science for a changing world

National Water-Quality Assessment Program

Methods for Processing and Summarizing Time-Series Temperature Data Collected as Part of the National Water-Quality Assessment Program Studies on the Effects of Urbanization on Stream Ecosystems

Data Series 330

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

Methods for Processing and Summarizing Time-Series Temperature Data Collected as Part of the National Water-Quality Assessment Program Studies on the Effects of Urbanization on Stream Ecosystems

By Thomas F. Cuffney and Robin A. Brightbill

National Water-Quality Assessment Program

Data Series 330

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

U.S. Department of the Interior
DIRK KEMPTHORNE, Secretary

U.S. Geological Survey
Mark D. Myers, Director

U.S. Geological Survey, Reston, Virginia: 2008

For product and ordering information:
World Wide Web: http://www.usgs.gov/pubprod
Telephone: 1-888-ASK-USGS

For more information on the USGS—the Federal source for science about the Earth, its natural and living resources, natural hazards, and the environment:
World Wide Web: http://www.usgs.gov
Telephone: 1-888-ASK-USGS

Suggested citation:
Cuffney, T.F., and Brightbill, R.A., 2008, Methods for processing and summarizing time-series temperature data collected as part of the National Water-Quality Assessment Program studies on the effects of urbanization on stream ecosystems: U.S. Geological Survey Data Series 330, 23 p. (available only online at *http://pubs.water.usgs.gov/ds330/*).

Foreword

The U.S. Geological Survey (USGS) is committed to providing the Nation with credible scientific information that helps to enhance and protect the overall quality of life and that facilitates effective management of water, biological, energy, and mineral resources (*http://www.usgs.gov/*). Information on the Nation's water resources is critical to ensuring long-term availability of water that is safe for drinking and recreation and is suitable for industry, irrigation, and fish and wildlife. Population growth and increasing demands for water make the availability of that water, now measured in terms of quantity and quality, even more essential to the long-term sustainability of our communities and ecosystems.

The USGS implemented the National Water-Quality Assessment (NAWQA) Program in 1991 to support national, regional, State, and local information needs and decisions related to water-quality management and policy (*http://water.usgs.gov/nawqa*). The NAWQA Program is designed to answer: What is the condition of our Nation's streams and ground water? How are conditions changing over time? How do natural features and human activities affect the quality of streams and ground water, and where are those effects most pronounced? By combining information on water chemistry, physical characteristics, stream habitat, and aquatic life, the NAWQA Program aims to provide science-based insights for current and emerging water issues and priorities. From 1991–2001, the NAWQA Program completed interdisciplinary assessments and established a baseline understanding of water-quality conditions in 51 of the Nation's river basins and aquifers, referred to as Study Units (*http://water.usgs.gov/nawqa/studyu.html*).

Multiple national and regional assessments are ongoing in the second decade (2001–2012) of the NAWQA Program as 42 of the 51 Study Units are reassessed. These assessments extend the findings in the Study Units by determining status and trends at sites that have been consistently monitored for more than a decade, and filling critical gaps in characterizing the quality of surface water and ground water. For example, increased emphasis has been placed on assessing the quality of source water and finished water associated with many of the Nation's largest community water systems. During the second decade, NAWQA is addressing five national priority topics that build an understanding of how natural features and human activities affect water quality, and establish links between sources of contaminants, the transport of those contaminants through the hydrologic system, and the potential effects of contaminants on humans and aquatic ecosystems. Included are topics on the fate of agricultural chemicals, effects of urbanization on stream ecosystems, bioaccumulation of mercury in stream ecosystems, effects of nutrient enrichment on aquatic ecosystems, and transport of contaminants to public-supply wells. These topical studies are conducted in those Study Units most affected by these issues; they comprise a set of multi-Study-Unit designs for systematic national assessment. In addition, national syntheses of information on pesticides, volatile organic compounds (VOCs), nutrients, selected trace elements, and aquatic ecology are continuing.

The USGS aims to disseminate credible, timely, and relevant science information to address practical and effective water-resource management and strategies that protect and restore water quality. We hope this NAWQA publication will provide you with insights and information to meet your needs, and will foster increased citizen awareness and involvement in the protection and restoration of our Nation's waters.

The USGS recognizes that a national assessment by a single program cannot address all water-resource issues of interest. External coordination at all levels is critical for cost-effective management, regulation, and conservation of our Nation's water resources. The NAWQA Program, therefore, depends on advice and information from other agencies—Federal, State, regional, interstate, Tribal, and local—as well as nongovernmental organizations, industry, academia, and other stakeholder groups. Your assistance and suggestions are greatly appreciated.

Robert M. Hirsch
Associate Director for Water

Contents

Figures

Tables

Acronyms, Abbreviations, and Definitions

Temperature in degrees Celsius (°C) may be converted to degrees Fahrenheit (°F) as follows:

$$°F = (1.8 \times °C) + 32$$

Temperature in degrees Fahrenheit (°F) may be converted to degrees Celsius (°C) as follows:

$$°C = (°F - 32) / 1.8$$

Project and metropolitan area abbreviations and definitions

ACFB	Apalachicola-Chattahoochee-Flint River basins
ALBE	Albemarle-Pamlico drainages
ATL	Atlanta, GA, metropolitan area
BIR	Birmingham, AL, metropolitan area
BOS	Boston, MA, metropolitan area
CITY	Metropolitan area studied
DEN	Denver, CO, metropolitan area
DFW	Dallas-Fort Worth, TX, metropolitan area
EUSE	Effects of Urbanization on Stream Ecosystems studies
GRSL	Great Salt Lake basins
MGB	Milwaukee-Green Bay, WI, metropolitan area
MOBL	Mobile River basin
NAWQA	National Water-Quality Assessment Program
NECB	New England Coastal basins
PORT	Portland, OR, metropolitan area
RAL	Raleigh-Winston Salem, NC, metropolitan area
SLC	Salt Lake City, UT, metropolitan area
SPLT	South Platte River basin
Study Units	Geographical areas representing major river basins in the NAWQA Program
SUID	A four-letter abbreviation used to identify NAWQA Program Study Units
TRIN	Trinity River basin
USGS	U.S. Geological Survey
WILL	Willamette River basin
WMIC	Western Lake Michigan drainages

Sampling abbreviations and definitions

Collection date	The date on which a temperature measurement was made
DATE	Collection date
oTime	The time (hour:minute) at which a temperature measurement was made (observation time)
SABBREV	Shortened or abbreviated station name
ShortName	Short version of station name
STAID	USGS station identification number
StationName	USGS station name
TZCD	Time zone code (for example, EST for Eastern Standard Time)
VALUE	Temperature measurement

Computer, programming, and statistical abbreviations and definitions

ADAPS	Automated Data Processing System. USGS software tools used to extract data from NWIS
Ave	Average (arithmetic mean) of temperature measurements
Batch mode	The simultaneous processing of data from multiple sampling sites
CumMin	Cumulative elapsed time in minutes since the start of data collection or POR
dDays, Degree days	The sum of daily mean temperatures for a specified time period
°C	Degree Celsius
°C/h	Degree Celsius per hour
Diff	The time difference (minutes) between the granularity time (TIME) and the time at which the temperature measurement was made (oTime)
DMIN	Time of day expressed in minutes past midnight
GRAN	Granularity; name of computer program developed to derive uniform granularity and calculate temperature statistics
Granularity	The time between temperature measurements in minutes
h	Hour
Max	Maximum of temperature measurements for a specified time period
Min	Minimum of temperature measurements for a specified time period
N	Number of temperature measurements used to calculate statistics for a specific time period
N_?h	Number of temperature rate changes based on measurements taken at ? h apart, where ? can be 1, 2, 4, 8, or 12 h
NWIS	National Water Information System. The database in which the USGS stores temperature data
POR	Period of record, which is the time over which temperature measurements were made or summarized
Range	Range (Max – Min) in temperature measurements for a specified time period
RDB	A tab-delimited data file format that can be read as a relational data base in UNIX using awk
SD	Standard deviation of temperature measurements for a specified time period
Spreadsheet	A series of rows and columns that holds information in Microsoft Excel® files
TIME	The time (hour:minute) that corresponds to the granularity selected for a data set

tRate?h	Rate of change (units/hour) between tVal? and VALUE where ? can be 1, 2, 4, 8, or 12 h
tRate?h_Ave	Mean temperature rate change (units/hour) for the date based on absolute values of differences in measurements taken ? h apart where ? can be 1, 2, 4, 8, or 12 h
tRate?h_Max	Maximum temperature rate change (units/hour) for the date based on absolute values of differences in measurements taken ? h apart where ? can be 1, 2, 4, 8, or 12 h
tRate?h_Min	Minimum temperature rate change (units/hour) for the date based on absolute values of differences in measurements taken ? h apart where ? can be 1, 2, 4, 8, or 12 h
tRate?h_n_%SD	Number of temperature rate changes (absolute values) that were ≥ %–1 and < % SD of the mean rate for the POR based on measurements taken ? h apart where % can be 1, 2, 3, 4, 5, or gt5 (> 5) and ? can be 1, 2, 4, 8, or 12 h
tRate?h_RANGE	Range in temperature rate changes (units/hour) for the date based on absolute values of differences in measurements taken ? h apart where ? can be 1, 2, 4, 8, or 12 h
tRate?h_SD	Standard deviation in temperature rate changes (units/hour) for the date based on absolute values of differences in measurements taken ? h apart where ? can be 1, 2, 4, 8, or 12 h
tTime?h	CumMin value for a time ? h prior to TIME where ? can be 1, 2, 4, 8, or 12 h
tVal?h	The temperature value at oTime corresponding to tTime?h where ? can be 1, 2, 4, 8, or 12 h
Visual Basic	A computer programming language for Microsoft Windows®
Workbook	A collection of spreadsheets contained within one Microsoft Excel® file
>	Greater than
≥	Greater than or equal to
<	Less than
≤	Less than or equal to

Ecological definitions

Assimilative capacity	The ability of a waterbody to cleanse itself
Biodiversity	The number and variety of organisms found within a specified region

Methods for Processing and Summarizing Time-Series Temperature Data Collected as Part of the National Water-Quality Assessment Program Studies on the Effects of Urbanization on Stream Ecosystems

By Thomas F. Cuffney and Robin A. Brightbill

Abstract

Temperature data and summary statistics are presented for 256 sites in 9 metropolitan areas as part of the U.S. Geological Survey National Water-Quality Assessment Program studies of the effects of urbanization on stream ecosystems. The computer program (GRAN) that was developed to derive uniform data granularity and calculate temperature statistics (means, standard deviations, rates of change, degree days) is described, as are the methods used to estimate missing daily mean temperatures, degree days (annual and summer periods), and 7-day running averages of daily mean temperatures.

