

Predicted Soil Protein

The Autoclaved Citrate Extractable (ACE) Protein Index is an indicator of the amount of protein-like substances that are present in the soil organic matter. Soil protein represents the largest pool of organically bound nitrogen (N) in the soil organic matter, which microbes can mineralize and make available for plant uptake.

Beginning in 2020, the Cornell Soil Health Lab moved to predict Soil Protein from a suite of measured parameters in the Standard comprehensive assessment of soil health (CASH) package. The lab measured ACE Soil Protein assay will remain available to researchers and others as an optional add-on to the standard CASH package.

Modeling Soil Protein

In 2020, the Cornell Soil Health Lab determined that soil protein, a valuable, but time-intensive measurement, could be accurately predicted. A model was developed using Random Forest to predict soil protein from a suite of measured parameters, including % sand, % silt, % clay, Organic Matter, Active Carbon, Total Carbon, Total N, Carbon-to-Nitrogen ratio, pH, and extractable Magnesium and Iron. This model was able to explain 85.2% of the variation in soil protein with a low average root mean square error (RMSE) value of between 1.0-1.3 (Fig. 1). RMSE is a measure of how much observed values deviate from the predicted values.

Random Forest (RF) is a robust machine learning algorithm that uses a decision tree approach to model variables. Machine learning algorithms such as RF have become increasingly popular techniques to model parameters that are difficult or costly to measure. Environmental scientists have developed models using routinely measured parameters to predict difficult or costly to measure soil properties such as soil nitrogen availability to crop plants.

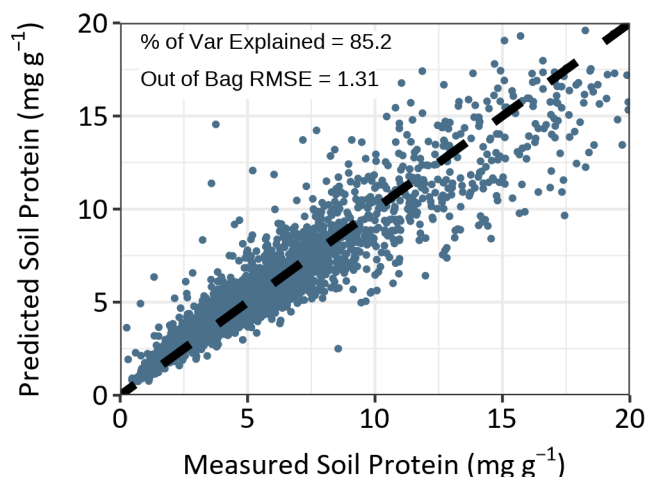


FIGURE 1. Measured Soil Protein vs. Predicted Soil Protein for the CASH database (n=4,171).

How soil protein relates to soil function

Plant residues are ultimately the source of much of the soil organic matter. Residues are made up of several types of compounds that are largely similar in composition (Fig. 2). Of these compounds, protein contains the largest fraction of N.

Protein content, as organically bound N, influences the ability of the soil to store N, and make it available by mineralization during the growing season. An active microbial population is responsible for this change from organic N to plant available N. Soil protein content has also been associated with soil aggregation and thus water storage and movement.

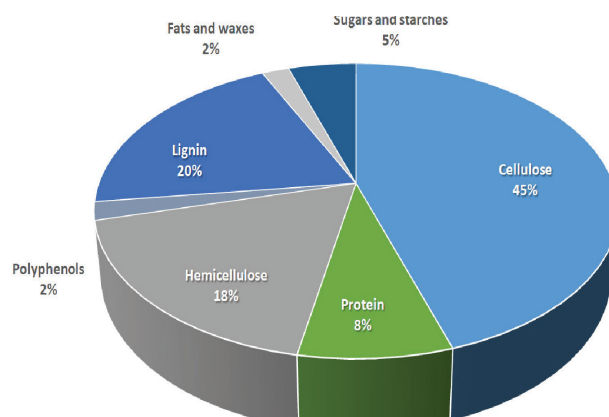


FIGURE 2. Types of compounds in plant residues. Proteins are found in high abundance and contain the largest N fraction.

