***Proposed definition for new section C. in Part 618.2***

1. Prior to this version of the National Soil Survey Handbook, no definition existed for low, high, and representative values. For recent and newly populated information in NASIS, the representative value (rv) is meant to approximate the 50th percentile (median) of the data set being described. The 50th percentile is the value where 50% of the data are less than this value. The low and high values are meant to approximate the 5th-10th and the 90th-95th percentiles, respectively. For example, the 5th percentile is the value where 5% of the data are below that value and the 95th percentile is the value where 5% of the data are above that value. While the percentile approach did not guide the low, high, and rv for data populated prior to this version of the National Soil Survey Handbook, they generally approximate the current definition. Where data used to populate the SSURGO database are not computationally-derived, the populated values are intended to approximate these percentiles for the property being described.

The percentile approach is ideal for typical soil survey data such as environmental and soil property data sets. Percentiles provide benchmarks for the spread and central tendency for both normal and non-normal distributions, and the values will always fall within the minimum and maximum of the observed data. Data managers can use any widely accepted method for calculating percentile values. For computationally derived percentile values, a data set of at least 5 values is required. Consider a hypothetical data set for field-described clay content from the A horizon from 10 pedons.:

clay content: 11, 10, 12, 23, 17, 16, 17, 14, 24, 22, 14

clay content sorted: 10, **11**, 12, 14, 14, **16**, 17, 17, 22, **23**, 24

Low/10th percentile = **11**

RV/50th percentile = **16**

High/90th percentile = **23**

Soil scientists have the flexibility to design data ranges that reflect the soil component and map unit concepts they observe on the ground and have developed based on the local soil-landscape model. The model is supported by a combination of field measurements, field observations, and laboratory data. For any computationally-derived low, high, and representative values, the soil scientist must thoroughly consider the source of the data and compare the derived values to knowledge gained in the field.