

## Part 648 – Digital Soil Mapping – Raster Products

### Subpart A – General Information

#### 648.0 Definition and Purpose

##### A. Definition

Raster soil survey is a reference to the products of soil survey work completed using digital soil mapping methodologies. Digital soil mapping is the production of georeferenced soil databases based on the quantitative relationships between soil measurements made in the field or laboratory and environmental data and may be represented as either discrete classes or continuous soil properties. Both digital and traditional soil mapping use a conceptual soil-landscape model as a means for organizing environmental information into discrete divisions. The primary difference between these two approaches is that digital methods exploit quantitative relationships of the environmental information, while traditional methods utilize a more subjective approach and the approximate relationships of the environmental information to spatially represent where the divisions are represented. Traditional soil mapping products are produced primarily through qualitative assessment of the landscape using aerial photography and other supporting digital environmental data and field sampling, and are delivered in a vector data structure (see section 647.4). Digital soil mapping exploits the quantitative relationships between soil observations and digital environmental data and products are delivered as raster data (i.e., rows and columns of pixels with geographic locations storing categorical or continuous data). Raster soil survey does not encompass the Gridded Soil Survey Geographic (gSSURGO) Database, which is a rasterized version of SSURGO polygon information.

##### B. Purpose

The purpose of this part is to provide standards regarding procedures, data development, responsibilities, mapping strategies, and products associated with raster soil survey efforts. These standards are used to consistently produce statewide raster soil survey databases corresponding to State raster soil survey areas and continuous, national soil property layers. The need for this guidance has been driven by significant advances in computer technology enabling soil survey information to be developed with an improved, quantifiable, and consistent representation of spatial variability. This improved information benefits conservation planners, modelers, policy makers, soil survey partners, and other stake holders making land inventory and management decisions. This part is not a step-by-step procedure for how to implement digital soil mapping methodologies to produce a raster soil survey. For more detailed information on digital soil mapping methodologies, see chapter 5 of the Soil Survey Manual (Soil Science Division Staff 2017).

##### C. Raster Soil Survey Strategies

- (1) The term “strategy” here refers to the approach and operations applied towards the goal of developing consistent soil information using digital soil mapping methods. The strategy used for raster soil survey is dependent on the availability and quality of existing soil and environmental data. Raster soil surveys can fill the voids in the current inventory or be used to refine or supplement the existing inventory by more explicitly and consistently identifying a soil class.
- (2) The continuous raster soil survey strategy, discussed below, is a national project focused on developing a spatially continuous, consistent set of soil properties to support large scope soil investigations.

(3) Project Mapping

- (i) This type of mapping identifies soil classes that are supported by expert knowledge, field observations, and associated geographic and environmental data. The quantitative relationships between these data are explored and discrete classes are established. Raster project mapping utilizes the SSURGO tabular data structure to support property and interpretation generation.
- (ii) Types of project mapping are—
- MLRA Raster Soil Survey Project Mapping (See Part 610).—Raster soil survey products improve or advance the currently existing official SSURGO inventories with detailed, refined, spatially explicit and seamless soil information across county, parish or State boundaries. These projects are often applied on a map unit, landform, or ecological unit potentially involving soil catenas. Data updates are delivered in statewide raster soil survey (RaSS) databases and published to the NRCS Data Gateway. Vector updates are discretionary for these projects.
  - Initial Raster Soil Survey Project Mapping (See Part 627).—Raster soil survey products are foundational to providing detailed soil information in unmapped areas as part of the non-MLRA progressive soil survey. Initial raster soil survey project mapping is analogous to a first generation soil survey. The soil survey area is part of a statewide raster soil survey (RaSS) area and the child objects (legend, map unit, data map unit, etc.) are populated in the same manner as a traditional soil survey. A traditional SSURGO tabular and vector product is created from the information collected during the project and final raster map. The products of initial raster soil surveys will become a part of the official SSURGO data.
- (iii) For raster project mapping, the target pixel value represents the map unit key and relates to tabular data stored in NASIS.

