



## **Posted on behalf of a colleague who would like to remain anonymous**

Dylan Beaudette

---

Community comment on "How well does Digital Soil Mapping represent soil geography? An investigation from the USA" by David G. Rossiter et al., SOIL Discuss., <https://doi.org/10.5194/soil-2021-80-CC3>, 2021

---

My only critical comment in the pdf concerns what appears to be a mis-characterization of POLARIS uncertainty estimates. Otherwise the paper is well written and provides a thoughtful comparison on the differences amongst soil datasets.

### **Specific comments:**

Line 60: Does it? Domain expertise seems critical to know when these ML models are overfit.

Line 242: PSP used a "classification" forest, thus how could it has also used a quantile "regression" forest? The PSP is based on a weighted average of soil components.

Line 405 / Table 1: What is with the units in this table? The RMSD is different by a pH of 4-6? You're multiplying by 10? Why?

Figure 3: I love visuals, but in this case a simple correlation matrix would be more informative.

Table 2: I would have figured that gNATSGO would have had the highest nugget because it is the most detailed, but I suppose this is an artifact of the polygons.

Figure 11: Red-Green is a bad color scheme for folks (like me) whose vision is color deficient.

Figure 12: Any thoughts on the use of a consistent color scheme across all thematic maps?

Figure 13: Shouldn't you have a legend for the gridded soil maps? Also, IMO I find the transparency and orientation distracting.