typical spectral signatures of photosynthetically active and non-photosynthetically active vegetation (Beeri et al., 2007)

The Leaf
Chloroplasts

“sieve effect”
Chlorophyll
Carotenoids
Chlorophyll A: green; Chlorophyll B: orange; Carotene: red; Structural proteins: yellow
Photosynthesis

$6\text{CO}_2 + 6\text{H}_2\text{O} = \text{C}_6\text{H}_{12}\text{O}_6 + 6\text{O}_2$

The Cornerstone of Life on this planet!
Dominant factors controlling leaf reflectance

- Leaf pigments in the palisade mesophyll:
  - chlorophyll \(a, b\)
  - \(\beta\)-carotene, etc.

- Scattering in the spongy mesophyll

- Leaf water content

Primary absorption bands

- Chlorophyll absorption bands

- Atmospheric water absorption bands

Reflectance (%)

Wavelength, \(\mu\)m

0.4 0.6 0.8 1.0 1.2 1.4 1.6 1.8 2.0 2.2 2.4 2.6

Visible

Reflective infrared

Near-infrared

Middle-infrared
Absorption Spectra of Chlorophyll $a$ and $b$

Pigment Absorption
Cell Wall Constituents
Protein Constituents
Cellular Water
Leaf biochemistry

pigments: chlorophyll a and b, α-carotene, and xanthophyll

absorb in blue (& red for chlorophyll)

Absorbed radiation converted into:

heat energy, fluorecence, or carbohydrates through photosynthesis

Stored carbon!
Chlorophyll Concentration

Reduced absorption due to decreasing chlorophyll concentration
Red or blue wavelength radiance, reflectance

Chlorophyll concentration

Chlorophyll Concentrations
What about estimating Phytomass?

First let's define phytomass:

leaf area x leaf mass per unit area (m² x kg/m² = kg)

Then let's introduce some surrogates for phytomass:

Leaf area Index (LAI)- one sided leaf area per unit ground area

Leaf area density (LAD) – leaf area per unit volume

Why use area instead of mass? Because estimating the leaf mass per unit area using remote sensing is very difficult. (How thick and heavy are the leaves?) Estimating area is more straightforward (remember areal mixtures)
Estimating Phytomass: Additive Reflectance

Additive reflectance from Leaf 1 and Leaf 2
\[ R_1 + T_3 = \frac{5}{8} \Phi_i = 62.5\% \]

Reflectance radiant flux, \( \Phi_r \)
\[ R_1 = \frac{1}{2} \Phi_i = 50\% \]

Transmitted radiant flux, \( \Phi_t \)
\[ T_3 = \frac{1}{2} R_2 \]
(or \( \frac{1}{8} \Phi_i = 12.5\% \))

Incident radiant flux, \( \Phi_i \)
\[ T_5 = \frac{1}{2} R_4 \]
(or \( \frac{1}{32} \Phi_i \))

Leaf 1

Transmitted radiant flux, \( \Phi_t \)
\[ T_1 = \frac{1}{2} \Phi_i \]

\[ R_2 = \frac{1}{2} T_1 \]
(or \( \frac{1}{4} \Phi_i \))

Leaf 2

\[ R_4 = \frac{1}{2} R_3 \]
(or \( \frac{1}{16} \Phi_i \))

\[ T_2 = \frac{1}{2} T_1 \]
(or \( \frac{1}{4} \Phi_i \))

\[ T_4 = \frac{1}{2} R_3 \]
(or \( \frac{1}{16} \Phi_i \))
Leaf Area Estimation

Most leaf area

- reflectance (%)
- wavelength
- density 1
- density 2
- density 3
- density 4
- density 5
- density 6
- sunlit soil
Both red and NIR bands carry complementary information. How can we use that info and reduce noise introduced by other sources?

- NIR = 75
- Red = 32
- NIR/R = 2.34

- NIR = 119
- Red = 49
- NIR/R = 2.38

Shaded area
Vegetation Indices

Vegetation indices (VI) are combinations of spectral measurements in different wavelengths as recorded by a radiometric sensor. They aid in the analysis of multispectral image information by shrinking multidimensional data into a single value. Huete (1994) defined vegetation indices as:

“
dimensionless, radiometric measures usually involving a ratio and/or linear combination of the red and near-infrared (NIR) portions of the spectrum. VI’s may be computed from digital counts, at satellite radiances, apparent reflectances, land-leaving radiances, or surface reflectances and require no additional ancillary information other than the measurements themselves...What VI's specifically measure remains unclear. They serve as indicators of relative growth and/or vigor of green vegetation, and are diagnostic of various biophysical vegetation parameters”.

