RULE BASED AGGREGATION OF RASTER IMAGE CLASSIFICATIONS INTO VECTOR GIS DATABASES WITH FIVE- AND FORTY-ACRE MINIMUM SIZE MAPPING UNITS *

Kenneth A. Stumpf
Geographic Resource Solutions
1125 16th Street, Suite 213
Arcata, CA 95521

John M. Koltun
Geographic Resource Solutions
1125 16th Street, Suite 213
Arcata, CA 95521

ABSTRACT

Pixel data developed using image classification techniques is frequently difficult to convert to a vector GIS format due to the heterogeneity of the pixel data. Conventional mathematically based filters often cause degradation of type boundaries and attributes. A rules based approach to aggregating pixels and resulting vegetation types based on the similarity of the data being mapped is presented. Aggregation was accomplished to both five- and forty-acre minimums for a 212,000 acre portion of the six-million acre Project Area. Preliminary comparative results and findings of this rules-based pixel aggregation to five- and forty-acre minimum type sizes are presented and discussed.

INTRODUCTION

In late 1990, the California Department of Forestry and Fire Protection (CDF) awarded Geographic Resource Solutions (GRS) the Klamath Province Mapping Pilot Study. This pilot study was to explore the methodologies and logistics required to utilize image classification and automated mapping techniques to map vegetation characteristics of six-million acres of rugged terrain in northern California. The study required that Landsat TM satellite imagery be used as the data source to identify and map vegetation types defined by the Wildlife Habitat Relations (WHR) classification system (Mayer and Laudenslayer, 1988).

A major objective of the project was to develop and use methodologies that would be accurate and repeatable. GRS developed a project work-flow utilizing vendor supplied software and GRS utilities and programs that have integrated the use of image processing techniques with grid modeling and GIS analysis. This methodology incorporates rules-driven pixel aggregation in a GIS environment and allows for the future modification of classification rules without having to reclassify the imagery. Also included in this project was a comparison of databases created using a five-acre minimum mapping unit (mmu) database versus a forty-acre mmu database.

* Presented at the ASPRS Conference in August 1992 in Washington, D.C.

GRS-designed pixel aggregation techniques based on decision weights were used to develop and maintain type boundary and attribute integrity of both these databases.

This paper describes the pixel aggregation methodologies used by GRS to create the WHR characteristic database from satellite image classification data (see companion paper by Brown and Fox, 1992) and the comparative results of mapping to five- and forty-acre minimum size mapping units.

AGGREGATION METHODOLOGY

The results of the image classification processes were raster grids that normally contained 150-200 classes and sometimes as many as 300 different classes. Large groups of homogeneous pixels that met the minimum size mapping unit criterion were rare. The pixel grids consisted of a very heterogenous mix of pixels that were frequently isolated or found in very small groups of pixels that did not form valid size types.

The typical method of resolving problems of isolated individual or small groups of pixels that differ from their neighboring pixels is to filter the pixel data and remove these undesirable pixels. The value of the undesirable pixel(s) is altered to smooth the image and produce a cleaner database. Modal filters, or other mathematically based filters, are frequently used to alter the value of the undesirable pixel. When using a modal filter, the pixel value is changed to reflect the pixel type that occurs most frequently in the immediate area (window) evaluated around the undesirable pixel. This approach may be appropriate for pixels completely surrounded by another type of pixel. However, modal filters may be quite inappropriate when used to smooth or clean pixel data representing small homogeneous groups of data or along the edges of different types, where mixed pixels are commonly found. In these situations we found that modal filtering had two negative impacts: the edges of stand boundaries were moved, and the stand characteristics or attributes were incorrectly changed. The mathematical filtering of a pixel(s) definitely effects the delineation of stand types, since the boundaries will be modified to include the modified pixel(s). Filtering may also alter the stands characteristics to reflect a different, less representative type than the previous stand type. This situation occurs in particular along the edges of different types, such as a conifer tree type and a grassland type. If enough mixed pixels, that represent some level of tree cover, are converted into the grassland type, a minimum percent tree cover (e.g. 10 percent) threshold required for a tree type definition may be exceeded and the grassland type suddenly becomes a low density conifer type. GRS developed filters and algorithms designed to overcome these inadequacies and instead reclassify pixels based on the similarity of the undesirable pixel(s) to its neighboring pixels.

Aggregation Rules

The project workplan required a minimum size mapping unit of forty acres for the entire six-million acre Project Area. In the 212,000 acre portion of the Project Area located in Humboldt County, referred to as the Wildlife Pilot Project Area, a separate map was developed using a minimum size mapping unit of five acres.

