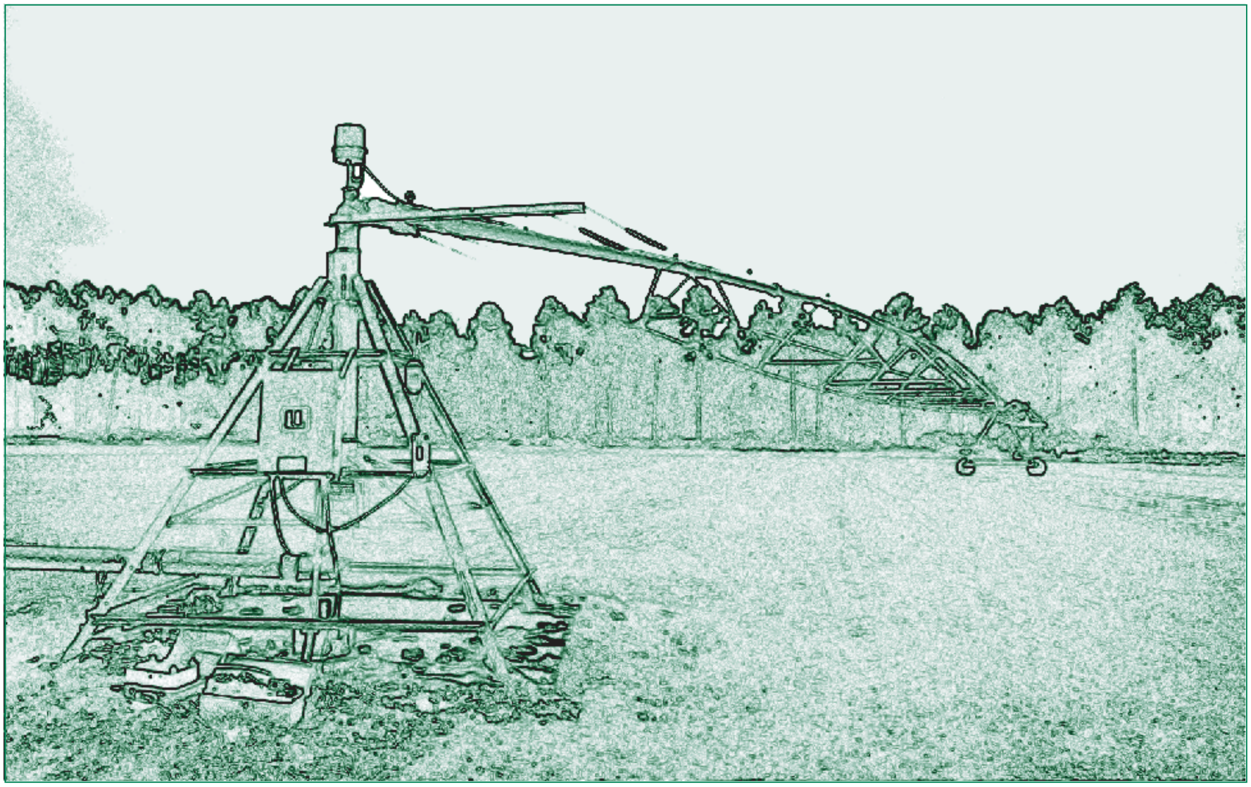
A Field and Statistical Modeling Study to Estimate Irrigation Water Use at Benchmark Farms Study Sites in Southwestern Georgia, 1995–96

Water-Resources Investigations Report 00-4292



Prepared in cooperation with

Georgia Department of Natural Resources
Environmental Protection Division
Georgia Geologic Survey

The University of Georgia
Agricultural Experiment Station

U.S. Department of the Interior
U.S. Geological Survey

A FIELD AND STATISTICAL MODELING STUDY TO ESTIMATE IRRIGATION WATER USE AT BENCHMARK FARMS STUDY SITES IN SOUTHWESTERN GEORGIA, 1995-96

By Julia L. Fanning, Gregory E. Schwarz, and William C. Lewis

---

U.S. GEOLOGICAL SURVEY

Water-Resources Investigations Report 00-4292

Prepared in cooperation with the

GEORGIA DEPARTMENT OF NATURAL RESOURCES
ENVIRONMENTAL PROTECTION DIVISION
WATER RESOURCES MANAGEMENT BRANCH

and

THE UNIVERSITY OF GEORGIA
AGRICULTURAL EXPERIMENT STATION

COVER: *Artist sketch of center-pivot irrigation system.*

U.S. DEPARTMENT OF THE INTERIOR
MARCH 3, 1849

Atlanta, Georgia
2001

U.S. DEPARTMENT OF THE INTERIOR
GALE A. NORTON, Secretary

U.S. GEOLOGICAL SURVEY
CHARLES G. GROAT, Director

For additional information, please write to:

State Representative
U.S. Geological Survey
Peachtree Business Center
3039 Amwiler Road, Suite 130
Atlanta, GA 30360-2824

Copies of this report can be purchased from:

U.S. Geological Survey
Branch of Information Services
Denver Federal Center
Box 25286
Denver, CO 80225-0286

CONTENTS

ILLUSTRATIONS

ILLUSTRATIONS—Continued

TABLES

ACRONYMS

CAES	The University of Georgia, College of Agriculture and Environmental Services
CES	The University of Georgia, Cooperative Extension Service
GaEPD	Georgia Department of Natural Resources, Environmental Protection Division
GWUP	Georgia Water-Use Program
USGS	U.S. Geological Survey
WRMB	Water Resources Management Branch

A FIELD AND STATISTICAL MODELING STUDY TO ESTIMATE IRRIGATION WATER USE AT BENCHMARK FARMS STUDY SITES IN SOUTHWESTERN GEORGIA,1995-96

by Julia L. Fanning, Gregory E. Schwarz, and William C. Lewis

ABSTRACT

A benchmark irrigation monitoring network of farms located in a 32-county area in southwestern Georgia was established in 1995 to improve estimates of irrigation water use. A stratified random sample of 500 permitted irrigators was selected from a data base—maintained by the Georgia Department of Natural Resources, Georgia Environmental Protection Division, Water Resources Management Branch—to obtain 180 voluntary participants in the study area. Site-specific irrigation data were collected at each farm using running-time totalizers and noninvasive flowmeters.

Data were collected and compiled for 50 farms for 1995 and 130 additional farms for the 1996 growing season—a total of 180 farms. Irrigation data collected during the 1996 growing season were compiled for 180 benchmark farms and used to develop a statistical model to estimate irrigation water use in 32 counties in southwestern Georgia. The estimates derived were developed from using a statistical approach know as "bootstrap analysis" that allows for the estimation of precision.

Five model components—whether-to-irrigate, acres irrigated, crop selected, seasonal-irrigation scheduling, and the amount of irrigation applied—compose the irrigation model and were developed to reflect patterns in the data collected at Benchmark Farms Study area sites.

The model estimated that peak irrigation for all counties in the study area occurred during July with significant irrigation also occurring during May, June, and August. Irwin and Tift were the most irrigated and Schley and Houston were the least irrigated counties in the study area. High irrigation intensity primarily was located along the eastern border of the study area; whereas, low irrigation intensity was located in the southwestern quadrant where ground water was the dominant irrigation source. Crop-level estimates showed sizable variations across crops and considerable uncertainty for all crops other than peanuts and pecans. Counties having the most irrigated acres showed higher variations in annual irrigation than counties having the least irrigated acres.

The Benchmark Farms Study model estimates were higher than previous irrigation estimates, with 20 percent of the bias a result of underestimating irrigation acreage in earlier studies. Model estimates showed evidence of an upward bias of about 15 percent with the likely cause being a misrepresented inches-applied model. A better understanding of the causes of bias in the model could be determined with a larger irrigation sample size and increased substantially by automating the reporting of monthly totalizer amounts.

INTRODUCTION

Until about 1970, minimal irrigation occurred in Georgia. During the late 1970's, irrigation of cropland increased rapidly—mainly as a result of increased usage of center-pivot irrigation systems—especially in southwestern Georgia (Pierce and others, 1984). By 1995, an estimated 722 million gallons per day (Mgal/d) of water was being withdrawn to irrigate about 1.1 million acres of cropland (Fanning, 1997). Of the total water used in the State in 1995, about 66 percent (479 Mgal/d) was ground water and about 34 percent (243 Mgal/d) was surface water (streams and ponds). During 1995, about 40 percent of all ground water in Georgia was used for irrigation.

In 1988, a law was enacted by the Georgia Legislature that requires a withdrawal permit for each irrigation water source exceeding 100,000 gal/d on a monthly average. The Georgia Department of Natural Resources, Environmental Protection Division (GaEPD), Water Resources Management Branch (WRMB), is responsible for issuing and monitoring these permits. The latest data (1997) show about 22,000 irrigators in Georgia of which about 19,000 have withdrawal permits.

Even though irrigation water-use permits are issued by the WRMB, irrigation water users are not required to meter or report water withdrawn for irrigation; and thus, irrigation water use is difficult to estimate. Irrigation water use is estimated by the Georgia Water-Use Program (GWUP)—a joint project between the U.S. Geological Survey (USGS) and the Georgia Geologic Survey—by multiplying acres of a crop irrigated by an estimated average application rate for that crop. The acreage in cropland and irrigation application rates are obtained from periodic surveys conducted by The University of Georgia Cooperative Extension Service (CES). The current estimation method, however, does not include types of irrigation systems, location, season, or varying water-application rates. A method based on site-specific data would be more useful to WRMB water managers.

The USGS, in cooperation with the WRMB and The University of Georgia College of Agriculture and Environmental Sciences (CAES), began a pilot study in 1994 to improve techniques for estimating agricultural water use. The pilot study of 50 farms was conducted during 1995; however, the original goal was to monitor 200 farms over a 5-year period. In 1996, additional farms were selected, bringing the total number of farms in the monitoring network to 180—20 farms less than planned at the origin of the study. At each of the 180 farms, irrigation frequency was monitored using vibration time-totalizers, and flow rate was measured using noninvasive flowmeters.

Purpose and Scope

This report presents irrigation data collected during 1995-96 in the study area and describes results of the statistical analysis. This report also describes the method(s) developed for the Benchmark Farms Study to monitor and collect site-specific irrigation water use in southwestern Georgia. A five-component statistical model developed to estimate total irrigation water use in the study area in 1996 is also presented. The five model components are ordered in a recursive structure—whether-to-irrigate, acres irrigated, crop selected, seasonal-irrigation scheduling, and the irrigation applied. Comparisons of model results to sampled data and to other methods of estimation also are included. The study area consists of 32 counties in southwestern Georgia (fig. 1).

The overall objective of the Benchmark Farms Study was to establish an irrigation-monitoring network in 32 counties in southwestern Georgia (fig. 1) to estimate irrigation water use in those counties. The Benchmark Farms Study is the most intensive study of irrigation water use conducted to date (1997) in Georgia. Results will be useful in evaluating and possibly improving current methods for estimating irrigation water use from both ground- and surface-water sources. The 180 farms are collectively "referred to" as "Benchmark Farms," that compose the Benchmark Farms Study network.

Figure 1. Counties included in the Benchmark Farms Study area (shaded), southwestern Georgia.

Previous Investigations

Initial USGS efforts to monitor irrigation water use in Georgia involved a field survey of 62 counties, conducted by the U.S. Soil Conservation Service (1982) (now the Natural Resources Conservation Service). That survey provided a data base of irrigation water users in 1979-80—and the data were used in conjunction with data from the CES for areas not included in the field survey—to estimate statewide and county irrigation water use for 1980 (Pierce and others, 1984). During 1988, legislation in Georgia was enacted requiring that water use from irrigation water sources that were pumped for more than 100,000 gal/d on a monthly average be reported to the WRMB. For several years, farmers filed reports under this system at county CES offices (Barber, 1983). The reporting was incomplete, however, because not all irrigation water use that met the criteria was reported and because many irrigation systems lacked working time/hour meters or flow meters. Because of incomplete irrigation water-use reports, this information was not used to estimate county or statewide irrigation water use. Various USGS reports used the periodic CES irrigation surveys, which showed acreage in cropland and irrigation application rates to estimate county and statewide irrigation water use for 1985, 1987, 1990, and 1995 (Turlington and others, 1987; Trent and others, 1990; Fanning and others, 1992; and Fanning, 1997).

Other studies involved the measurement of pipe flow to determine water use. Luckey and others (1980) investigated the suitability of using a propeller-type gated-pipe meter—a Reflective-Doppler flowmeter—and a transit-time flowmeter to determine flow measurements on large irrigation systems. Duerr and Trommer (1982) measured irrigation water use with a sonic flowmeter, a saddle (in-line) flowmeter, and a vibration time totalizer. Marella and Singleton (1988) described the use of invasive and noninvasive pipe flowmeters to collect water-use data. Arvin (1992) compared four different flowmeters for the monitoring of various kinds of water uses. Holland and Baker (1993) compared and evaluated the reported pumpage data of selected public suppliers against measured pumpage data, using two noninvasive pipe flowmeters—time-of-flight and the Reflective-Doppler.

Acknowledgments

The authors acknowledge the assistance provided by southwestern Georgia farmers who volunteered their time and effort to make the Benchmark Farms Study possible. The cooperation of these farmers, who allowed project personnel access to their farms with monitoring equipment, is greatly appreciated.

METHODS OF INVESTIGATIONS

Establishing a methodology to estimate irrigation water use was a major goal of the Benchmark Farms Study. Farms included in the study were randomly selected from the 32-county study area. Establishing a Benchmark Farms Study site was uncomplicated and required a minimum amount of the farmer's time. Data were collected and input into computer storage by project personnel in the USGS office in Atlanta and The University of Georgia (CES) Office in Tifton, Ga.

Criteria for Selecting a Benchmark Farms Study Site

In 1997, an estimated 22,000 farms in Georgia required an irrigation water-use permit. The WRMB, which maintains a data base of all permitted irrigators, issued about 19,000 permits. Nearly 14,000 of these 19,000 permitted farms are located within the 32-county Benchmark Farms Study area in southwestern Georgia (fig. 1). Agricultural permits are issued for each withdrawal location (well, pond intake, or stream intake), except for systems that pump water from a well to a pond having limited surface-water inflow—and individual permits are issued for these "well-to-pond" systems, although the water is from a mix of ground- and surface-water sources. Some farms have more than one irrigation source; and therefore, require multiple permits.

Participation in the Benchmark Farms Study was voluntary. Initial contacts were by a letter that solicited interest and willingness for farmers to participate in the study. A pamphlet that briefly described the Benchmark Farms Study (Appendix A, fig. A-1) and postage-paid reply cards (Appendix A, figs. A-2, A-3) were furnished to farmers. Since participation was voluntary, the farmer was asked to reply "yes or no" and verify or correct agricultural permit information (including specific questions about the irrigation system and source). Those farmers responding "no, not irrigating" were included in data analyses as zero pumpage. Incentives to participate—such as efficiency inspections of irrigation systems conducted by CAES personnel—were used to attract farmer interest. Responses were crucial, and every effort was made to contact all farmers from the random sample.

A stratified random sample of 500 permitted irrigators was chosen from the study area—using water source and acreage irrigated as the critical elements—to obtain 200 Benchmark Farms Study sites. The stratified random selection was used to eliminate bias in the data. Those 200 sites—or about 1 percent of the 22,000 irrigation sites in Georgia—is considered a statistically valid sample (Appendix B). Initially, the goal was to obtain a sample that was proportional by water source within each county, which was necessary to obtain a proportional distribution of water sources within the study area; however, because of the voluntary nature of the program, that was not always achieved. Only farms using ground- or surface-water sources for irrigation were selected for this study. Any permit indicating, by code, that well-to-pond systems were used for irrigation, was excluded. All permits identified by code where irrigation was used for aquaculture also were excluded.

Although identified well-to-pond systems were excluded from initial random sample procedures, some selected Benchmark Farms Study sites that originally reported ground- or surface-water sources were later determined to be well-to-pond systems—project personnel, however, decided to use the well-to-pond sites in the study. To meet the criteria established for selecting a Benchmark Farms Study site (stratified random sampling), an additional 300 potential Benchmark Farms Study sites were analyzed to find a sufficient number of permit holders who would agree to participate in the study. The number of selected sites in the Benchmark Farms Study, by water source used for irrigation in each county, are shown in table 1. Bleckley, Pulaski, and Schley Counties are in the study area; however, permit holders in these areas did not volunteer to participate in this study.

Any farm selected for the Benchmark Farms Study had to meet certain criteria. Selected Benchmark Farms Study sites were required to have a ground- or surface-water source that could be clearly identified, and the discharge pipe had to be accessible and suitable for installing field equipment. Selected sites also were required to have characteristics similar to other sites in the area (such as crop type, pump, or irrigation system). These criteria, combined with information compiled from the agricultural permits, were used to determine the final 180 benchmark farms.

Benchmark Farms Study site visits by project personnel were arranged by telephone with all farmers who responded “yes” on the reply card; these visits began in April 1995. Farmers were asked to accompany project personnel to the benchmark farm sites. During the visit, the farm owner or operator was interviewed by CAES and/or GWUP personnel to determine the correct agricultural permit number and other information regarding the water source—such as well or stream (Appendix A, fig. A-4). The site also was inspected to confirm the location of the water source and to determine the pump configuration. This was necessary because some agricultural permits did not always accurately describe the water source; for example, some permits issued for a surface-water source were actually well-to-pond systems. When a selected permit was determined to be unsuitable for the study, a replacement permit was assigned from the list of randomly selected sites to continue without data bias.

Table 1. Benchmark Farms Study sites, by county and water source, 1996

County	Water source		
	Ground water (number)	Surface water (number)	Well-to-pond system (number)
Baker	7	1	0
Ben Hill	6	0	0
Bleckley	0	0	0
Brooks	3	1	0
Calhoun	1	7	1
Clay	1	2	0
Colquitt	0	4	0
Cook	5	3	2
Crisp	0	1	0
Decatur	9	0	0
Dodge	1	1	4
Dooly	3	1	1
Dougherty	4	0	0
Early	7	3	0
Grady	7	0	0
Houston	2	0	0
Irwin	4	5	1
Lee	10	2	0
Macon	3	2	0
Miller	11	0	0
Mitchell	10	1	0
Pulaski	0	0	0
Randolph	2	1	0
Schley	0	0	0
Seminole	3	0	0
Sumter	6	3	0
Terrell	3	1	1
Thomas	0	0	1
Tift	5	6	3
Turner	1	1	0
Wilcox	0	1	1
Worth	4	0	0
Total	**118**	**47**	**15**

Establishment of a Benchmark Farms Study Site

Once selected, the process of establishing a Benchmark Farms Study site began by interviewing the farmer (owner or operator) to determine additional water-source and system information. Interviews provided stream names for surface-water sources; and well location, depth, and driller name for ground-water sources. Additional data obtained from farmers included type of pump, capacity of the pump, and irrigation system for surface- and ground-water sources (and any problems associated with irrigation systems). This information was used to cross-check differences between reported and measured flow rates at the irrigation system. Crop and acreage data, and water-source locations also were obtained. Water levels in all wells were measured, if possible. Project personnel used interview information to determine the usefulness of the site to the study and to correct any discrepancies that might occur later in the growing season. Standard USGS field procedures were followed in conducting field operations at each of the benchmark farm sites. Typically, the time spent at a benchmark farm site was about 1 hour. At the conclusion of each visit, farmers were given dated, pre-addressed, and postage-paid reply cards to report monthly time-totalizer and rain-gage readings (see Appendix A).

Benchmark Farms Study sites used center pivot, drip/trickle, traveller/cable tow, and solid-set sprinkler irrigation systems (table 2). The center pivot is a sprinkler line that operates on moderate to high pressures around a pivot point. Center pivot systems were used at 140 (77 percent) of the Benchmark Farms Study sites in 1996; these systems are adapted for high-growing crops such as corn and peanuts. Even though the center pivot system is the most commonly used irrigation system in the study area and statewide, using the random selection process in selecting Benchmark Farm Study sites provided a representative perspective of the irrigation systems used. Drip/trickle irrigation systems place the water at or below land surface and are designed for low pressure. Pecan orchards, nurseries, and vegetables are typical crops irrigated with drip/trickle systems. The traveller/cable tow—or large gun—is a device designed to spray water from a single nozzle for a long distance, and moves across a field by winding a cable or hose onto a reel; traveller/cable tow systems are commonly used with tobacco and sometimes with cotton. By contrast, the solid-set sprinkler system, which is designed to remain in one place during the irrigation season, uses distribution pipes or lines that are buried below ground and risers that extend up from the laterals to a sprinkler head.

Table 2. Benchmark Farms Study sites, by irrigation system type, 1996

Number of sites	Irrigation-system type
140	center pivot
24	drip/trickle
13	traveller/cable tow/large gun
6	solid-set sprinkler
[1]183	

[1] Total number of sites vary because multiple systems occasionally were used at one Benchmark Farms Study site or one system was used at more than one site.

Field Equipment

At each Benchmark Farms Study site, the irrigation flow rate was measured using a noninvasive flowmeter. The frequency of irrigation was measured using time totalizers. Rain gages were installed, if needed.

A noninvasive pipe flowmeter is an instrument that measures the flow of fluid through a pipe without being in direct contact with the fluid (Arvin, 1992). During this study, two Polysonic flowmeters were used to measure the flow rate—the Hydra Model DHT-P flowmeter (fig. 2) and the ISTT-P. Calibration tests were conducted on both flowmeters; results are given in Appendix C. The general operation of these two flowmeters is based on the Reflective-Doppler concept—disturbances in the water (such as sediment or air bubbles)—reflect an ultrasonic signal and both the Hydra Model DHT-P and the ISTT-P flowmeters use dual transducers.

The Hydra Model DHT-P flowmeter (fig. 2) was easy to operate, lightweight, and easily transported from site to site. The best flowmeter readings were obtained within the initial 5 to 10 minutes of fully pressurizing the pipe. Readings at center pivots and cable tows were obtained at pipes near the water source and generally were collected quickly and without problems. Readings on drip-trickle irrigation systems were difficult to obtain, possibly due to the pipe material. Most of the flow-rate measurements were within 50 gallons per minute (gal/min) of the flow rate reported on the agricultural permit, unless system configuration had changed.

Figure 2. Polysonic Hydra Model DHT-P portable flowmeter.

The flow rate (F), in gallons per minute, can be calculated using the formula:

$$F = (ID)^2 \times V \times C$$

where,

- ID is the inside diameter of the pipe, in inches;
- V is the velocity, in feet per second, and;
- C is a dimensionless conversion coefficient equal to 2.45,

A pick-up accelerometer attached by a 4-ft-long cable to an RTT-8 Universal Running-Time Totalizer (fig. 3) was used to determine the frequency of irrigation. The RTT-8 is activated by any mechanical vibration and records a cumulative total time, in hours, the equipment vibrates. The RTT-8 is battery powered in a waterproof aluminum enclosure and is capable of sensing vibrations over a range of frequencies.

The RTT-8 is attached to the irrigation system at a point where substantial vibration occurs during system operation. During this study, vibrations occurred most frequently either on the pipe or elbow joint near the pump, on the irrigation system itself, or on the engine that operated the pump. The vibration sensor was attached with nylon cable ties and the display unit was placed so that it could be easily read by the farmer.

Some Benchmark Farms Study sites required special attention or alternative equipment to properly measure water use. At four sites, a single pump was used to operate multiple systems individually or simultaneously. These benchmark farm sites were equipped with pulse meters rather than RTT-8 units because of the various combinations of withdrawal. Benchmark farm sites with well-to-pond systems required the use of two RTT-8 units—one at the well and one at the pond—to accurately track the amount of water withdrawn and the length of time the systems were in use.

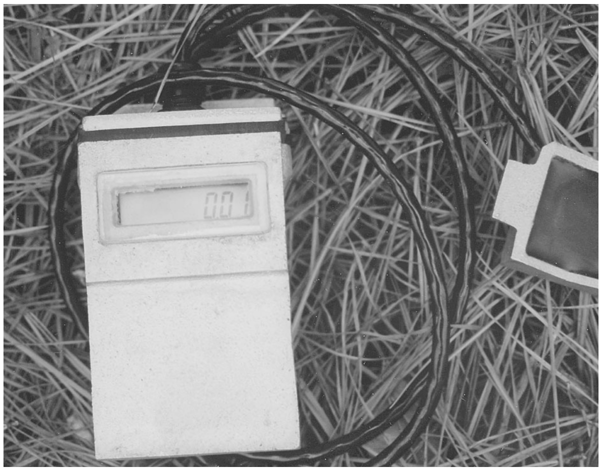


Figure 3. RTT-8 Universal running-time totalizer.

At most Benchmark Farms Study sites, rain gages also were installed at or near the irrigation system so that rainfall and irrigation rates could be measured—several Benchmark Farms Study sites were already equipped with rain gages or they were located near a weather station gage.

The latitude and longitude of the source was determined with a global positioning system (GPS) unit. The GPS unit receives signals from a network of satellites that transmit time and location information. The GPS uses information from three or more satellites to determine an object's position on the earth's surface. When these satellites are locked in, the latitude and longitude positions are displayed (Motorola, Inc., 1993). At Benchmark Farms Study sites with ground-water sources, the GPS unit was held directly over the well. Surface-water source locations were determined at a system intake point.

Equipment Problems

The most common problem during the 2-year study was RTT-8 unit failure (table 3). Data loss occurred more commonly in the first year (1995) of data collection than in the second year (1996). In 1995, 80 (about 38 percent) of the units were replaced at some point during the growing season. In 1996, 52 (about 27 percent) of all units installed were replaced, which was about 11 percent less than in 1995.

Table 3. Problems associated with running-time totalizer (RTT-8) units, 1995-96

Problem	Calendar year (percent)	
	1995	1996
Water leakage	23	7
Sensor/capacitor	10	13
Lightning	1	2
Miscellaneous	4	5
Total	**38**	**27**

Most equipment failure was due to water leaking into the RTT-8 unit, causing the unit to malfunction 23 percent of the time in 1995. Rubber gaskets were used to reduce water leakage into the units, and in 1996, the failure rate for water leakage decreased to about 7 percent.

Malfunctioning of the sensor, amplifier, or capacitor caused RTT-8 units to make continuous readings or to fail to pick up any vibration frequencies on the irrigation system. Attempts were made to repair the units; however, the problems persisted throughout the study period at a rate of as much as 13 percent.

Other problems, such as extreme heat and other weather conditions, were considered typical for this type of study. Lightning strikes accounted for 1 and 2 percent of RTT-8 unit failure in 1995 and 1996, respectively. Damage to the RTT-8 unit casing, cable wires, LCD display window, and batteries accounted for about 5 percent data loss for both years.

IRRIGATION WATER-USE DATA FOR SOUTHWESTERN GEORGIA

Irrigation water use in southwestern Georgia was monitored at 180 sites in 1996. Collection and compilation of the data was conducted by project personnel in the USGS office in Atlanta and The University of Georgia, Cooperative Extension Service Office in Tifton, Ga. A statistical model was developed to estimate irrigation water use in the study area for 1996. The results of the model were compared to the Benchmark Farms Study site-specific data to test the model accuracy.

Data Collection and Compilation

Data collection in the Benchmark Farms Study area also was part of the voluntary effort. Farmers were instructed to read the RTT-8 unit after the last irrigation application for the month and to record all the numbers shown in the unit's display window. Some farmers had double cropping (irrigated more than one crop) during the growing season; therefore, crop and acreage by crop also were collected for each month. Farmers also were instructed to record and date rain-gage readings and indicate whether the recorded readings were for rainfall or irrigation.

Monthly reminders were sent to all farmers participating in the study to encourage the farmers to record the RTT-8 data. Generally, farmers mailed in reply cards to the USGS office in Atlanta in a timely manner and called or made note of any problems with the RTT-8 units.

