

Review and pull soil data for test locations of corn - HWSD-global set

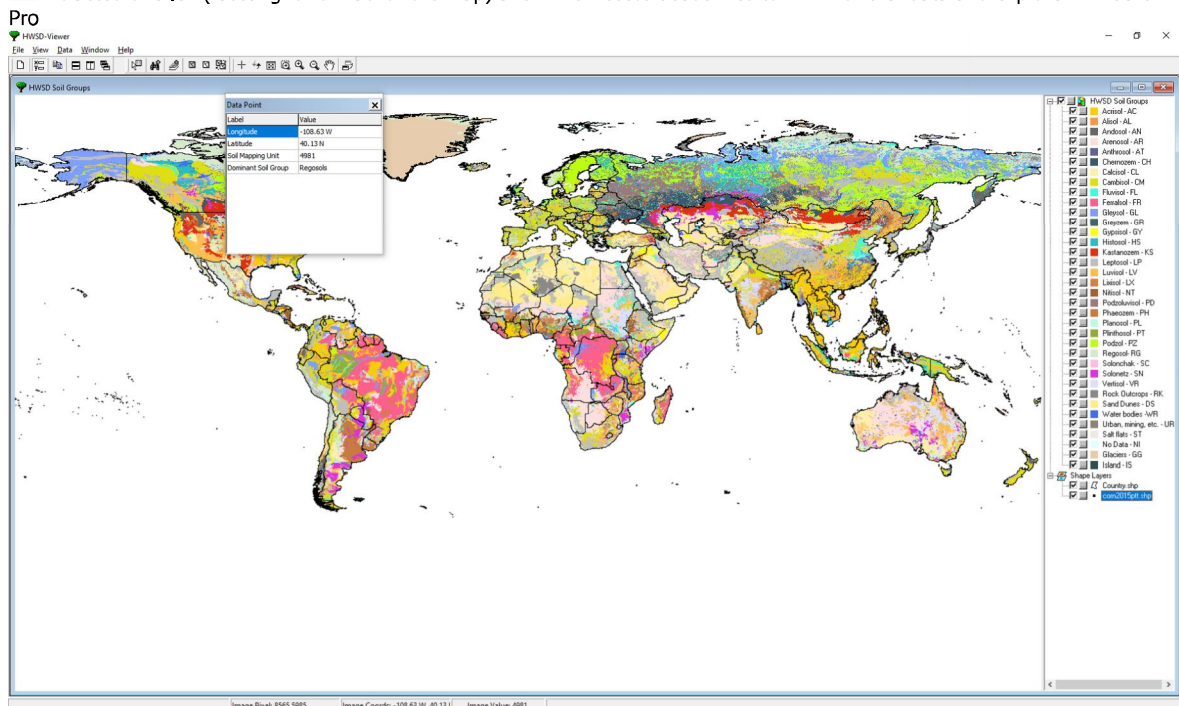
Notebook: Computomics
Created: 10/28/2020 8:57 PM
Author: Larisa Golovko

Updated: 10/29/2020 8:09 AM

- HWSD-Viewer

Workflow:

- ☒ Downloaded and extracted Soil Viewer and the datasets ~2 Gb
- ☒ Installed viewer to quickly explore data (note the coverage and list of soil types) - the data can be pulled as a table of soil properties under Lat/Lon - note: format of geographical coordinates is different than in given data set of 28 corn trial data.
- ☒ Extracted one .bil (raster grid format for the map) and 2 MS Access dbase files to link with trial data and explore in ArcGIS Pro



How HWSD works

The HWSD is composed of a raster image file and a linked attribute database. The raster database consists of 21600 rows and 43200 columns, of which 221 million grid cells cover the globe's land territory.

Each grid cell in the database is linked to commonly used soil parameters, namely, organic carbon, pH, water storage capacity, soil depth, cation exchange capacity of the soil and the clay fraction, total exchangeable nutrients, lime and gypsum contents, sodium exchange percentage, salinity, textural class, and granulometry. HWSD allows soil compositions to be displayed or queried in terms of user-selected soil parameters.

