

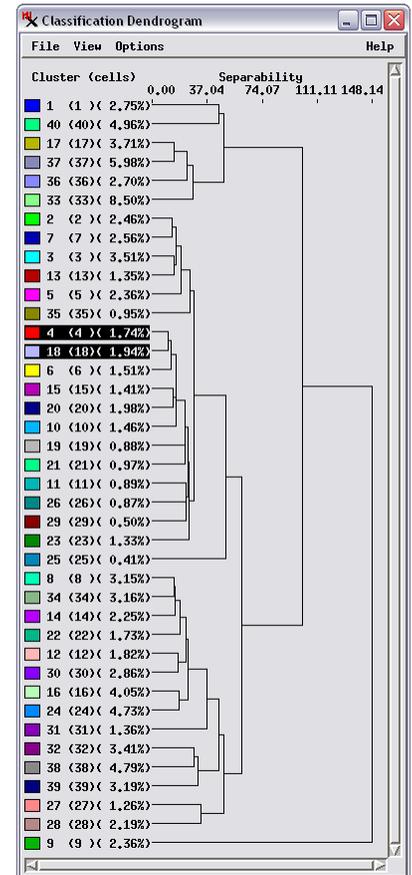
Analysis Tools for Automatic Classification

The Automatic Classification process in TNTmips allows you to classify multiband aerial or satellite images using either unsupervised or supervised methods. The process also provides a number of statistical analysis tools that you can use to assess the classification results. When used with the results from unsupervised classifications, these tools can help you identify the surface materials associated with the classes and to identify classes that are spectrally similar and/or spatially associated. You can simplify the class raster by merging similar classes and by reassigning isolated class cells to the class of their surroundings. The classification analysis tools are opened from the View menu in the Automatic Classification window (illustrated to the right).



Classification Dendrogram

The Classification Dendrogram is a hierarchical branching plot that shows the result of a progressive merging of pairs of classes (and/or previous class groups) with the closest class centers in the n-dimensional spectral space defined by the input image bands. With each merger a new joint class center is computed and all class-center distances are recalculated. Classes are shown with their class colors on the left, and are shown merging from left to right. Each pairing is shown by the lines that connect the classes or class groups. The position on the horizontal Separability axis of the vertical connecting line indicates the degree of difference between the classes or class groups being paired. Class pairs that join



The Classification Dendrogram (right) is a branching plot that shows the degree of relatedness of the output classes. The horizontal Separability axis shows a measure of the distance between class centers in spectral space. The two most similar classes have been selected in the window (black highlight)

Highest Cooccurrence:	4 (1.74%)	5 (2.36%)	6 (1.51%)	7 (2.56%)	8 (3.15%)	9 (2.36%)
14 (2.25%)	611,340	-123,424	98,640	-79,636	507,124	-139,506
15 (1.41%)	13,195	54,175	20,802	38,315	15,509	132,953
16 (4.05%)	97,012	-91,638	-81,279	334,950	-102,671	-111,158
17 (3.71%)	16,281	38,791	25,366	18,451	38,033	145,955
18 (1.94%)	-66,473	-178,649	-129,750	-172,519	268,736	-171,183
19 (0.88%)	48,079	88,573	52,112	171,650	25,711	100,943
20 (1.98%)	-85,220	235,976	-106,583	374,603	-185,316	-174,344
21 (0.97%)	50,921	18,193	52,254	27,364	76,619	187,146
22 (2.25%)	181,966	35,998	264,366	436,568	72,604	-130,673
23 (1.33%)	10,978	36,888	12,732	20,195	30,316	145,098
24 (4.73%)	-71,536	-63,853	-68,763	132,438	-97,973	-85,305
25 (2.5%)	32,250	37,855	35,709	20,532	50,727	149,093

The Co-Occurrence Analysis window shows the degree of spatial association of pairs of classes. The upper number in each box shows the frequency (raw or normalized) with which cells of the class pair occur adjacent to each other. The lower number is the separability value for the pair (as shown in the Dendrogram). The 10 highest co-occurrence values and 10 lowest separability values are highlighted in color. You can use the sliders at the top of the window to select one of these rank levels for either parameter; the grid automatically scrolls to the corresponding class pair and outlines the listing in the associated color.