Contacting the farmers was sometimes difficult because farmers work long hours. Every effort was made by project personnel to collect monthly readings from the farmers; however, because of many obstacles, irrigation and rainfall data were not always obtained for every site.

As reply cards were received in the USGS office, project personnel verified that the information was recorded properly, then the data were entered into a data base. The hours-per-month that the system operated was calculated from the RTT-8 readings. Water withdrawals (in million gallons) were calculated by multiplying the measured flow rate (gallons per minute) by the calculated hours of use. Using the reported acreage for each crop, the inches of water applied to each crop were determined for each Benchmark Farms Study site.

Quality-assurance checks were made on the data each month. When the calculated application rates (inches) seemed to be too large, attempts were made to contact the farmers to discuss possible problems. Farmers usually had an idea of the amount of water that had been applied to the crops in the previous month, and would confirm or reject the calculated application rates. Anomalous data that could not be explained were assumed to be erroneous because of equipment failure, and therefore, were not used in data analyses. At the end of each growing season, a final visit was made to most farms for final readings; at some sites in 1995, these were the only readings of the season. By the end of the 1996 growing season, all farms had reported at least 3 months of data.

At the end of the 1996 growing season, a comprehensive data base of all the benchmark farm sites was compiled for a seasonal data analysis. If multiple time-totalizers were required at a site, the information was compiled to produce a single set of data for each permit number. The agricultural permit data base for permits issued from 1988-94, was used to make various comparisons between field data collected at the Benchmark Farms Study sites and the information shown in the agricultural permit file, which was the same permit file used for site selection. Most of the agricultural permits were issued in 1991 (fig. 4).

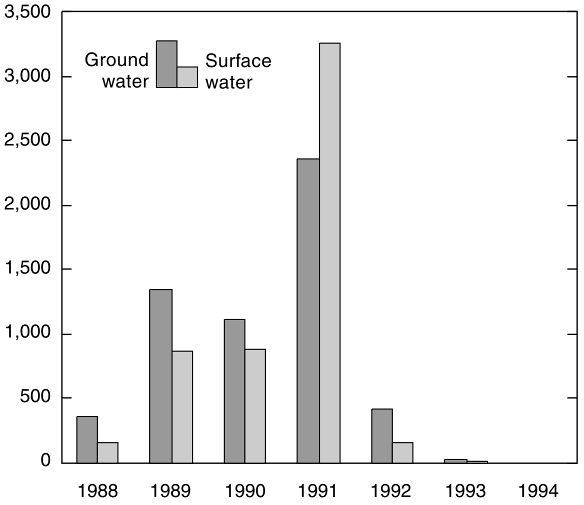


Figure 4. Number of agricultural permits issued in Georgia, by source, 1988–94.

The number of acres irrigated in 1996 at the Benchmark Farms Study sites were compared to the agricultural permit file of acres irrigated. As shown in figure 5, the permitted acres were generally greater than the actual number of acres irrigated in 1996. Project personnel found that farmers would report all possible acres to be irrigated or the total acres owned on the permit application form, rather than the actual number of acres irrigated during the permit issuance year.

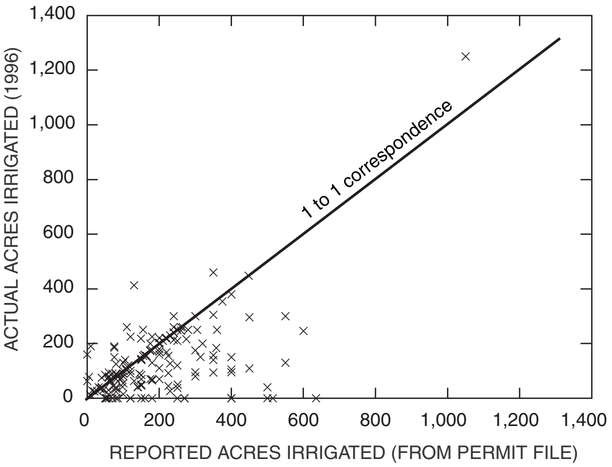


Figure 5. Reported and actual acres irrigated at Benchmark Farms Study sites, southwestern Georgia, 1996.

A variety of crops were irrigated in the Benchmark Farms Study area; however, cotton, peanuts, and corn were the most common (table 4). Although most farmers irrigated only one crop during the growing season, some farmers double and triple cropped acres at some benchmark farm sites during the 1996 growing season. Crop type was requested on the agricultural permit, allowing the farmer to list a variety of crops that might be irrigated during the life of the permit, but no comparisons were made between crop type in 1995 and 1996 and the crop information listed on the permit application.

Irrigation application rates varied by crop type and location in the study area. The 32 counties participating in the Benchmark Farms Study were included in three irrigation regions of the CES—southwest, west-central, and south-central. The calculated inches of water applied by crop were compared between the CES irrigation regions (table 4).

Well-pump capacity, in gallons per minute, for each well on a farm was furnished to the WRMB when a farmer applied for an agricultural permit. Discrepancies were found at a number of sites between the pump capacity on the permit and the actual flow rates measured during the Benchmark Farms Study (fig. 6).

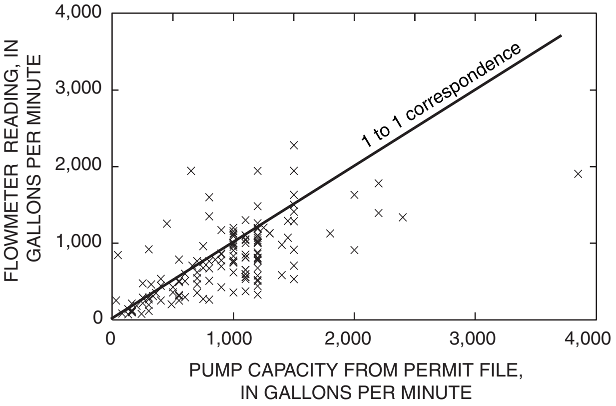


Figure 6. Reported pump capacity and flowmeter readings at Benchmark Farms Study sites, southwestern Georgia, 1996.

Table 4. Number of Benchmark Farms Study sites, by crop in 1996, and inches of water applied, by irrigation region, 1995
[number of sites in table exceed total number of sites because of multiple crops irrigated by some systems, either sequentially or simultaneously]

Crops irrigated	Number of sites	Irrigation, by region, 1995		
		Southwest[1/]	West-central[1/]	South-central[1/]
		(inches)	(inches)	(inches)
Cotton	74	7.2	6.31	6.19
Peanuts	62	6.55	5.61	7.56
Corn	46	8.1	5.76	5.97
Fruits and vegetables	22	8.35	6.16	9.94
Pecans	18	7.97	10.07	6.67
Turf grass	13	8.9	13.54	12.65
Small grains	6	3.46	3.33	0
Nursery	4	31.79	9.84	10.99
Soybeans	4	4.23	4.34	3.08
Tobacco	2	4.89	4.83	5.12
Specialty crops	1	6.63	5.82	5.53

[1/] Data from Georgia Cooperative Extension Service, written commun., 1995.

Modeling Irrigation Water-Use Data

Using data collected at Benchmark Farms Study sites, a statistical model was developed to estimate irrigation water use. Irrigation data collected during the 1996 growing season at the 180 Benchmark Farms Study sites (fig. 7), however, exhibited a number of patterns that complicated model development. Some permits listed multiple crops irrigated at different times, and some permits indicated that multiple pumps served separate irrigated plots, thus requiring multiple time totalizer records. Totalizers periodically malfunctioned resulting in 1 or 2 months of unrecorded (lost) data. Farmers sometimes failed to report a totalizer amount for a given month resulting in subsequent totalizer reports that exceeded a 1-month period (table 4). Months in which irrigation began and ended varied for agricultural permits. Some Benchmark Farms Study sites that had permits, never irrigated during the year. Once irrigation began at a Benchmark Farms Study site for the year, irrigation tended to continue at varying intensities until termination at the end of the growing season (that is, it was uncommon to observe repeated start-stop irrigation cycles over the growing season). The type of crop grown seemed to be important in determining when irrigation began and ended. For a given permit, the month-to-month irrigation intensity was correlated and the reported irrigated acreage listed on the permit application generally was not consistent with the number of actual acres irrigated.

The first modeling decision concerned the appropriate unit of analysis. Benchmark Farms Study data contained information for individual pumps servicing specific crops. Conversely, the information available to extrapolate Benchmark Farms Study results to the entire agricultural population existed only for individual permits. Since an individual agricultural permit could include multiple pumps, either multiple pump records had to be aggregated to the permit level or some method was needed to estimate the typical number of pumps for each permit. Aggregating pump information to the agricultural permit level was further complicated where pump records for multiple-pump permits were incomplete due to totalizer failure (table 3) or incomplete reporting by farmers. Additionally, multiple-pump permits were likely tied to multiple-crop permits. If

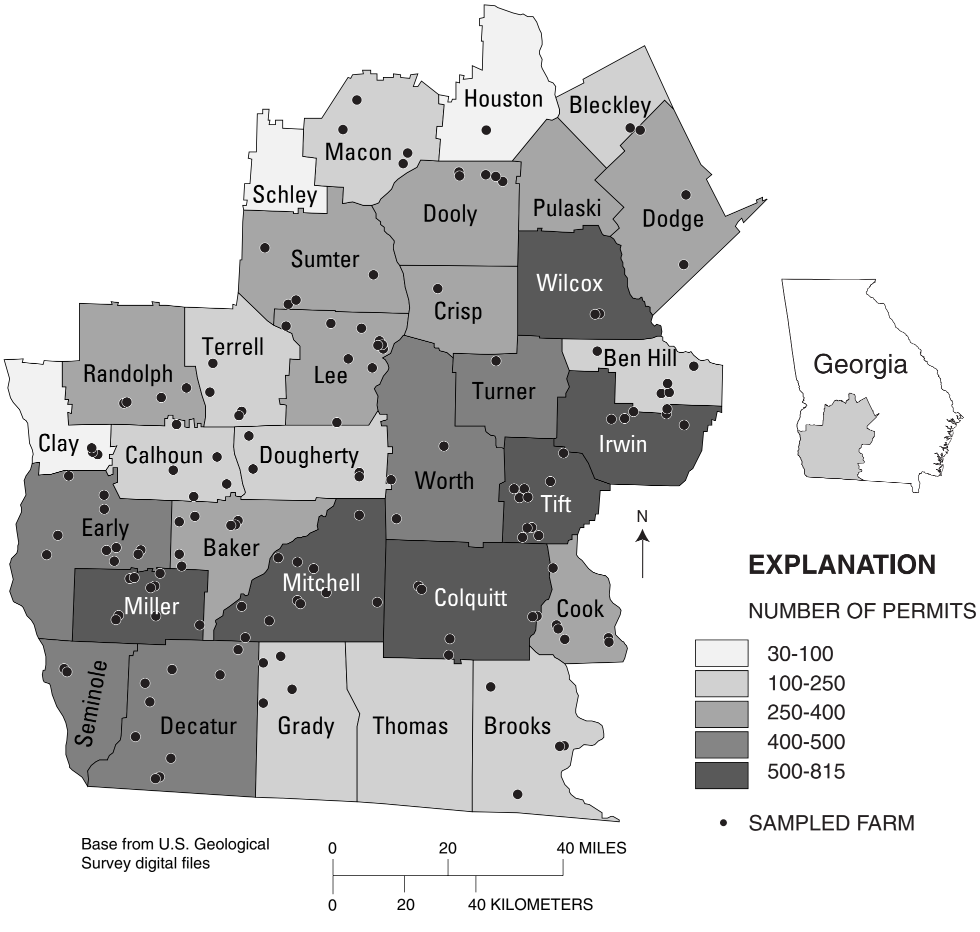


Figure 7. Number of permits for each county and locations of sites included in the Benchmark Farms Study area, southwestern Georgia.

crop type was important in determining irrigation practices, then the aggregated approach would require some assessment of the "average crop" for a permit. Conversely, using the individual pump as the unit of analysis required the development of a model to estimate the number of pumps associated with a given unsampled permit.

After carefully reviewing the sampled data, the permit level was selected as the unit of analysis. Most survey responses indicated a single-pump or single-crop arrangement for which aggregation was not an issue. For the few Benchmark Farms Study sites having permits requiring aggregation, missing data were estimated to make the aggregate record consistent. If most of the pumps had missing values for a given month, the permit was recorded as having a missing record. If most pumps did not have missing values, then the pump records were estimated assuming a straight-line interpolation between non-missing values. The assignment of crop type for each permit was based on the largest acreage of a single crop grown under the permit over the entire growing season (since multiple crops were sometimes grown for a single permit). For a given permit, the earliest month of irrigation for any pump was recorded as the irrigation start date and the last month of irrigation was recorded as the irrigation end date.

In order to accurately reflect other patterns in the data, it was necessary to break the irrigation model into five model components. Four of the model components address different aspects of the irrigation decision including whether or not to irrigate during the growing season, how much acreage to irrigate, when to irrigate during the growing season, and how intensively to irrigate. The fifth model component is the crop-selection model, which addresses a "hole" in farm-level information that must be filled in order to implement the other irrigation decision model components.

The five model components (fig. 8) are ordered in a recursive structure; that is, results of model components ordered first are passed to lower-ordered model components without the feedback of results to higher-ordered model components. The recursive model structure simplifies the process of calibration and prediction; but potentially imposes artificial constraints on the analysis. For example, in the structure adopted for this analysis, the decision of when to irrigate seemed higher in the ordering scheme than the decision of how intensively to irrigate. The implicit assumption of this structure was that the decision of when to irrigate was independent of factors affecting how much irrigation was applied—an assumption that probably was only approximately true.

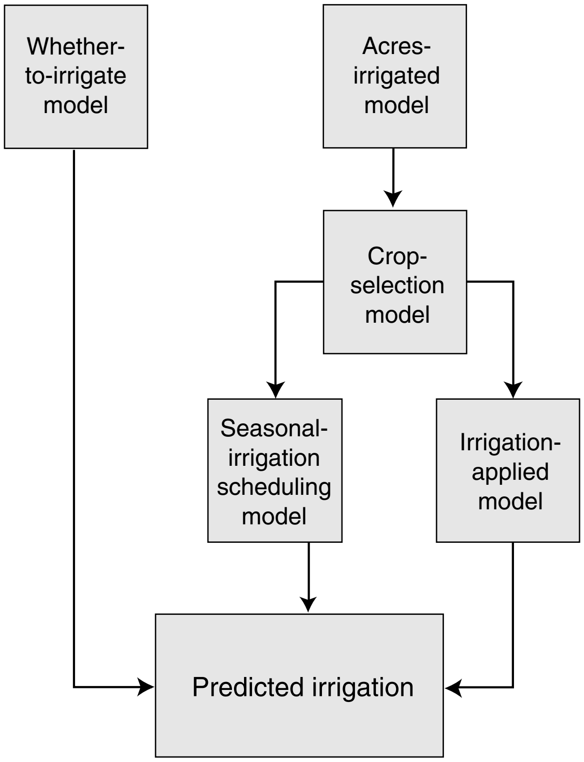


Figure 8. Benchmark Farms Study modeling components. The study data, agricultural permit application data, and rainfall data are input into each model. Results of the higher-ordered model components are passed to lower-ordered model components.

The first component in the model structure is the "whether-to-irrigate" model. This model component determined the probability that the permit holder will irrigate sometime during the growing season. The probability of irrigating is related to a number of readily available predictor variables, such as the irrigated acreage given by the permit application and the location of the farm in the study area. Processing of the other model components proceeded for a given permit only if the model component predicted the active irrigation at that permit location.

The second model component determines the total acres irrigated. Although acres irrigated is reported as part of the permit application, the information gathered from the Benchmark Farms Study sites showed that this information, in most cases, has changed since the permit was issued. Since acres irrigated is an important predictor of total water applied, and since the only study-wide information on acres irrigated came from the permit application, it was necessary to develop a model to translate the acreage reported on the permit to observed acres irrigated. The acres-irrigated model relates the observed acres irrigated to the number of acres reported on the permit and to other variables available for the population of permits in the study area. The predicted number of acres irrigated served as an important predictor for the crop-selection and irrigation-applied components of the model.

The third model component designates the major crop selected under the agricultural permit. Crop type is an important predictor of when farmers irrigate during the growing season and how much irrigation is applied. Since crop type is not specified in the permit application, it was necessary to predict crop type by using information such as acres irrigated and location. Explicit modeling of crop type was done only for crop types that accounted for most of the crops planted by Benchmark Farms Study participants.

Much of the month-to-month variation in withdrawals for a given permit is due to the decision of when to irrigate during the growing season. The Benchmark Farms Study did not collect information on daily irrigation scheduling; however, it was possible to determine which months of the year irrigation took place. A duration analysis was performed that related the probability that a farmer starts or stops irrigation in a given month, given that irrigation occurs at some time during the year, to various predictor variables including crop planted, precipitation, and permit location. Because only month-to-month variations in the scheduling decision were considered, this fourth model component is called the *seasonal* irrigation-scheduling model.

The fifth model component determines the amount of irrigation applied for each permit that is predicted to irrigate in a given month. This determination was based on a number of predictor variables including the crop grown, the number of acres irrigated, precipitation during a month, and site location within the study area. Because inches of water applied were correlated across months, the model component assumes that the error term consists of a permit-specific component that was constant across months but random across permits, and a purely random component that was random across both months and permits.

The irrigation modeling approach was conducted in two phases. The first phase was the calibration of the model where parameter estimates were computed using the data generated by the Benchmark Farms Study sites. The second phase was prediction, which used the parameter and error estimates from the calibration phase to predict each of the components of the irrigation decision for each permit in the study area. Permit-specific results from the prediction model were subsequently aggregated to produce county-level results.

An explicit concern of the model was the accuracy of the estimates. An important advantage of using a statistical approach is the ability to derive an estimate of the precision of any prediction made from the model. Unfortunately, because of the complexity of the model it was not possible to derive simple formulas that described this precision. An alternative method called “bootstrap analysis”, which allows for the estimation of precision in any model, regardless of complexity, was used (Appendix D).

Model Results

Bootstrap estimates of the total rates of irrigation, by month, and for all months during the March to October 1996 growing season for each of the 29 counties participating in the 32-county study area, are given in table 5. Irrigation is expressed in million gallons per day (Mgal/d). Thus, the estimate for March corresponds to the millions of gallons irrigated in March divided by 31 days; whereas, the estimate for the entire growing season March through October represents total million gallons applied divided by 245 days. Because little irrigation is assumed to occur in the months of January, February, November, and December, to obtain million gallons per day for the entire year, the million gallons per day for March through October should be multiplied by the factor 0.67.

Table 5 shows that peak irrigation for all counties seems to occur in July, with substantial irrigation also occurring in the months of May, June, and August. Irwin and Tift Counties have the most irrigation and Schley, Clay, and Houston have the least. A map of the counties in the study area and the total rate of irrigation throughout the growing season is shown in figure 9. A region of high volume irrigation extends diagonally from southwest to northeast, with the highest levels of irrigation located in the east-central part of the Benchmark Farms Study area (fig. 9). The 90-percent confidence intervals for the growing-season estimates for the entire region, expressed in terms of inches per month, and for each of the 32 counties in the study area, expressed in million gallons per day are shown in figures 10 and 11, respectively. The confidence intervals shown in figure 10 cover the average estimates computed from the Benchmark Farms Study sites for most months, but are distinctly above the Benchmark Farms Study sites averages for the peak irrigation months of June through August. The 90-percent confidence intervals shown in figure 11 demonstrate that counties having low rates of irrigation have relatively small confidence intervals—the converse is true for counties having high rates of irrigation. The imprecision of an annual county estimate, defined as the range of the 90-percent confidence interval divided by two times the estimate, generally is 25 to 30 percent. Average imprecision of the monthly estimates is denoted as “model” estimates in figure 12. These imprecisions generally are in the range of 60 to 80 percent. Thus, the growing-season estimates are more than twice as precise as the monthly estimates. This result partly follows from the “law of large numbers” applied to the growing-season estimate, which is an aggregation of individual monthly estimates.

Table 5. Model-estimated rate of irrigation, by county and by month; and for all months in the growing season (March through October) in the Benchmark Farms Study area, southwestern Georgia, 1996
[Mgal/d, million gallons per day over the identified period]

County	Model-estimated irrigation								
	March (Mgal/d)	April (Mgal/d)	May (Mgal/d)	June (Mgal/d)	July (Mgal/d)	August (Mgal/d)	September (Mgal/d)	October (Mgal/d)	March through October (Mgal/d)
Baker	3.1	19.9	48.5	88.4	119.1	63.5	21.1	3.4	45.9
Ben Hill	2.3	14.2	38.2	53.9	71.2	41.8	10.3	3.0	29.4
Bleckley	2.0	10.9	33	43.9	57.8	33.7	11.5	2.2	24.4
Brooks	2.6	11.7	22.7	28.5	37.3	20	5.8	2.1	16.3
Calhoun	1.2	7.8	43.9	78.2	115.9	68.9	24.8	3.4	43.1
Clay	0.3	2.6	10.7	18.3	32.3	18.3	4.6	0.6	11.0
Colquitt	6.0	37.6	93.6	140.4	188.3	103.2	24	7.2	75.1
Cook	5.1	23.6	42.2	54.2	69.1	33.7	11.2	4.3	30.4
Crisp	1.5	10.8	40	73.4	130.4	74.9	26.3	4.6	45.3
Decatur	4.2	28.3	80.2	112	145.5	84.8	24.3	4.6	60.5
Dodge	2.9	19.3	57.7	82.1	108.9	64.5	21.3	5.3	45.3
Dooly	1.4	8.9	54	88	131.3	74.3	29.1	4.8	49.1
Dougherty	0.7	4.2	21.9	38.5	59.1	34.4	12.5	2.2	21.7
Early	3.3	19.8	64.2	88.1	115.9	63.2	21.6	3.5	47.5
Grady	1.5	10.9	26.1	39.3	51.6	28.9	9.9	3.7	21.5
Houston	0.3	1.8	13	20.4	31	16.8	5.6	0.8	11.2
Irwin	7.8	48.9	119.2	178.5	247	142.8	49.4	11.5	100.7
Lee	1.5	11.9	58.3	104.9	161.6	96.3	39	5.5	60
Macon	0.7	4.5	29.4	48	70.3	43.2	13.2	1.9	26.5
Miller	4.8	26.3	85.5	119.2	157.1	86.9	30.1	5.1	64.4
Mitchell	5.3	34.8	77.4	121.8	177	96.9	30.6	5.0	68.7
Pulaski	3.4	18	61.4	84.8	122.2	59.8	17.2	3.9	46.4
Randolph	1.1	9.3	42.6	73.7	116.8	72.5	19.4	2.5	42.3
Schley	0.2	1.0	5.6	8.8	14	8.6	2.4	0.4	5.1
Seminole	3.5	25.2	69.9	95.5	130.3	74.3	20.3	4.3	53
Sumter	1.6	13.8	69.9	116.4	166.6	93	30	3.9	62
Terrell	0.9	8.7	30.4	61.9	97.5	58.1	19.9	2.6	35.1
Thomas	1.6	9.6	25.6	38.1	50.8	28	7.9	2.6	20.6
Tift	7.4	42.7	102.1	152.5	208.4	122.2	44.7	11.1	86.5
Turner	2.5	22.7	80.2	124.6	188.2	108.6	42.1	8.4	72.3
Wilcox	5.8	37.2	97.6	128.2	172.7	95.5	36.2	8.8	72.8
Worth	2.2	13	61.5	108.1	160	95.6	35.9	6.7	60.5
Study area	**88.6**	**559.6**	**1,706.3**	**2,612.2**	**3,705.2**	**2,106.8**	**701.9**	**139.6**	**1,454.5**

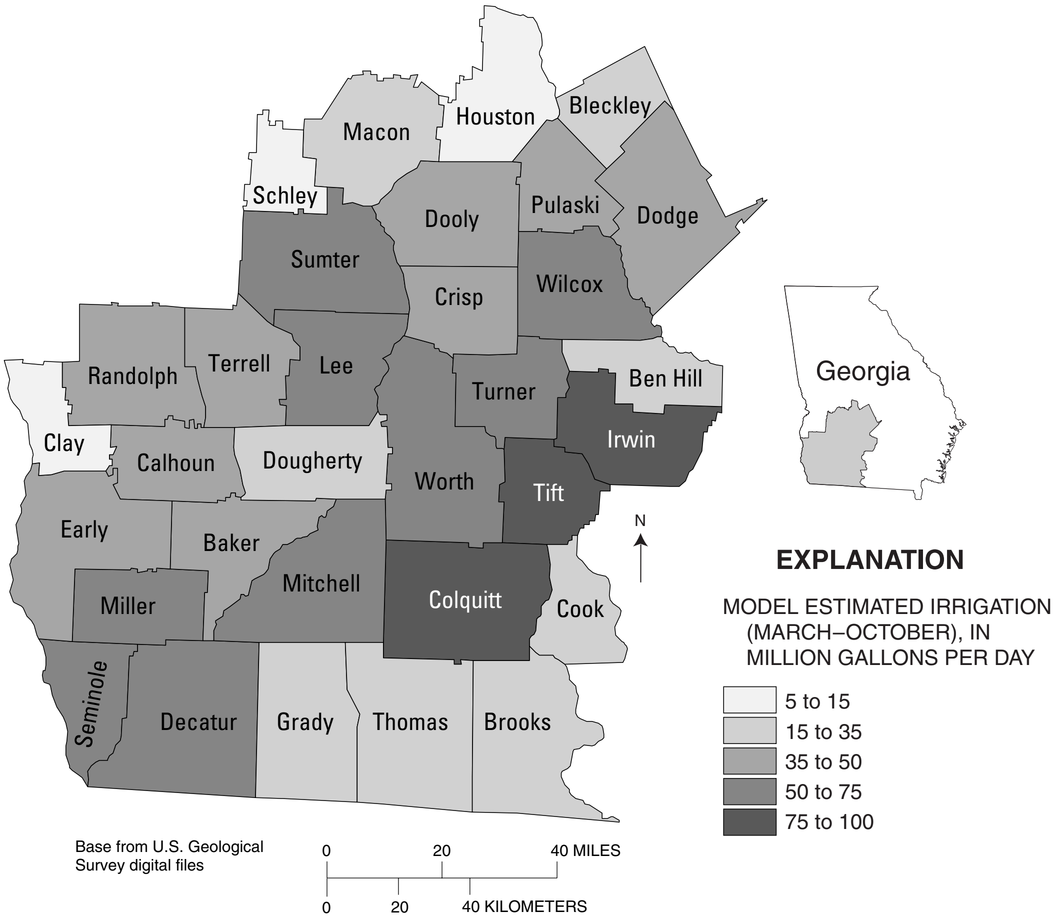


Figure 9. Model-estimated rate of irrigation during the growing season (March through October) in the Benchmark Farms Study area, southwestern Georgia.

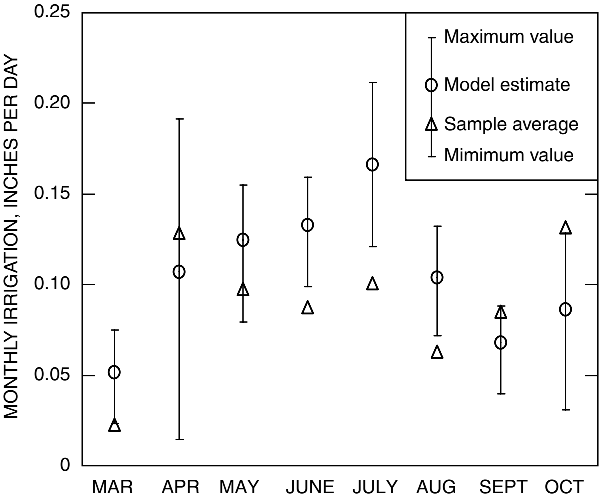


Figure 10. Model-estimated monthly irrigation rates with 90-percent confidence interval, and average monthly irrigation for the permits used in the Benchmark Farms Study area, southwestern Georgia.

