FINAL REPORT

Coastal Geodatabases
TWDB Contract Number 1000011103

April 4, 2011

Prepared for
Dharhas Pothina
Surface Water Resources Division
Texas Water Development Board

Prepared by
Jingqi Dong
Dr. Tim Whiteaker
Dr David Maidment

Center for Research in Water Resources
The University of Texas at Austin

FINAL REPORT

Coastal Geodatabases
TWDB Contract Number 1000011103

April 4, 2011

Prepared for
Dharhas Pothina
Surface Water Resources Division
Texas Water Development Board

Prepared by
Jingqi Dong
Dr. Tim Whiteaker
Dr David Maidment

Center for Research in Water Resources
The University of Texas at Austin

Introduction

This document is a final report for work performed by The University of Texas at Austin (CONTRACTOR) under the Coastal Geodatabases contract (TWDB contract number 1000011103). The CONTRACTOR was able to survey existing water data services for Texas to advise TWDB on data availability and suggest standard parameters and parameter names for a Texas Hydrologic Information System. The CONTRACTOR also investigated new approaches for water data sharing currently being researched by the Consortium of Universities for the Advancement of Hydrologic Science, Inc. (CUAHSI), the Open Geospatial Consortium (OGC), and the World Meteorological Organization (WMO).

Over the course of this contract, the CONTRACTOR submitted various documents and quarterly reports to TWDB'S CONTRACT MANAGER as deliverables were fulfilled. The following sections summarize products for all deliverables and indicate areas of future work based on the current status of the Texas Hydrologic Information System and new approaches for water data sharing which are still evolving.

This contract partially supported a Masters student named Jingqi Dong who helped complete work for the CONTRACTOR. Ms. Dong's thesis provides additional information about the work performed under this contract as it relates to her research. Once published online, her thesis can be accessed from the Center for Research in Water Resources (CRWR) online reports page (http://www.crwr.utexas.edu/online.shtml).

Deliverables

Five deliverables were defined in the contract. These deliverables are briefly described as:

1. A table containing the mutually agreed upon TxCentral Parameters with descriptions of each parameter and a TxCentral standard unit for each parameter.
2. A table listing all WaterOneFlow services that contain data within the geographical confines of Texas.
3. For each WaterOneFlow service that provides one or more TxCentral parameters, a table that maps the TxCentral Parameter to the variable code required to retrieve that parameter from the external web service and the units in which the data are provided by the web service.
4. TxCentral HydroDesktop plug-in along with documented source code and user's manual.
5. Updated websites as mentioned in the task description.

Deliverables one through three were completed. Deliverable four was altered in agreement with TWDB'S CONTRACT MANAGER due to new developments related to cyberinfrastructure for water data. Deliverable five could not be completed because new versions of WaterOneFlow services were not yet ready on TWDB's production server. The following sections provide more detail about the products related to each deliverable.

1. TxCentral Controlled Parameters

This deliverable consists of a list of parameters of interest developed by the CONTRACTOR in consultation with TWDB staff. This parameter list contains some parameters currently provided by TWDB, TPWD & TCOON. This list is called the TxCentral Controlled Parameter list and a standard unit is chosen for each parameter.

The CONTRATOR has completed this task and has proposed eleven parameters with standard units to TWDB. These parameters were chosen based on their presence in one or more Texas WaterOneFlow services, their general use in hydrologic applications, and in consultation with TWDB's CONTRACT MANAGER.

The parameter names are conformal with the CUAHSI hydrologic concept ontology, which provides a standard name for hydrologic parameters. Standard units are conformal to terms defined in the CUAHSI controlled vocabulary for units of measure. The parameters are shown in Table 1.

Table 1 TxCentral Controlled Parameter List

	Parameters	Units
Water Quality	Oxygen, dissolved	milligrams per liter
	Temperature, water	degree Celsius
	Salinity	parts per thousand
	Specific conductance	microsiemens per centimeter
	pH	dimensionless
	Turbidity	nephelometric turbidity units
Hydrology	Barometric pressure	inch of mercury
	Discharge, stream	cubic feet per second
	Evaporation, actual	inches per day
	Gage Height, stream	international foot
	Precipitation	international inch

2. TEXAS WATERONEFLOW SERVICES

This deliverable consists of a table listing all WaterOneFlow services that contain data within the boundaries of Texas. These include WaterOneFlow services hosted by the TWDB, CRWR and potentially other entities such as the USGS.

