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

# Factors Related to Well Yield in the Fractured-Bedrock Aquifer of New Hampshire

By Richard Bridge Moore, Gregory E. Schwarz, Stewart F. Clark, Jr., Gregory J. Walsh, and James R. Degnan

U.S. Geological Survey Professional Paper 1660

In cooperation with the
New Hampshire Department of Environmental Services

**U.S. Department of the Interior**

Gale A. Norton, Secretary

**U.S. Geological Survey**

Charles G. Groat, Director

2002

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

Library of Congress Cataloging-in-Publication Data

Factors related to well yield in the fractured-bedrock aquifer of New Hampshire / by Richard Bridge Moore ... [et.al.].
p. cm. -- (U.S. Geological Survey professional paper ; 1660)
Includes bibliographical references.
ISBN 0-607-98453-8 (alk. paper)
1. Groundwater--New Hampshire. 2. Aquifers--New Hampshire. 3. Wells--New Hampshire. I. Moore, Richard Bridge. II. Series.

GB1025.N45 F23 2002
553.7'9'09742--dc21 2002020252

# CONTENTS

## FIGURES

## PLATES

## TABLES

## CONVERSION FACTORS, ABBREVIATIONS, AND VERTICAL DATUM

| Multiply | By | To obtain |
|---|---|---|
| foot (ft) | 0.3048 | meter (m) |
| mile (mi) | 1.609 | kilometer |
| square mile ($mi^2$) | 2.590 | square kilometer |
| gallons per minute (gal/min) | 0.06309 | liter per second |
| gallons per minute per foot (gal/min/ft) | 0.2070 | liter per second per meter |

**Vertical Datum:** In this report "sea level" refers to the National Geodetic Vertical Datum of 1929 (NGVD of 1929)—a geodetic datum derived from a general adjustment of the first-order level nets of both the United States and Canada, formerly called Sea Level Datum of 1929.

# Factors Related to Well Yield in the Fractured-Bedrock Aquifer of New Hampshire

*By* Richard Bridge Moore, Gregory E. Schwarz, Stewart F. Clark, Jr.,
Gregory J. Walsh, *and* James R. Degnan

## Abstract

The New Hampshire Bedrock Aquifer Assessment was designed to provide information that can be used by communities, industry, professional consultants, and other interests to evaluate the ground-water development potential of the fractured-bedrock aquifer in the State. The assessment was done at statewide, regional, and well field scales to identify relations that potentially could increase the success in locating high-yield water supplies in the fractured-bedrock aquifer. statewide, data were collected for well construction and yield information, bedrock lithology, surficial geology, lineaments, topography, and various derivatives of these basic data sets. Regionally, geologic, fracture, and lineament data were collected for the Pinardville and Windham quadrangles in New Hampshire. The regional scale of the study examined the degree to which predictive well-yield relations, developed as part of the statewide reconnaissance investigation, could be improved by use of quadrangle-scale geologic mapping.

Beginning in 1984, water-well contractors in the State were required to report detailed information on newly constructed wells to the New Hampshire Department of Environmental Services (NHDES). The reports contain basic data on well construction, including six characteristics used in this study—well yield, well depth, well use, method of construction, date drilled, and depth to bedrock (or length of casing). The NHDES has determined accurate georeferenced locations for more than 20,000 wells reported since 1984. The availability of this large data set provided an opportunity for a statistical analysis of bedrock-well yields. Well yields in the database ranged from zero to greater than 500 gallons per minute (gal/min).

Multivariate regression was used as the primary statistical method of analysis because it is the most efficient tool for predicting a single variable with many potentially independent variables. The dependent variable that was explored in this study was the natural logarithm (ln) of the reported well yield. One complication with using well yield as a dependent variable is that yield also is a function of demand. An innovative statistical technique that involves the use of instrumental variables was implemented to compensate for the effect of demand on well yield.

Results of the multivariate-regression model show that a variety of factors are either positively or negatively related to well yields. Using instrumental variables, well depth is positively related to total well yield. Other factors that were found to be positively related to well yield include (1) distance to the nearest waterbody; (2) size of the drainage area upgradient of a well; (3) well location in swales or valley bottoms in the Massabesic Gneiss Complex and Breakfast Hill Granite; (4) well proximity to lineaments, identified using high-altitude (1:80,000-scale) aerial photography, which are correlated with the primary fracture direction (regional analysis); (5) use of a cable tool rig for well drilling; and (6) wells drilled for commercial or public supply. Factors negatively

related to well yields include sites underlain by foliated plutons, sites on steep slopes sites at high elevations, and sites on hilltops. Additionally, seven detailed geologic map units, identified during the detailed geologic mapping of the Pinardville and Windham quadrangles, were found to be positively or negatively related to well yields. Twenty-four geologic map units, depicted on the Bedrock Geologic Map of New Hampshire, also were found to be positively or negatively related to well yields.

Maps or geographic information system (GIS) data sets identifying areas of various yield probabilities clearly display model results. Probability criteria developed in this investigation can be used to select areas where other techniques, such as geophysical techniques, can be applied to more closely identify potential drilling sites for high-yielding (greater than 40 gal/min) bedrock wells.

To measure the added value of field-based methods for well-yield-probability forecasting, the model was run with and without the variables developed from the detailed geologic mapping for the Pinardville and Windham quadrangles. Four probability maps were produced for the two quadrangles, one with and one without the added variables. These maps clearly demonstrate the advantage of detailed geologic mapping when prospecting for new ground-water supplies in the fractured-bedrock aquifer of New Hampshire.

## INTRODUCTION

Considerable population growth has resulted in New Hampshire in the last 40-50 years. This growth has led to a need for additional water resources in many communities. Many of these communities have limited sand and gravel aquifers, which generally are the most favorable aquifers for constructing high-yield wells in New Hampshire. Additional water resources can be found in the fractured-bedrock aquifer, which generally produce low yields to wells (a few gallons per minute). However, through intensive site analyses, high-yielding zones can be located in the aquifer. Throughout New Hampshire, the bedrock is crystalline. Fractured-bedrock aquifers in crystalline bedrock are among the least understood and quantified ground-water resources in the Nation. A major problem in evaluating ground-water availability in fractured bedrock is its extreme variability in water-bearing properties.

The New Hampshire Bedrock Aquifer Assessment was done by the U.S. Geological Survey (USGS), in cooperation with the New Hampshire Department of Environmental Services (NHDES), Water Division, to provide information that can be used by communities, industry, professional consultants, and other interests to evaluate the ground-water development potential of the fractured-bedrock aquifer. This study was done at three scales—statewide, regional, and local—to quantify relations among well yields and bedrock data, physiographic setting, remotely sensed lineaments, and well characteristics in New Hampshire. The statewide scale was designed as a reconnaissance-level statistical analysis of the relations between bedrock well yield and bedrock type, lineament characteristics, topography, and other well-site and construction characteristics. The regional-scale analysis was similar, except that field-based geologic-data collection and additional lineament data were added to assess what effect quadrangle-scale (1:24,000) data had on well-yield relations. Local assessments (usually at a well-field scale) were done by use of geophysical tools to identify the location and orientation of discrete zones of fracture in the bedrock that are referred to as “fracture zones” in this report.

### Purpose and Scope

This report presents results of statistical analyses conducted for the statewide and regional scale investigations of the bedrock-aquifer assessment. Results of the well-field assessments are presented in a companion report by Degnan and others (2001).

The NHDES has a large data set of bedrock-well locations, yields, and other characteristics that are adequate for an exploratory, statistical analysis of hydrogeologic and construction factors associated with well yields. There are many possible factors (site and well characteristics) that relate to high-yield well sites. For the statewide analysis, available well data and remotely sensed data were compiled. For the regional analysis, fracture data and general geologic and hydrogeologic information were collected and compiled for two 1:24,000-scale New Hampshire quadrangles. The purpose of the more intensive regional-scale compilation was to determine the degree to which predictive well-yield relations, developed as part of the statewide reconnaissance investigation, could be improved with quadrangle-scale geologic mapping. Data were collected through bedrock mapping and supplemental hydrogeologic observation. The Pinardville and Windham quadrangles were chosen for regional analysis because of the large number of wells located in each quadrangle and the variety of geologic settings in the quadrangles.

This study was not intended to explain the occurrence and nature of water-bearing bedrock fractures in New Hampshire, nor the nature of water movement in the fracture system. Understanding the occurrence and distribution of fractures in crystalline bedrock and associated water movement are being intensively evaluated at research sites (Shapiro and Hsieh, 1991) and cannot be determined statewide or regionally with the present state of the science. The statistical identification of hydrologic and other factors that are related to bedrock well yields statewide and regionally, will, however, provide information to further the understanding of these processes.