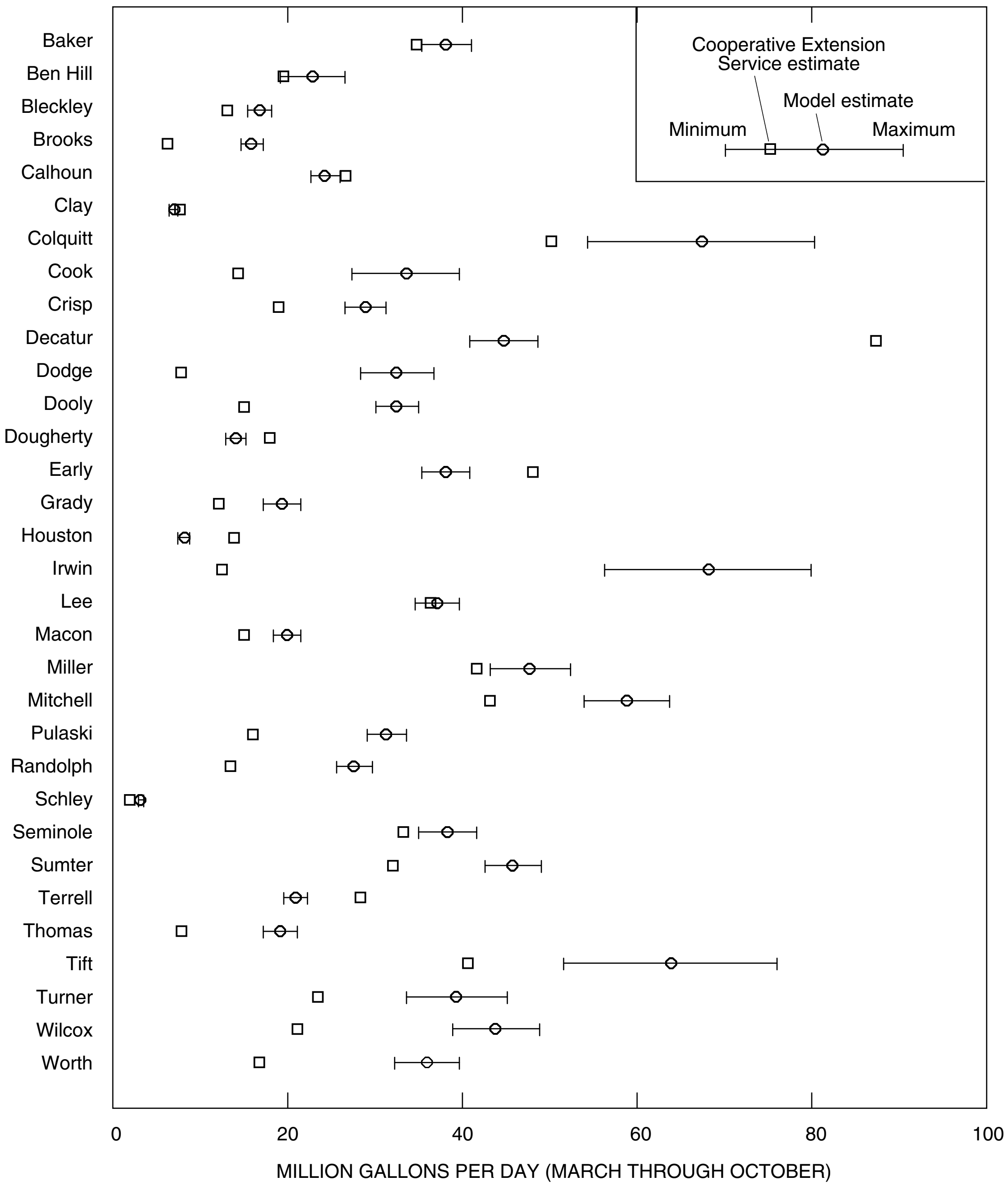


Figure 11. Sample-estimated rate of irrigation during the growing season (March through October) with 90-percent confidence interval and Cooperative Extension Service estimates in the Benchmark Farms Study area, southwestern Georgia.

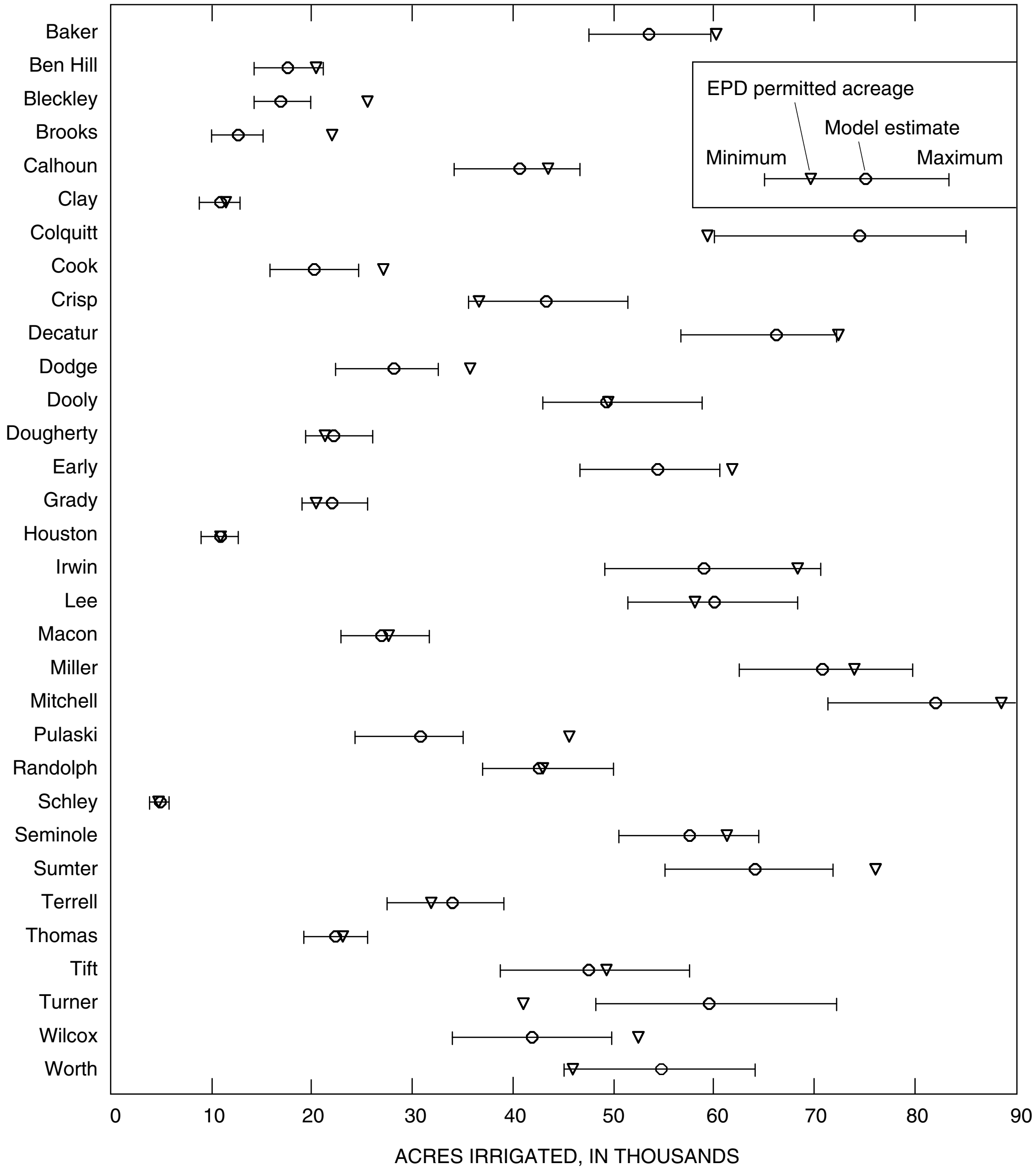


Figure 12. Model-estimated acres irrigated with 90-percent confidence interval and Environmental Protection Division (EPD) permitted acres irrigated in the Benchmark Farms Study area, southwestern Georgia.

Estimates of irrigation derived from ground water are listed in table 6 and shown in figure 12. Estimates of ground-water irrigation in the Benchmark Farms Study area are about half the amount of total irrigation for the entire study area. Counties having the most ground-water irrigation are Miller, Mitchell, Decatur, and Seminole located in the southwestern part of the study area (fig. 13).

Estimates of irrigation applied to land that has irrigated for at least 1 month during the growing season are given, by county and crop, in table 7. The highest irrigation occurred in the counties of Tift, Wilcox, Irwin, Ben Hill, and Dodge. Lowest irrigation occurred in the counties of Mitchell, Baker, Early, Decatur, Miller, Seminole, Sumter, and Thomas. As shown in figure 14, higher rates of irrigation principally occurred along the eastern border of the study area; whereas, lower rates of irrigation occurred in the southwestern quadrant of the study area where ground water is the dominant irrigation source. Confidence intervals for crop-level estimates of inches applied across the entire study area are given in figure 15; and for all crops according to county in figure 16. The crop-level estimates display sizable variation across crops and considerable uncertainty for all crops other than peanuts and pecans. Crops classified as "other," which include truck crops and nurseries, have the largest application rates, followed by turf, pecans, corn, cotton, and peanuts. From the 90-percent confidence intervals, it is possible to conclude that inches applied for "other" crops is statistically greater than application rates for all other crops, except turf and possibly corn; whereas, peanuts receive a statistically smaller rate than all other crops, except corn and cotton. Also depicted with the model estimates in figure 15 are the averages of inches applied for sampled permits. In computing these averages, only those permits that had no missing values for any month over the period were included. The sample average of inches applied for all crops is slightly below the 90-percent confidence interval determined for the model estimate, however, the variations in sample averages and model estimates across crops generally are closely related.

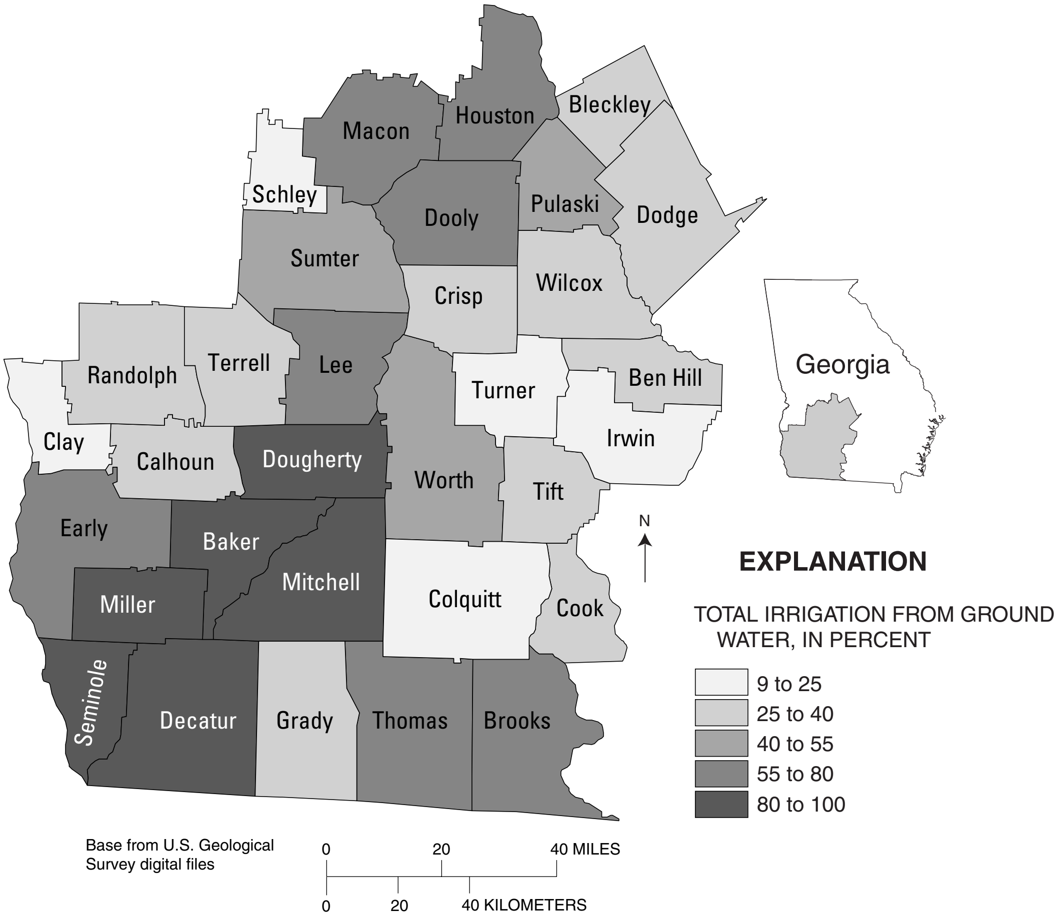


Figure 13. Percent of total irrigation from ground water in the Benchmark Farms Study area, southwestern Georgia.

Table 6. Model-estimated ground-water rate of irrigation, by county and by month; and estimated ground-water irrigation, by months in the growing season (March through October) for the Benchmark Farms study area, southwestern Georgia, 1996

[Mgal/d, million gallons per day over the identified period]

County	Model estimated ground-water irrigation								
	March (Mgal/d)	April (Mgal/d)	May (Mgal/d)	June (Mgal/d)	July (Mgal/d)	August (Mgal/d)	September (Mgal/d)	October (Mgal/d)	March through October (Mgal/d)
Baker	2.6	16.7	40.5	73.1	99.6	52.9	17.5	2.9	**38.2**
Ben Hill	0.6	3.8	9.9	14.7	20	11.2	2.6	0.7	**7.9**
Bleckley	0.8	4.3	12.7	17.6	22.6	13.3	4.5	0.8	**9.6**
Brooks	1.7	7.7	15.1	19.1	25.1	13.2	3.8	1.4	**10.9**
Calhoun	0.5	3.1	15.9	27.3	39.5	23.8	8.1	1.3	**15**
Clay	0.1	0.6	2.4	3.9	7.6	4.4	1.1	0.2	**2.5**
Colquitt	1.1	6.4	16.9	25.2	33.7	18.6	4.0	1.1	**13.4**
Cook	1.8	8.1	13.8	17.8	22.8	11.2	3.8	1.4	**10.1**
Crisp	0.5	3.6	13.8	25.2	46.1	26.4	9.2	1.6	**15.8**
Decatur	3.7	24.4	69.9	97.2	126.5	74	21.2	4.0	**52.7**
Dodge	0.9	6.2	19.5	28.2	36.3	21.9	7.0	1.7	**15.2**
Dooly	0.9	6.0	34.6	56.9	84.8	47.1	18.7	3.1	**31.6**
Dougherty	0.6	3.9	20.6	36.4	55.8	32.2	11.8	2.0	**20.5**
Early	2.4	14.3	46.9	64.7	84.5	45.9	15.6	2.6	**34.7**
Grady	0.5	3.8	9.6	13.7	17.6	9.7	4.5	1.8	**7.7**
Houston	0.3	1.4	9.8	15.8	23	12.7	4.2	0.6	**8.5**
Irwin	1.2	7.8	19.6	30.5	42.4	24.1	7.8	1.7	**16.9**
Lee	1.1	7.8	39.9	71	109.7	64.5	25.5	3.7	**40.5**
Macon	0.4	2.5	16.7	27	40.6	25	7.7	1.2	**15.2**
Miller	4.7	25.9	84.2	117.1	154.3	85.5	29.6	5.0	**63.3**
Mitchell	4.1	27.6	61.8	97.8	142.7	77.5	24.4	3.8	**55**
Pulaski	1.5	7.8	26.6	36.2	52.3	25.6	7.6	1.7	**19.9**
Randolph	0.3	2.5	11	19.1	30.6	19.1	5.0	0.6	**11**
Schley	0.0	0.1	0.5	0.8	1.2	0.8	0.2	0.0	**0.5**
Seminole	3.4	24.1	66.9	91.5	125.1	71.2	19.4	4.1	**50.8**
Sumter	0.9	7.8	38.7	63.5	90.1	51.3	16.4	2.2	**33.9**
Terrell	0.4	4.3	12.1	24	38.2	22.8	8.1	1.0	**13.9**
Thomas	0.9	5.7	14.8	21.7	29.5	15.9	4.6	1.5	**11.9**
Tift	2.3	12.8	31.1	46.1	63.9	37.6	13.9	3.4	**26.4**
Turner	0.6	5.1	18.6	28.8	42.1	25.4	9.6	1.8	**16.5**
Wilcox	2.1	13.5	35.3	45.9	62.5	33.9	12.8	2.9	**26.1**
Worth	1.1	6.4	32.4	55.9	82.8	49.1	18.4	3.3	**31.2**
Study area	**43.8**	**275.8**	**862.4**	**1,313.7**	**1,853.2**	**1,047.8**	**348.3**	**64.7**	**727.2**

Table 7. Model-estimated inches of irrigation applied, by crop and by county; and for all crops, by months in the growing season (March through October) in the Benchmark Farms Study area, southwestern Georgia, 1996

County	Model-estimated inches of irrigation applied						
	Corn (inches)	Cotton (inches)	Peanuts (inches)	Pecans (inches)	Turf (inches)	Other (inches)	All crops (inches)
Baker	9.2	6.4	5.6	8.4	15.2	23.1	**8.8**
Ben Hill	16.6	11.6	10.4	14.9	27.5	41.3	**17.1**
Bleckley	15	11	9.6	13.8	26.1	38.2	**14.9**
Brooks	11.3	8.5	7.4	10.6	18.9	27.7	**13.4**
Calhoun	14	9.3	8.6	11.4	21.1	31.8	**10.8**
Clay	14.7	8.8	8.2	12.1	24	34.7	**10.5**
Colquitt	9.9	6.7	5.8	8.7	16.1	24.3	**10.4**
Cook	11.8	8.8	7.3	10.7	18.1	30.9	**15.4**
Crisp	12.8	8.8	7.7	12	23	35	**10.8**
Decatur	9.9	6.8	5.9	9.3	15.6	24.4	**9.3**
Dodge	16.7	11.4	10.2	14.9	26.3	42	**16.4**
Dooly	12.7	8.6	7.4	11.5	22.5	34.1	**10.3**
Dougherty	11.9	8.3	7.4	11.6	22.5	30.3	**10**
Early	9.3	6.7	5.8	8.3	14.6	23.8	**9.0**
Grady	9.3	6.4	6.1	8.9	15.7	26.8	**10**
Houston	13.8	8.8	7.4	10.9	23.6	35.3	**10.4**
Irwin	17.4	11.9	10.5	15.8	29.2	43.8	**17.6**
Lee	12.7	8.9	7.4	10.9	21.4	32.3	**10.2**
Macon	12.6	8.8	7.3	10.6	19.7	31.2	**10.1**
Miller	9.5	6.9	5.8	8.7	15.4	24.2	**9.3**
Mitchell	9.1	6.1	5.2	7.7	14.9	22.2	**8.6**
Pulaski	16.8	11.5	9.7	15	28.7	43	**15.4**
Randolph	13.1	8.4	7.4	11.8	21.3	33.2	**10.2**
Schley	12.9	8.7	7.5	12.2	25.3	37.8	**10.6**
Seminole	9.9	6.9	6.1	8.9	15.8	22.5	**9.4**
Sumter	12.3	8.6	7.4	11.1	21.4	30.8	**9.9**
Terrell	12.8	8.8	7.9	11.4	22.4	29.9	**10.6**
Thomas	9.3	6.3	5.6	8.4	15.7	23.7	**9.4**
Tift	18.2	12.2	10.8	16.3	29.7	44.1	**18.7**
Turner	14.5	9.8	8.5	13.1	24.5	37.1	**12.5**
Wilcox	18.4	12.6	10.9	16.5	28.4	46.1	**17.8**
Worth	13.5	9.2	8.1	12.3	22.8	34.3	**11.3**
Study area	**11.6**	**8.6**	**7.4**	**12.2**	**18.9**	**31.7**	**11.4**

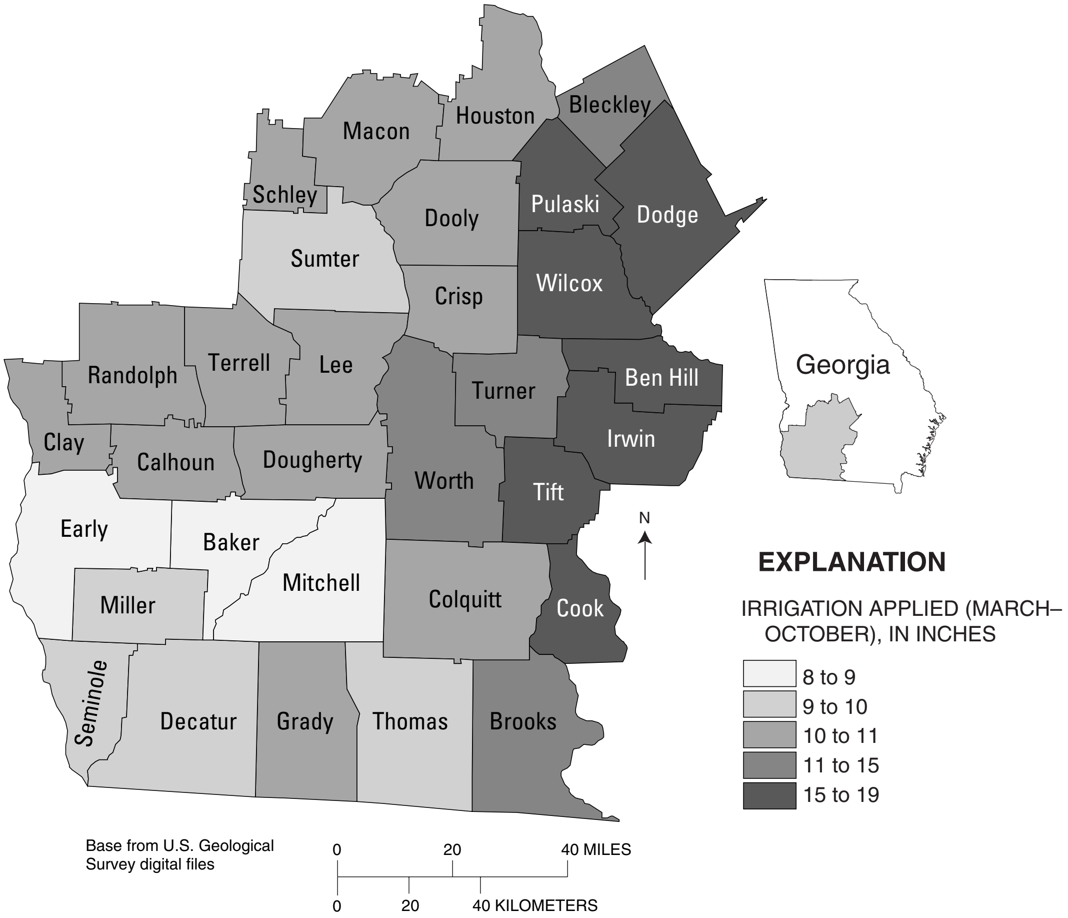


Figure 14. Inches of irrigation applied during the growing season ((March through October) in the Benchmark Farms Study area, southwestern Georgia.

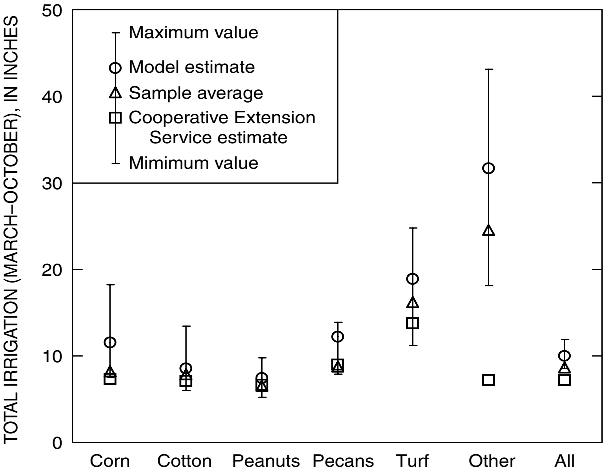


Figure 15. Model-estimated irrigation inches applied with 90-percent confidence interval, average irrigation inches applied, and Cooperative Extension Service estimated irrigation inches applied, by crop and all crops, in the Benchmark Farms Study area, southwestern Georgia.

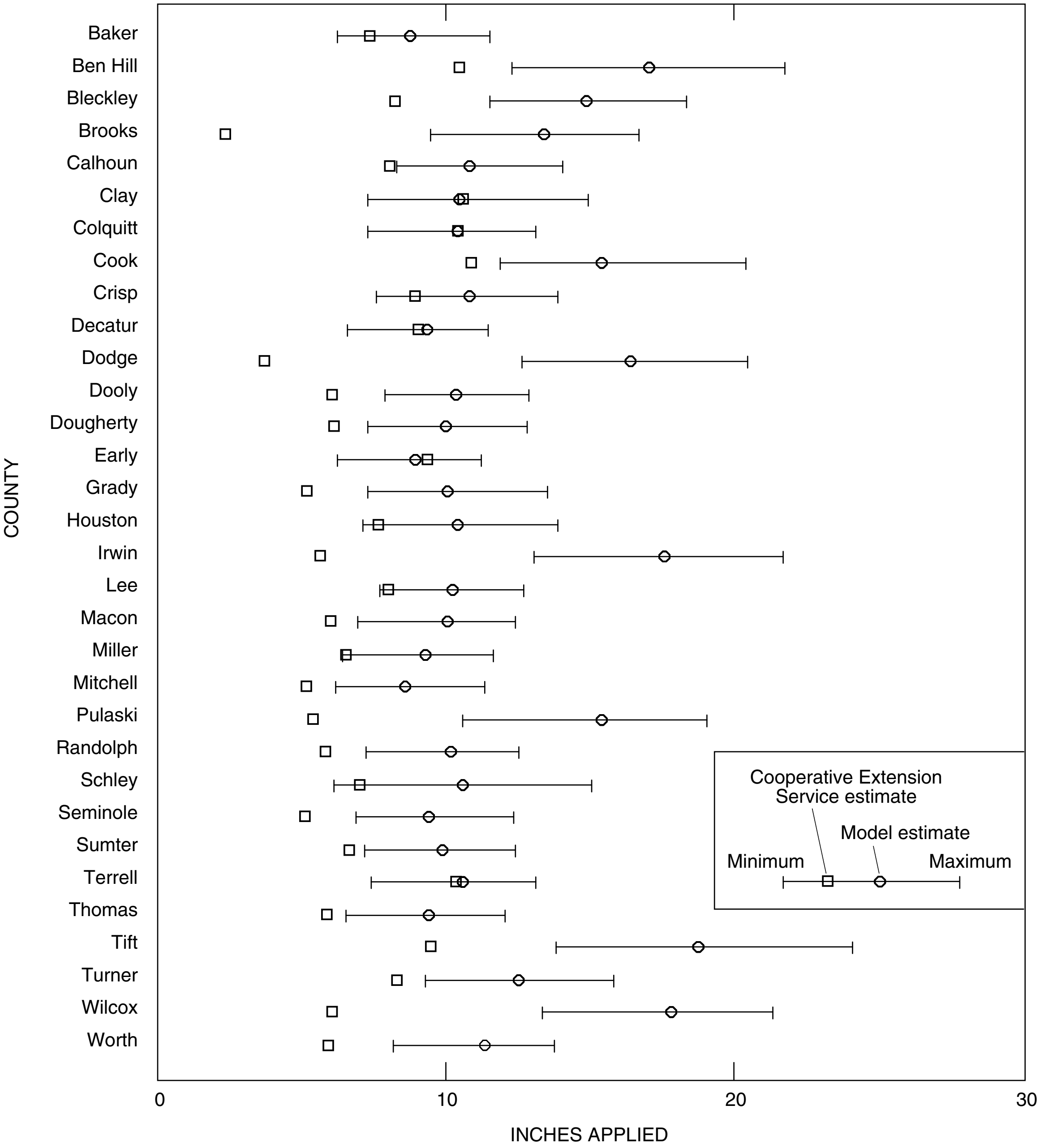


Figure 16. Estimated irrigation inches applied with 90-percent confidence interval in the Benchmark Farms Study area, southwestern Georgia (inches applied computed only for predicted or actual permits that irrigate).