Introduction

U.S. Geological Survey (USGS) National Water-Quality Assessment (NAWQA) Program studies of the effects of urbanization on streams across the conterminous United States are being conducted as part of the Effects of Urbanization on Stream Ecosystems (EUSE) (Tate and others, 2005; fig. 1). Urbanization has been shown to degrade streams by altering physical, chemical, and biological characteristics (Paul and Meyer, 2001; Coles and others, 2004; Brown and others, 2005). Alteration of the natural temperature regime (Pluhowski, 1970; Klein, 1979) is particularly important because water temperature has a critical influence on determining the distribution, growth, and survival of aquatic organisms (Vannote and Sweeney, 1980; Barthalow, 1989; Holmes and Regier, 1990; Armour, 1991). In addition, water temperature can influence migration patterns; egg maturation and incubation success; inter- and intra-specific competition of aquatic organisms; and resistance to parasites, diseases, and pollutants (Armour, 1991). Changes in water temperature also alter the rates of instream chemical reactions and decomposition (Paul and others, 1978). Collectively, these changes can affect the ability of streams to support aquatic life and to process pollutants. Alterations in the natural temperature regime can cause stream ecosystems to change from a stable state with high biodiversity and high assimilative capacity to an unstable state with low biodiversity and low assimilative capacity (LeBlanc and others, 1997).

Purpose and Scope

This report documents methods for processing and summarizing time-series water temperature data collected for the EUSE studies as part of the USGS NAWQA Program and introduces the software (GRAN) developed to process the data. Temperature statistics are presented for 256 EUSE

Figure 1. Locations of nine metropolitan areas from which water temperature data were collected across the conterminous United States.

sites (**Site_List.xls**) in 9 metropolitan areas (fig. 1) across the conterminous United States (ATL: Atlanta, GA; BOS: Boston, MA; BIR: Birmingham, AL; DFW: Dallas-Fort Worth, TX; DEN: Denver, CO; MGB: Milwaukee-Green Bay, WI; PORT: Portland, OR; RAL: Raleigh-Winston Salem, NC; SLC: Salt Lake City, UT). Data were collected using a variety of equipment, data-collection intervals, and study periods. Equipment problems, droughts, floods, and ice resulted in incomplete records for many sites. The methods set forth in this report provide a common approach for

1. transferring temperature data from USGS databases,
2. extracting data for annual and summer periods,
3. establishing a common measurement granularity (that is, a common time interval between temperature measurements),
4. characterizing temperature statistics for daily, annual, and summer periods,
5. estimating missing daily mean temperature values,
6. estimating total degree-days for annual and summer periods, and
7. characterizing temperature trends using 7-day running average daily mean temperatures.

These methods were coded into a computer program (GRAN) that provides an efficient method for processing large amounts of time-series temperature data (see appendix 1). These methods and the GRAN program can be used to process and summarize other time-series data (for example, stream stage or discharge) collected by USGS and non-USGS researchers. Files containing site information, temperature data, summary statistics, example calculations, and examples of file formats used to enter and process data are described in table 1.

Processing Water Temperature Data

Temperature data were collected at 256 gaged and ungaged sites across the conterminous United States. Standard USGS data-collection protocols were followed at gaged sites, whereas a variety of data loggers recorded temperature measurements at ungaged sites at intervals ranging from 1 to 60 minutes. Shorter measurement intervals (1 to 5 minutes) often were used when equipment was installed in order to evaluate the operation of the data logger and methods of data retrieval. Longer time intervals (30 to 60 minutes) were used when data could not be retrieved frequently and data-storage capacity had to be conserved. Most commonly, temperature data were logged at 15-minute intervals. The time period over which data were collected varied among studies, depending on how rapidly sites were selected for study and equipment was obtained and installed. Equipment malfunctions, floods, droughts, and winter ice resulted in loss of data and the generation of incomplete annual temperature records for some sites. Temperature data were processed to produce temperature records and data summaries that were comparable among study areas.

Acquiring Water Temperature Data

Water temperature data collected for the EUSE studies were processed and stored in the USGS National Water Information System (NWIS) Automated Data Processing System (ADAPS), which served as the database of EUSE temperature records. Data were exported from ADAPS as tab-delimited ASCII text files (RDB; table 2) that were either processed directly by the GRAN program or were imported into Microsoft Excel® and modified to a simpler file format (table 3) for use with the GRAN program. The GRAN program was used to create separate Access® databases for each of the nine metropolitan areas. The relevant water temperature records for each site (26–30 sites per area) were imported into the appropriate database using the “import” function of the GRAN program.

Selecting Annual and Summer Periods

The most complete annual period of record was extracted for each EUSE study using the automated annual period-of-record (POR) batch-extraction procedure in the GRAN program. This procedure combines data from all the sites, identifies all possible annual PORs starting with the earliest date that data were collected, and then determines the POR (or PORs) that maximizes the number of sites with data. If more than one annual period maximizes the number of sites with data, the program selects the period with the most temperature measurements. If more than one annual period meets these criteria, the program selects the period with the earliest start date as the best annual POR. The annual POR for each site then is extracted to a new file that is identified by appending the starting date for the annual period (for example, 02337395_7_30_2002). The resulting annual period is consistent for all sites within a metropolitan study area, but the beginning and ending dates for the annual POR may vary among metropolitan areas.

The summer period, which is defined as 1 June to 30 September, was extracted from the best annual POR using the manual option for selecting the POR in the GRAN program. The best annual POR for four of the metropolitan areas (ATL, RAL, BOS, DEN) began after 1 June, which required extracting the summer period from successive years. This was accomplished by creating separate data tables for each summer period (for example, RAL: 7/30/2002–9/30/2002 and 6/1/2003–7/29/2003) and then combining these tables using the “merge” function in the GRAN program (appendix 1) to produce a single data table containing temperatures for the summer period.

Table 1. Descriptions of data files containing site information, water temperature statistics, and examples of file formats used or generated by the GRAN program.

[GRAN, granularity; ADAPS, Automated Data Processing System; 1h, 1 hour; POR, period of record; stats, statistics; RDB, relational database]

File	Description
02335910.xls	An example of temperature data for Rottenwood Creek, GA, obtained using ADAPS and modified into 6-column format in Excel®. This is the simplest of the data formats that can be processed by the GRAN program (contrast with 02335910_ADAPS_data.txt).
02335910_ADAPS_data.txt	An example of temperature data (Rottenwood Creek, GA) in ADAPS RDB format (tab-delimited ASCII). This is one of several data formats that can be processed by the GRAN program.
7-Day_Ave_Temp.xls	An example of 7-day running averages of daily mean temperatures for sites in the nine metropolitan areas. Running averages are derived from the estimated daily mean temperature values (Annual_Daily_Temp_estimated.xls).
Annual_1h_Daily_stats.xls	Annual temperature statistics (mean, maximum, minimum, range, standard deviation, number of observations, rates of change over 1-h, 2-h, 4-h, 8-h, and 12-h periods) generated by the GRAN program for sites in the nine metropolitan areas based on daily mean temperature measurements.
Annual_1h_stats.xls	Daily temperature statistics (mean, maximum, minimum, range, standard deviation, number of observations, rates of change over 1-h, 2-h, 4-h, 8-h, and 12-h periods) generated by the GRAN program for sites in the nine metropolitan areas based on hourly temperature measurements (1-h granularity).
Annual_Daily_Temp_estimated.xls	Daily mean temperatures over an annual period for sites in the nine metropolitan areas. Missing values have been estimated using linear regressions that relate temperatures between sites.
Annual_dDays.xls	Annual degree days calculated for sites in the nine metropolitan areas. Annual degree days are calculated from estimated daily temperature values (Annual_Daily_Temp_estimated.xls).
Contents_of_Data_Examples.mdb.xls	Table explaining the contents of the file "Data_Examples.mdb," which is described below.
Data_Examples.mdb	An Access® database containing examples of data tables used and produced by the GRAN program.
Raw_data.xls	An example of a raw (ADAPS) data file (0208500600.xls) that has been imported into the GRAN program.
Raw_data_POR.xls	An example of a raw data file after it has been imported into the GRAN program and processed by the POR (period-of-record) function (30 July 2002 – 29 July 2003).
Site_List.xls	List of sites at which temperature data were collected in the nine metropolitan areas.
Summer_1h_Daily_stats.xls	Daily summer (1 June – 30 September) temperature statistics (mean, maximum, minimum, range, standard deviation, number of observations, rates of change over 1-h, 2-h, 4-h, 8-h, and 12-h periods) generated by the GRAN program for sites in the nine metropolitan areas based on hourly temperature measurements (1-h granularity).
Summer_1h_POR_stats.xls	Temperature statistics (mean, maximum, minimum, range, standard deviation, number of observations, rates of change over 1-h, 2-h, 4-h, 8-h, and 12-h periods) for the summer period of record (POR) generated by the GRAN program for sites in the nine metropolitan areas based on daily mean temperature measurements.
Temp_data_est.xls	Summary of temperature data that were used to calculate statistics for daily, summer, and annual periods and to estimate missing daily mean temperatures for the determination of annual degree days (Annual_dDays.xls) and 7-day running average daily temperatures (7-day_Ave_Temp.xls).

Table 2. Water temperature data file in tab-delimited ASCII (RDB) format produced by the U.S. Geological Survey's Automated Data Processing System (ADAPS).

```
# //UNITED STATES GEOLOGICAL SURVEY       http://water.usgs.gov/
# //NATIONAL WATER INFORMATION SYSTEM     http://water.usgs.gov/data.html
# //DATA ARE PROVISIONAL AND SUBJECT TO CHANGE UNTIL PUBLISHED BY USGS
# //RETRIEVED: 2005-05-04 08:06:51
# //FILE TYPE="NWIS-I UNIT-VALUES" EDITABLE=NO
# //DATABASE NUMBER=1 DESCRIPTION=" Georgia District Database     (ADAPS, GWSI, QW)"
# //STATION AGENCY="USGS " NUMBER="02335910       " TIME_ZONE="EST" DST_FLAG=N
# //STATION NAME="ROTTENWOOD C (INTERSTATE N PKWY) NR SMYRNA, GA."
# //DD DDID="   5" RNDARY="1112333331" DVABORT=120
# //DD LABEL="Temperature, water transducer, IN degrees C"
# //PARAMETER CODE="00010" SNAME="Temperature, water"
# //PARAMETER LNAME="Temperature, water, degrees Celsius"
# //TYPE CODE=C NAME=COMPUTED
# //RANGE START="00000000000000" END="99999999999999" ZONE="LOC"
```

DATE	TIME	TZCD	VALUE	PRECISION	REMARK	FLAGS	QA
8D	6S	6S	16N	1S	1S	32S	1S
20020719	85000	EST	24.1	3			W
20020719	90000	EST	24.1	3			W
20020719	93000	EST	24.2	3			W
20020719	100000	EST	24.5	3			W
:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:
20031203	90000	EST	8.5	2		@	W
20031203	93000	EST	8.5	2		@	W
20031203	100000	EST	8.6	2		@	W
20031203	103000	EST	8.6	2		@	W

Table 3. Example of a water temperature data file (table 2) converted to the six-column format used for importing data into the GRAN program from Excel® or Access®.

