

# UC Santa Barbara

## GIS Core Curriculum for Technical Programs (1997-1999)

### Title

Unit 41: Using Boolean Search Techniques

### Permalink

<https://escholarship.org/uc/item/2f9545dp>

### Authors

Unit 41, CCTP  
Allen, Chris

### Publication Date

1998

Peer reviewed

## Unit 41 - Using Boolean Search Techniques

Written by Chris Allen, The Clark Labs, Clark University

### Context

It is common to use a GIS to locate areas which satisfy a set of criteria. For example, we might be looking for all properties which are zoned for commercial development and which are located within 100 meters of a sewer line. A typical way to go about this is to prepare a data layer for each criterion in which only areas which satisfy the criterion are represented. In this case, we prepare one data layer which shows all commercially zoned properties and a second data layer showing only areas within 100 meters of a sewer line. By using an OVERLAY operation (see Units 34 and 35) between the two data layers we can determine which parts of our study area, if any, satisfy both criteria. Often there are many criteria to be satisfied and a series of such analytical steps is required.

In a raster system, it is not possible to create a new raster layer which contains ONLY the features of interest - one cannot "remove" or copy out a subset of pixels without changing the spatial characteristics of the dataset (e.g., number of rows and columns, and adjacency relationships). In a raster system, the differentiation of features of interest is performed by setting the attribute value of unselected feature pixels to zero or some other background value.

The situation where only two states of a variable are possible (in this case, criterion satisfied and criterion not satisfied) places us in the arena of boolean analysis. Boolean analysis is named after mathematician and logician George Boole, who devised rules and methodologies for combining such two-state variables. In boolean search we are generally most concerned with the AND operator (sometimes called the "logical AND" operator to indicate that we are referring to the word AND in this special context of specific rules defining the ways in which variables are combined). The logical AND operator produces a "true" result from the phrase "A AND B" only if both A and B are "true". In GIS this methodology is used in a multiplication overlay between layers containing only zeroes (representing areas where conditions are "false" or "criterion is not satisfied") and ones (representing areas where conditions are "true" or "criterion is satisfied"). When such boolean layers are multiplied together in an overlay operation, the only areas of ones in the output image occur where ones are present in both input images (i.e,  $1 \times 1 = 1$ , but  $1 \times 0 = 0$ ,  $0 \times 1 = 0$ , and  $0 \times 0 = 0$ ). The practical use of this technique is described more fully below.

In raster and vector GIS, the concepts underlying boolean search are similar but vary in the details. Conceptually, one can think of a data layer or series of data layers which describe different characteristics of the same spatial area; each layer represents the areas which satisfy

a different condition (in the example above, our two layers represented proximity to sewer lines and commercial zoning). The data layers are compared to produce a new data layer which depicts areas which are common to both input layers. In a raster GIS this analysis is performed cell-by-cell. Cells which correspond spatially are compared to determine the value of the corresponding cell in the output data layer. In a vector GIS the analysis is performed feature-by-feature or area-by-area. Areas or features which are common to both layers are represented in the output layer.

Boolean search techniques, as discussed in this unit, are fundamental to most GIS analyses. However, in many real-world applications, the decision making process can be more complicated. There are situations where a boolean analysis such as this does not produce a satisfactory solution. First, boolean analysis requires that each condition be defined by hard thresholds. In the sewer line example, all areas within 100 meters are treated as equally suitable, while all areas beyond 100 meters are treated as equally and totally unsuitable. Second, with the Boolean approach, all of the criteria carry equal importance in the solution. There is no distinction in the result between areas which fail one criterion and areas that fail multiple criteria. Some GIS systems, such as Idrisi, provide capabilities for coping with decision making problems such as this. These systems may provide tools for creating continuous suitability layers, assigning relative importance to data layers, and quantitatively analyzing the weighted layers to determine the best possible result. It is beyond the scope of this unit to discuss such applications, but it is important to note that some, but not all, GIS problems can be solved using boolean logic.

## [Detailed Comparison of Raster and Vector Approaches to Boolean Search](#)

### **Example Application - Raster GIS**

Let's work through the example described above. We are looking for properties which might be suitable for location of a new store. As a first step in locating a potential site, we have decided that we will search for all properties which are zoned for commercial development and which are located within 100 meters of a sewer line. A general outline of the procedure is as follows: we prepare one data layer in which all commercially zoned properties have the value 1 and everything else has the value 0 and a second data layer in which only areas within 100 meters of a sewer line have the value 1 and everything else has the value 0. By using an OVERLAY operation (see Units 34 and 35) between the two data layers we can determine which parts of our study area, if any, satisfy both criteria. For other problems there may be many criteria to be satisfied, in which case we would take the data layer which results from the first OVERLAY operation and overlay it with the next boolean criterion image, and take the result of that OVERLAY operation and overlay it with the next boolean layer and so forth (this procedure could be applied to as many criteria as necessary). When we have multiplied all of the boolean layers together, the resulting layer indicates areas where ALL of the conditions are satisfied (pixels with a value of 1 in the output layer) and areas where one or more conditions were not satisfied (pixels with a value of 0 in the output layer).

Let's work through the analytical steps for our example problem one-by-one:

We start with a vector data layer containing property parcel boundaries and a vector line file containing lines representing sewer line locations:



Figure 1

We rasterize our vector property parcels coverage to produce this raster parcels data layer; each pixel in the raster is given a value equal to the numeric parcel ID:



Figure 2

We rasterize the sewer line vector coverage and create a 100-meter buffer around it. We start with a data layer showing all sewer lines. Depending on the capabilities of the particular GIS we are using we could use a BUFFER operation (see Unit 33) to create a 100-meter buffer around the sewer lines. If necessary, we would then reclass all areas within the buffer (areas within 100 meters of a sewer line) to 1 and all areas outside the buffer (areas more than 100 meters from a sewer line) to 0. If buffering capability is not provided in the GIS one is using, we would create a data layer using a DISTANCE operator (see Unit 36) in which the attribute value of each pixel is its distance to the nearest sewer line. Reclassification would then be performed to convert all pixels with values less than 100 meters to 1 and all pixels with values greater than 100 meters to 0. Regardless of whether we use a buffering operator (and reclassification if necessary) or a distance operator and reclassification, the result is the same: all pixels representing areas within 100 meters of a sewer line are 1 and all other pixels are 0:



Figure 3

We need to create a boolean layer indicating the property parcels which satisfy the zoning

classification criterion. The polygon attribute table for our property parcels layer includes an attribute for zoning classification (called Zoning). We create a new field in the polygon attribute database called ZoneOK. For every parcel with a Zoning attribute of 3 (which means commercial zoning, in this example), we will set the value of ZoneOK to 1. Every parcel which has the Zoning attribute different from 3 gets a ZoneOK value of 0:

Parcel	Area	Perimeter	Volume	Max Area	Lat	Zoning	Percent
1	100.0000	100.0000	100.0000	100.0000	100.0000	3	100.0000
2	200.0000	200.0000	200.0000	200.0000	200.0000	3	200.0000
3	300.0000	300.0000	300.0000	300.0000	300.0000	3	300.0000
4	400.0000	400.0000	400.0000	400.0000	400.0000	3	400.0000
5	500.0000	500.0000	500.0000	500.0000	500.0000	3	500.0000
6	600.0000	600.0000	600.0000	600.0000	600.0000	3	600.0000
7	700.0000	700.0000	700.0000	700.0000	700.0000	3	700.0000
8	800.0000	800.0000	800.0000	800.0000	800.0000	3	800.0000
9	900.0000	900.0000	900.0000	900.0000	900.0000	3	900.0000
10	1000.0000	1000.0000	1000.0000	1000.0000	1000.0000	3	1000.0000
11	1100.0000	1100.0000	1100.0000	1100.0000	1100.0000	3	1100.0000
12	1200.0000	1200.0000	1200.0000	1200.0000	1200.0000	3	1200.0000
13	1300.0000	1300.0000	1300.0000	1300.0000	1300.0000	3	1300.0000
14	1400.0000	1400.0000	1400.0000	1400.0000	1400.0000	3	1400.0000

Figure 4

A few steps back, by rasterizing our vector property parcels layer, we created a property parcels raster data layer in which each pixel has a value equal to the numeric property ID (Userid in the database). We can use this property parcels raster layer to create the boolean layer we require to represent our property zoning criterion. To each pixel in the raster we assign a new value, the value from the ZoneOK field of the database, depending on the property parcel the pixel belongs to. For example, if a particular pixel has a value of 6168 in the parcels raster (that pixel is located in parcel number 6168 corresponding to the record in the database with a Userid of 6168), then it gets a new value of 0, which is the value of ZoneOK in the database record for that parcel. This process, assigning the ZoneOK value based on Userid, produces this boolean image which has zero-value pixels everywhere except where properties are zoned commercially. In the commercially zoned areas the pixels have a value of one:

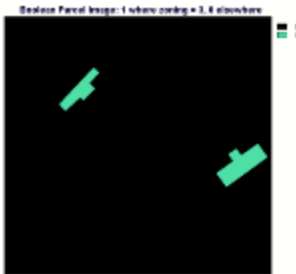


Figure 5

Now that we have our two boolean data layers, we use an overlay multiply operation to combine the data layers (data layer number one where all commercial pixels have a value of 1 and all non-commercial pixels have a value of 0; and data layer two where all pixels within 100 meters of a sewer line are 1 and all other pixels are 0). The result of the overlay is a data layer in which all non-zero pixels represent areas in which a 1 in one layer is multiplied by a 1 in the corresponding cell of the second layer, i.e., pixels which satisfy both criteria. Pixels where neither criterion is satisfied (have a value of 0 in both input layers) produce an overlay-multiply result of 0. Pixels where only one of the two criteria is satisfied (have a value of 1 in one data layer and 0 in the other) will also produce an overlay-multiply result of 0. The result of the overlay step is a data layer in which the areas we are looking for (within 100 meters of a sewer line and commercially zoned) are effectively differentiated from those areas we are not looking for. The pixels in areas we seek have a value of 1 in the output raster and the

pixels in areas we wish to ignore have values of 0:

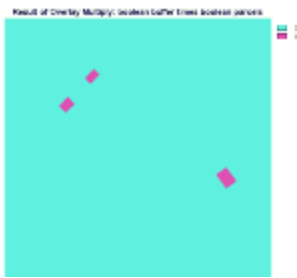


Figure 6

It is easier to envision where the suitable areas are if we superimpose our vector property outlines and IDs:

