SOIL HEALTH IN SMALLHOLDER COLOMBIAN COFFEE SYSTEMS AND ITS SOCIO-ECONOMIC IMPLICATIONS

A Thesis

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Master of Science

by

Fatma Rekik

January 2017

ABSTRACT

Coffee is a high value commodity in developing countries that is second in total commercial value only to petroleum. Understanding its agronomic characteristics including soil health (SH) is crucial for the environmental and socio-economic sustainability of coffee production. In this multidisciplinary study, SH and associated farmers' perceptions were explained by farmers' geo-demographic traits. Results revealed that female and co-op member-farmers have higher SH than their counterparts, correct SH perception is associated with farm SH conditions, and biological indicators are most related to farmers' perceptions. In the second chapter, the scoring framework created for assessing the Colombian soil samples was compared to other regional scoring frameworks. Results showed agreement in scoring SH indicators with disagreements arising from the contrasting inherent soil properties of the different localities. In the third chapter, the relationship between coffee quality and SH was assessed. Results showed that coffee quality is negatively correlated with key SH indicators.

BIOGRAPHICAL SKETCH

Fatma Rekik was born on August 24th, 1993 in Arlington, Virginia to Tunisian immigrant parents. At age nine, she and her family moved to Tunisia where she experienced an intense culture shock within the education system which hindered her from doing well in school. In January 2011, as a senior in high school, Fatma witnessed the birth of the Arab Spring in her hometown, and watched the people revolting against the old regime with the hopes of ensuring a better life for themselves and their families. Fatma saw a light at the end of the tunnel and soon experienced what she perceived as her own personal revolution. That summer, Fatma travelled back to the US and started college at Northern Virginia Community College (NVCC) where she excelled in both the academic and the extracurricular arenas. She then transferred her Associate's degree from NVCC to Cornell University's Agriculture Science major as a junior and continued her involvement in research and public service. Fatma joined a multidisciplinary coffee research project in her senior year as an undergraduate research assistant to Miguel Gomez while working in the soil health lab. That year, Fatma received the NSF-GRFP fellowship.

With her major advisor, Dr. Harold van Es, Fatma was able to expand the economics-focused project and create her own Master's research topic studying soil health in Colombian coffee farms and its social, environmental and economic factors.

To my

loving parents

Who have fought hard in life

That I may be where I am today

&

Who have

Encouraged me and believed in my capacities

Throughout my entire life.

Thank you

for

Believing in education

&

Women empowerment

ACKNOWLEDGMENTS

I would like to express my profound gratitude to my major advisor Dr. Harold van Es who has not only mentored me as a graduate student in his lab despite the demands of his positions as the President of the Soil Science Society of America, but also as an undergraduate student. I thank him for providing advice and mentorship on research and career issues and for opening up the doors to so many opportunities in professional and academic development. He is the best advisor one could ever hope for.

My gratitude further extends to my minor advisor, Dr. Miguel Gomez of the Charles Dyson School of Applied Economics and Management, for recruiting me into this project as an undergraduate research assistant in my senior year, and who, together with his exceptionally kind PhD student Juan Nicholas Hernandez-Aguilera, have generously shared their data so that my research could target other areas of this project and relate it to soil health. I am eternally grateful to both.

I am moreover grateful to my minor advisor, Dr. Peter Hobbs, who taught me many classes at Cornell University and whose invaluable international experience has brought my research to a whole new level. I thank him for his continuous successful attempts in connecting me with some of his prominent colleagues in international research institutions and for advising me on career different career paths.

I would like to gratefully acknowledge the Cornell Soil Health Team with special courtesy to Bob Schindelbeck and Kirsten Kurtz who championed the logistics of the pre- and post- phases

of soil sampling, and whose dedication, attention to detail and assertiveness had fueled the efficiency and successful outcome of the soil health assessment.

I owe deep appreciation to Stephen Parry of the Cornell Statistical Consulting Unit and Ahmad Rafiqi, PhD candidate in Cornell's Mathematics department who have provided invaluable statistical advice in data analysis throughout my time working on this project.

I am hugely grateful to all my lab- and office-mates; Leilah Krounbi, Becky Marjerison Rintaro Kinoshita, Phillip Frost and Shai Sela who have all provided in one way or another good company, engaging conversation and uplifting support. Special shout out to Becky who I always considered more than just an officemate but a friend, an acquaintance, a computer guru and someone to lean on when in need of support. I will forever cherish our conversations that ranged from science, to politics, to religion and culture. My dearest Leilah, who has been there for me more than I keep count of and who became the best friend I could ever ask for at Cornell. You are very special and I owe my success in some way to you. Thank you for believing in me and for being there for me. You are an exceptional human being.

I would like to express deep gratitude to the personnel at the Federación Campesina del Cauca especially Saul Agredo and Yimara Martinez, and the collaborating farmers without whom this study would not have been possible.

Finally, I could not have done this work without funding support from the National Science Foundation Graduate Research Fellowship Program. Thank you for your generosity in supporting scientific research.

To my love Houssem, thank you for being who you are and for your unconditional love and support. Your perfectionism and thorough attention to detail has opened my eyes to new ways in conducting and reporting research effectively. You're the best poster critic there is.

To my family, friends and colleagues, thank you all. Having you in my life is a blessing from God and has made my journey truly heavenly. Do know that your support- little or big- is significant to me and will forever be cherished.

Alhamdullilah!

الحمد لله!

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

INTRODUCTION

SH = soil health

CHAPTER 1

SH = soil health, RCM = relationship coffee model, FCC = Federación Campesina del Cauca, AWC = Available Water Capacity, WAS = Wet Aggregate Stability, OM = Organic Matter, AC = Active Carbon, Protein = Autoclaved Citrate Extractable Soil Protein Index, P = Extractable Phosphorus, K = Extractable Potassium, Mg = magnesium, Fe = iron, Mn = manganese, Zn = zinc, $(NaPO_3)_n$= sodium hexametaphosphate, ICP–OES = Inductively Coupled Plasma Optical Emission Spectroscopy, $KMnO_4$ = potassium permanganate solution, CND = Cumulative Normal Distribution function, ANOVA = Analysis of Variance, PCA = Principal Component Analysis

CHAPTER 2

SH = soil health, RCM = relationship coffee model, FCC = Federación Campesina del Cauca, AWC = Available Water Capacity, WAS = Wet Aggregate Stability, OM = Organic Matter, AC = Active Carbon, Protein = Autoclaved Citrate Extractable Soil Protein Index, P = Extractable Phosphorus, K = Extractable Potassium, Mg = magnesium, Fe = iron, Mn = manganese, Zn = zinc, $(NaPO_3)_n$= sodium hexametaphosphate, ICP–OES = Inductively Coupled Plasma Optical Emission Spectroscopy, $KMnO_4$ = potassium permanganate solution, CND = Cumulative Normal Distribution function, PCA = Principal Component Analysis, BSR = Best Subsets Regression, CCSHS-W = Weighted Colombian Coffee Soil Health Scoring, ICC = Intraclass correlation coefficient, CCSHS = Colombian Coffee Soil Health Scoring, CASH = Comprehensive Assessment of Soil Health

CHAPTER 3

SH = soil health, FCC = Federación Campesina del Cauca, AWC = Available Water Capacity, WAS = Wet Aggregate Stability, OM = Organic Matter, AC = Active Carbon, Protein = Autoclaved Citrate Extractable Soil Protein Index, P = Extractable Phosphorus, K = Extractable Potassium, Mg = magnesium, Fe = iron, Mn = manganese, Zn = zinc, $(NaPO_3)_n$= sodium hexametaphosphate, ICP–OES = Inductively Coupled Plasma Optical Emission Spectroscopy, $KMnO_4$ = potassium permanganate solution, CND = Cumulative Normal Distribution function, SCAA = Specialty Coffee Association of America, (SS) sensorial scores, MLR = Multiple Linear Regression, BSR = Best Subsets Regression

INTRODUCTION

As we consume one of the most sought after beverages in the world, we seldom stop to wonder what came about in making this cup of coffee possible, or better yet, to whom we owe this magical drink. In fact, almost all the coffee consumed globally is produced by poor, subsistence smallholder farmers (ICO, 2006). What makes this phenomenon alarming is that these resource-limited farmers are finding it increasingly challenging to keep their production at levels that will ensure a stable income for them and their families, which is largely due to drastic changes in global-coffee market prices and climate trends, which have negatively affected farmers' profitability. What is worth our attention is that soil health (SH) might be able to address the issue of income instability, and mitigate its repercussions. Aside from being a prime determinant of agricultural productivity, SH may contribute directly to the economic welfare of these farmers through the copious marketing advantages that it has to offer. Through soil conserving agricultural practices that are environmentally friendly and socially responsible, farmers are able to receive price premiums from local, national and international organization by promoting sustainability. Therefore, it is crucial to understand the characteristics and determinants of soil health on coffee farms and its effect on coffee cup quality which is another important source of premiums. This will help maximize the utility for farmers to better their chances of ensuring access to these price premiums.

CHAPTER 1. UNDERSTANDING SOIL HEALTH AND ASSOCIATED FARMERS' PERCEPTIONS IN COLOMBIAN COFFEE SYSTEMS

1.1. ABSTRACT

Soil health is important to the economics and environmental impacts of crop production, including coffee culture. This study was conducted to gain insights into farmers' perceptions related to soil health concepts and their realities on Colombian coffee farms. A total of 223 soil samples were collected from 145 coffee farms in Cauca, Colombia that vary by municipality, their membership status with a coffee co-operative (member; non-member), and the gender of farmer. Samples were analyzed for 10 soil health indicators including wet aggregate stability, available water capacity, respiration rate, pH, and contents of active carbon, organic matter, protein, phosphorus, potassium and minor elements. Farmer gender (females>males), municipality, and co-op membership (members> non-members) were significant factors for soil health status on farms. Farmer members of the co-op were also asked to identify from their farm those plots of perceived highest and lowest soil fertility, which allowed for the evaluation of (i) the correctness of farmers' soil fertility ranking, and (ii) which soil health indicators most influence farmers' perception of soil health. Results revealed that these coffee farmers were likely to correctly identify the level of soil health on their farms, and that organic matter content, respiration, and protein content were most correlated with farmers' perception of soil health. Farmers' soil health perception correctness was not significantly correlated with gender or municipality, but the perception correctness was greater with producers owning farms with higher soil health conditions.

1.2.KEYWORDS

Colombia, coffee, soil health, farmer perception, gender, co-op membership, Relationship Coffee Model

1.3.INTRODUCTION

Soil Health in Coffee Systems

Soil health can be a prime determinant of agricultural productivity both in terms of quality and quantity of yields. The ability to manage for soil health is crucial for environmental and economic reasons, especially for high value globally traded commodity crops like coffee (*Coffea spp.*) where actual or perceived sustainability may offer a marketing advantage. Coffee is a global commodity that is second in total commercial value only to petroleum (Haight, 2011) and has, in the aftermath of drastic changes in global coffee markets that negatively affected smallholder coffee growers, recently upgraded its value chain for specialty coffee (Hernandez-Aguilera *et al.,* 2015). This new model - Relationship Coffee Model (RCM) - ensures specific production standards that are socially and environmentally responsible, implicitly also incorporating soil health. Organizations such as coffee co-ops that have adopted RCM work to promote sustainability of both natural and human resources while emphasizing membership requirements and regulations. Hence, both parties benefit: the organization, by exporting high value specialty coffee, and the farmer, by receiving quality-based price premiums and having access to services provided by the RCM co-op (Hernandez-Aguilera *et al.,* 2015). These may include education and consultations; access to agricultural inputs; farm visits and recommendations by co-op personnel; workshops; etc. Understanding the factors that play a role

in soil health can be useful not only to researchers, educators and RCM personnel, but also to farmers who are looking for information to help them with decision making.

It is important for farmers to assess the health of their soils. In-field measurements are often the only option for soil assessment that subsistence farmers have due to the high costs of laboratory analyses. A practical approach can save time, money and energy if farmers are able to qualitatively assess their soil health and manage it in a timely manner. Previous perception studies in natural settings have predominantly revolved around environmental conscience and farmer climate-change awareness. One (Rahman, 2003) assessed farmers' awareness of adverse environmental impacts caused by agricultural technology and another study (Munyuli, 2011) addressed the key concepts in farmers' perception and management of natural resources (among others), but were not explicitly linked to soil health or fertility. The necessity to look at farmers' holistic perception of soil health will enable researchers, educators and extension workers to better communicate with farmers about soil health and help them overcome barriers established by differences in problem formulation (Karltun *et al.,* 2013). Few studies, however, have assessed farmer soil health perception. Among them, (Munyuli, 2011) conducted a gender-based farmer study on their perceptions of the importance of pollinators in coffee production in Uganda, which briefly touched upon the issue of soil health. It revealed that female farmers accepted the concept of soil fertility restoration as a basic component of coffee production enhancement more than male farmers. Other studies found that farmers typically associate soil health to organic matter content followed by crop appearance and biological activity (Romig *et al.* 1995); (Karltun *et al.,* 2013). However, findings by Ryder (2003) suggest that farmers' perceptions of soil fertility may vary regionally.

The objectives of this study are (i) to identify the demographic factors that play a role in farmers' perceptions of soil health in southern Colombian coffee systems; (ii) to identify the soil health parameters that most influence such farmers' perceptions; (iii) to assess whether Colombian coffee farmers have accurate perceptions of their soil health; and (iv) to identify the factors influencing farmers' perception accuracy of soil health.

1.4.MATERIALS AND METHODS

1.4.1. Project Location and Site Description

The study was performed in a predominantly coffee-growing region within the Department of Cauca, Colombia, situated at approximately 2.2°N and −76.4°W (Figure 1.1). The farm fields used in this study lie on elevations that range between 1269 and 1959 m as that provides favorable conditions for coffee cultivation in the tropics. Rainfall in Cauca ranges between 260.9-and 313.2- mm y^{-1} and has a bimodal distribution centered around the months of April and November (computed from: Promedios Climatológicos 1981 - 2010.xlsx)[1]. The soils in our project area are Andisols, of volcanic ash origin (Universidad del Cauca, unpublished data).

Coffee production in the region is mainly conducted by small-scale subsistence or less-resourced farmers in either monoculture or polyculture. Crops that accompany coffee trees in polyculture settings are typically grown for market or domestic purposes and mainly include maize (*Zea Mays*). A variety of shade tree species provide good canopy cover for the coffee and other ecosystem services (Hernandez-Aguilera *et al.,* 2015).

[1] http://www.ideam.gov.co/web/tiempo-y-clima/clima. Accessed: 2/21/2016

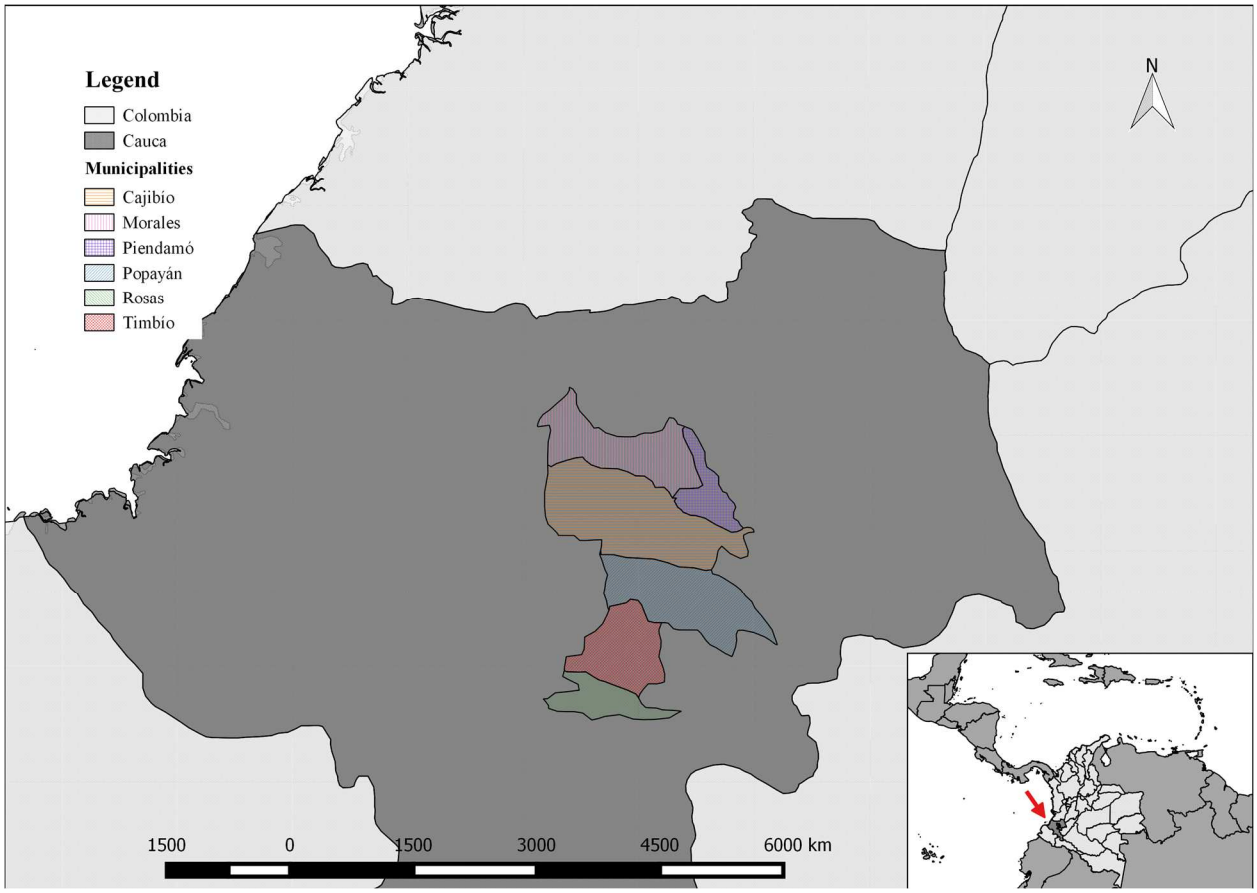


Figure 1.1: Project Location in Cauca Department showing the six selected municipalities

Coffee growers in Cauca, Colombia are largely split into two marketing groups: members of a Relationship Coffee Model (RCM) co-op - Federación Campesina del Cauca (FCC) headquartered in Popayan, and those who sell to the regular commodity market. As a result of their RCM membership, participating growers have access to coffee quality price-premiums, certification, loans, immunity to coffee price volatility, agricultural and commercial services, and access to the international market (Hernandez-Aguilera *et al.,* 2015).

