A DATA BASE OF NITRATE IN GROUND-WATER SAMPLES FROM THE CONTERMINOUS UNITED STATES

By Kenneth J. Lanfear

U.S. GEOLOGICAL SURVEY
Open-File Report 92-652

Reston, Virginia
1992

U.S. DEPARTMENT OF THE INTERIOR

MANUEL LUJAN, JR., Secretary

U.S. GEOLOGICAL SURVEY

Dallas L. Peck, Director

For additional information write to:

Assistant Chief Hydrologist, Water
Assessment and Data Coordination
U.S. Geological Survey
12201 Sunrise Valley Drive, MS 407
Reston, Virginia 22092

Copies of this report can be purchased from:

U.S. Geological Survey
Books and Open-File Reports Section
Box 25425, Mail Stop 517
Federal Center
Denver, CO 80225-0425

CONTENTS

A Data Base of Nitrate in Ground-Water Samples from the Conterminous United States

Kenneth J. Lanfear
U.S. Geological Survey

ABSTRACT

To facilitate studies of the distribution of nitrate in ground water in the conterminous United States, a data base was constructed that comprises all nitrate concentration data for ground-water samples recorded by the U.S. Geological Survey (USGS). The data comprise 474,484 analyses from 146,751 wells and represent samples taken from 1889 through 1990.

INTRODUCTION

Nitrate concentrations in ground water can be increased by fertilizer applications, discharges from septic tanks, or by surface contamination from poor well construction. High concentrations of nitrate could indicate human influence on the ground-water system. The U.S. Environmental Protection Agency standard for nitrate in drinking water is 10 milligrams per liter (mg/L) as nitrogen.

To facilitate studies of the distribution of nitrate in ground water in the conterminous United States, a data base was constructed that comprises all nitrate concentration data for ground-water samples recorded by the USGS. The U.S. Environmental Protection Agency provided partial funding to USGS for this project. This report describes how this data base was constructed. It is anticipated that other studies using this data base will include analysis and screening of these data and will add data from non-USGS sources including Federal and State agencies.

LINEAGE

USGS data on ground-water quality are stored in the National Water Information System (NWIS) of USGS. At the time of this retrieval (1991) different parts of the data base resided on a network of 46 computers located in USGS offices throughout the United States. The data on each computer pertain to the geographic area served by the local USGS office. All data on nitrate, along with relevant information about the wells themselves, were extracted from each computer and transferred to a single computer in Reston, Virginia. No limits were placed upon date of sampling or the type of well; if nitrate data resided in the data base when the local computer was queried, they were transferred to Reston.

The retrieval process began in July, 1991 and continued through October, 1991. Dates of samples retrieved range from 1889 to 1990. Data for years prior to computerization of the data base were retrieved only if they had been transferred to the computer. Whether any additional 1990 data being entered into NWIS during this period was retrieved depends upon when the local node was queried. A log was kept of all retrieval operations.

Values for concentrations of dissolved nitrate are stored under two parameter codes, each with different units. There also are parameter codes for different forms of nitrogen which may, in the absence of a dissolved nitrate value, be used to approximate the concentration of dissolved nitrate in ground water. Data stored under the following parameters were retrieved and stored in their original units:

Parameter Code	Description
600	Nitrogen, total, mg/L as nitrogen
602	Nitrogen, dissolved, mg/L as nitrogen
618	Nitrate, dissolved, mg/L as nitrogen
620	Nitrate, total, mg/L as nitrogen
630	Nitrite plus Nitrate, total, mg/L as nitrogen
631	Nitrite plus Nitrate, dissolved, mg/L as nitrogen
71850	Nitrate, total, mg/L as nitrate
71851	Nitrate, dissolved, mg/L as nitrate
71887	Nitrogen, total, mg/L as nitrate

All nitrate data were reformatted into an ARC/INFO[1] Geographic Information System (GIS) data base comprising four main files:

NITRATE.VALUE	-- Measured nitrogen concentrations.
NITRATE.EVENT	-- Date and time for each sample, and the depth of sampling.
NITRATE.SITE	-- Information about each site where data were collected.
NITRATE.PAT	-- GIS location for each site.

The relationship among these files is:

NITRATE.VALUE <<---> NITRATE.EVENT <<---> NITRATE.SITE <<---> NITRATE.PAT

where <<---> indicates a one-to-many relationship. That is, there can be many values for one event, many events at a site, and many sites at a location. The NITRATE.VALUE file is sorted by constituent code, site, and date. The file, NITRATE.SOURCE, records the information about the transfer of data to this data base. The data base files are described in more detail in Appendix A.

Depth of sampling can be critical to understanding the distribution of chemical species in ground water. A maximum and minimum depth of sampling was associated with each sample, where sufficient information was available. Ideally, this would be parameter codes 72015 and 72016, depth to the top and bottom of the sampling interval, in feet, or 82047 and 82048, the equivalent values in meters. Where these codes were missing, an inference procedure was used

1. Use of trade names is for identification purposes only and is not an endorsement by the U.S. Geological Survey.

whereby the missing values were estimated from (1) other depth measurements (e.g. top/bottom of water zone, depth of well) recorded for the event, (2) depth measurements made at the same site at different times, or (3) a default value associated with the site information. A remark code indicates the method used to infer depth.

