A NON-PARAMETRIC APPROACH
TO CLASSIFYING REMOTELY SENDED DATA

Jack Tubbs
NRC/JSC
MPAD, FM8
Houston, TX

John Engvall
JSC
MPAD, FM8
Houston, TX

ABSTRACT

A non-parametric classification procedure which categorizes a given pixel as wheat or non-wheat according to the trend it follows throughout the growing season is discussed. The cycle wheat makes during the growing season can be described by observing that wheat goes from "brown" (bare soil) to "green" (greening stage) to "brown" (Harvested or golden stage). Since wheat is the only significant U.S. crop which has this cycle during the growing season for winter wheat, we felt that if we could develop a procedure which would enable one to visually follow this type of trend, then perhaps one would be able to distinguish winter wheat from most non-wheat classes.

The following procedure has been developed. Let X be a four dimensional vector corresponding to the four LANDSAT bands 4, 5, 6, and 7 and define the linear transformation Y = AX + C where the matrix A and the vector C are chosen to best exhibit the desired trends. By normalizing the Y vector, we have projected a four-dimensional vector X into a two dimensional vector Y. This type of transformation is repeated for each of four MSS data acquisitions corresponding to the four biological phases of wheat. When each Y is plotted on triangular graph paper, we are able to visually follow the trend an observation makes over time. This procedure was applied to mean vectors of several fields where ground truth information was available. This ground truth information was taken over sites which differed both in geographical location and in growing seasons, but in all cases the following two things were observed: 1) one is able to distinguish wheat from non-wheat and 2) the trend an individual crop makes is fairly consistent regardless of geographical location.

After automating this procedure, we were able to classify entire sample segments and have compared our results with the ground truth sites, maximum likelihood estimates, and United States Department of Agriculture 1975 estimates. The above procedure has been surprisingly accurate and is currently being used as a quality assurance tool for flagging possible errors in maximum likelihood estimates.