Vegetation indices (VI’s) can be broken up into two basic categories:

Ratio based indices – VI’s based on the ratio of two or more radiance, reflectance, or DN values (or linear combinations thereof).

Difference indices – VI’s based on the difference between the spectral response of vegetation and the soil background.
Common Ratio Indices

Simple Ratio Index (SR) = \( \frac{\text{NIR}}{R} \)

Normalized Difference Vegetation Index (NDVI) =

\[
\frac{\text{NIR} - R}{\text{NIR} + R}
\]
The NDVI ratio capitalizes on the NIR and Red portions of the electromagnetic spectrum. The NIR portion of the spectrum is reflected by leaf tissue and recorded at the sensor, while the Red portion of the spectrum is absorbed by the chlorophyll present in the leaf tissue, thus reducing the reflectance of red light present at the sensor. The mathematical range of NDVI is -1 to 1. Thus, the contrast of reflectance and absorption by vegetation cover allows for the evaluation of vegetation present on the surface.
What bands should we ratio to reduce albedo effect and shadows?

- Assuming we want the vegetation to stand out, we should ratio a “bright” band with a “dark” band.
- Vegetation reflects darkly in the 400-700nm range, and brightly in the 700-1300nm range.
- TM - TM4/TM3 is traditional
- MSS - MSS7/MSS5 is traditional
LAI & NDVI

Leaf Area

NDVI vs. LAI
Coarse Spatial Resolution

Relationship between leaf area index of the primary sites and the monthly maximum value composite NDVI for July 1997.
Soil Reflectance Separation

**Soil Moisture Content**

- Percent reflectance vs. wavelength (μm)
  - Soil loam texture: sand 20%, silt 20%, clay 60%
  - Moisture content: 1%, 5%, 10%, 15%, 20%

**Soil Drainage**

- Percent reflectance vs. wavelength (μm)
  - Drainage types: Well, Moderately Well Drained, Somewhat Poorly Drained, Poorly Drained, Very Poorly Drained

**Organic Matter Content**

- Percent reflectance vs. wavelength (μm)
  - Mineral soils
  - Levels: Low, Medium, High
Soil and Vegetation Characteristics

- Soil and vegetation have very distinct reflectance curves.

Vegetation’s steep climb in the NIR makes that part of the spectrum stand out.

Which contrasts its dive in the red part of the spectrum

Soil stays relatively linear.
Soil line

- In a NIR:Red plot, soil’s reflectance is a straight line, which indicates the line where the vegetation begins.

Vegetation, as discussed earlier, gains significantly in the NIR, but descends in the Red.

PVI is a measure of the distance of a pixel from the soil line.

High soil moisture

Low soil moisture
Perpendicular Vegetation Index (PVI) for a single soil background:

\[ PVI = \sqrt{(R_{soil} - R_{veg})^2 + (NIR_{soil} - NIR_{veg})^2} \]

Where \( R_{soil} \) and \( NIR_{soil} \) are the red and NIR reflectance/radiance for the soil background.
Or the PVI for a multiple soil background:

\[
PVI = \frac{NIR_{\text{veg}} - aR_{\text{veg}} - b}{\sqrt{1 + a^2}}
\]

Where \(a\) and \(b\) are the slope and intercept respectively of the universal soil line for the area.
But what about other objects within the field of view (FOV) of the sensor other than vegetation?
Spectral Indices

Fractional Vegetation

NDVI values themselves can be used to calculate a vegetation index. Fractional Vegetation, or $F_v$, uses NDVI to derive a estimation of percent vegetation cover. The formula normalizes the resultant NDVI values to produce values that range from 0% vegetation cover to 100% vegetation cover.

$$f_v = \frac{(NDVI_{i,j} - NDVI_{min})}{(NDVI_{max} - NDVI_{min})}$$

Where:

$NDVI_{i,j}$ = The calculated NDVI for the pixel in row $i$ and column $j$ of the input NDVI image.

$NDVI_{min}$ = The NDVI value empirically determined to represent 0%.