The CONTRACTOR completed this task by querying known service registries. WaterOneFlow services are registered with a catalog so that they can be discovered by users. CUAHSI maintains a national catalog called HIS Central, while CRWR maintains a separate catalog called Data.CRWR which predominantly lists Texas services. By querying HIS Central and Data.CRWR, the CONTRACTOR found all registered WaterOneFlow web services that include Texas water data (Table 2). Note that because aquatic biology parameters are not within the scope of this project, no aquatic biological data services are included in the table below.

Table 2 WaterOneFlow Services with Data in Texas

Service Name	Service URL
TCOON	http://his.crwr.utexas.edu/tcoonts/tcoon.asmx?WSDL
TCEQ TRACS	http://his.crwr.utexas.edu/TRACS/cuahsi_1_0.asmx?WSDL
TPWD Coastal	http://his.crwr.utexas.edu/tpwd/cuahsi_1_0.asmx?WSDL
TWDB Sondes	http://his.crwr.utexas.edu/TWDB_Sondes/cuahsi_1_0.asmx?WSDL

TWDB ADCP	http://his.crwr.utexas.edu/twdb_adcp/cuahsi_1_0.asmx?WSDL
TWDB Quality	http://his.crwr.utexas.edu/twdb_quality/cuahsi_1_0.asmx?WSDL
TWDB Tides	http://his.crwr.utexas.edu/twdb_tides/cuahsi_1_0.asmx?WSDL
TAMU CC WQ (HRI)	http://ccbay.tamucc.edu/CCBayODWS/cuahsi_1_0.asmx?WSDL
Texas Evaporation	http://his.crwr.utexas.edu/TXEvap/cuahsi_1_0.asmx?WSDL
USGS (Daily Values) (Groundwater) (Real-time Values) (Instantaneous Values)	http://river.sdsc.edu/wateroneflow/NWIS/DailyValues.asmx?WSDL http://river.sdsc.edu/wateroneflow/NWIS/Groundwater.asmx?WSDL http://river.sdsc.edu/wateroneflow/NWIS/UnitValues.asmx?WSDL http://river.sdsc.edu/wateroneflow/NWIS/Data.asmx?WSDL
EPA	http://river.sdsc.edu/wateroneflow/EPA/cuahsi_1_0.asmx?WSDL
NOAA's National Weather Service	http://watershed.uta.edu/nws_wgrfc_daily_mpe/cuahsi_1_1.asmx?WSDL http://watershed.uta.edu/nws_wgrfc_daily_mpe_recent_values/cuahsi_1_1.asmx?WSDL http://watershed.uta.edu/wfo_ewx_daily_mpe/cuahsi_1_1.asmx?WSDL http://watershed.uta.edu/wfo_fwd_hourly_mpe/cuahsi_1_1.asmx?WSDL http://watershed.uta.edu/usgs_region12_hourly_mpe/cuahsi_1_1.asmx?WSDL

3. Parameter Mapping

This deliverable consists of populating a table that maps parameters and unit codes used by a WaterOneFlow service to the standardized parameter and unit codes used in the TxCentral Controlled Parameter List. The mapping includes variables from Texas WaterOneFlow services that provide one or more parameters of interest. The parameter mapping table facilitates concept-based search from web services provided at TxCentral.

The CONTRACTOR has completed this task. The mapping table is shown in Appendix A.

4. TxCentral HydroDesktop Plug-in

This deliverable consists of a TxCentral plug-in for HydroDesktop, a free and open source client for WaterOneFlow services. The plug-in was to enable using TxCentral web service hosted by TWDB to search for water data in Texas using the TxCentral Controlled Parameter List.

Based on an agreement between the TWDB's CONTRACT MANAGER and the CONTRACTOR, the HydroDesktop plug-in was not implemented due to the software being superseded by new technology in the form of a web-accessible viewer from TNRIS. Resources were instead allocated toward researching an OGC standards-based service and software stack for the publication of hydrologic data. This approach benefits from wider interoperability due to following published OGC standards and also from extensive testing and refinement from OGC processes such as interoperability experiments (IEs).

