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

1. For recent and newly populated information in NASIS, the representative values are meant to approximate the 50th percentile (median). The 50th percentile is the value where 50% of the data are both above and below that 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. The low, high, and representative values for data populated prior to this version of the National Soil Survey Handbook were not guided by the percentile approach, but generally approximate the current definition. Even where data used to populate the SSURGO database are not computationally derived, the populated values are designed to approximate the aforementioned percentiles for the data set being described.

The percentile approach is preferable to other measures of central tendency, such as the mean and standard deviation, because percentiles require no distributional assumptions and the percentile values fall within the bounds of the data set from which they are computed. This means that percentiles can 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. Consider a hypothetical data set for field-described clay content from the A horizon of the same taxa:

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**