[Time values also can be expressed in HHMM, HH:MM:SS, or HH:MM formats. STAID, station identification number; SABBREV, shortened or abbreviated station name; TZCD, Time zone code]

STAID	SABBREV	Date	TIME	TZCD	VALUE
02335910	Rottenwood Crk	7/19/2002	85000	EST	24.1
02335910	Rottenwood Crk	7/19/2002	90000	EST	24.1
02335910	Rottenwood Crk	7/19/2002	93000	EST	24.2
02335910	Rottenwood Crk	7/19/2002	100000	EST	24.5
:	:	:	:	:	:
:	:	:	:	:	:
02335910	Rottenwood Crk	12/3/2003	70000	EST	8.5
02335910	Rottenwood Crk	12/3/2003	73000	EST	8.5
02335910	Rottenwood Crk	12/3/2003	80000	EST	8.5
02335910	Rottenwood Crk	12/3/2003	83000	EST	8.5

Selecting a Common Data Granularity

Time intervals between temperature measurements (data granularity) varied from as little as 5 minutes to as much as 1 hour when interruptions in data collection, such as equipment loss or malfunction, are ignored. The "change granularity" function in the GRAN program was used to set data granularity to 1 hour for all sites to match the longest time interval between regular data-collection periods. The GRAN program assigns data to each hour by selecting the data-collection time that has the same hour value and the smallest deviation past the hour. For example, if data were available for 08:45 and 09:30, the data associated with 09:30 would be assigned to 09:00 even though 08:45 is closer to 09:00. The number of observations in which a deviation occurred between the desired time value and the actual time value was extremely small (less than 1 percent of observations) and occurred only at the initiation or termination of data collection.

Summarizing Water Temperature Data

Descriptive statistics (table 4) were used to summarize temperature data and rates of change (degree Celsius per hour, [°C/h]) over daily, annual, and summer (1 June–30 September) periods. Rates of temperature change were characterized as the absolute value of the difference between temperature measurements and were calculated for time intervals of 1 hour (h), 2 h, 4 h, 8 h, and 12 h. Variability in the rate of temperature change was characterized for each site by counting the number of rate measurements that fell within multiples of the standard deviation (SD) of the mean rate of change for the year (n_1SD: ≤ 1 SD; n_2SD: > 1 to ≤ 2 SD; n_3SD: > 2 to ≤ 3 SD; n_4SD: > 3 to ≤ 4 SD; n_5SD: > 4 to ≤ 5 SD; and n_gt5SD: > 5 SD). The number and distribution of missing values in the data varied by site (**Temp_data_est.xls** and **Annual_1h_Daily_stats.xls**); consequently the number of observations that underlie the summary statistics also varied among sites. Annual degree days (sum of daily mean temperatures for the year) and 7-day running average of daily mean temperatures over a consecutive 7-day period also were calculated after estimating missing daily mean temperatures using regressions with other nearby sites.

Summary Statistics for Annual and Summer Periods of Record

Summary statistics for annual (**Annual_1h_Daily_stats.xls** and **Annual_1h_stats.xls**) and summer (**Summer_1h_Daily_stats.xls** and **Summer_1h_POR_stats.xls**) periods were derived using the "POR statistics" function in GRAN. These summary statistics are derived after the GRAN program has converted the data to a common granularity (see **Data_Examples.mdb** for examples of this process and the generation of summary statistics). The data value that most closely approximates the desired time value (for example, data on the hour for 1-h granularity) is identified as are the data values and rates of change associated with time intervals that are multiples of the granularity (for example, 1 h, 2 h, 4 h, 8 h, and 12 h prior to the current time in the case of 1-h granularity). Because the summary statistics are closely linked to the data produced when setting the granularity, it is important to understand how the GRAN program sets data granularity. Table 5 contains 4 hours of data produced by setting the granularity to 1 h for a site where the temperature data were recorded at 30-minute intervals starting at 00:00 h (midnight). VALUE is the temperature associated with the observation time (oTime) that most closely approximates the TIME (hourly) value. Diff is the difference (in minutes) between the oTime and TIME values (30 minutes for the first value, zero for all subsequent values). CumMin is the elapsed time (minutes) since the start of the POR (0:00 in this example). The tTime (tTime1h, tTime2h, tTime4h) and tVal (tVal1h, tVal2h, tVal4h) data are time and temperature values corresponding to times 1 h, 2 h, and 4 h prior to the current time (TIME). These values are used to calculate the rate of change over time intervals of 1 h, 2 h, and 4 h in the example in table 5 as follows:

tTime is obtained by subtracting the time interval (minutes) from the current CumMin value, for example, tTime2h at 4:00 is 240–120 = 120.

tVal is the temperature (VALUE) at the value of CumMin that corresponds to the tTime value. For example, the value of tTime2h at 4:00 is 120, so tVal2h at 4:00 is the temperature (VALUE) at CumMin = 120, or 23.8 °C.

tRate values are calculated by subtracting tVal from VALUE and dividing by the difference between CumMin and tTime in hours for the respective TIME. For example, tRate2h at 4:00 is (23.5–23.8) / ((240–120)/60) = -0.15 °C/h.

Rates can only be calculated after sufficient time has elapsed such that there is a data value associated with the tTime value. For example, the tRate1h can be calculated after 1 hour of data collection (TIME 1:00) and tRate4h after 4 hours (TIME 4:00). The number of observations occurring within a day, or POR, is controlled by the granularity of the data and not by the time interval used to calculate the rate of change. The only exception to this is at the beginning of the data-collection period when the lack of antecedent data limits the number of rates that can be calculated. For example, if data collection started at 00:00 (midnight) the first rate based on a 12 time interval (tRate12h) could not be calculated until noon, and the number of tRate12h values that could be calculated on that first day would be 12 rather than 24. Statistics such as mean, maximum, minimum, standard deviations, range, and rates of change for the different time intervals (for example, 1 h, 2 h, 4 h) are calculated using the appropriate tVal data (tVal1h, tVal2h, tVal4h). Rate statistics are based on the absolute value

Table 4. Summary water temperature statistics calculated by the GRAN program for daily, annual, and summer periods of record.

[Statistics for rates of change are calculated for five time intervals where ?h corresponds to 1-hour (h), 2-h, 4-h, 8-h, and 12-h time intervals; Ave, average (mean) daily value; POR, period of record; Max, maximum value; Min, minimum value; SD, standard deviation; N, number of values; tRate, rate of change in temperature (°C/h); tRate?h_n_%SD, number of observations where the rate of change was greater than %–1 SD and less than or equal to % SD of the rates observed over the time period where ? h corresponds to 1-h, 2-h, 4-h, 8-h, and 12-h time intervals and % corresponds to 1, 2, 3, 4, 5, or > 5 intervals of SD; ≤, less than or equal to; >, greater than; gt, greater than; °C/h, degree Celsius per hour]

Statistics	Description
Ave	Average (arithmetic mean) temperature for the day (Daily) or over the POR.
Max	Maximum temperature for the day (Daily) or over the POR.
Min	Minimum temperature for the day (Daily) or over the POR.
Range	Range (MAX–MIN) in temperature for the day (Daily) or over the POR.
SD	Standard deviation of temperatures for the day (Daily) or over the POR.
N	Number of temperature measurements for the day (Daily) or for the POR.
tRate?h_Ave	Mean rate of change (°C/h) in temperature for the day (Daily) or POR based on absolute values of differences in measurements taken ? hours apart.
tRate?h_Max	Maximum rate of change (°C/h) in temperature for the day (Daily) or POR based on absolute values of differences in measurements taken ? hours apart.
tRate?h_Min	Minimum rate of change (°C/h) in temperature for the day (Daily) or POR based on absolute values of differences in measurements taken ? hours apart.
tRate?h_RANGE	Range (tRate?h_Max – tRate?hMin) in rates of temperature change (°C/h) for the day (Daily) or POR.
tRate?h_SD	Standard deviation in rates of temperature change (°C/h) for the day (Daily) or POR based on absolute values of measurements taken ? hours apart.
tRate?h_n_1SD	Number of temperature rate changes (absolute values) for the day (Daily) or POR that are ≤ 1 SD of the mean rate for the POR based on measurements taken ? h apart.
tRate?h_n_2SD	Number of temperature rate changes (absolute values) for the day (Daily) or POR that are > 1 SD and ≤ 2 SD of the mean rate for the POR based on measurements taken ? h apart.
tRate?h_n_3SD	Number of temperature rate changes (absolute values) for the day (Daily) or POR that are > 2 SD and ≤ 3 SD of the mean rate for the POR based on measurements taken ? h apart.
tRate?h_n_4SD	Number of temperature rate changes (absolute values) for the day (Daily) or POR that are > 3 SD and ≤ 4 SD of the mean rate for the POR based on measurements taken ? h apart.
tRate?h_n_5SD	Number of temperature rate changes (absolute values) for the day (Daily) or POR that are > 4 SD and ≤ 5 SD of the mean rate for the POR based on measurements taken ? h apart.
tRate?h_n_gt5SD	Number of temperature rate changes (absolute values) for the day (Daily) or POR that are > 5 SD of the mean rate for the POR based on measurements taken ? h apart.

Table 5. Time-interval data and water temperature values generated by the GRAN program in the process of setting data granularity.

[These values are used to derive summary statistics (for example, mean, maximum, minimum, and rates of change) for time intervals that are multiples of the selected granularity; oTime, time value for the observation closest to, but greater than or equal to, the hourly time value (TIME); Diff, difference (minutes) between the hourly value (TIME) and the oTime value; VALUE, temperature value at oTime; CumMin, cumulative minutes since the start of the time series data; tTime?h, the time value for ? h prior to the current TIME value where ? h can be 1, 2, 4, 8, or 12 hours; tVal?h, temperature value at tTime?h; tRate?h, the rate of change (°C/h) over the time interval ? h; negative time values indicate that the elapsed time since the start of data collection is less than the time interval; negative rates indicate that temperatures are falling over the time intervals]

	Hourly time value (TIME) as HH:MM				
	0:00	**1:00**	**2:00**	**3:00**	**4:00**
oTime	0:30	1:00	2:00	3:00	4:00
Diff	30	0	0	0	0
VALUE	24	23.9	23.8	23.7	23.5
CumMin	0	60	120	180	240
tTime1h	-60	0	60	120	180
tVal1h		24	23.9	23.8	23.7
tRate1h		-0.1	-0.1	-0.1	-0.2
tTime2h	-120	-60	0	60	120
tVal2h			24	23.9	23.8
tRate2h			-0.1	-0.1	-0.15
tTime4h	-240	-180	-120	-60	0
tVal4h					24
tRate4h					-0.125

of the rates; that is, they do not distinguish between rates associated with rising and falling temperatures. Variations in the rate of temperature change were assessed by counting the number of rate values that fall within multiples of the standard deviation of rates observed over the POR for each site (n_1SD: 1 ≤ SD; n_2SD: >1 SD to ≤ 2 SD; n_3SD: > 2 SD to ≤ 3 SD; n_4SD: >3 SD to ≤ 4 SD; n_5SD: > 4 SD to ≤ 5 SD; gt_5SD: > 5 SD).

