Basic Soil Survey Quality Assurance

Guiding Principles

1. All field estimated data in the soil survey database (texture, clay %, rock fragment %, depth to horizons, Ksat, LEP, pH, etc) must be consistent with field observations and lab data. Actual measured data (lab data) must not contradict the estimated data. Only modal or typical values, representing a central concept, are populated; data for anomalies, atypical soils, and similar soils are excluded.

2. There must be internal consistency among all data (particle-size data agrees with texture, rock fragment data agrees with textural modifier, etc.).

3. Soil properties that are derived from field-estimated data must be derived consistently and follow accepted guides (linear extensibility is consistent with clay content and clay mineralogy, available water is consistent with texture and bulk density, etc.).

4. All parts of a soil survey report must be consistent with each other. The database must be consistent with detailed map unit descriptions, general soil map unit descriptions, and taxonomic unit descriptions. Map unit component data must be within the allowed ranges of the taxonomic unit description. A user must never obtain data from one part of a report, and find it contradicted in another part.

5. All data displayed in taxonomic unit descriptions must agree with Soil Taxonomy and Official Series Descriptions. Exceptions are allowed for taxadjuncts and soils otherwise outside series ranges, but these should be documented in official correlation notes in NASIS.

A Systematic Process

For the above reasons, quality assurance is normally performed on a "package" of data and documents: the database is reviewed in conjunction with a review of related map unit descriptions, taxonomic unit descriptions, official series descriptions, and any available lab data. A systematic process should be followed: 1) the database and MUD are reviewed; 2) lab data is reviewed to ensure it is appropriately captured in the database and is consistent with the TUD; 3) the TUD is reviewed for technical accuracy and to ensure that it encompasses the MUD (DMU data); and 4) the OSD is reviewed to ensure it correctly encompasses the TUD. This type of review is not unique to the NASIS era; these same data checks have been the essence of progressive correlation for the past several decades. Because of the complex data interrelationships, one of these aforementioned documents should not be reviewed without the others. MUDs and TUDs must be generated from the database to ensure consistency.

Database Quality Assurance.

Quality Assurance is generally performed on a map unit (DMU) basis when the objective is map unit approval and progressive correlation. <u>All NASIS data</u> for a map unit is reviewed prior to the unit being designated as approved or correlated. This includes data for minor components.

1. <u>Data Review</u>. A check is made for data gaps, incorrectly populated data, and inconsistent/unreasonable data. An Edit Setup that displays only required data elements is useful. The use of a checklist of required data elements is strongly recommended. For each element on the checklist, major and minor component data in the subject DMU is checked to ensure the following:

   a) Required data is populated (the MO, state, or SSA minimum dataset. See MO6 Instruction 430-386 for minimum dataset for all MO6 soil surveys). An exception to this is the special list of required data elements for surface O horizons (duff layers); see MO6 Tech Note 9.

   b) The correct type of data is populated in accordance with the NSSH and MO Guidelines (correct units-of-measure are used, numbers rounded correctly, only class limits are

entered where required, values are calculated instead of manually entered where required (or vice versa) etc.)

c) Values for field-estimated or field-measured data are reasonable (ranges not too wide or too narrow, the RV is reasonable, data are consistent with a map unit concept and the series concept).

d) Values for derived soil properties are consistently derived from appropriate guides. Data inter-relationships are uniform and reasonable (strongly alkaline soils have elevated SAR values, diagnostic features [e.g. calcic horizon] are consistent with criterion data [CaCO3 content], parent material is consistent with landform, etc.).

2. NASIS Checking Tools. Numerous NASIS reports are available to display data for checking. These reports are particularly useful to display interrelated data side-by-side. Be aware it is impossible to find or write reports that will display all possible errors; nor is it possible to list here all reports that should be run to check a database and certify it as error-free. Many errors are uncovered simply by looking at data and thinking "is this correct and what other data must this be consistent with?" A review by a knowledgeable soil scientist who can examine data critically is the most effective manner to certify data; check reports, validations, and queries are merely NASIS tools to assist with this certification.

There are numerous reports available on the MO6 NASIS site, the Pangaea site, and other MO sites that are useful for checking data. Some reports are better suited to particular soils and populating techniques. Some of these reports display mainly the same data but in slightly different ways – choosing one report over the others is a matter of personal preference. Some reports are best suited for modern initial or update surveys, some are better suited to older correlated surveys with 3SD carryover data. Some reports are intended to check manually entered data and are not applicable for datasets where this data is calculated. Some reports work well when doing QA on an entire legend, but for QA on a single DMU it is often more efficient to visually observe the data elements rather than display or print a report. Some reports are useful to the less experienced individuals, but those reports may seem too elementary to the more experienced soil scientists.

So in short, there is no formally sanctioned list of reports that works best under all circumstances; each individual must develop a routine for checking data, using a set of reports that are best suited to the specifics of the dataset and personal preferences. Nevertheless some queries and reports stand out and are recommended for use in MO6:

The MO4 (Bozeman) Site has a number of queries with the prefix "CHECK:" that are valuable in selecting records with missing data or inconsistent data. Since only those records that contain gaps or errors are selected, global assign and global delete can often be used to fix the data.