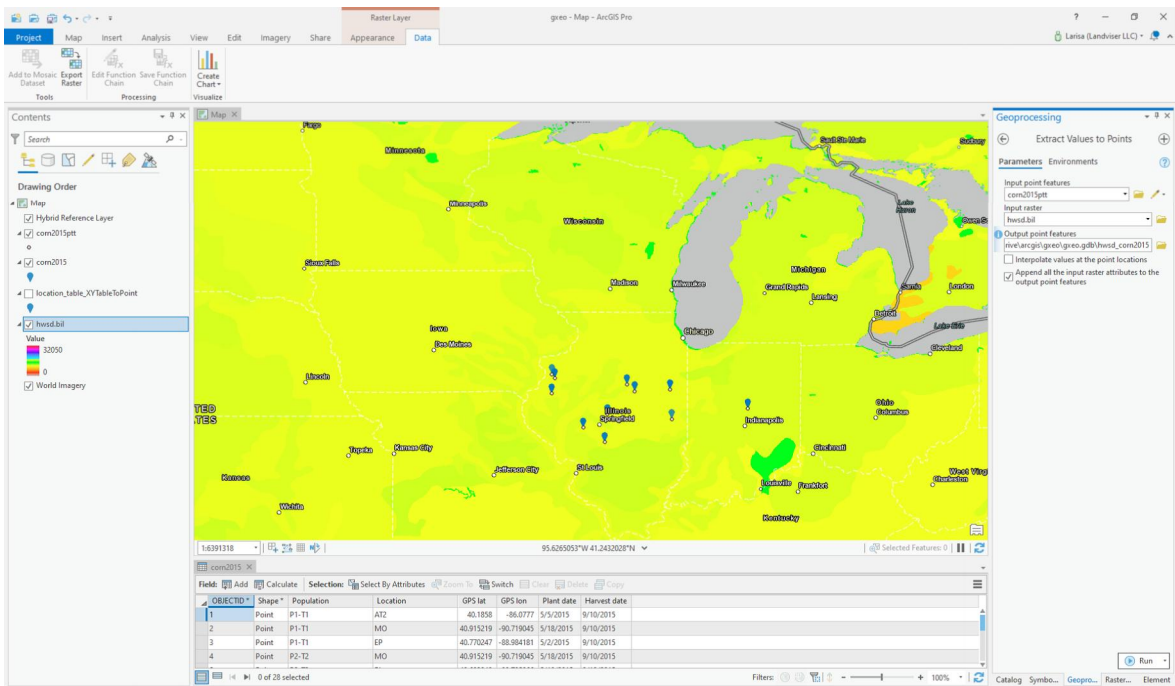
The HWSD Viewer allows soil association compositions to be displayed or queried in terms of user-selected soil parameters, and it provides a geographical tool to query and visualize the database. For modeling, the HWSD and its geographical layer can directly be read or imported by common GIS and Remote Sensing software.

The HWSD allows the soil components and attributes to be seen at a high level of spatial resolution at the global scale.