1.4.2. Sampling, Laboratory and Scoring Methods

A total of 223 soil samples were collected in January 2014 following the protocol described by Moebius-Clune *et al.* (2016) from 145 coffee farms across 6 municipalities (Cajibio, Timbío,

Rosas, Piendamò, Morales and Popayàn). All samples were collected from the 0- to15-cm depth range using a Dutch-style soil auger after surface residue removal. At each sampling location, five samples were collected and then composited to obtain a representative sample. Basic demographic information about the growers was collected to accompany each soil sample, including: gender of grower, status of association with RCM co-op, and location.

Fifty-four percent of coffee farms are members of the local RCM co-op, and the remaining are independent (Table 1.1). Farmers who belong to the cooperative were asked to designate on their farms areas that they perceived to be "most fertile" and "least fertile", and soil samples were subsequently collected from each.

Table 1.1. Demographic Information of Sampled Coffee Growers

Municipality	Male	Female	ns †
	Co-op Members		
Rosas	9	4	
Timbío	9	4	
Piendamó	7	3	3
Cajibio	7	6	
Morales	9	4	
Popayán	9	4	
	Co-op Non members		
Rosas	9	3	
Timbío	9	3	
Piendamó	8	4	
Cajibio	10	1	
Morales	6	4	
Popayán	9	1	
Total	**101**	**41**	**3**

† **ns** = not surveyed

Soil samples were sent to Cornell University in Ithaca, NY (USA) where processing and analysis of physical, chemical and biological soil health indicators were performed according to the Comprehensive Assessment of Soil Health (CASH) protocol (Moebius-Clune *et al.,* 2016). Briefly, this includes:

Physical Indicators: Available Water Capacity (AWC) between field capacity (-10kPa) and permanent wilting point (-1500kPa) was assessed gravimetrically by equilibrating saturated soil to each of 10 kPa and 1500 kPa on ceramic high pressure plates (Topp and Zebchuck 1979). The difference between soil water loss under 10 kPa and 1500 kPa pressures determined from calculating the difference in wet and dry weights was considered the AWC (Moebius-Clune *et al.,* 2016).

Wet Aggregate Stability (WAS) was assessed using a rainfall simulator adapted from Ogden *et al*. (1997) that allows particles of air-dried soil placed on a 0.25 mm mesh sieve to slake under 2.5 J of rainfall energy for 300 seconds, based on a total of 2.5 cm of rainfall. Wet Aggregate Stability was determined by subtracting the weight of slaked soil plus the remaining stones on the sieve (>0.25 mm) from total soil weight measured before rainfall (Moebius-Clune *et al.,* 2016).

Soil texture was determined using a rapid quantitative method developed by Kettler *et al.* (2001) where soil samples were fractionated with 3% sodium hexametaphosphate ($(NaPO_3)_n$) and a series of sieving and sedimentation steps was used to separate the different particle sizes.

Biological Indicators: Organic Matter content (OM) was analyzed by mass loss on ignition in a muffle furnace at 500 °C for two hours, with values corrected by multiplying percent loss on ignition by 0.7 and subtracting 0.23 (Moebius-Clune *et al.,* 2016). Active Carbon (AC) was measured by adding a dilute potassium permanganate solution ($KMnO_4$) to soil, which acts as an oxidant to AC, and measuring the solution's absorbance at 550 nm using a hand-held colorimeter (Hach, Loveland, CO) (Weil *et al.,* 2003).

Autoclaved Citrate Extractable Soil Protein Index (Protein) was measured by extracting proteins from the soil following a series of centrifugation and autoclaving steps using 0.02 M sodium citrate at pH 7. Soil protein concentration was determined by measuring bicinchoninic acid assay against bovine serum albumin standard curve for soil protein concentration (Walker, 2009; Wright and Upadhyaya, 1996). The soil Respiration test was performed by trapping and measuring CO_2 emitted by soil microorganisms over a 4-day room temperature incubation in a sealed chamber with a KOH trap (Haney and Haney, 2010). The dissolved CO_2 in the KOH trap quantifying microbial activity was measured using an electrical conductivity meter, and the change in conductivity before and after incubation quantifies the amount of CO_2 evolved.

Chemical Indicators: Soil pH was measured in a 1:1 water dispersed slurry determined by a pH electrode probe (SM802 Smart Combined Meter, Milwaukee Industries, Rocky Mount, NC). Soil nutrients, including P, K, Mg, Fe, Mn and Zn were extracted with a Modified Morgan solution (ammonium acetate - buffered at pH 4.8), and quantified by Inductively Coupled Plasma Optical Emission Spectroscopy (ICP–OES, Varian 730-ES, Mulgrave, Victoria, Australia).

Measured results for each individual SH indicator were translated into indicator scores following the approach of Andrews *et al.* (2004). It consisted of comparing each individual measurement to a standardized dataset for Cauca and Antioquia, Colombia. Scoring functions were used to establish scoring curves for each individual indicator in one of three forms ("more is better", "optimum range", and "less is better") without adjusting for texture. Scoring functions for the physical and biological indicators and pH followed cumulative normal distribution (CND) curves specific to each indicator: The Cumulative Normal Distribution function, CND (μ ,σ), gives the

probability (between 0 and 1) that a member of the distribution is less or equal to the SH indicator measurement, x (Eq. [1]):

$$CND\ (x, \mu, \sigma) = \frac{1}{2}\left(1 + erf\left[\frac{(x-m)}{s\sqrt{2}}\right]\right), \quad -\infty < x < \infty \qquad [1]$$

where parameters such as μ and σ were either estimated by the sample mean (m) and standard deviation (s) respectively, or were based on outcome-based universal thresholds as determined by Moebius-Clune *et al.* (2016) for the case of P, K and minor elements. Scoring functions were derived by multiplying CND functions by 100. From the individual indicator scores, an overall Soil Health Index Score is calculated as their unweighted arithmetic mean of each of the ten indicator scores (Moebius-Clune *et al.,* 2016).

1.4.3. Statistical Methods

Analysis of Variance (ANOVA) was performed on the collected demographic data along with the soil health results of the 223 soil samples to assess relationships among the factors affecting soil health. A separate analysis was performed on the paired 156 soil samples from the 78 farms that belong to the RCM co-op to assess farmers' perception accuracy of soil health which is defined as their ability to correctly differentiate between their high and low fertility plots. ANOVA assumptions were checked and mean separation was computed using Tukey's test at $\alpha = 0.05$.

Logistic Regression and Principal Component Analysis (PCA) were additionally performed to evaluate which indicators most affected farmers' soil health perceptions. The logistic regression was based on standardized soil health indicator values based on sample mean and

sample standard deviation to account for variation arising from the different indicator units (Eq. [2]).

$$y_{standardized} = \frac{y_i - \mu_y}{\sigma_y} \quad [2]$$

Finally, farmers' perception accuracy was analyzed relative to their gender, proximity to FCC headquarters in Popayan, and farm soil health conditions, using Fisher's Exact Test for Count Data - a more accurate test than chi-square or G–test when the expected numbers are less than 1000 (McDonald, 2009). For this, categorical data of farmers' soil ranking were translated into a continuous variable which was the difference of the overall Soil Health Index Score of the most fertile sample and the least fertile sample, as perceived by the farmer. This portrays the degree to which each farmer had accurately perceived their soil health. Accurate perceptions result in a positive number, and inaccurate perceptions in a negative number, with more extreme numbers indicating more intense accuracy or inaccuracy.

All statistical analyses were performed using the R-Project for Statistical Computing (Team, 2013).

1.5.RESULTS AND DISCUSSIONS

1.5.1. Soil Health Results and Description

Table 1.2 Summary of Soil Health Results (mean (sd)) for Cauca Department, including measured values and overall soil health index score.

	Rosas	Timbío	Piendamó	Cajibio	Morales	Popayán	**Total Averages**
WAS (%)	85.1(14.9)b†	93.5(5.0)a	96.9(2.8)a	97.5(1.7)a	97.0(2.1)a	96.1(6.0)a	**94.3(8.3)**
AWC (g. g^{-1})	0.2(0.0)d	0.3(0.08)bc	0.4(0.1)a	0.3(0.1)bc	0.3(0.1)ab	0.2(0.1)c	**0.3(0.1)**
OM (%)	10.1(4.5)c	18.5(4.3)ab	18.5(5.4)ab	18.8(6.0)ab	21.6(4.4)a	17.3(4.5)b	**17.4(6.0)**
AC (ppm)	561(178)c	852(175)b	854(234)b	856(219)b	1008(170)a	788(206)b	**818.3(237.7)**
Protein (mg. g^{-1})	8.7(2.8)a	9.3(1.9)a	9.7(2.7)a	9.7(2.5)a	9.6(2.0)a	8.5(2.7)a	**9.2(2.5)**
Respiration (mg. g^{-1})	0.9(0.3)c	1.0(0.1)bc	1.0(0.2)abc	1.0(0.3)ab	1.2(0.2)a	0.9(0.3)bc	**1(0.2)**
pH	4.9(0.3)a	4.8(0.3)abc	4.7(0.3)c	4.7(0.3)bc	4.8(0.4)abc	4.9(0.3)ab	**4.8(0.3)**
P (ppm)	6.9(4.3)d	10.1(2.2)abc	11.8(5.2)ab	9.4(3.7)bc	12.1(3.7)a	7.7(3.5)cd	**9.6(4.3)**
K (ppm)	136.0(75.7)a	118.4(88.1)a	101.6(58.4)a	93.5(59.3)a	119.0(81.9)a	125.0(109.8)a	**115.4(80.9)**
Mg (ppm)	487.0(520.3)a	82.7(74.3)b	61.4(49.6)b	180.5(562.2)b	61.9(72.1)b	119.3(238.0)b	**169.9(361.6)**
Fe (ppm)	25.7(17.0)a	17.3(8.3)b	21.3(8.7)ab	20.5(14.1)ab	22.2(5.9)ab	16.6(6.4)b	**20.6(11.2)**
Mn (ppm)	10.4(4.0)a	7.4(3.7)b	5.3(3.3)bc	6.4(6.0)b	2.9(1.7)c	5.6(4.6)bc	**6.4(4.7)**
Zn (ppm)	1.2(1.2)a	0.7(0.4)a	1.0(0.9)a	0.7(0.7)a	1.1(1.0)a	0.8(0.7)a	**0.9(0.9)**
Soil Health Index Score	**44.8(12.9)c**	**59.1(8.8)b**	**61.1(13.4)ab**	**58.6(10.9)b**	**67.1(9.7)a**	**54.9±(10.0)b**	**57.5(13.2)**

† a, b, c, d significant homogeneous groups among municipalities for each indicator using Tukey's HSD at $p<0.05$.

Table 1.2 shows measured values for the soil health indicators and the overall soil health index score (scale 0-100) for each municipality. Results reveal that indicator measurements for WAS, AWC, OM, AC, respiration, pH, P, Mg, Fe and Mn differed significantly across municipalities (Table 1.2). Among these indicators, Rosas consistently had the lowest measured values for all physical and biological indicators, and the highest measured values for the chemical indicators with the exception of P. Morales on the other had the highest measured values among the six municipalities for WAS (97%), OM (21.6 %), AC (1008 ppm), respiration (1.2 mg. g^{-1}) and P (12.1 ppm; Table 1.2). Consequently, Rosas scored the lowest overall soil health (44.8±12.9), while Morales scored the highest among the six municipalities (67.1±9.7), which highlights the importance of looking beyond chemical indicators during soil health assessment (Moebius-Clune *et al.,* 2016; Table 1.2). The overall soil health index score of the 223 soil samples from all municipalities averaged 57.5 with a standard deviation of ±13.2.

Results from the PCA analysis reveal that the first principal component, which explains 34% of the total variability, is strongly correlated with seven of the soil health indicator variables, where a correlation value >0.5 is deemed important. Principal Component 1 increases with increasing OM, AC, AWC, Respiration, P, WAS and Protein, suggesting that these seven indicators vary together (Figure 1.2; Table 1.3).

However, because we see that the first principal component correlates most strongly with OM (r=0.930), we could conclude that this principal component is primarily an indicator of broader benefits associated with higher OM levels in soil (Table 1.3).

The second principal component explains 17% of total variability, and is related to increasing levels of K, Zn, Mn, and P. This component can be viewed as a measure of the chemical

indicators of soil health, and this suggests that nutrient availability tends to be consistent across individual nutrients, and that nutrients are generally co-managed (Figure 1.2; Table 1.3). The third principal component does not show any strong correlations with any of the variables (Table 1.3).

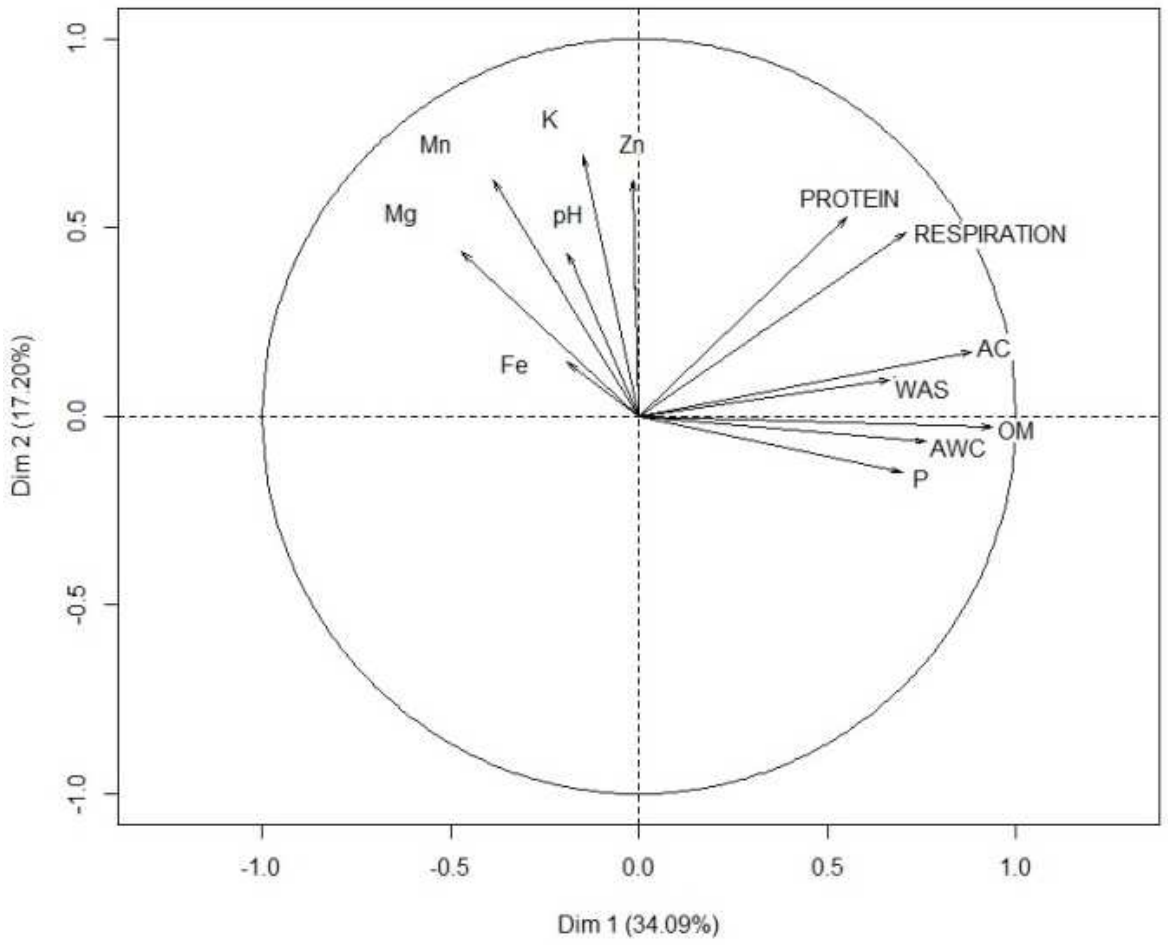


Figure 1.2: Variables Factor Map of PCA

Table 1.3 Dimension loadings and explanation of variance of first three principal components.

Indicator	PC 1	PC 2	PC 3
WAS	**0.662** †	0.097	0.171
AWC	**0.764**	-0.052	-0.268
OM	**0.931**	-0.026	-0.040
AC	**0.882**	0.175	-0.227
Protein	**0.570**	**0.511**	0.223
Respiration	**0.707**	0.488	-0.022
pH	-0.214	0.436	-0.777
P	**0.703**	-0.148	-0.300
K	**-0.149**	0.691	0.002
Mg	**-0.471**	0.435	0.201
K	-0.170	**0.690**	0.004
Mg	-0.474	0.432	-0.177
Fe	-0.182	0.139	0.392
Mn	-0.379	**0.610**	0.452
Zn	-0.023	**0.630**	0.049
Standard Deviation	2.1126	1.4852	1.13584
% of variance	34.333	16.967	9.924
Cumulative % of var.	34.333	51.299	61.223

†Values in bold represent strong correlation (r>0.5)

1.5.2. Factors Affecting Soil Health

The overall soil health index score was significantly different between farmer genders ($p=0.046$; Figure 1.3; Table 1.4). Fields managed by female farmers having higher soil health than those by male farmers (Figure 1.3; Table 1.4) is in line with findings by (Munyuli, 2011) that female farmers are more aware of certain aspects in their environments and are more understanding of the importance of soil fertility than males. Additionally, De Jalón et al. (2015) found that females are less "likely to deny the existence of climate change" and "are more likely to take up climate change action" and are perhaps also more prone to build soil health and thereby mitigate the consequences of climate change.

Overall soil health index score was also significantly different among municipalities ($p=1.05e15$; Figure 1.4; Tables 1.4; 1.5), which is likely due to inherent soil properties rather than large changes in soil management. The Morales municipality showed the highest average soil health, while Rosas has the lowest (Figure 1.4). Based on Hernandez-Aguilera et al. (2015), we foresee this to have a negative impact in the coffee production setting, both directly and indirectly.