Estimating Annual Degree Days

Annual degree days (**Annual_dDays.xls**) were calculated by summing daily mean temperatures for the annual POR (**Annual_1h_Daily_stats.xls**). Missing values were estimated using linear regressions that related water temperature between sites (**Temp_data_est.xls**). Though the GRAN program provides an option for estimating missing values using linear regressions, this feature was not used. The Microsoft® Excel program provided a more effective method of calculating and reviewing the large number of regressions (676–900) that had to be evaluated for each urban area. These regressions were based on EUSE sites except for BIR, BOS, DEN, and PORT where data from other USGS sites were needed to estimate missing values for some sites or time periods, such as during winter when data loggers in northern areas were removed to prevent ice damage. Prediction equations were selected by examining all regressions applicable to a site, generally, 26–33 regressions per site, and using the regression with the highest R^2 value and the best coverage of missing values. It was necessary to use predictor equations from multiple sites to fill in all missing data for some sites. Negative values in the original data or in the estimated data were set to 0.01 °C to ensure that degree days conformed to the progression of physiological time (Taylor, 1981); that is, it could slow to near zero but not run backward. Annual degree days were calculated as the sum of the daily mean temperatures over the annual POR. Annual degree days could not be calculated for all sites because of the limited data for the POR or poor correspondence among sites. Annual degree-day estimates are presented in **Annual_dDays.xls,** and the sites used to estimate missing values are summarized in **Temp_data_est.xls**.

Extracting 7-Day Running Average Daily Mean Temperatures

Seven-day running average daily mean temperatures (**7-Day_Ave_Temp.xls**) provide a time-series temperature record that minimizes short-term (daily) variability and facilitates comparisons among sites. These data were extracted from the daily mean temperature data generated during the process of estimating annual degree days and include estimates of missing values. The running average was calculated by averaging 7 consecutive days of daily mean temperatures. Because the 7-day running average is based on the annual POR, the first day of data occurs 7 days after the beginning of the POR. The results of estimating average daily mean temperatures and constructing the 7-day average temperature data are illustrated in figure 2.

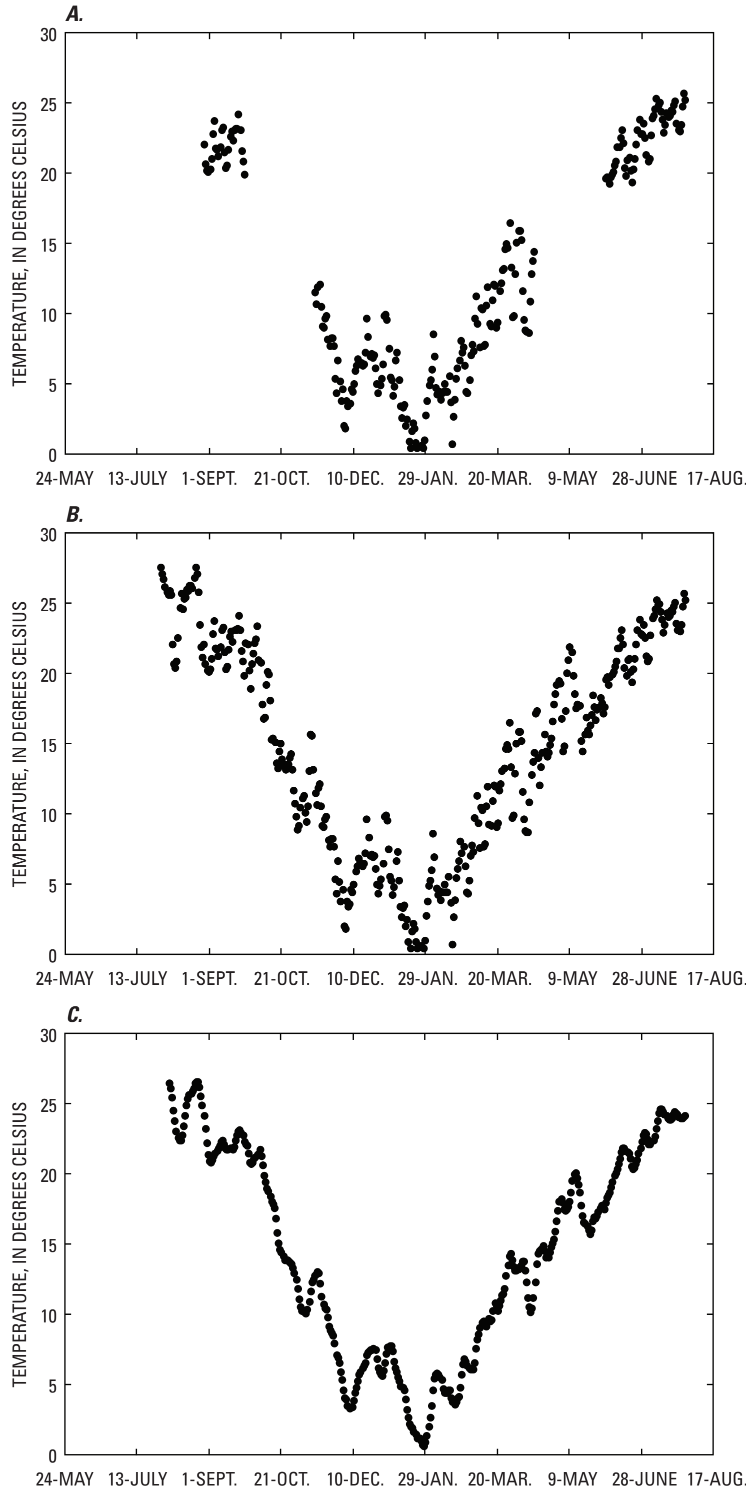

Figure 2. Annual daily mean water temperature data (degree Celsius) for Deep Creek (Raleigh-Winston Salem, NC, metropolitan area) showing *(A)* original data with missing values, *(B)* data with missing values estimated by linear regression, and *(C)* data expressed as 7-day running average temperatures.

Summary

This report describes methods used to process and summarize time-series water-temperature data collected from 256 sites associated with studies conducted in 9 metropolitan areas across the conterminous United States (Atlanta, GA; Boston, MA; Birmingham, AL; Dallas-Fort Worth TX; Denver, CO; Milwaukee-Green Bay, WI; Portland, OR; Raleigh-Winston Salem, NC, and Salt Lake City, UT) as part of the USGS NAWQA Program EUSE studies. Water-temperature data were collected over differing time periods using a variety of equipment. Measurement intervals (data granularity) differed among sites and over time. Equipment problems, floods, droughts, and winter ice resulted in the loss of data at some sites for variable periods of time. Comparison of temperature data within and among metropolitan areas required conversion of extremely large data sets (> 1,000,000 temperature measurements per study per year) to a uniform data granularity for comparable time periods (that is, annual and summer) before summary statistics could be calculated. Some temperature metrics such as annual degree days and 7-day moving averages required the estimation of missing temperature data.

A computer program (GRAN) was developed to automate and standardize the processing of temperature data and the calculation of summary statistics. Data were converted to common 1-h granularity, and summary statistics (mean, maximum, minimum, range, and standard deviation) were derived for daily, annual, and summer periods. Statistics on the rate at which temperature changed also were calculated and summarized. Rates of change were summarized on the basis of 1-h, 2-h, 4-h, 8-h, and 12-h intervals. Cumulative degree days were calculated for annual and summer periods. The annual period of record was derived separately for each metropolitan area by examining all possible annual periods and selecting the period that contained the most complete data set for all sites. The summer period of record (1 June–30 September) was extracted from the annual period and, in some cases, spanned consecutive years. Missing data were estimated by using linear regression to predict values based on temperature at other sites. Estimated values were used to calculate annual and summer degree days and 7-day moving average temperatures. Estimated values were not used in the derivation of summary statistics for daily, annual, and summer periods. Temperature data, summary statistics, and temperature metrics are provided as links to digital data files.

The GRAN program is described in the appendix. This program, which is available to USGS and non-USGS users (***ftp://ftpext.usgs.gov/pub/er/nc/raleigh/tfc/GRAN/***), can be used to process time-series data (for example, temperature, stage, or discharge data). This program provides a variety of tools for importing data, adjusting granularity, selecting periods of record, estimating missing values, and calculating summary statistics for daily records and user-defined periods of record.

References Cited

Armour, C.L., 1991, Guidance for evaluating and recommending temperature regimes to protect fish: Washington, DC, U.S. Fish and Wildlife Service Biological Report, Instream Flow Information Paper 28, v. 90, no. 22, 13 p.

Barthalow, J.M., 1989, Stream temperature investigations—Field and analytical methods: Washington, DC, U.S. Fish and Wildlife Service Biological Report, Instream Flow Information Paper 13, v. 89, no. 17, 139 p.

Brown, L.R., Gray, R.H., Hughes, R.M., and Meador, M.R., eds., 2005, Effects of urbanization on stream ecosystems: Bethesda, MD, American Fisheries Society Symposium 47, 423 p.

Coles, J.F., Cuffney, T.F., McMahon, Gerard, and Beaulieu, K.M., 2004, The effects of urbanization on the biological, physical, and chemical characteristics of coastal New England streams: U.S. Geological Survey Professional Paper 1695, 58 p.

Holmes, J.A., and Regier, H.A., 1990, Influence of temperature changes on aquatic ecosystems—An interpretation of empirical data: Transactions of the American Fisheries Society, v. 119, p. 374–398.

Klein, R., 1979, Urbanization and stream quality impairment: Water Resources Bulletin, v. 15. p. 948–963.

LeBlanc, R.T., Brown, R.D., and FitzGibbon, J.E., 1997, Modeling the effects of land use change on the water temperature in unregulated urban streams: Journal of Environmental Management, v. 49, p. 445–469.

Paul, R.W., Benfield, E.F., and Cairns, J.D., Jr., 1978, Effects of thermal discharge on leaf decomposition in a river ecosystem: Verhandlungen Internationale Vereinigung für Theoretische und Angewandte Limnologie, v. 20, p. 1759–1766.

Paul, M.J., and Meyer, J.L., 2001, Streams in the urban landscape: Annual Review of Ecology and Systematics, v. 32, p. 333–365.

Pluhowski, E.J., 1970, Urbanization and its effect on the temperature of streams on Long Island, New York: U.S. Geological Survey Professional Paper 627–D, 110 p.

Tate, C.M., Cuffney, T.F., McMahon, Gerard, Giddings, E.M.P., and Zappia, Humbert, 2005, Use of an urban intensity index to assess urban effects on streams in three contrasting environmental settings, *in* Brown, L.R., Gray, R.H., Hughes, R.M., and Meador, M.R., eds., Effects of urbanization on stream ecosystems: Bethesda, MD, American Fisheries Society, Symposium 47, p. 291–315.

Taylor, Fritz, 1981, Ecology and evolution of physiological time in insects: The American Naturalist, v. 117, p. 1–23.

Vannote, R.L., and Sweeney, B.W., 1980, Geographic analysis of thermal equilibria—A conceptual model for evaluating the effect of natural and modified thermal regimes on aquatic insect communities: The American Naturalist, v. 115, p. 667–695.

