National Soil Survey Center
Soil Survey Laboratory

Updates to the NCSS Soil Characterization Database
or
Integration of the NCSS Soil Characterization Database and NASIS

Henry Ferguson, soil scientist NSSC

Before I begin the main presentation
Some critical issues

Documentation!!! Why did you sample?

Original Source!!! Microfiche & PDF scans

User Pedon ids and Laboratory Sample numbers!!!

Depths of sampling!!! Horizons vs by depth

Why were changes made to NASIS and the Laboratory Data Mart?

To provide the laboratory data from KSSL and cooperating universities along with the latest classifications and descriptions

Changes to NASIS

1. Added the Taxonomic History Table
2. Added the NCSS Lab Pedon Table
3. Added the NCSS Lab Layer Table
4. Added depths to the Pedon Horizon Sample Table
5. Added the Pedon Horizon Lab Results Table

Changes to NASIS

6. Modified the Web Pedon Description Report
7. Added Ownership Groups by RO
8. Added Data Mining Reports
9. Added Calculations to assist in updating
    1. NCSS Lab Pedon table
    2. NCSS Lab Layer table
    3. Pedon Laboratory Sample Number
    4. Pedon Horizon Sample Numbers

Changes to Lab Data Mart

1. Added footnotes to the whole soil bulk density calculations
2. Modified results of the base saturation calculation not to exceed 100 percent
3. Added a link to the Microsoft Access Database from the NCSS Soil Characterization Website
4. Designed 6 new tables for future delivery

Support from the field

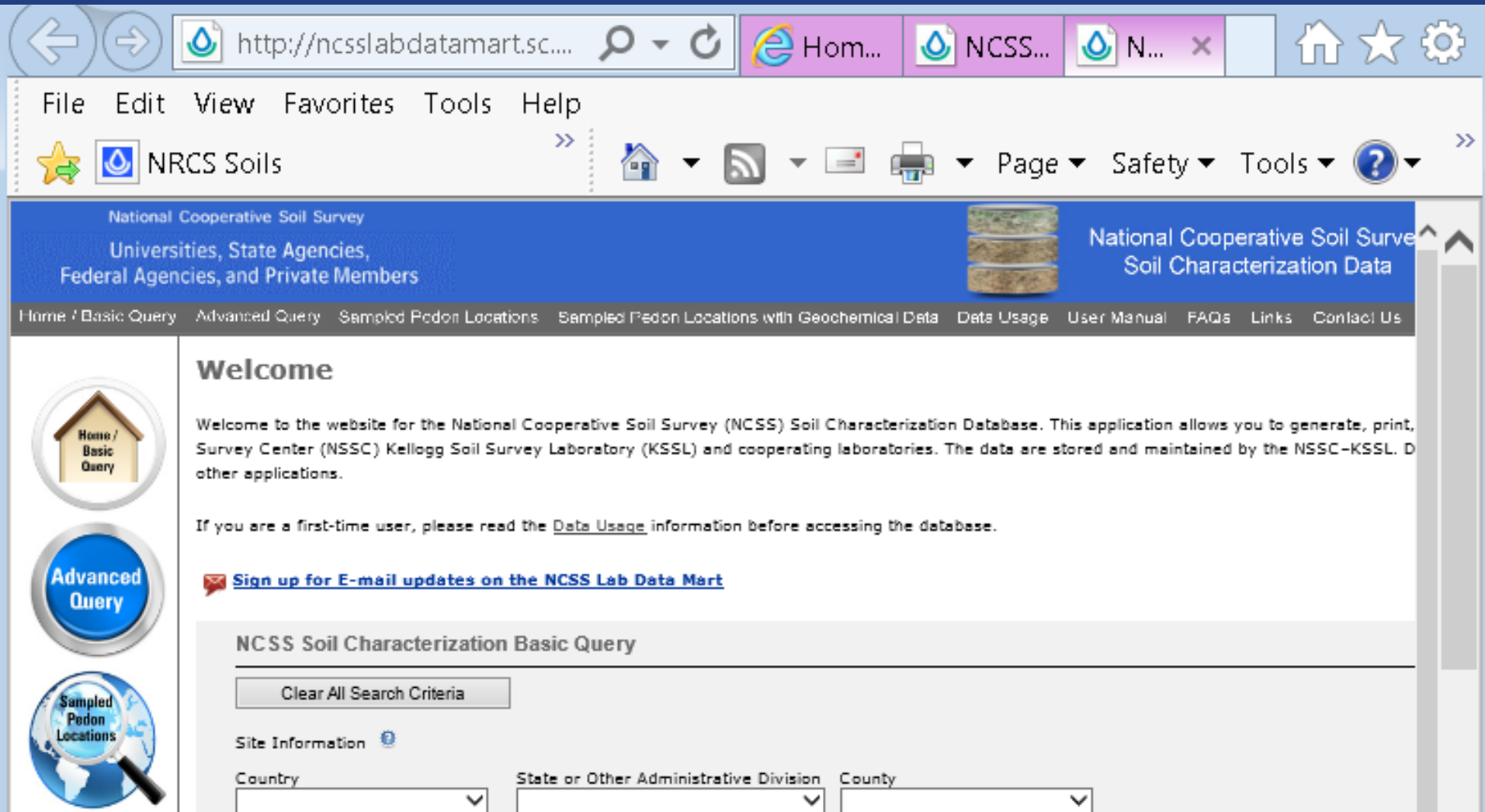
1. Kevin Godsey created the first download reports used for Analysis PC
2. Jason Nemecek has kept the pedon description report for the laboratory up to date
3. Adolfo Diaz and Jason Nemecek just finished creating a download for morphological and laboratory data from NASIS
4. Dylan Beaudette, Tom D Avello, Stephen Roecker and Jay Skovlin have supported the deployment and use of R to evaluate the laboratory data.
5. Kyle Stephens has updated a pedon descripton report to capture new tables and data elements added to NASIS

