STIF

04

III

IMPROVING ESTIMATES OF STREAMFLOW CHARACTERISTICS USING LANDSAT-1 (ERTS-1) IMAGERY

E7.6-10230
CR-146513

Este F. Hollyday
U.S.G.S., W.R.D.
A-413 Federal Bldg.
U.S. Courthouse
Nashville, TN 37203

(E76-10230) IMPROVING ESTIMATES OF STREAMFLOW CHARACTERISTICS USING LANDSAT-1 (ERTS-1) IMAGERY Final Report, 1 Jul. 1972 - 30 Jun. 1975 (Geological Survey, Reston, Va.) 64 p HC $4.50 CSCL 08H G3/43

N76-20580

Unclas
00230

31 August 1975

Type III Final Report for Period 1 July 1972 - 30 June 1975

Prepared for:

Goddard Space Flight Center
Greenbelt, Maryland 20771

TECHNICAL REPORT STANDARD TITLE PAGE

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Improving estimates of streamflow characteristics using Landsat-1 (ERTS-1) Imagery (1342F)		5. Report Date 31 August 1975 6. Performing Organization Code
7. Author(s) Este F. Hollyday (IN 389)		8. Performing Organization Report No.
9. Performing Organization Name and Address U.S. Geological Survey, WRD National Center 12201 Sunrise Valley Drive Reston, Virginia 22092		10. Work Unit No. 11. Contract or Grant No. S-70243-AG
12. Sponsoring Agency Name and Address Frederick Gordon, Jr. Code 902 Goddard Space Flight Center Greenbelt, Maryland 20771		13. Type of Report and Period Covered Type III Final Rept. 1 July 72 - 30 June 75 14. Sponsoring Agency Code

15. Supplementary Notes

Original photography may be purchased from:
EROS Data Center
10th and Dakota Avenue
Sioux Falls, SD 57198

16. Abstract Imagery from the first Earth Resources Technology Satellite was used to discriminate physical features of drainage basins in an effort to improve equations used to estimate streamflow characteristics.

Records of 20 gaged basins in Maryland and Delaware were analyzed for 40 statistical streamflow characteristics. Equations relating these characteristics to basin characteristics were obtained by a technique of multiple linear regression. A control group of equations used basin characteristics derived from maps. An experimental group of equations used basin characteristics derived from maps and imagery. Comparison of control group equations with those from the experimental group reveals that for 12 out of 40 equations the standard error of estimate was reduced more than 10 percent. As an example, the standard error of estimate of the equation for the 5-year recurrence-interval flood peak was reduced from 46 to 32 percent. Similarly, the standard error of the equation for the mean monthly flow for September was reduced from 32 to 24 percent, the standard error for the 7-day, 2-year recurrence low flow was reduced from 136 to 102 percent, and the standard error for the 3-day, 2-year flood volume was reduced from 30 to 12 percent.

It is concluded that data from Landsat imagery can substantially improve the accuracy of estimates of some streamflow characteristics at sites in the Delmarva Peninsula.

17. Key Words Suggested by Author	18. Distribution Statement
Remote Sensing, Satellites Watersheds (basins), ERTS Terrain analysis, Landsat Surface runoff Regression analysis Statistical models	No restriction on distribution

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	N?A	63	

Figure 2A. Technical Report Standard Title Page. This page provides the data elements required by DoD Form DD-1473, HEW Form OE-6000 (ERIC), and similar forms.

INT 2090-72

PREFACE

The image analysis phase of this investigation was designed to use the time-lapse analysis capabilities of ESIAC (Electronic Satellite Image Analysis Console) at the Stanford Research Institute. This equipment was being designed and tested under NASA Contract NAS 5-21841 concurrently with this investigation. In addition, this investigation used the photographic extractions of basin characteristics being produced from satellite imagery by the Topographic Division, U.S. Geological Survey, as part of a thematic mapping investigation under NASA Contract S-70243-AG.

The objective of the investigation was to test two hypotheses. First, it was postulated that indices of dynamic basin characteristics could be extracted from imagery acquired by the Earth Resources Technology Satellite (now Landsat). Second, it was postulated that these indices could be used quantitatively to improve equations for estimating streamflow characteristics.

In order to test the hypothesis that basin characteristics derived from Landsat imagery can be used to improve equations for estimating streamflow, it was necessary first to extract the data from the imagery and then to use the data in a simple analytic experiment. Several techniques were tried for accurately extracting and measuring basin characteristics; some were representative of state-of-the-art at the time they were tried. The analytic experiment involved computing and comparing two sets of multiple linear regression equations, where one set contained basin characteristics derived from Landsat imagery. This analytic experiment was considered the primary objective limiting the scope of work.

It is concluded that some basin characteristics can be extracted easily from a single set of Landsat imagery with the aid of multispectral image analysis systems that employ film-density discrimination techniques only. Two of the easily-extracted basin characteristics can be used to substantially improve the accuracy of equations used to estimate 12 out of 40 streamflow characteristics for streams on the Delmarve Peninsula of Delaware, Maryland, and Virginia. This improved accuracy could allow a reduction in effort needed to collect data that is used to define these streamflow characteristics at gaged sites. The savings in manpower and money could be applied to meeting other current and anticipated data needs.

New image analysis techniques need to be tested as to their capabilities for operational output of basin characteristics data. Techniques using digital data from computer compatible tapes are a likely source of data for expanding the matrices used to develop better equations for estimating streamflow.

Consideration might be given to testing basin characteristics derived from satellite data for improving estimates of streamflow in other physiographic regions in the nation.

Consideration might also be given to the possibility of applying results of this investigation nationally by developing an expanded basin characteristics file for correlation with the national streamflow data files.

CONVERSION FACTORS

Factors for converting English units to metric units are shown to four significant figures. However, in the text the metric equivalents are shown only to the number of significant figures consistent with the values for the English units.

English	Multiply by	Metric
ft^3/s (cubic feet per second)	2.832×10^{-2}	m^3/s (cubic metres per second)
ft (feet)	3.048×10^{-1}	m (metres)
in (inches)	$2.540 \times 10^{+1}$	mm (millimetres)
mi (miles)	1.609	km (kilometres)
mi^2 (square miles)	2.590	km^2 (square kilometres)

CONTENTS

ILLUSTRATIONS

TABLES

IMPROVING ESTIMATES OF STREAMFLOW CHARACTERISTICS USING LANDSAT-1 (ERTS-1) IMAGERY

By Este F. Hollyday, Nashville, TN.

Work done in cooperation with the National Aeronautics and Space Administration

Abstract.--Imagery from the first Earth Resources Technology Satellite (renamed Landsat-1) was used to discriminate physical features of drainage basins in an effort to improve equations used to estimate streamflow characteristics at gaged and ungaged sites. Records of 20 gaged basins in the Delmarva Peninsula of Maryland, Delaware, and Virginia were analyzed for 40 statistical streamflow characteristics. Equations relating these characteristics to basin characteristics were obtained by a technique of multiple linear regression. A control group of equations contains basin characteristics derived from maps. An experimental group of equations contains basin characteristics derived from both maps and imagery. Characteristics from imagery were forest, riparian (stream-bank) vegetation, water, and combined agricultural and urban land use. These basin characteristics were isolated photographically by techniques of film-density discrimination. The area of each characteristic in each basin was measured photometrically.

Comparison of equations in the control group with corresponding equations in the experimental group reveals that for 12 out of 40 equations the standard error of estimate was reduced by more than 10 percent. As an example, the standard error of estimate of the equation for the 5-year recurrence-interval flood peak was reduced from 46 to 32 percent. Similarly, the standard error of the equation for the mean monthly flow for September was reduced from 32 to 24 percent, the standard error for the 7-day, 2-year recurrence low flow as reduced from 136 to 102 percent, and the standard error for the 3-day, 2-year flood volume was reduced from 30 to 12 percent.

It is concluded that data from Landsat imagery can substantially improve the accuracy of estimates of some streamflow characteristics at sites in the Delmarva Peninsula.

INTRODUCTION

Nationally the U.S. Geological Survey operates more than 10,000 stations that are used for gaging streamflow. Records of stage collected continuously at gaging stations are converted to streamflow and are published annually. These data are useful to planners, engineers, and water managers for designing water supply reservoirs, controlling pollution, designing bridges, managing flood plains, forecasting and managing floods, producing power, and designing and maintaining navigation facilities. The purpose of this investigation is to determine if imagery from Landsat can improve on one aspect of this program, namely regionalization of streamflow information, or the transfer of streamflow records from gaged to ungaged sites.

In 1970 the Survey initiated a study to evaluate the national streamflow data program which had evolved over the previous 80 years (Benson and Carter, 1973). Existing data were evaluated in terms of newly-established goals, and a proposal for continuation of a revised program was released to the open file (Forrest and Walker, 1970).

As part of this evaluation, the records for each gage on unregulated streams were analyzed to derive statistical measures of flow termed streamflow characteristics, which include average flows, variability in average flows, flood peaks, flood volumes, and low flows. Using multiple linear regression techniques the streamflow characteristics were then correlated with basin characteristics, which are selected physiographic and climatic features of the corresponding drainage basins. This regression analysis was done in part to transfer the streamflow records from gaged to ungaged sites, and in part to evaluate the record. The statistical model that was used is:

$$Y = aX(1)^{b(1)} X(2)^{b(2)} \ldots X(n)^{b(n)},$$

where Y is a streamflow characteristic, X(1) to X(n) are basin characteristics, and a, b(1) to b(n) are the regression constant and coefficients.

