DATABASE OF WELL AND AREAL DATA,
SOUTH SAN FRANCISCO BAY AND
PENINSULA AREA, CALIFORNIA

by David A. Leighton, John L. Fio, *and* Loren F. Metzger

U.S. GEOLOGICAL SURVEY
Water-Resources Investigations Report 94-4151

Prepared in cooperation with the
BAY AREA WATER USERS ASSOCIATION

3015-15

Sacramento, California
1995

U.S. DEPARTMENT OF THE INTERIOR
BRUCE BABBITT, Secretary

U.S. GEOLOGICAL SURVEY
GORDON P. EATON, Director

CONTENTS

FIGURES

TABLES

CONVERSION FACTORS, VERTICAL DATUM, ABBREVIATIONS, ACRONYMS, AND WELL-NUMBERING SYSTEM

Conversion Factors

Multiply	By	To obtain
centimeter (cm)	0.3937	inch
meter (m)	3.281	foot
kilometer (km)	0.6214	mile
hectare (ha)	2.471	acre
square kilometer (km^2)	0.3861	square mile
liter per second per meter [(L/s)/m]	4.831	gallon per minute per foot

Temperature is in degrees Celsius (°C), which can be converted to degrees Fahrenheit (°F) by the following equation:

$$°F=1.8(°C)+32$$

Vertical Datum

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

Abbreviations

µS/cm   microsiemen per centimeter at 25 degrees Celsius
meq/L   milliequivalent per liter
mg/L   milligram per liter

Acronyms

ACWD   Alameda County Water District
ASCII   American Standard Code for Information Interchange
CADWR   California Department of Water Resources
DEM   digital elevation model
GIS   geographic information system
NAD 27   North American Datum of 1927
RMSE   root-mean-square error
SCVWD   Santa Clara Valley Water District
TIN   triangulated irregular network
UTM   Universal Transverse Mercator

Well-Numbering System

Wells are identified and numbered according to their location in the rectangular system for the subdivision of public lands. The identification consists of the township number, north or south; the range number, east or west; and the section number. Each section is further divided into sixteen 16.2-ha tracts lettered consecutively (except I and O), beginning with "A" in the northeast corner of the section and progressing in a sinusoidal manner to "R" in the southeast corner. Within the 16.2-ha tract, wells are sequentially numbered in the order they are inventoried. The final letter refers to the base line and meridian. In California, there are three base lines and meridians; Humboldt (H), Mount Diablo (M), and San Bernardino (S). All wells in the study area are referenced to the Mount Diablo base line and meridian (M) and the designation is not included in the well number. Well numbers consist of 13 characters and follow the format 004S001W19R02. Well numbers are abbreviated and written 4S/1W-19R2. Wells in the same township and range are referred to only by their section designation, 19R2. The following diagram shows how the number for well 4S/1W-19R2 is derived.
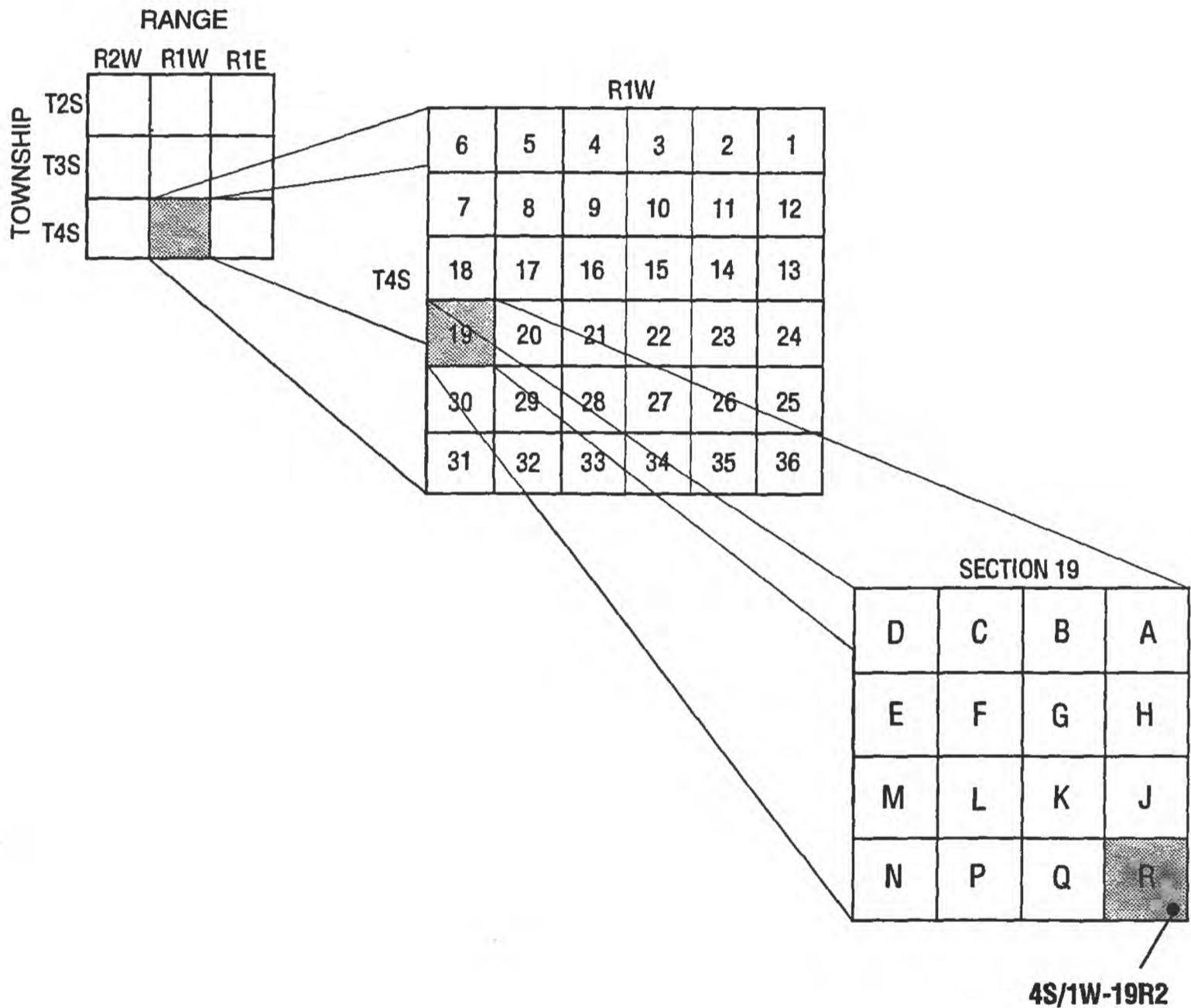
4S/1W-19R2

DATABASE OF WELL AND AREAL DATA, SOUTH SAN FRANCISCO BAY AND PENINSULA AREA, CALIFORNIA

By David A. Leighton, John L. Fio, *and* Loren F. Metzger

Abstract

A database was developed to organize and manage data compiled for a regional assessment of geohydrologic and water-quality conditions in the south San Francisco Bay and Peninsula area in California. Available data provided by local, State, and Federal agencies and private consultants was utilized in the assessment. The database consists of geographic information system data layers and related tables and American Standard Code for Information Interchange files. Documentation of the database is necessary to avoid misinterpretation of the data and to make users aware of potential errors and limitations.

Most of the data compiled were collected from wells and boreholes (collectively referred to as wells in this report). This point-specific data, including construction, water-level, water-quality, pumping test, and lithologic data, are contained in tables and files that are related to a geographic information system data layer that contains the locations of the wells. There are 1,014 wells in the data layer and the related tables contain 35,845 water-level measurements (from 293 of the wells) and 9,292 water-quality samples (from 394 of the wells). Calculation of hydraulic heads and gradients from the water levels can be affected adversely by errors in the determination of the altitude of land surface at the well. Cation and anion balance computations performed on 396 of the water-quality samples indicate high cation and anion balance errors for 51 (13 percent) of the samples. Well drillers' reports were interpreted for 762 of the wells, and digital representations of the lithology of the formations are contained in files following the American Standard Code for Information Interchange. The usefulness of drillers' descriptions of the formation lithology is affected by the detail and thoroughness of the drillers' descriptions, as well as the knowledge, experience, and vocabulary of the individual who described the drill cuttings.

Additional data layers were created that contain political, geohydrologic, and other geographic data. These layers contain features represented by areas and lines rather than discrete points. The layers consist of data representing the thickness of alluvium, surficial geology, physiographic subareas, watershed boundaries, land use, water-supply districts, wastewater treatment districts, and recharge basins. The layers were created by manually digitizing paper maps, acquisition of data already in digital form, or creation of new layers from available layers. The scale of the source data affects the accurate representation of real-world features with the data layer, and, therefore, the scale of the source data must be considered when the data are analyzed and plotted.

INTRODUCTION

The south San Francisco Bay and Peninsula area is part of the largest urban and industrial center in northern California (fig. 1). The population of the area is growing and currently (1994) exceeds 3 million people. Water needed to support this large population is supplied primarily by surface water from local and distant (greater than 200 km away) watersheds and is delivered directly for consumption or is used to supplement natural aquifer recharge. The Hetch Hetchy and South Bay Aqueducts are the primary conduits for water imported to
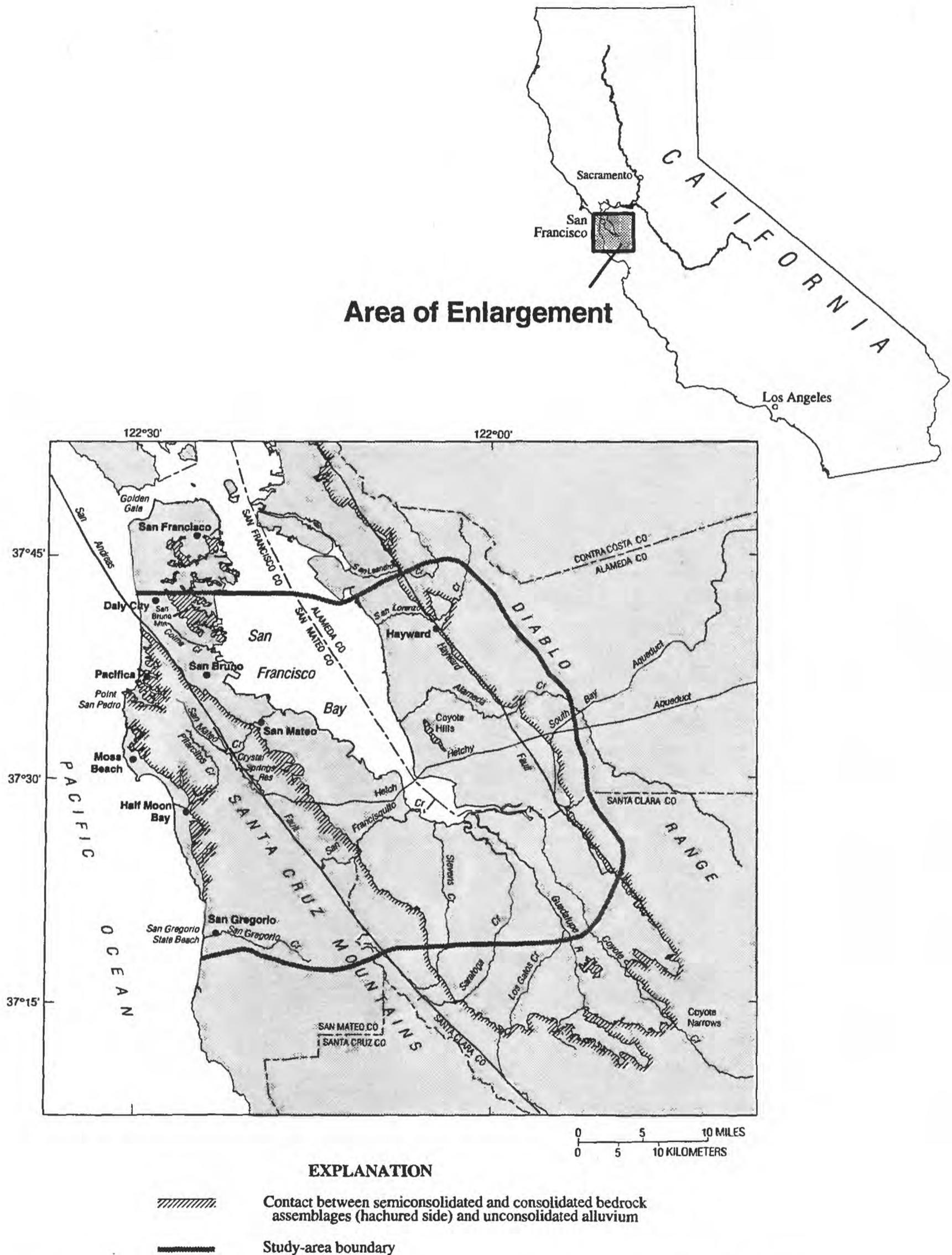
Figure 1. South San Francisco Bay and Peninsula area and study-area boundary, California.

the area. Water supply for the region can be reduced significantly during periods of below normal precipitation or disruptions to aqueduct deliveries as a result of earthquakes or other natural or man-made disasters. Increased withdrawals of ground water from areas that surround the South San Francisco Bay and coastal valleys on the west side of the Peninsula could be a viable source of water to augment available supplies. However, the hydrologic characteristics and boundaries of the aquifer systems underlying the area are not well understood.

The U.S. Geological Survey (USGS), in cooperation with the Bay Area Water Users Association, performed a regional assessment of ground-water resources in the south San Francisco Bay area during the period 1991-94. The study will aid local agencies and planners in understanding the geohydrologic system and in developing strategies and future studies aimed at improving water-use practices. Data from private consultants and local, State, and Federal agencies were provided in different formats, including internal reports, published reports and maps, other computer databases, and hard-copy records. This wide variety of formats and sources did not allow easy evaluation and analysis of the data on a regional scale. To facilitate the assessment, the data were compiled into a computer database consisting of geographic information system (GIS) data layers and related tables and files. A GIS is computer software used to store, manipulate, analyze, and display geographic data. The data layers contain the spatial location of the features, and the related tables and files contain descriptive information about the features. The database includes point-specific data on construction, water levels, water quality, pumping tests, and lithology collected from wells and boreholes (collectively referred to as wells in this report). Areal data layers contain data on the thickness of alluvium, surficial geology, physiographic sub-areas, watershed boundaries, land use, water-supply districts, wastewater-treatment districts, and recharge basins.

The purpose of this report is to describe the database used to assess regional ground-water resources in the south San Francisco Bay and Peninsula area. The report presents the sources and methods used in automating the data, database structure, summary statistics of database contents, uses of the data, and limitations and potential errors associated with the data. Documentation of the source, accuracy, scale, and method of automation of the data is provided to avoid misinterpretation and to make future users aware of data limitations.

STUDY AREA

The study area encompasses southwestern Alameda County, northern and central San Mateo County, and northern Santa Clara County (fig. 1). The predominant geohydrologic feature of the region is a large inland valley that surrounds the southern part of the San Francisco Bay. This inland valley, bordered by the Diablo Range on the east and the Santa Cruz Mountains on the west, is characterized by alluvium deposited by streams that drain the surrounding mountains to form the major aquifers of the region. Ground water in the alluvium is generally unconfined at higher altitudes and semiconfined to confined at lower altitudes. Precipitation, stream runoff, and infiltration from percolation ponds provide most of the recharge to the aquifer system at higher altitudes. Most pumpage in the study area is from the deeper, confined and semiconfined zones in the inland valley.

The part of the study area adjacent to the Pacific Ocean consists of small coastal plains and valleys. These lowlands are characterized by marine terraces created by historic changes in sea level and alluvial valleys deposited by streams that drain the Santa Cruz Mountains to the east. Ground water typically is unconfined in the coastal terraces. The aquifers in this area consist of semiconsolidated and unconsolidated marine and nonmarine sediments. The mountainous part of the study area consists primarily of consolidated bedrock and small alluvial valleys. Ground-water withdrawal in this area is variable and is used only locally.

Before 1965, most of the water supply in the region was ground water. Large declines in hydraulic head as a result of extensive ground-water pumping in parts of the study area resulted in land subsidence and saltwater intrusion. Since 1965, increasing quantities of surface water imported from distant watersheds (greater than 200 km away) have been used to reduce the demand for ground water and to supplement natural aquifer recharge. As a result of the spatial variability in pumping, subsidence, and saltwater intrusion, management of ground-water basins and data collection can vary from intensive to nonexistent (Fio and Leighton, 1994).

IMPLEMENTATION OF THE DATABASE

A computer database was developed to organize, manage, and analyze the data on a regional scale. This database contains information on the location, construction, water levels, water quality, pumping tests, and lithology of the formation for wells located throughout the study area. In addition, areal data describing the thickness of alluvium, surficial geology, physiographic subareas, watershed boundaries, land use, water-supply districts, wastewater-treatment districts, and recharge basins are included in the database.

The GIS utilized for this study was ARC/INFO Version 6.1.1 operated on a Data General Aviion 300 Series workstation with the UNIX operating system. Two primary types of geographic information are digitally stored in the GIS: (1) spatial information that describes the location and shape of a geographic feature, and (2) information about the geographic feature, known as attributes (Environmental Systems Research Institute, Inc., 1991a). A geographic feature is any physical or conceptual object that can be located spatially on the Earth's surface, such as wells, altitude contour lines, and water bodies. Attributes are data that describe the geographic feature, such as well depth, altitude along the contours, and names of lakes. The ability to store descriptive data about a geographic feature in a GIS as well as its spatial location, make the GIS a powerful tool to perform regional assessments. However, because data can be analyzed and displayed at any scale in a GIS, users must understand its limitations based on the accuracy, scale, and method of automation of the source data.

The geographic data are organized into data layers based on a common feature, such as wells, geology, or land use. Each layer contains spatial data referenced geographically to a common coordinate system, which allows the layers to be overlaid for geographic analysis and modeling as well as for the creation of maps. The descriptive attributes for the geographic features are stored in a series of attribute tables related to the spatial locations of the features contained in the data layers. Descriptive information about each layer, such as scale, source, and method of automation of the data is stored in a series of documentation files that remain with the layer throughout most of the processing performed by the GIS software.

Spatial Information

Geographic features are represented in a GIS in three primary ways: areas, lines, and points. Areas that represent the shape of homogeneous features, such as lakes, counties, and land use, are represented by one or more connecting lines that form a closed polygon. Each polygon contains a point used to identify its attributes. Lines represent linear features too narrow to have an area, such as streams, roads, political boundaries, and contours. Points represent features too small to be represented by an area at the scale of the source data (such as wells and cities), or features that have no area (such as points representing mountain peaks and altitude of land surface) (Environmental Systems Research Institute, Inc., 1991a).

For a spatial database to be useful for display and analysis, all feature locations must be stored in a common projection and coordinate system. A projection is a method of representation of the Earth's spherical surface on a flat, two-dimensional surface (such as a map). There are numerous types of map projections and each type produces different distortions; therefore, a projection that will minimize the type of distortion that most affects the intended use of the map is necessary. For example, the projection for a navigation map needs to have minimal distortion in direction and distance. The projection of a three-dimensional surface to two dimensions always creates some distortion in shape, area, distance, or direction.

The Universal Transverse Mercator (UTM) projection, a conformal projection, was used for this database. Conformal projections attempt to preserve local shape at the expense of distortions in area, distance, and direction. No projection can preserve accurately the shape of large features. The globe is divided into 60 zones in a UTM projection, each spanning 6° of longitude. Each zone has a central meridian and the zone extends 3° E and 3° W of the central meridian. Distortions of area, direction, and distance are minimal within each zone (Environmental Systems Research Institute, Inc., 1991c). The project study area is fully contained in UTM zone 10. Coordinates in the UTM projection have an accuracy of approximately 40 cm for every kilometer within a zone (Muehrcke and Muehrcke, 1978).

The coordinates of the features in the database are stored in meters and are based on the North American Datum of 1927 (NAD 27). The origin of the UTM coordinate system is defined by the equator and the central meridian of the UTM zone. The coordinates for the layers are stored as single-precision values that allow as many as seven significant digits for each coordinate. As a result of the distance from the equator to the study area, the coordinates in the north and south (y) direction are large and limits on the number of significant digits can cause a loss of precision for this coordinate. To prevent this loss of precision, a value of 4,000,000 m was subtracted from all coordinates in the y direction during the projection process.

Each data layer discussed in this report contains a projection file that stores information that describes the projection and coordinates used. The projection file includes information on the projection type (UTM), projection zone (10), the datum (NAD 27), the units of the coordinates (meters), and the coordinate shift in the y direction (-4,000,000 m). When a data layer is converted from one projection to another, the projection file is updated by the software to reflect the current projection. The projection file remains with the data layer when the layer is moved or copied to another computer location.

Data layers created from available maps were manually digitized using an ALTEK Datatab digitizing table with an AC40 Controller. The digitizing process can introduce errors in the digital representations of map features. Data digitized from paper maps can be distorted slightly because of changes in humidity, which can cause the location of the feature to be digitized incorrectly. The accuracy of the digitizing equipment, as well as the skill of the individual performing the digitizing, also can affect the accuracy of the spatial locations of the features. All manually digitized data in the database were checked against the source maps for potential digitizing errors.

The accuracy and resolution of features in the data layers are limited by the scale of the source data. The smaller the scale of the source map, the less likely the map features will represent accurately the real-world features. Utilization of data obtained at a small scale for analysis and plotting at a much larger scale can result in incorrect interpretation and location of the features. For example, a lake digitized from a large-scale map could represent accurately the true location of the lake. However, the overlay of a road that follows the shore of the lake digitized at a smaller scale could incorrectly show the lake transected by the road.