United States Department of Agriculture
Natural Resources Conservation Service

USDA NRCS

NCSS Soil Characterization Data

Web Site Released Feb 10, 2015, 7:00 PM



United States Department of Agriculture
Natural Resources Conservation Service

USDA NRCS

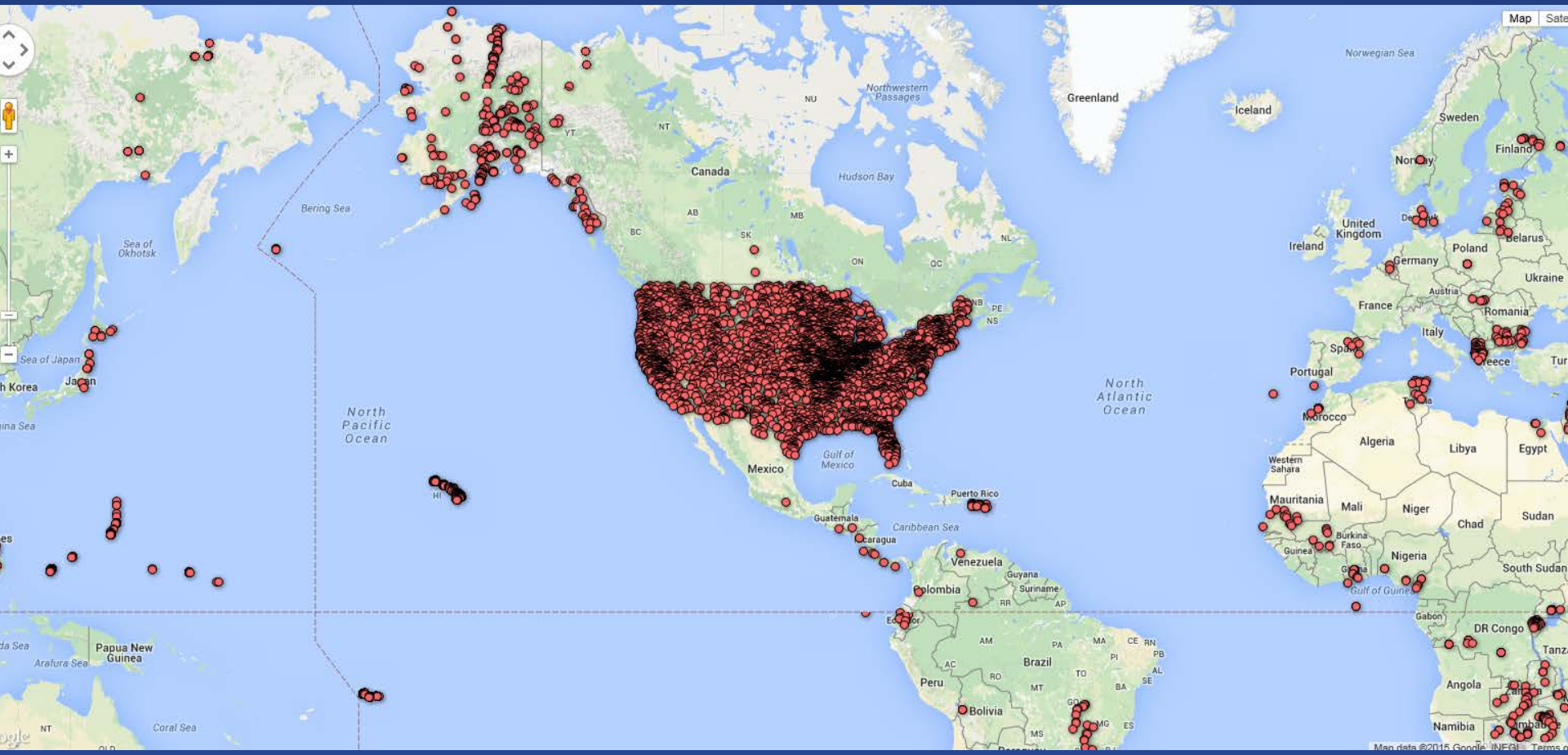
A Complete Copy of the National Database is Available in Microsoft ACCESS format from the Laboratory Website

National Cooperative Soil Survey Microsoft Access Database

A Microsoft Access database that contains the most commonly requested data addition to commonly requested data, the Access database includes metadata common columns. Users that wish to obtain the original data, which is separat

Soil Characterization Database (Zip file; 1.1 MB).

Google Fusion Map which can be used for Quality Assurance



NCSS Lab Pedon and Layer Tables

NCSS Layer Lab Data | NCSS Pedon Lab Data

Taxon Name	Layer...	Lab Sample #	Top Depth	Bottom...	Lab Text...	VC Sand - ...	Co. Sand -...	Med. Sand - ...	Fine Sand - ...
tifton	17	11N00278	0	2	s	0.8	9.0	34.2	31.9
tifton	18	11N00279	2	10	ls	1.8	8.7	30.8	33.3
tifton	19	11N00280	10	30	ls	1.2	7.4	31.0	34.1
tifton	20	11N00281	30	45	fsl	1.3	9.4	21.6	34.2
tifton	1	11N00263	0	0	ls	1.4	4.8	29.4	36.7
tifton	2	11N00264	0	0	s	1.6	7.5	26.0	37.0
tifton	3	11N00265	0	0	ls	1.4	7.3	27.0	37.4
tifton	4	11N00266	0	0	fsl	1.0	6.1	18.3	37.8
tifton	5	11N00267	0	0	scl	1.4	4.5	17.8	25.1
tifton	6	11N00268	0	0	scl	1.5	6.8	22.4	26.1
tifton	7	11N00269	0	0	scl	2.2	9.5	25.3	23.4
tifton	8	11N00270	0	2	ls	0.7	6.0	23.9	40.1
tifton	9	11N00271	2	16	ls	1.0	7.2	27.6	35.8
tifton	10	11N00272	16	26	ls	1.7	7.8	23.7	36.1
tifton	11	11N00273	26	45	ls	1.4	6.9	19.6	40.7
tifton	13	11N00274	0	2	ls	1.4	5.9	29.5	35.3
tifton	14	11N00275	2	18	ls	1.6	8.3	26.8	35.9
tifton	15	11N00276	18	29	ls	1.7	8.9	30.1	33.6
tifton	16	11N00277	29	45	fsl	1.1	7.2	20.4	35.1

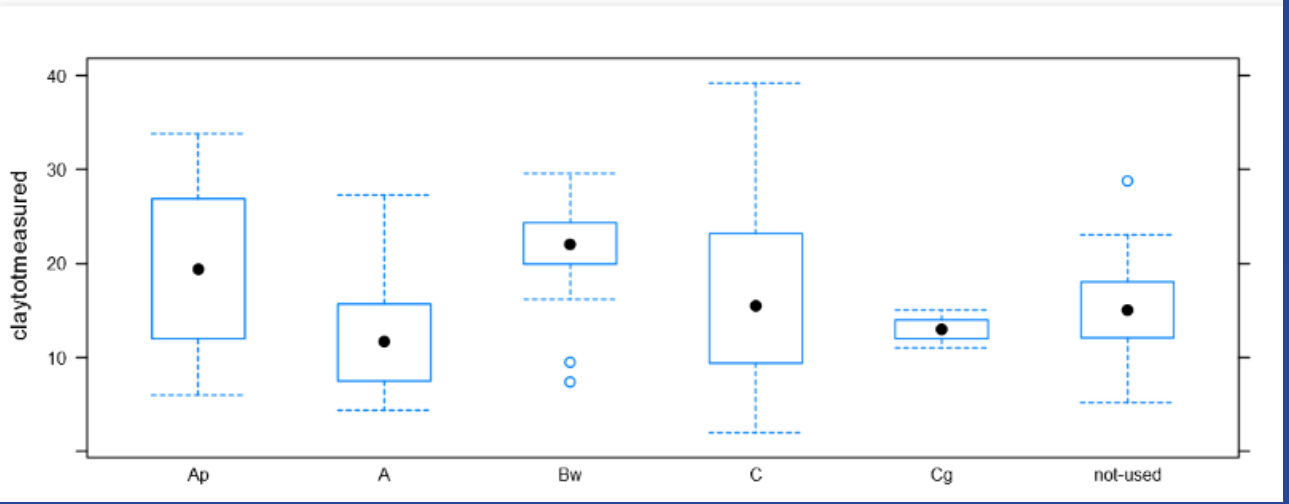
Use the fetchNASISLabData() in R

Summarizing labdata with R

```
quantile(lh$claytotmeasured, na.rm=TRUE)
```

```
  0%  25%  50%  75% 100%
 2.0 10.9 16.9 23.3 39.2
```

```
bwplot(claytotmeasured~genhz, data=lh)
```



For the benefit of individuals that have tuned into previous webinars, I am going to discuss developments in reverse. The newest to the oldest changes.