Highest Cooccurrence:	16 (4.05%)	17 (3.71%)	18 (1.94%)	19 (0.88%)	20 (1.98%)	21 (0.97%)
1 (2.75%)	-193,850	-155,112	-139,427	-76,540	-140,490	-99,972
2 (2.46%)	180,478	82,321	129,530	119,279	139,555	121,501
3 (3.51%)	-160,351	12,638	167,130	-86,040	-80,431	-41,792
4 (1.74%)	66,844	34,022	18,525	31,518	31,701	23,919
5 (2.36%)	-198,540	351,793	-11,593	-88,536	-88,995	-101,258
6 (1.51%)	82,814	18,212	34,093	35,613	44,863	34,847
7 (2.56%)	-66,473	-95,220	181,966	-71,536	65,066	-78,136
8 (3.15%)	48,079	50,921	10,978	32,250	16,977	28,126
9 (2.36%)	-178,649	235,976	35,998	-63,853	-110,667	292,470
10 (1.41%)	88,573	18,193	36,888	37,855	50,284	29,741
11 (1.1%)	-129,750	-106,583	264,366	-68,763	-75,036	144,641
12 (1.46%)	52,112	52,254	10,978	35,709	24,802	21,781

Classes that have been selected in any analysis window are highlighted in a different background shading in the Co-occurrence Analysis window.

near the left edge of the diagram are closely related in their spectral properties, and the degree of relatedness decreases to the right.

The class sizes are shown in parentheses; you can use the Options menu on the Dendrogram window to show class size as a percentage of cells or as cell counts. You can also use the Options menu to choose a separability measure. The default is Euclidean, which shows the Euclidean distance between class centers in the feature space. The other three choices (Bhattacharya, Jeffries-Matusita, and Transformed Divergence) are different statistical measures of the separation between means of two normal distributions. Separability in the latter two methods is expressed on a scale of 0 to 2, where 0 indicates complete overlap and 2 indicates complete separation.

Co-occurrence Analysis

The Co-Occurrence Analysis window presents a class matrix that is the result of an analysis of the spatial associations of pairs of classes. The upper value in the box representing a class pair is the co-occurrence value, the frequency with which cells of the two classes occur adjacent to each other. The co-occurrence value shown by default is the normalized frequency, which adjusts the raw adjacency frequencies to remove the bias related to differing class sizes. A positive value indicates that cells of the classes occur adjacent to each other more often than random chance would predict. A negative value indicates that the two classes tend not to occur together.

(over)

The lower value in each box in the co-occurrence matrix is the separability value for the class pair. The ten highest co-occurrence values and ten lowest separability values are highlighted in color. A pair of classes with low separability and high co-occurrence values have similar spectral properties and tend to occur together, making them a good candidate for merging.

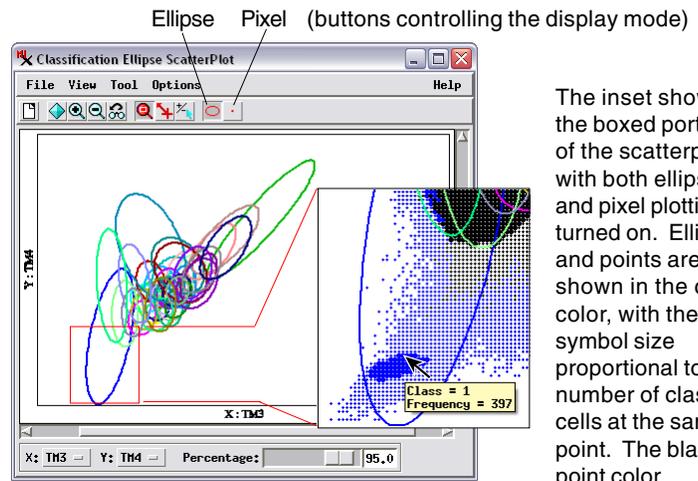
Ellipse Scatterplot

The Ellipse Scatterplot window (right) shows the distribution of classes projected onto a 2D plane through the n-dimensional spectral space. You determine this slice by choosing a pair of bands to assign to the X and Y axes of the diagram. In the default viewing mode each class is represented by an ellipse (drawn in the class color) that encompasses most of the scatter of points in the class. You can adjust the ellipse sizes using the Percentage slider, and you can view the scatter of cell value points in addition to or in place of the ellipse display.

The positions of classes in spectral space can provide important information about the identity of the materials making up the class. For example, a scatterplot of photoinfrared versus red bands (as in the illustration) can help identify classes representing bare soil, green vegetation, and water.

Output Statistics

You can view summary statistics in tabular form in the Output Statistics window (right). The window lists Cluster Counts (the number and percentage of cells assigned to each class) as well as Cluster Means and Cluster Standard Deviations for each input band. A Covariance Matrix for each class provides a relative measure of the degree of spectral correlation between each pair of input bands for the class.

