Orchard soil characterisation

New Zealand Avocado

NZ Avocado Growers' Association Inc.
NZ Avocado Industry Ltd

August 4, 2021

Authored by: Phillip West, Sarah Sorensen & Miguel Tapia

Content table

Executive summary

Supporting avocado growers to be more profitable is an industry goal, with a target of increasing yields to an average of at least 15 t/ha of high-quality fruit. More growers are looking to their soils to improve economic and environmental outcomes, particularly soil biology. This study aims to explore the role of different soil properties in relation to yield. Data from this study will also be incorporated into the Avovantage project looking at on-orchard practices that can help reduce the risk of fungal rots. The analysis relating to fruit quality is not included in this report but will rather be included in the Avovantage reporting channels.

Key variables relating to soil physical, biological and chemical composition were successfully collected from groups of high performing orchards (>15t/ha) and lower yielding orchards in the Bay of Plenty region. Results of this study provide a valuable comparison to growers wishing to test their own soil to see what variable may be influenced to potentially improve yields.

Nutrient testing on soil, leaf, fruit skin and flesh was conducted by Hill Laboratories. Commercially available biological soil tests from Hill's Laboratories, Linnaeus and Soil Foodweb were used to assess soil biology and a visual soil assessment system was used to assess soil physical properties. A comparison of soil biology results from Soil Foodweb and Linnaeus highlighted that common variable results did not agree with each other. The different tests use different methodologies but further investigation is required to understand why these different tests are providing different results.

Aspects of soil physical, biological and chemical composition all showed importance in classifying whether an orchard was high yielding or not. None of the leaf nutrient results showed a correlation with yield classification, but the sample size is small relative to most nutrient studies. Some of the correlations observed were contrary to what is expected to contribute to a high performing soil. It may be that there are common management practices among high performing orchards that negatively impact soil characteristics. Therefore, the correlations seen within the soil variables may be more related to the management practice rather than the high yields common to these orchards.

Higher levels of iron, low levels of aluminium, and high C/N ratio were chemical components of soil that were important predictors of yield and correlated with the group of orchards achieving over 15t/ha. Of the biological variables measured, higher ciliates, lower flagellate protozoa communities, higher dry weight, lower total bacteria and lower gram-negative bacteria correlated with higher yielding orchards. Several of these biological correlations are counter to what might be expected of a highly productive soil.

Soil worm counts was the only physical soil observation to correlate with orchard yield classification with lower worm counts present on higher yielding orchards. Again, this is counter to what is expected in a highly productive soil.

A combination of the Hill Laboratories chemical variables and the Soil Foodweb biological variables showed the highest classification accuracy of 79%. Nutrient variables had a higher importance than biological variables in classifying orchard performance within the orchards in this study. Testing the variables deemed as important for classification on a wider data set and across multiple seasons would be required before confident assumptions can be made that these variables truly correlate with higher yielding orchards.

The variables collected provide a valuable benchmark from high producing orchards in different regions for growers to compare their own soils to. This will support decision making on whether soil

amendments or changes to an orchards management may be the best avenue to pursue to improve yields. The results also highlight that there might be improvements that can be made to the soils of high yielding orchards that may further enhance yields.

Acknowledgements

NZ Avocado would like to thank the avocado growers who participated in this project.

Introduction

Supporting avocado growers to be more profitable is an industry goal, with a target of increasing yields to an average of 15 t/ha of high-quality fruit. However, this improvement needs to be both environmentally and economically sustainable. As a result, more growers are looking to their soils to improve economic and environmental outcomes, particularly soil biology. As well as playing direct roles in supporting orchard production, understanding soil biology can provide clues to other soil properties that influence productivity, such as physical structure and chemical composition.

Currently, there is limited information pertinent to soil characteristics and their influence on yield and fruit quality in New Zealand avocado orchards.

Historically, the main focus for avocado orchard soils has been the chemical aspect of the soil and what fertiliser should be applied to maintain or improve the health and yield of the tree. Literature shows that nutrient amounts and ratios are associated with tree yields and fruit quality; however, soil physical and biological characteristics also contribute to nutrient availability and tree performance (Crowley, 2007).

This project aims to collect a combination of biological, chemical, and physical avocado orchard soil data and identify which attributes influence yield. While the number of orchards being looked at is limited, outcomes may help establish biological benchmarks, nutrient target levels and prompt discussion about managing inputs better to achieve improved production and environmental outcomes. This project will also enhance knowledge about how soil characteristics interact in avocado orchards. Additionally, leaves and fruit nutrient data were collected and analysed to facilitate the interpretation of soil data and explore a possible influence on yield.

1.1 Soil chemistry (nutrients)

Nutrient amounts and nutrient ratios are essential for tree health, yield and fruit quality. Soil pH and relative amounts of some nutrients may limit the availability of uptake of key nutrients. As with any crop production, a portion of nutrients is removed with the crop that is harvested. Additional nutrients can be lost through leaching, erosion and volatilisation, and everything should be done to minimise this loss. Nutrients that are removed or lost need to be replaced to ensure crop production remains sustainable. As trees take macronutrients and micronutrients from the soil, it is crucial to monitor nutrient levels through soil testing to understand which nutrients need to be replaced. Below is a brief summary of the roles of some key nutrients (NZ Avocado Growers' Association Inc., 2018).

- Nitrogen is essential for the synthesis of proteins, including enzymes, DNA, RNA and hormones. Therefore, nitrogen has one of the most significant effects on tree behaviour. The use of nitrogen fertiliser is an important management tool for growers, but must be used carefully as there is a narrow optimal range. Excessive nitrogen can promote excessive leafy growth to the detriment of fruit production and quality. Traditionally spring applications of nitrogen have been deemed important for supporting vegetative flush, especially if flowering is heavy. More recently, this is coming under greater scrutiny in potentially limiting calcium uptake into fruit, negatively impacting fruit quality.
- Phosphorous is used to synthesise ATP (a cell's source of energy), DNA, RNA, cell membranes and sugar storage. Phosphorous is particularly important when trees are becoming established as it is key to good root development.
- Potassium is important for many key functions within the plant, including opening and closing of stomata and the ionic balance of cells. It can make up approximately 6% of the plant by dry

weight, and international literature suggests that it is removed in fruit at harvest at a greater rate than other nutrients.

- Calcium is a key element for plant growth and cell wall development. It is involved with hormone signalling pathways as well as cell wall integrity and permeability. Although it is thought to be a key element for disease management in avocados, there is little research to support this in New Zealand conditions; however, some international studies have correlated high levels of soil calcium with lower levels of post-harvest fruit rots.
- Boron is important for pollen germination, pollen tube growth, ovule viability and fruit set. It is also involved in carbohydrate metabolism and cell division. New Zealand research supports the importance of sufficient boron for fruitset, and boron is frequently applied as a foliar spray at the cauliflower stage of flowering to maximise fruitset.
- Magnesium is an essential component of chlorophyll, involved in photosynthesis and the conversion of sunlight into carbohydrates for the plant to use.
- Sulphur is a key element in protein synthesis.
- Trace elements are a range of other nutrients that are essential for normal plant growth and function. These include iron, manganese, zinc, copper, molybdenum and nickel. They are often essential as cofactors for enzymes. It is important to recognise that many of these are removed from the orchard with the fruit, and therefore, their replacement should be considered in any good fertiliser plan.

1.2 Soil biology

Avocado trees have shallow roots that are predominantly in the first 30 cm of topsoil (Figure 1). The zone where tree roots and microorganisms interact is called the rhizosphere, and several symbiotic relationships happen between tree roots and microorganisms.

Figure 1. Exposed feeder roots in the avocado tree (The Mid North, April 2021).

Some of the most relevant roles/functions of soil biology for avocado production are listed below.

1.2.1 Nutrient harvesting

Nutrient harvesting refers to the trees ability to uptake nutrients. This capability can be increased by root association with particular microorganisms. Mycorrhizae are among the most successful

associations between fungi and plant roots, including the avocado tree. The fungus colonises the roots and creates a network of fungal mycelia in the root cortex. The mycelium extends outside the roots into the soil and increases the plant's ability to absorb nutrients (Sullia, 1991). The tree supply's sugars to the mycorrhizae fungi in exchange for phosphorus and other nutrients that mycorrhizae help to collect.

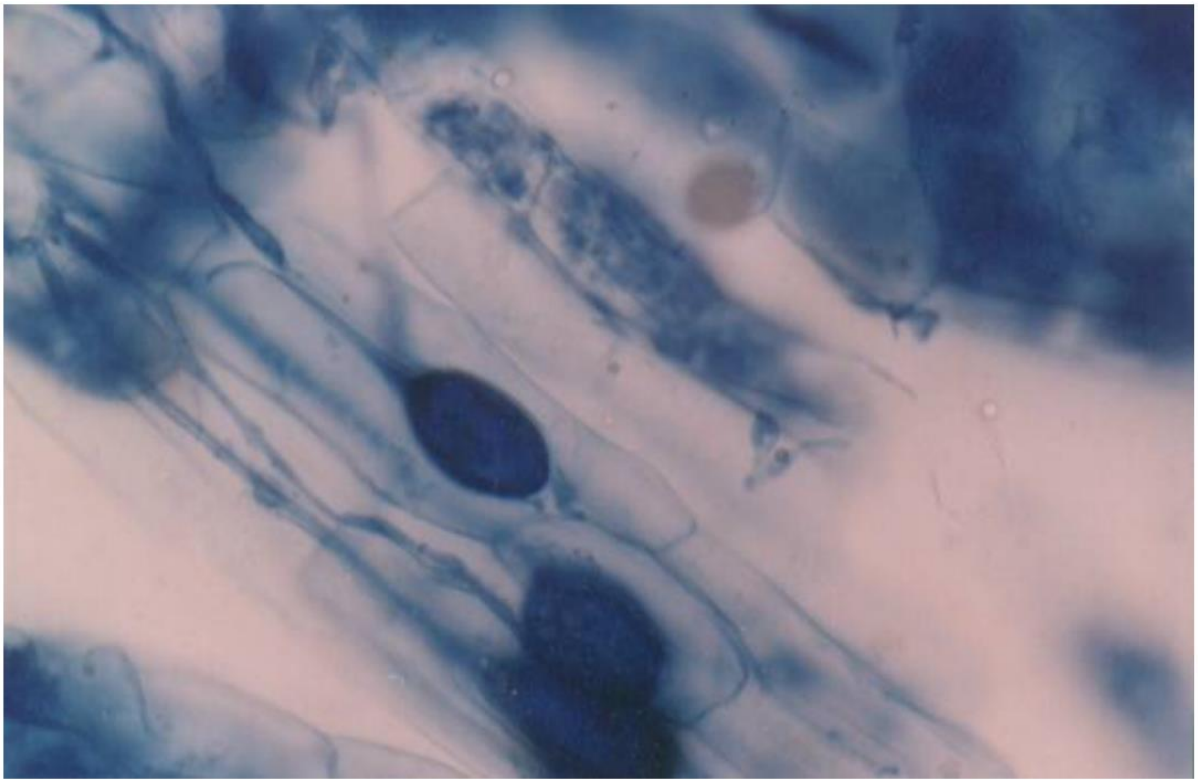

Figure 2. Mycorrhiza colonising the avocado roots (Herrera and Marcelo n.d.).

1.2.2 Nutrient cycling

Nutrient cycling refers to the flux of nutrients from one pool to another (e.g. carbon and nitrogen cycles). In the context of an orchard, the focus is generally on nutrients transitioning into a form that becomes available to the plant.

Carbon plays an essential role in the active and passive (hummus) fraction of soil organic matter—for instance, some of the soil's physical characteristics include water holding capacity and cation exchange capacity are influenced by carbon content. The first step in carbon cycling is typically done by saprophytic fungi that can catabolise some sources of carbon that are difficult to break down, such as lignin and cellulose. In addition, these fungi make bioproducts (e.g. organic acid) available for other microorganisms (Ingham, 2021). Another relevant microorganism genus is the *Actinomyces*. These hyphal bacteria can metabolise difficult to break down compounds such as chitin, cellulose and hemicellulose (Pepper, Gentry, and Gerba, 2015).

Nitrogen can be in different forms. Some are very easily lost from the soil through leaching or from volatilisation into gas. The nitrogen cycle is dominated by reduction and oxidation reactions where certain species of bacteria can help transform nitrogen into different forms. For example, certain species of bacteria are involved in the fixation of nitrogen (N_2) gas from the air into ammonia that plants can use. Nitrifying bacteria convert ammonia (NH_3) to nitrate (NO_3^-), which plants can use as well but is more easily leached from the soil than ammonia. Denitrifying bacteria convert nitrate to gaseous forms

of nitrogen like nitrous oxide (N_2O) and N_2, removing nitrogen from the soil. Denitrifying bacteria are more likely to convert nitrogen to a gaseous form under waterlogged or compact soil environments where oxygen is less available. *Pseudomonas* (Kumar et al., 2017) and protozoans also generate ammonia as part of their metabolism, adding to the nitrogen cycle.

1.2.3 Disease suppression

Among all the diseases that affect the avocado tree, root rot is the most important. This pathology is frequently associated with the oomycete *Phytophthora cinnamomi*. There are several root rot management methods, and when they are applied in combination, the control efficacy increases (Ramirez-Gil, Castaneda-Sanchez, and Morales-Osorio, 2016).

Some microorganisms can also reduce the impact of this disease. For example, some isolates from avocado roots closely related to *Bacillus acidiciler* have been shown to produce volatile metabolites that can reduce *P. cinnamomi* mycelial growth by 76%. (Méndez-Bravo et al., 2018).

Pseudomonas also play an essential role in disease suppression due to their ability to produce antibiotics and cell wall-degrading enzymes that target pathogenic microorganisms (Kumar et al., 2017). An interesting example is *Pseudomonas pseudoalcaligenes,* which controls white root rot disease in avocados caused by *Rosellinia necatrix* by competing for nutrients and spots in the roots (Pliego et al., 2019).

Ciliates are protozoans that can propel themselves and have predatory behaviours. This group feeds preferably on anaerobic bacteria, which can predominate in prolonged waterlogged soils.

1.2.4 Tree health

Soil biology can support tree health by making more nutrients available to the plant, suppressing pathogens, and producing phytohormones (e.g. *Pseudomonas,* which colonise roots, and *Actinomyces*, which can form hyphae and make phytohormones that stimulate the plant immune system (Kumar et al., 2017; Pepper, Ian L. and Gerba, Charles P., 2015). The benefit of mycorrhizae in tree health has been well known for decades, as it helps the tree absorb nutrients directly from soil (Menge et al., 1978).

1.2.5 Drought tolerance/soil structure

It is well known from other crops that certain microorganisms can increase drought tolerance; this is the case for *Actinomyces* (Grover et al., 2016; Pepper, Ian L. and Gerba, Charles P., 2015) and Mycorrhiza fungi (Li et al., 2019). These microorganisms colonise the soil by expanding their hyphae, opening the soil structure and effectively increasing the soil volume that the roots can extract water and nutrients from.

1.2.6 Earthworm numbers

Earthworms can be overlooked in soil analysis, but their presence is beneficial to tree health by making essential nutrients available to the tree (e.g. nitrogen, phosphorus, potassium and magnesium). Earthworms also improve soil structure improving porosity, aeration and water mobility. In pastures, soils with significant earthworms can have up to threefold more microorganisms and up to 6-7 times more *Actinomyces* than low worm soils (Shepherd, 2019). Different earthworms work at different depths, and a healthy community has a mixture of species (see table below).

Table 1. Common earthworms species in agricultural lands in New Zealand and the zone where they inhabit (Shepherd, 2019).

Earthworm species	Zone
Lumbricus rubellus	Superficial litter
Aporectodea caliginosa	Topsoil
Aporectodea longa	Subsoil

Earthworms also play an essential role in transforming organic matter into humus by bonding carbon to clay particles. Since this fraction is stable, it adds to the total organic carbon pool of the soil. However, there is limited information on the relationship between avocado orchards and earthworm's numbers. One study showed earthworms in avocado orchard soils avoided areas of copper contamination (>34 mg/kg)(Van Zwieten et al., 2004)

1.3 Soil physical properties

The physical structure of soil influences water dynamics, water holding capacity, root penetration and aeration. Key soil physical characteristics are described below.

1.3.1 Soil texture

Soil texture is determined by the proportion of different particles sizes in the soil. These particles are sand (>0.06 mm), silt (0.06-0.002 mm) and clay (<0.002 mm). Different proportions of these particles influence fundamental soil properties such as water-holding capacity, aeration, soil structure drainage, and nutrient retention.

1.3.2 Soil structure

Soil structure relates to the compaction of the soil, aggregates and clods of soil and the proportion of macro and micropores. A good soil structure is dominated by friable and fine aggregates with sub-rounded shape and no significant clodding. These soils have excellent water mobility, aeration, gas exchange capacity, soil temperature management and potential for root development. Conversely, a poor soil structure increases susceptibility to drought, ponding, and a higher risk of erosion. It also decreases the supply of oxygen to the roots, and therefore, can limit the availability of some nutrients such as nitrogen, phosphorus, calcium, magnesium, zinc, and boron (Shepherd, 2019).

Figure 3. Left: Avocado soil with good soil structure which is friable. Right: Avocado soil with a poor structure where aggregates are predominant.

1.3.3 Potential rooting depth

Potential rooting depth is the depth that roots can potentially explore before a physical barrier such as a hardpan prevents further root expansion. This indicator influences some essential elements in tree health, such as water holding capacity, availability of nutrients, and resilience against drought. Although avocado trees have the majority of their roots relatively shallow in the soil, deeper roots help support the tree and having a deep potential rooting depth is beneficial.

2 Methodology

Twenty nine orchards were sampled in this project with all production data, nutritional, biological and physical tests carried out on each orchard. Figure 4 depicts the number of orchards per region and the type of information collected for this project.