Appendix 1. Granularity Program (GRAN)

The GRAN program provides a flexible, efficient, and standardized approach for preparing time-series data, such as temperature, stage, or discharge, for analysis. The program was designed to process data collected at multiple sites with varying periods of record and measurement intervals (that is, varying data granularity). The program provides tools that summarize the amount of missing data and categorize the number of measurements made at different data granularities for a single site or group of sites. Based on this information, the program assists the user with the selection of appropriate periods of record (for example, an annual period that maximizes the amount of data) and data granularity. The program calculates basic statistics for daily and user-defined periods of record, such as annual or seasonal, and can be used to estimate missing data based on the correspondence (linear regression) of data among sites.

Capabilities

The GRAN program has the following capabilities:

1. Creates or opens an existing Access® database of time-series data
2. Imports time-series data that are in Excel® workbooks, Access® data tables, tab-delimited ASCII, or comma-delimited ASCII formats
3. Copies, deletes, merges, or renames Access® data tables
4. Summarizes the amount of missing time-series data for a site or group of sites
5. Summarizes the data granularity associated with a site or group of sites
6. Identifies the start and end dates for data collected at a site or group of sites
7. Extracts a period of record from data collected at a site or group of sites
    a. Allows manual selection of a common period of record for a site or group of sites
    b. Automatically selects an annual period of record that is common to all sites within a group and maximizes the amount of data in the annual period
    c. Automatically selects an annual period of record that is unique to each site within a group and maximizes the amount of data associated with each site
8. Estimates missing data by using relations among sites (linear regressions)
9. Calculates summary statistics for daily values and user-defined periods of record
10. Exports data as tab-delimited ASCII files

Sources of Data Used by GRAN

The GRAN program can import data in comma-delimited ASCII, tab-delimited ASCII, Excel®, and Access® formats (***ftp://ftpext.usgs.gov/pub/er/nc/raleigh/tfc/GRAN/Examples/***). Comma- and tab-delimited data can be imported in one of two formats: (1) RDB-format produced by the USGS ADAPS (table 1-1) or (2) a generalized file format (table 1-2) consisting of six columns of data. The data columns are as follows: STAID (USGS station identification number as a text string); SABBREV (shortened station name or abbreviation as text); DATE (date measurement was made, in M/D/YYYY format); TIME (time that the measurement was made in HHMM, HHMMSS, HH:MM or HH:MM:SS format); TZCD (time-zone code as a text string); and VALUE (temperature measurement in numeric format, integer or decimal). Data in Excel® or Access® files must be in the standard six-column format (STAID, SABBREV, DATE, TIME, TZCD, VALUE).

Installation

The GRAN program is a Visual Basic 6® program designed to run on a laptop or desktop microcomputer with a Microsoft Windows® operating system. The GRAN software can be obtained through an anonymous ftp site (***ftp://ftpext.usgs.gov/pub/er/nc/raleigh/tfc/GRAN/Installation/***) or by contacting the authors (*tcuffney@usgs.gov*, *rabright@usgs.gov*) for a compact disc (CD). The ftp site and installation CD contain the GRAN program (Gran.exe), program documentation, and examples of data files. Even though the GRAN software was written specifically to work with NAWQA Program EUSE temperature data, it can work with any time-series data that can be converted to the six-column format described in table 1-2.

The installation package consists of three files that can be obtained separately or combined into a 4.6-megabyte (MB) compressed file (Gran.Zip):

1. Gran.cab—a 4.6-MB file that contains the GRAN program, forms, and support files as packaged by the Visual Basic 6® software packaging and distribution tool
2. setup.exe—a 137-kilobyte (KB) program that installs the GRAN software on the host computer
3. SETUP.LST—a 5-KB file that contains setup information used by setup.exe

Table 1-1. Example of a temperature data file in tab-delimited (RDB) format produced by an automated data processing system (ADAPS), which can be processed by the GRAN program.

```
# //UNITED STATES GEOLOGICAL SURVEY       http://water.usgs.gov/
# //NATIONAL WATER INFORMATION SYSTEM     http://water.usgs.gov/data.html
# //DATA ARE PROVISIONAL AND SUBJECT TO CHANGE UNTIL PUBLISHED BY USGS
# //RETRIEVED: 2005-05-04 08:06:51
# //FILE TYPE="NWIS-I UNIT-VALUES" EDITABLE=NO
# //DATABASE NUMBER=1 DESCRIPTION=" Georgia District Database     (ADAPS, GWSI, QW)"
# //STATION AGENCY="USGS " NUMBER="02335910       " TIME_ZONE="EST" DST_FLAG=N
# //STATION NAME="ROTTENWOOD C (INTERSTATE N PKWY) NR SMYRNA, GA."
# //DD DDID="   5" RNDARY="1112333331" DVABORT=120
# //DD LABEL="Temperature, water transducer, IN degrees C"
# //PARAMETER CODE="00010" SNAME="Temperature, water"
# //PARAMETER LNAME="Temperature, water, degrees Celsius"
# //TYPE CODE=C NAME=COMPUTED
# //RANGE START="00000000000000" END="99999999999999" ZONE="LOC"
```

DATE	TIME	TZCD	VALUE	PRECISION	REMARK	FLAGS	QA
8D	6S	6S	16N	1S	1S	32S	1S
20020719	85000	EST	24.1	3			W
20020719	90000	EST	24.1	3			W
20020719	93000	EST	24.2	3			W
20020719	100000	EST	24.5	3			W
:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:
20031203	90000	EST	8.5	2		@	W
20031203	93000	EST	8.5	2		@	W
20031203	100000	EST	8.6	2		@	W
20031203	103000	EST	8.6	2		@	W

Table 1-2. Example of a six-column temperature data file that can be read by the GRAN program.

[The file shown here is a tab-delimited text file; however, the GRAN program accepts comma-delimited text files, Excel® files, or Access® databases. Time values can be expressed as HHMM, HH:MM:SS, or HH:MM; STAID, station identifier; SABBREV, short station name or abbreviation; TZCD: time zone code]

STAID	SABBREV	Date	TIME	TZCD	VALUE
02335910	Rottenwood Crk	7/19/2002	85000	EST	24.1
02335910	Rottenwood Crk	7/19/2002	90000	EST	24.1
02335910	Rottenwood Crk	7/19/2002	93000	EST	24.2
02335910	Rottenwood Crk	7/19/2002	100000	EST	24.5
:	:	:	:	:	:
:	:	:	:	:	:
02335910	Rottenwood Crk	12/3/2003	70000	EST	8.5
02335910	Rottenwood Crk	12/3/2003	73000	EST	8.5
02335910	Rottenwood Crk	12/3/2003	80000	EST	8.5
02335910	Rottenwood Crk	12/3/2003	83000	EST	8.5

These files should be copied to a temporary directory in preparation for installation. Running setup.exe will install the program and add GRAN to the startup menu. By default, GRAN and supporting files are installed in a folder called "Granularity" within the Program Files directory of the boot drive, although the user has the option of specifying a different directory during the installation process. The installation program also will create an EcoTools directory that can be used to hold data and output files. In the event that the installation program signals that it is attempting to replace a newer version of a support file with an older version, select the option to keep the newer version of the support file. Once the installation process is completed, the files in the temporary directory can be deleted. USGS users will need administrator rights to install GRAN.

Updates

The GRAN software is updated periodically based on requests from users for new features or the discovery of bugs in the software. Users do not have to go through the installation process each time the GRAN software is updated. The GRAN software can be updated simply by downloading the latest version of the program executable file (Gran.exe) from the GRAN ftp site and copying it over the previous copy of Gran.exe in the folder "Program Files/Granularity" or wherever the program was installed.

Help and Documentation

The primary sources for help when installing and using the GRAN software are the program documentation, the program developer, and other GRAN users. The GRAN program does not have a fully implemented help system. Help within the GRAN program is limited to explanatory text provided within the program and message boxes that appear when the program encounters an error in file content or requires a decision on the part of the user. The following sources provide information and documentation that may be useful to GRAN users.

Email the program developer at:
tcuffney@usgs.gov

Electronic copies of the GRAN program and documentation can be obtained from:
ftp://ftpext.usgs.gov/pub/er/nc/raleigh/tfc/GRAN/

Information about the National Water-Quality Assessment (NAWQA) Program can be obtained from:
http://water.usgs.gov/nawqa/

Using GRAN

Data processing in the GRAN program is controlled by a series of program menus (fig. 1-1, table 1-3). Processing begins by using the **File** menu to open an existing database (**Open database**) or to create a new database (**Create new database**). The databases created by GRAN are standard Access® databases that can be processed using GRAN or Access®. Three types of data are stored in the GRAN database: (1) raw data (table 1-4), (2) processed data (table 1-5), and (3) statistics (table 1-6). Raw data are data that are imported into the database using the **Import data** submenu. Processed data are raw data that have been processed to a common data granularity using the **Change** submenu in the **Granularity** menu. Statistics are generated from processed data by calculating daily or POR statistics using the **Stats** menu. The **File** menu also is used to close databases (**Close**). The **Exit** menu is used to close the GRAN program. The **Options** menu contains the **Confirm deletions** submenu that, when set to "Yes," requires the user to confirm the deletion of a data table. When set to "No" the program will delete data tables without asking for confirmation. The selection that is grayed-out is the selection that is currently active. The **About** menu provides the version and contact information for the GRAN program.

The GRAN program is designed to process data in which each data table contains data for a single site. The normal progression for data processing is to import the data, set the period of record, set the granularity, calculate statistics, merge POR statistics (if desired), and export statistics to ASCII text files for further processing. GRAN can process data for one site at a time (single data tables) or for a group of sites (batch processing of multiple data tables). GRAN is not designed to process data tables that contain data from more than one site per table. Processing such data tables can lead to unreliable results.

The **View** menu is critical to the operation of the GRAN program. It is used to display data in the ***Database contents*** window (❶, fig. 1-1) based on data type (**Raw**, **Processed**, **Statistics**, **Other**, and **All**). The ***Database contents*** window is used to select one or more data tables for processing. Multiple data tables can be selected using standard methods for selecting multiple items (ctrl + left click of mouse) or ranges of items (shift + left click of mouse). The type of data displayed in the ***Database contents*** window and the number of items selected determine the types of data-processing menus that are available (table 1-7). The status bar located along the bottom of the main window (❻, fig. 1-1) provides information on the database that is open, the data tables that are being processed or created, and information on the status of data processing.