## Previous Investigations

Large-scale analyses of bedrock-aquifer yields are few, and especially are limited in crystalline bedrock settings in the northeastern United States. Investigators have successfully related geohydrologic factors, particularly lineaments, to well yields in the carbonate rocks of Pennsylvania (for example, Knopman, 1990; Siddiqui and Parizak, 1971). One of the largest studies (Daniel, 1989; Daniel and others, 1983) assessed hydrogeologic characteristics and yields of 6,200 bedrock wells in the Piedmont and Blue Ridge provinces of North Carolina. These investigations found that wells in draws or valleys have average yields three times those of wells on hills and ridges. Wells in the most productive hydrogeologic units had average yields twice those of wells in the least-productive units. One of the investigations, directly applied to fractured-bedrock aquifers in the Northeast, was performed by Mabee and others (1994) on a 17-mi$^2$ island on the Maine coast. This study examined 35 bedrock wells and found that aquifer transmissivity, normalized by well depth, was positively related to well proximity to fracture-correlated lineaments, which are lineaments trending in the same direction as fractures observed in the bedrock. A number of these studies (table 1) measured the degree to which wells near lineaments represented a population of high yields.

**Table 1.** Summary of well yield and lineament investigations considered for this study

[>, greater than; NHDES, New Hampshire Department of Environmental Services; NA, not available]

| Study/imagery used | Region/units tested | Measure of well productivity | Total number of wells | Number of wells near lineaments | Number of wells beyond buffer zone around lineaments | Confidence the samples represent, two separate populations (in percent) |
|---|---|---|---|---|---|---|
| Siddiqui (1969); Siddiqui and Parizek (1971 and 1974) Low-altitude aerial photography | Pennsylvania carbonate rocks | Normalized transmissivity | 45<br>80 | 18<br>53 | 27 | >99 |
| Mabee and others (1994) High-altitude areal photography and side-looking radar | Maine metasedimentary | Normalized transmissivity | 35 | 20 | 15 | 30 |
| Mabee and others (1994) High-altitude areal photography and side-looking radar | Maine "Fracture-correlated" lineaments (involving extensive additional fracture-fabric field work) | Normalized transmissivity | 35 | 7 | 28 | 89 |
| Preliminary report of present investigation (Moore and others, 1998) Low- and high-altitude areal photography | Plates 1-6, covering southern New Hampshire, "migmatites" and unfoliated plutons (involving no field work beyond NHDES locating wells) | Reported yield | 11,212 | NA | NA | 99.9 |

Most of the investigations mentioned in table 1 used lineament analysis to identify potential bedrock-fracture zones or bedrock-solution-channel enhancements in carbonate bedrock. Numerous investigations refer to lineament analysis as “fracture-trace analysis”; however, field observations of fractures is desirable for fracture-trace analysis. A lineament is defined as a linear pattern, seen on aerial photographs and other remotely sensed imagery, that meet established criteria for features that may be the result of underlying zones of fractured bedrock (Clark, Moore, and others, 1996). Despite the use of rigid criteria, any given lineament may not be underlain by a fracture zone. Anthropogenic remnants such as an abandoned woods road, rail lines, or right-of-ways, could be identified as a lineament but may not be related to fractures. Lineament data also are limited because of the accuracy of lineament locations, which varies with the scale of the imagery. A lineament analysis, is therefore, a preliminary analysis whereby lineaments are identified on remotely sensed imagery, and can only be confirmed as fracture related through subsequent field observation. Blanchet (1957) described patterns observed on aerial photographs, which were related to fracturing in the Earth’s crust. One of the first thorough treatments of lineament analysis was described by Lattman (1958) who provided much of the framework for techniques often applied currently (2001). In New England, many consulting firms commonly use some form of lineament analysis, in combination with other investigative techniques, to locate high-yield well sites in fractured-bedrock aquifers. The specific techniques applied, however, generally are considered to be proprietary and descriptions are often incomplete or not available.

Clark, Moore, and others (1996) describe the lineament-analysis methods used in the New Hampshire Bedrock Assessment. Some of these methods previously were used by Blanchet (1957), Lattman (1958), Brown (1961), Daniel (1989), and Mabee and others, (1994), but no studies integrated results of statewide, regional, and local-scale investigations based on the variety of data compiled for the New Hampshire bedrock-aquifer statistical analysis.

## Acknowledgments

The authors thank the individuals who contributed their time and effort offering advise and assistance. Thanks go especially to the following individuals: from New Hampshire Department of Environmental Services, Frederick Chormann, John Cotton, Philip O’Brien, Eugene Boudette, and David Wunsch; from the University of New Hampshire, Mathew Davis, Wallace Bothner, and Lawrence Dingman; from Dartmouth College, John Lyons; from the Maine Geological Survey, Marc Loiselle; from the University of Massachusetts, Stephen Mabee; from the private sector, Gary Smith and Bradford Caswell; and from the U.S. Geological Survey, Thomas Armstrong, Lawrence Drew, James Borchers, Paul Hsieh, Ben Morgan, Jack Schuenemeyer, Allen Shapiro, Charles Daniel and Jeffrey Starn.

## HYDROGEOLOGIC SETTING

New Hampshire’s landscape is underlain by bedrock, which makes up the crust of the Earth, and by surficial deposits, most of which were left behind by the continental ice sheet during glaciation. The trend of major bedrock units that form the regional structural pattern in New Hampshire generally is north-northeast. The bedrock that underlies New Hampshire is composed of metamorphic and plutonic rocks (Lyons and others, 1997). The geologic history of New Hampshire’s bedrock materials, their structure, and the tectonic forces and processes that formed them is part of the history of the Appalachian Mountains. Details of this history can be found in Hatcher and others (1989) and Boudette (1990).

Plutonic rocks, locally common in New Hampshire, are composed of interlocking minerals and have little or no primary pore space. Likewise, primary porosity of the parent materials of metamorphic rock in New Hampshire has been eliminated by metamorphism. Sedimentary and igneous rocks are the parent materials for many of the metamorphic rocks in New Hampshire. These rocks have been subjected to a wide variety of conditions of temperature and pressure, which formed differing assemblages of minerals and textures and eliminated voids by compaction and recrystallization.

Tectonic processes, which accompanied metamorphism and mountain building, folded and faulted New Hampshire's rocks. Tectonic processes created many fractures, faults, and voids and, at the same time, locally sealed some of these structures (Swanson, 1988).

Despite a complex history of sedimentation and multiple events of deformation, metamorphism, and intrusion, New Hampshire's bedrock can be viewed from a broad hydrologic perspective. Water moves through, and is stored in, open fractures. The size, number, distribution, and degree of interconnection of fractures are highly variable; in general, however, fractures are few and, when present, generally are presumed to decrease in size and number with depth. Thus, the overall storage capacity of bedrock is small and tends to decrease with depth (U.S. Geological Survey, 1984, p. 304). Wells that penetrate bedrock commonly yield dependable supplies of water suitable for single-family domestic needs, and, for this purpose, bedrock is a principal aquifer. Zones where bedrock is extensively fractured may yield large quantities of water. Many small water systems that serve residential developments use bedrock wells, and in the past two decades, the application of exploration technology has enabled more municipal water-supply systems to use the bedrock aquifer.

Because ground-water flow processes in fractured crystalline bedrock are complex, developing a more complete understanding of these processes, and methods to characterize them, is a priority of USGS research. The USGS has been conducting investigations at its national fractured-bedrock research site at the Hubbard Brook Experimental Forest in Thorton, N.H., since 1990 (Shapiro and Hsieh, 1991; Hsieh and others, 1993). Research conducted at this site has included developing and utilizing state-of-the-art geologic, geochemical, geophysical, and hydrologic methods to gain a better understanding of ground-water-flow processes in fractured crystalline bedrock.

## METHODS AND APPROACH

The approach of this study was to relate bedrock well yields in New Hampshire to factors known to affect well yields in other parts of the country. These factors include bedrock type, lineament characteristics, physiography, and other well-site and construction characteristics on a statewide scale. This analysis was done by use of a multivariate-regression model with instrumental variables. Emphasis was placed on the use of available data and remotely sensed information, with some additional field-data collection involving two geologically diverse 7.5-minute quadrangles. Lineament investigations, as previously discussed, were useful indicators of high-yield wells in crystalline bedrock. Other geologic and site characteristics that were significantly related to well yield also were used in a predictive mode to identify areas with various probabilities of obtaining a selected yield. The approach incorporates the use of statewide and regional information to evaluate the added benefit of additional data collection. The following sections describe the sources and treatment of data and development of the statistical approach.

In investigating factors that relate to high-yield well sites, many possible factors (referred to as variables in this report) were examined. Statewide data sets were compiled from existing sources or were created for this study. The primary information that was analyzed included well information, bedrock lithology, surficial geology, lineaments, topography, and various derivatives of information in those data sets. In addition, data were collected at a more detailed scale for the areas contained in the Pinardville and Windham quadrangles, New Hampshire (fig. 1). These data are discussed in detail in the following paragraphs.

