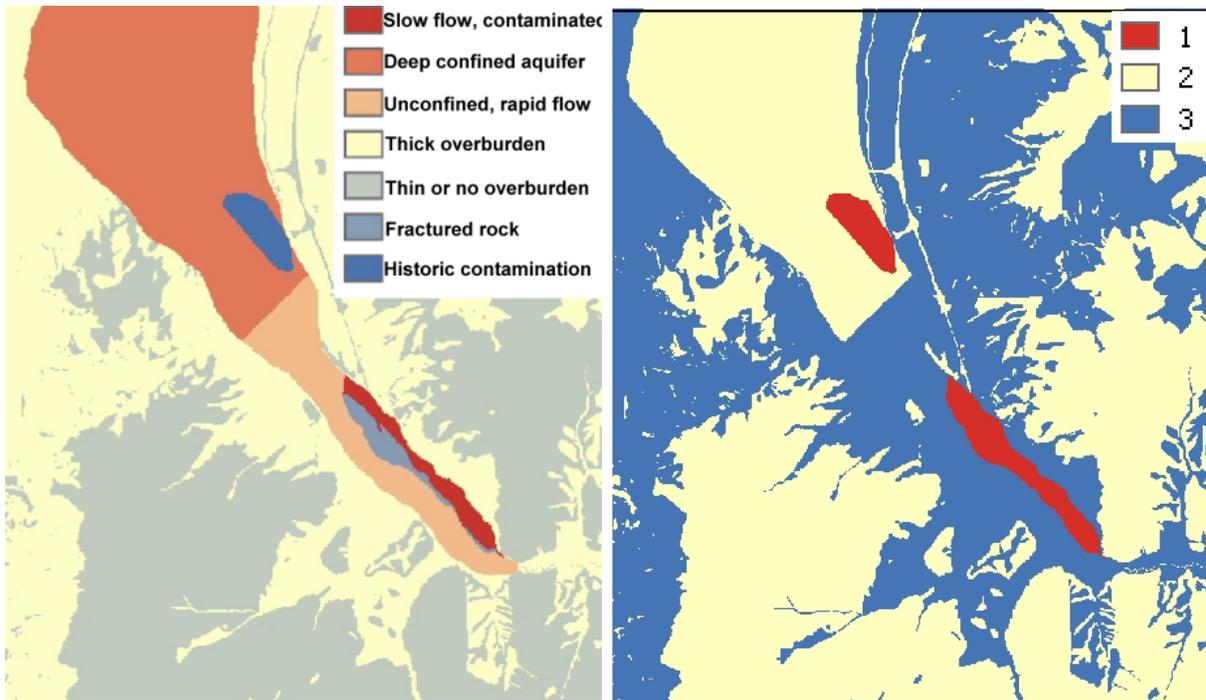


Figure 12. Uncalibrated (left) and calibrated (right) hydrogeologic influences. Note that only three of the original seven classes are statistically justified (1 = least suitable; 3 = most suitable).

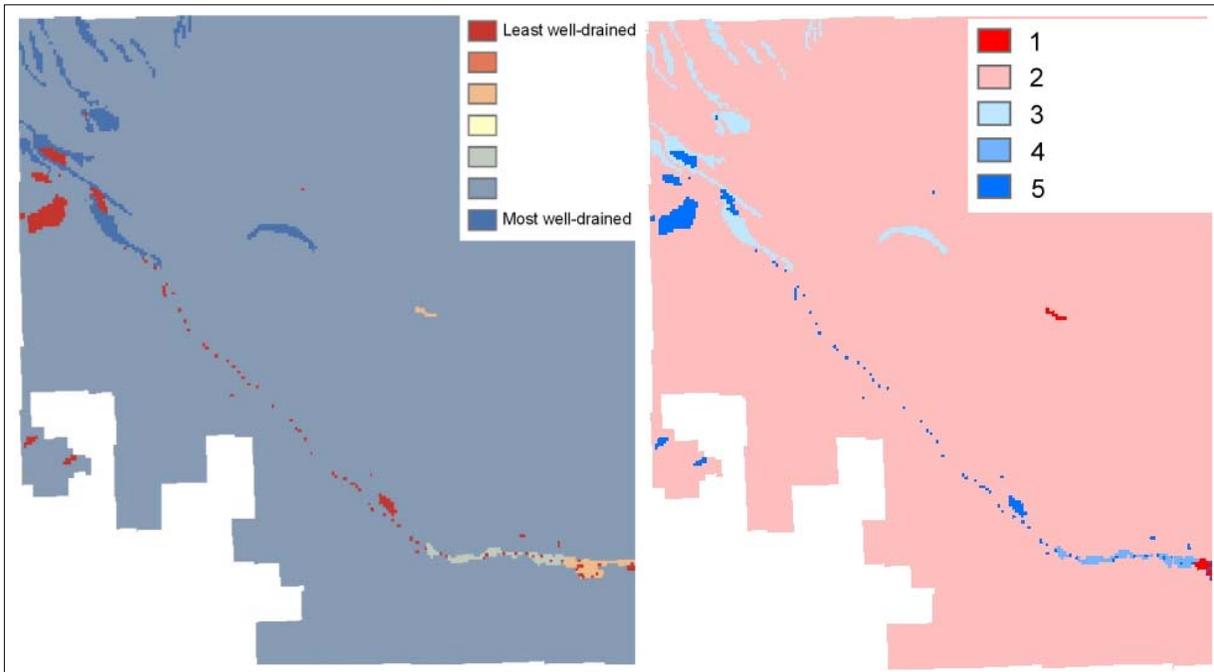


Figure 13. Uncalibrated (left) and calibrated (right) soil drainage classes. Five of the original seven classes are statistically justified (1 = least suitable; 5 = most suitable).

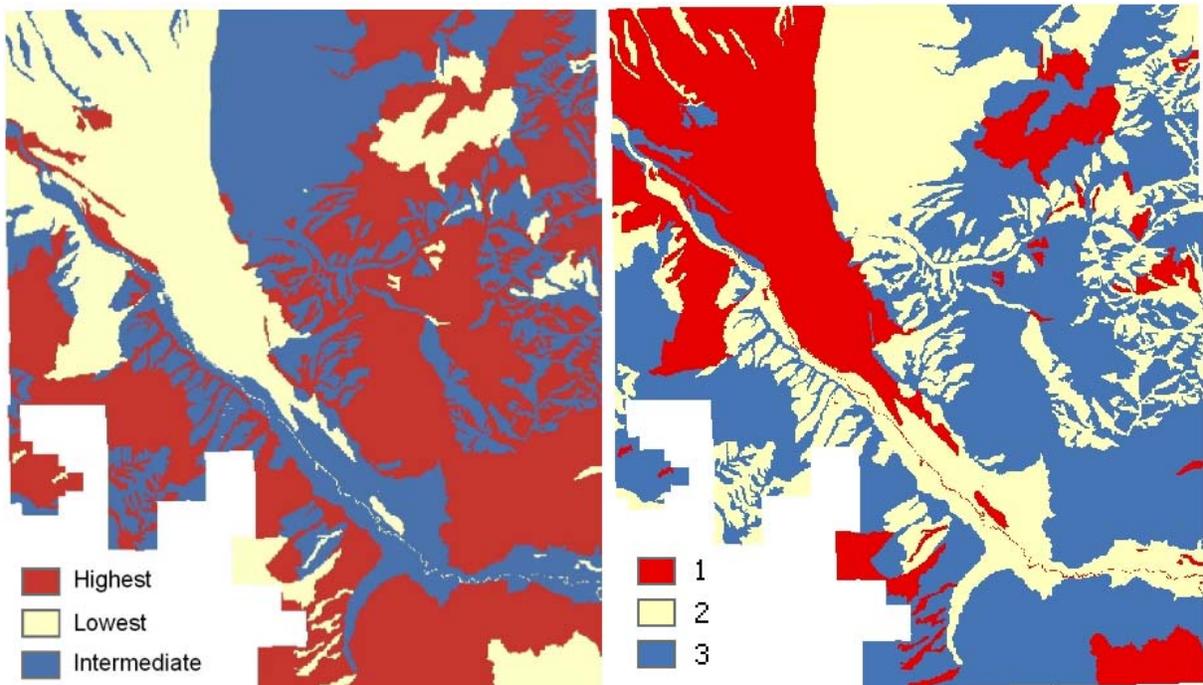


Figure 14. Uncalibrated (left) and calibrated (right) soil permeability classes. Note that statistical calibration reorders arbitrarily ranked classes (1 = least suitable; 3 = most suitable).

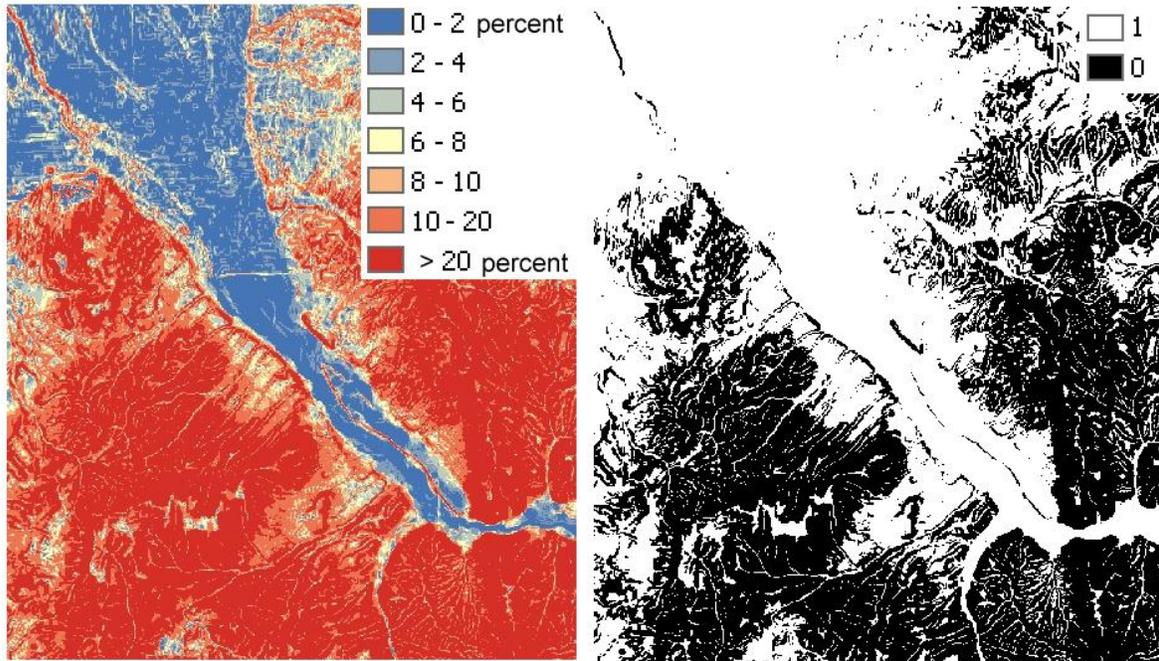


Figure 15. Arbitrarily classified slope map (left) and a binary mask created from it (right) showing areas classified as suitable (=1, slopes < 20%) and unsuitable (=0, slopes >20%).

In a similar way, the sum of the weighted products for Soil Permeability and Soil Drainage were equated to the weighted product of DRASTIC's Soil Media layer to calculate a combined adjusted weight for the two LPRV layers, which was then split equally among the two layers. That is,

$$\text{Combined Weight} = [K + Dr] * (w_K + w_{Dr}) / 2 = S * w_S \quad \text{Eq'n (4)}$$

where K and Dr represent the midpoints of the class ranks for Soil Permeability and Soil Drainage (Table 3), and S and w_S are, respectively, the midpoint of DRASTIC's Soil Media class ranks and its layer weight (Table 7 in Aller and others, 1987). The layer weights for Soil Permeability and Soil Drainage in Table 4 were obtained by solving Equation (4) for $(w_K + w_{Dr})$ and dividing by 2.

The spreadsheet "Table 4, 7 layer weights.xls" calculates adjusted layers weights in Table 4 and Table 7 (below). It is included on the CD-ROM and can be used as a template to accommodate any number or type of layers that might be considered in specific situations.

3.3. CALIBRATED SEPTIC SUITABILITY

Figure 16 shows the septic suitability rating map that was generated from the calibrated rasters in Figures 10 to 14 and the layer weights in Table 4, using Equation (1). It also shows an example of how relative suitability could be further classified on the basis of a fatal criterion like excessive slope, using Equation (2).

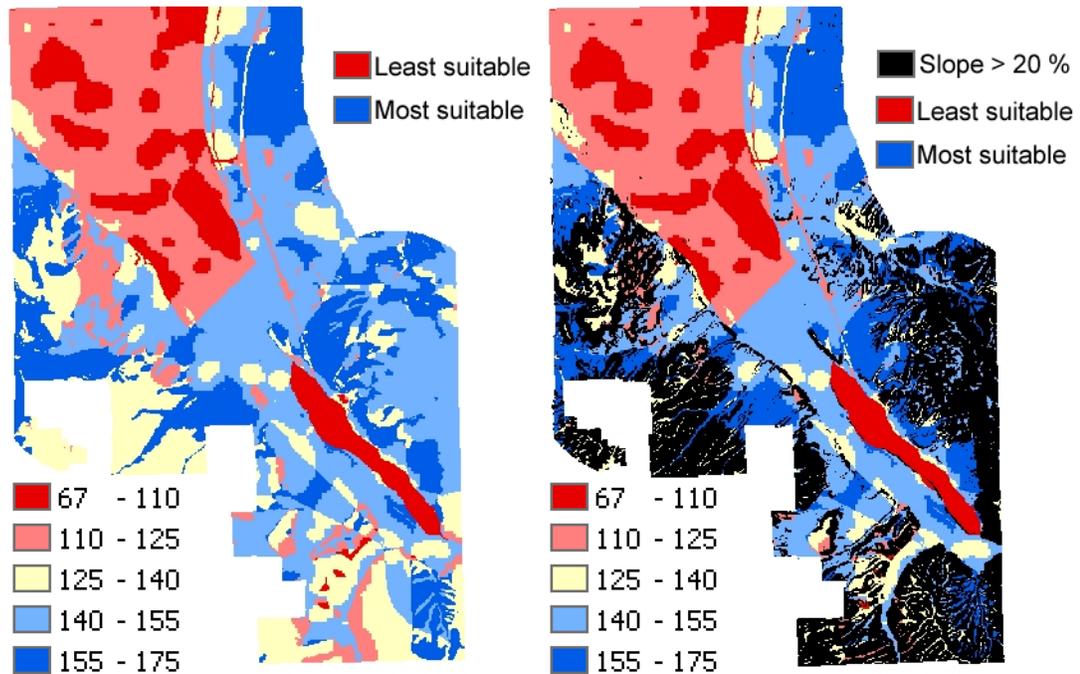


Figure 16. Septic suitability ratings for the study area (left) based only on physical factors and (right) incorporating a policy-driven factor (in this case, areas classified as unsuitable where slopes exceed 20 percent).

3.4. SENSITIVITY ANALYSIS AND FINAL CLASSIFICATION

3.4.1. Using calibrated information

The methodology in the preceding sections relies on statistical calibration to rank the class indices and maintains compatibility and technical objectivity by constraining the relative layer weights to conform to those in the original DRASTIC formalism. Unlike the original DRASTIC ratings, these suitability ratings are completely objective because they reflect class ranks that are spatially correlated with ground water that is progressively more contaminated with nitrate. However, the ratings are still only a measure of relative suitability. In the absence of other information, it is impossible to classify relative ratings into a map of suitable and unsuitable categories that are objective: the manner in which the ratings are classified is subjective. For example, within the ratings range of Figure 16 (67 to 175), should ratings less than 110 (the lowest 25 percent of cell-wise values) be considered unsuitable? Or should values less than 127 (the lowest 50 percent) or 145 (the lowest 75 percent) be classified as unsuitable? There is no objective answer because classification of a relative scale is inherently subjective. A more important question is: where are we most and least confident in the predicted ratings? This question can be answered objectively--and also provides an objective method for classifying the septic suitability ratings into suitable vs. unsuitable categories.

3.4.1.1. Sensitivity analysis protocol

The ArcMap script described in Section 2.7 was designed to perform a sensitivity analysis of a septic suitability map in order to determine where the suitability ratings vary the most as layer weights change. One hundred permutations of Figure 16 were generated, with slightly different values of the layer weights used in each permutation. The weights were drawn at random from a normal distribution centered on the values shown in Table 4, with standard deviations of one-quarter those values. For example, [11.1, 13.4, 12.3, 8.7, 9.6, 14.3, . . . , 12.9] are random values drawn from a normal distribution with a mean of 12 (septic recharge weight) and a standard deviation of 3. For a normal distribution having a mean of 31.5 (hydrogeologic influences weight) and a standard deviation of 7.8, the random values are [29.3, 24.9, 38.1, 35.4, 20.6, 41.6, . . . , 37.7]. These random weights were multiplied by the corresponding rasters (Figures 10 to 14) to produce 100 random permutations of the septic suitability ratings.

Figure 17 summarizes the cell-wise standard deviations and means of these permutations (that is, how much the calculated suitability rating varied at each raster cell location and its average, respectively). For a very large number of permutations (e.g., $N = 10,000$), the map of cell-wise means would be indistinguishable from Figure 16. In this example, with $N = 100$, small cell-wise differences averaging about 1.4% are discernible, the result of too small a sample size to fully average out the high and low variations in the calculated ratings. However, the cell-wise coefficients of variation (= standard deviation / mean) are much less sensitive to sample size, so that a map of their values accurately represents areas where the predicted ratings are most sensitive to the choice of layer weights.

Figure 18 shows the geographic areas that correspond to the lowest (least suitable) 10, 25 and 50 percent of the cell-wise means. **Figure 19** shows the areas with the 25 percent highest values of the coefficient of variation, indicating where the predicted ratings are least confident. Most prominent in this regard are the southern valley floor from Portneuf Gap to Red Hill, the East and West Benches north of Red Hill, and parts of the Mink Creek and Gibson Jack drainages.

3.4.1.2. Setting an objective suitability threshold

As discussed in Section 3.4.1, an objective classification cannot be created from a map of relative ratings in the absence of other information. The statistical results generated in Section 3.4.1.1 provide the necessary information to classify the relative ratings in Figure 16 into areas that are suitable and unsuitable for septic-based development. The details of the procedure are provided in **Appendix F** and will only be outlined here. In essence, the coefficient of variation (CV, Figure 19) provides a quantitative basis for setting a classification threshold around which the septic suitability ratings of Figure 16 can be objectively classified into suitable and unsuitable categories. The process involves finding the threshold that optimizes the distinction between suitable and unsuitable areas relative to areas whose classification is deemed uncertain at a specified CV threshold. The classification threshold at different CV levels in order to identify the value at which all areas of the suitability rating map are uniquely classified as either Suitable or Unsuitable and areas classified as Uncertain do not overlap with areas classified as Unsuitable.

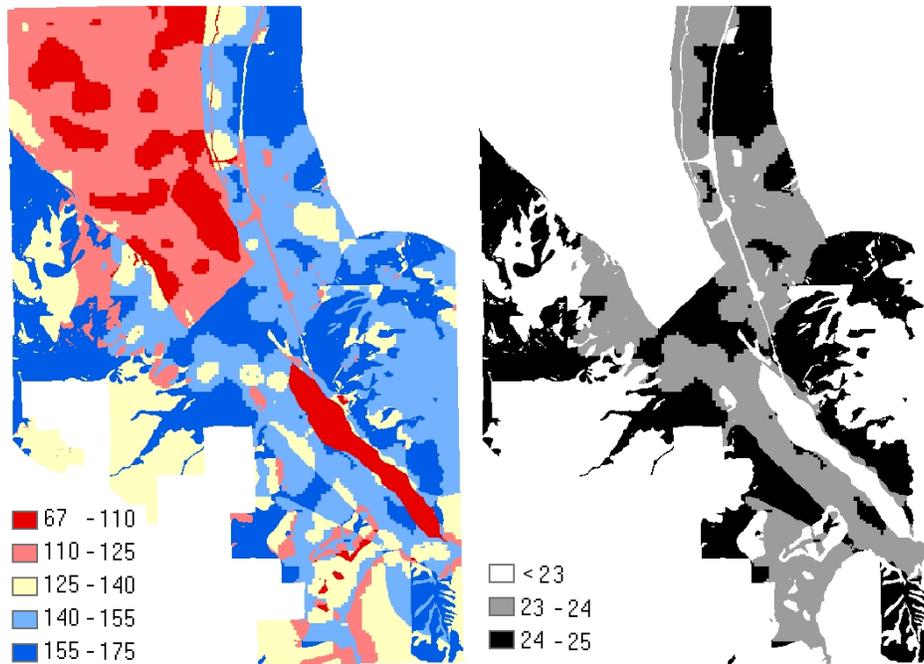


Figure 17. Cell-wise means (left) and standard deviations (right) of 100 permutations of the suitability rating map created by varying the layer weights, as described in the text.