(4) Continuous Raster Mapping

- (i) This type of mapping predicts soil physical or chemical properties in horizontal and vertical dimensions. The soil properties are represented across a continuous range of values.
- Raster layers of key soil properties are predicted at specified depth intervals.
  - Depth intervals will be 0-5 cm, 5-15 cm, 15-30 cm, 30-60 cm, 60-100 cm, and 100-200 cm.
  - Soil properties must include all of the following, at a minimum:
    - Total profile depth (cm)
    - Plant exploitable (effective) soil depth (cm)
    - Organic carbon (g/kg)
    - pH (x10)
    - sand (g/kg)
    - silt (g/kg)
    - clay (g/kg)
    - gravel ( $m^3 m^{-3}$ )
    - ECEC ( $cmol_c/kg$ )
    - Bulk density of fine earth (<2 mm) fraction (excluding gravel) ( $Mg/m^3$ )
    - Bulk density of whole soil (includes gravel) ( $Mg/m^3$ )
    - Available water holding capacity (mm)
  - Each property will have an associated uncertainty at each depth interval, representing the 90-percent prediction interval.
- (ii) The target for a continuous raster is a pixel that stores a quantitative value for the respective property. The standards for continuous raster mapping shadow the GlobalSoilMap.net standard 2.4, initially. This includes depth intervals and soil

properties investigated. The one deviation of continuous raster mappings standards from the Global Soil Map standards is the horizontal resolution of the products. The Global Soil Map Standard is 100 meters, whereas the continuous raster mapping standard is 30 meters.

- (iii) Work for the continuous raster mapping will not interfere with the day-to-day activities of the majority of soil scientists working within the National Cooperative Soil Survey. The work will be completed by a few agency employees and include cooperators and partners.

#### D. Procedure

- (1) MLRA Raster Soil Survey Project Mapping.—A NASIS project plan is developed by the Soil Survey Office and follows the project approval process (430-NSSH-610-A-610.1A(3)).
- (2) Initial Raster Soil Survey Project Mapping.—A signed memorandum of understanding (MOU) or interagency agreement is required (430-NSSH-606). A NASIS project is created and follows the project approval process (430-NSSH-610-A-610.1A(3)).
- (3) Initial and MLRA raster soil survey project mapping resolution is determined by the available data, intended use, and is defined on project basis. Project spatial resolution is established in the project proposal, MOU, or interagency agreement. Projects (MLRA and initial raster soil survey) completed at resolutions other than 10 meter are resampled to 10 meter to fit seamlessly within a State raster soil survey database. Resolution for continuous raster mapping will initially be 30 meter but may vary with the version of the product and be seamless for the continental United States (CONUS) and noncontiguous States and territories.
- (4) Spatial data resulting from MLRA and initial raster soil survey project mapping are aligned to the National Land Cover Database (NLCD) and delivered using the following projected coordinate systems:
  - (i) CONUS: USA Contiguous Albers Equal Area Conic USGS version
  - (ii) Puerto Rico and U.S.V.I: USA Contiguous Albers Equal Area Conic USGS version
  - (iii) Hawaii: Hawaii Albers Equal Area Conic
  - (iv) Alaska: WGS 1984 Albers (MRLC - Alaska)
  - (v) PAC Basin: Western Pacific Albers Equal Area Conic
  - (vi) American Samoa: Hawaii Albers Equal Area Conic
- (5) The tabular information collected during MLRA project or initial raster soil survey mapping is stored and maintained in NASIS. The accompanying spatial data are stored in a statewide geodatabase with a corresponding tabular export and posted to the NRCS Data Gateway.

Raster soil survey efforts are pursued using accepted exploratory and prediction methods, such as unsupervised (e.g., ISODATA, k-means) and supervised classification (e.g., predictive modeling, knowledge-based) and geostatistics. These methods are found in geographic information system (GIS) and statistical software, peer-reviewed literature, and are discussed in chapter 5 of the Soil Survey Manual (Soil Science Division Staff 2017). A consistent approach is desired for projects across similar landscapes or landforms such as MLRAs or floodplains.

- (6) SSURGO vector data reflects the findings of MLRA or initial raster project mapping in coincident areas. MLRA raster project mapping is not carried out as a means to require the development of a more detailed SSURGO product. Regional discretion, with concurrence from State conservationists (or designated appointee), the board of advisors (430-NSSH-609-A-609.1B(3)), technical teams, and cooperators, is used to determine the degree to which findings from MLRA raster soil survey projects are incorporated into existing SSURGO spatial and tabular information. The intent is to refine and improve upon the information which has already been inventoried. Raster mapping products must support consistent soil survey products across political boundaries, including correlation of existing SSURGO

- products into MLRA map units across natural landforms over broad areas (430-NSSH-610-A-610.1A(4)).
- (7) MLRA and initial raster soil survey mapping is quantitatively validated as part of the modeling process using accepted validation methods. These estimates of accuracy and uncertainty are published alongside the raster soil maps and published in the metadata. Examples are overall map accuracy, standard error, uncertainty, or probability maps. Continuous raster property maps must include uncertainty measures as specified in the GlobalSoilMap.net standard.
  - (8) Legends for MLRA and initial raster soil survey mapping projects are built and managed in a raster soil survey area corresponding to a respective State or territory. Legends for initial raster soil survey mapping projects are also built and managed on a non-MLRA soil survey area extent. Publication symbols are the same as the national map unit symbol.

