1-1-1977

Effect of the Size of Training Samples on Classification Accuracy

R. Kumar

M. Niero

M. S. S. Barros

L. A. M. Lucht

A. P. Manso

Reprinted from

Symposium on

Machine Processing of

Remotely Sensed Data

June 21 - 23, 1977

The Laboratory for Applications of
Remote Sensing

Purdue University
West Lafayette
Indiana

IEEE Catalog No.
77CH1218-7 MPRSD

Copyright © 1977 IEEE
The Institute of Electrical and Electronics Engineers, Inc.

EFFECT OF THE SIZE OF TRAINING SAMPLES ON CLASSIFICATION ACCURACY

R. KUMAR, M. NIERO, M.S.S. BARROS, L.A.M. LUCHT AND A.P. MANSO

Instituto de Pasquisas Espaciais (INPE)
Conselho Nacional de Desenvolvimento
Cientifico e Tecnologico (CNPq) 12.200 -
Sao Jose dos Campos, SP-Brasil

Cloud free multispectral scanner
(MSS) data of LANDSAT were analysed for
studying the effect of the size of train-
ing samples on classification accuracy over
the following test sites: Sao Jose dos
Campos (23° 10' S, 45° 50' W), Cachoeira
Paulista (22° 40' S, 45° W) Jardinopolis
(21° S, 47° 50' W).

With the help of ground observations
and aerial photography, Sao Jose dos Campos
and Cachoeira Paulista were divided into
their respective land use classes whereas
Jardinopolis was divided into its agri-
cultural classes. LANDSAT MSS data were
used to divide each of these classes into
two independent data sets -- training
fields (about 20 percent of the area)
and test fields. Using training fields of
each class, test fields were classified
using the single-cell option of the Image-
100 as well as a sample classifier based
on Bhattacharyya distance. An identical
analysis will be done using about ten
percent, as well as one percent of the
area, for training. This work has many
applications since the training phase of
the data analysis of earth resources is
expensive in the digital as well as the
analog mode. If time permits, the effect
of the distance of the training samples
on the classification accuracy of the
test samples will be investigated.