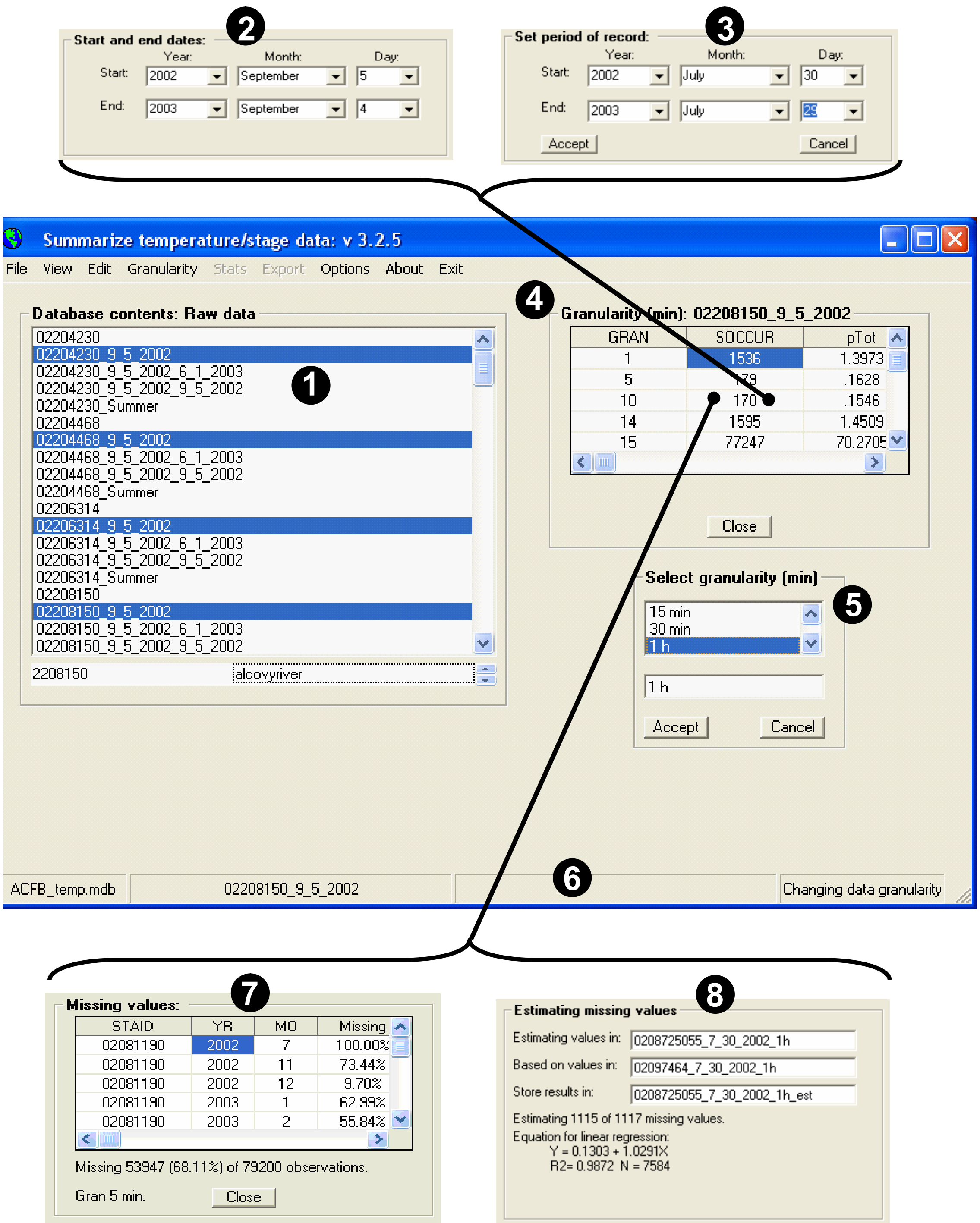

Figure 1-1. The main window of the GRAN program showing the the various information windows (1–8) that appear when data are processed.

[❶ window that lists database contents for a particular data type (raw, processed, statistics); ❷ window that shows the starting and ending dates for the selected data set; ❸ window that allows the user to select a start and end date for extracting data for a period of record (POR); ❹ window that displays the granularity (GRAN in minutes), and the sum of occurrences (SOCCUR) and percentage of total observations (pTot) at specific granularities; ❺ window that displays the available granularities and allows the user to select a common granularity for processing data; ❻ status bar that displays the file and tables that are being processed and the current operation and status; ❼ window that displays a summary of the percentages of missing values in the dataset; ❽ window that displays missing value estimations for one site based on a linear regression of known values for another site]

Table 1-3. List of menus and submenus contained in the GRAN program and the actions carried out by invoking the menu or submenu.

Menu and submenus	Action
File	
Open database	opens an existing Access® database
Create new database	creates a new Access® database
Import data	import data from ASCII, Access®, or Excel® files
Close	closes an open Access® database
View	
Raw data	displays names of raw data tables
Processed data	displays names of processed data tables
Statistics	displays names of tables of statistics
Other	displays names of tables containing data other than raw, processed, or statistical
All	displays all tables in the database
Edit	
Copy table	copies a data table to a new table
Delete table	deletes a data table
Merge tables	merges data tables of the same data type
Missing values	summarizes the amount of missing data
Period of record	
View start/end dates	displays the starting and ending dates
Extract POR – manual	manually extracts data for a period of record
Extract "best" annual POR for group – auto	automatically extracts the most complete annual POR for a group of sites; start and end dates are the same for all sites in the group
Extract "best" annual POR for each site – auto	automatically extracts the most complete annual POR for a group of sites; start and end dates are determined separately for each site
Rename	changes the name of an existing table
Granularity	
Descibe	describes the granularity in one or more data tables
Change	changes granularity in one or more data tables to create a common granularity
Stats	
Daily stats	calculates daily statistics
Estimate missing	estimates missing data values by linear regression with data from another site
POR stats	calculates statistics over the period of record
Export	exports data as ASCII text files
Options	
Confirm deletions	sets option requiring confirmation before deleting data tables
About	contact information for the program
Exit	exits the program

Table 1-4. Structure of the raw-data table created by the GRAN program when it imports data (table 1-2) using the File and Import data menus.

[STAID, USGS station identification number; SABBREV, shortened or abbreviated station name; DATE, date on which measurement was made; TIME, time (HHMMSS) at which measurement was made; TZCD, time zone code; VALUE, a measured value such as temperature or discharge; DMIN, elapsed time (minutes) in since midnight; CumMin, cumulative time (minutes) since midnight of the first day of data collection; GRAN, time interval between successive measurement]

STAID	SABBREV	DATE	TIME	TZCD	VALUE	DMIN	CumMin	GRAN
02335910	Rottenwood Crk	7/19/2002	85000	EST	24.1	530	530	10
02335910	Rottenwood Crk	7/19/2002	90000	EST	24.1	540	540	10
02335910	Rottenwood Crk	7/19/2002	93000	EST	24.2	570	570	30
02335910	Rottenwood Crk	7/19/2002	100000	EST	24.5	600	600	30
:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:
02335910	Rottenwood Crk	12/3/2003	70000	EST	8.5	420	721,440	30
02335910	Rottenwood Crk	12/3/2003	73000	EST	8.5	450	721,470	30
02335910	Rottenwood Crk	12/3/2003	80000	EST	8.5	480	721,500	30
02335910	Rottenwood Crk	12/3/2003	83000	EST	8.5	510	721,530	30

Table 1-5. Structure of the processed-data tables used in the GRAN program.

[Time values (?) for tTime?, tVal?, and tRate? can be 15 minutes (min), 30 min, 1 hour (h), 2 h, 4 h, 8 h, or 12 h depending on the granularity of the processed data]

Variable	Data type	Definitions
STAID	Text	USGS station identification number
SABBREV	Text	Shortened station name or abbreviation
DATE	Date	Date on which a measurement was made
TIME	Long integer	Time that corresponds to the selected granularity
TZCD	Text	Time zone code
oTime	Long integer	Time at which a measurement was made (observation time)
Diff	Long integer	Time difference, in minutes, between TIME and oTime
VALUE	Double	Value of time-series data (temperature, stage, discharge)
CumMin	Long integer	Cumulative elapsed time (minutes) since midnight of the start of the period of record
tTime?h	Long integer	CumMin value for a time ? hours prior to TIME
tVal?h	Double	The VALUE at oTime corresponding to tTime?h
tRate?h	Double	Rate of change (units/hour) between tVal? and VALUE

Table 1-6. Structure of the daily and period-of-record (POR) tables of summary statistics used in the GRAN program.

[Time values (?) for tTime?, tVal?, and tRate? can be 15 minutes (min), 30 min, 1 hour (h), 2 h, 4 h, 8 h, or 12 h depending upon the granularity of the processed data; AVE, daily mean value; MAX, maximum value; MIN, minimum value; SD, standard deviation; N, number of values; tRate, rate of change (measurement units per hour); N_1SD, number of observations where the rate of change was greater than or equal to 1SD of the rates observed over the time period (daily, annual, or summer); POR, period of record; ≤, less than or equal to; >, greater than]

Variable	Data type	Definition
STAID	Text	USGS station identification number
SABBREV	Text	Shortened or abbreviated station name
DATE	Date	Date on which measurement was made (daily statistics only)
dDaysPOR	Double	Degree-days in POR (POR statistics only)
nDaysPOR	Long integer	Number of days in POR for which there were data (POR statistics only)
dDaysAnnual	Double	Annual degree-days (POR statistics only)
Ave	Double	Mean value of measurements for the day (Daily) or over the POR
Max	Double	Maximum value of measurements for the day (Daily) or over the POR
Min	Double	Minimum value of measurements for the day (Daily) or over the POR
Range	Double	Range in values (MAX-MIN) for the day (Daily) or over the POR
SD	Double	SD of values for the day (Daily) or over the POR
N	Long integer	Number of measurements for the day (Daily) or for the POR
tRate?h_Ave	Double	Mean rate change (units/hour) for the day (Daily) or POR based on absolute values of differences in measurements taken ? hours apart
tRate?h_Max	Double	Maximum rate of change (units/hour) for the day (Daily) or POR based on absolute values of differences in measurements taken ? hours apart
tRate?h_Min	Double	Minimum rate of change (units/hour) for the day (Daily) or POR based on absolute values of differences in measurements taken ? hours apart
tRate?h_RANGE	Double	Range (tRate?h_Max – tRate?h_Min) in rates of change (units/hour) for the day (Daily) or POR
tRate?h_SD	Double	SD in rates of change (units/hour) for the day (Daily) or POR based on absolute values of differences in measurements taken ? hours apart
tRate?h_n_1SD	Long integer	Number of rate changes (absolute values) that are ≤ 1 SD of the mean rate for the POR based on measurements taken ? h apart
tRate?h_n_2SD	Long integer	Number of rate changes (absolute values) that are > 1 SD and ≤ 2 SD of the mean rate for the POR based on measurements taken ? h apart
tRate?h_n_3SD	Long integer	Number of rate changes (absolute values) that are > 2 SD and ≤ 3 SD of the mean rate for the POR based on measurements taken ? h apart
tRate?h_n_4SD	Long integer	Number of rate changes (absolute values) that are > 3 SD and ≤ 4 SD of the mean rate for the POR based on measurements taken ? h apart
tRate?h_n_5SD	Long integer	Number of rate changes (absolute values) that are > 4 SD and ≤ 5 SD of the mean rate for the POR based on measurements taken ? h apart
tRate?h_n_gt5SD	Long integer	Number of rate changes (absolute values) that are > 5 SD of the mean rate for the POR based on measurements taken ? h apart

Table 1-7. Features of the GRAN program that are available for different data types for single- and batch-mode processing.