#### E. Mapping Data Development

##### (1) Spatial Data

- Spatial data used for raster soil survey map development meets the National Standard for Spatial Data Accuracy, set forth by the Federal Geographic Data Committee (<https://www.fgdc.gov/standards/projects/accuracy>). This organized structure of Federal geospatial professionals and constituents provide executive, managerial, and advisory direction and oversight for geospatial decisions and initiatives across the Federal Government (<https://www.fgdc.gov/organization>).
- (2) The data delivered to project offices is coordinated and validated by regional GIS specialists to ensure the consistency of coordinate systems, proper selection of datum transformations, resampling techniques, resolution, data types, co-registration, etc. Examples of data used include those listed below, although many options exist for data and derivatives beyond those presented here.
    - (i) Digital Elevation Models (SRTM, NED, IFSAR, LiDAR) and derivatives
      - Slope
      - Aspect
      - Topographic wetness index
      - Curvature
    - (ii) Spectral (LANDSAT, ASTER, Orthophotography) and Derivatives
      - Band ratios
      - Principal components
    - (iii) Thematic Data
      - National Land Cover Database
      - Gap Analysis Program
      - Landforms
      - National Wetland Inventory
    - (iv) Climatic Data
      - PRISM
      - Annual water balance
      - Incoming solar radiation

##### (3) Data Capture

Data capture relates to the process of maintaining, storing, and delivering information from a raster mapping project into a useable data product capable of generating soil properties and interpretations.

- MLRA Raster Soil Survey Project Mapping

Resources guiding MLRA raster soil survey projects include information already present within map unit descriptions (soil survey publications), historical documentation (transects, pedons, general soil maps), environmental data (digital elevation, spectral, thematic), and local expert knowledge.

- Suitable extent is identified for improvement of official data (map units, landform, ecological sites, etc.). The corresponding extent is coincident with SSURGO polygon information and serves as the area over which the MLRA raster project is completed. Projects extending across very large areas can be further subdivided if soil forming factors, such as climate, vary considerably across the identified extent. Refined raster classes are identified and proposed using digital soil mapping methods, image analysis, field observations, and expert knowledge.
- Correlation occurs on the basis of articulating the class concepts through model training data (field observations or ancillary data), model outputs, environmental data (digital elevation and spectral data), independent validation data (field observations/ground truthing), field reviews, visual analysis, estimates of accuracy or uncertainty, and documentation of the correlation in the NASIS project object.
- Every State has a raster soil survey area and an accompanying raster soil survey legend in NASIS, identified with an area symbol using the State abbreviation and an area of 000 (e.g., MA000 (section 648.12)). Raster classes identified in MLRA raster soil survey projects are populated at the map unit level and below within the State raster soil survey area.
- Field observations (training and validation data) are entered as NASIS pedons and related data.
- Map layers of associated accuracy and uncertainty values are produced and will accompany the class layer (figure 648-A1). Validation measures for overall model accuracy or performance, as well as individual class accuracy or performance, are reported separately in metadata. Measures for overall model accuracy or performance will meet a minimum target of 60 percent. Validation using an independent data set is highly recommended, but not required, for all projects. Cross-validation or other appropriate data-splitting methods may be used in the absence of an independent validation data set.
- A file geodatabase is established for each State raster soil survey area (section 648.12). Results of raster soil survey project mapping (MLRA or initial) accumulate in this geodatabase, including raster layers and tabular exports. SSURGO products are developed or updated from MLRA raster soil survey mapping efforts, but require considerations of correlation and map unit design to create a more generalized representation of spatial and tabular data.
- The State soil scientist certifies the data for publication in the NASIS “Legend Export Certification History” table.
- The State raster soil survey area geodatabase, which includes the State raster class and accuracy or uncertainty spatial layers, and tabular data are published to the NRCS Data Gateway as two data packages. The first is a file geodatabase system containing both rasters and tables, and the second is a directory containing rasters in compressed tagged image file format (.tif) and tabular data as text files in pipe delimited format.

