From:	Paul Benedict
To:	Beaudette, Dylan - NRCS, Sonora, CA; Wood, Jennifer - NRCS, Davis, CA; Monger, Curtis - NRCS, Lincoln, NE
Cc:	Scheffe, Kenneth - NRCS, Lincoln, NE
Subject:	RE: RV Defnition NSSH Proposal_171117 - PB comments
Date:	Friday, December 22, 2017 12:49:25 PM

Hi Folks,

Sorry I didn't respond earlier. I fully understand Jennifer's emphasis that repopulation of legacy values is not needed or wanted. I also share Ken's concern that we need to be aware of how data was populated, prior to having NASIS.

Most of my field career was in the days of SOI Form 5s and Form 6s. We were so limited in our soils data. It wasn't just database limitations, it was the constantly changing philosophy of soil survey from year to year and how it was interpreted from state to state. I remember being very frustrated in my early years of mapping (1978 to 1980 or so) when some of my mapping, where I had found significant differences to existing map units in the legend, were correlated into oblivion by a regional correlator who had never been in the county before. This might have been done because we didn't have the 800 acres to justify a taxadjunct or 2,000 acres to justify a series, or because somebody thought "yeah, it's different, but too similar to another series to be meaningful for the current land use". My field notes were put into the manila folder for the particular map unit and probably have not been looked at since, even if they still exist. Taxonomic puritanism was very strong in those days. Sometimes we had a policy of cutting off ranges at class limits and sometimes we didn't, and how they did it in Arkansas might have nothing to do with how they did it in Idaho.

Correlation has always been more of an art than a science and has led to subjective and sometimes biased interpretations. I mapped in four states and they all did it differently. Some people, who I will keep anonymous, felt concepts were more important than actual measurements. I just discovered that some of these biases still exist after all these years. In Kern County, California, Southwest Part (CA691) my crew mapped thousands of acres of Hesperia. The series was setup in the Mojave Desert (elevations above 2000 feet), but because our soil classified the same and at that time we only looked at mean annual soil temperatures, and didn't worry about differences between summer and winter, we called our soil at 300 feet elevation in the San Joaquin Valley Hesperia too. Hesperia is an aridic soil bordering on xeric, so our correlator wanted to apply his concept of xeric intergrades to our area and he made us remove any mention of calcium carbonate in the surface layer for the whole survey area, even though every single auger note and profile description showed slight effervescence when treated with HCl. His concept that slightly increased precipitation in areas bordering on xeric should leach carbonates out of the surface and overruled the facts. He ignored that dried lake beds where we had a wind erosion c factor of 300 provided a ready source of carbonates. I just looked on Web Soil Survey and it still shows 0% $CaCO_3$ equivalent in the surface of Hesperia soils in Southwest Kern County.

I also collected monthly soil moisture and temperature data in the hills of SW Kern for three years. My data showed that north facing slopes were xeric and mesic and south facing soils were aridic and thermic. It wasn't even close. They were completely different. Nearly every month of the year south facing slopes were 10 degrees F warmer than north facing slopes. North slopes had oak forests,

south slopes had sparse yucca, cacti, and small desert shrubs. That same correlator chose to ignore one of the primary soil forming factors (topography) and told me three years of data was meaningless and he wouldn't allow associations with two moisture regimes and two temperature regimes. These were just concepts and I needed to decide which way to go. Well, that mistake still exists in the database. Map unit 403 Loslobos-Calleguas association, 30 to 100 percent slopes is a good example. It is shown to be Xeric and Thermic now, but that is incorrect. If you read Soil Taxonomy it doesn't say moisture and temperature regimes are intended to be conceptual. In Fact, it gives excruciating detail on how to determine them.

When SSSD was converted to NASIS a lot of assumptions were made, some of them wrong, and some still not fixed.

I like the statistical approach of your proposal, but knowing what I know about existing data it implies an accuracy that does not exist, at least in old mapping.

On that thought, I will shut up. Merry Christmas if you celebrate that holiday and Happy New Year to all!