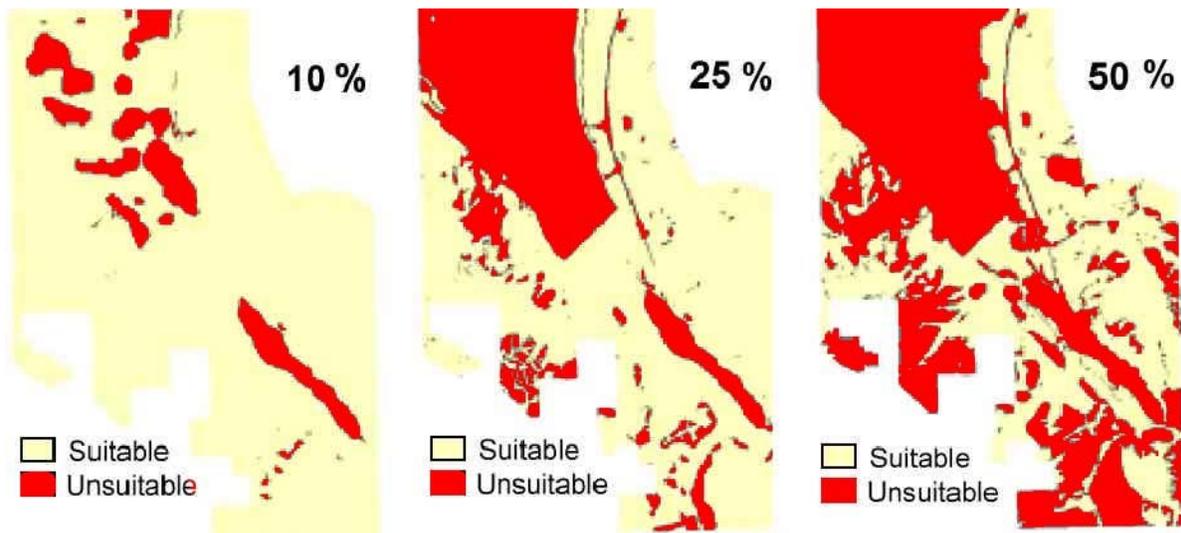


Figure 18. Fraction of the study area classified as unsuitable (red), based on average septic suitability ratings generated in the sensitivity analysis (Figure 17) and various suitability cutoffs: left, suitability rating <110 (10th percentile); center, rating <115 (25th percentile); right, rating <142 (50th percentile).

Figure 20 shows the result of applying this classification procedure. In this case, cell-wise mean DRASTIC index values above 139 (the 45th percentile of values mapped in Figure 17) represent suitability ratings that can be confidently rated as Suitable for septic-based development. At this threshold, 45 percent of the study area (in red) is classified as unsuitable for septic-based development. Areas shown in black have the highest 25 percent CV values and cannot be confidently classified. Such areas are not unsuitable for septic development, but their status has been objectively flagged at a specified confidence level to assist the decision-maker: For example, a developer proposing to install septic drain fields in such areas might be required to provide site-specific geotechnical information (e.g., nutrient-pathogen or other data) before the location is approved or disapproved for septic sewage disposal.

3.4.2. Using uncalibrated information

The methodology outlined in the preceding sections emphasizes the use of statistical correlations of ordinal and nominal information with water quality data to eliminate subjectivity in the assignment of class ranks, one of the major limitations of the original DRASTIC method. However, in areas that lack a substantial historical record of ground water quality monitoring or whose geographic sampling coverage is limited, statistical calibration may not be possible or justifiable. On the other hand, the use of uncalibrated layer information could do more harm than good if the information is classified and applied incorrectly to produce an erroneous suitability map. The accuracy, utility and defensibility of a map generated from information whose class ranks have been assigned subjectively depends entirely on how well those class ranks reflect the physical likelihood that surface contamination will compromise aquifer water quality.

This section provides a template for situations where statistical calibration cannot be performed. The rationale for what follows is based on a simple proposition: If uncalibrated layers are assigned class ranks that accurately reflect their relative contribution to nitrate pollution vulnerability, and the range of class ranks within each layer are considered when scaling the layer weights, then the septic suitability rating that is generated from such information is technically sound.

After the required information layers have been assembled, as described below, the assistance of a professional hydrogeologist or environmental engineer should be enlisted. To ensure that the suitability rating map will be technically defensible, the class ranks must be assigned using best professional judgment. Layer weights will need to be recalculated as in Section 3.2 to adjust for the range of class ranks in each uncalibrated information layer.

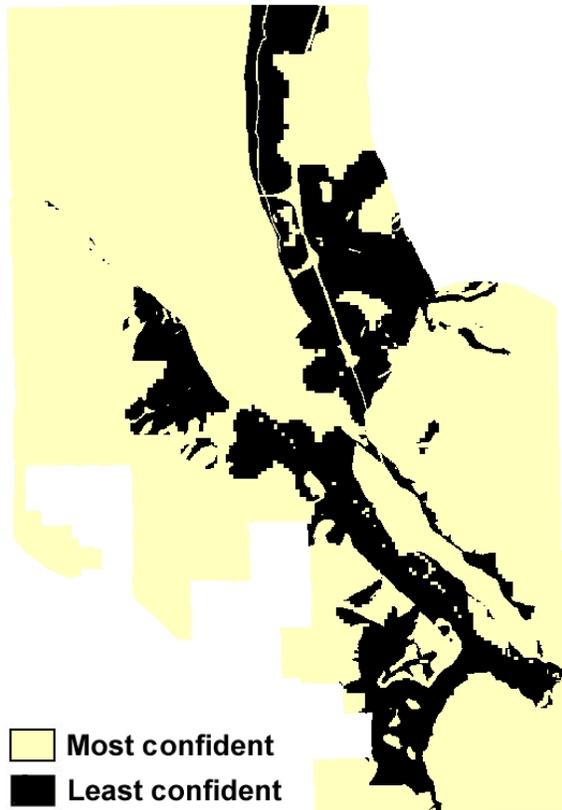


Figure 19. Areas with coefficients of variation at or above the 75th percentile as identified from sensitivity analysis (least confident predicted ratings)

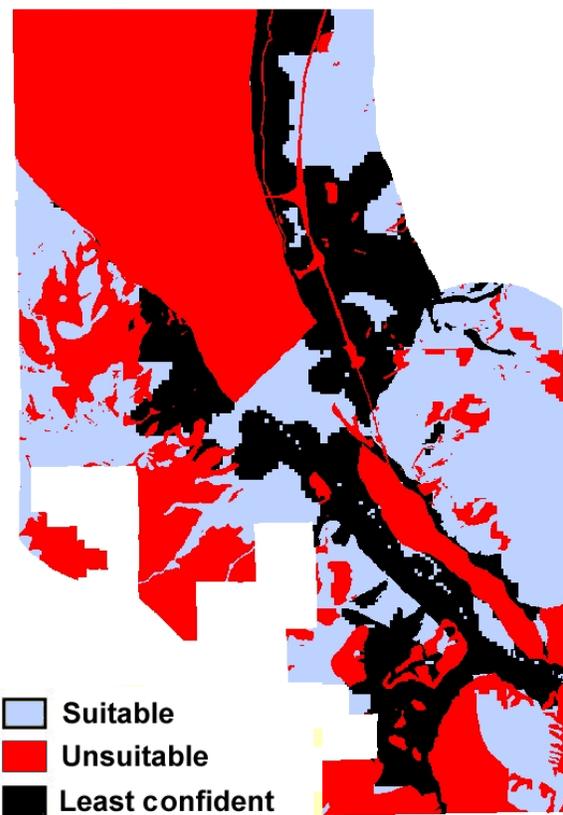


Figure 20. Final septic suitability map showing areas that are unsuitable for septic-based development at a confidence level of 75%. Areas in which more information is needed to make a decision at the 75% confidence level are shown in black.

The Hydrogeologic Influences layer was the most difficult to rank manually because its information is classified nominally and because the justification for both the original classes and their relative order was subjective. The seven nominal categories in Figure 8 could be reordered in many different ways: **Table 5** shows two possible rankings, both of which are based on sound professional judgment. Both classification schemes differ from the one based on statistical calibration (Figure 12). In this case, Alternative 2 was considered more defensible (fewer classes, hence fewer ranking decisions to justify). **Table 6** summarizes the resulting classes and ranks in each uncalibrated layer.

Table 7 summarizes the layer weights that were calculated as described in Section 3.2, using the "Table 4, 7 layer weights.xls" spreadsheet from the manually assigned ranks in Table 6. Based on these, a septic suitability rating map was generated, a sensitivity analysis was performed, and the classification protocol described in Section 3.3.1.2 was applied.

Figure 21 compares the septic suitability map generated from purely objective calibrated information (see Figure 20) and the map created with uncalibrated information that was evaluated and ranked using professional hydrogeologic judgment and classified using the procedure described in Sections 3.4.1.1 and 3.4.1.2. In this case, the suitability classification threshold is very low (at the 8th percentile of cell-wise means) and, more significantly, the level of classification uncertainty is high. As discussed in Appendix F, an objective classification is possible at no more than a 60% confidence level.

Table 5. Two possible ways of converting the nominal classes in the Hydrogeologic Influences layer to ordinal classes based on best professional judgment so that the ordinal ranks reflect relative suitability for septic-based development and are also technically defensible.

Alternative 1			
<u>Nominal Class</u>	<u>Class</u>	<u>Rank</u>	<u>Technical Justification</u>
6	Fractured rock	1	Both conditions are equally unsuitable
5	Thin or no soil	1	
7	Severe historic contamination	2	Confined aquifer, but nitrate is getting through somehow
1	Slow flow	3	Severe contamination, but this condition is not directly relevant to vertical transport
3	Very rapid flow	4	Unconfined aquifer, shallow water table but good dilution
4	Thick overburden	5	Of clear benefit, other things being equal
2	Confined aquifer	6	Of clear benefit, superior to thick overburden

Table 5. (continued)

Alternative 2			
<u>Nominal Class</u>	<u>Class</u>	<u>Rank</u>	<u>Technical Justification</u>
1	Slow flow and	1	In both areas, severe contamination is getting through somehow
7	Severe historic contamination	1	
6	Fractured rock	2	Both conditions are equally unsuitable
5	Thin or no soil	2	
3	Very rapid flow	3	Unconfined aquifer, shallow water table but good dilution
4	Thick overburden	4	Of clear benefit, and equal importance
2	Confined aquifer	4	

Table 6. Ordinal class ranks assigned to uncalibrated layers on the basis of professional judgment.

<u>Information Layers</u>	<u>Range of uncalibrated class ranks</u>
Depth to water	1 - 7 (shallowest to deepest)
Septic recharge	1 - 3 (most dense to least dense spatial arrangement)
Soil drainage	1 - 3 (see text)
Soil permeability	1 - 3 (least to most suitable; see Section 2.4.3)
Hydrogeologic influences	1 - 4 (see text and Table 5, Alternative 2)

Table 7. Layer weights assigned to uncalibrated information layers so as to preserve, as much as possible, Aller and others' (1987) relative weight assignments while accounting for differences in the number and ranks of calibrated layer classes.

<u>DRASTIC layer and weight</u>		<u>Equivalent layer and weight in this study</u>	
Depth to water	5	Depth to water	6.3
Net recharge	4	Septic recharge	12
Aquifer media and conductivity	3	Hydrogeologic influences	25.2
Unsaturated zone	5		
Soil media	2	Soil drainage	3
		Soil permeability	3

Figure 21 shows that although much more area is classified as suitable, a much larger area is classified as uncertain, particularly in the northern LPRV, where the presence of unseen physical factors negates the protective influence of the confined aquifer's aquitards. Essentially, lacking water quality information it is impossible to know that nitrate has contaminated the deep aquifer in spite of its protective clay units, so that the Hydrogeologic Influences layer cannot be appropriately classified, regardless of professional judgment.

3.5. COMPARISON OF RESULTS

Figure 22 compares nitrate-N concentration data with final classified maps derived from calibrated and uncalibrated information, showing areas that were classified as Suitable for septic-based development. **Figure 23** includes areas that were classified as Uncertain, and **Figure 24** summarizes results for the southern LPRV in greater detail.

The results indicate that a larger proportion of the study area was classified as Suitable in the uncalibrated analysis than in the calibrated analysis, mostly in the northern LPRV where the aquifer is confined. The uncalibrated analysis did not have the benefit of water quality data to rank layer classes, particularly for the Hydrogeologic Influences layer, so the northern aquifer was misclassified because the confining aquitard was not nearly as protective of aquifer water quality as expected for such geologic conditions. However, a large part of the misclassified area was deemed too uncertain to be reliably classified, suggesting that misclassification and uncertainty go hand in hand and that the use of uncalibrated information may be less of a concern to accuracy and technical defensibility.

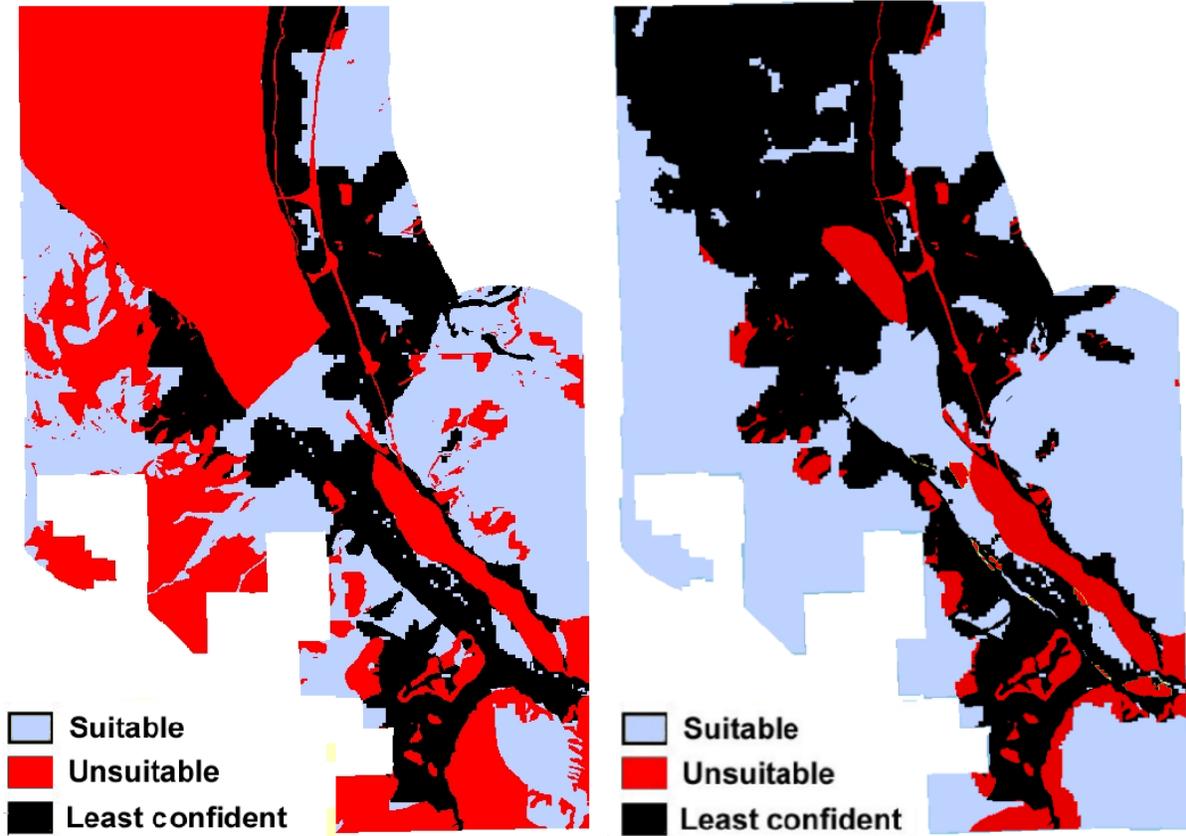


Figure 21. A comparison of septic suitability ratings generated with calibrated information (left, same as Figure 20) at a 75% confidence level and with uncalibrated information using best professional judgment (right) at a 50% confidence level. Both maps were classified using the objective classification protocol described in Appendix F.

As a final check on overall performance of the classification created from calibrated information, nitrate-N data were compared among areas classified as Suitable, Unsuitable and Uncertain. As shown in **Figure 25**, areas classified as Suitable have much less water quality data associated with them than areas classified as Unsuitable or Uncertain. **Table 8** compares the descriptive statistics of these three groups of information. Median nitrate beneath areas classified as Suitable is much lower than areas classified as Unsuitable. Although the mean and median nitrate concentrations in areas classified as Suitable and Uncertain are similar, a Mann-Whitney test returned a p-value of 0.00, indicating that median nitrate is significantly higher in areas classified as Uncertain. However, these areas tend to have much lower nitrate levels than areas that are classified as Unsuitable.



Figure 22. Areas classified as Suitable (blue), with nitrate-N concentrations overlain. Maps were created from calibrated (left) and uncalibrated (right) information.

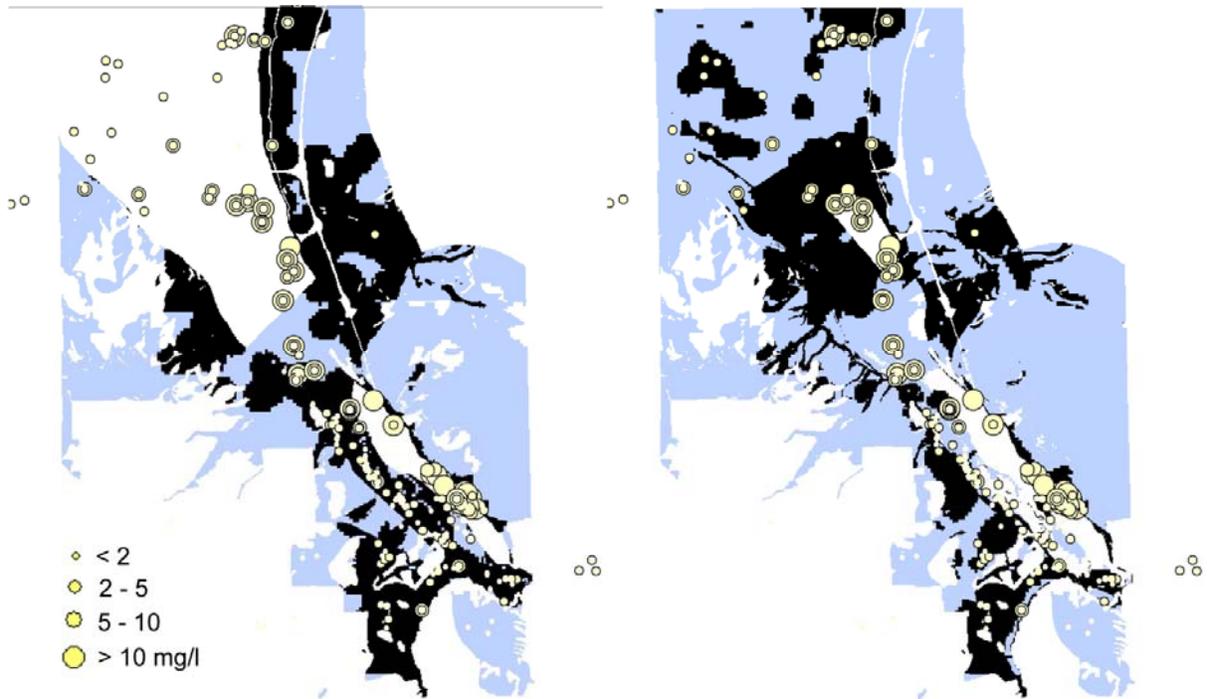


Figure 23. Areas classified as Suitable (blue) and Uncertain (black), with nitrate-N concentrations overlain, as determined using calibrated (left) and uncalibrated (right) information.

