UNITED STATES
DEPARTMENT OF THE INTERIOR
U.S. GEOLOGICAL SURVEY

PREDICTIVE SPATIAL MODELING OF NARCOTIC CROP GROWTH PATTERNS

By Frederick A. Waltz and Donald G. Moore

Open-File Report 86-007

Sioux Falls, South Dakota

1986

PREDICTIVE SPATIAL MODELING OF NARCOTIC CROP GROWTH PATTERNS

By Frederick A. Waltz
TGS Technology, Inc.

and

Donald G. Moore
U.S. Geological Survey

ABSTRACT

Spatial models for predicting the geographic distribution of marijuana crops have been developed and are being evaluated for use in law enforcement programs. The models are based on growing condition preferences and on psychological inferences regarding grower behavior. Experiences of local law officials were used to derive the initial model, which was updated and improved as data from crop finds were archived and statistically analyzed. The predictive models are changed as crop locations are moved in response to the pressures of law enforcement.

The models use spatial data in a raster geographic information system. The spatial data are derived from the U.S. Geological Survey's US GeoData, standard 7.5-minute topographic quadrangle maps, interpretations of aerial photographs, and thematic maps. Updating of cultural patterns, canopy closure, and other dynamic features is conducted through interpretation of aerial photographs registered to the 7.5-minute quadrangle base. The model is used to numerically weight various data layers that have been processed using spread functions, edge definition, and categorization.

The building of the spatial data base, model development, model application, product generation, and use are collectively referred to as the Area Reduction Program (ARP). The goal of ARP is to provide law enforcement officials with tactical maps that show the most likely locations for narcotic crops.

INTRODUCTION

Narcotic crops are an increasing problem for law enforcement personnel. Marijuana (Cannabis) is a major crop and has an estimated street value of up to $18 billion annually. This untaxed and unregulated flow of cash attracts violent crime and supports many criminal enterprises. Traditionally, the major supply of marijuana in this country has been illegally imported, but the domestic crop has recently increased due to law enforcement emphasis on foreign suppliers and to the development of seedless marijuana known as sinsemilla. Various levels of government are involved in the prevention, identification, and eradication of marijuana plots, but there are also formal organizations that promote and lobby for legalization of marijuana.

Other Federal agencies requested the U.S. Geological Survey's EROS Data Center to determine if Cannabis cultivation on Federal lands could be predicted using routinely available data and spatial data analysis techniques. This prediction, provided as probability estimates, has several objectives in existing law enforcement programs:

1. Eliminate the cultivation of marijuana on Federal lands.

2. Promote the safe use of Federal lands by the public by discouraging potential marijuana cultivation.

3. Reduce criminal activity on Federal lands.

4. Increase public awareness of the problem and the associated dangers.

This paper describes the techniques used to model and predict the distribution of marijuana plots in different landscapes under different law enforcement pressures. The model uses natural resource data, cultural data, terrain data, and human psychological responses. The current parameters and weightings of the model will not be presented, but general criteria and methods for using the geographic information system are documented.

SYSTEM CRITERIA

The Area Reduction Program (ARP) produces maps with probability estimates obtained by modeling topographic, natural, and cultural resources measurements. The models are derived initially from the experiences of local personnel together with known environmental restrictions to Cannabis growth. The models are continually refined as crop locations are statistically related to features and classes contained within the spatial data base. The program is a multistage approach with detailed output at the 7.5-minute quadrangle level and flight planning output at the 1° quadrangle level. Other targeting procedures are available for use at the State level.

The program of information analysis and management serves as institutional memory from year to year, and is field oriented from initial location of marijuana plots to surveillance and finally to arresting the offender and destroying the crop. The program is based on a dynamic statistical model whereby parameters may be changed from year to year as law enforcement pressures indicate the need to do so.

PROCEDURES

ARP is a multistage geographic information system in which information management is controlled at the regional level and actual area reduction and information management are carried out at the local level. Historic marijuana finds are registered into the data base and statistical descriptors are derived to associate crop locations with other layers of the data base. Roads and water sources for irrigation are included in the data base. Current land use and land cover information are derived through photographic interpretation and are then digitized and georeferenced for use with the other spatial information. Rationale for including the variety of data layers is related to known plant growth needs and to the ability of growers to conceal a crop until harvest.

The data base components are terrain data, land ownership, cultural features, transportation routes, water sources, canopy closure, and cultivation history. The terrain data (elevation, sun angle, aspect, and slope) are derived from the Survey's US GeoData and are related to various crop growth requirements. Natural and manmade features such as streams, ponds, roads, and dwellings are surrounded by buffers of sixteen 30-meter pixels for model weighting purposes.

The predictive model takes the following form:

$$P = (P_1 + P_2 \ldots + P_N)\ (L_1)\ (L_2)\ \ldots\ (L_N)$$

where,

P = probability of occurrence, and

P_N = distances, such as water, transportation, dwellings, and canopy edge.