Estimates of acres irrigated by crop for each of the 32 counties and the entire study area are listed in table 8. An acre is counted as irrigated if it is predicted to be irrigated for at least 1 month during the growing season. Counties having the most irrigated acreage are Mitchell, Colquitt, and Miller. Low irrigated acreage counties include Schley, Clay, and Houston. The map of acreage irrigated (fig. 17) shows a band of counties having the largest amount of irrigated acreage, trending east-west across the middle of the study area. Counties having the smallest amount of irrigated acreage generally are located on the north and northwestern borders of the Benchmark Farms Study area. The 90-percent confidence intervals for irrigated acreage for the region, by crop, and for individual counties are shown in figures 18 and 19, respectively. Most acreage is planted with cotton, followed by peanuts, corn, pecans, and others.

The results for inches applied and acres irrigated presented in tables 7 and 8, respectively, help explain variations in annual irrigation across counties described in table 5. Irwin and Tift—the two counties having the largest amounts of irrigation in the study area—were only the eighth and fourteenth largest counties in terms of acreage irrigated, respectively. These two counties were third and first in inches of irrigation applied. Conversely, Schley, Clay, and Houston have the least irrigation and also the fewest irrigated acres. Thus, large amounts of irrigation was mostly a result of high rates of irrigation coupled with moderate irrigation to large numbers of acres with a minimum of labor and cost (*effort*); whereas, low amounts of irrigation primarily was due to low irrigated effort.

Comparison of Model Results to Estimates Based on Sampling and Other Methods

The results from the irrigation-statistical model can be partially verified by comparison with estimates from other estimation methods and independent sources. An independent set of irrigation estimates was obtained for all counties in the study area in 1995, through a survey of CES agents conducted by The University of Georgia Extension Service (K.E. Harrison, Cooperative Extension Service, Tifton, Ga., written commun., 1995). In addition to providing a check on the total irrigation estimates, the CES estimates were useful for validating the estimates from the models.

Additionally, it was useful to check for bias in the statistical model by comparing the results with model-free or robust (that is, statistical design) estimates derived from the original sample of permits. The identification of sites in the Benchmark Farms Study was originally based on a stratified random design. In principle, sample-based estimates derived from these data should be unbiased, however, decisions made by farmers on whether to participate in the study could systematically bias the results. The model has the potential to overcome such a selection bias by conditioning predictions on numerous predictor variables. Additionally, by imposing structure on the estimation process, the model has the potential to raise the precision of the predictions. On the other hand, the model could be misrepresented, resulting in biased predictions; and prediction with the model involves the introduction of numerous independent errors that could result in less precision.

The merits of the model compared to a robust sampling method were investigated in two ways. First, robust sampling methods were applied to the original Benchmark Farms Study sites data to obtain robust estimates of total irrigation for counties and the study area. Second, the structural statistical model (described in Appendix D) was combined with Monte Carlo methods to simulate crop choice, acres irrigated, irrigation duration, and irrigation inches applied for the set of farms comprising the original Benchmark Farms Study sites. Robust sampling methods then were applied to these simulated farm data to obtain estimates of county and study-area irrigation. The aggregated irrigation estimates, derived from the actual and simulated data, were used to identify relative bias between the model and robust estimation methods.

The first comparison relates the amount of acreage available for irrigation, as estimated from the model, to acreage reported on the permit applications. County estimates of the amount of acreage predicted from the acres-irrigated model, with a 90-percent confidence interval, plotted with the amount of acreage recorded on the permit application are shown in figure 12. The model estimates were obtained by predicting acreage for each non-sampled permit in the permit file, regardless of whether or not the permit holder was predicted to irrigate during the growing season. As shown in figure 12, the permit values lie within the model confidence intervals for 21 of the 32 counties. Overall, total acreage from the permit applications for the study area was about 1,370,717 acres, which was only slightly more than the model estimate of 1,304,323 acres. These results show good consistency between the two estimates.

To assess the validity of the whether-to-irrigate model, the acres irrigated as estimated by the model are compared to estimates derived from the 1995 CES survey (K.E. Harrison, Cooperative Extension Service, Tifton, Ga., written commun., 1995). The CES estimates lie within the model-derived confidence intervals for only 11 of the 32 counties (fig. 19). For most counties, the model produced a larger estimate than the reported acres irrigated in the CES survey. Total acreage irrigated in the study area was estimated by the model to be 1,155,214 acres, compared to only 977,974 acres from the CES, an 18-percent difference. Given that potential irrigated acreage was closely approximated by permitted acreage, an amount derived from surveying the

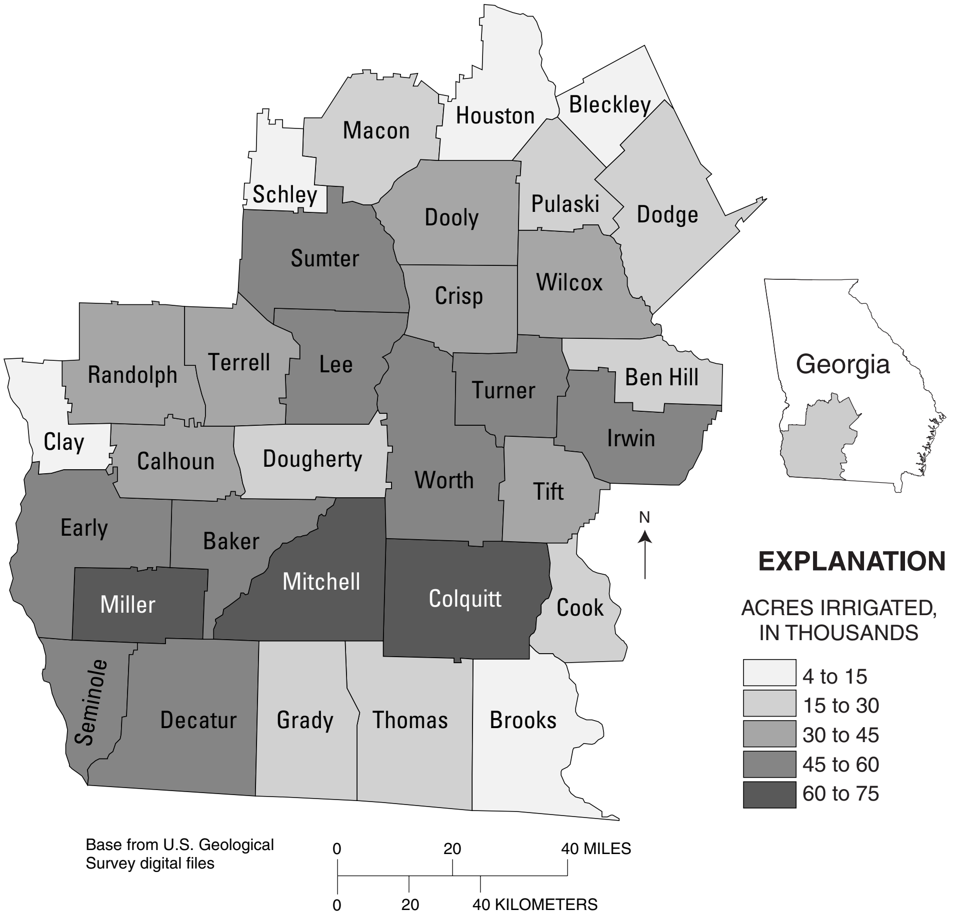


Figure 17. Acres irrigated in the Benchmark Farms Study area, southwestern Georgia.

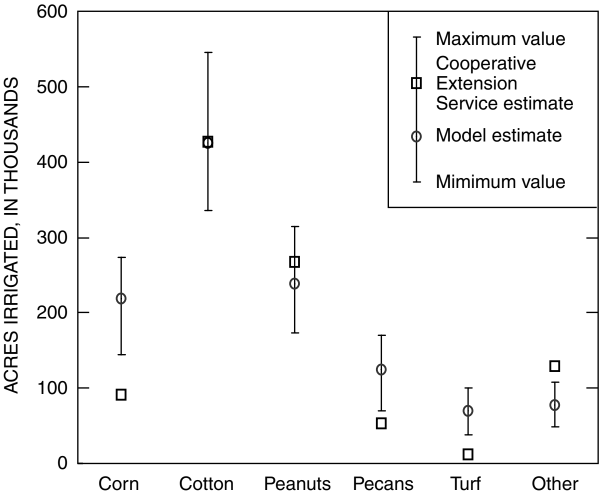


Figure 18. Model-estimated acres irrigated with 90-percent confidence interval, and Cooperative Extension Service estimated acres irrigated by crop, in the Benchmark Farms Study area, southwestern Georgia.

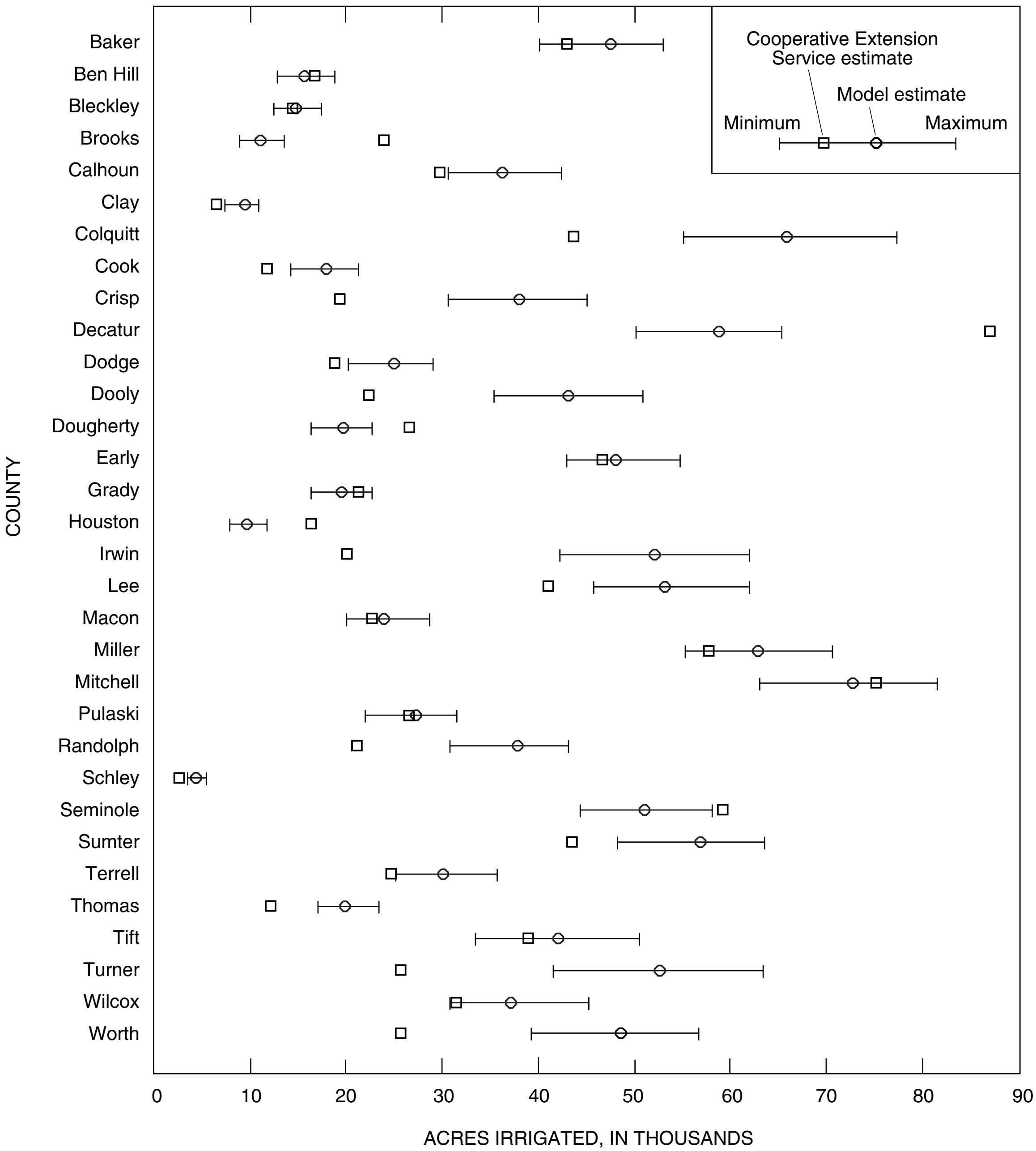


Figure 19. Estimated acres irrigated with 90-percent confidence interval in the Benchmark Farms Study area, southwestern Georgia.

Table 8. Model-estimated irrigated acreage, by crop and by county; and total irrigated acres for all crops in the Benchmark Farms Study area, southwestern Georgia, 1996

County	Estimated irrigated acres, by crop						
	Corn	Cotton	Peanuts	Pecans	Turf	Other	Totals
Baker	13,040	15,701	9,451	2,525	3,917	2,940	**47,574**
Ben Hill	2,395	4,868	2,821	2,870	917	1,865	**15,736**
Bleckley	2,163	5,013	2,931	2,637	765	1,410	**14,918**
Brooks	2,941	2,684	1,677	791	1,391	1,650	**11,135**
Calhoun	4,508	17,223	8,431	4,087	909	1,115	**36,273**
Clay	1,110	4,272	2,591	1,075	247	267	**9,562**
Colquitt	17,396	18,706	11,482	4,144	7,096	7,052	**65,874**
Cook	4,217	3,481	2,248	1,273	3,406	3,396	**18,022**
Crisp	5,148	16,517	8,871	4,906	1,173	1,497	**38,112**
Decatur	15,507	20,164	11,647	3,122	4,667	3,695	**58,801**
Dodge	3,622	8,466	4,499	4,406	1,496	2,607	**25,097**
Dooly	5,654	19,461	10,089	5,276	1,211	1,529	**43,221**
Dougherty	2,591	8,548	4,450	2,932	576	649	**19,746**
Early	12,824	16,043	9,750	2,549	3,990	3,042	**48,199**
Grady	5,076	5,814	3,655	1,127	1,950	1,889	**19,510**
Houston	1,228	4,361	2,202	1,386	298	320	**9,794**
Irwin	7,644	16,570	9,552	9,645	3,038	5,780	**52,230**
Lee	6,926	23,738	13,021	6,592	1,431	1,624	**53,331**
Macon	3,019	10,593	5,761	3,056	680	811	**23,919**
Miller	17,009	20,551	12,349	3,366	5,365	4,251	**62,892**
Mitchell	19,161	23,788	14,175	4,363	6,288	4,995	**72,770**
Pulaski	3,975	9,852	5,605	4,600	1,321	2,061	**27,414**
Randolph	4,924	17,231	9,043	4,467	1,042	1,148	**37,855**
Schley	578	1,950	1,053	564	127	158	**4,431**
Seminole	13,537	17,229	10,117	2,624	4,148	3,484	**51,140**
Sumter	7,323	26,361	13,716	6,407	1,478	1,683	**56,967**
Terrell	4,200	13,070	6,883	3,866	887	1,236	**30,143**
Thomas	5,181	6,302	3,773	1,144	1,893	1,663	**19,956**
Tift	5,944	12,759	7,250	8,092	2,601	5,482	**42,128**
Turner	7,391	21,370	11,587	7,600	2,025	2,756	**52,729**
Wilcox	5,248	12,493	7,065	6,566	2,232	3,562	**37,166**
Worth	6,642	20,811	10,858	6,536	1,668	2,054	**48,569**
Study area	**218,121**	**425,989**	**238,605**	**124,598**	**70,231**	**77,670**	**1,155,214**

entire population, either the CES survey missed about 20 percent of the irrigated acreage, or the rate that farmers decided not to irrigate was 26 percent rather than the 11 percent estimated by the model (see the discussion of results from the whether-to-irrigate model). Because the rate at which farmers decide to irrigate was estimated with relatively high precision (see Appendix D, table D-1), it seems likely that the CES estimate was biased downward.

The crop-selection model can be assessed by comparing model estimates of acreage by crop for the entire study area with similar estimates made from the CES survey. The two estimates depict a similar pattern of acreage across crops (fig. 18) (the correlation of ranks is 0.83). Estimates of cotton acreage—the most commonly irrigated crop—were virtually identical for the two methods, and the CES estimate for peanuts, the second most commonly irrigated crop, were within the model-estimated confidence interval. With the exception of "other" crops, however, the CES estimates for the remaining crops generally were less than half the estimate obtained from the model. This suggests that the CES estimates of irrigated acreage probably was biased downward explaining at least some of the shortfall. A general downward bias, however, should be reflected in a uniform shortfall across all crops, which was not the case as shown in figure 16. Therefore, the model may have over-represented cotton and peanut acreage without consideration for the influence of corn, pecans, and turf.

The CES survey data could not produce estimates of irrigation by month; therefore, it was not possible to use these data to validate the when-to-irrigate model. It was possible, however, to determine if the model created any biases relative to 167 of the 180 sites in the original sample used for the model. The methods used the 200 sets of parameter estimates for the acre, crop, and when-to-irrigate models, generated as part of the bootstrap calibration, to obtain 200 predictions of when irrigation took place for each permit in the irrigated sample (predictions from the acres-irrigated model were used to predict crop planted, which was incorporated into the when-to-irrigate model). Irrigation frequency distribution of these simulated sample permits could be compared to the original sample to determine relative bias—a comparison of the two frequency distributions is given in table 9. These months represent the range of months for which the when-to-irrigate model produced estimates. Permits from the original sample that have starting or ending months outside this range were assigned to the closest month in the defined range. The results showed close agreement between the sample and model-derived estimates; the correlation between frequencies was equal to 0.94. The when-to-irrigate model was not an apparent source of significant bias in the irrigation estimation process.

The CES provided crop-level estimates of irrigation in inches applied for 1995. These estimates, aggregated to the study area, by county, are shown in relation to the model estimates in figure 15. Across crops, the CES estimates show a general correspondence with the model estimates, except for the crop categories of corn, other, and all crops. The CES estimates lie within the confidence intervals of the model estimates. This is primarily the consequence of large uncertainty surrounding the model estimates. The county-level estimates, shown in figure 16, indicate a relative

Table 9. Comparison of irrigation starting and ending month frequencies between the sample and 200 simulations of the sample using the irrigation model, in the Benchmark Farms Study area, southwestern Georgia, 1996

[—, not applicable; sample frequencies are in bold text; model-simulated sample frequencies are in normal text]

Irrigation ending month (percent)		Irrigation starting month				
		March (percent)	April (percent)	May (percent)	June (percent)	July (percent)
—	—	**10.14**	**16.22**	**40.54**	**24.32**	**8.78**
		9.75	17.96	38.81	24.54	8.94
July (percent)	**7.43**	**0.00**	**2.70**	**3.38**	**1.35**	**0.00**
	8.40	0.56	1.38	3.24	2.35	0.85
August (percent)	**40.54**	**2.03**	**8.11**	**16.89**	**9.46**	**4.05**
	40.25	2.96	6.60	15.43	10.96	4.30
September (percent)	**41.22**	**2.03**	**4.73**	**16.22**	**13.51**	**4.73**
	41.24	4.09	7.40	16.53	9.73	3.48
October (percent)	**10.81**	**6.08**	**0.68**	**4.05**	**0.00**	**0.00**
	10.11	2.12	2.58	3.60	1.51	0.30

downward bias of the CES estimates compared to irrigation model estimates. Only 11 counties have a CES estimate within the model-derived confidence interval, the remaining 21 counties all have an CES estimate that was less than the model estimate. The average inches applied for each crop estimated directly from the Benchmark Farms Study sites is shown in figure 15. The averages were computed by summing the amount of irrigation over all sites for a given crop and dividing by the total amount of acreage among those permits. The summations included only sites that did not have missing monthly values, although a site was included if multiple monthly values were temporally aggregated. The average of the Benchmark Farms Study sites was in close agreement with the CES estimates for all crops except turf and other crops, implying that the inches-applied model was biased high. The Benchmark Farms Study sites average of inches applied over all crops was about 15 percent below the model estimate; and the CES estimate was about 30 percent below the model estimate.

Due to the low estimates of inches applied and acres irrigated reported by the CES, the estimate of total irrigation for the study area, shown in figure 20, was about half of the model estimate. The discussion above provides some evidence that the CES estimate of irrigated acres is biased low by about 20 percent. Conversely, the evidence describing the relation between the sample average and model estimates of irrigation inches applied suggests that the model was biased high by about 15 percent. Thus, there appears to be a convergence towards an unbiased estimate of total irrigation in the range of 15 to 30 percent below the model estimate.

There were some subtle issues in the use of the sample average estimates that needed to be addressed. First, the original sample was selected on the basis of a stratified random design with preferential weight given to agricultural permits with large amounts of permitted acres. This was important because there was evidence that the simple correlation between inches of irrigation applied and irrigated acreage (that is, without controlling for other factors such as pump capacity) was strongly negative. Thus, if large farms were more irrigation efficient and large farms were over-represented in the sample, a simple (that is, unweighted) average estimate of irrigation inches applied was biased downward. Correction of this bias could only be achieved through a weighted averaging of observations according to acreage or by the calibration of a model that controls for acreage irrigated.

A second consideration affecting the interpretation of the sample average estimates concerns the potential for bias in the final sample. A condition of the Benchmark Farms Study gave farmers the option to decline participation, an option that was exercised by a great many of the original contacts. Extensive lack of participation by farmers opened the possibility that the resulting sample contained significant selection bias, a bias that was difficult to predict and correct. Such bias could have similar effects on both sample average and model-derived estimates. If the unknown selection criteria, however, were sufficiently correlated with predictor variables used in the model specification, then the resulting model estimates may have been relatively bias free.

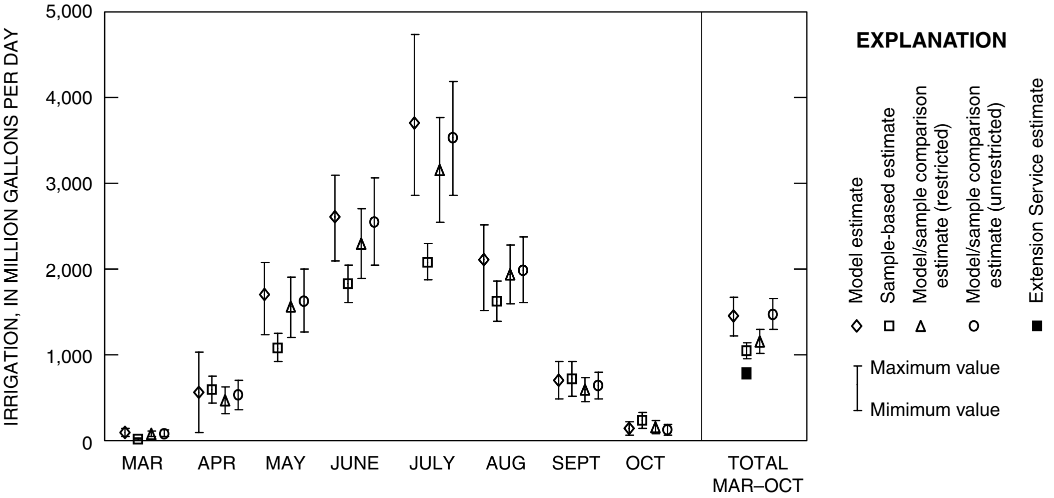


Figure 20. Model, sample-based, model/sample comparison (restricted and unrestricted), and Cooperative Extension Service estimated irrigation with 90-percent confidence interval, by month, and total estimated irrigation during the growing season (March through October) in the Benchmark Farms Study area, southwestern Georgia.

To assess the potential for relative bias between sample and model estimates, a third estimate of total irrigation for the Benchmark Farms Study area was developed, based on robust methods (Cochran, 1977) applied to the original stratified random sample. The sample was selected using optimal stratification criteria to delineate five permitted acreage classes; less than 43 acres, between 43 and 105 acres, between 105 and 230 acres, greater than 230 acres, and acreage not reported. Analysis of total irrigation in the sample showed that the distinction of these classes is statistically significant at the 5-percent significance level. The resulting county-level estimates of total irrigation and irrigation taken from ground water are given in tables 10 and 11, respectively. A comparison of sample- and model-based monthly estimates for the entire study area is presented in figure 21. The comparison shows model estimates were larger than sample-based estimates for most months, with significant differences occurring for the peak irrigation months May through August. Over the entire growing season, the sample-based estimate of total irrigation was 28 percent less than the model estimate. Thus, this analysis confirmed the presence of relative bias between the sample- and model-based methods.

Although the sample-based estimate was low relative to the model estimate, the sample-based estimate was still more than 30 percent greater than the CES estimate (fig. 21). This relative bias was greater than the suspected 20 percent downward bias associated with the CES estimate of irrigated acreage. Thus, there was an implied additional 10 percent relative bias in inches of irrigation applied between the sample-based and CES estimates.

A second observation gleaned from the comparisons in figure 19, and also depicted in figure 22, concerned the relative error associated with the estimates. The sampling-based method produced estimates that were decidedly more precise than the model estimates. This is partly explained by the model estimate being an aggregation of independently random component estimates for each permit whereas the sample estimate was derived from a single permit datum. For example, to estimate annual irrigation the model aggregated permit-level monthly estimates consisting of random simulations of various model components. In contrast, the sampling method simply aggregated annual irrigation statistics compiled at the permit level. Further investigation, however, showed this was not a major source of differential precision. The disparity in precision was even greater for many of the monthly estimates; and a bootstrap simulation of the model, in which only the inches applied component was repeatedly randomized across iterations (all other components are randomly simulated only once, making them effectively fixed) showed only a small effect on precision. A more likely explanation concerns the estimation of the inches-applied model in logarithmic space, which resulted in a less-precise prediction than estimation in real space. But a complete understanding of the precision differences across methods has yet to be obtained.

Given the existence of relative bias between the model and sample-based methods, it remained to determine which method was the cause. To help determine the cause, an experiment was attempted in which the model's bootstrap calibration results were used to obtain 200 simulated realizations of the 167-permit sample. The simulated samples were used to compute 200 stratified sample-based estimates of irrigation. The average of those estimates then were compared to the model and original sample-based estimates. If the simulated sample estimates were close to the original sample estimates, then this implied the model was providing an unbiased estimate of the sample. Consequently, the source of the bias must have been the selection of permits used to generate the sample-based estimate. On the other hand, if the simulated sample estimates were close to the model estimates, then this implied the sampling method, by itself, was unbiased and the source of the relative bias was the model.

Two variants of the experiment were conducted. In the first approach, the simulated sample was restricted to include only months for which the original sample provided an estimate. This implied that no observations were generated for months in which a permit had a zero, missing, or temporally aggregated value. This approach implicitly used the original survey data to determine whether to irrigate and when a permit holder irrigated; no simulations for the whether-to-irrigate or when-to-irrigate model components were generated (simulations of the acres irrigated, crop selection, and irrigation-applied model were included). Under the second approach, no restrictions were placed on the simulated sample. A concern with the first approach was that it did not adequately test the whether-to-irrigate and when-to-irrigate model. Given the discussion of results earlier in this section, however, these components showed no evidence of bias with respect to the sample. A concern with the second approach was that it included observations that were not in the original sample, thereby raising the possibility that the sample-based results were sensitive to the particular observations included in the analysis.