A joint effort by the WMO and OGC in the form of a Hydrology Domain Working Group is currently developing the next generation of WaterML, called WaterML 2, where WaterML is the encoding used by WaterOneFlow to describe water data. Surface water and groundwater IEs are currently underway to test WaterML 2. Concurrently, a Water Information Services Concept Development Study (Water ISCD) has been undertaken to explore the overall architecture that will be required. This architecture will utilize a combination of standard OGC services such as Catalogue Service for the Web (CSW) and Web Feature Service (WFS) in lieu of the HIS Central and original

WaterOneFlow services developed by CUAHSI. In other words, the next generation of CUAHSI services and encodings will be compliant with OGC standards.

While the OGC has yet to endorse a particular implementation pattern, one pattern that has already emerged is to use CSW and WFS together to access a service catalog instead of the HIS Central service. The HIS Central service includes a method to return a list of time series matching search criteria. The search is carried out on a database at HIS Central created by harvesting metadata from all registered WaterOneFlow services. Keeping this database current has been problematic. With an alternate OGC approach, HIS Central only maintains overall descriptions of services rather than the more granular approach of maintaining descriptions of each time series within a service. A client queries a CSW at HIS Central to discover services that match the search criteria. The results include pointers to WFS services typically hosted by the organization providing the time series data such as the USGS. The WFS includes a layer where each feature in the layer describes a time series. The client then queries each WFS from the search results to identify individual time series that match the search criteria. Because the client is querying the WFS from the data provider directly, the client is much more likely to get up to date results. Once time series of interest have been identified, the client can request to download the time series.

This approach has been tested by the CONTRACTOR. Note that because the OGC mechanism for downloading time series data has yet to be formalized, WaterOneFlow version 1.0 services were used for this experiment. To simulate the data publishing pattern, the CONTRACTOR installed ESRI's Geoportal, a free and open source software package supporting a catalog of services with a CSW endpoint. The CONTRACTOR worked with CUAHSI to install two other Geoportals on CUAHSI servers. The CONTRACTOR registered some Texas services with the CRWR Geoportal while CUAHSI made some existing services registered with HIS Central available through its Geoportal at the San Diego Supercomputer Center (SDSC). In the second CUAHSI Geoportal (housed at the CUAHSI Program Office), the Geoportals at SDSC and CRWR were registered. What this demonstrates is the ability to support multiple catalogs of services, where each catalog is discoverable through a metacatalog hosted by CUAHSI.

To search these catalogs, the CONTRACTOR developed a lightweight desktop client called MetaCatalog Search which is capable of searching for and downloading time series data (Figure 1). The data access workflow using MetaCatalog Search consists of these steps:

1. Specify Search Criteria:
    a. Choose catalog(s) to search.
    b. Choose from a list of variables to search. For this research, the list was constrained to the TxCentral Controlled Parameter List.
    c. Use the interactive map to zoom to an area of interest.
2. Click a button to run the search. At this point, the application performs these steps without user interaction:
    a. Search each catalog for WFS services that match the search criteria.
    b. Search each resulting WFS to find time series that match the search criteria.
    c. Display locations of matching time series as dots in the interactive map, color coded by data provider.
3. Click on a dot in the map to download a time series. The application uses information from the WFS records to make the connection to the WaterOneFlow service and retrieve the data. The data are saved to a text file on the user's computer.