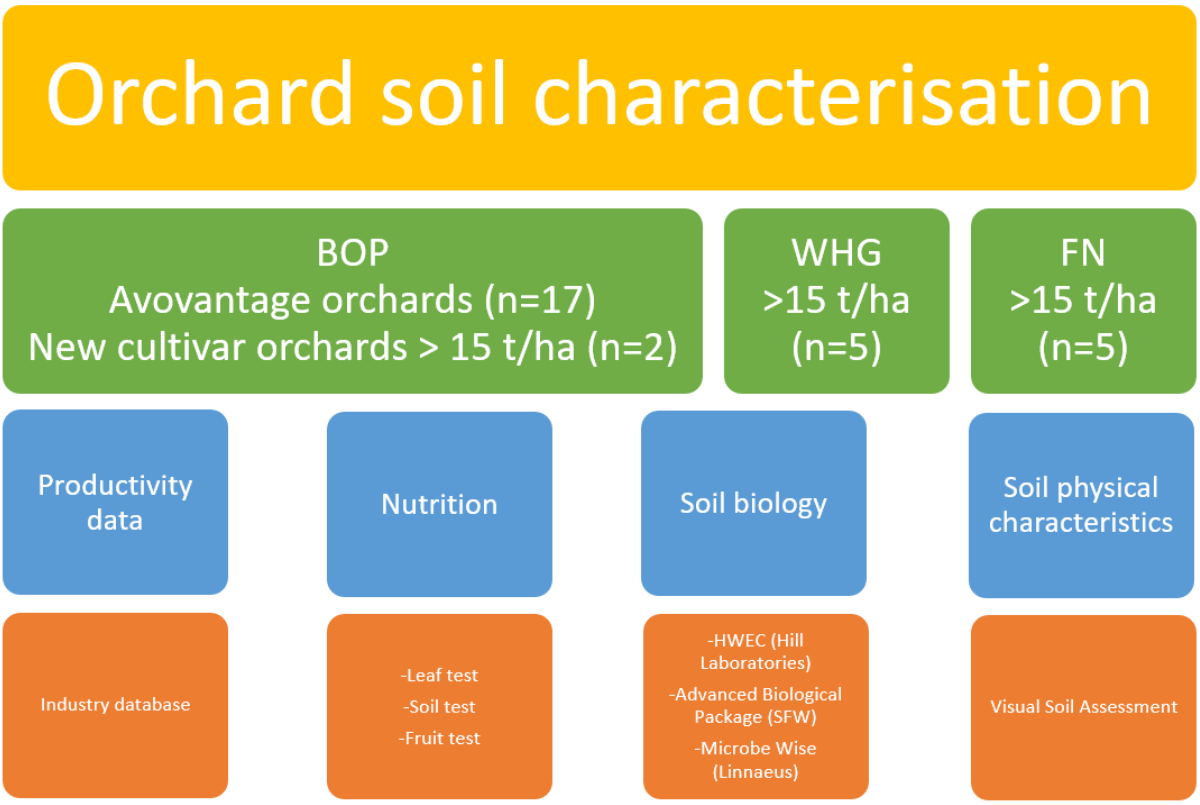

Figure 4. Overview of the Soil characterisation project with the different sources of information.

2.1 Orchard selection

Orchards from the three main growing regions of the Bay of Plenty (BOP), Mid North (MN) and Far North (FN) were selected:

- In the Bay of Plenty, 17 of the chosen orchards were part of the Avovantage project as there is a robust amount of data already collected from these trees relating to fruit quality, and the orchards represent a range of yields for comparison purposes.
- Additionally, in the Bay of Plenty, two orchards that are part of the New cultivar trials were also included to provide information on orchards achieving over 15t/ha.
- In the Mid and Far North, ten high yielding orchards (five in each region) producing above 15 t/ha over four years were included.

2.2 Tree selection

Ten trees were assessed in each orchard. Soil, root, leaf and fruit samples were collected during autumn 2021 for testing. Trees on the BOP orchards were pre-selected as they have been monitored as part of the Avovantage or New cultivar trials.
In the Mid and Far North, ten representative trees were selected in the best yielding block. The criteria for selecting the trees were:

- Represent the average condition (health and crop load) of trees in the block.
- Not edge/end of row trees.
- Planted in similar soil.
- Similar age.
- Hass scion with Zutano rootstock. If unavailable, this was recorded.
- Even distribution over the sampling area.

2.3 Sampling collection and delivery

Soil was sampled in a quadrant around each tree to ensure that the samples were a reflective aggregation from the trees and block; a similar protocol for leaves and fruit was implemented (Appendix 5.1). Additional to soil, leaf and fruit samples, roots were collected for mycorrhiza analysis in the Soil Foodweb test. (Appendix 5.2).

Samples were tested as follows:

- **Soil microbiology:** Soil samples were tested using three commercially available soil biology tests. Hot Water Extractable Carbon (HWEC), Advanced Biological Package, and Microbe Wise tests provided by Hill Laboratories, Soil Foodweb, and Linnaeus Laboratory, respectively.

- **Nutrient testing:** Soil, leaf and fruit samples underwent nutritional analysis by Hill Laboratory, including Basic soil, Organic soil profile, Total copper, Mehlich 3 profile, fruit nutrients and leaf nutrients.

2.4 Visual Soil Assessment (VSA)

The VSA was performed following the Scorecard's instructions (see Appendix 5.3.2) that reference the book Visual Soil Assessment, Vol 1 by Graham Shepherd Assessment includes a range of visual indicators relating to soil structure, soil texture, number of earthworms and potential rooting depth. It was assumed that the grower was uniformly treating the area within the dripline of each tree. Three VSAs per orchard/block were undertaken. See Appendix 5.3 for methodology.

2.5 Productivity data

The orchard 4-year average yields were taken from the industry database that gathers information directly from registered packhouses. Orchard selection was based on data from the 2016-17 to 2019-20 seasons, and yield data from the 2017-18 to 2020-21 seasons was used in the analysis as it became available later in the project. This did not change the orchard selection criteria for the high yielding orchards in the Mid and Far North.

2.6 Statistical analysis

Results were statistically analysed by NZ Avocado and Plant & Food Research using a variety of statistical analysis and machine learning tools. Machine learning potentially provides predictive variables and combinations of variables for yield outside the constraints of traditional statistical analysis where results are only significant if they have an accuracy of 90, 95 or 99%. See appendix 5.4 for methodology.

3 Results

3.1 Data summary

3.1.1 Productivity

Orchards were classified as either low or high yielding based on a four year average from the 2017-18 to 2020-21 seasons. Of the 29 selected orchards, 13 were classified as low yielding, and 16 were classified as high yielding. (Table 2).

Table 2. Selected orchards of high (H) and low (L) yielding as well as a yield information source. The orchard code start with the region's initial and then a random position.

Orchard code	Average yield (t/ha)	Category (yield)	Related project	Yield information; level
BOP-1	8.2	L	Avovantage	Last 4 seasons; whole orchard
BOP-2	7.6	L	Avovantage	Last 4 seasons; whole orchard
BOP-3	8.7	L	Avovantage	Last 4 seasons; whole orchard
BOP-4	9.4	L	Avovantage	Last 4 seasons; whole orchard
BOP-5	10.1	L	Avovantage	Last 4 seasons; whole orchard
BOP-6	11	L	Avovantage	Last 4 seasons; whole orchard
BOP-7	11.1	L	Avovantage	Last 4 seasons; whole orchard
BOP-8	11.5	L	Avovantage	Last 4 seasons; whole orchard
BOP-9	11.9	L	Avovantage	Last 4 seasons; whole orchard
BOP-10	11.9	L	Avovantage	Last 4 seasons; whole orchard
BOP-11	12	L	Avovantage	Last 4 seasons; whole orchard
BOP-12	12.3	L	Avovantage	Last 4 seasons; whole orchard
BOP-13	14.6	L	Avovantage	Last 4 seasons; whole orchard
BOP-14	18.1	H	New cultivar trial	Last 4 seasons; block level
BOP-15	18.1	H	Avovantage	Last 4 seasons; whole orchard
BOP-16	18.3	H	Avovantage	Last 4 seasons; whole orchard
BOP-17	19.6	H	Avovantage	Last 4 seasons; whole orchard
BOP-18	17.6	H	New cultivar trial	Last 3 seasons; block level
BOP-19	26.5	H	Avovantage	Last 4 seasons; whole orchard
FN-1	15	H	NA	Last 4 seasons; whole orchard
FN-2	15	H	NA	Last 4 seasons; whole orchard
FN-3	16.6	H	NA	Last 4 seasons; whole orchard
FN-4	17.9	H	NA	Last 4 seasons; whole orchard
FN-5	19	H	NA	Last 4 seasons; whole orchard
MN-1	15	H	NA	Last 4 seasons; whole orchard
MN-2	16.4	H	NA	Last 4 seasons; whole orchard
MN-3	16.6	H	NA	Last 4 seasons; whole orchard
MN-4	18.9	H	NA	Last 4 seasons; whole orchard
MN-5	21.4	H	NA	Last 4 seasons; whole orchard

3.1.2 Soil biology, nutrients and visual soil assessment

There were large ranges for some biological variables. For example, flagellates and amoebae (Table 4), and Hot Water Extractable Carbon (Table 6). There was also insufficient variation in some variables, such as the number and colour of soil mottles (Table 3). Initial analysis and observations highlighted variables with insufficient variation and variables with large missing data gaps. These were excluded from further statistical analysis because variables with low variability provide no benefit to the model, and missing data decreases the dataset's reliability.

Table 3 to Table 7 summarises variables measured by soil visual assessment, Soil Foodweb, Linnaeus and Hill Laboratories. Ranges are the minimum and maximum of results, not recommended ranges.

The visual soil assessment methodology scores from 0 (poor condition) to 2 (good condition) across all the variables. All the orchards assessed had total scores over 28 that classified their soil in the best possible classification of good quality according to Graham Shepherd methodology (see Appendix 5.3).

Table 3. Regional summary of soil visual assessment variables.

Name	Unit	BOP <15t/ha Average (min-max)	BOP >15t/ha Average (min-max)	MN Average (min-max)	FN Average (min-max)	Target
Soil texture	Indicator	1.3(1.0-1.5)	1.5(1.3-2.0)	1.3(0.8-1.5)	1.8(1.5-2.0)	2.0
Soil structure	Indicator	1.6(0.9-2.0)	1.9(1.7-2.0)	2.0(1.8-2.0)	1.9(1.5-2.0)	2.0
Number and colour of soil mottles	Indicator	2.0(1.8-2.0)	2.0(2.0-2.0)	2.0(2.0-2.0)	2.0(2.0-2.0)	2.0
Soil colour	Indicator	1.9(1.0-2.0)	2.0(2.0-2.0)	2.0(2.0-2.0)	2.0(2.0-2.0)	2.0
Earthworms	score	0(0-0)	0(0-0)	0(0-0)	0(0-0)	2.0
Soil smell	Indicator	1.8(1.3-2.0)	1.9(1.3-2.0)	2.0(2.0-2.0)	1.9(1.3-2.0)	2.0
Potential rooting depth	Indicator	2.0(1.8-2.0)	2.0(2.0-2.0)	1.7(0.8-2.0)	1.7(1.3-2.0)	2.0

Table 4. Regional summary of Soil Foodweb biological variables with labels used in this report.

Name	Label	Unit	BOP <15t/ha Average (min-max)	BOP >15t/ha Average (min-max)	MN Average (min-max)	FN Average (min-max)	Target
Dry Weight	SF01	---	0.61(0.57-0.68)	0.62(0.5-0.73)	0.61(0.57-0.64)	0.79(0.71-0.85)	0.45-0.85
Active Bacteria (AB)	SF02	mg/kg	30.1(15.1-38.1)	25.2-13.3-35.2)	37.0(29.1-46.0)	20.8(16.9-24.0)	>30
Total Bacteria (TB)	SF03	mg/kg	361.5(245.4-482.6)	319.7(236.9-535.2)	391.0(246.2-487.8)	282.6(237.9-345.7)	>300
Actinobacteria	SF04	mg/kg					<20
Active Fungi (AF)	SF05	mg/kg	3.61(0.08-10.39)	3.07(0.07-9.02)	11.28(3.52-16.35)	1.36(0.06-3.01)	>150
Total Fungi (TF)	SF06	mg/kg	384.4(204.8-513.4)	468.6(136.6-911.0)	227.7(140.9-445.9)	259.0(176.7-412.8)	>1,500
Hyphal Diameter	SF07	µm	2.92(2.75-3.00)	3.0(3.0-3.0)	2.85(2.75-3.00)	2.90(2.75-3.00)	
Flagellates	SF08	number/g	17,276(2,038-55,500)	11,630(782-44,249)	13,909(4,450-43,325)	7,918(3,266-18,588)	>20,000
Amoebae	SF09	number/g	10,613(2,038-24,426)	7,111(2,237-9,183)	5,182(733-9,376)	2,150(371-7,247)	>20,000
Ciliates	SF10	number/g	177(43-472)	391(79-836)	563(433-808)	138(38-329)	<334
Endo (colonization)	SF11	%	0.55(0.07-0.81)	0.64(0.29-0.9)	0.45(0.25-0.65)	0.45(0.10-0.64)	>40
TF/TB	SF12	ratio	1.09(0.61-1.52)	1.47(0.53-2.39)	0.58(0.31-1.01)	0.97(0.51-1.73)	5-10
AF/TF	SF13	ratio	0.01(0-0.02)	0.00(0.00-0.01)	0.05(0.02-0.11)	0.01(0.00-0.01)	>0.10
AB/TB	SF14	ratio	0.08(0.04-0.11)	0.09(0.03-0.13)	0.10(0.08-0.15)	0.07(0.07-0.08)	>0.10
AF/AB	SF15	ratio	0.13(0-0.33)	0.17(0.00-0.68)	0.30(0.12-0.45)	0.07(0.00-0.17)	5-10

Table 5. Linnaeus biological variables with labels used in this report.

Name	Label	Unit	BOP <15t/ha Average (min-max)	BOP >15t/ha Average (min-max)	MN Average (min-max)	FN Average (min-max)	Guide
Actinomycetes	LN01	mg/kg	2.9(1.1-5.4)	2.8(1.8-3.5)	2.5(2.0-3.0)	1.7(1.09-2.56)	1.0
Ammonium (NH4) N Before Incubation	LN02	mg/kg					
Ammonium (NH4) N to Nitrate NO3) N Conversion	LN03	%/month					
Bacteria Stress Indicator	LN04	Indicator	0.5(0.4-0.8)	0.5(0.4-0.6)	0.6(0.5-0.8)	0.46(0.40-0.50)	<0.5
Carbon to Nitrogen Ratio	LN05	Ratio					
Disease Resistance - MWSE only	LN06	Indicator	88.8(79.9-100)	90.0(81.6-100)	88.1(80.5-100)	89.5(77.6-100)	70-100
Drought Resistance - MWSE only	LN07	Indicator	83.1(69.8-100)	85.0(72.4-100)	82.1(70.7-100)	86.0(74.9-100)	70-100
Fungi to Bacteria Ratio	LN08	Ratio	2.7(2.0-3.6)	2.9(2.5-3.6)	2.5(2.3-3.0)	3.4(2.9-4.5)	2.3
Gram Negative Bacteria	LN09	mg/kg	7.3(3.6-14.3)	7.5(4.4-8.8)	5.8(4.6-8.3)	4.88(2.83-8.38)	11.0
Gram Positive Bacteria	LN10	mg/kg	9.3(3.9-15.9)	9.1(6.0-11.7)	8.0(6.0-10.7)	6.8(4.9-9.8)	4.0
Methane Oxidising Bacteria	LN11	mg/kg					
Microbial Balance	LN12	Indicator	84.4(73.3-93.6)	85.9(74.8-91.3)	80.7(72.3-91.2)	80.3(71.7-90.7)	70-100
Microbial Diversity Indicator	LN13	Indicator	35.9(32.8-38.6)	34.8(30.8-37.9)	37.3(36.2-38.8)	31.8(28.3-34.2)	80.0
Mycorrhizal Fungi (AMF)	LN14	mg/kg	7.0(4.0-13.6)	7.1(4.5-10.7)	8.0(4.1-17.8)	7.4(5.0-10.8)	10.0
Nitrate (NO3) N Before Incubation	LN15	mg/kg					
Nitrogen (N) Fixed	LN16	mg/kg/month					
Nitrogen Mineralised - Estimated by Indices	LN17	mg/kg/month					
Nutrient Accessibility (VAM) - MWSE only	LN18	Indicator	66.5(39.7-100)	70.0(44.8-100)	64.3(41.4-100)	71.9(49.9-100)	70-100
Nutrient Cycling Rate - MWSE only	LN19	Indicator	89.7(78.4-100)	91.1(80.0-95.0)	82.2(71.2-93.9)	85.2(80.1-93.5)	70-100
Nutrient Solubilisation Rate - MWSE only	LN20	Indicator	83.2(69.8-100)	85.0(72.4-100)	82.1(70.7-100)	84.3(66.4-100)	70-100
Organic N (%)	LN21	%					
Percentage of Total N Mineralised	LN22	% of Total N/month					
Protozoa	LN23	mg/kg	2.0(1.1-3.1)	2.2(1.4-2.9)	1.7(0.8-3.8)	1.65(1.3-1.9)	1.3
Pseudomonas	LN24	mg/kg	2.0(1.1-4.1)	2.2(1.1-3.0)	1.7(1.2-2.9)	1.4(0.83-2.34)	1.0
Residue Breakdown Rate - MWSE only	LN25	Indicator	98.8(90.4-100)	98.4(90.1-100)	94.4(89.7-100)	98.4(92.3-100)	70-100
Sulphur Reducing Bacteria	LN26	mg/kg					
Total Bacteria	LN27	mg/kg	16.6(7.5-30.3)	16.5(10.3-20.3)	13.8(10.6-19.0)	11.6(7.7-18.1)	15.0
Total Fungi	LN28	mg/kg	44.0(27.3-77.7)	48.8(27.1-72.2)	35.7(26.9-56.6)	38.7(28.6-60.8)	33.8
Total Microorganisms	LN29	mg/kg	62.7(36.4-108.9)	67.5(38.8-94.4)	51.2(38.3-79.4)	52.0(40.4-80.2)	50
Total Nitrogen (N) Before Incubation	LN30	%					
Total Organic Carbon (C) Before Incubation	LN31	%					
True Anaerobic Bacteria	LN32	mg/kg	0.5(0.1-2.2)	0.5(0.3-0.8)	0.4(0.3-0.5)	0.2(0.1-0.3)	<0.05

Table 6. Hill Laboratories soil variables with labels used in this report. The highlighted variable (HS13) is considered biological for this report; the remaining variables are nutrient levels and ratios.

