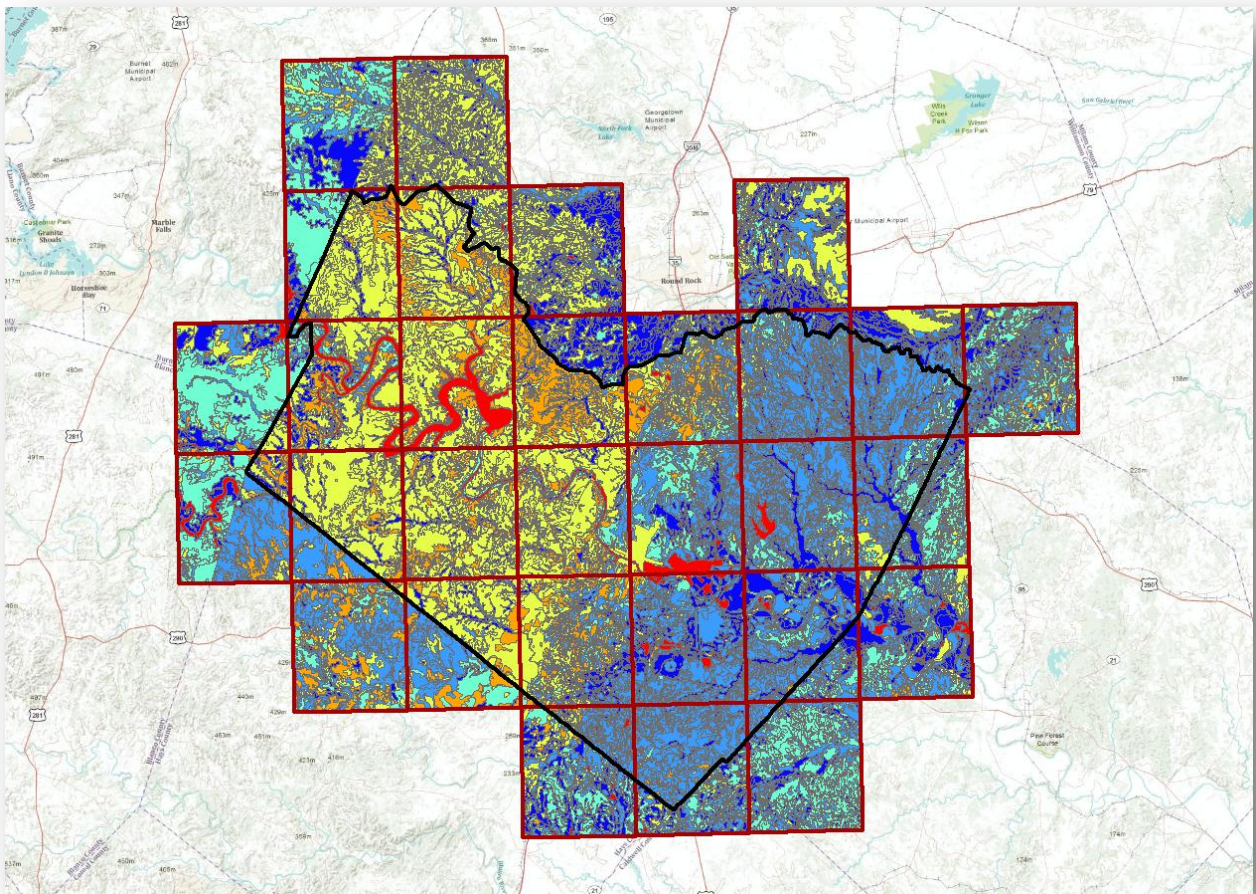


# SOIL MOISTURE MAPPING OF DROUGHT IN TRAVIS COUNTY, TX



Prepared by: Johnny Sullivan

CE 394K – GIS in Water Resources

December 2, 2011

## **I. Background**

While there is consensus in the scientific community that anthropogenic climate change exists, the locations and magnitudes of its effects are far less certain. Some regions will experience wetter climates, whereas other regions will receive less rainfall than ever before. This summer Texas suffered – and continues to suffer – one of the worst droughts on record. This primarily affects the agriculture industry, destroying crops and weakening or killing livestock. When such a large industry is affected, however, the repercussions are felt in countless other areas; prices for many commodities people depend upon will go up – everything from steak to a pair of socks (Hylton 2011). It is not known whether this drought can be attributed directly to climate change, but it follows that having a system in place to understand the drought situation better would improve future preparedness for such events, whether climate change-related or otherwise.

One way to consider drought is to measure the stage and flow of important rivers and water bodies. This is just one part of the picture, though. Another aspect is the amount of water present in the soils, for this relates directly to plant health, which in turn preserves soil quality and prevents erosion. Drought could be understood better if a complete picture of soil moisture content could be constructed. It is not enough just to know how much water is present; the current water content as well as the maximum available water storage must be known. These two pieces of data can be combined to present a picture of the current drought status because they will describe how much water is contained in the soils compared to how much could be present.

## **II. Project Goal**

The goal of this project to create a system that is able to provide information about the available water content in Texas soils and update in real-time. In this way it will make available critical drought information to the numerous agencies in Texas who would find this data useful, such as the Texas Natural Resources Information System (TNRIS). This will be accomplished through a combination of the available water storage data provided by the Soil Survey Geographic Database (SSURGO) and the continuously-updating current soil water content data made available by the North-American Land Data Assimilation System (NLDAS). The SSURGO data divides land areas into polygons with unique values describing the maximum

potential water content for the given area. The NLDAS data, on the other hand, details how much moisture that soil is currently holding. The SSURGO data can be thought of as a bucket and the NLDAS data as the amount of water in the bucket. Knowing both of these pieces of information thus describes how full the bucket is, i.e. the extent of the current drought.