The relation of multiple variables to bedrock-well yields initially was examined by simple regression techniques to compare preliminary analyses with other investigations and to help streamline more detailed regression-model analyses involving the use of instrumental variables. The relation of selected individual variables to yields also was examined by simple, bivariate, statistical analyses.

Regression was selected as the primary method of analysis. As such, the established model is a stochastic (statistical) model rather than a process-oriented geologic model. A process-oriented model (modeling flow in known fractures) was not possible because the necessary detailed geologic information, such as a statewide data set on fractures, is not available. Variables examined included categorical variables (such as bedrock type), as well as continuous numerical variables such as the slope of the land surface. A nested model was used, where detailed geologic data from the quadrangles were applied

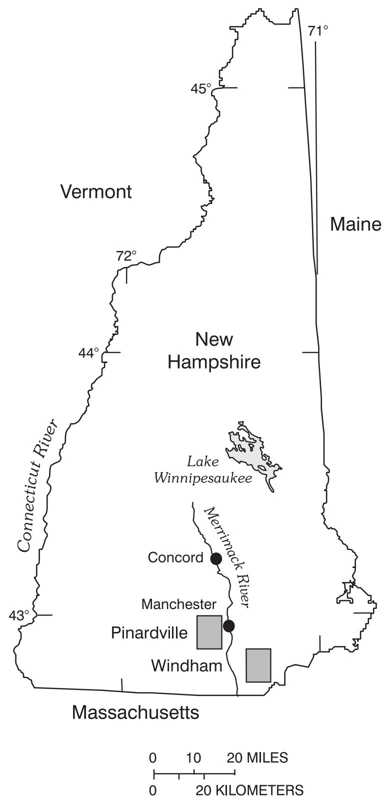


**Figure 1.** Location of the Pinardville and Windham 1:24,000-scale quadrangles in southern New Hampshire where additional geohydrologic, fracture, and lineament data were collected.

where available or nested in the statewide model. The model was used to evaluate the simultaneous effect of many variables and the interaction of those variables, and provided confidence intervals. These intervals can be used to predict the probabilities of obtaining a given yield given different site and well characteristics.

## Descriptions and Sources of Data

This study required the development of geographic information system (GIS) compatible data sets by use of ARC/INFO. All data sets used were the most detailed available statewide. Many variables were considered, however, only 43 were found to be significant and are included in the final regression model.

### Well Information

Detailed information on water wells constructed since 1984 is collected and maintained in a computer database by NHDES. Many of these wells have been field-inventoried to obtain accurate locations. Wells are categorized as to whether they are used for domestic, public supply, commercial, or other uses. Nearly 98 percent of the all wells in the database were drilled for domestic use. Basic well data are reported by drillers and include six characteristics used in this study—well yield, well depth, well use, method of construction, date of construction, and depth to bedrock (or length of casing). Well diameter was not included in this database. However, nearly all of the domestic wells are 6 inches in diameter. A georeferenced digital database is maintained by the State where location coordinates were determined either by plotting and digitizing from a map at 1:24,000 or 1:25,000 scale or by differentially corrected Global Positioning System (GPS) measurement. Well locations are accurate to ±100 ft.

The well data set collected by NHDES and used in this study contains records of 20,308 wells drilled from 1984 to 1998, and subsequently field located. A randomly selected subset (20 percent or 4,050 wells) of the available well data were reserved for verification tests of the model. The availability of this large data set provided an opportunity for statistical analysis of variables considered to be related to well yields. Specific components of the database used in this study are described in the following paragraph with respect to their use for statistical analysis.

Well yields in the database ranged from zero to greater than 500 gal/min. Yields are reported by well drillers and are determined by a variety of methods at the time the well is drilled. Yields generally are determined as the rate of water that can be airlifted on a continuous, short-term (generally tens of minutes) basis. Low yields may be accurately measured by timing the pumpage of water into a known volume. Moderate- to high-reported yields often are quantified by discharging water through a rated channel and are likely to be rough approximations of the true yield. Such methods introduce an unquantifiable variance in measured yields but are acceptable for statewide or regional evaluations involving large numbers of wells. Yields equal to or greater than 40 gal/min were considered high in this investigation. High-yielding wells constitute nearly 10 percent of the well population (fig. 2). The yields of public-supply wells typically are quantified by an aquifer test of hours to days in duration and represent a more rigorous estimate of yield than that reported for domestic wells by the drillers. The dependent variable that was statistically explored in this study was the natural log of the reported well yield. Other possible dependent variables considered were the yield and yield per foot of open hole; however, the natural log of well yield was selected because the population of well yields is log normally distributed.

A problem in using well yield as the dependent variable in a multivariate-regression analysis is that yield is a function of a well owner's water needs. Generally, a well only is drilled as deep as is needed to meet a specific targeted yield. Thus, the reported yield is not the maximum potential yield for that site. As well yield is, in part, a function of demand, an advanced statistical technique involving instrumental variables must be used for a more accurate estimate of yield.

The year each well was drilled can be included as a surrogate for demand in the model because demand for water (and the depths to which well owners have been willing to pay to drill to obtain more water) has increased in New Hampshire. The effect of time was observed using data from the Pinardville quadrangle (Lawrence Drew, U.S. Geological Survey, oral commun., 2000). Drillers confirmed that well depth increased to meet this demand (Terry Swain, Capital Well Co., oral commun., 2000).

Depth, reported by the drillers, likely is accurate to the nearest foot. Depth also is a function of demand and is inversely related to well yield. Wells are drilled deeper at low-yield sites and drilling is stopped when a desired yield is reached. This action makes depth an "endogenous variable," and, for this reason, depth is inversely related to well yield.

Data on the method of well construction essentially are limited to two general categories; cable tool and rotary. Bedrock wells drilled with a cable-tool drill rig can be identified in the NHDES database, but all other methods, such as the various forms of rotary drilling, were not differentiated. As cable-tool construction is different from the other techniques, it is reasonable to test the effects of cable tool in comparison to other methods of construction in the regression model. With cable-tool construction, the drill bit is pounded into the bedrock perhaps opening up or enhancing fractures. An indicator variable was used to identify whether or not each well was constructed by use of a cable-tool drill rig.

## Bedrock Lithology

Bedrock lithology was assigned using a digital data set of the Bedrock Geologic Map of New Hampshire (Lyons and others, 1997). This data set contains 174 mapped bedrock units and was compiled at a scale of 1:250,000.

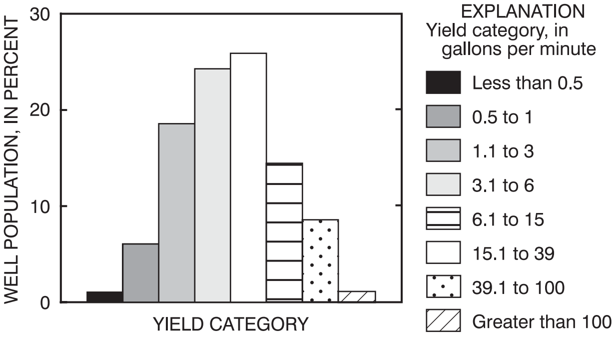


**Figure 2.** Percent of statewide bedrock well yields in New Hampshire.

In addition to the 174 mapped bedrock units, 7 generalized categories, covering all the rocks of New Hampshire, were created with input from Dr. Wallace Bothner (University of New Hampshire, written commun., 1998). Variables identifying these groupings were created to broadly categorize similar rock types. Some mapped geologic units depicted on the State bedrock map have limited areal extent, yet have properties such as origin, texture, and degree of metamorphism common to other mapped units. Grouping these lithologies together into common categories allows for statistical analysis of areas that would otherwise have had too few well-data points to be included individually as independent variables. These seven major lithologic units with similar lithologic characteristics are (1) nonfoliated to weakly foliated igneous rocks, (2) well-foliated igneous rock (well-foliated rocks of the New Hampshire Plutonic Suite and Oliverian Plutonic Suite), (3) fine-grained, tightly bonded metasedimentary rocks, (4) fine to medium-grained granular-metasedimentary rocks, (5) gray-black graphitic micaceous metapelite locally interbedded with calcareous or siliceous rock, (6) volcanic rock and associated metapelite, and (7) rocks of the Massabesic Gneiss Complex and Breakfast Hill Granite (fig. 3).

