

Final Technical Report for NGGDPP 2010-2011

USGS Award Number G10AP00125

**Development of a Document Management System  
For Drill Logs by the  
South Carolina Geological Survey**

Prepared  
11 January 2012

by  
C. Scott Howard  
Erin E. Koch  
South Carolina Department of Natural Resources  
Geological Survey  
5 Geology Rd  
Columbia, SC 29212  
(803) 896-7712 – office  
(803) 896-7695 – fax  
howards@dnr.sc.gov – email

Project Term Period: 1 August 2010 to 31 October 2011

**Development of a Document Management System  
For Drill Logs by the  
South Carolina Geological Survey**

**Abstract**

The FY10 NGGDPP project consisted of one major action: scanning existing paper drill-logs into a Document Management System. Tasks included inventorying existing drill logs, organizing logs by county, assigning proper site IDs to logs, verifying drill-hole locations, dealing with oversized sheets, and developing a database system to catalog scanned logs. At the end of the project, 3,341 drill logs, as well as 217 miscellaneous maps and reports accompanying logs, were scanned. Of the 46 counties in the state, logs from 42 counties were located. To date, 715 logs have made it through final QA/QC, which includes verification of drilling location.

Final Technical Report for NGGDPP 2010-2011  
Development of a Document Management System  
For Drill Logs by the  
South Carolina Geological Survey

Award Number G10AP00125

C. Scott Howard  
Erin E. Koch  
South Carolina Department of Natural Resources  
Geological Survey

## **Introduction**

This report documents the major work and results of the South Carolina Geological Survey's (SCGS) participation in the FY 10 National Geologic and Geophysical Data Preservation (NGGDP) program. The project consisted of two tasks: 1) collecting and scanning all the known records of power augered, hand augered, and other non-core and non-well related drilling records, and 2) inputting the digital files of the drill-logs into a Document Management System (DMS) to create a digital index that would interface with a developing Oracle database of water wells.

## **Task 1**

This task consisted of scanning drill-log records and producing pdf files. The county in which the hole was drilled was chosen as the major identifier of a drill log. Initial work began with organizing the records by the county. To establish a unique ID for each record, a second number was added to the county identifier (numeric). The second number is a 4-place sequential

number. The earliest record was identified as 0001, and the most recently drilled record sequentially evolved from the series. This process was done for each county in which drill records were located. Subsidiary issues associated with the task included: drill-site verification and correcting inconsistencies in original drill-log IDs.

### **Scanning Procedures**

1. It was decided to use the county as the primary dividing unit. Historic drill records were filed this way so it seemed an obvious choice, and it was a method used by other sections in the South Carolina Department of Natural Resources (SCDNR) developing statewide data sets.

2. Counties that had the fewest number of records and did not have any recent drilling were scanned first. This approach presented the best opportunity to complete a county with little problem of finding missing or recently added records.

3. Given the S.C. Geological Survey's early work in site-specific projects in the 1960's, 70's, and 80's, it was decided to scan site maps and reports that accompanied drill logs because those products also contained useful information. Commonly site maps and well-written location descriptions were used to locate drill holes prior to the introduction of GPS.

4. One person was dedicated to produce all of the initial record scanning and data entry into DMS.

5. Initial QA/QC procedures were carried out by several trained individuals and involved the following steps:

a) Check scanned image for problems: too short (re-scan), cut-off sides (re-scan), color when not necessary (re-scan in black and white), extra pages captured (back sides scan sometimes, delete those),

b) Check accuracy of the data entry: collar elevation, depths, dates,

c) Locate the hole using GIS layers/tools, historic maps, and location descriptions. The point of this exercise was to locate the hole as accurately as possible using all the information available,

d) Select a location method or accuracy level, e.g. 24K topographic map, unknown, county map,

e) Number the log: The log ID consisted of a county code plus a four digit sequential number. The first (oldest) drill record in a county begins with (0001). Most importantly, the unique number for each hole needed to be verified. Verification was needed because in counties

where hundreds of holes have been drilled over many decades, there was a good chance for number duplication as a result of multiple workers,

f) Route the document into the Geology Complete Workflow for final QA/QC.

6. The final QA/QC process was carried out by the senior GIS project leader. Problems found with the location of a record were routed back to the Geology QC Workflow for corrective action. The records were QA/QC'ed by a combination of exporting the grid from ImageNow into CSV files that could be easily sorted and manipulated in Excel and visual checked in the program itself. A final geographic check took place when the records were exported as a shapefile through an Oracle/GIS interface.