Thank you,
Paul Benedict
USDA-NRCS Contractor
202-578-8226
soilwarden@verizon.net

From: Beaudette, Dylan - NRCS, Sonora, CA [mailto:Dylan.Beaudette@ca.usda.gov]
Sent: Wednesday, December 20, 2017 1:03 PM
To: Wood, Jennifer - NRCS, Davis, CA <Jennifer.Wood@ca.usda.gov>; Monger, Curtis - NRCS, Lincoln, NE <Curtis.Monger@lin.usda.gov>
Cc: Scheffe, Kenneth - NRCS, Lincoln, NE <Kenneth.Scheffe@lin.usda.gov>; Paul Benedict <soilwarden@verizon.net>
Subject: RE: RV Defnition NSSH Proposal_171117 - PB comments

Thanks everyone for the great commentary. I am going to integrate my comments with Jennifer's into a single document to keep the record as simple as possible.

Dylan

From: Wood, Jennifer - NRCS, Davis, CA
Sent: Tuesday, December 19, 2017 1:44 PM
To: Monger, Curtis - NRCS, Lincoln, NE <Curtis.Monger@lin.usda.gov>

Cc: Beaudette, Dylan - NRCS, Sonora, CA <Dylan.Beaudette@ca.usda.gov>; Scheffe, Kenneth - NRCS, Lincoln, NE <Kenneth.Scheffe@lin.usda.gov>; Paul Benedict <soilwarden@verizon.net>
Subject: RE: RV Defnition NSSH Proposal_171117 - PB comments

Dear Curtis, Ken, and Paul,

I thought I would also write separately, to work through something I think we want to communicate. Either we are off-base or not communicating clearly – equally likely to be true!

We want to say that there is no need to specifically repopulate legacy values to accord with this proposed, approximate definition, because they already do. Also want to communication that it is not necessary to computationally derive values from a lot of data in order to populated the database with this approach in mind. I believe that how we have populated l-rv-h values in the past is actually pretty close, philosophically, to the proposed approach - even if we were eyeballing data from individual pedons and field notes:

We used values from the observed data set
We picked rv valuse that generally fell within the middle of the ranked values observed
We didn't use the full range, when very high and very low values were not commonly observed.

This is pretty close to the percentile approach.

To further include legacy data under the umbrella of the percentile approach:
Even class limits could be common low/high values in a ranked set of observed values.
Values of 0 could easily end up as a 5th or 10th percentile value if that is a commonly observed value in the soil/on the landscape
Component data can always be subsets of mapunit values - field observations and knowledge help create custom ranges

Fun stuff right?

Jennifer Wood
Soil Data Quality Specialist
USDA-NRCS Pacific Soil Survey Region (SSR-2)
Office 530-792-5643 Cell 707-845-1216
Telework Mondays and Fridays
www.soils.usda.gov

From: Wood, Jennifer - NRCS, Davis, CA
Sent: Tuesday, December 19, 2017 12:41 PM
To: Monger, Curtis - NRCS, Lincoln, NE <Curtis.Monger@lin.usda.gov>
Cc: Beaudette, Dylan - NRCS, Sonora, CA <Dylan.Beaudette@ca.usda.gov>; Stiles, Cynthia - NRCS,

Davis, CA <Cynthia.Stiles@ca.usda.gov>; Tallyn, Ed - NRCS, Davis, CA <Ed.Tallyn@ca.usda.gov>; Scheffe, Kenneth - NRCS, Lincoln, NE <Kenneth.Scheffe@lin.usda.gov>; Paul Benedict <soilwarden@verizon.net>
Subject: RE: RV Defnition NSSH Proposal_171117 - PB comments

Thank you Paul and Ken,

Those are great comments, including ones that had been made in Boise that I had neglected to record and accommodate.

There were more people than just me and Dylan who put this together, so I am circulating this to them so we can respond to/modify the proposal.

I appreciate your time and insights.