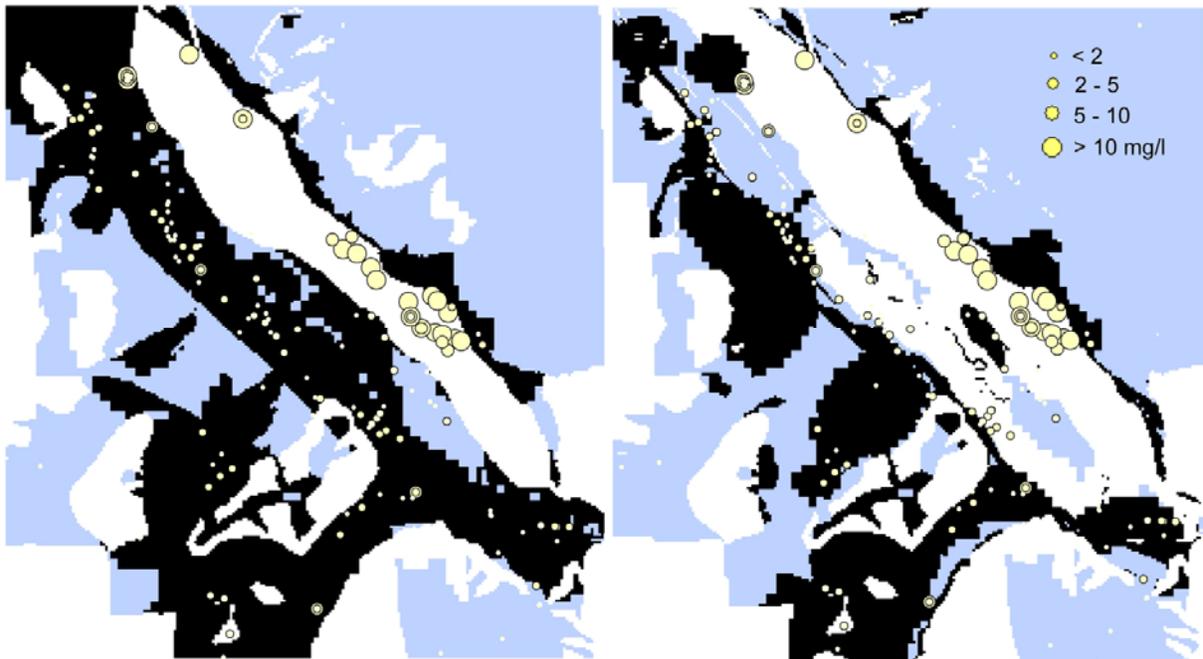


Figure 24. Enlarged view of southern aquifer showing nitrate-N concentrations in areas classified as Suitable (blue) and Uncertain (black), as determined using calibrated (left) and uncalibrated (right) information.

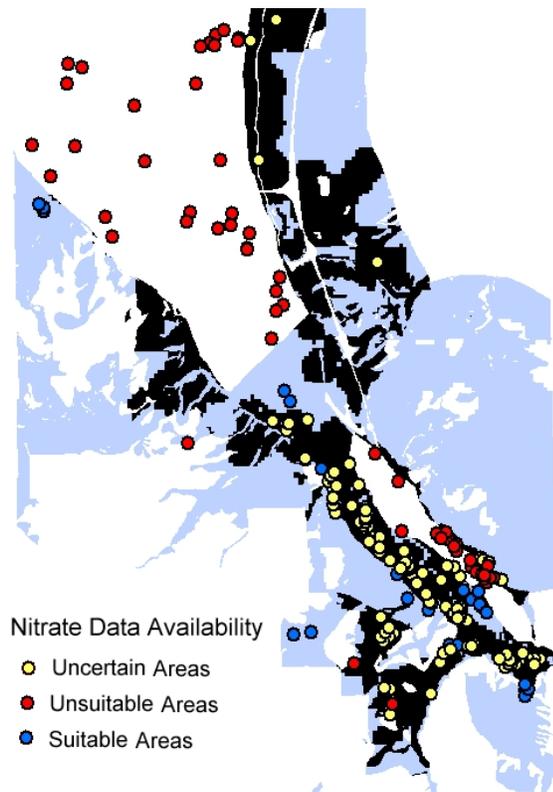


Figure 25. Availability of nitrate-N data in areas classified as Suitable, Unsuitable and Uncertain.

Table 8. Comparison of observed nitrate-N concentrations (mg/l) in areas of the LPRV that were classified on the basis of calibrated layer information as Suitable, Unsuitable and Uncertain for septic development.

	Suitable	Unsuitable	Uncertain
Sample size	131	581	481
Mean	2.25	6.37	2.86
Median	1.79	4.20	2.28
Std. deviation	1.92	11.40	2.13
Coeff. of variation	0.85	1.79	0.74

4. CONCLUSIONS AND RECOMMENDATIONS

The goal of this study was to develop an improved methodology for rating the relative suitability of land for septic-based sewage disposal. As such, it emphasized the use of readily available GIS data in an objective and technically defensible manner that is easy to implement. The method was developed and tested with data from the lower Portneuf River valley. The method is based on a modification of the EPA's DRASTIC layer-based formalism and relies on a sensitivity analysis to classify the resulting relative suitability scale into objective categories (Suitable / Unsuitable / Uncertain) at a specified level of statistical confidence. Water quality information is essential to correlate each physical risk-determining factor with historic nitrate impacts (Rupert, 2001) and help eliminate subjectivity in the results.

Information on depth to water, local recharge (septic throughput), soil drainage and soil permeability was utilized to create the rankings, as well as a "geohydrologic influences" layer that was proposed in recognition of the fact that quantitative information on aquifer media, hydraulic conductivity and/or the unsaturated zone may be sparse or nonexistent. This "composite" layer also permits the incorporation of anecdotal or other relevant proxy information when available.

The results of the LPRV case study are encouraging. Septic suitability index values were generated from readily available information on (i) depth to ground water, (ii) spatial density of existing septic drain fields, (iii) several relevant geohydrologic influences (soil thickness, fractured bedrock, ground water flow rate, , and (iv) soil characteristics considered most relevant to nitrate mobility (relative drainage and permeability). The methodology developed in this study can be used in situations where water quality information is unavailable, but as shown in a comparative analysis may deliver less than satisfactory results. Lacking water quality information to evaluate historic impacts on ground water, even the best professional hydrogeologic judgment may not account for hidden system vulnerability. In the LPRV, the resulting suitability map could only be classified at a 50 or 60% confidence level.

County, city or regulatory personnel who have basic ArcMap skills should be able to perform the data compilation, statistical calibration, and sensitivity analysis relatively easily. Statistical calibration is the most time consuming step in the process. The IGS satellite office in Pocatello can provide limited consultation and technical advice should users require assistance.

The methodology developed in this study consists of the following steps:

- 1) evaluate the quantity, reliability and spatial and temporal coverage of water quality data in the study area. Do sufficient data exist to identify temporal trends and justify a statistical calibration based on geographic rather than temporal variability of nitrate-N ?
- 2) if so, proceed with Steps 3 to 13; if not, go to Step 14
- 3) obtain the following data from state agencies and the Internet in GIS format:
 - SSURGO soil characteristics
 - IDWR well drillers' depth-to-water information
 - IGS geologic mapping information
 - IDEQ, IDWR, ISDA and USGS water quality data (nitrate-N concentrations)
 - USGS topographic data (digital elevation models); and
 - protected and sensitive lands (USFS, BLM, IDL, IDEQ maps)
- 4) acquire county land-parcel information and municipal sewer service maps and convert to a GIS map of existing septic drain field locations
- 5) assemble information on hydrogeologically relevant features in the study area, including but not limited to
 - depth of overburden (from geologic maps and/or drillers' logs)
 - principal aquifers and their geographic extents
 - areas in which silt/clay layers offer a degree of protection to an aquifer
 - areas where ground water is suspected to move very fast or very slow
 - areas where ground water has noxious odors, taste or iron staining
 - areas in which ground water is known to be, or has been, chronically contaminated (from any source)
- 6) Create GIS raster versions of the following information layers that are most relevant to nitrate mobility and accumulation:
 - (i) depth to ground water (feet below surface)
 - (ii) spatial density of septic drain fields (number per square mile)
 - (iii) hydrogeologic influences (map areas identified in Step 5)
 - (iv) soil drainage (categories of low to high)
 - (v) soil permeability (categories of low to high)

7) if nonphysical suitability factors are to be incorporated into a septic suitability rating (e.g., slope, protected lands, sensitive areas), then create additional GIS rasters representing that information

8) categorize the information in each of layers (i) to (v) into classes, the number of which should be sufficient to ensure that physical conditions that may correlate with water quality variations are adequately represented but not so large as to make subsequent statistical analysis more labor intensive than it need be (typically no more than five, plus or minus two, classes). Assign ordinal ranks to each class, where the ranks increase from conditions deemed least likely to prevent nitrate contamination to most likely. Class ranks for the Hydrogeologic Influences layer (nominal information) are arbitrary

9) for each raster of nonphysical information in Step 7, create a binary mask that represents conditions that are unsuitable (ranked 0) and suitable (ranked 1) for septic-based development

10) perform a statistical correlation of class ranks and nitrate-N concentrations in ground water for each of the raster layers (i) to (v); use the results to objectively rank the classes in each layer in contiguous order from least suitable (ranked 1) to most suitable (ranked highest)

11) using Equation (1), generate a suitability raster based solely on physical suitability factors (calibrated layers [i] to [v])

12) perform a sensitivity analysis of the suitability rating map, based on N=100 permutations (or more); using the method in Appendix F, reclassify the resulting map of cell-wise means into areas that are suitable and unsuitable for septic sewage disposal and areas that are uncertain

13) if any binary masks were created in Step 9, use Equation (2) to mask the reclassified suitability raster. The resulting map depicts areas that are suitable for septic development as well as areas that are unsuitable and areas that require additional site-specific information before a decision is possible.

14) if statistical calibration is not possible, perform Steps 3 to 9 and skip Step 10; instead, use best professional hydrogeological judgment to convert the Hydrogeologic Influence layer's nominal classes to an ordinal ranking that reflects increasingly suitable conditions for septic-based development and that is technically defensible, then proceed with Steps 11 to 13 and generate a classified map of septic suitability.

The principal difference between the calibrated and uncalibrated approaches, aside from the greater uncertainty associated with the latter, is their level of technical defensibility. An uncalibrated suitability map ignores the potential of ground-water quality information to objectively assign relative suitability to individual layer classes. The accuracy of a suitability map based on uncalibrated information is entirely dependent on the ranking decisions imposed by hydrogeological judgment. As shown in Section 3.5, however, even the best professional judgment is unable to compensate for hidden factors that only water quality information can reveal. A suitability map based on calibrated information reflects class ranks that incorporate unknown factors that may affect contaminant mobility, so its suitability ratings are as accurate as possible. Together with a sensitivity analysis of layer weights, a suitability classification based

on calibrated information will always be completely objective and technically defensible and is the preferred option in all situations.

ACKNOWLEDGEMENTS

The authors are indebted to Ted Dunsford, whose knowledge and expertise were crucial in the development of the ArcToolbox script for generating permutations during sensitivity analysis, and to Joe Sheffield for debugging the code to run reliably on ArcMap 9.3 and 10.0. We wish to thank IDEQ Pocatello Regional Office Ground Water Manager Tom Hepworth for his review of the draft report and for his enthusiastic support of this research effort from its inception, Shannon Ansley for her advice and encouragement throughout the project, and Toni Mitchell for her thorough review of the final draft.

This research was funded by the Idaho Department of Environmental Quality under Contract C751, and we are most grateful to IDEQ that the project was deemed worthy of funding in a very tight state budget situation.

REFERENCES CITED

- Al-Adamat, R.A.N., I.D.L. Foster, and S.M.J. Baban, 2003, Groundwater vulnerability and riskmapping for the basaltic aquifer of the Azraq basin of Jordan using GIS, remote sensing and DRASTIC: *Applied Geography*, vol. 23, p. 303–324.
- Aller, L., T. Bennett, J.H. Lehr, R.J. Petty, and G. Hackett, 1987, DRASTIC: A standardized system for evaluating ground-water pollution potential using hydrogeologic settings: Report EPA/600/2-87/035, U.S. Environmental Protection Agency, 641 p.
- Fetter, C.W., 1994, *Applied Hydrogeology*: MacMillan College Publishing, New York, 691 p.
- Focazio, M.J., T.E. Reilly, M.G. Rupert, and D.R. Helsel, 2002, Assessing ground-water vulnerability to contamination: Providing scientifically defensible information for decision-makers: U.S. Geological Survey Circular 1224, 33 p.
- Howarth, R., B. Burnell, and B. Wicherski, 2002, Nutrient-Pathogen Evaluation Program for On-Site Wastewater Treatment Systems: http://www.deq.idaho.gov/media/493564-nutrient_pathogen_eval_guide.pdf with mass balance spreadsheet for Level-1 evaluation at http://www.deq.idaho.gov/media/492856-nutrient_pathogen_eval_guide_calc.xls
- McLay, C.D.A., R. Dragten, G. Sparling, and N. Selvarajah, 2001, Predicting groundwater nitrate concentrations in a region of mixed agricultural land use: a comparison of three approaches: *Environmental Pollution*, vol. 115, p. 191–204.

Meehan, C., 2005, Geochemistry of the southern Pocatello aquifer and its implications for aquifer recharge and contaminant potential: unpublished Master's thesis, Idaho State University, 201 p. plus CD-ROM

Rahman, A., 2008, A GIS-based DRASTIC model for assessing groundwater vulnerability in shallow aquifer in Aligarh, India: *Applied Geography*, vol. 28, p. 32–53.

Rupert, M.G., 2001, Calibration of the DRASTIC ground-water vulnerability mapping method: *Ground Water* vol. 39, p. 625-630.

Thapinta, A. and P.F. Hudak, 2003, Use of geographic information systems for assessing groundwater pollution potential by pesticides in Central Thailand: *Environment International*, vol. 29, p. 87– 93.

Thirumalaivasan, D., M. Karmegam, and K. Venugopal, 2003, AHP-DRASTIC: software for specific aquifer vulnerability assessment using DRASTIC model and GIS: *Environmental Modeling & Software*, vol. 18, p. 645–656.

Welhan, J.A., C.W. Meehan, and T.V. Reid, 1996, The lower Portneu River Valley aquifer: A geologic / hydrologic model and its implications for wellhead protection strategies: Final Report to EPA Wellhead Protection Demonstration Project, 48 p. plus appendices.

Wicherski, B., 2006, Treasure Valley Nutrient-Pathogen Validation Study: Idaho Department of Environmental Quality, Ground Water Quality Technical Report No. 28, 62 p.

Appendix A

Examples of DRASTIC and DRASTIC-like Assessments in Various Jurisdictions in Montana and Wyoming

Wyoming Ground Water Vulnerability Assessment Handbook Vol. 1 – Ground Water Vulnerability Mapping Project (1998)

Retrieved from: <http://www.sdvc.uwyo.edu/groundwater/report.html> (see Chapter 4. Aquifer sensitivity mapping methodology)

Authors: Wyoming DEQ, Wyoming Water Resources Center, Wyoming Geological Survey, Wyoming Department of Agriculture, and the US Environmental Protection Agency.

Summary: Handbook describing the mapping project using the DRASTIC model to predict ground water sensitivity to pesticides. The sensitivity is based upon a modification of the DRASTIC model and potential for surface contamination from known land uses such as agriculture, landfill, and septic systems.

Note: Though septic systems are mentioned, I do not believe they are directly addressed in this study but rather focus more toward agricultural contamination. Even so, the document gives a good description of the DRASTIC model with nice flowcharts and illustrations as overviews to the input and processes.

Septic System Impact Study, Goose Creek watershed, Sheridan County, WY (2006)

Retrieved from: http://www.city-sheridan-wy.com/info/assets/pwd/ssis_final_report.pdf

Authors: Report was prepared by HKM Engineering with input and support provided by many sources including Sheridan Country, City of Sheridan, Sheridan Area Water Supply, and the University of Wyoming.

Summary: “HKM Engineering Inc. entered into an agreement with the City of Sheridan and Sheridan County in June 2006 to “assess the impact of septic systems in the Goose Creek Watershed, evaluate alternative treatment technologies and determine criteria for implementing various alternatives in high impact or high risk zones.” The project is divided into phases. The

first phase is the inventory of existing septic systems and mapping those systems, along with other relevant information, on a Geographic Information System (GIS) map. The outcome of Phase I is the identification of zones of high risk for impacts to groundwater. Phase II is intended to develop options to mitigate impacts from conventional septic systems. It includes identifying appropriate alternative technologies and methods by which those technologies could be applied within the Sheridan County septic permitting process.” For the purposes of the LPRV Septic Suitability Study, only Phase I is summarized.

A base map containing the general information and features within the watershed was compiled from the following input: Goose Creek watershed boundary, cities within Sheridan County (including year 2000 demographic data), creeks, lakes, county roads, national forest, parcel boundary, Sheridan Area Water Supply (SAWS) service area boundary, existing and potential sewer service boundaries, existing sewer system, and various aerial photography of the study area.

Input data included water quality monitoring stations and impaired streams (specifically fecal coliform discharge data), City of Sheridan Wastewater Treatment Plant outfall and fecal coliform discharge data, septic permits, data for number of SAWS accounts/taps, and the FEMA/FIRM flood plain boundary.

In addition, the aquifer sensitivity and contributing layers (Wyoming Groundwater Vulnerability Assessment Handbook, Vol. 1) were used to assist in the identification of areas where groundwater is more sensitive to contamination due to naturally occurring conditions within the aquifer. The contributing layers included depth to initial groundwater, geohydrologic setting (characteristics of the uppermost aquifer related to aquifer media and hydraulic conductivity), soils, aquifer recharge, land surface slope, vadose zone, and aquifer sensitivity (derived using the six prior layers). “The procedure used for defining Sensitivity within the Wyoming Ground Water Vulnerability Assessment Handbook has been modified from the original DRASTIC model in the following ways:

1. The hydraulic conductivity and aquifer media layers have been combined within the Wyoming model to form a "geohydrologic setting" layer. Therefore, the Wyoming model only contains six model parameter layers while the DRASTIC model contains seven.
2. The Wyoming model uses a different method for assigning rating values based on the unique nature of Wyoming’s hydrogeologic environment.
3. Equal weights are assigned to the parameters due to lack of scientific information describing the weight relationships between these parameters.”

Impact Zones of critical, high, medium, and regular were designated and mapped according to parcel boundaries. The Impact Zone designations are intended to provide a key map which will show landowners, septic installers and the permitting authority in which areas alternative systems might be appropriate.

Evaluation of Unsewered Areas in Missoula, Montana (1996)

Retrieved from:

<http://www.co.missoula.mt.us/WQ/FAQs/Reports/UnseweredReport/pdfs/unsew.pdf>

Authors: Missoula Valley Water Quality District, Environmental Health Division, Missoula City-County Health Department

Summary: “The study analyzes, and ultimately ranks, unsewered areas according to environmental and public health factors. To evaluate relative health risks associated with the use of septic systems in unsewered areas, data related to septic systems, water supply wells and hydrogeology were compiled. This included septic system densities, types of septic systems, public and private well locations, soil types, depth to groundwater, groundwater flow direction, hydraulic conductivity and land use. The hydrogeologic properties (excluding groundwater flow direction) of the unsewered areas were evaluated using an aquifer sensitivity method called DRASTIC (EPA, 1987). The DRASTIC method was used to map the relative sensitivity of groundwater under the unsewered areas to degradation by septic systems.

The unsewered areas were evaluated and ranked based on the following eight factors:

1. Percentage of commercial properties
2. Overall septic system density
3. Total sewage loading
4. Percentage of septic systems using seepage pits
5. Percentage of septic systems replaced since 1967
6. Average current groundwater nitrate-N concentration
7. Overall water well density
8. Overall aquifer sensitivity based on DRASTIC analysis

For each unsewered area, the eight factors listed above were compared on a relative scale by 7 ranking each area from 1 to 8, for each factor considered. The highest ordinal ranking score of 8 was assigned to the unsewered area with the highest (or worst) value for the factor. For example, the area with the highest density of septic systems was assigned the ordinal ranking score of 8. The final prioritization of the unsewered areas was determined by summing the ordinal ranking scores for each of the eight factors to obtain a total score as follows:

Total Score = scores for (% commercial units + septic system density + % replacement septic systems + % seepage pits + average nitrate-N concentration in groundwater + well density + sewage loading + average DRASTIC value).