### **III. Acquisition of SSURGO Data**

The Soil Survey Geographic Database, administered by the Natural Resources Conservation Service (NRCS) – which, in turn, is an arm of the U.S. Department of Agriculture – makes available country-wide soil water storage data down to the county level. The data can be accessed from SSURGO’s Soil Data Mart website and is then most easily processed using the Soil Data Viewer (SDV), a tool distributed by the NRCS. This tool was developed as an ArcMap extension that facilitates the creation of soil-based thematic maps. The soil survey attribute database, which is behind the spatial soil data that is distributed by the Soil Data Mart, is complex and users can find it difficult to cull from it the specific data needed. The SDV was created to make this process easier; it acts as an interface between the user and the database (NRCS 2011). The soil maps, displayed as connected polygons, contain numerous types of information about the soil. The Soil Data Viewer focuses on a single variable for the map as a whole and displays that information so that the user does not need to be concerned with querying the database and then linking processed data to spatial data (NRCS 2011).

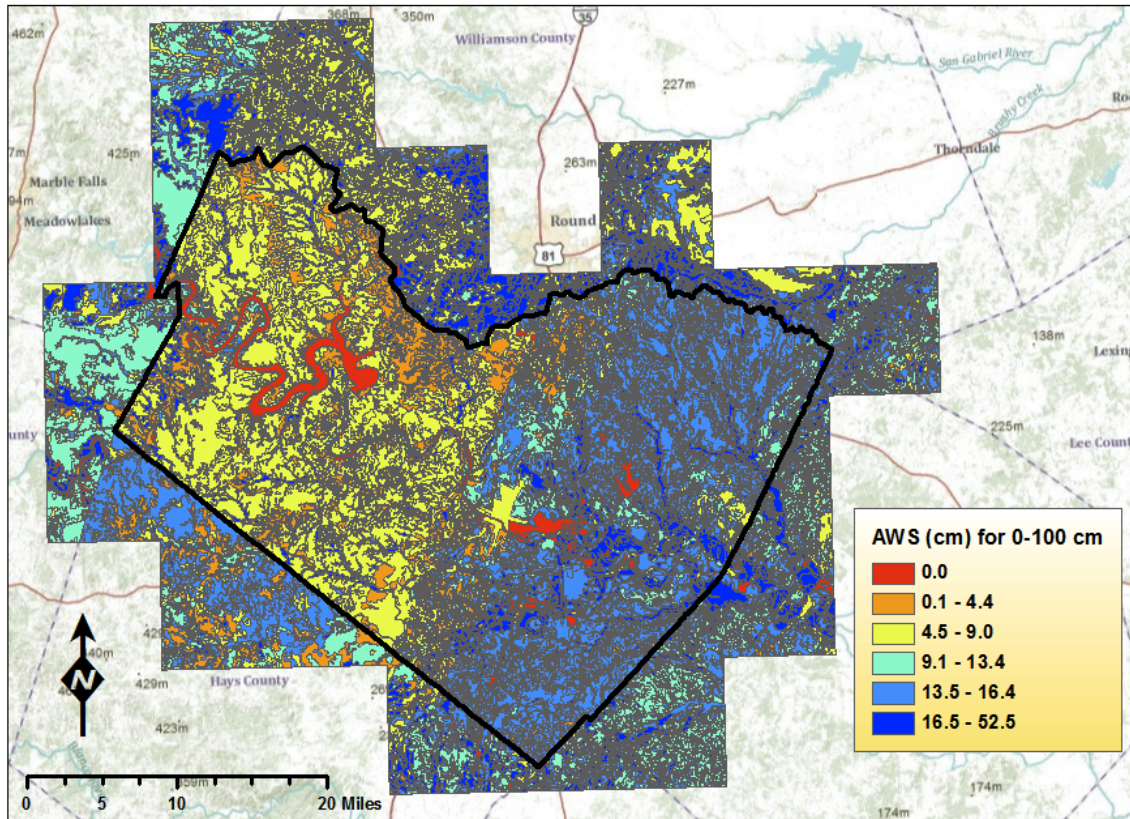
To generate a thematic map that contains available water storage (AWS) values in a layer of adjoining polygons for Travis County, one first goes to the Soil Data Mart (<http://soildatamart.nrcs.usda.gov>). Here, select the state and county of interest and then look at the “Available Survey Areas” category; a value of zero indicates no soil information, while a value of one or more indicates that some soil information is available. After this, one is able to see whether both tabular and spatial data is accessible, or just one or the other. Both types of data are required in order to create a thematic map. Finally, choose the spatial format (shapefile, in this case), the coordinate system, and either a custom database for the state of interest (if available) or one of the default national databases.

A (usually) short time after clicking “Submit,” the user will receive an email containing a link to the requested data. Unzip this file and the resulting folder will contain various files: .txt

and .xml files containing the Federal Geographic Data Committee metadata for the survey area in ASCII and XML file formats, respectively; a folder titled “spatial” that contains the spatial soil data requested, with the important file entitled “soilmu\_a\_xxXXX.shp” and xxXXX being two letters and three numbers describing the soil survey area of interest; a folder titled “tabular” containing the requested tabular data; and finally a template .zip file, the beginning of which will be named similar to “soildb\_.zip.” Unzip the template .zip file and, after navigating to the location of the now-unzipped template database, open the database file using MS Access. Enter into the Import form the entire pathname of the directory where the tabular data is located. Access will then complete the database import process. When that is finished the database can be closed and the data is ready to be used with the Soil Data Viewer. (The process described above is presented in further detail in the Soil Data Viewer User guide, available at [ftp://ftp-fc.sc.egov.usda.gov/NSSC/Soil\\_Data\\_Viewer/Soil\\_Data\\_Viewer\\_6\\_0\\_User\\_Guide.pdf](ftp://ftp-fc.sc.egov.usda.gov/NSSC/Soil_Data_Viewer/Soil_Data_Viewer_6_0_User_Guide.pdf))

The SDV User Guide should be consulted for specific instructions on thematic map generation, but the basic process is as follows. The user loads a valid “Soil Map Layer” that specifies the thematic map area; this is often one of the shapefiles included in the “spatial” folder discussed earlier. This is then specified when the SDV tool is run, generating the thematic map.