Overall soil health index score was also significantly different between members and non-members of the farmer co-op ($p=0.0276$; Figure 1.5; Table 1.4). This implies that the RCM-provided agricultural services on the average result in a measurable increase in soil health on member farms.

Table 1.4. ANOVA of factors affecting overall Soil Health Index Score.

Factor	Df	F value	Pr(>F)
Co-op Association	1	0.248	0.0276 *
Gender or Grower	1	4.032	0.0462 *
Municipality	5	19.018	4.15e-13 ***
Significance codes: 0 '***' , 0.001 '**', 0.01 '*', 0.05			

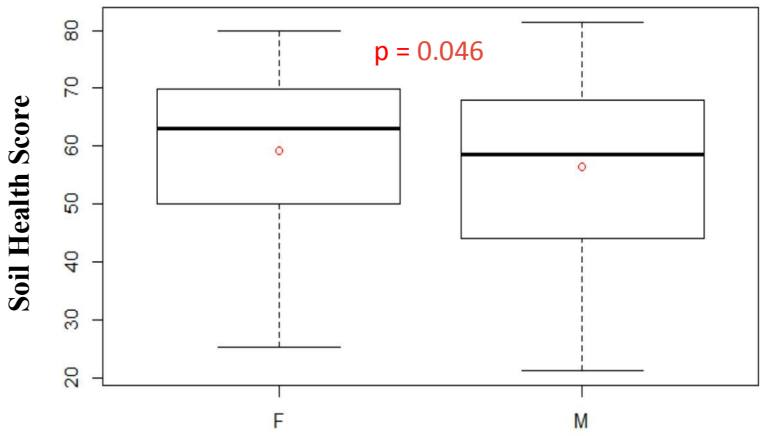


Figure 1.3: Overall Soil Health Index Score for female (F), and male (M) coffee farmers.

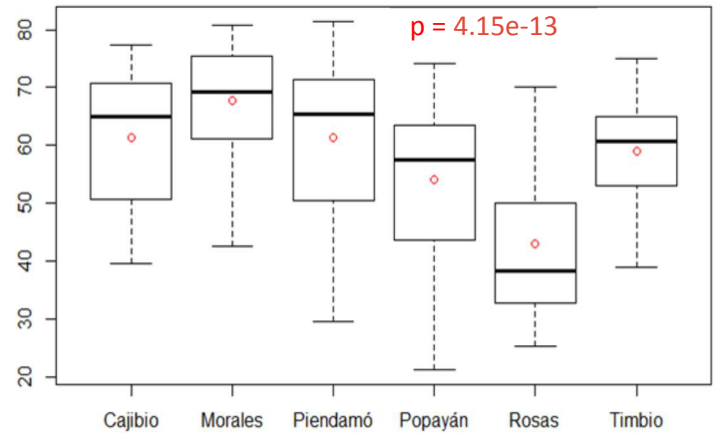


Figure 1.4: Overall Soil Health Index Score by municipality.

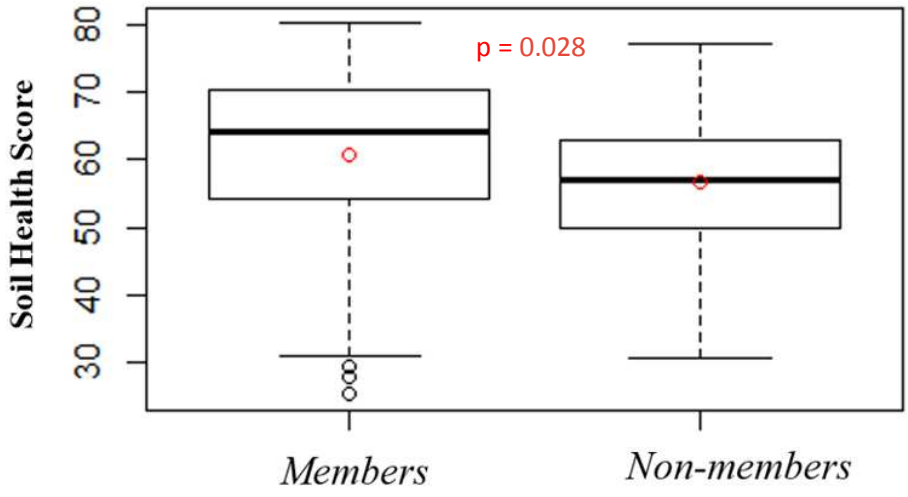


Figure 1.5: Overall Soil Health Index Score for members and non-members of the coffee farmer cooperative.

Table 1.5. Average overall soil health index scores of member-farms by municipality, perceived plot fertility and gender.

Municipality	Soil Health "most fertile"		Soil Health "least fertile"		Average Overall Soil Health Index
	Male	*Female*	Male	Female	
Rosas	44.3	50.5	40.3	38.8	**43.5c†**
Timbío	63.9	60.8	55.4	53.9	**58.5ab**
Piendamó	55.7	71.2	54.1	66.2	**61.8ab**
Cajibio	65.0	61.7	64.1	53.5	**61.1ab**
Morales	71.3	69.5	62.0	71.1	**68.5ba**
Popayán	57.4	64.0	45.7	56.1	**55.8b**
Average	**59.6a**	**63.0a**	**53.6b**	**56.6b**	

† Average male vs. female is significant at α=0.05; a, b, c significant homogeneous groups

1.5.3. Soil Health Indicators Influencing Farmers' Soil Health Perceptions

Given the high variability in the first dimension (34%) in contrast to the other dimensions (17 and 10%) as shown in the Principal Component Analysis (Table 1.3), we focused only on the seven indicators highlighted by the first principal component in the logistic regression analysis.

Results from the Logistic Linear Regression (Table 1.6) further corroborates that, out of the seven soil health indicators, farmers' perception of soil health shows a significantly positive correlation with Protein ($p = 2.48e{-}05$) and OM ($p=0.003$), validating that farmers often perceive high OM as a sign of good soil health (Knutson *et al.,* 2011), and that it is a commonly used indicator of soil fertility when measured data are lacking (Karltun *et al.,* 2013). Farmers' perception of soil health shows significant negative correlation with Respiration ($p=0.02$). It is unclear why this might be, although it may have to do with collinearity as was seen in some studies reported in Hurisso *et al.* (2016). In general, farmers may have an implicit understanding

that OM contributes nutrients and water to coffee trees, that it promotes biological activity and nutrient cycling, and that protein - a nitrogen based compound - boosts coffee yields.

Table 1.6. Soil health parameters that influence farmers' perceptions

	Estimate	**Std. Error**	**z value**	**Pr(>|z|)**
(Intercept)	-0.016464	0.182472	-0.090	0.92811
OM	1.539695	0.522987	2.944	0.00324**
AC	-0.923761	0.524993	-1.760	0.07848.
AWC	-0.001368	0.304239	-0.004	0.99641
RESPIRATION	-0.727225	0.312635	-2.326	0.02001∗
P	-0.167848	0.258682	-0.649	0.51643
WAS	-0.239705	0.233130	-1.028	0.30385
PROTEIN	1.256712	0.298058	4.216	2.48e−05 ∗∗∗

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' '

1.5.4. Farmers' Perception Accuracy

Coffee growers in general tend to have correct perceptions of relative soil health on their farms. Table 1.7 shows that 75% of coffee farmers were correct in ranking their soils. Figure 1.6 also shows that the average Soil Health Index Score is higher for plots that growers identified as the most fertile compared to the least fertile plots on their farms (61.3 and 55.1, respectively, p=0.005). This confirms the finding of (Karltun *et al.,* 2013) who concluded that "there is good agreement between farmers' knowledge (of soil health) and scientific indicators of soil fertility".

Table 1.7. Count for farmers with in/correct perception

	Correct Perception	Incorrect Perception	Total
Number of farmers	56	19	75
Proportion of farmers	75%	25%	100%

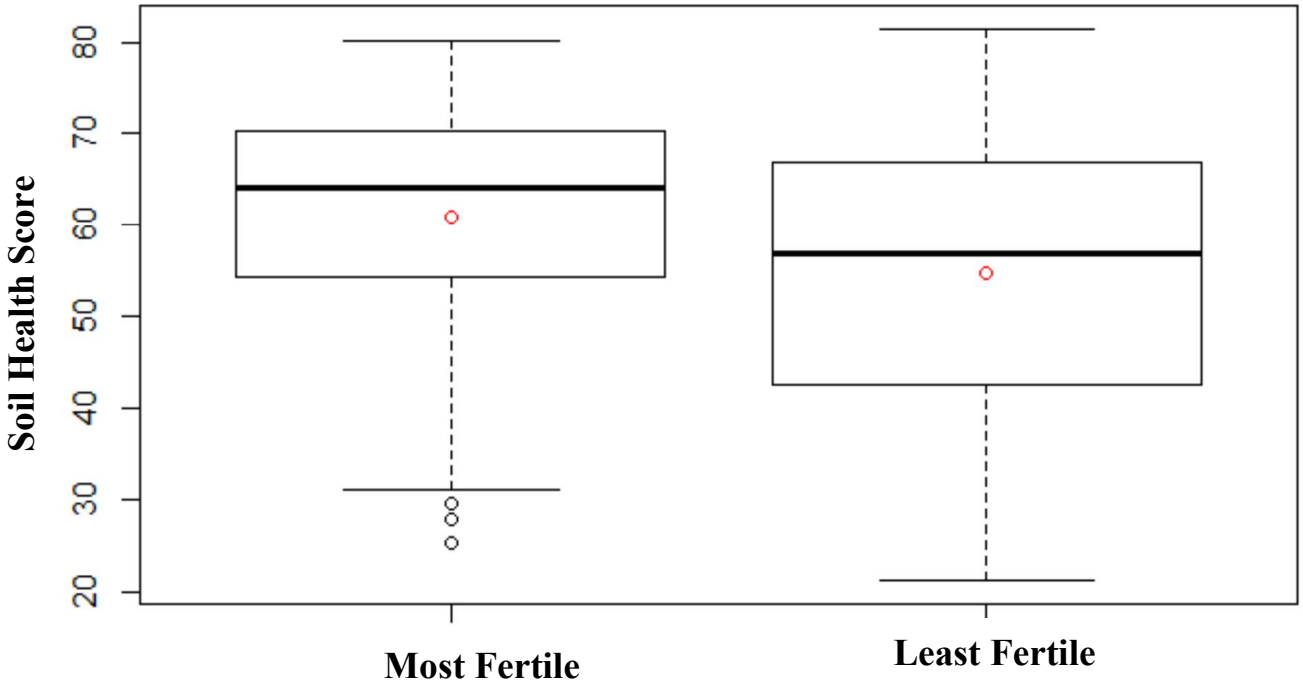


Figure 1.6: Perceived versus actual soil health of Colombian coffee farmers

1.5.5. Farmers' Perception Accuracy based on Gender, Municipality and Actual Soil Health

Given that coffee farmers tended to correctly perceive relative soil health on their farms, interest lies in knowing whether their gender, affiliation to a specific municipality, or their actual soil health played a role in their perceptions. Fisher's Exact Test for Count Data could not prove the presence of statistical difference between male and female farmers in their ability to accurately perceive their soil health meaning that there is no gender effect on a farmer's ability to accurately perceive their soil health ($p \sim 1$). Similarly, Fisher's Exact Test for Count Data could not prove

the existence of statistical difference between farmers from different municipalities (p= 0.78). Interestingly, the relationship between perception accuracy at different soil health levels (Figure 1.7) shows an upward trend suggesting that farmers tend to be more capable of correctly perceiving their soil health when their soils are actually healthier, and that farmers with the healthiest soils have a better understanding of their soil health (p = 0.04).

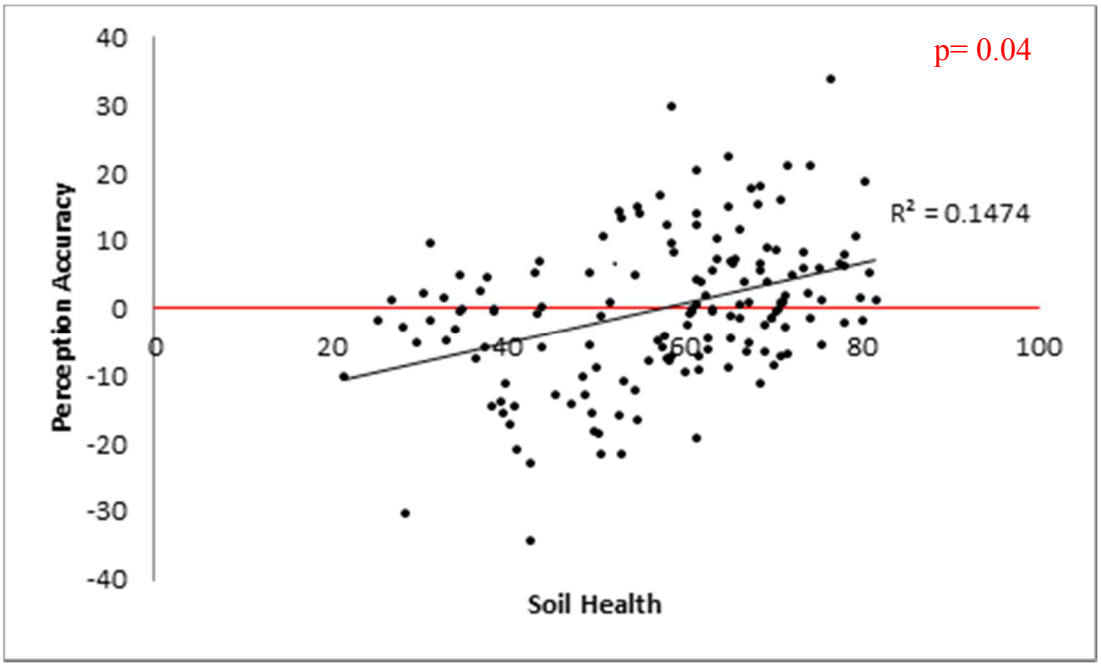


Figure 1.7: Degree of Perception Accuracy at Different SH Levels

1.6. CONCLUSIONS

This study was conducted to evaluate the demographic factors that affect soil health, which soil health indicators influence Colombian coffee farmers' perceptions of soil health, whether farmers correctly ranked their soils based on the actual soil health score, and which factors affect farmers' perception accuracy. Our findings suggest that soil health varies across the 6

municipalities, presumably due to genetic soil differences. Co-op member farms had significantly higher soil health than non-member farms, which suggests that the co-op services have measurable impacts on farm soil health and that the co-op is effectively addressing soil health issues. On average, soil health is higher on female-managed farms than with male farmers, which may be related to overall higher environmental consciousness among female farmers, which was measured in other studies.

Coffee farmers in the region appear to have a reasonably correct perception of their soil health and were most likely successful in accurately ranking soils that were “most fertile” and “least fertile” on their farms. Their correct perception is not associated with which municipality they belong to or what their gender is, but with how healthy their soil actually is.

Organic matter, Protein and Respiration are indicators that are most related to the soil health perception of farmers, as also seen in other studies. Five other variables (AC, AWC, P, WAS, Protein) are also strongly related to soil health.

1.7. ACKNOWLEDGEMENTS

The authors would like to thank the Cornell Statistical Consulting Unit and Ahmad Rafiqi for their insightful statistical and mathematical advice, the personnel at local co-op (FCC) for their contribution to this study and the collaborating farmers.

This study is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE-1650441.

CHAPTER 2. CHARACTERIZATION OF SOIL HEALTH STATUS ON COLOMBIAN COFFEE FARMS

2.1. ABSTRACT

Developing global standards for soil health (SH) is important for overcoming barriers established by differences in assessment protocols. This study was conducted to develop local SH standards specific to Colombian coffee farms, and to compare them to standards in the Northeast USA and Western Kenya. A total of 223 soil samples were collected from coffee farms in Cauca, Colombia. Samples were analyzed for 10 SH indicators including wet aggregate stability (WAS), available water capacity (AWC), respiration rate, pH, and contents of active carbon (AC), organic matter (OM), protein, phosphorus (P), potassium (K) and minor elements (Mg, Mn, Fe and Zn). Samples were scored for SH using four different scoring methods: the Comprehensive Assessment of Soil Health (CASH), the Western Kenyan scoring functions, the Colombian Coffee Soil Health Scoring (CCSHS) method and the Weighted Colombian Coffee Soil Health Scoring (CCSHS-W) method, with the latter two being the research-specific regionally-adjusted scoring functions, developed from a combination of average local conditions and indicator thresholds found in the literature. While the CCSHS computed the overall SH score as the unweighted mean of all the individual indicator scores, the CCSHS-W method assigned weights for each individual indicator as determined by a principal component analysis (PCA) for the SH indicator measurements. We also defined a more simplified version of the CCSHS- through a Best Subsets Regression (BSR) on the measured soil health indicators in addition to % sand, silt and clay, to determine the indicators with the most predictive power of overall soil health. Results from a t-test revealed that the four scoring mechanisms generated significantly different

mean indicator and overall SH scores but that there was a general agreement in scoring the overall SH Index, as well as WAS, AWC, AC, Protein, Respiration, K and Minor Elements as specified by the Intraclass Correlation Coefficient (ICC). Results from BSR analysis revealed that AC was the best single predictor of soil health, and that AC combined with protein, P and pH offered additional predictability, suggesting them for a simplified and less expensive SH assessment.

2.2. KEYWORDS

Soil health, scoring functions, Colombia, coffee, global soil health standards.