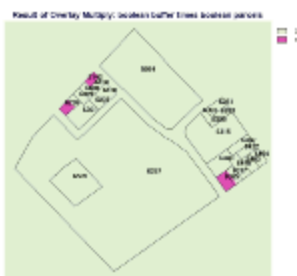


Figure 7

In this example, where we have determined that there are only several suitable property parcels, it would be possible to look at the last figure and write down the IDs of the suitable parcels. If there were a large number of suitable parcels, we might wish to determine their identities analytically. In Idrisi we could use the CROSSTAB module, which compares two raster layers and determines the number of pixels in each combination of values present in the two images. In the present case, we are looking for all of the combinations where pixel values of 1 are present in either of the previous two figures. In this example, all or portions of 4 property parcels satisfy our criteria, as indicated in this table produced by CROSSTAB (with comments added by the author):

```

CROSS TABULATION OF PARCEL NUMBER AND VALUE = 1 BY PROPERTY AND VALUE GROUP
NUMBER OF PIXELS
-----
PARCEL #   0       1       TOTAL
-----
1004      10000    1000    11000
1005      10000    1000    11000
1006      10000    1000    11000
1007      10000    1000    11000
1008      10000    1000    11000
1009      10000    1000    11000
1010      10000    1000    11000
1011      10000    1000    11000
1012      10000    1000    11000
1013      10000    1000    11000
1014      10000    1000    11000
1015      10000    1000    11000
1016      10000    1000    11000
1017      10000    1000    11000
1018      10000    1000    11000
1019      10000    1000    11000
1020      10000    1000    11000
1021      10000    1000    11000
1022      10000    1000    11000
1023      10000    1000    11000
1024      10000    1000    11000
1025      10000    1000    11000
1026      10000    1000    11000
1027      10000    1000    11000
1028      10000    1000    11000
1029      10000    1000    11000
1030      10000    1000    11000
1031      10000    1000    11000
1032      10000    1000    11000
1033      10000    1000    11000
1034      10000    1000    11000
1035      10000    1000    11000
1036      10000    1000    11000
1037      10000    1000    11000
1038      10000    1000    11000
1039      10000    1000    11000
1040      10000    1000    11000
1041      10000    1000    11000
1042      10000    1000    11000
1043      10000    1000    11000
1044      10000    1000    11000
1045      10000    1000    11000
1046      10000    1000    11000
1047      10000    1000    11000
1048      10000    1000    11000
1049      10000    1000    11000
1050      10000    1000    11000
1051      10000    1000    11000
1052      10000    1000    11000
1053      10000    1000    11000
1054      10000    1000    11000
1055      10000    1000    11000
1056      10000    1000    11000
1057      10000    1000    11000
1058      10000    1000    11000
1059      10000    1000    11000
1060      10000    1000    11000
1061      10000    1000    11000
1062      10000    1000    11000
1063      10000    1000    11000
1064      10000    1000    11000
1065      10000    1000    11000
1066      10000    1000    11000
1067      10000    1000    11000
1068      10000    1000    11000
1069      10000    1000    11000
1070      10000    1000    11000
1071      10000    1000    11000
1072      10000    1000    11000
1073      10000    1000    11000
1074      10000    1000    11000
1075      10000    1000    11000
1076      10000    1000    11000
1077      10000    1000    11000
1078      10000    1000    11000
1079      10000    1000    11000
1080      10000    1000    11000
1081      10000    1000    11000
1082      10000    1000    11000
1083      10000    1000    11000
1084      10000    1000    11000
1085      10000    1000    11000
1086      10000    1000    11000
1087      10000    1000    11000
1088      10000    1000    11000
1089      10000    1000    11000
1090      10000    1000    11000
1091      10000    1000    11000
1092      10000    1000    11000
1093      10000    1000    11000
1094      10000    1000    11000
1095      10000    1000    11000
1096      10000    1000    11000
1097      10000    1000    11000
1098      10000    1000    11000
1099      10000    1000    11000
1100      10000    1000    11000
1101      10000    1000    11000
1102      10000    1000    11000
1103      10000    1000    11000
1104      10000    1000    11000
1105      10000    1000    11000
1106      10000    1000    11000
1107      10000    1000    11000
1108      10000    1000    11000
1109      10000    1000    11000
1110      10000    1000    11000
1111      10000    1000    11000
1112      10000    1000    11000
1113      10000    1000    11000
1114      10000    1000    11000
1115      10000    1000    11000
1116      10000    1000    11000
1117      10000    1000    11000
1118      10000    1000    11000
1119      10000    1000    11000
1120      10000    1000    11000
1121      10000    1000    11000
1122      10000    1000    11000
1123      10000    1000    11000
1124      10000    1000    11000
1125      10000    1000    11000
1126      10000    1000    11000
1127      10000    1000    11000
1128      10000    1000    11000
1129      10000    1000    11000
1130      10000    1000    11000
1131      10000    1000    11000
1132      10000    1000    11000
1133      10000    1000    11000
1134      10000    1000    11000
1135      10000    1000    11000
1136      10000    1000    11000
1137      10000    1000    11000
1138      10000    1000    11000
1139      10000    1000    11000
1140      10000    1000    11000
1141      10000    1000    11000
1142      10000    1000    11000
1143      10000    1000    11000
1144      10000    1000    11000
1145      10000    1000    11000
1146      10000    1000    11000
1147      10000    1000    11000
1148      10000    1000    11000
1149      10000    1000    11000
1150      10000    1000    11000
1151      10000    1000    11000
1152      10000    1000    11000
1153      10000    1000    11000
1154      10000    1000    11000
1155      10000    1000    11000
1156      10000    1000    11000
1157      10000    1000    11000
1158      10000    1000    11000
1159      10000    1000    11000
1160      10000    1000    11000
1161      10000    1000    11000
1162      10000    1000    11000
1163      10000    1000    11000
1164      10000    1000    11000
1165      10000    1000    11000
1166      10000    1000    11000
1167      10000    1000    11000
1168      10000    1000    11000
1169      10000    1000    11000
1170      10000    1000    11000
1171      10000    1000    11000
1172      10000    1000    11000
1173      10000    1000    11000
1174      10000    1000    11000
1175      10000    1000    11000
1176      10000    1000    11000
1177      10000    1000    11000
1178      10000    1000    11000
1179      10000    1000    11000
1180      10000    1000    11000
1181      10000    1000    11000
1182      10000    1000    11000
1183      10000    1000    11000
1184      10000    1000    11000
1185      10000    1000    11000
1186      10000    1000    11000
1187      10000    1000    11000
1188      10000    1000    11000
1189      10000    1000    11000
1190      10000    1000    11000
1191      10000    1000    11000
1192      10000    1000    11000
1193      10000    1000    11000
1194      10000    1000    11000
1195      10000    1000    11000
1196      10000    1000    11000
1197      10000    1000    11000
1198      10000    1000    11000
1199      10000    1000    11000
1200      10000    1000    11000
    
```

Figure 8

Alternatively, we could determine the identities of the parcels satisfying our conditions by taking the boolean image representing the solution derived above and overlay-multiplying it by the property parcels raster layer we created early on. The only non-zero pixels in the resulting raster layer will occur in areas which satisfy the criteria, and the values of those pixels will be the property parcel identifiers of the suitable properties. We could run HISTO in Idrisi to create a tabular histogram listing the suitable property identifiers.

## Learning Outcomes

The following list describes the skills which students are expected to master at each of the three levels of training, Awareness, Competency, and Mastery.

### Awareness:

The learning goal of this section is to understand the concept of creating boolean data layers in which areas of interest, or which satisfy some criterion, are differentiated from areas to be excluded, or which do not satisfy the criterion.

Students should understand that two or more of these boolean layers can be combined to produce new data layers in which areas satisfying both/all criteria are differentiated from areas which satisfy some/none of the criteria.

Students should understand that, in a vector system, the determination of which features satisfy multiple criteria is typically by query operations on an attribute database or by overlaying data layers and, in a raster system the determination is made using overlay multiplication between data layers.

### Competency:

The learning goal of this section is to understand how boolean data layers can be prepared from complex data layers to differentiate features of interest from features to be excluded from further analysis.

Students should understand how to combine boolean layers, each of which represents a single criterion, to determine areas which satisfy multiple criteria.

If a vector GIS is used, students should understand that selection of features can sometimes be performed from attribute tables. At other times performance of spatial operations, such as buffering, is required to derive a selected subset of features. The subset of features can be used in an intersect overlay to derive a solution to multiple criterion search, or the results of such spatial operations can be used to create attributes for use in deriving a database query solution to the search problem.

### Mastery:

Students should be capable of determining a set of criteria necessary to implement a multiple-criterion spatial selection problem.

Students should be capable of determining the required data layers or attribute fields necessary to represent the relevant criteria.

Students should be capable of applying available tools within a GIS to implement the analysis.

If a vector system is used, students should be able to translate the results of a spatial operation

such as buffering into a new attribute field and to perform attribute query based on that field.

---

## **Preparatory Units**

### **Recommended**

What is GIS?

Unit 19 - Planning a Tabular Database

Unit 22 - Merging Tabular Data with Spatial Data

Unit 31 - Managing Database Files

Unit 34 - Types of Overlay Operators/Using Overlay Operators

Unit 38 - Data Expansion - Deriving New Attributes/Fields/Layers

### **Complementary**

Unit 33 - Using Buffering Operators

Unit 36 - Using Distance and Connectivity Operators

Unit 42 - Using Map Algebra

Unit 47 - On-Screen Visualization

---

## **Awareness**

### **Learning Objectives**

1. Student can demonstrate that he/she understands the concept of data layer overlay.
2. Student can break down a complex search task into separate boolean criteria.
3. Student can explain how each of the criteria would be represented by a boolean data layer.
4. Student can explain the expected result of combining the boolean criterion layers to produce a data layer which contains features which satisfy all of the criteria.

### **Vocabulary**

Attribute



Boolean

Buffer

Criterion

Overlay-Intersect

Overlay-Multiply

Query

Reclassification

Reselection

### **Awareness Topics**

1. In a vector system, features of interest can be selected from a dataset. This can be based on spatial location or on one or more attribute values contained in an attribute table.
2. Features can have one of two states with respect to a given criterion, selected or unselected. This is a boolean representation.
3. In a vector system, the selected features can form a new discrete data layer or, in some systems, the selection can be virtual - only the selected features are displayed or processed, but the unselected features are not actually removed from the data set.
4. In a raster system, it is not possible to create a new raster layer which contains ONLY the features of interest - one cannot "remove" or copy out a subset of pixels without changing the spatial characteristics of the dataset (e.g., number of rows and columns, and adjacency relationships). In a raster system, the differentiation of features of interest is performed by setting the attribute value of unselected feature pixels to zero or some other background value.
5. Although the methods differ between raster and vector GIS, both have the capability of determining the features which have the same boolean state (e.g., selected or unselected) in two input datasets.
6. By taking the result of one such determination and evaluating it against a third boolean dataset, the features which satisfy all three selection conditions can be determined. This process can be applied as many times as necessary until all criteria or selection conditions have been evaluated.

---

## **Competency**

### **Learning Objectives:**

1. Student can explain the difference between raster and vector systems in the way selected features are represented; in a vector system, a data layer which can contain only features of interest, but in a raster system, features of interest can be differentiated by setting all pixels to zero in areas to be excluded from consideration.
2. Student can create boolean data layers or RESELECT features based on criteria.
3. Student can combine the boolean images which represent multiple criteria to produce a result which contains only those features which satisfy all of the criteria.

### **Generic List of Tasks**

1. Starting with a data layer in which categorical data is represented, e.g., land use, zoning type, use reclassification to create a boolean data layer in which two categories are combined to form one class with an attribute value of 1 and all other categories are combined to form another class with an attribute value of 0.
2. Take a second data layer which covers the same spatial extent as the first layer. If a raster system is being used, choose a data layer in which attribute values vary continuously across the study area, e.g. the result of a distance operator. Apply a reclassification operator to this data layer to produce a new boolean data layer based on a given threshold.
3. Overlay-intersect (vector system) or overlay-multiply (raster system) the data layers resulting from steps one and two.

---

## **Mastery**

### **Learning Objectives:**

1. Student can explain how multiplicative overlay in the boolean case (data layers containing zeroes and ones) is analogous to a spatial intersection overlay.
2. Student can apply a buffering or distance operation to derive the following:

In a raster system: a boolean data layer in which all pixels within a buffer have an attribute value of 1 and all pixels outside the buffer have an attribute of 0.

Or:

In a vector system: create a data layer containing a buffer area and for a data layer covering the same extent (e.g., property parcels) to add a field to the attribute table and populate the new field with values which differ based on whether the feature is inside or outside the buffer (in some vector GIS this is performed automatically during buffering).

3. Student can use the buffered data layer created in (2) in an OVERLAY-MULTIPLY operation (raster) or an OVERLAY-INTERSECT operation (vector) to exemplify boolean

search using spatially derived criteria.

### **Mastery Tasks:**

1. Describe a scenario in which a GIS analyst is searching for a set of suitable sites, but for which boolean search would be an inappropriate method to use.
2. Discuss the following scenario: Your boolean search for line features, roads for example, based on a certain set of criteria yields seven road features satisfying all of the criteria, but you can only use one feature for the intended purpose, how can you decide which to use? Can boolean search methods help?
3. Develop a methodology to address the following situation: You are trying to decide which lake should be used for an experiment in water quality. You have just performed a boolean search for one or more lakes which satisfy criteria such as size larger than 10 hectares, no more than 500 people living within a mile of the lake shore, swimming is permitted, etc. Upon completion of your analysis, you find that there are no lakes in the study area which satisfy all of the criteria. However, this is an unacceptable result. At least one lake must be selected for the experiment. How can you use GIS to determine which lake to use?



### **Example Implementation (with lots of graphics)**

---

### **Follow-Up Units:**

Unit 22 - Merging Tabular Data with Spatial Data

Unit 31 - Managing Database Files

Unit 33 - Using Buffering Operators

Unit 34 - Types of Overlay Operators/Using Overlay Operators

Unit 36 - Using Distance and Connectivity Operators

Unit 38 - Data Expansion - Deriving New Attributes/Fields/Layers

Unit 42 - Using Map Algebra

Unit 47 - On-Screen Visualization

---

### **Resources:**

*[Outdated links have been removed.]*

Idrisi Users' Guide and Tutorial Exercises

Idrisi Salzburg Resource Center

---

## **Acknowledgments**

The author wishes to thank Michele Fulk of Clark Labs for the invaluable expert advice and editorial assistance she provided in the development of this unit.

---

Created: May 14, 1997. Last updated: October 5, 1998.