Jennifer

Jennifer Wood
Soil Data Quality Specialist
USDA-NRCS Pacific Soil Survey Region (SSR-2)
Office 530-792-5643 Cell 707-845-1216
Telework Mondays and Fridays
www.soils.usda.gov

From: Monger, Curtis - NRCS, Lincoln, NE
Sent: Tuesday, December 19, 2017 11:28 AM
To: Wood, Jennifer - NRCS, Davis, CA <Jennifer.Wood@ca.usda.gov>
Cc: Beaudette, Dylan - NRCS, Sonora, CA <Dylan.Beaudette@ca.usda.gov>; Stiles, Cynthia - NRCS, Davis, CA <Cynthia.Stiles@ca.usda.gov>; Tallyn, Ed - NRCS, Davis, CA <Ed.Tallyn@ca.usda.gov>; Scheffe, Kenneth - NRCS, Lincoln, NE <Kenneth.Scheffe@lin.usda.gov>; Paul Benedict <soilwarden@verizon.net>
Subject: FW: RV Defnition NSSH Proposal_171117 - PB comments

Hello Jennifer,

Pasted below are comments by Ken Scheffe and Paul Benedict on the RV Definition proposal. Also see the attached document with Track Changes.

Please address the comments and send a revised proposal (or rebuttal). If you would like to discuss any of the comments by phone, please let me know. Once received, I will post the RV proposal for its 30-day comment period.

Thank you and let me know if you have any questions.

Curtis

Curtis Monger
National Leader for Soil Survey Standards
USDA-NRCS National Soil Survey Center
100 Centennial Mall
Lincoln, NE 68508
O-402-437-4010
Professor Emeritus, New Mexico State University
www.soils.usda.gov

ALSO, need to remember we populated map unit component data, which may have been wider, or narrower than the series named in the map unit. Similarly, there were certainly real world occurrences, for example, where we had a soil with a fine-loamy argillic (upper 50cm for the family class) that very often was a sandy clay loam, however, it was known that on occasion in the 'real world' there were layers in the argillic that were actually heavy fine sandy loams at 18-20% clay. This is still in the fine-loamy family, but we might have only entered the weighted average for he control section (sandy clay loam) to avoid potential confusion. When the clay 'low value' was assessed and entered, we probably used 20% for the sandy clay loam class limit.

As you stated, similar situations and accommodations were likely made for all the entries of low, representative value, and high. And also, sometimes the value of zero is appropriate for such things as salinity in leached soils, slope on level soils, etc. I don't know that using the 5 or 10 percentile is appropriate when values of zero real. You also mentioned in your track changes the need to avoid creating 'gaps' between classes if we restricted the current class limits to the 5th and 95th percentiles of that class.

Perhaps this is not so easy to do with existing data for (l, rv, h) as some would like to think.

Kenneth F. Scheffe
National Soil Classification Specialist (Soil Taxonomy)
USDA-NRCS-NSSC-Standards
(402) 437-5698 (O)
(402) 318-1149 (M)

From: Paul Benedict [mailto:soilwarden@verizon.net]
Sent: Tuesday, December 19, 2017 11:10 AM
To: Scheffe, Kenneth - NRCS, Lincoln, NE <Kenneth.Scheffe@lin.usda.gov>
Subject: RE: New RV definition proposal for NSSH

Yes, your points were what I was thinking (although not in the detail you used) when I said I wondered if the same rules need to apply to all data elements. Sometimes for example, the lower value really should be a zero, even if the 10th percentile is a 0.1.

Paul

From: Scheffe, Kenneth - NRCS, Lincoln, NE
Sent: Tuesday, December 19, 2017 9:55 AM
To: Monger, Curtis - NRCS, Lincoln, NE <Curtis.Monger@lin.usda.gov>
Subject: RE: New RV definition proposal for NSSH

Good morning Curtis,

I read the proposal by Dylan and Jennifer regarding the population of low, rv, and high values in the NASIS database using percentiles. It is a good proposal and would allow our scientists to increase the accuracy/precision in the database. I believe this is the correct way to move forward, but I have a few caveats and perhaps warnings that need to be considered.

The original population for low, rv, high values in NASIS most often was not with observed values, but actually represented the numeric limits of descriptive classes for the property or feature. In some circumstances, the low and high actually represented the limits spanning two adjacent classes (soil reaction or pH was one of these).