## Lineaments

A statewide data set of lineaments was created for this project following the methods presented in Clark, Moore, and others (1996). The imagery used came from four remotely sensed platforms. These platforms included Landsat imagery (1:1,000,000 enlarged to 1:250,000), side-looking airborne radar (SLAR) (1:250,000), high-altitude aerial photography (approximately 1:80,000), and low-altitude

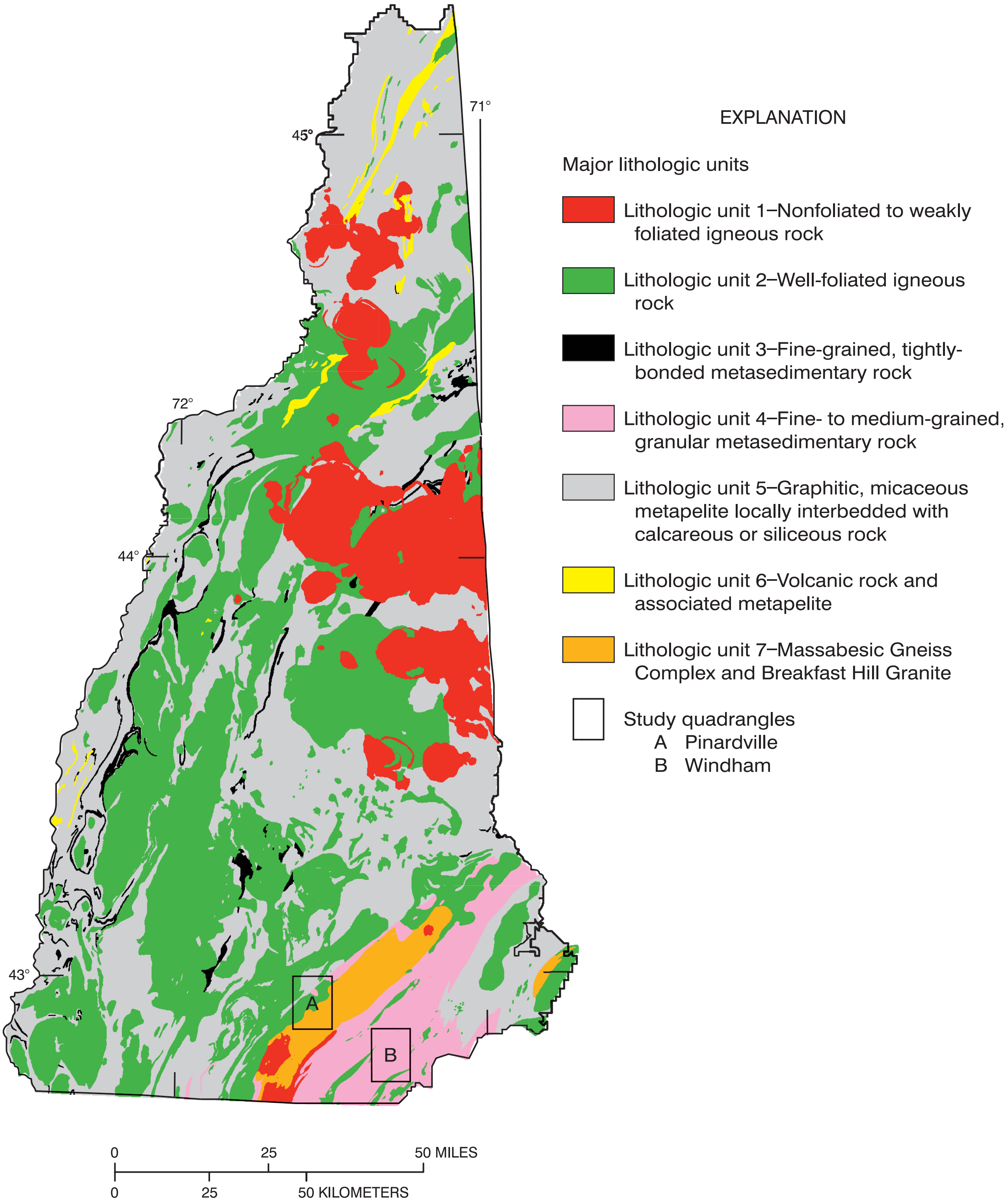


**Figure 3.** Major lithologic units in New Hampshire adapted from Lyons and others (1997).

photography (approximately 1:40,000, enlarged to 1:20,000). In addition to the imagery mentioned above, the regional-scale analysis included color-infrared imagery (approximately 1:40,000) and topography from standard (7.5 minute) topographic maps (1:24,000). Data at different scales were analyzed by different analysts in order to maximize the reproducibility of lineament identifications (Clark, Moore, and others, 1996). This analysis was done by selecting just those lineaments identified by two different observers. The imagery was examined and lineaments were recorded on mylar overlays by the observers working independently. Results were compared and placed in the following three categories: (1) blind comparisons—matching lineaments drawn by independent observers; (2) confirmed lineaments—lineaments drawn by one observer and confirmed by another observer; and (3) rejected lineaments—lineaments drawn by one observer and not confirmed by another observer. All blind or confirmed lineaments were digitized into a GIS database for use in the statistical analysis. These features were either plotted on scale-stable base maps that were later mounted on a digitizing tablet and digitized, or were plotted directly on imagery overlays that were scanned later and rectified to allow on-screen digitizing.

Lineament data are limited because (1) the accuracy of lineament locations varies with the source and scale of the imagery, (2) lineaments can be caused by cultural rather than geologic features, (3) lineaments related to bedrock features are most easily observed in areas where the surficial material is thin (areas where overburden is thin are known to have low-yielding bedrock), and (4) lineaments theoretically are restricted to steeply dipping fractures or fracture zones that produce straight-line intersections with the earth surface. As lineaments are more readily found in the hilly shallow-to-bedrock areas, establishing a relation to well yield is difficult because of inaccuracies in the georeferenced locations of the wells and the lineaments, and because these shallow-to-bedrock areas also are associated with relatively low yields.

For each well site, distance to lineaments and lineament characteristics were compiled for use in the statistical analyses. Distance from wells to lineaments was coded by identifying whether or not the wells were in a "buffer zone" surrounding each lineament. A buffer zone of 100 ft was selected by trial-and-error, considering the accuracy of the well and lineament locations. The accuracy of lineament locations is approximately ±80 for aerial photography and ±150 ft for the Landsat imagery. Well locations are estimated to have an accuracy of ±100 ft. Selecting a buffer zone less than 100 ft would, thus, have little or no meaning. The buffer zone selected (100 ft) is consistent with the findings of Mabee (1992, p. 106), who selected (by trial-and-error) an optimal buffer zone, for statistical analysis, of 98 ft (30 m).

Lineaments identified from SLAR were excluded from the analysis because of difficulties in positioning the lineaments even to an accuracy of ±100 ft. The SLAR data, compiled on 1:250,000 scale maps, were located to approximately ±200 ft. Additionally, because SLAR relies on shadowing effects, there may be a tendency for the apparent location of the lineament (the shadow) to be off-center from the bottom of the valleys or swales that cause the feature. Linear cultural features, such as road cuts, rail lines, right of ways, and tree lines at the edges of open spaces, can be prominent on SLAR imagery (more so than on other imagery) and could be mistaken for geologic lineaments. The SLAR imagery was carefully examined to avoid misidentification; however, the difficulties in positioning the SLAR lineaments precluded its use.

Additional variables were created to evaluate interactions between lithology and lineaments. Well yields in relation to lineaments or fractures can differ between lithologies because the character of fractures and the extent of the affected area near fracture traces may differ between lithologic units. For example, weathered fracture zones in plutons may contain permeable "grus," a granular fragmental product of weathering of granitic rocks (Bates and Jackson, 1980), but in metasediments may contain clay-rich low-permeability saprolite. The character of fractures and the extent of affected areas near fracture traces also can differ between bedrock types. In plutonic rocks, for example, there tend to be more sheeting fractures, which could be hydrologically connected to vertical fracture zones.

The effect of the orientation of the lineaments was examined late in the process of building the regression model to minimize the number of possible variables. The strike of the nearest lineament within 100 ft of each well was computed and binned into eighteen 10-degree-angle categories. Eighteen angle categories multiplied by numerous lithologies would result in an inordinate number of variables with the likely possibility of spurious results.

A lineament-density database for New Hampshire was created by identifying the density of lineaments within a 1,000-foot radius around each well for each lineament platform (Clark, Moore, and others, 1996, p. 10, plates 1-11. Plate areas 1-10 include all of New Hampshire south of 44 degrees latitude and plate 11 is just to the north and includes most of the eastern part of the White Mountains). Lineament density was equal to the total length of all lineaments within 1,000 ft of each well divided by the area of this circle. Contrary to expectations, lineament density was found to be negatively correlated to well yield in these preliminary analyses. Thus, areas with high lineament densities are most likely in areas of thin overburden, which are inversely related to well yield. (Fracture zones beneath thick overburden are less apt to be observed). Because of this negative relation, and the availability of better predictors of areas that have shallow depths to bedrock, lineament density was not used in further well-yield analyses. Similarly, the distances from the wells to where lineaments intersect one another were examined for five test quadrangles. Unlike positive relations found elsewhere in carbonate bedrock, no significant relations were found in the New Hampshire setting; therefore, intersections were not used in further analyses.