Some vegetation types were recognized as significantly different types that should not be aggregated with each other, if possible. For example, non-tree types such as small clearcut areas, brushfields, prairies, or bodies of water should not always be merged into surrounding vegetation types such as coniferous forest, montane hardwood/conifer, and hardwood types. Preservation of distinctly different types helps to maintain the accuracy of the mapping effort since fewer stands are generated that represent a mixture of significantly different types, solely for the purpose of satisfying a minimum size mapping unit. The minimum size mapping unit obviously affects the accuracy of any map to represent what is present on the ground. The larger the minimum size mapping unit, the greater the probability that the polygon represents a diverse grouping of heterogeneous types that could have been represented by smaller, more homogeneous polygons if the minimum size mapping unit were smaller.

GRS attempted to maintain polygon boundaries between distinctly different types of data. Each major vegetation type could be assigned its own minimum size mapping acreage thereby enabling the continued separation of significantly dissimilar types. For the final forty-acre minimum size map, tree types were assigned a forty-acre minimum size while non-tree types, such as shrub, herbaceous/grassland, urban, water, and barren were assigned a minimum mapping size of only ten acres. This enabled the continued segregation of small but significant types that could not be merged without a potential loss or degradation of information due to inclusion within another stand.

The aggregation of pixels into polygons is based on two concepts:

1. Data, as both pixels and sub-minimum size mapping units or stands, should be aggregated based on a hierarchial set of rules that merges sub-minimum area groups of pixels and polygons into the most similar neighboring type.

2. Data aggregation should be compensating. The sum of the pixel acreages by WHR characteristic should be equal to the sum of the polygon acreages by WHR characteristic. This would tend to indicate that bias is not introduced during the aggregation process. The hierarchial rules must be developed and implemented without bias towards any particular type or class. Note that this principle cannot always be demonstrated as there may be new types that evolve from the

aggregation of pixels. For example, there are individual pixels that represent the canopy structure type, as either even or uneven-structure. However, there may be polygons that are characterized as uneven-structure that contain no uneven-structure pixels. The uneven-structure characteristic is determined from the canopy cover by size class distribution at the polygon level rather than the structure classes of the pixels within the polygon. In addition, there may be small types often less than the minimum size mapping unit which will be lost if they must always be aggregated with other types to form forty-acre minimum stands.

The initial unfiltered image classification results indicate vegetation type, average size, canopy cover, and structure and are in the form of individual data elements or pixels approximately 0.2 acres in size. Each of these individual pixels should be viewed as an estimate of the WHR characteristics of the area represented by that pixel. These individual data elements must be aggregated to form polygons or stands representative of groups of pixels having similar WHR characteristics. Most often, the homogeneous groups of pixels that are identified are below the minimum size mapping limitations and do not constitute a valid size vegetation type by themselves. These sub-minimum size mapping units must be merged or absorbed into one of their adjacent cover types. **GRS** uses an aggregation program, called **GRS_aggregate**, that estimates which of the adjacent groups of pixels (stands) is most similar to the sub-minimum size mapping unit. The subject stand is then merged with the adjacent stand estimated to be most similar. This aggregation process is used to filter individual pixels during the initial stages of aggregation, as well as to aggregate sub-minimum size mapping units into valid size polygons.

Pixel Level Aggregation

At the individual pixel level, aggregation follows the hierarchial order of significance as defined by the WHR classification rules. Individual pixels and small groups of pixels (isolated blocks no larger than 3 by 3 pixels), that are obviously below the minimum size mapping unit are aggregated with the neighboring group of pixels of the most similar type of vegetation characteristics. During the estimation of similarity, the following order of vegetation forms is recognized: conifers, montane hardwood/conifer, hardwood, shrub/brush, herbaceous/prairie, barren, urban, and water. At this level of aggregation, the estimated size of the trees, percent crown closure, percent conifer, and predominant species are also evaluated.

In order to apply the guided filtering routines, the unfiltered image classification is first reclassified to reflect the WHR characteristics represented by each pixel. The pixel filtering is applied to the reclassified image in multiple stages using two separate filtering routines. The first filter reclassifies the single isolated pixels into the most similar immediately adjacent type. This filter considers pixels that have already been processed to determine the optimal reclassification. The second filter removes isolated islands of like pixels that do not extend beyond a 3 by 3 pixel window. Pixels that have already been changed are also considered by this filter. The pixels of the isolated islands and groups are compared to adjacent groups and reclassified into the most appropriate group. The nature of these filters requires two to three applications of each filter depending on the complexity of the classification and the number of groups created. As these are grid processes, it is important to reclassify those pixels most similar to surrounding pixels before reclassifying those pixels that are fairly dissimilar to all neighboring pixels.