The scale of the source data needs to be considered to prevent inappropriate application of the data.

Attribute Tables

Every data layer contains a feature-attribute table with one row (record) of data in the feature-attribute table for every feature in the data layer (fig. 2). Descriptive information about the feature is stored in columns (fields) within the feature attribute table. The fields in the feature-attribute tables contain data generated by the GIS software and optional fields added by the user. The user must define the type of data that each optional field will contain.

The definition of all fields, optional or required, consists of the field name, type, and width, in bytes (the number of characters or numbers that can be contained in the field). The four types of fields used in the feature-attribute tables in the database are character (C), integer (I), real (N), and date (D). Character fields can contain any alphanumeric characters such as letters, numbers, and punctuation; integer fields can contain numeric characters without decimal places; real fields can contain numeric characters with decimal places; and date fields must contain a complete and valid date. The definition for real fields also must include the number of decimal places allowed in the field. Dates can be stored in any of the four field types, but applicability of dates for processing within the GIS is limited when dates are stored in character, integer, and real fields.

In addition to the feature-attribute tables, attributes also can be stored in related-attribute tables. Related-attribute tables are associated with the features in the data layer if they contain at least one field that also is contained in the feature-attribute table. This common field is used to relate the related attribute table to the feature-attribute table and, hence, to the specific feature in the data layer (fig. 2). Related-attribute tables differ from feature-attribute tables in the following ways: (1) unlike feature-attribute tables that contain only one record for each feature in the data layer, related-attribute tables can contain one or more records for features in the data layer; (2) some features in the data layer can contain no records in the related-attribute table; (3) all fields in related-attribute tables are user-defined; and (4) related-attribute tables can be sorted on any field, but feature-attribute tables should not be sorted.
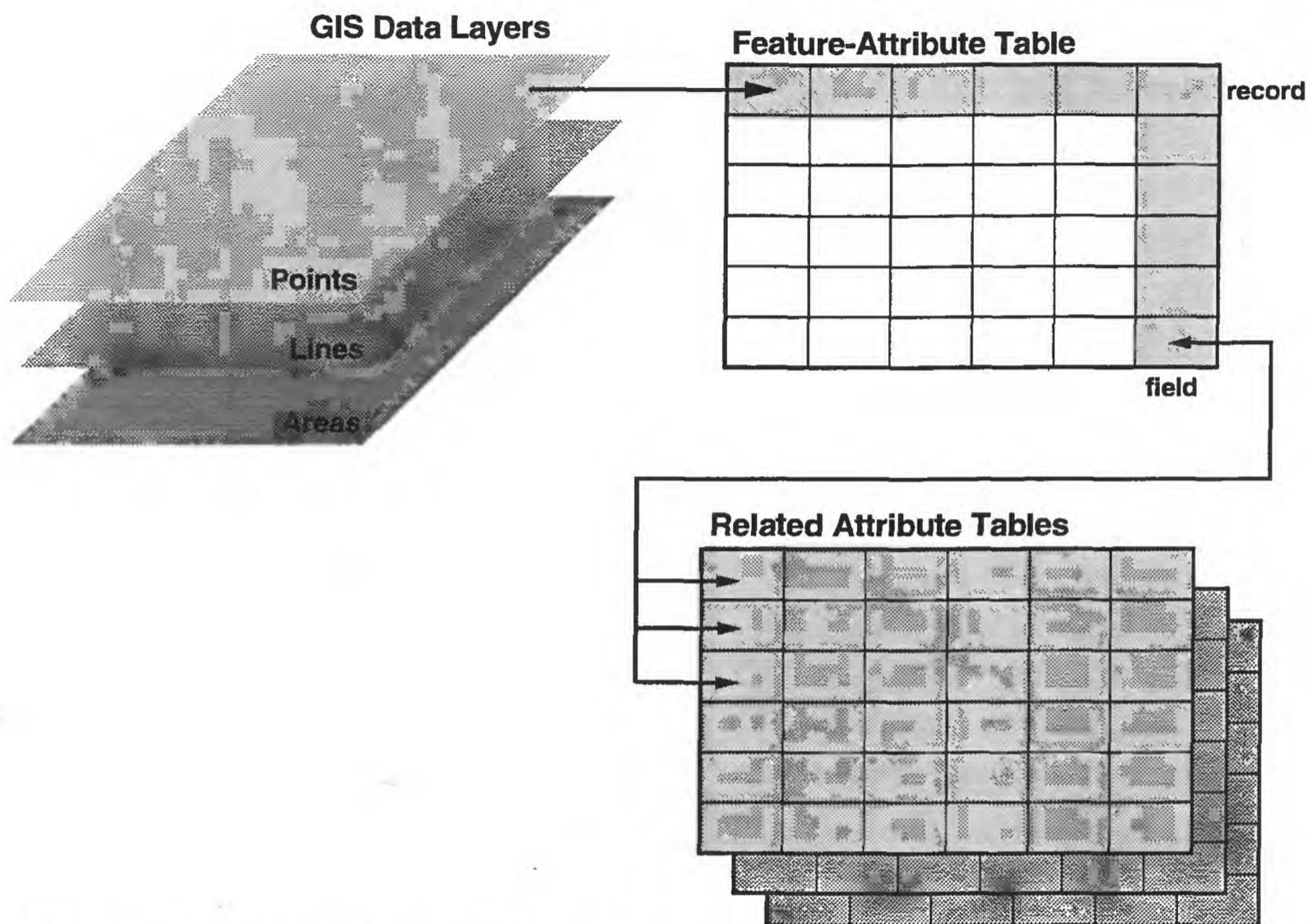
Figure 2. Relation of geographic information system (GIS) data layers, feature-attribute tables, and related attribute tables. (Modified from Environmental Systems Research Institute, Inc., 1991a).

Data Layer Metadata Files

In addition to the feature-attribute tables and projection file described previously, there are three files associated with each layer that contain descriptive information, or metadata, about the layer. The three files are: (1) the documentation file (DOC); (2) the attribute description file (ATT); and (3) the narrative file (NAR). The files are not a standard part of the ARC/INFO software and were created with a program developed by the USGS with the ARC/INFO programming language (Mark Negri, U.S. Geological Survey, written commun., 1994). However, the files can be readily accessed and modified using the GIS software. The files contain much of the information presented in this report, as well as additional information that describes the data layer. These files remain with the data layers as the layers are projected, moved, or copied to another computer location and can be updated periodically as the data layers are modified.

The DOC file contains the basic information about the layer (table 1). Several fields are generated by the software such as the name of the layer, the spatial ex-

tent of its features, and the number and type of features in the layer. Additional fields supplied by the user include the theme and general description of the layer and the scale and citation for the source of the data.

The ATT file contains a description of the attributes in the feature-attribute tables as well as any related-attribute tables associated with the data layer (table 2). Fields generated by the software include the attribute names and definitions. Fields provided by the user include a short description of the attributes, the range of acceptable values to be stored for the attributes, and information about the source and accuracy of the attribute data.

The NAR file contains text supplied by the user that further describes the data layer. This includes a description of the data layer, applications utilizing the data layer, discussion of feature attributes, the procedures used to create or automate the data layer, and revisions and reviews performed on the data layer. The typical format of the NAR file is presented in table 3.

Table 1. Data layer documentation (DOC) file in the database

Description	Field
Revision number of program used to create the documentation	DOC-REV
Documentation file creation date	CREATE-DATE
Individual that updated the documentation file	UPDATE-PERSON
Documentation file update date	UPDATE-DATE
Name of data layer	COVER
Pathname of data layer location	WORKSPACE
Minimum and maximum coordinates of data layer	EXTENT
Precision of data	PRECISION
Tolerances used by software in creation of the data layer	TOLERANCES
Number of arcs in the data layer	NUM-ARCS
Number of segments in the data layer	NUM-SEGS
Number of polygons in the data layer	NUM-POLYS
Number of points in the data layer	NUM-POINTS
Number of tics in the data layer	NUM-TICS
Number of annotation strings in the data layer	NUM-ANNOS
General theme of the data layer	THEME
Short description of the data layer	DESCRIPTION
Contact person for the data layer	CONTACT-PERSON
Contact phone number	CONTACT-INSTRUC
Organization of contact person	ORGANIZATION
Data layer revision number	COVER-REV
Geographic location of the data layer	LOCATION
Resolution of the data layer	RESOLUTION
Scale of the data layer	SCALE
Archive information	ARCHIVE
Publication status	PUB-STATUS
Line 1 of source citation	CITATION-1
Line 2 of source citation	CITATION-2
Line 3 of source citation	CITATION-3
Line 4 of source citation	CITATION-4
Line 5 of source citation	CITATION-5

Table 2. Attribute table documentation (ATT) file in the database

Description	Field
Pathname of data layer location	PATHNAME
Type of items being described	TYPE
Name of file being described	FILENAME
Name of attribute field	ITEMNAME
Width of attribute field	ITEMWIDTH
Width of attribute field for output of reports	OUTPUTWIDTH
Field type (character, integer, real, date)	ITEMTYPE
Number of decimal places for real field type	NUMDECIMAL
Individual that updated the ATT file	USERNAME
Description of attribute	SHORTDEF
Description of values allowed for attribute field	DATADOMAIN
Description of source for attributes	DATASOURCE
Accuracy of values for attribute	ATTACCURACY

Table 3. Data layer narrative (NAR) file in the database

Abstract
   Descriptors and keywords
   Applications that use the data layer
      Intended use of the data layer
      Limitations of the data layer
Feature attribute discussion
Procedures used to create or automate the data layer
Revisions made to the data layer
Reviews applied to the data layer
Related spatial and tabular data layer sets and programs
References cited
Notes

WELL-DATA LAYER

Construction, water-level, water-quality, pumping-test, and lithologic data for the wells were compiled from numerous sources. These data will be used in describing and analyzing the physical and chemical properties of the aquifer system in the study area. The wells were identified with records and databases provided by local, State, and Federal agencies; drillers' reports obtained from the California Department of Water Resources (CADWR); internal and published reports and maps. Because the time frame of the study prevented compilation of all available well data, a subset of 1,014 wells was selected for entry into the well-data layer (fig. 3). The selection criteria were based on the availability of construction, water-level, water-quality, or lithologic data, and the ability to locate accurately the points on a U.S. Geological Survey 7.5-minute 1:24,000-scale topographic map.

The wells were plotted on topographic maps on the basis of written descriptions and maps in the State drillers' reports, maps provided by the well owners (such as water districts), maps published in reports, or parcel numbers and parcel maps. Wells located by parcel number were plotted in the center of the parcel when the exact location was not known. Field validation of the well locations was beyond the scope of this study. Latitude and longitude coordinates determined for each well from the topographic maps were used to create the points that represent the geographic location of the wells. The well identification number was stored in the field WELLNUM in the feature-attribute table to relate the point location to related-attribute tables that contain the descriptive data for each point.

Structure of the Attribute Tables and Files

The points identified in the well-data layer contain the spatial location of the wells that compose the well layer. Well-attribute data are contained in four related-attribute tables and a series of American Standard Code for Information Interchange (ASCII) files. The construction, water-level, water-quality, and pumping-test data are in the related-attribute tables. Every record in the related-attribute tables contains the well-identification number of the well from which the data were collected. Lithologic data are stored in files that follow the ASCII format. Each ASCII file contains the well identification number and lithologic data for a single well. The ASCII files were determined to be the most effective method for managing the lithologic data because of the type of analysis performed on the data, and because the amount of data for each point varies with depth. By convention, construction, water-level, pumping-test, and lithologic data were reported and entered into the database in inch-pound units. All data were converted to metric units for summarizing in this report and in the geohydrologic assessment of the study area.

Construction

The construction table (table 4) lists 21 fields of data on the construction and use of each well and identifiers that indicate if water-level, water-quality, pumping-test, and lithologic data are available for the well. The field types listed in the construction table are character (C), integer (I), and real (N). The number of decimal places allowed in field type real is given in parentheses. The no-data identifier for all fields is 0, except where not applicable. All units of measurement for data in the construction table are in feet except casing diameter, which is in inches.

The WELLNUM field contains the well identification number and is used to relate each record in this table to the geographic location of the well identified by the well-data layer. Each of the 1,014 points in the well-data layer has one record in this table. The HOLE-DEPTH field contains the depth to which the borehole was drilled. The COMP-DEPTH field contains the depth of the completed well and is considered to be the depth to which the casing was installed. The TPERFS and BPERFS fields contain the depths to the top and bottom,
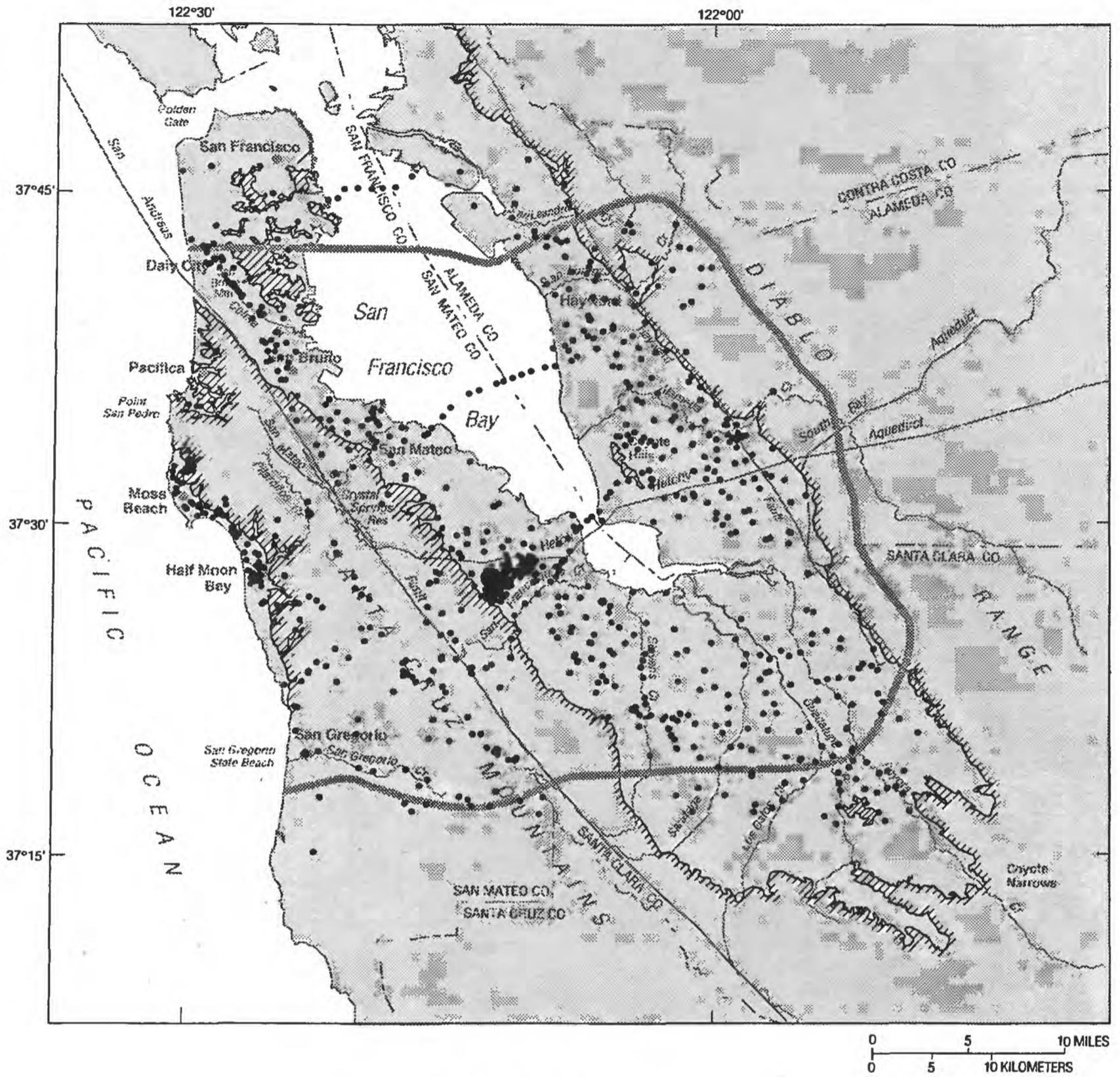
Figure 3. Location of well and borehole data in the database, south San Francisco Bay and Peninsula area, California.

respectively, of the perforated interval. For wells with multiple perforated intervals, TPERFS contains the depth to the top of the uppermost interval and BPERFS contains the depth to the bottom of the lowest interval. The DIAM field contains the diameter of the casing installed in the well. For wells with casings that vary in diameter, the diameter of the casing at land surface was used.

Table 4. Construction information in the database

[Width, in bytes; type: C, character; I, integer; N, real; (2), number in parentheses is number of decimal places in N; na, not applicable; 0, no data]

Description	Field	Width	Type	No data
Well-identification number	WELLNUM	13	C	na
Depth to which the borehole was drilled, in feet below land surface ..	HOLE-DEPTH	4	I	0
Depth of the completed well, in feet below land surface	COMP-DEPTH	4	I	0
Depth to the top of the perforations, in feet below land surface	TPERFS	4	I	0
Depth to the bottom of the perforations, in feet below land surface ...	BPERFS	4	I	0
Diameter of the casing, in inches	DIAM	2	I	0
Identifies if well was gravel packed	GPACK	1	I	0
Depth to top of gravel pack, in feet below land surface	GTOP	4	I	0
Depth to bottom of gravel pack, in feet below land surface	GBOT	4	I	0
Identifies if surface seal was installed	SSL	1	I	0
Date well was completed	DATE-DRLD	6	I	0
Altitude of land surface at well, in feet above sea level	ALT	7	N(2)	0
Possible error in land-surface altitude, in feet	ALT-ACC	5	N(2)	0
Source of land-surface altitude	ALT-SOURCE	2	I	0
Primary use of well	USE	2	I	0
Identifies if drillers' lithologic report is on file	DLOG	1	I	0
Source of data	SOURCE	3	I	na
Identifies if database contains water-level data	WL	1	I	0
Identifies if database contains water-quality data	QW	1	I	0
Identifies if database contains pump-test data	PT	1	I	0
Identifies if database contains lithologic data	LITH	1	I	0

Table 5. Sources of land-surface altitude at well locations in the database

Code	Source
1	California Department of Water Resources drillers' report
2	U.S. Geological Survey 7.5-minute topographic map, 1:24,000 scale
3	Geoconsultants, Inc., 1991, appendix A (altitude of top of well)
4	Santa Clara Valley Water District, written commun., 1991
5	U.S. Geological Survey National Water Information System (altitude determined by leveling)
6	Calculated from water-level measurements (water-level altitude + depth to water)
7	Computer generated from digital-elevation data

The GPACK field is used to identify if a gravel pack was used in the well construction. A code of 1 indicates that a gravel pack was used and a code of 0 indicates that a gravel pack was not used or that there are no data for this field. The GTOP and GBOT fields contain the depths to the top and bottom of the gravel pack, respectively. The SSL field is used to identify if a surface seal was installed in the well during construction. A code of 1 indicates that a surface seal was installed and a code of 0 indicates that a surface seal was not installed or that there are no data for this field.

The DATE-DRLD field contains the date that well construction was completed. Three pairs of digits identify the date. The first pair of digits identifies the year, the second pair identifies the

month, and the third pair identifies the day of well completion. Any pair of digits that contains only zeros indicates that the year, month, or day of completion is not known. The field type normally used for dates requires that a complete and valid date be entered into the table. Because the date of construction for many of the wells was incomplete or unknown, an integer field was required in order to store partial dates. The ALT field contains the altitude of land surface (in feet above sea level) at the well. The ALT-ACC field contains an estimate of the accuracy, in terms of the possible error, in feet, of the land-surface altitude and the ALT-SOURCE field contains a code that identifies the source of the land-surface altitude data (table 5). A code of 0 in the ALT-ACC field indicates that the accuracy of the land-surface altitude measurement is

Table 6. Use of wells in the database

Code	Description
0	Unknown
1	Domestic
2	Irrigation
3	Stock
4	Industrial
5	Commercial
6	Institutional
7	Municipal
8	Cathodic protection
9	Test borehole
10	Observation
11	Dry
12	Other

unknown. Digital-altitude data and the GIS software were used to generate land-surface altitudes for the locations of all wells that did not have land-surface altitude determined from the other sources. The USE field contains a code that identifies the primary well use. Table 6 lists the codes and a description of the types of well use.

The DLOG field is used to identify if a driller's lithologic report for the well is on file. A code of 1 indicates that a report is on file and a code of 0 indicates that no report is on file. The SOURCE field is used to identify the source of the data contained in the construction table. The codes and a description of the sources of data are listed in table 7. Many of the codes listed in table 7 apply only to data in the water-level and water-quality tables (tables 8 and 9) and do not appear in the SOURCE field in the construction table (table 4). The fields WL, QW, PT, and LITH are used to identify if there are data for a well in the water-level, water-quality, or pumping-test tables or lithologic files. A code of 1 indicates that data are available and a code of 0 indicates that no data are available.