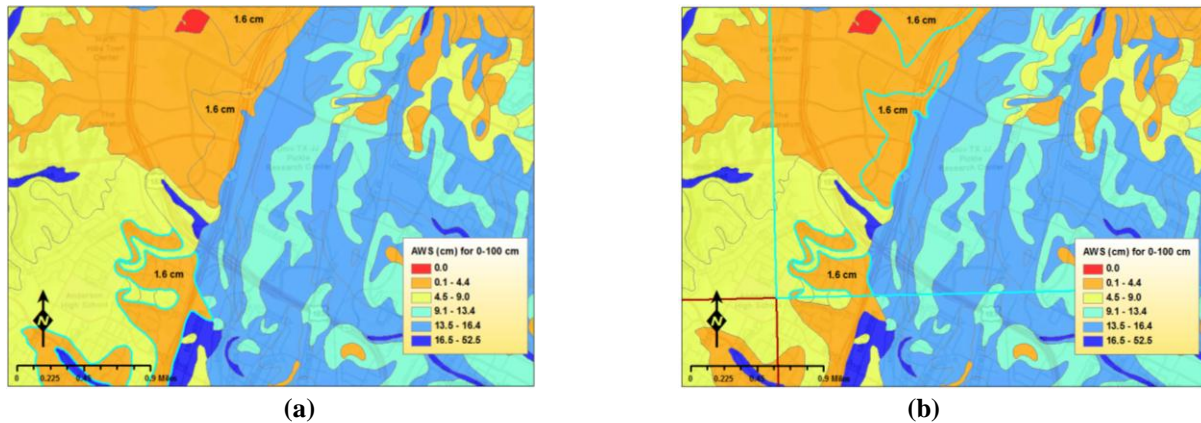
The SDV was not initially used to access SSURGO data for this project. Unaware of the existence of the SDV, SSURGO data for 0-100cm was obtained from Michael Dangermond at ESRI, who clipped it to the 30 1/8<sup>th</sup> degree quads that intersect Travis County (the clipping was done so that the SSURGO data could be combined with NLDAS data in the future, which is distributed as one value per 1/8<sup>th</sup> degree quad). This data can be seen in Figure 1. In the future, though, the SDV is the best tool by which to access SSURGO data for AWS.



**Figure 1.** Initial SSURGO data obtained from Michael Dangermond at ESRI (symbology changed and outline of Travis County added).

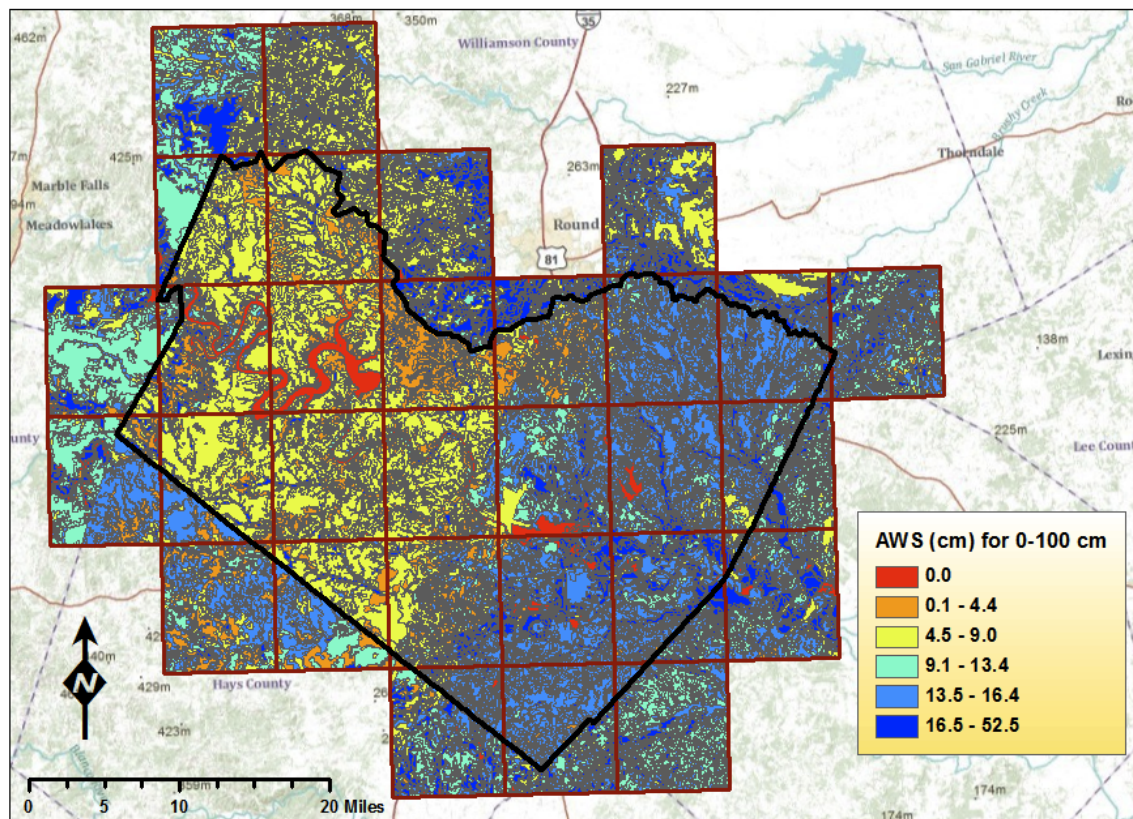
#### IV. AWS Analysis for Travis County

Once the SSURGO data was obtained, the task was to determine the total number of soil polygons and unique AWS values that exist for the 30 quads intersecting Travis County in order to understand the dimension of the problem of mapping computed soil moisture (NLDAS) down to the polygon level. With this initial data the polygons were not divided by quad, and there were many polygons for the same AWS value. The first step was to combine the SSURGO data with the borders of the 30 quads, obtained from the USGS. The “Intersect” tool was used here to create a new feature where soil polygons that spanned two quads were divided into two different polygons. The “Dissolve” tool was used next, with the “dissolve field” defined as the AWS value. Figure 2 displays the difference between selecting a polygon with an AWS value of 1.6cm (a) before and (b) after the “Intersect” and “Dissolve” tools were utilized.