Why 6 new tables for delivering data?

1. Fewer joins
2. Easier to find desired data
3. Clearer links to metadata/methods
4. Clearer distinction between calculations based upon measured results vs those based on estimates or default bulk densities and particle densities

What are the 6 table names?

1. Physical Properties
2. Chemical Properties
3. Major and Trace Elements and Total Oxides
4. Optical Mineralogy Glass Counts
5. X-Ray And Thermal
6. Calculations Including Estimates and Default Values

What is the time table?

1. The scripts have been written
2. Sample tables have been generated
3. The data has been reviewed
4. Additional review will be needed
5. The tables can then be generated for distribution (1 to 4 days per table)
6. Possibly create a web service to point to the tables

What lead up to the creation of new tables?

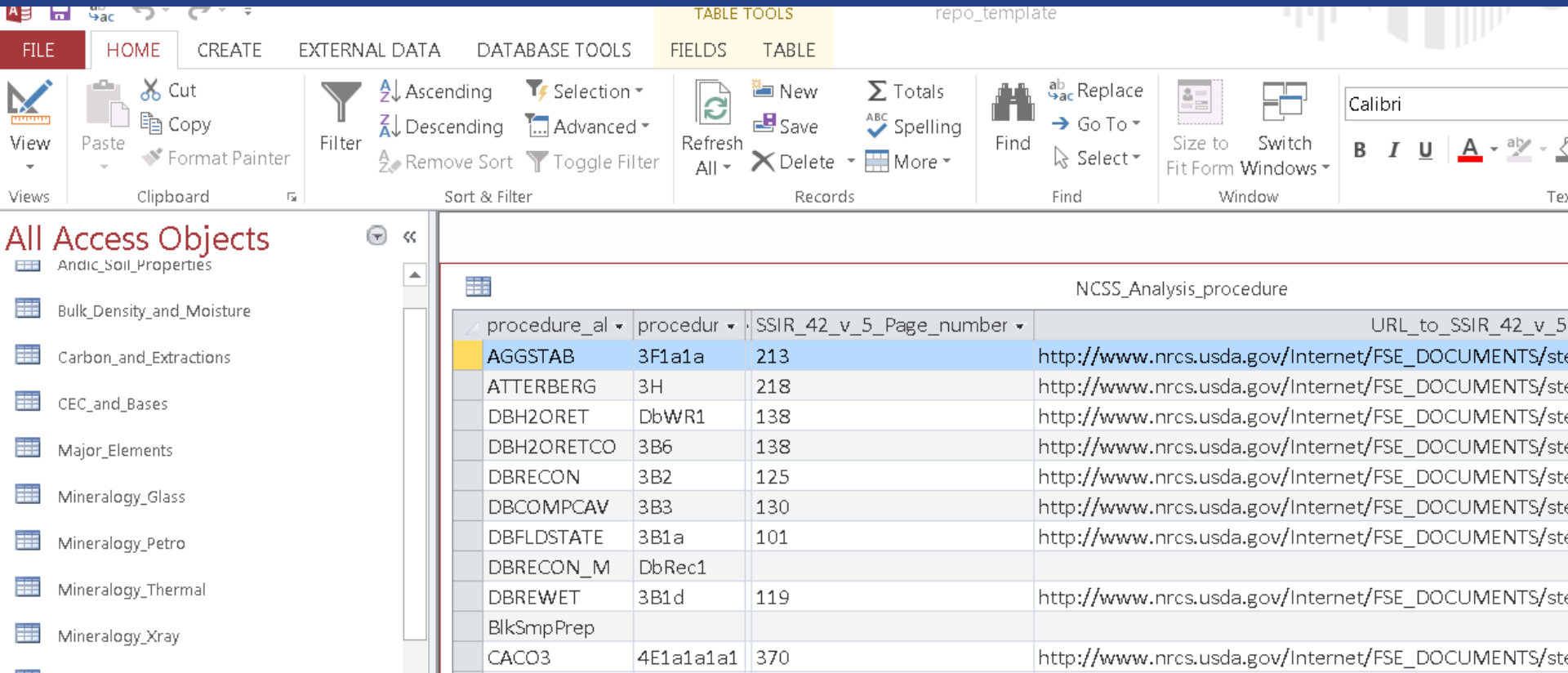
1. Misuse of results due to misunderstandings regarding methods used.
2. Confusion regarding joins between tables
3. Inability to match each result with a method
4. Desire for a Web Service
5. Need for better documentation of the database distribution product

What do the tables look like?

natural_key	result_sourc	prep_code	texture	particle_size_method	clay_total	silt_total	sand_total
MU043427	15774230	S	sicl	3A1a1a	39.100	55.000	5.900
MU059151	15795032	S	sil	3A1a1a	26.200	67.900	5.900
95P00399	137160	S	fsl	3A1a1a	18.000	26.600	55.400
04N02790	193378	S	s	3A1a1a	0.200	2.500	97.300
MU055247	15803842	S	sicl	3A1a1a	27.200	59.400	13.400
80P02432	56045	S	l	3A1a1a	16.800	34.600	48.600
89P00464	98081	S	sil	3A1a1a	23.800	71.400	4.800
IL22039	15718811	S	sil	3A1a1a	20.000	73.000	5.000
IL29143	15724540	S	sil	3A1a1a	16.000	75.000	8.000
UAR0003460	15860724	S	c	hyd1	47.100	34.500	18.400
MU019497	15767535	S	C	3A1a1a	49.900	8.900	41.200

Method code columns have been added and column headers expanded from abbreviations to full text

NCSS_Analyte_Procedure table includes page reference and url to the Soil Laboratory Information Manual 42 Version 5



What are the disadvantages

natural_key	result_sourc	prep_code	texture	particle_size_method	clay_total	silt_total	sand_total
MU043427	15774230	S	sicl	3A1a1a	39.100	55.000	5.900
MU059151	15795032	S	sil	3A1a1a	26.200	67.900	5.900
95P00399	137160	S	fsl	3A1a1a	18.000	26.600	55.400
04N02790	193378	S	s	3A1a1a	0.200	2.500	97.300
MU055247	15803842	S	sicl	3A1a1a	27.200	59.400	13.400
80P02432	56045	S	l	3A1a1a	16.800	34.600	48.600
89P00464	98081	S	sil	3A1a1a	23.800	71.400	4.800
IL22039	15718811	S	sil	3A1a1a	20.000	73.000	5.000
IL29143	15724540	S	sil	3A1a1a	16.000	75.000	8.000
UAR0003460	15860724	S	c	hyd1	47.100	34.500	18.400
MU019497	15767535	S	C	3A1a1a	49.900	8.900	41.200

More columns (Very wide tables)

Very long column headers

Examples of Some Column Names

bulkdensity3rdbarforcalculations
Bulk_Density_Third_Bar_Less_Than_2_mm_for_Calculations

percentpassingnumber200sieve
Percent_Passing_Number_200_Sieve

waterretentiondifferencewholesoil
Water_Retention_Difference_Whole_Soil

More Advantages:

Data that was never distributed before has been added to the tables.