### Surficial Material

Surficial material such as stratified-drift aquifers possibly affect bedrock well yield through (1) increased permeability or the ability to supply water to the bedrock aquifer, (2) saturated overburden at the bedrock interface, and (3) greater occurrence of stratified-drift aquifers in valleys. New Hampshire lacks a statewide GIS data set of overburden types; however, a statewide stratified-drift aquifer data set is available and can be applied to the regression model by use of a dichotomous indicator (0 or 1 binary) variable. Overburden type was obtained from maps of the extent of stratified-drift aquifers in New Hampshire, which were mapped as part of a statewide series of investigations (Medalie and Moore, 1995; Moore and others, 1999). Areas outside the mapped stratified-drift aquifers can be classified as till and (or) bedrock. An indicator data set (1 or 0) was used to identify whether a well penetrates through stratified drift or not. No further differentiation of surficial materials is available on a statewide basis.

The thickness of non-stratified-drift deposits is not mapped statewide but can be determined or estimated at individual well sites. Most well drillers' reports contain the depth to bedrock determined during drilling. Where these data are incomplete, the length of casing minus 12 ft can be used as a surrogate for depth of overburden. Twelve feet is the median depth that casings are set into bedrock based on 18,330 wells statewide. Drillers prefer to end the casing well in competent bedrock to avoid the upper few feet, which may be weathered or fractured. Because there is no statewide GIS data set of overburden thickness, this variable could not be used in a predictive mode.

### Topographic Settings and Characteristics

Data were compiled to provide various sets of topographic and physiographic information that may be associated with bedrock well-yield characteristics. Topographic settings can be derived from analysis of USGS Digital Elevation Models (DEMs). DEMs consist of a sampled array of elevations for a number of ground positions at regularly spaced intervals. The DEM data for 7.5-minute units correspond to the USGS 1:24,000 and 1:25,000-scale topographic quadrangle map series. Each 7.5-minute DEM is based on 30 x 30-m data spacing with the Universal Transverse Mercator (UTM) projection. The 1-degree DEMs, with 3 x 3-arc-second data spacing (or roughly 100-m spacing) provides coverage in 1 x 1-degree blocks. The 1-degree DEMs also are referred to as 1:250,000-scale DEM data. Land-surface elevations of wells determined from the 7.5-minute DEMs were included in the model. Lower well yields are expected in high elevations as a result of greater depths to water, less saturated thicknesses, and(or) the predominance of competent rock with few fractures.

Topographic setting also can be analyzed regionally to identify, for example, the effects of major river valleys. This analysis is done by use of 1:250,000 scale DEMs. Effects of more localized features, such as swales or cuts in ridgelines, can be assessed using 1:24,000-scale DEMs. Slope, or the rate of change in elevation surrounding a well, was tested at both scales. Slope at the 1:24,000-scale can be used to compare the change in elevations between neighboring DEM elevation points about 100 ft away from one another. Slope at the 1:250,000-scale DEM is a more regional measurement. Slope expressed as a percent is calculated in ARC/INFO relative to the direction of maximum change in elevation. Assuming similar results to those for North Carolina (Daniel, 1989), slopes or hillsides are expected to have average well yields, and valley bottoms or depressions are expected to have above-average well yields. Wells on hills and ridges are expected to have below-average yields.

Curvature, calculated as the second derivative of the land-surface elevation, provides another quantifiable measure of topographic setting. The curvature function used in this study fits a fourth order polynomial surface to each 3 x 3 block of DEM data points (referred to as GRID cells in ARC/INFO) surrounding and including the central DEM point for the calculated curvature. Negative values indicate that the land surface is concave upward (for example, a swale or valley bottom) and positive values indicate the land surface is concave downward (for example, a hilltop). Again, use of the 1:250,000-scale DEM results in a coarse resolution with a much more generalized regional depiction of the land surface, and cannot be used to detect small features such as depressions in open areas or small gaps in ridges. Curvature at the 1:24,000-scale, however, will not show large physiographic features at the well site. Sites where the topography was concave upward (especially with gentle slopes) represented locations that were eroded more than the surrounding rock and may be underlain by fractures. Conversely, hilltops, which are concave downward, are more resistant to erosion and may be related to a decrease in fracture density and well yield.

Categorical variables were created for combinations of slope and curvature. These variables are “dichotomous,” in that they indicate with a 0 or a 1 whether a given condition is met. For example, the quartile of the well population with steepest site slopes was identified as being on hill slopes (dichotomous variable is set equal to 1). The cross between the most negative quartile of the population, for curvature, with the quartile of the population with the most gentle slopes, was used to identify swales or valley bottoms.

Drainage area, upgradient from each well site, was determined by use of the 1:24,000-scale DEM and a computer process known as “flow accumulation.” First, isolated low points (“sinks”) surrounded by high points are filled in. These points typically represent an artifact of the DEM. Next, the downslope direction between all DEM cells was determined, and the number of cells upgradient of every cell was determined (including that cell). This flow accumulation was done statewide to measure the drainage area of each well site.

The distance of each well to the nearest waterbody was determined by use of the 1:24,000-scale Digital Line Graphs (DLG) of hydrography. For this method, only shorelines and perennial streams were used to compute the distance. Wetlands and intermittent streams were excluded because they are depicted less accurately and consistently on the DLGs than are open perennial waterbodies.

### Additional Variables Investigated at the Quadrangle Scale

Additional geologic, fracture, and lineament data were collected for the Pinardville and Windham quadrangles (fig. 1). These quadrangles were selected because they represent different geohydrologic settings, and because these quadrangles contain the largest number of georeferenced bedrock wells anywhere in the State of New Hampshire. Different degrees of metamorphism have affected the rocks present in these quadrangles with much of the Pinardville quadrangle containing the Massabesic Gneiss (migmatitic) Complex. In the Pinardville quadrangle, there are 1,682 bedrock wells with a complete set of well information in the database, and for the Windham quadrangle, there are 1,504 wells. These data provided a large enough data set to evaluate the effects of additional variables, such as fracture-correlated lineaments, lineaments identified from topographic maps and color infrared imagery, and additional variables identified by use of quadrangle-scale geologic mapping.

Topographic lineaments, drawn on a 1:48,000-scale topographic base reduced from the 1:24,000-scale topographic maps, and lineaments drawn on 1:58,000-scale color infrared photographs were added to the lineament database for the Windham and Pinardville quadrangles. Topographic maps are a readily available platform commonly used for lineament identification. Color-infrared photography especially is sensitive to anomalies in vegetation growth, which may be related to underlying fracture zones. Vegetation type and intensity can vary with ground-water recharge and discharge zones. The same procedures described in Clark, Moore, and others (1996) were followed in this process. Color-infrared photographs with overlays were rectified and both sets of lineaments were digitized into GIS data sets for use in the statistical analysis.

Geologic mapping at the 1:24,000 scale in the Pinardville quadrangle (Thomas R. Armstrong and William C. Burton, U.S. Geological Survey, written commun., 2000) and Windham quadrangle (Walsh and Clark, 1999) provided an accurate representation of lithologies and data on brittle fracture and ductile structure orientation for correlation with lineaments. These maps accurately depict geologic units, contacts, fault, axial trace, metamorphic isograd, dike and other features, which were tested relative to well yields. Simplified versions of the maps showing geologic units are shown in figures 4 and 5.

The geologic units mapped at the quadrangle scale include numerous subdivisions of the geologic units shown on the Bedrock Geologic Map of New Hampshire (Lyons and others, 1997). The digital GIS coverage of the State map, used for development of the regression model, was adjusted where contacts mapped at the quadrangle scale separate units shown on the State map.

Other geologic features in the quadrangles similarly were assessed as the lineaments in the regression model. These features include contacts, faults, dikes, and axial planes of folds large enough to be mapped at a scale of 1:24,000. Indicator variables were created to indicate whether a well was within 100 ft of one of these features. Geologic contacts, faults, axial traces, metamorphic isograds, and dikes also were matched with lineaments in the Windham and Pinardville quadrangles. Lineament segments that fall on, or are parallel within 200 ft of these features, were identified and tested to determine if they were associated with high-yield wells.