The MO6 Site contains a number of reports, beginning with the prefix "CHECK". All are useful, some are particularly useful in that they comply with MO6 data population guides or display data in a manner that is not known to be duplicated in reports elsewhere:

CHECK AASHTO and Unified calculations vs database

CHECK CE Activity vs CEC and Clay

CHECK Carbonate Clay Estimation

CHECK Geomorphic Properties

The NSSC Pangaea Site has many data checking reports with the prefix "UTIL". New reports are added frequently, particularly to test draft calculations. One report, titled UTIL – Component Data Population Issues" is a quick listing of data voids in the nationally required minimum dataset. This is the closest thing to one report that checks for all data gaps.

Also on the Pangaea Site are reports with the prefix "VAL" (validations). These reports are copies of reports from various MO Sites that were copied to the Pangaea site because they were particularly useful.

3. NASIS Validations. Validations must be run, not only to validate data but to ensure that the field soil scientist has run them. Limitations of the validations must be well understood, as the current validations check for only a small fraction of the errors that are possible in a dataset. For example the particle-size validation does not check for all possible errors in particle-size data, and the Unified class validation checks only for errors of commission, not for errors of omission.

4. Scope of the QA. The QA that is performed is intensive and comprehensive for a new soil survey project or with MLRA Soil Survey Leaders or Subset Leaders with unproven abilities. As trust in the individual's competency is developed the QA review can become less comprehensive and may approach random spot-checking. However, some of the most complex and error-prone data elements may always warrant a thorough review. The field soil scientist's competency can be improved by using the QA as a training tool. For this reason, feedback to the field soil scientist is imperative. The SDQS must not act as a generous and accommodating "data fixer" for the field soil scientist. This may do more harm than good.

Map Unit Description Quality Assurance

Map unit descriptions (MUDs) must be generated from the database to ensure consistency. Displaying the MO6 MUD report with nulls as "unspecified" will flag possible data voids. There is usually very little data in a MUD that is manually entered; for that reason a fully populated dataset will produce a fully described and accurate MUD. A review of the MUD therefore involves more of an evaluation of the map unit concept and how clearly it is portrayed to the user. Nevertheless, in doing such a review data values should be evaluated critically; values that look perfectly reasonable in a database may actually be incorrect when evaluated in the context of a description of a map unit. Review of map unit descriptions is performed with the following principles in mind:

a) The map unit concept is meaningful and valid for the intensity of the survey and for the projected use and management of the land. Map unit design complies with decisions made during progress reviews and other guidance. The map unit is clearly different from other map units on the legend that have been approved.

b) Major and minor components are dissimilar soils. The soil properties that make the components dissimilar should be apparent in the component descriptions. The composition of components agrees with NSSH specifications for kinds of map units and allowable inclusions.

c) A user can clearly understand the physiographic characteristics of the map unit. The pattern of components on the landscape is clearly portrayed. Both major and minor components are clearly located on landforms and/or on landform positions, or are otherwise distinguished by slope gradient, slope aspect, or slope shape.

d). Miscellaneous land types are briefly described in a short narrative.

Taxonomic Unit Quality Assurance

During approval of a map unit, it must be verified that the map unit component data fits within, and agrees with, the taxonomic unit description (TUD). The TUD is an aggregation of the individual component data for all map units that contain the taxon. Data ranges for a map unit component may be narrower than the ranges of the TUD, but must not be wider. These data relationships are ensured when the TUD is generated from the database as much as possible.

The taxonomic unit description (TUD) is also reviewed for consistency with the NCSS soil classification system (Soil Taxonomy and official series descriptions). The ranges in characteristics must clearly portray a "series concept". All data in the TUD must be internally consistent. Unlike the MUD, the TUD contains a number of manually-entered data, which requires manual cross-checking. The most common required cross checks are listed below.

Pedon Description

Format and Internal consistency check *within* pedon description

(Some of the following checks will not be necessary if the pedon description was generated form NASIS pedon.)

Format is correct (punctuation, horizon first-line indent)

Horizon properties described in correct sequence

Rock fragment modifiers agree with rock fragment percentages

Numerical pH is within reaction range (i.e. if neutral, pH value is 6.6-7.3)

Horizon symbols are correct

Description consistent with taxonomic class (Classifies correctly)

Consistency *between* Pedon Description and Range in Characteristics

The following typical pedon properties are within RIC ranges:

- Colors
- Texture
- Rock frags
- Carbonates
- Reaction
- Depth to/Thickness of taxonomic features

Range in Characteristics

Evaluate the horizon headings (layers that are diagnostic for the series) within the Range in Characteristics.

Layers given (horizon headings) clearly portray the series concept; transitional horizons are ranged only when significant to the series concept.

