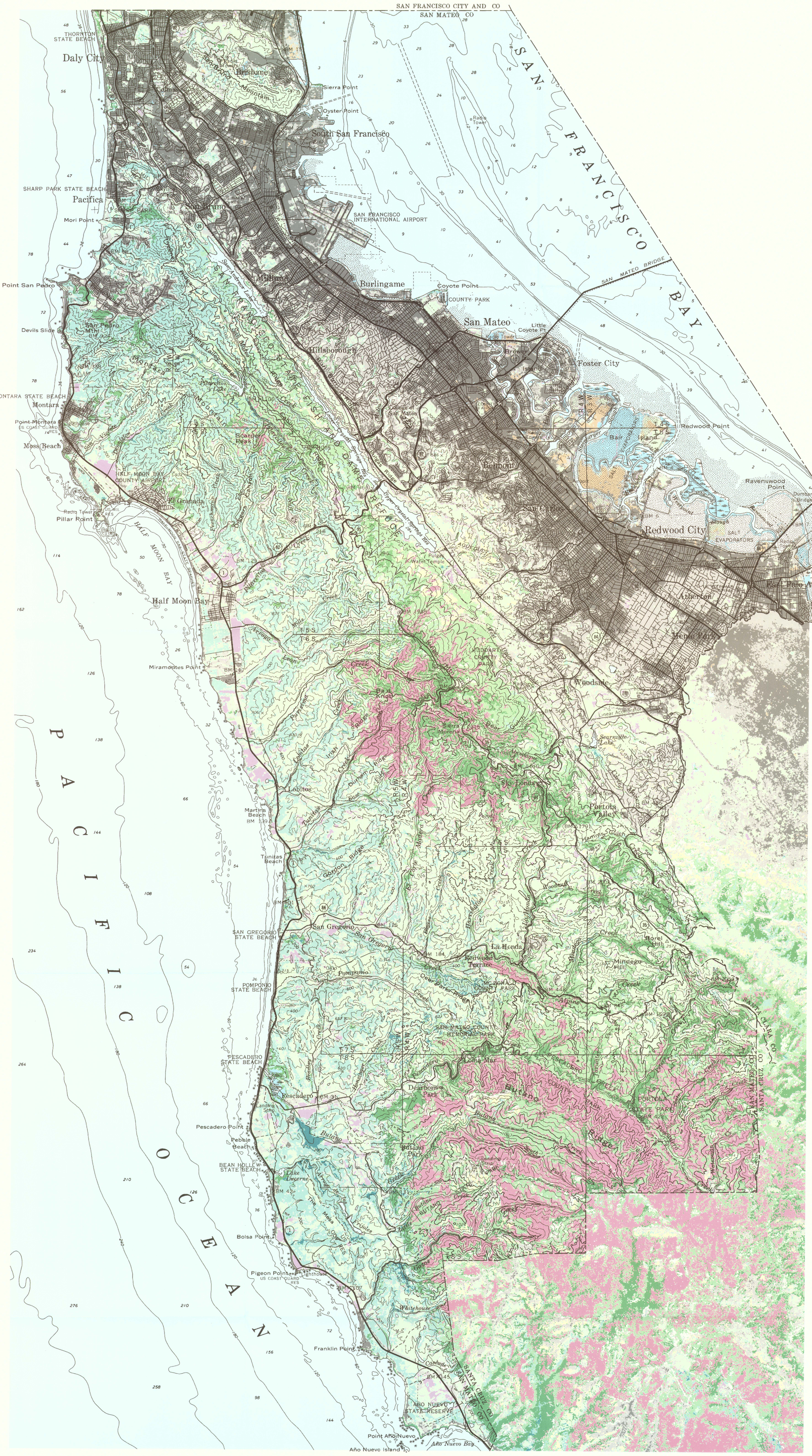

EXPLANATION

Grassland—Includes pasture, rangeland, natural grasslands, and recreational turf

Coastal scrub—Characteristic species include *Artemisia californica* and *Baccharis pilularis*. *Baccharis* domination in disturbed area is included in this class

Chamise chaparral—Characteristic species is *Adenostoma fasciculatum*. Includes, however, areas of serpentine chaparral

Redwoods—Characteristic species is *Sequoia sempervirens*

Other conifers—Predominantly Douglas fir (*Pseudotsuga menziesii*); includes isolated occurrences of Monterey pine

Mixed evergreen forest—Characteristic species include oaks (*Quercus* spp.), California bay (*Umbellularia californica*), madrone (*Arbutus menziesii*), California buckeye (*Aesculus californica*), and scattered conifers

Riparian woodland—Characterized by moisture-loving species such as willow (*Salix* spp.) and alder (*Alnus* spp.)