$NDVI_{max}$ = The NDVI value empirically determined to represent 100% vegetation cover.

Note that change is more detectable over longer periods of time. Also, prior to multi-temporal image change detection, images should be radiometrically corrected to account for atmospheric, sensor, and sun-angle anomalies.

By subtracting the two $F_v$ images, we get a relative difference between years.
Composite Canopy Reflectance

- Red flux radiation
- NIR flux radiation

Sensor

Leaf area Density

Shadow vs Sunlit Background
Composite Canopy Reflectance

- 0% veg. Cover; LAI = 0
- 50% veg. Cover; 2 leaf layers; LAI = 1
- 33% veg. cover; 3 leaf layers; LAI = 1
- 100% veg. Cover; 1 leaf layer; LAI = 1

Are the reflectances for these 3 pixels the same?
In this region, there is complete vegetation cover and differences are due to increasing canopy density. Additive Reflectance (multiple scattering)
How to separate....
Separability between classes can be evaluated by computing the $M$ statistic. To compute the $M$ statistic you must find the mean pixel value for each vegetation type you want to test and also the standard deviation.

$$M = \frac{\mu_1 - \mu_2}{\sigma_1 + \sigma_2}$$

where

- $\mu_1$ = mean value of the reflectance for vegetation type 1
- $\mu_2$ = mean value of the reflectance for vegetation type 2
- $\sigma_1$ = the standard deviation of the reflectance for veg 1
- $\sigma_2$ = the standard deviation of the reflectance for veg 2

$M > 1$ indicates adequate separability

You can merge classes that do not have good separability.
Using Remote Sensing to Map Vegetation Density on a Reclaimed Surface Mine
Michael Shank

Abstract. The West Virginia Department of Environmental Protection, in a cooperative agreement with the Office of Surface Mining’s Charleston Field Office, is evaluating the utility of high resolution satellite images for characterizing vegetation patterns on reclaimed surface mines. This paper details the results of the first phase of this project, which sought to determine whether satellite images could be used to estimate percentage vegetation cover.

This paper details a simple technique for estimating percent vegetation cover based on the widely-used Normalized Difference Vegetation Index (NDVI). NDVI exhibited a 0.96 correlation with percent vegetation cover for 34 reference samples collected on a 94 acre study area in southern West Virginia. Based on this relationship, a technique was developed that produced a mean error of 6.41% (+/- 2.68% at the 90% confidence level) when estimating percent cover for the 34 field sites.
Sample locations were identified from the satellite image by examining Normalized Difference Vegetation Index (NDVI) statistics calculated for a 5x5 moving window. NDVI is a common tool for identifying and characterizing vegetation. In this instance, NDVI served as a surrogate measure of vegetation density and homogeneity in the neighborhood surrounding an image pixel. The average NDVI value for the 5x5 window was used to stratify image pixels into ten groups representing conditions ranging from bare earth to fully revegetated.

The graph depicts the relationship between NDVI and percent cover for the 34 field sites. The solid line traces the best fit equation calculated using simple linear regression:

\[
PCT\_COVER = -0.140224 + 2.5886 \times NDVI
\]

The equation produces an R² value of 0.9195. Residual errors for this model ranged from 0 to 20.27%, averaging 6.78% for the entire sample set (RMSE was 8.66%). A linear regression established a relationship between NDVI and Percent cover. With this relationship they were able to map % Cover.
Using Normalized Difference Vegetation Index (NDVI) as an Indicator of Cheatgrass (*Bromus tectorum*) Infestations in Skull Valley, Utah

(Photograph: Copyright © 1971 Roy D. Tea. Taken from http://www.images.google.com)
A comparative method using Normalized Difference Vegetation Index (NDVI) has been developed for monitoring the presence and spread of cheatgrass. This method was applied on Landsat-7 ETM data for Skull Valley, UT. NDVI values of the area from June 4, were subtracted from NDVI values from May 3 of the same year in order to indicate the presence and area of cheatgrass infestations in the valley.