Name	Label	Unit	BOP <15t/ha Average (min-max)	BOP >15t/ha Average (min-max)	MN Average (min-max)	FN Average (min-max)
pH	HS01	pH	5.7(4.8-6.3)	5.7(5.1-6.2)	5.8(5.3-6.2)	6.2(6.1-6.4)
Olsen Phosphorus	HS02	mg/L	117(41-346)	108(42-213)	96(55-186)	47(24-75)
Potassium	HS03	MAF	16.5(10-33)	13(7-18)	21.2(14-29)	9(7-10)
Calcium	HS04	MAF	11.5(6-17)	12.7(8-17)	14.6(12-19)	9(5-14)
Magnesium	HS05	MAF	37.6(22-57)	36.7(22-67)	55.2(40-78)	31.6(18-44)
Sodium	HS06	MAF	3.3(2-9	3.4(2-6)	5(3-8)	8(6-10)
Potentially Available Nitrogen (15cm Depth)	HS07	kg/ha	129.2(85-159)	159.7(125-237)	230.6(204-253)	120.4(82-150)
Anaerobically Mineralisable N	HS08	µg/g	127.6(76-169)	163.8(120-286)	203.4(172-242)	80(51-101)
Anaerobipcally Mineralisable N/Total N Ratio	HS09	%	2.0(1.4-2.4)	2.2(1.7-2.8)	3.1(2.6-3.6)	2.2(1.8-2.7)
Organic Matter	HS10	%	12.7(8.6-15.7)	15.2(12.2-23)	12.8(11.7-14.9)	9.9(7.5-13.7)
C/N Ratio	HS11	---	11.3(10.2-12.1)	12.0(10.4-13.1)	11.4(10.3-13.2)	16.7(12-24.7)
Total Carbon	HS12	%	7.4(5-9.1)	8.8(7.1-13.3)	7.4(6.8-8.6)	5.7(4.3-7.9)
Total Nitrogen	HS14	%	0.65(0.48-0.8)	0.73(0.55-1.02)	0.65(0.63-0.67)	0.35(0.25-0.43)
Phosphorus (Mehlich 3)	HS15	mg/L	138(49-434)	142.3(69-244)	122.6(71-232)	154.6(67-309)
Sulphur (Mehlich 3)	HS16	mg/L	71(26-153)	57(25-98)	72(20-139)	42(29-61)
Potassium (Mehlich 3)	HS17	mg/L	294(172-587)	230(125-303)	367(242-492)	151(119-176)
Calcium (Mehlich 3)	HS18	mg/L	1,845(933-2,680)	1,966(1,239-2,740)	2,094(1,806-2,710)	1,392(785-2,310)
Magnesium (Mehlich 3)	HS19	mg/L	195.8(112.1-300)	190.9(103.8-321)	282.4(220-410)	171.5(96.1-247)
Sodium (Mehlich 3)	HS20	mg/L	16(10-43)	16(10-28)	23.8(13-37)	39(28-48)
Iron (Mehlich 3)	HS21	mg/L	74.5(52-156)	77.7(67-88)	108.2(97-132)	133(87-189)
Manganese (Mehlich 3)	HS22	mg/L	20.7(10.1-34)	21.4(8.3-42.2)	79.5(55.7-98.1)	8.7(3.5-12.2)
Zinc (Mehlich 3)	HS23	mg/L	41.8(7.8-71.2)	51.8(17.4-81.8)	26.2(11.5-43.5)	34.4(5.7-117.8)
Copper (Mehlich 3)	HS24	mg/L	18.9(2.9-41.9)	32.6(18.3-55.1)	13.5(3.6-28.3)	15.2(0.3-54.2)
Boron (Mehlich 3)	HS25	mg/L	4.93(1.58-10.24)	6.78(1.98-11.65)	6.8(1.8-16.9)	2.17(1.24-4.52)
Cobalt (Mehlich 3)	HS26	mg/L	<0.1(<0.1-<0.1)	<0.1(<0.1-0.1)	0.2(0.1-0.4)	<0.1(<0.1-<0.1)
Aluminium (Mehlich 3)	HS27	mg/L	1,551(1,388-1,704)	1,496(1,379-1,652)	913(739-1,174)	1,609(1,441-1,926)
Total Copper	HS28	mg/kg	103(27-193)	163(101-280)	87(43-142)	55(<4-158)
Potassium	HS29	me/100g	1.17(0.71-2.1)	0.97(0.45-1.31)	1.37(0.85-1.91)	0.41(0.36-0.46)
Calcium	HS30	me/100g	13.8(5.9-21.2)	15.4(8.7-21.2)	15.5(12.1-19.6)	6.7(3.4-11.3)
Magnesium	HS31	me/100g	2.47(1.45-3.95)	2.46(1.38-4.38)	3.22(2.2-4.44)	1.41(0.74-2.03)
Sodium	HS32	me/100g	0.1(0.06-0.27)	0.10(0.06-0.18)	0.14(0.08-0.22)	0.17(0.12-0.21)
Potassium	HS33	%BS	3.9(2.6-7.9)	2.95(2.2-3.6)	1.3(3-6)	3.3(2.3-4.4)
Calcium	HS34	%BS	45(23-59)	47(38-59)	48(38-59)	50.6(37-64)
Magnesium	HS35	%BS	8.0(4.8-11.1)	7.5(4.6-11.9)	10.1(7-13.5)	10.7(8-12.9)
Sodium	HS36	%BS	0.3(0.2-0.9)	0.3(0.2-0.6)	0.5(0.3-0.7)	1.4(0.8-2)
CEC	HS37	me/100g	31(23-45)	33(21-46)	32(28-35)	13(9-18)
Total Base Saturation	HS38	%	57(38-74)	58(45-71)	63(50-77)	66(51-79)
Volume Weight	HS39	g/mL	0.68(0.62-0.79)	0.67(0.55-0.73)	0.76(0.69-0.8)	1.01(0.97-1.07)
Hot Water Extractable Carbon	HS13	mg/kg	1,560(882-2,422)	1,892(1,331-2,887)	1,840(1,124-2,496)	942(747-1,240)

Table 7. Hill laboratories leaf tissues soil variables.

Name	Units	BOP <15t/ha Average (min-max)	BOP >15t/ha Average (min-max)	MN Average (min-max)	FN Average (min-max)
Nitrogen	%	2.2(1.8-2.6)	2.2(1.8-3)	2.4(2.3-2.6)	2.5(2.1-2.7)
Phosphorus	%	0.11(0.08-0.15)	0.12(0.1-0.18)	0.15(0.12-0.16)	0.15(0.11-0.19)
Potassium	%	1.1(0.7-1.3)	1.0(0.9-1.2)	1.2(1.0-1.4)	1.0(0.8-1.1)
Sulphur	%	0.26(0.2-0.32)	0.30(0.25-0.38)	0.31(0.27-0.41)	0.29(0.23-0.36)
Calcium	%	2.05(1.62-2.65)	2.40(2.03-2.92)	2.23(1.56-2.83)	1.50(1.28-1.75)
Magnesium	%	0.38(0.31-0.47)	0.43(0.35-0.56)	0.48(0.37-0.53)	0.47(0.43-0.56)
Sodium	%	0.003(0.002-0.005)	0.005(0.003-0.008)	0.004(0.003-0.006)	0.007(0.005-0.01)
Iron	mg/kg	48(36-62)	54(38-83)	57(48-66)	49(41-65)
Manganese	mg/kg	209(71-630)	163(84-240)	330(220-570)	68(22-104)
Zinc	mg/kg	39(27-55)	43(35-51)	36(21-53)	38(21-59)
Copper	mg/kg	206(32-910)	126(32-197)	232(9-590)	129(6-230)
Boron	mg/kg	35(20-48)	31(21-35)	37(20-48)	53(32-79)
Chloride	%	0.17(0.07-0.24)	0.20(0.09-0.34)	0.15(0.09-0.21)	0.24(0.18-0.39)

3.2 Statistical analysis

3.2.1 Benchmarking of the variables between biological soil tests.

The Soil Foodweb, Linnaeus and Hill Laboratories soil biological tests use different methodologies to measure soil biological variables. Soil Foodweb use microscopy-based observations and counts to quantify different microbial populations; Linnaeus uses molecular markers to quantify the amounts of different microbial species, and the Hill Laboratories Hot Water Extractable Carbon test quantifies the labile fraction of soil carbon that has been shown to correlate with soil microbial biomass and is sensitive to changes in soil quality. While the Hill's Hot Water Extractable Carbon test provides a single value as a measure of the total microbial biomass in the soil, some of the variables are shared between the Soil Foodweb and Linnaeus tests. A comparison between these shared variables was undertaken to see if growers may be able to use these tests interchangeably. The two common and comparable variables were Total fungi and Total bacteria, where the values of Soil Foodweb were on average 9 and 25 times higher than Linnaeus, respectively. The ranges or thresholds for a "good" level were also different (see Table 8). Since both laboratories use different techniques, some variances were expected, but the variances were not proportional to each other. For instance, the Linnaeus result for orchard BOP-3 indicated a good level of Total fungi while Soil Foodweb results indicated low levels, and BOP – 14 had the highest value for the Soil Foodweb Total fungi test but one of the lowest for the Linnaeus test.

Table 8. Compare results from Soil Foodweb and Linnaeus for the 29 orchards in the only two common variables.

Lab	Name	Units	Average (min-max) High yield	Average (min-max) Lower yield	Target
SFW	Total fungi	mg/kg	328(137-911)	384(205-513)	>1500
Linnaeus	Total fungi	mg/kg	41.5(26.9-72.2)	44.0(27.3-77.7)	34
SFW	Total bacteria	mg/kg	330(237-565)	362(245-483)	>300
Linnaeus	Total bacteria	mg/kg	14.2(7.7-20.3)	16.6(7.5-30.3)	15

Figure 5 and Figure 6 show Total fungi and Total bacteria values for Soil Foodweb (SFW) and Linnaeus. The subsamples send to the laboratories came from the same original homogeneous sample taken from the orchard.

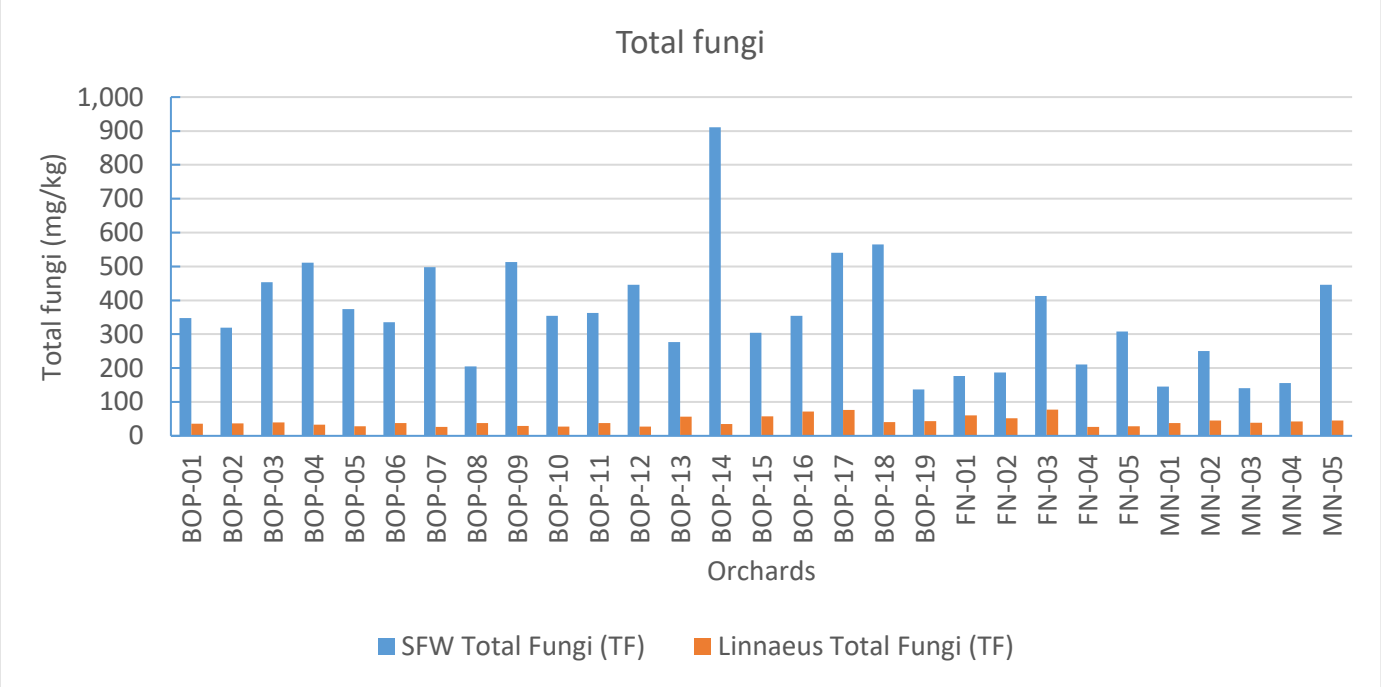

Figure 5.Values of Total fungi quantified by Soil Foodweb (SFW) and Linnaeus.

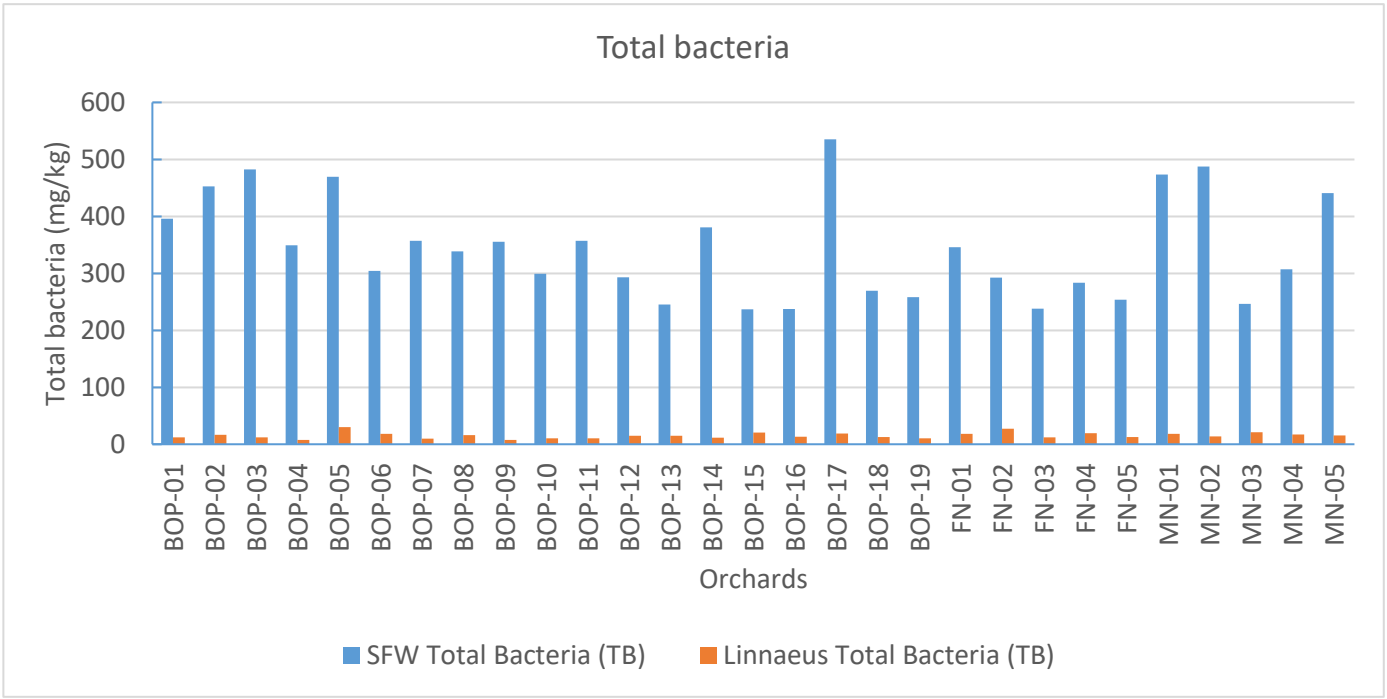

Figure 6. Values of Total bacteria quantified by Soil Foodweb (SFW) and Linnaeus.

Both laboratories also quantified mycorrhizae and *actinomycetes*. However, Soil Foodweb provides a percentage of mycorrhizae colonisation while Linnaeus reports mycorrhizae biomass in mg/kg; therefore, a comparison was not possible. Soil Foodweb did not detect *actinomycetes* in any of the soil samples (0 mg/kg in all samples); in contrast, Linnaeus provided a range from 1.1 to 5.4 mg/kg.

3.2.2 Yield predictability by individual leaf and fruit variables

Linear regression was also carried out to look for trends in nutrient concentration in leaf and fruit samples with orchard yield but poor correlations was observed (R^2= 0.005-0.04, leaf macronutrients). Table 9 summarises all the nutrient values. All the charts are available in Appendix 5.5.

Table 9. Summary of all R-square values of linear regression of selected nutrient averages between the low and high yield classification in the different plant tissues tested.

Plant tissue	Name	Label	Units	Average (min-max) High yield	Average (min-max) Lower yield	R-square
Flesh	Nitrogen	HST1	g/100g	1.28(0.22-1.88)	1.4(0.86-1.96)	0.11285
Flesh	Potassium	HST2	g/kg	15.6(12.2-19.9)	14.(11.4-16.9)	0.00360
Flesh	Phosphorous	HST3	g/kg	1.71(1.44-2.00)	1.56(1.05-1.82)	0.07482
Flesh	Boron	HST4	mg/kg	115(46-186)	99(52-147)	0.00008
Flesh	Calcium	HST5	mg/kg	926(550-1340)	1051(210-1450)	0.00368
Skin	Nitrogen	HST6	g/100g	0.98(0.85-1.18)	1.02(0.85-1.25)	0.05349
Skin	Potassium	HST7	g/kg	10.7(8.8-12.7)	11.7(9.3-21.0)	0.09077
Skin	Phosphorous	HST8	mg/kg	914(820-1050)	976(790-1330)	0.09477
Skin	Boron	HST9	mg/kg	59(30-99)	57(33-130)	0.02492
Skin	Calcium	HST10	mg/kg	1141(790-1750)	1379(610-2100)	0.06928
Leaves	Nitrogen	HST11	%	2.4(1.8-3.0)	2.2(1.8-2.6)	0.04173
Leaves	Potassium	HST12	%	1.1(0.8-1.4)	1.1(0.7-1.3)	0.03016
Leaves	Phosphorous	HST13	%	0.14(0.1-0.19)	0.11(0.08-0.15)	0.05180
Leaves	Boron	HST14	mg/kg	39.8(20-79)	35.3(20-48)	0.00713
Leaves	Calcium	HST15	%	2.06(1.28-2.92)	2.05(1.62-2.65)	0.00581

3.2.3 Yield predictability by individual soil variables

Statistically significant differences were detected in some of the soil chemical and physical variables measured. HS09 (Anaerobically Mineralisable N/Total N Ratio) and HS11 (C/N Ratio) were both higher for orchards in the high yielding classification, but ranges of values in each overlapped (both p-values = 0.03). Earthworm counts were also significantly different with lower counts from the high yielding orchard group. There was no statistically significant variable across all the other chemical, biological, and physical soil variables measured (p-values > 0.05). A full table of results is listed in Appendix 5.6.

The number of orchards in the Bay of Plenty were considerably higher than the other two regions, and all low yielding orchards were located in the Bay of Plenty. Statistical analysis of a regional effect was not undertaken due to data limitations (i.e. small sample size and absence of low yielding orchards in the Mid and Far North).

3.2.4 Inter-correlation among variables and multivariate analysis

Some intercorrelation of variables was seen within the reports from the different laboratories. If many different variables are correlated with each other, it may be possible to exclude some to save on testing costs while using other variables as proxies for the correlated variable. Intercorrelated variables are also less likely to help explain variability in yield across orchards beyond what one of these variables can indicate. Statistical analysis was undertaken to identify the potential correlation and combination of variables that predict yield.

Linnaeus variables had the highest intercorrelation, as indicated by the more clustered points on the Biplot (Figure 7). Conversely, the intercorrelation among the Hill Laboratories soil nutrient and Soil Foodweb variables were comparatively lower (Figure7).