**Figure 2.** Comparison of selection of the same 1.6cm AWS polygon (a) before the SSURGO data was intersected with the quads and dissolved by AWS value, and then (b) after Intersect and Dissolve were used.

This process produced an updated feature, shown in Figure 3, that contained within each quad only one polygon for each unique AWS value. SSURGO data is a static value, unlike the continuously-updating NLDAS data, so this map will not change for the duration of the analysis. Returning to the analogy used earlier, it describes how big the bucket is.



**Figure 3.** SSURGO data shown after being intersected with quads and dissolved by AWS value

The number of unique AWS values that exist in each quad and the maximum value among these were determined. These results are displayed in Table 1.

**Table 1.** Summary of AWS values after Intersecting with Travis County quads and Dissolving by AWS value. Count shows the total number of unique AWS values for the quad. Max shows the maximum AWS value.

Row Labels	Count of aws0100wta	Max of aws0100wta
AUSTIN EAST	46	18.48
AUSTIN WEST	35	18.48
BEE CAVE	24	18.48
BERTRAM	41	52.5
BUDA	63	18
COUPLAND	38	19.89
CREEDMOOR	39	19.26
DRIPPING SPRINGS	30	18
ELGIN WEST	42	39.64
HAMMETTS CROSSING	48	52.5
HUTTO	30	19.89
JOLLYVILLE	29	21.15
LEANDER	36	21.15
LIBERTY HILL	39	40
LYTTON SPRINGS	40	39.64
MANOR	28	21.44
MANSFIELD DAM	14	18
MONTOPOLIS	51	18.48
NAMELESS	38	21.15
OAK HILL	57	18
PACE BEND	41	52.5
PFLUGERVILLE EAST	35	19.89
PFLUGERVILLE WEST	61	21.15
SHINGLE HILLS	32	18
SIGNAL HILL	28	18
SPICEWOOD	34	52.5
STRUCTURE	62	19.89
TRAVIS PEAK	34	40
UTLEY	57	39.64
WEBBERVILLE	52	39.64
<b>Grand Total</b>	<b>1204</b>	<b>52.5</b>

Additionally, it was possible to determine from this analysis the total soil water storage available in Travis County. Simply for visualization's sake this was calculated for Travis County

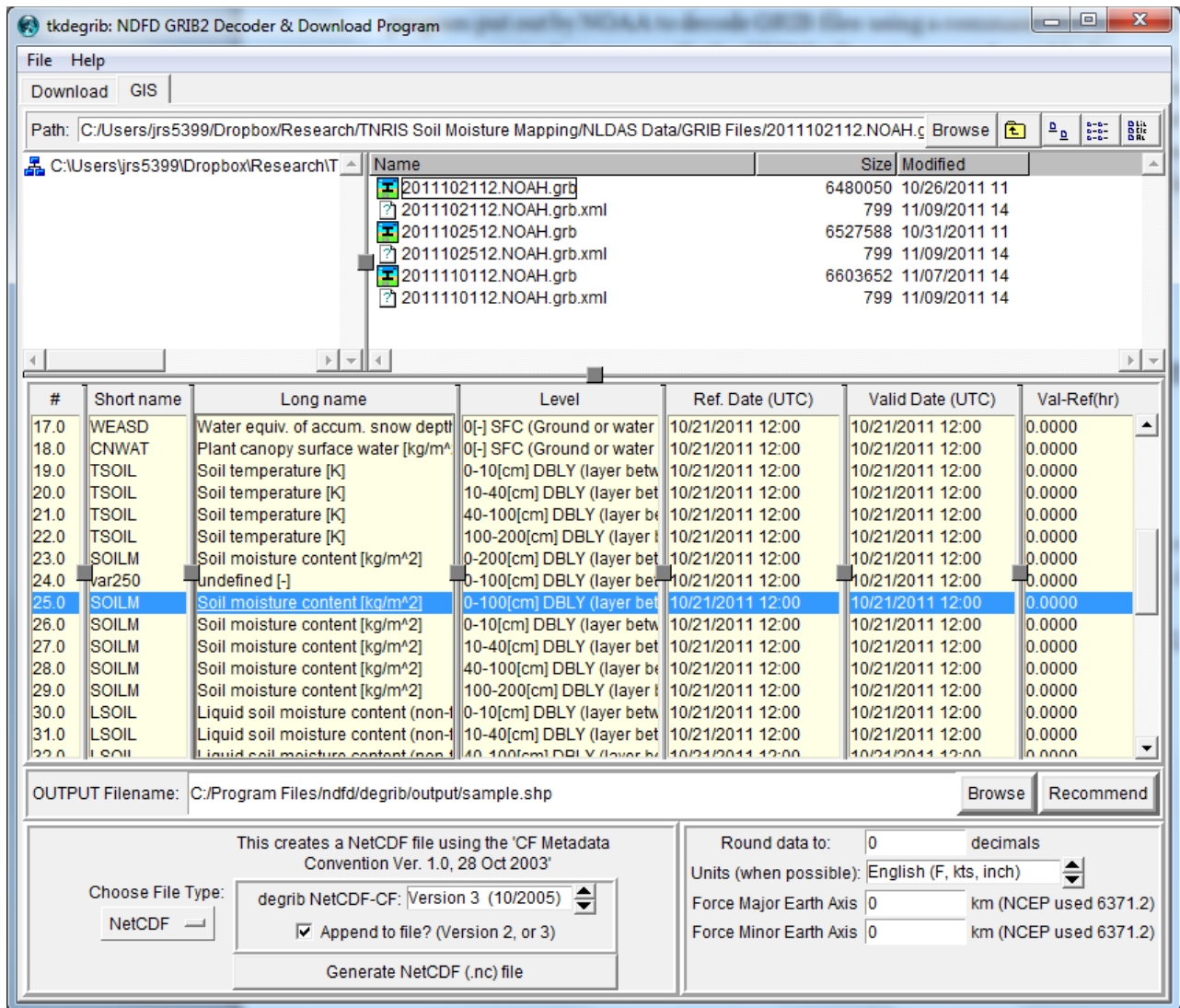
only, not for the 30 quads that intersect the county. This total storage was calculated to be 0.263 cubic kilometers. To put this in perspective, the total storage of Lake Travis is 1.44 cubic kilometers (Saunders 1996), so the available water storage in soil is roughly 20% of the storage in the lake.