Predicted Soil Protein

Managing constraints and maintaining optimal soil protein content

To store and maintain N in SOM, we need to accumulate compounds that are relatively stable, rich in N (low C:N ratio), microbially degradable, and potentially abundant in amendments, crops, cover crops, or residues. Building and maintaining healthy, biologically active soil with large reserves of decomposing plant tissue in organic form is a good approach to provide a crop with its N needs over time as opposed to applying synthetic forms of N that plants may not use immediately and be lost. Organic N reserves are built over years and should be maintained to the extent possible.



FIGURE 3 a - d. SOM building practices. (a) Manure (amendment), (b) Red clover (green manure), (c) Soybeans (crop rotation), (d) No-till drill (reduced tillage).

Soil protein content can be increased by adding biomass such as manure, fresh green biomass, well finished compost high in N (Fig. 3). Cover crop mixtures using diverse species can add large amounts of biomass and maintain the presence of actively growing roots. Well nodulated legumes are important sources of root zone N. Most of these sources provide a slow release of N over time, which can reduce environmental losses. Use careful planning to reduce tillage intensity and to enable timely no-till drilling of cover crops. Protein content tends to decrease with increasing soil disturbance such as tillage.

Scoring function

Figure 4 below depicts **predicted** Soil Protein scoring functions and upper value limits for coarse, medium, and fine textured soils. It is important to note that extremely high N mineralization could increase losses of N to the environment and thus harm air and water quality.

The red, orange, yellow, light green and dark green shading reflects the color coding used for the ratings on the soil health report summary page.

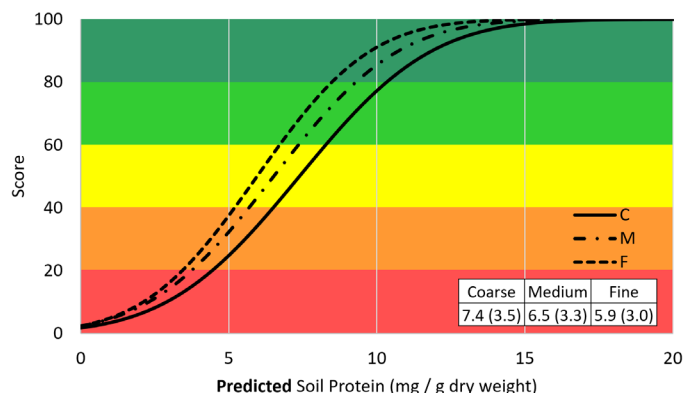


FIGURE 4. Predicted Soil Protein scoring functions and upper value limits for Coarse (C), Medium (M) and Fine (F) textural classes. Mean and standard deviation (in parenthesis) for each class are provided. In this case more is better. Higher protein index scores indicate a larger pool of organically-bound soil N.

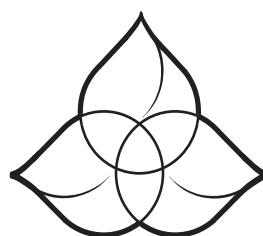
Acknowledgment

The Cornell Soil Health Lab thanks the School of Integrative Plant Science and the College of Agriculture and Life Sciences (CALS) at Cornell University. We also thank the Northeast Sustainable Agriculture Research & Education Program, the New York Farm Viability Institute, the USDA-NRCS and Cornell Cooperative Extension.

Cornell University Soil Health Laboratory

Bob Schindelbeck, Kirsten Kurtz

The Soil Health Manual Series of Fact Sheets is based off work published in the Comprehensive Assessment of Soil Health – The Cornell Framework (Version 3.2). <https://soilhealth.cals.cornell.edu/manual/> We thank all co-authors for their contributions to these fact sheets. Thanks to Joseph Amsili for his contribution to predicting results via the Random Forest Model. Details about the models used to predict Soil Protein can be found in the [peer-reviewed literature](#).



Citation: Schindelbeck, R.R., A.J. Ristow, K.S. Kurtz, L.F. Fennell, H.M. van Es., January 2017. Cornell University Comprehensive Assessment of Soil Health Laboratory Soil Health Manual Series