Additional definitions have been added to the documentation to help with the understanding of the measured values.

The tables themselves can help with education as they are clearer

A closer look at the calculations table

The source of the 1/3 bar bulk density used for whole soil calculations is explicitly displayed.

bulk_density_3rd_bar_for_calc	bulk_density_3rd_bar_source
1.44	db_13b
1.47	db_13b
1.47	db_13b
1.68	db_13b
1.68	db_13b
1.81	db_13b
1.81	db_13b
1.74	db_13b
1.74	db_13b
1.45	default
1.45	default
1.45	default
1.45	default
1.45	default
1.45	default
1.45	default

1.47 is a measured 1/3 bar bulk density value

1.45 is a default value that was not measured

Some values may be estimated and are designated as such

Footnotes added to Supplemental Characterization Sheet Tier 2

Tier 2				-26-	-27-	-28-	-29-	-30-	-31-	-32-	-33-	-34-	-35-	-36-	-37-	-38-	-39-	-40-	-41-	-42-	-43-	-44-	-45-
				(---------- Weight Fractions ----------)														(---------- Weight Per Unit Volu				
				Whole Soil (mm)								<75 mm Fraction						Whole Soil				<2
Layer	Depth (cm)	Horz	Prep	>2	250 -UP	250 -75	75 -2	75 -20	20 -5	5 -2	<2	75 -2	75 -20	20 -5	5 -2	<2		Soil 33 kPa	Sur Oven -dry	Engineering Moist	Satur -ated	Soil 33 kPa	S 15 kP
				(---------- % of Whole Soil ----------)								(---- % of <75 mm ----)											
													3B1	3B1	3B1								
40A10064	0-18	Ap	S	–	–	–	–	–	–	–	100	–	–	–	–	100		1.45[1]					
40A10065	18-26	Bhs1	S	–	–	–	–	–	–	–	100	–	–	–	–	100		1.45[1]					
40A10066	26-36	Bhs2	S	–	–	–	–	–	–	–	100	–	–	–	–	100		1.45[1]					
40A10067	36-47	Bs1	S	–	–	–	–	–	–	–	100	–	–	–	–	100		1.45[1]					
40A10068	47-64	Bs2	S	–	–	–	–	–	–	–	100	–	–	–	–	100		1.45[1]					
40A10069	64-180	C	S	–	–	–	–	–	–	–	100	–	–	–	–	100		1.45[1]					

[1]The whole-soil bulk density was not measured. It was calculated using 1.45 g/cc as the 1/3-bar bulk density of the less-than-2-mm fraction for mineral soils.

The whole soil bulk density was not measured . It was calculated using 1.45 g/cc as the 1/3 bar bulk density of the less than 2 mm fraction for mineral soils.

The particle density is usually a default value of 2.65

particle_density_for_calc	particle_density_calc_sour
2.65	pd_g2
2.65	pd_g2
2.65	pd_g2
2.65	pd_g2
2.65	pd_g2
2.65	pd_g2
2.65	pd_g2
2.65	default
2.65	default

If a particle density was measured the source would indicate that it was measured.

Improvements made to the database:

The default value of 1.45 for 1/3 Bar Bulk Density was removed for most organic soils.

Calculation results using a default bulk density of 1.45 for organic soils were removed from the database.

This is a continuing maintenance issue …

For the Future!!!

I recommend that we consider the use of pedotransfer functions to make better estimates for 1/3 bar bulk density and recalculate appropriate results in the Calculations and Estimates Table when a measured value for 1/3 bar bulk density is lacking.

More Future Problems to Tackle

Coarse Fragments/Rock Fragments has been a major topic of discussion for many meetings.

How to solve the problem in the future

If the laboratory submission sheet is properly filled out and rock fragments are assigned to specific size classes, NASIS and the laboratory database can be reconciled.

What would be the process?

1. The field properly fills out the sample submission sheet.
2. Rock fragments are entered into the laboratory database
3. The volumes of rock fragments for classes greater than 75 mm are estimated in the field.
4. Rock fragments 20 mm to 75 mm are weighed in the field and discarded.
5. Fragments greater than 2 mm and less than 20 mm are weighed by the laboratory

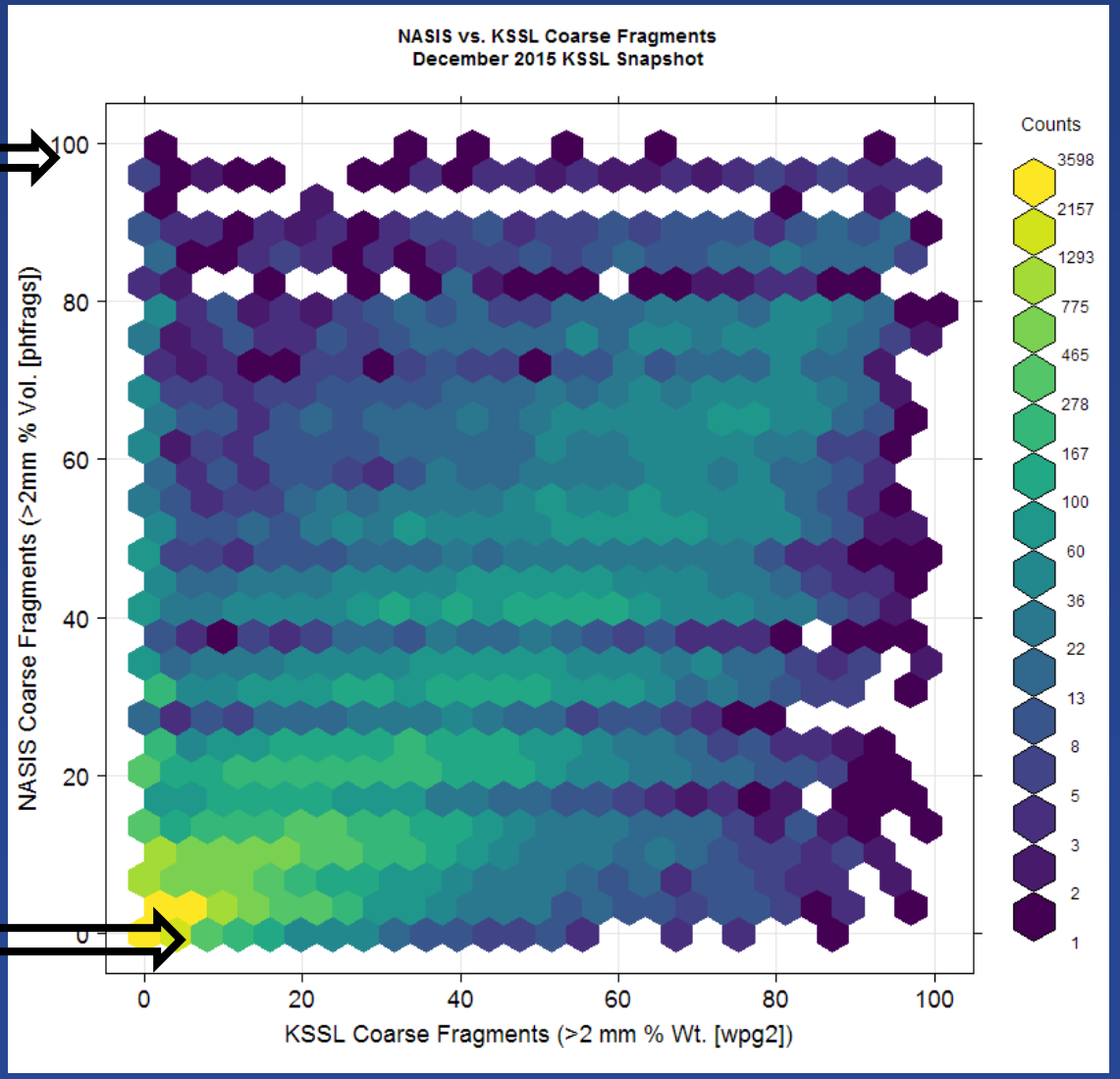
What would be the process? Cont.