DATA QUALITY

Positional Accuracy

Location of each well was determined from the latitude and longitude reported in the Ground-Water Site Inventory (GWSI) file of NWIS, which records these values to the nearest second. The accuracy code in GWSI is recorded in the NITRATE.SITE file. If the GWSI file did not record the location, the location was inferred from the site identification number, which is based upon latitude and longitude to the nearest second.

Attribute Accuracy

Except for depth of sampling, which was inferred as described previously, all attributes were recorded as found in NWIS.

Topological Consistency

All nitrate values in the NITRATE.VALUE file are associated, through the relational structure described previously, with a location in the NITRATE coverage.

Completeness

The data base represents a complete picture of USGS ground-water nitrate computerized data holdings through 1990. The data include 474,484 nitrogen samples taken for 303,251 events at 146,751 sites from 1889 through 1990.

AVAILABILITY

A copy of the data base in its native ARC/INFO format, approximately 64 megabytes in size, is available from the author.

APPENDIX A -- Descriptions of Files.

[COL is the starting position in the file, ITEM NAME is the name of the item, WDTH is the internal length of the item, in bytes, OPUT is the length when printed, TYP is the item type (B=binary, F=floating point, I=integer, C=character), and N.DEC is the number of digits following the decimal when printed)]

DATAFILE NAME: NITRATE.VALUE
6 ITEMS: STARTING IN POSITION 1

COL	ITEM NAME	WDTH	OPUT	TYP	N.DEC	Description
1	EVENT#	4	7	B	-	Pointer to NITRATE.EVENT file.
5	EVENT-ID	4	7	B	-	Event number.
9	PARM	4	5	B	-	Parameter code.
13	REMARK	1	1	C	-	Remark code.
14	VALUE	4	9	F	3	Value recorded.
18	KEEP	1	1	I	-	If not = 0, ignore record.

DATAFILE NAME: NITRATE.EVENT
9 ITEMS: STARTING IN POSITION 1

COL	ITEM NAME	WDTH	OPUT	TYP	N.DEC	Description
1	EVENT-ID	4	7	B	-	Event number.
5	SITE#	4	7	B	-	Pointer to NITRATE.SITE file.
9	SITE-ID	4	7	B	-	Site number.
13	DATE	8	8	C	-	Date of sampling (-1=missing).
21	TIME	4	4	C	-	Time of sampling (-1=missing).
25	DTOP	4	8	F	2	Depth to top of sampling, ft.
29	DTOP_R	1	1	C	-	Depth remark code.
30	DBOT	4	8	F	2	Depth to bottom of sampling, ft.
34	DBOT_R	1	1	C	-	Depth remark code.
	** REDEFINED ITEMS **					
13	YEAR	4	4	I	-	Year of sampling.
17	MONTH	2	2	I	-	Month of sampling.
19	DAY	2	2	I	-	Day of sampling.
21	HOUR	2	2	I	-	Hour of sampling.
23	MINUTE	2	2	I	-	Minute of sampling.

DATAFILE NAME: NITRATE.SITE
44 ITEMS: STARTING IN POSITION 1

COL	ITEM NAME	WDTH	OPUT	TYP	N.DEC	Description
1	SITE-ID	4	7	B	-	Site number.
5	NITRATE#	4	5	B	-	Pointer to NITRATE.PAT file.
9	NITRATE-ID	4	7	B	-	Location number.
13	SAGNCY	5	5	C	-	C0004[2] Source agency.
18	SID	15	15	C	-	C0001 Site identification.
33	SNAME	24	24	C	-	C0012 Site name.
57	SLAT	7	7	C	-	C0009 Latitude.
64	SLONG	8	8	C	-	C0010 Longitude.
72	SCORAC	1	1	C	-	C0011 Lat-long accuracy.
73	SDIST	3	3	C	-	C0006 District code.
76	SSTATE	2	2	C	-	C0007 State code.
78	SCNTY	3	3	C	-	C0008 County code.