## Detailed Comparison of Raster and Vector Approaches to Boolean Search

In a raster GIS, this type of analysis is typically performed by creating a boolean data layer (a layer which contains only 0 and 1, with 1 signifying areas where the criterion is satisfied, and 0 signifying areas where the condition is not satisfied) for each criterion. After all of the boolean data layers have been prepared, a series of overlay steps is performed to produce a final result. Generally multiplication overlay is used; a boolean image for one criterion overlays a second boolean image representing the second criterion. Multiplication overlay produces an output pixel value of 1 only if both input pixel values are 1 (1 times 1 equals 1), i.e., if both criteria are satisfied. An output pixel value of 0 is produced if one or the other or both input pixels is 0 (1 times 0 equals 0, or 0 times 0 equals 0), i.e., if one/both criteria is/are NOT satisfied.

An overlay operation is applied to evaluate the areas which satisfy the first two criteria. The resulting raster layer is then overlaid with the third boolean image. The raster resulting from that overlay operation is overlaid with the fourth boolean image and so forth until all of the criteria have been evaluated. Note that some GIS allow multiple boolean images to be overlaid in a single step. In the final result all non-zero pixels will be those in which the corresponding pixels in ALL of the boolean criterion images have a value of one. If a pixel in any of the criterion images is zero, i.e., if any of the criteria are not satisfied at the location represented by that pixel, then the result for that pixel will be zero, regardless of how many criteria WERE satisfied at that location.

The order in which the overlay operations are performed is not important as long as each of the boolean criterion images is used in one of the overlay steps.

In a vector GIS, there are two commonly used methods of conducting a search for features which satisfy multiple criteria. The first method uses database query to perform the analysis. The second method involves an intersection type of overlay operation.

In the database query method, one typically starts with a data layer which contains a number of features, e.g., property parcel polygons. The data layer has an associated attribute table which contains attributes for each feature, e.g., parcel size, assessed property value, zoning classification, etc. The essential procedure is to query the attribute database to determine the subset of features which satisfy the criteria.

Imagine, for example, that we are trying to locate commercial properties located less than 100 meters from a sewer line. For this example, we assume that we have a data field in our

property parcels attribute database called SEWERDIST indicating the minimum distance (in meters) of the property to the nearest sewer line. [It is beyond the scope of this Unit to discuss the method for creating the SEWERDIST field in the attribute table and filling the field with the correct distance value for each polygon (see Unit 36 Using Distance and Connectivity Operators)]. We also have an attribute field that indicates whether the property is zoned for commercial use, residential use, agricultural use, and so forth. For this example we will imagine that an attribute value of "3" in the ZONING field indicates that a property is commercially zoned. Our goal is to find the properties which have both commercial zoning (i.e., ZONING = 3) and a distance-to-sewer less than 100.

First, we apply a filter, sometimes called SELECT or RESELECT, to the attribute database to extract the subset of polygons which satisfy one of the criteria, e.g., "RESELECT ZONING = 3". This creates a data layer of all polygons for which the ZONING field in the attribute table has a value of 3, which, in our example indicates a zoning classification of "Commercial". We then RESELECT SEWERDIST < 100 to restrict our selected set to the subset of commercial properties for which the distance to sewer line is less than 100 meters.

Alternatively, we could perform the extraction of suitable polygons in one step by RESELECT ZONING = 3 AND SEWERDIST < 100. It is sometimes the case that a data layer (e.g., parcel polygons) may have several separate attribute tables and the different attributes one needs to RESELECT on are contained in different tables. In such a case, it is necessary to RELATE the attribute tables on an identifier which is common to both tables, e.g., parcel ID.

Note that this only works when we are considering multiple attributes for the same spatial layer (for example, when we are working with a number of attributes all of which pertain to the same property parcels feature layer). When we are considering multiple spatial layers (for example, towns and drainage basins) each with its own attributes, a spatial overlay is required, as is described next.

In vector GIS, the second common method of searching for areas which satisfy multiple criteria is the use of the overlay intersection operator. In the overlay-intersect operation, two data layers are overlaid and only the areas/features common to both are present in the output layer. A data layer displaying the 100-meter buffer around sewer lines could be INTERSECTed with the layer containing only commercial properties to derive an output data layer containing only those areas common to both, i.e., those areas which are commercially zoned and which are located within 100 meters of a sewer. In some vector GIS, only entire parcels will be selected and the user decides if a parcel must be completely within, or only intersected by, the buffer or selection area. In other vector GIS, the software can create new polygons which contain only the portions of parcels within the buffer or selection area.

[Return to Unit 41 - Using Boolean Search Techniques](#)



# Property Parcels

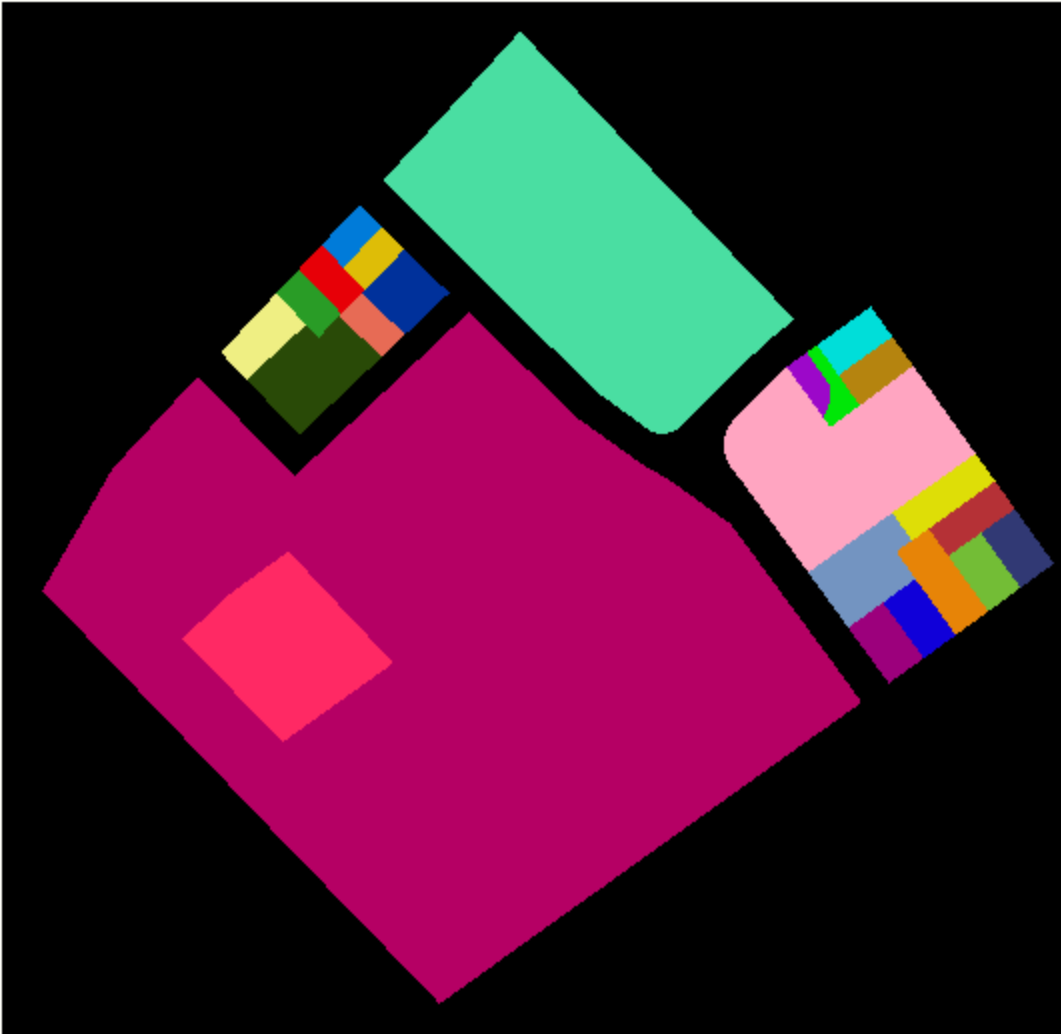


Figure 2



**Boolean Buffer Image: 1 within 100m of sewer; 0 elsewhere**

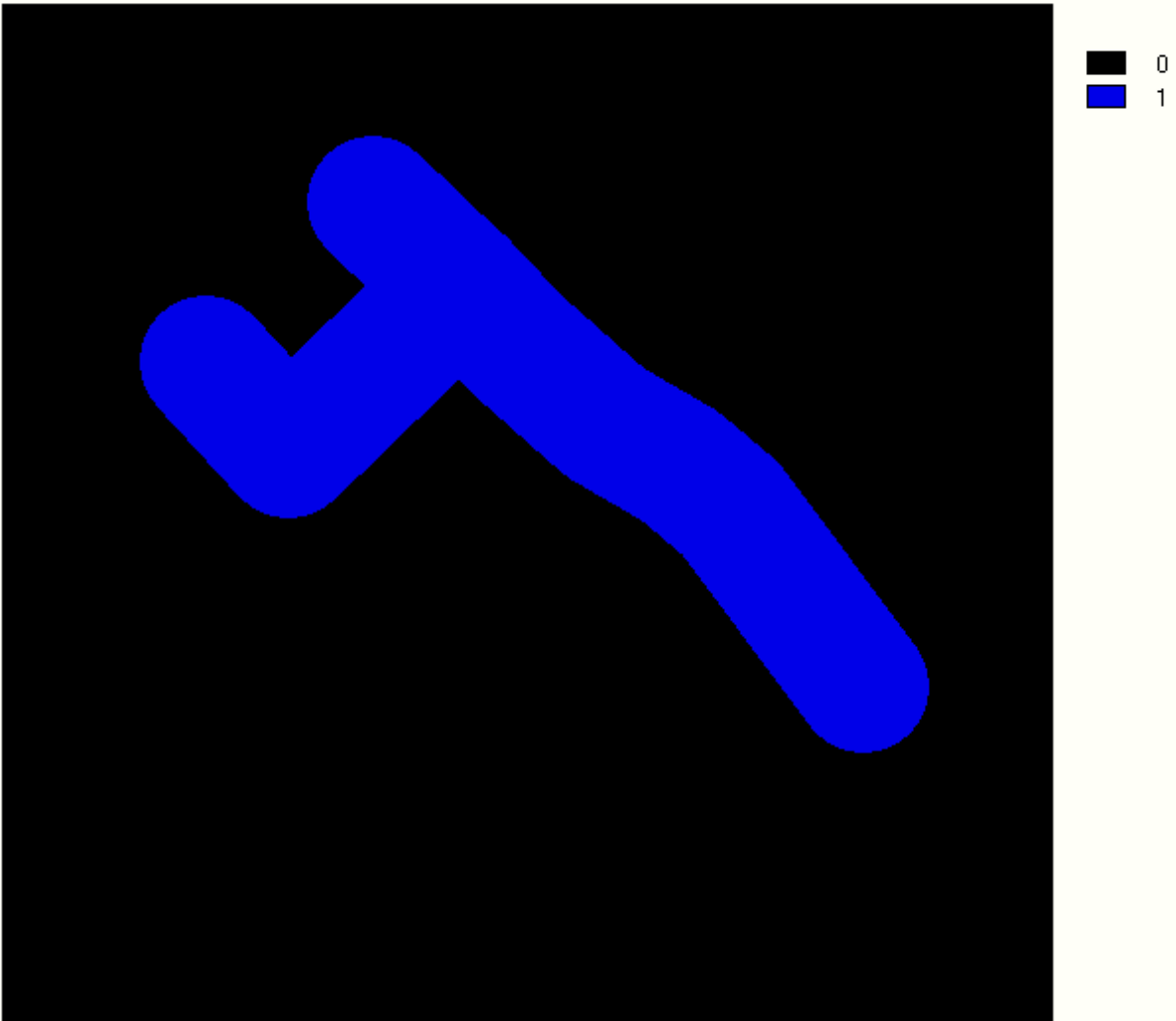


Figure 3

Parcels								
Polyid	Area	Perimeter	Userid	Map	Block	Lot	Zoning	Zoneok
1	170877.429688	1735.414225	5964	08	002	00003	8	0
2	3908.976563	260.140799	6142	08	004	00003	3	1
3	4008.359375	264.176054	6168	08	004	00016	1	0
4	4650.328125	286.370855	6180	08	004	00015	3	1
5	8378.335938	366.261552	6188	08	004	00014	1	0
6	4573.093750	283.992740	6209	08	004	00010	3	1
7	7365.648438	364.536181	6236	08	004	00001	3	1
8	4313.781250	275.108855	6237	08	004	00013	1	0
9	5487.406250	308.901886	6251	06	042	00002	1	0
10	632085.001512	4301.701522	6257	08	003	00001	2	0
11	17915.406250	566.695461	6261	08	004	00002	1	0
12	5420.414063	308.004269	6283	06	042	00003	1	0
13	3232.312500	314.498076	6295	06	042	0004A	1	0
14	2951.015625	245.131314	6303	06	042	00004	1	0
15	64058.562500	1121.342424	6315	06	042	08-11	1	0
16	7841.796875	409.077476	6402	06	042	00029	1	0
17	6183.812500	346.060547	6427	06	042	00028	1	0
18	6268.218750	326.124906	6454	06	042	00027	3	1
19	14217.070313	511.539785	6460	06	042	00007	1	0
20	9216.148438	426.636608	6478	06	042	00024	3	1
21	6120.937500	324.194438	6483	06	042	00026	3	1
22	46195.984375	856.890136	6509	08	007	00012	2	0
23	6179.250000	326.548360	6537	06	042	00025	3	1
24	6682.820313	335.771149	6565	06	042	00023	3	1