gxco - Map - ArcGIS Pro

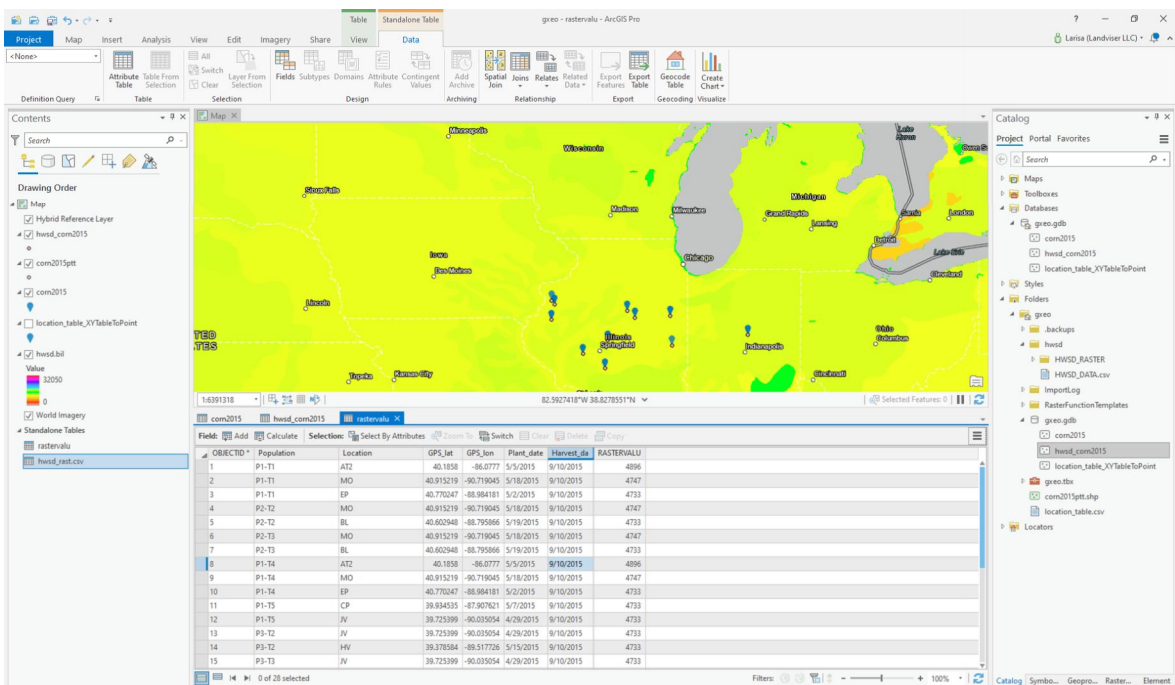
✓ see below soil grid data and mapped locations of corn trial supplied (there are 28 entries, but only 10 unique locations - there 2015-2016 data, which is relevant to weather data, but not for soil)

bigger and more diverse spatial dataset is desired to really capture the EO effect, the workflow can accept as many geo-located data as available in one table and link with soil .bil file (manual, but API can be developed)



✓ HWSD .bil file has arbitrary numbers - ids of global soil mapping units to link with Access dbase of soil properties

✓ spatial join performed in GIS software - only 3 different soil IDs present for 10 location in test



✓ test location file with RASTERVALU linked **geographically** is in GitLab soils/ folder (.csv file)

MS Access - HWSD : Database-
C:\Users\laris\OneDrive\arcgis\gxco\hwsd\HWSD.mdb
(Access 2000 file format)