The equations so generated were used to determine data collection needs by comparing the accuracy of the equations with the accuracy goals specified for estimates of flow characteristics at ungaged sites. These accuracy goals were given in terms of equivalent years of record. This means it was specified that information provided for any ungaged point on a stream should be equivalent in accuracy to that which would have been attained by an actual record of a selected number of years (10 or 25) at that point. Accuracy goals in terms of equivalent years of record in a given state or region were converted to standard error in percent of mean using the methods described by Hardison (1969). Independent of specifying these accuracy goals, a value called the standard error of estimate, was computed for each regression equation.

This value is a general index of the accuracy of estimates obtained by use of the equation. By comparing the standard error of estimate of each streamflow characteristic with the accuracy goal for that characteristic, it was possible to judge the degree to which some goals had already been achieved by existing data and to judge the need for continued data collection. According to the concepts used in the evaluation study (Benson and Carter, 1973), about 10 percent of the present streamflow program effort should be redirected to areas of higher priority.

Since 1970 the streamflow data program is being reevaluated periodically, and this investigation is pertinent to the continuing evaluation effort in that it tests the impact of additional basin characteristics on estimates of streamflow characteristics. If basin characteristics derived from Landsat data are added to an equation and if this addition results in a reduced standard error of estimate of the streamflow characteristic then the new standard error may be less than the accuracy goal. In such a case, a reduction in effort to collect data on that characteristic would be in order. In this way, additional basin characteristics derived from satellite data could have a substantial impact on the streamflow information program.

The basin characteristics used in the regression analysis that was part of the evaluation were compiled from U.S. Geological Survey topograpic maps and from National Weather Service Climatological Data. The maps are not the most suitable source of information on basin characteristics. The maps vary in scale and detail; they also vary in age of photography used for compiling land-cover information. In order to have maximum ground visibility, mapping photography is usually taken in early spring before leafing-out of trees which generally coincides with conditions of high water in the eastern half of the nation. The scale is commonly 1:20,000 or about 3 mi (5 km) on a photograph edge.

On July 23, 1972, the U.S. National Aeronautics and Space Administration (NASA) launched the first Earth Resources Technology Satellite (renamed Landsat-1) capable of repeatedly and uniformly imaging the Earth. The opportunity is presented to evaluate basin characteristics extracted from satellite images. Although the resolution of the Landsat-1 system cannot compare with standard mapping photography, several advantages are foreseen in using Landsat imagery as a source of basin characteristics rather than topographic maps or the photographs from which they were compiled. Landsat can provide seasonal information on land-cover conditions rather than early spring conditions only. A single Landsat image covers an area of 13,200 mi^2 (34,200 km^2) or 115 mi (185 km) on an image edge. The view has nearly uniform lighting conditions allowing more uniform application of criteria for extracting basin characteristics than a photographic mosaic of comparable coverage. An interpreter can select those characteristics believed to be most closely related to hydrology rather than accepting the standard cultural information on topographic maps.

Many previous applications of Landsat imagery to water problems have emphasized the inventory or mapping of land-cover types or conditions that should be related to hydrology. This investigation sought to test quantitatively the inferred relationships between streamflow and selected land-cover types. This report describes results of testing the hypotheses that hydrogically-significant basin characteristics can be extracted from Landsat-1 imagery, and that these characteristics can be used quantitatively to improve equations for estimating streamflow characteristics.

PHYSIOGRAPHY OF STUDY AREA

The study area selected for testing Landsat imagery in this investigation encompasses most of the Delmarva Peninsula and part of the adjacent mainland covered by Landsat-1 image 1079-15133 and succeeding images with comparable coverage (fig. 1). The area lies within the Central Atlantic

---

Figure 1.--(caption on next page) belongs near here.

Regional Ecological Test Site (CARETS) of the U.S. Geological Survey. It is part of the Chesapeake Bay group of Landsat investigations funded by the National Aeronautics and Space Administration.

Twenty gaged drainage basins were selected within the study area (table 1). Drainage areas range from 3.85 mi^2 (9.97 km^2) to 113 mi^2 (293 km^2) and average 24.6 mi^2 (63.7 km^2). From 11 to 31 years of streamflow data were available for each gage when the data were evaluated in 1970.

Mean annual precipitation ranges from 46 in (1170 mm) to 48 in (1220 mm). Mean annual temperature ranges from 12 ° to 14°C, and extremes are moderated by nearby Chesapeake Bay and Atlantic Ocean. During winter there is a 45 percent probability that the area will be cloud covered during a satellite overpass. By using parts from as many as five successive images of the same scene, however, complete coverage each season is practically assured.

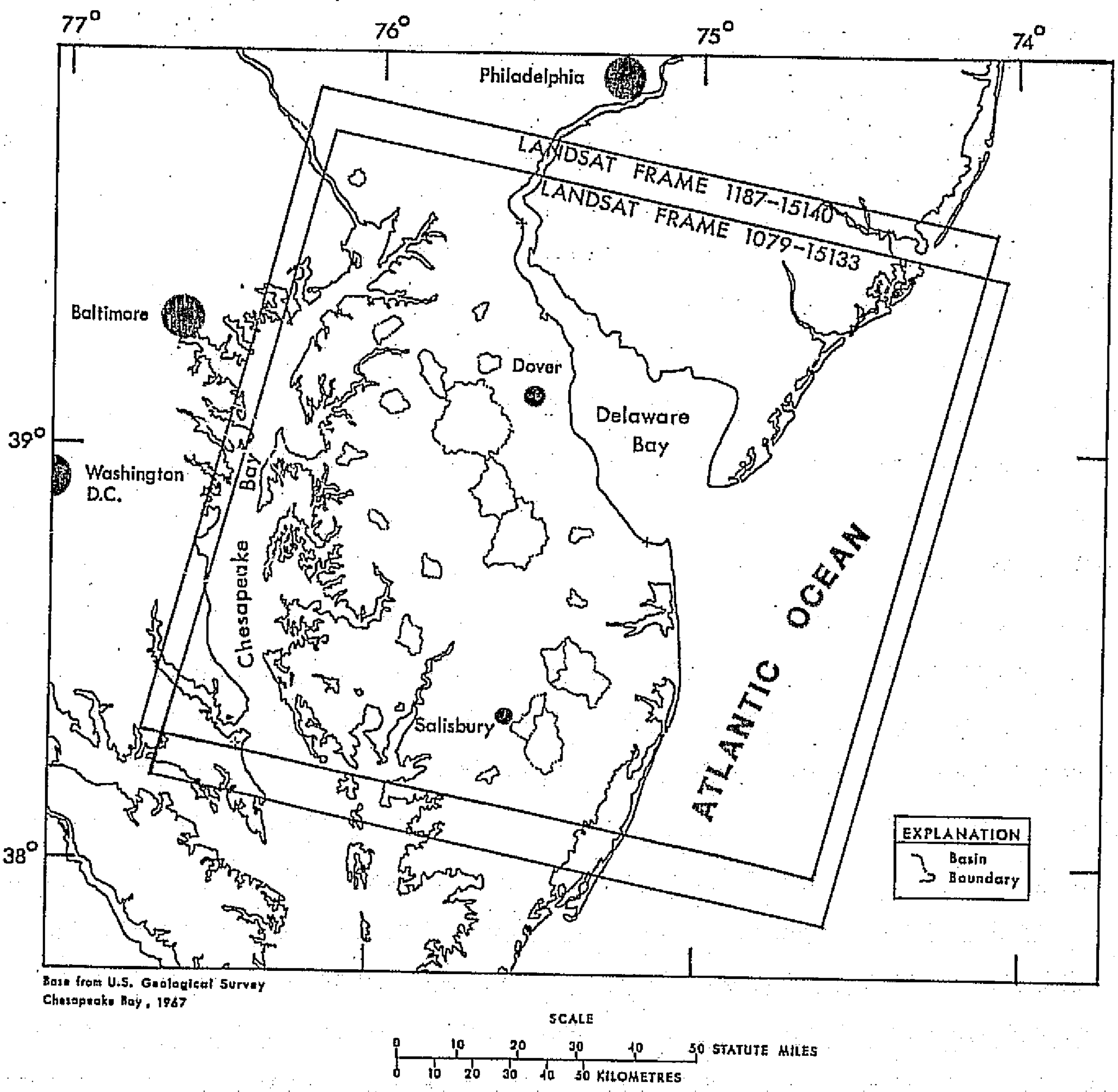

Figure 1.--Study area showing location of Landsat frames and 20 gaged drainage basins.

Table 1.--Gaging stations (drainage basins) used in multiple regression analysis.