Pixel Group/Polygon Aggregation

Pixel level aggregation results in groups of pixels that range from small groups of as few as three or four pixels, to large groups that already exceed the minimum area requirements and form valid stands. The sub-minimum size groups, or types must be aggregated with other sub-minimum area groups, or with valid size groups to form a thematic database of valid size types. When sub-minimum size groups of pixels (stands) are recognized, the sub-minimum size mapping group is aggregated with the adjacent group estimated to be most similar to the subject area. The characteristics of each neighboring group are evaluated to develop an index of similarity. All of the WHR characteristics (vegetation type, canopy density, size class(QMD), and canopy structure) are used in this evaluation. In addition, several other attributes are used in this evaluation of ecologically associated types: these values are the major vegetation type class (i.e. conifer, hardwood, shrub, herbaceous, and so forth), the estimated percent tree canopy closure, the estimated size (quadratic mean diameter(QMD)), the predominant species, and the percent conifer composition.

Aggregation is based on the premise that if all the characteristics of adjoining groups are similar except for one, such as the tree canopy cover, then the most similar adjoining group is the group with the most similar (least different) tree canopy cover. Differences between the subject group's characteristics and the adjacent groups' characteristics are numerically estimated to enable a quantitative estimate of similarity. The differences of stand characteristics are estimated as absolute values and then summed to estimate a similarity index.

Aggregation choices that involved only one different characteristic are relatively easy to make as compared to choices involving multiple differences. Most often, multiple differences, such as tree canopy cover, QMD, and WHR type, are observed. Some of these types of differences (WHR type versus QMD) are more significant than others. Levels of significance were estimated and represented by assigning weights or factors to the type of difference being estimated. The factors and relationships used to develop similarity indices were attempts to recognize differences between WHR types and characteristics and reflect the WHR classification rules. These weights and factors can be modified to reflect other interpretations of ecological associations and the significance of the vegetation characteristics.

The implementation of a rules based aggregation of isolated pixels and sub-minimum area groups of pixels or stands using estimates of similarity is a relatively new capability and is heavily reliant on ecological relationships and concepts. For example, the associations of Oregon white oak with grassland types and tanoak with coniferous types must be known and included in the estimation of similarity. Hardwood pixels should be merged with the type that the specific hardwood type is associated with in its natural range rather than generalizing and always putting a hardwood type with a conifer type rather than a grassland type. Wildlife habitat requirements should also play a significant role in the determination of the significance of different characteristics and the role any characteristic plays in the aggregation of data. If wildlife respond to tree size more than canopy cover, then the aggregation process should preserve groups by size representing a variety of densities rather than groups by density representing a wide range of sizes. Size groupings/class boundaries may also be significant with respect to mapping wildlife habitat characteristics. These types of relationships may be true for some wildlife species and not others which indicates that different maps of wildlife characteristics could be developed for different species that each may have their own set of aggregation rules. Similarly, a botanist may be more interested in species purity or diversity and may develop a different set of rules that accentuates species differences.

In this pilot project, the hierarchial rules tended to place greater emphasis on the similarity of major vegetation types, tree vegetation classes, the predominant conifer species type, the canopy structure, the percent conifer composition of the stand, and the estimated QMD of the trees present. Canopy cover was also considered, but size (QMD) appeared to be a more significant factor in the aggregation process than was the canopy cover. Thus, areas of similar seral stage of different densities would be grouped rather than grouping areas of different seral stages of the same density class.

Aggregation was performed in several steps, starting with low minimum acreage limits and progressing to limits of five- and forty-acres as identified in the objectives of the project. The aggregation process performed in one step, from the initial groups to the final limits is difficult to process and appeared to result in larger, more generalized types, than a step-wise aggregation process. The step-wise aggregation process involved smaller size increases and tended to merge smaller numbers of stands at each step and thereby maintain stands of similar characteristics rather than merging many small stands at once into a few large generalized stands.

An advantage of the step-wise approach is that maps reflecting different minimum size mapping units are developed as intermediate products of the aggregation process. The five-acre map was generated from the intermediate results of aggregation using the intermediate five-acre limits. The forty-acre map was the result of continuing the aggregation process using intermediate ten- and twenty-acre limits and the final forty-acre limit. These databases were then vectorized using standard vectorization routines. The vectors were then smoothed, using GRS software, ***segjoin***, to remove the stair-step appearance characteristic of vectors derived from pixel(raster) databases and to reduce the size of the database. The characteristics of the final polygons were then determined and loaded into the relational database tables.

Estimation of Polygon Characteristics

The final estimate of each stand's WHR characteristics is based on the summarization of all the different types of unfiltered image classification pixels found within each of the final stand boundaries. The pixels of all sub-minimum size mapping units that are merged into a final polygon are included in the polygon summaries. Therefore, inclusions of sub-minimum size mapping units are contained in the polygon pixel summaries.