The naming convention for the State raster workspace or directory is the product, area symbol, and year (e.g., RaSS\_MA000\_2017). Within the directory is an unsigned 32-bit raster class layer (MA000\_class\_2017) and 32-

bit floating-point accuracy or uncertainty layers (MA000\_a or MA000\_u). Both layers have a NoData value of 0.

- Initial Raster Soil Survey Mapping

Information aiding the development of initial raster soil surveys include adjoining SSURGO soil survey areas, STATSGO2, geological maps, environmental data (digital elevation, spectral, thematic data), soil sampling, and expert knowledge.

- Map units or soil classes are identified and proposed based on the needs of the project set forth in the MOU or interagency agreement using digital soil mapping methods, image analysis, field observations, and expert knowledge.
- Correlation occurs on the basis of articulating the class concepts through model training data (field observations or ancillary data), environmental data (digital elevation and spectral data), model outputs, independent validation data (field observations), field reviews, visual analysis, estimates of accuracy or uncertainty, and documentation of the correlation in the NASIS project object.
- Every State has a raster soil survey area and an accompanying raster soil survey legend in NASIS, identified with an areasymbol using the State abbreviation and an area of 000 (e.g., MA000 (section 648.12)). Raster classes identified in initial raster soil survey projects are populated as map units and below within the State raster soil survey area.

Additionally, initial raster soil survey projects have a non-MLRA soil survey area legend and below in populated NASIS for incorporating into SSURGO and Web Soil Survey.

- Field observations (training and validation data) are entered as NASIS pedons and related data.
- Map layers of associated accuracy and uncertainty values are produced and will accompany the class layer (figure 648-A1). Validation measures for overall model accuracy or performance, as well as individual class accuracy or performance, are reported separately in metadata. Measures for overall model accuracy or performance will meet a minimum target of 60 percent. Validation using an independent validation data set is highly recommended, but not required, for all projects. Cross-validation or other appropriate data-splitting methods may be used in the absence of an independent validation data set.
- A file geodatabase is established for each raster soil survey area (section 648.12). Results of raster soil survey project mapping (MLRA or initial) accumulate in this geodatabase, including raster layers and NASIS tabular exports. SSURGO products are developed from initial raster soil survey efforts, but require considerations of correlation and map unit design to create a more generalized representation of spatial and tabular data.
- The State soil scientist creates a SSURGO export of the non-MLRA soil survey area and certifies the data for publication in the NASIS “Legend Export Certification History” table.
- Vector spatial and tabular data are hosted on Web Soil Survey.
- The State raster soil survey area geodatabase, which includes the State raster class and accuracy or uncertainty spatial layers, and tabular data are published to the NRCS Data Gateway as two data packages. The first is a file geodatabase system containing both rasters and tables, and the second is a directory containing rasters in compressed .tif and tabular data as text files in pipe delimited format.

The naming convention for the State raster workspace or directory is the product, area symbol, and year (e.g., RaSS\_MA000\_2017). Within the directory is an unsigned 32-bit raster class layer (MA000\_class\_2017) and 32-bit floating-point accuracy or uncertainty layers (MA000\_a or MA000\_u). NoData values are 0.

- Raster Property Mapping
  - Soil properties are predicted from soil observations using digital soil mapping methods. For more detailed information on digital soil mapping, see chapter 5 of the Soil Survey Manual (Soil Science Division Staff 2017).
  - Field observations (training and validation data) are entered into NASIS.
  - A map of uncertainty values is produced for each property and depth interval representing the 90-percent prediction interval.

The initial, primary spatial entity is a point location with defined X, Y coordinates. Points are located at the cell centers of a raster grid, initial spatial resolution will be 30 m.