Correlations between the NDVI and Fire Finder outputs of the area, which were both generated in ERDAS Imagine’s Model Maker, are visually apparent resulting from the mutualistic relationship between cheatgreass and wildfire.
The higher the NDVI value, the more green, or photosynthetically active, is the vegetative cover (Burgan 1993). Note the visually apparent reduction in brightness from the May 3 image to the June 4 image (fig. 7, fig. 8).
Figure 8 – NDVI output from the image taken on June 4, Skull Valley, UT.
Figure 9 – Model which subtracts NDVI values of June 4 image from NDVI values of May 3 image.
Figure 10 – Map of Skull Valley indicating areas of cheatgrass infestation.
- A noticeable correlation exists between the cheatgrass and Fire Finder images in Skull Valley.
Reality
• Though these transformations show us trends we could not see, they are far from perfect.
• Researchers have found that none of these indices effectively deals with atmospheric error, and some critics question whether interpretable results can even be obtained without a pixel by pixel atmospheric correction.
• Understanding basic Vegetation indices does help to solidify the concept of the soil line.
Phenology

Time Sequence of Hyperion Images
Coleambally Irrigation Area Farm 33

EO-1 Hyperion Hyperspectral Sensor http://rs.csr.utexas.edu/rsinfo/sensors/sensor.php?n_id=13
### TM vs SPOT in Crop Identification

<table>
<thead>
<tr>
<th>Band</th>
<th>Bandwidth (μm)</th>
<th>IFOV (m)</th>
<th>Quantization (bits)</th>
<th>Off Nadir Viewing</th>
<th>Temporal Resolution (days)</th>
<th>Altitude (km)</th>
<th>Total Data Rate (Mbits/s)</th>
<th>Number Pixels per Line</th>
<th>Swath Width (km)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Landsat Thematic Mapper (TM) on Landsat 4 and 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.45–0.52</td>
<td>30 × 30</td>
<td>8</td>
<td>No</td>
<td>16</td>
<td>705</td>
<td>85</td>
<td>3000</td>
<td>185</td>
</tr>
<tr>
<td>2</td>
<td>0.52–0.60</td>
<td>30 × 30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.63–0.69</td>
<td>30 × 30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.76–0.90</td>
<td>30 × 30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1.55–1.75</td>
<td>30 × 30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>10.4–12.5</td>
<td>120 × 120</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>2.08–2.35</td>
<td>30 × 30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>French SPOT High Resolution Visible Sensor Systems (HRV) 1, 2, and 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Multispectral Mode</strong></td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td><strong>Panchromatic Mode</strong></td>
</tr>
<tr>
<td>1</td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th>Products (sensors)</th>
<th>Features</th>
<th>Vegetation mapping applications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Landsat TM</td>
<td>Medium to coarse spatial resolution with multispectral data (120 m for thermal infrared band and 30 m for multispectral bands) from Landsat 4 and 5 (1982 to present). Each scene covers an area of 185 x 185 km. Temporal resolution is 16 days.</td>
<td>Regional scale mapping, usually capable of mapping vegetation at community level.</td>
</tr>
<tr>
<td>Landsat ETM+ (Landsat 7)</td>
<td>Medium to coarse spatial resolution with multispectral data (15 m for panchromatic band, 60 m for thermal infrared and 30 m for multispectral bands) (1999 to present). Each scene covers an area of 185 km x 185 km. Temporal resolution is 16 days.</td>
<td>Regional scale mapping, usually capable of mapping vegetation at community level or some dominant species can be possibly discriminated.</td>
</tr>
<tr>
<td>SPOT</td>
<td>A full range of medium spatial resolutions from 20 m down to 2.5 m, and SPOT VGT with coarse spatial resolution of 1 km. Each scene covers 60 x 60 km for HRV/HRVIR/HRG and 1000 x 1000 km (or 2000 x 2000 km) for VGT. SPOT 1, 2, 3, 4 and 5 were launched in the year of 1986, 1990, 1993, 1998 and 2002, respectively. SPOT 1 and 3 are not providing data now.</td>
<td>Regional scale usually capable of mapping vegetation at community level or species level or global/national/regional scale (from VGT) mapping land cover types (i.e. urban area, classes of vegetation, water area, etc.).</td>
</tr>
<tr>
<td>MODIS</td>
<td>Low spatial resolution (250–1000 m) and multispectral data from the Terra Satellite (2000 to present) and Aqua Satellite (2002 to present). Revisit interval is around 1–2 days. Suitable for vegetation mapping at a large scale. The swath is 2330 km (cross track) by 10 km (along track at nadir).</td>
<td>Mapping at global, continental or national scale. Suitable for mapping land cover types (i.e. urban area, classes of vegetation, water area, etc.).</td>
</tr>
<tr>
<td>AVHRR</td>
<td>1-km GSD with multispectral data from the NOAA satellite series (1980 to present). The approximate scene size is 2400 x 6400 km</td>
<td>Global, continental or national scale mapping. Suitable for mapping land cover types (i.e. urban area, classes of vegetation, water area, etc.).</td>
</tr>
<tr>
<td>IKONOS</td>
<td>It collects high-resolution imagery at 1 m (panchromatic) and 4 m (multispectral bands, including red, green, blue and near infrared) resolution. The revisit rate is 3–5 days (off-nadir). The single scene is 11 x 11 km.</td>
<td>Local to regional scale vegetation mapping at species or community level or can be used to validate other classification result.</td>
</tr>
<tr>
<td>QuickBird</td>
<td>High resolution (2.4–0.6 m) and panchromatic and multispectral imagery from a constellation of spacecraft. Single scene area is 16.5 x 16.5 km. Revisit frequency is around 1–3.5 days depending on latitude.</td>
<td>Local to regional scale vegetation mapping at species or community level or used to validate vegetation cover extracted from other images.</td>
</tr>
<tr>
<td>ASTER</td>
<td>Medium spatial resolution (15–90 m) image with 14 spectral bands from the Terra Satellite (2000 to present). Visible to near-infrared bands have a spatial resolution of 15 m, 30 m for short wave infrared bands and 90 m for thermal infrared bands.</td>
<td>Regional to national scale vegetation mapping at species or community level.</td>
</tr>
<tr>
<td>AVIRIS</td>
<td>Airborne sensor collecting images with 224 spectral bands from visible, near infrared to short wave infrared. Depending on the satellite platforms and latitude of data collected, the spatial resolution ranges from meters to dozens of meters and the swath ranges from several kilometers to dozens of kilometers.</td>
<td>At local to regional scale usually capable of mapping vegetation at community level or species level. As images are carried out as one-time operations, data are not readily available as it is obtained on an ‘as needs’ basis.</td>
</tr>
<tr>
<td>Hyperion</td>
<td>It collects hyperspectral image with 220 bands ranging from visible to short wave infrared. The spatial resolution is 30 m. Data available since 2003.</td>
<td>At regional scale capable of mapping vegetation at community level or species level.</td>
</tr>
</tbody>
</table>