## **Task 2**

The SCDNR has chosen ImageNow as its DMS. The incorporation of scanned logs into DMS was designed by the IT Section's DMS administrator. Index keys for filing and searching were chosen by SCGS and include county, quadrangle, date drilled, depth drilled, and field ID. Although not as robust as a full-fledged database, this system is capable of integrating information with an Oracle database. Presently, efforts are underway to develop an Oracle database of water wells throughout the state. The final component will be to develop links from the DMS to the database enabling retrieval of logs from database searches.

## **Conclusions**

Three major problems arose during this project. The location of drill sites for much of the data had to be recreated manually. In the absence of GPS data or accurate measured locations, drill sites had to be manually transferred from original county drill maps to UTM coordinates. Transferring information was a time-consuming process; but eventually, it will get done. The upside of this problem is that once located, site information is immediately transferable to the Oracle water-well database.

The second problem encountered involved assigning unique IDs to all the records. In heavily drilled counties, for example, there could be gaps or overlaps of already assigned IDs. This problem was created by multiple workers and sloppy record keeping. The data preservation project allowed us to address these problems in one single effort. Now there is an established routine for accepting new information into our data collection that will prevent those errors from re-occurring.

The third problem was associated with the second problem. In the larger counties, if there was an overlap of older drilling studies followed by more recent studies in the same vicinity e, the likelihood for ID problems was much higher. Again, this project allowed us to correct those problems.

At the completion of the project, 3,341 drill logs, as well as 217 miscellaneous maps and reports accompanying logs, were scanned (Table 1). Of the 46 counties in the state, logs from 42 counties were located. To date, 715 logs have made it through final QA/QC, which includes verification of drilling location, and 761 records have passed the first level of QA/QC. Because this project has been integrated with an ongoing project to develop an Oracle database, we are continuing our initial data preservation efforts until the DMS and the Oracle water-well database are up-to-date. New drill-log information then will be entered as part of a project.

**Table 1 Scanning records and status by county.**

<b>County</b>	<b>Logs Scanned</b>	<b>Misc. Reports/Maps</b>	<b>1st QA/QC Complete</b>	<b>Final QA/QC</b>	<b>Total Records in County</b>	<b>Complete Y/N</b>
Abbeville	6	2	0	8	8	Y
Aiken	164	1	0	0	165	N
Allendale	115	0	0	0	115	N
Anderson	57	23	0	80	80	Y
Bamberg	0	0	0	0	0	N
Barnwell	1	1	0	0	2	N
Beaufort	164	0	0	0	164	N
Berkeley	269	2	0	0	271	N
Calhoun	186	1	89	0	187	N
Charleston	41	0	0	0	41	N
Cherokee	10	6	0	16	16	Y
Chester	13	8	0	21	21	Y
Chesterfield	101	5	0	106	106	Y
Clarendon	107	0	0	0	107	N
Colleton	16	0	0	0	16	N
Darlington	124	14	0	138	138	Y
Dillon	50	5	0	55	55	Y
Dorchester	6	0	0	0	6	N
Edgefield	80	0	0	80	80	Y
Fairfield	10	2	0	0	12	N
Florence	117	13	21	0	130	N
Georgetown	151	0	0	0	151	N
Greenville	30	12	0	42	42	Y
Greenwood	9	6	0	15	15	Y
Hampton	21	0	0	0	21	N
Horry	322	8	329	0	330	N
Jasper	186	0	0	0	186	N
Kershaw	0	0	0	0	0	N
Lancaster	28	14	0	0	42	N
Laurens	21	8	0	29	29	Y
Lee	31	6	0	0	37	N
Lexington	102	0	0	0	102	N
Marion	54	11	65	0	65	N
Marlboro	61	10	60	0	71	N

<b>County</b>	<b>Logs Scanned</b>	<b>Misc. Reports/Maps</b>	<b>1st QA/QC Complete</b>	<b>Final QA/QC</b>	<b>Total Records in County</b>	<b>Complete Y/N</b>
McCormick	17	2	0	19	19	Y
Newberry	0	0	0	0	0	N
Oconee	12	4	0	16	16	Y
Orangeburg	211	1	0	0	212	N
Pickens	10	6	0	16	16	Y
Richland	274	19	187	0	293	N
Saluda	10	0	10	0	10	N
Spartanburg	0	0	0	0	0	N
Sumter	35	0	0	0	35	N
Union	27	14	0	41	41	Y
Williamsburg	72	0	0	0	72	N
York	20	13	0	33	33	Y
<b>Totals</b>	<b>3341</b>	<b>217</b>	<b>761</b>	<b>715</b>	<b>3558</b>	