#### F. Responsibilities

- (1) National Headquarters provides overall direction, policy, guidance, and leadership for the National Cooperative Soil Survey within NRCS. See part 608, section 608.1, of this handbook for more detailed information on the responsibilities of National Headquarters and the other NRCS offices mentioned in this section.
- (2) The National Soil Survey Center is responsible for—
  - (i) National Standards.
  - (ii) NASIS database support.
  - (iii) Training.
  - (iv) Research.
  - (v) Analysis.
  - (vi) Coordination of NCSS raster property mapping.
  - (vii) Raster property metadata development.
  - (viii) Packaging and delivering raster property data to the NRCS Data Gateway.
  - (ix) Packaging and delivering raster property data to web and map services which are to be developed.
- (3) The regional office is responsible for providing the leadership for—
  - (i) The production and quality assurance of raster project spatial and tabular data.
  - (ii) Ensuring soil scientists receive digital soil mapping training.
  - (iii) The qualitative review of raster property data.
  - (iv) Correlation of complete raster mapping projects to legends.
  - (v) Acquiring and distributing raster project spatial data (elevation, spectral, others).
  - (vi) Raster project metadata development.
  - (vii) Packaging and delivering raster class and tabular data to the NRCS Data Gateway.
  - (viii) Packaging and delivering SSURGO data to the staging server.
- (4) The State office (430-NSSH-609-A-609.1B(3)) is responsible for the following:
  - (i) State Conservationist (or Designated Appointee)
    - Certifying initial raster soil surveys, incorporated into SSURGO, are official soil survey data
    - Participating or appointing a member to the board of advisors for MLRA soil survey activity
  - (ii) State Soil Scientist
    - Serves as a member of SSOmanagement/tech team for MLRA soil survey activity, identifying priorities for MLRA raster projects

- Review and concur with MLRA raster projects and recommendations
  - Exporting tabular data and committing vectorized versions of MLRA or initial raster mapping projects to Web Soil Survey
  - Providing certification of State Raster Soil Survey Area data and interpretations
  - Promoting raster soil survey products to internal and external customers and providing technical assistance in their use
- (5) The MLRA soil survey office is responsible for—
- (i) Coordinating technical and management team communication.
  - (ii) Identifying and proposing MLRA raster projects.
  - (iii) Completing digital soil mapping training.
  - (iv) Managing and populating MLRA and initial raster projects in NASIS.
  - (v) Providing supporting documentation and detailed logic on raster class development using pedons and detailed text notes linked to the project and component data in NASIS.
  - (vi) Conducting quality control for raster projects and spatial and tabular data.
- (6) The National Geospatial Center for Excellence (NGCE) is responsible for—
- (i) Providing technical assistance to regional offices in spatial and metadata development to meet raster project and SSURGO specifications.
  - (ii) Assisting regions with resolving problems related to submitting raster data to the Geospatial Data Gateway.
  - (iii) Providing technical assistance on metadata development.

G. Metadata

- (1) The format for metadata may be either FGDC CSDGM or ISO19115. The FGDC is moving towards the ISO format.
- (2) Metadata development avoids the use of escape characters or the abbreviation is spelled out in text narrative. These characters include the following:
  - (i) " double quotes
  - (ii) ' single quote or apostrophe
  - (iii) > greater than
  - (iv) < less than
  - (v) & ampersand
- (3) Additional information on metadata can be found here:  
<https://www2.usgs.gov/datamanagement/describe/metadata.php>

**648.1 Accuracy and Uncertainty**

A. Significance.—The quality control and quality assurance (QC/QA) process for traditional soil mapping is based on many forms of evaluation, accumulates several degrees of subjectivity, and relies heavily on qualitative information evaluation (field experience, institutional knowledge, etc.). In contrast, validation of raster soil survey is often claimed to be free of subjective interference and entirely quantitative. However, simply reporting “right or wrong” can distort the reality of accounting for variability inherent in soil and associated environmental systems. It is recommended more than one method of both accuracy and uncertainty be used to evaluate MLRA and initial raster soil survey projects in order to explore and communicate results. Methods which allow for weighting class distance (e.g., “degrees” of right or wrong) are encouraged.