6. The lab converts the rock fragments weights to percent weight on a 75 mm base.
7. Using the bulk density, the lab converts the rock fragment weights to volume on a 75 mm base and a whole soil base.
8. The volume of rock fragments from 2 to 75 mm are subtracted from the total rock fragment volume as entered in NASIS
9. The volumes of rock fragment classes > 75 mm are adjusted accordingly in NASIS and the laboratory database

The past/present reality as depicted in a histogram

NASIS estimates did not make it into the laboratory database

Laboratory Measurements did not make it into NASIS



How are the worst cases created:

1. The field relying upon the lab to measure the fragments and not entering any fragments into NASIS at all
2. The field entering fragments in NASIS using breaks that do not fit the laboratory data model, so no fragments made it into the laboratory database.
3. Sieved samples were submitted with no rock fragment information.

Why haven't these cases been fixed?

1. It is a matter of ownership.

2. No one at NSSC should be messing with descriptions. The state soil scientists mandated that the "sampled as" or "original" pedon description should be maintained as is.

My recommendation

1. The Sampled As pedons (KSSL owned) pedons should remain as is

2. The copies of the KSSL pedons could be updated based upon the 9 steps previously described.

3. Data mining could be used to extract the revised rock fragment volumes from NASIS and then update the calculations table.

Realistically

1. 90 percent of the estimates are probably good enough

2. The most important pedons to catch are those with 0 for fragments in the laboratory and a large estimate in NASIS.

3. Or a measured value in the laboratory data and 0 in NASIS.

Why hasn't the data mining taken place already?

1. Sometimes NASIS has duplicate rock fragment data amounting to 150 or 200% estimates

2. The first step is awareness, to educate for the future.

3. The next step would be to data mine once pedons were properly certified

Other issues with the laboratory database

1. There are some “impossible results” in the database.
    i. Negative values due to detection limits of instruments.
    ii. Some values that do not match the common units such as peak heights.
    iii. Typos

In the past researchers ran scripts against the database to "clean it up" before they used it.

1. The goal is to clean the database so that it is suitable to distribute to individuals that are not familiar with the cleaning techniques previously used.
    i. Examples would be to convert all Base Saturation values > 100 to 100 %
    ii. Convert all negative values to 0's or null

In the NCSS Soil Characterization Website scripts clean the data

1. We need to use some of the same scripts to clean the exported database.
    i. Users of the website are given the choice to convert TR and "-" dashes to null or zero
    ii. Currently we have allowed the negative values to be exported and delivered to the public, that policy may remain or it may be changed.

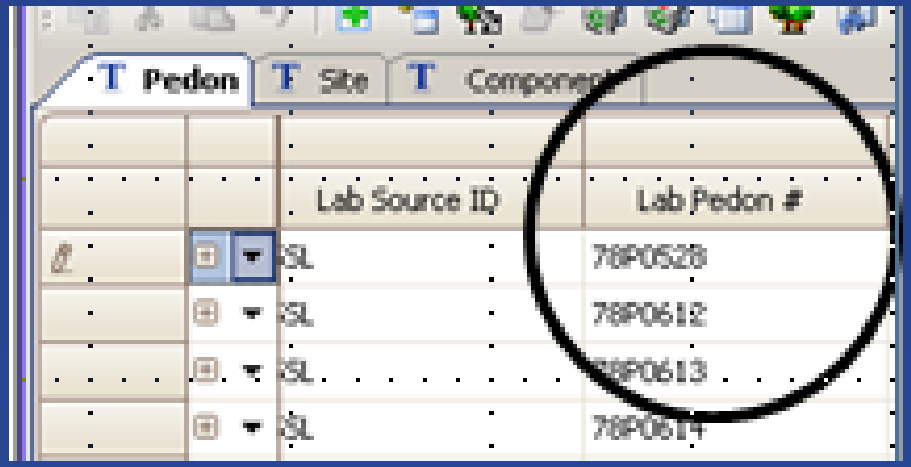
Some history related to recent changes

Prior to 2011 the report "SSL WEB Pedon Description (landscape)" delivered all pedon descriptions from NASIS with a given pedon sample number.

In 2011 the report was changed to deliver a single pedon description report – preferentially choosing the RO pedon over the KSSL owned pedon

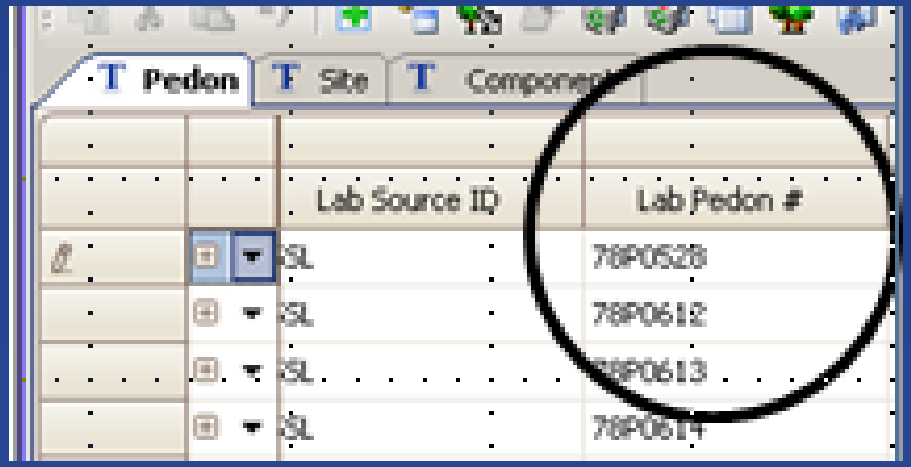
This increased the importance of the RO-owned pedon

The pedon sample number has to be in the pedon table for the report to find the RO-owned description.