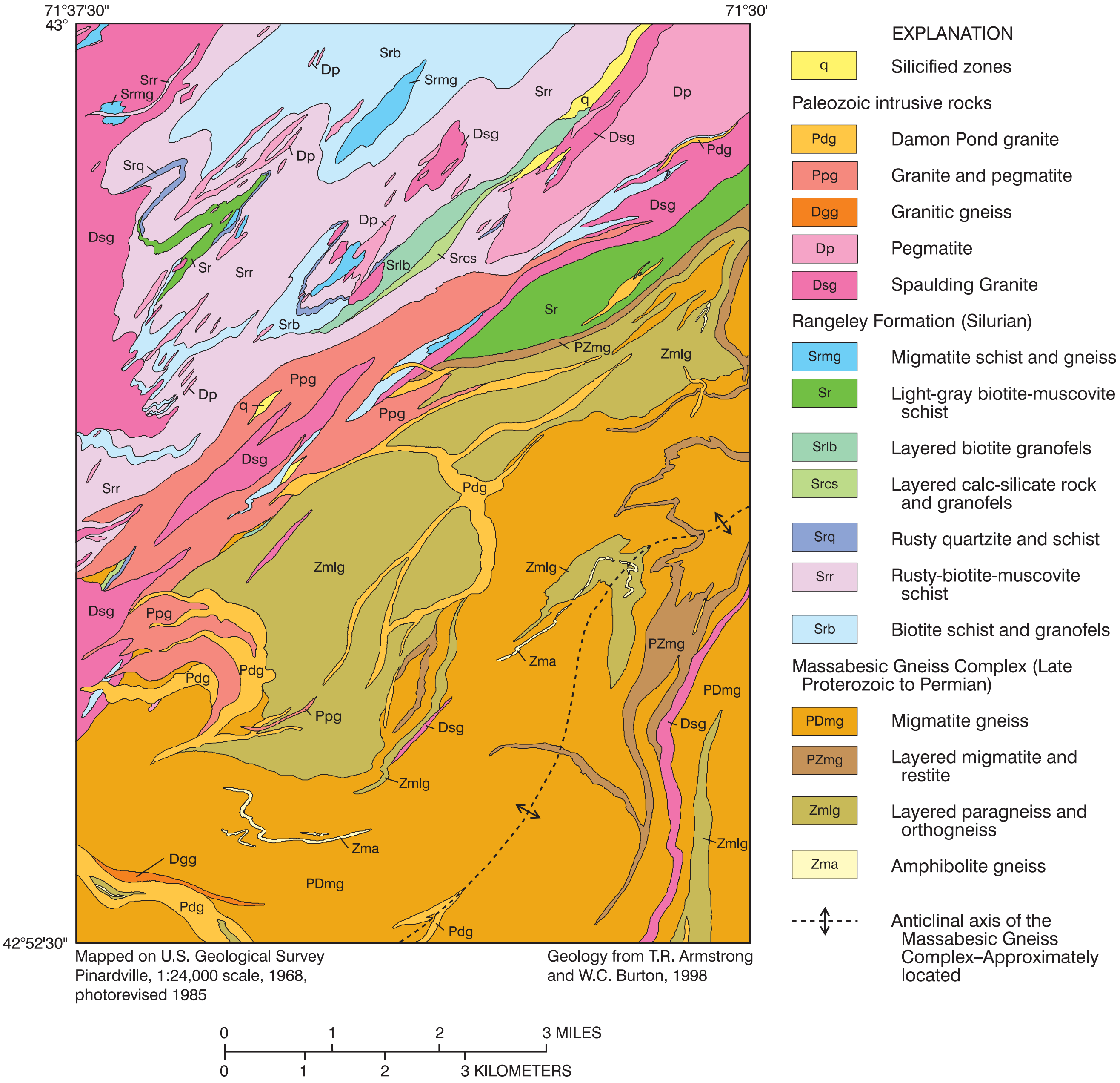
Information on low-angle jointing was not available statewide but was collected for the two quadrangles. The occurrence of low-angle joints near the anticlinal axis of the Massabesic Gneiss Complex (fig. 4) is a prime example of where these features can be related to well yield. An indicator variable was created to test whether or not wells within 2,000 ft of the mapped axis of this broad anticlinal feature had high yields.

Whether some lineaments are more likely to be related to fractures than others has prompted the search for lineaments that are "fracture correlated." Mabee and others (1994) recognized the need to filter, or reduce, lineament data sets before correlating these remotely sensed features with well yield. Their method for identifying fracture-correlated lineaments is based on identifying fracture domains. Every fracture at a few select outcrops (fracture stations) was sampled and the geographic extent of the fractures was identified by comparing fracture families at each fracture station. Fracture families, as defined by Mabee and Hardcastle (1997, p. 24), are "a set of fractures that have similar orientations." A comparison of studies done at different scales by Mabee and others (1994) and Mabee and Hardcastle (1997) shows that the sizes of fracture domains appear to be related to the spacing of fracture stations. Fracture domains, therefore, appear to be scale-dependent and the domain boundaries can be arbitrary if they cannot be correlated to a specific geologic feature.

In the New Hampshire Bedrock Aquifer Assessment, fracture families were derived using fracture data from the entire mapped area rather than just selected stations. Walsh and Clark (2000) compared mapped fracture orientations in the Windham quadrangle to fracture orientations at selected stations (large outcrops in the Windham quadrangle where they collected additional data). The results show a correlation between station fractures and mapped fractures for the quadrangle. However, about one-ninth of the quadrangle could not be correlated and the mapped fractures surrounding a given fracture station may contain fracture families not represented by the fracture-station data. For the development of the regression model, fracture-correlated lineaments are based on the geographically, more extensively mapped fracture data.

Lineament data were analyzed to see if geologic mapping at the quadrangle scale could identify a set of lineaments with improved correlation to high-yield wells. The lineaments were categorized in the model in the following three ways: (1) unfiltered lineaments that include data from the entire statewide database; (2) filtered, domain-based fracture-correlated lineaments; and (3) filtered, discrete-analysis-based fracture-correlated lineaments (fig. 6).

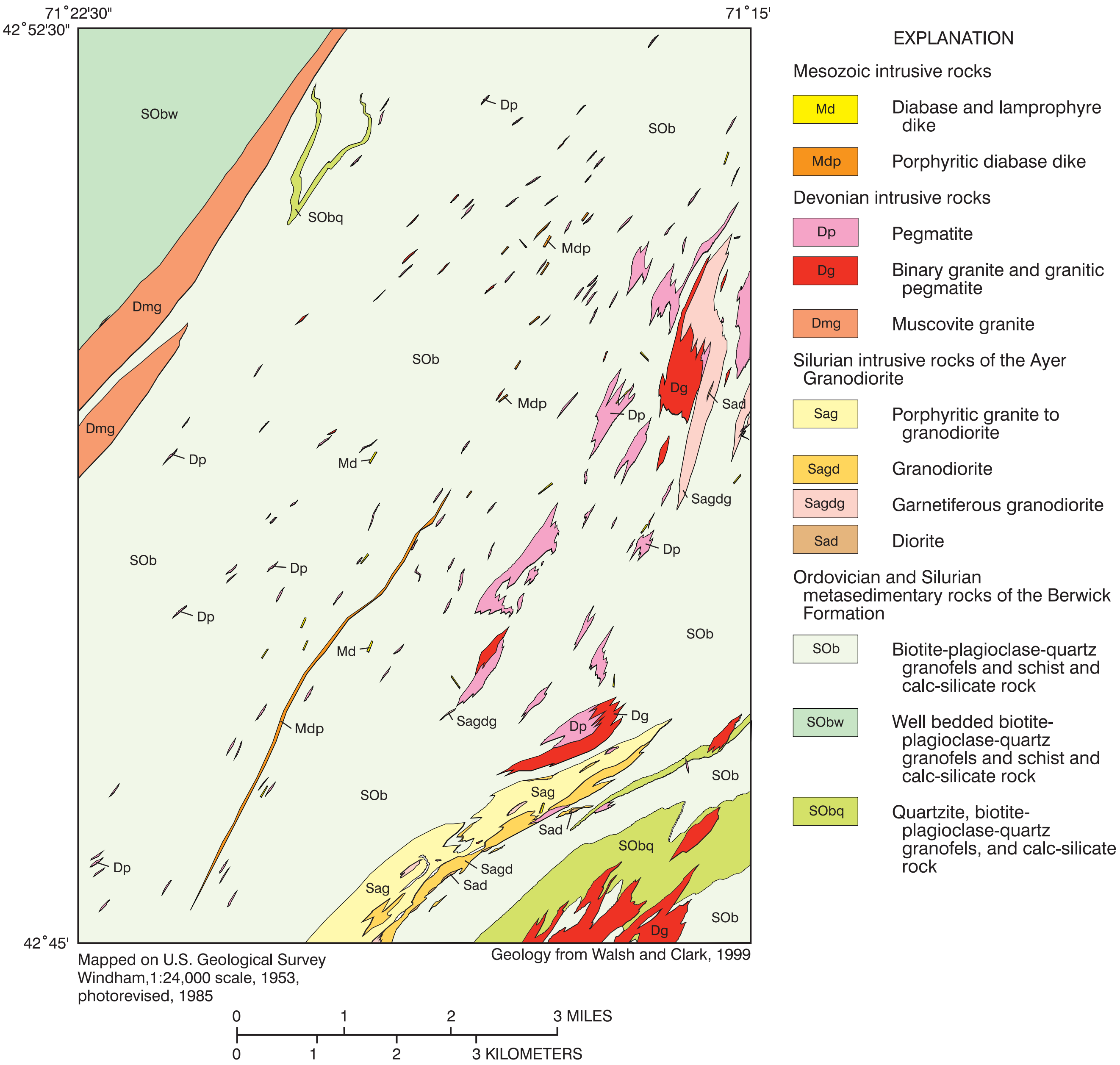
Domain-based fracture-correlated lineaments were determined by defining fracture families for each of the two quadrangles. A square-cell sampling grid, or domain, was developed so that no cell would have less than five sample points. The optimal grid cell measured 3,300 x 3,300 m. Fracture families for each cell were derived from spatial analysis of mapped fracture data for each of the two quadrangles by plotting frequency-azimuth (rose) diagrams in the Structural Data Integrated System Analyzer software (DAISY 2.19) by Francesco Salvini, Dipartimento