2. The Cxxxx codes are the parameters as found in the Ground-Water Site Inventory System (GWIS) file of NWIS.

COL	ITEM NAME	WDTH	OPUT	TYP	N.DEC	Description
81	SDATUM	4	8	F	2	C0016 Land-surface datum.
85	SALTMT	1	1	C	-	C0017 Method altitude deter.
86	SALTAC	3	3	C	-	C0018 Altitude accuracy code
89	STYPE	7	7	C	-	C0802 Station type codes.
96	SUSE	1	1	C	-	C0803 Agency use of site code.
97	SCLASS	1	1	C	-	C0003 Data reliability code.
98	SGWFLE	20	20	C	-	C0712 Data-other GW files.
118	SGWTYPE	1	1	C	-	C0002 Type of ground-water site.
119	SUSE1	1	1	C	-	C0023 Primary use of site.
120	SUSE2	1	1	C	-	C0301 Secondary use of site.
121	SUSE3	1	1	C	-	C0302 Tertiary use of site.
122	SWUSE1	1	1	C	-	C0024 Primary use of water.
123	SWUSE2	1	1	C	-	C0025 Secondary use of water.
124	SWUSE3	1	1	C	-	C0026 Tertiary use of water.
125	SAQTYP	1	1	C	-	C0713 Aquifer-type code.
126	SAQUFR	8	8	C	-	C0714 Aquifer code.
134	SHDEPT	4	8	F	2	C0027 Hole depth.
138	SWDEPT	4	8	F	2	C0028 Well depth.
142	SWDSRC	1	1	C	-	C0029 Source of depth data.
143	SINVWL	4	7	F	2	C0030 Inventory water level.
147	SIWLDT	8	8	C	-	C0031 Date of water-level meas.
155	SIWLMT	1	1	C	-	C0034 Method water level deter
156	SIWLST	1	1	C	-	C0037 Site status.
157	SIWLSC	1	1	C	-	C0033 Source of water-level data
158	DDTOP	4	8	F	2	Default depth to top of interval, ft.
162	DDTOP_R	1	1	C	-	Depth remark code.
163	DDBOT	4	8	F	2	Default depth bottom of interval, ft.
167	DDBOT_R	1	1	C	-	Depth remark code.
168	SITE_STATUS	1	1	I	-	Site status.
169	SOURCE-ID	4	5	B	-	Source of data.
173	DNOTE	20	20	C	-	Note.
193	NUMEVENT	4	7	B	-	Number of events at site.
	** REDEFINED ITEMS **					
13	FULL_SITE_ID	20	20	C	-	Agency-SID combination.
57	LATD	3	3	I	-	Latitude, degrees.
60	LATM	2	2	I	-	Latitude, minutes.
62	LATS	2	2	I	-	Latitude, seconds.
64	LOND	4	4	I	-	Longitude, degrees (West is negative)
68	LONM	2	2	I	-	Longitude, minutes.
70	LONS	2	2	I	-	Longitude, seconds.
57	LATLON	13	13	C	-	Latitude-longitude
76	FIPS	5	5	I	-	County code.
76	FIPSST	2	2	I	-	State code.
13	IDLATD	2	2	I	-	Inferred lat/lon from
15	IDLATM	2	2	I	-	SID information. Use
17	IDLATS	2	2	I	-	this if lat/lon is not
19	IDLOND	3	3	I	-	explicitly given.
22	IDLONM	2	2	I	-	
24	IDLONS	2	2	I	-	
13	L01	68	68	C	-	Fields used for data entry.
81	L02	4	8	F	2	
85	L03	49	49	C	-	
134	L04	4	8	F	2	

COL	ITEM NAME	WDTH	OPUT	TYP	N.DEC	Description
138	L05	4	8	F	2	
142	L06	1	1	C	-	
143	L07	4	7	F	2	
147	L08	11	11	C	-	

DATAFILE NAME: NITRATE.PAT
7 ITEMS: STARTING IN POSITION 1

COL	ITEM NAME	WDTH	OPUT	TYP	N.DEC	Description
1	AREA	4	12	F	3	ARC/INFO required items
5	PERIMETER	4	12	F	3	
9	NITRATE#	4	5	B	-	
13	NITRATE-ID	4	5	B	-	
17	WORKB	4	5	B	-	Working variables for
21	SYMBOL	4	5	B	-	plotting and analysis
25	VALUE	4	12	F	3	

DATAFILE NAME: NITRATE.DEPTH_REMARKS
2 ITEMS: STARTING IN POSITION 1

COL	ITEM NAME	WDTH	OPUT	TYP	N.DEC	Description
1	DTOP_R	1	1	C	-	Depth remark code.
2	DEPTH_REMARK	70	70	C	-	Meaning of code (see below)
	** REDEFINED ITEMS **					
1	DBOT_R	1	1	C	-	Depth remark code.
1	DDTOP_R	1	1	C	-	"
1	DDBOT_R	1	1	C	-	"

The following are the values for DEPTH_REMARKS codes:

DTOP_R	DEPTH_REMARK
D	Value is inferred from sampling depths recorded at the site
E	Value is inferred from hole depth recorded at the site
G	Value is inferred from hole depth reported in GWSI
H	Value is inferred from hole depth recorded for event
M	Assumed to match the corresponding top or bottom sample
S	Sampling data
X	No data

DATAFILE NAME: NITRATE.SOURCE
5 ITEMS: STARTING IN POSITION 1

COL	ITEM NAME	WDTH	OPUT	TYP	N.DEC	Description
1	SOURCE-ID	4	5	B	-	Key to file.
5	SOURCE_FILE_GWSI	50	50	C	-	File name for GWSI data.
55	SOURCE_FILE_QW	50	50	C	-	File name for QW data
105	SOURCE_WHEN	40	40	C	-	Date entered into data base
145	SOURCE_NOTE	60	60	C	-	Note on source material.