This increased the importance of the RO-owned pedon

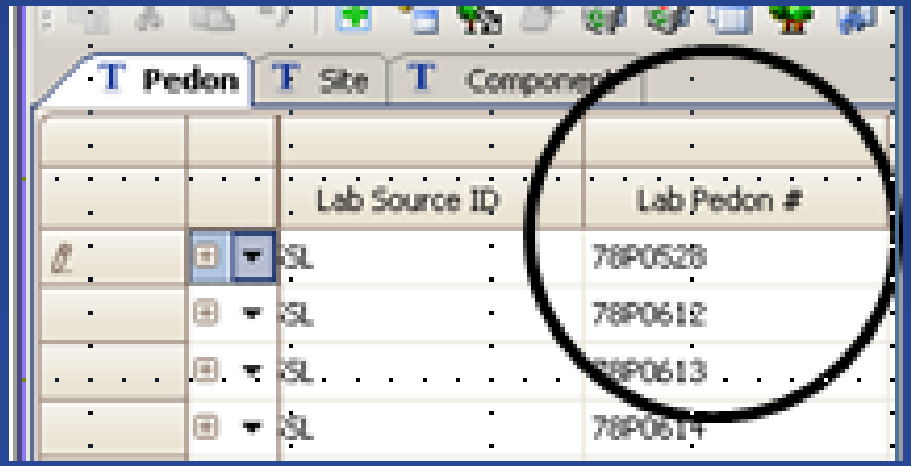
If the pedon sample number is in the wrong pedon, then the wrong pedon will be associated with the laboratory data!!



SDJR and challenges using NASIS queries - Do no harm!!

The laboratory database is “hidden” of sorts

User pedon ids were created in NASIS independently of the laboratory database



If we distribute morphological data

The horizon sample numbers have to be in the pedon horizon sample table or users can not join the laboratory data to the horizons.

Pedon Horizon Sample

Lab Sample #	Field Sample ID
95P03521	S1995MN137537-1
95P03522	S1995MN137537-2
95P03523	S1995MN137537-3
95P03524	S1995MN137537-4

Challenges to populating the horizon sample numbers

Identical pedons with different user pedon ids

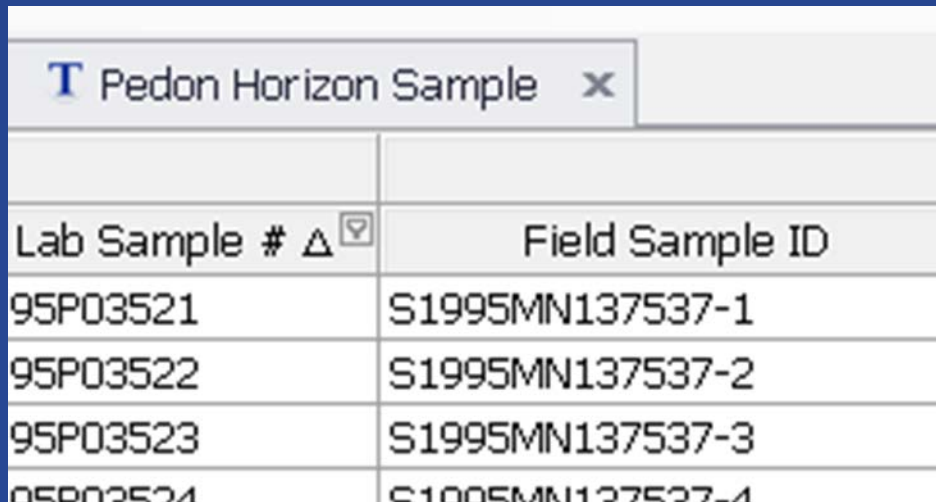
Different pedons with identical user pedon ids

Horizons with different horizon designations

Horizons with different depths

The NASIS interface

Forms will help



What Steps Does it Take to Ensure Sample Numbers are 100% Correct in the Future?

1. The field soil scientist inputs the pedon and site information into NASIS
   a) That includes the Field Sample Numbers and Sample Depths from the Sample Bags

Next Steps

2. Change ownership of the Site to the RO group named KSSL_Sites_Rox
3. Change ownership of the Pedon to the KSSL_PedonX group

Why?

1. Assigning sites to the KSSL_Sites_Rox group gives local ownership and control to the sites.
2. Assigning ownership to the KSSL_Pedons groups allows KSSL to add the laboratory sample numbers to the pedon table and pedon horizon sample table.

Why?

3. Management has decided that it is important to maintain an original sampled as pedon in NASIS.
4. In reality are there any/many original sampled as pedons in NASIS?

a) Probably not, except for the most recent pedons submitted.

Why?

1. Older pedons were described using English units which have been converted to metric.
2. Depth to the top of the soil was changed.
3. Over the years horizon designations have been modified to keep up with standards.
4. Many attributes were previously recorded as classes and are stored in NASIS as numbers (roots/pores/features)
5. Redoximorphic features were added and many mottles were reclassified

6. Some pedons were “correlated/updated” before they were ever entered into NASIS

7. Prior to 2010 there was no taxonomic history table in NASIS so if classifications were updated the previous classification was overwritten.

Recommendations/actions

1. The taxonomic history has been maintained in the laboratory database and much of it has been imported into NASIS.

2. The NCSS Lab Layer table was added to NASIS.

The ability to store two horizon designations per sample is possible in NASIS

NCSS Layer Lab Data | Pedon

Lab Sample #	Top De...	Bottom Depth	Layer Type	Designation	Designation - Orig
09N04172	0	5	horizon	E	E
09N04173	5	15	horizon	A	A
09N04174	15	48	horizon	Btk1	Btn1
09N04175	48	88	horizon	Btk2	Btn2
09N04176	88	175	horizon	C	C
09N04409	0	2	horizon	A	A
09N04410	2	6	horizon	A	A
09N04411	6	21	horizon	Btk	Btk
09N04412	21	36	horizon	Btk	Btk
09N04413	36	40	horizon	BCky	BCky

Currently the website does not necessarily reflect the horizon designations as populated in NASIS

pe	Designation	Designation - Orig
	E	E
	A	A
	Btk1	Btn1
	Btk2	Btn2
	C	C
	A	A
	A	A
	Btk	Btk
	Btk	Btk
	BCky	BCky

From the Website

Layer	Horizon	Orig Hzn	Depth (cm
09N04172	E	E	0-5
09N04173	A	A	5-15
09N04174	Btk1	Btk1	15-48
09N04175	Btk2	Btk2	48-88
09N04176	C	C	88-175

We have Scanned Project files that can be used for QA/QC

*** Primary Characterization Data ***

(Tooele County, Utah)