Figure 4

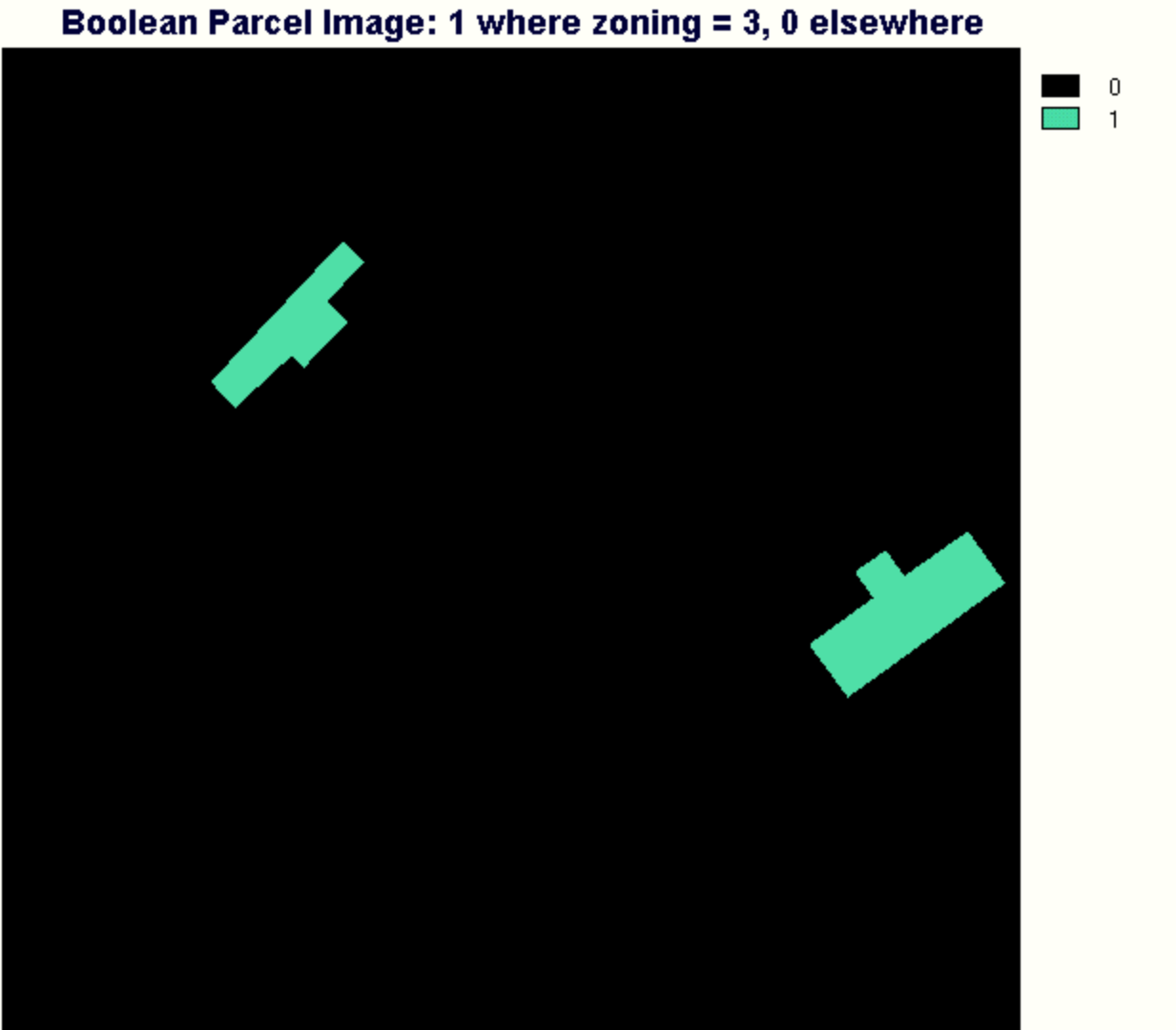


Figure 5

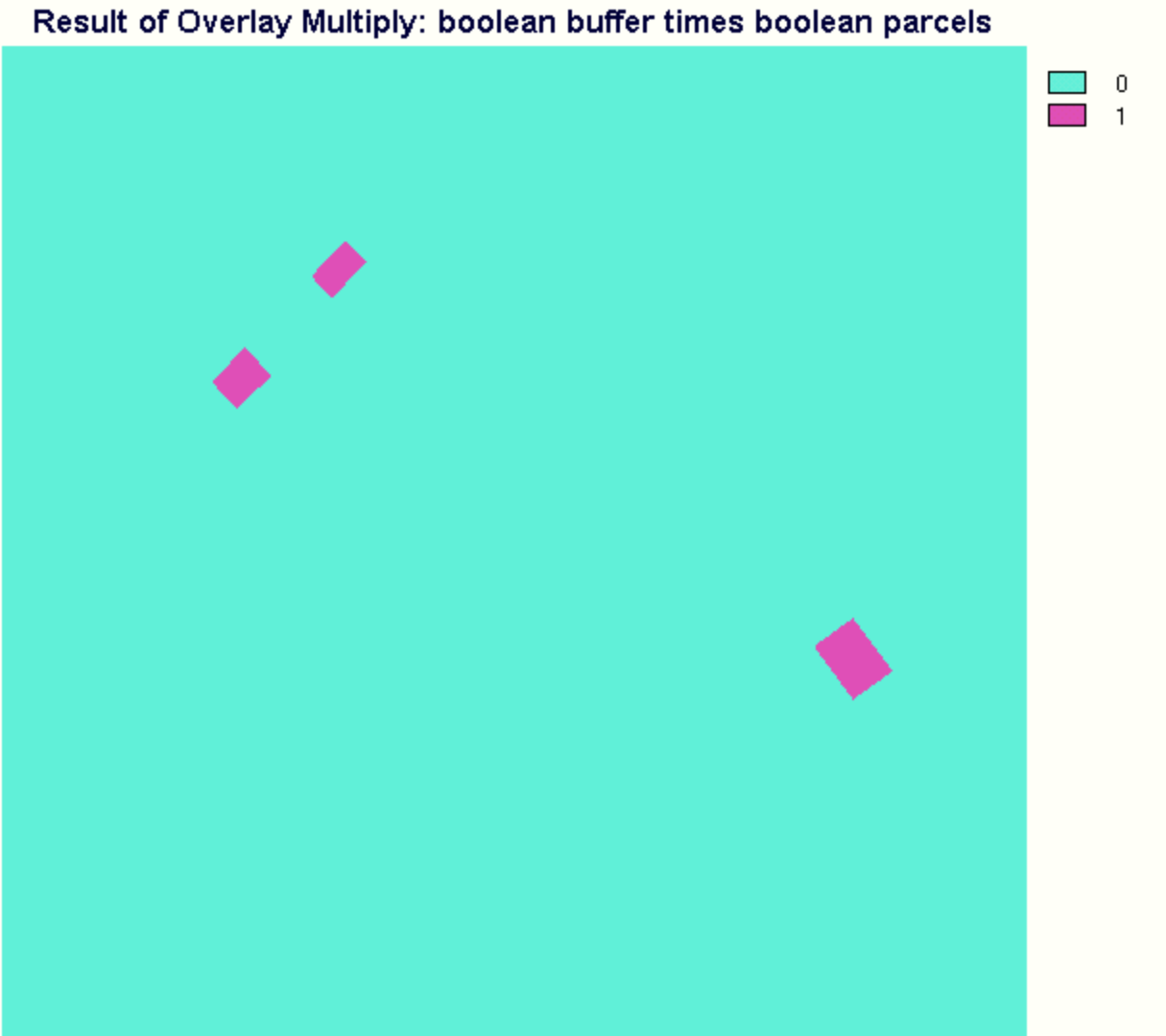


Figure 6

### Result of Overlay Multiply: boolean buffer times boolean parcels



Figure 7

Cross-tabulation of (within buffer AND zoning = 3) (columns) against parcelID (rows) (number of pixels)

Parcel ID	0	1	Total	
0	159631	0	159631	
5964	18971	0	18971	
6142	27	409	436	(409 pixels of parcel 6142 satisfy the criteria)
6168	448	0	448	
6180	515	0	515	
6188	925	0	925	
6209	507	0	507	
6236	329	491	820	(491 pixels of parcel 6236 satisfy the criteria)
6237	479	0	479	
6251	613	0	613	
6257	70228	0	70228	
6261	1991	0	1991	
6283	602	0	602	
6295	359	0	359	
6303	328	0	328	
6315	7124	0	7124	
6402	871	0	871	
6427	689	0	689	
6454	697	0	697	
6460	1580	0	1580	
6478	1022	0	1022	
6483	679	0	679	
6509	5134	0	5134	
6537	559	128	687	(128 pixels of parcel 6537 satisfy the criteria)
6565	0	742	742	(742 pixels of parcel 6565 satisfy the criteria)
<b>Total</b>	<b>274308</b>	<b>1770</b>	<b>276078</b>	

Figure 8

## The NCGIA GIS Core Curriculum for Technical Programs

### Example Implementation for [Unit 41](#)

Imagine that we have been contracted to locate an area within our study region which satisfies the following criteria for possible residential development:

The area must be 50 hectares or larger.

The area must be outside the 50-meter protective buffer zone around water bodies and streams.

The area must not already be developed. Acceptable current uses for the area are crop land, forest land, pasture land, or open undeveloped.

The land must not contain slopes in excess of 15%.

We are provided with the following data layers to work with:

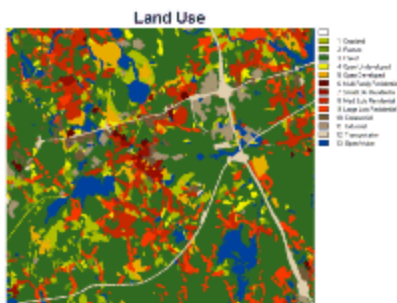
A raster land use map

A raster digital elevation model (DEM)

A vector streams layer

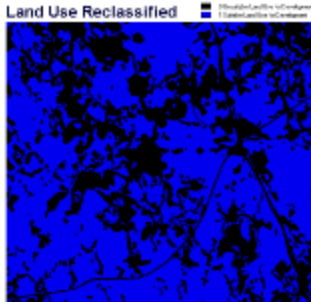
A vector roads layer

Let's start by taking a look at our land use map:



**Land Use**

We start by reclassifying our land use categories to create a boolean image, pixel values of 1 for all suitable land uses (crop land, forest land, pasture land, open undeveloped land), and pixel values of 0 for all other classes. This is the first of our boolean images representing one of the criteria.