2.3.INTRODUCTION

Soil health (SH) is critical to sustainable agricultural production. The ability to quantitatively assess SH is of increasing importance. Proper interpretation of SH measurements requires benchmarks to assess where a sample lies on the SH spectrum (Arshad and Martin, 2002). The Comprehensive Assessment of Soil Health (CASH) approach developed at Cornell University measures biological, chemical and physical soil properties that are key indicators of SH (Moebius-Clune *et al.,* 2016). It converts raw laboratory and field measurements into generally recognized and easily interpretable scores that aid in management decisions. These scores are derived from scoring functions that were developed following the approach of the Soil Management Assessment Framework by Andrews et al. (2004) that uses logic statements to assess the relationship between a set of empirical values and indicator measurements of soils from Georgia, Iowa, California, and the Pacific Northwest, thereby assigning a normalized score. Thus, the scoring functions developed in the CASH approach consisted of comparing individual

measurements to a standardized dataset from the Northeastern (NE) United States- a region that is characterized by a temperate climate with diverse production systems including grain, livestock, vineyards and vegetable production. Scoring functions were used to establish scoring curves for each individual indicator which come in one of three forms ("more is better", "optimum range", and "less is better"), and are sometimes adjusted for soil texture where it affects the dynamic properties of the soil.

Scoring functions for the physical and biological indicators follow cumulative normal distribution (CND) curves specific to each indicator. Others are based on thresholds determined in the literature which are outcome-based in terms of crop response to different levels of an indicator, as in the case of P, K, pH, and minor elements (Gugino *et al.,* 2009; Moebius-Clune *et al.,* 2016).

All scoring functions are scaled to values between 0 and 100. Thus, indicator scores fall in one of three ranges: "high" (between 70 and 100), "medium" (between 30 and 70) and "low" (between 0 and 30; Gugino *et al.,* 2009). From the individual indicator scores, an overall SH score is calculated as their unweighted arithmetic mean and is interpreted as "very low" if SH scores below 40%, "low" if SH scores between 40- and 55%, "medium" if SH scores between 55- and 70%, "high" if SH scores between 70- and 85%, and "very high" if SH scores higher than 85% (Gugino *et al.,* 2009; Moebius-Clune *et al.,* 2016).

Regional, climatic and soil characteristic differences have a significant impact on the standardization of SH (Congreves et al. 2015). Similar to work done by Moebius-Clune (2010) who developed scoring functions for the assessment of soil health in western Kenya from a chronosequence experiment on smallholder farms, it is important to make further advancements

in “testing the Test” in other ecosystems, and to assess differences between scoring methods which will ultimately help verify whether a widely-standardized SH assessment protocol is feasible. Because the scoring functions used in CASH represent data collected from the Northeast USA, their use in the assessment of SH status in Colombian coffee smallholder farms may not be applicable (Moebius-Clune, B. N., 2010; Schindelbeck *et al.,* 2008; Moebius-Clune, 2010; Congreves et al. 2015; Idowu *et al.,* 2008).

The objectives of this study were to (i) develop a set of adjusted scoring functions for Cauca, Colombia coffee farms and (ii) assess the differences and similarities between other regional scoring functions (Northeast USA and Western Kenya).

2.4. MATERIALS AND METHODS

2.4.1. Site Description

The study was performed in a coffee-growing region within the Department of Cauca, Colombia, situated at approximately 2.2°N and −76.4°W (Figure 2.1), with farm fields on elevations ranging between 1269 and 1959m. Rainfall in Cauca ranges between 313.2- and 260.9- mm y^{-1} and has a bimodal distribution centered around the months of April and November (computed from: Promedios Climatológicos 1981 - 2010.xlsx)[2] .The soils in the project area are Andisols, of volcanic ash origin (Universidad del Cauca, unpublished data).

[2] http://www.ideam.gov.co/web/tiempo-y-clima/clima. Accessed: 2/21/2016

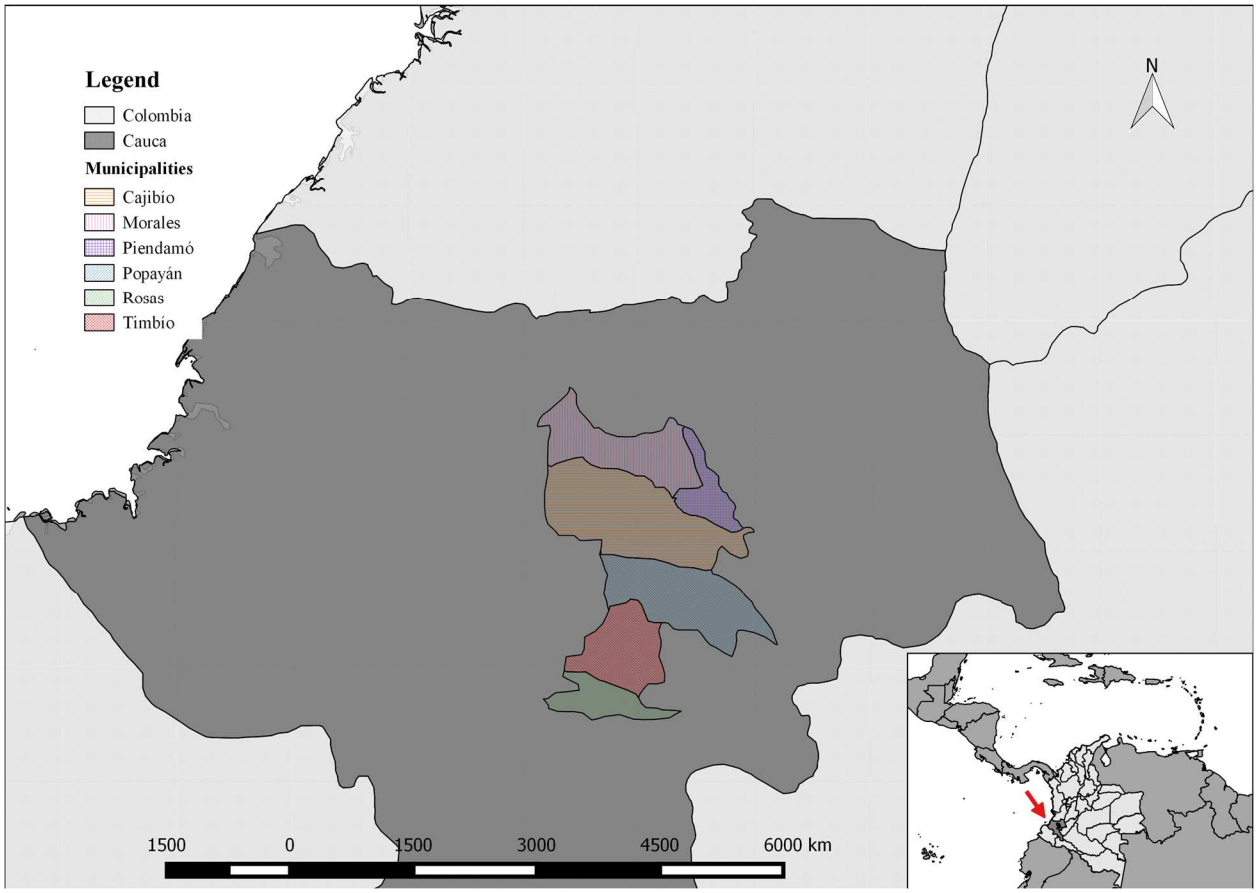


Figure 2.1: **Study Area in Cauca Department, Colombia and six selected municipalities**

Coffee production in the region is mainly conducted by small-scale subsistence or less-resourced farmers in either monoculture or polyculture. Crops that accompany coffee trees in polyculture settings are typically grown for market or domestic purposes and mainly include maize (*Zea -mays*). A variety of shade tree species are grown which provide canopy cover for the coffee and perform other ecosystem services (Hernandez-Aguilera *et al.,* 2015).

Coffee growers in Cauca, Colombia are largely split into two marketing groups: Those who are members of a Relationship Coffee Model (RCM) co-op - Federación Campesina del Cauca (FCC) headquartered in Popayan- and those who sell to the regular commodity market. As a result of their RCM membership, participating growers have access to coffee quality price-premiums, certification, loans, immunity to coffee price volatility, agricultural and commercial services, and access to the international market (Hernandez-Aguilera *et al.,* 2015).

2.4.2. Soil Sampling

A total of 223 soil samples were collected in January 2014 following the CASH Protocol (Gugino *et al.,* 2009) from 145 coffee farms across 6 municipalities (Cajibio, Timbío, Rosas, Piendamò, Morales and Popayàn) within Cauca. All samples were collected from the 0- to15-cm depth range using a Dutch-style soil auger after surface residue removal. At each sampling location, five samples were collected and then composited to obtain a representative sample. Two representative soil samples were collected from 78 farms: One from the farm's most fertile plot and the other from the least fertile, as indicated by the farmer. One representative sample was collected from the remaining 67 farms. Collected soil samples were air-dried and passed through a 2-mm sieve (Gugino *et al.,* 2009).

2.4.3. Soil Health Measurements

Soil samples were sent to Cornell University in Ithaca, NY (USA) where processing and analysis of physical, chemical and biological soil health indicators were performed according to the Comprehensive Assessment of Soil Health (CASH) protocol (Moebius-Clune et al., 2016). Briefly, this includes:

Physical Indicators: Available Water Capacity (AWC) between field capacity (-10kPa) and permanent wilting point (-1500kPa) was assessed gravimetrically by equilibrating saturated soil to each of 10 kPa and 1500 kPa on ceramic high pressure plates (Topp and Zebchuck 1979). The difference between soil water loss under 10 kPa and 1500 kPa pressures determined from calculating the difference in wet and dry weights was considered the AWC (Moebius-Clune *et al.,* 2016).

Wet Aggregate Stability (WAS) was assessed using a rainfall simulator adapted from Ogden *et al*. (1997) that allows particles of air-dried soil placed on a 0.25 mm mesh sieve to slake under 2.5 J of rainfall energy for 300 seconds, based on a total of 2.5 cm of rainfall. Wet Aggregate Stability was determined by subtracting the weight of slaked soil plus the remaining stones on the sieve (>0.25 mm) from total soil weight measured before rainfall (Moebius-Clune *et al.,* 2016).

Soil texture was determined using a rapid quantitative method developed by Kettler *et al.* (2001) where soil samples were fractionated with 3% sodium hexametaphosphate ($(NaPO_3)_n$) and a series of sieving and sedimentation steps was used to separate the different particle sizes.

Biological Indicators: Organic Matter content (OM) was analyzed by mass loss on ignition in a muffle furnace at 500 °C for two hours, with values corrected by multiplying percent loss on ignition by 0.7 and subtracting 0.23 (Moebius-Clune *et al.,* 2016). Active Carbon (AC) was measured by adding a dilute potassium permanganate solution ($KMnO_4$) to soil, which acts as an oxidant to AC, and measuring the solution's absorbance at 550 nm using a hand-held colorimeter (Hach, Loveland, CO) (Weil *et al.,* 2003).

Autoclaved Citrate Extractable Soil Protein Index (Protein) was measured by extracting proteins from the soil following a series of centrifugation and autoclaving steps using 0.02 M sodium citrate at pH 7. Soil protein concentration was determined by measuring bicinchoninic acid assay against bovine serum albumin standard curve for soil protein concentration (Walker, 2009; Wright and Upadhyaya, 1996). The soil Respiration test was performed by trapping and measuring CO_2 emitted by soil microorganisms over a 4-day room temperature incubation in a sealed chamber with a KOH trap (Haney and Haney, 2010). The dissolved CO_2 in the KOH trap quantifying microbial activity was measured using an electrical conductivity meter, and the change in conductivity before and after incubation quantifies the amount of CO_2 evolved.

Chemical Indicators: Soil pH was measured in a 1:1 water dispersed slurry determined by a pH electrode probe (SM802 Smart Combined Meter, Milwaukee Industries, Rocky Mount, NC). Soil nutrients, including P, K, Mg, Fe, Mn and Zn were extracted with a Modified Morgan solution (ammonium acetate - buffered at pH 4.8), and quantified by Inductively Coupled Plasma Optical Emission Spectroscopy (ICP–OES, Varian 730-ES, Mulgrave, Victoria, Australia).

2.4.4. Soil Health Scoring

2.4.4.1.Comprehensive Assessment of Soil Health (CASH)

Measured results for each SH indicator were translated into indicator scores on a scale of 0-to-100 following the CASH protocol. It relies on three scoring types: “Less is better” (, “Optimum Range”, and “More is better”; Moebius-Clune *et al.,* 2016; Gugino *et al.,* 2009). The scoring functions used are specific to the NE USA and are texture adjusted for WAS, AWC, AC, OM, Protein and Respiration. For indicators with unestablished absolute thresholds, scoring functions were developed using a Gaussian distribution function, as suggested by Arshad and Martin (2002):

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{(x-\mu)^2}{2\sigma^2}}, -\infty < x < \infty \qquad [3]$$

where μ was estimated by the sample mean (m) and σ by the sample standard deviation(s). The associated Cumulative Normal Distribution function, CND (m, s), gives the probability (between 0 and 1) that a member of the distribution is less or equal to the SH indicator measurement, x (Eq. 4). It then serves as the SH scoring function, after multiplying by 100:

$$CND\ (x, m, s) = \frac{1}{2}\left(1 + erf\left[\frac{(x-m)}{s\sqrt{2}}\right]\right) \times 100, -\infty < x < \infty \qquad [4]$$

where *erf* denotes the error function.

For other SH indicators (P, K, Mg, Zn, Fe, Mn, pH) scoring functions were based on existing critical values from the literature (Gugino *et al.,* 2009). Micronutrients including Mn, Zn, Fe and Mn are scored on a deficiency/ sufficiency level as given by the respective scores 0 or 100, and a Minor Elements Score is generated as the average of the four scores i.e. 100% if all 4 micronutrient levels are sufficient, 75% if three micronutrient levels are

sufficient, 50% if two micronutrients are sufficient, 25% if only one micronutrient level is sufficient, and 0% if none of the micronutrient levels are sufficient.

Overall SH was calculated as the unweighted arithmetic mean of the individual indicator scores including the minor elements score.

2.4.4.2.Colombian Coffee Soil Health Scoring (CCSHS)

The Colombian Coffee Soil Health Scoring (CCSHS) approach was developed from our dataset (n=223) to establish region-specific SH standards for locally appropriate interpretation. The approach used in developing the CCSHS was similar to the one used in CASH in terms of the use of both linear and non-linear functions to interpret the measured indicator values (Gugino *et al.,* 2009). The scoring range used in CASH was also implemented in the CCSHS (0-100), and similarly relied on the three scoring types: "Less is better" (Mn), "Optimum Range" (pH, Zn, Fe), and "More is better" (WAS, AWC, AC, OM, Respiration, Protein, P, K, Mg; Moebius-Clune *et al.,* 2016; Gugino *et al.,* 2009). Some indicators required accounting for textural groupings (fine, medium; Dexter, 2004; Moebius *et al.,* 2007), i.e., those that showed significantly different mean measured values among textural groups (AWC, OM, AC and Respiration; Table 2.1). Specific scoring functions for the coarse textural class were not possible due to the absence of data.

Table 2.1. Descriptive statistics of the measured soil health indicators (n=223)

Indicator	Weights	Min	Max	Median	Mean	SD	Fine textured Mean Values	Medium Textured Mean Values	p-value
Sand %	0	4.1	58.0	15.2	16.5	8.4			
Silt %	1.051	31.8	79.6	60.5	58.5	9.5			
Clay%	0.677	4.0	60.0	24.0	25.0	7.9			
WAS (%)	0.976	34.1	100.0	96.84	94.3	8.3	94.42[a]	94.19[a]	0.84
AWC (m^3/m^3)	0.946	0.1	0.6	0.3	0.3	0.1	0.20[a]	0.30[b]	1.2e-14
OM (%)	0.970	4.7	29.8	17.4	17.4	6.0	14.79[a]	18.92[b]	7.9e-7
AC (ppm)	1.103	302.5	1256.0	877.2	818.3	237.7	688.6[a]	893.3[b]	1.4e-09
Protein (mg/g soil)	1.413	2.5	16.8	9.3	9.2	2.5	9.65[a]	9.05[a]	0.09
Respiration (mg/g soil)	1.347	0.3	1.7	1.0	1.0	0.2	0.95[a]	1.03[b]	0.03
pH	0.561	3.8	6.2	4.8	4.8	0.3			
P (ppm)	1.320	1.5	25.2	9.4	9.6	4.3			
K (ppm)	0.801	22.7	550.5	90.9	115.4	80.9			
Mg (ppm)	0.493	4.6	3172.6	52.3	169.9	361.6			
Fe (ppm)	1.093	4.1	93.8	17.7	20.6	11.2			
Mn (ppm)	1.029	0.7	26.9	5.1	6.4	4.7			
Zn (ppm)	0.964	0.1	6.9	0.6	0.9	0.9			
Total:	14.744								

Where thresholds, ranges or critical values for SH indicators were not established (AWC, OM, AC, Protein, Respiration and P), scoring was based on CND functions (Eq. 4) with parameters estimated from local conditions, as implemented in the CASH approach (Table 2.2). Distinctively, because WAS is a percent-based measurement (0 -100), the measured values were directly used as scores.

Where thresholds, ranges or critical values for SH indicators did exist (K, Mg, Zn, Mn, and Fe), scoring was based on normal distribution functions with set baselines found in literature specific to coffee soils (Table 2.2). Individual indicator scores were averaged to calculate the overall SH score (Gugino *et al.,* 2009). Although critical P values were available in the CASH approach for the NE USA, P in the CCSHS was approached using distribution-based scoring (Eq. 4), because there is a lack of information that agrees on the optimal soil P levels for coffee production using Modified Morgan extraction, and there are inconsistent optimal P levels reported for other extractions. Furthermore, the volcanic nature of these soils causes P to be one of the most recognized limiting nutrient for coffee production due to fixation (Melke and Ittana, 2015).

Critical values and optimal ranges for K, Mg, Zn, Fe and Mn related to coffee production have been established in the literature, and were used to define the scoring curves for these indicators, again using a CND (m,s) function (Eq. 4). Since these critical values fell within the normal distribution range of measured values for each indicator (Figure 2.2), log transformations were not necessary for Mg, Mn, Zn, Fe, and K. We relied on graphical interpretation of the population distribution where a bell-shaped curve (or an approximation

thereof) was assumed normal, since the Shapiro-Wilk test is considered overly sensitive to minor departures from normality- particularly in large sample sizes (Ahad *et al.,* 2011).