U.S.G.S. station number	Station name	Latitude	Longitude	Drainage area, in mi^2	Period of record analyzed
01483200	Blackbird Creek at Blackbird, Del.	392158	0754010	3.85	1956-67; Annual max. 1952-56
01483500	Leipsic River near Cheswold, Del.	391358	0753757	9.35	1933; 1943-57; Annual max. 1958-67
01484300	Sowbridge Branch near Milton, Del.	384851	0751939	7.08	1956-67
01484500	Stockley Branch at Stockley, Del.	383819	0752031	5.24	1943-67
01485000	Pocomoke River near Willards, Md.	382320	0751930	60.5	1949-67
01485500	Nassawango Creek near Snow Hill, Md.	381344	0752819	44.9	1949-67
01486000	Manokin Branch near Princess Anne, Md.	381250	0754018	5.8	1951-67
01486500	Beaverdam Creek near Salisbury, Md.	382105	0753411	19.5	1936-67
01487000	Nanticoke River near Bridgeville, Del.	384345	0753341	75.4	1943-67
01487500	Trap Pond Outlet near Laurel, Del.	383140	0752858	16.7	1951-67
01488500	Marshyhope Creek near Adamsville, Del.	385059	0754024	44.8	1943-67
01489000	Faulkner Branch at Federalsburg, Md.	384244	0754734	7.10	1950-
01490000	Chicamacomico River near Salem, Md.	383045	0755250	15.0	1951-
01491000	Choptank River near Greensboro, Md.	385950	0754709	113	1948-
01492000	Beaverdam Branch at Matthews, Md.	384841	0755815	5.85	1950-
01492500	Sallie Harris Creek near Carmichael, Md.	385755	0760630	8.09	1951-56; Annual max. 1957-67
01493000	Unicorn Branch near Millington, Md.	391459	0755140	22.3	1948-67
01493500	Morgan Creek near Kennedyville, Md.	391648	0760054	10.5	1951-67
01494000	Southeast Creek at Church Hill, Md.	390757	0755851	12.5	1951-56; Annual max. 1957-65
01579000	Basin Run at Liberty Grove, Md.	393930	0760610	5.31	1948-58; Annual max. 1965-67

The basin for station number 01579000 lies within the Piedmont Plateau physiographic province; all other basins are situated on the Peninsula within the Atlantic Coastal Plain province. This part of the Coastal Plain is a flat, low almost featureless plain. Maximum elevations throughout the Peninsula rarely exceed 80 feet (24 m) above mean sea level. Maximum elevations in the Piedmont may exceed 400 feet (120 m) near the north end of the Peninsula. Relief within any one square mile may have a value as much as half the local maximum elevation. The shoreline of the Peninsula is extremely broken and sinuous along the Chesapeake Bay and is characterized by very small relief and tidal marshes. The shore of the Atlantic Ocean is composed of a long line of barrier beaches with lagoons on the landward side. All gaging stations are located away from the shore and above tidal influence. The Coastal Plain is underlain by a series of southeasterly dipping layers of unconsolidated sand and clay with a subordinate amount of gravel. These layers overlie the weathered crystalline rock of the Piedmont and thicken to the southeast from the northwest edge of the Coastal Plain. These materials are drained by sluggish rivers, many of which have been or are being channelized above tide water for the purpose of improving the drainage of agricultural land.

Before colonial time, the area was completely covered with forest. The amount of land cleared for agriculture has varied throughout history. Currently, second- and third-growth forest is largely restricted to river flood plains, swamps, and wet upland depressions in the northern two-thirds of the area and also to state forest preserves in the southern one-third of the area. Upland forest is composed predominantly of pine, mostly Virginia pine in the northern part of the Peninsula and loblolly pine in the south. Lowland areas have a distinctive swamp hardwood forest composed of pin, willow, and swamp oak, red and black gum, red maple, river birch, yellow poplar, sycamore, beech, and walnut (Vokes, 1957). Permanently-flooded or tidal, fresh-water areas in the south have extensive stands of bald cypress. Agricultural land in the northern half of the study area, is used predominantly for dairy farming, livestock, and feed grain, particularly corn. In the southern half it is used for truck farming and poultry production.

Level 1 land use (Anderson and others, 1972) within the 20 selected drainage basins was predominantly forest and agriculture in 1970 with the latter category being most common in the northern half of the study area (E. J. Pluhowski, written commun. 1974).

Level 1 land-use category	Land use within basins, in percent range	average
Urban and built up	0.0 - 5.1	0.5
Agriculture	20.2 - 96.2	58.4
Forest	3.8 - 79.6	40.9
Water	0.0 - 0.8	0.1

LANDSAT-1

Landsat-1 is an experimental satellite for demonstrating that remote sensing from space with unmanned satellites is feasible and can provide valuable data to assist efficient management of water and other earth resources (National Aeronautics and Space Administration, 1972). Although the satellite was launched in July 1972 with a design life of 1 year, it has continued to provide imagery and to relay data through June 1975. In order to allow systematic, repetitive imaging of the Earth under nearly constant lighting conditions, it was launched and maneuvered into a circular, near-polar orbit that allows the satellite to repeat its ground trace at the same local solar time every 18 days. This is 0942 hours at the equator and about 10 minutes earlier for most of the lower 48 states and Hawaii.

The satellite contains two multispectral imaging systems, a data relay system, and support systems. The Multispectral Scanner (MSS) is a line scanning device with arrayed detectors rather than film or sensitized-phosphor plates. It detects daylight solar energy reflected from the Earth's surface in the visible and near (non-thermal) infrared region of the spectrum. This energy passes through a single optical system that allows it to be recorded in four spatially-registered spectral bands.

Multispectral Scanner band	Wavelength, in micrometres
4	0.5 - 0.6
5	0.6 - 0.7
6	0.7 - 0.8
7	0.8 - 1.1

The human eye is sensitive to an 0.4 to 0.7 micrometre wavelength band. The satellite system scans from west to east at right angles to the path of travel. Forward motion of the satellite provides continuous coverage along the orbital track. The detector response is sampled, bit encoded, and transmitted to a ground station. During image processing, the continuous data stream is located geographically, corrected geometrically and radiometrically, and framed to produce a 55-mm image on 70-mm film. Each frame covers an area of 34,200 km^2. Further processing provides a variety of products and formats including black and white or color-composite images with a scale of 1:1,000,000 on 9-1/2-inch (240-mm) film. Each image has 10-percent forward lap with immediately previous and succeeding images and about 20 percent side-lap for imagery of most of the nation. Spatial resolution for the images averages about 80 m, but may be considerably higher or lower depending upon scene contrast.

BASIN CHARACTERISTICS

CHARACTERISTICS FROM LANDSAT IMAGERY

Spectral Reflectance

A basin characteristic observed in Landsat imagery absorbs and reflects solar radiation in its own way. The incoming solar radiation striking an object may be compared with the resulting outgoing or reflected radiation for small increments of wavelength. The resulting data can be used to produce spectral reflectance curves for the selected object (fig. 2). Dry snow reflects greater than 50 percent and clear

---

Figure 2.--(caption on next page) belongs near here.

---

water reflects less than 10 percent of solar radiation over the band width 0.3 to 1.1 micrometres. As a result dry snow appears very bright in cloud-free imagery for all four MSS bands, and water appears very dark, especially in bands 6 and 7. Color films that have emulsion layers sensitive to blue, green, and red light in the visible spectrum portray forest as green because maximum reflected solar radiation occurs in the green part of the visible spectrum. An infrared color composite of Landsat bands 4, 5, and 7, however, portrays forest as red because maximum reflected radiation occurs in the near-infrared or non-visible part of the spectrum between 0.8 and 1.1 micrometres, band 7.

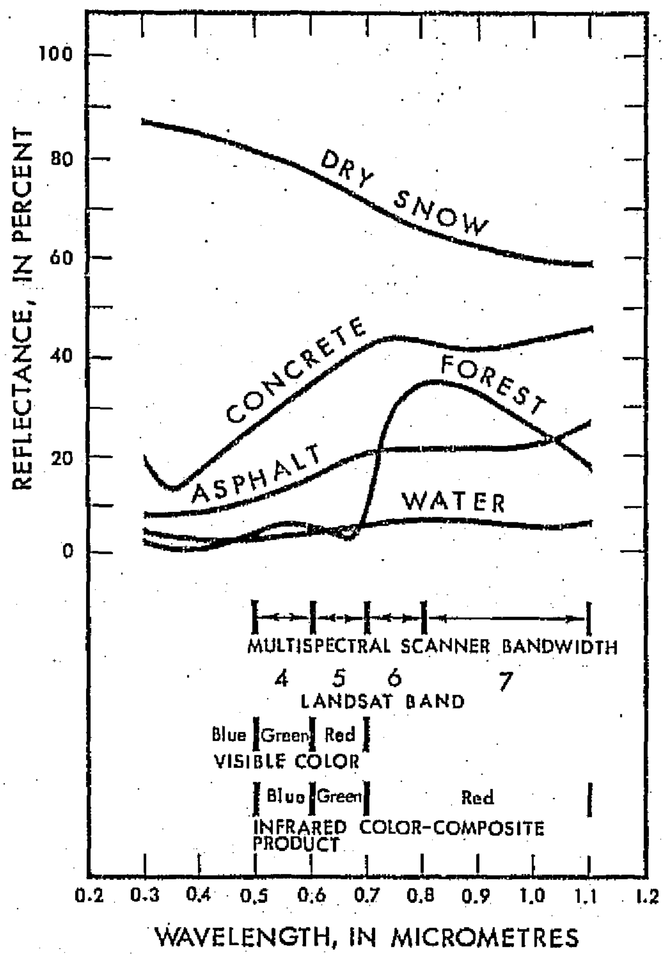

Figure 2.--Generalized spectral reflectance curves for five land-cover categories compared to Landsat spectral bands. Curves for concrete, asphalt, forest, and water are modified from Root and Miller (1971); for dry snow from Serebreny and others (1974).