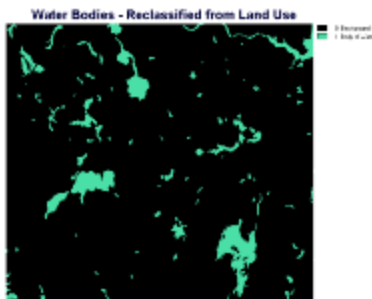
**Land Use Reclassified**

Our land use map contains water bodies such as ponds and lakes, but does not contain streams and rivers as a separate land use. Here is the land use map again. This time, the vector streams layer is superimposed for display:



**Land Use with Streams**

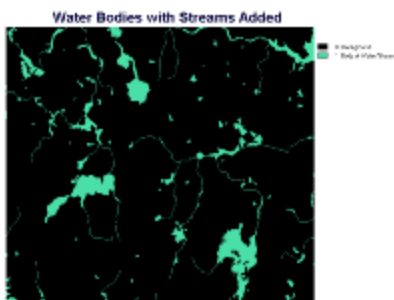
We need to create a suitability layer based on buffers around both lakes and streams. The first step in creating such a layer is to reclassify our land use map (using the RECLASS module in Idrisi) to create a data layer where pixels representing water bodies have a value of 1 and all other pixels have a value of 0:



**Water Bodies - Reclassified from Land Use**

#### **Water Bodies - Reclassified from Land Use**

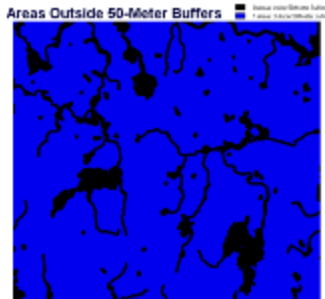
Now we will rasterize our vector streams into this layer (using the LINERAS module in Idrisi) to create a raster layer with streams and other water bodies all represented as non-zero pixels:



**Water Bodies with Streams Added**

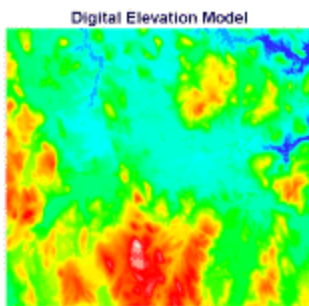


We now create buffers around all of the water features. The suitable areas we are looking for are those areas OUTSIDE the 50-meter buffers around water features. We can use the BUFFER module in Idrisi and specify that the areas within the buffers will have pixel values of 0 and areas outside the buffers will have values of 1. This image is the second of our boolean criterion images:



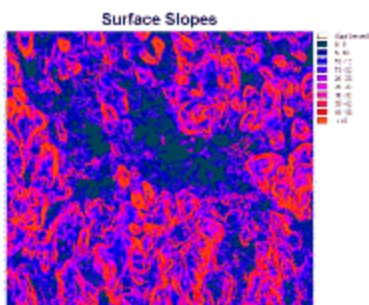
**Areas Outside 50-meter Buffers**

The third criterion we wish to work with is the restriction that slopes in excess of 15% are not permissible within our target areas. We start with a DEM of the study area:



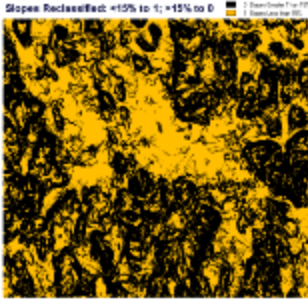
**Digital Elevation Model**

From the DEM we can determine the slopes throughout the study area. In Idrisi this is done using the SURFACE module. This is the result of running SURFACE to create a slope image. The value of each pixel is the slope of that pixel in percent, which is calculated from the DEM based on the elevation value of that pixel relative to its nearest neighbors. The continuous slope image has been reclassified into slope categories for display:



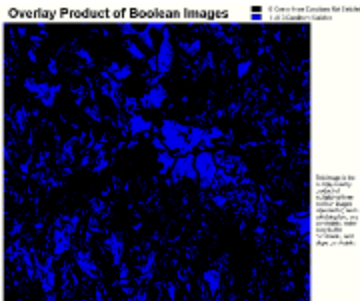
**Surface Slopes**

We reclassify the slope image to create a boolean image to represent our slope criterion; areas with slope less than or equal to 15 percent get a pixel value of 1, areas where slopes exceed 15 percent get a value of 0:



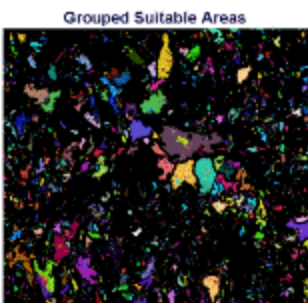
**Slopes Reclassified: <15% to 1; >15% to 0**

We have created boolean images to represent three of our criteria: (1) areas outside 50-meter buffers around water bodies and streams, (2) areas with current land use of crop land, forest land, pasture land, or open undeveloped, and (3) areas with slopes less than or equal to 15%. At this point we perform overlay-multiply steps between the three images. This can be done by performing overlay-multiply between two of the images and then overlay-multiplying the result of that operation by the third image. In Idrisi, Image Calculator permits the three images to be overlay-multiplied in one step. Regardless of the sequence of operations, the following boolean raster image is produced:



**Overlay Product of Boolean Images**

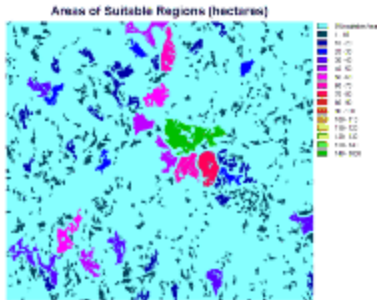
Now that we know what parts of our study area satisfy the three criteria, we need to evaluate which of the satisfactory areas are greater than 50 hectares. In Idrisi we would use the GROUP and AREA modules. The GROUP module looks for contiguous blocks of pixels with the same value. For each block that it finds, each pixel within that block is given a unique ID value. GROUP produces this image:



**Grouped Suitable Areas**

The AREA module is now applied to the raster image created by GROUP. AREA creates an image in which each pixel value is equal to the area of the group to which that pixel belongs. One thing to note is that GROUP also considered the background areas, areas which did not satisfy our criteria, as possible groups. Now that we have run AREA, the AREA module calculated areas for non-target portions of our study area. To keep these areas, which we

rejected from consideration based on our boolean analysis, from confusing our analysis, we overlay-multiply the result from AREA by the previous boolean multiplication product. This essentially "zeroes-out" the areas we are not interested in. After performing this "zeroing-out" step and reclassifying the image into categories for display the result of our area analysis looks like this:



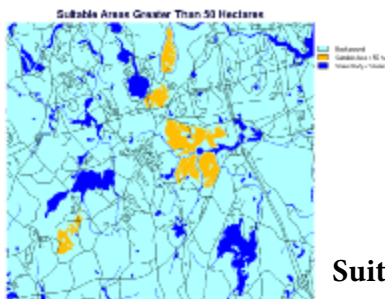
**Areas of Suitable Regions (hectares)**

We can reclassify this image to exclude all areas which do not exceed 50 hectares. When we do so the following areas remain for consideration:



**Suitable Areas Greater Than 50 Hectares**

If we were to continue with our attempt to evaluate a potential site for our proposed activity, we would probably be prudent to consider other contextual variables of the candidate locations. If we display water bodies, streams, roads, and our candidate properties together, it may provide some perspective on the areas we are considering.

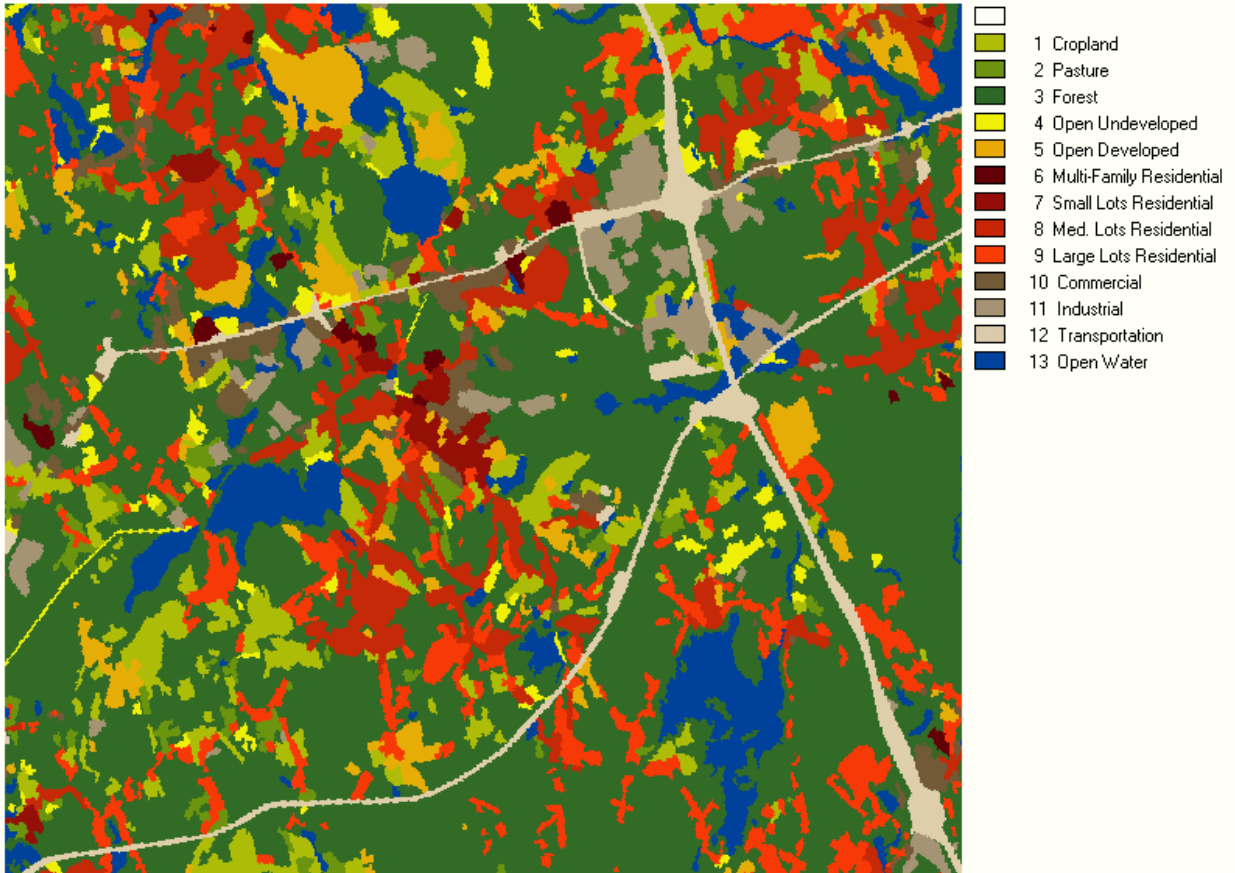


**Suitable Areas Greater Than 50 Hectares**

---

Created: May 14, 1997. Last updated: October 5, 1998.