Results of the experiment are shown in figure 20 and summarized below. Points identified as "Model/Sample Comparison Estimate (restricted)" and "Model/Sample Comparison Estimate (unrestricted)" correspond to the sample-based estimates of irrigation using the restricted and unrestricted simulated samples. The restricted and unrestricted simulated sample estimates were both between the model and sample-based estimates for all peak irrigation months, with the restricted estimate about halfway between and the unrestricted estimate close to the model estimate. The estimates for the entire growing season, which were not

Table 10. Sample-based estimated rate of irrigation, by county and by month; and for the growing season (March through October) in the Benchmark Farms Study area, southwestern Georgia, 1996

[Mgal/d, million gallons per day over the identified period]

County	Sample-based irrigation estimates								
	March (Mgal/d)	April (Mgal/d)	May (Mgal/d)	June (Mgal/d)	July (Mgal/d)	August (Mgal/d)	September (Mgal/d)	October (Mgal/d)	March through October (Mgal/d)
Baker	0.8	21.5	37	69.4	84.9	65.4	18.8	4.6	38.2
Ben Hill	0.5	12.9	25.3	37.3	38.9	30.3	21.9	8.4	23
Bleckley	0.4	11.3	17.6	32.9	34.8	27.7	10.2	2.8	16.9
Brooks	0.3	9.0	16.3	28.5	32.9	25.7	9.4	2.8	15.9
Calhoun	0.5	12.8	23.3	42.7	54.1	41.4	12	3.0	24.4
Clay	0.1	3.3	6.3	12	16.7	12.5	2.8	0.6	7.0
Colquitt	1.6	38.5	74.6	109.2	111.6	87.2	66.8	26	67.5
Cook	0.7	16.9	37	50.7	55	42.5	33.4	13.4	33.6
Crisp	0.6	15.6	29.3	50.1	59.9	45.9	18.9	6.0	29
Decatur	0.6	26.7	46.2	85.4	97.6	77.4	20.3	4.4	44.9
Dodge	0.7	19.1	34.9	56	59.7	46.7	27	9.6	32.6
Dooly	0.7	16.8	31.3	55.4	70	53.2	18.9	5.6	32.6
Dougherty	0.3	7.8	14	25.5	30.9	23.9	7.3	1.9	14.2
Early	0.7	22.9	38	72.4	84.8	66.3	17.6	3.8	38.2
Grady	0.4	11.1	20.6	32.8	35.6	27.7	16	5.7	19.4
Houston	0.2	3.7	7.6	13.1	18.1	13.5	4.5	1.4	8.2
Irwin	1.6	41.7	76.1	116.6	115.4	91.5	64.1	24.1	68.3
Lee	0.7	20.2	36.1	67.2	83.5	64.4	17.3	4.0	37.3
Macon	0.4	8.9	19.1	31.7	42.8	32.2	11.5	3.6	20
Miller	0.8	30.7	49.8	93.2	101.8	81.1	24.7	6.0	47.9
Mitchell	1.0	35.3	61.1	109.4	122.7	96.8	33.4	9.3	59
Pulaski	0.6	17.5	31.2	55.6	66.4	51.3	18.2	5.3	31.4
Randolph	0.5	14.1	26.6	47.9	61.4	47	13.4	3.5	27.7
Schley	0.1	1.7	3.0	5.6	7.1	5.4	1.8	0.5	3.2
Seminole	0.6	23.4	39.3	73.8	84	66.3	18	4.0	38.4
Sumter	0.9	18.8	40.4	72.3	106.4	78.7	20.7	5.4	45.9
Terrell	0.5	12.3	20.5	38.4	45.5	35.2	11.5	3.1	21
Thomas	0.4	9.9	19.8	31.4	37.2	28.5	14.4	5.1	19.3
Tift	1.5	35.7	71.9	101.1	101.4	79.3	66.8	26.6	64
Turner	1.0	25.8	43.6	70.8	69	54.9	35.4	12.8	39.4
Wilcox	1.0	27.7	47.1	79.2	82	64.7	34.8	11.9	43.9
Worth	0.8	21.4	38	63.8	69.7	54.6	26.2	8.7	36.1
Study area	**21.1**	**595**	**1,082.7**	**1,831.5**	**2,082**	**1,619.4**	**718**	**233.8**	1,048.4

Table 11. Sample-based ground-water estimated rate of irrigation, by county and by month; and for the growing season (March through October) in the Benchmark Farms Study area, southwestern Georgia, 1996
[Mgal/d, million gallons per day over the identified period]

County	Sample-based ground-water irrigation estimates								
	March (Mgal/d)	April (Mgal/d)	May (Mgal/d)	June (Mgal/d)	July (Mgal/d)	August (Mgal/d)	September (Mgal/d)	October (Mgal/d)	March through October (Mgal/d)
Baker	0.6	18.3	31	58.9	70.8	54.9	14.8	3.3	31.7
Ben Hill	0.2	4.2	8.0	12.9	14.7	11.4	5.7	2.0	7.7
Bleckley	0.1	4.8	7.9	15.1	17.3	13.6	3.7	0.8	7.9
Brooks	0.2	6.2	10.8	19.9	23.4	18.3	5.3	1.3	10.7
Calhoun	0.2	6.2	11	19.5	22.6	17.7	6.0	1.7	10.8
Clay	0.0	0.8	1.6	3.1	4.6	3.4	0.7	0.1	1.9
Colquitt	0.2	7.8	13.5	23.4	25.6	20.2	8.5	2.7	12.9
Cook	0.2	6.1	11.6	18.5	20.2	15.8	8.6	3.1	10.9
Crisp	0.2	5.7	10.3	18.6	22.6	17.4	5.7	1.5	10.4
Decatur	0.5	23.3	39.8	74.3	84.4	67.1	16.7	3.4	38.6
Dodge	0.3	7.3	12.7	22.8	26.6	20.6	7.7	2.2	12.7
Dooly	0.5	11.4	21.5	37.8	47.8	36.4	12.9	3.9	22.3
Dougherty	0.2	7.4	13.1	24	28.9	22.4	6.6	1.7	13.2
Early	0.5	17.1	28.2	53.7	62.1	48.8	12.8	2.7	28
Grady	0.2	4.0	8.4	12.6	14.7	11.3	6.6	2.5	8.0
Houston	0.1	2.9	5.9	10.1	13.2	10	3.6	1.1	6.2
Irwin	0.3	7.8	13.8	23.7	26.1	20.5	8.7	2.8	13.1
Lee	0.4	14.6	25.9	47.8	57.6	44.8	12.3	2.9	26.1
Macon	0.2	5.4	11.9	19.4	26.2	19.6	7.5	2.5	12.4
Miller	0.8	30.1	48.8	91.5	100.1	79.7	24	5.8	47
Mitchell	0.7	28.4	48.6	89.7	102.5	80.8	23.4	5.6	47.5
Pulaski	0.3	7.7	15	25.6	31.5	24.2	9.0	2.7	15.1
Randolph	0.1	3.9	7.7	14.5	19.9	15.1	2.8	0.4	8.4
Schley	0.0	0.2	0.3	0.6	0.8	0.6	0.1	0.0	0.3
Seminole	0.6	22.3	37.3	70.4	80.3	63.5	16.6	3.5	36.5
Sumter	0.5	11.1	24	42.7	63.1	46.6	12.5	3.3	27.3
Terrell	0.2	5.2	8.8	16.3	19.2	14.9	4.8	1.3	8.9
Thomas	0.2	6.1	11.8	19.1	22.1	17.1	8.2	2.8	11.3
Tift	0.6	12.6	21.2	35.2	36.1	28.2	17.4	6.2	19.8
Turner	0.2	6.4	10.8	19.1	20.9	16.4	6.7	2.0	10.3
Wilcox	0.4	10.5	17.4	31.6	35	27.5	10.3	2.9	16.9
Worth	0.4	11	19	35	41.8	32.4	10	2.6	19.2
Study area	**10**	**316.6**	**557.9**	**1,007.2**	**1,182.5**	**921.2**	**300.3**	**81.3**	554.1

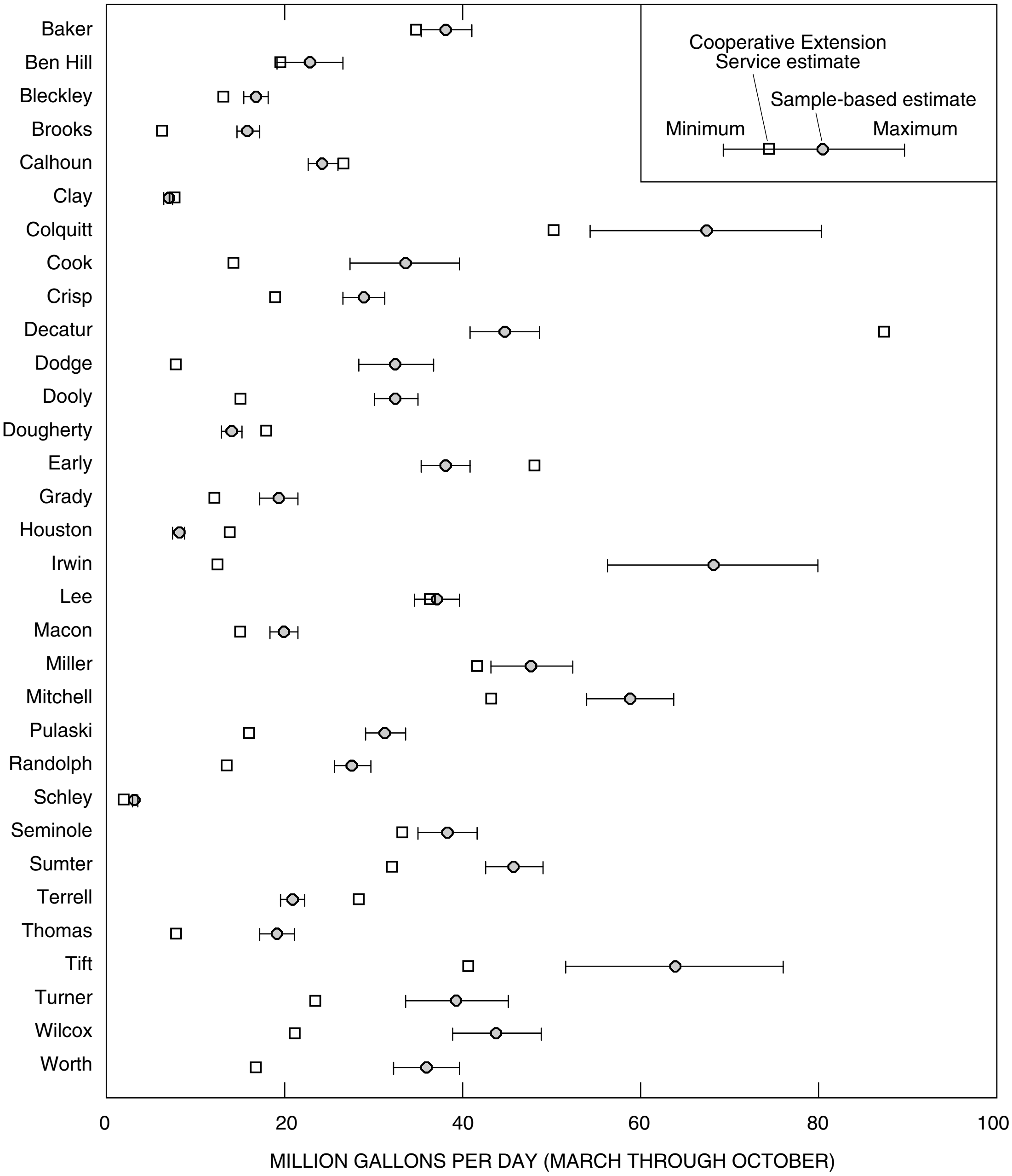


Figure 21. Sample-based estimated irrigation during the growing season (March through October) with 90-percent confidence interval in the Benchmark Farms Study area, southwestern Georgia.

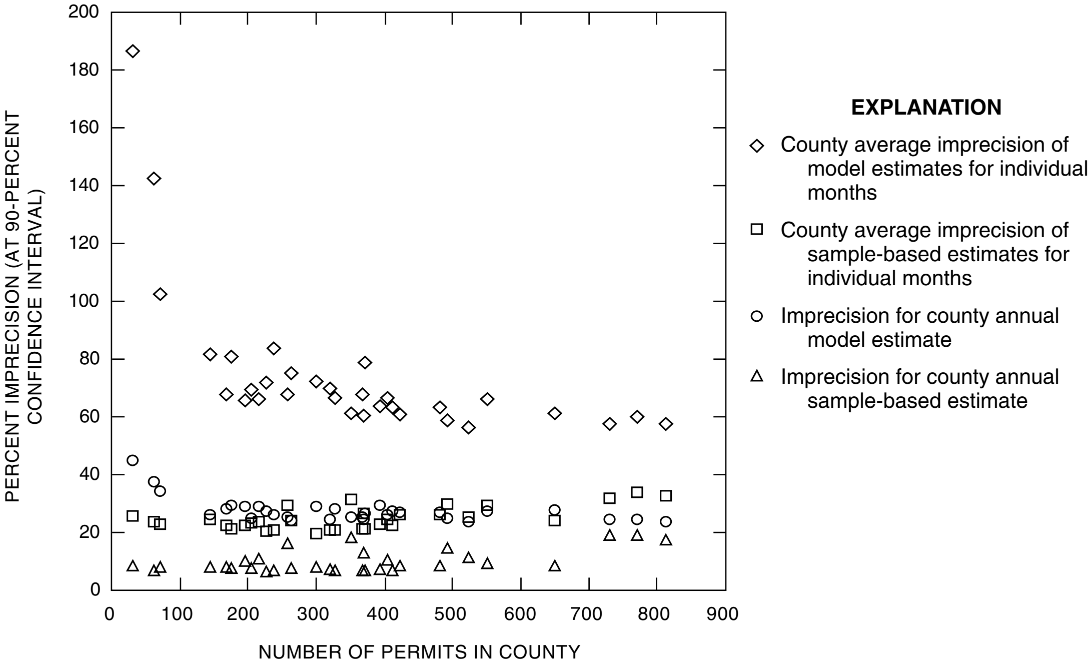


Figure 22. Averages of percent imprecision for monthly estimates and imprecision for annual estimates from model and sample-based methods, by number of permits issued in each county, in the Benchmark Farms Study area, southwestern Georgia.

a simple aggregation of the monthly estimates because they included only permits that have no missing monthly observations, showed the restricted estimate to be close to the sample-based estimate and the unrestricted estimate to be close to the model estimate. Although these results are arguable, they generally support the concept that approximately half of the relative bias between the model and sample-based estimates was due to the model. This bias probably originated in the irrigation-applied model. The remaining relative bias could be due to the sample-based method and was likely the result of sensitivity of the method to a few unrepresentative observations. Confirmation of this assessment would consist of identifying the particular observations to which the analysis was sensitive. At present, such an identification has not been made.

In conclusion, there was evidence that the CES irrigation estimates were low by at least 30 percent and possibly by as much as 40 percent, with 20 percent of the bias due to an underestimate of irrigated acreage. The model estimate showed evidence of an upward bias of about 15 percent, with the likely cause being a misrepresented inches-applied model. A better understanding of the causes of bias in the model could be obtained with a larger sample size. Given the necessisty to eliminate monthly observations that were aggregated due to non-reporting, sample size from further studies could be increased substantially by automating the reporting of monthly totalizer amounts.

SUMMARY AND CONCLUSIONS

In 1994, the U.S. Geological Survey, in cooperation with the Georgia Department of Natural Resources, Environmental Protection Division, Water Resources Management Branch; and The University of Georgia College of Agriculture and Environmental Sciences, began a pilot study of irrigation water use to establish a benchmark irrigation monitoring network of farms in a 32-county area of southwestern Georgia. A stratified random sample of 500 permitted irrigators from the study area was selected, using water source and acres irrigated as the critical elements, in order to obtain 180 participants.

A portable noninvasive pipe flowmeter (an instrument that can measure the flow of fluid through a pipe, without having to be in direct contact with the fluid) was attached to a portion of discharge pipe while the pump was running. The RTT-8 Universal Running Time Totalizer has a vibration pick-up accelerometer attached and an 8-digit display, which is activated by the vibration of the fluid in the pump and records a cumulative total of the pump operation time in hours. Data were collected and compiled for 50 farms for 1995 and 130 additional farms for the 1996 growing season—a total of 180 farms.

Estimating irrigation water use was accomplished by developing a statistical model. To accurately reflect patterns in the data, however, it was necessary to break the irrigation model into five model components. The five model components are ordered in a recursive structure—whether to irrigate, acres irrigated, crop selected, seasonal-irrigation scheduling, and the amount of irrigation applied.

The irrigation modeling approach was accomplished in two phases. The first phase was the calibration phase and the second phase was prediction. "Bootstrap analysis," the version of the method used in this report, relies on repeated resampling of the original study observations, considered to be distinct permits containing multiple monthly irrigation values, to construct multiple bootstrap pseudo-samples. Bootstrap analysis is a method that allows for the estimation of precision in any model—regardless of complexity.

The results for irrigation applied and acres irrigated help explain variations in annual irrigation across the counties described. Irwin and Tift, the two largest irrigating counties in the study area, were only the eighth and fourteenth largest counties, respectively, in terms of acreage irrigated, but were third and first, respectively, in inches of irrigation applied. Conversely, Schley, Clay, and Houston had the least irrigation and also the fewest irrigated acres. Thus, higher rates of irrigation mostly are a result of high rate of irrigation coupled with moderate irrigation effort; whereas, lower rates of irrigation primarily are due to low irrigation effort.

In conclusion, there was evidence that the Cooperative Extension Service irrigation estimates were low by at least 30 percent and possibly by as much as 40 percent with 20 percent of the bias due to an underestimate of irrigated acreage. The model estimate showed evidence of an upward bias of about 15 percent, with the likely cause being a misrepresented irrigation-applied model. A better understanding of the causes of bias in the model could be obtained with a larger sample size. Additionally, the estimates could be greatly improved by eliminating the need to exclude temporally aggregated monthly observations. Given the necessisty to eliminate monthly observations that were aggregated due to non-reporting, sample size from further studies could be increased substantially by automating the reporting of monthly totalizer amounts.

REFERENCES CITED

Arvin, D.V., 1992, Feasibility of using portable non-invasive pipe flowmeters and time totalizers for determining water use: U.S. Geological Survey Water-Resources Investigations Report 91-4110, 65 p.

Barber, N.L., 1983, Integrating irrigation water-use information into Georgia's water management program, *in* Proceedings of the National Water Well Association, Eastern Regional Conference on Ground Water Management, Orlando, Fla., October 31-November 2, 1983: Worthington, Ohio, National Water Well Association, p. 641-651.

Cochran, W.G., 1977, Sampling Techniques: New York, John Wiley & Sons, *3rd ed.*, 428 p.

Duerr, A.D. and Trommer, J.T., 1982, The Benchmark Farm Program—A method for estimating irrigation water use in southwest Florida: U.S. Geological Survey Water-Resources Investigations Report 82-17, 49 p.

Fanning, J.L., 1997, Water use in Georgia by county for 1995: Georgia Geologic Survey Information Circular 101, 110 p.

Fanning, J.L., Doonan, G.A., and Montgomery, L.T., 1992, Water use in Georgia by county for 1990: Georgia Geologic Survey Information Circular 90, 98 p.

Holland, T.W. and Baker, N.T., 1993, Evaluation of pumpage data furnished by selected public water suppliers in Arkansas, May 1990 Through March 1991: U.S. Geological Survey Water-Resources Investigations Report 93-4104, 80 p.

Judge, G.G., Griffiths, W. E., Hill, R. Carter, Lütkepohl, Helmut, and Lee, Tsoung-Chao, 1985, The theory and practice of econometrics: New York, John Wiley & Sons, *2nd ed.*, 1,019 p.

Lancaster, Tony, 1990, The econometric analysis of transition data: Cambridge, Cambridge University Press, 352 p.

Luckey, R.R., Heimes, F.J. and Gaggiani, N.G., 1980, Calibration and testing of selected portable flowmeters for use on large irrigation systems: U.S. Geological Survey Water-Resources Investigations Report 80-72, 21 p.

Maddala, G.S., 1983, Limited-dependent and qualitative variables in econometrics: Cambridge, Cambridge University Press, 401 p.

Marella, R.L. and Singleton, V.D., 1988, Metering methods and equipment used for monitoring irrigation in the St. Johns River: Palatka, Fla., St. Johns River Water Management District, 17 p.

Motorola, Inc., 1993, Traxar GPS Navigator Owner's Manual For Models: Traxar, Traxar+, and Traxar MG+,.

Pierce, R.R., Barber, N.L., Stiles, H.R., 1984, Georgia irrigation, 1970-80—A decade of growth: U.S. Geological Survey Water-Resources Investigations Report 83-4177, 29 p.

Trent, V.P., Fanning, J.L., Doonan, G.A., 1990, Water use in Georgia by county for 1987: Georgia Geologic Survey Information Circular 85, 112 p.

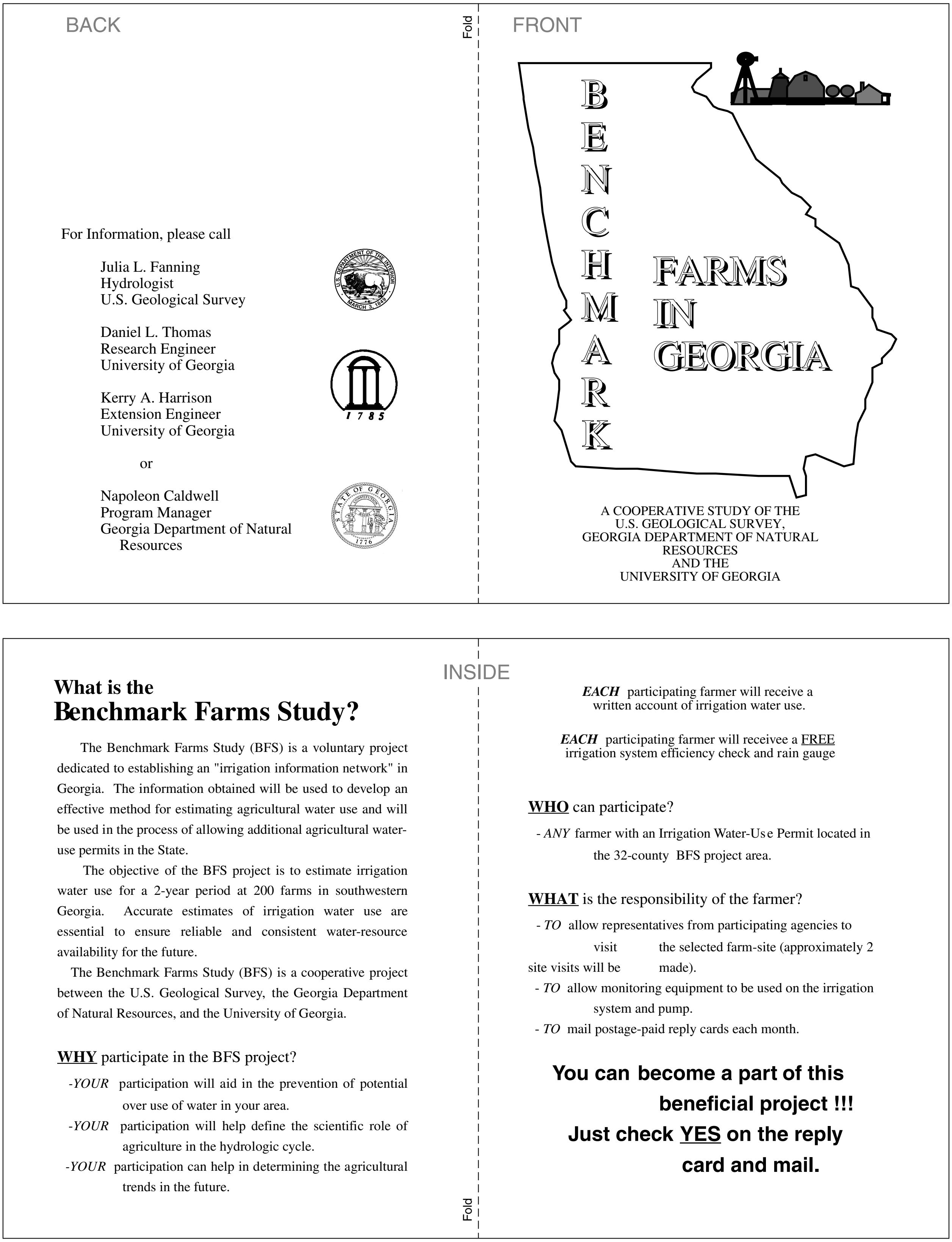
Turlington, M.C., Fanning, J.L., Doonan, G.A., 1987, Water use in Georgia by county for 1985: Georgia Geologic Survey Information Circular 81, 110 p.

U.S. Soil Conservation Service, 1982, Southwest Georgia land and water resource cooperative study: U.S. Department of Agriculture, U.S. Soil Conservation Service, variously paged.

APPENDIX A—Benchmark Farms Study Pamphlet, Cards, and Forms

By Julia L. Fanning

Examples of the pamphlet, cards, and forms used by project personnel during the Benchmark Farms Study are shown in this appendix. The pamphlet, which briefly describes the study, was used to inform and encourage farmer participation (fig. A-1). The reply cards (figs. A-2, A-3) were used to determine the willingness of the farmer to participate, and provided project personnel with source and system information. The field-collection form used during each site visit to record detailed farm descriptions is shown in figure A-4. An example of the postage-paid reply cards completed by farmers with monthly time-totalizer and rain-gage readings is shown in figure A-5. The reminder cards were mailed by project personnel each month (fig. A-6) during the growing season to each participating farmer. At the end of each growing season, an annual report (fig. A-7) was prepared that contained all collected information.



For Information, please call

Julia L. Fanning
Hydrologist
U.S. Geological Survey

Daniel L. Thomas
Research Engineer
University of Georgia

Kerry A. Harrison
Extension Engineer
University of Georgia

or

Napoleon Caldwell
Program Manager
Georgia Department of Natural Resources

BENCHMARK FARMS IN GEORGIA

A COOPERATIVE STUDY OF THE
U.S. GEOLOGICAL SURVEY,
GEORGIA DEPARTMENT OF NATURAL
RESOURCES
AND THE
UNIVERSITY OF GEORGIA

What is the
Benchmark Farms Study?

The Benchmark Farms Study (BFS) is a voluntary project dedicated to establishing an "irrigation information network" in Georgia. The information obtained will be used to develop an effective method for estimating agricultural water use and will be used in the process of allowing additional agricultural water-use permits in the State.

The objective of the BFS project is to estimate irrigation water use for a 2-year period at 200 farms in southwestern Georgia. Accurate estimates of irrigation water use are essential to ensure reliable and consistent water-resource availability for the future.

The Benchmark Farms Study (BFS) is a cooperative project between the U.S. Geological Survey, the Georgia Department of Natural Resources, and the University of Georgia.

<u>WHY</u> participate in the BFS project?

-*YOUR* participation will aid in the prevention of potential over use of water in your area.
-*YOUR* participation will help define the scientific role of agriculture in the hydrologic cycle.
-*YOUR* participation can help in determining the agricultural trends in the future.

EACH participating farmer will receive a written account of irrigation water use.

EACH participating farmer will receivee a <u>FREE</u> irrigation system efficiency check and rain gauge

<u>WHO</u> can participate?