In theory, a specific basin characteristic can be isolated in Landsat imagery from others by comparing film densities (which are related to spectral reflectance); however, difficulties are frequently encountered when attempts are made to isolate a characteristic using film densities only. For a variety of reasons there may be significant overlap among spectral curves. For example, a thin cloud cover between ground and satellite tends to increase the apparent spectral reflectance of all terrain features. In this way water with a thin cloud cover may easily be confused with asphalt in another part of the same scene where there is no cloud cover. Suspended sediment also increases the spectral reflectance of water. Thus water with sediment may be confused with asphalt in coastal cities even in cloud-free imagery. Conversely, shadow from either cloud or terrain tends to decrease the apparent spectral reflectance. In this way, asphalt in shadow may be confused with water in open sunlight. In addition to these problems of isolating a specific basin characteristic in a single Landsat image, seasonally, the angle of solar illumination changes, trees lose or gain their leaves, and soil moisture varies. Accordingly, criteria for isolating basin characteristics must also change from scene to scene throughout the year.

These difficulties limit the success of isolating a basin characteristic that has been selected solely on the basis of hydrologic significance. Some preselected characteristics do not have a unique spectral response that is always detectable with the Landsat system. As a result the selection of basin characteristics to be derived from Landsat imagery is a process involving compromise between hydrologic significance and ease of extraction.

Selected Basin Characteristics

Four basin characteristics were finally selected after giving consideration to hydrologic significance, ease of extraction, and availability of hydrologic and remote-sensing data in the study area. The four characteristics and their symbol designations are: forest (Uf), water (Uw), riparian vegetation (Urv), and combined agricultural and urban land use (Uau).

Forest (Uf) is the relative area of a drainage basin that is covered by trees, expressed in percentage of total basin area. In the study area this includes upland pine forest as well as lowland, swamp hardwood forest. Forest appears to affect streamflow indirectly by affecting rates of evapotranspiration, precipitation interception, and the accumulation and melting of snow. Forest cover was found significant at no less than the 5 percent level in estimating 19 out of 41 streamflow characteristics (Forrest and Walker, 1970). Forest has a unique spectral response detectable with the Landsat system.

Water (Uw) is the relative area of a drainage basin that is covered by a water surface visible from above, expressed in percentage of total basin area. In the study area it is the area of mill ponds, industrial waste lagoons, and flooded fields. Water is a measure of surface water storage which affects evaporation, peak flows for short- and intermediate-recurrence intervals, and low flows of natural streams. Storage was found significant at no less than the 5 percent level in estimating 2 out of 41 streamflow characteristics (Forrest and Walker, 1970). Clear water has a unique spectral response that is easily detected with the Landsat system.

Riparian vegetation (Urv) is the relative area of a drainage basin that is covered by vegetation situated on or near the bank of a stream or other water body, expressed in percentage of total basin area. In the study area it is area of shrubs and trees in swamps, marshes, and seasonally-flooded depressions. The roots of riparian vegetation have constant access to water and thereby affect streamflow, particularly intermediate and low flow, by affecting rates of evapotranspiration. The roots can either withdraw water directly from a stream or intercept water moving through the ground to a stream. Riparian vegetation was not used previously in estimating streamflow characteristics. In some areas riparian vegetation has a unique spectral response, that can be detected in Landsat imagery for winter months.

Combined agricultural and urban land use (Uau) is fields, pasture, buildings, roads, and sand pits, expressed in percentage of total basin area. In the study area it is all areas not covered by either forest or water. No attempt was made to extract this basin characteristic from Landsat imagery. It was arbitrarily measured as the difference between total basin area and the combined area of forest (Uf) and water (Uw).

Image Analysis

To make a quantitative test of the inferred relationships between streamflow and selected basin characteristics derived from Landsat imagery the basin characteristics are first measured and then used in an analytic experiment. Figure 3 shows the key steps in this analysis.

Figure 3.--(caption on next page) belongs near here.

Cloud-free imagery for each season is selected from catalogued Landsat data. The imagery is inspected to estimate the ease of isolating the selected characteristics from a single band, or a combination of bands, over one or more seasons. Characteristics such as forest are then isolated and extracted. The relative area of each characteristic in each basin is then measured. After image analysis, the characteristics derived from Landsat data are merged with an available matrix of map-derived basin characteristics. The new matrix is then correlated with a matrix of streamflow characteristics by a technique of multiple regression.

During image analysis, effort was concentrated within the 20 gaged basins. A mask was used to black-out all of an image except the areas within the basins. The basin mask was prepared by transferring the basin boundaries from topographic maps to opaque film. The opaque layer of this film was then removed from the area enclosed by the boundaries. Geographic control points were then added to allow registration of the mask with imagery having a scale of 1:1,000,000.

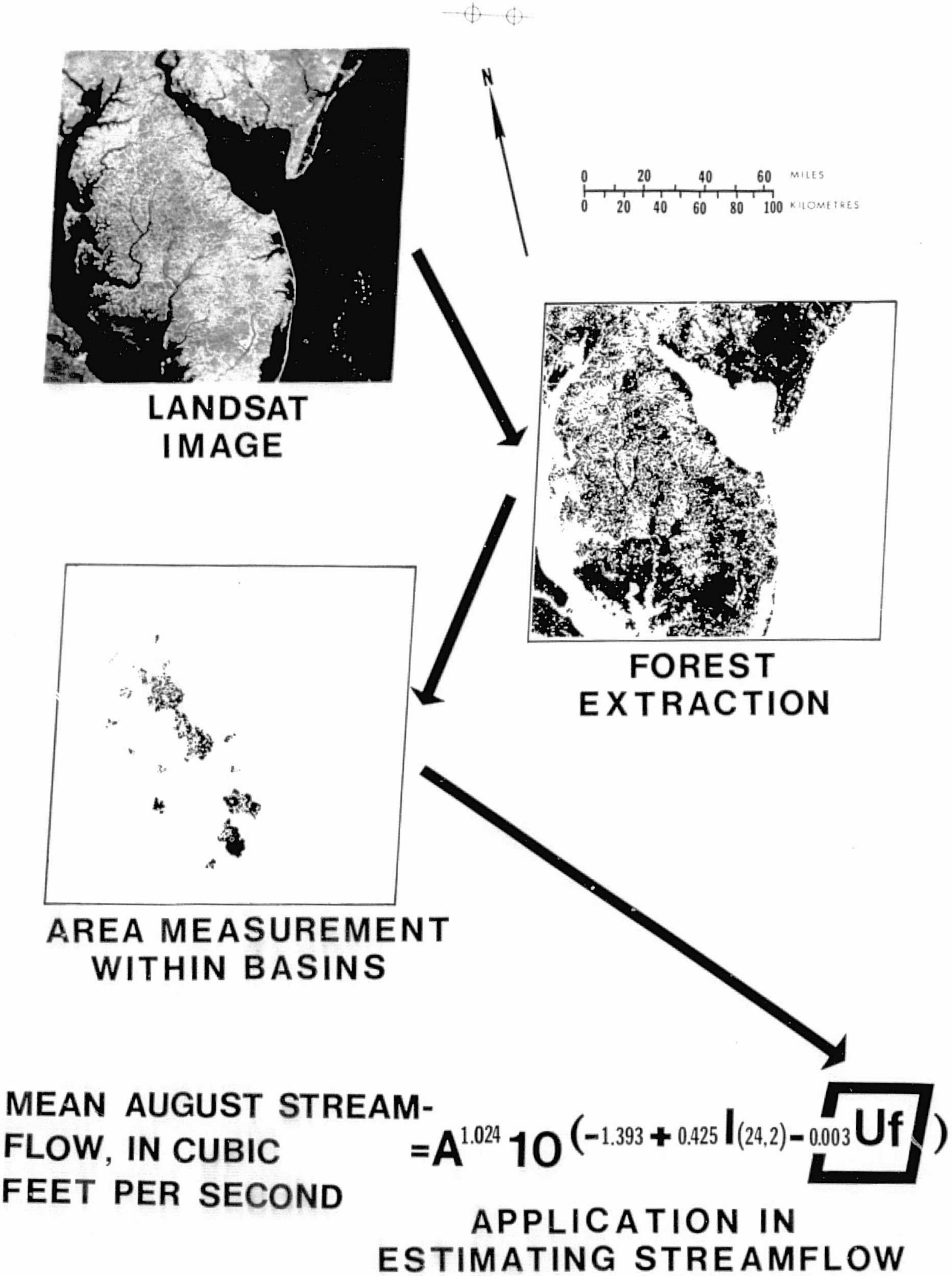

Figure 3.--Flow diagram of technique for improving equations used to estimate streamflow characteristics. Forest areas are extracted from Landsat image 1079-15133, are measured basin by basin, and are added to characteristics used previously in developing regression equations.

Manual extractions of forest and water were prepared as guides to automated image analysis using multispectral image analysis systems and using photography. A positive film transparency of band 5 of image 1079-15133 was overlain with the basin mask and a clear film was placed over the mask. The distribution of the selected basin characteristic was estimated visually and drawn on the clear film by subjectively applying criteria of film density, texture, shape, and terrain position relative to the drainage pattern. Areas believed to meet the criteria were inked-in to complete the extraction.