[POR, period of record; auto, automatic]

Menu	Raw data		Processed data		Statistics	
Edit	Single	Batch	Single	Batch	Single	Batch
Copy	X		X		X	
Delete	X	X	X	X	X	X
Merge	X	X	X	X	X	X
Rename	X		X		X	
Missing values	X	X	X	X	X	
Period of record	X	X	X	X		
View start/end dates	X	X	X	X		
Extract POR – manual	X	X	X	X		
Extract "best" annual POR for group – auto		X		X		
Extract "best" annual POR for each site – auto		X		X		
Granularity						
Describe	X	X				
Change	X	X				
Stats						
Daily stats			X	X		
Estimate missing values			X			
POR stats			X	X		
Export	X		X		X	
Options (Confirm deletions)	X	X	X	X	X	X
About	X	X	X	X	X	X

Importing Data

Data can be imported as tab- or comma-delimited ASCII data in ADAPS RDB format (table 1-1) or as tab- or comma-delimited ASCII, Access®, or Excel® versions of the generalized format shown in table 1-2. The import function (**File/Import data**) creates a new file format (table 1-4) that contains information about the USGS station identification number, USGS shortened station name or abbreviation, collection date, collection time, time zone code, data value, elapsed time for each day as minutes past midnight, cumulative minutes since midnight of the first day of the time series, and granularity, which is the time interval in minutes between consecutive measurements. The GRAN value for the first observation is undefined because there are no prior measurements. For this reason, the GRAN value for the first observation is set to the granularity associated with the second observation in the data set. The program checks the structure of the data file that is being imported and alerts the user if any problems occur with the data file formats such as missing columns of data or multiple values associated with a combination of STAID, DATE, and TIME. In the latter case, the program allows the user the option of correcting the data outside of the GRAN program or having the program automatically fix the duplicates by selecting the duplicate record that occurs first in the data file.

Edit Menu Options

The **Edit** menu allows the user to copy, delete, merge, and rename data tables as well as summarize missing values and select a POR. Selecting one or more data tables in the ***Database contents*** window (❶, fig. 1-1) activates the **Edit** menu choices that are appropriate for the data type and number of data tables selected (table 1-7). The **Copy** and **Rename** submenus are available for all data types but are not active in the batch mode. The **Delete** and **Merge** submenus are available for all data types and for the single and batch modes. Selecting the **Merge** submenu calls up the ***Select tables to merge*** window (fig. 1-2), which provides a list of source tables and tables that are to be merged. Tables can be transferred between the source and merge windows by double clicking on the table name. The program checks that the data tables being merged are of the same type (raw, processed, daily statistics, or POR statistics) and checks to see if there are multiple values for a common reporting period (STAID+DATE+TIME for raw and processed data, STAID+DATE for Daily statistics, and STAID for POR statistics). If multiple data types are detected, GRAN alerts the user and returns to the main program window without merging the data. If it detects multiple values for a reporting period, it alerts the user and presents five options for handling the duplicate data (fig. 1-3). The **Options** menu controls whether or not the program prompts the user to confirm deletions.

Missing Values

The **Missing values** submenu examines one or more tables of raw data, processed data, or daily statistics and assesses the amount of missing data based on the minimum granularity in the data set. This information is displayed in the ***Missing values*** window that appears in the upper right-hand corner of the main window of the GRAN program (❼, fig. 1-1). The percentages of data that are missing are listed by STAID, year, and month. Time periods (year, month) that do not have missing data are not displayed. The total number and percentage of missing values also are listed along with the granularity used to determine the number of missing values. In batch mode, the total number and percentage of missing values are for all data combined.

Period of Record

The **Period of record** menu allows the user to identify the dates encompassed by one or more data tables and to select a range of dates over which to extract data from one or more data tables. The **Period of record** menu is only applicable to raw and processed data types (table 1-7). The **View start/end dates** submenu is used to identify the minimum and maximum dates for the data tables selected and displays these dates in the ***Start and end dates*** window that appears in the upper right-hand corner of the main program window (❷, fig. 1-1).

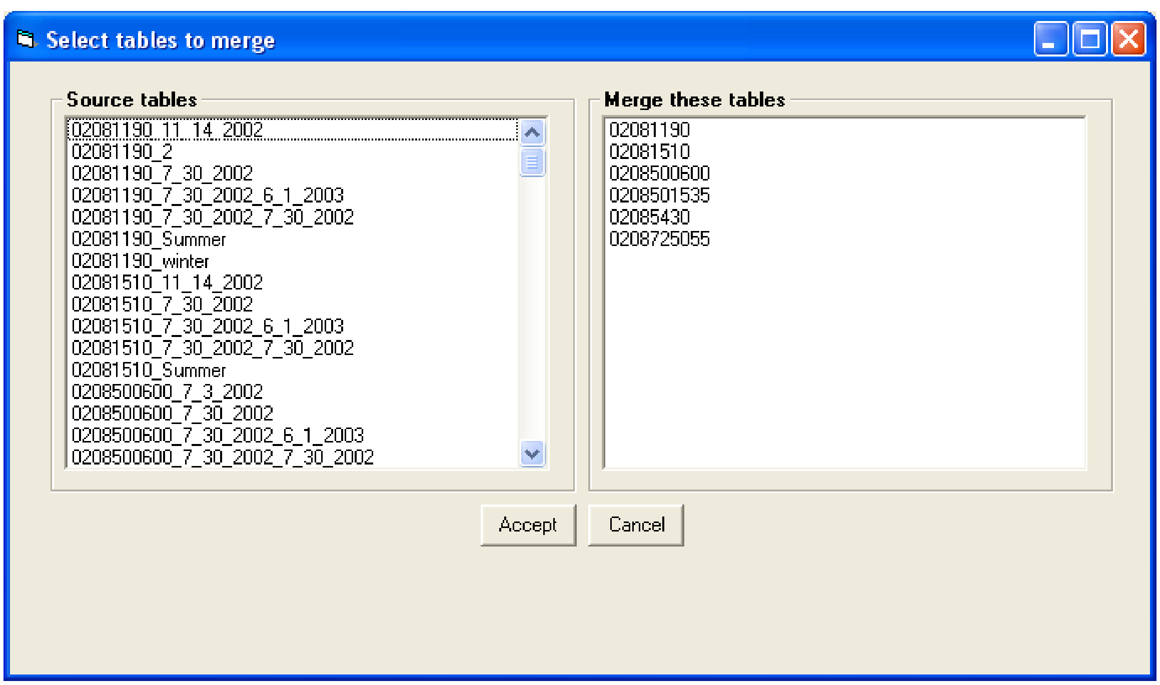

Figure 1-2. The Merge submenu of the GRAN program allows users to select data tables (source tables) that will be merged.

[Data tables are transferred between the two windows by double clicking on the table names]

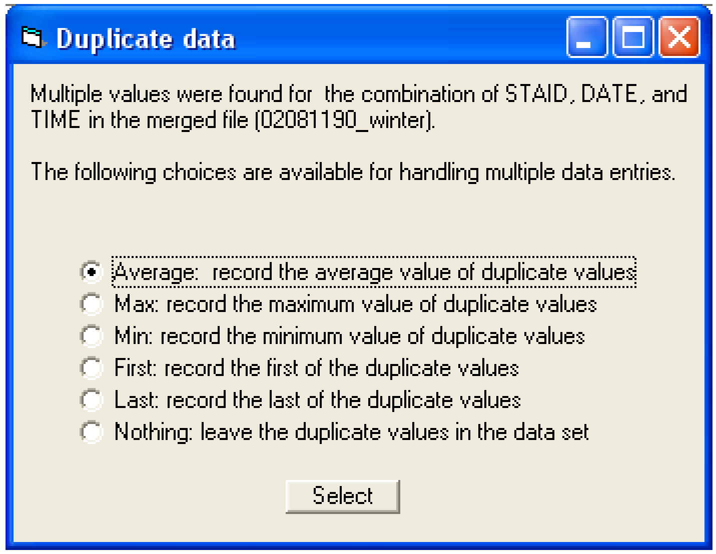

Figure 1-3. The duplicate data window of the GRAN program appears when the program detects that merging data tables will result in multiple data values for a common reporting period.

[The user can choose to resolve the problem outside of GRAN or to have GRAN resolve the problem using one of six methods for removing duplicate data]

Three submenus allow the user to extract data that conform to a period of record (that is, start and end dates) derived using one of three methods:

1. **Extract POR – manual**, which allows the user to manually select the start and end dates that define a POR. This method is applicable to both single and batch modes.

2. **Extract "best" annual POR for group – auto**, which automatically selects the annual POR that provides the most data for a group of sites (batch mode).

3. **Extract "best" annual POR for each site – auto**, which automatically selects an annual POR that provides the most data for each site (batch mode).

Methods 2 and 3 differ in that method 2 uses an annual POR that is common to all sites, but method 3 chooses an annual POR that is specific to each site and may or may not be common among sites.

Method 1 (**Extract POR – manual** submenu) examines all the data tables that have been selected, extracts the dates encompassed by the data, and displays these in the ***Set period of record window*** that appears in upper right-hand corner of the main program window (❸, fig. 1-1). The user selects start and end dates from the dates provided in the pull-down lists that include year, month, and day. Selecting the year calls up appropriate entries for month, and selecting the month calls up appropriate entries for days. The POR should be selected by first selecting the start date and then the end date. A period that falls outside of the selected start and end dates can be obtained by first selecting the closest time period using the pull-down lists and then going back and typing in the year, month, and day information. For example, the earliest time period contained in the data (❸, fig. 1-1) is 2002 July 30. This period could be altered to 2002 June 1 by typing "June" in the month-selection window and "1" in the day-selection window. The GRAN program automatically checks that the day and month selections are valid. Once the desired POR has been selected, clicking on the "Accept" button will extract the data for the selected POR to a new data table. If the selected start and end dates fall within the dates for which data are available, the program will extract this subset of data to a new data table. If the selected start or end dates fall outside of the period for which data are available, the program will add a data record for the start or end date using a TIME value of 0:00 h for the start date and 23:45 h for the end date. CumMin and DMIN values will be adjusted to reflect the new start and end dates and times. If the POR is being extracted from a single data table, the user will be prompted to provide a name for the data table that will contain the POR data. If the POR is being extracted for a batch of data tables, the program automatically creates new data table names by appending the start date (month, day, and year) to the data table name (for example, 02089110_6_1_2002). If the file name already exists, the user is given the option of replacing the existing data table with the new one. The "Cancel" button terminates the POR procedure and returns the user to the main program window.

Method 2 (**Extract "best" annual POR for group – auto** submenu) examines the POR encompassed by all sites that were selected. The earliest start date within the combined data is used to define all possible annual periods of record within the combined data set (for example, 6/1/2002–5/31/2003, 6/2/2002–6/1/2003, 6/3/2002–6/2/2003, etc.). The number of sites that have data for each date within the annual POR are summed together as are the number of measurements that fall within each annual POR. The program determines the "best" annual POR by selecting the POR that maximizes the number of sites with data. If more than one POR maximizes this value, the POR with the maximum number of measurements is selected. If multiple PORs that match these criteria remain, the POR with the earliest start date is selected. Once the start and end dates for the "best" annual period of record have been determined, these dates are used to extract data from each site. If the start or end dates for the annual POR fall outside of the POR encompassed by the data, the program adds data records for the start date (VALUE = null, TIME = 0:00 h) and(or) end date (VALUE = null, TIME = 23:45 h). The DMIN and CumMin data are adjusted to include the new start and end dates. This method does not allow the user to select time periods outside of the time period encompassed by the combined data. This restriction can be overcome by using method 1 (**Extract POR – manual** submenu) to manually set the POR for one or more data tables to the desired dates and then including this data table as part of the data submitted for analysis by method 2. Since method 2 operates in batch mode only, the user does not have the option of selecting a name for each new table. Instead, the program automatically assigns names to each new data table by appending the start date to the data table name (for example, 02089110_6/1/2002). The **Rename** submenu can be used to change the names that are automatically assigned.