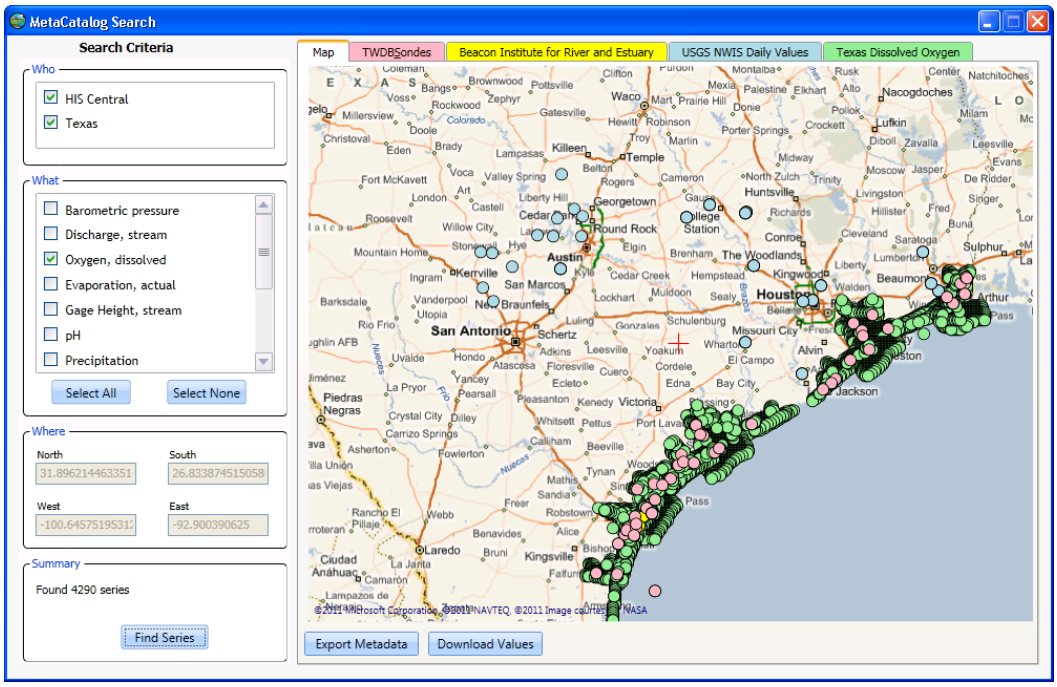

Figure 1 MetaCatalog Search Client for Data Search and Retrieval

Using MetaCatalog Search, the CONTRACTOR determined that the performance and capabilities of this proposed OGC approach are fairly comparable with the current approach used by CUAHSI for the limited number of services registered in the CRWR and CUAHSI Geoportals. Limitations include the handling of ontology concepts (which must now be incorporate into HydroServers instead of solely at HIS Central) and limitations on querying capabilities of Geoportal's CSW implementation. **Based on this research, the CONTRACTOR has informed the Water ISCD team accordingly.** Scalability tests and user interface refinement are left for future work once the OGC has endorsed a complete implementation pattern.

The CONTRACTOR can provide the source code for MetaCatalog Search to TWDB if requested. However, because this application was merely used to test a potential OGC approach for data discovery and access, the CONTRACTOR recommends that the TWDB wait until the OGC approach is more concrete, and then take advantage of any software libraries developed by OGC or CUAHSI at that time.

5. Website Updates

The TWDB is currently developing their own WaterOneFlow services to augment and replace the prototypes hosted by CRWR. The fifth deliverable for this contract consists of updating known registries to swap out the prototype services in favor of the production TWDB services. This includes services for TCEQ, TPWD, TCOON, and TWDB datasets which are registered at the Data.CRWR or HIS Central websites.

This deliverable was provisional on the TWDB finishing their production services in time for the CONTRACTOR to make the website updates. Because production versions of these services are not yet available from TWDB, this deliverable has not been completed.

Units Conversion Factors

At the request of TWDB'S CONTRACT MANAGER, the CONTRACTOR determined conversion factors required to transform from a given WaterOneFlow service's output to the standard unit defined by TxCentral. Note that conversions cannot be performed when units are not provided by the WaterOneFlow services, as is the case with some results from the EPA service. A diagram illustrating the conversion factors is attached in Appendix B.

Note:
Usually only one offset factor is necessary in a unit conversion process; however, some conversions need two factors such as the conversion from Fahrenheit to Celsius, which require two factors. Additionally, NTU and JTU of turbidity are measured using different approaches and instruments; therefore, conversion between these units is not recommended.

Recommendations for Future Work

In the short term, when TWDB finalizes their production WaterOneFlow services to replace the prototypes developed by CRWR, TWDB should register their services with HIS Central. CRWR can provide assistance to TWDB with this step if necessary. Concurrent with this step, CRWR will unregister the prototypes from HIS Central and update its Data.CRWR page accordingly.