Each polygon summary of unfiltered image classification pixels yields an estimated distribution of canopy cover by species and size class. An example of one of these distributions is shown in Table 1. Each polygon's characteristics are then estimated by evaluating the distribution of cover by species and size class, using the WHR classification rules and definitions.

Data Definition

The characteristics estimated for each polygon are listed in Table 2. These characteristics included the WHR vegetation type, the canopy closure class, the size class, and the canopy structure class, as well as other data items that provided additional descriptive information about each stand. The estimation of these additional values, such as the estimated QMD or the specific percent of canopy cover enables future reclassification of these stands using modified or alternative class definitions. These additional data items were used during the aggregation process to estimate the similarity of adjacent stands.

Table 1
Polygon Cover Density Summary

Stand: 1010
Total Number of Pixels: 498
Contributing Pixels: 498

WHR Type: Douglas-fir
Canopy Structure: UNEVEN
Size Class: 5 (36"+)
Density Class: SPARSE (12.7%)

Size Class: / Species	0-5"	6-10"	11-23"	24-35"	36"+	Tree Cover	Non-Tree Cover	Total Cover
Douglas-fir	3.8%	8.2%	10.4%	3.4%	6.5%	32.3%		32.3%
redwood	2.6%	2.0%	2.9%	0.7%	6.2%	14.3%		14.3%
clos cone pine	0.1%	4.1%	6.1%	0.0%	0.0%	10.3%		10.3%
misc. conifer	0.0%	0.8%	0.1%	0.0%	0.0%	0.9%		0.9%
hardwoodC	3.6%	6.2%	5.7%	0.5%	0.0%	15.9%		15.9%
hardwood	0.0%	0.4%	0.7%	0.6%	0.0%	1.7%		1.7%
shrub/brush							4.2%	4.2%
cham/art							4.5%	4.5%
misc shrub							0.9%	0.9%
forb/herbac							1.9%	1.9%
duff/debris							6.1%	6.1%
noncontributor							4.1%	4.1%
Total Cover	10.2%	22.0%	26.4%	5.3%	**12.7%**	76.6%	23.5%	100.0%
Total Tree Cover						76.6%		

Stand Tree Density Summary:
Stand: 1010

Size Class: / Species	0-5"	6-10"	11-23"	24-35"	36"+	Tree Cover
Douglas-fir	5.0%	10.7%	13.6%	4.4%	**8.5%**	42.1%
redwood	3.3%	2.7%	3.8%	0.9%	8.1%	18.7%
clos cone pine	0.1%	5.4%	8.0%	0.0%	0.0%	13.4%
misc. conifer	0.0%	1.0%	0.1%	0.0%	0.0%	1.1%
hardwoodC	4.7%	8.1%	7.4%	0.6%	0.0%	20.8%
hardwood	0.1%	0.5%	1.0%	0.8%	0.0%	2.3%
Total Tree Cover	13.4%	28.7%	34.4%	6.9%	16.6%	100.0%

Stand Quadratic Mean DBH Summary:
Stand: 1010

Size Class: / Species	0-5"	6-10"	11-23"	24-35"	36"+	Tree Cover
Douglas-fir	3.8"	7.9"	16.2"	28.8"	48.9"	25.4"
redwood	4.1"	8.2"	16.0"	29.4"	96.8"	64.0"
clos cone pine	4.2"	8.4"	13.6"	0.0"	0.0"	11.8"
misc. conifer	0.0"	6.9"	13.5"	0.0"	0.0"	7.7"
hardwoodC	4.0"	8.0"	16.1"	28.8"	0.0"	12.1"
hardwood	3.8"	6.5"	14.3"	31.4"	0.0"	20.6"
Quad Mean DBH	3.9"	8.0"	15.6"	29.1"	76.5"	33.2"
Quad Mean DBH - Con	3.9"	8.0"	15.4"	28.8"	**76.5"**	37.3"
Quad Mean DBH - Hwd	4.0"	8.0"	15.9"	30.3"	0.0"	13.1"

Table 2
Stand Attribute Definitions

Data Item	Format	Description
stand_id	integer	Stand (polygon) ID Number
whrtype	char(3)	WHR Type
pr_species	char(14)	Predominate Species
closure_class	char(1)	WHR Closure Class
density	smallfloat	Pct Tree Canopy Cover (TC)
other_cover	smallfloat	Pct Other Veget. Cover
pct_conifer	smallfloat	Pct Conifer Cover of TC
pct_hdwood	smallfloat	Pct Hardwood Cover of TC
size_class	smallint	WHR Size Class
qmdbh	smallfloat	Quadratic Mean Diameter
structure	char(1)	WHR Structure Class
forest_land	char(1)	Forestland Type Flag
acreage	smallfloat	Acreage

RESULTS AND DISCUSSION

5/40 Acres - Minimum Mapping Unit Size

The choice of the appropriate minimum size acreage used to develop a thematic database, such as a WHR database, must include consideration of the quality and accuracy of the database as well as the usefulness and ease of use of the information. Very detailed specific stand data for very small size stands may be very accurate but unusable due to the magnitude of the database. Similarly, a less detailed more generalized database may be of the appropriate size, but it may not include the specificity required to develop reliable information. These issues of size and specificity must be balanced to provide a usable accurate database that meets the analytical needs of the assessment.