Water Levels

The water-level table consists of six fields that contain all water-level data collected for wells in the data layer (table 8). Each record in the table

Table 7. Sources of construction, water-level, water-quality, and lithologic information in the database

Code	Source
1	U.S. Geological Survey, National Water Information System database
2	California Department of Water Resources drillers' reports
3	Alameda County Flood Control and Water Conservation District, written commun., 1991
4	Alameda County Water District, written commun., 1991
5	Santa Clara Valley Water District, written commun., 1991
6	California Department of Health Services, written commun., 1991
7	California Department of Water Resources, written commun., 1992
8	City of San Bruno, written commun., 1993
9	California Water Service Company, written commun., 1992
10	San Mateo County Health Department, written commun., 1992
11	City of San Mateo, written commun., 1992
12	Green Hills Country Club, written commun., 1992
13	Citizen's Utilities Company of California, written commun., 1992
14	Palo Alto Park Mutual Water Company, written commun., 1992
15	Howard Oliver, U.S. Geological Survey, written commun., 1991
16	Coastside County Water District, written commun., 1992
17	City of Milpitas, written commun., 1992
18	Skyline Water District, written commun., 1992
19	City of Cupertino, written commun., 1992
20	Averett and others, 1971
21	Earth Science Associates and Ludhorff and Scalmanini, 1987
22	Geoconsultants, Inc., 1991
23	Earth Science Associates, 1991
24	Atwater and others, 1977
25	Applied Consultants, 1991

Table 8. Water-level information in the database

[Width, in bytes; type: C, character; D, date; I, integer; N, real; (2), number in parentheses is number of decimal places in N]

Description	Field	Width	Type
Well-identification number	WELLNUM	13	C
Date of measurement	DATE	8	D
Depth to water surface in well, in feet below land surface	DTW	6	N(2)
Altitude of water surface in well, in feet above sea level	WLALT	7	N(2)
Source of data	SOURCE	3	I
Reference point for water-level data	REF	1	I

Table 9. Water-quality information in the database

[Width, in bytes; type: C, character; D, date; I, integer; N, real; (2), number in parentheses is number of decimal places in N; na, not applicable; μS/cm, microsiemen per centimeter at 25 degrees Celsius; mg/L, milligram per liter; no data, value stored in field when no data was available for that constituent]

Description	Field	Units	Width	Type	No data
Well-identification number	WELLNUM	na	13	C	na
Date of sample	DATE	na	8	D	na
Specific conductance	SC	μS/cm	5	I	0
Total dissolved solids	TDS	mg/L	5	I	0
pH	PH	na	5	N(2)	0
Calcium	CA	mg/L	5	I	0
Magnesium	MG	mg/L	5	I	0
Sodium	NA	mg/L	5	I	0
Potassium	K	mg/L	5	N(1)	0
Bicarbonate alkalinity	HCO_3	mg/L	5	I	0
Sulfate	SO_4	mg/L	5	I	0
Chloride	CL	mg/L	5	I	0
Silica	SIO_2	mg/L	5	I	0
Nitrate, as nitrate	NO_3	mg/L	5	N(1)	0
Boron	B	mg/L	5	N(2)	0
Bromide	BR	mg/L	5	N(2)	0
Trace metal data	TRACE	na	1	I	0
Organic data	ORG	na	1	I	0
Isotope data	ISO	na	1	I	0
Source of data	SOURCE	na	3	I	na

contains one water-level measurement. Unlike the construction table (table 4), the water-level table can contain multiple records of data for a specific well. The WELLNUM field contains the well-identification number. The DATE field contains the date that the water-level measurement was made. For measurements with only the month and year noted, the 15th day of that month was used as the date of the measurement. The DTW field contains the depth to water, in feet below land surface. The WLALT field contains the altitude of the water level, in feet above sea level. The SOURCE field contains a code that identifies the source of the water-level measurement. The REF field identifies the point from which the original water-level measurement is referenced. A code of 1 indicates that the measurement was reported as the depth to water, in feet below land surface, and that the altitude of the water level above sea level was calculated. A code of 2 indicates that the measurement was reported as the altitude of the water level, in feet above sea level, and that the depth to water was calculated. Refer to table 7 for the codes and sources of data.

WATER QUALITY

The water-quality table consists of 20 fields that contain all water-quality data collected for wells in

Table 10. Pumping-test information in the database

[Width, in bytes; type: C, character; N, real; (1), number in parentheses is number of decimal places in N]

Description	Field	Width	Type
Well-identification number	WELLNUM	13	C
Pumping rate, in gallons per minute	RATE	6	N(1)
Water-level drawdown, in feet	DRWDN	5	N(1)
Time of pump test, in hours	TIME	6	N(1)
Specific capacity, in gallons per minute per foot of drawdown	SC	8	N(4)
Remarks on pump-test data	REMARKS	40	C
Source of data	SOURCE	3	1

the database (table 9). Each record in the water-quality table contains the result of one water-quality analysis. Like the water-level table, the water-quality table can contain multiple records of data for a specific well.

For fields that contain the results of chemical analyses, a code of 0 indicates that no data were available for that constituent. A negative number indicates that the concentration of the constituent was below the detection limit of the analyses method. The value that follows the negative sign is the detection limit. The TRACE, ORG, and ISO fields indicate if trace metal, organic, or isotope data, respectively, are available for the well. A code of 1 indicates that data are available and a code of 0 indicates that no data are available. These data were not entered into the database, but hard copies are maintained for each well. The SOURCE field contains a code that identifies the source of the water-quality measurement. Refer to table 7 for the codes and sources of data.

Pumping Tests

The pumping-test table lists seven fields that contain pump-test data for wells in the database (table 10). The WELLNUM field contains the well-identification number. The RATE field contains the pumping rate, in gallons per minute, during the test. For pumping tests with more than one pumping rate reported, the highest rate was entered into the database. The DRWDN field contains the drawdown of the water level in the wells, in feet, during the pump test. For pumping tests with more than one drawdown reported, the largest drawdown was entered into the database. The TIME field contains the length, in hours, of the pumping test.

Specific capacity was calculated from each pump test by dividing the discharge rate by the drawdown. The specific capacity, in gallons per minute per foot of drawdown, is contained in the SC field. The REMARKS field contains information that could further qualify the pumping-test data. The SOURCE field contains a code that identifies the source of the pumping-test data. Refer to table 7 for the codes and sources of data.

Lithology

Drillers' reports contain written descriptions of the rocks and sediments encountered during well drilling. In most cases, these written descriptions provide the only source of information to evaluate the nature of the subsurface material throughout the study area; however, the subjectivity and variability of these descriptions make analyses difficult. To facilitate analysis of the subsurface lithology, the descriptions from the drillers' reports were converted into numeric codes that describe the material encountered in 1-ft (0.3048-m) intervals to the bottom of the well. For further analysis, the lithologic files stored these numeric codes in a form accessible to the GIS and other software.

The structure of the lithologic data files is different from the construction, water-level, water-quality, and pumping-test tables. Because of the variability in the amount of data for each well (dependent on the well depth) and the type of analysis performed on the data, the lithologic data are stored in ASCII files and organized in system directories rather than in the related-attribute tables. The files are organized in 36 system directories, each representing a township in the area. Each directory contains the lithologic data files for wells in the township represented by that directory. The

Line Number	Well ID Number	Depth of Well	Detail of Log
1	006S003W04E01	250	2
	Depth of Interval	DEP	DESCR
2	1	1	4
3	2	1	4
4	3	1	4
5	4	1	4
6	5	1	4
·	·	·	·
·	·	·	·
·	·	·	·
·	·	·	·
·	·	·	·
·	·	·	·
249	248	1	6
250	249	1	6
251	250	1	6

EXPLANATION

Shaded areas are shown for clarity only and are not part
of the file

Depth is in feet below land surface

DEP--Unconsolidated or consolidated sediments

DESCR--Lithologic descriptor

Figure 4. Example of a file of lithologic data in the database.

naming convention for the files is based on the section, 16.2-ha subdivision, and sequence number of the well-identification number. For example, the lithologic data for well 006S003W04E01 is contained in the file W04E01 in the directory T6SR3W. All file names begin with the letter W, which signifies that the data are from a specific well. The letter W was used to accommodate many computer systems that require file names to begin with an alphabetic character.

An example of a file that contains lithologic data is shown in figure 4. The first line of each file consists of three fields. The first field contains the well-identification number. The second field contains the total depth of the well. The value in this field is equal to the value in the HOLE-DEPTH field in the construction table (table 4). The third field contains a code that identifies the relative detail of the drillers' lithologic descriptions. The remaining records in the file contain interpreted lithologic data for the well. The first field contains the depth of the interval, in feet below land surface. The second field (DEP) contains a code identifying the degree of consolidation of the interval. The third field (DESCR) contains the code that identifies the lithologic description and oxidation state of the interval.

Table 11. Lithologic descriptions and characteristics in the database

[DEP, (1) unconsolidated or (0) consolidated; DESCR, lithologic descriptor and oxidation state. Negative value, sediment is reduced; positive value, sediment is oxidized. Reduced colors, blue, green, and gray; oxidized colors, yellow, brown, tan, red, or no color]

Description	DEP	DESCR
Unknown	0	0
Rock, stone, shale, fractured shale . . .	0	1,-1
Hard or cemented clay, hardpan	0	2,-2
Mud, clay, silt, silty clay, peat, clayey silt	1	3,-3
Sandy or gravelly clay, sandy silt, top soil	1	4,-4
Tight, hard, or cemented sand, sandstone, tight, hard or cemented gravel . . .	0	5,-5
Clayey or silty sand, clayey or silty gravel	1	6,-6
Sand, gravelly sand	1	7,-7
Gravel, sandy gravel	1	8,-8

Typical lithologic descriptions were grouped by similar geohydrologic characteristics and were assigned numeric codes (table 11). The DEP field was used to identify if the material was unconsolidated or consolidated. The interval was assigned a code of 1 if it was unconsolidated and a code of 0 if it was semiconsolidated to consolidated. The DESCR field was used to identify the lithologic description of the interval as well as the oxidation state of the sediment.

Summary of Data Layer Attributes

The attribute data in the related tables and ASCII files are summarized to provide a quantitative description of the contents of the database. The availability of construction data for each well varies because of incomplete records and because many wells were drilled for exploratory purposes and a casing was not installed or was not perforated. Spatial variability in the available data is evident and is due to variability in ground-water management and data collection throughout the region. Geohydrologic analyses performed with the data are affected by limitations of the data.

Contents of the Tables and Files

The HOLE-DEPTH is known for 1,010 of the 1,014 wells, and the distribution of these points based on HOLE-DEPTH is shown in figure 5.
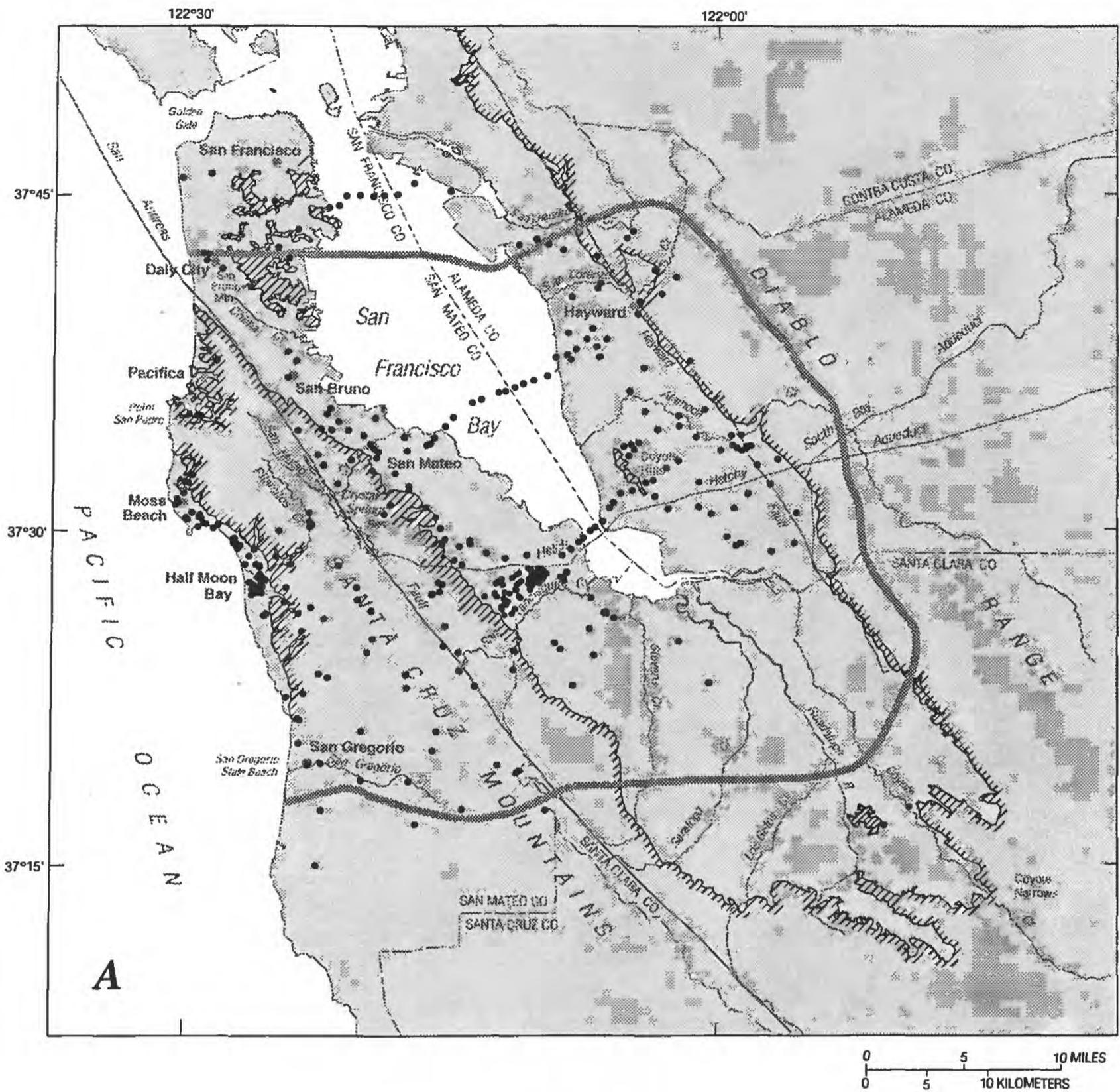
EXPLANATION
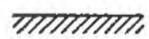
/////////// Contact between semiconsolidated and consolidated bedrock
            assemblages (hachured side) and unconsolidated alluvium

▬▬▬▬▬▬ Study-area boundary

•   Well or borehole location

Figure 5. Location of wells and boreholes included in database that are drilled into *(A)* the shallow zone and *(B)* the deep zone, south San Francisco Bay and Peninsula area, California.

Figure 5A shows the location of the wells drilled into the shallow zone (387 wells); the shallow zone generally is considered to be from land surface to a depth of 60 m. Figure 5*B* shows the wells drilled into the deep zone (623 wells); the deep zone is considered to be greater than 60 m in depth and is the zone from which most ground water is pumped. Many of the wells drilled in the deep zone could be partly or fully perforated in the shallow zone.

Table 12 contains statistics for selected fields in the construction table (table 4). Table 12 identifies

EXPLANATION

/////////   Contact between semiconsolidated and consolidated bedrock
                 assemblages (hachured side) and unconsolidated alluvium

▬▬▬▬▬   Study-area boundary

•   Well or borehole location
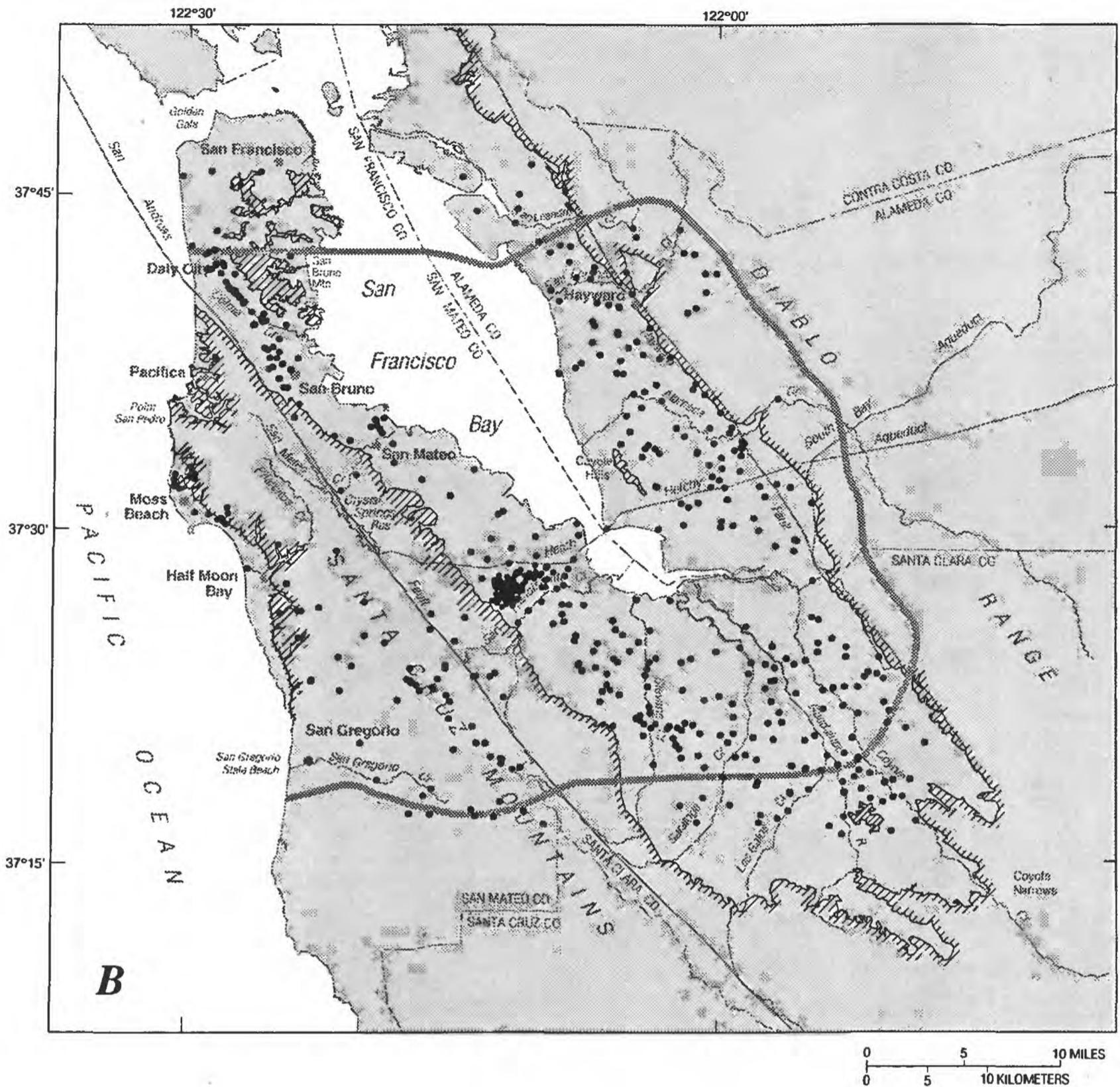
Figure 5. *Continued.*

the number of wells that contain data for a given field and the minimum, maximum, mean, and median values. The HOLE-DEPTH ranges from 2 to 457 m below land surface with a mean of 96 m and a median of 69 m. The perforations range in depth from 1 to 454 m below land surface. The diameter of the casing ranges from 3 to 51 cm with a mean of 22 cm and a median of 20 cm. Typically, wells designed to supply large volumes of water have casing diameters larger than 20 cm. Wells with diameters between 10 and 20 cm generally are used for domestic water supply, and wells with diameters smaller than 10 cm generally are used for observation and monitoring.

Table 12. Statistics for selected fields in the construction table (table 4)

[Values in meters below land surface, except DIAM, in centimeters]

Description	Field	Number of wells	Minimum	Maximum	Mean	Median
Depth to which the well was drilled	HOLE-DEPTH	1,010	2	457	96	69
Depth of the completed well	COMP-DEPTH	827	5	457	91	67
Depth to the top of the perforations	TPERFS	810	1	442	41	24
Depth to the bottom of the perforations . .	BPERFS	810	5	454	87	64
Diameter of the casing	DIAM	824	3	51	22	20

Some wells are perforated as deep as 454 m below land surface; however, most wells are perforated at much shallower depths. Figure 6 shows the distribution of wells perforated at 5-m depth intervals. The interval from 20 to 25 m contains the largest number of perforated wells. Of the 810 wells with perforation data, 325 are perforated at this interval.

The predominance of wells perforated at shallower depths is explained by the analysis of well use and diameter. Of the 325 wells perforated from 20 to 25 m below land surface, 261 (80 percent) are used for domestic water supply and irrigation. Most of the irrigation wells are in an area where irrigation of private residences is the major use of ground water. Large volumes of water, which are typically withdrawn from agricultural irrigation and municipal wells, are not required for these domestic and irrigation wells; therefore, the wells are not usually drilled to great depths. The mean and median diameter of the wells in the interval is 15.5 and 12.7 cm, respectively. This is a further indication that these wells are designed to pump relatively small quantities of ground water.

The water-level table (table 8) contains 35,845 water-level measurements for 293 wells collected from July 1, 1936, to October 15, 1992. HOLE-DEPTH is known for 290 wells, and the distribution of these points on the basis of HOLE-DEPTH is shown in figure 7. Figure 7A shows the location of 92 wells that are drilled into the shallow zone and figure 7B shows the location of 198 wells that are drilled into the deep zone. Many of the wells drilled into the deep zone could be partly or fully perforated in the shallow zone.

Statistics for selected fields in the construction table (table 4) for wells with water-level measurements are presented in table 13. The number of wells with data for a given field and the minimum,

maximum, mean, and median values are listed. The HOLE-DEPTH ranges from 8 to 457 m below land surface with a mean of 125 m and a median of 114 m. The perforations range from 4 to 454 m below land surface. The diameter of the casing ranges from 3 to 51 cm with a mean of 27 cm and median of 30 cm.

The number of water-level measurements available for a well varies from a minimum of 1 for 21 of the wells to a maximum of 965 for well 4S/1W-19R2. The number of wells that have water-level measurements within 10-year periods and the number of measurements in each period is listed in table 14. The period from 1930 through 1939 contains the minimum number of wells with measurements (5) as well as the minimum number of measurements (264). The period from 1980 through 1989 contains the maximum number of wells with measurements (265), and the period from 1970 through 1979 contains the maximum number of measurements (10,539).