a Many sensors provide imagery for producing VI (e.g. NDVI) that is calculated from the bands in the visible and near-infrared regions.
Preprocessing of satellite images prior to vegetation extraction is essential to remove noise and increase the interpretability of image data. This is particularly true when a time series of imagery is used or when an area is encompassed by many images since it is essentially important to make these images compatible spatially and spectrally. The ideal result of image preprocessing is that all images after image preprocessing should appear as if they were acquired from the same sensor. Two types of preprocessing include: 

**Radiometric and Geometric.**

**Radiometric correction** of remote sensing data normally involves the process of correcting radiometric errors or distortions of digital images to improve the fidelity of the brightness values. Factors such as seasonal phenology, ground conditions and atmospheric conditions can contribute to variability in multi-temporal spectral responses that may have little to do with the remote sensed objects themselves.

It is mandatory to differentiate real changes from noises through radiometric correction in cases where the spectral signals are not sufficiently strong to minimize the effects of these complicating factors. Several methods are available to make radioactive corrections. Some of them are based on complex mathematical models that describe the main interactions involved. However, the values of certain parameters (i.e. the atmospheric composition) must be known before applying them. Other radiometric correction methods are based on the observations of reference targets (e.g. water or desert land) whose radiometry is known.
Geometric correction is aimed to avoid geometric distortions from a distorted image and is achieved by establishing the relationship between the image coordinate system and the geographic coordinate system using the calibration data of the sensor, the measured data of position and altitude and the ground control points. Therefore, geometric correction usually includes the selection of a map projection system and the co-registration of satellite image data with other data that are used as the calibration reference.

The outcome of geometric correction should obtain an error within plus or minus one pixel of its true position, which allows for accurate spatial assessments and measurements of the data generated from the satellite imagery.