Agriculture—Bare fields and crops

Low-density built—Rural/suburban, highly vegetated

Medium-density built—Suburban, often residential, moderate vegetation

High-density built—Downtown urban, little vegetation

Salt ponds

Wetlands—Includes both coastal and inland marshlands

Brush—Weedy growth on disturbed ground

INTRODUCTION

Vegetation and land cover information are helpful for the understanding and assessment of many issues in the Earth-science and planning fields. Examples of such issues include the study of vegetation root strength in relation to landslide distribution and potential and the selection and management of open space. Regional vegetation information has historically been obtained through aerial photo interpretation. The advent of polar-orbiting satellites carrying multi-spectral scanners offers an alternative mapping approach providing frequent land-cover information over large areas. Advantages of this methodology include a digital raster format that fits easily into a computerized geographic information system (GIS), the potential for automated rather than manual classification of the data, and the ease of updating and modifying the map, for example, through the regrouping of vegetation classes for different map applications.

LANDSAT THEMATIC MAPPER

The satellite data used to create this map came from the Landsat 5 Thematic Mapper, a multi-spectral scanner in polar orbit around the Earth. This family of satellites has been recording data around the world since 1982. Digital data is recorded by the scanner in seven wavelength intervals across the spectrum of incoming light (table 1). Data are received, within a spectral band, as an integrated value of radiation over the sensor field of view and converted into the "pixel" values of the final digital tapes. These final values, known as digital counts, represent the radiance as influenced by the sensing and recording mechanisms. The sensor field of view determines the final resolvable ground element, which is, in the case of the Thematic Mapper 30 by 30 meters. The sensor records a digital count for each 30 x 30 meter unit of the underlying landscape, scanning east/west as the satellite orbits the Earth. Data are recorded over the same ground location every 16 days.

Table 1. *Landsat Thematic Mapper "band" description*

Thematic Mapper	Wavelength (micrometers)	Description of incoming light
Band 1	.45 - .52	Blue
Band 2	.52 - .60	Green
Band 3	.63 - .69	Red
Band 4	.76 - .90	Near infrared
Band 5	1.55 - 1.75	Mid infrared
Band 6	10.40 - 12.50	Thermal
Band 7	2.08 - 2.35	Mid infrared

DATA CLASSIFICATION

Computer-based classification of remotely sensed data may involve consideration of temporal, spectral, and spatial information. This map classification incorporates techniques from each of these methods. In each case, the objective is to identify consistent, quantitative differences between the desired classes of information.

TEMPORAL

Data from two different dates were selected to take advantage of the strong temporal differences in vegetation in the Bay Area, a result of the dry summer/wet winter regime. Thematic mapper scenes from March 1988, when vegetation was actively photosynthesizing and green, and July 1985, when much of the nonwoody vegetation was senescent, were used to characterize vegetation by differing phenological response from spring through the summer drought.

SPECTRAL

A "supervised classification" was used to differentiate land-cover classes by the character of the radiance returned to the sensor. Areas of known cover type were selected from the imagery, and the pixel values were used to quantitatively define that class of cover. The spectral characteristics of these representative samples, for example, the digital counts of the ground targets at the wavelengths recorded by the satellite, were compared to assess the separability of the classes. This was done using the mean and covariance of the pixel values of the sites, for each wavelength band. The bands that provided the best separation were selected for use in the classification. Eight bands were used (see table 1 for band wavelength ranges): Band 1, Band 6, and the ratios of Band 4 to Band 3 and Band 5 to Band 4 from each of the two dates. A "maximum likelihood classification" was then performed on the entire county-wide data set. In this procedure each pixel in the image was assigned to the class which it was statistically most likely to belong to. The rule used in this case, which compares the mean vector of each pixel to the mean and covariance descriptions of the known classes, is called the "maximum likelihood" decision rule (Richards, 1986).

The resulting preliminary classification was visually assessed for areas where misclassification was apparent. These classes were spectrally recharacterized using an unsupervised clustering algorithm to separate, where possible, the overlapping component of the class statistics. The new statistics were used to reassign the data cluster to the correct class.

SPATIAL

Spatial processing was used selectively to eliminate commission error in cases of small isolated pixels surrounded by another class. Commission error refers to the assignment of a pixel to a class to which it does not belong. For classes where such "noise" was a problem, a 3- by 3-pixel mode filter was used to assign each image pixel to the most frequent class of the nine surrounding pixels.