Seasonal changes that might be useful in identifying and extracting basin characteristics were detected by time-lapse processing of two or more images of the same scene (Serebreny and others, 1974). Color displays made of a magnified part of image 1079-15133 of October 10, 1972, and of image 1187-15140 of January 26, 1973, using ESIAC (Electronic Satellite Image Analysis Console at Stanford Research Institute) revealed that riparian vegetation could be distinguished from upland hardwood forest in the winter image of basin 01485000. Values of scene radiance along a cursor trace through the image of the basin indicated that the variation in film density (directly related to scene radiance) was just large enough to isolate riparian vegetation by equidensitometry.

Photomechanical extractions were made of forest, riparian vegetation, and water using one or more bands of images 1079-15133 and 1205-15141 of February 13, 1973. The film density of a selected basin characteristic throughout the image was measured along with the density of selected steps in the image gray scale. A suitable range in density was selected for isolating the characteristic and was specified for photochemical control in a high-quality graphic arts photographic laboratory. The laboratory used suitable density-isolating films to produce binary extractions in the form of photographic transparencies. Two or more of these transparencies can form a composite or sandwich that reduces unwanted or spurious data while isolating the desired basin characteristic. The photomechanical extractions used in this investigation were produced in the laboratory that is part of the Autographic Theme Extraction System of the Earth Resources Observation Systems program of the Department of Interior (Smith, 1973).

Measurement of the relative area of each characteristic in a gaged basin was made photometrically. Light from a photographic enlarger was focused upon an integrating photometer. The basin mask was introduced between the light source and photometer. The total illumination passing through the transparent area inside a basin boundary on the mask is a measure of the total area of the basin. A photomechanical extraction with transparent areas representing the basin characteristic was then registered with the mask. The total illumination passing through both basin mask and extraction is a measure of the area of the basin characteristic only. The ratio of illumination measurements was expressed as percent of basin covered by the selected basin characteristic. The photometric equipment was calibrated by introducing circular holes of known diameter between the light source and the photometer.

The ratio of illuminations for different hole diameters agreed within 5 percent with the ratio of areas calculated for the corresponding holes. The equipment selection and setup were conceived by W. E. Evans, Stanford Research Institute. Basin characteristics determined by this method are summarized in table 2.

Table 2.--Basin characteristics derived from Landsat-1 imagery.

Station number	Area of basin covered by given characteristic, in percent of total basin area			
	Forest (Uf)	Water (Uw)	Riparian vegetation (Urv)	Combined agr. and urban (Uau)
01483200	43.6	1.1	52.8	55.3
01483500	14.2	0.0	21.0	85.8
01484300	29.6	1.8	6.1	68.6
01484500	23.9	0.0	3.9	76.1
01485000	40.8	0.0	11.9	59.2
01485500	79.4	0.0	5.4	20.6
01486000	67.4	0.0	10.8	32.6
01486500	39.4	0.7	0.8	59.9
01487000	29.3	0.0	3.5	70.7
01487500	59.7	1.0	2.4	39.3
01488500	32.3	0.0	18.0	67.7
01489000	19.5	0.0	3.0	80.5
01490000	55.2	0.9	3.6	43.9
01491000	40.4	0.4	39.1	59.2
01492000	36.2	0.0	3.0	63.8
01492500	31.3	0.0	13.8	68.7
01493000	23.3	0.8	35.6	75.9
01493500	6.0	0.0	6.7	94.0
01494000	20.2	0.0	18.0	79.8
01579000	21.3	0.0	32.2	78.7

CHARACTERISTICS FROM MAPS AND CLIMATOLOGICAL RECORDS

The twelve physiographic and climatic characteristics used in the regression analysis that was part of the evaluation for Maryland and Delaware (Forrest and Walker, 1970) include the following:

A, drainage area, in square miles, contributing to surface runoff, derived from topographic maps and shown in the latest U.S. Geological Survey streamflow report,

S, main channel slope, in feet per mile computed by the 85- to 10-percent method (Benson, 1962),

L, stream length, in miles, measured along the main channel from gage to basin divide,

E, mean basin elevation , in feet above mean sea level, measured from topographic maps by the grid sampling method (20 to 80 points in basin were sampled),

St, area of lakes, ponds, and swamps, in percent of contributing drainage area, measured by the grid sampling method,

F, forest area, in percent of contributing drainage area, measured by the grid sampling methods,

Si, soil index, a relative measure of potential maximum infiltration capacity in inches, estimated from data provided by the U.S. Soil Conservation Service,

P, mean annual precipitation in inches, from The National Weather Service, "Climates of States;" grid sampling method used on isohyetal maps, I24,2, precipitation intensity; maximum 24-hour rainfall, in inches, expected on the average of once every 2 years; estimated from U.S. Weather Bureau Technical Paper 29,

Sn, mean annual snowfall, in inches, from The National Weather Service,"Climates of States,"

T1, mean minimum January temperature, in degrees Fahrenheit, from The National Weather Service, "Climates of States," and

T7, mean maximum July temperature, in degrees Fahrenheit, from The National Weather Service, "Climates of States."

Values for these twelve basin characteristics for each of the 20 gaged basins used in this investigation are available from the U.S. Geological Survey computer file, Streamflow/Basin Characteristics, and are published in previous reports (Forrest and Walker, 1970). The hydrologic significance of each basin characteristic is discussed in previous reports (Benson, 1962; Thomas and Benson, 1970). Identical values were used in this investigation and in the original streamflow data evaluation (Forrest and Walker, 1970).

REGRESSION ANALYSIS

Relations between streamflow characteristics (dependent variables) and drainage basin characteristics (independent variables) were derived by multiple linear regression analysis. This analysis provides an equation of the statistical relation between a streamflow characteristic and selected basin characteristics. It also provides a measure of the accuracy of the relation defined for the sample population (known as the standard error of the estimate of the dependent variable, or simply standard error of estimate). The standard error of estimate is a measure of the spread of the data about the line of relation. It is a statistical parameter such that the value of the streamflow characteristic estimated with the equation for 2 out of 3 gaging stations will, on the average, plot within one standard error of the curve of relation. In addition the value for 19 out of 20 gaging stations will plot within two standard errors of the curve.

If equations are calculated for each of the 40 streamflow characteristics using the 20 gaged basins in the study area and 12 basin characteristics derived from maps only, then these equations are comparable to the equations derived in streamflow program evaluation for Maryland and Delaware (Forrest and Walker, 1970) and constitute a control group for experimentation. The standard error of estimate will provide a measure of the accuracy of the equation. If equations are then calculated for each of the 40 streamflow characteristics from the same 20 gaged basins, the same 12 map-derived basin characteristics, and 4 basin characteristics derived from Landsat imagery in the same analysis, then these equations constitute an experimental group for testing characteristics derived from Landsat. The new standard error of estimate will provide a measure of the accuracy of the new equation. Any reduction in the standard error of estimate between a new or experimental equation and an old or control equation represents an increase in the accuracy of the estimate of that particular streamflow characteristic. The improvement is due solely to including basin characteristics derived from Landsat imagery.

Calculations required for the stepforward regression analysis were performed by digital computer using a group of computer programs known as STATPAC. The program group eliminated any indefinite values from the dependent variables, added a very small constant (0.0001) to those variables which might be expected to go to zero and transformed all dependent variables and four independent variables to their logarithms. Program D0094, Multiple Linear Regression (stepforward), first provided the following statistical parameters: means, standard deviation, and correlation matrices of all variables.

The Stepforward program computed a set of equations by starting with the most effective independent variable then adding the next most effective variable and additional variables until the accuracy of the equation was not significantly improved by any additional variables. After computing each equation it provided several parameters including regression constant and coefficients, multiple correlation coefficient, standard error of estimate of the dependent variable, and percent of the total sums of squares of the dependent variable that are explained. In addition, the program tabulated the observed, the calculated, and the residual of each streamflow characteristic (dependent variable) for each of the 20 basins.

The streamflow characteristics (dependent variable) used in the regression analysis for the streamflow data evaluation for Maryland and Delaware (Forrest and Walker, 1970) were defined at 105 gaging stations and include the full range of flow. In computing these characteristics, frequency curves were not extrapolated beyond twice the length of record. These characteristics include the following:

PT, annual flood peak, in cubic feet per second, of T-year recurrence interval; the recurrence intervals of 2, 5, 10, 25, and 50 years are denoted in this report as P2, P5, P10, P25, and P50 respectively,

QA, mean annual discharge, in cubic feet per second, defined as the mean of the annual means,

SDA, standard deviation of mean annual discharge, in cubic feet per second,

QM, mean discharge for the M-calendar-month, in cubic feet per second; the M refers to the numerical order of the month beginning with January as 1,

SDM, standard deviation of mean discharge for M-calendar-month, in cubic feet per second, the M refers to the numerical order of the month beginning with January as 1,

MD,T, low-flow characteristics are the annual minimum D-day mean flow for T-year recurrence interval, in cubic feet per second,

VD,T, flood volume characteristics are the annual maximum D-day mean flow for T-year recurrence interval, in cubic feet per second, and

D50, discharge, in cubic feet per second, exceeded 50 percent of the time.