Mapped on U.S. Geological Survey Pinardville, 1:24,000 scale, 1968, photorevised 1985

Geology from T.R. Armstrong and W.C. Burton, 1998

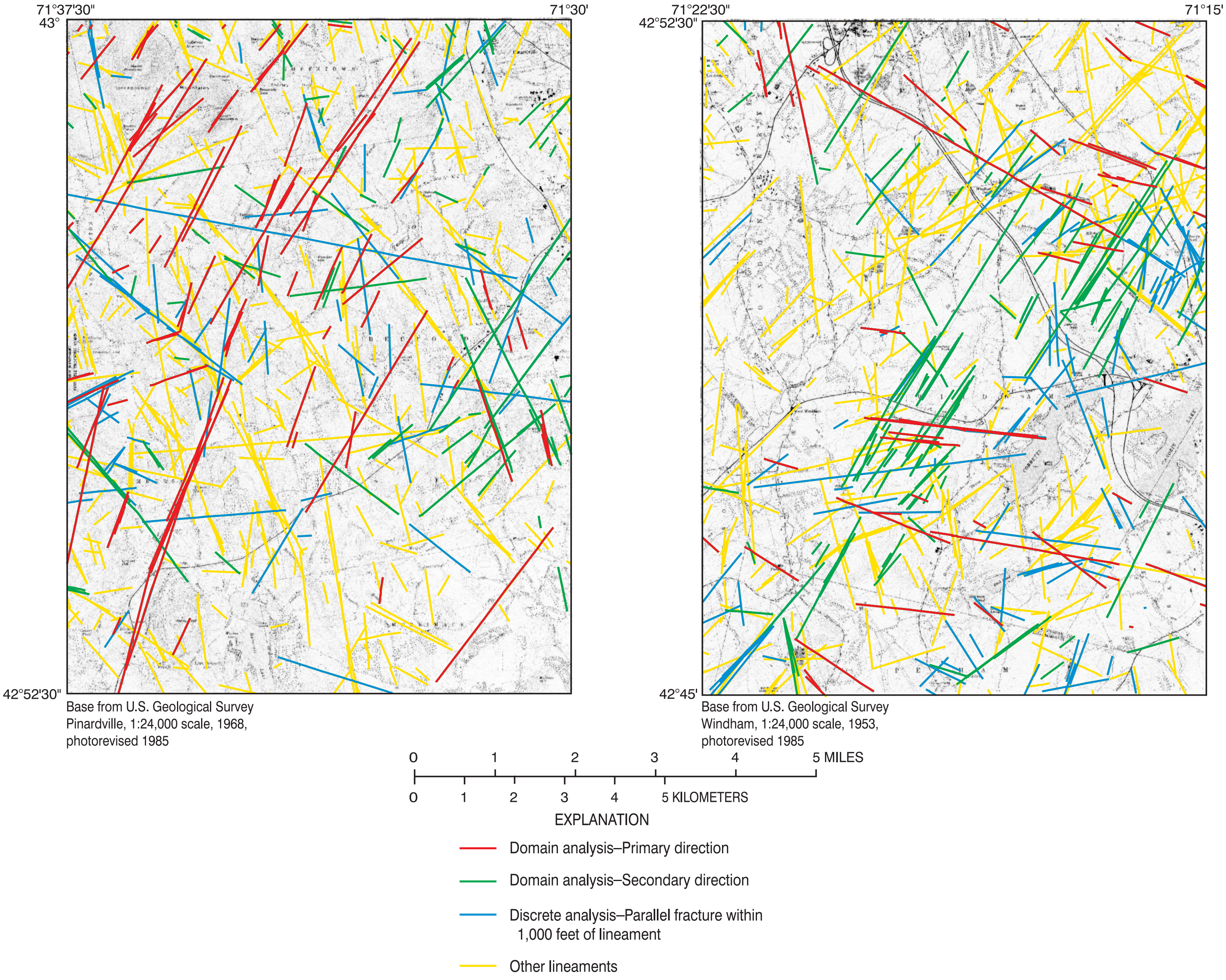
**Figure 4.** Simplified bedrock geologic map of the Pinardville quadrangle, New Hampshire (Location of quadrangle shown in figure 1.)



Mapped on U.S. Geological Survey Windham,1:24,000 scale, 1953, photorevised, 1985

Geology from Walsh and Clark, 1999

**Figure 5.** Simplified bedrock geologic map of the Windham quadrangle, New Hampshire (Location of quadrangle shown in figure 1.)



Base from U.S. Geological Survey
Pinardville, 1:24,000 scale, 1968,
photorevised 1985

Base from U.S. Geological Survey
Windham, 1:24,000 scale, 1953,
photorevised 1985

**Figure 6.** Fracture-correlated lineaments in the Pinardville and Windham quadrangles, New Hampshire.

di Scienze Geologiche, Universita degli Studi di “Roma Tre”, Rome, Italy (computer software, 2000). The DAISY software uses a Gaussian curve-fitting routine for determining peaks in directional data (Salvini and others, 1999) that was first described by Wise and others (1985). The rose diagrams included strike data for steeply dipping fractures (dips greater than 45°, after Mabee and Hardcastle, 1997). In this method, primary fracture families on the rose diagrams were defined as the highest peak, and secondary fracture families were defined as normalized peaks between 30 and 99 percent of the highest peak. Lineaments that are in or pass through a given cell, and are parallel to the fracture family for that cell, are identified as fracture correlated. Parallelism is defined by lineament trends that match primary and secondary fracture families within one standard deviation as determined by the Gaussian curve-fitting routine. The term “fracture correlated” in the domain method follows the definition by Mabee and others (1994) in that the correlated lineaments cannot be used to unconditionally identify fractures on the ground, but merely to imply that fractures are likely in the given area. Results from the domain analysis indicate that, by lineament length, 30.5 percent of the lineaments in the two quadrangles correlate with fracture trends, and 14.6 percent correlate with the primary fracture trend. These filtered, domain-based fracture-correlated lineaments were tested in the regression model.

Discrete-analysis-based fracture-correlated lineaments were determined by use of a more rigorous and conservative method than domain analysis (fig. 6). Domain analysis assumes that a statistically identified fracture family is applied to an entire area or domain, and, thus, only can be used to filter groups of lineaments that match the fracture trend in that domain. The domain approach does not address spatial variability in the domain and spatial variability is inherent in quadrangle-scale fracture-trend analysis (Walsh and Clark, 2000). This discrete-analysis method examines each lineament individually, or discretely, by comparing the strike of each lineament in a data set with the strike of nearby steeply dipping planar features or the trend of linear features in a structural geology data set. The discrete method identified lineaments that correlated with many types of structures (including fractures) measured during geologic mapping. The discrete method identified individual lineaments that matched the strike or trend within ±5° (after Mabee and others, 1994) of the structures observed in outcrops up to 1,000 ft (305 m) from a lineament. Lineaments that match these structures are called “structure-correlated lineaments,” and may be parallel to ductile or brittle features identified during quadrangle-scale geologic mapping. A subset of these lineaments includes the brittle fracture-correlated lineaments, and this subset was tested for inclusion in the regression model. Results of the discrete method indicate that, by lineament length, 30.9 percent of the lineaments were identified as fracture correlated.

## Creation of a Verification Data Set

Twenty percent of the well data were randomly selected statewide and set aside as a verification data set. Sequestered data are used for verification of testing results of the model developed by use of the primary data set. The primary data set contained data from 16,302 wells. The verification data set contained data from 4,050 wells.

## Development of the Statewide Regression Model

Multivariate regression was selected as the primary analytical technique because it is the most efficient tool for predicting a single variable (such as the natural log of well yield) with many potentially independent variables. Multivariate-regression analysis, unlike simple bivariate analyses, examines the added contribution of each parameter in explaining the variance of the dependent variable. The regression model was structured to account for drillers targeting well yields to specific demands. For a group of wells drilled to meet similar demands, such as domestic needs, the result is a minimized range of yields. If well depths were random, or if a uniform depth was drilled at each well site, there would be a greater range, or variance, among the well yields.