A subset of the Project Area located in Humboldt County and covering approximately 212,400 acres was mapped at both the five-acre and forty-acre minimum size mapping units to develop comparative databases. These maps of this area are preliminary in nature, as of the time of this writing, as they were developed at a point in the project prior to finalization of the image classification, pixel filtering, and aggregation processes used to map the entire Project Area. Final five-acre and forty-acre maps of the Study Area will be developed consistent with the final classification of the entire Project Area. The data contained in these preliminary maps are useful to project trends and relationships associated with aggregation and different minimum size mapping limits.

The forty-acre map was not based on an absolute forty-acre minimum as the aggregation rules allowed ten-acre minimum size units for non-tree vegetation types. The five-acre map does represent a five-acre minimum for all types. This map was the result of interrupting the aggregation process after reaching the five-acre minimum size limit and saving the intermediate results. The forty-acre map was generated by allowing the aggregation process to continue until the forty-acre minimum size limit was reached. This process took an additional four hours of processing time. Subsequent vectorization and attribution processes required an additional three to four hours for the five-acre map as compared to the forty-acre map so the end result was that each map took about the same amount of total time to generate.

As was expected, the five-acre minimum unit map resulted in a larger database than the forty-acre minimum unit map. The acreage, number of stands, and average stand acreage by WHR Class are shown in Tables 3, 4, 5, and 6. There were 7897 polygons in the five-acre map and the average polygon size was 26.9 acres. There were 1311 polygons in the forty-acre map and the average size polygon was 162.1 acres. There were roughly six times as many polygons in the five-acre map as the forty-acre map. Interestingly, the average acreages of the stands in each of these preliminary maps are substantially higher, by at least a factor of four, than the minimum sizes used in the formulation of these maps.

The five-acre map was 3.8 megabytes in size whereas the forty-acre map was only 1.2 megabytes in size, or 31.6 percent of the size of the five acre map, as measured in an Intergraph GIS format. Files sizes in an ARC/INFO format are projected to be nearly twice as large due to the different data formats of the two systems. Both of these test databases are of reasonable size, however, mapping the entire Klamath Province (approximately 19 million acres) at these minimum size limits would result in extremely large databases of approximately 342 megabytes and 108 megabytes and would require significantly more processing time to query, model, analyze, and report. The WHR database, for an area representing California, or even the Klamath Province, would represent a very large database if maintained as an individual coverage on a GIS.

Table 3
Acreage by Map and WHR Type

WHR TYPE	5-Acre Map Acreage	5-Acre Map Count	5-Acre Map Average	40-Acre Map Acreage	40-Acre Map Count	40-Acre Map Average
AGS	2,187.2	27	**81.0**	2,158.7	18	**119.9**
BAR	2,088.8	67	**31.2**	1,946.1	31	**62.8**
CPC	811.7	66	**12.3**	317.0	5	**63.4**
DFR	79,523.7	3,409	**23.3**	84,703.2	628	**134.9**
KMC	16,610.3	813	**20.4**	14,044.2	89	**157.8**
MCP	736.9	30	**24.6**	651.7	13	**50.1**
MHC	70,265.0	2,377	**29.6**	76,916.4	411	**187.1**
MHW	14,881.1	748	**19.9**	9,875.5	65	**151.9**
RDW	4,306.0	172	**25.0**	3,708.3	26	**142.6**
UNDF	49.2	7	**7.0**	10.9	1	**10.9**
WAT	14,455.1	18	**803.1**	14,443.4	15	**962.9**
WFR	6,720.8	163	**41.2**	3,728.9	9	**414.3**
Totals	212,635.7	7,897	**26.9**	212,504.2	1,311	**162.1**

AGS = Herbaceous/forb
BAR = Barren
CPC = Closed-cone pine
DFR = Douglas-fir
KMC = Klamath mixed conifer
MCP = Shrub
MHC = Montane hardwood/conifer
MHW = Montane Hardwood
RDW = Redwood
UND = Undefined
WAT = Water
WFR = White fir