The water-quality table (table 9) contains the results from 9,292 water-quality samples collected from 394 wells from August 16, 1937, to August 31, 1992. HOLE-DEPTH is known for 392 of the wells, and the distribution of these wells, on the basis of HOLE-DEPTH, is shown in figure 8. Figure 8A shows the location of 151 wells drilled into the shallow zone and figure 8B shows the location of 241 wells drilled into the deep zone. Many of the wells drilled into the deep zone could be partly or fully perforated in the shallow zone.

Statistics for selected fields in the construction table (table 4) for wells with water-quality data are presented in table 15. The number of wells with data for a given field and the minimum, maximum, mean, and median values are listed. The HOLE-DEPTH ranges from 6 to 457 m below land surface with a mean of 108 m and a median of 82 m. The perforations range from 3 to 384 m below land
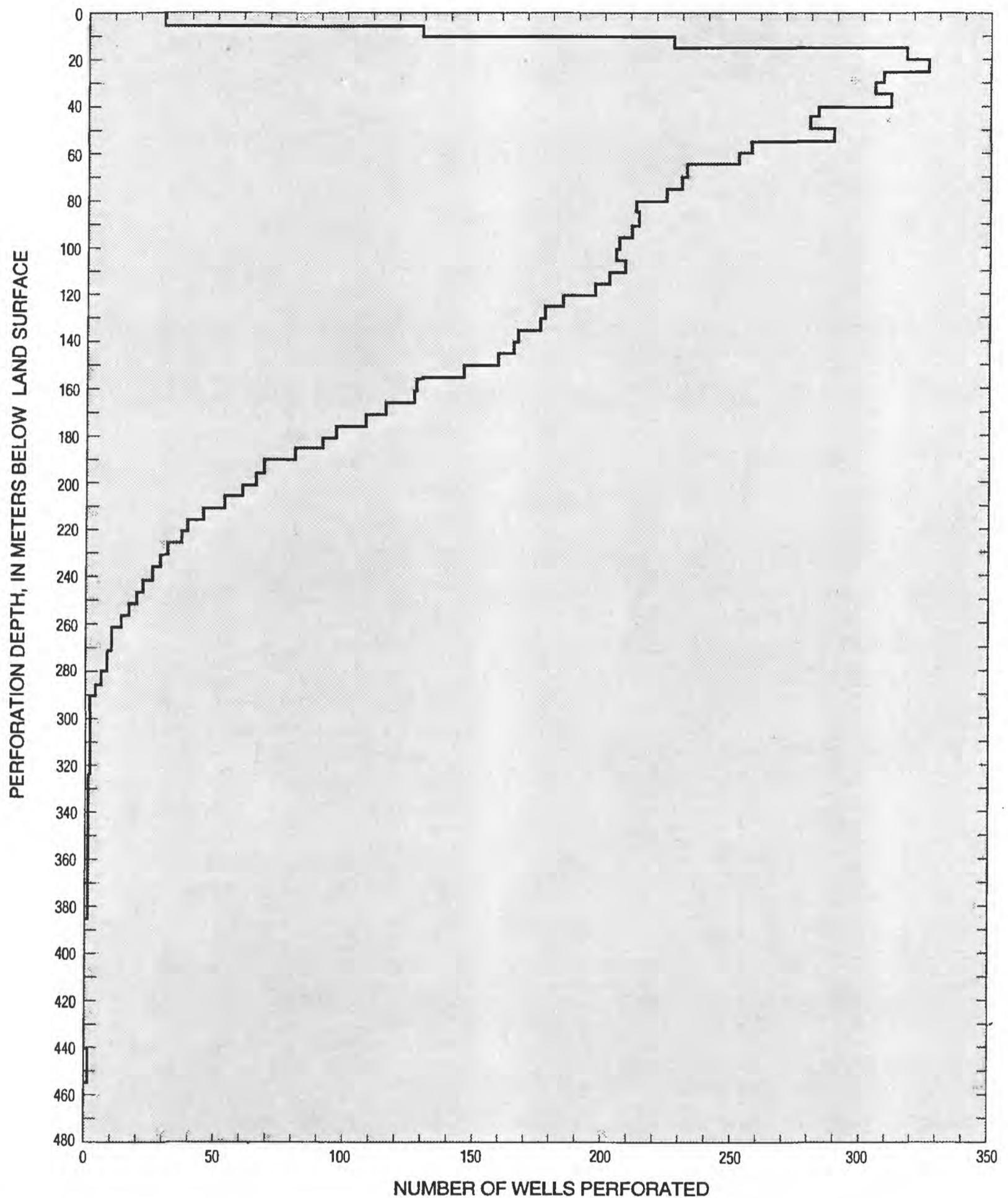
Figure 6. Perforation depths (in 5-meter depth intervals) of wells in the database.

surface. The diameter of the casing ranges from 3 to 48 cm with a mean of 24 cm and median of 25 cm.

The distribution of the number of times each constituent was reported, the total number of samples collected, and the number of wells sampled for 10-year periods are presented in table 16. The constituent most often reported is chloride; this is the result of a large number of samples that were collected to detect the intrusion of saltwater inland from San Francisco Bay. The period from 1930 through 1939 contains the minimum number of wells sampled (1) and the minimum number of
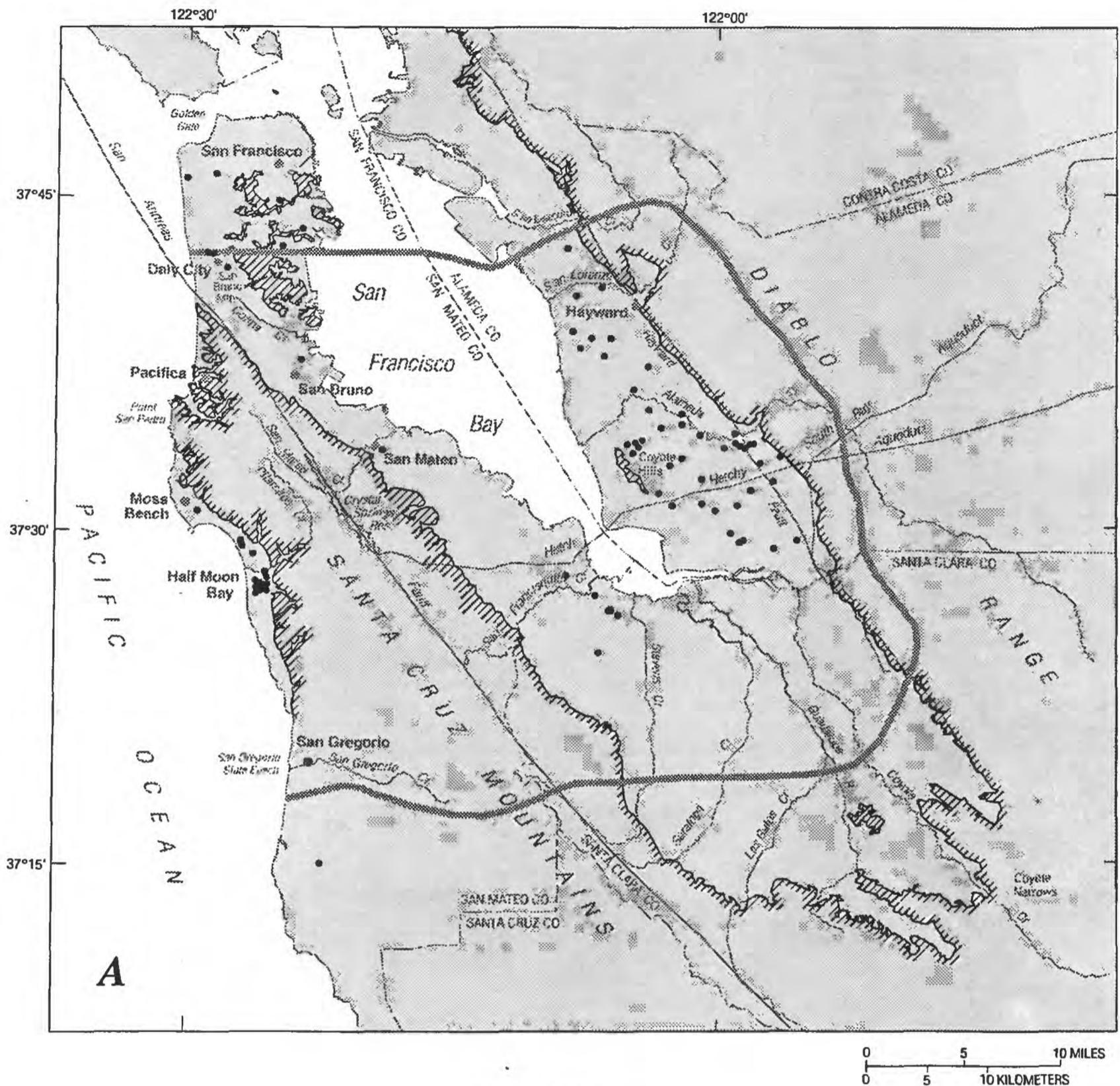
Figure 7. Location of wells with water-level data that are drilled into *(A)* the shallow zone and *(B)* the deep zone, south San Francisco Bay and Peninsula area, California.

samples collected (2). The period from 1980 through 1989 contains the maximum number of wells sampled (309), and the period from 1970 through 1979 contains the maximum number of

samples collected (3,587). The number of water-quality samples reported for a given well ranges from a minimum of 1 for 153 of the wells to a maximum of 606 for well 4S/1W-28D9.
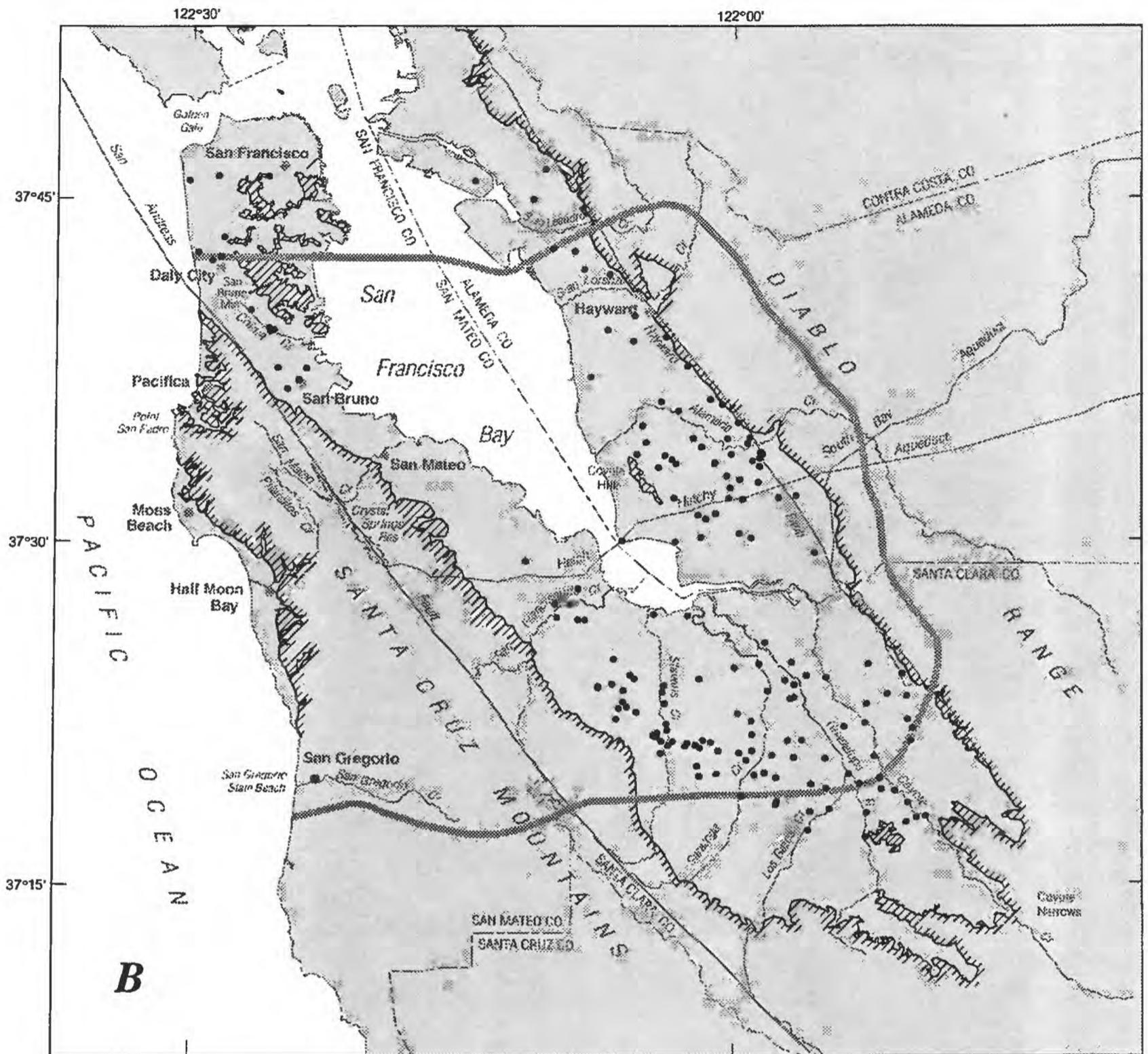
EXPLANATION
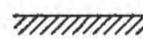
///////// Contact between semiconsolidated and consolidated bedrock assemblages (hachured side) and unconsolidated alluvium

———— Study-area boundary

• Well location

Figure 7. *Continued.*

The pumping-test table (table 10) contains the data from 471 pump tests for wells in the database. Specific capacity calculated from the pumping-test data ranges from 0.0002 to 122 (L/s)/m and the mean and median are 4.6 (L/s)/m and 1.1 (L/s)/m, respectively. HOLE-DEPTH is known for 350 of the wells, and their distribution is shown in figure 9. Figure 9A shows the location of 99 wells drilled into the shallow zone and figure 9B shows the location of 251 wells drilled into the deep zone. Many of the wells drilled in the deep zone could be partly or fully perforated in the shallow zone.

Table 13. Statistics for selected fields in the construction table (table 4) for wells with water-level data in the database

[Values in meters below land surface, except DIAM, in centimeters]

Description	Field	Number of wells	Minimum	Maximum	Mean	Median
Depth to which the well was drilled	HOLE-DEPTH	290	8	457	125	114
Depth of the completed well	COMP-DEPTH	269	11	457	119	113
Depth to the top of the perforations	TPERFS	262	4	442	61	57
Depth to the bottom of the perforations ...	BPERFS	262	9	454	113	107
Diameter of the casing	DIAM	272	3	51	27	30

Table 14. Wells with water-level data for 10-year periods in the database

Period	Number of wells	Number of measurements
1930-39	5	264
1940-49	8	576
1950-59	45	2,520
1960-69	155	9,700
1970-79	199	10,539
1980-89	265	9,462
1990-92	215	2,784

Statistics for selected fields in the construction table (table 4) for wells with pumping-test data are presented in table 17. The number of wells with data for a given field and the minimum, maximum, mean, and median values are listed. The HOLE-DEPTH ranges from 6 to 373 m below land surface with a mean of 111 m and a median of 74 m. The perforations range from 2 to 325 m below land surface. The diameter of the casing ranges from 5 to 48 cm with a mean of 24 cm and a median of 15 cm.

The lithologic files contain descriptions for 762 wells. The distribution of the wells based on HOLE-DEPTH is shown in figure 10. Figure 10A shows the location of 275 wells drilled in the shallow zone, and figure 10B shows the location of 487 wells drilled in the deep zone. The depth of these wells ranges from 2 to 457 m below land surface with a mean of 100 m and a median of 70 m.

Limitations of the Attribute Data

The well-data layer and attribute tables are not a complete compilation of all well data available for the study area. Over 25,000 points that could con-tain useful well data were identified from the sources used to complete the database. Inclusion of all the available data for the study area was not possible within the timeframe of the study; therefore, only a subset of the available data was used to create this database. Additional efforts to compile well data would result in a more complete distribution of data points for use in conducting geohydrologic investigations. For example, in a small area north of San Francisquito Creek, only seven wells were entered in the database during the initial data compilation. Nearly 200 additional wells were located and added to the database for a local ground-water study in this area.

The primary use of water-level data is to calculate hydraulic-head values and gradients. Ideally, the water-level measurements are used to determine the hydraulic-head value for a specific point in the aquifer. Because the water levels in this database usually were collected from wells with perforations over wide intervals, the head values calculated from these water levels do not accurately represent the head value for a specific point in the aquifer. The water-level measurements can be used only to represent the average head over the length of the perforated interval. Also, because the water levels are affected by hydraulic pressure over the complete perforated interval, these water levels may not accurately represent the altitude of the water table.

Hydraulic heads and gradients calculated from water levels are affected by two sources of error. The first source of error is the measuring equipment used to collect the water-level measurement. The three primary methods used to collect the water-level measurements reported in this database are steel tape, electric water-level detector, and the air line. The steel tape generally is considered to be the most accurate method and a measurement can be collected within 0.5 cm of the actual water level.
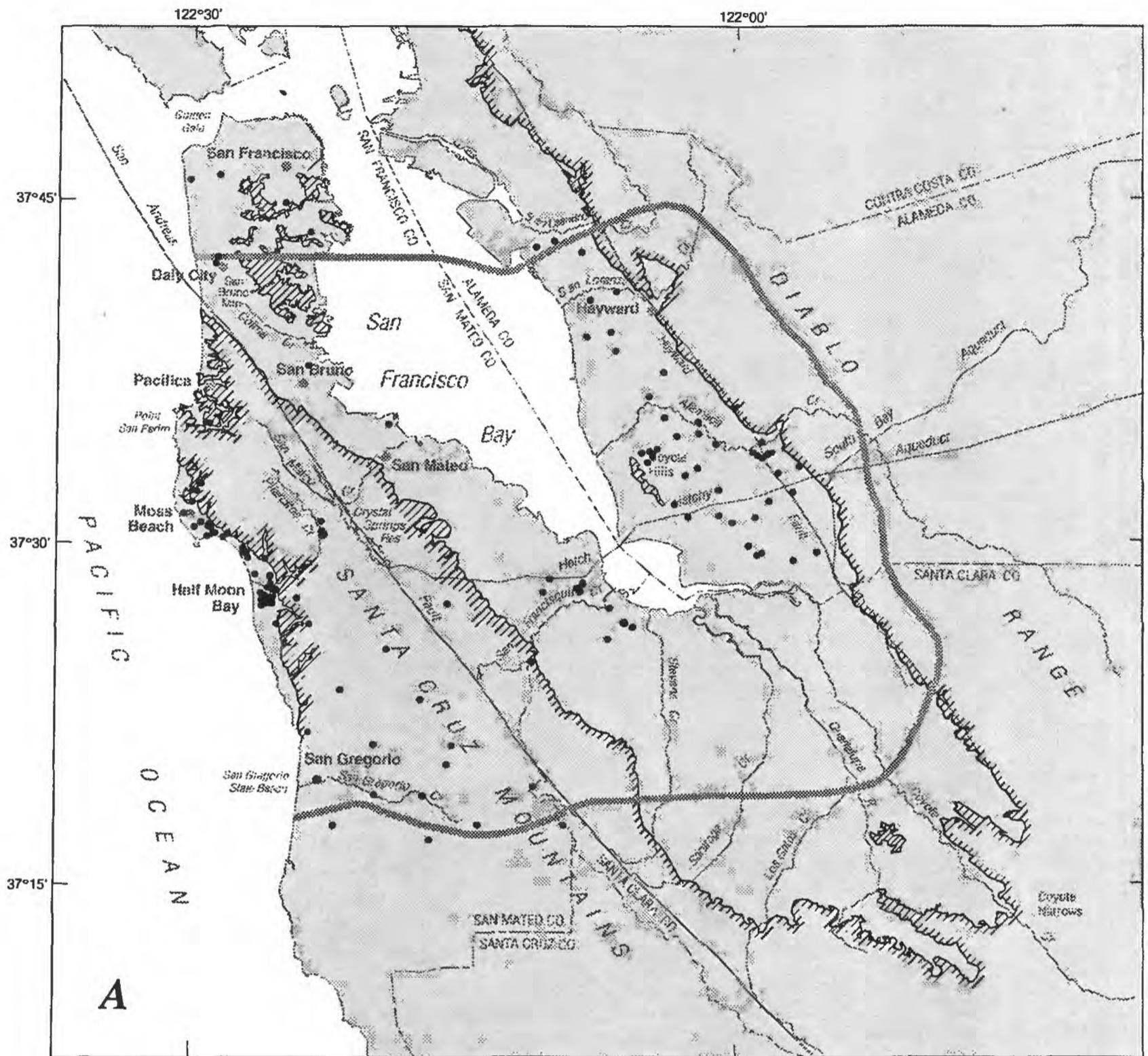
Figure 8. Location of wells with water-quality data that are drilled into *(A)* the shallow zone and *(B)* the deep zone, south San Francisco Bay and Peninsula area, California.

The electric water-level detector typically is not as accurate as the steel tape and a measurement can be collected within 0.5 to 3.0 cm of the actual water level. The air line is the least accurate method and usually a measurement can be collected within 15 to 30 cm of the actual water level. The accuracy of these three methods can be affected by the depth to

water in the well, the condition of the well, and the condition of the measuring equipment.