- *ANY* farmer with an Irrigation Water-Use Permit located in the 32-county BFS project area.

<u>WHAT</u> is the responsibility of the farmer?

- *TO* allow representatives from participating agencies to visit the selected farm-site (approximately 2 site visits will be made).
- *TO* allow monitoring equipment to be used on the irrigation system and pump.
- *TO* mail postage-paid reply cards each month.

You can become a part of this beneficial project !!!
Just check <u>YES</u> on the reply card and mail.

Figure A-1. Benchmark Farms Study phamphlet.

FARM IRRIGATION WATER USE PERMIT INFORMATION

EPD LOG NUMBER — A96-160-0000

WELL INFORMATION: **Corrected Information**

Depth **DEPTH** feet
Well casing diameter **DIAMETER** inches ________________
Depth of pump intake below ground surface **DEPTH** feet ________________
Depth of well casing **DEPTH** feet ________________
Design pumping capacity of well pump **CAPACITY** gallons/minute ________________

CROP INFORMATION:

Number of acres irrigated from this water source **ACRES** ________________
Average number of inches of water applied per year **INCHES** ________________
Do you inject chemicals, fertilizers, fungicides, herbicides, insecticides or nematicides into the irrigation water? _____Yes _____ No ________________

- -

Are you interested in participating in the **Georgia Benchmark Farm Study?** _____ Yes _____ No
If **YES** please answer the following questions

Are you ______ **Owner** ______ **Operator**?
Kind of **crop(s)** to be irrigated ____
How many **systems** does the pump supply? (circle one) 1 2 3 4 5 6
What kind of **system(s)** is (are) in use? ____
Power source ______ **Diesel** ______ **Electric**?

Figure A-2. Benchmark Farms Study ground-water source reply card.

FARM IRRIGATION WATER USE PERMIT INFORMATION
EPD LOG NUMBER — A96-160-0000

SURFACE WATER: **Corrected information**

Name of Pond, Lake or Stream: **SOURCE** **BRANCH POND** ________________
Number of pumps withdrawing from this source: **PUMPS** ________________
Total design pumping capacity of pumps: **CAPACITY** gallons/minute ________________

CROP INFORMATION:

Number of acres irrigated from this water source **ACRES** ________________
Average number of inches of water applied per year **INCHES** ________________
Do you inject chemicals, fertilizers, fungicides, herbicides, insecticides or nematicides into the irrigation water? ____ Yes ____ No ________________

Are you interested in participating in the **Georgia Benchmark Farm Study?** _____ Yes _____ No

If **YES** please answer the following questions

Are you ______ **Owner** ______ **Operator**?
Kind of **crop(s)** to be irrigated ____
How many **systems** does the pump supply? (circle one) 1 2 3 4 5 6
What kind of **system(s)** is (are) in use? ____
Power source______ **Diesel** ______**Electric**?

Figure A-3. Benchmark Farms Study surface-water source reply card.

1996

COUNTY: ______________________________

FARM NAME: ______________________________

OWNER/OPERATOR NAME: ______________________________

ADDRESS: ______________________________

CITY/STATE: ______________________________

PHONE: ____________

SOURCE INFORMATION:

WELL: ________ WATER LEVEL: ____ FT.

SURFACE WATER: ____

WELL-TO-POND: ____

TYPE OF PROBLEM(S) ENCOUNTERED (CHECK ALL THAT APPLY):

____ ENGINE (MOTOR) OVERHEATS

____ PRESSURE TOO LOW

____ EXCESSIVE FUEL CONSUMPTION

____ QUESTIONABLE FLOW RATE FROM PUMP

____ OTHER - EXPLAIN ________________

IRRIGATION SYSTEM INFORMATION:

PRIME MOVER: DIESEL ______

ELECTRIC ______

LP GAS ______

OTHER ______

ESTIMATED HORSEPOWER: ______

TYPE OF SYSTEM: CENTER PIVOT ______

CABLE TOW ______

TRAVELER ______

OTHER ______

FLOWMETER CALCULATIONS:

MEASURED FLOW RATE: ______ GPM

PUMP INFORMATION: (VERY NECESSARY)

BRAND NAME: ________________

SERIAL NUMBER: ________________

MODEL NUMBER: ________________

Figure A-4. Benchmark Farms Study field-data-collection form.

Benchmark Farms Study

JANUARY - 1996

NAME: JOHN DOE
EPD LOG NUMBER:__________

RAIN GAGE READINGS:

Date	Amount	Rain or Irrigation
_______	________	________
_______	________	________
_______	________	________
_______	________	________
_______	________	________

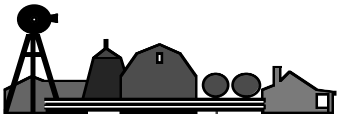
TIME-TOTALIZER READING: _ _ _ _ _ **.** _ _ **Date of reading:**_________
(time-totalizer reading should be taken after the last irrigation for the month)

CROP(S) IRRIGATED: _____________ **ACRES IRRIGATED:**___________
_____________ ___________
_____________ ___________

SIGNATURE: ____________________________________ **Date:**___________

Figure A-5. Benchmark Farms Study monthly readings reply card.

BENCHMARK FARMS STUDY



JUST A REMINDER.....
IT'S TIME TO...
RECORD YOUR TIME-TOTALIZER READINGS
INCLUDE THE CROPS AND ACRES
MAIL AT YOUR EARLIEST CONVENIENCE

THANK-YOU FOR YOUR PARTICIPATION

Figure A-6. Benchmark Farms Study monthly reminder card.

BENCHMARK FARMS STUDY
1996 DATA COLLECTION REPORT

Identification number:
Irrigation Water-Use Permit Number:
Owner/Operator:
Farm/System name:
County:
Address:

Water source: Latitude:
Water level: Ft. Longitude:

Measured flowrate: 500 gallons/minute

<u>Irrigation system:</u>	<u>Pump information:</u>
Type of system:	Brand name:
Power source:	Serial number:
Horsepower:	Model number:

	Pumping time Hours	**Irrigation Inches**	**CROP**
JANUARY			
FEBRUARY			
MARCH			
APRIL			
MAY			
JUNE			
JULY			
AUGUST			
SEPTEMBER			
OCTOBER			
NOVEMBER			
DECEMBER			
TOTAL			

Figure A-7. Annual report of a Benchmark Farms Study site with monthly readings, 1996.

APPENDIX B—Irrigation Water Use in Southwestern Georgia
Description of Sampling Methodology

APPENDIX B—Irrigation Water Use in Southwestern Georgia Description of Sampling Methodology

by Gregory E. Schwarz, U.S. Geological Survey, Reston, Virginia

The methodology used to derive the irrigation water-use sample is described in this appendix. The principal objective of the Benchmark Farms Study was to estimate irrigation water use by permitted irrigators in selected counties in southwestern Georgia. Total irrigation water use was disaggregated into two components representing the respective quantities of water from ground- and surface-water sources. Two features of this objective were important to the study design. First, only aggregated estimates of water use by source were desired. The study was not intended to estimate water use by source on a county-by-county basis. Second, the study was conducted by USGS and CES personnel. This made it desirable to sample irrigators proportionally at the county level for each source type.

Sampling was proportional at the county/source-type level of aggregation and the efficiency of the study was improved by stratifying water users according to some variable that was likely to correlate with water-use quantity. The agricultural permit application contained information useful for this purpose, including the permit holder's reported number of acres to be irrigated, an estimate of average inches applied, and the capacity of the irrigation pump.

Factors that affected the choice of a stratifying variable include the quality of the data and number of missing observations. Although pump capacity had the fewest missing values, the pump capacity poorly correlated with irrigated acres and with average quantity of water applied (irrigated acres multiplied by average inches applied). Quantity of water applied had the most missing values (approximately 10 percent of the population) and included many questionable values (average annual application exceeding 100 inches). Acres irrigated had about 6-percent missing values (most were due to matching problems described below), and was considered by permit writers to be fairly reliable; acres irrigated correlated reasonably with quantity of water applied for those observations having application rates below100 inches per year (the correlation coefficient was about 0.5). Although no variable clearly dominated all of the criteria, acres irrigated was chosen as the stratifying variable because it correlated highly with irrigation water use and contained relatively few missing values.

The population of irrigation permit holders selected for analysis consisted of all agricultural permit holders, located in one of the 32 selected counties, identified as withdrawing water for irrigation (those withdrawing water for aquaculture were omitted), and having ground or surface water as the source. Permits indicating that the water flows from a ground-water source to a holding pond were omitted, as were any permits showing non-zero values for both ground- and surface-water pump capacity. In forming the final population list, a few permits appeared more than once. It was assumed that these permits pertained to erroneous observations where multiple sources were placed on the same permit number. These multiple sources were retained in the final list as separate observations, resulting in a total population of 9,840 permits. Finally, the population list contained 396 observations corresponding to permit numbers that could not be matched with the permit report file—these permits were missing important information such as name and address, source of water, pump capacity, and acres irrigated. These incomplete observations were combined with observations having missing values for acres to form a separate stratum within the population.

The first step in the sample selection process was to determine the number of samples to be taken from each county/source-type group. As explained above, this allocation was based on the proportion of permits within the county/source-type group relative to the total number of permits. The proportional allocation is given by the formula

$$T_{ij}^{TOT} = \left(\left(\frac{N_{ij}^{TOT}}{N^{TOT}}\right)\cdot 500\right) \qquad (1)$$

where T_{ij}^{TOT} is the total number of samples drawn from county i with source-type j (where $j = 1$ or 2, and corresponds to a ground- or surface-water source, respectively), N_{ij}^{TOT} is the population of irrigation permits in county i with source-type j, N^{TOT} is the total number of permits in the population, and 500 is the total number of samples to be taken. The allocation given was rounded to the nearest integer, totaled over all the county/source-type groups, and then compared to the desired total of 500 samples. Any discrepancy from the desired 500 samples (due to rounding error) was corrected by adding to or subtracting from those allocations having the largest or smallest (most negative) rounding errors.

The observations then were organized into strata. All “missing-value” observations were placed in a separate stratum that was labeled stratum 0. The remaining “non-missing-value” observations were delineated by the number of irrigated acres into four additional strata. Separate strata delineations were applied to ground- and surface-water users, although the delineation boundaries were the same in all of the selected counties.

To determine the optimal stratum delineations, a technique described by Cochran (1977, p. 127-131) was adopted. First, the non-missing observations were organized into 100 bins of equal size, where the bin size is given by:

$$\mathit{Bin\ size} = \begin{array}{c}\dfrac{\mathit{Maximum\ Value\ for\ Acres} -}{100}\\ \dfrac{\mathit{Minimum\ Value\ for\ Acres}}{100}\end{array} \qquad (2)$$

The population frequency counts, by source type, were determined for each bin. For each source type, the square root of these frequency counts was summed and divided by four—the number of strata to be identified. The bins were delineated into four groups (corresponding to the four strata), such that the sum of the squares of the frequency counts for each group were approximately equal. The frequency distributions for irrigated acres for ground- and surface-water users are shown in figures B-1 and B-2, respectively.

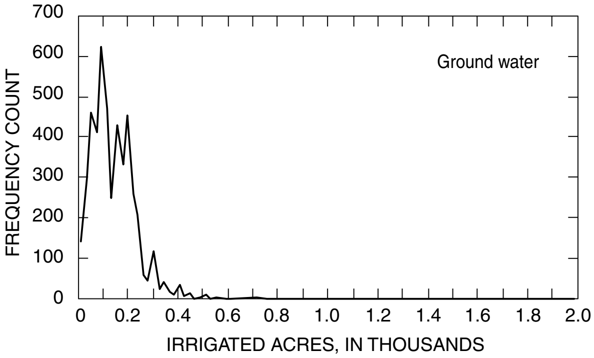


Figure B-1. Frequency distribution for irrigated acres from a ground-water source.

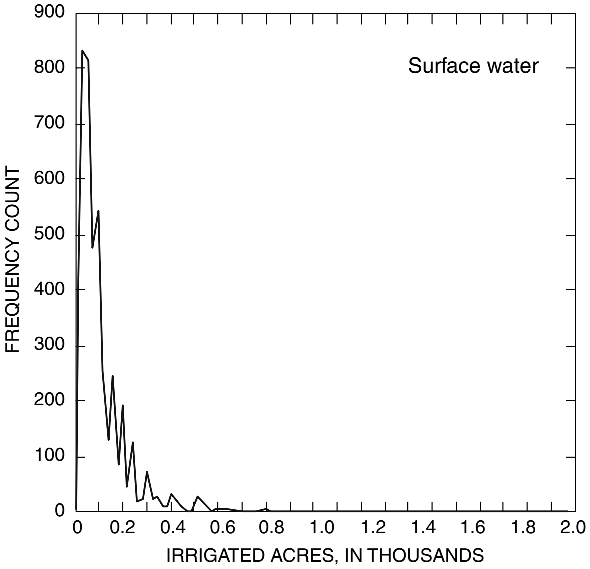


Figure B-2. Frequency distribution for irrigated acres from a surface-water source.

This procedure allowed the user to assign each bin, and consequently, each “non-missing-value” observation, to one of the four strata. The delineations for ground- and surface-water users are listed in table B-1.

Table B-1. Strata delineations for ground- and surface-water users
[—, not applicable; do., ditto]

Source type	Strata	Lower bound	Upper bound
Ground water	1	0.00	84.04
Do.	2	84.04	146.32
Do.	3	146.32	229.36
Do.	4	229.36	—
Surface water	1	0.00	42.52
Do.	2	42.52	104.80
Do.	3	104.80	229.36
Do.	4	229.36	—

The next step was to determine the sampling rates for each of the strata. For the "missing-value" stratum, sampling was proportional, implying an allocation rule:

$$T_{oij} = \left(\frac{N_{oij}}{N_{ij}^{TOT}} \right) T_{ij}^{TOT} \, , \qquad (1)$$

where T_{0ij} is the number of samples in stratum 0 for county *i* with source-type *j*, N_{0ij} is the population of irrigation permits in stratum 0 for county *i* with source-type *j*, and N_{ij}^{TOT} and T_{ij}^{TOT} are defined above.

The sampling rates for the four "non-missing-value" strata were based on an optimal sampling rule that allocated samples according to each stratum's size and variance of irrigated acres. The allocation formula was discussed in Cochran (1977, p. 96-99) and is given by:

$$T_{hij} = \left| \frac{N_{hj} V_{hj}^{1/2}}{\sum_{h=1}^{4} d_{hij} N_{hj} V_{ij}^{1/2}} \right| \left(T_{ij}^{TOT} - T_{oij} \right) \, , \qquad (2)$$

where T_{hij} is the number of samples drawn from "non-missing-value" stratum *h* in county *i* with source-type *j*, N_{hj} is the sum of the number of permitted irrigators over all 32 counties in "non-missing-value" strata *h* with source *j*, V_{hj} is the variance of irrigated acres (corrected for degrees of freedom) in "non-missing-value" strata *h* with source type *j*, and d_{hij} is a 0,1 dummy variable equal to 0 if there are no permits in stratum *h* of county *i* with source type *j* or 1 otherwise. The second term of equation (4) corresponds to the number of "non-missing-value" samples taken in county *i* with source-type *j*. The first term of equation (4) represents a Neyman allocation of the county/source-type apportioned "non-missing-value" samples. The Neyman allocation is given by the standard error of irrigated acres for a given stratum, multiplied by the stratum's population, divided by the sum of population multiplied by standard errors for all strata represented in county *i* with source type *j*.

Since the Neyman allocation is determined by variances and shares evaluated at the aggregate level (that is, sums over counties), the allocation generally will not be optimal at the county level. Furthermore, in a given stratum/county/source-type group, the allocation formula may tend to allocate more samples than there are permits. In this case, all of the permits were sampled in that group, and the remaining samples in the county/source-type group were allocated according to the share of population multiplied by the standard error for the remaining strata. Because of these types of constraints, the aggregate sample may have deviated somewhat from the optimum. Some statistics on sampling proportions that show stratum constraints were not major problems for this sample.

As a final adjustment to the allocation formula in equations (3) and (4), it was necessary to compensate for rounding errors that resulted in a difference between the number of selected samples in a given county/source-type group and the desired level given by T_{ij}^{TOT}. The resulting sample was approximately optimal at the aggregate level and proportional at the county/source-type level.

Summary information of the selected sample is shown in tables B-2 and B-3. The share of samples selected from each county/source-type group compared with the population shares is shown in table B-2. The sample and population shares by strata for each source type are shown in table B-3. The first column in table B-3 shows the source-type group and the second column indicates the stratum number. The third column shows the aggregate shares of "non-missing-value" samples allocated to the various strata, if there were no rounding or proportional sampling constraints. Thus, this column gives the shares determined by the variances and relative number of the population within each stratum. The fourth column shows the share of samples selected by stratum. Stratum 0 in table B-3 refers to the "missing-value" stratum. Shares shown for this stratum were determined relative to the total number of samples selected for that source type. Shares reported in the remaining rows are the number of samples in a "non-missing-value" stratum relative to the total number of "non-missing-value" samples selected for a given source type. Thus the sum of the shares for strata 1 to 4 should equal one (except for rounding error). The last column of table B-3 gives the population shares. Again, the shares given in the stratum 0 rows are relative to the total population of that source type; whereas, shares in the remaining rows are relative to the total population in the "non-missing-value" strata for the relevant source type.

The close agreement between the third and fourth columns is evident from table B-3, indicating that the constraint that sampling be proportional at the county/source-type level did not greatly inhibit the ability to reach desired sampling rates across the strata at the aggregate level. Thus, aggregate estimates derived from the sample had variances that were approximately minimized for the specified sample size.

Table B-2. Number of permits in the population and sample, by county, source type, and strata
[do., ditto]

County number	Source type	Strata	Number of permits in the population	Number of permits in the sample
4	ground water	0	22	1
4	do.	1	54	3
4	do.	2	108	2
4	do.	3	84	3
4	do.	4	42	7
4	surface water	0	2	0
4	do.	1	7	0
4	do.	2	21	1
4	do.	3	8	1
4	do.	4	14	1
9	ground water	0	39	2
9	do.	1	10	0
9	do.	2	7	0
9	do.	3	5	0
9	do.	4	4	1
9	surface water	0	98	5
9	do.	1	24	0
9	do.	2	20	1
9	do.	3	9	1
9	do.	4	2	1
12	ground water	1	18	1
12	do.	2	25	0
12	do.	3	24	1
12	do.	4	9	2
12	surface water	1	15	1
12	do.	2	59	1
12	do.	3	31	1
12	do.	4	7	3
14	ground water	0	3	0
14	do.	1	19	1
14	do.	2	42	1
14	do.	3	23	1
14	do.	4	12	2
14	surface water	0	2	0
14	do.	1	14	0
14	do.	2	17	1
14	do.	3	16	1
14	do.	4	4	1
19	ground water	0	32	2
19	do.	1	26	1
19	do.	2	14	0
19	do.	3	21	1
19	do.	4	7	1
19	surface water	0	44	2
19	do.	1	2	0
19	do.	2	25	1
19	do.	3	21	1
19	do.	4	21	2

Table B-2. Number of permits in the population and sample, by county, source type, and strata—Continued
[do., ditto]

County number	Source type	Strata	Number of permits in the population	Number of permits in the sample
30	ground water	1	1	0
30	do.	2	3	0
30	do.	3	2	0
30	do.	4	6	1
30	surface water	1	1	0
30	do.	2	12	1
30	do.	3	20	0
30	do.	4	13	1
35	ground water	0	2	0
35	do.	1	31	1
35	do.	2	26	1
35	do.	3	20	1
35	do.	4	11	2
35	surface water	0	13	1
35	do.	1	198	2
35	do.	2	197	6
35	do.	3	41	5
35	do.	4	18	10
37	ground water	0	3	0
37	do.	1	51	1
37	do.	2	20	1
37	do.	3	23	1
37	do.	4	9	2
37	surface water	0	9	0
37	do.	1	103	1
37	do.	2	64	3
37	do.	3	23	2
37	do.	4	9	5
40	ground water	0	16	1
40	do.	1	16	1
40	do.	2	11	0
40	do.	3	24	1
40	do.	4	5	1
40	surface water	0	40	2
40	do.	1	31	1
40	do.	2	44	1
40	do.	3	36	1
40	do.	4	12	3
43	ground water	0	7	0
43	do.	1	66	4
43	do.	2	108	3
43	do.	3	175	4
43	do.	4	37	9
43	surface water	0	1	0
43	do.	1	8	0
43	do.	2	18	1
43	do.	3	20	1
43	do.	4	8	1
45	ground water	0	9	0
45	do.	1	41	1
45	do.	2	30	1
45	do.	3	23	1
45	do.	4	12	3

Table B-2. Number of permits in the population and sample, by county, source type, and strata—Continued
[do., ditto]

County number	Source type	Strata	Number of permits in the population	Number of permits in the sample
45	surface water	0	12	1
45	do.	1	64	1
45	do.	2	86	2
45	do.	3	42	2
45	do.	4	8	5
46	ground water	0	7	0
46	do.	1	59	2
46	do.	2	31	1
46	do.	3	44	2
46	do.	4	34	4
46	surface water	0	6	0
46	do.	1	12	1
46	do.	2	38	1
46	do.	3	24	1
46	do.	4	19	2
47	ground water	0	6	0
47	do.	1	28	1
47	do.	2	31	1
47	do.	3	35	1
47	do.	4	18	3
47	surface water	1	2	0
47	do.	2	3	0
47	do.	4	1	0
49	ground water	0	9	0
49	do.	1	53	3
49	do.	2	96	2
49	do.	3	86	3
49	do.	4	34	6
49	surface water	0	1	0
49	do.	1	5	1
49	do.	2	36	1
49	do.	3	38	1
49	do.	4	16	2
65	ground water	0	13	1
65	do.	1	29	0
65	do.	2	6	0
65	do.	3	4	1
65	do.	4	5	1
65	surface water	0	34	2
65	do.	1	20	0
65	do.	2	36	1
65	do.	3	17	1
65	do.	4	5	2
76	ground water	0	8	0
76	do.	1	4	0
76	do.	2	2	0
76	do.	3	4	0
76	do.	4	7	1
76	surface water	0	2	0
76	do.	2	1	0
76	do.	3	1	0
76	do.	4	2	0
77	ground water	0	4	0
77	do.	1	47	1
77	do.	2	38	1
77	do.	3	28	1
77	do.	4	11	3

Table B-2. Number of permits in the population and sample, by county, source type, and strata—Continued
[do., ditto]

County number	Source type	Strata	Number of permits in the population	Number of permits in the sample
77	surface water	0	6	0
77	do.	1	204	3
77	do.	2	285	7
77	do.	3	100	7
77	do.	4	18	14
88	ground water	0	13	1
88	do.	1	54	2
88	do.	2	58	2
88	do.	3	79	2
88	do.	4	34	5
88	surface water	0	3	0
88	do.	1	3	1
88	do.	2	33	1
88	do.	3	34	1
88	do.	4	23	2
94	ground water	0	7	0
94	do.	1	31	1
94	do.	2	19	1
94	do.	3	22	1
94	do.	4	23	2
94	surface water	1	8	0
94	do.	2	14	1
94	do.	3	23	1
94	do.	4	16	1
100	ground water	0	7	0
100	do.	1	120	4
100	do.	2	168	4
100	do.	3	154	6
100	do.	4	42	11
100	surface water	1	1	0
100	do.	2	3	0
100	do.	3	4	0
101	ground water	0	13	1
101	do.	1	124	4
101	do.	2	110	3
101	do.	3	186	5
101	do.	4	49	11
101	surface water	0	4	0
101	do.	1	34	1
101	do.	2	56	1
101	do.	3	26	1
101	do.	4	8	3
116	ground water	0	30	1
116	do.	1	14	1
116	do.	2	34	1
116	do.	3	26	1
116	do.	4	19	2
116	surface water	0	23	1
116	do.	1	13	1
116	do.	2	51	2
116	do.	3	51	1
116	do.	4	25	3
120	ground water	1	3	1
120	do.	2	20	0
120	do.	3	29	1
120	do.	4	18	2

Table B-2. Number of permits in the population and sample, by county, source type, and strata—Continued
[do., ditto]

County number	Source type	Strata	Number of permits in the population	Number of permits in the sample
120	surface water	1	22	1
120	do.	2	54	2
120	do.	3	70	2
120	do.	4	29	4
123	ground water	3	1	0
123	do.	4	1	0
123	surface water	0	1	0
123	do.	1	2	0
123	do.	2	9	0
123	do.	3	9	0
123	do.	4	6	1
125	ground water	0	29	1
125	do.	1	76	3
125	do.	2	94	3
125	do.	3	156	4
125	do.	4	35	9
125	surface water	0	4	0
125	do.	1	4	0
125	do.	2	9	0
125	do.	3	4	0
125	do.	4	1	1
129	ground water	0	11	1
129	do.	1	31	1
129	do.	2	30	1
129	do.	3	46	2
129	do.	4	53	4
129	surface water	0	1	0
129	do.	1	9	1
129	do.	2	28	2
129	do.	3	51	1
129	do.	4	48	3
135	ground water	0	4	0
135	do.	1	21	1
135	do.	2	26	0
135	do.	3	17	1
135	do.	4	11	2
135	surface water	0	4	0
135	do.	1	10	1
135	do.	2	46	1
135	do.	3	27	1
135	do.	4	19	2
136	ground water	0	8	0
136	do.	1	42	1
136	do.	2	21	1
136	do.	3	21	1
136	do.	4	11	2
136	surface water	0	2	0
136	do.	1	21	0
136	do.	2	23	1
136	do.	3	16	1
136	do.	4	11	2
137	ground water	0	26	1
137	do.	1	128	2
137	do.	2	55	2
137	do.	3	12	2
137	do.	4	12	5

Table B-2. Number of permits in the population and sample, by county, source type, and strata—Continued
[do., ditto]

County number	Source type	Strata	Number of permits in the population	Number of permits in the sample
137	surface water	0	10	0
137	do.	1	219	3
137	do.	2	166	7
137	do.	3	61	7
137	do.	4	6	6
142	ground water	0	10	0
142	do.	1	31	1
142	do.	2	25	1
142	do.	3	23	1
142	do.	4	7	2
142	surface water	0	7	0
142	do.	1	77	1
142	do.	2	177	4
142	do.	3	49	4
142	do.	4	8	7
156	ground water	0	9	0
156	do.	1	61	2
156	do.	2	39	1
156	do.	3	46	2
156	do.	4	16	4
156	surface water	0	12	1
156	do.	1	82	1
156	do.	2	151	4
156	do.	3	52	3
156	do.	4	11	7
159	ground water	0	4	0
159	do.	1	23	1
159	do.	2	44	1
159	do.	3	26	1
159	do.	4	12	3
159	surface water	0	1	0
159	do.	1	31	1
159	do.	2	48	1
159	do.	3	25	1
159	do.	4	3	2

Table B-3. Comparison among desired, sample, and population shares, by strata and source type
[—, not applicable; SHARET, computer program used to show share of samples selected by stratum; SHAREN, computer program used to show population shares; do., ditto]

Source type	Strata	Sample share	SHARET	SHAREN
Ground water	0	—	0.05019	0.06926
Do.	1	0.17737	0.18699	0.27738
Do.	2	0.14235	0.14228	0.28351
Do.	3	0.22700	0.21138	0.31057
Do.	4	0.45328	0.45935	0.12854
Surface water	0	—	0.06224	0.07188
Do.	1	0.09137	0.10177	0.28216
Do.	2	0.24259	0.24779	0.41440
Do.	3	0.21744	0.22124	0.21490
Do.	4	0.44860	0.42920	0.08854

APPENDIX C—1995 Flowmeter-Test Description and Brief Analysis

Appendix C—1995 Flowmeter-Test Description and Brief Analysis

by Daniel L. Thomas and Kerry A. Harrison, The University Georgia,
College of Agriculture and Environmental Sciences

This appendix describes the results of a flowmeter test conducted at the Rainbow Irrigation Facilities in Fitzgerald, Ga., in March 1995. Two flowmeters—the ISTT-P (Peek Measurements) and the DHT-P (Polysonic)—were tested. Full lengths of individual irrigation pipe (planned for use in field center pivot systems) were set in series and the flow rate adjusted. The measured flow rate was based on outlet pressure of 1- to 2-inch-diameter irrigation guns of specific diameter (Rainbow Irrigation Facilities, Fitzgerald, Ga., written commun., 1982). The flow rate determined the need for one or two guns. The original data and values representing the measured and estimated flows for the different meters are listed in table C-1.