The unsewered areas were then prioritized based on the total score, with the area with the highest total score being assigned the highest priority.

Geographical Information System (GIS) software was used to create a series of maps which are included as attachments to this report. The maps summarize the data collected on septic system density, locations of wells, results of nitrate and bacteria sampling, groundwater flow direction and the results of the DRASTIC aquifer sensitivity analysis. While groundwater flow direction was not used directly to rank the areas, it is presented as additional information to consider along with the final prioritization.”

Bitterroot Valley Groundwater Vulnerability Mapping Project (2008)

Retrieved from:

<http://www.ravallcounty.mt.gov/environmental/documents/BitterrootValleyGWVulnerabilityMapping.pdf>

Authors: presentation prepared by PBS&J

Summary: “The project is intended to identify areas that are most vulnerable to impacts from septic system discharge and provide a planning tool to be used to help prevent or reduce nutrient loading to groundwater. Phase I uses the standard DRASTIC methodology was used to identify areas of groundwater vulnerability. Phase II uses a refined/calibrated DRASTIC analysis to predict nitrate impacts:

1. Update the depth to groundwater map to reflect depth of water
2. Plot Water Quality (Nitrates) to Identify Impacted Areas
3. Correlation Analysis of DRASTIC Parameters to Nitrates
4. Adjust Parameters to Reflect Best Indicators of Nitrate in Groundwater.

Maps were prepared for Septic Density, Nitrate Concentration, Statistical Correlations of DRASTIC Inputs to Nitrate Concentration, and a Revised Groundwater Vulnerability Map. Conclusions were as follows:

1. Phase 1 DRASTIC did not appear to explain (correlate) Nitrates in Groundwater
2. Reversing Recharge and Soil Parameter Values provided better fit
3. Highest Vulnerability appears along eastside terraces
4. Tool that can be refined as better data becomes available

Ravalli County Land Suitability Analysis (2008)

Authors: Ravalli County Planning Department, GEUM Environmental Consulting, Inc., and DTM Consulting, Inc.

Retrieved from:

<http://www.ravallcounty.mt.gov/planning/documents/RavalliCountyLandSuitabilityAnalysisReport.pdf> . A presentation of the report is also available at:
<http://www.ravallcounty.mt.gov/planning/documents/LandSuitabilityAnalysisPresentation012608.pdf>

Summary: “The Land Suitability Analysis (LSA) described in this document is a Geographic Information Systems (GIS) based tool for evaluating the relative suitability of land for development in Ravalli County, Montana. The purpose of this analysis is to support the development of baseline zoning in Ravalli County. The end product is a generalized map

showing areas of the County that are more or less suitable for development. The resulting suitability map is not intended to be used at a site specific scale and does not constitute a zoning map.

For purposes of this analysis, *suitability* can be defined in terms of physical limitations, existing regulatory restrictions, and the community goals expressed in the Ravalli County Growth Policy (Ravalli County 2002, rev. 2004). Physical limitations such as steep slopes, poorly drained soils or high risk of wildfire make the land less suitable for development. Features subject to existing regulatory restrictions, such as a mapped floodplain or wetland, also pose challenges to development. The goals in the Ravalli County Growth Policy were used as a guide in determining areas of the County that are important community resources, such as open space and wildlife habitat, as well as areas where development should be encouraged, such as near existing towns and infrastructure.

The results of the LSA show areas within the County that are more or less suitable for development based on six categories: (1) existing infrastructure; (2) water resources; (3) wildlife resources; (4) working lands; (5) open lands; and (6) public health and safety. These categories are referred to as "*submodels*." Category (6) public health and safety considers septic suitability based on the NRCS SSURGO data.

Note: Appendix E – ModelBuilder Schematics and Data Descriptions contains modelbuilder images and descriptions of the model processes.

Appendix B

GIS Information Layers Compiled for This Study

All GIS data developed in this study were compiled in Idaho Transverse Mercator (NAD 1983) projected coordinates, in meters. Data were converted to ESRI raster format (floating type) with a cell size of 30.42 x 30.42 meters (100 x 100 feet).

Digital geologic mapping data was obtained from the Idaho Geological Survey (<http://www.idahogeology.org/>) and soils data from the Natural Resources Conservation Service (<http://datagateway.nrcs.usda.gov/NextPage.aspx?HitTab=1&Progress=0>). Depth to ground water was interpolated from the Idaho Dept. of Water Resources' well construction database (<http://www.idwr.idaho.gov/ftp/gisdata/GISScripts/downloadform.asp?path=Spatial/Wells/WellConstruction&package=wells.pkg>). Topographic slope was calculated from 30-meter resolution U.S. Geological Survey digital elevation models available as 7.5-minute quadrangles at <http://www.geocommunity.com/>.

The septic recharge layer was constructed by IDEQ's Pocatello regional Office. M. Byrd (written comm., 2002) assembled Bannock County's parcel database and coverages of the City of Pocatello's and City of Chubbuck's municipal sewer systems. County land parcel polygons were geo-coded to create a map of addressed parcels. Parcels that lacked an address were assumed to be undeveloped and parcels with addresses were compared to a buffered map of municipal sewer lines. An addressed parcel boundary more than 200 feet from a sewer line was assumed to be serviced by a private septic system whose location was assigned as the centroid of that land parcel. The resulting map (**Figure A.1**) portrays more than 3000 septic drain field locations in the LPRV and reflects the geographic extent of septic-based development as of 2002.

A copy of the data is available by contacting the author at weljohn@isu.edu.

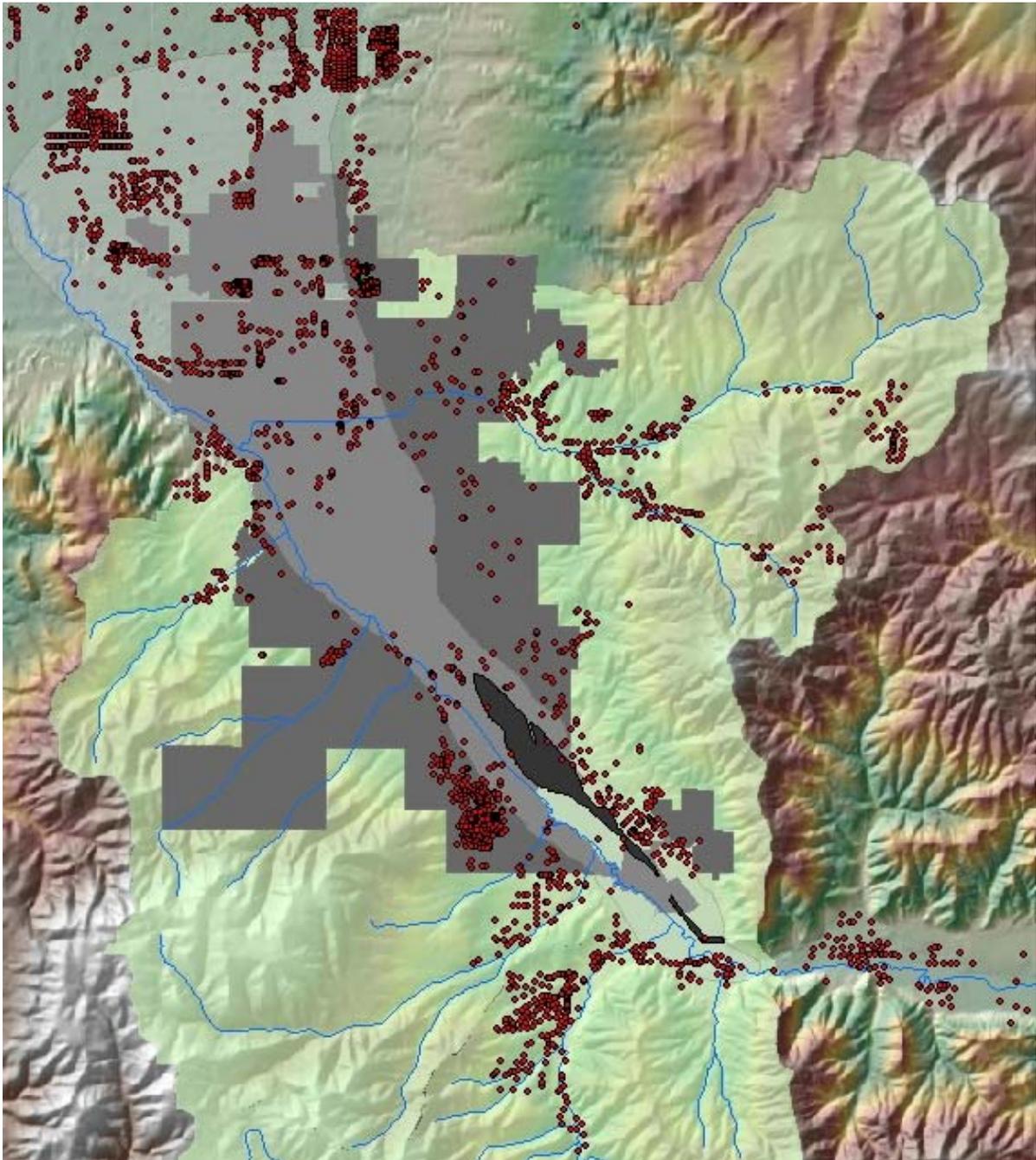


Figure B.1. Approximate locations of septic drain fields in the LPRV, as estimated from county parcel maps and the geographic extent of municipal sewer services.

Appendix C

Compilation of Water-Quality Data for the Lower Portneuf River Valley

Ground-water chemistry data were taken from Meehan's (2004) compilation that represents all available nitrate information in the LPRV from ca. 1952 to 2004. Those data were organized in four ESRI shapefiles derived from a) several ISU sampling campaigns, b) IDEQ's database (including IDWR's Statewide Monitoring Network data), c) the City of Pocatello's municipal well sampling, and d) Bannock County's ground-water monitoring of the Fort Hall Canyon landfill. Each data set contained different data fields and attribute information according to its origins so that in concatenating the files into a single data set, data fields had to be rearranged, renamed and/or created. In addition, the data locations had to be projected into IDTM NAD 83 and unnecessary fields and duplicate data were removed.

Wells outside the LPRV study area were deleted from the compilation. Location information was unavailable for some wells so well locations were assigned by digitizing over NAIP rectified aerial imagery. Any fields that indicated source type (e.g., "well" vs. "spring") were retained, and information on the source of the chemistry data was retained in the "SOURCE" and "SOURCE_NAM" fields. Meehan had assigned region designators to indicate where in the LPRV the sampling site was located. These designators are:

- MA ("main aquifer") - wells in the high-permeability municipal aquifer and City of Inkom wells; domestic wells that are located in this aquifer in the vicinity of the Portneuf Gap are designated "SG"
- SW ("southwest aquifer") – includes wells in the west bench (WB) and Mink Creek (MC) areas
- EA ("east aquifer") – wells in the eastern aquifer (beneath South 5th St.); a subset of these, in the Black Cliffs Trailer Park, are designated "BC"
- NE ("northeast aquifer") - well locations in the KOA Campground area of Pocatello Creek and wells below Chink's Peak that do not fall in the EA category
- FHC ("Fort Hall Canyon") - monitoring wells below the Bannock County landfill
- OAI ("outside area of interest"): 13 wells south and southeast of the Mink Creek area

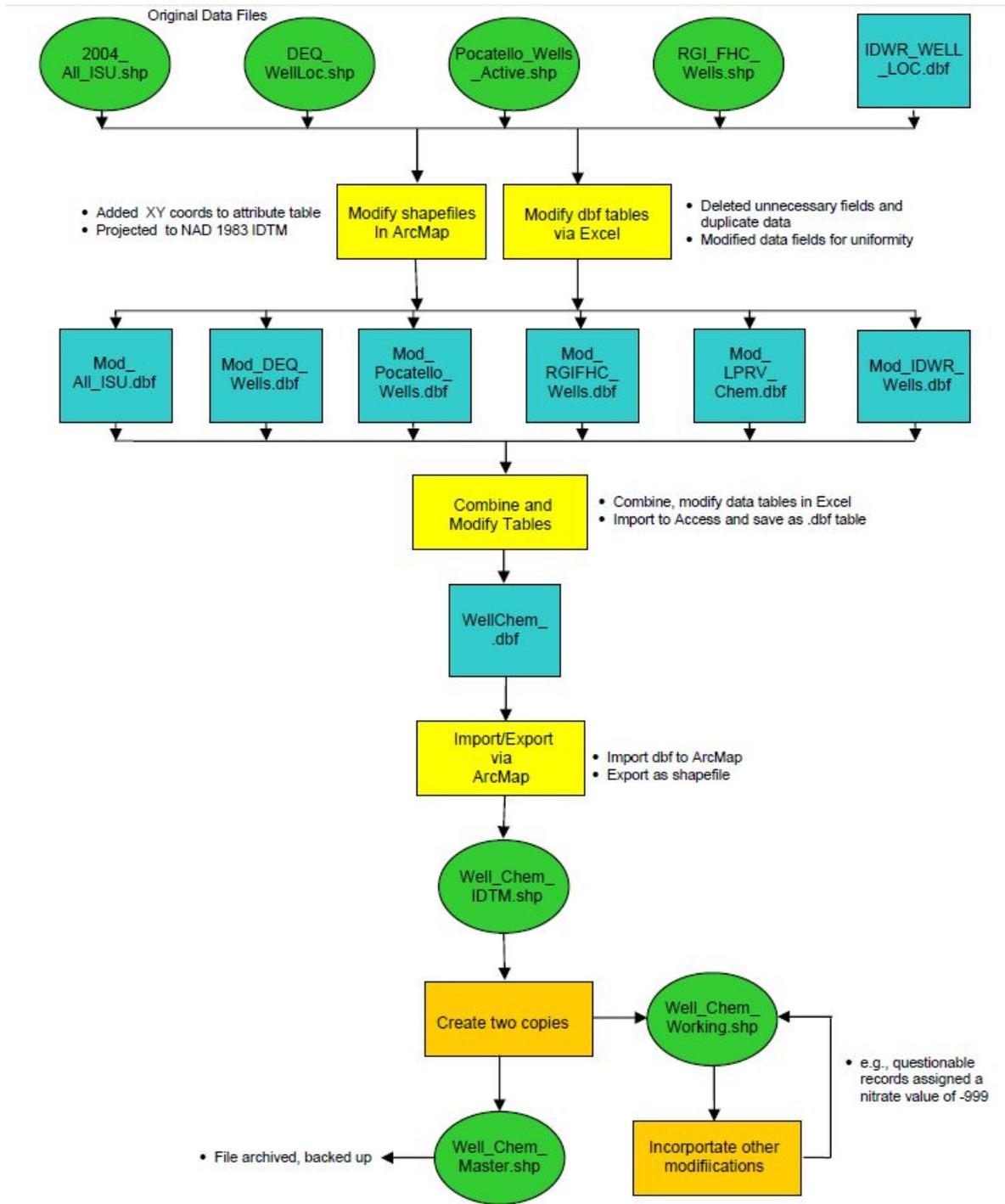


Figure C.1. Schematic of the procedure used to convert multiple shapefiles from different sources into a single data file.

Nitrate data were composited from several fields in the original data sets. Data in fields representing (i) nitrate-nitrogen concentration and (ii) nitrate- plus nitrite-nitrogen concentration were composited as three fields: "NO3_N", "NO3_NO2_N", and "N". The "N" field was assigned the all the values in the "NO3_N" field except where an entry's decimal information was missing, in which case the "NO3_NO2_N" value was assigned to the "N" field. Missing nitrate data were assigned a value of -999; a value of 0 indicates missing data for all other analytes. The final step was to assign a data quality flag ("DQ_Flag" field) to indicate the presence of suspected outliers outliers (spatial and statistical), erroneous values, or other anomalies.

At this point, a master copy of the composited data set was created and any other modifications were made on a working copy. For example, in Bannock County's Fort Hall Canyon landfill data, a nitrate concentration of "99.0" in the context of two other samples collected the same day ("16.0" and "2.0") was not considered to be a reliable representation of the true nitrate concentration in the aquifer. In this and similar instances, the measured nitrate concentrations represent replicate samples collected during purging of the well. Short of checking the original data sheets to verify the assumption, in all such cases only the lowest value was assumed to be representative of the aquifer at that location on that date. Therefore, the high and intermediate values in replicate samples were flagged as unreliable and their nitrate concentrations were set to "-999".

A copy of the data is available by contacting the author at weljohn@isu.edu.

Appendix D

Worked Example of the Statistical Calibration Process

As Rupert (2001) has shown, the statistical calibration of information layers is an indispensable step in developing an objective and technically defensible risk / suitability map. To minimize subjective influences on how an information layer is geographically subdivided and ranked, each sub-area's relationship to nearby septic source impacts must be quantified and ranked in a completely objective manner. Because nitrate is the most sensitive indicator of septic impact on ground water quality, its concentration in different parts of an aquifer system serves as a convenient measure of the relative impact on the aquifer.

The calibration process consists of four steps: 1) create individual polygon layers, one for each category in the information layer; 2) segregate all available nitrate data into groups that correspond to individual subareas; 3) compare nitrate concentrations among these groups to identify areas whose nitrate concentrations are statistically indistinguishable and combine these areas into new polygon areas; and 4) rank the parts of the aquifer system that have distinctly different nitrate concentrations using an ordinal classification ranking (e.g., 1, 2, . . . , 5) that quantitatively indicates where ground water quality has been affected the most (1) and the least (5) by nitrate contamination. The resulting calibrated ranking reflects each geographic area's relative suitability for siting a source of nitrate, such as septic drain fields, within that area.

In this example, all data were processed using ArcMap and Statmost statistical analysis software, but any software could be utilized. To illustrate the calibration process, data pertaining to the "hydrogeologic influences" data layer will be evaluated in a step-by-step fashion. This particular information layer is an example of a nominally categorized data layer that was arbitrarily subdivided into areas indicating such factors as "thin overburden", "fractured rock" and "slow flow" that affect how easily surface sources of nitrate could enter and spread through the underlying aquifer. However, the overall process is the same for any information layer, whether it is ranked ordinally or nominally.

1) Create Individual Polygon Shapefiles for Each Subarea in a Data Layer

Seven geologically and hydrologically important subareas within the hydrogeologic influences layer were identified that could affect how nitrate travels to the water table. In this case, the subareas were defined based on expert judgment but could have been arbitrarily drawn. In any case, the subareas at this point are subjective and need to be statistically evaluated to define objective boundaries and class rankings (the calibration process). Therefore, seven subarea polygons were created using four geological coverages - shapefiles

representing (i) the aquifer boundary, (ii) municipal aquifer, and (iii) the area of the Portneuf lava flow and a raster representing (iv) areas of exposed bedrock and overburden cover – plus an arbitrary subdivision of the municipal aquifer into northern and southern segments (in this case, because the southern segment is unconfined whereas the northern is confined), plus an area within the northern segment that has seen some of the highest nitrate concentrations in the LPRV and therefore may indicate that hydrogeologic conditions in the vicinity make the aquifer susceptible to contamination from surface nitrate sources. Note that raster (iv) comprises three areas: no overburden thickness (“bedrock” in Figure D.1.1), a minimal thickness (ca. <100 feet) of overburden (“other sediments” in Figure D.1.1) and considerable overburden thickness (>100 feet) in the valley (the municipal aquifer in Figure D.1.1).