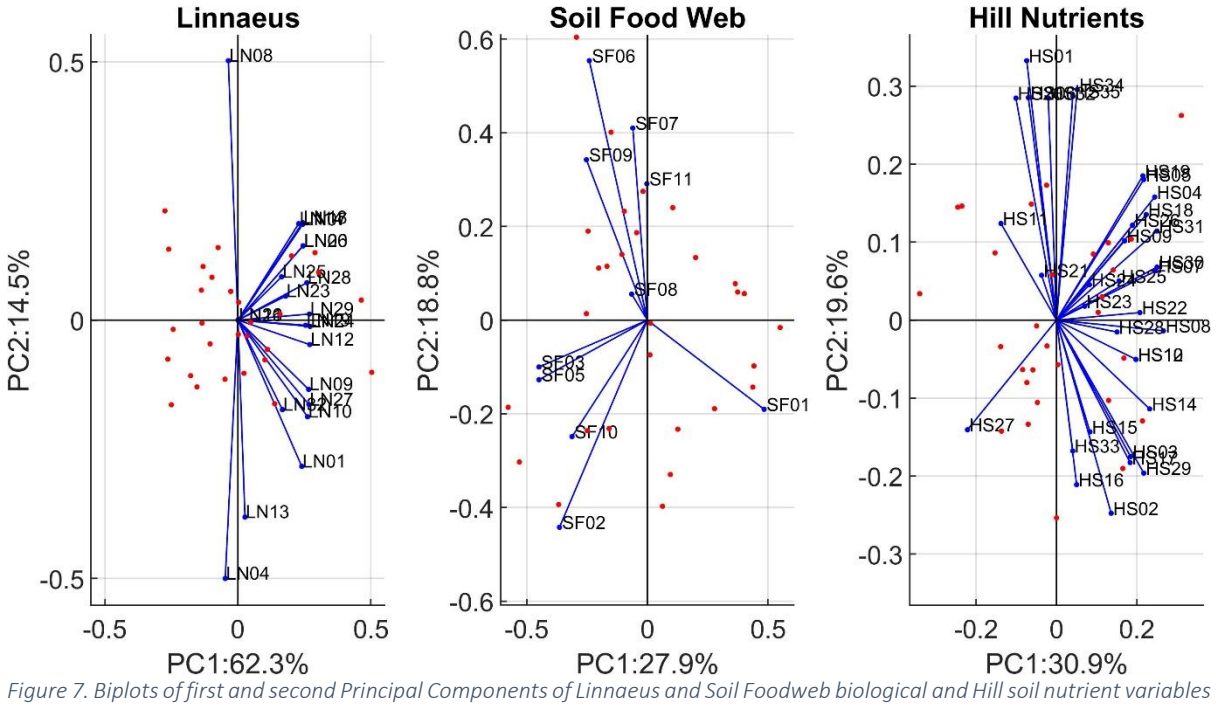

Figure 7. Biplots of first and second Principal Components of Linnaeus and Soil Foodweb biological and Hill soil nutrient variables for all orchards. The scatter dots are approximate locations of the data points (orchards) in the PC coordinate system.

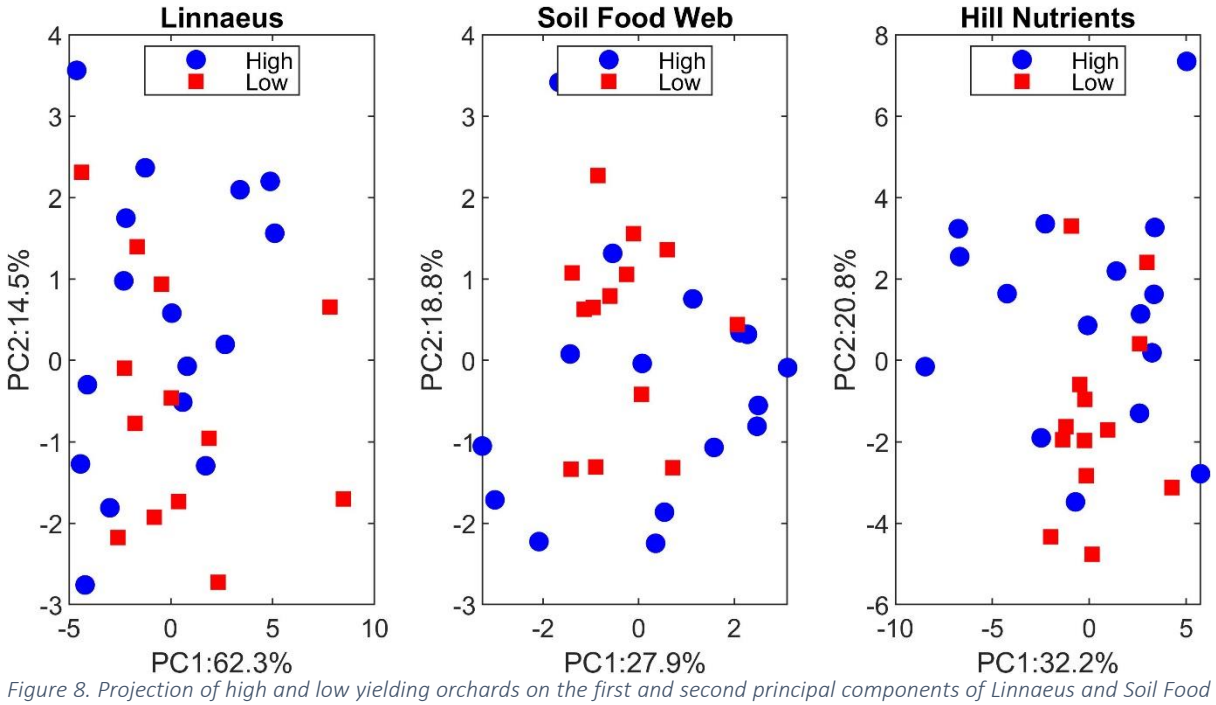

Figure 8. Projection of high and low yielding orchards on the first and second principal components of Linnaeus and Soil Food Web biological, and Hill soil nutrient variables. All variables were normalised before projection.

The variables from each test did not correlate particularly well with the high and low yield classifications. A strong correlation between the variables and the low and high yielding orchard class would be observed by clear clusters of high yielding and low yielding orchards in the panels of Figure 8. The panels show considerable overlap between the two classes in all datasets, indicating that each test does not correlate well with the high and low yield classes. No set of variables predicted orchard yield with 95% confidence.

3.3 Machine learning

3.3.1 Classification trees (yield predictability by sets of variables)

Machine learning was undertaken to identify any potential predictors and combinations of predictors for low and high yielding orchards. The algorithm was run 100 times on 24 orchard results and then tested against the remaining five orchards in the data set to check accuracy. Importance estimates identified variables that were predictors of yields, and classification trees were developed as a potential decision-making tool for growers to try to identify thresholds for the important variables influencing yield.

3.3.1.1 Soil Foodweb

Ten Soil Foodweb biological variables were considered as predictors for classifying between high and low yielding orchards. From the 100 iterations of the decision trees generated with data from 24 orchards, the maximum classification accuracy on the remaining five orchards was 76% by the Soil Foodweb variables.

Ciliates (SF10) was the most important variable in predicting yield, followed by Total fungi (SF06), Flagellates (SF08), Dry weight (SF01) and Total bacteria (SF03) (Figure 6 and Appendix 5.7.15.6). It was observed that high yielding orchards tended to have higher levels of ciliates, higher dry weight and lower levels of total bacteria compared to the low performing orchards.

The importance of each variable in classifying whether an orchard is high or low yielding is shown in Figure 9.

Appendix 5.8.1 shows one of the 100 decision trees tested as part of the machine learning process that helps to define what variables are important and what thresholds may be relevant. Unfortunately, thresholds could not be defined with any level of confidence that incorporated all the important variables identified.

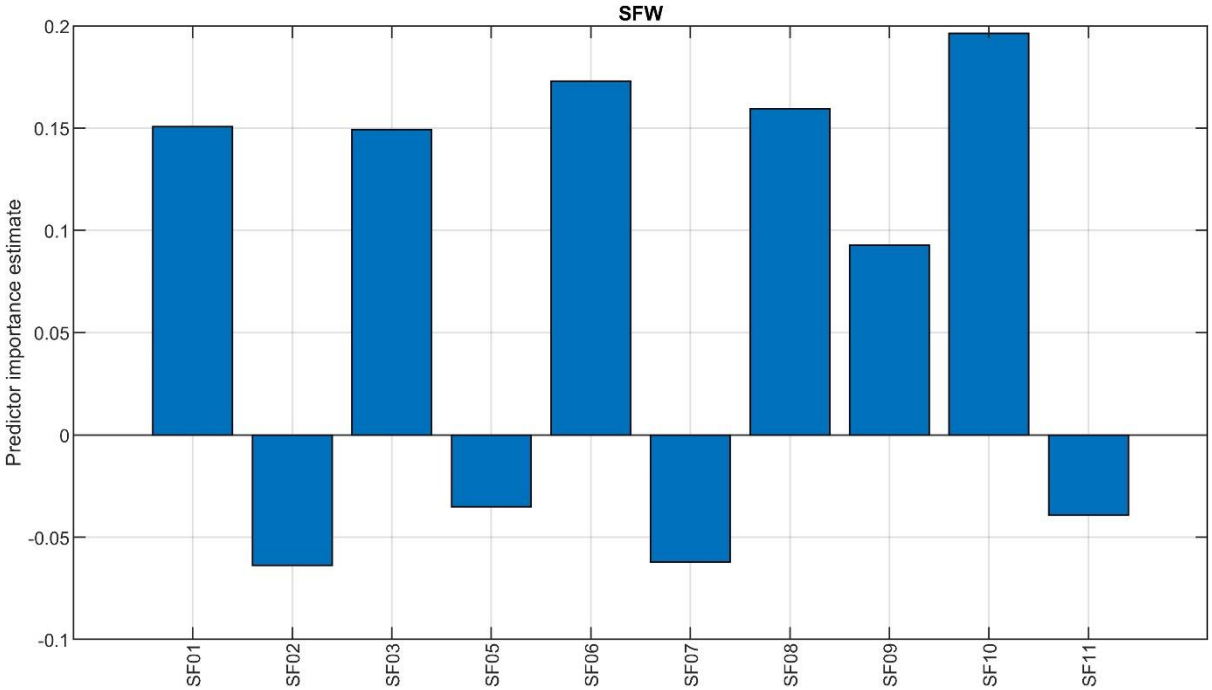

Figure 9. Importance estimate of Soil Foodweb biological variables in classifying between low and high yielding orchards. The higher the 'predictor importance estimate' (y-axis), the better the variable is at predicting yield.

3.3.1.2 Linnaeus laboratories

Twenty Linnaeus biological variables were considered as predictors for classifying between high and low yielding orchards. From the 100 iterations of the decision trees generated with data from 24 orchards, the average classification accuracy on the remaining five orchards was 62% by the Linnaeus variables.

Gram-negative bacteria (LN09) was the most important variable in classification between high and low yielding orchards, with high yielding orchards having lower gram-negative bacteria counts. (Figure 10 and Appendix 5.7.2). The importance metric of the other variables was either negative or close to zero.

Appendix 5.8.2 shows one of the 100 decision trees tested as part of the machine learning process that helps to define what variables are important and what thresholds may be relevant. Unfortunately, thresholds could not be defined with any level of confidence that incorporated all the important variables identified.

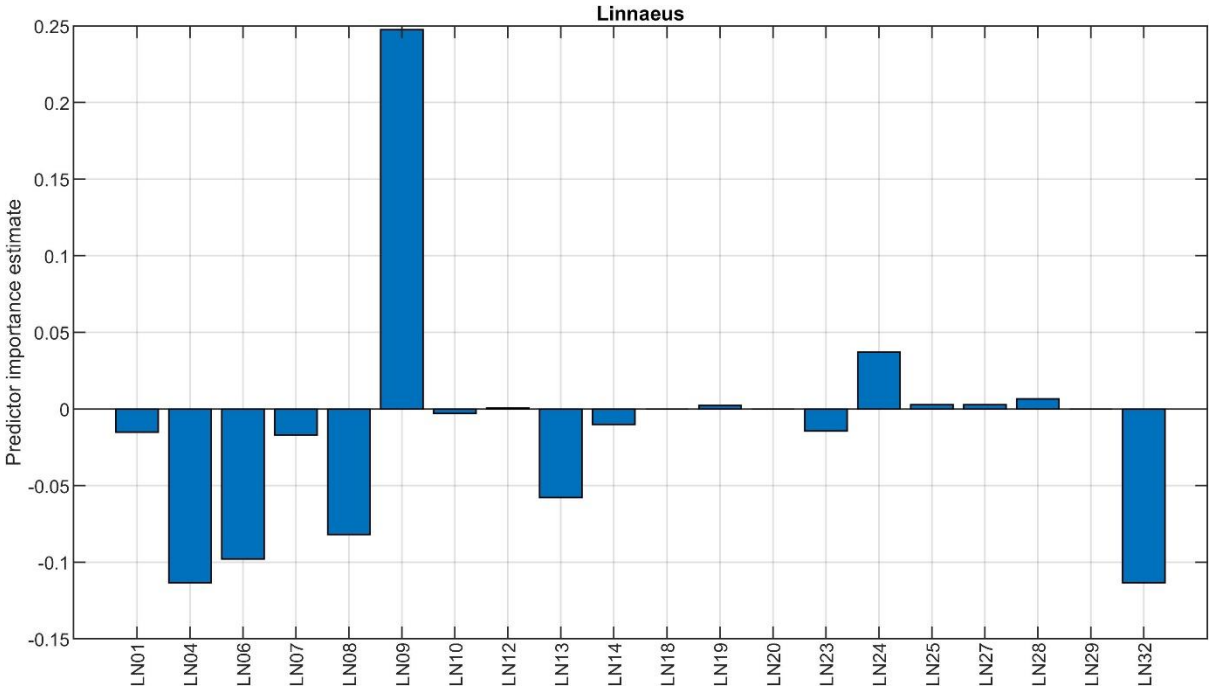

Figure 10. Importance estimate of Linnaeus variables in classifying between high and low yielding soils. The higher the 'predictor importance estimate' (y-axis), the better the variable is at predicting yield.

3.3.1.3 Hill Laboratories

38 Hill Laboratories soil nutrient variables[1] were considered as predictors for classifying between high and low yielding orchards. From the 100 iterations of the decision trees generated with data from 24 orchards, the maximum classification accuracy on the remaining five orchards was 76% by the Hill Laboratories variables.

Iron-Mehlich 3 (HS21), followed by Aluminium-Mehlich 3 (HS27) and C/N ratio (HS11), were the most important variables in classification between the high and low yielding orchards (Figure 11 and

[1] HWEC (HS13) was not include in this set of variable because is biological variable and was analysed. The univariate analysis showed no significant difference between high and low yielding orchards (p-value 0.22).

Appendix 5.7.3 Results showed that high yielding orchards had higher Iron-Mehlich 3 and C/N ratio levels and lower levels of Aluminium-Mehlich 3.

Appendix 5.8.3 shows one of 100 decision trees tested as part of the machine learning process that helps to define what variables are important and what thresholds may be relevant. Unfortunately, thresholds could not be defined with any level of confidence that incorporated all the important variables identified.

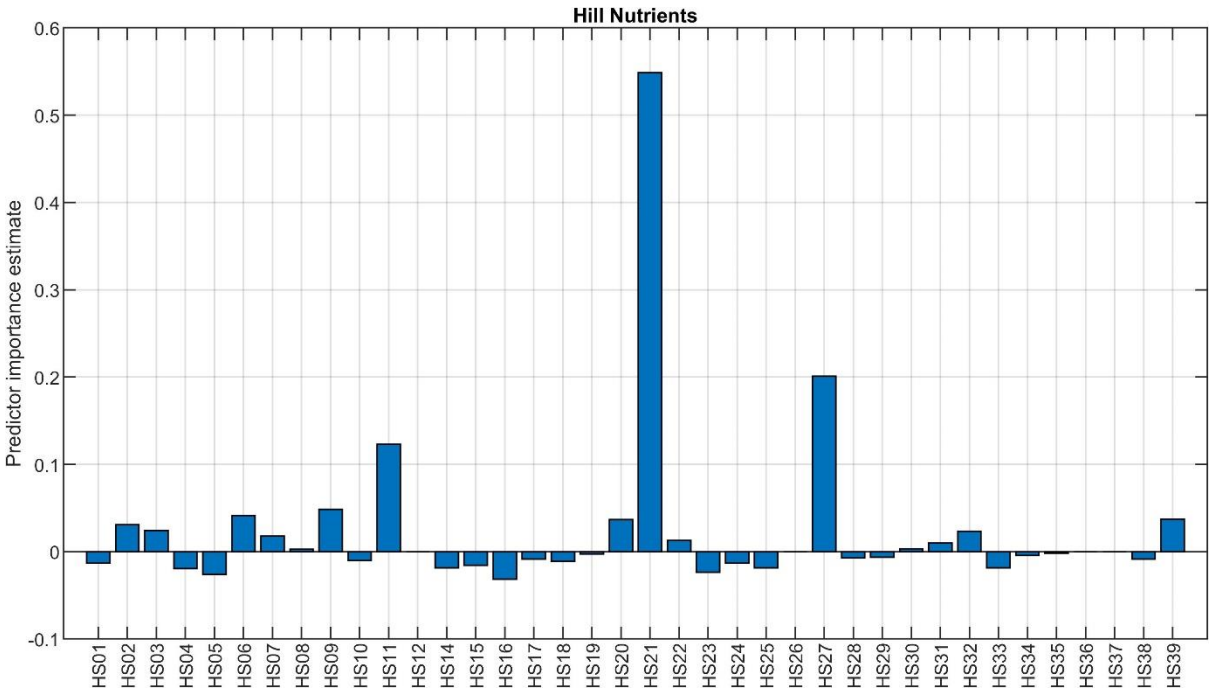

Figure 11. Importance estimate of Hill nutrient variables in classifying between high and low yielding soils. The higher the 'predictor importance estimate' (y-axis), the better the variable is at predicting yield.

3.3.1.4 Combination of Soil Foodweb and Hill Laboratories Soil Nutrients

49 Hill Laboratories nutrient and Soil Foodweb variables were analysed as predictors for classifying high and low yielding orchards. From the 100 iterations of the decision trees generated with data from 24 orchards, the maximum classification accuracy on the remaining five orchards was 79% using the combined Hill Laboratories and Soil Foodweb. Slightly higher than the 76% accuracy each set of variables achieved on their own.

The important variables selected in the analysis of Soil Foodweb and Hill Laboratories variables had relatively high importance when combined (Figure 12 and Appendix 5.7.4). However, the Hill Laboratories soil nutrient variables had stronger importance in classification than the Soil Foodweb variables. One example decision tree showed that classification based on Iron-Mehlich 3 (HS21) could provide 67% accuracy without any biological variables.

Appendix 5.8.4 shows one of the 100 decision trees tested as part of the machine learning process that helps to define what variables are important and what thresholds may be relevant.