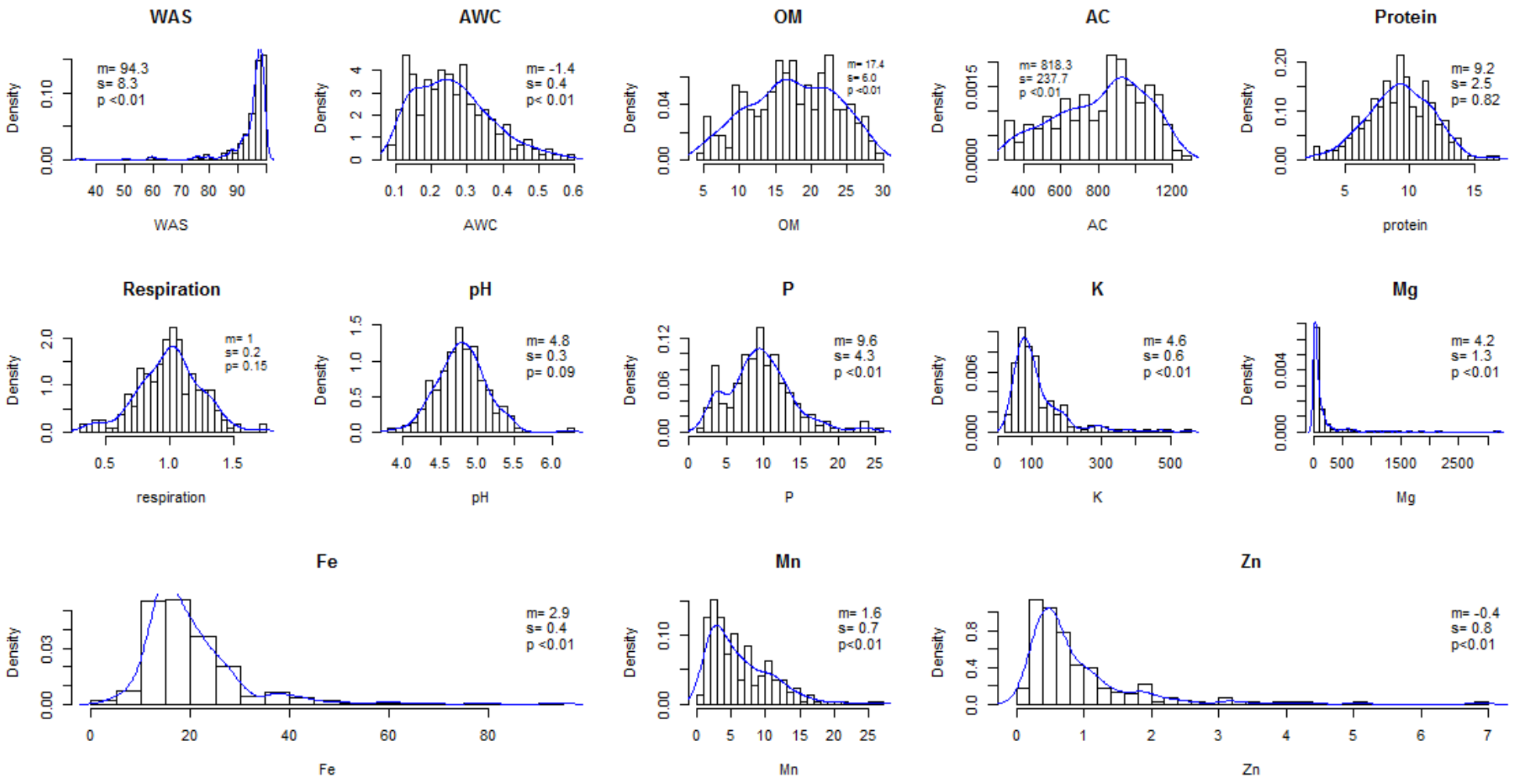


Figure 2.2: Frequency distributions for the soil health indicators; p-values based on the Shapiro-Wilk test for normality.

We scored the chemical indicators by assigning m to be the midpoint between boundary thresholds for each portion of the scoring curve, and s the midpoint between the mean and each threshold. This is equivalent to m being a score of 50% and s a score of 25% in our rating system. Scoring curves for the chemical indicators came in one of three forms: "More is better", "Optimum Range" and "Less is better" (Table 2.2).

The K scoring curve was based on a threshold specific to coffee soils (National Coffee Federation of Colombia, FNCC), where soil K is deemed deficient at levels below 156 ppm. Therefore, we assigned K values of above 160 ppm scores approaching 100%, to result in a scoring curve of CND (80,40); Table 2.2; Figure 2.3). The Mg scoring curve was also based upon a FNCC threshold, and is deemed deficient at levels below 108 ppm. Therefore, we assigned Mg values of above 110 scores approaching 100%, to result in a scoring curve of CND (55,27.5); Table 2.2; Figure 2.3).

The Mn scoring curve was approximated based on literature thresholds specific to coffee soils discussed in terms of DTPA-extracted Mn by Melke and Ittana (2015), which was converted to its equivalent modified Morgan levels using a regression equation by Kreij *et al.* (1993). Modified Morgan extracted Mn is considered optimal at levels below eight ppm (Melke and Ittana, 2015), in which case a 1-CND (m, s) function ("less is better") is appropriate. Mn values below eight ppm were assigned scores given by the function 1-CND (4,2) (Table 2.2; Figure 2.3).

and function and sources of critical values and conversion equations for each indicator

Type of Function	Sources of critical values	Sources of extractant conversion factors †	Scoring Function (0-100) ‡
More is better	laboratory measurements	na	scores are measurement values
More is better	average local conditions	na	CND(0.30, 0.10)*100§ CND(0.20, 0.08)*100
More is better	average local conditions	na	CND(893.26, 203.63)*100 CND(688.63, 237.99)*100
More is better	average local conditions	na	CND(18.92, 5.59)*100 CND(14.79, 5.84)*100
More is better	average local conditions	na	CND(1.03, 0.24)*100 CND(0.95, 0.25) *100
More is better	average local conditions	na	CND(9.24, 2.5)*100
More is better	average local conditions	na	CND(9.62, 4.26)*100
More is better	FNCC ¶	na	CND(80, 40)*100
More is better	FNCC	na	CND(55, 27.5)*100
Optimum range	Melke and Ittana (2015); Alloway (1995)	Rodriguez-Suarez *et al.* (2007)	CND(0.05, 0.025)*100 [1-CND(4.3, 2.15)]*100
Optimum range	Melke and Ittana (2015); Abrey *et al.* (2005)	Al-Mustafa *et al.* (2001)	CND(6.5, 3.25)*100 [1-CND(146.5, 73.25)]*100
Optimum range	Winston *et al.* (2005); Bitterbender; Malavolta and Netto (1989); Kuit *et al.* (2004); South Africa Department of Agriculture, Forestry and Fisheries (2012)	na	CND(5, 2.5) *100 [1-CND(7.1, 3.55)]*100
ess is better	Melke and Ittana (2015)	Kreij *et al.* (1993)	[1-CND(4, 2)]*100

tical values found in literature given by extractants other than modified Morgan
ere scored separately by soil texture, the function for medium-textured soils is listed first, followed by the function

tive normal distribution, where m = mean, s = standard dev.

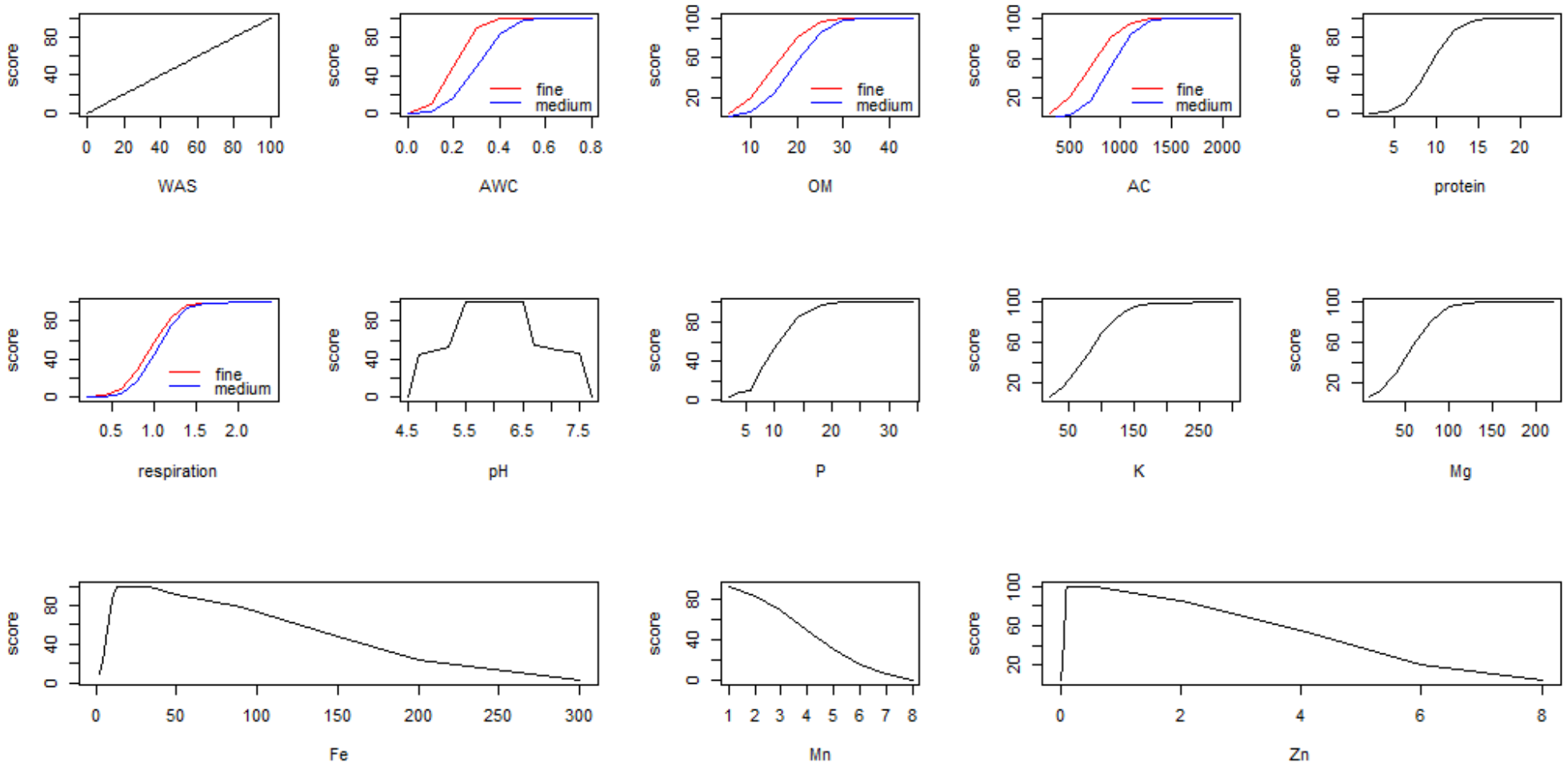


Figure 2.3. Scoring curves developed using the Colombian Coffee Soil Health Scoring method

The Zn values were interpreted using an "optimum range" scoring curve based on literature thresholds specific to coffee soils discussed in terms of DTPA-extracted Zn by Melke and Ittana (2015) and Alloway (1995), which were converted to their equivalent modified Morgan values using regression equations by Rodriguez-Suarez *et al.* (2007). Modified Morgan extractable Zn is considered optimal at levels in the range 0.1- 0.6 ppm, where values were assigned a score of 100%, and phyto-toxic in general at levels above 8 ppm at which values were assigned a score of 0% (Melke and Ittana, 2015; Alloway, 1995). We further assigned Zn values below the lower threshold (0.1 ppm) scores given by the function CND (0.05,0.025), and values above the upper threshold (0.6) scores given by the function 1- CND (4.3,2.15; Table 2.2; Figure 2.3).

The Fe scoring curve, also "optimum range" type, was approximated based on literature thresholds specific to coffee soils discussed in terms of DTPA-extracted Fe by Melke and Ittana (2015) and Abrey *et al.* (2005), which were converted to their equivalent modified Morgan values using regression equations by Al-Mustafa *et al.* (2001). Hence, modified Morgan extractable Fe is considered optimal at levels between 13 and 33 ppm (Melke and Ittana, 2015), where values within this range were assigned a score of 100%, and phyto-toxic in general at levels above 260 ppm (Al-Mustafa *et al.,* 2001) where values were scored 0%. We further assigned Fe values below 13 ppm scores given by the function CND (6.5,3.25), and values above 33 ppm scores given by the function 1- CND (146.5,73.25) (Table 2.2; Figure 2.3).

The pH scoring curve was based on literature thresholds specific to coffee soils discussed by Winston *et al.* (2005), Bitterbender; Malavolta and Netto (1989), Kuit *et al.* (2004), and

in a guidebook by South Africa's Department of Agriculture, Forestry and Fisheries (2012).

Soil pH ranging between 5.5 and 6.5, is considered the optimal range for coffee production, and where values were assigned a score of 100%. Based on the same literature, samples with a pH $\leq$4.5 and $\geq$ 7.7 were deemed suboptimal and were scored 0%. We further assigned pH values in the range 4.5<pH< 5.5 scores given by the function CND (5,2.5), and values in the range 6.5<pH<7.7 scores given by the function 1-CND (7.1,3.55) (Table 2.2; Figure 2.3).

The Minor Elements indicator score was calculated as the unweighted arithmetic mean of Zn, Mn, Fe and Mg scores, and the Overall SH score was computed as the arithmetic mean of all individual SH indicators.

2.4.4.3. Weighted Colombian Coffee Soil Health Scoring (CCSHS-W)

Since different indicators may not have equal impact on overall soil health, the new SH scoring framework for Colombian samples was adapted from a study by Congreves *et al.* (2015), and consisted of scoring the individual soil indicators according to the CCSHS protocol, but with an overall SH score computed as a weighted mean of the individual soil indicator scores plus soil texture (% sand, silt and clay). The weights were determined from a Principal Component Analysis (PCA) of the SH indicators, where weights were assigned as the sum of the eigenvectors of the first four principal components (Table 2.1), based on Kaiser's cut-off principle (eigenvalues > 1.00; Kaiser, 1960). Negative eigenvalues were

assumed as 0. Hence, the new overall SH score for each sample was computed using the following equation

$$Weighted\ Overall\ SH\ Index = \frac{(s_1 \times w_1) + (s_2 \times w_2) + (s_3 \times w_3) \ldots (s_n \times w_n)}{w_1 + w_2 + w_3 \ldots + w_n}, \qquad [5]$$

where s represents the adjusted SH indicator score as developed in the CCSHS method and % sand, silt and clay; and w the weighting factors determined by the PCA analysis (Table 2.1). Scores that exceeded 100 were reset to 100.

2.4.5. Regional Soil Health Scoring Comparisons

In an attempt to assess differences between regional SH scoring approaches that would enable proper soil monitoring and help overcome differences in the SH interpretive framework (Arshad and Martin,2002; Barrios *et al.,* 2006), a comparison of regional scoring functions was performed between those developed for the Northeast USA; the Colombia coffee system; and the Western Kenya smallholder system by Moebius-Clune (2010) which was developed from a chronosequence experiment on smallholder farms at the Kakamega and Nandi Forest margins in Kenya that were converted from forest between 1930 and 2000 (Table 2.3).

ons of regional scoring functions developed from average local conditions in the Northeast (2,200 from the ssment of Soil Health (CASH)), Cauca, Colombia (n=223) from the Colombian Coffee Soil Health Scoring rn Kenya (n=227) from the scoring method in Moebius-Clune (2010) and categorized by textual grouping.

Fine			Medium			Coarse
auca, ombia	**Western, Kenya‡**	**Northeast, USA**	**Cauca, Colombia**	**Western Kenya**	**Northeast, USA**	**Northeast, USA**
= 94.3 = 8.3	CND(58, 19)	CND(43.3,17.1)	m= 94.3 s=8.3	CND(49, 19)	CND(41.2, 24.6)	CND(58.5, 25.3)†
D(0.20, 0.10)	CND(0.13,0.04)	CND(0.18, 0.08)	CND(0.30, 0.10)	CND(0.16,0.04)	CND(0.16, 0.05	CND(0.13,0.07)
D(14.8, 5.8)	CND(5.9, 2.9)	CND(4.1, 1.6)	CND(18.9, 5.6)	CND(5.8, 2.9)	CND(3.5, 1.3)	CND(3.1, 1.6)
D(689, 38)	CND(333, 214)	CND(616, 192)	CND(893, 238)	CND(427, 214)	CND(561, 180)	CND(494, 219)
D(0.95, 0.25)	**na§**	CND(0.53,0.32)	CND(1.03, 0.24)	na	CND(0.60, 0.32)	CND(0.64,0.42)
(9.2, 2.5)	na	CND(4.9, 1.3)	CND(9.2, 2.5)	na	CND(6.7, 3.3)	CND(8.9, 4.7)

lative normal distribution, where m = mean, s = standard dev.
d by Moebius-Clune (2009) for Western Kenyan soils.

Measured SH indicators for n=223 were converted into their respective scores using each of the three scoring functions and were subject to comparative analysis using a t-test and an inter-rater reliability test where the Intraclass correlation coefficient (ICC) represents the level of agreement between the different scoring systems. Negative ICC's suggesting no agreement at all were capped at 0.

2.4.6. Simplified Assessment of Soil Health for Colombian Coffee Systems

The CCSHS uses the scores of 13 physical, biological and chemical indicators to compute an overall SH score using an unweighted arithmetic mean. A Best Subsets Regression (BSR) was used to evaluate whether overall SH could be predicted with a lower number of indicators, and which indicator(s) are most predictive of overall SH that could be used in a low-cost test.

2.4.7. Statistical Analyses

All statistical analyses including PCA and BSR were performed using the R-Project for Statistical Computing (R Core Team, 2013). Scoring of indicators using the CASH, CCSHS, CCSHS-W and Western Kenyan functions were done on Excel (Microsoft Office, 2013) using the NORMDIST function.