The second source of error is from the determination of the altitude of land surface at the well. Land-surface altitude at each well in the database was obtained from several sources, such as
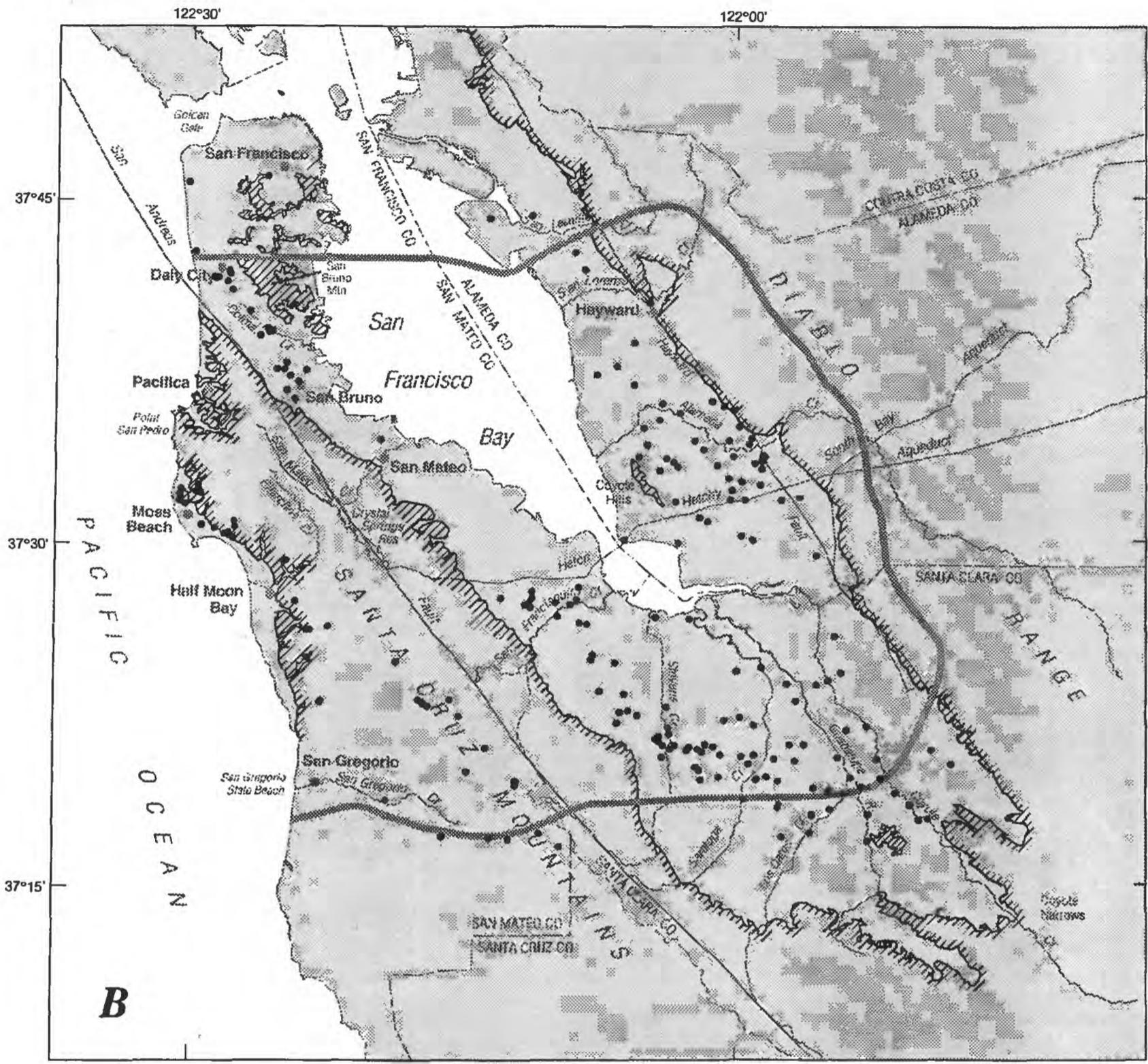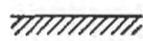
Figure 8. *Continued.*

EXPLANATION

⧄⧄⧄⧄⧄	Contact between semiconsolidated and consolidated bedrock assemblages (hachured side) and unconsolidated alluvium
▬▬▬▬	Study-area boundary
•	Well location

leveling, topographic maps, and interpolated values from digital representations of the land-surface altitude (table 5). The ALT-ACC field in the construction table (table 4) has a range in potential errors in altitude of 0.3 to 610 cm for the 242 wells with a known accuracy. To obtain an estimate of the potential errors associated with altitudes determined with the GIS software through the use of digital representations of land surface, the software-generated altitudes were compared with the altitudes from the other sources for 351 of the wells in the database.

Table 15. Statistics for selected fields in the construction table (table 4) for wells with water-quality data in the database

[Values in meters below land surface, except DIAM, in centimeters]

Description	Field	Number of wells	Minimum	Maximum	Mean	Median
Depth to which the well was drilled	HOLE-DEPTH	392	6	457	108	82
Depth of the completed well	COMP-DEPTH	371	5	389	101	73
Depth to the top of the perforations	TPERFS	358	3	171	47	33
Depth to the bottom of the perforations	BPERFS	358	5	384	97	76
Diameter of the casing	DIAM	367	3	48	24	25

Table 16. Number of wells with water-quality data and number of samples for 10-year periods by constituent in the database

Description	Field	1930-39	1940-49	1950-59	1960-69	1970-79	1980-89	1990-92	Total
Specific conductance	SC	0	0	23	108	688	372	197	1,388
Total dissolved solids	TDS	0	0	18	44	606	278	115	1,061
pH	PH	1	2	27	106	675	202	93	1,106
Calcium	CA	1	2	27	107	150	343	123	753
Magnesium	MG	1	2	27	107	149	341	123	750
Sodium	NA	1	2	33	107	130	348	122	743
Potassium	K	0	0	16	32	69	212	95	424
Bicarbonate alkalinity	HCO$_3$	2	6	27	84	131	345	124	719
Sulfate	SO$_4$	1	2	23	33	79	289	114	541
Chloride	CL	2	6	97	2,586	3,558	2,701	159	9,109
Silica	SIO$_2$	0	0	9	11	28	50	35	133
Nitrogen, nitrate	NO$_3$	0	2	22	33	660	330	127	1,174
Boron	B	0	0	0	0	0	13	6	19
Bromide	BR	0	0	0	0	0	14	6	20
Trace metal data	TRACE	0	0	11	19	42	138	106	316
Organic data	ORG	0	0	0	0	45	28	21	94
Isotope data	ISO	0	0	0	0	4	12	14	30
Total number of samples		2	6	97	2,600	3,587	2,763	237	9,292
Wells sampled		1	2	31	71	127	309	160	394

The difference between the software-generated altitudes and the altitudes from other sources (see table 5) ranged from -79.95 to 27.86 m with a mean of -0.94 m and median of 0.01 m. Spatial analysis of the magnitude of the errors indicates that the computer-generated altitude is generally an underestimate of the altitude in areas of high topographic relief and an overestimate of the altitude in areas of low relief. To confirm this analysis, altitudes generated by the software were compared with altitudes determined by leveling at 14 sites in San Mateo County. These sites were in an area of low topographic relief and were leveled as part of a local ground-water study in the area north of San Francisquito Creek. The altitudes for 11 of the 14 sites were overestimated by application of the software with a mean and median difference for all 14 sites of 1.2 and 0.8 m, respectively.

The altitudes of land surface generated by the GIS software potentially have the greatest deviation from the actual altitude. The errors in the altitude of land surface are greater than the errors associated with the water-level collection methods and are more likely to adversely affect the calculation of hydraulic-head gradients and estimates of ground-water flow between points.

The data entered into the water-quality table (table 9) are the result of samples collected and analyzed by various agencies and laboratories. The accuracy of the constituent values reported could vary substantially due to the application of analytical techniques and quality-control procedures at the time of analysis. To obtain a measure of the accuracy of the water-quality data, cation- and anion-balance computations were performed on samples
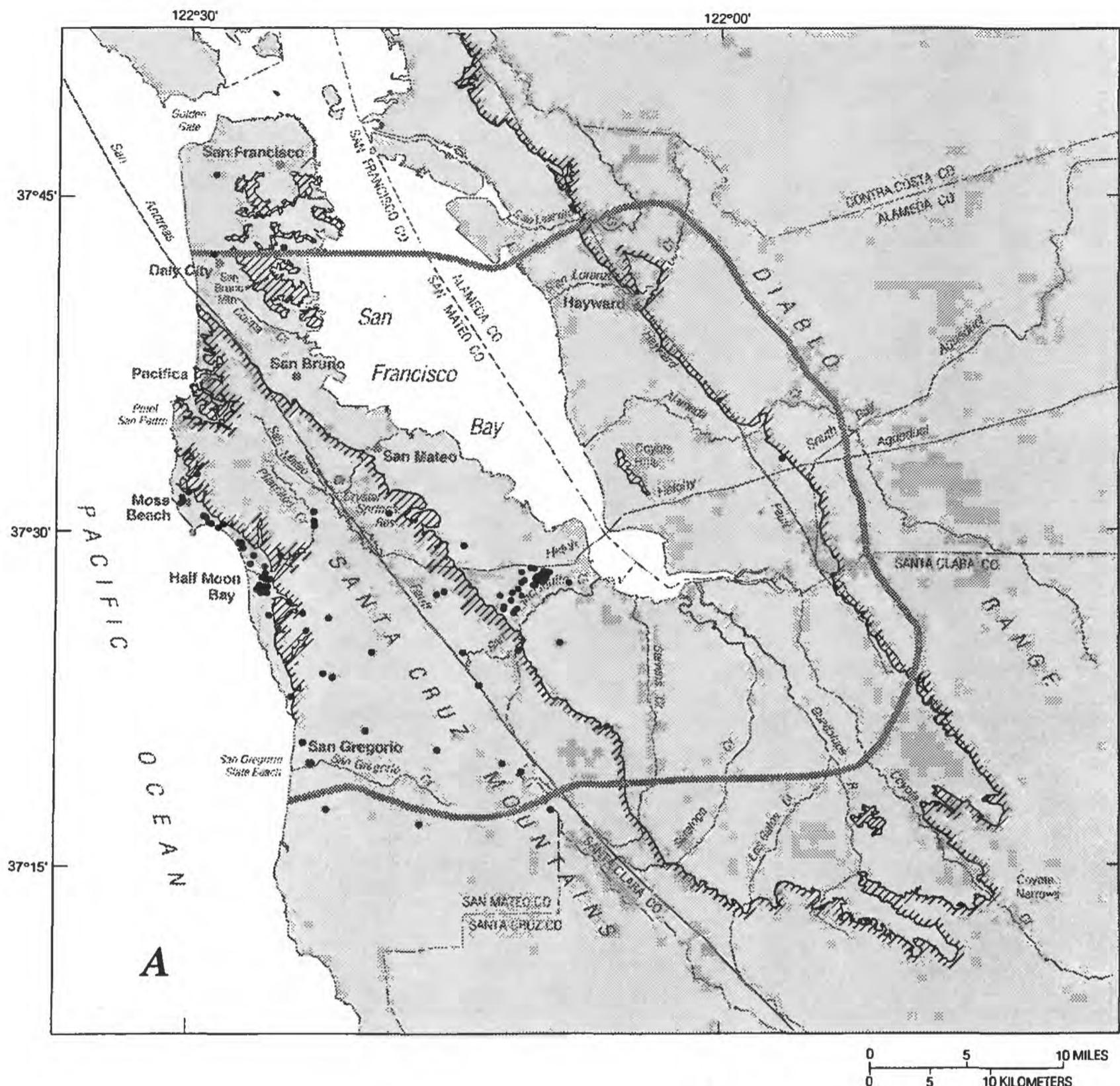
**Figure 9.** Location of wells with pumping-test data that are drilled into **(A)** the shallow zone and **(B)** the deep zone, south San Francisco Bay and Peninsula area, California.

that reported all major ions. Major ions needed for these computations are bicarbonate alkalinity, calcium, chloride, magnesium, nitrate, potassium, sodium, and sulfate. To perform this analysis, the concentration of each constituent is converted from milligrams per liter to milliequivalents per liter, a unit of measurement that describes the chemical equivalence of the ions. For an analysis expressed in milliequivalents per liter, the total milliequivalents per liter of cations theoretically should be
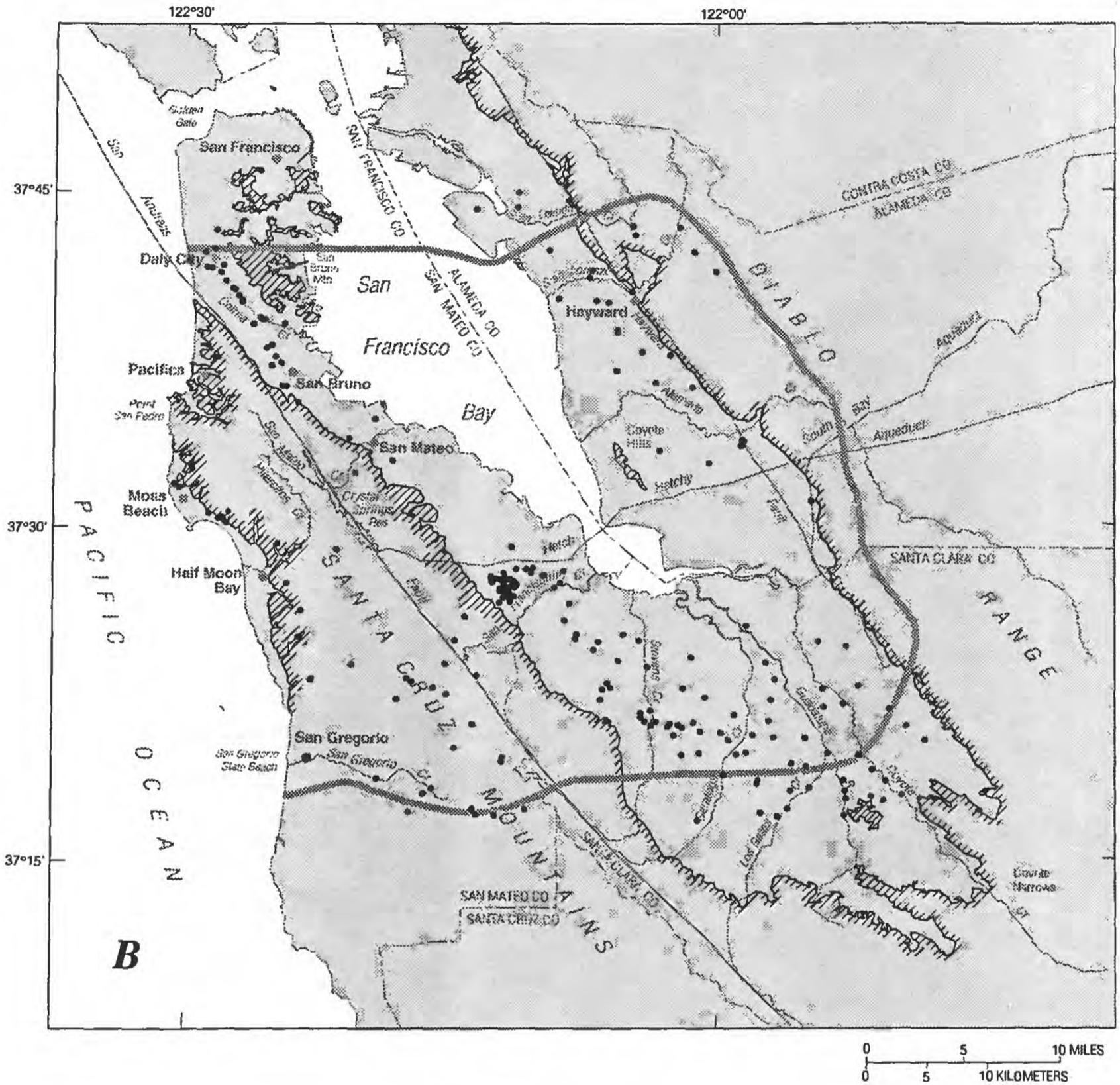
EXPLANATION
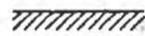
///////////   Contact between semiconsolidated and consolidated bedrock
             assemblages (hachured side) and unconsolidated alluvium

━━━━━━━   Study-area boundary

•   Well location

Figure 9. *Continued.*

equal to the total milliequivalents per liter of anions (Hem, 1985). The accuracy of an analysis can be determined by the evaluation of the percent difference between the sum of the cations and the anions.

Of the 396 samples that were analyzed for the required constituents, 51 samples (13 percent) had a a cation and anion percent difference greater than 5.5. An analysis with a cation and anion difference

Table 17. Statistics for selected fields in the construction table (table 4) for wells with pumping-test data in the database

[Values in meters below land surface, except DIAM, in centimeters]

Description	Field	Number of wells	Minimum	Maximum	Mean	Median
Depth to which the well was drilled	HOLE-DEPTH	350	6	373	111	74
Depth of the completed well	COMP-DEPTH	345	5	325	105	73
Depth to the top of the perforations	TPERFS	344	2	171	41	25
Depth to the bottom of the perforations .	BPERFS	344	5	325	100	72
Diameter of the casing	DIAM	340	5	48	24	15

greater than 5.5 percent can be considered suspect whenever the sum of the cations and anions is greater than 1.71 meq/L (Edward Pustay, U.S. Geological Survey, written commun., 1993). Because all 396 samples analyzed had cation and anion sums greater than 1.71 meq/L, the use of samples with a cation and anion percent difference greater than 5.5 for any analysis of water-quality conditions should be considered carefully. The water-quality data for the remainder of the samples without the constituents needed for the cation and anion computations were not checked for analytical accuracy.

The time of determination of alkalinity can have an effect on the cation and anion balance of an analysis. Alkalinity determined in the laboratory rather than in the field can underreport the true alkalinity value due to the precipitation of carbonate minerals in the sample bottle. This precipitation results in a positive error in the cation and anion balance. In other words, the sum of the cations is greater than the sum of the anions. Of the 51 samples with unacceptably high cation and anion balance results, 43 (84 percent) had a positive error. Only the alkalinity values reported by the USGS are known to have been determined in the field. The time of determination of the alkalinity for the rest of the samples was not identified.

The methods of measuring the pumping-test data were not identified. The accuracy of the drawdown measurement is dependent on the method used to collect the data, as described above in the discussion on the water-level data. There are several methods of measuring the discharge of a well, and each method has an associated level of accuracy. Errors in the measurement of the pumping-test data have a greater effect on results of those tests with a small drawdown or discharge rate. The results from pumping tests can be applied to

Table 18. Lithologic descriptions and quality ratings of drillers' reports in the database

Description	Rating	Code
Clay	Below average	1
Yellow clay with sand streaks	Average	2
Clay, brown to light brown, soft, cohesive, hydrated with minor fine to medium sand . .	Above average	3

estimate the transmissivity of aquifers (Lohman, 1972). Specific capacity is nearly proportional to transmissivity for a given pumping-test duration and gradually decreases as the duration of the pumping test increases. The relation between specific capacity and transmissivity also is affected by the efficiency of the well, which is dependent on such factors as age and construction.

Because of the subjective nature of describing the lithology of the formation during drilling, the detail of the written descriptions on the drillers' reports was variable. To obtain an estimate of the variability in the descriptions, each report was assigned a rating to describe the relative quality with respect to detail and thoroughness. Examples of lithologic descriptions and their associated quality ratings are presented in table 18. Most of the reports were considered to be of average quality, and only those that varied significantly from the typical description were rated above or below average.