Table 4
Acreage by Map and Density Class

Size Class	5-Acre Map Acreage	5-Acre Map Count	5-Acre Map Average	40-Acre Map Acreage	40-Acre Map Count	40-Acre Map Average
UNDF	16,543.9	85	**194.6**	16,389.5	46	**356.3**
DENSE	164,025.1	5,822	**28.2**	167,554.7	983	**170.5**
MODERATE	20,537.7	1,189	**17.3**	21,286.6	187	**113.8**
OPEN	5,571.7	452	**12.3**	4,792.5	62	**77.3**
SPARSE	5,957.4	349	**17.1**	2,480.9	33	**75.2**
Totals	212,635.7	7,897	**26.9**	212,504.2	1,311	**162.1**

Table 5
Acreage by Map and Size Class

Size Class	5-Acre Map Acreage	5-Acre Map Count	5-Acre Map Average	40-Acre Map Acreage	40-Acre Map Count	40-Acre Map Average
UNDF	19,467.9	142	**137.1**	19,199.9	77	**249.3**
0-5"	679.9	61	**11.1**	47.4	1	**47.4**
6-11"	15,363.8	990	**15.5**	8,280.6	81	**102.2**
12-23"	106,484.7	3,743	**28.4**	107,146.8	580	**184.7**
24-35"	29,600.0	1,206	**24.5**	29,964.2	210	**142.7**
36"+	41,039.3	1,755	**23.4**	47,865.4	362	**132.2**
Totals	212,635.7	7,897	**26.9**	212,504.2	1,311	162.1

Table 6
Acreage by Map and Structure Class

Size Class	5-Acre Map Acreage	5-Acre Map Count	5-Acre Map Average	40-Acre Map Acreage	40-Acre Map Count	40-Acre Map Average
UNDF	19,467.9	142	**137.1**	19,199.9	77	**249.3**
EVEN	182,928.0	7,172	**25.5**	173,825.0	1,076	**161.5**
UNEVEN	10,239.7	583	**17.6**	19,479.3	158	**123.3**
Totals	212,635.7	7,897	**26.9**	212,504.2	1,311	**162.1**

A comparison of the five-acre and forty-acre maps was performed by spatially overlaying these two maps and generating reports of how many acres were typed by WHR characteristic. Four tables were generated representing the WHR types, canopy density classes, size classes, and structure classes present in both maps. These tables are identified as Tables 7, 8, 9, and 10 and are presented in matrix form so that it is possible to see how the acres typed in the five-acre map were typed in the forty-acre map and visa versa. For example, in Table 7, 79,475 acres were typed as Douglas-fir type in the five-acre map and 84,659.3 acres were typed as Douglas-fir in the forty-acre map. 66,887.9 acres were typed as Douglas-fir in both efforts. The other 15,000 acres typed as Douglas-

Table 7

WHR TYPE ACREAGE CORRELATION MATRIX

WHR TYPE 40 ACRE TYPES	5 ACRE TYPES													
	AGS	BAR	CPC	DFR	KMC	MCP	MHC	MHW	RDW	UNDF	WAT	WFR	TOTAL	PERCENT
AGS	**2131**	0	0	1	0	0	27	0	0	0	0	0	2159	**98.7%**
BAR	0	**1892**	4	8	1	16	4	0	0	0	18	0	1943	**97.3%**
CPC	0	0	**219**	57	32	0	9	0	0	0	0	0	317	**69.1%**
DFR	7	34	194	**66888**	7377	36	7809	407	1041	8	8	850	84659	**79.0%**
KMC	0	41	271	2728	**6786**	0	984	38	754	6	0	2428	14036	**48.3%**
MCP	13	18	0	0	0	**614**	6	0	0	0	0	0	652	**94.3%**
MHC	29	78	116	8759	1785	70	**59584**	6067	142	11	1	222	76864	**77.5%**
MHW	6	6	0	60	1	0	1438	**8348**	0	13	0	0	9872	**84.6%**
RDW	0	0	0	789	463	0	79	0	**2354**	0	0	21	3706	**63.5%**
UNDF	0	0	0	0	0	0	0	0	0	**11**	0	0	11	**100.0%**
WAT	0	15	0	0	0	0	1	0	0	0	**14267**	0	14442	**99.9%**
WFR	0	0	6	185	157	0	159	15	11	0	0	**3194**	3727	**85.7%**
TOTALS	2186	2085	810	79475	16601	736	70098	14876	4302	49	14454	6714	212389	
PERCENT	**97.5%**	**90.7%**	**27.0%**	**84.2%**	**40.9%**	**83.4%**	**85.0%**	**56.1%**	**54.7%**	**22.2%**	**99.8%**	**47.6%**		