2.5.RESULTS AND DISCUSSIONS

2.5.1. Summary of Soil Health Results

Table 2.1 shows a summary of the indicator measurements for the 223 soil samples. Physical indicators WAS and AWC measured between 34-100% and 0.1- 0.6 g g^{-1}, with 94(8)% and 0.3(0.1) g g^{-1} average and standard deviation, respectively. Biological indicators including

Organic matter, AC, protein and respiration measurements ranged between 4.7- 29.8%, 302.5-1256 ppm, 2.5-16.8 mg g^{-1} soil and 0.3-1.7 mg g^{-1} with averages and standard deviations of 17.4(6), 818.3(237.7), 9.2(2.5) and 1.0(0.2) respectively. Chemical indicators including pH, P and K measurements ranged between 3.8- 6.2, 1.5- 25.2 ppm and 22.7- 550.5 ppm, with averages and standard deviations 4.8(0.3), 9.6(4.3) and 115.4(80.9) respectively. Minor elements including Mg, Fe, Mn and Zn ranged in measurements between 4.6- 3172.6 ppm, 4.1- 93.8 ppm, 0.7-26.9 ppm, and 0.1-6.9 ppm, with averages and standard deviations 169.9(361.6), 20.6(11.2), 6.4(4.7), and 0.9(0.9) respectively.

2.5.2. Comparison of Scores

Table 2.4 shows average scores (scale 0-to-100) for each SH indicator and overall SH index of the 223 soil samples, and the ICC of each set of comparisons.

Wet aggregate stability scores averaged 99, 94 and 96% using Northeast USA, Colombian Coffee System and Kenyan Smallholder functions respectively, while AWC scores averaged 77, 46 and 89 %. As indicated by the t-test, average scores generated for the Northeast USA for WAS and AWC were significantly different from scores generated by Colombian Coffee System and Kenyan Smallholder functions. There was also significant difference between Colombian Coffee System and Kenyan Smallholder functions average scores in AWC but none in WAS. Although results from the t-test corroborate that the three scoring mechanisms yield significantly different average scores, the ICCs indicate that there is strong agreement in scoring WAS and fair agreement in scoring AWC (ICC=0.711; ICC=0.314 respectively), corroborating that the three regions have similar standards for WAS and less similar standards for AWC, and that the differences identified by the t-test for WAS are minimal in terms of rating given that the

indicator was consistently scored high by the three scoring methods, whereas AWC presented some variation in scoring ("High" by Northeast USA and Kenyan smallholder functions versus "Medium" by Colombian Coffee System functions; Table 2.4).

Biological scores averaged by Northeast USA, Colombian Coffee System and Kenyan Smallholder functions show stinking differences: 100, 50 and 96 % for OM, and 69, 50, and 93 % for AC, respectively. Based on the t-tests and ICCs, there was no agreement whatsoever between the three scoring methods in scoring OM (ICC= 0), and fair agreement in scoring AC (ICC= 0.298). This scoring discrepancy between the three scoring methods can be conceptualized by the fact that coffee systems, in contrast to grain and vegetable systems which are characteristic of the Northeast USA temperate zones, are typically grown in tropical agro-climatic zones that induce OM accumulation from the forested surroundings and the no-till practices (mean values for OM in Colombian Coffee systems= 17.4%; Table 2.4), thereby increasing AC content (818 ppm; Table 2.4). Because of this, OM was scored remarkably high on average (100%; Table 2.4) by the Northeast USA scoring function due to its lower standards for OM content, followed by Kenyan Smallholder scoring function (96%; Table 2.4) which reflects recently deforested agricultural land that is depleted of OM. The Colombian Coffee scoring function assigned an average score of 50% for OM since it was based on average local conditions thereby reflecting the local norms. This however did not translate similarly for AC. Although the Northeast USA scoring function did score AC higher on average than Colombian coffee scoring, the difference was not as grand as the difference seen with average OM scores (69% vs. 50%; Table 2.4). In agreement with the OM scoring, the Kenyan smallholder function did on average score AC accordingly very high (93%).

Table 2.4: Mean comparisons between indicator scores generated by the Comprehensive Assessment of Soil Health (Northeast USA), the Adjusted Colombian Soil Health Scoring Method (Colom.Coffee), and Kenyan Smallholder scoring functions. Mean comparison between overall soil health scores generated by Northeast USA, Colombian Coffee System and the Weighted Colombian Soil Health Scoring Method (Colomb. Coffee Weighted) for n=223.

Indicator	Average Values	Average Scores		Intraclass Correlation Coefficient †
WAS	94.3 %	Northeast USA	99(5) a‡	**0.711** (p = 8.5e-80)
		Colom. Coffee	94(8) b	
		Kenyan Small.	96(9) b	
AWC	0.3 m/m	Northeast USA	77(29) b	**0.314** (p = 7.01e-15)
		Colom. Coffee	47(29) c	
		Kenyan Small.	89(20) a	
OM	17.4 %	Northeast USA	100(2) a	**0** (p=1)
		Colom. Coffee	50(29) c	
		Kenyan Small.	96(12) b	
AC	818.3 ppm	Northeast USA	69(34) b	**0.298** (p = 1.33e-13)
		Colom. Coffee	50(27) c	
		Kenyan Small.	93(13) a	
Protein	9.2 (mg/g soil)	Northeast USA	60(21) a	**0.869** (p = 3.23e-70)
		Colom. Coffee	50(29) b	
Respiration	1.0 (mg/g soil)	CASH	49(18) a	**0.887** (p = 4.66e-77)
		Colom. Coffee	50(28) a	
pH	4.8	Northeast USA	0(6) b	**0** (p=1)
		Colom. Coffee	40(20) a	
P	9.6 ppm	Northeast USA	96(14) a	**0** (p=1)
		Colom. Coffee	49(28) b	
K	115.4 ppm	Northeast USA	90(21) a	**0.239** (p = 0.000154)
		Colom. Coffee	62(29)b	
Mg	169.9 ppm	Northeast USA	54(39)	
Fe	20.6 ppm	Northeast USA	97(10)	
Mn	6.4 ppm	Northeast USA	93(14)	
Zn	0.9 ppm	Northeast USA	94(10)	
Minor Elements Score		Northeast USA	74(29)b	**0.22** (p = 0.000453)
		Colom. Coffee	84(8)a	
Overall Soil Health Score		Northeast USA	71(10)a	**0.636** (p = 2.19e-60)
		Colom. Coffee	64(11)b	
		Colom. Coffee-Weighted	60(10)c	

†If: 0<ICC≤0.2: poor agreement; 0.3≤ICC≤0.4: fair agreement; 0.5≤ICC≤0.6: moderate agreement; 0.7≤ICC≤0.8 strong agreement; ICC >0.8: almost perfect agreement.
‡ a, b, c indicates statistically significant difference at α=0.05 between different scoring Methods. Bonferroni correction (α=0.05/3) was performed for comparisons of WAS, AWC, OM, AC and Overall Soil Health Score.

Average protein scores generated by Northeast USA and Colombian Coffee System functions were 60 and 50%, and average respiration scores were 49 and 50% respectively. Average scores generated by the Northeast USA and Colombian Coffee System functions were significantly different from one another for protein but not for respiration as indicated by the t-test, but there was an almost perfect agreement between the different methods in scoring both protein and respiration (ICC=0.869; 0.887 respectively; Table 2.4), corroborating that the standard protein and respiration content in Colombian tropical coffee systems is comparable to that in Northeast USA temperate agricultural systems.

Average scores for chemical indicators generated by Colombian Coffee System functions for pH, P and K were 40, 49, 62% respectively, while the average scores generated by the Northeast USA for the same indicators were 0, 96 and 90% respectively (Table 2.4). Average scores generated by the two regional functions for these three indicators were significantly different from each other with no agreement in scoring pH and P (ICC= 0; ICC= 0 respectively; Table 2.4), and poor agreement in scoring K (ICC=0.239; Table 2.4). Average scores for minor elements generated by the Northeast USA and Colombian Coffee System functions were 74 and 84 respectively and were significantly different from each other with poor agreement in their scoring (ICC=0.22, Table 2.4). These results corroborate that the standards for optimal crop growth in terms of pH and nutrient levels are different for Colombian coffee systems than they are in the Northeast USA, and that the average pH characterized by the Colombian coffee soils (4.8; Table 2.4) are low relative to Northeast USA agricultural system standards thus assigning a pH score of 0% (Table 2.4).

Overall SH index scores averaged 71, 64, and 60% as given by the Northeast USA, Colombian Coffee System and Weighted Colombian Coffee functions respectively, and were significantly different from each other as shown by the t-test. However, ICC showed a moderately strong agreement between the three scoring methods (ICC= 0.636; Table 2.4). This implies that for the same set of samples, SH is scored similarly by the different scoring methods, and that although differences do exist between the different scoring functions for certain indicators which is in part due to the conflicting primary goal of the different scoring methods (e.g. identifying constraints using CASH versus determining optimality by CCSHS and Kenyan functions), the overall SH index scores are ultimately similar.

2.5.3. Best Subsets Regression

When considering the single most predictive variable, AC had the most power in predicting overall soil health with a 0.67 coefficient of determination, followed by OM (0.54; Table 2.5). This suggest that more than half of the variability in overall SH can be explained by either AC or OM, which is in accordance with conclusions by Weil *et al*. (2003) and Culman *et al.* (2012) that AC and OM are key indicators of soil health.

When considering the best pair in predicting soil health, the combination of AC with each of protein and respiration had approximately similar predictive power of overall SH (R2= 0.78 and 0.75 respectively, Table 2.5). This further highlights the important role that AC plays in soil health assessment and how combining it with another biological indicator can increase predictive power by 10%.

When predicting SH using groups of 3 and 4 variables, chemical indicators including P, K and pH are combined with AC and protein to represent as much as 84% of the variability in overall SH (Table 2.5). Incorporating two chemical indicators (P and pH) in the assessment of a highly predictive pair of biological indicator (AC and protein) can increase the predictive power by approximately 5%. This creates potential for a simplified and less expensive soil health assessment using only AC, which may be more attractive to small-holder farmers in Colombian coffee systems, especially given that AC is an easily adopted, in-field test that uses inexpensive equipment.

Table 2.5: Results of Best Subsets Regression using 13 soil health indicators plus Sand, Silt and Clay (n=223)

variables	*R-sq*	*Sand*	*Silt*	*Clay*	*WAS*	*AWC*	*AC*	*OM*	*protein*	*respiration*	*pH*	*P*	*K*	*Mg*	*Fe*	*Mn*	*Zn*
1(1)	*0.67*						*										
1(2)	*0.54*							*									
2(1)	*0.78*						*		*								
2(2)	*0.75*						*			*							
3(1)	*0.81*						*		*						*		
3(2)	*0.8*						*		*				*				
4(1)	*0.84*						*		*		*		*				
4(2)	*0.83*						*		*				*		*		

2.6. CONCLUSIONS

This study was conducted to develop a SH assessment protocol for Colombian coffee systems using scoring functions with parameters derived from average local conditions for the assessment of physical and biological indicators (with the exception of WAS), and with those found in the literature for the assessment of chemical indicators (with the exception of P). A set of adjusted weighted and unweighted SH scoring functions standardized to local conditions in Colombian coffee farms were developed and compared to two other existing scoring methods specific to the NE USA and Western Kenya. Our findings suggest that different scoring methods yield significantly different average indicator and overall SH scores for the same set of soil samples, but that there actually is agreement between the different scoring systems based on the ICC which testifies that the differences identified by t-test are minimal in terms of scoring considering they all generate scores that fall within the same ranges. Such was the case for WAS, AWC, AC, Protein, K, Minor Elements and Overall SH Index score. Despite the strong agreement in scoring the Overall SH Index, the Northeast, USA generally has the lowest soil health standards thus generating the highest scores for the Colombian samples, followed by Colombian Coffee Scoring method, and the Weighted Colombian Coffee scoring method, respectively. Moreover, the standards for scoring WAS, AWC, AC and OM were generally lower in the Northeast, USA than Cauca, Colombian followed by Western, Kenya, thereby generating higher scores by using the Northeast, USA scoring functions. The exceptions were for pH and respiration for which the Northeast, USA has higher standards than Cauca, Colombia, thus scoring them more conservatively. In general, we conclude that adaptation of scoring functions for certain indicators based on regional conditions and cropping systems is necessary to identify specific

constraints. Other indicator scoring functions such as those for Respiration and Protein can be used interchangeably among different localities.

The Best Subsets Regression analysis revealed that AC followed by OM were the best single predictors of overall soil health. Protein and respiration were additionally informative, followed by macronutrients P and K, and pH. Biological indicators seem to generally have highest predictability of overall soil health and should be considered for simplified SH assessments in Colombia.

2.7. ACKNOWLEDGEMENTS

The authors would like to thank the Cornell Statistical Consulting Unit for valuable services, the personnel at the local co-op (FCC) for their contribution to this study and the collaborating farmers without whom this study would not be complete.

This study is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE-1650441.

CHAPTER 3. SOIL HEALTH; A KEY DETERMINANT OF COFFEE CUP QUALITY

3.1. ABSTRACT

Assessing the effect of soil health (SH) on the quality of high value crops such as coffee is important in enabling smallholder farmers to benefit from commodity price premiums as a result of product differentiation. This study was conducted to assess the existence and nature of the relationships between coffee cup quality and SH. Sixty-eight co-op member farms located in six municipalities in Cauca, Colombia were selected, where soil and coffee bean samples were collected. Soil samples were tested for 10 SH indicators including wet aggregate stability (WAS), available water capacity (AWC), active carbon (AC), organic matter (OM), Protein, Respiration, pH, phosphorus (P), potassium (K) and micronutrients (Mg, Mn, Fe and Zn). Elevation and tree density on a 20x50 m quadrant were also recorded on each farm. Coffee samples were tested for a set of physical, granulo-metric and sensorial attributes including fragrance/aroma, flavor, aftertaste, acidity, body, uniformity, sweetness, clean cup and balance. An overall coffee quality score is calculated by summing individual attribute scores minus the defects. A correlation matrix and a full linear model were developed for all SH indicators, elevation, number of trees, as they relate to coffee cup quality. A reduced linear model was developed aided by a best subsets regression (BSR) that identified the best 5 predictors of coffee quality, which were used for parsimony. Best subsets regression was also performed to identify the best three predictors of coffee sensorial traits. Results from the correlation test reveal that coffee quality is in fact negatively associated with AWC, OM, P and Fe. This was also confirmed by the reduced linear model. The full model did not show any significance in predicting coffee quality. We infer that coffee might be similar to grapevines in a way that good

soils for the production of high coffee quality may not have the characteristics of a conventional healthy soil. Identifying the optimal formula that maximizes coffee cup quality without significantly jeopardizing production or environmental welfare should be a focus in future research.

3.2. KEYWORDS

Soil health, Colombia, coffee cup quality

3.3. INTRODUCTION

Since soil health (SH) emerged as a notion, a practice and a belief, there have been countless studies proving the existence of a direct link between SH and plant health. Though, what hasn't been well-established is the link between SH and the quality of plant-derived products. As we know, a rich soil is characterized by higher biological activity thus higher nutrient mineralization, cycling and availability. The question is whether this translates directly into better plant and human nutrient availability and higher product quality. There is growing evidence that fruits harvested from SH-investing organic production have higher levels of vitamins, minerals and antioxidants and better fruit quality and taste marks (Mitchell *et al.,* 2007; Weibel *et al.,* 1998; Reganold *et al.,* 2010). However, the socio-economic implications of smallholder farmers producing superior quality produce in the developing world -especially when it comes to high value export crops such as coffee- remains largely unknown.

Twenty-five million small coffee farmers and their families produce 90% of the world's coffee (ICO 2006). Although some of the poorest rural communities in the tropics rely on coffee production as a way of living, the number of coffee-producing farmers is expected to drop as a

result of the wide fluctuations in the coffee market price, jeopardizing the global coffee supply. At the beginning of the 21st century, the annual prices for *Arabica* beans plummeted by 70% and reached a 30-year low, marking the lowest pay farmers had seen during the 20th century (Gresser and Tickell, 2002). In today's era, coffee production is challenged by not only the effects of climate change but also the inconsistent and unpredictable returns. This has forced farmers to abandon their farms and leave the coffee business in pursuit of a more secure source of income.

In an attempt to mitigate this problem, small farmers must focus their attention on ways to increase their revenues in the coffee business. This could be achieved by quality differentiation and the production of high quality specialty coffee that is worth price premiums. Superior coffee cup quality starts at the farm level, and honing farm management practices that help achieve the traits desired in a cup of coffee is as important as maximizing yield. We believe it should be on the forefront of a grower's goals.

It has been previously reported that variation in the method of cultivation such as cultivar, tree density, shade coverage, and processing method all affect coffee quality (Läderach *et al.,* 2011; Oberthür *et al.,* 2011; Vaast *et al.,* 2006; Vaast *et al.,* 2005, Daniels, 2009, Decazy *et al.,* 2003). From a biochemical standpoint, certain compounds within the coffee bean such as proteins, carbohydrates, amino acids, peptides and phenolic compounds are thought to play a role in determining coffee quality (Clifford, 1997; De Amorim *et al.,* 1968; Montavon *et al.,* 2003). From a soil management standpoint, Castro-Tanzi *et al.* (2012) reported that soil Calcium (Ca) depletion as a result of excessive NPK fertilizer use and increased aluminum (Al3+) toxicity could be involved in the reduction of cupping quality. They further revealed that coffee cup quality increased with increasing CaO application.

To our knowledge, this is the only study that has attempted to establish a relationship between soil characteristics and coffee quality. Therefore, the objectives of this study are to assess whether there is a relationship between coffee quality and SH, and to determine which SH indicators influence coffee cup quality score and sensorial attributes while accounting for the effect of elevation and number of trees.

3.4. MATERIALS AND METHODS

3.4.1. Soil Health Sampling and Laboratory Methods

Soil samples were collected in January 2014 following the CASH protocol (Moebius-Clune *et al.,* 2016) from 68 farms across 6 municipalities (Cajibio, Timbío, Rosas, Piendamò, Morales and Popayàn) in Cauca, Colombia. At each farm, two soil samples were collected: One from the farm's most fertile plot and the other from the least fertile, as indicated by the farmer- producing a total of 136 soil samples. All samples were collected from the 0-to-15-cm depth range using a Dutch-style soil auger after surface residue removal. At each sampling location, five samples were collected and then composited to obtain a representative sample. The 136 samples were sent to Cornell University in Ithaca, NY (USA) where the samples were air-dried and passed through a 2-mm sieve in preparation for the analysis of physical, chemical and biological soil properties following the Comprehensive Assessment of Soil Health (CASH) protocol (Moebius-Clune *et al.,* 2016).