L_N = restrictions, such as from elevations, ownership, and land surfaces.

The P's are additive, whereas the L's are multiplicative; the model is non-zero only where not restricted by multiplying by zero. This allows a reduction of the total search area and increases the efficiency of the program. The final model result is obtained by adding and multiplying the layers; interim results may be observed at each step on the display.

Because marijuana plots in the United States tend to be hidden on the south side of groves of trees, it is necessary to do edge detection on dense canopy and then to delineate the south side. This border layer is then given a high weight. Likewise, ownership lines are drawn within the image and, if necessary, land outside or inside specific ownership categories may be excluded from consideration. Another requirement may be to assign a higher weight to government land adjacent to private land, as such areas have been favored by many growers. Decision rules vary, however, with location and grower preferences. The model uses geographic analysis based upon these or other decision rules. The initial model is defined by collaboration with local land managers or law enforcement officers. Results are then correlated with known occurrences, and adjusted accordingly. This interactive approach between local experts and the geographic information system allows for the first level model definition. As actual plot locations are geographically entered into the data base, statistical relationships of the known locations and their individual data layers are derived and are used to update the model.

The final map may take several forms. For example, the output may be level sliced and the highest probability assigned to 15 percent of the area. Thus, 15 percent of the spatial area of a quadrangle would be colored red. The second slice may also be 15 percent and colored green. Usually no more than the top 50 percent of the probability model is shown on the map. Roads and streams may be added for geographic orientation and field location of the user. The preferred method is to present the model result on transparent mylar, which can be overlaid on the 7.5-minute quadrangle.

RESULTS

In the 1984 model for a test quadrangle, the highest probability area (30 percent of the total area), contained 88 percent of the 1984 finds; the medium 20 percent of the spatial area contained 9 percent of the finds, and the bottom 50 percent contained 3 percent of the finds. Thus, a search program could be instituted to cover only 30 percent of the area with the probability of finding 88 percent of the plots. Although finds were not used to build the model for 1984, the experience of enforcement officers was used in developing the parameter weights.

The 1985 model was developed for all or parts of 36 quadrangles, which were largely U.S. Forest Service land. Historic (1984 and before) finds were plotted on the finished product for reference. Sixty-eight percent or two-thirds of the historic finds fell into the high probability slice which was 30 percent of the total area. Another 15 percent of the finds were found in the medium probability slice, which was 20 percent of the spatial area. Nine percent of the finds were in the 50 percent area of low probability. An interesting fact was noted in that 8 percent of the finds were outside the high probability area, but were on public land adjacent to private land. This may be a result of the U.S. Supreme Court ruling that private land is subject to confiscation when illegal crops are found.

The percentage of 1985 finds on the high probability area of the 1984 model was considerably lower (49 percent) than the historic finds. This was due in part to law enforcement pressure. For example, the preferred terrain aspect prior to 1985 was southwest slopes. In 1985, growers moved to northwest slopes by a 2-to-1 margin. The overall aspect distribution of the area was uniform for this test site (no one aspect predominated). Likewise, pre-1985 finds were within 90 feet of wells, ponds, and springs, but the 1985 finds were located up to 400 feet away. The largest single change was the type of cover in which gardens were located, with 80 percent of the 1985 finds located in old clearcut, regeneration areas. Fifty percent of the gardens were hidden from aerial view, making remote sensing techniques ineffective. As an example of updating by statistical analysis, the finds for 1985 were heavily concentrated around unimproved roads. A spread function was used to mark distances from roads (each pixel being 30 m by 30 m). A Chi-square function was then used to measure the observed versus the expected number of finds at each distance away from the roads. Pixels on the roads and those beyond the 16-pixel limit were labeled zero. Table 1 gives the number of pixels from a road and its Chi-square value. The data indicate most gardens were within 7 pixels of a road. However, care must be taken when interpreting the Chi-square test. The pixel value zero would seem to indicate plots, but the opposite is true because the value is squared; deviations in either direction give large values. In fact, the zero pixel observed value was a great deal less than the expected value.

Table 1.--Significance of distance to roads from plots

The pixel (30 m) distance normal to the road with the Chi-square value representing the observed, minus the expected plots quantity squared, divided by the expected plots. Large values indicate deviations from the expected values, with the astericked (*) values significant at the 95-percent level of probability.

Distance Pixel from Road	Chi-square
0	*7.53
1	2.33
2	*10.58
3	.02
4	*5.15
5	.84
6	1.91
7	*4.20
8	.68
9	.03
10	.02
11	.84
12	.58
13	.84
14	.08
15	.00
16	.00

CONCLUSIONS

The use of spatial predictive models in locating marijuana crops has proved to be of value to law enforcement personnel. The accuracy of the model depends heavily on proper interpretation of all input factors, especially the human psychological factors. The models are adaptive from year to year as parameters change and can be interactively modified on a daily basis as new information on changes in cultural practices is obtained.