Method 3 (**Extract "best" annual POR for each site – auto** submenu) differs from method 2 in that the extraction process considers each site separately. That is, the earliest start date for a site is used to define all possible annual periods of record for that site (for example, 6/1/2002–5/31/2003, 6/2/2002–6/1/2003, 6/3/2002–6/2/2003, etc.). The "best" annual POR is selected by considering the number of sites for which there are data, the number of measurements in the POR, and the POR with the earliest start date, as was done for method 2. The DMIN and CumMin values are updated to reflect the new start and end dates and times. Method 3 does not allow the user to select time periods outside of the time period encompassed by the combined data. Unlike method 2, this restriction cannot be overcome by using method 1 to manually set the POR for one or more data tables that are then submitted for analysis by method 3. Since method 3 operates in batch mode only, the program automatically assigns names to new data tables by appending the start date to the table name (for example, 02089110_6_1_2002). The user does not have the option of selecting a name for each new table; however, the **Rename** submenu can be used to change default data table names. Once an annual POR is selected for a site, the process is repeated for each of the remaining sites.

Granularity Menu Options

The **Granularity** menu can be used to display a summary of the data granularity (**Describe** submenu) or to change the data granularity (**Change** submenu) for one or more data tables. The **Describe** submenu examines the time interval (granularity) between consecutive time-series measurements for one or more data tables and displays the ***Granularity*** window, which summarizes data granularity (❹, fig. 1-1). The ***Granularity*** window lists the granularity in minutes (GRAN), the sum of occurrences (SOCCUR), and percentage (pTot) of measurement intervals associated with each granularity (GRAN). The "Close" button can be used to remove the granularity display.

The **Change** submenu is used to set the data granularity to a common value for one or more data tables. Selecting the **Change** submenu brings up the ***Granularity*** window (❹, fig. 1-1) and the ***Select granularity*** window (❺, fig. 1-1). The ***Granularity*** window provides information on the distribution of measurement intervals among different granularities that can be used to determine the best granularity to use for summarizing a data table or group of data tables. Very large granularity values typically are associated with gaps in the data that result from equipment problems. The most common data granularities are 5, 10, 15, 30, and 60 minutes. The ***Select granularity*** window allows the user to select granularity values (5 min, 10 min, 15 min, 30 min, 1 h, 2 h, 4 h, 8 h, or 12 h) that are greater than or equal to the minimum granularity value for the dataset. The **Change** menu converts raw data (table 1-4) to processed data (table 1-5), from which daily and POR statistics can be calculated.

Once the user has selected a common data granularity to apply to the selected data table(s), clicking on the "Accept" button of the ***Select granularity*** window will start the data conversion. The process begins by creating a data table for the processed data. This data table starts at 0 h on the start date and progresses to 24:00 h minus the granularity value (for example, 23:45 if granularity is 15 min) on the end date by adding successive rows of data where the time values (TIME and CumMin) are incremented by the data granularity value. Data values (VALUE) are assigned to the TIME value by searching for the data value with a TIME value that most closely approximates the desired time value without passing into the next day. That is, TIME values for 6/1/2002 will be matched by considering data with TIME values that are greater than or equal to 0 h on 6/1/2002 and less than 0 h on 6/2/2002. The actual time value for the data value assigned to TIME is recorded as oTime (observed time), and the difference between the observed and actual time (minutes) is stored in the Diff column. If no VALUE can be associated with TIME, the VALUE is set to null (missing value). The program assigns time-zone code values (TZCD) to missing data based on the

TZCD associated with each month. This is only an approximation of daylight saving time and does not take into account the actual day when daylight saving time is invoked. None of the calculations performed by the GRAN program makes adjustments for changes in TZCD.

The **Change** submenu not only sets data to a common granularity, it also calculates the rate of change between measurements whose CumMin values differ by multiples of the data granularity. For example, if data granularity is set to 15 minutes, rates of change would be calculated for intervals of 15 min, 30 min, 1 h, 2 h, 4 h, 8 h, and 12 h. That is, for a CumMin value of 1200 minutes, rates would be calculated by comparing the VALUE at 1200 minutes with VALUEs at 1185, 1170, 1140, 1080, 960, 720, and 480 minutes, respectively. The rate of change for each interval is calculated by subtracting the previous value (for example, 8.7 °C at CumMin = 960) from the current value (for example, 5.6 °C at CumMin = 1200) and dividing by the difference in time, expressed as hours [(1200–960)/60 = 4], to get the rate of change per hour [(5.6–8.7)/4 = -0.775 °C/h]. The CumMin, data values, and rates of change are stored in columns labeled tTime?, tVal?, and tRate?, where ? represents the time interval (for example, tTime15m, tVal15m, tRate15m for 15-minute granularity and tTime1h, tVal1h, and tRate1h for 1-hour granularity). The program will prompt the user to provide a name for the table of processed data if the granularity is being changed for a single data table. If granularity is being changed for a batch of data tables, the program will automatically assign names to the new data tables by appending the granularity designation to the file name (for example, 02089110_6_1_2002_30m and 02089110_6_1_2002_1h for data granularities of 30 minutes and 1 hour, respectively). The **Rename** submenu can be used to change the names that are automatically assigned.

Statistics Menu Options

The **Stats** menu is used to calculate daily (**Daily stats** submenu) and POR (**POR stats** submenu) summary statistics and to estimate missing values (**Estimate missing** submenu) based on linear regressions that use data from one site to estimate missing data at another site. The statistics functions operate only on processed data (table 1-7). The **Daily stats** and **POR stats** submenus provide similar statistical summaries (table 1-6), but differ in that the daily statistics function produces statistical summaries for each day (for example, daily mean temperature), whereas the POR statistics function produces a statistical summary for the entire POR (for example, annual mean daily temperature). The daily and POR stats provide mean, maximum, minimum, range, standard deviation, and number of observations for each day or for the entire POR. In addition, these functions provide statistical summaries (mean, maximum, minimum, range, standard deviations, and number of observations) for the rate changes over the time intervals produced by the change granularity function (**Granularity** and **Change** menus) and count of the number of rate measurements that fall within increments ($0 \leq SD < 1$; $1 \leq SD < 2$; $2 \leq SD < 3$; $3 \leq SD < 4$; $4 \leq SD < 5$; $SD \geq 5$) of the standard deviation (SD) over the POR. These values can be used to compare the variability in the rates at which temperature changes among sites. The POR statistics function also produces a summary of degree days (dDay) for the POR (dDaysPOR – degree days in POR and number of days with data in POR – nDaysPOR) and for the year (dDaysAnnual). Degree days in POR are the sum of the daily mean temperatures (Ave) for the POR. Annual degree days are estimated by dividing dDaysPOR by the proportion of the year represented by nDaysPOR (for example, 2,516 / (315/365) = 2,916, where dDaysPOR = 2,516 and nDaysPOR = 315).

The **Estimate missing** submenu uses linear regression to predict missing temperatures at one site based on temperatures measured at another site. The **Estimate missing** submenu is available only when working with data in the processed format (table 1-7). When this menu is invoked, the user is asked to select a data table with missing values, a data table that will be used to predict these missing values, and a data table in which to store the data after the missing values have been estimated. The names of these data tables are displayed in the ***Estimating missing values*** window that appears in the upper right-hand corner of the main program window (❽, fig. 1-1). Once the three data tables have been selected, the program determines how many of the missing values can be predicted and reports this information in the ***Proceed with estimations*** window (fig. 1-4). The user then can decide whether to continue with the estimation (Yes) or select a different data table (No) to estimate missing values. If the user elects to continue with the estimation procedure, the program then derives the linear regression and displays the regression equation and regression coefficient (R^2) for review in the ***Regression statistics*** window (fig. 1-5). The user then can decide whether to proceed with the estimation process (Yes) or terminate the procedure and select a new estimator table (No). Once the estimation procedure begins, the number of missing values, the number

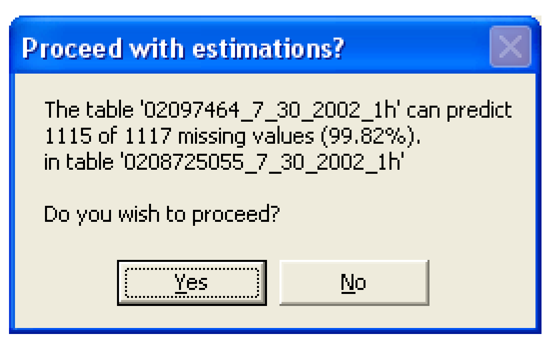

Figure 1-4. The ***Proceed with estimations*** window of the GRAN Program informs the user about the number of missing values that can be estimated based on values from another site.

[The user can elect to proceed (Yes) or return to the main program window (No)]

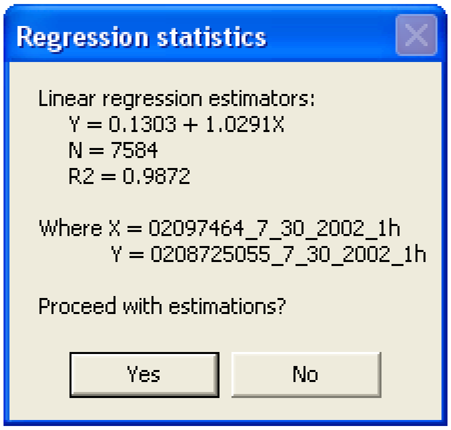

Figure 1-5. The ***Regression statistics*** window of the GRAN program displays the linear regression equation and regression coefficient (R^2) to be used in estimating missing values.

[The user can elect to proceed or not, in which case a different estimator site can be selected]

of values to be estimated, the regression equation, and the R^2 value are displayed in the ***Estimating missing values*** window along with the names of the data tables (❽, fig. 1-1). The data with the estimated values are stored in the new data table.

Export Menu

The **Export** menu saves raw, processed, or statistics (daily or POR) data as tab-delimited (*.txt) or comma-delimited (*.csv) ASCII text files that can be used with other statistical, graphics, or word-processing programs. Data tables are exported by selecting the desired table in the ***Database contents*** window (❶, fig. 1-1) and clicking on the **Export** menu. If only a single data table is being exported, the program will prompt the user to supply a file name for the exported data. If a batch of data files is being exported, the program will name the exported data by appending the appropriate suffix for the file type (.txt or .csv) to the data table name.

Cuffney and Brightbill—**Processing and Summarizing Time-Series Temperature Data Collected as Part of the NAWQA Studies on EUSE**—Data Series 330