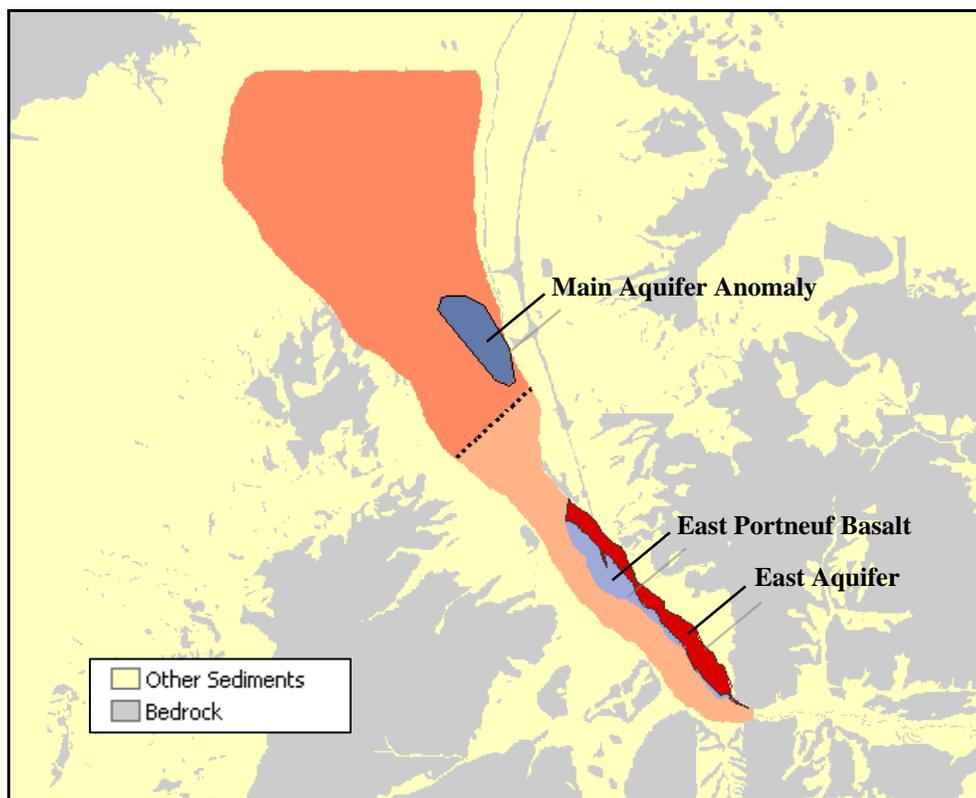


Figure D.1.1. Hydrogeologic Influences layer consisting of seven arbitrarily defined subareas. Individual polygon shapefiles of each subarea were created from this to select the nitrate data that correspond to each polygon.

a) East Aquifer Subarea polygon shapefile:

As shown in figure D.1.1 above, a polygon shapefile representing the east aquifer subarea was created from the southeastern part of the municipal aquifer. This was accomplished using shapefiles (i), (ii) and (iii) as follows:

To begin, the Union Tool (Figure D.1.2) was used to combine these three shapefiles into a single shapefile. Since this tool allows the union of only two shapefiles at a time, the tool was used twice to combine the aquifer boundary, main aquifer, and Portneuf Basalt into one shapefile.

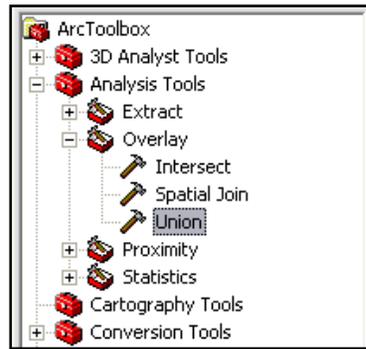


Figure D.1.2. Arctoolbox Union tool.

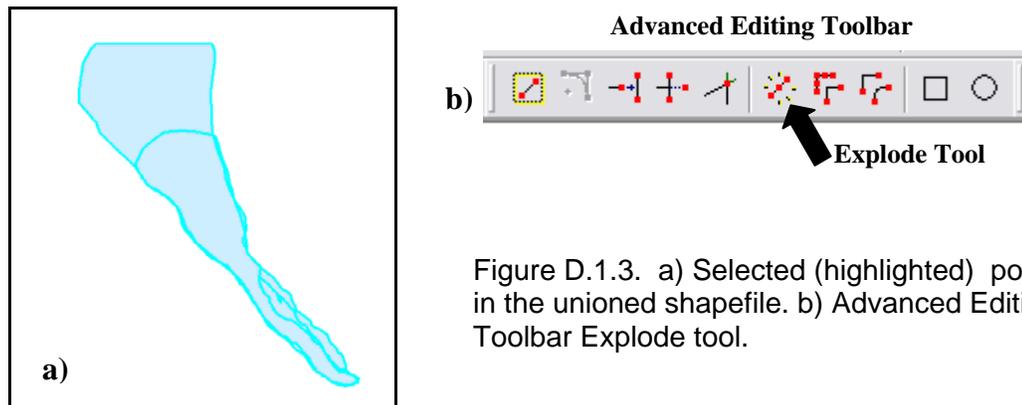


Figure D.1.3. a) Selected (highlighted) polygon in the unioned shapefile. b) Advanced Editing Toolbar Explode tool.

Next, all attributes in the unioned shapefile were selected and an editing session was begun. From the Advanced Editing Toolbar, the Explode Tool was used to separate the polygon into individual polygons, the editing session was stopped and edits were saved. (If the Advanced Editing Toolbar is not visible, add it by selecting from the dropdown menu Editor → More Editing Tools → Advanced.)

Selecting only the southeastern polygon (Figure D.1.4), it was then exported by performing a right click on the unioned shapefile in the table of contents and selecting Data → Export Data, selected features, same coordinate system as this layers source data, and output: East_Aquifer.shp.

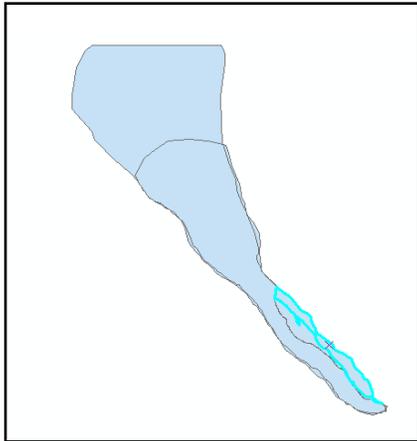


Figure D.1.4. Selected polygon representing the East Aquifer subarea.

b) East Portneuf Basalt Subarea

The Portneuf lava was already defined by its own shapefile (portneuf_lava_mod.shp) which was renamed East_PB.shp.

c) Municipal Aquifer Anomaly Subarea

To begin, a graphics polygon was drawn to define the subgroup of anomalously high nitrate values in the northern part of the municipal aquifer (Figure D.1.1). Using ArcCatalog, a new, empty polygon shapefile was created and named MA_Anomaly.shp. This was opened in ArcMap and an editing session was begun by selecting Editor → Start Editing → selecting the path for MA_Anomaly.shp, Target: MA_Anomaly.shp, Task: Create New Feature. A polygon was then defined for the new shapefile by zooming in over the graphic and using the Editor sketch tool to trace the shape of the graphics polygon. The editing session was then stopped and edits saved.

d) Municipal Aquifer North and South Subareas

A dotted graphics line was drawn over the main aquifer to illustrate the division of the aquifer into the North confined and South unconfined aquifer subareas as shown in Figure D.1.1. To split the aquifer shapefile into North and South regions, an editing session was begun by clicking on Editor → Start Editing → selecting the path for aq_boundsCopy.shp, Target: aq_boundsCopy.shp, Task: Modify features, Cut Polygon Features. Using the Editor pointer, the aq_boundsCopy polygon was selected.

Next, the Editor sketch tool (pencil) was selected and a line was traced over the dotted graphics line. The northern polygon was selected and exported as a shapefile

by right-clicking on the aq_boundsCopy.shp in the table of contents and selecting Data → Export Data and naming the output aq_bounds_N.shp.

The selection was then cleared and the process was repeated by selecting the southern polygon and exporting it as aq_bounds_S.shp.

To complete the Municipal Aquifer North shapefile, it was necessary to remove (cut out) the anomalous area. This was done by combining the two shapefiles (aq_bounds_N.shp and MA_Anomaly.shp) then selecting the anomalous area polygon and deleting it from the attributes as follows:

1. Analysis Tools → Overlay → Union tool; specifying MA_Anomaly.shp.
2. All attributes were then selected and an editing session was begun.
3. The polygons were separated using the Explode Tool on the advanced editing toolbar as shown in Figure D.1.3.
4. Selecting only the northern polygon, it was then saved by Right clicking on the unioned shapefile in the table of contents and selecting Data → Export Data, selected features, same coordinate system as this layers source data, and output: MA_North.shp.

The Municipal Aquifer South shapefile was completed via the Union Tool described above to combine the Aq_Bounds_S.shp, East_Aquifer.shp, and Portneuf_lava_mod.shp. Since only two files may be unioned in a single execution, it was performed twice to combine the three shapefiles. Next, the Explode Tool was used to separate the polygons. Lastly, only the southern polygons were selected and exported to MA_South.shp.

e) **Bedrock and Other Sediments Subarea**

The depth to bedrock raster dtb_fm was used to create the bedrock and the other sediments subareas shapefiles.

The first step was to convert the raster to a polygon shapefile. Our initial attempts at this we received the error “invalid field type; the input is not within the defined domain”. We speculated that the error might be a result of the conversion tool expecting 2 values in the raster rather than 3 values as ours had (0, 1, 100). Since we were only interested at this point in the cells with a value of 0 (bedrock), the cells > 0 were all changed to contain 1's. This was accomplished by reclassifying the raster values by clicking on Spatial Analyst Tools → Reclass → Reclassify (Fig D.1.5):

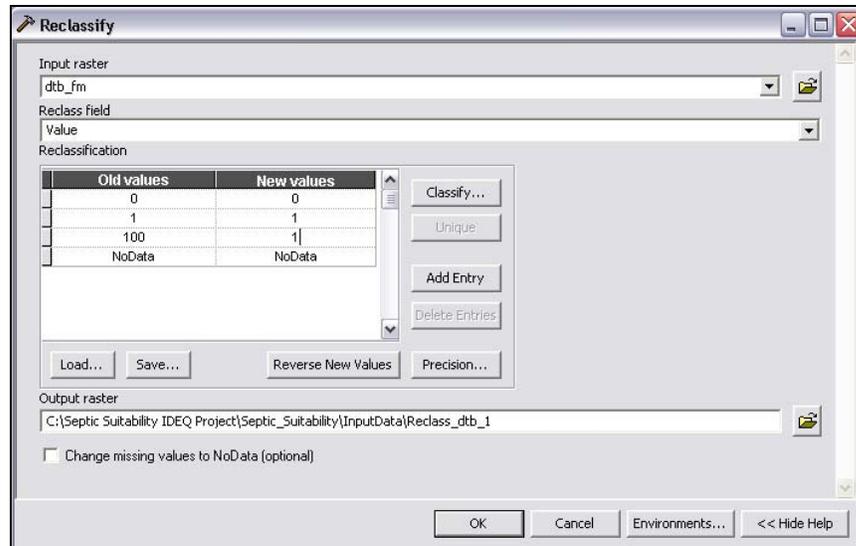


Figure D.1.5. Reclassification of the depth to bedrock raster values.

The raster was then converted to polygons successfully using the Raster to Polygon tool (Figure D.1.6) and selecting dtb_fm as the input raster.

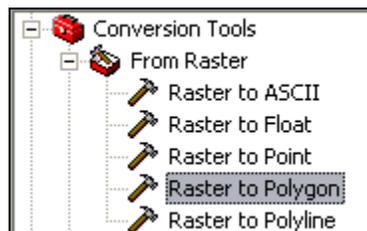


Figure D.1.6. Raster to Polygon tool.

The bedrock shapefile was created by selecting the bedrock polygons from the new depth to bedrock polygon shapefile and exporting them to the Bedrock shapfile.

1. The bedrock polygons were selected by clicking Selection → Select by Attribute on the main toolbar dropdown menu and specifying Layer as the new polygon shapefile, Method as Create New Selection, clicking on “gridcode” from the list of fieldnames, = 0, and OK.
2. Right click on the polygon shapefile in the Table of Contents and select Data → Export Data, selected features, same coordinate system as this layers source data, and output: Bedrock.shp

After clearing the selection, the Other Sediments shapefile was created by repeating steps 1 and 2 above but selecting “gridcode”=1 for other sediments in step 1 and exporting it as Other.shp in step 2.

Because the Other Sediments Subarea shapefile was reclassified (Figure D.1.5), it was necessary to remove (cut out) the Aquifer Boundary area to recover the subarea

representing minimal overburden thickness. This was done by combining the two shapefiles (aq_boundsCopy.shp and Other.shp) then selecting the aquifer boundary area polygon and deleting it from the attributes as follows:

1. Analysis Tools → Overlay → Union tool; specifying output Other_Sed.shp.
2. All attributes in Other_Sed.shp were then selected and an editing session was begun.
3. The polygons were separated using the Explode Tool on the advanced editing toolbar as shown in Figure D.1.3.
4. Selecting only the aquifer boundary polygon, it was then deleted using the delete key.
5. The editing session was stopped and all edits to Other_Sed.shp were saved.

2) Segregate the Nitrate Data into Subarea Groups

All nitrate concentration data were plotted on the Hydrogeologic Influences layer then segregated into seven subsets corresponding to the seven polygons created above. Instructions are given here for only one of the seven subarea polygons, the East Aquifer subarea (Figure D.2.2a).

First, from the Selection menu, choose Select by Location to display the dialog box shown in Figure D.2.1

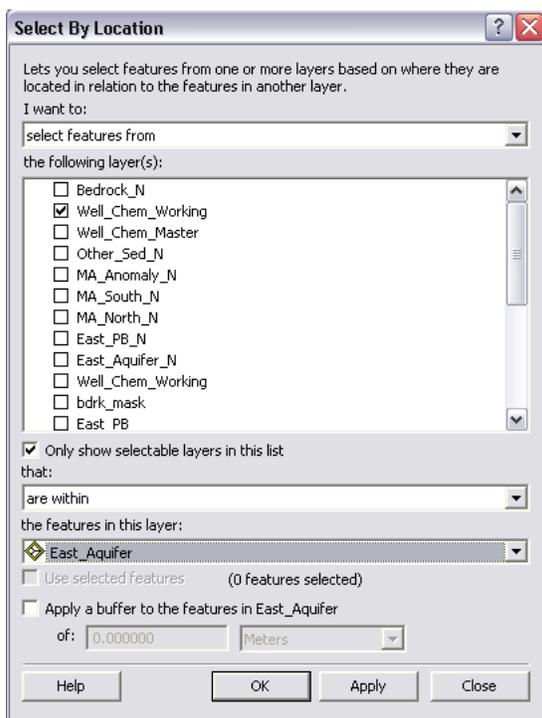


Figure D.2.1. Select By Location dialog box.

- specify the “select features from” option in the top drop-down box
- check the appropriate point data layer (“Well_Chem_Working” in this case)
- choose the selection option “are within” from the second drop-down box
- select the appropriate polygon layer from the third drop down menu
- click Apply, then OK

In the table of contents, right click on “Well_Chem_Working.shp” and select **Data | Export Data**. In the resulting dialog box:

- Specify “Selected Features” and the same coordinate system as “this layer’s data source”
- Enter an output filename such as “East_Aquifer_N”
- Click OK

The resulting shapefile contains point data only for wells within the East Aquifer subarea (Figure D.2.2b).

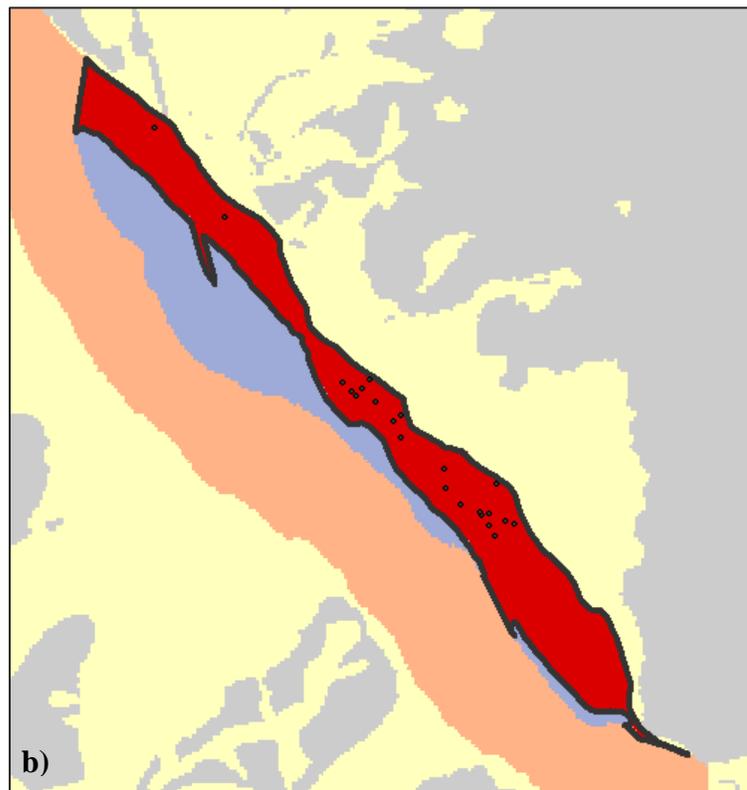
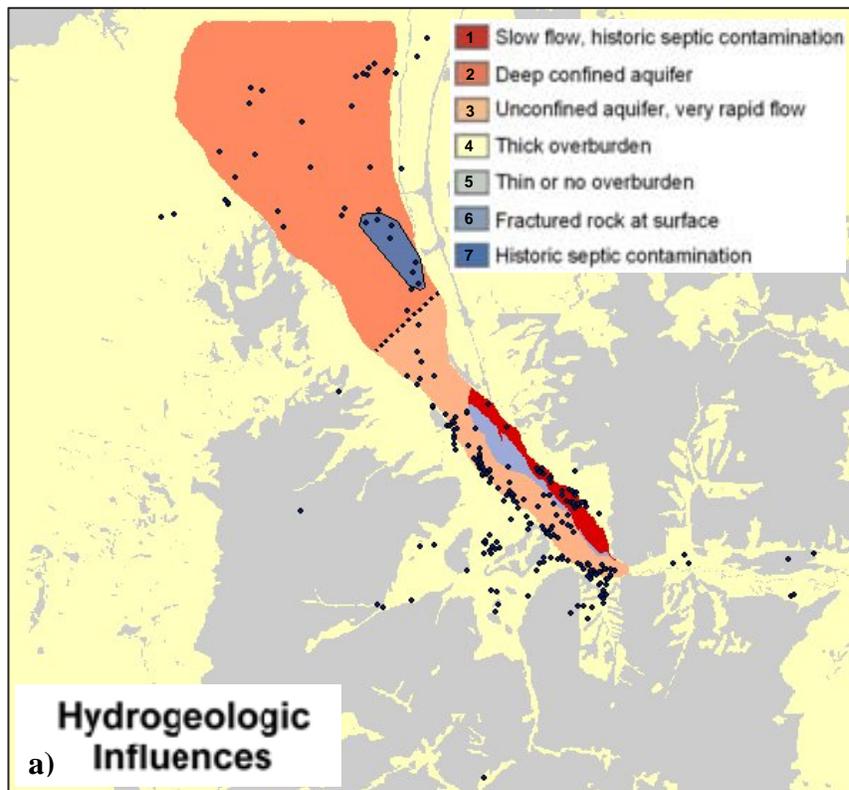


Figure D.2.2. a) Hydrogeologic Influences layer with seven arbitrarily defined subareas (colored polygons). Wells with chemistry data are shown as filled circles. B) Closeup of a single subarea (the "East_Aquifer" polygon, in red) and the end result of applying the procedure outlined in this appendix to isolate well chemistry data that correspond to that polygon.

3) Statistical Evaluation of Nitrate Concentrations Within Each Subarea

Once individual nitrate shapefiles were created for all subareas, a spreadsheet containing nitrate data for each subarea was created by opening its nitrate shapefile's .dbf table in Excel and saving the data in spreadsheet form. The spreadsheet was imported into a statistical analysis software package like Minitab and the nitrate data values were plotted as histograms and box plots for each individual subarea (Figure D.3.1).