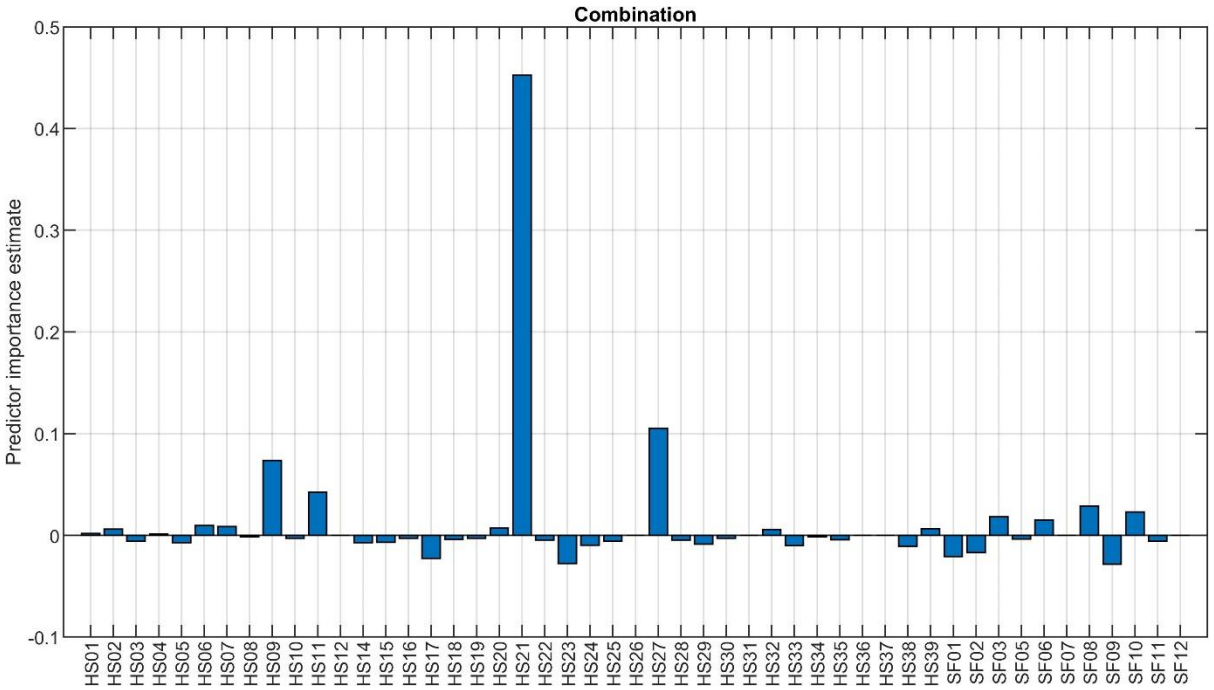

Figure 12. Importance estimate of Hill Laboratories nutrient variables combined with Soil Foodweb biological variables in classifying between high and low yielding soils by Random Forest. The metric for importance was the average OOB Delta error over 100 ensembles.

4 Conclusion and discussion

Key variables relating to soil physical, biological and chemical composition were successfully collected from groups of high performing orchards in the three main avocado producing regions of New Zealand and lower yielding orchards in the Bay of Plenty region. The information collected in this study will provide a valuable comparison across the different tests for avocado growers to compare their own soil test results to. Aspects of soil physical, biological and chemical composition all showed importance in classifying whether an orchard was high yielding or not. Some of the correlations observed are contrary to what is expected to contribute to a high performing soil. Caution is needed when interpreting these observations as management and environment also influence yield, and these variables were not captured in this study. It may be that there are common management practices among high performing orchards that negatively affect soil characteristics. Therefore, the correlations seen within the soil variables may be more related to management practices rather than the high yields common to these orchards. Further investigation may reveal that changes to management practices to enhance soil conditions may further enhance yields beyond what these already successful orchards are currently achieving. The machine learning approach taken when analysing the data means additional data can be added as it becomes available and the analysis re-run to improve the predictive model accuracy in identifying important variables.

Common variables within the commercial biological tests investigated showed poor correlation with each other. This suggests that biological tests should not be used interchangeably if a time series of data is desired. Further investigation is needed into the repeatability of results from these tests to ensure they are reliable and provide similar results for growers. Since only two tests were compared, at a single time point, it was impossible to identify which one was the most accurate. There may be possible improvements to sampling, sample handling and processing that could be highlighted from further investigation to improve consistency between the tests.

There was little variation in the soil physical variables scores across the orchards in this study. A wider selection of orchards may capture a wider range of soil properties and highlight the importance of soil physical properties with yield to a greater extent than seen in this study.

Although the leaf and fruit nutrient results show little correlation with yield in this study, greater correlation may be seen to fruit quality fruit quality. As more data is collected across the physical, biological and chemical composition (nutrients) of soils of avocado orchards, key variables and ideal ranges for these key variables will likely become clearer. Increasing the range of yields to include more lower yielding orchards would increase the statistical power and may better highlight correlations between variables and orchard yield.

Each of the variables identified as important in classifying an orchard as high or lower yielding are summarised and discussed below.

4.1 Soil chemistry (nutrition)

Soil chemical variables had the greatest importance when classifying orchard yield performance. This is not surprising as many of the benefits of good soil physical and biological properties are related to nutrient availability and enhanced uptake, and therefore, are likely secondary to the actual nutrient content of the soil.

Iron (Mehlich 3) and Aluminium (Mehlich 3) were identified as chief predictors when the Hill Laboratories and Soil Foodweb variables were combined. The table below is a summary of the Hill nutrient variables across the two yield classifications.

Table 10. The nutritional variables with the high importance estimate that were highlighted with machine learning analysis.

Sample type	Name	Label	Units	Average (min-max) High yield	Average (min-max) Lower yield
Soil	Iron (Mehlich 3)	HS21	mg/L	105(67-189)	75(52-156.0)
Soil	Aluminium (Mehlich 3)	HS27	mg/L	1349(739-1926)	1551(1338-1704)
Soil	C/N ratio	HS11		13.3(10.3-24.7)	11.3(10.2-12.1)

Iron (Mehlich 3): Higher iron levels (Mehlich 3) correlated with higher yielding orchards. The high and lower yielding orchards had 105 mg/l and 75 mg/l of iron on average, respectively. The range used by industry consultants is from 40-400 ppm of iron - Mehlich 3 (NZ Avocado Growers' Association Inc., 2018) and all orchards results were well within this range. For this reason, it appears unusual that the small average difference seen in this study is a predictor for yield. Iron is not a nutrient that is deemed difficult to manage in New Zealand; therefore, this result warrants further investigation. Iron is a structural component of enzymes in electron transport chains and is required to synthesise chlorophyll (Lovatt, 2015). Both of these are essential processes for the tree.

Aluminium (Mehlich 3): Lower aluminium levels (Mehlich 3) correlated with high yielding orchards. High and low yielding orchards had on average 1349 mg/l and 1551 mg/l, respectively. Aluminium toxicity can become a problem but only at soil pH below 5.5. The average soil pH for lower yielding orchards was 5.7 with a range of 4.8 to 6.3, and higher yielding orchards had an average of 5.9 with a range of 5.1 to 6.2. High yielding orchards had two orchards with soil pH below 5.5, and three lower yielding orchards had three. Soil pH may also be contributing to this correlation of lower Aluminium levels in higher yielding orchards.

Carbon/Nitrogen (C/N) ratio: The average C/N ratio for high yielding orchards in this study in the Bay of Plenty was 12.0, whereas the average for low yielding orchards was 11.3. This ratio is the quotient between the percentage of organic carbon and total nitrogen. It is related to mineralisable nitrogen as it relates to the availability of nitrogen to the tree. A C/N ratio greater than 35 results in microbial immobilisation of nitrogen as the microbes scavenge nitrogen from the soil to utilise the carbon that is available via respiration (microorganisms release carbon dioxide). This means nitrogen is not available to the plant. When the C/N ratio is below 25, soil microbes generate mineralised nitrogen that is available for the plant (McLaren and Cameron n.d.). Once C/N ratios are between 1 and 15, nitrogen is rapidly mineralised. These results indicate that the nitrogen that is in the soil had a high availability to the plant and little chance of nitrogen draw down by microbes scavenging nitrogen from the soil. This ratio should be interpreted in context with total organic carbon and total nitrogen. The difference between the two groups classification is minimal, but a larger sample size with a broader yield range along with broader C/N ratios may be helpful to identify an ideal ratio in different regional soils. Ideal mulch for avocado has C/N ratios of between 25 and 100 to ensure some carbon residues are retained to protect the soil from sun and provide an environment for the roots to grow. The average organic matter of soils tested were 13.5, 12.8 and 9.9 for the Bay of Plenty, Mid North and the Far North, respectively. The orchards that fit into the higher yielding classification in the Bay of Plenty had an organic matter value of 15.2 compared to 12.7 for the lower yielding orchards.

4.2 Soil biology

Soil biological variables were also able to classify orchard performance with a reasonably high degree of accuracy. The table below is a summary of the important predictive results.

Table 11. The biological variable with the high importance estimates that come up from machine learning analysis. The range is a threshold for good levels provided by the laboratories.

Lab	Name	Label	Units	Average (min-max) High yield	Average (min-max) Lower yield	Target
SFW	Dry weight	SF01	--	0.67(0.50-0.85)	0.61(0.57-0.68)	0.45-0.85
SFW	Total bacteria (TB)	SF03	mg/kg	330(237-535)	362(245-483)	>300
SFW	Total fungi (TF)	SF06	mg/kg	328(137-911)	384(205-513)	>1500
SFW	Flagellates	SF08	number/g	11,182(782-44,249)	17,276(2038-55,500)	>5000
SFW	Ciliates	SF10	number/g	366(38-836)	177(43-472)	<334
Linnaeus	Gram negative bacteria	LN09	mg/kg	6(3-9)	7(4-14)	11

Ciliates (SF10) and Flagellate (SF08): Protozoans play a role in nutrient cycling and have some predatory behaviours that control the size and composition of microbial communities. Flagellates and amoeba are desired in high numbers as they work aerobically to digest bacteria and make their nutrients available to the plant. However, ciliates feed on anaerobic bacteria and are more tolerant of anaerobic environments meaning high levels of ciliates can indicate a compact of water logged soil. Higher levels of ciliates were found on high yielding orchards compared to lower yielding orchards, which is not the expected association. Studies have shown higher ciliate abundance correlates with higher soil moisture, organic matter, available nitrogen, phosphorous, copper, zinc, nickel and total microbial biomass (Acosta-Mercado and Lynn 2004; Luu 2019). The cause of the higher ciliate counts from the higher yielding orchards warrants further investigation but may be related to recent rain conditions or other soil conditions that favour ciliates. Higher counts of flagellates were seen on the lower yielding orchards, but both the higher yielding and lower yielding orchards had average counts well above the targets. High levels of flagellates are not assumed to be negative as they play a key role in making nutrients available to trees.

Dry weight (SF01): High yielding orchards had high dry weight compared to low yielding orchards, correlating with the higher organic matter seen on higher yielding orchards. Both high and lower yielding orchard levels were within normal range, according to the Soil Foodweb range (see Table 11). These results could be quite variable depending on sampling conditions as rain and irrigation will likely influence the result.

Total bacteria (SF03): High yielding orchards had low levels of total bacteria. In both cases, the level was above the target range, according to the Soil Foodweb range, indicating that probably soils were high in simple carbon forms from root exudates and fresh plant litter. Avocado trees generally have a lot of litter under the dripline; therefore, the high levels seen may be indicative of this environment. Why higher yielding orchards may have lower levels of bacteria is unclear but the higher protozoa levels in the form of ciliates may relate to this.

Gram negative (LN09): High yielding orchards had lower levels of gram-negative bacteria compared to the low yielding orchards, but both are below good levels, according to Linnaeus range, of microbial biomass (see Table 11). Some gram-negative bacteria are associated with nitrogen fixation but this requires an association with the root system that does not occur with avocado. This group may have other positive microbial roles in agriculture as some members produce plant growth hormones, fight

pathogens, or make nutrients available for plants such as some Pseudomonas spp, *Azotobacter spp.* and *nitrosomas*. Therefore, it is not clear why lower levels of gram negative bacteria might be associated with higher yielding orchards.

A Comparison of Soil Foodweb and Linnaeus variables of the same description showed little correlation for Total bacteria and Total fungi. This suggests that the tests should not be used interchangeably and using the same test over time will be the best way to track changes in soil biology.

Hill Laboratories Hot Water Extractable Carbon is a biological measure of soil health. A single variable was not a strong predictor of orchard yield classification in this study and so was not analysed separately. Other studies have been proven to be closely correlated with soil biomass and so would still offer a reasonable measure of soil biology over time (Ghani, Dexter, and Perrott, 2003). It was the cheapest biological method used in this study and would be an effective tool to monitor soil biology over time to understand degradation or enhancement. It does not differentiate between fungi and bacteria; therefore, it may be difficult to understand how best to influence the biology of the soil from this test alone.

4.3 Soil physical properties (Visual soil assessment)

Worm count was the only statistically significant variable in visual soil assessments that acted as a predictor of orchard yield in this study. Soil physical properties such as compaction and poor structure are known to impact avocado tree health negatively, so it was unexpected not to identify some more of these issues in the lower yielding orchards. If a larger number of orchards were investigated, these physical limitations may be encountered more, resulting in more variable physical soil assessment scores. This in turn would give a more accurate assessment of the importance of these parameters on orchard yield. However, it is encouraging to see from orchards in this project that soil physical properties are unlikely to be limiting yields. All 29 orchards included in this study had a soil quality index of over 28 points. According to visual soil assessment methodology (Shepherd, 2019), all orchards were in the top soil quality category.

Table 12. The total number of earthworm counts. Four holes on the different cardinal points of the tree were taken so counts are from a total of 60 cm^2 of soil.

Name	Unit	BOP <15t/ha Average (min-max)	BOP >15t/ha Average (min-max)	MN Average (min-max)	FN Average (min-max)
Earthworms	count	15.2(0-48)	7.5(1-18)	**17.2(1-53)**	**10.6(0-41)**

Earthworm count: Although earthworms are a biological variable, they were included under the visual soil assessment methodology. All sampled orchards had scored zero for the earthworms count under the scoring system used because the average of the four reps (200 mm cube of soil each) was below 15 earthworms. The total sum of earthworms in the four reps was significantly correlated with orchard yield, where orchards with lower earthworms total count had higher yields. This was an unexpected result as earthworms have a beneficial role. The lower counts on higher producing orchards may be related to other cultural practices common amount the more successful orchards and warrants further investigation.

4.4 Leaf and fruit nutrients

Leaf nutrient testing is a common method to check the nutrient status of a tree and identify any deficiencies that may be limiting production. While no correlation was seen in this study between leaf nutrient levels and yield, the published nutrient ranges of high producing orchards will still be

informative to growers. The impact of other environmental and management practices on yield suggest higher sample numbers are needed to elucidate the impact of nutrients. A correlative study in the Western Bay of Plenty across over a hundred orchards showed different averages between the yield classes stepping up in 5t/ha increments from 0 to >30t/ha but overlapping ranges between the lowest and highest groups (Dixon, 2008). A Californian study that used novel data analysis to refine leaf nutrient targets and fertiliser application utilised over 3500 observations for the analysis (Crowley 2016).

4.5 Potential future work

4.5.1 Understanding how soil variables, as well as leaf and fruit nutrient concentrations, influence fruit rots

17 of the orchards used in this study are also part of the Avovantage project looking to understand the on-orchard factors that contribute to avocado fruit rots. Nutrition has always been associated with fruit rot potential with the nitrogen to calcium ratio of the fruit of particular interest. Data from this study will be incorporated into the Avovantage project to look for correlations between soil, leaf and fruit properties with fruit rot potential.

4.5.2 Potential partnership with commercial soil biology and nutrient testing laboratories

Both soil nutrients and biological variables acted as predictors of orchard yield in this study. Some of the correlations observed were contrary to the current understanding of what makes a 'good' soil. Increasing the number of orchards involved in the analysis would help build confidence in the results of this study or identify alternative predictor variables. By partnering with growers and the commercial laboratories providing the testing services, additional data points could be obtained to re-run this analysis with ever increasing sample size. Running across multiple years would also account for seasonal fluctuations that may influence results.

4.5.3 Incorporate additional management and environmental variables into the analysis to identify the relative importance

Machine learning allows multiple variables to be analysed collectively and a hierarchy of importance to be built. By capturing additional management and environmental data, it should be possible to understand better the relevance of soil composition variables in the context of management and environmental influences (e.g. is a $10,000 investment in a compost application going to have more or less influence on yield than $10,000 spent on extending an irrigation system).

References

Acosta-Mercado, Dimaris, and Denis H. Lynn. 2004. "Soil Ciliate Species Richness and Abundance Associated with the Rhizosphere of Different Subtropical Plant Species." *The Journal of Eukaryotic Microbiology* 51(5):582–88. doi: 10.1111/j.1550-7408.2004.tb00295.x.

Crowley, D. E. 2007. "Managing Soil for Avocado Production and Root Health." *California Avocado Society Yearbook* 90:107–30.

Crowley, David. 2016. *Decision Support Tools for Management of Avocado Nutrition and Chloride Toxicity*. UC Grant#: 12034311.

Dixon, J. 2008. "Leaf Mineral Norms of 'hass' Avocado Trees from the Western Bay of Plenty of New Zealand." *NZ Avocado Gowers Association Research Report.* Vol 8.:1–14.

Ghani, A., M. Dexter, and K. W. Perrott. 2003. "Hot-Water Extractable Carbon in Soils: A Sensitive Measurement for Determining Impacts of Fertilisation, Grazing and Cultivation." *Soil Biology and Biochemistry* 35(9):1231–43. doi: 10.1016/S0038-0717(03)00186-X.

Grover, Minakshi, Shrey Bodhankar, Maheswari Mandapaka, and Srinivasrao Ch. 2016. "Actinomycetes as Mitigators of Climate Change and Abiotic Stress." Pp. 203–12 in *Plant Growth Promoting Actinobacteria: A New Avenue for Enhancing the Productivity and Soil Fertility of Grain Legumes*.

Herrera, Aravena, and Cristián Marcelo. n.d. "Efecto de la micorrización en plantas de vivero de palto y cítricos bajo diferentes dosis de fertilización." 102.

Ingham, Elaine R. 2021. "Soil Fungi | NRCS Soils." *United States Department of Agriculture*. Retrieved May 6, 2021 (https://www.nrcs.usda.gov/wps/portal/nrcs/detailfull/soils/health/biology/?cid=nrcs142p2_053864).

Kumar, Ajay, Hariom Verma, Vipin Singh, Prem Singh, Sandeep Singh, Waquar Ansari, Akhilesh Yadav, Pawan Singh, and Kapil Pandey. 2017. "Role of Pseudomonas Sp. in Sustainable Agriculture and Disease Management." Pp. 195–215 in *Agriculturally Important Microbes for Sustainable Agriculture*.

Li, Junqin, Bo Meng, Hua Chai, Xuechen Yang, Wenzheng Song, Shuixiu Li, Ao Lu, Tao Zhang, and Wei Sun. 2019. "Arbuscular Mycorrhizal Fungi Alleviate Drought Stress in C3 (Leymus Chinensis) and C4 (Hemarthria Altissima) Grasses via Altering Antioxidant Enzyme Activities and Photosynthesis." *Frontiers in Plant Science* 10. doi: 10.3389/fpls.2019.00499.

