

Summary: How censored, estimated, and missing data were handled by the Glacial Aquifer Team for regional analysis of ground-water data from the glacial aquifer system and by NST for national synthesis of cycle I (1992-2001) trace element and nutrient data

Compiled by Terri Arnold

Treatment of Data by the Glacial Aquifer Team (extracts from metadata of the datasets prepared by Terri Arnold 05/20/2004)

VOC, radionuclides, stable isotopes, and dissolved organic carbon

data retrieved from the NAWQA Internal Data Warehouse on May 20, 2004. The data will be used in a series of reports to characterize the Glacial Principal Aquifer as part of the U.S. Geological Survey, National Water-Quality Assessment Program. For detailed information about this data set, see

http://hexe.er.usgs.gov/usgs/uirb/gapt/data/GLAC_vocradchemisotope.README

The data retrieved had multiple censoring levels for most constituents. Censored data is indicated by a '<' or '>' in the remark column (remark columns begin with a 'R_' in the final data set.) Due to the large number of censored values, it was determined by those team members that would use the VOC data that the censored values would not be treated in this data set.

Individual studies will be responsible for treating the censored values or for using them to indicate percent detection.

Field parameter, major ion, trace element, and nutrient data retrieved from the NAWQA Internal Data Warehouse on May 20, 2004. The data will be used in a series of reports to characterize the Glacial Principal Aquifer as part of the U.S. Geological Survey, National Water-Quality Assessment Program. For detailed information about this data set, see

http://hexe.er.usgs.gov/usgs/uirb/gapt/data/GLAC_fieldmajortenuts.README

The data retrieved had multiple censoring levels for most constituents.

Censored data is indicated by a '<' or '>' in the remark column (remark columns begin with a 'R_' in the final data set.) Glacial data will be compared with data for the Cambrian-Ordovician and York-New England aquifers so when dealing with censored values, data for these aquifers was also considered. For analysis, censored values were handled similar to the method employed by the NAWQA Trace Elements National Synthesis Team as follows (original data prior to handling censored values is stored in columns with a title beginning with a 'P'. Post handling of censored values is stored in columns whose titles begin with an 'A_'):

10/19/2004

1) For any constituent that has censored values, insert an additional column to the right of the value column (titled with the PCODE prefixed by 'P') and label the new column with the PCODE number prefixed by an 'A_'.

2) Look at all the censored values for each constituent. Identify the most common detection or reporting.

i) Values that are censored above the most common detection or reporting limit: delete them and place a '-' in the 'A_' column.

ii) Censored values including the E-marked values below the common detection or reporting limit: convert them to the most common detection or reporting limit in the 'A_' column.

iii) E-remarked values that are above the common reporting limit: keep that E-remarked value in the 'A_' column.

#

#	PCODE	censored value used
---	-------	---------------------

#	90410	--
---	-------	----

#	01106	1
---	-------	---

#	00623	0.2
---	-------	-----

#	00608	0.04
---	-------	------

#	01095	1
---	-------	---

#	01000	1
---	-------	---

#	01005	1
---	-------	---

#	01010	1
---	-------	---

#	01020	12
---	-------	----

#	71870	0.01
---	-------	------

#	01025	1
---	-------	---

#	00915	--
---	-------	----

#	00940	0.1
---	-------	-----

#	01030	0.8
---	-------	-----

#	01035	1
---	-------	---

#	01040	1
---	-------	---

#	72019	--
---	-------	----

#	00300	0.1
---	-------	-----

#	00950	0.1
---	-------	-----

#	01046	10
---	-------	----

#	01049	1
---	-------	---

#	01130	6
---	-------	---

#	00925	0.01
---	-------	------

#	01056	1
---	-------	---

#	01060	1
---	-------	---

#	01065	1
---	-------	---

#	00631	0.05
---	-------	------

#	00613	0.01
---	-------	------

#	00671	0.01
---	-------	------

#	0090	--
---	------	----

#	00666	0.01
---	-------	------

#	00935	0.1
---	-------	-----

10/19/2004

70300 --
01145 1
00955 --
01075 1
00930 0.2
90095 --
00095 --
01080 0.2
00945 0.1
00010 --
01057 0.04
22703 1
01085 10
01090 1
00400 --
00403 --

Pesticide data retrieved from the NAWQA Internal Data Warehouse on May 20, 2004. The data will be used in a series of reports to characterize the Glacial Principal Aquifer as part of the U.S. Geological Survey, National Water-Quality Assessment Program. For detailed information about this data set, see http://hexe.er.usgs.gov/usgs/uirb/gapt/data/GLAC_pest2001_03_LCAA.README

Due to the large number of censored values, it was determined by those team members that would use the Pesticide data that the censored values would not be treated in this data set. Individual studies will be responsible for treating the censored values or for using them to indicate percent detection.

Treatment of Data by NST

Trace Elements Data (NST data set) (extracts from InfoTEAnalysis.doc provided by Lori Apodaca Sept-2003)

Trace elements and radionuclides analyzed include: Antimony (Sb), Arsenic (As), Beryllium (Be), Lead (Pb), Molybdenum (Mo), Radon (Rn), Selenium (Se), Thallium (Th), Uranium (U), and Vanadium (V).

Data that are identified as being less than the detection or reporting limits are censored data. Because of changes in the development of analytical methods for trace elements in Cycle I several changes in the reporting levels for most TEs has occurred. Also, beginning in October 1999, the NWQL provided semi-quantified (E-remarked) values below the RL for selected trace elements. In order to statistically compare the water-quality data for a given constituent and not exclude all of the censored data, these

data needed to have a common detection or reporting limit (CRL). For the data used in this report, the most common detection or reporting limit was kept. However, values that were censored above the most common detection or reporting limit were deleted and censored values including the E-marked values below the common detection or reporting limit were converted to the most common detection or reporting limit before statistically analyzing the data. In addition E- remarked values that were above the common reporting limit were kept and that value was used in the statistical summaries.

Nutrients Data (NST data set)
(extracts from InfoNutrientAnalysis.doc provided by Lori Apodaca Sept-2003)

Nutrient analyses reported here consist of dissolved concentrations of ammonia (P00608), nitrite-plus-nitrate (P00631) (designated “nitrate” because nitrite contribution generally is negligible), and orthophosphate (P00671).

Censored values were set to one-half the detection limit before calculating network medians. In the case of ammonia, which has two reporting levels, the higher one (0.015 mg/L) was used. Exceedance percentages of the nitrate maximum contaminant level (MCL) were determined for each land-use study and subunit survey based on the number of wells in the network with nitrate concentration greater than 10 mg/L.