# Land Use

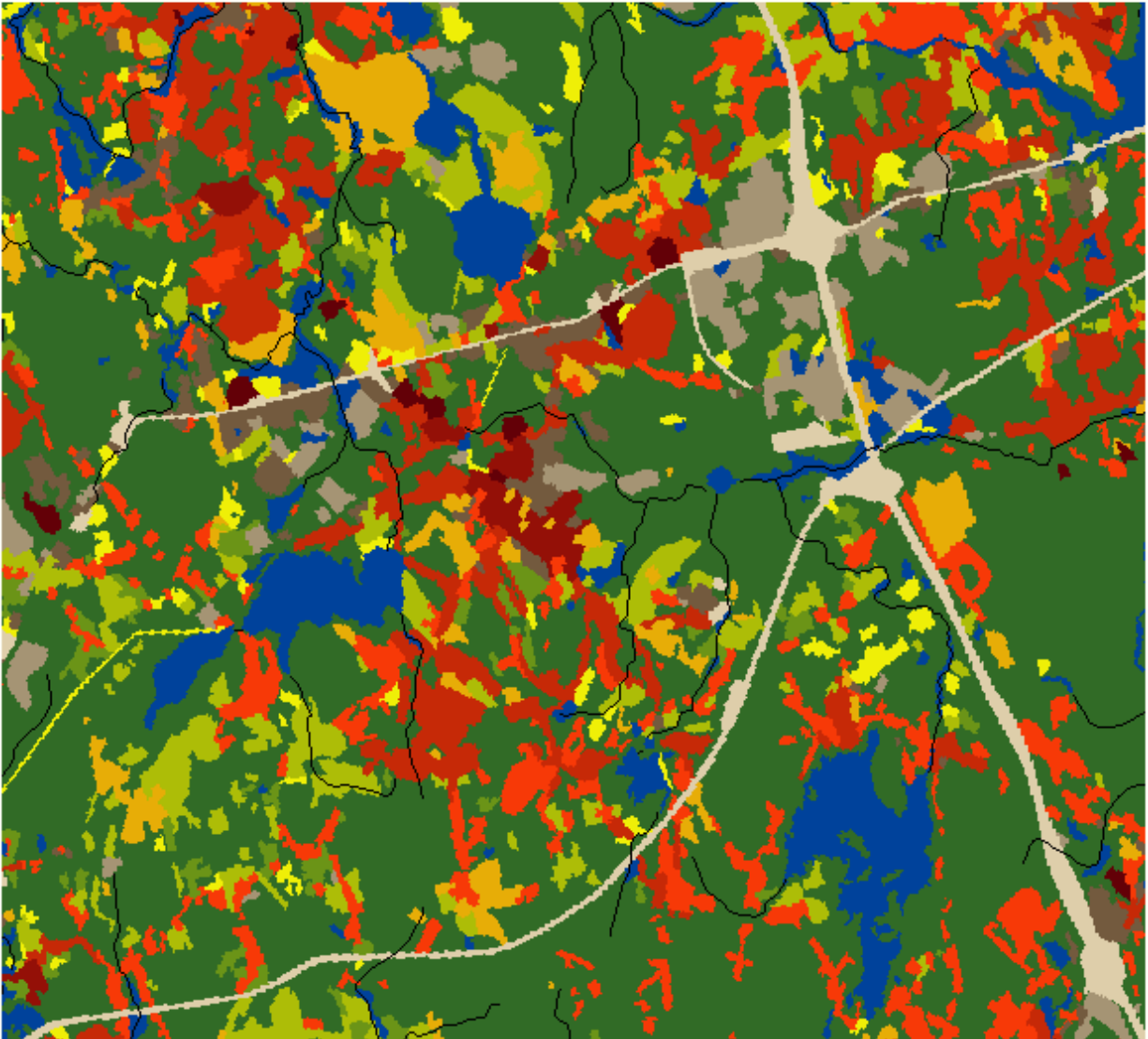


# Land Use Reclassified

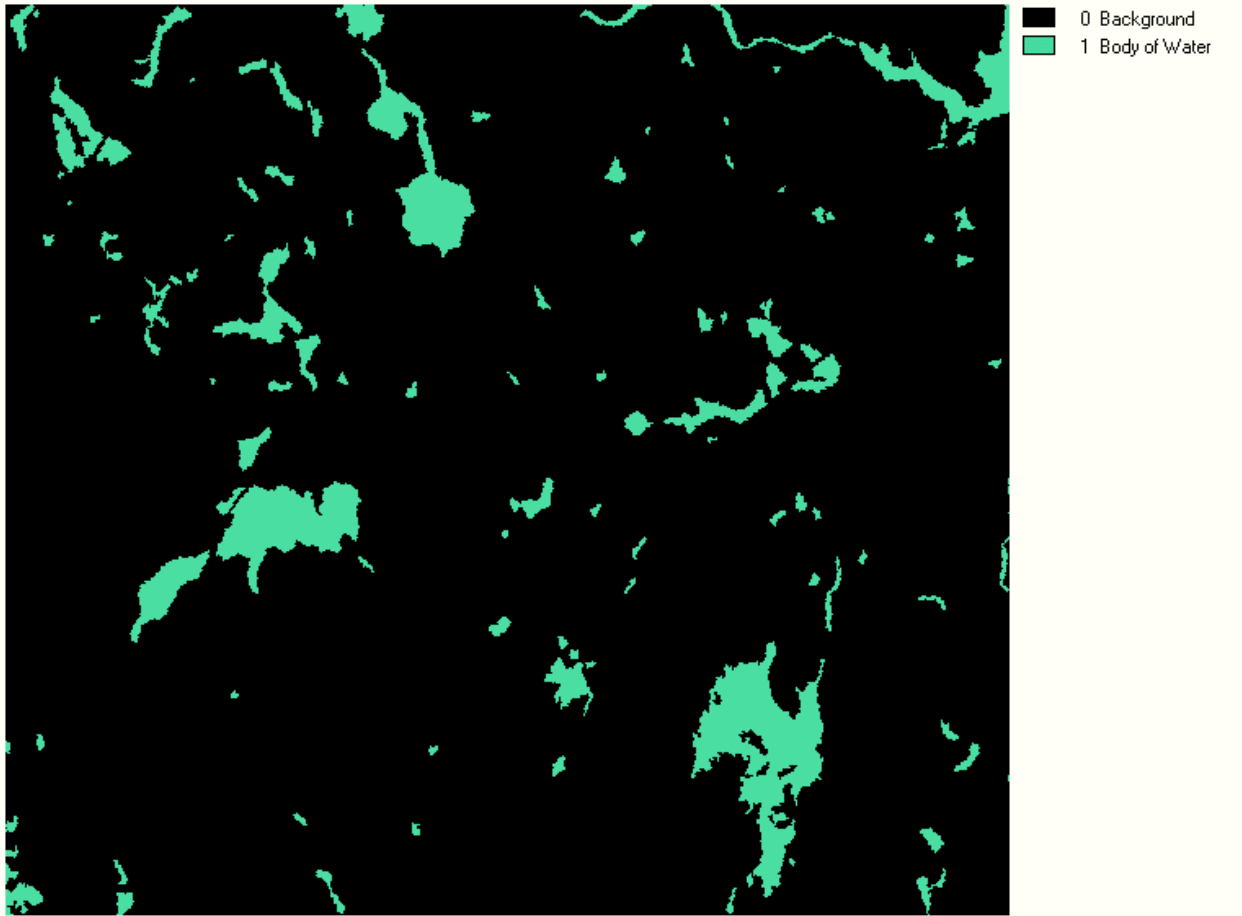
- 0 Unsuitable Land Use for Development
- 1 Suitable Land Use for Development