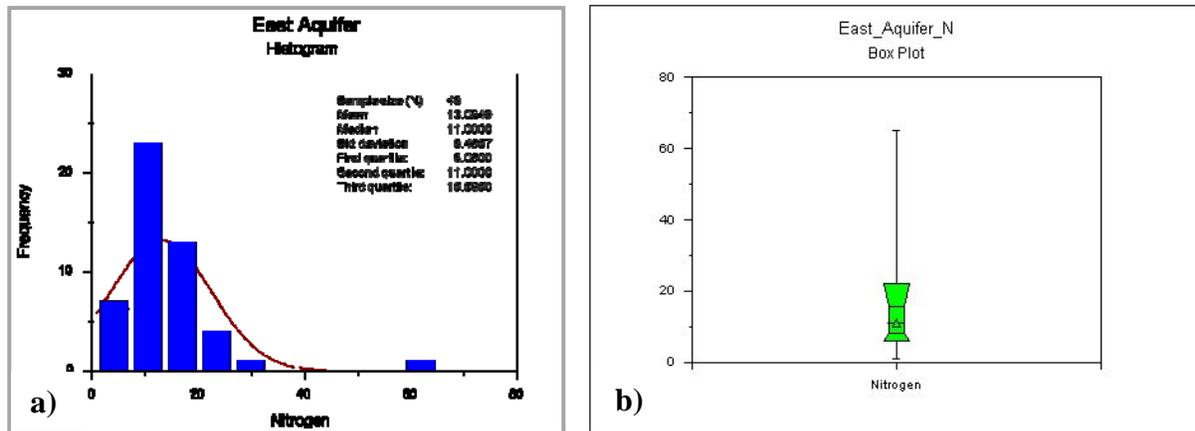


Figure D.3.1. Subarea nitrate statistics. Histogram a) and box plot b) of nitrate concentrations in the Eastern Aquifer subarea show that out of 49 samples, 50% had concentrations ranging from approximately 8-15.6 mg/l with a median value of 11 mg/l.

A series of non-parametric tests of difference were then run to identify which groups have median nitrate levels that are statistically indistinguishable and those whose medians are different at a specified confidence level. In this study, a two-sample Wilcoxon signed-rank test was employed, but a Mann-Whitney test would also be appropriate. Multiple two-sample tests were run to compare the median nitrate concentrations between all subareas; results for five of the seven subareas are summarized in Figure D.3.2.

At a 99% confidence level ($p > 0.01$), the median nitrate levels in three subareas are statistically indistinguishable, so there is no defensible basis for continuing to treat them as separate subareas. Therefore, these four subareas (and their nitrate data) were regrouped into two larger subareas comprising (i) the North main aquifer (confined) and Bedrock (thin or no soil) areas, and (ii) the South main aquifer (unconfined) and Other Sedimentary (thick soil) areas. Following regrouping of the corresponding nitrate data, the Wilcoxon test was repeated on the new data groups. This process of statistical comparison and regrouping was repeated until all comparisons among subareas resulted in p -values < 0.05 , indicating that the resulting regrouped subareas' median nitrate levels are all significantly different at 95% confidence (Figure D.3.3). The median nitrate concentrations of these final groupings provides the basis for ranking the data layers' suitability for septic siting.

In Figure D.4, a rank of 1 (least suitable/highest risk) was assigned to the subarea with the highest median nitrate concentration (East subarea); incrementally higher ranks for the other subareas indicate a progressive increase in suitability. In this case

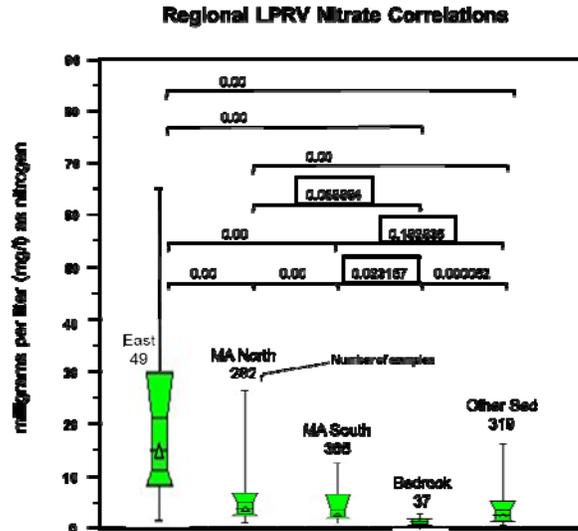


Figure D.3.2. Comparison of statistical distributions among subareas. P-values > 0.01 (boxed) indicate no difference in nitrate concentrations at 99% confidence. Data for such subareas was subsequently regrouped and treated as a single subarea for further testing.

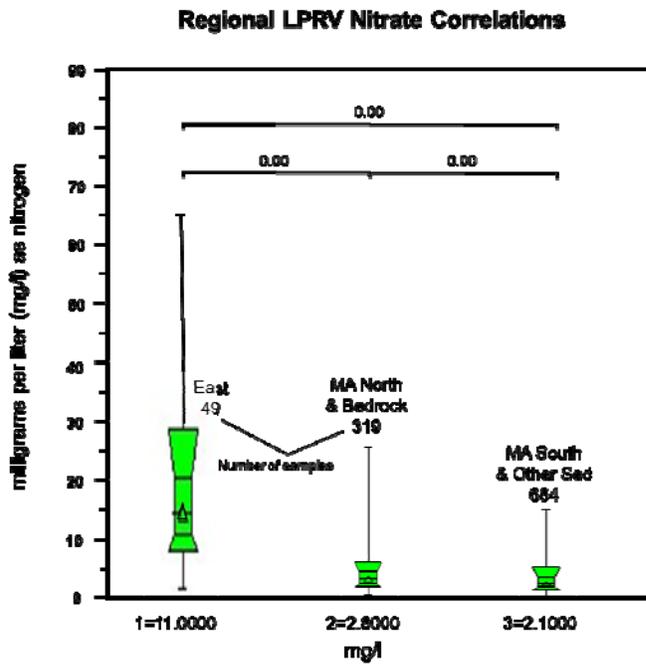


Figure D.3.3. Final results of statistical comparisons. Resulting p-values of < 0.01 indicate real statistical differences in nitrate concentrations between these groups. The three areas are ordinaly ranked according to their median nitrate concentrations.

4) Create Final Raster with Suitability Rankings

Creation of the final ordinally ranked data layer required several steps to combine the individual subarea polygons that remain following statistical calibration. The following steps describe the detailed process.

Combining the seven original subarea polygon shapefiles to create a single, new polygon shapefile was accomplished by 1) merging the 5 aquifer subarea polygon shapefiles, 2) converting it to raster, 3) combining it with the existing Depth to Bedrock raster (dtb_fm), 4) assigning ordinal ranks to the combined raster, and then 5) clipping and projecting the final raster. (Initially, all 7 polygon shapefiles were merged and converted to raster but the resulting raster contained random “no data” cells within the bedrock and other sediments areas. Various methods were tried to resolve this with all failing except the method described below.)

i. Merge the five Aquifer Subarea Shapefiles

Begin by adding two new fields to each of the subarea polygon shapefiles; 1) SubArea, string, length of 5 and 2) Rank, numeric. Populate the fields by opening each attribute table with data as shown in Table D.4.1. Start an editing session, select the entire SubArea column, click on the options button and select “Find and Replace”. Specify to find “”, whole field, and replace with the associated value for the subarea then click OK. Repeat this process for the Rank field. Stop and Save edits.

Table D4.1. New data fields were added to the shapefile’s attribute tables and populated with new subarea names and corresponding ordinal ranks.

Aquifer Subareas	East Aquifer	Main Aquifer North (Confined)	Main Aquifer South (Unconfined)	East PB	Main Aquifer Anomaly
Median N	11	3	2.2	-	-
Rank	1	2	3	6	7
Subarea	EA	MAN	MAS	EPB	MAA

The five shapefiles were merged as shown in Figure D.4.1.i (Note: during this process, it is important to remove (delete) all fields except subarea and rank in the field map section after the addition of each input dataset).

Data Management → General → Merge

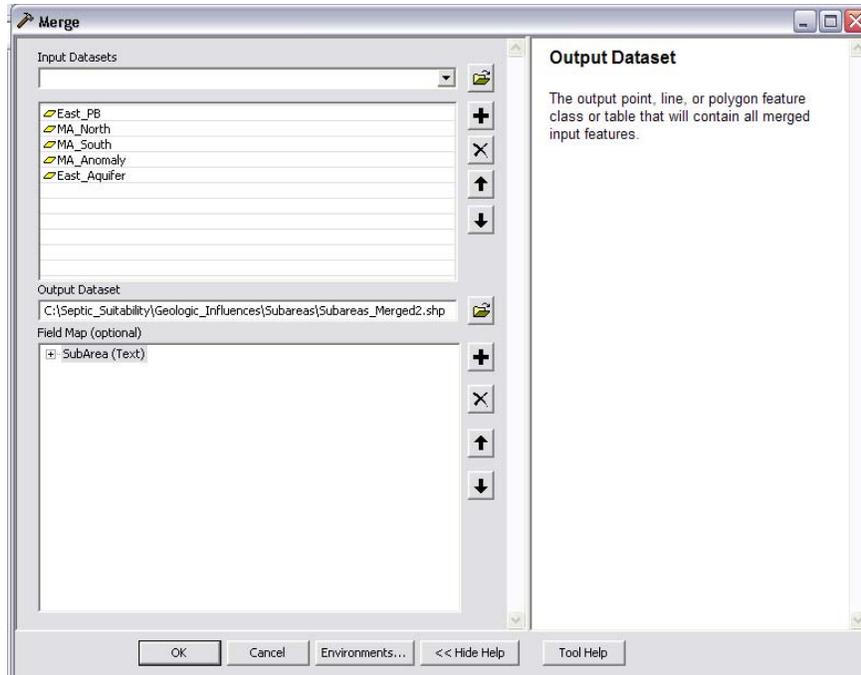


Figure D.4.1. Merging aquifer subarea shapefiles.

ii. **Convert the Merged Aquifer Shapefile to Raster**

The merged shapefile was converted to a raster based on the rank field as shown below (Figure D.4.2) with care taken to specify the environment (Figure D.4.3).

Conversion Tools => Polygon to Raster =>

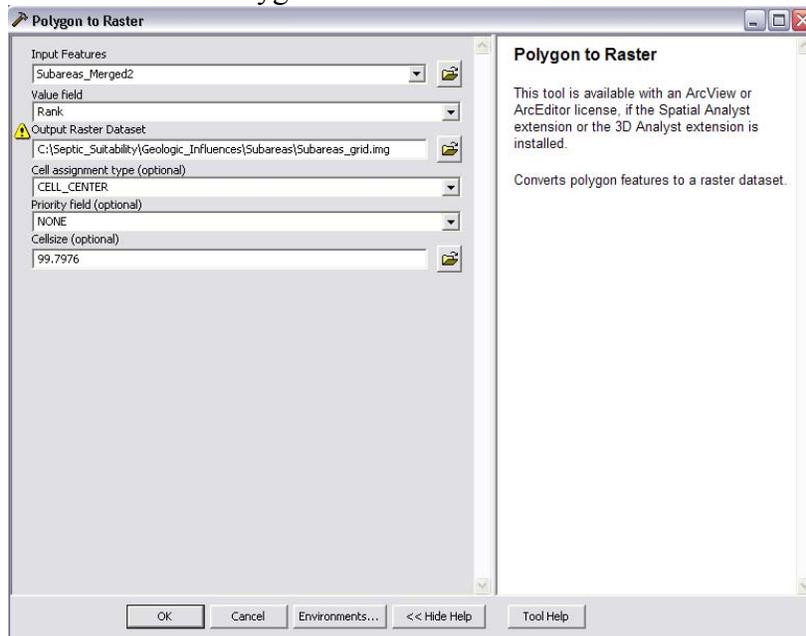


Figure D.4.2. Converting merged shapefiles to raster.

Now click on the Environments button at the bottom and set the extent to be the same extent as the Depth to Bedrock raster (Figure D.4.3).

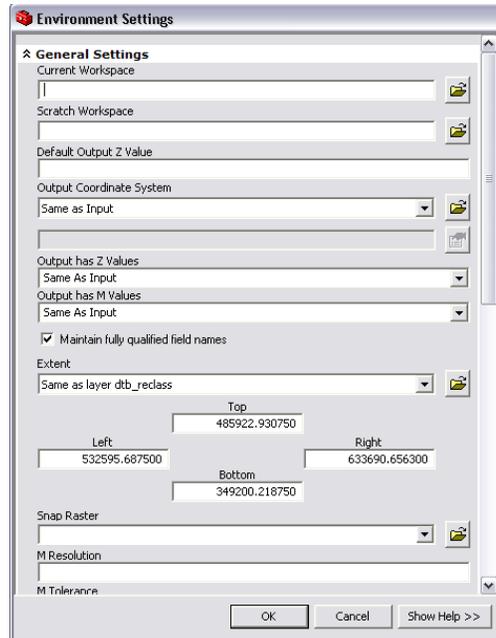


Figure D.4.3. Converting merged shapefiles to raster.

Next, reclassify the new subareas raster to set “no data” cells to a value of zero (Figure D.4.4) so the full extent of both rasters may be added together properly in subsequent steps.

Spatial Analyst Tools → Reclass → Reclassify

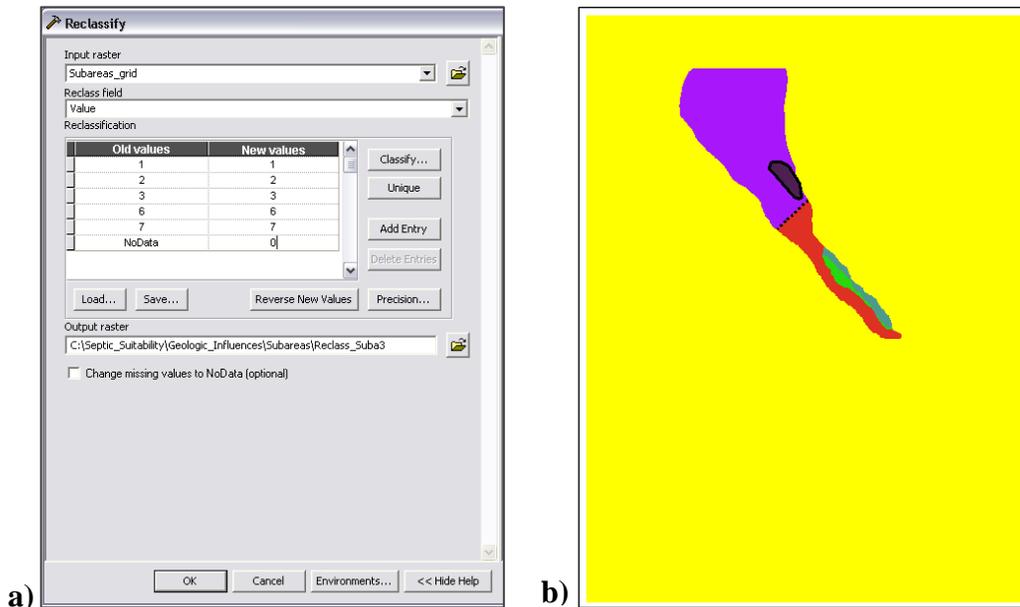


Figure D.4.4. Reclassify the raster: a) Assigning “no data” cells with a value of zero; b) resulting raster of the five merged aquifer subareas (yellow = no data).

iii. Combine the Rasters

In preparation for combining the Depth to Bedrock raster (dtb_fm) with the new aquifer subareas raster, a working copy was created by copying the dtb_fm raster, renaming it dtb_fmcopy, and importing symbology from the original file.

Next, the working copy was reclassified (Figure D.4.5) to replace the value of 0, 1, and 100 with values of 10 (bedrock), 20 (other sediments), 30 (region where the other 5 aquifer subareas will occupy). This was done to make it easier to understand how the values in the summed raster were obtained in the raster so we can correct them if needed.

Spatial Analyst Tools → Reclass → Reclassify

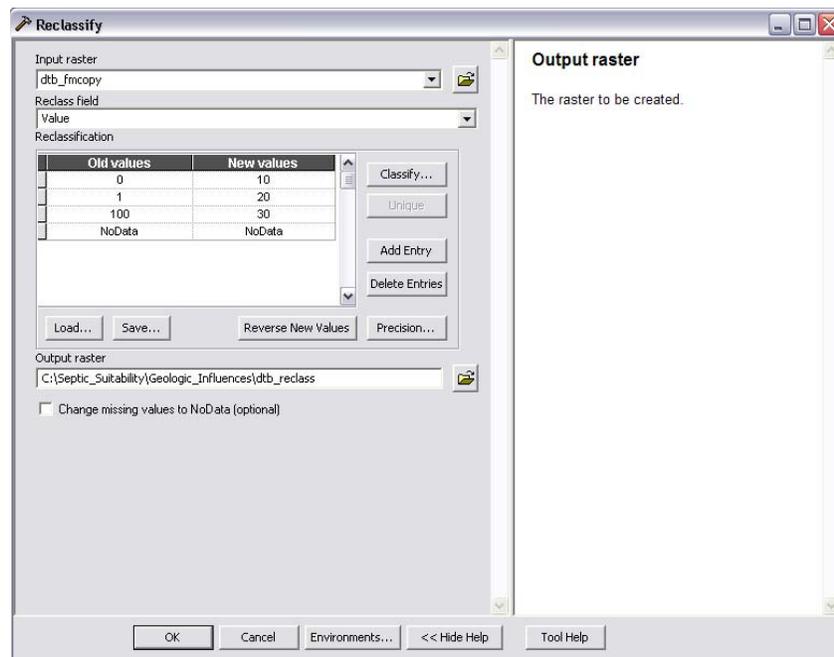


Figure D.4.5. Reclassify the Depth To Bedrock raster for tracking purposes.

The rasters were then combined using the the Spatial Analyst Math tool (Figure D.4.6).

Spatial Analyst Tools → Math → Plus

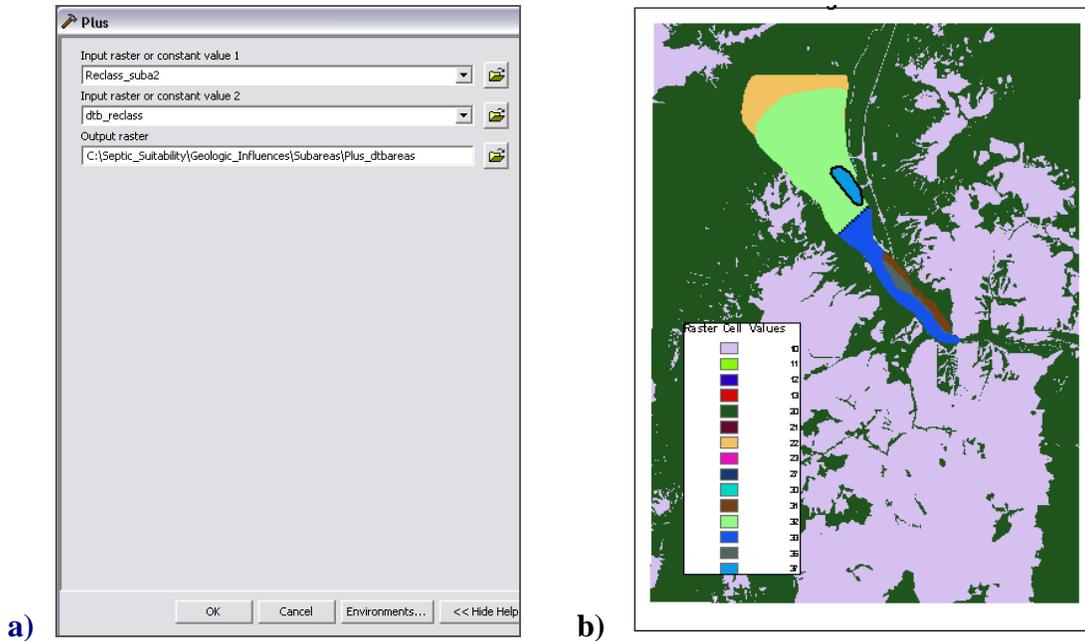


Figure D.4.6. Adding the rasters: a) the grid values of the two rasters were summed on a cell-by-cell basis using Spatial Analyst’s “Plus” tool; b) the resulting hydrogeologic raster.

5) Assign Ordinal Ranks

Grid values of the combined raster were evaluated and reclassified to reflect the ordinal ranking for the seven subareas as determined by the statistical analysis (Figure D.5.1 and Tables D.5.1 and D.5.2). The East_PB, East_Aquifer, and Anomaly subareas were all ranked as unsuitable because the argument that the high N values in the anomaly area indicate unsuitable conditions is tempered by the knowledge that hydro-geologically, it is uncertain whether the nitrogen in this area originates from the surface or from another area outside the classified anomaly. For the sake of simplicity, it was classified with the same suitability as the East_PB and East_Aquifer class because they are all unsuitable.

