What is your idea?

Stakeholders related to agriculture (including the farmers) and development often undermine the importance of soil health, especially below ground soil microbial diversity (in terms of diversity and abundance). More often, assessing soil microbial diversity and other soil indicators sounds as having nothing to do with crop health and yield. However, this enormous diversity of microorganisms provides "support and regulation services" to the biogeochemical cycling of elements such as Carbon, Nitrogen, Phosphorous and Sulphur. Soil microbes are "hidden" under the soil surface, but they play a key role for sustaining soil fertility, ecosystem functioning, crop health and productivity. Extensive literature suggests that various above ground topological, climatological, biophysical and land use features, as well as ground level soil type and soil moisture, shape soil microbial diversity, further influencing soil health and crop productivity. In this project, **we propose to take advantage of rapidly reducing genome sequencing costs to generate new geocoded soil data (by characterizing microbial diversity), fuse this with available earth observation data, and leverage the transformative potential of artificial intelligence, to identify a combination of key above ground and ground level features that can best describe soil microbial diversity and hence provide an estimate and an indicator for soil health.** The outcome will enable placing agricultural landscapes on a continuum of soil health. This information will enable stakeholders (including farmers) to obtain a "sense" of soil health (by circumventing the need to sample soils on a large scale), which effectively can translate into influencing and adoption of sustainable farming practices, thereby incorporating data into farming. Essentially, the model when provided with a new set of above ground and ground level features will "crunch" through the data, and provide an estimate of soil health associated with that particular agro-ecosystem.

By merging different sources of datasets, and using artificial intelligence, we aim to provide stakeholders a "peephole" into soil health, based on what is present above and in the ground. This is extremely essential, since in order to sustainably intensify, it is important to monitor and achieve ecologically diverse agricultural systems, which includes both functional above and below ground biodiversity. However there has always been a "data gap" in terms of below ground biodiversity, which we attempt to fill through the proposed activity.

How will you pilot it?

Rapid transition from small-scale subsistence farming to market oriented and high input farming are rendering northern Vietnam's agricultural landscapes unsustainable. This transitionary state provides a timely opportunity to pilot this approach in this region. Northern Vietnam is also home to one of the three climate-smart villages in Vietnam: through this pilot, we want to enable processes that achieve scale, and positively impact several farmers, moving from "climate-smart villages" into "climate-smart regions". Additionally, the site is also home to several ongoing CGIAR related research for development projects, which could benefit from the outcomes of the project. Therefore northern Vietnam becomes pivotal to this study. In addition, the availability of a high quality gridded soil map also drives our choice of northern Vietnam to pilot this project. Agricultural area of northern Vietnam will be classified, based on a typology driven by a set of above ground (topological, climatological, biophysical and land use) and ground level (soil type and soil moisture) features obtained from sources such as LANDSAT, SENTINEL, FAO's Soil type grids, NASA's SMAP. Classes that significantly differ in these features will be subsampled and selected for further soil sampling and microbial diversity characterization.

To be used in defining diagnosis of soil health status, biological indicators need to take the functioning of the soil health into account and be sensitive to modifications within the soil. Microbial communities offer great potential in this regard because (1) they are present in high density and diversity in all environments, (2) they are actively involved in biological functioning and in the services rendered by ecosystems, and (3) they are sensitive to changes in biomass and soil conditions. Moreover, the use of microorganisms as indicators also satisfies practical and economic criteria (i.e. to be quick and easy to use and interpret, reproducible, inexpensive, and readily accessible to users). They must also be associated with appropriate reference systems so that the desired diagnosis can be obtained by positioning the measured values within a range (weak/normal/high) of variability. To identify soil sampling locations for microbial diversity characterization within the diverse classes of agricultural landscapes, a randomized pooled soil sampling strategy will be deployed, that balances between capturing enough soil variation and sampling resolution. Subsequently, relationship between soil microbial diversity and its influencing factors (above and ground level features) will be obtained to calibrate a predictive model for soil microbial diversity of the agricultural lands of northern Vietnam. As a result, these areas will be placed on a continuum of soil microbial diversity, which will enable ecological diversity monitoring and site prioritization, further facilitating site-specific agricultural advisory. As this approach enables identification of farmer communities in sites with differing levels of soil microbial diversity, learning alliances will be established to enable a farmer knowledge diffusion scheme to enable timely adoption of sustainable farming measures.

The budget will primarily be used to cover soil sampling, processing and metagenomics (sequencing) analysis activities. The remaining will be used to cover certain percentage of time of people involved. However if there is a need to buy high-resolution digital imagery, processing power and storage space that will be prioritized. CIAT will contribute to data analytics and model development; CATAS (Chinese Academy of Tropical Agriculture Sciences) will bring its expertise in developing sampling strategies for this activity, and contribute to processing and sequencing of soil samples. NOMAFSI (Northern Mountainous Agriculture and Forestry Sciences Institute) has nuanced understanding of agricultural systems in the region and has a reputable relationship with the farmers. They will coordinate sampling activities and development of farmer learning alliances. NOMAFSI will be the primary beneficiary of this project as they are involved in extension activities and piloting of eco-efficient practices. The first quarter of the year will be dedicated to development of typology and identification of sampling sites. The second quarter will be dedicated for sampling activities and metagenomics analysis of the samples. The third quarter will be dedicated to model development and validation. In the fourth quarter, priority sites will be identified and learning alliances for farmers will be set up. In conclusion, through this pilot we aim to fill the existing data gap that describes diversity of agricultural landscapes, by developing a model that provides an estimate of soil microbial diversity based (as an indicator of soil health) on a set of above ground and ground level features. The outcome will strategically inform and influence existing farming practices at multiple scales to meet market demands (both quantity and quality), hence sustaining livelihoods and restoring degraded lands of small holder farmers.