The first-order transformation and the nearest neighbor resampling of the uncorrected imagery are among those popularly adopted methods in geometric correction. The first-order transformation, also known as the linear transformation, applies the standard linear equation \((y = mx + b)\) to the X and Y coordinates of the ground control points. The nearest neighbor resampling method uses the value of the closest pixel to assign to the output pixel value and thus transfers original data values without averaging them. Therefore, the extremes and subtleties of the data values are not lost.
The COST Atmospheric Correction tool creates a spatial model by converting the images digital numbers to reflectance and performs an image-based atmospheric correction using the Chavez (1996) COST method (www.gis.usu.edu). The equation is thus:

\[ \rho_{\text{BandN}} = \frac{\pi((L_{\text{BandN}} \times Gain_{\text{BandN}} + Bias_{\text{BandN}}) - (H_{\text{BandN}} \times Gain_{\text{BandN}} + Bias_{\text{BandN}})) \times D^2}{E_{\text{BandN}} \times (\cos((90 - \theta) \times \pi / 180)) \times \tau} \]

Where, \( \rho_{\text{BandN}} = \) Reflectance for Band N; \( L_{\text{bandN}} = \) Digital Number for Band N; \( H_{\text{bandN}} = \) Digital Number representing Dark Object for Band N; \( D = \) Normalized Earth-Sun Distance; \( E_{\text{bandN}} = \) Solar Irradiance for Band N; \( \tau = \) Atmospheric Transmittance expressed as \( (\cos((90 - \theta) \times \pi / 180)) \)

**Figure 5** – Equation used by the COST Atmospheric Correction tool along with explanation of variables.
It is very common that the same vegetation type on ground may have different spectral features in remote sensed images.

Also, different vegetation types may possess similar spectra, which makes very hard to obtain accurate classification results either using the traditional unsupervised classification or supervised classification. Searching for improved classification methods is always a hot research topic.

All classification methods are derived from the traditional methods which provide the basic principles and techniques for image classification. Unsupervised, supervised, etc…
Expert Knowledge-Based Classification
One of the major disadvantages to most of the techniques discussed above is that they are all per-pixel classifiers. Each pixel is treated in isolation when using the technique to determine which feature or class to assign it to – there is no provision to use additional cues such as context, shape and proximity, cues which the human visual interpretation system takes for granted when interpreting what it sees. One of the first commercially available attempts to overcome these limitations was the IMAGINE Expert Classifier.

Figure 3: The Knowledge Engineer showing a decision tree leading to land use classes.
•The expert classification software provides a rules-based approach to multispectral image classification, post-classification refinement and GIS modeling. In essence, an expert classification system is a hierarchy of rules, or a decision tree that describes the conditions for when a set of low level constituent information gets abstracted into a set of high level informational classes. The constituent information consists of user-defined variables and includes raster imagery, vector layers, spatial models, external programs and simple scalars.

•A rule is a conditional statement, or list of conditional statements, about the variable’s data values and/or attributes that determine an informational component or hypotheses. Multiple rules and hypotheses can be linked together into a hierarchy that ultimately describes a final set of target informational classes or terminal hypotheses. Confidence values associated with each condition are also combined to provide a confidence image corresponding to the final output classified image.

LIMITATIONS: While the Expert Classification approach does enable ancillary data layers to be taken into consideration, it is still not truly an object based means of image classification (rules are still evaluated on a pixel by pixel basis). Additionally, it is extremely user-intensive to build the models – an expert is required in the morphology of the features to be extracted, which also then need to be turned into graphical models and programs that feed complex rules, all of which need building up from the components available. Even once a knowledge base has been constructed it may not be easily transportable to other images (different locations, dates, etc).
Lab 2. Decision (Classification) Tree

- Represented as a set of hierarchically-arranged decision rules (i.e., tree-branch-leaf)

- Could be generated by knowledge engineering, neural network, or statistic methods.

- S-Plus:
  - Tree Models: successively splitting the data to form homogeneous subsets.
Classification Example

Tree Assessment:
(Proportion Correctly Classified):

- Leaf 1: 100%
- Leaf 2: 74%
- Leaf 3: 69%
- Leaf 4: 100%
- Overall: 82%

Rules:
- if \(X_1 < 2.8\) and \(X_2 < -10\) then class = MAX
- if \((X_1 < 2.8 \text{ and } X_2 > -10)\) or \((X_1 > 2.8 \text{ and } X_3 < 0.003)\) then class = MED
- if \(X_1 > 2.8\) and \(X_3 > 0.003\) then class = MIN