Table D.5.1. Assign ordinal suitability ranks to subareas based on their median nitrate concentration, where a value of 1 is assigned to the subareas with the highest nitrate-N (= least suitable) and higher ranks indicating greater suitability (lower nitrate).

	East Aquifer	East_PB	Anomaly	Main Aquifer North (Confined) & Bedrock	Main Aquifer South (Unconfined) & Other Sediments
Median N	11	--	--	2.8	2.1
Rank	1	1	1	3	2

Table D.5.2. Grid values of the combined raster. Cell values resulting from the combination of the two rasters, origin of derived values, and assigned ordinal ranking.

Cell Value	How Derived	How Derived	Reclassify As
10	10 or 10 + 0	Bedrock + no data	2
11	10 + 1	Bedrock + East	1
12	10 + 2	Bedrock + North	2
13	10 + 3	Bedrock + South	3
20	20 or 20 + 0	Other Sed + no data	3
21	20 + 1	Other Sed + East	1
22	20 + 2	Other Sed + North	2
23	20 + 3	Other Sed + South	3
27	20 + 7	Other Sed + Anomaly	1
30	30 or 30 + 0	Aquifer + no data	3
31	30 + 1	Aquifer + East	1
32	30 + 2	Aquifer + North	2
33	30 + 3	Aquifer + South	3
36	30 + 6	Aquifer + EPB	1
37	30 + 7	Aquifer + Anomaly	1

Spatial Analyst Tools → Reclass → Reclassify

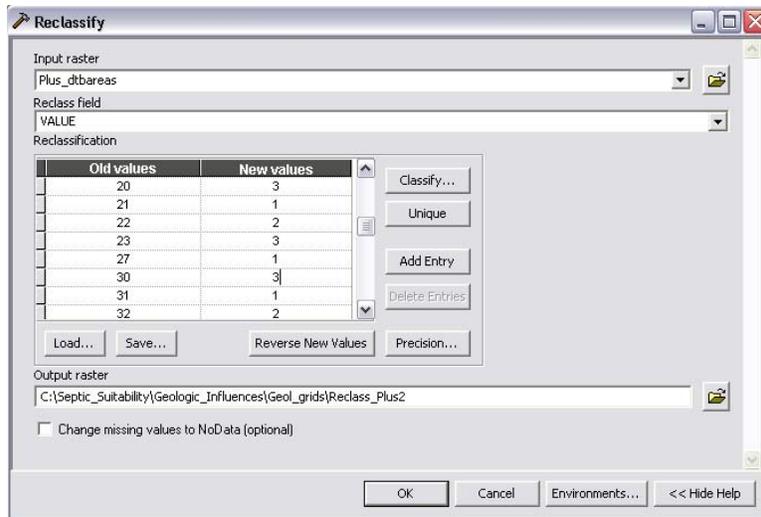


Figure D.5.1. Reclassify combined raster with ordinal ranks. A value of 1 indicates low suitability whereas; a value of 3 indicates high suitability.

6) Finalize the Raster

To ensure uniformity amongst the various DRASTIC layers, this raster was clipped (extract by mask) to the bdrk_mask extents and projected again with the IDTM projection and 30m cell size (Figure D.6.1) resulting in the final hydrogeological raster as shown in Figure D.6.2.

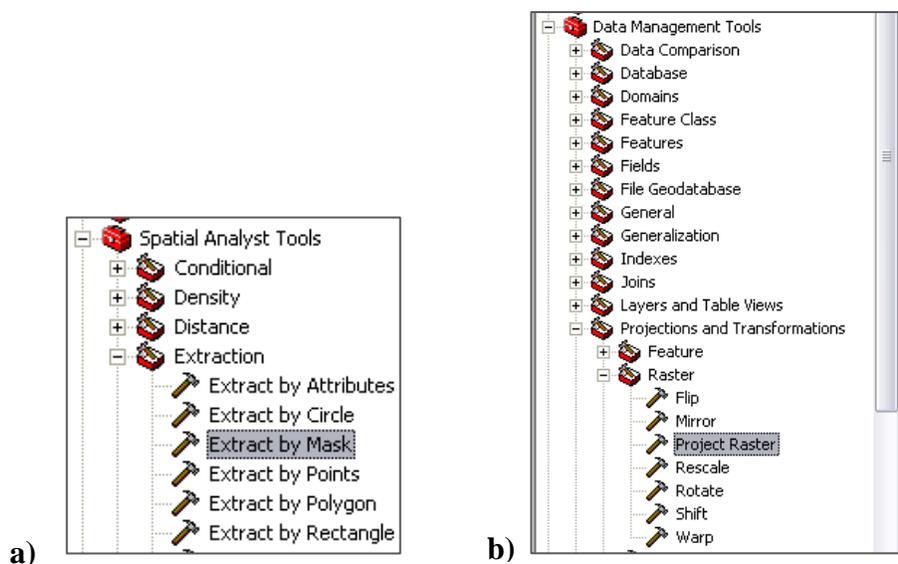


Figure D.6.1. Finalizing the raster. a) The Extract by Mask tool was used to clip the final raster to the same extents as the other layers. b) The Project Raster tool was used to project the final raster to that of the other layers.

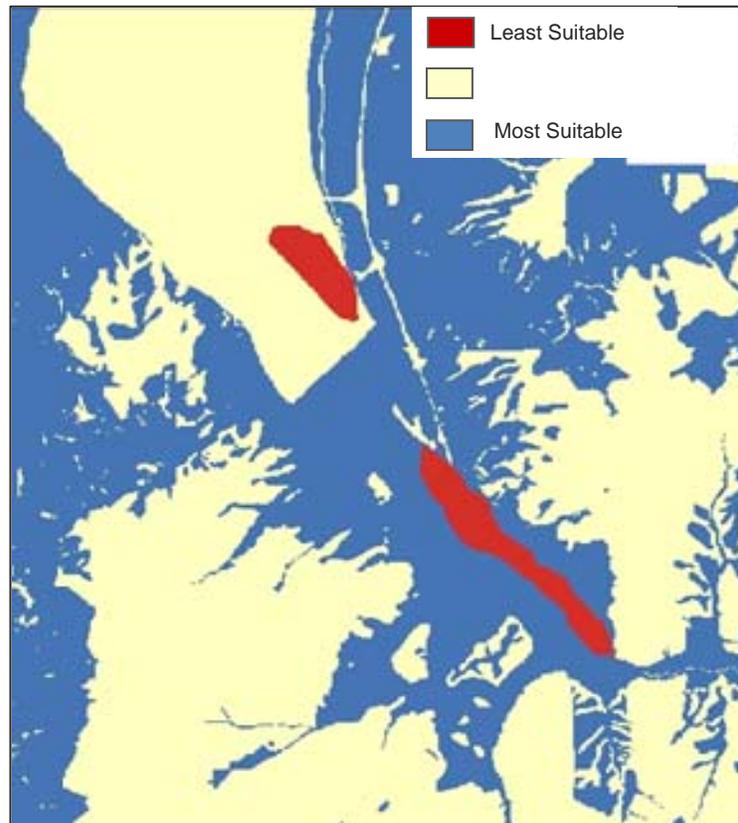


Figure D.6.2. Final ranked categories of the Hydrogeologic Influences information layer. The layer now contains three ordinal classes that are ranked according to the statistically based correlation with nitrate contamination that indicates these subareas' relative suitability for septic-based development.

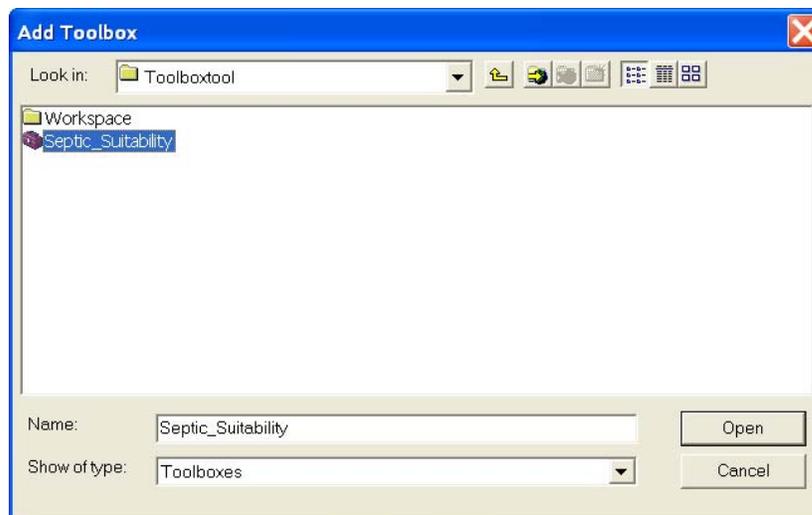
Appendix E

ArcToolbox Script for Permutation / Sensitivity Analysis

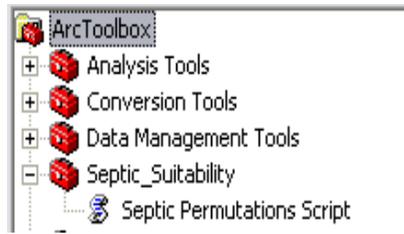
To facilitate the analysis of the suitability rating map's sensitivity to the choice of layer weights, a Python script was created (T. Dunsford, written comm., 2010) to generate any number of permutations of the suitability map using many different layer weights. These values were drawn at random from specified normal distributions of possible layer weights centered around the adjusted layer weights determined in Table 4 and with specified standard deviations about those values. The script is listed at the end of this appendix and included on the CD-ROM that accompanies this report; prior to using the script, ensure that the toolbox file path (highlighted in red in line 16 of the script) is correct for your ArcMap installation.

To install the script

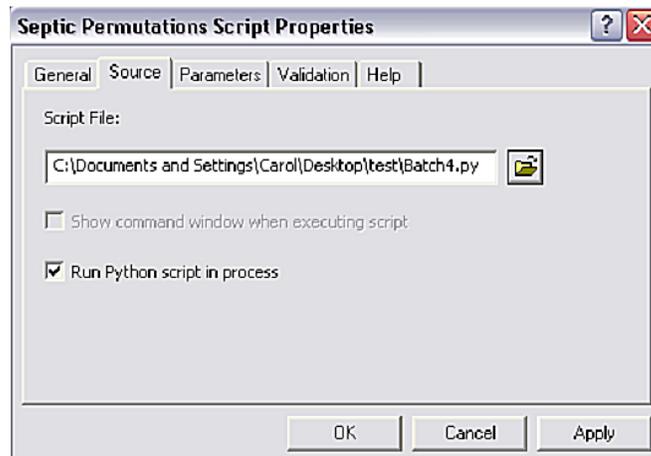
1. Set up a workspace folder called "\\Workspace" and create three empty subfolders in it named \\Randomized, \\FinalPermutations, and \\CellStats. Copy the files "septic_suitability.tbx" and "Batch4.py" into \\Workspace.
2. Open the ArcToolbox window in ArcMap and right-click anywhere in the blank space below the listed tools. Select "Add Toolbox" and point to its location:



3. Click Open Double-click on the Septic Suitability toolbox.in the ArcToolbox window:



4. If the script is not already present (as shown above), go to Step 5. Otherwise, right-click on it and select Properties, then click on the Source tab as shown below. Browse to the appropriate location of the script file (Batch4.py) and click Okay:

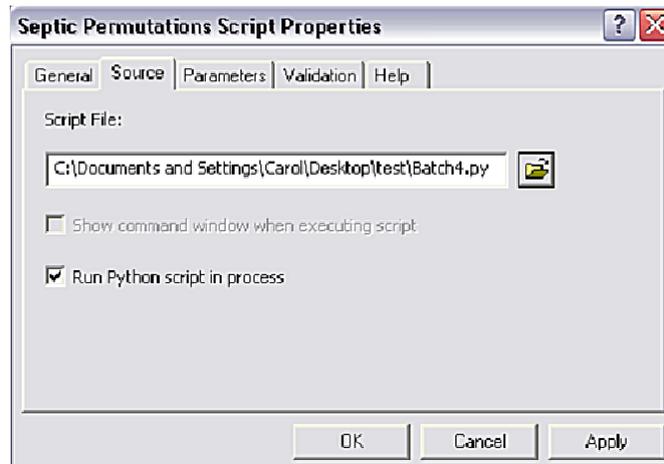


The tool is now ready to use..

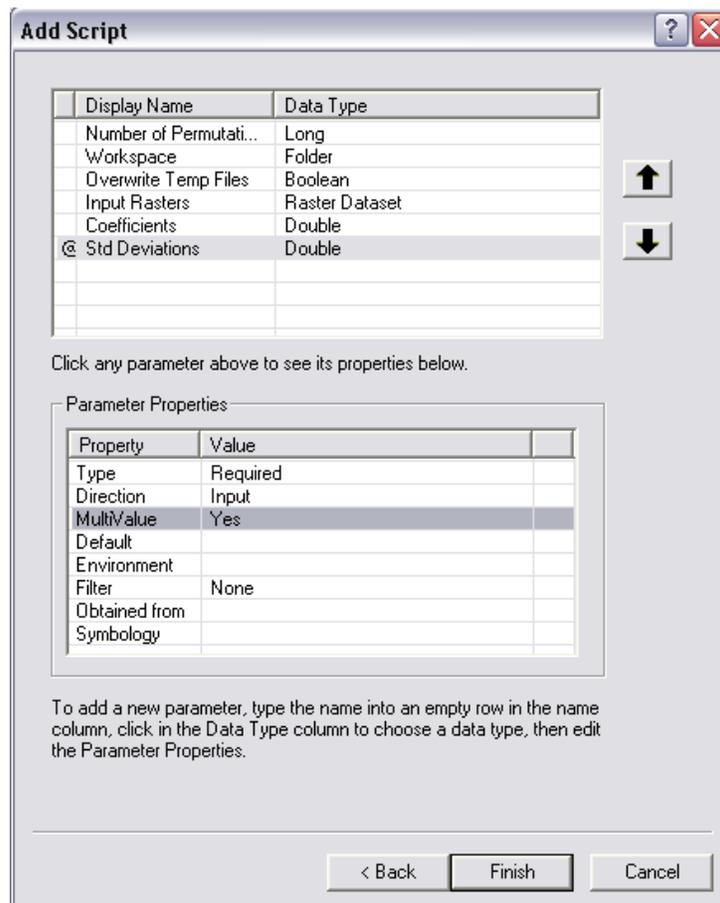
5. If the script is not already showing in the toolbox then right click on the Septic_Suitability toolbox, point to Add, and click on Script. Enter data in the dialog boxes as shown below:



Click Next and browse to the script file. Click Next.

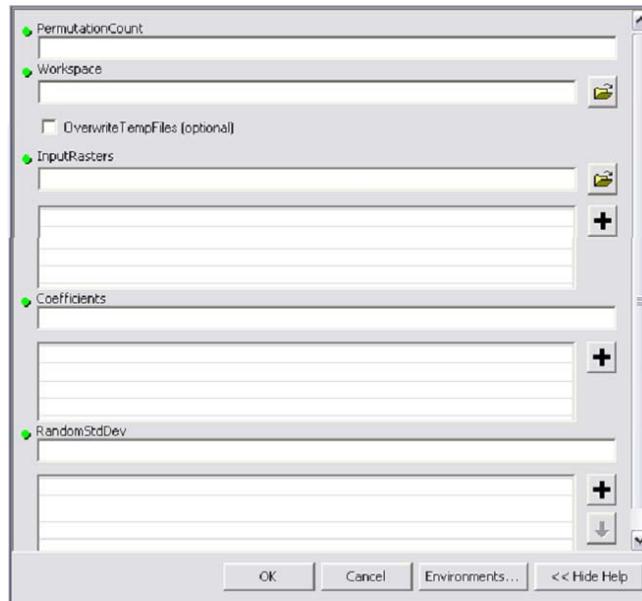


As this is a new tool, it will need to have its parameters set up. Add parameters as shown below, taking care to specify the parameter properties for each of Input Rasters, Coefficients, and Std Deviations as Multivalue=yes (all others “no”) and Overwrite Temp Files as Type=optional. (all others “required”). Click Finish. The script will be added to the new toolbox

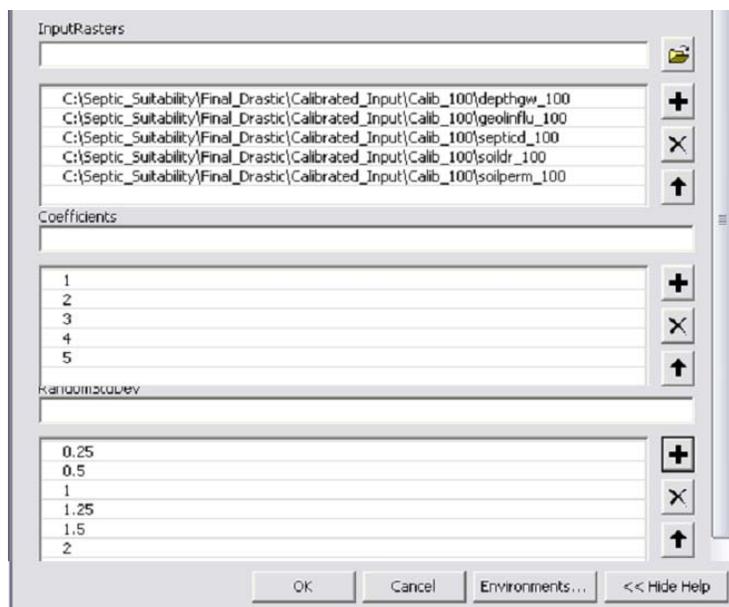


To run the script

Double-click on the script in ArcToolbox. The following window will be displayed:



Ignore the Overwrite TempFiles option in the dialog window (it is not an option). For the remaining inputs, enter the required information and click Okay to run the script. **It is critically important to maintain the same number of, and the same input order for, the input rasters, their layer weights ("coefficients"), and their standard deviations.** For example:



Note: The above run will fail because too many std. deviation values have been specified.

The outputs are written to the \Workspace\CellStats subfolder (Mean, Median, Stddev and Range rasters, which summarize the statistical variability of the permutations that were created) and the \Workspace\FinalPermutations subfolder (containing all individual permutation rasters that were created). Normally, these will be the only files of interest (the \Workspace\Randomized subfolder is a scratch folder that contains only intermediate results) Following a run, the contents of the first two subfolders should be copied in their entirety to a results folder, otherwise the results will be overwritten by the next run (if copying individual raster files, do so only with ArcCatalog)

Some helpful hints

- Any number of permutations can be specified, but execution time can exceed an hour
- Combining input rasters with different cell sizes is not good GIS practice and may create unexpected results
- The extents of the output rasters will be that of the smallest input raster, so input rasters should be roughly of the same x,y size
- The standard deviations can be any size relative to the specified layer weights, but if they are too large, many of the randomly generated layer weights will be negative; although these will be ignored, too many will bias the output statistics, possibly skewing the output distribution and causing problems with the classification procedure described in Appendix F.
- In most cases, standard deviations of 1/4 to 1/3 of the layer weight will produce a sufficient range of permutations without skewing the output statistics.