Lovatt, C. 2015. "Optimising 'Hass' Avocado Tree Nutrient Status to Increase Grower Profit-an Overview." *World Avocado Congress Proceedings.VII*.

Luu, Hai Thi Thuy. 2019. “Biodiversity of Ciliated Protozoa in Soil Ecosystems and Assessment of Their Potential as Bio-Indicators of Soil Quality.” PhD, Bournemouth University.

McLaren, R. G., and K. C. Cameron. n.d. “Soil Science. An Introduction to the Properties and Management of New Zealand Soils.” Oxford University Press.

Méndez-Bravo, Alfonso, Elvis Marian Cortazar-Murillo, Edgar Guevara-Avendaño, Oscar Ceballos-Luna, Benjamín Rodríguez-Haas, Ana L. Kiel-Martínez, Orlando Hernández-Cristóbal, José A. Guerrero-Analco, and Frédérique Reverchon. 2018. “Plant Growth-Promoting Rhizobacteria Associated with Avocado Display Antagonistic Activity against Phytophthora Cinnamomi through Volatile Emissions” edited by S. Jogaiah. *PLOS ONE* 13(3):e0194665. doi: 10.1371/journal.pone.0194665.

Menge, John A., R. Michael Davis, Edward L. V. Johnson, and George A. Zentmyer. 1978. “Mycorrhizal Fungi Increase Growth and Reduce Transplant Injury in Avocado.” 2.

NZ Avocado Growers’ Association Inc. 2018. *Avocado Growers Manual*.

Pepper, Ian L., Jerry J. Gentry, and Charles P. Gerba. 2015. *Environmental Microbiology*. Third Edition. Elsevier.

Pliego, Clara, José Ignacio Crespo-Gómez, Adrián Pintado, Isabel Pérez-Martínez, Antonio de Vicente, Francisco M. Cazorla, and Cayo Ramos. 2019. “Response of the Biocontrol Agent Pseudomonas Pseudoalcaligenes AVO110 to Rosellinia Necatrix Exudate.” *Applied and Environmental Microbiology* 85(3):e01741-18. doi: 10.1128/AEM.01741-18.

Ramirez-Gil, J. G., D. A. Castaneda-Sanchez, and J. G. Morales-Osorio. 2016. “Production of Avocado Trees Infected with Phytophthora Cinnamomi under Different Management Regimes.” *Plant Pathology*. doi: 10.1111/ppa.12620.

Shepherd, Graham. 2019. *Visual Soil Assessment*. Vol. 1. Second Edition. Palmerston North, New Zealand: Horizons Regional Council.

Sullia, S. B. 1991. “Use of Vesicular - Arbuscular Mycorrhiza (VAM) as Biofertilizer for Horticultural Plants in Developing Countries.” Pp. 49–53 in *Horticulture — New Technologies and Applications: Proceedings of the International Seminar on New Frontiers in Horticulture, organised by Indo-American Hybrid Seeds,Bangalore, India, November 25–28, 1990*, *Current Plant Science and Biotechnology in Agriculture*, edited by J. Prakash and R. L. M. Pierik. Dordrecht: Springer Netherlands.

Van Zwieten, L., J. Rust, T. Kingston, G. Merrington, and S. Morris. 2004. “Influence of Copper Fungicide Residues on Occurrence of Earthworms in Avocado Orchard Soils.” *Science of the Total Environment* 329.:29–41.

5 Appendix

5.1 Sampling collection

5.1.1 Soil samples collection

i. Divide the area under the tree drip lines into four quadrants.
ii. Choose two sample sites in NW and SE quadrants and take pictures from the tree tag and quadrants starting with the NW.
iii. Choose and spot halfway between the tree trunk and the dripline edge (in the rooting zone) and screenshot GPS coordinates.
iv. Remove any vegetation or any other element (e.g. mulch) from the top of the sample sites.
v. Use the auger to take the profiles through the top 15cm, starting with the NW and then SE.
vi. Place the two cores in the bucket.
vii. Move to the following tree
viii. Once all the cores have been taken, mix the sample in a clean bucket.
ix. Label them and place into the cooler box with the ice pack.

5.1.2 Leaf sample collection

i. Identify shoots that are not flushing nor fruiting.
ii. Take 4 to 8 random leaves of the youngest mature leaf (blade plus petiole) at shoulder height.
iii. Ensure leaves from each tree are taken evenly from the sunny and shaded sides.
iv. Do not mix cultivars or trees of different ages in the samples.
v. Label the sample and place it into the cooler box with the ice pack.

5.1.3 Fruit sample collection

i. Sample one fruit per tree. If there is no fruit in the pre-selected trees, then take fruit from the nearby trees in the same block, and note that.
ii. Placed directly into the sampling bag with minimum hand contact.
iii. Label the sample and place it into the chilli box with the ice pack

5.2 Sampling delivery

To ensure the integrity of the samples, Fridays were avoided, and samples were sent on Thursdays to reduce the chance that they spent the weekend in the courier storage. Hill laboratory provide bags for their samples. The other samples were packed in re-sealable zipped bags. Table 13 summarises the sample amount and consideration that every laboratory require.

Table 13. The total sample size and storage requirements per laboratory.

Type	Laboratory	Minimum sample size	Storage
Soil	Hill Laboratory	600 g	Samples were held in a refrigerator overnight (5 °C), then couriered to the lab.
Soil	Linnaeus	400 g	All samples were frozen (-18°C) for at least 24 hours, wrapped in a freezer bag and courier to the lab.
Soil	Soil Food Web	400 g	Samples were held in a refrigerator overnight (5 °C), then couriered to the lab.
Leaf	Hill Laboratory	20-40 leaves	Samples were held in a refrigerator overnight (5 °C), then couriered to the lab.
Fruit	Hill Laboratory	10 pieces of fruit	Samples were held in a refrigerator overnight (5 °C), then couriered to the lab wrapped in bubble wrap.

5.3 Visuals soil assessment

5.3.1 Special considerations

The VSA was performed following the instructions in the book Visual Soil Assessment, Vol 1 by Graham Shepherd Assessment and posterior errata and addenda o the book. Additionally, the following considerations were included:

- Earthworm number is the most variable visual indicator; therefore, four earthworm replications per VSA (12 replications per orchard/block) were done under the advice of Graham Shepherd.
- Soil texture was determined utilising the guidance "Determining soil texture" by Irrigation New Zealand since this procedure is well-known by growers. This document is available on: https://www.irrigationnz.co.nz/news-resources/irrigation-resources/soil-texture-water/Attachment?Action=Download&Attachment_id=104
- Potential rooting depth the maximum deep of the hole was restricted to 1 meter because digging a hole more than one-meter depth was considered unpractical and time-consuming. According to that, the criteria for unirrigated orchards were scored in the same fashion that irrigated orchards.
- Surface ponding - orchards were assessed according to time availability and not always after significant rain; therefore, this indicator was scored based on growers' feedback or our know-how.

5.3.2 VSA Scorecard

SOIL INDICATORS

Land owner: Land use:
Site location: GPS ref:
Sample depth: Topsoil depth:
Soil type: Soil classification:
Drainage class (p. 19): Date:

Textural group: (upper 1m) ☐ Sandy ☐ Coarse loamy ☐ Fine loamy ☐ Coarse silty ☐ Fine silty ☐ Clayey ☐ Peaty
Moisture condition: ☐ Dry ☐ Slightly moist ☐ Moist ☐ Very moist ☐ Wet
Seasonal weather conditions: ☐ Dry ☐ Wet ☐ Cold ☐ Warm ☐ Average

Visual Indicators of Soil Quality	Visual Score (VS) 0 = Poor condition 1 = Moderate condition 2 = Good condition	Weighting	VS Ranking
Soil texture (p. 16)		× 3	
Soil structure (p. 17)		× 3	
Number and colour of soil mottles (p. 19)		× 2	
Soil colour (p. 20)		× 2	
Earthworms (Number =) (Non-herbicide area, p.22; herbicide strip, p.76)		× 3	
Soil smell (p.24)		× 2	
Potential rooting depth (m) (see Table below)		× 3	
Surface ponding (Non-herbicide area, p.28; herbicide strip, p.82)		× 3	
SOIL QUALITY INDEX (Sum of VS rankings)			

Soil Quality Assessment	Soil Quality index
Poor	< 14
Moderate	14–28
Good	> 28

VSA score	Potential rooting depth (m)	
	Rain-fed. Deep rooting stock	Irrigated. Shallow rooting stock
2.0 (Good)	>2	>0.8
1.0 (Moderate)	1.0–1.5	0.4–0.6
0 (Poor)	<0.5	<0.2

5.4 Statistical analysis description

5.4.1 Yield predictability by individual variables (Univariate analysis)

Linear Mixed Model (LMM) was employed to investigate the difference in every single soil variable measured by the laboratories or visually assessed, except the earthworm. That is because all other variables have normal or lognormal distributions except earthworms that had Poisson distribution.

The fixed effect was the binary classification of high/low yielding, and the block effect was the region. The model assumptions were checked for each variable prior to fit, and the F-test was used to report on the statistical significance of the fixed effect. The predicted means and 95% confidence intervals were reported.

Generalised Linear Mixed Model (GLMM) was employed to investigate the difference in earthworm counts between the binary classifications of high/low yielding orchards. The fixed effect was the binary classification, and the random effect was the region. The model was fit by Poisson distribution at log scale. Wald test was used to investigate the statistical significance of the fixed effect, and the back-transform means and confidence intervals were reported.

The computations were performed by the Statistics and Machine Learning toolbox, MATLAB 2020b.

5.4.2 Inter-correlation among variables (Multivariate analysis)

Degrees of inter-correlation or inter-dependencies were expected among three groups of variables: The 15 biological variables measured by Soil Food Web, the 32 biological variables measured by Linnaeus and 38 soil nutrient variables measured by Hill. This expectation was because these three groups of variables were collected from the same soils samples and belonged to the same category (biological or chemical). Strong inter-correlation among each group of variables would potentially impact the overall relationship between the variables and the binary classifications of high performing/low yielding orchards. If there were a strong (or weak) correlation between a particular variable and the binary classifications, a similar relationship would have been expected between all other variables that are strongly correlated with the particular variable and the binary classifications.

Principal Component Analysis (PCA) was used to investigate inter-correlation among the three groups, and each group was analysed independently from the other two. PCA is a geometric transformation that transforms the n-dimensional matrix of correlated variables to another n-dimensional matrix whose columns are uncorrelated. Here, the dimension of the matrices was 15, 32 and 38 variables by 29 orchards for SFW, Linnaeus and Hill soil nutrients accordingly. PCA transformed these matrices into matrices of the same dimension. The transformed dimensions no longer represent every single measured variable but contain overall information about the variables mean, variance and higher orders statistics. The transformed dimensions are called the Principal Components (PCs).

Typically, we would expect the PCs to be negligible from PC4 onwards. In other words, PCA transforms a high dimensional matrix of correlated variables into a low dimensional matrix of PCs (typically 3 dimensional) as the PCs are negligible for higher dimensions. This provides a simple visualisation tool for high dimensional data as it is summarised into 2 or 3 dimensions.

The 15x29 matrix of Soil Food Web, 32x29 matrix of Linnaeus and 38x29 matrix of Hill soil nutrient variables were normalised by deduction of the mean of every column and division by each column standard deviation prior to the analysis to meet the PCA assumptions. The percentage of explained

variation by each PC was reported, and the overall correlation between the first two PCs and the binary classes or high/low yielding orchards was investigated.
The computations were performed by Statistics and Machine Learning toolbox, MATLAB 2020b.

5.4.3 Classification trees

Classification trees are group of supervised machine learning algorithms, addressing the following objective: Given the datasets, which biological, nutrient or combination variables could classify between high yielding/low yileding classes of orchards with maximum accuracy. What is the maximum accuracy, what are the cutting thresholds of the classifying variables, and how generalisable the conclusions are?

Therefore, unlike traditional statistical methods such as univariate or multivariate models, there is no fixed target statistical significance (for example, 95%) for classification trees; the accuracy of classification trees is a relative metric and usually calculated by Out-Of-Bag (OOB) error estimation. The OOB error estimate is the misclassification rate in the dataset on which the classification tree was trained. The consistency is another metric for classification accuracy. It is the correct-classification rate in the dataset on which the classification tree was tested. It is important to consider both OOB error estimate and consistency in the performance assessment, as the classification trees tend to overfit the training set and perform poorly in the test set, which can hinder their generalisability.

The Random Forest algorithm was used to fit trees to classify between high/low yielding orchards based on the 15 SFW biological, 32 Linnaeus biological and 38 Hill soil nutrient variables, as well as combinations of the variables. Random Forest fits bags of random classification trees on the training dataset to avoid problems such as overfitting. The bag of trees per Random Forest run is called an ensemble.

There are several measures of the predictors (e.g. laboratory variables) importance in classification. The OOB Permuted Predictor Delta error (or OOB Delta error in short) was used. For each predictor (variable), the OOB Delta error is the increase in OOB error if that variable is left out of the bag among the observations. This measure is computed for every tree, then normalised by the average and standard deviation of the ensemble.

Before fitting Random Forest trees, the algorithm should be tuned to the dataset. Therefore, two hyper-parameters of trees minimum leaf size and number of predictors to sample were tuned for each dataset separately. The tuning was performed by Bayesian optimisation with OOB error estimate as the objective function. The tuned hyper-parameters were then used in the Random Forest algorithm for each dataset.

In each round of Random Forest running, the (biological, chemical or combination) variables of 24 orchards were randomly selected from all regions as the training set. The algorithm bagged 50 classification trees on the training set and tested the performance on the remaining five orchards was recorded. This procedure was repeated 100 times per dataset. The mean OOB error estimate, the mean consistency, and the mean importance of the variables in classification were reported.

The computations were performed by *Statistics and Machine Learning toolbox, MATLAB 2020b*

5.5 Fruit and leaves linear regression against yield

5.5.1 Leaves

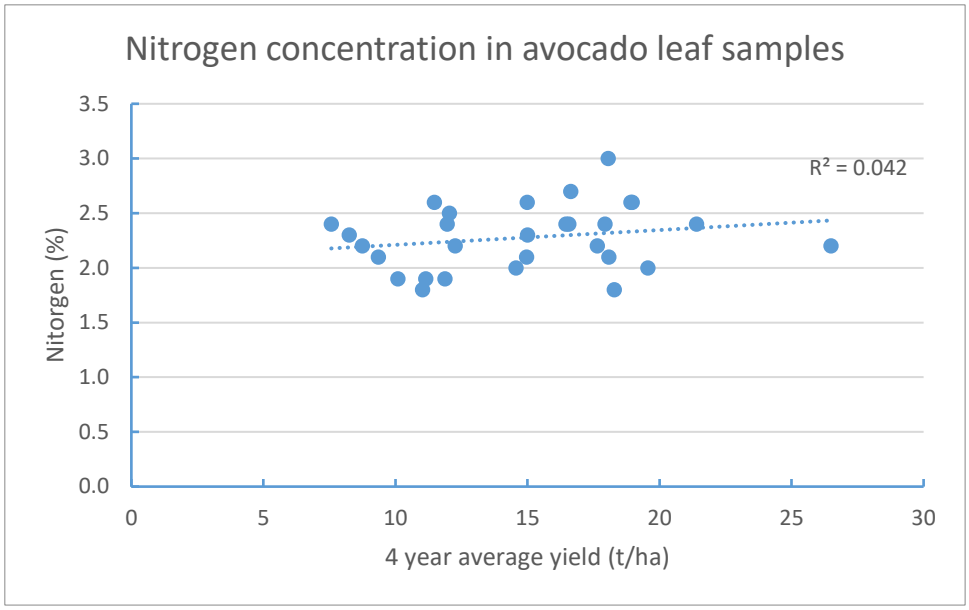

Figure 13. Nitrogen levels versus 4-years average yield in avocado leaf samples.

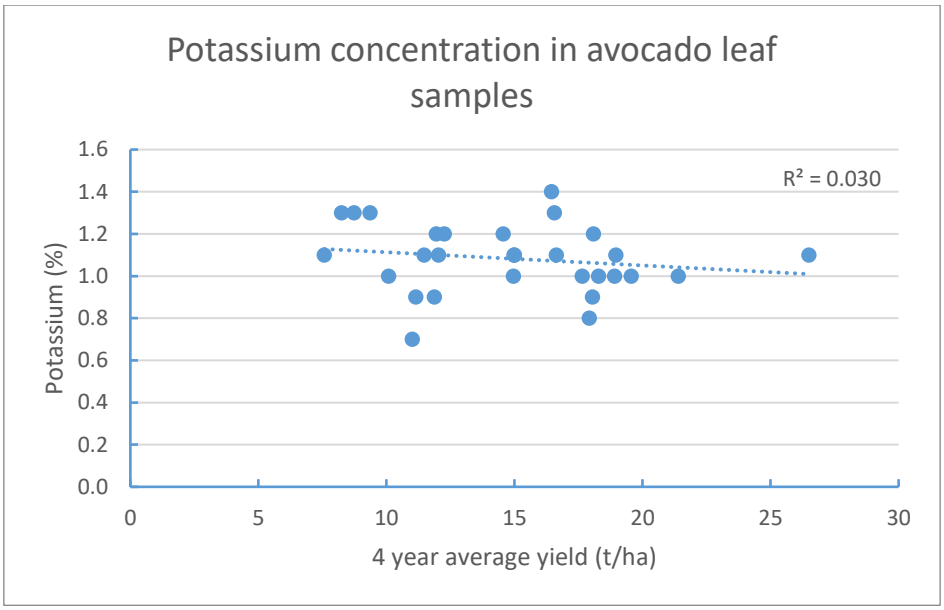

Figure 14. Potassium levels versus 4-years average yield in avocado leaf samples.

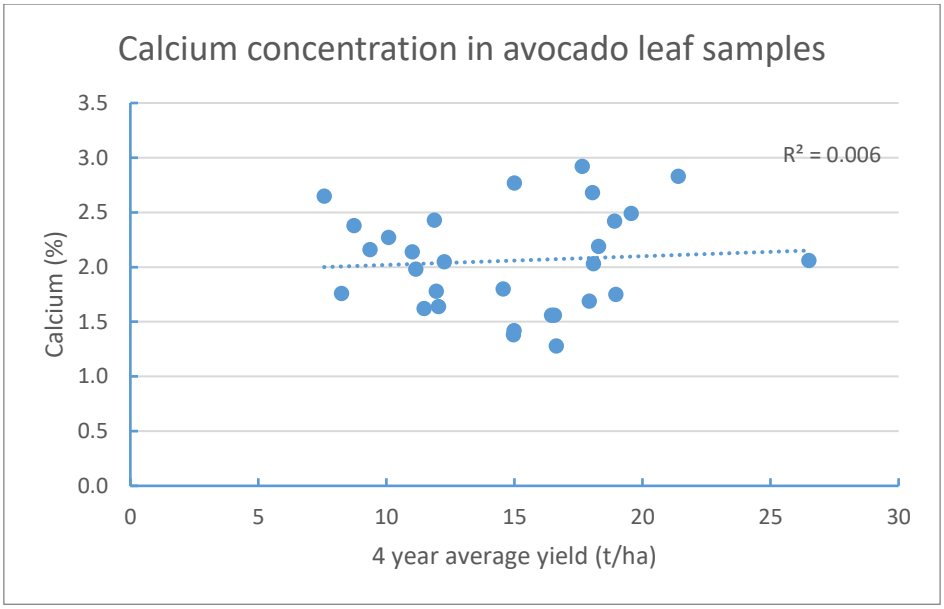

Figure 15.Calcium levels versus 4-years average yield in avocado leaf samples.