Pedon ID: S09UT045002

Sampled As : Skumpah Fine-silty, mixed, active, mesic Typic Natrargid

USDA-NRCS-NSSC-National Soil Survey Laboratory ; Pedon No. 09N1057

CEC & Bases				-1-	-2-	-3-	-4-	-5-	-6-	-7-	-8-	-9-	-10-
				(- - - - - - NH_4OAC Extractable Bases - - - - -)								CEC8	CEC7
								Sum	Acid-	Extr	KCl	Sum	NH_4
	Depth			Ca	Mg	Na	K	Bases	ity	Al	Mn	Cats	OAC
Layer	(cm)	Horz	Prep	(- - - - - - - - - - - - - - - - - cmol(+) kg^{-1} - - - - - - - - - - - - - - - - - - -)							mg kg^{-1}	(- - - - cmol(+) k	
				4B1a1a	4B1a1a	4B1a1a	4B1a1a						4B1a1
09N04172	0-5	E	S	49.1*	3.5	9.8	3.2	65.6					7.6
09N04173	5-15	A	S	52.5*	5.9	12.1	3.4	73.9					11.1
09N04174	15-48	Btk1	S	47.8*	6.7	17.6	4.0	76.1					14.3
09N04175	48-88	Btk2	S	54.5*	8.2	20.7	3.5	86.9					15.3
09N04176	88-175	C	S	51.2*	6.1	34.1	2.7	94.1					15.5

Today I updated the NCSS Lab Layer Table to reflect the current state of the database

*** Primary Characterization Data ***

(Tooele County, Utah)

Pedon ID: S09UT045002

Sampled As : Skumpah — Fine-silty, mixed, active, mesic Typic Natrargid

USDA-NRCS-NSSC-National Soil Survey Laboratory ; Pedon No. 09N1057

CEC & Bases -1- -2- -3- -4- -5- -6- -7- -8- -9- -10-

(- - - - - - NH_4OAC Extractable Bases - - - - -) CEC8 CEC7

Layer	Depth (cm)	Horz
09N04172	0-5	E
09N04173	5-15	A
09N04174	15-48	Btk1
09N04175	48-88	Btk2
09N04176	88-175	C

NCSS Layer Lab Data | Pedon | Local Database | Pedon Text

Δ	Bottom Depth	Layer Type	Designation	Designation - Orig
0	5	horizon	E	E
5	15	horizon	A	A
15	48	horizon	Btn1	Btk1
48	88	horizon	Btn2	Btk2
88	175	horizon	C	C

How to modify the horizon designations and depths

Load your pedons into NASIS and run these reports

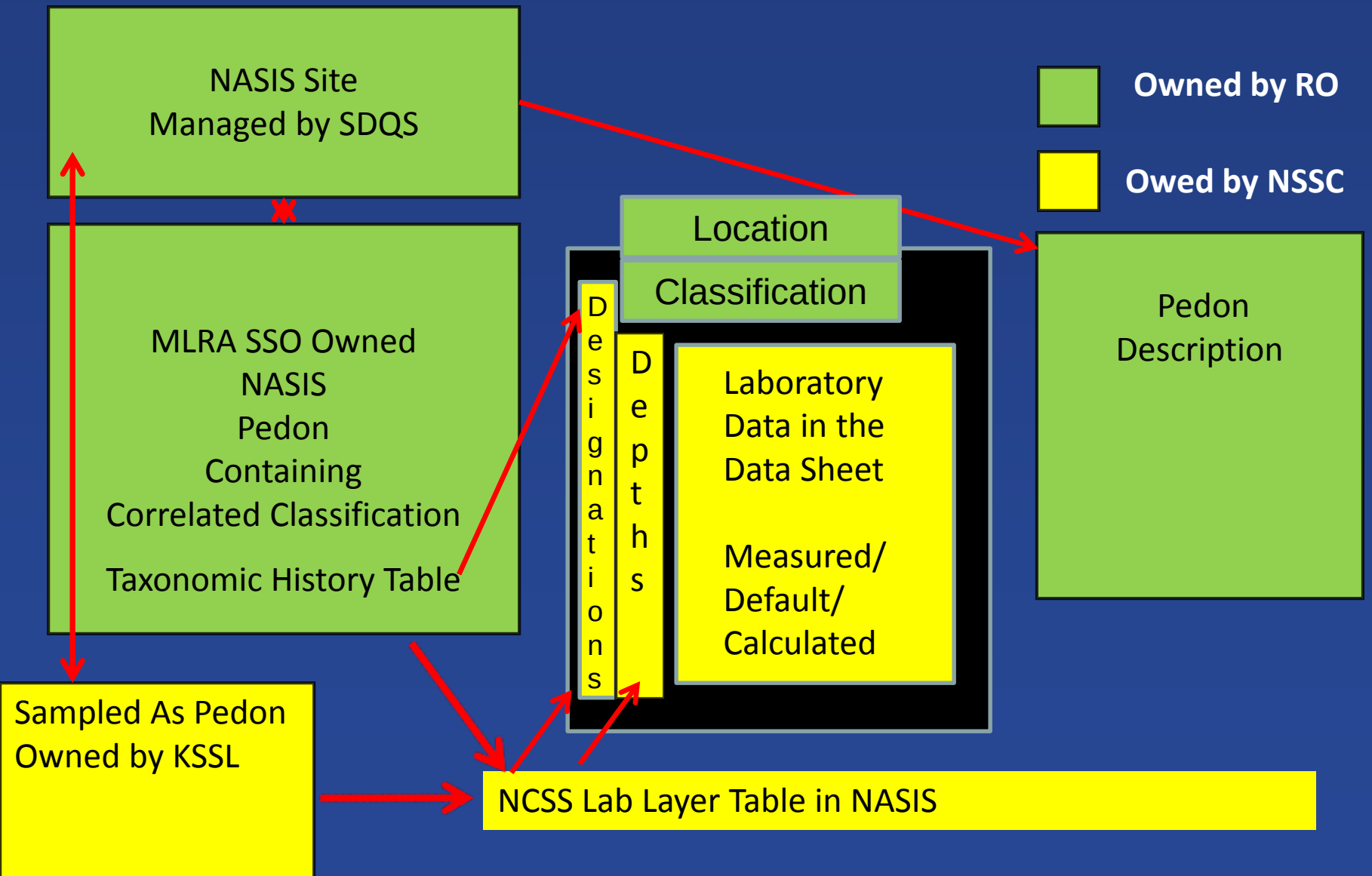

NCSS_Lab_Layer_hznamecurrent_update


NCSS_Lab_Layer_hznameoriginal_update


NCSS_Layer_laboratory_table_depths_update

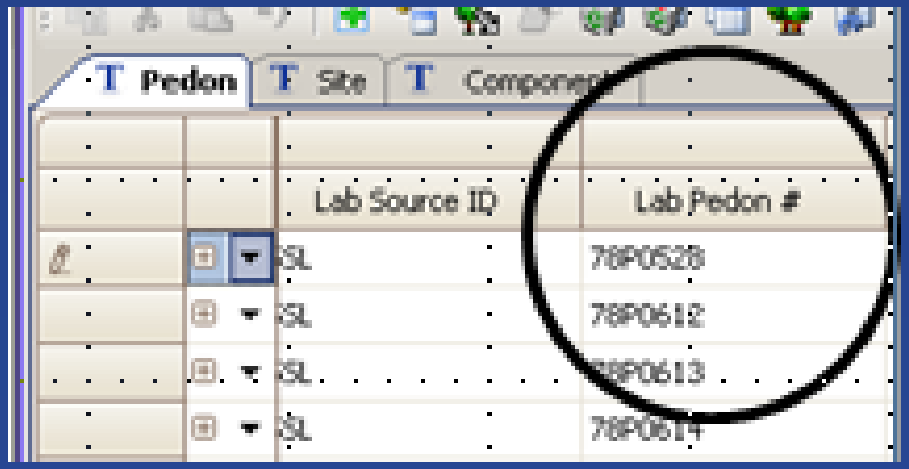
Send the results to the owner of the NCSS Lab Layer table in NASIS

The Current State of Affairs



Data Mining Reports in NASIS

- Site_Laboratory_Locations_Overlaps_Classifications
- Update_NCSS_Lab_Pedon_recordIDS_from_NASIS



How fast do the data mining reports work??