Script documentation

Note: ensure that the toolbox file path (highlighted in red, below) is correct for your ArcMap installation

```
# Batch4.py
# Created on: Thu Jan 14 2010 02:56:30 PM
# (generated by ArcGIS/ModelBuilder)
# Usage: Batch2 <DepthGW_100> <Map_Algebra_expression> <Map_Algebra_expression_2_>
<Map_Algebra_expression_3_> <GeolInflu_100> <Map_Algebra_expression_4_>
<Map_Algebra_expression_5_> <Map_Algebra_expression_6_> <Map_Algebra_expression_7_>
<septic_d100> <SoilDR_100> <soilperm_100> <Map_Algebra_expression_8_> <Map_Algebra_expression_9_>
<Map_Algebra_expression_10_>
# Import system modules
import sys, string, os, arcgisscripting, random
# Create the Geoprocessor object
gp = arcgisscripting.create()
# Check out any necessary licenses
gp.CheckOutExtension("spatial")
# Load required toolboxes...
gp.AddToolbox("C:/Program Files/ArcGIS/ArcToolbox/Toolboxes/Spatial Analyst Tools.tbx")
# Set the Geoprocessing environment...
gp.XYResolution = ""
gp.scratchWorkspace = "C:\\Septic_Suitability\\Stats_Experiment5\\Stats5_Intermed"
gp.MTolerance = ""
gp.randomGenerator = "0 ACM599"
gp.outputCoordinateSystem = ""
gp.outputZFlag = "Same As Input"
gp.qualifiedFieldNames = "true"
gp.extent = "DEFAULT"
gp.XYTolerance = ""
gp.outputZValue = ""
gp.outputMFlag = "Same As Input"
gp.geographicTransformations = ""
gp.ZResolution = ""
gp.workspace = "C:\\program files\\program files\\program files\\program files\\Septic_Suitability"
gp.MResolution = ""
gp.ZTolerance = ""
workspace = sys.argv[2]
# Local variables...
RandomizedFolder = workspace + "\\Randomized\\"
FinalPermutationsFolder = workspace + "\\FinalPermutations\\"
CellStatsFolder = workspace + "\\CellStats\\"
Mean_Stats = workspace + "\\CellStats\\Mean"
Median_Stats = workspace + "\\CellStats\\MEDIAN"
StdDev_Stats = workspace + "\\CellStats\\StdDev"
Range_Stats = workspace + "\\CellStats\\Range"
# Script arguments...
# Splits the string at the semicolons
rasterpaths = sys.argv[4].split(";")
# Create an empty list to store the truncated filenames
rasterfiles = []
# Start the index at 0 since the list is 0 based index
i = 0;
# Loop through the list of full path names until you reach the length of that list
while(i < len(rasterpaths)):
    # Add the truncated filename to the rasterfiles list
```

```

rasterfiles.append(os.path.basename(rasterpaths[i]))
# increment i to the next member
i = i + 1
coeffs = sys.argv[5].split(";")
dispersions = sys.argv[6].split(";")
from random import Random
g = Random()
permut = 0
StatInputs = ""
numPermut = int(sys.argv[1])
gp.AddMessage("total permutations: " + str(numPermut))
while(permut < numPermut):
    i = 0
    gp.AddMessage("Permutation: "+ str(permut+1) + "/" + str(numPermut))
    sumExpression = ""
    while(i < len(rasterpaths)):
        rand = g.gauss(0, float(dispersions[i]))
        exp = str(float(coeffs[i]) + rand)
        while(exp < 0):
            rand = g.gauss(0, float(dispersions[i]))
            exp = str(float(coeffs[i]) + rand)
        gp.AddMessage(exp)
        RandomExpression = rasterpaths[i] + " * " + exp
        gp.AddMessage(RandomExpression)
        outfile = RandomizedFolder + rasterfiles[i]
        if not sys.argv[3]:
            outfile = outfile + str(permut)
        gp.SingleOutputMapAlgebra_sa(RandomExpression, outfile)
        if i == 0:
            sumExpression = outfile
        else:
            sumExpression = sumExpression + " + " + outfile
        i = i + 1
    gp.AddMessage("Sum of rasters")
    gp.SingleOutputMapAlgebra_sa(sumExpression, FinalPermutationsFolder + "Permut" + str(permut))
    if permut > 0:
        StatInputs = StatInputs + ";" + "" + FinalPermutationsFolder + "Permut" + str(permut) + ""
    else:
        StatInputs = "" + FinalPermutationsFolder + "Permut" + str(permut) + ""
    permut = permut + 1
# Process: Mean...
gp.AddMessage("Mean")
gp.CellStatistics_sa(StatInputs, Mean_Stats, "MEAN")
# Process: Median...
gp.AddMessage("Median")
gp.CellStatistics_sa(StatInputs, Median_Stats, "MEDIAN")
# Process: Std Dev...
gp.AddMessage("STD")
gp.CellStatistics_sa(StatInputs, StdDev_Stats, "STD")
# Process: Range...
gp.AddMessage("Range")
gp.CellStatistics_sa(StatInputs, Range_Stats, "RANGE")

```

Appendix F

A New Objective Classification Procedure Based on Sensitivity Analysis Results

Any map of DRASTIC suitability index ratings can be classified into areas that are suitable and unsuitable for septic development by selecting an arbitrary classification threshold (for example, the median index rating, the 75th percentile of index rating values, etc.). However, the resulting classification is then entirely subjective, being wholly dependent on the selected value of the classification threshold. The problem can be avoided if the classification threshold is selected on the basis of an objective criterion that makes sense in the context of the classification problem. The sensitivity analysis described in Section 3.4.1.1 yields a statistical measure that provides such a criterion: the coefficient of variation of the suitability ratings, a measure of classification uncertainty, can be used to determine an entirely objective and defensible classification threshold based on a level of classification uncertainty that is deemed acceptable. The DRASTIC rating map can then be objectively classified into Suitable and Unsuitable categories, as well as a category in which the classification is deemed too uncertain to be acceptable.

The following methodology expands on the information outlined in Section 3.4.1.2 and provides additional insight into the method's rationale. It is presented as a series of step-wise decisions based on statistical confidence criteria:

(1) Specify the level at which the coefficient of variation (CV) derived from the sensitivity analysis is deemed too high to be acceptable. For example, if the 90th percentile is chosen, then this will be the confidence level of the final suitability classification (e.g., 90% confidence).

(2) Classify the CV raster into 10 uncertainty quantiles using ArcMap's classification dialog (Layer Properties | Symbology | Classification Method = Quantile, with 10 classes). Note the threshold value, U_{90} , that corresponds to the 90th percentile. In Raster Calculator, classify the CV raster around this value of U_{90} so that binary values of "0" correspond to areas that are "certain" and values of "1" correspond to areas that are "uncertain". For example, consider the Raster Calculator expression "[coeffvar] > 0.158899218", where 0.158899218 is the 90th percentile of the CV raster's values. This expression assigns a value of "1" (uncertain) to 10% of all cells whose CV values exceed this threshold.

(3) Classify the MEANs raster values derived from the sensitivity analysis (Section 3.4.1.1) into 20 suitability quantiles (Classification Method = Quantile, with 20 classes) and then classify this raster into several suitable/unsuitable maps each around a different suitability threshold. For example, on the suitability quantiles raster, first note the threshold value that corresponds to the $X_i = 80$ th percentile and classify the MEANs raster into suitable and unsuitable areas around this threshold. If 156.7529364 is the 80th percentile value in the MEANs raster (Figure 17), then the Raster Calculator expression "[means] < 156.7529364" classifies 80% of the map area as "unsuitable" and assigns a value of "1" to those cells. Repeat this for several lower X_i (e.g., $i = 40, 50, 60, 70$) and save each new binary raster as "mean_Xi" .

(4) Multiply the corresponding "CV_U_i" raster (e.g., CV_U₉₀) created in step (2) with each "mean_X_i" binary raster created in step (3). The resulting set of output rasters identifies areas of cell-wise overlap between "unsuitable" and "uncertain" zones at several different suitability classification thresholds. Open the attribute table of each output raster and record the COUNT value; this represents the number of overlapping cells that are classified as both "unsuitable" and "uncertain"). Plot values of COUNT vs. X. From this graph estimate the value, X₀, at which COUNT drops to zero. For example, the calibrated ratings map created in Section 3.4.1.1 has a classification threshold at the X₀ = 44th percentile of the ratings distribution (**Figure F.1a**). This is the minimum ratings threshold at which areas classified as "unsuitable" and "uncertain" do not overlap at this confidence level. Note that the same threshold applies at confidence levels down to at least the 60th CV percentile (**Figures F.1b, c**). At and below the 50th percentile, however, areas classified as "unsuitable" and "uncertain" always overlap and X₀ is undefined (**Figure F.1d**).

(5) If classification thresholds are undefined for all X_i, repeat steps (1) to (4) at a lower U_i level (e.g., 75th percentile).

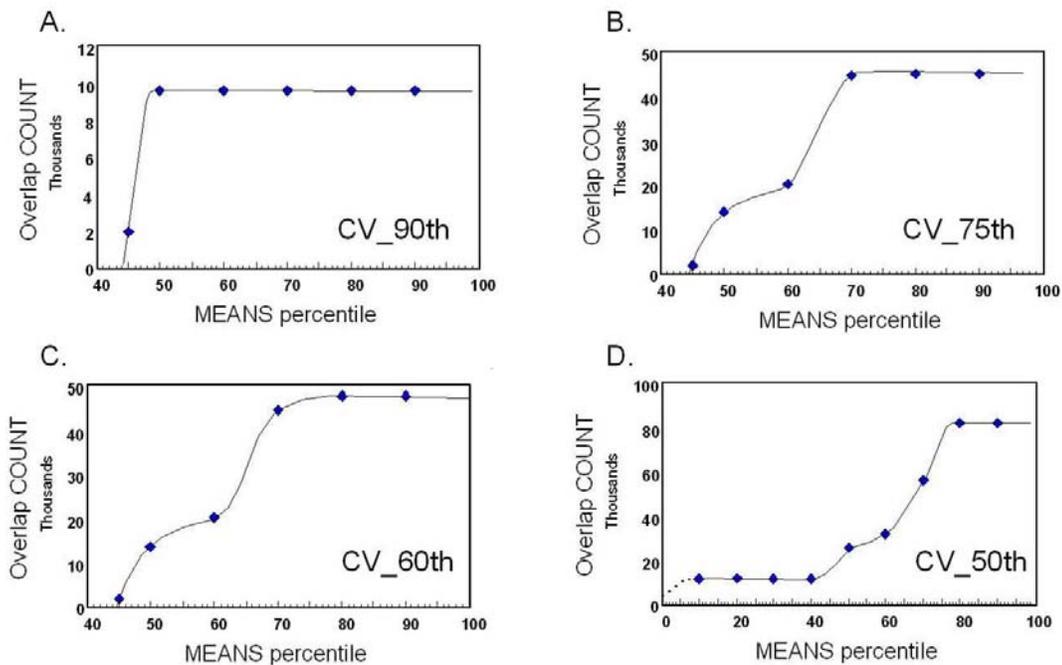


Figure F.1. Number of raster cells classified as "unsuitable" and "uncertain" that overlap at various classification thresholds of the MEANS raster at four different confidence thresholds, U_i, of the coefficient of variation (CV_U_i). The classification threshold (X₀) at which overlap begins is at the 44th percentile for a range of CV thresholds at and above the 60th percentile (a, b, c) and is undefined below this CV confidence level (d).

In comparison, **Figures F.2a** and **F.2b** indicate that the suitability classification threshold for the uncalibrated LPRV suitability map (Section 3.4.2) is undefined at confidence levels higher than about 50% (where the MEANS intercept approaches zero or is negative; Figures F.2a and b). In other words, the highest confidence level at which the uncalibrated suitability map can be classified objectively is only 50%. At this confidence level, the MEANS classification threshold is about at the 44th percentile (Figures F.2c and d).

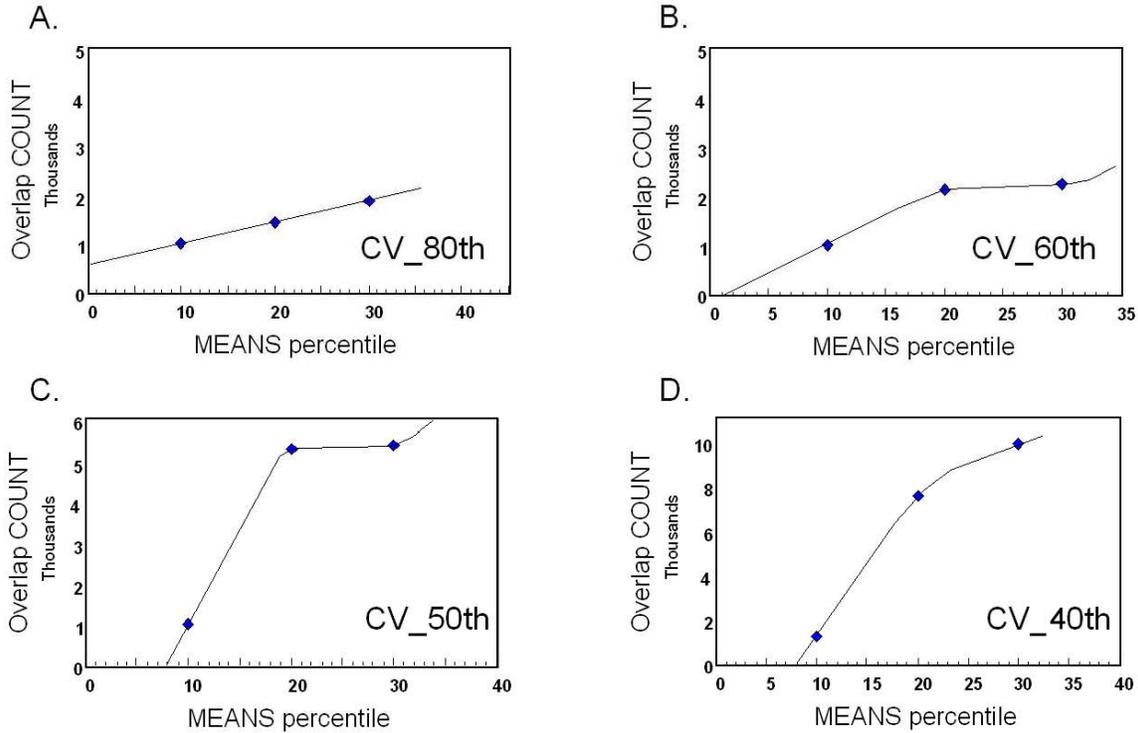


Figure F.2. The classification threshold ($X_o = 8$) for the uncalibrated septic suitability map is defined up to about the 60% confidence level ($U_i < 60$), as shown in b, c, and d whose MEANS intercepts are positive. In contrast, at higher confidence levels (a), the intercept is negative and the classification threshold is undefined.

The value of X_o so determined defines the threshold at which the DRASTIC ratings can defensibly be classified into "suitable" and "unsuitable" zones at a specified level of confidence. In effect, this threshold represents the suitability rating value at which areas that are classified as "unsuitable" and "uncertain" do not overlap at a specified level of uncertainty. The most conservative classifications are achieved at lower confidence levels (lower U_i values, larger area classified as "uncertain"). The classification threshold determined via this methodology is free of any subjective assumptions about what constitutes a "suitable" rating. It is entirely objective because it relies only on the results of the sensitivity analysis and the choice of a classification confidence level based on the sensitivity analysis.

In practice, the simplest way to create a final classified suitability map is by classifying the MEANs raster in ArcMap around the 5-percent quantile that is nearest to X_0 and then overlaying the "CV_Ui" binary raster as a mask over it. **Figure F.3a** shows the LPRV's calibrated MEANs raster classified around the 45th percentile ratings threshold rather than the 44th percentile threshold (as determined in Figure F.1), and overlain with the "CV_U75" raster; **Figure F.3b** shows the same suitability map but with the "CV_U90" raster overlain. Note the small extent of misclassification error that arises by using the nearest 5-percent quantile rather than the exact quantile.

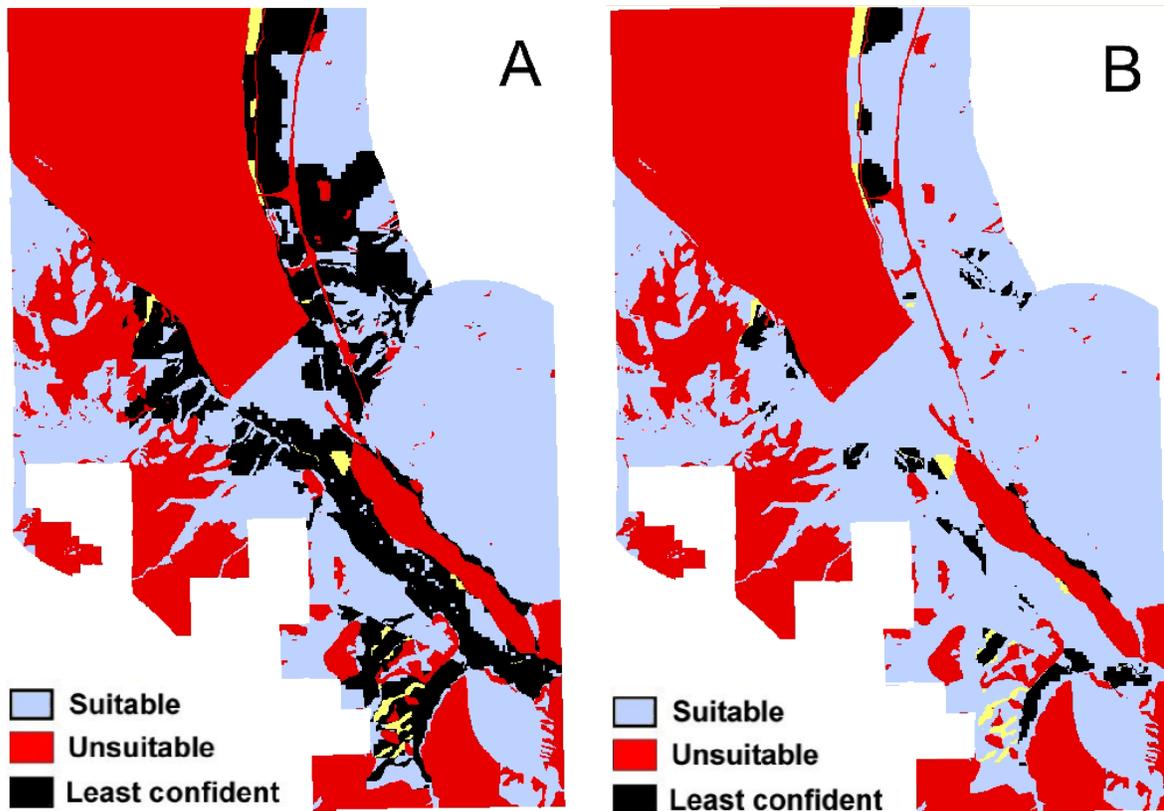


Figure F.3. LPRV calibrated septic suitability map classified at the 45th percentile threshold (blue = suitable; red = unsuitable), with areas (in black) whose classification is uncertain at the (a) 75th and (b) 90th percentile CV thresholds, respectively. Areas in yellow indicate the extent of the misclassification error associated with using the 45th percentile threshold rather than the 44th, as determined in Figure F.1.

For comparison, **Figure F.4** shows the uncalibrated MEANs raster classified around the 10th percentile ratings threshold (nearest to the 8th percentile threshold that was determined in Figure F.2), overlain with the "CV_U₅₀" raster.

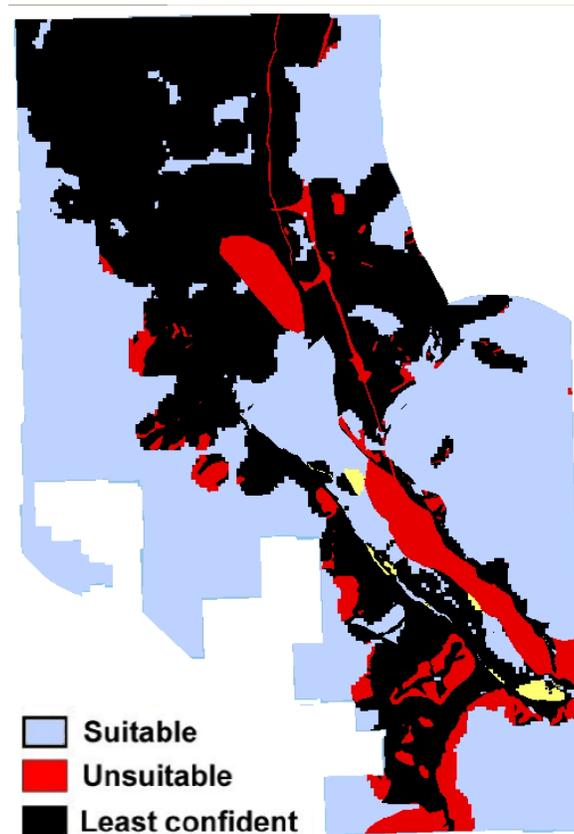


Figure F.4. LPRV uncalibrated septic suitability map classified at the 10th percentile threshold (blue = suitable; red = unsuitable), together with areas classified as uncertain at the 50th percentile CV threshold. Areas in yellow indicate the extent of the classification error incurred by using the 10th percentile threshold rather than the 8th, as determined in Figure F.2.