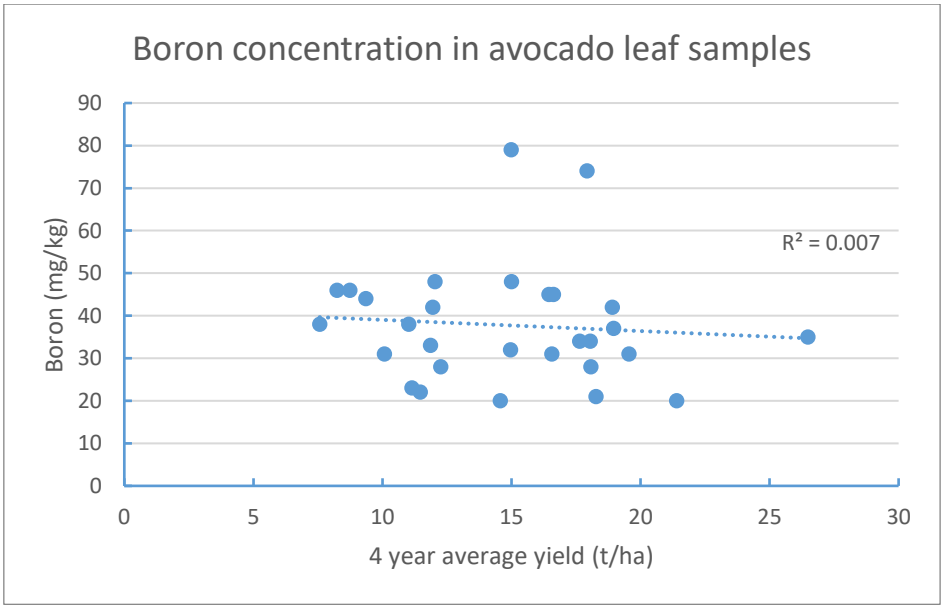

Figure 16. Boron levels versus 4-years average yield in avocado leaf samples.

5.5.2 Fruit skin

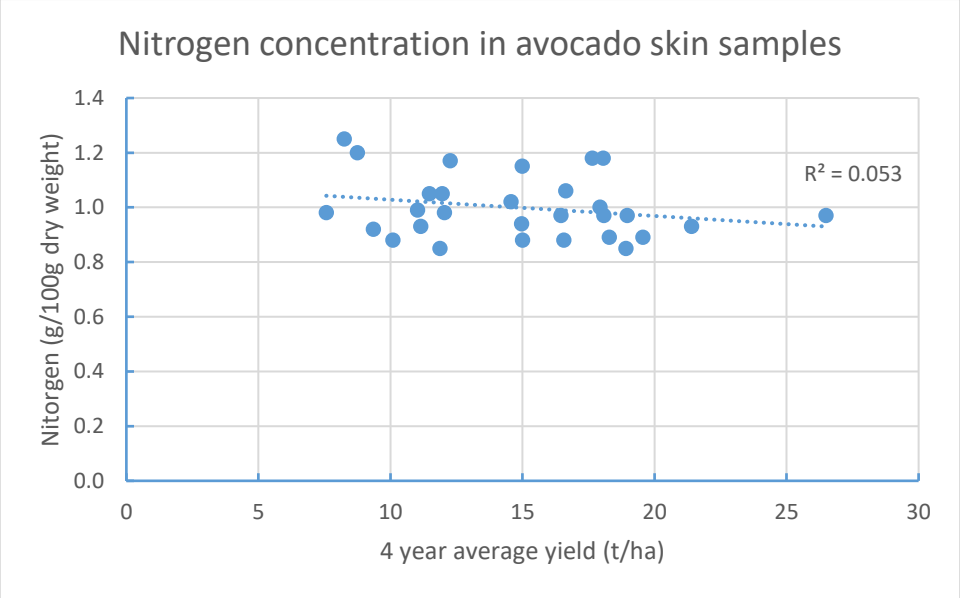

Figure 17. Nitrogen levels versus 4-years average yield in avocado skin samples.

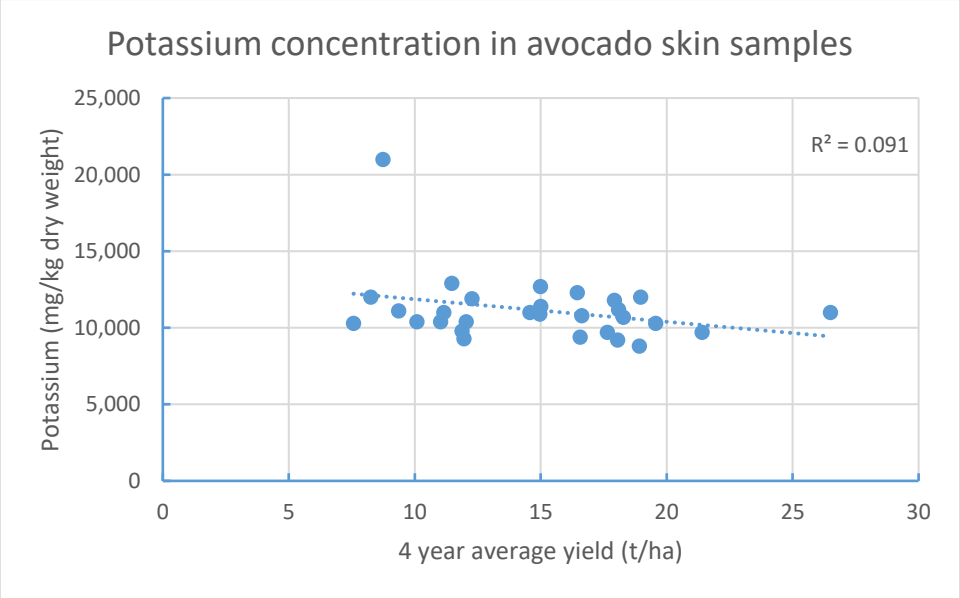

Figure 18. Potassium levels versus 4-years average yield in avocado skin samples.

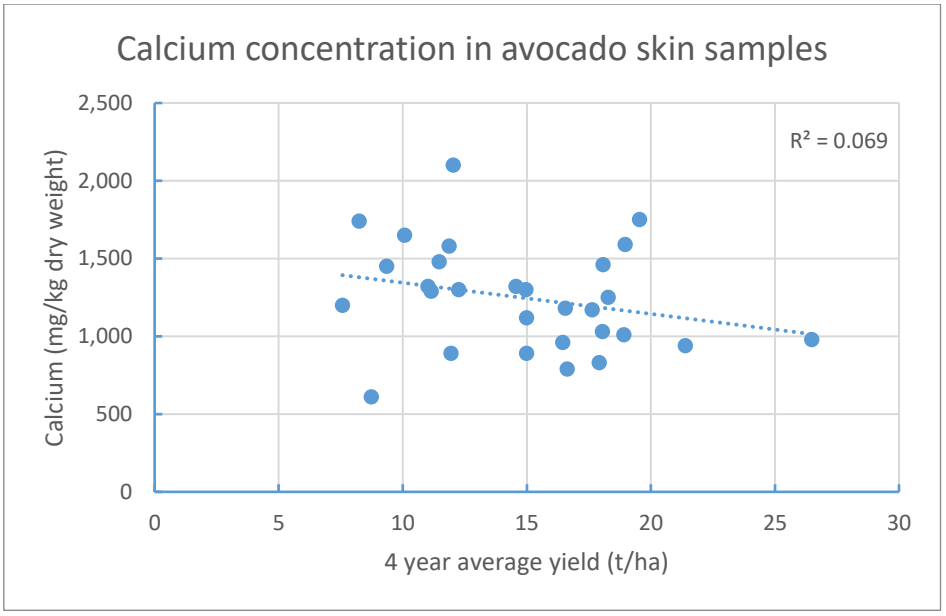

Figure 19. Calcium levels versus 4-years average yield in avocado skin samples.

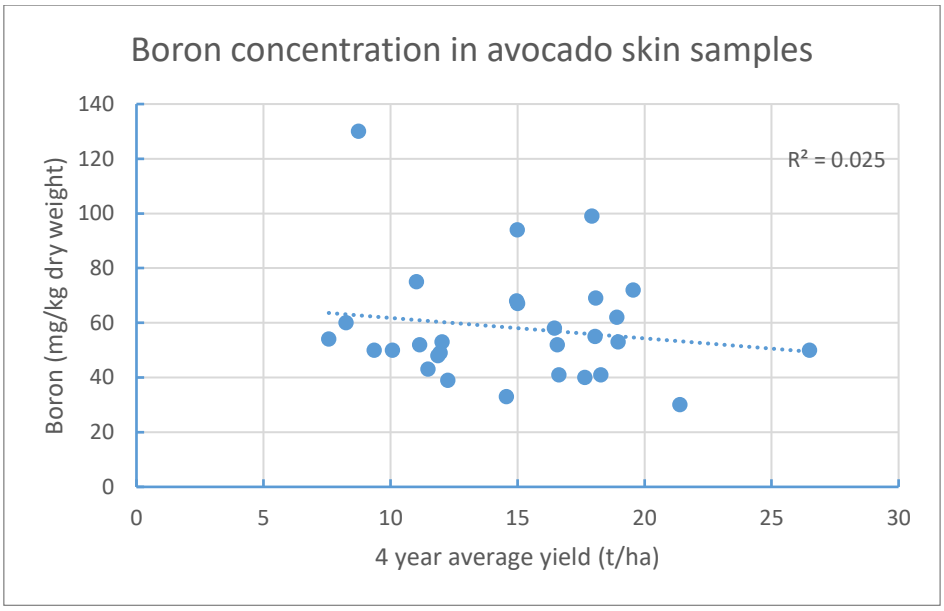

Figure 20. Boron levels versus 4-years average yield in avocado skin samples.

5.5.3 Fruit flesh

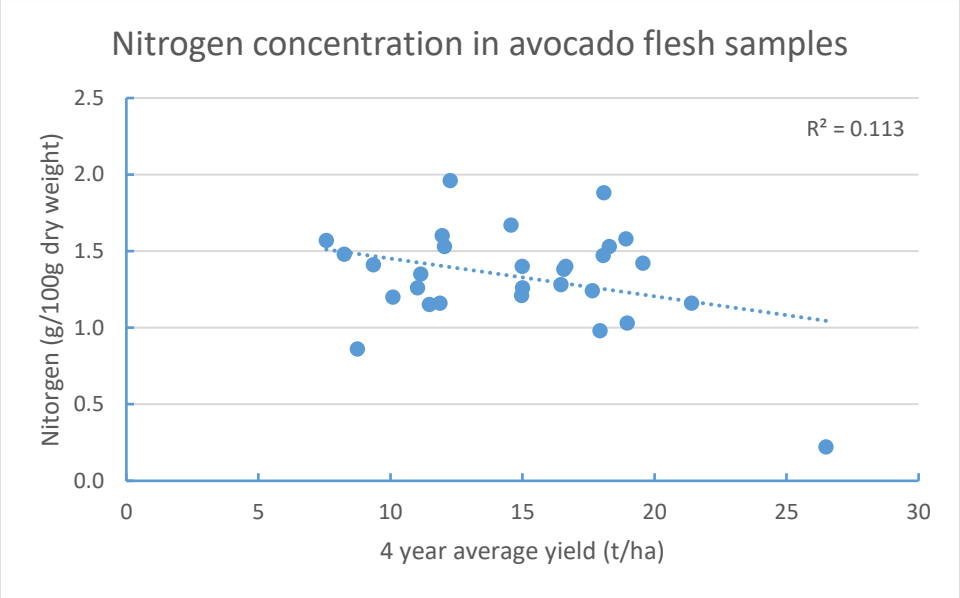

Figure 21. Nitrogen levels versus 4-years average yield in avocado flesh samples.

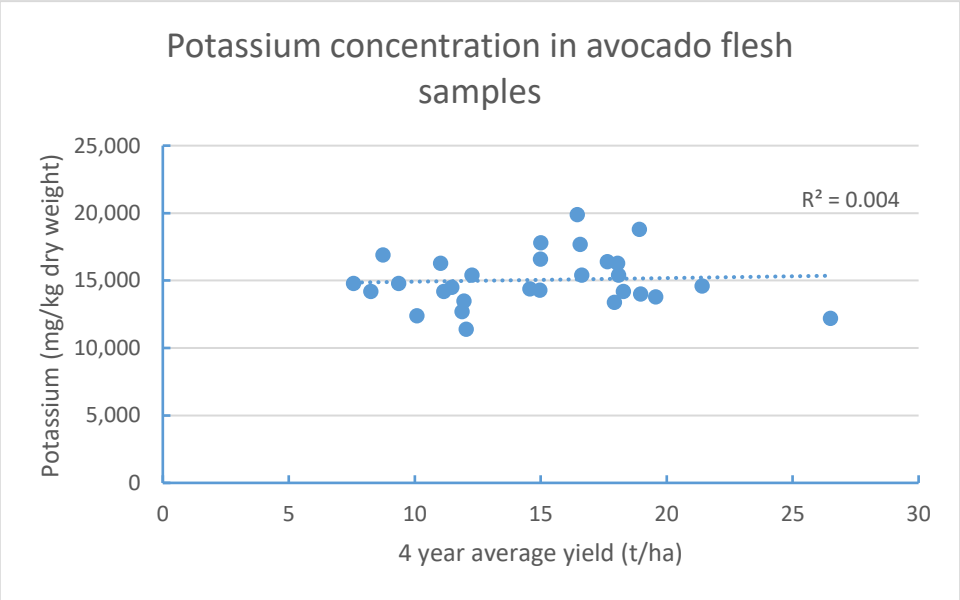

Figure 22.Potassium levels versus 4-years average yield in avocado flesh samples.

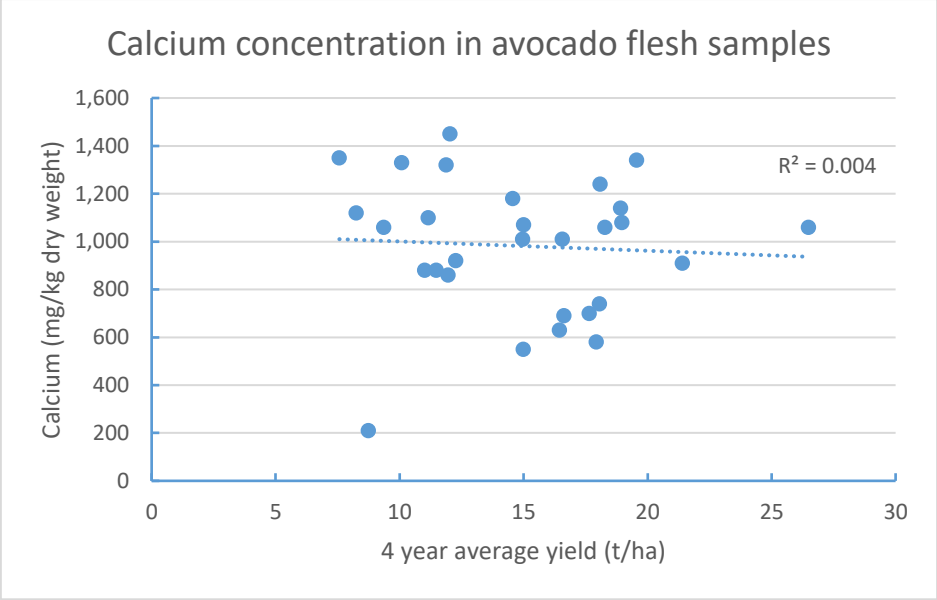

Figure 23.Calcium levels versus 4-years average yield in avocado flesh samples.

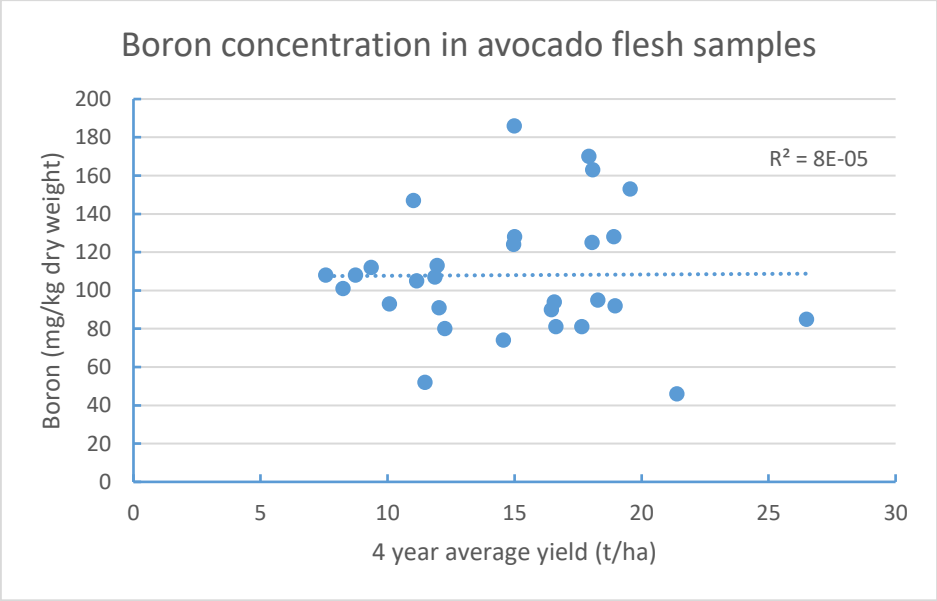

Figure 24.Boron levels versus 4-years average yield in avocado flesh samples.

5.6 Predicted means and confidence intervals of individual indicators

Table 14.Predicted means following by confidence intervals in parenthesis of individual indicators for low and high yielding orchards.