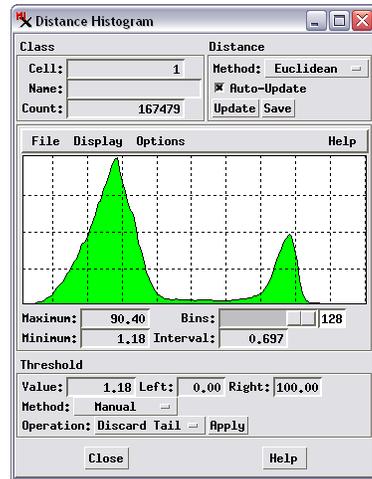


Use the X and Y option buttons to assign a band to each axis of the Ellipse Scatterplot. The Percentage slider adjusts the sizes of the ellipses (percentage of cells included in each ellipse).

The inset shows the boxed portion of the scatterplot with both ellipses and pixel plotting turned on. Ellipses and points are shown in the class color, with the symbol size proportional to the number of class cells at the sample point. The black point color indicates superimposed points from different classes.

Cluster	TH1	TH2	TH3	TH4	TH5	TH7
1	5.62	4.08	7.71	21.71	27.26	12.02
2	3.09	1.69	3.78	5.63	4.95	4.39
3	3.48	2.13	4.58	4.41	5.01	3.94
4	3.36	2.06	4.38	4.42	4.20	3.61
5	3.91	2.38	4.44	5.58	6.47	5.84
6	3.80	2.09	4.15	5.65	6.79	5.70
7	5.22	1.85	3.91	5.24	5.51	4.42
8	3.36	2.14	4.14	5.97	6.99	5.27
9	13.95	6.45	17.42	19.22	14.49	17.08
10	3.34	1.97	4.71	5.90	5.52	4.01
11	3.75	2.71	6.76	8.00	6.78	4.03
12	3.71	2.49	5.38	7.16	6.88	4.00
13	3.76	2.16	4.78	6.66	6.65	5.21
14	3.66	2.17	4.45	5.87	5.37	4.32
15	3.49	2.06	4.43	5.09	4.59	4.06
16	4.48	3.02	5.04	6.95	7.42	5.44
17	3.55	2.15	4.14	6.49	7.03	4.24
18	3.34	1.88	3.82	4.85	5.16	4.53
19	6.75	3.40	4.58	5.69	7.66	5.24
20	3.83	2.22	4.45	5.00	4.68	4.23
21	5.03	2.98	4.85	5.35	7.14	6.49

The Output Statistics window shows tabular listings of class statistics.



Distance Histogram for a class with a bimodal distribution of point distance.

Distance Histogram

The Distance Histogram window shows the spread of points relative to the class center for any class you select. A compact class with a small spread of cell values has a narrow distance histogram. A diffuse class with many cells far from the class center has a histogram with a tail extending to higher distance values, or perhaps even a second mode (peak) in the histogram, as illustrated to the left. These outlier cells may represent distinctly different materials than the cells near the class center. You can remove outlier cells from a class by using the histogram crosshairs to set a distance threshold (vertical crosshair position). The numeric fields in the Threshold box show the distance value of the threshold and the percentage of points to the left and right of the threshold. Press the Apply button to discard the points at greater distances (note that there is no “Undo” for this operation).

Error Matrix

	Alfalfa	Beans	Corn	Grass &	Potatoes	Sugar Be	Summer F	Wheat-Ir	Winter W	Total	Accuracy
Alfalfa	21	0	0	0	0	0	0	0	0	21	100.00%
Beans	0	402	1	0	6	7	0	0	0	416	96.63%
Corn	0	2	163	0	0	1	0	0	0	166	98.19%
Grass &	0	0	0	431	0	0	1	0	0	432	99.77%
Potatoes	0	53	0	0	42	1	0	0	0	96	43.75%
Sugar Be	0	0	7	0	163	0	0	0	0	170	95.88%
Summer F	0	15	0	0	0	15	316	0	0	346	91.33%
Wheat-Ir	0	0	0	0	0	0	0	208	6	214	97.20%
Winter W	0	1	0	17	0	3	8	35	447	511	87.48%
Total	21	473	171	448	48	190	325	243	453	2372	
Accuracy	100.00%	84.99%	95.32%	96.21%	87.50%	85.79%	97.23%	85.60%	98.68%		
Overall Accuracy = 92.45% Khat Statistic = 91.11%											

The Error Matrix compares the supervised classification result to the sample areas in a training set raster or ground truth raster.

The Error Matrix window can be used to assess the results of a supervised classification. You select a “ground truth” raster containing sample areas of known class, which can be either the training set raster used in the classification or an independent ground truth raster. The class of each sample area cell is compared to the class assignment produced by the supervised classification. Each row in the matrix represents a result class and each column represents a ground truth class. The boxes on the diagonal (with red background color) show counts of correctly classified cells. The values in the off-diagonal matrix boxes are counts of incorrectly-classified image cells. Accuracy percentage values are shown for each result class and for each ground truth class, along with overall accuracy values.