The flowmeters used to measure flow rate for irrigation pumping plants were capable of reproducing flow rates to within ± 0.5 percent of the velocity for the ISTT-P, Portable Transit Time Ultrasonic Flowmeter based on manufacturer specifications (Polysonics, written commun., 1995). Accuracy characteristics and repeatability for both the ISTT-P and the DHT-P, however, needed to be sufficiently compatible for the meters to be used interchangeably at any of the test sites. The ability for both flowmeters to perform reliably and within the specifications of the manufacturers is required since these results would be extrapolated over a large area. Both flowmeters used in the study to estimate flow rates in the irrigation pipes were non-invasive; that is, they are designed to measure flow rates without being placed into the pipe.

The ISTT-P flowmeter underestimated the flow rate in the 6-inch-diameter pivot pipe (average percent difference of 4.8 from measured). For the 6-5/8 inch-diameter, the ISTT-P underestimated the flow rate at the low-flow rates and overestimated the flow rate at the high rates (average 9.9 percent difference). For the 8-inch-diameter pipe, the ISTT-P performed satisfactorily throughout the range, with some slight overestimation at the last flow rate tested of 8.7 feet per second (average 2.1 percent difference). On an overall comparison basis, the error in measurement was within 5.6 percent for all pipes.

The DHT-P flowmeter performed well on the 6-inch-diameter pipe (average 3.9 percent difference). Some deviation in measurements were evident on the 6 5/8-inch pipe at the high-flow rates, which resulted in less accuracy overall (average 4.8 percent difference). On the 8-inch-diameter pipe, the DHT-P underestimated the flow rate at the low rates, but was more accurate at the high rates (average 6.9 percent difference). Overall, the DHT-P performed within 5.2 percent of the measured flow rate for all pipes.

Table C-1. Flowmeter test at Rainbow Irrigation Facilities, Fitzgerald, Georgia, March 22, 1995
[ISTTP, Peak Measurements; DHT-P, Polysonic; USGS, U.S. Geological Survey; —, not measured]

Estimated flow (gallons per minute)	Nozzle number	Used size (inches)	Pressure (pounds per square inch)	Expected (feet per second)	Velocity for pipe size of 6 inches (5.78)				Percent difference	
					Pitot tube (feet per second)	ISTT-P (feet per second)	DHT-P, USGS (feet per second)	Calibration Coefficient (percent)	ISTT-P versus expected	DHT-P versus expected
400	1	1-3/4	32	4.89	—	4.5	5.4	5.6	7.98	10.43
500	1	1-3/4	51	6.11	—	5.9	6.6	5.57	3.44	8.02
600	1	1-3/4	73	7.33	—	6.8	7.6	5.36	7.23	3.68
700	1	1-3/4	99	8.55	—	7.3	8.8	5.27	14.62	2.92
800	2	1-3/4	32	9.78	—	9.5	10	—	2.86	2.25
900	2	1-3/4	41	11	—	10.8	11	—	1.82	0.00
1,000	2	1-3/4	51	12.22	—	12	12.1	—	1.80	.98
1,100	2	1-3/4	61	13.44	—	12.6	13.6	—	6.25	1.19
1,200	2	1-3/4	73	14.66	—	14.6	14.2	4.89	0.41	3.14
1,300	2	1-3/4	85	15.88	—	15.6	14.8	4.68	1.76	6.80
1,330	2	1-3/4	90	—	—	16.2	—	—	—	—
							Average	**5.34**	**4.82**	**3.94**

Estimated flow (gallons per minute)	Nozzle number	Used size (inches)	Pressure (pounds per square inch)	Expected (feet per second)	Velocity for pipe size of 6-5/8 inches (6.41)				Percent difference	
					Pitot tube (feet per second)	ISTT-P (feet per second)	DHT-P USGS (feet per second)	Calibration Coefficient (percent)	ISTT-P versus expected	DHT-P versus expected
400	1	1-3/4	32	3.97	3.5	2.5	4	—	37.03	0.76
500	1	1-3/4	51	4.97	4.2	4.1	5	—	17.51	.60
600	1	1-3/4	73	5.96	5.4	4.8	6.4	5.45	19.46	7.38
700	1	1-3/4	99	6.96	6.7	5.9	6.7	4.94	15.23	3.74
800	2	1-3/4	32	7.95	7.2	7.8	8	5	1.89	.63
900	2	1-3/4	41	8.94	8.5	9	9	5	.67	.67
1,000	2	1-3/4	51	9.94	9.5	10	10	5	.60	.60
1,100	2	1-3/4	61	10.93	10.7	10.9	10.8	5.03	.27	1.19
1,200	2	1-3/4	73	11.92	11.7	12.4	10.2	4.2	4.03	14.43
1,300	2	1-3/4	85	12.92	12.2	13.7	11.4	4.38	6.04	11.76
1,330	2	1-3/4	90	13.22	12.4	14	11.8	4.42	5.90	10.74
							Average	**4.83**	**9.88**	**4.77**

Estimated flow (gallons per minute)	Nozzle number	Used size (inches)	Pressure (pounds per square inch)	Expected (feet per second)	Velocity for pipe size of 8 inches (7.78)				Percent difference	
					Pitot tube (feet per second)	ISTT-P (feet per second)	DHT-P, USGS (feet per second)	Calibration Coefficient (percent)	ISTT-P versus expected	DHT-P versus expected
400	1	1-3/4	32	2.7	—	—	—	—	—	—
500	1	1-3/4	51	3.37	—	—	—	—	—	—
600	1	1-3/4	73	4.05	—	—	—	—	—	—
700	1	1-3/4	99	4.72	—	—	—	—	—	—
800	2	1-3/4	32	5.4	4.7	5.4	5.2	4.34	0.00	3.70
900	2	1-3/4	41	6.07	5.5	6.2	5.2	4.15	2.14	14.33
1,000	2	1-3/4	51	6.74	6.2	6.8	6	4.2	.89	10.98
1,100	2	1-3/4	61	7.42	6.8	7.5	6.8	4.51	1.08	8.36
1,200	2	1-3/4	73	8.09	7.2	8.3	7.8	4.72	2.60	3.58
1,300	2	1-3/4	85	8.77		9.3	8.8		6.04	.34
							Average	**4.43**	**2.12**	**6.88**

APPENDIX D—Statistical Analysis of Irrigation Water-Use Data in Southwestern Georgia

Appendix D—Statistical Analysis of Irrigation Water-Use Data in southwestern Georgia

by Gregory E. Schwarz, U.S. Geological Survey, Reston, Virginia

Model Calibration

Model calibration proceeded in five steps corresponding to the five model components. All model components were estimated using Statistical Analysis System/Econometric Time Series (SAS/ETS) software. Two versions of the results are presented for each model component. Under the first version of results, parameter estimates were computed for an exploratory model which included all the considered predictors. The second version reports results for a final model containing a subset of the predictors determined to be the most robust. The exploratory model estimates were obtained from the original Benchmark Farms Study data; whereas, the final model estimates were generated using the bootstrap algorithm with 200 randomly-selected pseudo-samples. The bootstrap estimate of a parameter was the mean of the 200 estimates generated from the 200 pseudo samples and the bootstrap standard deviation was the standard deviation of the 200 estimates. The bootstrap 90-percent minimum confidence interval is defined as the minimum interval containing 90 percent of the 200 parameter estimates.

The first model component determined whether or not the permitted site actively irrigated at some time during the current growing season. This model component was specified as a logit model which related the probability of a permitted site irrigating to a number of predictor variables (Maddala, 1983, p. 22-27). The assumed relation is:

$$P_{\mathrm{irr}} = \left(\frac{\exp(\mathbf{a}'_{\mathrm{irr}} \boldsymbol{Z}_{\mathrm{irr}})}{1 + \exp(\mathbf{a}'_{\mathrm{irr}} \boldsymbol{Z}_{\mathrm{irr}})} \right), \qquad (1)$$

where P_{irr} is the probability the permitted site irrigates at some time during the growing season, $\mathbf{a}_{\mathrm{irr}}$ is a vector of coefficients, $\boldsymbol{Z}_{\mathrm{irr}}$ is a vector of predictor variables affecting the decision whether or not to irrigate.

Because this model component is at the top of the recursive structure, only predictors that were available for every permit in the study area were considered. The $\boldsymbol{Z}_{\mathrm{irr}}$ variables that were tried included an intercept, the natural logarithm of the permit-reported irrigated acres, the natural logarithm of the permit-reported pump capacity, a 0/1 variable indicating if the permit was located in a western county of the study area, a 0/1 variable indicating if the permit was located in a southern county, and a 0/1 variable indicating if ground water was used to supply the permit's irrigation.

The logit model was estimated using maximum likelihood methods; results are reported in table D-1. The results of the preliminary model show that none of the candidate predictor variables were successful in predicting variations in the probability that the permitted site irrigates during the growing season. Apparently, the decision to irrigate was not related to the size or amount of invested irrigation capital or location of the farm but is likely a result of crop selection that follows a fixed rotation strategy.

Table D-1. Calibration results for the whether-to-irrigate model
[—, not applicable]

Coefficient	Exploratory model (estimates based on original sample)			Final model (estimates based on 200 bootstrap Pseudo-samples)			
						90-percent confidence interval	
	Estimate	Standard error	Significance	Estimate	Standard error	Lower bound	Upper bound
Intercept	2.614	2.537	0.305	2.065	0.237	1.689	2.390
Log of pump capacity (permit)	–0.239	0.472	0.613	—	—	—	—
Log of irrigated acres (permit)	0.141	0.430	0.743	—	—	—	—
If west	–0.336	0.595	0.573	—	—	—	—
If south	0.336	0.607	0.581	—	—	—	—
If ground water	0.606	0.678	0.373	—	—	—	—
Pseudo R^2	2.015	—	—	—	—	—	—
Number of observations	165	—	—	167	—	—	—

The only variable appearing in the final model was the intercept, which was computed such that substituting its value into (1) results in a probability that equals the share of permits in the sample reporting irrigation at some time during the year. That share is 0.89. The intercept's bootstrap 90-percent minimum confidence interval implied a confidence interval for the share of 0.84 to 0.92.

The second model component was the acres-irrigated model. This model related the amount of total irrigated acreage for the permit, as reported by the Benchmark Farms Study site, to the irrigated acreage recorded in the agricultural permit application. The form of the model is a log-linear regression

$$1n(A) = \mathbf{a}'_{\mathrm{acre}}\mathbf{Z}_{\mathrm{acre}} + e_{\mathrm{acre}'} \qquad (2)$$

where **l**n is the natural logarithm function, A is the total irrigated acres for the permit as reported in the Benchmark Farms Study site, $\mathbf{a}_{\mathrm{acre}}$ is a vector of coefficients, $\mathbf{Z}_{\mathrm{acre}}$ is a vector of predictor variables, and e_{acre} is a random error term assumed to be independent and identically distributed across permits.

The acres-irrigated model was estimated only for those permitted sites that actively irrigate. The method of estimation is least squares. Predictor variables $\mathbf{Z}_{\mathrm{acre}}$ used in the regression were the same as those tried in the whether-to-irrigate model and included an intercept, the natural logarithm of the permit-reported irrigated acres, the natural logarithm of the permit-reported pump capacity, a 0/1 variable indicating if the permit is located in a western county of the study area, a 0/1 variable indicating if the permit is located in a southern county, and a 0/1 variable indicating if groundwater is used to supply irrigation to the permitted site.

The results of the acres-irrigated model are reported in table D-2 and a graph of the residuals plotted against the predicted value is shown in figure D-1. The results of the preliminary model show that all the predictor variables except the southern county and ground-water indicators are significant at the 5-percent level or better. Although insignificant, the southern county indicator did show some explanatory power at the 15-percent level and so was retained in the final model. The bootstrap estimates of the parameters reported for the final model were very similar to the conventional estimates. The estimates show that 98.5 percent of the variation in the natural logarithm of Benchmark Farms Study sites-reported acres can be explained by the model component. Although a significant predictor, the permit application-recorded acreage has a coefficient that is significantly lower than one implying the application information was either systematically biased or subject to reporting error. The significant positive coefficients for pump capacity and the western county indicator could arise if these variables were positively correlated with permit application-reported irrigated acreage and application-reported acreage is subject to measurement error. Regardless of the bias associated with the coefficients, the prediction of the natural logarithm of the acres-irrigated model was unbiased and had the lowest standard error of any other conceivable linear prediction based on the same available information.

Table D-2. Calibration results for the acres-irrigated model
[—, not applicable]

Coefficient	Exploratory model (estimates based on original sample)			Final model (estimates based on 200 bootstrap Pseudo-samples			
						90-percent confidence interval	
	Estimate	Standard error	Significance	Estimate	Standard error	Lower bound	Upper bound
Intercept	1.231	0.480	0.011	1.228	.554	0.347	2.034
Log of pump capacity (permit)	0.159	0.077	0.040	0.165	0.085	0.006	0.294
Log of irrigated acres (permit)	0.424	0.066	<0.001	0.418	0.096	0.280	0.596
lf west	0.457	0.113	<0.001	0.481	0.119	0.300	0.680
If south	–0.162	0.106	0.130	-0.164	0.109	-0.342	0.003
If ground water	0.027	0.122	0.825	—	—	—	—
MSE	0.357	—	—	—	—	—	—
R^2	0.985	—	—	—	—	—	—
Number of observations	146	—	146	—	—	—	—

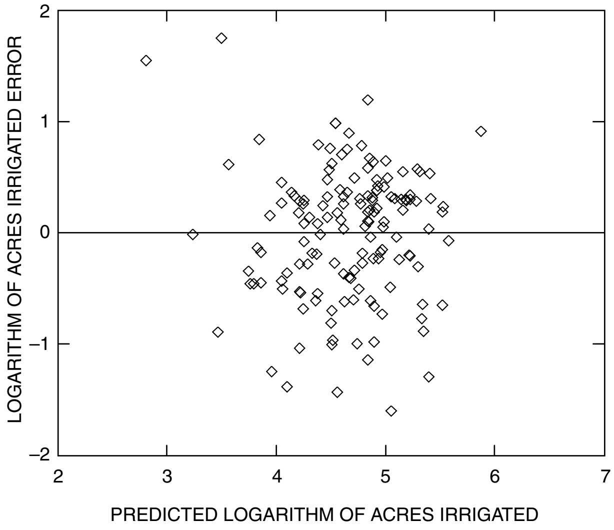


Figure D-1. Natural logarithm of error from the acres-irrigated model relative to the predicted natural logarithm of acres irrigated in the Benchmark Farms Study area, southwestern Georgia.

The third model component was the crop-selection model. This model component determines the probability that a given permitted site will grow a particular crop. The list of crops the model explicitly models was determined by the crops most often grown as reported in the data collected at the study sites. The shares of crops grown among the Benchmark Farms Study sites are given in table D-3. The five principal crops grown according to the Benchmark Farms Study sites data are cotton, corn, peanuts, pecans, and turf grass. These crops—together with a class called "other" consisting of all the other crops—were modeled using a multinomial logit model for unordered variables (Maddala, 1983, p. 34-37).

Table D-3. Number and frequency of crops grown for 148 permits in the Benchmark Farms Study area, southwestern Georgia

Crop	Number of permits in samples	Percent of sample
Cotton	45	30.4
Corn	28	18.9
Peanuts	27	18.2
Pecans	18	12.2
Turf grass	12	8.1
Other:	18	12.2
Vegetables	*10*	*6.8*
Nurseries	*3*	*2.0*
Melons	*2*	*1.4*
Tobacco	*2*	*1.4*
Sorghum	*1*	*0.7*

The third model component relates the probability Pj of selecting crop j (out of a total of J crops) to a vector of explanatory variables $\mathbf{Z}_{\text{crop}}$ according to:

$$P_j = \begin{cases} \dfrac{\exp\left(\mathbf{a}'_{\text{crop},j}\,\mathbf{Z}_{\text{crop}}\right)}{1+\sum\limits_{k=1}^{J-1}\exp\left(\mathbf{a}'_{\text{crop},k}\,\mathbf{Z}_{\text{crop}}\right)} & (j=1,\ldots,J-1) \\[2ex] \dfrac{1}{1+\sum\limits_{k=1}^{J-1}\exp\left(\mathbf{a}'_{\text{crop},k}\,\mathbf{Z}_{\text{crop}}\right)} & (j=J) \end{cases} \qquad (3)$$

where $\mathbf{a}_{\text{crop},j}$ is a vector of crop-specific coefficients. Note that the specification designates one of the crops, crop J, to serve as a residual choice, thereby insuring the sum of the probabilities is one. Consequently, the residual choice crop has no crop-specific coefficients. Here, turf grass was chosen to be the residual. Additionally, the specification of the crop specific coefficients was chosen such that each crop had a common set of coefficients and crop-specific *differential* coefficients so that

$$\mathbf{a}_{\text{crop},\,j} = \begin{cases} \mathbf{a}_{\text{crop},\,0} + \tilde{\mathbf{a}}_{\text{crop},\,j} & (j = 1,\ldots,J-2) \\ \mathbf{a}_{\text{crop},\,0} - \sum_{k=1}^{J-2} \tilde{\mathbf{a}}_{\text{crop},\,k} & (j = J-1) \end{cases} \tag{4}$$

where $\tilde{a}_{\text{crop},j}$ is the vector of crop-specific differential coefficients. Given this specification, one crop, chosen here to be the crop "other," has no explicit differential coefficients. Thus, the crop-selection model has a set of universal coefficients, $\mathbf{a}_{\text{crop},0}$, and explicit crop-specific differential coefficients only for cotton, corn, peanuts, and pecans.

The crop-selection model was estimated only for permitted sites that actively irrigate. The method of estimation is maximum likelihood. The set of predictor variables $\mathbf{Z}_{\text{crop}}$ appearing in the preliminary model included an intercept, acreage irrigated (according to the survey), and the 0/1 variable indicating the permit was located in a southern county. Attempts to estimate the model including a western county indicator failed due to colinearity with crop types. Results of estimation are reported in table D-4. The final model retains all the predictors used in the preliminary model. Seven of the fifteen coefficients are significant at the 10-percent level.

Two intercept terms, the cotton and peanut intercept differentials, were significant and negative implying that the probability of selecting these crops was lower holding acreage and location fixed. The acreage coefficient for all crops was positively significant as are the acreage coefficients for cotton and peanuts. The positive acreage coefficient for all crops indicated that the probability of selecting the residual crop, turf grass was smaller. Positive acreage coefficients for cotton and peanuts implied the probability of selecting those crops was higher if reported acreage was increased. The southern county indicator was positively significant for corn and negatively significant for pecans implying the probability of selecting corn was higher

Table D-4. Calibration results for the crop-selection model

Coefficient	Exploratory model (estimates based on original sample)			Final model (estimates based on 200-bootstrap Pseudo-samples)			
						90-percent confidence interval	
	Estimate	Standard error	Significance	Estimate	Standard error	Lower bound	Upper bound
Intercept	-2.114	1.887	0.264	-2.232	2.176	-6.038	0.923
Log of acres	0.722	0.434	0.099	0.760	0.468	0.068	1.584
If south	-1.006	0.659	0.129	-1.053	0.765	-2.003	0.316
Corn effects:							
Intercept	-1.125	1.355	0.408	-1.279	1.414	-3.376	1.137
Log of acres	0.230	0.289	0.428	0.267	0.285	-0.160	0.742
If south	0.711	0.370	0.057	0.754	0.402	0.234	1.556
Cotton effects							
Intercept	-3.239	1.157	0.006	-3.587	1.385	-5.999	-1.649
Log of acres	0.857	0.241	<0.001	0.945	0.284	0.546	1.440
lf south	-0.204	0.339	0.547	0.219	0.361	-0.733	0.423
Peanut effects							
Intercept	-3.336	1.602	0.039	-3.560	1.534	5.930	-1.559
Log of acres	0.764	0.320	0.018	0.820	0.316	0.296	1.228
lf south	-0.087	0.390	0.823	-0.087	0.390	-0.751	0.466
Pecan effects							
Intercept	1.598	1.300	0.221	1.714	1.682	-0.668	4.183
Log of acres	-0.313	0.282	0.269	-0.342	0.365	-0.869	0.238
If south	O.958	0.491	0.053	-0.974	0.545	-1.789	-0.070
Pseudo R^2	30.235						
Number of observations	148			148			

and selecting pecans was lower in the southern counties. The pseudo R^2 (see Maddala, 1983, p. 40) for this likelihood model was only 30 percent, indicating considerable potential for improving the explanatory power of the model using additional predictors—if they were available.

The fourth model component determined the irrigation starting and ending months for permitted sites that irrigate at some time during the growing season. This model component used the analysis of transition data (Lancaster, 1990, p. 23-32) to estimate the probability of starting and stopping irrigation in specific months. The essential concept driving this analysis was that of conditional probability. To see this, first consider the problem of determining the probability that irrigation begins in some month *S*. This probability is likely to depend on conditions that have prevailed in all the months leading up to *S*. In particular, if there is interest in knowing why irrigation begins in month *S*, it must also be known why irrigation did not begin in some earlier month. Untangling this causal structure can be difficult. A construct that greatly simplifies the analysis is to define the problem in terms of the conditional probability of starting irrigation in a given month assuming irrigation did not start in any earlier month. This probability generally can be stated without having to specify the entire history of causal factors leading up to period *S*. That is, the simple assumption was made that the conditional probability that irrigation begins in month *S* (given that it did not begin in an earlier month) was a function of period-*S* factors and non-varying factors only. Then j is defined to be the period in which irrigation begins. A mathematical notation for this assumption is:

$$P_S(j=S\,|\,j>S-1)=F_S\left(\overline{\mathbf{Z}}_{\text{start}},\mathbf{Z}_{\text{start},S}\right) \tag{5}$$

where $P_S(j = S|j`>S-1)$ is translated as the conditional probability that irrigation begins in period *S* (given that irrigation begins in a period that is greater than *S*-1 and given the history of predictors through period *S*) and F_S is the period-*S* probability function depending on a vector of non-varying factors, $\overline{Z}_{\text{start}}$, and period-*S* factors, $\mathbf{Z}_{\text{start},S}$.

The *unconditional* probability that irrigation begins in period *S* (but conditioned on the history of the predictor variables) can be built up from a series of the conditional probabilities given in (5). To see this, use Bayes' Law to write the unconditional probability as:

$$P_S(j=S)=P_S(j=S\,|\,j>S-1)P_{S-1}(j>S-1) \tag{6}$$

The unconditional probability $P_{S-1}(j>S-1)$ can be written as

$$\begin{aligned}P_{S-1}(j>S-1)&=1-\sum_{s=1}^{S-1}P_s(j=s)\\&=1-P_{S-1}(j=S-1\,|\,j>S-2)P_{S-2}(j>S-2)-\sum_{s=1}^{S-2}P_s(j=s)\\&=\left(1-P_{S-1}(j=S-1\,|\,j>S-2)\right)\left(1-\sum_{s=1}^{S-1}P_s(j=s)\right)\end{aligned} \tag{7}$$

By recursion, the unconditional probability is

$$1-\sum_{s=1}^{S-1}P_s(j=s)=\prod_{s=1}^{S-1}\left(1-P_s(j=s\,|\,j>s-1)\right) \tag{8}$$

where by definition the conditional probability $P_1(j = 1|j>0) = P_1(j = 1)$.

Finally, by substituting (7) and (8) into (6), the unconditional probability was expressed as a series of current and preceding unconditional probabilities

$$P_S(j=S)=P_S(j=S\,|\,j>S-1)\prod_{s=1}^{S-1}\left(1-P_s(j=s\,|\,j>s-1)\right) \tag{9}$$

The model component for determining the month in which irrigation terminates during the growing season was similar in construction to the seasonal-irrigation scheduling model. The one difference was that the probabilities for the end of the month in which irrigation terminates, denoted here by the random variable *k*, have to be made conditional on when irrigation began. Using the methods described above, the probability of terminating irrigation at the end of month *T*, given that irrigation began in month *S* (with *T*þ≥ *S*) is

$$\begin{aligned}&P_T(k=T\,|\,j=S)=\\&\quad P_T(k=T\,|\,k>T-1,j=S)\prod_{t=1}^{T-1}\left(1-P_t(k=t\,|\,k>t-1,j=S)\right)\end{aligned} \tag{10}$$

The conditional probability $P_t(k = t|k > t, j = S)$ possibly depends on *S* in two ways. First, because irrigation cannot end before it begins, we know this probability is zero for $t < S$. Second, the month in which irrigation begins can affect the termination decision in a relative sense. For example, a farmer may be less likely to stop irrigating if irrigation just began this month than if irrigation began four months ago. This is equivalent to assuming the conditional probability depends on a factor Z_S which is a simple function of the month in which irrigation begins. Additionally, the termination conditional probability depends on other fixed and time-dependent factors which were denoted by the vectors $\bar{\mathbf{Z}}_{\text{end}}$ and $\mathbf{Z}_{\text{end},T}$. Thus, the conditional probability for terminating irrigation is

$$P_T\left(k = T \mid k > T-1, j = S\right) \quad \begin{cases} G_T\left(\mathrm{Z}_S, \bar{\mathbf{Z}}_{\text{end}}, \mathbf{Z}_{\text{end},T}\right) & (T \geq S) \\ 0 & (T < S) \end{cases} \tag{11}$$

where G_T is a probability distribution function.

The above considerations lead to the specification of the joint probability of beginning irrigation in period S and terminating irrigation at the end of month T in terms of underlying conditional probabilities

$$\begin{aligned} P_T\left(j = S, k = T\right) &= P_S\left(j = S\right) P_T\left(k = T \mid j = S\right) \\ &= P_S\left(j = S \mid j > S-1\right) \left[\prod_{s=1}^{S-1}\left(1 - P_s\left(j = s \mid j > s-1\right)\right)\right] \mathrm{x} \\ &\quad P_T\left(k = T \mid k > T-1, j = S\right)\left[\prod_{t=S}^{T-1}\left(1 - P_t\left(k = t \mid k > t-1, j = S\right)\right)\right] \end{aligned} \tag{12}$$

The simplifying assumption was made that the conditional probability of terminating irrigation in month T, given that irrigation began and did not terminate prior to T, depends only on period-T factors and on the natural logarithm of the month in which irrigation began.