Subhorizons are grouped ("Bt horizon" for Bt1…Bt5) as long as ranges in properties are narrow; when range is too wide subhorizons are split (Bt horizon, upper part; Bt horizon, lower part).

Diagnostic layers are generally equivalent to diagnostic horizons (the A horizon is not ranged to include both mollic and non-mollic colors; a Bk horizon is not ranged to include calcic and non-calcic carbonate content.

Horizons are combined when their properties are the same ("Bw and BC horizons").

Verify the Internal Consistency of the Range in Characteristics

All textures in RIC agree with the % clay, %sand given;

("clay 10-20%" and "sandy loam" is incorrect- either cut off clay at 18% or add a texture [sandy clay loam or loam]; requires NASIS edits).

All textural modifiers in RIC agree with the given RIC rock fragment ranges

(if "gravelly", total fragment content is >15%, there is >2 times as much gravel as cobble, etc); (only "gravelly' and 15-40% total fragments is incorrect, either cut off fragments at 35% or add "very gravelly".

Verify Agreement Between Range in Characteristics and Taxonomic Class

The weighted average values for clay content, rock fragment content, and sand content are within the limits for the particle-size family. (Individual horizons within the PSCS may be outside the family limits, as long as the weighted average is within the limits.)

Organic matter range is valid for mollic, umbric

Carbonate content consistent with calcic or non-calcic, or with carbonatic mineralogy

Gypsum content is consistent with gypsic or non-gypsic

Sodium content (as ESP and/or SAR) supports natrics and Sodic subgroups.

Depth to diagnostic features (e.g. depth to argillic, calcic, petrocalcic) agrees with taxonomic class; entire range is within the class limits)

Colors given agree with mollic epipedon, albic horizon, Ustollic subgroups, etc. (color range for dry value cannot be "4 to 6" for a Mollisol.)
In general, all soil properties support and do not contradict the diagnostic horizons and the taxonomic class.

Official Series Description Quality Assurance

The content and format of the Official Series Description must be evaluated against specifications given in the NSSH, MO6 Instruction 430-81, and the MO6 OSD template.

All consistency cross-checks discussed above for taxonomic unit descriptions must be similarly performed on the OSD. In addition the following must be verified:

Information given for the typical pedon setting (slope, elevation, landform, vegetation) must agree with information given in the Geographic Setting section.

The diagnostic horizons and features given in the Remarks section must agree with the typical pedon and the Range in Characteristics.

Proposed revisions to an existing OSD must be closely examined. When ranges are expanded (clay content, parent material, climate data, horizon thicknesses, etc.) revisions are acceptable only if the changes are within the series concept and do not conflict or overlap with the differentiating characteristics of competing series. Furthermore, ranges must not be expanded so broadly that the series loses its interpretive meaning and usefulness.

A Short List
The main QA steps for approving and progressively correlating a map unit are outlined below. See the preceding discussion for more detail.

For map unit to be approved, the following must be obtained from the MLRA Soil Survey Office:

- Assurance from the MLRA Soil Survey Leader that the subject map unit is fully populated in NASIS and ready to be certified; a thorough quality control review has been completed;
- A MUD for the subject map unit;
- A TUD for each major component of the subject map unit;
- A draft OSD, or an existing OSD with proposed revisions indicated with Track Changes.

Database QA
1. Within NASIS, visually check all NASIS data elements for the subject map unit to ensure all required data is correctly populated.
2. Run NASIS check reports, queries, and validations to further evaluate and cross-check the data.

Map Unit Description QA
3. Read generated map unit description critically; ensure map unit concept is valid and clearly described (it makes sense).
4. Ensure all required map unit data is correctly displayed; generate a MUD from NASIS with nulls displayed as "unspecified" if needed.

Taxonomic Unit Description QA
5. Read the generated taxonomic unit description critically to ensure it accurately and clearly describes a series concept (it makes sense).

6. Ensure the TUD data encompasses the map unit component data; regenerate a TUD from NASIS to confirm this if necessary.
7. Review the narrative pedon description and verify the taxonomic class.
8. Check pedon narrative description for internal consistency and cross check for consistency with Range in Characteristics.
9. Check Range in Characteristics for internal consistency and for consistency with taxonomic class and diagnostic features.

Official Series Description QA

10. If OSD is a new draft for a proposed series, ensure the format and content of the entire OSD follows the MO6 template and MO6 Instructions, as well as NSSH specifications.
11. Read through all physiographic attributes and soil characteristics given in the OSD, cross-check with TUD to ensure those values include the entire range from the TUD. All TUD data must fit within the OSD, or be explained in a correlation note.
12. Verify that proposed revisions to existing OSDs do not violate the series concept, nor overlap with competitors. TUD data that cannot be accommodated by revising the OSD must be explained with a correlation note.

13. Summarize errors and deficiencies discovered in the above QA review and communicate them to the field soil scientist for revision.
14. When follow-up revisions are satisfactorily completed, process the OSD(s), change map unit status to “approved” and populate Map Unit History accordingly.