Of the 762 wells contained in the lithologic files, 621 (82 percent) were rated as average, whereas 79 (10 percent) were rated below average and 62 (8 percent) were rated above average. The ratings apply only to the detail of the descriptions and do not reflect the variability in the geologic knowledge, experience, and vocabulary of the individuals who completed the reports.
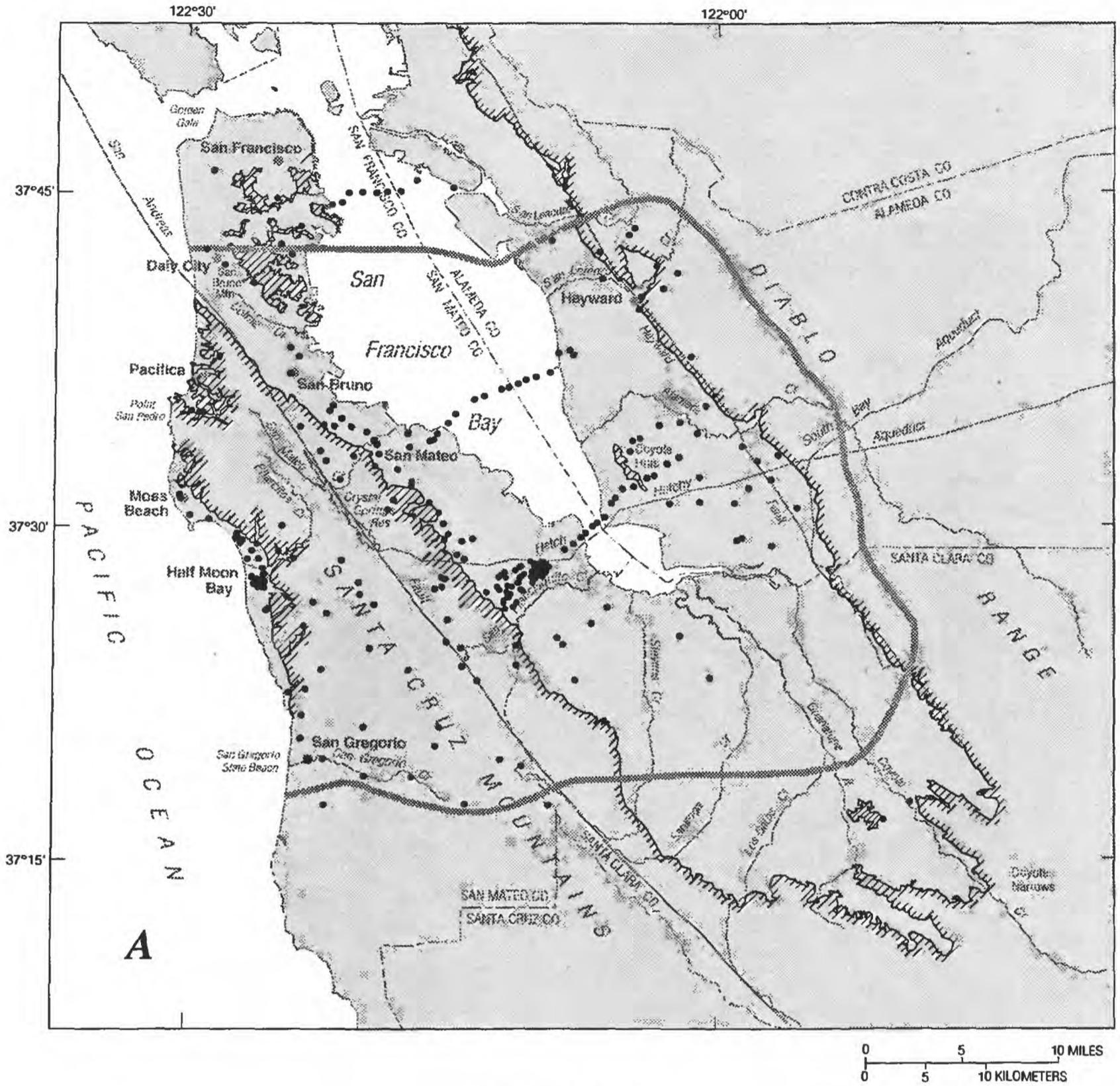
EXPLANATION
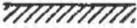
⁄⁄⁄⁄⁄⁄⁄⁄⁄⁄   Contact between semiconsolidated and consolidated bedrock
             assemblages (hachured side) and unconsolidated alluvium

▓▓▓▓▓▓▓▓   Study-area boundary

•   Well or borehole location

Figure 10. Location of wells and boreholes with lithologic data that are drilled into *(A)* the shallow zone and *(B)* the deep zone, south San Francisco Bay and Peninsula area, California.
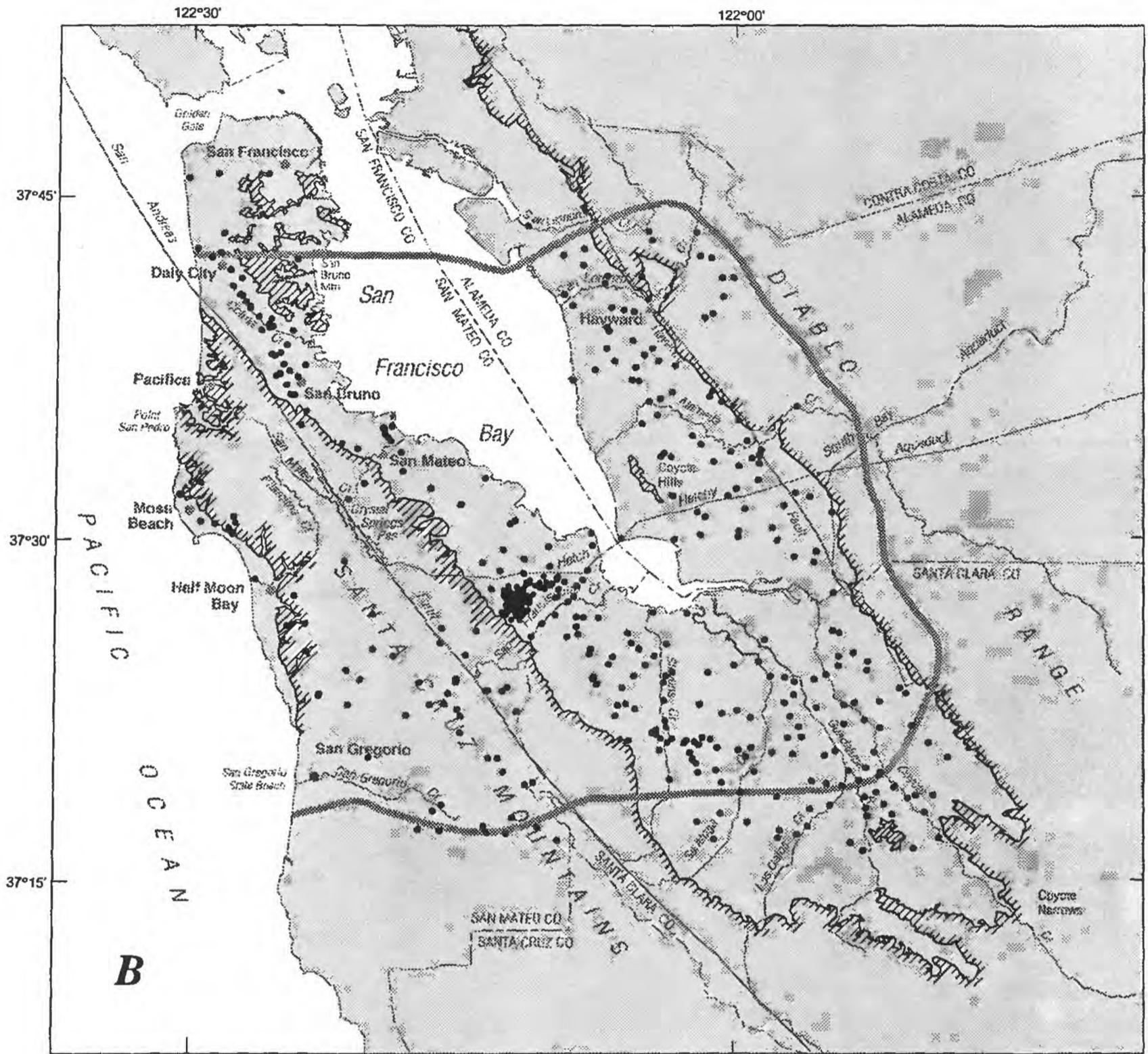
EXPLANATION

//////// Contact between semiconsolidated and consolidated bedrock
assemblages (hachured side) and unconsolidated alluvium

━━━━ Study-area boundary

• Well or borehole location

Figure 10. *Continued.*

AREAL-DATA LAYERS

The areal-data layers contain features that extend over much of the study area and are represented by lines and areas rather than discrete points. The layers contain political, geohydrologic, and other geographic features used in the assessment of aquifer conditions in the study area. These layers include thickness of alluvium, surficial geology, physiographic subareas, watershed boundaries, land use, water-supply districts, wastewater-treatment districts, and recharge basins. The layers were created by digitizing paper maps, acquiring data already in digital format, or using the GIS software to create new layers from available layers. The following sections describe the sources, scale, and contents of the areal layers.

Thickness of Alluvium

The ability of the aquifer system to store and yield ground water is affected by the areal extent and thickness of the saturated alluvium. A layer that contains contours that represent the thickness of the alluvium (fig. 11) was created with data available on the altitude of land surface and the altitude of the bedrock surface.

Continuous surfaces, such as altitude of land surface and bedrock, are represented by the GIS with triangulated irregular networks (TIN) and lattices. A continuous surface that is defined by irregularly spaced data points can be represented with a TIN. Each data point represents an x and y coordinate value and a z value that describes the altitude of the surface at that point. A series of connected triangles is created with the triangle vertices represented by the irregularly spaced data points. The x, y, and z values of the three vertices of each triangle are used to determine the slope and aspect of the triangle. The triangles act as a unit and can accurately describe the represented surface (Environmental Systems Research Institute, Inc., 1991b). For algebraic manipulations on a continuous surface, the surface is converted into a lattice with the GIS software. A lattice represents a continuous surface with values at regularly spaced points.

Several sources of data were required to develop complete coverage of the land-surface altitude of the study area. Digital-elevation models (DEM) for areas that correspond to 26 U.S. Geological Survey 7.5-minute 1:24,000-scale

topographic quadrangles were obtained. These models provided altitude data for most of the terrain above sea level. Digital elevation models are records of terrain altitudes at regularly spaced horizontal intervals. The locations represented by the 7.5-minute DEM data are spaced 30 m apart in both the x and y directions. The altitude data for DEMs can be in units of feet or meters, and the coordinates can be based on either the North American Datum of 1927 or the North American Datum of 1983.

The altitude data contained in the DEM are acquired from map contour overlays or from automated or manual scanning of high-altitude aerial photography. The accuracy of a DEM is determined by computing the root-mean-square error (RMSE) for altitudes in the DEM when compared with corresponding altitudes from published maps. The desired RMSE for 7.5-minute DEM data is 7 m, and 15 m is the maximum permitted (U.S. Geological Survey, 1990).

To use the data in the GIS, each DEM file was converted into a lattice. During the conversion process, DEMs that contained altitude data, in feet, were converted to meters, and coordinates based on the North American Datum of 1983 were converted to the North American Datum of 1927. The transformation of coordinates between datums was accomplished with NADCON, a procedure incorporated into the GIS that was developed by the National Oceanic and Atmospheric Administration (Environmental Systems Research Institute, Inc., 1991c).

The lattices were merged to form one lattice that represents the land-surface altitude for the study area. To decrease storage requirements and improve processing speed, a filtering procedure was performed to select the most important points to represent land-surface altitude. The procedure evaluates each point in the lattice to determine its significance in representing the land surface based on the rate of change in the slope in the immediate vicinity of the point. All lattice points are then ranked based on their respective significance value, and a percentage of the points with the highest ranking are chosen for the resulting output data layer. The resulting layer contains irregularly spaced points that represent topographic highs and lows as well as points of significant change in slope. A selection criteria of 5 percent was determined to provide enough points to represent land surface without resulting in an excessively

Figure 11. Thickness of alluvium in the south San Francisco Bay and Peninsula area, California.
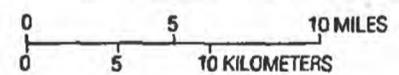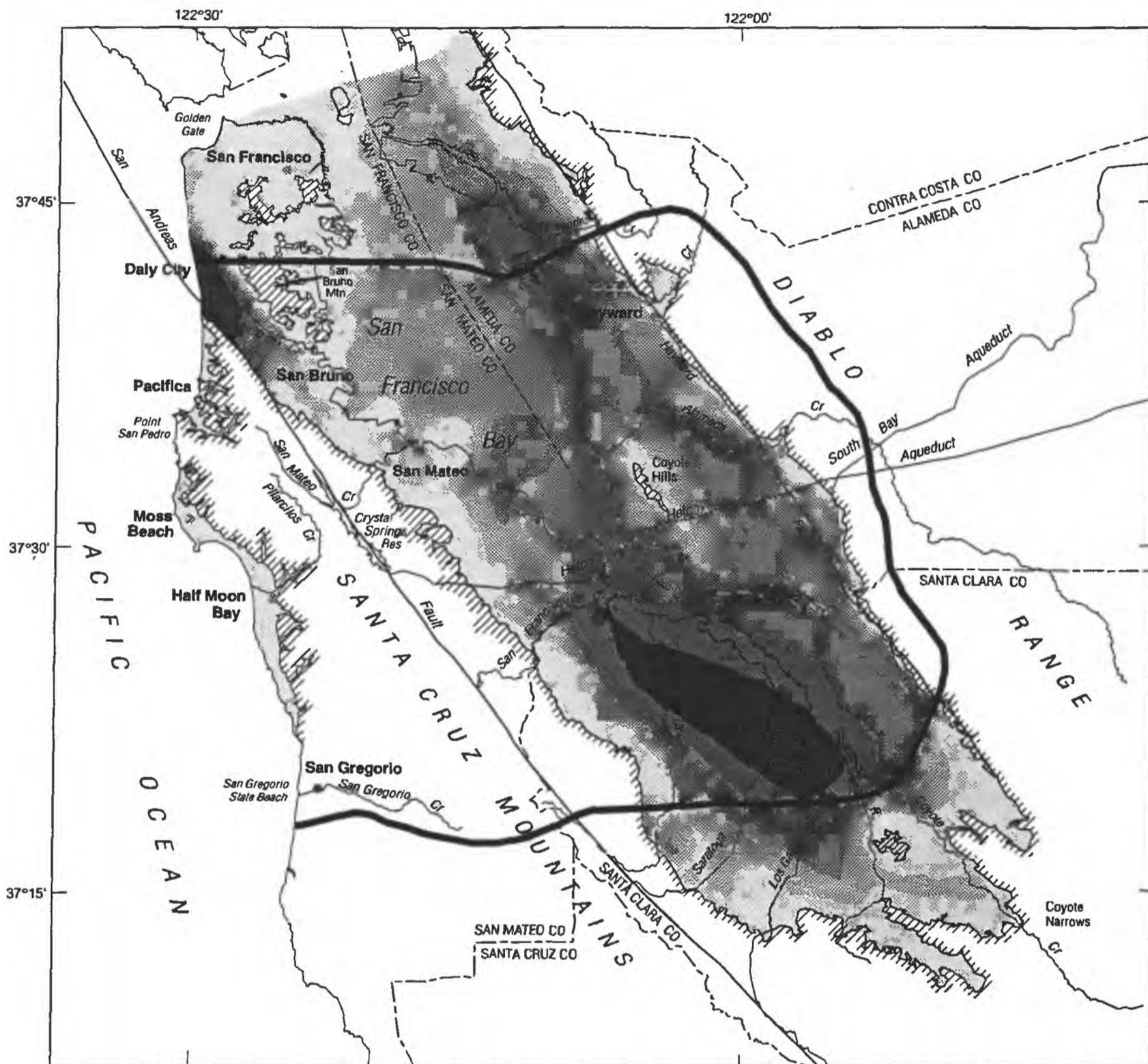
EXPLANATION

Thickness of alluvium, in meters
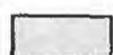
0 to 100

101 to 200

201 to 300

301 to 400

greater than 400

Contact between semiconsolidated and consolidated bedrock assemblages (hachured side) and unconsolidated alluvium

Study-area boundary

large data layer. The 5-percent criteria resulted in an output layer with 247,955 points.

Altitude data for small areas not covered by the DEMs, generally along the coast north of Half Moon Bay, were digitized manually from the U.S. Geological Survey 7.5-minute 1:24,000-scale topographic quadrangles. Altitude data for terrain beneath San Francisco Bay were obtained from a database of points that represent the depth of the bay (Burau and Cheng, 1989). The points in this database were digitized from National Oceanic and Atmospheric Administration-National Ocean Service 1:40,000-scale charts. Points from these two sources were used, along with the points output from processing the DEM data, to create a TIN that represents the land-surface altitude of the entire study area. A lattice to calculate the thickness of the alluvium was created from the TIN.

The bedrock altitude surface was developed with data obtained from a 1:62,500-scale map of the altitude of the bedrock in San Mateo County (Hensolt and Brabb, 1990) and from preliminary bedrock altitude data at a scale of 1:62,500 for the south San Francisco Bay region (Earl E. Brabb, U.S. Geological Survey, written commun., 1992). Contours of the altitude of bedrock were manually digitized and converted to a TIN. The TIN was converted to a lattice, and the lattice was subtracted from the land-surface altitude lattice to create a lattice of the thickness of alluvium. Contours of the thickness of alluvium were then generated to create the thickness-of-alluvium data layer (fig. 11). The thickness of alluvium, in meters, for each line feature is contained in the field THICKNESS in the feature-attribute table.

Surficial Geology

Rocks and sediments in the study area can be grouped by age to assess their relative significance to ground-water resources (California Department of Water Resources, 1967). A data layer delineating four age-related rock and sediment types (fig. 12) was created from the surficial geology digital map database by Wentworth (1993). A discussion of the relation between rock and sediment age and ground-water resources in the study area is presented by Fio and Leighton (1994).

The rocks and sediments in the study area were grouped into the following four categories: (1) marine deposits of Holocene age (MR); (2) alluvium of Pleistocene and Holocene ages (PR); (3) alluvium of Pliocene and Pleistocene age (PP); and (4) the Franciscan Assemblage of Jurassic and Cretaceous age and rocks of Tertiary age (FT). A fifth category identifying areas of open water (OW) also is contained in the data layer. The two-letter codes listed above are contained in the field AGE in the feature-attribute table. Data on the geologic unit, general lithology, and age contained in the source database (Wentworth, 1993) were used to delineate these categories.

Physiographic Subareas

To facilitate the evaluation of the geohydrology of the region, physiographic subareas with similar hydrologic and geologic properties were identified. The boundaries of the subareas are contained in the regional physiographic-subarea data layer (fig. 13) and the feature-attribute table contains a 30-byte character field with the name of each subarea.

The northern third of the data layer, and the part along the Pacific Coast, was delineated based on the 1:125,000-scale surficial geology digital map database by Wentworth (1993). The southern two thirds of the layer were digitized from 1:145,000-scale maps showing physiographic regions (California Department of Water Resources, 1967) and were further modified based on the surficial geology (Wentworth, 1993). The boundaries of the subareas were determined as follows: (a) the Exposed Bedrock subarea is the surficial exposures of the oldest, compressed-rock formations; (b) the Uplands are the surficial exposures of continental sediments deposited during late Pliocene and early Pleistocene age; (c) the East and West Side Alluvial Aprons are the surficial exposures of continental sediment deposited during the Pleistocene and Holocene ages; (d) the Niles subarea, a part of the East Side Alluvial Apron subarea, is delineated by the boundaries of the Niles Cone; (e) the Coastal subarea is the semiconsolidated and unconsolidated sediment in exterior areas adjacent to the Pacific Ocean; (f) the Merced subarea is a 3- to 5-km-wide valley between the Santa Cruz and San Bruno Mountains, underlain by predominantly marine sediments deposited during Pliocene and Pleistocene age; (g) the San Jose Plain subarea is the distal portions of alluvial fans deposited by Coyote Creek, Guadalupe River, and Los Gatos Creek; and (h) the Bay Plain subarea is delineated by the tidal flats, marshlands, and bay-fill areas adjacent to San Francisco Bay. A more detailed discussion of each subarea is presented by Fio and Leighton (1994).
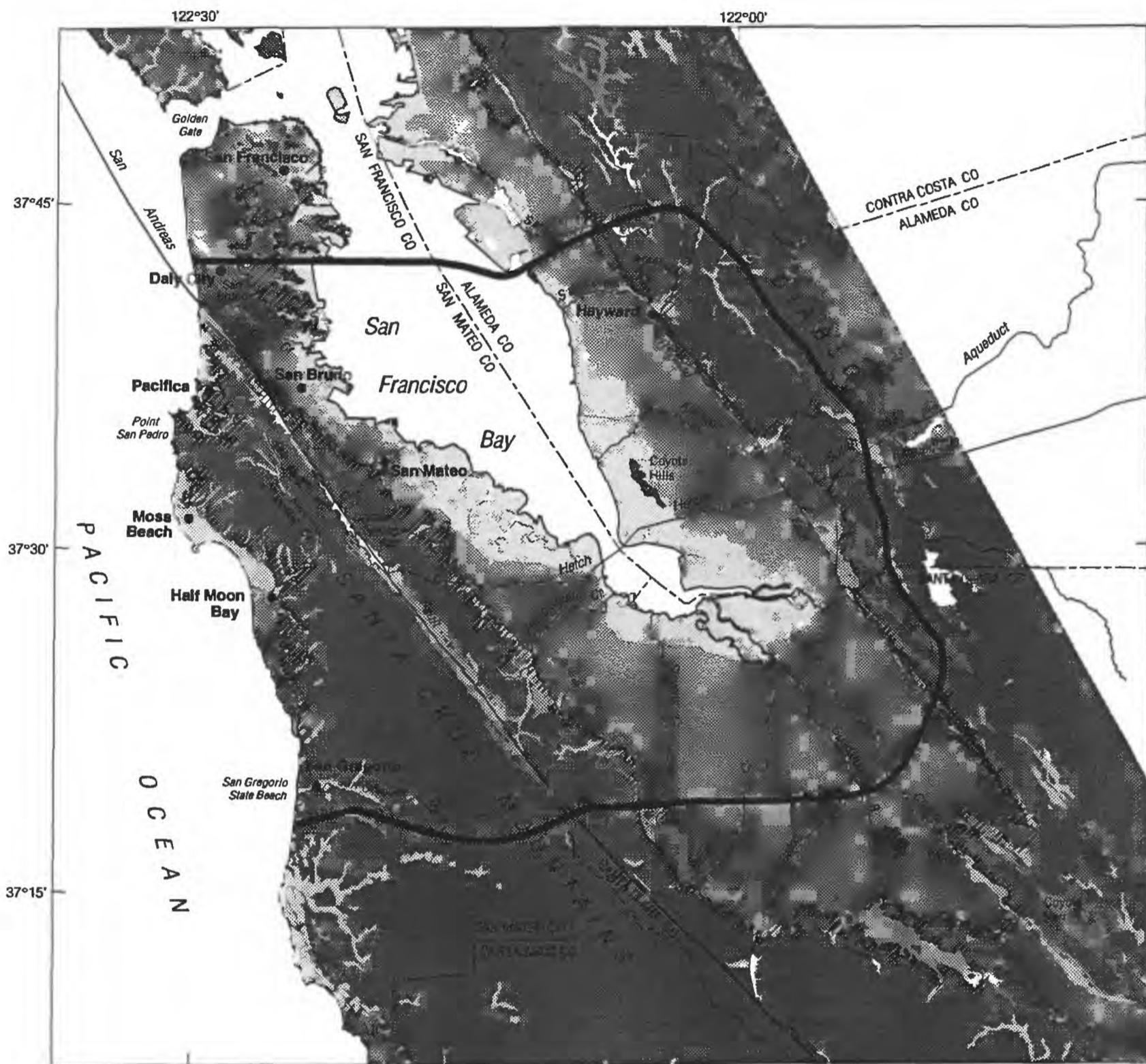
EXPLANATION
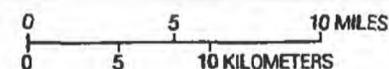
Marine deposits of Holocene Age--Modified from Wentworth (1993)