For example, if the texture was recorded in SSSD or the Soils-5 as 'sandy clay loam' the population of low, and high values for clay percentage in NASIS would have been low=20%, high=35% --- the class limits for clay for the texture 'SCL'. We were permitted to modify the range with observed limits and the rv was to be the most commonly observed value based upon the experience or other supporting data. However, most often, there was no one with field experience, nor time, nor desire to populate individual layers with tailored data. We were 'instructed' by National Bulletin to populate l, rv, h – just get it done because it was needed now in order to generate new calculated interpretations (every property, every horizon, every component, every map unit, every survey in the state – and do it this year!)

In NM, we were fully engaged in initial and update projects under reimbursable agreements during this time, so we had limited time/staff to research manuscripts, laboratory data, or field notes, so we developed routines to query for properties, such as texture = scl, and would populate in-mass using NASIS calculations, the low, rv, and high values at the class limits, and the rv as an approximation of the middle of the class. I don't believe NM was different than many other state, so much of the original population of l, rv, h in NASIS was by algorithms which generated class limits and a median value.

Not all is lost. It has been 15 years since this was done. Survey offices have done a lot of update work, and hopefully during this time, have adjusted the low, rv, and high values away from strict class limits with observations, lab data, and descriptive narratives gleaned from available sources to better approximate the real world. Kudos to those who tailored their data. The problem is knowing which values for low, rv, and high have been tailored and which have not. Perhaps the most telling thing would to be look for populated values that correspond to descriptive class limits as they likely have not been appropriately adjusted.

My concern, in addition to those expressed in the proposal by Dylan and Jennifer, as we go into this percentile representation of low, rv, and high would be a direct conversion, or simply a redefinition of data elements to the 5 or 10 percentiles and 90 or 95 percentiles. I don't believe it to be wise to simply redefine the terms low, representative value, and high to the percentiles without actually validating the numbers. It would give the impression of a more statistically evaluated entry, without actually being true. I didn't see anything in the proposal by Dylan and Jennifer that indicating an evaluation method to assure appropriateness or validity would be performed as part of the redefinition. I'm not worried about new data population based upon observations or measurements, but rather the legacy entries that simply represent the lower and upper limits of descriptive classes and the median value (rv).

I think we should move forward on the proposal, but we probably need to talk with Dylan and Jennifer to see how this would be done. It is not simply or only a redefinition of terms in the NSSH. I believe the only honest way to do this would be to design and effort or initiative to systematically evaluate whether the current entries represent class limits, or are tailored values based upon observations, measurements, and documentation. If the entries represent class limits, we need to evaluate and adjust them to more appropriate values which represent the reality on the ground for the appropriate percentile being used.

Kenneth F. Scheffe
National Soil Classification Specialist (Soil Taxonomy)
USDA-NRCS-NSSC-Standards
(402) 437-5698 (O)
(402) 318-1149 (M)

From: Paul Benedict [mailto:soilwarden@verizon.net]
Sent: Tuesday, December 19, 2017 10:30 AM
To: Monger, Curtis - NRCS, Lincoln, NE <Curtis.Monger@lin.usda.gov>; Scheffe, Kenneth - NRCS, Lincoln, NE <Kenneth.Scheffe@lin.usda.gov>
Subject: RV Defnition NSSH Proposal_171117 - PB comments

Hi Curtis,

I had some comments and suggestions on this proposal. A clear guide on low, RV and high values is probably needed, but I also wonder if the same rules need to apply to all data elements. I hope my comments don't come across as too negative as I know this issue is important to those that proposed it and they are the ones dealing with this issue in their daily work and not me.

Thank you,
Paul Benedict
USDA-NRCS Contractor
202-578-8226

soilwarden@verizon.net

This electronic message contains information generated by the USDA solely for the intended recipients. Any unauthorized interception of this message or the use or disclosure of the information it contains may violate the law and subject the violator to civil or criminal penalties. If you believe you have received this message in error, please notify the sender and delete the email immediately.