The last step was to assume a probability distribution for the start- and terminate-irrigation conditional probabilities F_S and G_T. We used the logistic distribution and assumed the predictors were combined linearly to obtain

$$F_S\left(\bar{\mathbf{Z}}_{\text{start}}, \mathbf{Z}_{\text{start},S}\right) = \frac{\exp\left(\bar{\mathbf{a}}'_{\text{start},S}\,\bar{\mathbf{Z}}_{\text{start}} + \mathbf{a}'_{\text{start},S}\,\mathbf{Z}_{\text{start},S}\right)}{1 + \exp\left(\bar{\mathbf{a}}'_{\text{start},S}\,\bar{\mathbf{Z}}_{\text{start}} + \mathbf{a}'_{\text{start},S}\,\mathbf{Z}_{\text{start},S}\right)}$$

$$G_T\left(\mathrm{Z}_S, \bar{\mathbf{Z}}_{\text{end}}, \mathbf{Z}_{\text{end},T}\right) = \frac{\exp\left(a_{\text{end},T}\,\mathrm{Z}_S + \bar{\mathbf{a}}'_{\text{end},T}\,\bar{\mathbf{Z}}_{\text{end}} + \mathbf{a}'_{\text{end},T}\,\mathbf{Z}_{\text{end},T}\right)}{1 + \exp\left(a_{\text{end},T}\,\mathrm{Z}_S + \bar{\mathbf{a}}'_{\text{end},T}\,\bar{\mathbf{Z}}_{\text{end}} + \mathbf{a}'_{\text{end},T}\,\mathbf{Z}_{\text{end},T}\right)} \tag{13}$$

where $\bar{\mathbf{a}}_{\text{start},S}$ and $\mathbf{a}_{\text{start},S}$ are time-dependent vectors of coefficients associated with the fixed and time-dependent predictors of the when-to-start irrigation decision, and $a_{\text{end},T}$, $\bar{\mathbf{a}}_{\text{end},T}$ and $\mathbf{a}_{\text{end},T}$ are time-dependent coefficients and vectors of coefficients associated with the fixed and time-dependent predictors of the when-to-end irrigation decision. Additionally, maximum months can be imposed in which irrigation can begin, $S_{\max}$, and end, $T_{\max}$. To ensure probabilities sum to one, the conditional starting and ending probabilities for their respective maximum months were set to one (that is, $F_{S_{max}}$ and $G_{T_{\max}}$ each equal one). Thus, the model did not include time-dependent starting coefficients for the period $S_{\max}$ or ending coefficients for the period $T_{\max}$.

The seasonal irrigation-scheduling model was estimated using the statistical method of maximum likelihood. The fixed predictors for both the starting- and ending-irrigation transitions included an intercept, the acreage irrigated during the year (from the survey), 0/1 indicator variables identifying the principal crop planted, and the 0/1 indicators identifying permits in western or southern counties. The only time-dependent predictor, included in both the starting- and ending-irrigation probabilities, was the natural logarithm of rainfall. The predictor for determining when irrigation began, $\mathbf{Z}_S$, was specified as the natural logarithm of S, the number corresponding to the irrigation starting month. Time-dependent coefficients for a given predictor were specified as a fixed effect common to all periods and a month-specific effect expressed as a deviation from the fixed effect. Under this specification, the month-specific deviations for the next to last months, $\mathrm{S}_{\max}$-1 for the irrigation starting probability and $T_{\max}$-1 for the irrigation ending probability, were restricted to equal the negative of the sum of the other month-specific deviations. Thus, no month specific deviations were included for the periods $S_{\max}$-1 or $T_{\max}$-1. Similarly, the crop indicators also were specified as deviations from the mean intercept. Consequently, the differential effect of the residual crop “other” was given by the negative of the sum of the differential effects for the explicitly identified crops.

To estimate the seasonal-irrigation scheduling model, it was necessary to have a sufficient number of permits starting and ending irrigation by the first and last months included in the specification. A compilation of irrigation starting and ending months for permits included in the survey is presented in table D-5. The criteria for designating the earliest and latest months for computing transitional probabilities was that at least six permits change transitional state in that month. By this criteria, the earliest month in which the start-irrigation probability was computed is March and the last month is July (that is, $S_{\max}$ is 7). For purposes of model calibration, the six permits beginning irrigation in January and February were reclassified as starting in March and the five permits beginning irrigation in August were reclassified as starting in July. The earliest month in which the end-irrigation probability is computed is July and the last month is October (that is, $T_{\max}$ is 10). The five permits ending irrigation in June were reclassified as terminating irrigation in July and the three permits terminating irrigation in November were reclassified as ending in October.

Table D-5. Frequency distribution of months irrigation begins and terminates among respondents of the Benchmark Farms Study area, southwestern Georgia

Month	Permits beginning irrigation			Permits terminating irrigation		
	Number	Percent	Cumulative percent	Number	Percent	Cumulative percent
January	4	2.7	2.7	0	0.0	0.0
February	2	1.4	4.1	0	0.0	0.0
March	9	6.1	10.1	0	0.0	0.0
April	24	16.2	26.4	0	0.0	0.0
May	60	40.5	66.9	0	0.0	0.0
June	36	24.3	91.2	5	3.4	3.4
July	8	5.4	96.6	6	4.1	7.4
August	5	3.4	100.0	60	40.5	48.0
September	0	0.0	100.0	61	41.2	89.2
October	0	0.0	100.0	13	8.8	98.0
November	0	0.0	100.0	3	2.0	100.0
December	0	0.0	100.0	0	00.0	100.0

Results of parameter calibration for the seasonal-irrigation scheduling model are reported in table D-6. Preliminary regressions indicated a lack of robustness if all coefficients were specified to be time dependent. Robust estimates were obtained by limiting time dependence to the intercept coefficients alone. The exploratory model and final model had the same specification and yielded similar results. Most of the coefficients were significant at the 10-percent level. Estimates show that the probability of starting irrigation in the months of March and April is lower than in May and June (the implied June effect is given by the negative sum of the March, April, and May effects). Additional rainfall in a given month between March and June causes the probability of starting irrigation in that month to increase. Permits growing cotton and peanuts have a statistically significant higher probability of starting irrigation in any given month between March and June. Permits located in the western counties of the study area have a lower probability of starting irrigation in any given month between March and June; whereas, permits in southern counties have a higher probability of starting irrigation.

The probability of terminating irrigation in July is significantly lower than in other months; whereas, the implied probability of terminating irrigation in September is higher. There was no statistically significant effect of August on the termination probability. More rainfall in any given month lowers the probability of terminating irrigation in that month for the months of July through September. The crops of corn, cotton, peanuts, and pecans have significantly higher probabilities of ending irrigation in any given month between July and September; whereas, the residual crop "other" has a lower probability of termination. Finally, permits located in the southern counties of the study area have a higher probability of terminating irrigation in any given month between July and September and a corresponding lower probability of ending irrigation in October.

The fifth model component related the natural logarithm of the inches of irrigation applied in a given month, W, to a vector of predictor variables, Z_{applied}. To account for month-to-month correlation in the amount of irrigation applied on a given farm, the model was specified to include a permit-specific error component, u, that is independent across farms, along with an idiosyncratic component, v, that is independent across both farms and months. The resulting error component model takes the form

$$\ln\left(W_{i,t}\right) = \mathbf{a}'_{\text{applied}}\,\mathbf{Z}_{\text{applied}\,i,t} + u_i + v_{i,t} \tag{14}$$

where i and t index the permit and month of the observation, $\mathbf{a}_{\text{applied}}$ is a vector of coefficients associated with the predictor variables, and u and v are independent normal random variables with zero means and permit-specific standard deviations $\sigma_{u,i}$ and $\sigma_{v,i}$. Generally, a better method for estimating a model with a specific permit-level effect would be to include a separate 0/1 indicator variable for each permit. That was not feasible in this case because the list of predictors included permit-specific variables—making model identification impossible.

An issue in the estimation of (14) concerns the inclusion of observations composed of aggregated monthly withdrawals arising from a respondent's failure to report. A modified method was adopted for estimating the error components model (Judge and others, 1985), a method that accounts for both temporally aggregated data and for permits that contain different numbers of monthly observations.

The first step of the method was to compute the mean-monthly irrigation over the period in which temporal aggregation occurs and substitute this value for the unobserved individual monthly values. Applying the logarithmic transformation to this mean estimate results in an estimate of

Table D-6. Calibration results for the seasonal-irrigation scheduling model
[—, not applicable]

Coefficient	Exploratory model (estimates based on original sample)			Final model (estimates based on 200 bootstrap Pseudo-samples)			
						90-percent confidence interval	
	Estimate	Standard error	Significance	Estimate	Standard error	Lower bound	Upper bound
Begin irrigation decision							
Intercept	-3.140	2.357	0.185	-0.785	0.581	-1.736	0.055
March effect	-2.681	0.507	<0.001	-2.795	0.421	-3.659	-2.236
April effect	-1.086	0.257	<0.001	-1.120	0.242	-1.500	-0.735
May effect	1.321	0.288	<0.001	1.343	0.284	0.999	1.919
Log of rainfall	0.703	0.350	0.047	0.775	0.348	0.246	1.326
Log of acres	-0.001	0.242	0.996	—	—	—	—
Log of pump capacity	0.412	0.366	0.262	—	—	—	—
If corn	0.169	0.548	0.758	0.146	0.571	0.790	1.055
If cotton	-1.109	0.457	0.017	-1.107	0.546	-2.006	-0.215
If peanuts	1.216	0.473	0.011	-1.131	0.597	-2.167	-0.248
If pecans	0.882	0.822	0.285	0.280	0.624	-0.694	1.264
If turf	1.033	0.603	0.089	1.038	0.957	-0.488	2.608
If west	-1.064	0.489	0.031	-0.899	0.362	-1.540	-0.406
If south	0.762	0.346	0.030	0.745	0.351	0.117	1.193
If ground water	-0.318	0.326	0.332	—	—	—	—
Terminate irrigation decision							
Intercept	5.313	2.795	0.060	-3.149	1.167	-4.983	-1.306
July effect	-2.330	0.349	<0.001	2.431	0.310	-3.070	-2.098
August effect	0.092	0.273	0.736	0.079	0.181	-0.213	0.358
Log of rainfall	-1.185	0.601	0.051	-1.111	0.389	-1.727	-0.542
Log of acres	0.430	0.317	0.178	—	—	—	—
Log of pump capacity	0.088	0.356	0.805	—	—	—	—
If corn	2.791	0.751	<0.001	3.361	0.842	2.181	4.876
If cotton	1.540	0.709	0.032	2.250	0.726	0.851	3.197
If peanuts	2.241	0.779	0.005	2.838	0.774	1.764	4.237
If pecans	2.039	0.862	0.020	2.037	0.849	0.713	3.431
If turf	0.897	1.211	0.460	1.180	0.744	0.039	2.471
If west	-0.904	0.528	0.090	-0.571	0.488	-1.289	0.311
If south	1.360	0.445	0.003	1.118	0.458	0.414	1.880
If ground water	-0.462	0.430	0.284	—	—	—	—
Log of start month	1.780	0.757	0.020	1.383	0.572	0.599	2.296
Pseudo R^2	52.516						
Number of observations	148			148			

mean logarithmically transformed irrigation that is biased upwards due to the convexity of the logarithm function. An attempt was made to account for this bias by including in the list of predictors two indicator variables that identify temporally aggregated observations. The first indicator simply identified those monthly observations that belonged to a group of temporally averaged observations. The second indicator identified those temporally aggregated observations that occurred at the beginning of the permit's irrigation season. These latter observations are identified because irrigation at the beginning of the season tended to be more variable resulting in a larger convex-transformation bias.

The next step was to obtain estimates of the variances of the permit-specific and idiosyncratic errors. Although these variances were assumed to be permit specific, it was not possible to estimate individual variances for individual permits. Instead, permits were classified into a few distinct types, indexed by m, and determine separate variances for each type. This was done by estimating (14) using ordinary least squares and computing the residuals $\hat{e}_{i,t}$ (in large samples, $\hat{e}_{i,t}$ equals $u_i + v_{i,t}$). The estimated residuals were inflated by dividing by the square root of 1 minus the observation's leverage. Let $\tilde{e}_i$ be the average of the N_i monthly residuals for permit i (N_i includes temporally aggregated months but not months with zero or missing values for irrigation), let $\bar{e}_i{}^*$ be the permit mean residual weighted by the square root of N_i, and let M_m be the number of permits in the sample of type m. Define $S_{\bar{e}}^{2}{}_{*m}$ to be the estimated variance of the M_m permit-specific weighted mean residuals. The expectation of this estimated variance is

$$E\left(s^2_{\bar{e}*\,m} \right) = \bar{N}_m\, \sigma^2_{um} + \sigma^2_{vm} \tag{15}$$

where $\mathbf{\bar{N}}_m$ is the average of the M_m values of N_i for permits of type m. Let $\tilde{e}_{i,\,t}$ be the difference $e_{i,t}$ - $\bar{e}_i$, let $n_{i,t}$ be the number of temporally aggregated months used to compute irrigation for permit i in month t ($n_{i,t}$ equals one if the monthly value is not temporally aggregated), and let $\tilde{e}^{\,*}_{i,\,t}$ be the residual difference $\tilde{e}_{i,\,t}$ weighted by the square root of the factor $N_i n_{i,t}/(N_i - n_{i,t})$. The estimated variance of the $\mathbf{\bar{N}}_m M_m$ weighted differential residuals $\tilde{e}^{\,*}_{i,\,t}$ for permits of type m, denoted $S^2_{\tilde{e}*m}$, the expectation

$$E\left(s^2_{\tilde{e}*\,m} \right) = \sigma^2_{vm} \tag{16}$$

Therefore, $s_{\tilde{e}}^{2}{}_{*m}$ forms an estimate of σ^2_{vm} and σ^2_{um} can be estimated by $\max\left(s^2_{\bar{e}*m} - s^2_{\tilde{e}*m}, 0\right) / \bar{N}_m$.

The original data are transformed according to the equations

$$1\mathrm{n}\left(W_{i,t}\right)^* = \frac{\left(1\mathrm{n}\left(W_{i,t}\right) - k_i\, \overline{1\mathrm{n}\left(W\right)}_i\right)}{\sqrt{n_{i,t}}}$$

$$\mathbf{Z}^*_{\mathrm{applied}\,i,t} = \frac{\left(\mathbf{Z}_{\mathrm{applied}\ i,t} - k_i\, \mathbf{\bar{Z}}_{\mathrm{applied}\ i}\right)}{\sqrt{n_{i,t}}} \tag{17}$$

where $\overline{\boldsymbol{1n(W)}}_i$ and $\mathbf{\bar{Z}}_{\mathrm{applied}\ i}$ are permit i averages of $1\mathrm{n}(W_{i,t})$ and $\mathbf{Z}_{\mathrm{applied}\ i,t}$, and

$$k_i = 1 - \sqrt{\frac{\bar{N}_{m(i)}\, s^2_{\tilde{e}*\,m(i)}}{N_i \max\left(s^2_{\bar{e}*\,m(i)} - S^2_{\tilde{e}*\,m(i)}, 0\right) + \bar{N}_{m(i)}\, s^2_{\tilde{e}*\,m(i)}}} \tag{18}$$

with $m(i)$ denoting the variance type of the ith permit. Lastly, the transformed data were temporally re-aggregated so that there was only one observation for each distinct temporally-aggregated value in the original sample. The re-aggregated transformed data were weighted by the factor $1/\sqrt{n_{i,\,t}\sigma^2_{vm(i)}}$, substituted into (14) and the coefficients estimated using ordinary least squares. The resulting residuals were, in large samples, independent and identically distributed with unit variance.

Equation (14) was estimated with and without the temporally aggregated data. For both samples, there was greater model error variance associated with permits that grew corn or cotton. Additionally, for the sample that included temporally aggregated data, model error variance was greater for permits with aggregated observations. These cases of heteroscedasticity were corrected using the procedure described above. The coefficient estimates obtained with either sample were similar but the model mean-square error was much larger and the crop variables insignificant with the temporally aggregated data included. Consequently, estimates were chosen from the sample that excluded the temporally aggregated data. These results are presented in table D-7. The predictor variables include an intercept, month-specific indicators, the natural logarithms of rainfall and acres irrigated (taken from the survey), crop indicators, and west and south location indicators. Only indicators for the months of March through October were specified, irrigation in other months were allocated to the nearest specified month. The monthly indicators were expressed in differential form so that the October effect was given by the negative sum of the effects for the months March through September. An exploratory regression including monthly effects for rainfall was found to be very sensitive to individual observations included in the regression, particularly observations having irrigation early and late in the season. Consequently, these effects were dropped from the model and only a single variable representing rainfall was included in the preliminary model.

The estimates for the monthly effects suggest a systematic pattern with low irrigation in March, rising consistently to a peak in July, followed by a moderate decline in August and steep decline in September. Rainfall has a negative effect on irrigation but its influence is not statistically strong. This may be due to the fact that rainfall in the area was very localized and not well captured by the rainfall statistics compiled at the county level. It also may be due to a lack of sufficient variation in rainfall across the study area making the rainfall signal difficult to detect.

Alternatively, it could be that irrigation was most sensitive to deviations from average rainfall for a given month. In that case, the rainfall variable appearing in the model could be measured with error; thereby suppressing the estimated effect on irrigation. However, for this effect to be important, it was necessary to observe that average monthly rainfall in a county is poorly represented by the month and location indicator variables included in the model. Regardless of cause, the rainfall variable lacked statistical significance and was excluded from the final bootstrap model. The coefficients for acres irrigated and pump capacity were negative and positive, respectively, and highly statistically significant. The coefficients were similar in magnitude implying that if pump capacity scales with acres irrigated the net effect on irrigation is zero. The crop effect coefficients were all negative implying that irrigation on "other" crops—the category unspecified in the regression—was generally larger than for the specified crops. The turf coefficient was insignificant from zero, making turf and "other" crops statistically similar in terms of irrigation applied.

Farms located in southern counties irrigated significantly less intensively than those in other counties—perhaps because they received more average rainfall. The variance of the permit-specific error was similar across crop types but the idiosyncratic error variance for corn and cotton was more than twice that for all remaining crops. Permit-specific variance was only about one-tenth of the idiosyncratic variance for farms growing corn or cotton, but was about one-third the idiosyncratic variance for the remaining farms. Because the model was estimated in logarithmic space, the percent accuracy of the regression model at the 90-percent level of confidence could be expressed as 165 times the square root of the magnitude of the error variance. If only the idiosyncratic variance was considered, this calculation implied the model was capable of predicting irrigation on a given farm within an accuracy of about 250 percent for corn and cotton and 160 percent for the remaining crops.

Model Prediction

The second phase of the analysis predicts monthly irrigation withdrawal for each of the non-sampled permits in the study area. The predictions, standard errors of the predictions, and confidence intervals were obtained using bootstrap analysis applied to the same five model components described for the calibration phase. The approach involves generating 200 individual predictions—or simulated realizations using the 200 sets of parameter and error estimates produced from the calibration phase in conjunction with information on the 11,305 individual permits within the study area and recorded in the permit file. The following provides details on the bootstrap methodology as it was applied to each model component.

The version of bootstrap analysis used herein relied on repeated resampling of the original survey observations, considered herein to be distinct permits containing multiple monthly irrigation values, to construct multiple bootstrap pseudo-samples. Each resample, consisting of the same number of permits as in the original sample, was produced randomly using replacement. Thus, in any pseudo-sample, a given permit from the original sample may appear multiple times or not at all. The model was recalibrated for each pseudo-sample to obtain multiple estimates of the parameters and associated model errors. The distribution of these estimates derive from the underlying characteristics of the sample to form an *empirical* distribution (as opposed to assuming the distribution was of a particular type—for example, normal). Note that because the random sampling was done at the observation level (as opposed to other bootstrap techniques that randomly select estimated errors to construct pseudo-dependent variables), parameter estimates and their estimated errors were robust to heteroscedasticity (model variance varies by observation) in the data. Moreover, the estimate of uncertainty produced by the method included uncertainty arising from sampling problems associated with missing, aggregated, or incomplete data in the case that a sampled permit holder did not irrigate.

The bootstrap method was extended to the prediction phase by taking the estimated parameters and randomly selected model errors generated from each pseudo-sample to compute *simulated realizations* of the predicted variables of interest for each month and each permit. The aggregated permit-level simulated realizations for each pseudo-sample constitute multiple predictions of the model, the distribution

Table D-7. Calibration results for the irrigation-applied model
[—, not applicable]

Coefficient	Exploratory model (estimates based on original sample)			Final model (estimates based on 200-bootstrap Pseudo-samples)			
						90-percent confidence interval	
	Estimates	Standard error	Significance	Estimate	Standard error	Lower bound	Upper bound
Intercept	-3.199	0.924	<0.001	-3.169	0.969	4.756	-1.546
March effect	-0.741	0.369	0.045	-0.680	0.340	-1.121	-0.050
April effect	-0.071	0.259	0.783	-0.056	0.587	1.080	0.763
May effect	0.315	0.184	0.087	0.322	0.144	0.038	0.547
June effect	0.431	0.149	0.004	0.443	0.136	0.276	0.735
July effect	0.677	0.141	<0.001	0.678	0.154	0.431	0.909
August effect	0.209	0.147	0.157	0.193	0.167	-0.102	0.421
September effect	0.341	0.158	0.032	-0.376	0.207	-0.645	-0.001
Log of rainfall	-0.020	0.I33	0.881	—	—	—	—
Log of acres	-0.477	0.139	<0.001	-0.503	0.135	-0.701	-0.284
Log of pump capacity	0.440	0.136	<0.001	0.431	0.149	0.200	0.687
Corn effect	-0.922	-0.314	0.004	-0.920	0.307	-1.411	0.449
Cotton effect	-1.078	0.304	<0.001	-1.074	0.308	-1.479	-0.496
Peanut effect	0.779	0.309	0.012	-0.791	0.287	-1.256	-0.360
Pecans effect	-0.697	0.356	0.051	0.789	0.355	-1.421	-0.279
Turf effect	-0.414	0.378	0.274	-0.342	0.353	-0.886	0.212
If west	-0.147	0.231	0.526	—	—	—	—
If south	-0.364	0.188	0.054	-0.426	0.197	-0.720	-0.076
If ground water	-0.083	0.200	0.680	—	—	—	—
Variance components Corn and Cotton							
σ_s^2 (permit)	0.241	—	—	0.229	0.179	0.000	0.483
σ_v^2 (idiosyncratic)	2.328	—	—	2.244	0.422	1.665	2.935
Other crop							
σ_u^2 (permit)	0.330	—	—	0.282	0.101	0.122	0.439
σ_v^2 (idiosyncratic)	1.064	—	—	1.054	0.183	0.766	1.341
Number of observations	458			458			

of which could be used to construct bootstrap mean predictions, standard deviations of the predictions, and associated confidence intervals.

The first step in each iteration of the bootstrap prediction model was to determine which permitted sites irrigated during the growing season. The probability that a given permit irrigated was determined from the iteration's estimate of the intercept term for the calibrated whether-to-irrigate model. This probability was used in conjunction with a binomial random number generator to simulate an independent realization of the irrigation decision for each of the 11,138 permits that were not included in the survey (11,305 total permits minus 167 sampled permits). Sampled permits were simulated to irrigate or not irrigate in accordance with their actual decision. If the simulated realization results in no irrigation, then irrigation withdrawal for that permit and bootstrap iteration was set to zero. Otherwise, the process proceeded to simulate the remaining model components.

The simulated value for acres irrigated for a non-surveyed permit was determined by first predicting the logarithm of acres irrigated using the log-linear acres-irrigated model (2) with coefficients set to the current iteration's estimates. A randomly selected error estimated from the current iteration's regression model is added to the prediction and the resulting sum was converted back into real space by applying the exponential transformation to obtain a "simulated realization" of acres irrigated. The actual acres irrigated was used as the simulated value for all sampled permits.

The crop-selection decision was simulated by substituting the current iteration's coefficient estimates for the crop selection model into the probability equation (3) to predict the probability of selecting each crop. In computing these predicted probabilities, simulated acres from the acreage irrigated prediction model were used. A multinomial random number generator used the estimated probabilities to randomly assign the crop for each non-surveyed permit. The actual crop planted was used as the simulated crop for all sampled permits.

A similar procedure was used to simulate the irrigation starting and ending months for a given bootstrap iteration. The coefficient estimates for an iteration were entered into the probability equations (12) and (13), along with the permit's attributes and simulated selected crop, to compute the predicted irrigation start/stop unconditional probabilities. These probabilities then were used in conjunction with a multinomial random number generator to randomly assign the irrigation starting and stopping months for each non-surveyed permit. Sampled permits were simulated to start and stop irrigation according to their observed operation.

The last simulation component determined the amount of irrigation applied. For each month the permit was simulated to irrigate, the logarithm of inches of irrigation applied was predicted using the irrigation-applied model (14) with coefficients evaluated at the current bootstrap iteration's regression estimates and permit-specific and idiosyncratic errors randomly selected from inferred values of these errors associated with the iteration's regression residuals (fig. D-2). The computed errors were inferred because even if the coefficients of the model were known with certainty, the only error that was observed was the gross error—a combination of the permit-specific and idiosyncratic errors. With a large number of time periods, the permit's average gross error would converge on the true permit-specific error. However, for these datam, a large sample consisted of a large number of permits, not a large number of time periods making asymptotic approximations of the permit-specific component infeasible. Note also that an acceptable bootstrap technique would be to forego the distinction between permit-specific and idiosyncratic errors and select gross errors as a block corresponding to all the errors available for a given permit. However, this method also was intractable because permits have differing numbers of time periods.

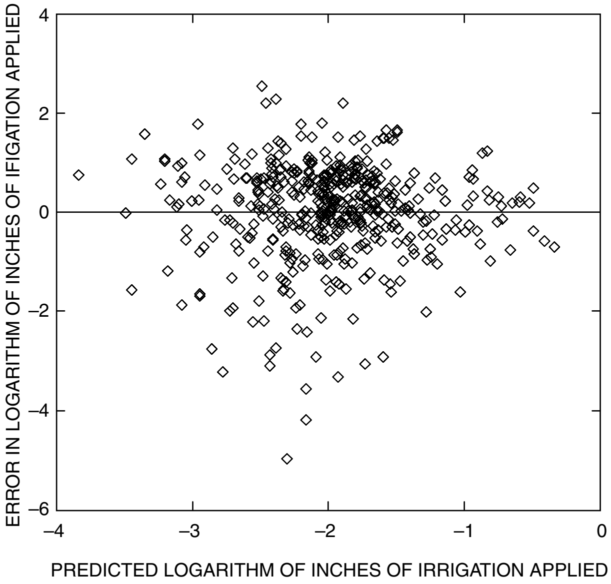


Figure D-2. Natural logarithm of error from the irrigation inches applied model relative to the predicted natural logarithm of irrigation inches applied in the Benchmark Farms Study area, southwestern Georgia.

The determination of permit-specific and idiosyncratic errors was accomplished as follows. The errors computed from the estimation of (14) used the transformed data (17), denoted $\tilde{e}_{i,t}$ were asymptotically independent with variance σ_{vm}^2. Errors were used to estimate the idiosyncratic errors v_i. The proposed estimate of the permit-specific error u_i is

$$\hat{u}_i = \sqrt{\frac{\frac{H_i}{N_i}\left(1-\left(1-k_i\right)^2\right)}{1-k_i}} \; \frac{\sum_{t=1}^{H_i} \tilde{e}_{i,t}}{H_i} \tag{19}$$

where $\tilde{e}_{i,t}$ is the error from the transformed variable regression and H_i is the number of distinct, possibly temporally aggregated, observations for permit i. From the definition of k_i, it can easily be shown that the variance of $\hat{u}_i$ is σ_{um}^2.

Using the above methodology, simulated realizations for all model components for each of the 11,305 permits were obtained. Estimates of total water applied on a farm were computed by multiplying the simulated inches applied by the farm's simulated acres irrigated (with a conversion factor included to convert the product to million gallons per day). These realizations then were aggregated to the county and study unit level and stored. The process was repeated 200 times using the 200 separate estimates of the coefficients and sets of errors generated by the calibration bootstrap analysis. The mean and standard deviation of the 200 aggregated estimates determined the bootstrap estimate and standard deviation of the estimate. The smallest range encompassing 180 of the 200 iteration estimates determined the upper and lower bounds to the 90-percent confidence interval for each bootstrap estimate.