Alluvium of Pleistocene and Holocene Age--Modified from Wentworth (1993)

Alluvium of Pliocene and Pleistocene Age--Modified from Wentworth (1993)

Franciscan Complex of Jurassic and Cretaceous Age and rocks of Tertiary Age -- Modified from Wentworth (1993)

Contact between semiconsolidated and consolidated bedrock assemblages (hachured side) and unconsolidated alluvium

Study-area boundary

Figure 12. Surficial geology in the south San Francisco Bay and Peninsula area, California.
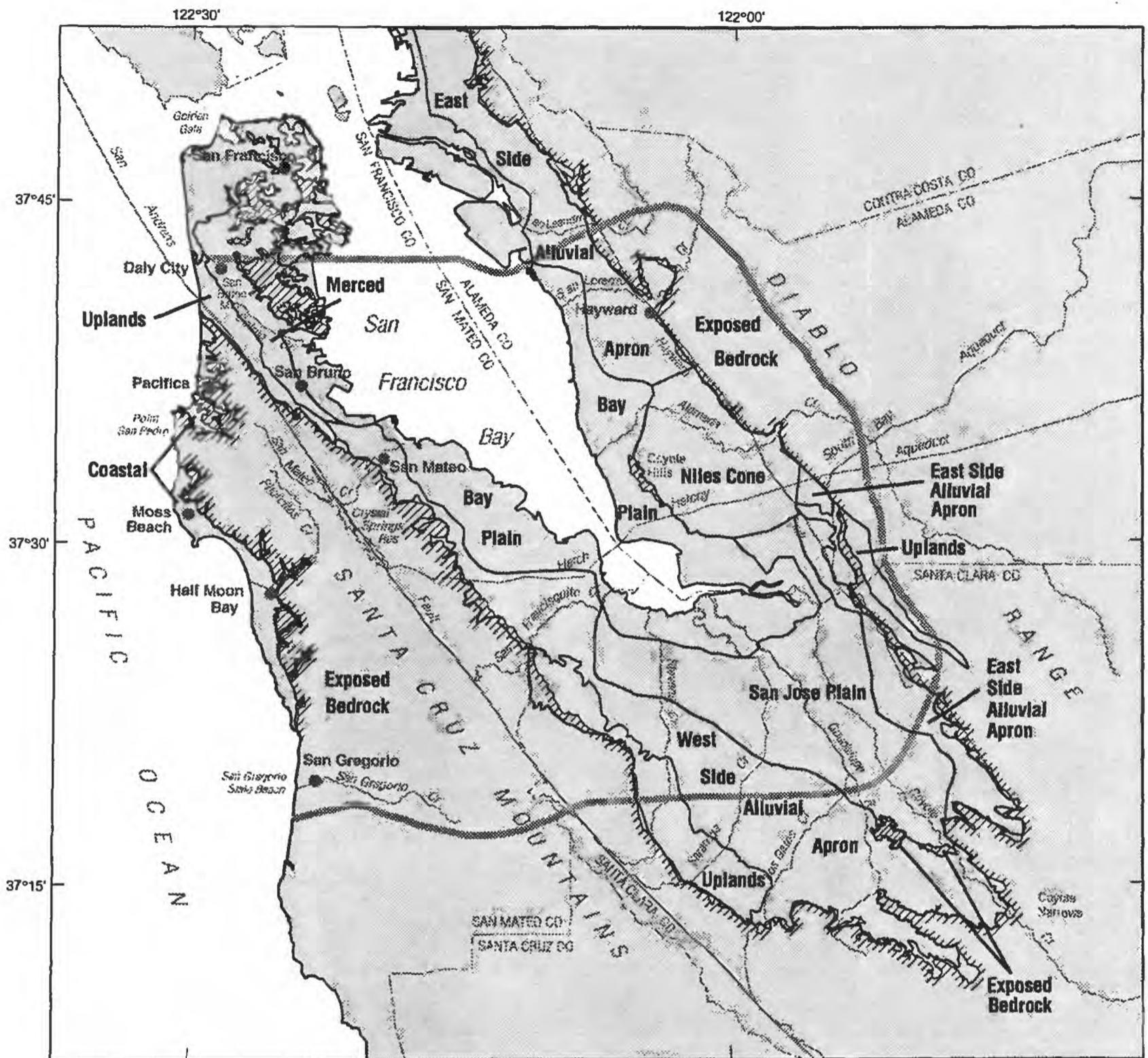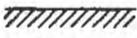
Figure 13. Boundaries of regional physiographic subareas in the south San Francisco Bay and Peninsula area, California.

EXPLANATION

/////////	Contact between semiconsolidated and consolidated bedrock assemblages (hachured side) and unconsolidated alluvium
▬▬▬▬▬	Study-area boundary
———	Physiographic subarea boundary

Watershed Boundaries

Estimation of ground-water recharge requires the analysis of surface-water runoff in relation to precipitation. Surface-water runoff data were collected in 1990 from 19 streamflow-gaging stations located throughout the study area. To evaluate the distribution of surface-water runoff and

precipitation, a data layer of watershed boundaries was created for both gaged and ungaged streams in the study area (fig. 14). The watershed boundaries were delineated on U.S. Geological Survey 7.5-minute 1:24,000-scale topographic maps and then manually digitized. The digital-map sheets were appended to form one data layer for the study area.

Three fields in the feature-attribute table contain information about each watershed. The field NAME contains the name of the stream draining the watershed, the field DSWS contains the name of the stream draining the watershed immediately downstream, and the field DEST identifies the final destination of water as the San Francisco Bay or the Pacific Ocean. Relatively flat areas along the bay and ocean that do not drain into a specific stream were given a watershed name of either San Francisco Bay or Pacific Ocean.

An additional layer was created to contain the points where the stream that drains a watershed discharges into the watershed downslope. These points were digitized manually for the outlet of each watershed. Three fields in the feature-attribute table contain data for watershed outlets that have a streamflow-gaging station and streamflow data for 1990. The field NUMBER contains the gaging station identification number, the field AGENCY identifies the agency that maintains the gaging station, and the field NAME contains the name of the gaging station. All gaging stations identified in this layer are maintained by either the USGS or the Santa Clara Valley Water District (SCVWD). The locations of the streamflow-gaging stations are shown in figure 14.

Land Use

Land-use data were compiled for the study area to aid in the estimation of consumptive water use based on land-use categories. To evaluate the temporal land-use changes, land-use data were obtained for two periods, 1974-76 and 1986-90. The two periods selected were determined by the availability of land-use data. The data for the two periods are based on different classification systems. The data from 1974-76 are based on the classification system described by Anderson and others (1976). The

system consists of nine general land-use categories (Level I) subdivided into 37 more specific categories (Level II) (table 19). The classification system used for the 1986-90 data is based on four general categories (urban, agriculture, native vegetation, and recreation) that are further divided into detailed land-use types (Dean Reynolds, California Department of Water Resources, written commun., 1992). For the 1986-90 data, the detailed land-use categories were combined into seven categories to simplify the digitizing process while providing the land-use data necessary for analysis of current land- and water-use conditions (table 20).

1974-76

Digital land-use data are available from the U.S. Geological Survey for the continental United States and Hawaii at scales of 1:250,000 and 1:100,000 (U.S. Geological Survey, 1986). The primary sources used in compiling the land-use data were aerial photographs and other remotely sensed data. Previously compiled land-use maps and limited field verification also were used. The smallest area mapped for all urban areas, bodies of water, surface mines, quarries, gravel pits, and certain agricultural areas is 4 ha. A minimum size of 16 ha is used for all other categories.

The 1974-76 land-use layer for the study area is a composite of four digital land-use maps (fig. 15). Due to the complexity of this data layer, only Level I land-use types are shown, but Level II land-use types are maintained in the data layer. Level I land-use types tundra and perennial snow and ice were not found in the study area and were not used. The four maps used are San Francisco (scale 1:250,000, 1975), San Francisco (scale 1:100,000, 1976), San Jose (scale 1:100,000, 1974), and Stockton (scale 1:100,000, 1976). The maps are provided in the form of digital data contained in ASCII files. The GIS software contains a procedure for converting the ASCII files into the data layers used by the GIS for analysis and display. The four maps were appended to form the 1974-76 land-use layer. The LUCODE field is a 2-byte integer field in the feature-attribute table, which contains the code that describes the land-use type (table 19).
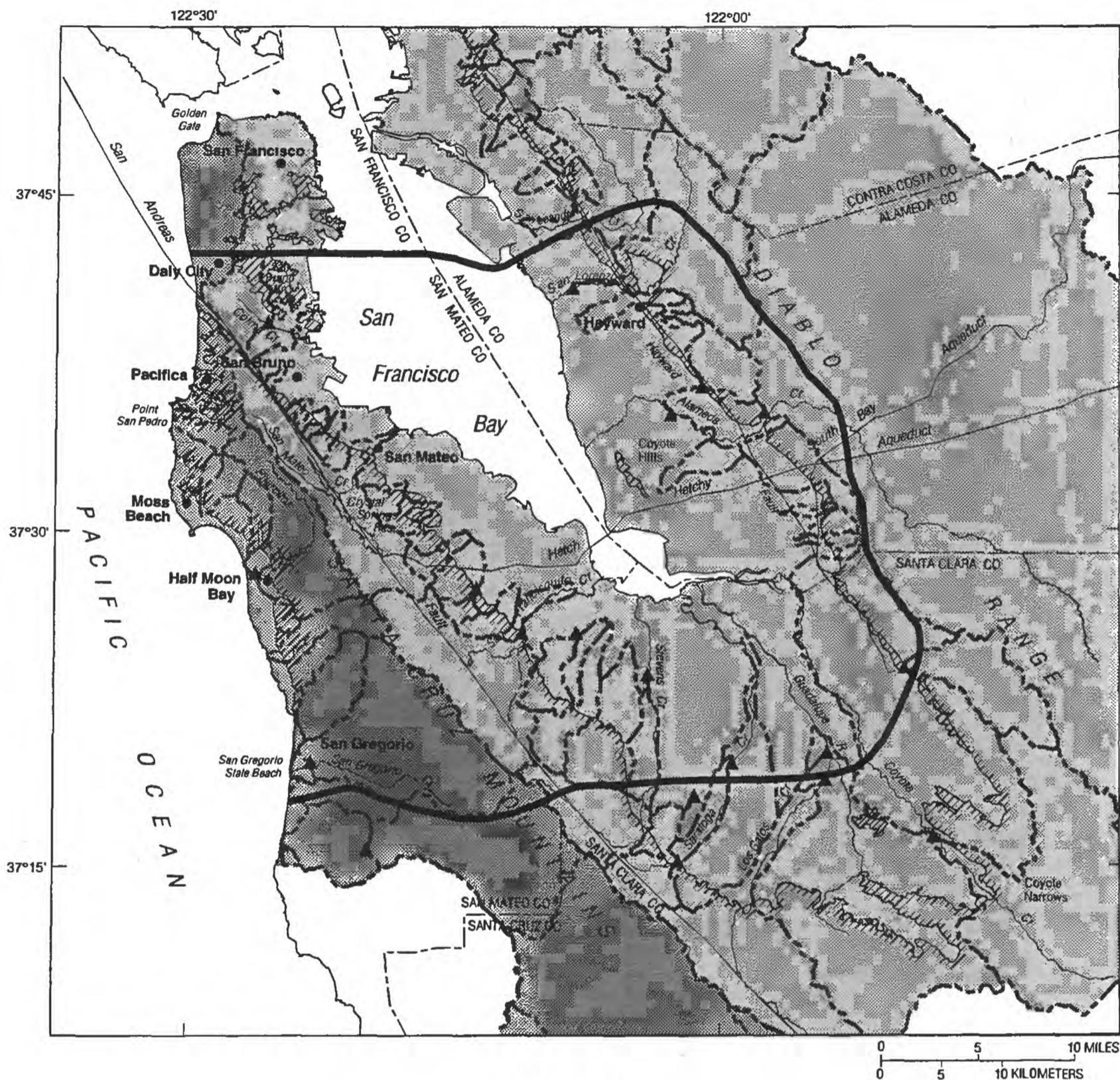
Figure 14. Watershed boundaries and location of streamflow-gaging stations, south San Francisco Bay and Peninsula area, California.
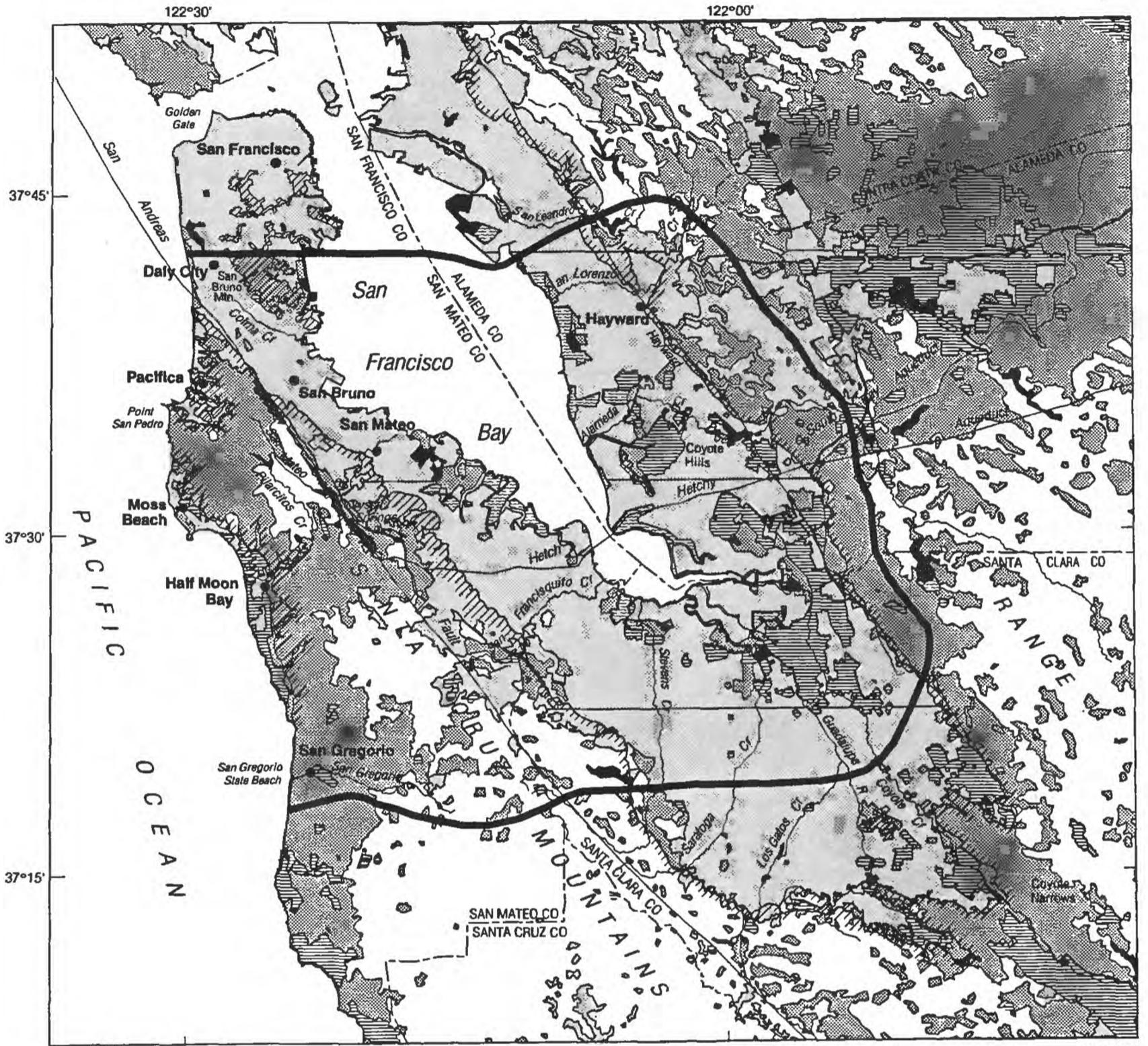
Figure 15. Land use, south San Francisco Bay and Peninsula area, California, 1974-76.

Table 19. Land-use categories for 1974-76 land-use data in the database

[Land-use codes 1-9 are general Level I categories and land-use codes 11-92 are more specific Level II categories (Anderson and others, 1976)]

1 Urban or Built-up Land
 11 Residential
 12 Commercial
 13 Industrial
 14 Transportation, commercial, and utilities
 15 Industrial and commercial complexes
 16 Mixed urban or built-up land
 17 Other urban or built-up land
2 Agricultural Land
 21 Cropland and pasture
 22 Orchards, groves, vineyards, and nurseries
 23 Confined feeding operations
 24 Other agricultural land
3 Rangeland
 31 Herbaceous rangeland
 32 Shrub and brush
 33 Mixed rangeland
4 Forest Land
 41 Deciduous forest land
 42 Evergreen forest land
 43 Mixed forest land
5 Water
 51 Streams and canals
 52 Lakes
 53 Reservoirs
 54 Bays and estuaries
6 Wetland
 61 Forested wetland
 62 Nonforest wetland
7 Barren Land
 71 Dry salt flats
 72 Beaches
 73 Sandy areas other than beaches
 74 Bare exposed rock
 75 Strip mines, quarries, and gravel pits
 76 Transitional areas
 77 Mixed barren land
8 Tundra
 81 Shrub and brush tundra
 82 Herbaceous tundra
 83 Bare ground tundra
 84 Wet tundra
 85 Mixed tundra
9 Perennial snow and ice
 91 Perennial snowfields
 92 Glaciers

Table 20. Land-use categories for 1986-90 land-use data in the database

[LUCODE, a 2-byte character field that contains the code that describes land use]

Description	LUCODE
Urban	UN
Irrigated lawn	LI
Non-irrigated lawn	LN
Irrigated agriculture	AI
Non-irrigated agriculture	AN
Native vegetation	NV
Water	WR

the estimation of agricultural land-use practices. The maps were created using a combination of aerial photography and field checking. Maps covering Alameda County were last updated in 1986, San Mateo County in 1987, and Santa Clara County in 1990. The CADWR land-use data layer (fig. 16) was digitized directly from 21 maps obtained from the CADWR and appended to form one data layer for the study area. The LUCODE field is a 2-byte character field in the feature-attribute table, which contains the code that describes the land-use type (table 20).

A second data layer was created by subdividing the native vegetation land-use categories. This was accomplished by overlaying the 1974-76 data on the 1986-90 data. Areas identified as native vegetation in the 1986-90 data were changed to the land-use category identified by the 1974-76 data. Land-use categories identified by the 1986-90 data as other than native vegetation were left unchanged. Because of differences in the scale and source of the two layers and changes in land use over time, boundaries between major land-use categories did not always align. Native-vegetation polygons for the 1986-90 data that lacked a corresponding vegetation type for the 1974-76 data were assigned a specific vegetation type based on the predominant vegetation type in the immediate vicinity. The modified land-use layer is shown in figure 17. The land-use categories (table 21) are contained in a 2-byte character field called LUCODE in the feature-attribute table.