Values for the 40 streamflow characteristics for each of the 20 gaged basins used in this investigation are available from the U.S. Geological Survey computer file, Streamflow/Basin Characteristics. Identical values were used in this investigation and in the original streamflow data evaluation (Forrest and Walker, 1970).

REGRESSION EQUATIONS

Tables 3, 4, and 5 summarize the results of the multiple regression analyses. These analyses defined mathematical equations of the form:

$$\log Y = b(1) \log X(1) + b(2) \log X(2) \ldots + b(n) \log X(n) + a + b(n+1) X(n+1) + b(n+2) X(n+2) \ldots + b(m) X(m)$$

or its equivalent form:

$$Y = X(1)^{b(1)} X(2)^{b(2)} \ldots X(n)^{b(n)} 10^{(a + b(n+1) X(n+1) + b(n+2) X(n+2) \ldots + b(m) X(m))}$$

where Y represents a streamflow characteristic, X(1) to X(m) represent basin characteristics (where numbers in parentheses designate numeric order in a series), a represents the regression constant, b(1) to b(m) represent regression coefficients. In the program group, X(1) through X(n) were log transformed while X(n+1) through X(m) were not log transformed prior to calculating equations. Basin characteristics (independent variables) such as drainage area (A) and stream length (L) which have a larger range in values than other independent variables for the 20 basins were log transformed. Those with a small range were not log transformed. No consideration was given in this study to the appropriateness of a regression model using solely logarithmic transforms, or to the reasonableness of the exponents of characteristics included in derived equations. Rather, a simple test to isolate effects of characteristics derived from Landsat by adding them to an existing data matrix and observing the changes in derived equations was performed.

In the tables the first column indicates the streamflow characteristic. In tables 3 and 4, the second column presents the number of basins out of 20 that were used in the regression analysis. The next set of columns show the computed regression constant and the regression coefficients for those independent variables found to be statistically significant at the 95 percent level, and which significantly improved the accuracy of the equation.

Table 3.--Control group equations relating streamflow characteristics to physiographic and climatic characteristics of drainage basins as determined from maps and climatological records.

$$Y = A^{b(1)} S^{b(2)} L^{b(3)} E^{b(4)} 10^{(a + b(5)St + b(6)F + b(7)Si + b(8)P + b(9)I24,2 + b(10)Sn + b(11)T1 + b(12)T7)}$$

Flow characteristic Y	Number of basins used	Regression constant a	Regression coefficient for indicated basin characteristic											
			A	S	L	E	St	F	Si	P	I24,2	Sn	T1	T7
P2	20	3.666	0.806			0.618					-0.963			
P5	20	10.065	0.754							-0.183				
P10	20	10.931	0.727							-0.198				
P25	18	-21.546	0.447					-0.007						0.280
P50	9	-40.155	0.499											0.492
QA	18	-1.581	0.984			0.188					0.387			
SDA	18	-0.415	1.019											
Q10	18	-0.209	0.908											
Q11	18	-2.003	0.961								0.432	0.030		
Q12	18	-1.690	1.022			0.272					0.364			
Q1	18	-1.074	1.031			0.131					0.289			
Q2	18	-0.891	1.014							0.025				
Q3	18	-1.159	1.018							0.032				
Q4	18	-0.763	0.980								0.291			
Q5	18	0.036	0.973											

Table 3.—Control group equations relating streamflow characteristics to physiographic and climatic characteristics of drainage basins as determined from maps and climatological records–Continued.

$$Y = A^{b(1)}S^{b(2)}L^{b(3)}E^{b(4)}10^{(a + b(5)St + b(6)F + b(7)Si + b(8)P + b(9)I24,2 + b(10)Sn + b(11)T1 + b(12)T7)}$$

Flow characteristic Y	Number of basins used	Regression constant a	Regression coefficient for indicated basin characteristic											
			A	S	L	E	St	F	Si	P	I24,2	Sn	T1	T7
Q6	18	-0.141	0.987											
Q7	18	-0.162	0.917											
Q8	18	-0.078	1.028											
Q9	18	-2.866	0.860								0.598	0.043		
SD10	18	-3.457	0.951										0.112	
SD11	18	-0.318	1.069											
SD12	18	-0.308	1.116											
SD1	18	-0.185	1.089											
SD2	18	0.228	1.083									-0.023		
SD3	18	-2.529	1.049							0.089	-0.466			
SD4	18	-0.162	1.040											
SD5	18	-0.261	1.035											
SD6	18	-0.935			2.052									
SD7	18	-1.016	0.973									0.042		
SD8	18	-0.154	1.150											

Table 3.--Control group equations relating streamflow characteristics to physiographic and climatic characteristics of drainage basins as determined from maps and climatological records-Continued.

$$Y = A^{b(1)} S^{b(2)} L^{b(3)} E^{b(4)} 10^{(a + b(5)St + b(6)F + b(7)Si + b(8)P + b(9)I24,2 + b(10)Sn + b(11)T1 + b(12)T7)}$$

Flow charac-teristic Y	Number of basins used	Regress-ion constant a	Regression coefficient for indicated basin characteristic											
			A	S	L	E	St	F	Si	P	I24,2	Sn	T1	T7
SD9	18	-0.342	0.914			0.059						-0.478		
M7,2	17	-1.389			1.953									
M7,10	16	-4.857			4.957									
M7,20	15	-24.151			4.279						5.704			
V3,2	10	0.770	1.080											
V3,25	10	1.001	1.327				-0.068							
V7,2	17	0.626	1.071											
V7,10	17	0.919	1.058											
V7,25	10	0.900	1.131											
D50	11	-0.451			1.691									

Table 4.--Experimental group equations relating streamflow characteristics to physiographic and climatic characteristics of drainage basins as determined from maps, climatological records, and Landsat imagery.

$$Y = A^{b(1)} S^{b(2)} L^{b(3)} E^{b(4)} 10^{(a + b(5)St + b(6)F + b(7)Si + b(8)P + b(9)I24,2 + b(10)Sn + b(11)T1 + b(12)T7 + b(13)Uau + b(14)Uf + b(15)Urv + b(16)Uw)}$$

Flow characteristic Y	Number of basins used	Regression constant a	Regression coefficient for indicated basin characteristic															
			A	S	L	E	St	F	Si	P	I24,2	Sn	T1	T7	Uau	Uf	Urv	Uw
P2	20	4.491	0.717			0.514		-0.008		-0.083						0.009		-0.148
P5	20	5.349	0.964	0.570						-0.092								-0.254
P10	20	-5.331	0.569							-0.126				0.153				-0.246
P25	18	-21.546	0.447					-0.007						0.280				
P50	9	-40.155	0.499											0.492				
QA	18	-1.581	0.984			0.188					0.387							
SDA	18	-0.415	1.019															
Q10	18	-0.209	0.908															
Q11	18	0.049	0.999													-0.004		
Q12	18	-1.690	1.022			0.272					0.364							
Q1	18	-1.074	1.031			0.131					0.289							
Q2	18	-0.891	1.014							0.025								
Q3	18	-1.159	1.018							0.032								
Q4	18	-0.763	0.980								0.291							

Table 4.--Experimental group equations relating streamflow characteristics to physiographic and climatic characteristics of drainage basins as determined from maps, climatological records, and Landsat imagery-Continued.

$$Y = A^{b(1)} S^{b(2)} L^{b(3)} E^{b(4)} 10^{(a + b(5)St + b(6)F + b(7)Si + b(8)P + b(9)I24,2 + b(10)Sn + b(11)T1 + b(12)T7 + b(13)Uau + b(14)Uf + b(15)Urv + b(16)Uw)}$$

Flow characteristic Y	Number of basins used	Regression constant a	Regression coefficient for indicated basin characteristic															
			A	S	L	E	St	F	Si	P	I24,2	Sn	T1	T7	Uau	Uf	Urv	Uw
Q5	18	0.036	0.973															
Q6	18	-0.122	1.048					0.007								-0.011		
Q7	18	-0.076	0.998													-0.009		
Q8	18	-1.393	1.024								0.425					-0.003		
Q9	18	-0.022	0.964													-0.008		0.154
SD10	18	-3.457	0.951										0.112					
SD11	18	-0.318	1.069															
SD12	18	-0.308	1.116															
SD1	18	-0.140	1.074															0.077
SD2	18	-0.268	1.073													0.004		
SD3	18	-2.529	1.049							0.089	-0.466							
SD4	18	-0.162	1.040															
SD5	18	-0.261	1.035															
SD6	18	-0.935			2.052													
SD7	18	-0.819	0.998												0.007			
SD8	18	-0.154	1.150															

Table 4.—Experimental group equations relating streamflow characteristics to physiographic and climatic characteristics of drainage basins as determined from maps, climatological records, and Landsat imagery-Continued.

$$Y = A^{b(1)} S^{b(2)} L^{b(3)} E^{b(4)} 10^{(a + b(5)St + b(6)F + b(7)Si + b(8)P + b(9)I24,2 + b(10)Sn + b(11)T1 + b(12)T7 + b(13)Uau + b(14)Uf + b(15)Urv + b(16)Uw)}$$

Flow characteristic Y	Number of basins used	Regression constant a	Regression coefficient for indicated basin characteristic A	S	L	E	St	F	Si	P	I24,2	Sn	T1	T7	Uau	Uf	Urv	Uw
SD9	18	-0.342	0.914			0.059						-0.478						
M7,2	17	-95.603			2.121											-0.015		
M7,10	16	4.760			5.543			0.005								-0.008		
M7,20	15	-24.151			4.279						5.704							
V3,2	10	0.850	1.085															-0.183
V3,25	10	1.001	1.327				-0.068											
V7,2	17	0.691	1.047															-0.092
V7,10	17	0.919	1.058															
V7,25	10	1.048	1.148					-0.002										-0.097
D50	11	-0.451			1.691													

Table 5.--Comparison of equations as to change in standard error of estimate of the streamflow characteristic resulting from use of Landsat-derived basin characteristics.