In the long term, TWDB should prepare to implement an OGC approach for sharing water data once the approach is established. The approach is expected to be more widely used than the current implementation of CUAHSI data services due to the new approach's conformance to international standards. As for timing, the OGC has just initiated a WaterML 2.0 Standards Working Group whose mission is to define and publish WaterML 2.0 as an official OGC encoding standard. This process typically takes up to a year, after which the service and software stack surrounding the WaterML 2.0 encoding will also need to be established.

When the time is right to make the transition, there are some free and open source options that work in both Windows and Linux environments. For example, ESRI's Geoportal software includes a CSW endpoint. OWSLib is a library of Python modules which facilitate interaction with OGC web services. GeoServer is geospatial server software that includes a WFS endpoint. These programs may not have WaterML 2.0 support when WaterML 2.0 is officially released; therefore, TWDB could even contribute to the development of these open source products to serve both its own interests and those of the community.

Appendix A: Parameter Mapping Table

Parameter Mapping Table

Parameters & Units / Web Services	Oxygen, dissolved	Temperature, water	Salinity	Specific conductance	pH	Turbidity	Barometric pressure	Discharge, stream	Evaporation, actual	Gage Height, stream	Precipitation
	milligrams per liter	degree celsius	parts per thousand	microsiemens per centimeter	dimensionless	NTU	inch of mercury	cubic feet per second	inches per day	international foot	international inch
TCOON	do milligrams per liter	wtp degree celcius								pwl meter	
TCEQ TRACS	89855~89858 mg/L 20388~20391, 00301,00303 %Saturation 00300 mg/L 00389 Not Specified	00209~00211, 00221, 0010 degree centigrade 0011 degree Fahrenheit	00480, 00217~00220 PPT (parts per thousand)	00094~00095, 00212~00214 CM (millisiemens per centimeter) 00222 # of measurements	00215,00216,0022 3,00400,00403,82 583,89505,70310 standard units	82078, 82079 nephelometric turbidity units		00060~00061 cfs 00056 gallons/day 30209, 50051 instantaneous flow 74069 89835-89836 89848,89863 20480 cubic m/day			82554 inch
TPWD Coastal	OXY001 milligrams per liter	TEM001 degree celcius	SAL001 parts per thousand			TUR001 nephelometric turbidity units					
TWDB Sondes	DIO001 milligrams per liter	TEM001 degree celcius	SAL001 parts per thousand	CON001 millisiemens per centimeter	PEH001 dimensionless					GAH001 international foot	
TWDB ADCP								Flow cubic feet per second			
TWDB Quality	DO milligrams per liter	Temp degree celsius	Sal parts per thousand	Cond millisiemens per centimeter	pH dimensionless					Depth international foot	
TWDB Tides										GaH001 international foot	
TAMU CC WQ(HRI)	DOConcCon; DOBottomCon; DOConGrab; DOBottomGrab; milligrams per liter DOSatGrab; PctDOGrab; percent	TempCon; TempGrab; degree celcius	SalinityCon; SalinityGrab; practical salinity units	ConductGrab microsiemens per centimeter	pHCon pH Unit					DepthCon; DepthGrab; meter	
TWDB									Evap001 inches per day		
EPA	201-7,201-1,201- 2,201-6 Not Specified	481-1,2,7,5,6,4 Not Specified	434-1,5,4,6,7 Not Specified	139-1,5,2,4,7,6 Not Specified	1-2,4,7,1,6 Not Specified	514-1,2,7,6 Not Specified	42-5,1 Not Specified				
USGS	00300 mg/L 00301 %Saturation	00010 degree celsius 00011 degree Fahrenheit	00096 grams per liter 00480 ppt	00095 microsiemens per centimeter	00400 dimensionless	JacksonTU 61028 NTU 63680 FNU 63684 FNMU	62602,62603; 00025 inch of mercury; millimeter of mercury	00060 cubic feet per second 00061 instantaneous cfs	00050 inches/day	00065 international foot	00045 international inch
NOAA's National Weather Service											MPE international inch

NOTE: Within the table, the first line is the Parameter Code in the Remote Service; The second line is the unit used in the Service.

Appendix B: Unit Conversion Diagram