Appendix 3. Importing Data to FoxDB from USGS and EPA Databases

The U.S. Geological Survey (USGS), the U.S. Environmental Protection Agency (USEPA), and the Illinois Environmental Protection Agency (IEPA) have long-term, routine water quality data collection programs. The USGS and the USEPA maintain standard databases to archive data. This appendix provides information on how data from these national databases were imported into the FoxDB. Not all fields in the national databases are listed in the following tables. Some fields in the Legacy STORET records did not contain any data for the Fox River watershed stations, some code values were not used, and some information was not relevant to the current study; thus, equivalent fields or codes in the FoxDB do not exist. This appendix is not meant to provide instructions for accessing the data, those are provided at the respective Web sites, but rather to show the relationship between data fields used in the national databases and tables and fields in the FoxDB database.

The USGS collects water quality data that may be accessed through the National Water Information System Web site (NWISWeb, http://waterdata.usgs.gov/nwis/nwis). Water quality sample information may be retrieved on the basis of user-selected search criteria, such as hydrologic unit or station number. Each data record includes, but is not limited to the USGS station number, the date, the parameter, and the result value. Listed in Table A3.1 are the fields and descriptions that document USGS data fields imported to the FoxDB, and how these fields appear in the FoxDB. The USGS field names and their descriptions are listed in the first two columns of Table A3.1, and the corresponding table and field name in the FoxDB are listed in the last two columns.

There are some special cases. Both the USGS database and the FoxDB have a field to enter information qualifying a result value. The USGS database field *remark_cd* has a lookup table with definitions of the codes used in this field. Somewhat different codes are used in the FoxDB. Table A3.2 lists the USGS code and the equivalent code in the FoxDB. Likewise, the field *medium_cd* has a corresponding lookup table with code definitions; however, all data retrieved for import to the FoxDB had a value of 9 in this field. The USGS NWIS defines a 9 in the *medium_cd* field as: “surface water, water on the surface of the earth stored or transported in rivers, streams, estuaries, lakes, ponds, swamps, glaciers or other aquatic areas. It also may refer to water in urban drains and storm-sewer systems.” In the FoxDB this information is documented by entering a value of W, for *Medium* in TBLSample (W is the code for water in the FoxDB).

The USEPA and the IEPA collect a variety of water quality data. Eventually these data will be regularly posted at the USEPA Website (http://www.epa.gov/storet/dbtop.html). Water quality sample information provided to the USEPA prior to 1999 may be retrieved from the STORET Legacy Data Center (Legacy). Data submitted to the USEPA from 1999 on will be posted in the “modernized” STORET (also called the new STORET database). However, at the time of this study, data collected in 1999 and later in the Fox River watershed were not available through the “new” STORET database and were acquired directly from the IEPA in various formats. The following discussion relates only to the electronic data retrieved from the Legacy Data Center.

Table A3.1. USGS NWIS Data Fields Imported to the FoxDB

USGS field name	*USGS field description*	*FoxDB table name*	*FoxDB field name*
agency_cd	Agency code	TBLOrganization	Organization_Code
site_no	Site identification number	TBLStation_Information	USGS_Station_Code
station_nm	Site name	TBLStation_Information	Place_Name_Description
dec_lat_va	Decimal latitude	TBLStation_Information	Latitude
dec_long_va	Decimal longitude	TBLStation_Information	Longitude
coord_acy_cd	Latitude-longitude accuracy	TBLStation_Information	Lat_Long_Accuracy_Code
huc_cd	Hydrologic unit code	TBLStation_Information	Hydrologic_Unit_Code
drain_area_va	Drainage area	TBLStation_Information	Total_Area
sample_dt	Date of sample	TBLSample	Start_Date (date and time)
sample_tm	Time of sample	TBLSample	Start_Date (date and time)
parameter_cd	Parameter Code	TBLResults	Parameter_Code
result_va	Value	TBLResults	Result_Value
remark_cd	Remark Code	TBLResults	Remark_Code
qa_cd	Quality Assurace Code	Used to assign QAPP grade	
qw_method_cd	Quality Assurance Method	Used to assign QAPP grade	
result_sg	Results significant figure	TBLParameter_Codes	Decimal_Point
medium_cd	Sample medium code	TBLSample	Medium

Table A3.2. USGS NWIS *remark_cd* and FoxDB *Remark_Code* Equivalents

USGS NWIS remark_cd	*Description*	*FoxDB Remark_Code*
<	Actual value is known to be less than the value shown.	K
>	Actual value is known to be greater than the value shown.	L
A	Average value	A
E	Estimated value	J
M	Presence of material verified but not quantified	M
N	Presumptive evidence of presence of material	N
U	Analyzed for, not detected	U

Through the Legacy Data Center option at the USEPA Web site given above, users may identify search criteria, such as hydrologic unit or station number to retrieve data on line. Each data record includes, but is not limited to, USEPA/IEPA station number, station descriptors, date, parameter, and result value. This appendix is not meant to provide instructions for accessing the Legacy data, but rather to show how data retrieved from this source were imported to the FoxDB.