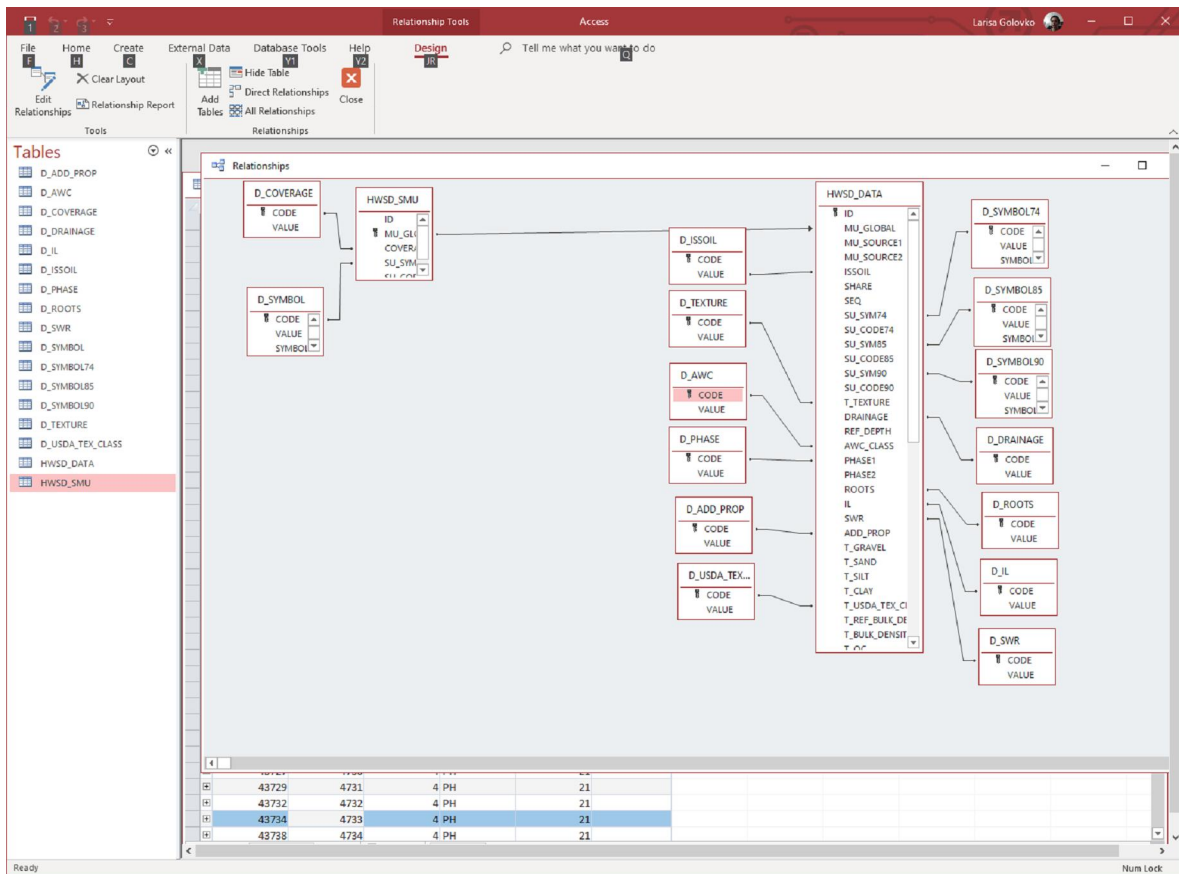
Overview of MS Access soil dbase:

☒ links with GIS map by MU_Global (mapping unit)

Mapping Unit Code (Global)

ID	MU_Global	COVERAGE	SU_SYMBOL	SU_CODE
43635	4692	4 PD	20	20
43638	4693	4 PD	20	20
43640	4694	4 LP	16	16
43644	4697	4 GL	11	11
43650	4698	4 GL	11	11
43653	4699	4 GL	11	11
43656	4700	4 GL	11	11
43659	4701	4 GL	11	11
43662	4702	4 GL	11	11
43665	4703	4 GL	11	11
43667	4704	4 GL	11	11
43669	4705	4 GL	11	11
43671	4706	4 GL	11	11
43673	4707	4 GL	11	11
43676	4708	4 GL	11	11
43678	4710	4 GL	11	11
43682	4711	4 GL	11	11
43686	4713	4 GL	11	11
43687	4714	4 GL	11	11
43689	4715	4 GL	11	11
43691	4716	4 GL	11	11
43693	4717	4 GL	11	11
43695	4718	4 PH	21	21
43697	4719	4 PH	21	21
43701	4720	4 PH	21	21
43705	4721	4 PH	21	21
43706	4722	4 PH	21	21
43707	4723	4 PH	21	21
43709	4724	4 PH	21	21
43712	4725	4 PH	21	21
43714	4726	4 PH	21	21
43717	4727	4 PH	21	21
43719	4728	4 PH	21	21
43723	4729	4 PH	21	21
43727	4730	4 PH	21	21
43729	4731	4 PH	21	21
43732	4732	4 PH	21	21
43734	4733	4 PH	21	21
43738	4734	4 PH	21	21

☒ review graphical relationships between tables in dbase



☒ key table with all soil properties -> extracted as standalone Excel and posted on GitLab in /soils folder