Physical Indicators: Available Water Capacity (AWC) between field capacity (-10kPa) and permanent wilting point (-1500kPa) was assessed gravimetrically by equilibrating saturated soil to each of 10 kPa and 1500 kPa on ceramic high pressure plates (Topp and Zebchuck 1979). The

difference between soil water loss under 10 kPa and 1500 kPa pressures determined from calculating the difference in wet and dry weights was considered the AWC (Moebius-Clune *et al.,* 2016).

Wet Aggregate Stability (WAS) was assessed using a rainfall simulator adapted from Ogden *et al*. (1997) that allows particles of air-dried soil placed on a 0.25 mm mesh sieve to slake under 2.5 J of rainfall energy for 300 seconds, based on a total of 2.5 cm of rainfall. Wet Aggregate Stability was determined by subtracting the weight of slaked soil plus the remaining stones on the sieve (>0.25 mm) from total soil weight measured before rainfall (Moebius-Clune *et al.,* 2016).

Soil texture was determined using a rapid quantitative method developed by Kettler *et al.* (2001) where soil samples were fractionated with 3% sodium hexametaphosphate ($(NaPO_3)_n$) and a series of sieving and sedimentation steps was used to separate the different particle sizes.

Biological Indicators: Organic Matter content (OM) was analyzed by mass loss on ignition in a muffle furnace at 500 °C for two hours, with values corrected by multiplying percent loss on ignition by 0.7 and subtracting 0.23 (Moebius-Clune *et al.,* 2016). Active Carbon (AC) was measured by adding a dilute potassium permanganate solution ($KMnO_4$) to soil, which acts as an oxidant to AC, and measuring the solution's absorbance at 550 nm using a hand-held colorimeter (Hach, Loveland, CO) (Weil *et al.,* 2003).

Autoclaved Citrate Extractable Soil Protein Index (Protein) was measured by extracting proteins from the soil following a series of centrifugation and autoclaving steps using 0.02 M sodium citrate at pH 7. Soil protein concentration was determined by measuring bicinchoninic acid assay

against bovine serum albumin standard curve for soil protein concentration (Walker, 2009; Wright and Upadhyaya, 1996). The soil Respiration test was performed by trapping and measuring CO_2 emitted by soil microorganisms over a 4-day room temperature incubation in a sealed chamber with a KOH trap (Haney and Haney, 2010). The dissolved CO_2 in the KOH trap quantifying microbial activity was measured using an electrical conductivity meter, and the change in conductivity before and after incubation quantifies the amount of CO_2 evolved.

Chemical Indicators: Soil pH was measured in a 1:1 water dispersed slurry determined by a pH electrode probe (SM802 Smart Combined Meter, Milwaukee Industries, Rocky Mount, NC). Soil nutrients, including P, K, Mg, Fe, Mn and Zn were extracted with a Modified Morgan solution (ammonium acetate - buffered at pH 4.8), and quantified by Inductively Coupled Plasma Optical Emission Spectroscopy (ICP–OES, Varian 730-ES, Mulgrave, Victoria, Australia).

3.4.2. Elevation and Tree Density Index

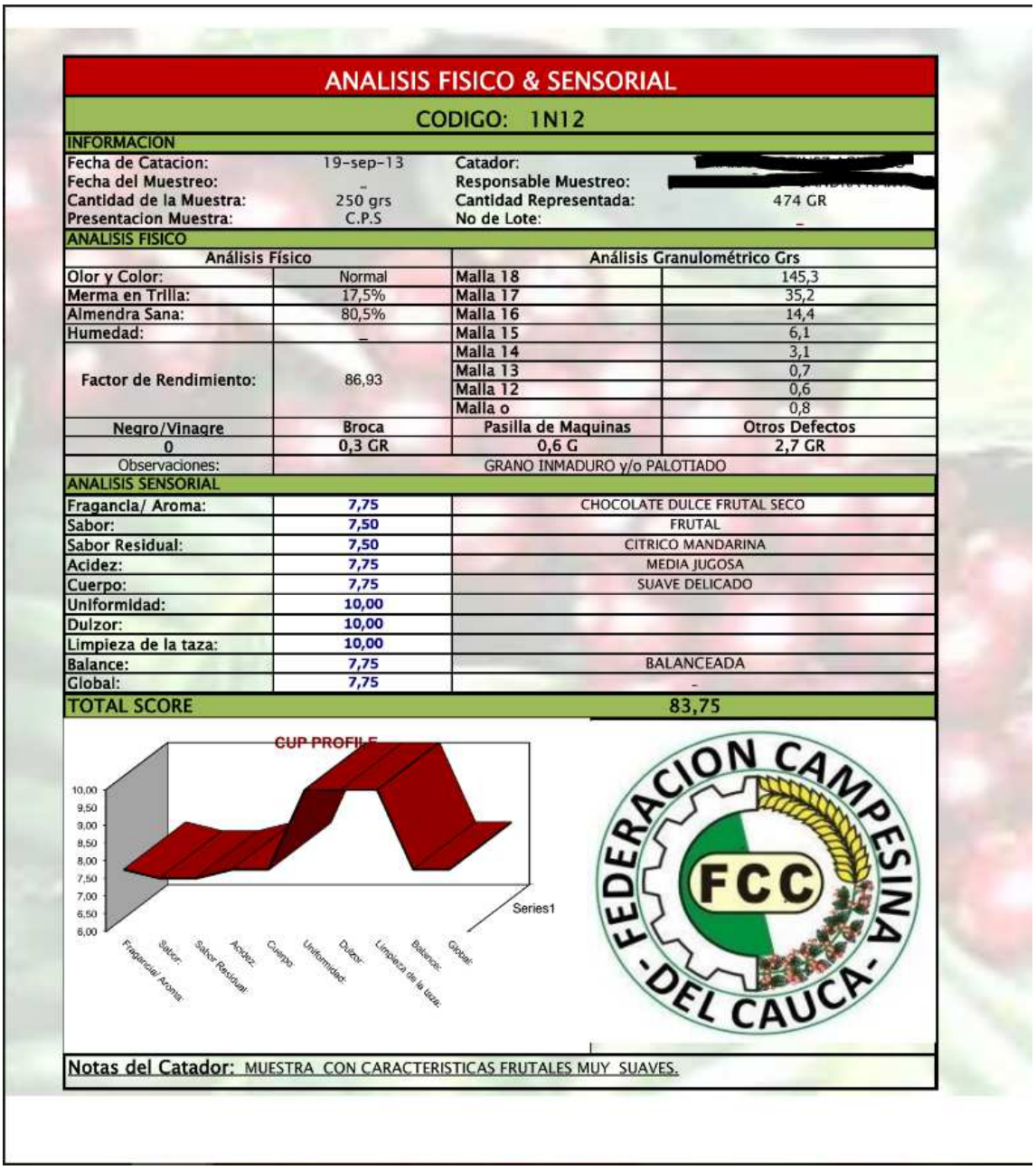
Each farm was demarcated by a 20x 50 m quadrant where tree species richness was measured by counting the number of non-coffee trees within the entire transect area and compositing them into a tree density index as a proxy for both shade coverage and biodiversity. Elevation was recorded with a hand-help G.P.S device (Garmin-Etrex 10) in the southwestern-most point of the quadrant.

3.4.3. Coffee Sampling and Quality Scoring

Approximately 1 kg of green coffee was collected from each farm in the first quarter of September 2013. Samples were processed, roasted and ground, and were analyzed for quality by a professional certified cupper at the coffee quality laboratory within the Coffee Growers

Federation of Cauca (FCC) headquarters in Popayàn. The professional cupper scored a set of physical (odor and color, decline in threshing, shape, humidity), granule (a series of sieving tests), and sensorial quality attributes (aroma, flavor, aftertaste, acidity, body, uniformity, sweetness, cup cleaning and balance; Fig. 3.1), following the Specialty Coffee Association of America (SCAA) Standards which are adopted by the multi-national coffee wholesalers. The scores were composited, and a quality index score representing the overall coffee quality was derived. (Figure 3.1 shows a sample of a coffee quality report generated by FCC). This study only focused on the overall quality index and the sensorial scores (SS) which were categorized as "Low" ($SS < 3$), "Medium" ($3 \leq SS \leq 7$) and "High" ($SS > 7$).

Because the coffee samples were not collected and analyzed at the peak of the harvest season, and given the fact that quality of beans degrades within hours of being picked (Daniels, 2009), scores were corrected for quality using weighted quality averages for each municipality (retrieved from the FCC historical data archives), where samples coming from the same farm were given more weight. Thus, we adjusted the observed means by evaluating the distance of each farm's coffee quality score from the weighted average for its respective municipality and applying these deviations to the observed averages. Such, adjustments were made to the means to comply with historic averages without altering the distribution.



ANALISIS FISICO & SENSORIAL

CODIGO: 1N12

INFORMACION			
Fecha de Catacion:	19-sep-13	Catador:	[redacted]
Fecha del Muestreo:	-	Responsable Muestreo:	[redacted]
Cantidad de la Muestra:	250 grs	Cantidad Representada:	474 GR
Presentacion Muestra:	C.P.S	No de Lote:	-

ANALISIS FISICO			
Análisis Físico		Análisis Granulométrico Grs	
Olor y Color:	Normal	Malla 18	145,3
Merma en Trilla:	17,5%	Malla 17	35,2
Almendra Sana:	80,5%	Malla 16	14,4
Humedad:	-	Malla 15	6,1
Factor de Rendimiento:	86,93	Malla 14	3,1
		Malla 13	0,7
		Malla 12	0,6
		Malla o	0,8
Negro/Vinagre	Broca	Pasilla de Maquinas	Otros Defectos
0	0,3 GR	0,6 G	2,7 GR
Observaciones:	GRANO INMADURO y/o PALOTIADO		

ANALISIS SENSORIAL		
Fragancia/ Aroma:	7,75	CHOCOLATE DULCE FRUTAL SECO
Sabor:	7,50	FRUTAL
Sabor Residual:	7,50	CITRICO MANDARINA
Acidez:	7,75	MEDIA JUGOSA
Cuerpo:	7,75	SUAVE DELICADO
Uniformidad:	10,00	
Dulzor:	10,00	
Limpieza de la taza:	10,00	
Balance:	7,75	BALANCEADA
Global:	7,75	-
TOTAL SCORE		83,75

Notas del Catador: MUESTRA CON CARACTERISTICAS FRUTALES MUY SUAVES.

Figure 3.1: An example coffee quality report generated by the Federation Campensina del Cauca (the Cauca Growers Federation, FCC)

3.4.4. Statistical Methods

Multiple Linear Regression (MLR) and Best Subsets Regression (BSR) were performed on the coffee quality scores as they relate to each of the individual SH indicators plus elevation and number of trees. A reduced MLR model was performed based on the BSR selection of the five best predictors of coffee quality. A simple correlation test was also performed between coffee quality score and each of the SH indicators, elevation and tree density. Best Subsets Regression was further performed to determine which individual and sets of SH indicators are best at predicting each of the nine sensorial coffee quality attributes. Respiration was excluded from these models due to collinearity.

All statistical analyses were performed using the R-Project for Statistical Computing (Team, 2013).

3.5. RESULTS AND DISCUSSIONS

3.5.1. Variability of Coffee Quality Scores Between Municipalities

Tables 3.1 and 3.2 show the SH and coffee quality differences among municipalities. All SH indicator measurements except K, number of trees, elevation, and overall SH differed significantly among municipalities (Table 3.1). Coffee quality scores did not differ by municipality suggesting that the six regions produce similar grades of coffee quality (Table 3.2; Figure 3.2).

Table 3.1: Descriptive statistics (mean (sd)), including average measurements of soil health indicators, %clay, elevation and number of trees by municipality (n=68)

	Rosas	Timbío	Piendamó	Cajibio	Morales	Popayán	**Total**
WAS- values (%)	82.2(17.2)^{b}	94.2(4)^{a}	96.7(3.5)^{a}	97.3(1.7)^{a}	97.1(2.1)^{a}	95.6(6.9)^{a}	**93.8(9.7)**
AWC (g/g)	0.2(0.1)^{c}	0.3(0.1)^{b}	0.4(0.1)^{a}	0.3(0.1)^{b}	0.3(0.1)^{ab}	0.2(0.1)^{b}	**0.27(0.1)**
OM- values (%)	9.3(4.6)^{c}	19.0(3.6)^{ab}	19.5(5.4)^{ab}	19.8(5.7)^{ab}	22.7(3.4)^{a}	17.8(4.8)^{b}	**17.9(6.3)**
AC- values (ppm)	527.9(183.1)^{c}	856.2(133.2)^{ab}	860.2(242.5)^{ab}	894.0(219.5)^{ab}	1000(174.3)^{a}	764.4(208.3)^{b}	**813.4(244.8)**
Protein- values (mg/g soil)	7.7(2.7)^{b}	9.5(2.3)^{ab}	9.4(2.5)^{ab}	10.2(2.8)^{a}	9.8(2.1)^{ab}	8.2(3.0)^{ab}	**9.1(2.7)**
Respiration- values (mg/g soil)	0.8(0.3)^{c}	1.0(0.2)^{abc}	1.0(0.2)^{ab}	1.1(0.3)^{ab}	1.2(0.2)^{a}	0.9(0.3)^{bc}	**1.0(0.3)**
pH- values	5.0(0.3)^{a}	4.8(0.2)^{b}	4.7(0.3)^{b}	4.7(0.3)^{b}	4.8(0.3)^{ab}	4.9(0.3)^{ab}	**4.8(0.3)**
P- values (ppm)	5.5(3.1)^{c}	11.0(2.1)^{ab}	12.9(5.6)^{a}	8.6(3.9)^{bc}	12.9(3.8)^{a}	8.1(3.7)^{bc}	**9.6(4.6)**
K- values (ppm)	124.6(62.7)^{a}	95.9(37.6)^{a}	118.8(74.6)^{a}	103.6(66.9)^{a}	103.4(52.2)^{a}	126.6(121.3)^{a}	**112.5(75.2)**
Mg- values (ppm)	562.0(530.6)^{a}	58.3(35.9)^{b}	70.3(58.4)^{b}	273.2(689.9)^{ab}	52.6(30.8)^{b}	127.7(266.7)^{b}	**199.0(421.1)**
Fe- values (ppm)	26.3(18.8)^{a}	17.3(8.9)^{a}	23.5(10.8)^{a}	20.8(16.9)^{a}	22.3(4.6)^{a}	17.9(6.5)^{a}	**21.3(12.5)**
Mn- values (ppm)	9.5(3.3)^{a}	6.8(3.1b)^{a}	5.5(3.2)^{bc}	6.4(5.0)^{b}	3.1(2.0)^{c}	5.7(4.3)^{bc}	**6.2(4.1)**
Zn- values (ppm)	1.4(1.4)^{a}	0.5(0.3)^{b}	1.1(1.1)^{ab}	0.9(0.8)^{ab}	1..0(0.9)^{ab}	0.9(0.7)^{ab}	**1.0(0.9)**
Clay (%)	25.6(5.9)^{ab}	19.6(4.1)^{c}	27.9(9.2)^{a}	28.8(8.1)^{a}	24(4.4)^{abc}	21.7(7.3)^{bc}	**24.6(7.3)**
Elevation	1628(196) ^{c}	1778(31) ^{a}	1751(70) ^{abc}	1651(74)^{bc}	1699(15)^{abc}	1742(52) ^{ab}	**1705(107)**
Number of Trees	15(8)^{a}	12(4)^{ab}	7(6)^{ab}	9(8)^{ab}	12(12)^{ab}	4(5)^{b}	**10(9)**

† a, b, c, d significant homogeneous groups among municipalities (horizontal comparison) for each indicator measured value using Tukey's HSD at $p<0.05$.

Table 3.2: Descriptive statistics of soil health scores and coffee quality scores by municipality

Municipality	Soil Health Score			Coffee Quality Score		
	Min.	Max.	Mean	Min.	Max.	Mean
Rosas	45	77	57.3[c]	74.4	87.6	82.5[a]
Timbío	54	80	68.6[ab]	68.9	92.9	84.0[a]
Piendamó	49	87	71.4[ab]	74.3	83.2	79.9[a]
Cajibio	49	86	68.3[ab]	78.2	82.1	80.6[a]
Morales	59.0	90.0	74.2[a]	68.6	87.9	80.2[a]
Popayán	38	77	63.2[bc]	67.9	86.4	81.7[a]
Total	**38.0**	**90**	**66.9**	**67.9**	**92.9**	**81.6**

† a, b, c significant homogeneous groups among municipalities for each indicator using Tukey's HSD at $p<0.05$.

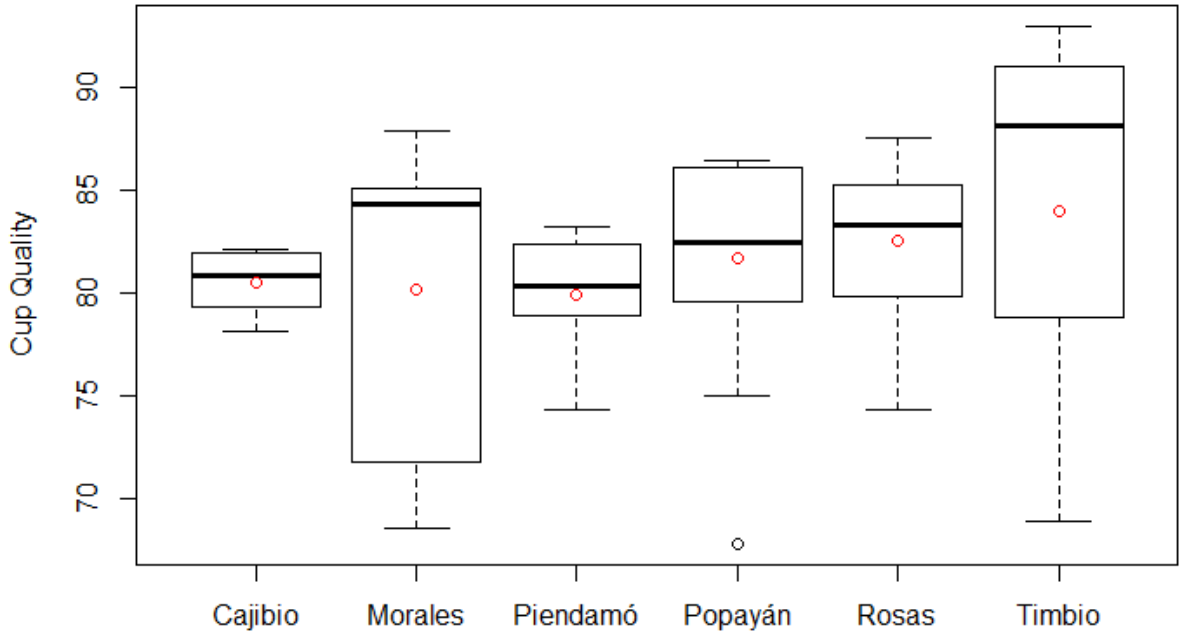


Figure 3.2: Coffee quality scores across the six municipalities in our project location in Cauca, Colombia (n=68). Red dots indicate means.