## **V. Acquisition of NLDAS Data**

The NLDAS project is a collaboration between numerous governmental and academic institutions, a full list of which is available on the NLDAS website (<http://www.emc.ncep.noaa.gov/mmb/nldas/>). NLDAS publishes GRIB files in real-time that contain many types of information, one of which is the soil moisture content, in  $\text{kg/m}^2$ , for a depth of 0-100cm (the same depth examined with the SSURGO data). The data is available from an ftp server that contains multiple datasets; the Noah model was used for this analysis based on instructions from my advisor, Dr. Maidment. This model outputs data on an hourly basis, so in accessing it the most recent day was chosen and the file for 12:00pm on that day was downloaded. (In the future, this download process will be automated so that the most current data is being used.) The data is downloaded as a file that is in the GRIB file format. GRIB is a format commonly used for the storage of weather data, and, unfortunately, ArcMap does not support it. Instead, the file must first be converted to an ArcMAP-supported format before it can be opened in the mapping software.

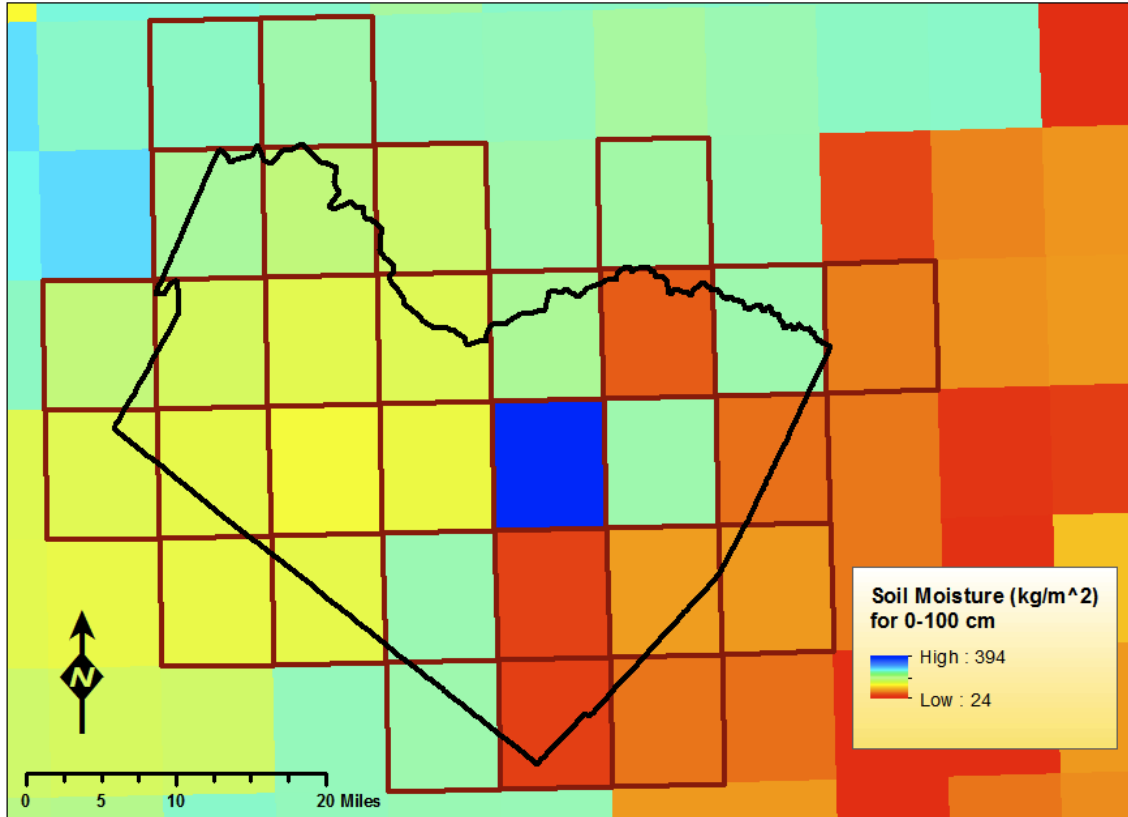
Numerous methods were attempted to accomplish this task. Much time was spent with `wgrib`, a program put out by NOAA to decode GRIB files using a command-line interface. Use of this program required one to specify a GRIBTAB environmental variable, however, and it proved difficult to succeed in getting this to run properly. Ultimately, the Degrib program was found to work well. It is a decoder program published by the National Digital Forecast Database with a simple graphical user interface. The user browses to the GRIB file and is then shown a list of all data attributes included in the file from which he is able to select the attribute of interest – in this case, soil moisture from 0-100cm. The user then selects the output file format he would like to export into. NetCDF was chosen here, as it is a common format used in ArcMap for gridded data. A screen capture of the Degrib program is shown in Figure 4 below.





**Figure 4.** Screen capture showing the Degrib program with the "Soil moisture content" attribute selected and output file type chosen to be NetCDF.