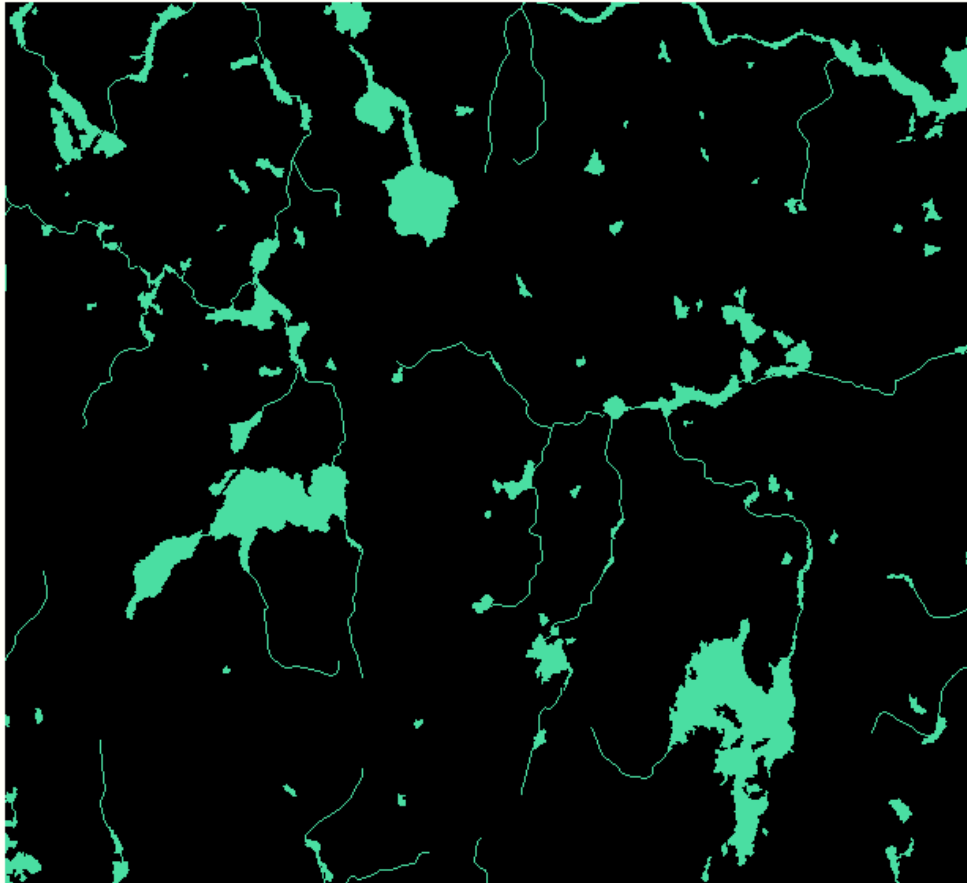
## Land Use with Streams



## Water Bodies - Reclassified from Land Use



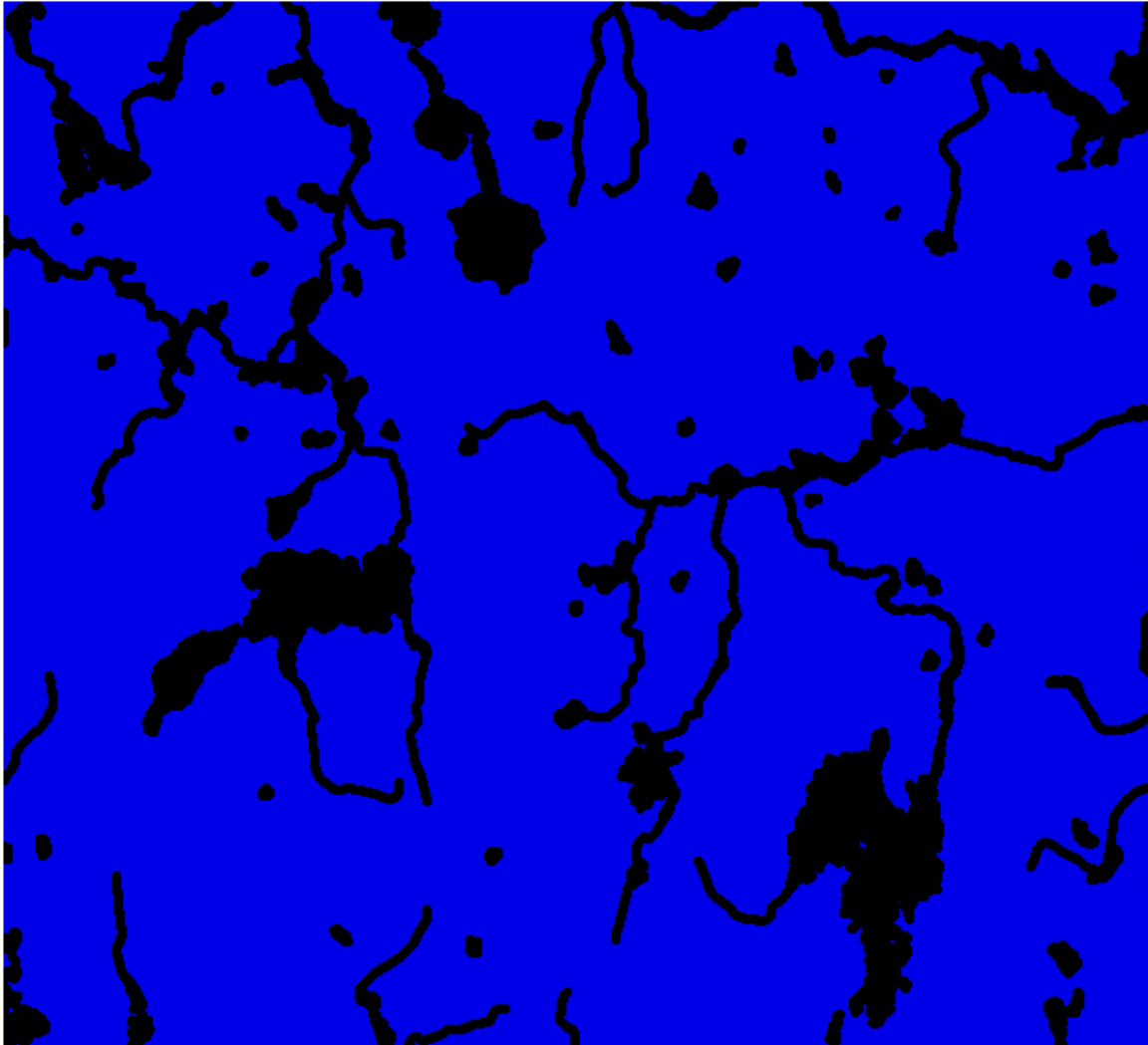
## Water Bodies with Streams Added



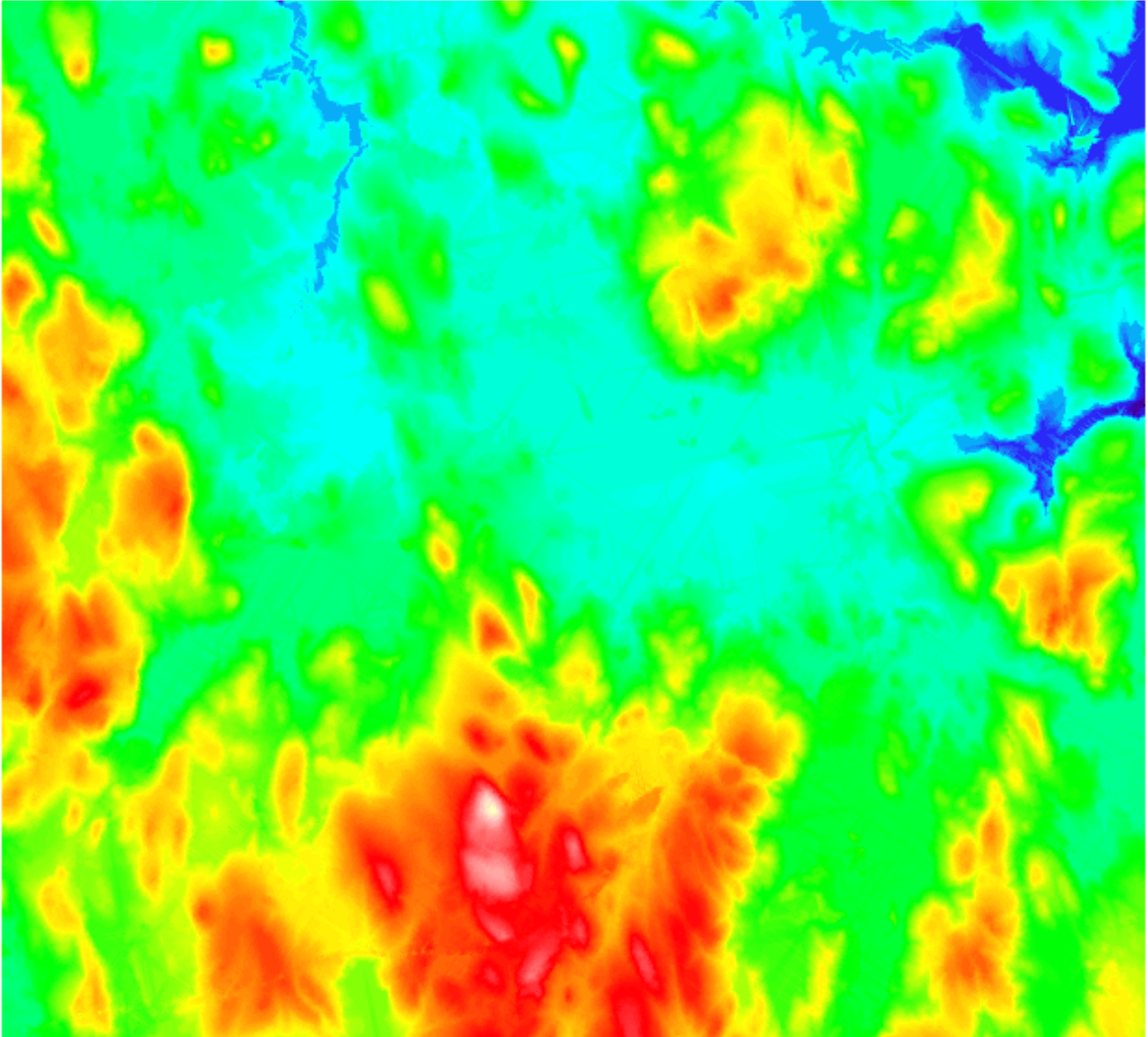


## Areas Outside 50-Meter Buffers

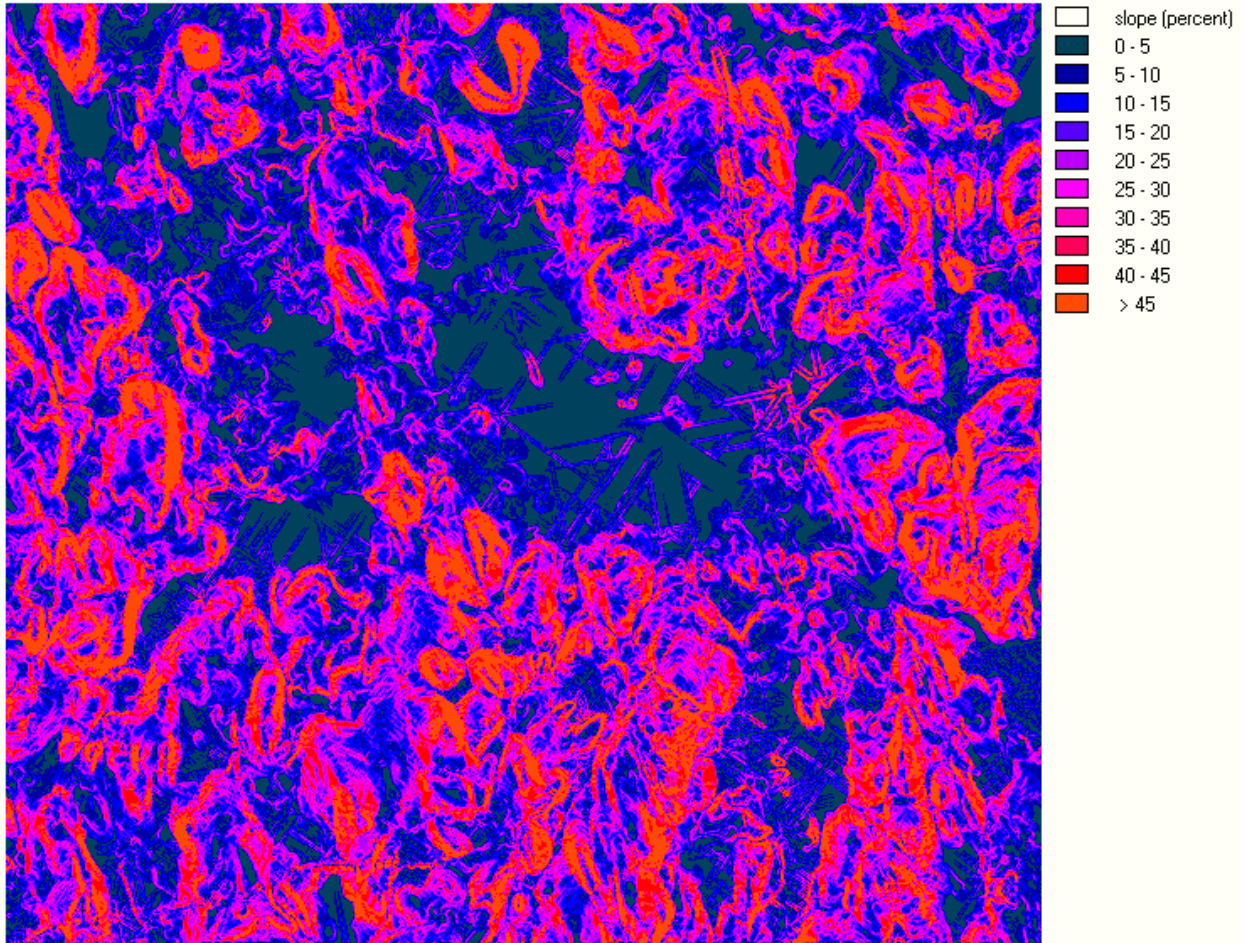
- 0 Areas Inside 50-Meter Buffers
- 1 Areas Outside 50-Meter Buffers





# Digital Elevation Model

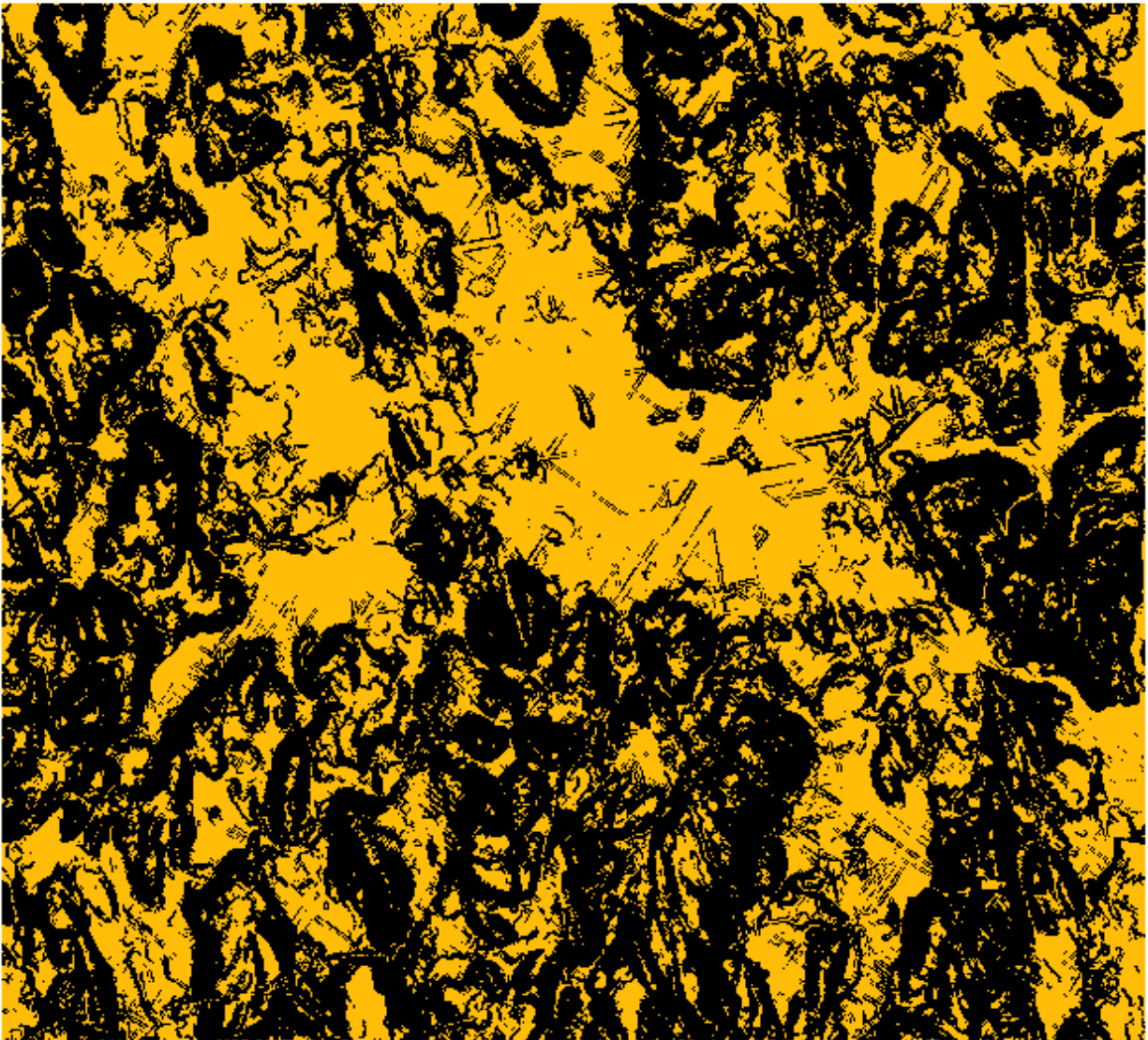


# Surface Slopes



**Slopes Reclassified: <15% to 1; >15% to 0**

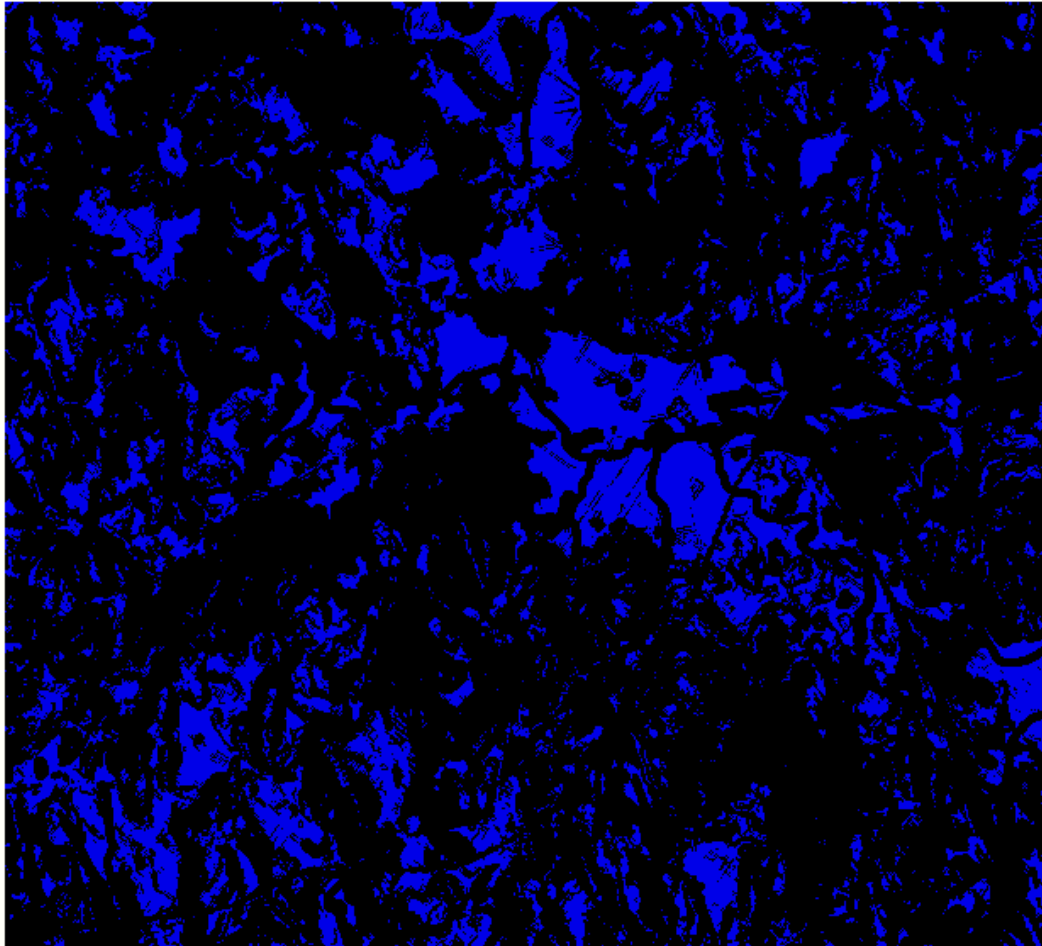
-  0 Slopes Greater Than 15%
-  1 Slopes Less than 15%