AGS = Herbaceous/forb
BAR = Barren
CPC = Closed-cone pine
DFR = Douglas-fir
KMC = Klamath mixed conifer
MCP = Shrub

MHC = Montane hardwood/conifer
MHW = Montane Hardwood
RDW = Redwood
UND = Undefined
WAT = Water
WFR = White fir

fir in the five-acre map were aggregated into other types in the forty-acre map. 2,728 acres became Klamath Mixed Conifer, 8,758.8 acres became Montane Hardwood/Conifer, 789 acres became redwood, and so forth. Of the approximately 18,000 acres that was not typed as Douglas-fir in the five-acre map, but was aggregated into stands typed as Douglas-fir in the forty-acre map, 7,808.7 acres came from the Montane Hardwood/Conifer type, 7377.4 acres came from the Klamath Mixed Conifer type, 1040.7 acres came from the redwood type, and so forth.

Overall, 78.4 percent of the acres by WHR vegetation type were the same type in both the five-acre and forty-acre maps. 80.0 percent of the acres by Size Class were the same class in each map, 89.7 percent of the acres by Density Class were the same class in each map, and 92.1 percent of the acres by Structure Class were the same class in each map. These figures might be viewed as an indicator of stability with respect to the change in the minimum size mapping unit. A very high level, such as 99 percent would indicate that the change in minimum size mapping unit had little impact on the mapped characteristics of that type and that the definition of that rule was insensitive to changes of the minimum size mapping unit. A low level, such as 25 percent, would indicate the opposite, that the mapped characteristics were very much dependent on the minimum size mapping unit and sensitive to its change. The 'uneven structure' rule is an example of a rule that appears to be quite sensitive to the minimum size mapping level used in this comparison as only a third of the 'uneven structure' acres of the 40-acre map were classified as 'uneven structure in the 5-acre map.

Review of the changes in acreage by type indicates that the aggregation process worked well, as small sub-minimum size areas normally tended to be aggregated into a similar type rather than a completely dissimilar type. Two trends related to the aggregation process and the minimum mapping size are evident in the information in these tables. One trend is related to the percent of the map that a class or type comprises. Types accounting for small portions of the overall acreage tend to decrease in total acreage during the aggregation process whereas the types that are most abundant tend to increase in overall acreage. The sub-minimum size stands of less frequent types are more likely adjacent to stands of the more frequent types than to stands of the less frequent types. Therefore the more frequent types tend to dominate the aggregation process as these types absorb acres of sub-minimum size stands of the less frequent types. For example, there were 6,714.1 acres typed as white fir in the five-acre map. There were only 3,726.8 acres typed as white fir in the forty-acre map. This decrease would tend to indicate that a large portion of white fir stands are less than the forty-acre minimum. These acres were aggregated into other WHR types, most notably the Klamath Mixed Conifer and Douglas-fir types.

Table 8
WHR DENSITY CLASS ACREAGE CORRELATION MATRIX

WHR TYPE 40 ACRE TYPES	5 ACRE TYPES UNDF	DENSE	MODERATE	OPEN	SPARSE	TOTAL	PCT
UNDF	**16351**	12	4	0	18	16386	99.8%
DENSE	48	**158041**	7226	1154	994	167463	**94.4%**
MODERATE	59	5468	**12216**	2216	1312	21272	**57.4%**
OPEN	18	230	976	**1945**	1618	4788	**40.6%**
SPARSE	62	51	104	252	**2010**	2479	**81.1%**
TOTALS	16538	163802	20526	5568	5953	212388	
PCT	**98.9%**	**96.5%**	**59.5%**	**34.9%**	**33.8%**		

Table 9
WHR SIZE CLASS ACREAGE CORRELATION MATRIX

WHR TYPE 40 ACRE TYPES	5 ACRE TYPES UNDF	0-5"	6-11"	12-23"	24-35"	36"+	TOTAL	PCT
UNDF	**19144**	6	20	18	8	0	19196	**99.7%**
0-5"	0	47	0	0	0	0	47	**100.0%**
6-11"	54	291	**6276**	1647	8	2	8278	**75.8%**
12-23"	217	285	8360	**92106**	4577	1535	107081	**86.0%**
24-35"	38	38	621	8054	**17009**	4182	29942	**56.8%**
36"+	8	12	79	4599	7853	**35293**	47844	**73.8%**
TOTALS	19462	680	15356	106424	29456	41012	212388	
PCT	**98.4%**	**7.0%**	**40.9%**	**86.5%**	**57.7%**	**86.1%**		

Table 10
WHR STRUCTURE CLASS ACREAGE CORRELATION MATRIX

WHR TYPE 40 ACRE TYPES	------- 5 ACRE TYPES ------- UNDF	EVEN	UNEVEN	TOTAL	PERCENT
UNDF	**19144**	52	0	19196	**99.7%**
EVEN	316	**169842**	3565	173723	**97.8%**
UNEVEN	1	12801	**6667**	19469	**34.2%**
TOTALS	19462	182695	10231	212388	
PERCENT	**98.4%**	**63.0%**	**65.2%**		

There was not a corresponding shift of acres of other types into the white fir type, indicating that sub-minimum area types of other conifer species adjacent to the larger white fir types were aggregated into other more similar conifer types, such as Douglas-fir and Klamath Mixed Conifer.