Fields and descriptions that document the Legacy data fields that were imported to the FoxDB, and how they appear in the FoxDB are listed in Table A3.3. The Legacy field names are listed in the first column of Table A3.3, and the corresponding table and field name in the FoxDB are listed in the last two columns.

Information given in the Legacy Station Location Name 1, Station Location Name 2, and Station Location Name 3 fields were reviewed and used to populate the FoxDB fields in TBLStation_Information, Place_Name_Description and Water_Body_Name. Entries in the Legacy STORET Station Type Code field are a combination of codes. Information recorded for stations in the Fox River watershed and subsequently included in the FoxDB are listed in the first column of Table A3.4, and the corresponding entries in the FoxDB tables and fields are listed in the remaining columns.

Many fields were not populated for any record; thus, many fields that are part of the Legacy database are omitted from the lists in the following tables. Some fields had a variety of possible values, but only one value occurred in the retrieved data. For example, in the only value found (other than blank, no information) the Primary Activity Category was T, temporal composite. Table A3.5 lists other values that Legacy STORET defines for this field and how these fields appear in the FoxDB. The Secondary Activity Category is another example where only one value occurred: water. The information coded was sometimes ambiguous; for example, the Composite Method Code field was blank (no information) or had values of B: “Samples are not composited. Sample is a simple grab sample. STORET also used this code for noncomposite replicate samples.” If B was recorded for a sample, the results listed for the sample were reviewed to determine if, in fact, they were replicates, and then imported to the FoxDB accordingly. There are some duplications of information in the Legacy scheme, and data entry in some fields is not consistent.

Table A3.3. USEPA Legacy STORET Data Fields Imported to the FoxDB

Legacy STORET Field	*FoxDB Table*	*FoxDB Field*
Organization Code	TBLOrganization	Organization_Code
Organization Name	TBLOrganization	Organization_Name
Primary Station ID	TBLStation_Information	EPA_Station_Code
Secondary ID #1	TBLStation_Information	stationcode1
Secondary ID #2	TBLStation_Information	Stationcode2
Station Location Name 1	TBLStation_Information	Place_Name_Description and Water_Body_Name
Station Location Name 2	TBLStation_Information	Place_Name_Description and Water_Body_Name
Station Location Name 3	TBLStation_Information	Place_Name_Description and Water_Body_Name
Latitude	TBLStation_Information	Latitude
Longitude	TBLStation_Information	Longitude
Hydrologic Unit Code	TBLStation_Information	Hydrologic_Unit_Code
Legacy STORET Station Type Code (See Table A2B.4)	TBLStation_Information	Station_Type See Table A3.4
Sample Code	TBLSample	Sample_Code
Start Date	TBLSample	Start_Date (date and time)
End Date	TBLSample	End_Date (date and time)
Composite Method Code	TBLSample	Sample_Type
Sample Depth	TBLSample	Sample_Depth
Start Time	TBLSample	Start_Date (date and time)
End Time	TBLSample	End_Date (date and time)
	TBLSample	
Primary Activity Category	TBLProjects_Programs	See Table A3.5
Secondary Activity Category	TBLSample	Medium
Parameter Code	TBLResults	Parameter_Code
Parameter Long Name	TBLParameter_Codes	Full_Name
Result Value	TBLResults	Result_Value
Remark Code	TBLResults	Remark_Code
Composite Statistic Code	TBLSample	Composite_Statistic_Code

Table A3.4. Legacy STORET Station Type and FoxDB Translation

	FoxDB Translation			
	TBLStation_ Information			*TBLSample*
Legacy STORET Station Type	*Station_ Primary_ Type*	*Station_ Secondary_ Type*	*Ambnt*	*Medium*
TYPA/AMBNT/STREAM/BIO	River/Stream	N/A	Y (yes)	Biological
SEWER/TYPA/MUN/OUTFL/ NONAMB/PIPE	Facility	(1)	N (no)	(2)
AMBNT/STREAM	River/Stream	N/A	Y (yes)	(2)
SEWER/TYPA/MUN/OUTFL/ NONAMB	Facility	(1)	N (no)	(2)
TYPA/AMBNT/STREAM	River/Stream	N/A	Y (yes)	(2)
TYPA/AMBNT/LAKE	Lake	N/A	Y (yes)	(2)

Notes:
1. Use station description to determine Municipal Sewage or Municipal Water supply.
2. Value determined from parameters sampled.
N/A not applicable (no secondary type available).

Table A3.5. Interpretation of Legacy Primary Activity Category to FoxDB Tables and Fields

	Legacy STORET	*FoxDB*		
Code	*Primary_Activity_ Category*	*TBLSample Medium*	*TBLSample Sample_Type*	*TBLProject_Programs Project_Purpose*
C	Effluent Permit Condition			effluent monitoring
L	Biological Sample	Biological		
J	Tissue	Biological	Fish tissue	
S	Spatial Composite		Spatial composite	
T	Temporal Composite		Temporal Composite	
B	Both Spatial and Temporal Composite			
F	Flow Proportional		Flow proportional	
G	Grab Sample		Grab sample	
D	Replicate		Moved to Replicate table	