1986-90

The California Department of Water Resources (CADWR) maintains paper copies of land-use maps based on the U.S. Geological Survey 7.5-minute 1:24,000-scale topographic quadrangles for use in

Water-Supply Districts

Water-supply districts provide water for much of the study area, including all heavily populated areas. The locations of the water-supply districts are essential for determining potential sources of
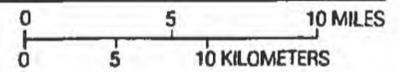
Figure 16. Land use, south San Francisco Bay and Peninsula area, California, 1986-90.
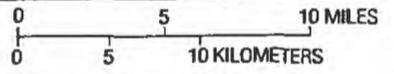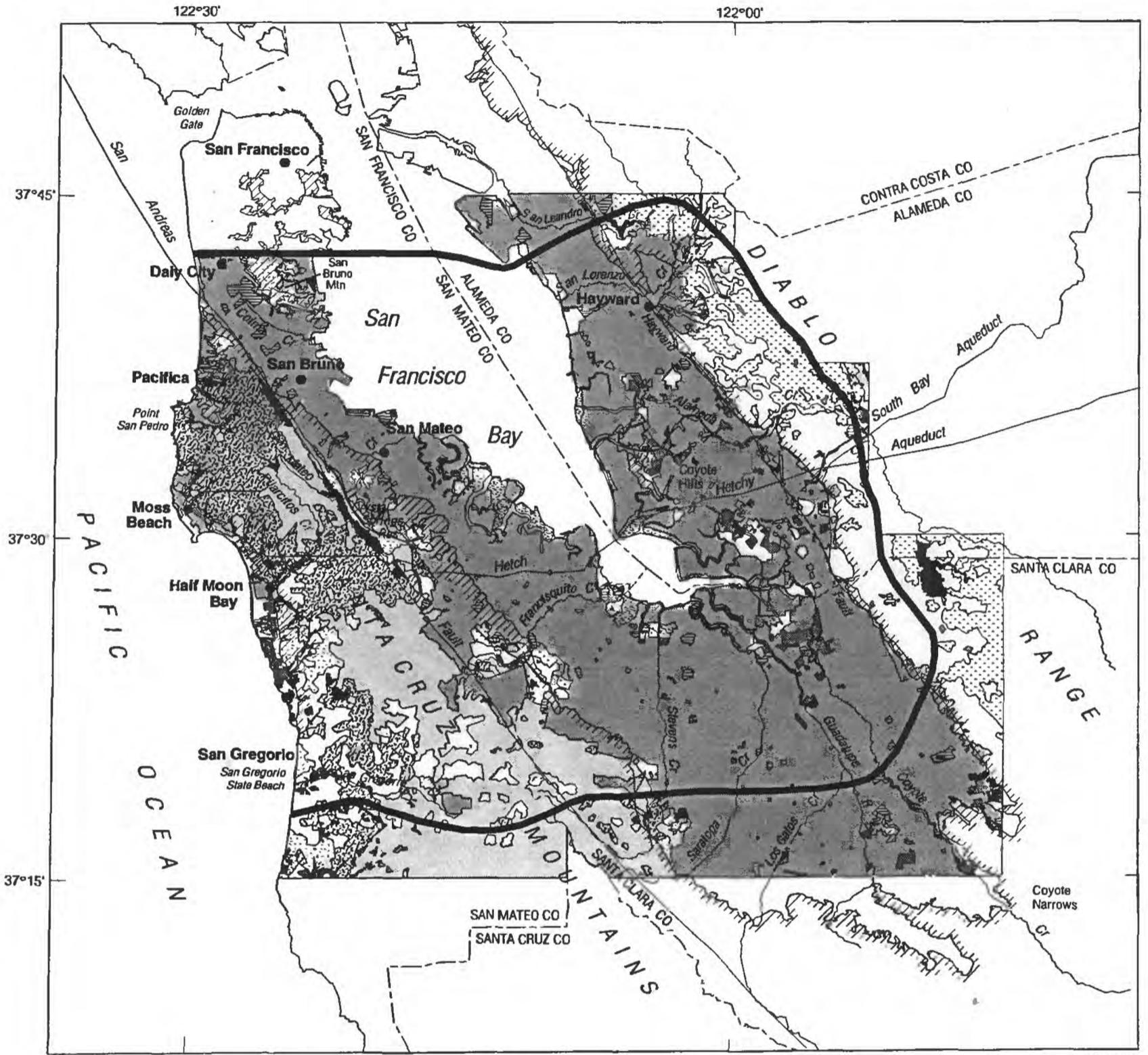
EXPLANATION

LAND-USE CATEGORIES

Urban/developed land		Scrub/chaparral	
Lawn		Evergreen forest	
Irrigated agriculture		Mixed forest	
Non-irrigated agriculture		Marsh/wetlands	
Herbaceous rangeland		Freshwater	
Mixed rangeland			

////////// Contact between semiconsolidated and consolidated bedrock assemblages (hachured side) and unconsolidated alluvium

▬▬▬ Study-area boundary

Figure 17. Modified land use, south San Francisco Bay and Peninsula area, California, 1986-90.

Table 21. Land-use categories for modified 1986-90 land-use data in the database

[LUCODE, a 2-byte character field that contains the code that describes land use]

Description	LUCODE
Urban/developed land	UN
Lawn (parks, cemeteries)	LN
Irrigated agriculture	AI
Non-irrigated agriculture	AN
Herbaceous rangeland (grasslands)	HR
Mixed rangeland (grass and scrub)	MR
Scrub/chaparral (brushland)	SB
Evergreen forest (conifers)	EF
Mixed forest (woodland-conifer)	MF
Marsh/wetlands	MW
Freshwater (lakes, reservoirs)	FW

Table 22. Water-supply districts larger than a square kilometer in the database

Code	Water-supply district
1	San Francisco Water Department
2	Daly City Municipal Water District
3	Brisbane Water Department
4	Guadalupe Valley Water District
5	California Water Service Company - South San Francisco
6	Westborough County Water District
7	North Coast County Water District
8	San Bruno Municipal Water District
9	San Francisco International Airport
10	Millbrae Municipal Water District
11	Burlingame Municipal Water Department
12	Hillsborough Municipal Water Department
13	California Water Service Company - San Mateo
14	Citizens Utilities Company
15	Coastside County Water District
16	Estero Municipal Improvement District
17	Belmont County Water District
18	Redwood City Municipal Water Department
19	California Water Service Company - San Carlos
20	San Mateo County Water Works District No. 3
21	Menlo Park Municipal Water District
22	East Palo Alto Water Works District
23	Skyline County Water District
24	California Water Service Company - Bear Gulch
25	Stanford University
26	City of Palo Alto
27	Los Trancos County Water District
28	Purissima Hills County Water District
29	California Water Service Company - Los Altos
30	City of Moutain View
31	Moffett Field Naval Air Station
32	City of Sunnyvale
33	City of Cupertino
34	San Jose Water Company
35	Great Oaks Water District
36	City of San Jose-Evergreen
37	City of Santa Clara
38	City of San Jose-Alviso
39	City of Milpitas
40	Alameda County Water District
41	City of Hayward
42	East Bay Municipal Utility District

geohydrologic data and for evaluating the distribution of reported surface-water supply, groundwater pumpage, and consumptive use in the study area. The boundaries of these districts are contained in the water-supply-districts data layer (fig. 18). The water-supply districts shown in figure 18 are labeled with a numeric code, and the codes, along with the corresponding district names, are listed in table 22. Water-supply districts that cover areas of less than 1 km² are too small to be identified in the figure, but are included in the data layer. These small water-supply districts are listed in table 23. The name of each water-supply district is contained in a 50-byte character field in the feature-attribute table.

The boundaries were obtained from maps provided by Alameda County Water District (1991), East Bay Municipal Utility District (1986), San Mateo County Planning and Development Division (1986), and Santa Clara Valley Water District (1990) and were plotted on and manually digitized from U.S. Geological Survey 7.5-minute 1:24,000-scale topographic quadrangles. Water-supply districts too small to be represented by areas on the source maps were not included in this data layer.

Wastewater-Treatment Districts

Wastewater is collected and treated throughout much of the study area by wastewater-treatment districts. The boundaries of these districts are contained in the wastewater-treatment-districts data layer (fig. 19). The wastewater-treatment districts shown in figure 19 are labeled with a numeric code and the codes along with the district names are

listed in table 24. The name of each district is contained in a 50-byte character field in the feature-attribute table.

The boundaries were obtained from paper maps provided by the City of San Jose, Environmental Services Department (scale unknown) (Roland Sun, written commun., 1993), the East Bay Municipal Utility District (1:276,000 scale) (Bill Meckel, written commun., 1993), San Leandro Sanitary District (1:32,000 scale) (written commun., 1993),
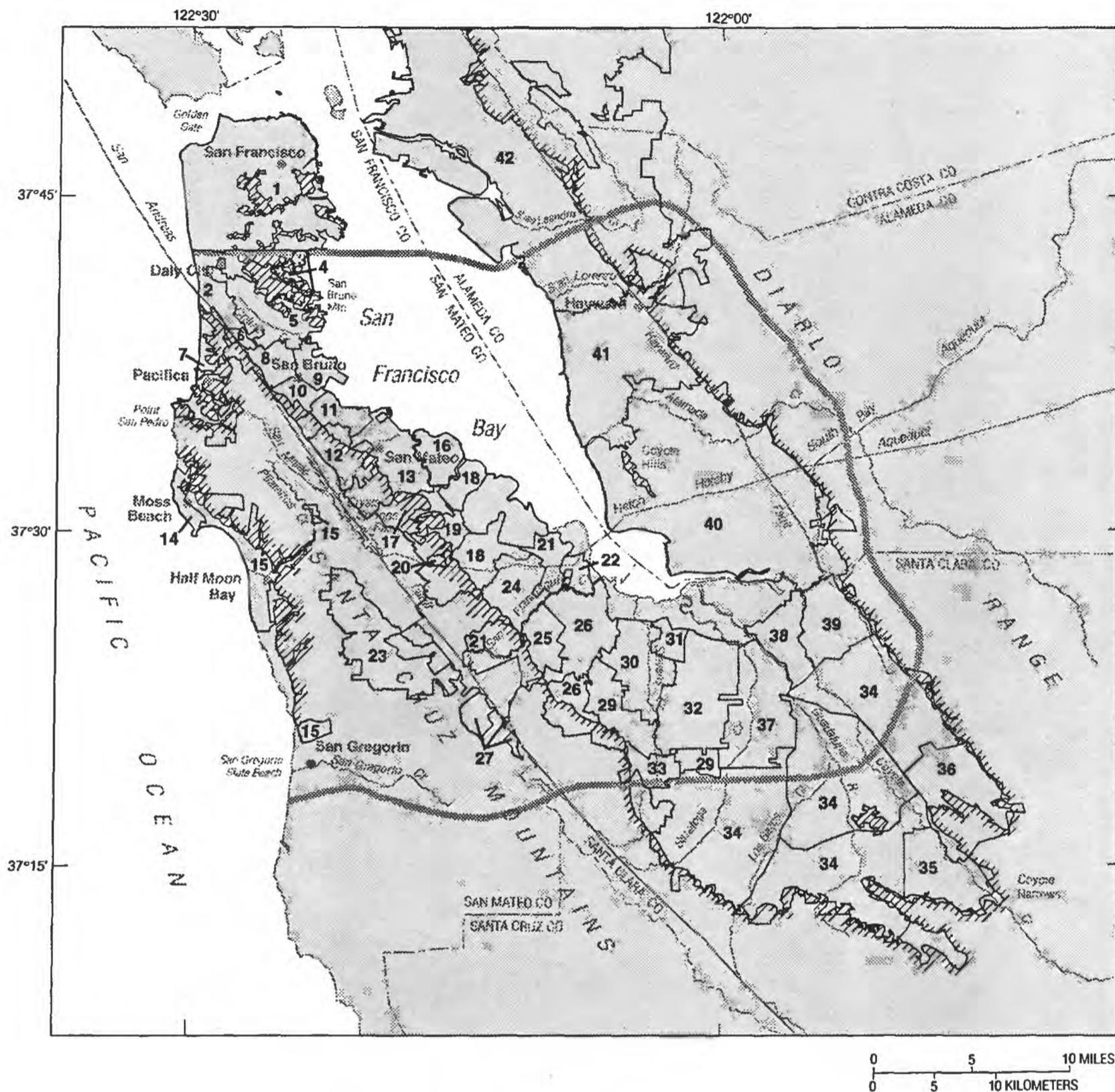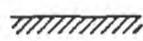
Figure 18. Boundaries of water-supply districts, south San Francisco Bay and Peninsula area, California.

EXPLANATION

//////	Contact between semiconsolidated and consolidated bedrock assemblages (hachured side) and unconsolidated alluvium
~~	Study-area boundary
——	Water-supply district boundary
27	Water-supply district (see table 22 for name of district)

the Sunnyvale Department of Public Works (1:25,000 scale) (Norman Linn, written commun., 1993), Union Sanitary District (1:24,000 scale) (Richard Davis, written commun., 1993), Castro Valley Sanitary District (1988), Oro Loma Sanitary District (1991), San Francisco Department of Public Works (1977), and San Mateo County Planning and Development Division (1991). The boundaries were manually digitized from the source maps and appended to form one data layer.

Figure 19. Boundaries of wastewater-treatment districts; south San Francisco Bay and Peninsula area, California.

EXPLANATION

//////////   Contact between semiconsolidated and consolidated bedrock assemblages (hachured side) and unconsolidated alluvium

▬▬▬▬   Study-area boundary

─────   Wastewater-treatment district boundary

13   Wastewater-treatment district (see table 24 for name of district)

Recharge Basins

Alameda County Water District and Santa Clara Valley Water District manage ground-water recharge programs to replenish water pumped from aquifers underlying their service areas. Imported and locally collected surface water is held in small recharge basins in highly permeable areas and is

EXPLANATION

/////////////// Contact between semiconsolidated and consolidated bedrock
assemblages (hachured side) and unconsolidated alluvium

▓▓▓▓▓▓▓▓▓▓ Study-area boundary

Figure 20. Location of ground-water recharge basins, south San Francisco Bay and Peninsula area, California.

allowed to percolate into the underlying aquifers. The recharge basin layer contains the boundaries of these basins (fig. 20). The layer was prepared by manually digitizing polygons that were identified as recharge basins on the U.S. Geological Survey 7.5-minute 1:24,000-scale topographic quadrangles. Information provided by Alameda County Water District (Earl Lenahan, written commun., 1991) and Santa Clara Valley Water District (1988) was used to identify the basins on the topographic maps.

Table 23. Water-supply districts smaller than a square kilometer in the database

Water-supply district
Palo Alto Park County Municipal Water Works District
O'Conner Tract Mutual Water District
Sky Londa Mutual Water Company
La Honda Vista Mutual Water Company
Dearborn Water System
Loma Mar Mutual Water Company
San Mateo County Service Area No. 7
Redwood Terrace Mutual Water System
Cuesta La Honda Guild
Portola Improvement Association Mutual Water Company

Table 24. Wastewater-treatment districts in the database

Code	Wastewater-treatment district
1	City of San Francisco - Richmond/Sunset
2	City of San Francisco - North Point
3	City of San Francisco - Southeast
4	North San Mateo County Sanitation District
5	South San Francisco/San Bruno Wastewater Treatment Plant
6	City of Pacifica
7	San Francisco International Airport
8	City of Millbrae
9	City of Burlingame
10	Sewer Authority Mid-Coast
11	City of San Mateo/Estero Municipal Improvement District
12	South Bayside System Authority
13	Palo Alto Regional Wastewater Treatment District
14	City of Sunnyvale
15	San Jose-Santa Clara Wastewater Facility
16	Union Sanitary District
17	City of Hayward
18	Oro Loma Sanitary District
19	Castro Valley Sanitary District
20	City of San Lorenzo
21	East Bay Municipal Utility District Special District No. 1

SUMMARY

A regional assessment of the ground-water resources of the south San Francisco Bay and Peninsula area was performed from 1991-94 by the U.S. Geological Survey in cooperation with the Bay Area Water Users Association. A large amount of available geohydrologic and water-quality data were obtained from local, State, and Federal agencies, as well as private consultants. The data were acquired in many formats, including internal reports, published reports and maps, and computerized databases. To assist in the evaluation of the data on a regional scale, a database was created that consists of geographic information system data layers and tables and a series of system directories and American Standard Code for Information Interchange (ASCII) files.

The database includes 11 data layers that represent geographic features in the study area. The layers consist of features that represent point-specific data collected from wells and features that represent areal data such as thickness of alluvium, surficial geology, physiographic subareas, watershed boundaries, land use (1974-76, 1986-90, and modified 1986-90), water-supply districts, waste-water-treatment districts, and recharge basins. Each layer has a feature-attribute table that contains descriptive information about the features in the data layer. Each layer also contains three files that document the sources, scale, methods of auto-mation, accuracy, and attributes of the data layer. The Universal Transverse Mercator (UTM) pro-jection was utilized for each layer in the database. The study area is completely within UTM zone 10 and projection related distortions are considered minimal. Projection data are contained in a projection file that is stored with each data layer.

Most of the data collected for this study were obtained from wells located throughout the study area and are stored in the well-data layer. The well-data layer has 1,014 points representing wells for which data on the subsurface geohydrologic and water-quality conditions of the study area are available. The 1,014 points in the well-data layer are a subset of more than 25,000 points identified in the study area. Limitations in the timeframe of the study prevented the inclusion of additional data into the database. Data compilation efforts on a local scale have indicated that a greater density of data points can be achieved by using available data.

The well-data layer has four related tables that store data on construction, water levels, water quality, and pumping tests. The construction table consists of data on the construction and use of each well in the data layer. The water-level table con-sists of 35,845 water-level measurements from 293 wells. The water-level measurements were col-lected from July 1, 1936, to October 15, 1992. The water-quality table consists of 9,292 water-quality samples from 394 wells collected from August 16, 1937, to August 31, 1992. The water-quality data reported for each sample vary from a single constit-uent to a complete analysis of major ions, trace

metals, and organic contaminants. The pumping-test table consists of the results of pumping tests performed on 334 wells. The pumping-test data were used to calculate specific capacity. The lithologic data are stored in ASCII files and organized in system directories. Written lithologic descriptions for 762 wells were converted into numeric codes to be stored in the lithology files.

Interpretation of geohydrologic and water-quality conditions based on the data in the database could be affected by several potential errors in the data. Calculations of hydraulic heads and gradients from the water-level data can be affected by errors in the collection of the water-level data and errors in the determination of the altitude of land surface at the well. Errors in the land-surface altitude at the well generally are expected to be larger than errors in the collection of the water-level data. Analytical methods and quality-assurance procedures used by the laboratories that analyzed the water samples are unknown. Evaluation of the cation and anion balance for the 396 samples with the required constituents indicates that 51 samples (13 percent) have unacceptably high cation and anion balances results. Analysis of alkalinity in the laboratory rather than in the field can be a major cause of high cation and anion balance errors. Interpretation of the lithologic data is affected by the knowledge, experience, and vocabulary of the individual describing the drill cuttings.

In addition to providing a means to collect, manage, and analyze the large amount of data acquired from numerous sources, the GIS database will provide water-management agencies in the region the means to perform additional ground-water investigations. Because the contents of the database documented in this report were determined by the needs and available data at the time of this study, the value of the database to water managers can be increased by the addition of new data to the database. The scale of the source data could provide some limitations when applying the database at the local level.

References

Alameda County Water District, 1991, Water supply facilities: Alameda County Water District, 1 sheet, scale 1:24,000.

Anderson, J.R., Hardy, E.E., Reach, J.T., and Witmer, R.E., 1976, A land use and land cover classification system for use with remote sensor data: U.S. Geological Survey Professional Paper 964, 28 p.

Applied Consultants, 1991, Daly City ground-water investigation and model: Prepared for the City of Daly City, 149 p.

Atwater, B.F., Hedel, C.W., and Helley, E.J., 1977, Late quaternary depositional history, holocene sea-level changes, and vertical crustal movement, southern San Francisco Bay, California: U.S. Geological Survey Professional Paper 1014, 15 p.

Averett, R.C., Wood, P.R., and Muir, K.S., 1971, Water chemistry of the Santa Clara Valley, California: U.S. Geological Survey Open-File Report, 24 p.

Burau, J.R., and Cheng, R.T., 1989, A general method for generating bathymetric data for hydrodynamic models: U.S. Geological Survey Open-File Report 89-28, 45 p.

California Department of Water Resources, 1967, Evaluation of ground-water resources, south San Francisco Bay. Appendix A: Geology, Bulletin 118-1, 153 p.

Castro Valley Sanitary District, 1988, Castro Valley Sanitary District, 1 sheet, scale 1:19,200.

Earth Science Associates, 1991, Water resources evaluation of the north coast county water district service area: Earth Science Associates, 17 p.

Earth Science Associates and Ludhorff and Scalmanini, 1987, Half Moon Bay Airport/Pillar Point Marsh ground-water basin: Phase I study report, 18 p.

East Bay Municipal Utility District, 1986, Facility locations: East Bay Municipal Utility District, 1 sheet, scale 1:120,000.

Environmental Systems Research Institute, Inc., 1991a, ARC/INFO users guide: ARC/INFO Data Model, Concepts, and Key Terms: Environmental Systems Research Institute, Inc.

_____1991b, ARC/INFO users guide: Surface Modeling with TIN: Environmental Systems Research Institute, Inc.

_____1991c, ARC/INFO users guide: Map projections and coordinate management: Environmental Systems Research Institute, Inc.

Fio, J.L., and Leighton, D.A., 1994, Geohydrologic framework, historical development of the ground-water system and general hydrologic and water quality conditions in 1990, south San Francisco Bay and Peninsula area, California: U.S. Geological Survey Open-File Report 94-357, 47 p.

Geoconsultants, Inc., 1991, Annual report 1990-91, ground-water resources: Half Moon Bay, California: Geoconsultants, Inc., 23 p.

Hem, J.D., 1985, Study and interpretation of the chemical characteristics of natural water: U.S. Geological Survey Water-Supply Paper 2254, 263 p.

Hensolt, W.H., and Brabb, E.E., 1990, Maps showing elevation of bedrock and implications for design of engineered structures to withstand earthquake shaking in San Mateo County, California: U.S. Geological Survey Open-File Report 90-496, 11 p.

Lohman, S.W., 1972, Ground-water hydraulics: U.S. Geological Survey Professional Paper 708, 70 p.

Muehrcke, P.C., and Muehrcke, J.O., 1978, Map use: Reading, analysis, and interpretation: Madison, Wisc., J.P. Publications, 469 p.

Oro Loma Sanitary District, 1991, Sanitary sewage system: Oro Loma Sanitary District, 1 sheet, scale 1:12,000.

San Francisco Department of Public Works, 1977, San Francisco hydrologic and hydraulic data acquisition and recording system location map. Drainage boundaries, level monitors, and rain gages: San Francisco Department of Public Works, 1 sheet, scale 1:22,500.

San Mateo County Planning and Development Division, 1986, Water suppliers: San Mateo County Planning and Development Division, 1 sheet, scale 1:48,000.

_____1991, Wastewater service areas: San Mateo County Planning and Development Division, 1 sheet, scale 1:48,000.

Santa Clara Valley Water District, 1988, Water supply facilities: North valley: Santa Clara Valley Water District, 1 sheet, scale 1:60,000.

_____1990, Preliminary water service areas: Santa Clara Valley Water District, 1 sheet.

U.S. Geological Survey, 1986, Land use and land cover digital data from 1:250,000- and 1:100,000-scale maps: National Mapping Program Technical Instructions Data Users Guide 4, 36 p.

_____1990, Digital elevation models: National Mapping Program Technical Instructions Data Users Guide 5, 51 p.

Wentworth, C.M., 1993, General distribution of geologic materials in the southern San Francisco Bay region, California: A digital map database: U.S. Geological Survey Open-File Report 93-693, 10 p.