ID	MU_GLOBAL	MU_SOURCE	MU_SOURCE1	ISSOIL	SHARE	SEQ	SU_SYM74	SU_CODE74	SU_SYMB5	SU_CODE5	SU_SYM90	SU_CODE90	T_TEXTURE	DRAINAGE	REF_DEPTH	AWC_CLASS
1	7001 1			0	100	1	UR	228 UR								
2	7002 2			2	0	100	1	HD	227 HD		202					
3	7003 3			3	0	100	1	WR	230 WR		198					
4	7004 4			4	0	100	1	Od	156 HSF		89	3	1	100	1	
5	7005 5			5	0	100	1	GG	231 GG		199					
6	7006 70001		70001	1	70	1	Tv	197 ANz		35	1	5	100	4		
7	7006 70001		70085	1	20	2	Th	194 ANH		32	1	5	100	4		
8	7006 70001		70086	1	10	3	Tv	197 ANi		37	2	4	100	4		
9	7007 70002		70001	1	80	1	Tv	197 ANz		35	1	5	100	4		
10	7007 70002		70085	1	20	2	Th	194 ANH		32	1	5	100	4		
11	7008 70003		70001	1	90	1	Tv	197 ANz		35	1	5	100	4		
12	7008 70003		70011	1	10	2	Ok	159 HSi		91	2	1	100	1		
13	7009 70004		70001	1	90	1	Tv	197 ANz		35	1	5	100	4		
14	7009 70004		70012	1	10	2	Ok	158 HSi		91	2	1	100	1		
15	7010 70005		70001	1	90	1	Tv	197 ANz		35	1	5	100	4		
16	7010 70005		70050	1	10	2	Gm	84 GLm		14	2	2	100	1		
17	7011 70006		70001	1	100	1	Tv	197 ANz		35	1	5	100	4		
18	7012 70007		70002	1	100	1	Qa	174 ARa		43	1	6	100	3		
19	7013 70008		70003	1	80	1	Qc	175 ARb		40	1	6	100	3		
20	7013 70008		70099	1	10	2	Xk	237 ClJ		59	2	4	100	1		
21	7013 70008		70102	1	10	3	Zo	223 Sch		163	2	4	100	1		
22	7014 70009		70003	1	90	1	Qc	175 ARb		40	1	6	100	3		
23	7014 70009		70049	1	10	2	Gm	84 GLm		14	2	2	100	1		
24	7015 70010		70003	1	90	1	Qc	175 ARb		40	1	6	100	3		
25	7015 70010		70047	1	5	2	Gd	66 GLd		12	2	2	100	1		
26	7015 70010		70052	1	5	3	Gx	90 GLi		17	2	2	100	1		
27	7016 70011		70003	1	90	1	Qc	175 ARb		40	1	6	100	3		
28	7016 70011		70065	1	10	2	Kh	124 KSh		93	2	4	100	1		
29	7017 70012		70003	1	90	1	Qc	175 ARb		40	1	6	100	3		
30	7017 70012		70075	1	10	2	Pof	170 Pzf		151	1	6	100	6		
31	7018 70013		70003	1	90	1	Qc	175 ARb		40	1	6	100	3		
32	7018 70013		70076	1	10	2	Qc	175 ARb		40	1	6	100	3		
33	7019 70014		70003	1	90	1	Qc	175 ARb		40	1	6	100	3		
34	7019 70014		70079	1	10	2	Rx	189 Rgi		161	2	3	100	6		
35	7020 70015		70003	1	90	1	Qc	175 ARb		40	1	6	100	3		
36	7020 70015		70102	1	10	2	Zo	223 Sch		163	2	4	100	1		
37	7021 70016		70003	1	100	1	Qc	175 ARb		40	1	6	100	3		
38	7022 70017		70004	1	90	1	Bk	32 CMc		65	1	5	100	6		
39	7022 70017		70033	1	10	2	Ck	45 CMk		53	3	4	100	1		
40	7023 70018		70005	1	80	1	Bh	31 CMu		64	2	4	100	1		
41	7023 70018		70026	1	10	2	Bd	34 CMd		63	1	5	100	6		
42	7023 70018		70042	1	10	3	EO	66 LPk		100	1	4	30	5		
43	7024 70019		70005	1	80	1	Bh	31 CMu		64	2	4	100	1		