Because domestic wells make up most of the well-yield data set, the demand criteria by homeowners needs to be factored into the analysis. Domestic needs generally are met by a yield of a few (usually less than 10) gallons per minute, and the wide range of yields is decreased by targeting similar yields. At a potentially high-yield site, drilling typically is stopped at a shallower depth than at an average site, which effectively reduces the number of deep high-yielding wells in that population. At low-yield sites, drilling is continued to a greater depth to obtain a desired yield, which consequently reduced the number of low-yielding shallow wells. Applying an ordinary least squares model to the relation between yield and depth would result in a model that was biased downward; that is, the coefficient of the depth would be lower than its true value in an unbiased data set.

Assuming that drillers are targeting yield, the variable “DEPTH” was “endogenous” (not independent of yield but developed with yield) and correlated with the residual of the yield regression. A regression of yield on depth violates the regression assumption that predictor variables and the error term are independent. To overcome this problem, the statistical technique of using instrumental variables (Davidson and MacKinnon, 1993) was applied in the New Hampshire Bedrock Aquifer Assessment. By this technique, the model evaluates the effects of demand. The bias, caused by the endogenous depth variable, can be corrected if there are variables (instrumental variables) that are correlated with yield demand but uncorrelated with random variation in the physical yield-depth relation. Initially, an attempt was made to include water use in the model because it is related to demand requirements. The water-use variables tested were the indicator variables that identify if the well is a commercial well or a public supply well. Other factors, likely related to well-construction techniques, such as well diameter, however, cause higher well yields in public supply and commercial wells than yields in domestic wells and water-use variables were, thus, not useful. Other more useful instrumental variables were (1) year drilled, (2) well driller, and (3) median household income based on the 1990 census (U.S. Census Bureau digital data, 1990). These variables are correlated with desired yield but are uncorrelated with random variation in the physical yield-depth relation.

Well data from the Pinardville quadrangle (Lawrence Drew, U.S. Geological Survey, oral commun., 2000) indicated that the yields of newly constructed wells, measured by the drillers, increased with time because of an increase in demand for water. Local drillers confirmed that they gradually are drilling new wells deeper to meet this increased demand. Year of drilling can be used as a surrogate for increased demand, and it was introduced into the regression model as an instrumental variable.

A two-stage regression model with instrumental variables was used to control the depth variable. The principle is depicted in figure 7. In this example, only year of drilling is used as an instrumental variable. Conceptually, it is included in the first stage of the model, which predicts lnDEPTH, the natural log of depth. The predicted lnDEPTH then is used in the second stage as a predictor. By excluding year from the second stage and substituting the predicted lnDEPTH from the first stage, a depth coefficient describing the physical relation of depth to yield, rather than the demand relation, is computed (fig. 7). The independent variables in the model are accounted for and the differences from year to year define the lnYIELD to lnDEPTH relation, providing a coefficient for depth that represents the physical relation.

A statistical analysis of well depth by drilling company, for the 25 major drillers in New Hampshire, indicated that 21 drillers tend to drill deeper or shallower than average. As a result, drilling company also was used as an instrumental variable that is correlated with desired yield but uncorrelated with random variation in the physical yield/depth relation.

The spatial distribution of median annual household income, determined from the TIGER data for the 1990 census (U.S. Census Bureau, 1990), also was used as an instrumental variable (fig. 8). Generally, wells in areas of high-median income are assumed to be drilled deeper to meet a somewhat higher demand for water than in areas of low-median

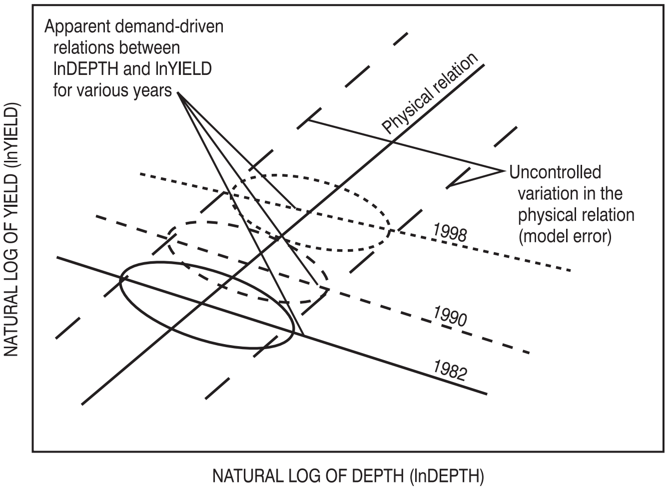


**Figure 7.** Use of the instrumental variable “YEAR,” and a two-stage regression model to separate demand-driven relations from physical relations between well depth (lnDEPTH) and yield (lnYIELD) used in this study of New Hampshire.

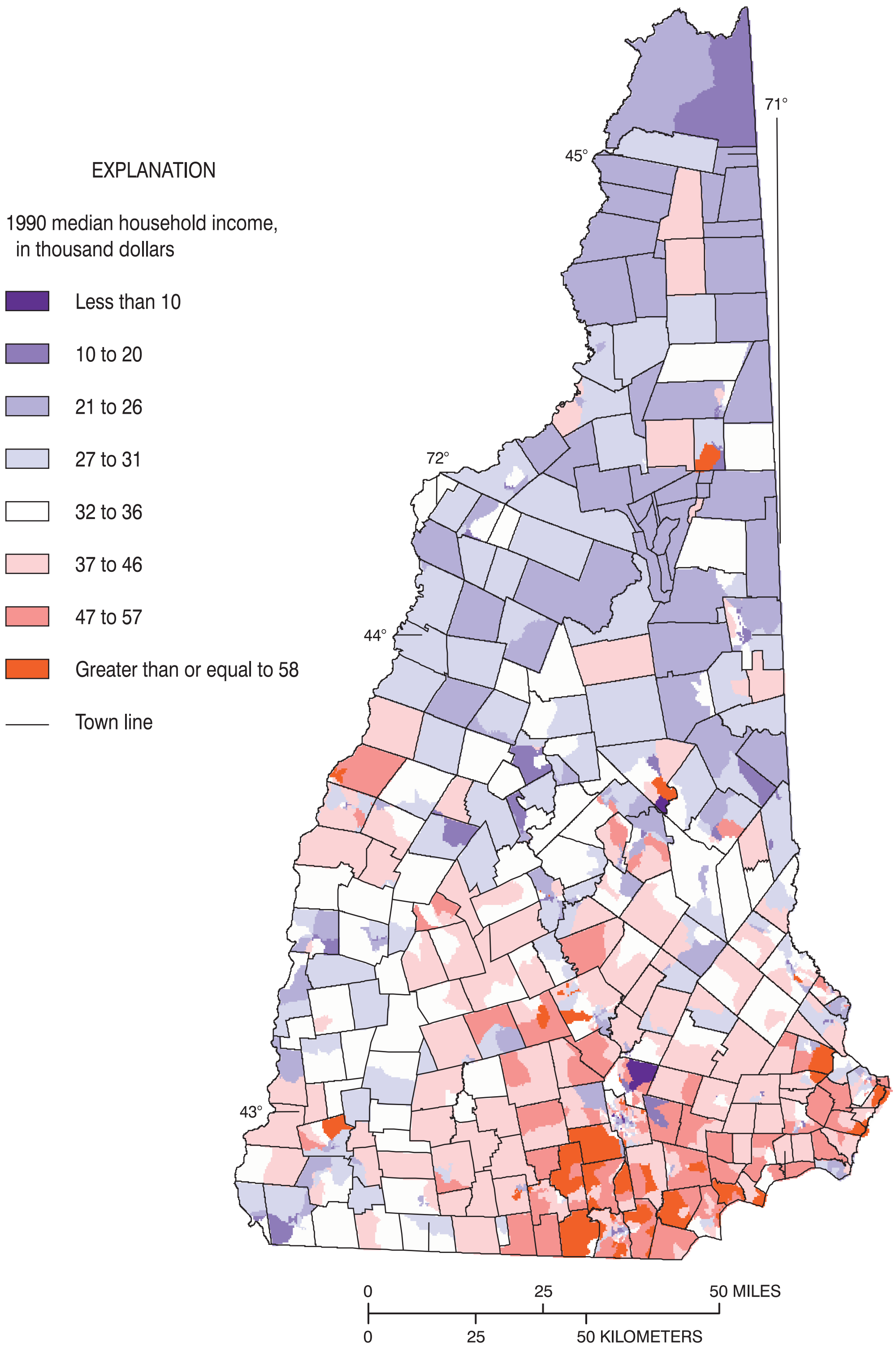


**Figure 8.** Median household income for New Hampshire using the 1990 U.S. Bureau of Census data. Median income was used as a surrogate demand variable (instrumental variable) in the model.