S. E.: Standard error, in percent.

△ S. E.: Improvement in standard error, in percent, caused by addition of Landsat-derived characteristics.

Flow characteristic Y	Independent variables included		S.E., in log units		S.E., in percent [a]		△ S.E., in percent
	Control group	Experimental group	Control	Exper.	Control	Exper.	
P2	A; E; I24,2	A; E; F; P; Uf; Uw	0.159	0.092	37.4	21.2	43.3
P5	A; P	A; S; P; Uw	0.197	0.134	46.5	31.5	32.2
P10	A; P	A; P; T7; Uw	0.205	0.149	49.0	34.9	28.8
P25	A; F; T7	A; F; T7	0.154	0.154	36.3	36.3	
P50	A; T7	A; T7	0.229	0.229	51.1	51.1	
QA	A; E; I24,2	A; E; I24,2	0.054	0.054	12.8	12.8	
SDA	A	A	0.088	0.088	20.1	20.1	
Q10	A	A	0.149	0.149	28.8	28.8	
Q11	A; I24,2; Sn	A; Uf	0.098	0.111	23.0	25.8	12.2 [b]
Q12	A; E; I24,2	A; E; I24,2	0.045	0.045	10.2	10.2	
Q1	A; E; I24,2	A; E; I24,2	0.047	0.047	10.8	10.8	
Q2	A; P	A; P	0.054	0.054	12.8	12.8	
Q3	A; P	A; P	0.062	0.062	14.3	14.3	

Table 5.--Comparison of equations as to change in standard error of estimate of the streamflow characteristic resulting from use of Landsat-derived basin characteristics-Continued.

S. E.: Standard error, in percent.

Δ S. E.: Improvement in standard error, in percent, caused by addition of Landsat-derived characteristics.

Flow characteristic	Independent variables included		S.E., in log units		S.E., in percent[a]		Δ S.E., in percent
Y	Control group	Experimental group	Control	Exper.	Control	Exper.	
Q4	A; I24,2	A; I24,2	0.078	0.078	17.1	17.1	
Q5	A	A	0.120	0.120	28.0	28.0	
Q6	A	A; F; Uf	0.153	0.097	36.2	22.8	37.0
Q7	A	A; Uf	0.238	0.168	57.2	39.6	30.8
Q8	A	A; I24,2; Uf	0.134	0.106	31.5	24.6	21.9
Q9	A; I24,2; Sn	A; Uf; Uw	0.136	0.102	31.9	23.8	25.4
SD10	A; T1	A; T1	0.159	0.159	37.2	37.2	
SD11	A	A	0.165	0.165	39.0	39.0	
SD12	A	A	0.139	0.139	32.5	32.5	
SD1	A	A; Uw	0.083	0.073	19.3	17.0	11.9
SD2	A; Sn	A; Uf	0.102	0.101	23.8	23.5	1.3

Table 5.--Comparison of equations as to change in standard error of estimate of the streamflow characteristic resulting from use of Landsat-derived basin characteristics-Continued.

S. E.: Standard error, in percent.

Δ S. E.: Improvement in standard error, percent, caused by addition of Landsat-derived characteristics.

Flow characteristic Y	Independent variables included		S.E., in log units		S.E., in percent[a]		Δ S.E., in percent
	Control group	Experimental group	Control	Exper.	Control	Exper.	
SD3	A; P; I24,2	A; P; I24,2	0.084	0.084	19.4	19.4	
SD4	A	A	0.088	0.088	20.1	20.1	
SD5	A	A	0.106	0.106	24.6	24.6	
SD6	L	L	0.158	0.158	37.0	37.0	
SD7	A; Sn	A; Uau	0.225	0.213	54.0	50.8	5.9
SD8	A	A	0.160	0.160	37.5	37.5	
SD9	A; E; Sn	A; E; Sn	0.158	0.158	39.5	39.5	
M7,2	L	L; Uf	0.482	0.388	136	102	25.0
M7,10	L	L; F; Uf	1.220	0.825	No meaningful equation		
M7,20	L; I24,2	L; I24,2	1.264	1.264	No meaningful equation		
V3,2	A	A; Uw	0.127	0.051	29.5	12.0	59.3

Table 5.--Comparison of equations as to change in standard error of estimate of the streamflow characteristic resulting from use of Landsat-derived basin characteristics-Continued.

S. E.: Standard error, in percent.

△ S. E.: Improvement in standard error, in percent, caused by addition of Landsat-derived characteristics.

Flow characteristic Y	Independent variables included		S.E., in log units		S.E., in percent[a]		△ S.E., in percent
	Control group	Experimental group	Control	Exper.	Control	Exper.	
V3,25	A; St	A; St	0.112	0.112	26.1	26.1	
V7,2	A	A; Uw	0.088	0.073	20.1	17.0	15.4
V7,10	A	A	0.114	0.114	26.4	26.4	
V7,25	A	A; F; Uw	0.094	0.048	21.9	10.8	50.7
D50	L	L	0.181	0.181	42.8	42.8	

[a] Standard error, in approximate equivalent percent, calculated from standard error in logarithmic units (Thomas and Benson, 1970, p. 31).

[b] Standard error increased by 12.2 percent; equation not improved.

Table 5 summarizes the differences between independent variables included in the control and in the experimental equations. The standard error of each equation is given in logarithmic units and also in approximate equivalent percentage. The percentages are actually the arithmetic averages of the plus and minus percentages about the mean calculated from the standard error in log units. Thus, an average standard error of 24.4 percent represents a deviation of 27.3 percent on the plus side and 21.5 percent on the minus side of the mean of the streamflow characteristic (Benson, 1962). The last column shows the change in standard error where one or more variables derived from Landsat were included in the experimental group of equations. The change is expressed as a percentage of the old (control group) standard error in log units.

The value of the change is given for all equations computed except the 7-day, 10-year low flow, and the 7-day, 20-year low flow. The standard errors for these equations were not significantly less than the standard deviations of the values of the dependent variables.

An improvement in the accuracy of the control group equations is considered substantial when there is at least a 10-percent reduction in the standard error of estimate of the control group equations by including basin characteristics derived from Landsat imagery. Table 5 shows that the standard errors of 15 out of 40 equations were changed. Fourteen equations were improved and one equation was not improved. Among the 14 equations improved, 12 were improved by at least 10 percent.

OCCURRENCE OF VARIABLES IN EQUATIONS

Table 5 shows that variables derived from Landsat were included 18 times in the experimental group equations. The most often used variables were forest and water as indicated below:

Streamflow characteristic group	Total Number of equations	Number of times that indicated variable occurred in equations			
		Uau	Uf	Urv	Uw
High	10	0	1	0	6
Average	14	0	5	0	1
Low	3	0	2	0	0
Variability	13	1	1	0	1
All characteristics	40	1	9	0	8

Forest, Uf, was included in 5 out of 14 equations for streamflow characteristics describing average flow. The coefficients (table 4) for forest are negative and imply an inverse relation between forest cover and mean monthly streamflow for summer and early fall months in the study area. Forest was also included in 2 out of 3 equations for low flows. In both cases the coefficients are also negative, implying an inverse relation between forest cover and annual minimum 7-day mean flow.

Water, U_w, was included in 6 out of 10 equations for high flows. In all 6 occurrences, the coefficients for water are negative, implying an inverse relation between water and flood peaks and volumes with less than 25-year recurrence. Water as defined and measured in this investigation is considered to be a measure of surface water storage.

Combined agricultural and urban land use, U_{au}, was included in only 1 out of 13 equations for flow variability. Riparian vegetation, U_{rv}, as defined and measured in this investigation was not included in any of the equations for streamflow characteristics.

Inferred hydrologic signific.nce was used in the initial selection of all basin characteristics. However, the basis for including any characteristic in an equation is primarily statistical. The inter-relations between the basin characteristics along with the inability of the characteristics to completely describe a drainage basin makes tenuous any assertions about the physical effects of the basin characteristics on streamflow. Despite the inability of the relations to describe the fundamental causes of streamflow variations, the basin characteristics frequently included in the equations are numerical measures that are related to the flow variations.

Regression analyses were performed with both measures of forest, F (map-derived) and Uf (Landsat-derived), and surface water storage, St (map-derived) and Uw (Landsat-derived), in the data matrix. This expedient was expected to result in one measure being replaced by the other where significance was obtained. In 3 cases (P2, Q6, and M7,10) however, the inclusion of Uf generated the inclusion of F. In all 3 cases the exponents of F and Uf are of opposite sign. Because the simple correlation coefficient of F versus Uf is 0.82, inclusion of both in the derived equation tends to cancel their effect and makes the computed standard error suspect. A conclusion which is still valid for the 3 cases, however, is that Uf is a more powerful characteristic than F; otherwise, F would have appeared in the control group equations. Where only Uf appears, the standard error improvement criterion is valid as a demonstratable effect of using Landsat-derived data.