Lab	Variable	High performing soil	Low performing soil	pValue
Soil Foodweb	SF01	0.68(0.56-0.79)	0.66(0.54-0.79)	0.66
	SF02	27.62(17.76-37.47)	32.3(21.39-43.2)	0.15
	SF03*	317.85(197.18-512.37)	363.36(225.41-585.73)	0.22
	SF05	5.22(-1.04-11.48)	5.59(-1.19-12.37)	0.84
	SF06	320.62(163.28-477.95)	269.5(82.69-456.31)	0.48
	SF07	2.92(2.83-3.01)	2.87(2.76-2.99)	0.37
	SF08*	6840.11(993.96-47071.19)	12090.9(1756.98-83205.21)	0.12
	SF09*	2939.44(535.7-16129)	4467.78(814.23-24515.16)	0.32
	SF10*	245.95(55.11-1097.65)	123.07(27.58-549.26)	0.07
	SF11	0.52(0.38-0.65)	0.49(0.32-0.66)	0.74
Linnaeus	LN01	2.35(1.69-3)	2.68(1.85-3.51)	0.39
	LN04	0.53(0.43-0.62)	0.57(0.45-0.68)	0.43
	LN06	89.26(85.77-92.74)	88.82(84.96-92.69)	0.87
	LN07	84.41(79.39-89.44)	83.15(77.57-88.72)	0.73
	LN08	2.96(2.45-3.46)	2.73(2.14-3.33)	0.32
	LN09	6.13(4.83-7.44)	7.33(5.88-8.78)	0.22
	LN10	8.02(6.62-9.41)	9.32(7.77-10.86)	0.21
	LN12	82.5(79.2-85.8)	84.41(80.74-88.07)	0.43
	LN13	34.62(31.39-37.86)	35.75(32.14-39.36)	0.31
	LN14	7.46(5.88-9.04)	6.97(5.21-8.73)	0.67
	LN18	68.83(58.81-78.85)	66.46(55.34-77.58)	0.75
	LN19	86.28(80.77-91.78)	86.58(79.73-93.44)	0.92
	LN20	83.88(78.65-89.11)	83.22(77.42-89.03)	0.86
	LN23	1.86(1.5-2.22)	2.04(1.64-2.44)	0.5
	LN24	1.79(1.4-2.19)	2.03(1.59-2.47)	0.41
	LN25	97.07(94.37-99.77)	98.12(94.71-101.53)	0.53
	LN27	14.15(11.48-16.81)	16.65(13.69-19.61)	0.21
	LN28	41.53(33.91-49.15)	44(35.55-52.45)	0.66
	LN29	57.55(47.29-67.8)	62.69(51.31-74.07)	0.5
	LN32	0.35(0.15-0.55)	0.47(0.25-0.69)	0.4
Visual Soil Assessment	Soil Texture	1.5(1.35-1.65)	1.35(1.18-1.52)	0.2
	Soil Structure	1.8(1.64-1.97)	1.74(1.55-1.92)	0.59
	Earth Worms	11.19(6.81-18.39)	21.66(12.61-37.18)	**<0.001**
Hill soil nutrients	HS01	5.89(5.59-6.2)	5.8(5.42-6.19)	0.61
	HS02	85.44(50.4-120.47)	117(78.13-155.87)	0.23
	HS03	14.31(6.81-21.81)	17.47(9.08-25.86)	0.23
	HS04	11.96(8.38-15.55)	10.89(6.7-15.09)	0.49
	HS05*	37.83(19-75.34)	39.1(19.64-77.87)	0.84
	HS06	5.26(2.18-8.34)	5.25(1.84-8.66)	0.99
	HS07*	161.28(106.86-243.43)	131.17(86.91-197.99)	0.05

	HS08*	134.46(80.33-225.06)	108.23(64.66-181.16)	0.11
	HS09	2.51(1.89-3.13)	2.26(1.59-2.93)	0.15
	HS10	12.66(9.56-15.77)	10.61(7.04-14.19)	0.11
	HS11	13.35(9.91-16.78)	12.54(8.78-16.3)	0.44
	HS12	7.33(5.52-9.14)	6.14(4.05-8.22)	0.1
	HS13**	1440.67(826.51-2511.21)	1213.08(695.94-2114.49)	0.22
	HS14	0.58(0.34-0.82)	0.51(0.26-0.76)	0.17
	HS15	140(93.01-186.99)	138(85.87-190.13)	0.95
	HS16	56.81(38.75-74.87)	71.23(51.19-91.27)	0.28
	HS17	249.25(120.88-377.63)	311.06(166.62-455.49)	0.18
	HS18	1820.87(1385.47-2256.26)	1749.02(1210.51-2287.53)	0.76
	HS19	214.55(145.19-283.91)	213.94(131.15-296.73)	0.98
	HS20	26.03(12.11-39.95)	25.47(10.32-40.63)	0.89
	HS21	105.98(72.93-139.02)	98.15(60.09-136.21)	0.55
	HS22	36.5(-8.24-81.25)	35.68(-9.68-81.04)	0.87
	HS23	38.38(24.28-52.49)	41.76(26.11-57.41)	0.74
	HS24*	11.75(1.46-94.68)	6.84(0.85-55.09)	0.29
	HS25	5.27(2.15-8.38)	3.86(0.06-7.67)	0.38
	HS26	0.07(0-0.2)	0.05(0-0.19)	0.47
	HS27*	1290.59(1059.99-1571.35)	1342.29(1102.45-1634.3)	0.44
	HS28*	98.61(26.69-170.53)	50.01(0-133.01)	0.1
	HS29	0.92(0.35-1.49)	1.13(0.5-1.75)	0.24
	HS30	12.58(6.63-18.53)	11.22(4.61-17.83)	0.5
	HS31	2.36(1.3-3.43)	2.39(1.17-3.62)	0.95
	HS32	0.14(0.09-0.18)	0.12(0.07-0.18)	0.63
	HS33	3.5(2.69-4.32)	4.13(3.1-5.15)	0.23
	HS34	48.44(43.54-53.33)	44.62(39.18-50.05)	0.29
	HS35	9.4(7.32-11.48)	9.35(6.81-11.89)	0.96
	HS36*	0.55(0.24-1.25)	0.53(0.23-1.2)	0.83
	HS37	26.02(12.69-39.35)	24.11(10.1-38.12)	0.51
	HS38	61.94(56.49-67.39)	56.69(50.64-62.74)	0.2
	HS39	0.81(0.6-1.03)	0.83(0.61-1.04)	0.69

5.7 Importance estimate of the set of variables in predicting high and low yield

5.7.1 Soil Foodweb

Table 15. Importance estimate of Soil Foodweb variables in classifying between high and low yielding soils. If the average OOB delta error was 0.1 or greater, the variable is considered a good predictor; otherwise, the predictors were considered of low importance.

Ave. OOB delta error	Name	Label	Units
≥ 0.1	**Dry Weight**	**SF01**	**---**
≥ 0.1	**Total Bacteria (TB)**	**SF03**	**mg/kg**
≥ 0.1	**Total Fungi (TF)**	**SF06**	**mg/kg**
≥ 0.1	**Flagellates**	**SF08**	**number/g**
≥ 0.1	**Amoebae**	**SF09**	**number/g**
≥ 0.1	**Ciliates**	**SF10**	**number/g**
< 0	Active Bacteria (AB)	SF02	mg/kg
< 0	Active Fungi (AF)	SF05	mg/kg
< 0	Hyphal Diameter	SF07	µm
< 0	Endo (colonization)	SF11	%

5.7.2 Linnaeus

Table 16. Importance estimate of Linnaeus variables in classifying between high and low yielding soils. If the average OOB delta error was 0.1 or greater, the variable is considered a good predictor; otherwise, the predictors were considered of low importance.

Ave. OOB delta error	Name	Label	Units
≥ 0.1	**Gram Negative Bacteria**	**LN09**	**mg/kg**
0-0.1	Microbial Balance	LN12	Indicator
0-0.1	Nutrient Cycling Rate - MWSE only	LN19	Indicator
0-0.1	Pseudomonas	LN24	mg/kg
0-0.1	Residue Breakdown Rate - MWSE only	LN25	Indicator
0-0.1	Total Bacteria	LN27	mg/kg
0-0.1	Total Fungi	LN28	mg/kg
0-0.1	Total Microorganisms	LN29	mg/kg
0-0.1	Nutrient Solubilisation Rate - MWSE only	LN20	Indicator
< 0	Actinomycetes	LN01	mg/kg
< 0	Bacteria Stress Indicator	LN04	Indicator
< 0	Disease Resistance - MWSE only	LN06	Indicator
< 0	Drought Resistance - MWSE only	LN07	Indicator
< 0	Fungi to Bacteria Ratio	LN08	Ratio
< 0	Gram Positive Bacteria	LN10	mg/kg
< 0	Microbial Diversity Indicator	LN13	Indicator
< 0	Mycorrhizal Fungi (AMF)	LN14	mg/kg
< 0	Protozoa	LN23	mg/kg
< 0	True Anaerobic Bacteria	LN32	mg/kg

5.7.3 Hill Laboratories

Table 17. Importance estimate of Hill laboratories variables in classifying between high and low yielding soils. If the average OOB delta error was 0.1 or greater, the variable is considered a good predictor; otherwise, the predictors were considered of low importance.

Ave. OOB delta error	Name	Label	Units
≥ 0.1	**C/N Ratio**	**HS11**	**---**
≥ 0.1	**Iron (Mehlich 3)**	**HS21**	**mg/L**
≥ 0.1	**Aluminium (Mehlich 3)**	**HS27**	**mg/L**
0-0.1	Olsen Phosphorus	HS02	mg/L
0-0.1	Potassium	HS03	MAF
0-0.1	Sodium	HS06	MAF
0-0.1	Potentially Available Nitrogen (15cm Depth)	HS07	kg/ha
0-0.1	Anaerobically Mineralisable N	HS08	µg/g
0-0.1	Anaerobically Mineralisable N/Total N Ratio	HS09	%
0-0.1	Total Carbon	HS12	%
0-0.1	Sodium (Mehlich 3)	HS20	mg/L
0-0.1	Manganese (Mehlich 3)	HS22	mg/L
0-0.1	Cobalt (Mehlich 3)	HS26	mg/L
0-0.1	Calcium	HS30	me/100g
0-0.1	Magnesium	HS31	me/100g
0-0.1	Sodium	HS32	me/100g
0-0.1	Sodium	HS36	%BS
0-0.1	CEC	HS37	me/100g
0-0.1	Volume Weight	HS39	g/mL
< 0	pH	HS01	pH
< 0	Calcium	HS04	MAF
< 0	Magnesium	HS05	MAF
< 0	Organic Matter	HS10	%
< 0	Total Nitrogen	HS14	%
< 0	Phosphorus (Mehlich 3)	HS15	mg/L
< 0	Sulphur (Mehlich 3)	HS16	mg/L
< 0	Potassium (Mehlich 3)	HS17	mg/L
< 0	Calcium (Mehlich 3)	HS18	mg/L
< 0	Magnesium (Mehlich 3)	HS19	mg/L
< 0	Zinc (Mehlich 3)	HS23	mg/L
< 0	Copper (Mehlich 3)	HS24	mg/L
< 0	Boron (Mehlich 3)	HS25	mg/L
< 0	Total Copper	HS28	mg/kg
< 0	Potassium	HS29	me/100g
< 0	Potassium	HS33	%BS
< 0	Calcium	HS34	%BS
< 0	Magnesium	HS35	%BS
< 0	Total Base Saturation	HS38	%

5.7.4 Hill Laboratories and Soil Foodweb combination

Table 18. Importance estimate of Hill Laboratories and Soil Foodweb combination variables in classifying between high and low yielding soils. If the average OOB delta error was 0.1 or greater, the variable is considered a good predictor; otherwise, the predictors were considered of low importance.

Ave. OOB delta error	Name	Label	Units
≥ 0.1	**Iron (Mehlich 3)**	**HS21**	**mg/L**
≥ 0.1	**Aluminium (Mehlich 3)**	**HS27**	**mg/L**
0-0.1	pH	HS01	pH
0-0.1	Olsen Phosphorus	HS02	mg/L
0-0.1	Calcium	HS04	MAF
0-0.1	Sodium	HS06	MAF
0-0.1	Potentially Available Nitrogen (15cm Depth)	HS07	kg/ha
0-0.1	Anaerobically Mineralisable N/Total N Ratio	HS09	%
0-0.1	C/N Ratio	HS11	---
0-0.1	Total Carbon	HS12	%
0-0.1	Cobalt (Mehlich 3)	HS26	mg/L
0-0.1	Magnesium	HS31	me/100g
0-0.1	Sodium	HS32	me/100g
0-0.1	Sodium	HS36	%BS
0-0.1	CEC	HS37	me/100g
0-0.1	Volume Weight	HS39	g/mL
0-0.1	Total Bacteria (TB)	SF03	mg/kg
0-0.1	Total Fungi (TF)	SF06	mg/kg
0-0.1	Hyphal Diameter	SF07	µm
0-0.1	Flagellates	SF08	number/g
0-0.1	Ciliates	SF10	number/g
0-0.1	TF/TB	SF12	ratio
0-0.1	Sodium (Mehlich 3)	HS20	mg/L
<0	Potassium	HS03	MAF
<0	Anaerobically Mineralisable N	HS08	µg/g
<0	Manganese (Mehlich 3)	HS22	mg/L
<0	Calcium	HS30	me/100g
<0	Magnesium	HS05	MAF
<0	Organic Matter	HS10	%
<0	Total Nitrogen	HS14	%
<0	Phosphorus (Mehlich 3)	HS15	mg/L
<0	Sulphur (Mehlich 3)	HS16	mg/L
<0	Potassium (Mehlich 3)	HS17	mg/L
<0	Calcium (Mehlich 3)	HS18	mg/L
<0	Magnesium (Mehlich 3)	HS19	mg/L
<0	Zinc (Mehlich 3)	HS23	mg/L
<0	Copper (Mehlich 3)	HS24	mg/L
<0	Boron (Mehlich 3)	HS25	mg/L
<0	Total Copper	HS28	mg/kg

<0	Potassium	HS29	me/100g
<0	Potassium	HS33	%BS
<0	Calcium	HS34	%BS
<0	Magnesium	HS35	%BS
<0	Total Base Saturation	HS38	%
<0	Dry Weight	SF01	---
<0	Amoebae	SF09	number/g
<0	Active Bacteria (AB)	SF02	mg/kg
<0	Active Fungi (AF)	SF05	mg/kg
<0	Endo (colonization)	SF11	%

5.8 Decision trees

These trees are one out hundred trees produce by machinery learning analysis; they are not a conclusive methodology for reading laboratory results since the dataset was limited to 29 orchards. To read the decision tree, follow the logic in the example below:

Figure 25, if total bacteria is less than 292, then it is a high yielding orchard. Otherwise, if Total bacteria is equal to or greater than 471, you have a high yielding orchard. Otherwise, if Endomycorhyizae is lower than 0.425, you have a high yielding orchard, otherwise low yielding orchard.

5.8.1 Soil Foodweb

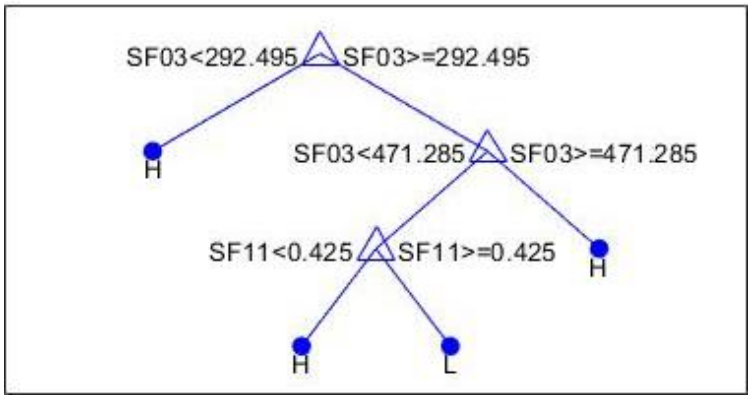

Figure 25. An example of a decision tree for classifying between low and high yielding orchards by Soil Foodweb biological variables Total bacteria (SF03) and Endomycorhyizae (SF11).

5.8.2 Linnaeus

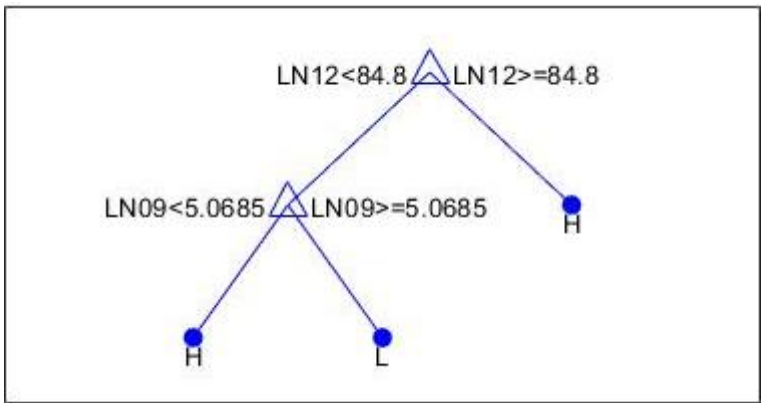

Figure 26. An example of a decision tree classifying between high performing and low yielding orchards by Linnaeus biological variables. LN12 – Microbial balance and LN09 – Gram negative bacteria.

5.8.3 Hill Laboratories

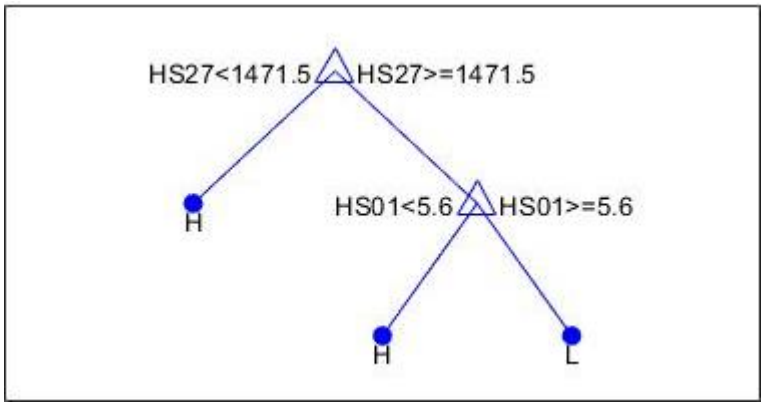

Figure 27. An example decision tree for classifying between high and low yielding soils by Hill laboratories nutrient variables, HS27 – Aluminium Mehlich 3 and HS01 – pH.

5.8.4 Combination of Soil Foodweb and Hill laboratories soil nutrients

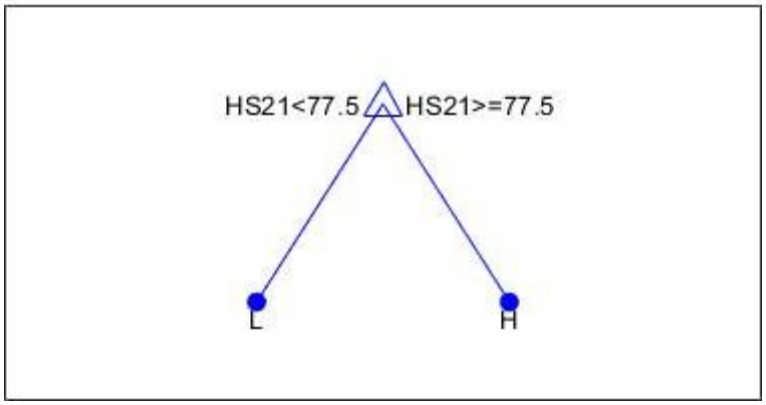

Figure 28. An example decision tree for classifying between high and low yielding soils by Hill laboratories nutrient variables combined with Soil Food Web biological variables.