B. Suggested Methods of Accuracy and Uncertainty

Figure 648-A1: Suggested Methods of Accuracy

Accuracy				
Method	Description	Pros	Cons	Reference

Title 430 – National Soil Survey Handbook

<b>Overall Accuracy (Cohen’s Kappa)</b>	The proportion of correctly classified observations as tabulated in the cross-classification matrix.	Simple to calculate and interpret, based entirely on the cross-classification matrix.	Does not account for mistakes due to chance. Does not incorporate prior knowledge of class proportions. Does not incorporate class similarity. Does not incorporate information contained within vector of predicted probabilities.	
<b>Tau index</b>	An index of accuracy that accounts for agreement by chance—effectively a replacement for Cohen’s Kappa.	The index is more informative when appropriate class proportions are supplied. Index values can be referenced to concepts such as “better” or “worse” than random allocation.	Appropriate prior class proportions are method dependent. Interpretation requires some training.	Rossiter et al., 2017
<b>Weighted Tau index</b>	Alternative version of the tau index that accommodates class similarity.	The index is more informative when appropriate class proportions are supplied. Class similarity down-weights mistakes between similar classes.	Appropriate prior class proportions are method dependent. There is no universal method for estimating class similarity weights. Interpretation requires some training.	Rossiter et al., 2017
<b>Brier Score</b>	The Brier score is an index of agreement between an observed (actual) class and vector of predicted probabilities. Lower values denote higher accuracy.	Integrates more information about predictions (all probabilities) vs. the most likely class used by overall accuracy or tau index.	Interpretation requires some training. Does not incorporate prior knowledge of class proportions. Does not incorporate class similarity.	Brier, 1950 Harrell, 2011

Figure 648-A2: Suggested Methods of Uncertainty

Uncertainty				
Method	Description	Pros	Cons	Reference
<b>Shannon Entropy</b>	Quantitative measure of “information” contained within a vector of probabilities. Larger values denote less information.	Integrates “confusion” over all probabilities. Entropy values from widely different models (e.g., differing numbers of classes) can be directly compared.	May not be implemented in all software packages, but simple enough to compute manually.	Shannon, 1949
<b>Normalized Shannon Entropy</b>	An alternate version of Shannon Entropy that is constrained to the interval [0,1].	Integrates “confusion” over all probabilities. Constrained to [0,1].	Cannot be used to compare models with differing numbers of classes. May not be implemented in all software packages, but simple enough to compute manually.	Kempen et al., 2009
<b>Confusion Index</b>	An index of confusion among the top two most likely classes.	Simple calculation and constrained to [0,1].	Cannot be used to compare models with differing numbers of classes. Only describes uncertainty of top two classes. May not be implemented in all software packages, but simple enough to compute manually.	Burrough et al., 1997

Notes:

1. The “overall accuracy” and unweighted Tau index are computed from the cross-classification (confusion) matrix. Rows within the cross-classification matrix record counts of predicted classes and columns record counts of actual classes.

2. The row-wise proportion of correct entries in the cross-classification matrix is typically described as the “user’s accuracy” (map unit purity). The column-wise proportion of correct entries in the cross-classification matrix is typically described as the “producer’s accuracy” (or soil class representation).
3. Most predictions from a model of soil classes are represented as a vector of class-wise probabilities typically associated with each pixel in the modeling domain. Entries within the cross-classification matrix use only on the most likely (highest probability) class.
4. Differences between Shannon Entropy and the Confusion Index are more pronounced when there are a large number of classes or median class probability is low (e.g., there is no clear dominant class in the predictions).

## 648.2 References

- A. Brier, G.W. 1950. Verification Of Forecasts Expressed In Terms Of Probability. *Monthly Weather Review* 78(1): 1–3. [https://doi.org/10.1175/1520-0493\(1950\)078<0001:VOFEIT>2.0.CO;2](https://doi.org/10.1175/1520-0493(1950)078<0001:VOFEIT>2.0.CO;2)
- B. Burrough, P., P.V. Gaans, and R. Hootsmans. 1997. Continuous classification in soil survey: spatial correlation, confusion and boundaries. *Geoderma* 77(2-4): 115–135. [https://doi.org/10.1016/S0016-7061\(97\)00018-9](https://doi.org/10.1016/S0016-7061(97)00018-9)
- C. Harrell, F.E. 2001. *Regression Modeling Strategies*. Springer, New York, NY.
- D. Kempen, B., D.J. Brus, G.B. Heuvelink, and J.J. Stoorvogel. 2009. Updating the 1:50,000 Dutch soil map using legacy soil data: A multinomial logistic regression approach. *Geoderma* 151(3-4): 311–326. <https://doi.org/10.1016/j.geoderma.2009.04.023>
- E. Rossiter, D.G., R. Zeng, and G.-L. Zhang. 2017. Accounting for taxonomic distance in accuracy assessment of soil class predictions. *Geoderma* 292: 118–127. <https://doi.org/10.1016/j.geoderma.2017.01.012>
- F. Shannon, C.E. 1949. *A mathematical theory of communication*. University of Illinois Press, Urbana.
- G. Soil Science Division Staff. 2017. *Soil survey manual*. C. Ditzler, K. Scheffe, and H.C. Monger (eds.). USDA Handbook 18. Government Printing Office, Washington, DC.