Access - HWSD_META : Database- C:\Users\laris\OneDrive\arcgis\gxco\hwsd\HWSD_META.mdb (Access 2000 file format)

- ☒ whole another MS Access dbase is given for explaining METADATA about soils
- ☒ exported as Excel file and posted in Gitlab in **soils/**

ID	FIELD	UNIT	DESCRIPTION	DATATYPE	DOMAIN
0	ID		Database ID	Number	
1	MU_GLOBAL		Soil Mapping Unit	Number	
2	MU_SOURCE1			String	
3	MU_SOURCE2			Number	
4	COVERAGE		Coverage	Number	D_COVERAGE
5	ISSOIL			Number	D_ISSOIL
6	SEQ		Sequence	Number	
7	SHARE	%	Share in Soil Mapping Unit	Number	
8	SU_SYMBOL		Dominant Soil Group Symbol	String	
9	SU_SYM74		Soil Unit Symbol (FAO 74)	String	
10	SU_SYM85		Soil Unit Symbol (FAO 85)	String	
11	SU_SYM90		Soil Unit Symbol (FAO 90)	String	
12	SU_CODE		Dominant Soil Group	Number	D_SYMBOL
13	SU_CODE74		Soil Unit Name (FAO74)	Number	D_SYMBOL74
14	SU_CODE85		Soil Unit Name (FAO 85)	Number	D_SYMBOL85
15	SU_CODE90		Soil unit Name (FAO 90)	Number	D_SYMBOL90
16	T_TEXTURE		Topsoil Texture	Number	D_TEXTURE
17	REF_DEPTH	cm	Reference Soil Depth	Number	
18	DRAINAGE		Drainage class (0-0.5% slope)	Number	D_DRAINAGE
19	AWC_CLASS	mm	AWC	Number	D_AWC
20	PHASE1		PHASE1	Number	D_PHASE
21	PHASE2		PHASE2	Number	D_PHASE
22	ROOTS	cm	Obstacles to Roots (ESDB)	Number	D_ROOTS
23	IL	cm	Impermeable Layer (ESDB)	Number	D_IL
24	SWR		Soil Water Regime (ESDB)	Number	D_SWR
25	ADD_PROP			Number	D_ADD_PROP
26	T_GRAVEL	%	Topsoil Gravel Content	Number	
27	T_SAND	%	Topsoil Sand Fraction	Number	
28	T_SILT	%	Topsoil Silt Fraction	Number	
29	T_CLAY	%	Topsoil Clay Fraction	Number	
30	T_USDA_TEX_C		Topsoil USDA Texture Classification	Number	D_USDA_TEX_C
31	T_REF_BULK_D	kg/dm3	Topsoil Reference Bulk Density	Number	
32	T_BULK_DENSITY	kg/dm3	Topsoil Bulk Density	Number	
33	T_OC	% weight	Topsoil Organic Carbon	Number	
34	T_PH_H2O		Topsoil pH (H2O)	Number	
35	T_CEC_CLAY	cmol/kg	Topsoil CEC (clay)	Number	
36	T_CEC_SOIL	cmol/kg	Topsoil CEC (soil)	Number	
37	T_BS	%	Topsoil Base Saturation	Number	
38	T_TEB	cmol/kg	Topsoil TEB	Number	

Next steps:

- ☐ I can link those files together in Pandas, or you can grab those yourself, but we would only have 3 different soils (and properties) for this test set -
- ☐ looking for different sets/crops, approaching my contacts for:
 - ☒ Columbia: oil palms, rice
 - ☐ USA: cranberries, blueberries
 - ☐ Russia: rice?
- ☐ Do you have bigger sets?
- ☐ review HWSD data set in more details - which properties would be of importance to different crops/scenarios, literature review
- ☐ explore SSURGO data for more detail soils info (only applicable to US)
- ☐ develop API for soil data spatial pull aside from ESRI software (although landviser now is a partner and actually the use /incorporation of their modules could be beneficials for exposure and marketing our solution to other companies)