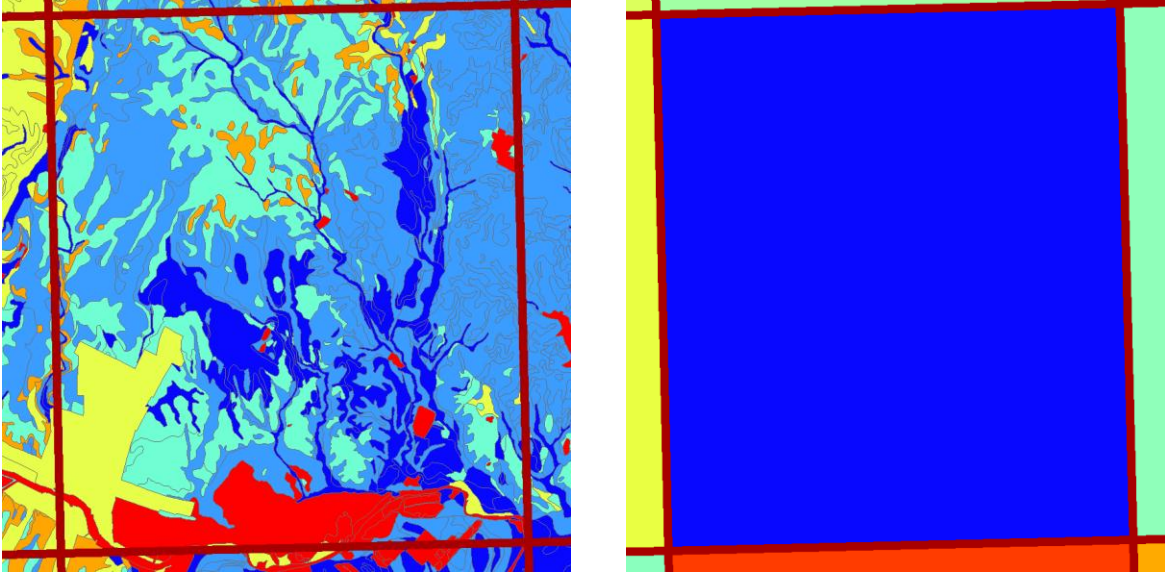
With the netCDF file obtained, the "Make netCDF Raster Layer" tool was used. Longitude and Latitude were chosen for X Dimension and Y Dimension, respectively. Once the NLDAS data is exported to a raster feature, it is ready for analysis. Figure 5 shows the final prepared data.



**Figure 5.** NLDAS data for Travis County and surrounding area.

## VI. Synthesis of SSURGO and NLDAS Data

SSURGO and NLDAS data is presented in two different structures making it necessary to develop a methodology for the combination of the two. NLDAS provides one value for every 7.5 minute quad across the country; the SSURGO data operates on a much smaller scale than that, with values for numerous soil polygons within each quad. Figure 6 illustrates this difference. Thus there is one NLDAS value for many different SSURGO soil polygons. In order to provide meaningful drought information, it must be determined what fraction of the NLDAS value (i.e., the current water storage) should be apportioned to each soil polygon; allotting the same value across the quad does not make sense.



**Figure 6.** Comparison of SSURGO and NLDAS data; SSURGO has many soil polygons with different values, whereas NLDAS has just one value for the whole quad.

It was decided that the NLDAS value should be distributed based on the ratio of current water storage volume to total volume of available water storage. In other words, if one imagines the total available water storage (the SSURGO data) as a bucket, the NLDAS data describes how much water is in that bucket. The ratio of the two, then, is the extent to which the bucket is full. It is then further assumed that the bucket is the same percentage full across the quad.

## VII. Model Methodology

The maximum available water storage (AWS in cm) and the current water storage (CWS' in kg/m<sup>2</sup>) are known, as is the area occupied by each AWS polygon and each quadrangle. As the CWS' value for each quad is in kg/m<sup>2</sup>, it is first divided by the density of water and then multiplied by the area of the quad to obtain the current water storage volume of the quad in cubic meters:

$$CWS'_{vol} (m^3) = CWS' \left( \frac{kg}{m^2} \right) \times \frac{m^3}{1000 kg} \times A_{quad} (m^2) \quad (1)$$

Dividing this current water storage volume by the total AWS volume yields the fraction of available storage that is currently occupied. AWS is a depth in cm for each polygon, however, so this must be converted to a volume by multiplying by the area of the polygon:

$$AWS_{vol} (m^3) = AWS (cm) \times \frac{m}{100 \text{ cm}} \times A_{polygon} (m^2) \quad (2)$$

The AWS volume for each polygon is then known, and the sum of these values for n polygons describes the total AWS volume for the entire quad:

$$AWS_{vol_{tot}} (m^3) = \sum_{i=1}^n AWS_{vol_i} (m^3) \quad (3)$$

Thus, the fraction of available storage that is currently occupied is:

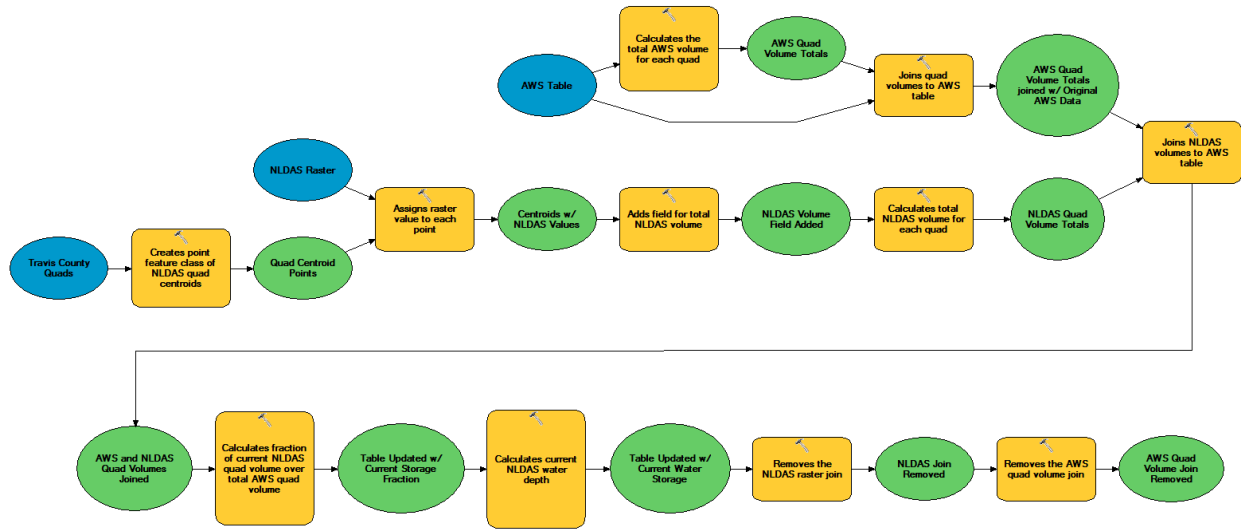
$$X = \frac{CWS_{vol} (m^3)}{AWS_{vol_{tot}} (m^3)} \quad (4)$$

Multiplying each AWS value by this fraction, X, yields the current water storage, CWS (in cm):

$$CWS_i (cm) = AWS_i (cm) \times X \quad (5)$$

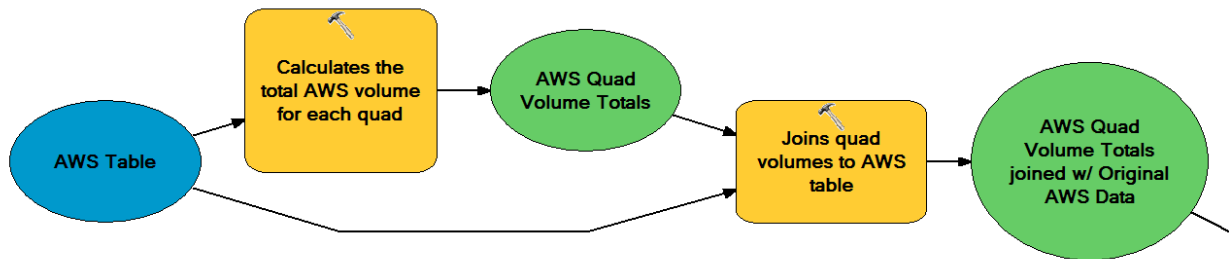
Returning to the bucket analogy, AWS describes the size of a bucket and X describes the percentage to which that bucket is full, so multiplying the two together results in the current height of water in the bucket.

Because this process would need to be carried out whenever a new NLDAS file is published (e.g. every hour, once the system is automated), a program was built using the Modelbuilder feature in ArcMap to conduct these calculations automatically. The model is shown in Figure 7.



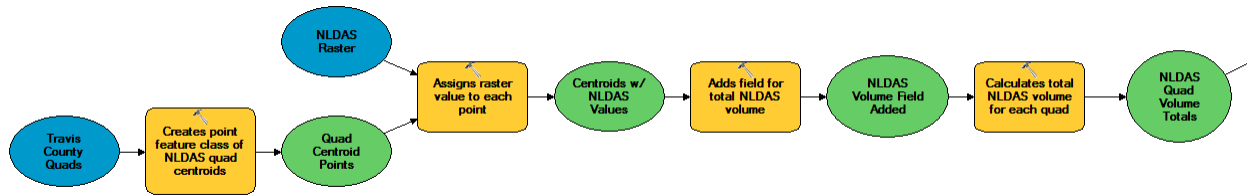
**Figure 7.** Model constructed using Modelbuilder to carry out NLDAS distribution calculations.

The first section, shown in Figure 8, first calculates the total AWS volume for each quad, Equation (3), using the “Summarize” tool. The individual polygon volumes, Equation (2), were calculated using the Field Calculator within the attribute table. They are summarized in the Statistics Field of the tool, and then the Case Field specifies the Quad Number (a number, 1-30, assigned to each quad) so that the sum of the volumes is calculated for each quad, not the entire 30 quad area. This is sent to an output table, which is then joined to the original AWS data feature using the “Add Join” tool.



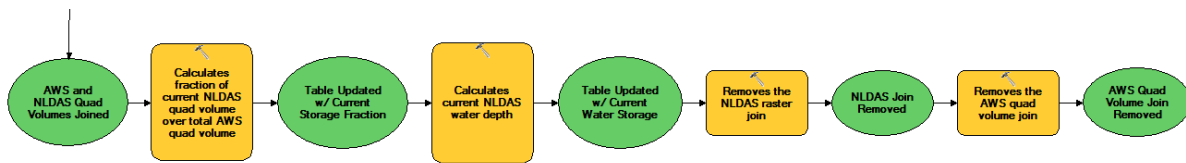
**Figure 8.** First section of model used to calculate total AWS volume for each quad.

While the model is carrying out the operations shown in Figure 8, it is also performing those depicted in Figure 9. Using the “Feature to Point” tool, a point feature class is created from the Travis Quads feature; the points are the centroids of the quads. The “Extract Values to Points” tool is then used to assign raster values to each quad centroid. “Add Field” makes sure that there is a field in which to add the NLDAS volumes that will be calculated. Finally, “Calculate Field” carries out the calculation shown in Equation (1).



**Figure 9.** Section of model run in parallel to section shown in Figure 3; calculates total NLDAS quad volumes.

In the last series of steps, displayed in Figure 10, Equations (4) and (5) are carried out. The features containing the AWS and NLDAS quad volumes are first joined, and then the “Calculate Field” tool is used twice: first to find the fraction of the NLDAS quad volume over the total AWS volume for the quad, Equation (4); and second to calculate the current water storage depth in cm, Equation (5). Lastly, the NLDAS raster join and AWS quad volume join are both removed so that these values can be updated when the model is run the next time.