MAP ERROR ASSESSMENT

To assess the accuracy of the resulting classification, information on ground data was collected from USGS topographical quadrangle maps (roadways, urbanized areas), field visits, and color infrared (1:40,000) photography flown of the region in 1985. All available ground-confirmed data was collated and a comparative matrix was produced in which the diagonal elements represent pixels for which the classification and the ground data agree. The off diagonal elements indicate the amount and particular type of disagreement that occurred. The Kappa coefficient of agreement (Bishop and others, 1975; Dixon, 1990) was calculated for the matrix. This measurement of agreement between two variables takes into account that part of the accuracy resulted from chance alone:

$$\text{Kappa} = \frac{p_{\text{observed}} - p_{\text{expected}}}{1 - p_{\text{expected}}}$$

where

$p_{\text{observed}} = \Sigma\, x_{ii}/N \qquad p_{\text{expected}} = \Sigma\, x_{\text{row}_i}\, x_{\text{column}_i}/N^2$

p = proportion of agreement

x_{ii} = number of observations in a category which agree

N = total number of observations

The Kappa for the map, using all visually checked data available, indicated an overall accuracy of 82 percent.

The ground data collection was limited to areas accessible by roads, so it was not possible to take a truly random sample, a requirement of most statistical tests. In order to calculate a random test, it was necessary to rely on aerial photography for comparative information. For this error assessment, the minimum number of points required to calculate the accuracy at a 95-percent confidence level was determined (Fitzpatrick-Lins, 1981). Five hundred nine points were checked, which exceeded the minimum required to sample with a 95-percent confidence interval assuming an 85-percent accuracy. The points were randomly sampled from the area defined as the class in the final map. The corresponding points were located on the aerial photography when not available from the ground checks. A matrix was constructed as before. The resulting Kappa coefficient of agreement was found to be 95 percent.

CONSIDERATIONS FOR USE OF THIS DATA

Due to the instantaneous field of view of the satellite, which limits the pixel resolution size, the minimum mapping unit is limited to approximately 90 by 90 m, the size at which at least one pure pixel can be obtained. Smaller features cannot be assumed to be accurately mapped.

Disturbed ground with bare soil, such as dirt trails, fire roads, and stream banks were, in some places, mapped into the low-density-built category.

Stands of eucalyptus trees were not separated as a class.

The classification did not separate transitional classes such as mixed rangeland (coastal scrub and grassland), or mixed Douglas fir and Redwood areas.

ACKNOWLEDGMENTS

Funding for this work was provided by National Aeronautics and Space Administration through the Earth Observation Commercialization/Applications (EOCAP) Program.

We are grateful to Earl Brabb, who provided the concept for the project, advised on content and structure, and participated in field checking. Len Gaydos facilitated project funding, consulted on remote sensing techniques, and kept the project spirit alive. Brian Bennett graciously made two early "working" versions of this map.

REFERENCES CITED

Bishop, Y., Fienberg, S., Holland, P., 1975, Discrete Multivariate Analysis: Theory and Practice: MIT Press, Mass., p. 395-398.

Dixon, W. (ed.), 1990, BMDP Statistical Software Manual: University of California Press, Berkeley, Calif.

Fitzpatrick-Lins, Katherine, 1981, Comparison of sampling procedures and data analysis for a land-use and land-cover map: Photogrammetric Engineering and Remote Sensing v. 47, no. 3, p. 343-351.

Richards, J.A, 1986, Remote Sensing and Digital Image Analysis: An Introduction: Springer-Verlag, N.Y. 281 p.

Base from U.S. Geological Survey 1960
San Francisco Bay Region, Sheet 3
5000-meter Universal Transverse Mercator grid ticks, zone 10

APPROXIMATE MEAN DECLINATION, 1996

SCALE 1:62 500

CONTOUR INTERVALS 200 FEET
DOTTED LINES REPRESENT 40-FOOT CONTOURS
NATIONAL GEODETIC VERTICAL DATUM OF 1929
DEPTH CURVES IN FEET–DATUM IS MEAN LOWER LOW WATER
SHORELINE SHOWN REPRESENTS THE APPROXIMATE LINE OF MEAN HIGH WATER
THE MEAN RANGE OF TIDE IS APPROXIMATELY 4 FEET IN THE PACIFIC OCEAN AND 5 FEET IN SAN FRANCISCO BAY

AREA OF MAP

Mapped in 1990
Edited by Carol Ostergren; cartographic design by Kevin Ghequiere; cartographic production by Alan Vaughn
Manuscript approved for publication, March 9, 1994

[1]Terra-Mar Resource Information Services, Inc.
1710 South Amphlett Blvd. Suite 215,
San Mateo, CA 94402

Any use of trade, product, or firm names in this publication is for descriptive purposes only and does not imply endorsement by the U.S. Geological Survey

For sale by the U.S. Geological Survey Information Services, Box 25286, Federal Center, Denver, CO 80225

INTERIOR—GEOLOGICAL SURVEY, RESTON, VA—1996

MAP SHOWING VEGETATION IN SAN MATEO COUNTY, CALIFORNIA

By
Sally W. Buechel[1] and Anne F. Wagner

1996

ISBN 0-607-85479-0