The second trend is related to the difference between pure types, such as Hardwood, Douglas-fir, or Even Canopy Structure, and mixed types, such as Montane Hardwood/Conifer and Uneven Canopy Structure. As types are aggregated, the purer types tend to be diluted when merged with other similar but slightly different types. The resulting types are more representative of mixed types or conditions. The five-acre map contains many small, more homogeneous types that often are representative of purer types. The variance of the five-acre map is primarily between the different stands and not within them. For example, the small types generally have a more limited range of canopy heights than the larger types, and each stand generally tends to represent a fairly similar, less variable, group of canopy heights. However, as the smaller stands are aggregated into larger stands, the variance of the map data shifts. The variance between stands tends to decrease and the variance within stands increases, resulting in more stands of mixed types. The diversity of the map is now also contained within stands rather than primarily between the stands. For example, only 10,231.5 acres were typed as 'uneven structure' in the five-acre map and 182,695.2 acres were typed as 'even structure' (see Table 10). The 'even structure' is characteristic of more homogeneous types. Nearly twice the acreage, 19,468.9 acres was typed as 'uneven structure' in the forty-acre map and 173,723.0 acres were typed as 'even structure'. In this case, stands with increased diversity were created as the sub-minimum area stands were aggregated. Resulting stands have sufficient diversity of canopy cover by size class to be classified as 'uneven structure', although none of the original small types might have been that structure type. The mixed type is derived from the aggregation of the small groups of homogeneous but different types into larger types that meet the minimum area requirements. A similar situation exists regarding the increased acreages of the Montane Hardwood/Conifer class and the decrease of the acreage of the Montane Hardwood class.

It is apparent that changes in the minimum size mapping unit can influence the resulting database. Small, less common types may tend to be lost as the aggregation proceeds. If these smaller less frequent types are the basis for critical wildlife habitat requirements, then it appears that the larger minimum mapping size may tend to underestimate the presence of the less frequent critical habitat types. In addition, distinct characteristics necessary for critical wildlife habitat requirements, such as large diameter trees, may be lost in the averages of the aggregate of the smaller types. Note that as the uneven structure class increased by 9,200 acres, there was not a corresponding increase in the acres of size classes 4 and 5. These size classes increased only 7,300 acres indicating a shift of larger size class acres to the mid-size class acres in those types that did not shift from the even to the uneven structure canopy class. Size classes 1 and 2 also reflect a shift in acreage toward the mid-size class. These classes total 16035.6 acres in the five-acre map but only 8325.3 acres in the forty-acre map. More than half of the size class 2 acres in the five-acre map are classified as size class 3 acres in the forty-acre map.

If diversity is indicated by high between stand variance rather than high within stand variance, then measures of diversity may also be influenced by the selection of the minimum size mapping unit. Small minimum size mapping units would tend to indicate higher diversity in stand typing than would larger minimum size limits. Effects such as these must be evaluated to determine their impact(s) on modeled wildlife habitat relationships. The minimum mapping size is partly the result of the capability to accurately describe stand characteristics, but it is also due to a concern regarding the size of these databases. Tradeoffs between database size and ease of processing must be balanced with the loss of specificity in the database.

Summary

Complex heterogenous pixel grids generated through image classification processes can be filtered using rules-based aggregation processes. Rule definition allows grouping of pixels with respect to the type information the pixels represent rather than the mathematical value of any individual pixel value. Rules may be altered to reflect significant aspects or relationships of the data being processed, including species type and composition, cover, tree size, and canopy structure. Data tends to be more generalized as types are aggregated to larger and larger minimum size specification.

REFERENCES

Mayer, K.E. and W.F. Laudenslayer, Jr. (eds) 1988. *A Guide to Wildlife Habitats of California*. California Department of Forestry and Fire Protection, Sacramento, California. 166pp.

Brown, G.K. and L. Fox 1992. Digital classification of thematic mapper imagery for recognition of wildlife habitat characteristics. *Proc:* 1992 ASPRS/ACSM Convention, American Society for Photogrammetry and Remote Sensing, Bethesda, Maryland, In press.