**Figure 10.** Final section of model that calculates the fraction of current over available water storage volume, and then also calculates the current storage in cm.

### VIII. Problem with Model Output

The model was used to distribute the NLDAS soil moisture volume to the SSURGO polygons across the quad. When the model was run, however, it produced NLDAS depths that were frequently two to three times the depth of the SSURGO data values. This is not possible;

the current depth of soil moisture cannot be greater than the available soil moisture depth. To use the bucket analogy again, the model reported that the bucket was overflowing.

An investigation into the cause of this problem was then conducted. The model was checked over many times to verify that the methodology on which it was constructed was sound. This was all correct, and so the problem lay elsewhere. The discrepancy between the two models was solved with the help of Matthew Rodell and David Mocko at NASA. It was found that soil moisture is characterized using two parameters: field capacity and wilting point. The field capacity of a soil is the maximum amount of moisture it can hold, while the wilting point is the amount of moisture at which plants begin to die. SSURGO defines available water storage as “the amount of water that the soil...can store that is available to plants” (SSURGO Metadata and Columns Report). It is calculated from the available water capacity through an aggregation process. The available water capacity is defined as the difference between the field capacity and wilting point, i.e. the dynamic range of moisture after gravity drainage; the low point is not zero soil moisture, but rather the lowest moisture content plants can withstand. The NLDAS model, on the other hand, gives the volumetric soil moisture content, and this can range from zero to saturation. As such, it is theoretically possible for the NLDAS value to be greater than the amount of available storage. Given that Texas has been suffering from drought, however, this is not likely. It was further found that the datasets on which the NLDAS model is based are from at least twenty years ago, whereas the SSURGO data is current. This could be resulting in a significant inconsistency between the SSURGO-defined soil properties and those used by NLDAS.

## **IX. Solution to Problem with Data Disagreement**

Although the accuracy of the specific NLDAS values was called into question by the discovery of the outdated parameters, a solution was suggested by Matthew Rodell wherein the NLDAS data is used to obtain a sense of “dry” or “wet.” As an example, if NLDAS ranges from 10-50cm and SSURGO from 20-40, and the current NLDAS reading is 15, then it can be said that the soil is rather “dry” and that the current moisture content (based on the SSURGO data) is just above 20. A more technical way to describe this, also suggested by Rodell, is to define a soil wetness index (SWI) as:

$$SWI = \frac{CWS-WP}{FC-WP} \quad (6)$$

In the above equation, CWS is the current water storage from NLDAS, WP is the wilting point, and FC is the field capacity. Both WP and FC are from the NLDAS model. This will produce a value between 0 and 1 that will be multiplied by the SSURGO available water storage. The result of this will be the current water storage, with SSURGO defining the range and NLDAS providing how much moisture is present in the soil.

## **X. Continuing Work**

The strategy described above for utilizing the NLDAS model has not yet been implemented; this is the task that will be accomplished in the immediate future. The parameter values for field capacity and wilting point will be found, and the further required calculations will be added to the model. With that in place, the model will produce current water storage values consistent with the maximum available water storage identified by SSURGO, i.e. the bucket will no longer be overflowing. The current water storage will be divided by the SSURGO AWS to calculate a new soil wetness index, or in other terms a drought severity index, that is a combination of both NLDAS and SSURGO data.

The next step will be to automate the entire process. The goal is to create a map that is continuously updated in real-time via a program that downloads the current NLDAS run and then executes the model to apportion the current water storage to the SSURGO polygons. This will be accomplished through the use of Unidata's LDM and THREDDS server. LDM stands for Local Data Manager, and it is a program published by Unidata that fetches data and is then able to share it with other computers on a network. LDM has been installed on a computer on the CRWR network. In the near future it will be programmed to acquire new NLDAS runs, likely on a daily basis, and then transmit that data to a THREDDS server.

The THREDDS project, Thematic Realtime Distributed Data Services, is a way by which gridded data can be processed. The NLDAS runs are in the GRIB format, and THREDDS is one method that can be used to convert these files to netCDF; it is far more efficient than using Degrib because, as it does not rely on a graphical user interface, it can be integrated into an automated workflow. THREDDS ingests the GRIB files and converts them to its own native file



format. Data can then be queried from the server via, for example, a web coverage service request, at which time THREDDS converts the data from its native format to numerous other formats, such as netCDF.

In the last step, these netCDF files will be processed as input to the previously discussed model, in the form of a Python script. This will achieve the goal of producing a continuously updated map that displays current drought information. The map will be published online so that it can be accessed by anyone interested in the current extent of drought in Travis County. Eventually, this analysis will be expanded to include all of Texas.

## **References:**

Hylton, Hilary. Forget Irene: The Drought in Texas Is the Catastrophe That Could Really Hurt. Time U.S. 31 August 2011 <http://www.time.com/time/nation/article/0,8599,2091192-2,00.html>

Saunders, Geoffrey P. Sandy Creek/Lake Travis NPS Water Quality Study: Final Report. Lower Colorado River Authority. 1996. [http://www.twdb.state.tx.us/rwpg/rpgm\\_rpts/92483335.pdf](http://www.twdb.state.tx.us/rwpg/rpgm_rpts/92483335.pdf)

Soil Data Viewer User Manual

[ftp://ftp-fc.sc.egov.usda.gov/NSSC/Soil\\_Data\\_Viewer/Soil\\_Data\\_Viewer\\_6\\_0\\_User\\_Guide.pdf](ftp://ftp-fc.sc.egov.usda.gov/NSSC/Soil_Data_Viewer/Soil_Data_Viewer_6_0_User_Guide.pdf)

SSURGO Metadata and Columns Report.

<http://soildatamart.nrcs.usda.gov/documents/SSURGO%20Metadata%20-%20Tables%20and%20Columns%20Report.pdf>