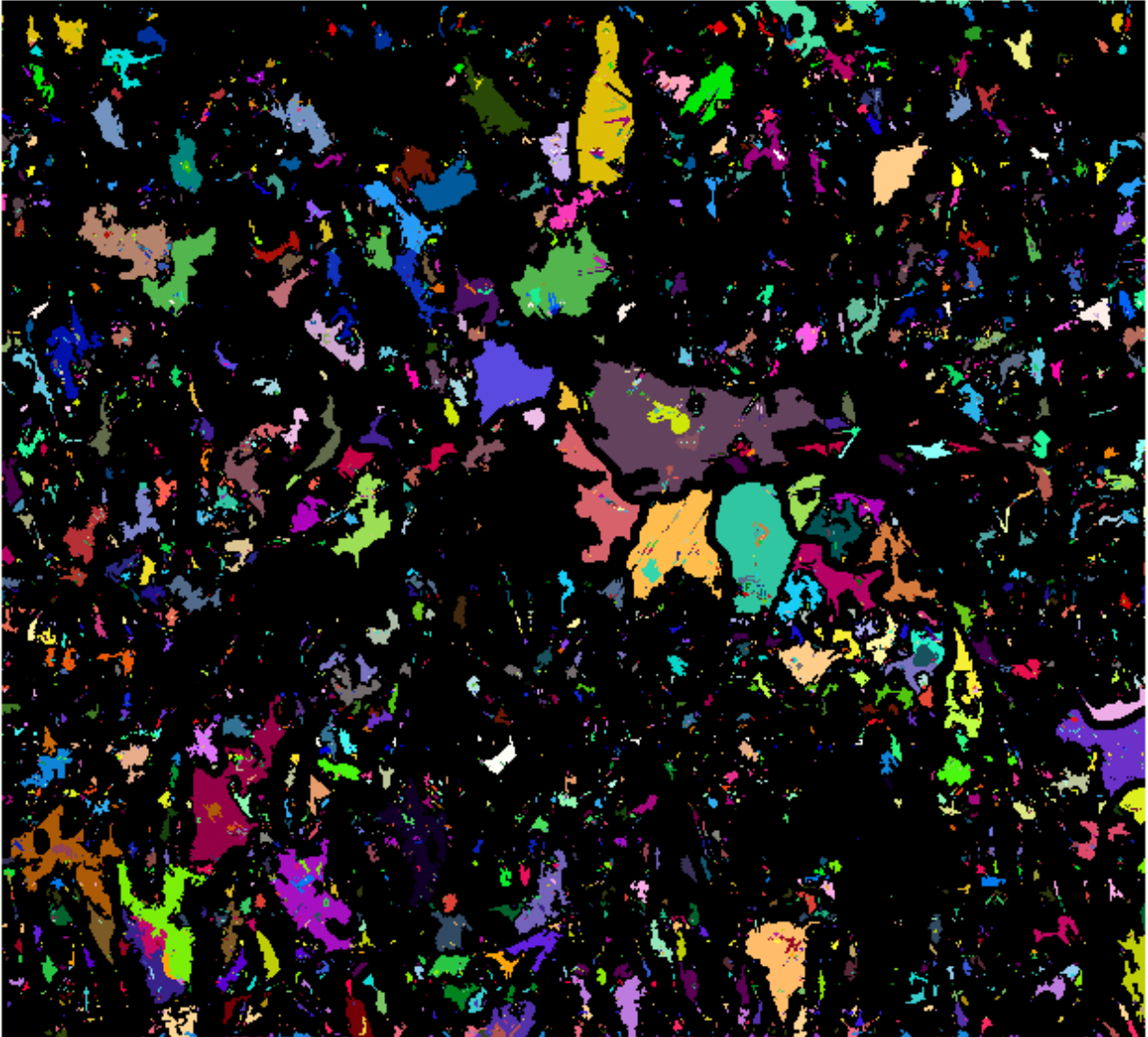
## Overlay Product of Boolean Images

- 0 One or More Conditions Not Satisfied
- 1 All 3 Conditions Satisfied

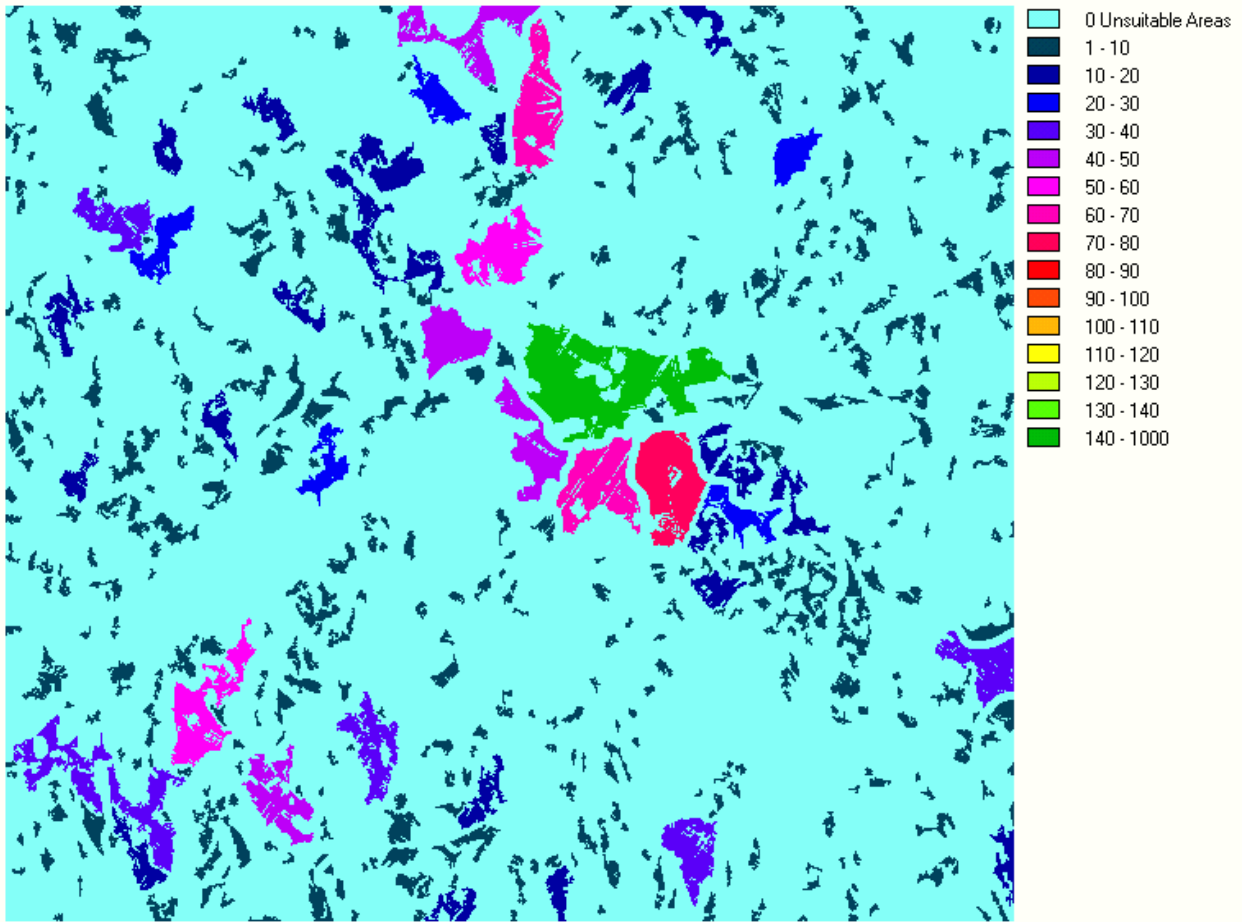


This image is the multiply-overlay product of multiplying three boolean images representing areas satisfying land use constraints, water body buffer constraints, and slope constraints.

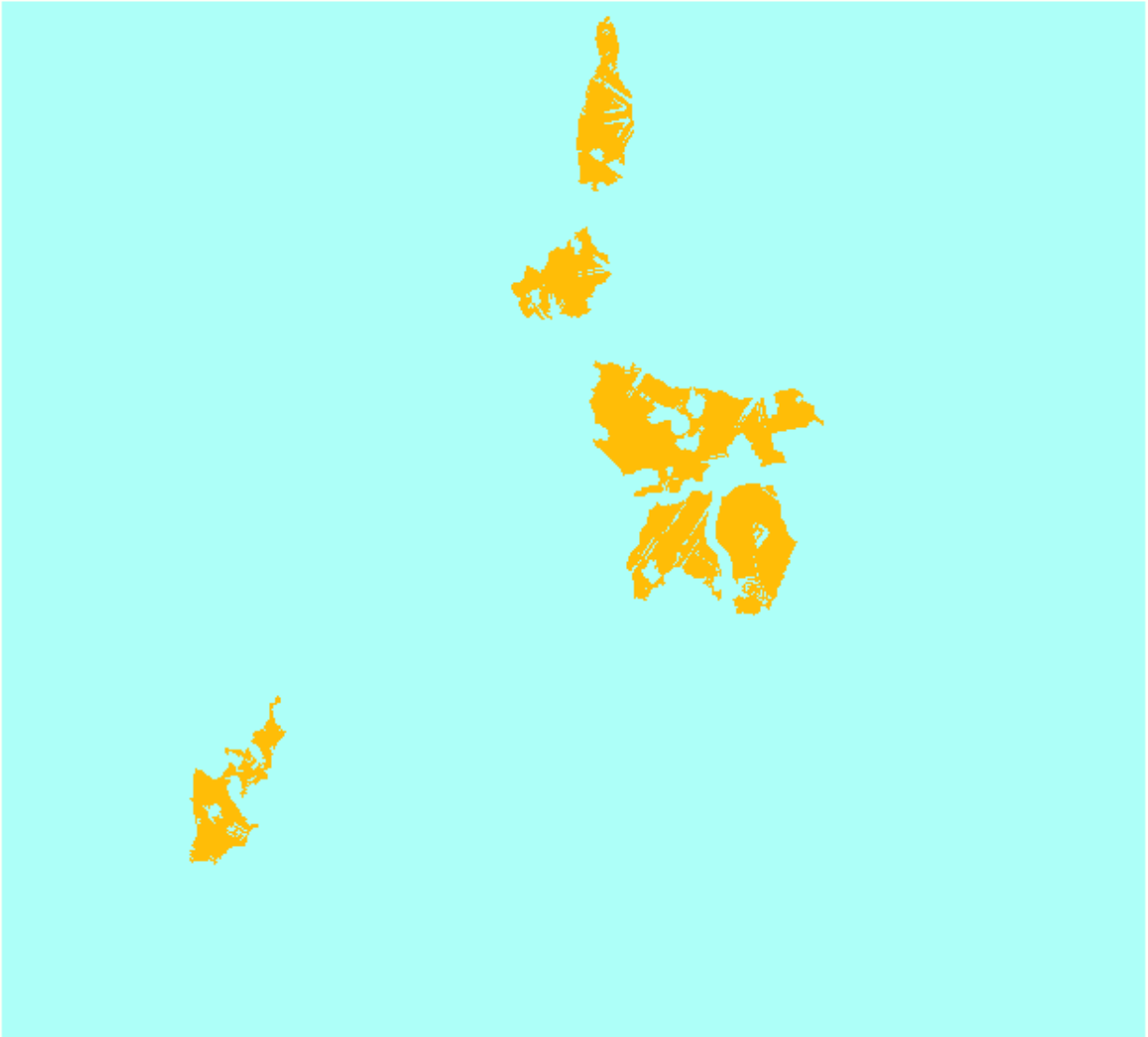
## Grouped Suitable Areas



### Areas of Suitable Regions (hectares)



## Suitable Areas Greater Than 50 Hectares





### Suitable Areas Greater Than 50 Hectares