Concurrent appearance of St and Uw did not occur. The simple correlation coefficient of St versus Uw is 0.02, and the two determinations were completely independent.

F and St were included in only one control group equation each. In contrast to this, Uf and Uw were included in 9 and 8 experimental group equations, respectively. Accordingly, Uf may be a more reliable measure of forest cover in the study area than F, and Uw may be a more reliable measure of basin storage than St. If so, the frequent inclusion of Uf and Uw in the experimental group equations is consistent with the frequent inclusion of F and St in equations for the Potomac River basin immediately west of the study area (Thomas and Benson, 1970).

In the Potomac River basin study, however, F was included in 11 out of 14 high flow equations and in only 2 out of 15 average flow equations. These results appear inconsistent with results of this investigation unless differences in the average physiography of the two study areas are considered. In the flat, sandy terrain of the Delmarva Peninsula, forest should be more closely related to evapotranspiration and therefore to low flows as well as mean monthly flows for summer and fall months. In the hilly, rocky terrain of the Potomac River basin, forest should be more closely related to steep untillable slopes and therefore to flood flows.

The basin characteristic riparian vegetation, Urv, was not accurately isolated during image analysis, and it contains considerable upland vegetation. This may be the principal reason why riparian vegetation was not included in any of the experimental group equations.

SUMMARY AND CONCLUSIONS

This study tested the usefulness of basin characteristics derived from Landsat imagery for improving equations used to estimate streamflow characteristics. The Delmarva Peninsula of Maryland, Delaware, and Virginia is a study area representative of rural land use in areas of low topographic relief on the humid east coast of the United States. Basin characteristics derived from Landsat imagery, especially forest (Uf) and water (Uw), are representative of characteristics which most readily can be extracted from Landsat imagery by multispectral image analysis systems employing film density-discrimination techniques. Other hydrologically-significant characteristics may be extracted as remote sensing technology improves.

The basin characteristics derived from maps and climatological records and used in the control group of equations computed by the multiple-regression program were the same characteristics used in the original streamflow data program evaluation (Forrest and Walker, 1970). They cover a wide range of characteristics found to be significant in several previous studies and include, contributing drainage area, main channel slope, stream length, mean basin elevation, area of lakes, ponds, and swamps, forest area, soil infiltration index, mean annual precipitation, precipitation intensity, snowfall, and mean minimum January and mean maximum July temperature.

The streamflow characteristics derived from the records of daily discharge of 20 gaged basins in the study area are representative of the full range in flow conditions and include all of those commonly used for design or planning purposes. They include annual flood peaks with recurrence intervals of 2, 5, 10, 25, and 50 years, mean annual discharge, standard deviation of the mean annual discharge, mean monthly discharges, standard deviation of the mean monthly discharges, low-flow characteristics, flood volume characteristics, and the discharge equalled or exceeded 50 percent of the time.

These streamflow characteristics were related to the basin characteristics of the corresponding 20 drainage basins by a technique of multiple regression using a digital computer. A control group of equations was computed using basin characteristics derived from maps and climatological records. An experimental group of equations was computed using basin characteristics derived from Landsat imagery as well as from maps and climatological records. The standard error of estimate of the two groups of equations was compared to see if any reduction in standard error could be considered a substantial improvement upon the original equations.

Based on a reduction in standard error of estimate equal to or greater than 10 percent, the equations for 12 streamflow characteristics were substantially improved by adding to the analyses basin characteristics derived from Landsat imagery. These improvements are summarized in table 6.

Table 6.--Twelve streamflow characteristics and corresponding standard errors reduced by at least 10 percent.

Streamflow characteristic	Standard error, in percent: Control group	Standard error, in percent: Experimental group	Change in standard error, in percent of control group
P2, annual flood peak of 2-year recurrence	37.4	21.2	43.3
P5, annual flood peak of 5-year recurrence	46.5	31.5	32.2
P10, annual flood peak of 10-year recurrence	49.0	34.9	28.8
Q6, mean monthly flow for June	36.2	22.8	37.0
Q7, " " " " July	57.2	39.6	30.8
Q8, " " " " August	31.5	24.6	21.9
Q9, " " " " September	31.9	23.8	25.4
SD1, standard deviation of mean flow for January	19.3	17.0	11.9
M7,2, annual minimum 7-day mean flow for 2-year recurrence	136	102	25.0
V3,2, annual maximum 3-day mean flow for 2-year recurrence	29.5	12.0	59.3
V7,2, annual maximum 7-day mean flow for 2-year recurrence	20.1	17.0	15.4
V7,25, annual maximum 7-day mean flow for 25-year recurrence	21.9	10.8	50.7

Improvements occurred in all flow regimes. The basin characteristics derived from Landsat imagery that were included most frequently in the experimental group equations are forest, Uf, and water, Uw. It is possible that these two characteristics are more reliable measures of forest cover and basin storage than F and St used in the previous streamflow data evaluation for Maryland and Delaware (Forrest and Walker, 1970).

By comparing the reduced standard error of each streamflow characteristic with the accuracy goal for that characteristic, it is possible to judge the degree to which some goals can be achieved by using basin characteristics derive. from Landsat imagery. In those cases where the goals can thus be achieved, the streamflow data collection effort can be redirected to areas of higher priority.

REFERENCES CITED

Anderson, J. R., Hardy, E. E., and Roach, J. T., 1972, A land-use classification system for use with remote-sensor data: U.S. Geol. Survey Circ. 671, [illegible] p.

Benson, M. A., 1962, Factors influencing the occurrence of floods in a humid region of diverse terrain: U.S. Geol. Survey Water-Supply Paper 1580-B, 64 p.

Benson, M. A., and Carter, R. W., 1973, A national study of the streamflow data-collection program: U.S. Geol. Survey Water-Supply Paper 2028, 44 p.

Forrest, W. E., and Walker, P. N., 1970, A proposed streamflow data program for Maryland and Delaware: U.S. Geol. Survey open-file report, 55 p.

Hardison, C. H., 1969, Accuracy of streamflow characteristics, in Geol. Survey Research, 1969, Chap. D: U.S. Geol. Survey Prof. Paper 650-D, p. 210-214.

National Aeronautics and Space Administration, 1972, ERTS data users handbook: ERTS Project Office, NASA Goddard Space Flight Center, Greenbelt, MD, p. 2-1 to 2-9, A-7 to A-14.

Root, R. R., and Miller, L. D., 1971, Identification of urban watershed units using remote multispectral sensing: Colorado State Univ., Environmental Resources Center Completion Report for OWRR Project A-012-COLO, 51 p.

Serebreny, S. M., Evans, W. E., and Wiegman, E. J., 1974, Study of time-lapse processing for dynamic hydrologic conditions: Stanford Research Institute Final Report, Project 2165, 109 p.

Smith, D. G., 1973, Autographic theme extraction system: Proceedings of 7th United Nations Regional Cartographic Conference for Asia and the Far East, Tokyo, Japan, October 15-27, 1973.

Thomas, D. M., and Benson, M. A., 1970, Generalization of streamflow characteristics from drainage basin characteristics: U.S. Geol. Survey Water-Supply Paper 1975, 55 p.

Vokes, H. E., 1957, Geography and geology of Maryland: Maryland Geol. Survey Bull. 19, 243 p.

PROJECT REPORTS

IN PROGRESS

Hollyday, E. F., 1976, Improving estimates of streamflow characteristics using Landsat-1 (ERTS-1; Earth Resources Technology Satellite) imagery: U.S. Geol. Survey Jour. of Research.

PUBLISHED

Hollyday, E. F., 1972, Drainage basin characteristics from ERTS data, 1 July 1972-31 Aug. 1972: U.S. Geol. Survey Admin. Rept. to EROS Program.

_________1973, Drainage basin characteristics from ERTS data, 1 Sept. 1972 - 31 Oct. 1972: U.S. Geol. Survey Admin. Rept. to EROS Program.

_________1973, Preliminary test of ERTS-1 imagery for improving definition of natural streamflow (Abs) in Significant results obtained from ERTS-1--A Symposium: National Aeronautics and Space Adm., New Carrollton, Md., Mar. 5-9, 1973, p. 77.

_________1975, Improving estimates of streamflow characteristics with ERTS-1 imagery (Abs.): Amer. Society Photogrammetry Preprints to Southeastern Meeting Papers, Athens, GA.

Hollyday, E. F., and Cook, J. G., 1975, Tornado tracks in Williams, R. S., Jr., and Carter, W. D., Editors, ERTS-1, A new window on our planet: U.S. Geol. Survey Prof. Paper 929, p. 300-301. (In press)

Hollyday, E. F., and Pluhowski, E. J., 1975, Improving estimates of streamflow characteristics in Williams, R. S., Jr., and Carter, W. D., Editors, ERTS-1, A new window on our planet: U.S. Geol. Survey Prof. Paper 929, p. 136-138. (In press)

Moore, G. K., and Hollyday, E. F., 1975, Discovery and significance of the Beech Grove Lineament of Tennessee in Williams, R. S., Jr., and Carter, W. D., Editors, ERTS-1, A new window on our planet: U.S. Geol. Survey Prof. Paper 929, p. 164-168. (In press)