SDQS - Owned

Location

NASIS Site

Location

Laboratory
Data in the
Data Sheet

KSSL_Sites_MO_1
KSSL_Sites_MO_10
KSSL_Sites_MO_11
KSSL_Sites_MO_12
KSSL_Sites_MO_2
KSSL_Sites_MO_3
KSSL_Sites_MO_4
KSSL_Sites_MO_5
KSSL_Sites_MO_6
KSSL_Sites_MO_7
KSSL_Sites_MO_8
KSSL_Sites_MO_9

Location

Pedon
Description

You can immediately see a location change in the pedon description report.

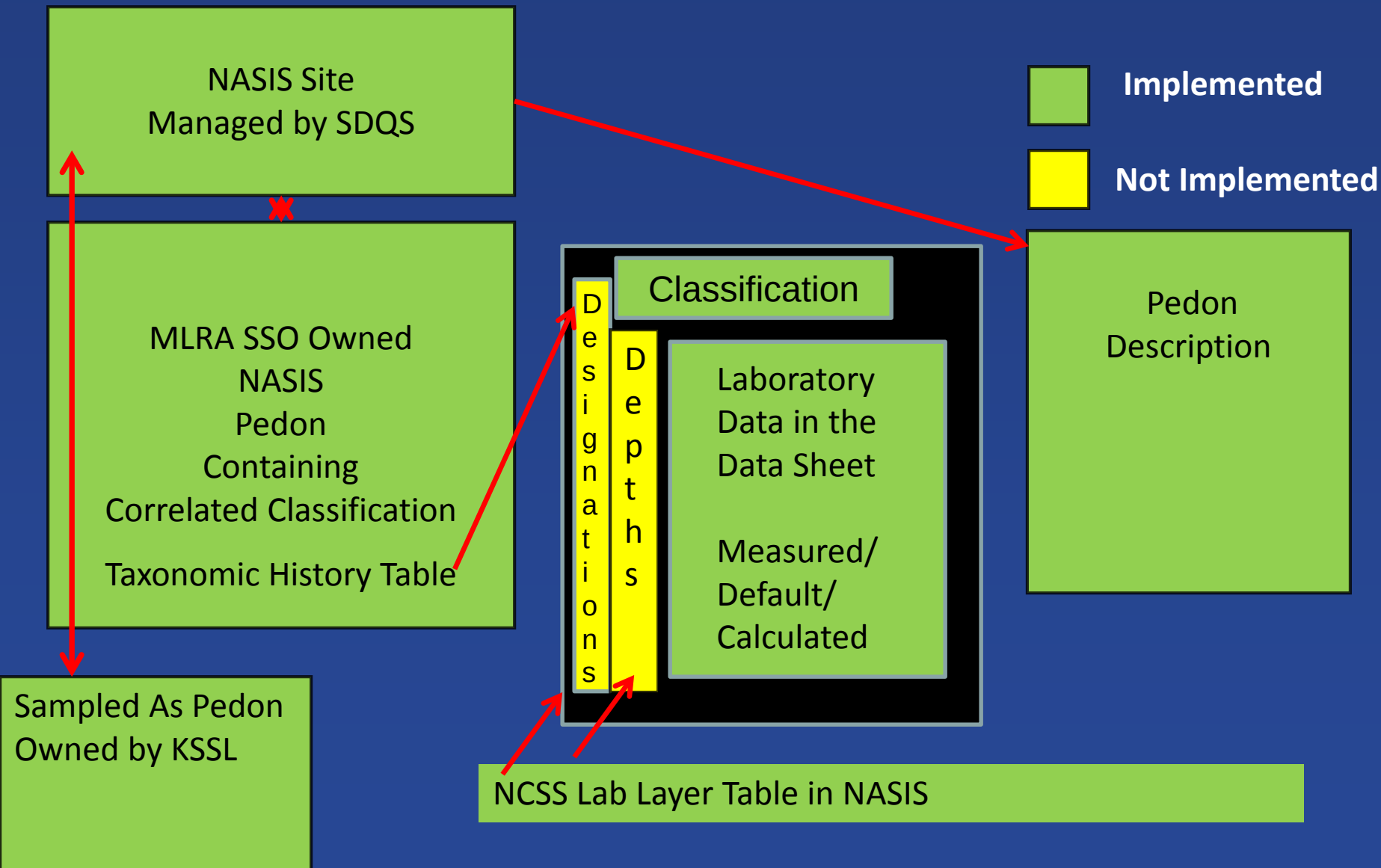
How Fast Does the location Get Updated In All of the Databases?

It can take from 3 months to a year for the updates to ripple through all of the products.

Why might a location fail to be updated?

The update process includes QC checks against the classification table. If the parts of the classification do not match the calculated classification name the update fails. That includes the location as well as the classification.

What was the original vision?



What is the new vision?

- 1. Scanned project files linked the the appropriate data and available through the Enterprise Content Management System
- 2. Method codes associated with each piece of data and linked to a description of the methods manual

What is the new vision?

3. Morphological database available and easily linked to the laboratory database.
4. A cleaner database that has been scrubbed of values that do not make sense.
5. An easier way to upload university data and correct data that has previously been uploaded if errors are identified

What kind of tools are available to help people use the databases

Calculations to help update NASIS

1. Load Pedon Laboratory Numbers from a Text File
2. Load Seq No Field and Lab Sample Numbers and depths Append
3. Pedon Horizon Designation from a text file

What kind of tools are available to help people use the databases

Calculations to help update NASIS

1. Update NCSS Pedon Lab Data Peiidref
2. Update Particle Size Control Section from a Text File
3. LE 0-100 cm
4. Wt. Ave Wt. Fraction 0.1 to 75mm &Vol >2mm
5. Wt. Ave – Noncarbclay & clay
6. Wt. Ave – CEC7 to clay ratio

What kind of tools are available to help people use the databases

Calculations to help update NASIS

1. Load Seq No Field and Lab Sample Numbers and depths Append
2. Update current horizon des NCSS Layer Laboratory Data Table
3. Update original horizon des NCSS Layer Laboratory Data Table
4. Update sample depths in NCSS Layer Laboratory Data Table

Reports and Queries

Queries to load NCSS Lab Pedon table in increments of 5,000 records (hopefully obsolete)

Analysis PC download (hopefully obsolete)

However it is also used to upload university laboratory data so do not delete it yet.

Many queries to load lists of sites and pedons by user site id, siteiid, user pedon id, peiid, horizon record id, pedon sample number, horizon sample number…

Ask NSSC or your local RO for Assistance

- Questions?