Although coffee quality does not differ significantly between municipalities (p=0.637; Table 3.2), there is a wide range within most municipalities (Timbío, Morales, Rosas, Popayàn). For example, the distribution of coffee quality in Morales shows that it has the highest median of coffee quality, with the 25th percentile starting at the lowest level compared to all municipalities. Contrastingly, the variability within Cajibio seems to be consistent and uniform despite the relatively low median (Figure 3.2).

3.5.2. Effect of Soil Health Indicators on Coffee Cup Quality

Results from the correlation test revealed that coffee quality seems to be negatively correlated with AWC and OM using the significance level α=0.1 (r= -0.25, -0.23; Table 3.3). Although the results from the MLR of the full model did not show any significance in predicting coffee quality by the explanatory variables due to a weak model (Adj. R^2= -0.151 Table 3.4), the reduced model containing OM, Protein, Respiration, pH and Fe, as determined by the BSR to be the best subset for predicting coffee quality, showed that at a significance level of α=0.1 these SH indicators, with the exception of Protein, were significant in effecting coffee quality, and that OM, pH and Fe have an inverse relationship with coffee quality (β= -0.5; -4.9; -0.2 respectively; Adj. R2= 0.13; Table 3.5).

Table 3.3: Correlation Coefficients of coffee quality to all SH indicators (n=68)

	Coffee Quality
Coffee Quality	1
WAS	-0.02
AWC	**-0.25***
OM	**-0.23***
AC	-0.2
Protein	0.05
Respiration	0.06
pH	-0.05
P	-0.18
K	0.09
Mg	0.04
Fe	-0.17
Mn	0.18
Zn	-0.09
Clay	0.04
Sand	0.10
Elevation	-0.01
Number of Trees	-0.12

'*': significant at α=0.1

Table 3.4: Multiple Linear Regression (MLR) of all individual soil health (SH) indicators plus Number of Trees and Elevation as they relate to coffee cup quality score (n=68)

	Estimate	Standard Error	t value	p value	Model Adjusted R-squared	R-squared of Best subset of 5 Predictors based on BSR
(Intercept)	119.30	26.63	4.48	0.00 ***	-0.151	0.2
WAS	-0.06	0.17	-0.36	0.72		
AWC	-5.09	18.26	-0.279	0.78		
AC	0.00	0.01	0.139	0.89		
†OM	-0.25	0.56	-0.439	0.66		
Protein	-0.64	0.62	-1.017	0.32		
Respiration	12.12	7.84	1.546	0.13		
pH	-6.68	5.84	-1.144	0.26		
P	-0.28	0.36	-0.769	0.45		
K	0.01	0.02	0.331	0.74		
Mg	0.00	0.00	0.001	1.00		
Mn	-0.05	0.38	-0.123	0.90		
Fe	-0.12	0.11	-1.07	0.29		
Zn	-0.44	1.29	-0.343	0.73		
Number of Trees	-0.06	0.13	-0.43	0.67		
Elevation	0.00	0.02	0.152	0.88		

'***': significant at α=0

†: indicators in bold signify the five best predictors of coffee quality

Table 3.5: Reduced Linear Model of the Top Soil Health Predictors of Coffee Quality (n=68)

	Estimate	Standard Error	t value	p value	Model Adjusted R-squared
(Intercept)	111.9243	13.92386	8.038	8.56E-11 ***	0.1271
OM	-0.47162	0.14061	-3.354	0.00146 **	
Protein	-0.35907	0.3823	-0.939	0.35179	
Respiration	8.55549	4.40757	1.941	0.05747 .	
pH	-4.88291	2.53991	-1.922	0.05983 .	
Fe	-0.17113	0.07525	-2.274	0.02696 *	

".": significant at α=0.1; '*': significant at α=0.05; "**": significant at α=0.01; "***": significant α=0.001

Results from the reduced linear model and the correlation test confirm that coffee quality is in fact higher with lower AWC, OM, P and Fe, and higher Mn. Coffee quality's negative correlation with AWC and OM could suggest that it is enhanced when there is less supply of water in the soil or under water deficits. This suggests that coffee might be similar to grapevines in that good quality coffee is usually derived from poorer soils that tend to cause water stress during certain stages of the plant's growth. This phenomenon is explained by the more efficient photosynthate partitioning and sugar accumulation that happens under moderate water stresses (Prichard, 2004). These results are in accordance with a study by Silva *et al.* (2005) on coffee beans which revealed that reducing sugars, phenols, proteins, nitrogen, protease and polyphenoloxidase were higher under non-irrigated or suspended-irrigated conditions than under continuously irrigated conditions in some regions of São Paulo, Brazil. It is important to consider however that according to Camargo and Camargo (2001), a severe drought during coffee's fruit setting period can result in fruit drop. Therefore, a thorough understanding of induced water deficits and the underlying mechanisms by which higher quality coffee is achieved should be one priority in specialty coffee research. It is important to understand how and when to optimize superior cup quality through carefully-managed water shortages without jeopardizing quantity produced. Farmers must subsequently learn how to effectively balance tradeoffs between quality, quantity, and environmental consequences in their management decisions for specialty coffee production. Moreover, the fact that coffee quality is higher with lower pH could hint to a direct link between acidity in the soil and in the coffee bean. However, once more, farmers must compromise between quality and quantity since it is shown in the literature that coffee is more productive under pH approaching acid-neutral levels (Winston *et al.*, 2005; Bitterbender;

Malavolta and Netto, 1989; Kuit *et al.*,2004; South Africa Department of Agriculture, Forestry and Fisheries, 2012).

Finally, coffee quality appears to favor higher levels of respiration. This could hint to the possibility that biological activity in the soil may play a role in mediating some of the important biochemical pathways in the cherry that are important in triggering quality gene expression. However, Respiration may be confounded by OM since there is no correlation using respiration alone (Table 3.3).

3.5.3. Soil Health Indicators Affecting Coffee Sensorial Attributes

The BSR of a set of three indicators affecting each of the nine sensorial coffee attributes revealed that Mn was most often one of the three best indicators in affecting coffee sensorial attributes (5 out of 9 attributes; Table 3.6), followed by a tie between sets that included WAS, AWC, AC and Protein, which were equally affecting four out of the nine coffee sensorial attributes. Therefore, combining Mn with two of these indicators would offer increased predictive capability of coffee sensorial attributes. Only one of the two non-soil indicators (number of trees) was associated with flavor and aftertaste possibly due to the lack of variation in farm elevation (i.e. all coffee was grown in high elevations). However, looking at the coefficients of determination for each subset, there is limited predictability (R^2= 0.17-0.36; Table 3.6). Uniformity was the exception which had a relatively high degree of predictability by OM, K and Fe (R^2=0.61; Table 3.6). Ultimately, it does appear that some SH indicators are more important in terms of affecting sensorial traits than others. These mainly are the Physical and Biological indicators. As for the chemical indicators, Mn appears to be the highest influencer.

Table 3.6: Best subset of three variables affecting coffee sensorial traits (n=23)

Indicator	Sensorial Attributes									Number of times the indicator appeared in the best subset affecting the different sensorial attributes
	Aroma	Flavor	Aftertaste	Acidity	Body	Uniformity	Sweetness	Bowl Cleaning	Balance	
WAS		*	*				*	*		**4**
AWC	*			*	*				*	**4**
OM						*				**1**
AC	*			*	*				*	**4**
Protein				*	*		*	*		**4**
pH										**0**
P										**0**
K	*					*				**2**
Mg										**0**
Fe						*				**1**
Zn										**0**
Mn		*	*				*	*	*	**5**
Sand										**0**
Elevation										**0**
Number of Trees		*	*							**2**
R-squared	0.17	0.21	0.21	0.34	0.36	0.61	0.21	0.27	0.28	

3.6. CONCLUSIONS

This study was conducted to assess the existence and nature of a relationship between SH and coffee cup quality. The full linear model containing all SH indicators, number of trees and elevation showed that none of the predictor variables had significant effect on coffee quality, whereas the reduced linear model containing OM, Protein, Respiration, pH and Fe -as determined by the BSR- showed that all, with the exception of Protein, have a significant effect on coffee quality. The correlation test between coffee quality and SH indicators further confirmed that the coffee quality tends to be significantly higher with lower OM and AWC levels, suggesting that coffee, like grapevines, tends to yield better quality product under conditions where the plants experience some water and nutrient stresses. However, these results remain constrained by the observational nature of our study, and would need to be confirmed in controlled trials. In hindsight, the subjective nature of cupping is one aspect of this study that could have hindered the establishment of a strong relationship between SH indicators, number of trees and elevation and the quality of coffee. It was reported by Silva et al. (2005) that statistical studies revealed that sensorial classification done by trained tasters is susceptible to errors. Hence, by means of their suggestions, a more objective method involving artificial sensors such as the "electronic tongue" (Riul *et al.,* 2003; Ferreira *et al.,* 2003) could be a more reliable and accurate method in differentiating coffee quality in future studies.

3.7. ACKNOWLEDGEMENTS

This study is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE-1650441.

CONCLUSIONS

The purpose of this multidisciplinary study was threefold: to identify the factors that influence a farm's soil health and the farmers' perception thereof in smallholder Colombian coffee systems, to develop a set of adjusted scoring functions specific to Cauca, Colombia coffee farms and compare it to other regional scoring functions (Northeast USA and Western Kenya), and to determine the existence and nature of the relationship between coffee quality and soil health. Results from the first chapter showed that female farmers and co-op members have significantly higher SH than their counterparts, that farmers appear to have correct perception of their SH which is not associated with their gender or municipality, but with how healthy their soil actually is, and that OM, respiration and protein are indicators that are most related to farmers' perceptions of their soil health. Results from the second chapter in regards to comparison of regional scoring functions revealed that although discrepancies between the different scoring functions for the same indicator exist (which is due to the agro-climatic differences between the regions) achieving global soil health standards is possible when only considering soil health index scores or certain individual indicators. Results from the third chapter showed that certain soil health indicators including OM, AWC and pH are significant determinants of coffee cup quality score and that mainly physical and biological indicators affect individual coffee sensorial traits. In conclusion, we emphasize that the observational nature of this study has allowed us with constraints to only lightly unveil the surface of what's more to be discovered. We encourage others to challenge or elaborate on our findings, especially through controlled trials. Until then, there will remain more questions than answers.

REFERENCES

Abreu, Cleide Aparecida de, van Raij, Bernardo, Abreu, Mônica Ferreira de, & González, Antonio Paz. (2005). Routine soil testing to monitor heavy metals and boron. Scientia Agricola, 62(6), 564-571

Ahad, N. A., Yin, T. S., Othman, A. R., & Yaacob, C. R. (2011). Sensitivity of normality tests to non-normal data. Sains Malaysiana, 40(6), 637-641.

Al-Mustafa, A., Abdallah, E., & Falatah, M. (2001). Assessment of Five Extractants for their Ability to Predict Iron Uptake and Response of Sorghum Grown in Calcareous Soils. Communications in Soil Science and Plant Analysis, 32(5-6), 907–919.

Andrews, S.S., D.L. Karlen, and C.A. Cambardella. (2004). The soil management assessment framework: A quantitative soil quality evaluation method. Soil Science of America Journal 68: 1945-1962

Arshad, M.A., and S. Martin. 2002. Identifying critical limits for soil quality indicators in agro ecosystems. Agric. Ecosyst. Environ. 88:153–160. doi:10.1016/S0167-8809(01)00252-3

Barrios, E., R.J. Delve, A. Bekunda, J. Mowo, J. Agunda, J. Ramisch, M.T. Trejo, and R.J. Thomas. (2006). Indicators of soil quality: A South-South development of a methodological guide for linking local and technical knowledge. Geoderma 135:248-259.

Bitterbender, HC. Coffee Growing Basics: Fertilizer, Disease, Insects [PowerPoint slides]. Retrieved from: http://www.jatropha.pro/PDF%20bestanden/Coffee%20basics-%20pests%20disease%20nutrition.pdf.

BJ Alloway. (1995). Heavy metals in soils. Springer Science & Business Media.

Camargo, A. D., & CAMARGO, M. D. (2001). Definição e esquematização das fases fenológicas do cafeeiro arábica nas condições tropicais do Brasil. Bragantia, 60(1), 65-68.

Castro-Tanzi, S., Dietsch, T., Urena, N., Vindas, L., & Chandler, M. (2012). Analysis of management and site factors to improve the sustainability of smallholder coffee production in Tarrazú, Costa Rica. Agriculture, ecosystems & environment, 155, 172-181.

Clifford, M. N. (1997). The nature of Chlorogenic acids. Are they advantageous compounds in coffee?. In *COLLOQUE Scientifique International sur le Café, 17. Nairobi (Kenya), Juillet 20-25, 1997...*

Congreves, K. A., Hayes, A., Verhallen, E. A., & Van Eerd, L. L. (2015). Long-term impact of tillage and crop rotation on soil health at four temperate agroecosystems. Soil and Tillage Research, 152, 17-28.

Culman, S.W., Snapp, S. S., Freeman, M. A., Schipanski, M. E., Beniston, J., Lal, R., Drinkwater, L. E., Franzluebbers, A. J., Glover, J. D., Grandy, A. S., Lee, J. Six, J. Maul, J. E., Mirsky, S. B., Spargo, J. T., Wander, M. M. (2012) Permanganate oxidizable carbon reflects a processed soil fraction that is sensitive to management. Soil Sci. Soc.Am. J. 76:494-504.

Daniels, N. (2009). Variations in Coffee Processing and Their Impact on Quality and Consistency (Doctoral dissertation, Michigan Technological University).

De Kreij, C., Benchaalal, K., & Van Elderen, C. W. (1993). Comparison of water, ammonium acetate, and DTPA to assess the availability of manganese in peat substrate. Communications in Soil Science & Plant Analysis, 24(9-10), 927-937.

De Amorim, H. V., & Silva, D. M. (1968). Relationship between the polyphenol oxidase activity of coffee beans and the quality of the beverage. Nature, 219, 381-382.

De Jalón, S. G., Silvestri, S., Granados, A., & Iglesias, A. (2015). Behavioural barriers in response to climate change in agricultural communities: an example from Kenya. Regional Environmental Change, 15(5), 851-865.

Decazy, F., Avelino, J., Guyot, B., Perriot, J. J., Pineda, C., & Cilas, C. (2003). Quality of different Honduran coffees in relation to several environments. Journal of Food Science, 68(7), 2356-2361.

Department of Agriculture, Forestry and Fisheries. (2012.) *Production Guideline: Coffee (Coffea spp.)* [Brochure]. Republic of South Africa: Department of Agriculture, Forestry and Fisheries

Dexter, A.R. (2004). Soil physical quality - Part I. Theory, effects of soil texture, density, and organic matter, and effects on root growth. Geoderma 120:201-214.

Ferreira, M., Riul, A., Wohnrath, K., Fonseca, F. J., Oliveira, O. N., & Mattoso, L. H. (2003). High-performance taste sensor made from Langmuir-Blodgett films of conducting polymers and a ruthenium complex. Analytical Chemistry, 75(4), 953-955.

Gresser, C., S. Tickell (2002) Mugged: Poverty in Your Coffee Cup. Oxfam International, Boston, 60p

Gugino, B. K., Abawi, G. S., Idowu, O. J., Schindelbeck, R. R., Smith, L. L., Thies, J. E., ... & Van Es, H. M. (2009). Cornell soil health assessment training manual. Cornell University College of Agriculture and Life Sciences.

Haight, C. (2011). The problem with fair trade coffee. Stanford Social Innovation Review, 3, 74-79.

Haney, R. L., & Haney, E. B. (2010). Simple and rapid laboratory method for rewetting dry soil for incubations. Communications in soil science and plant analysis, 41(12), 1493-1501.

Hernandez-Aguilera, J. N., Gómez, M. I., Rodewald, A. D., Rueda, X., Anunu, C., Bennett, R., ... & van Es, H. M. (2015, May). Impacts of smallholder participation in high-quality coffee markets: The Relationship Coffee Model. In 2015 AAEA & WAEA Joint Annual Meeting, July 26-28, San Francisco, California (No. 205650). Agricultural and Applied Economics Association & Western Agricultural Economics Association.

Hurisso, T. T., Culman, S. W., Horwath, W. R., Wade, J., Cass, D., Beniston, J. W., ... & Lucas, S. T. (2016). Comparison of permanganate-oxidizable carbon and mineralizable carbon for assessment of organic matter stabilization and mineralization. Soil Science Society of America Journal, 80(5), 1352-1364.

Idowu, O. J., van Es, H. M., Abawi, G. S., Wolfe, D. W., Ball, J. I., Gugino, B. K., ... & Bilgili, A. V. (2008). Farmer-oriented assessment of soil quality using field, laboratory, and VNIR spectroscopy methods. Plant and Soil, 307(1-2), 243-253.

International Coffee Organization (2006) Annual Review (http://www.ico.org/documents/review8e.pdf) accessed 10/25/2016.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. Educational and Psychological Measurement, 20, 141-151.

Karltun, E., Lemenih, M., & Tolera, M. (2013). Comparing Farmers' Perception of Soil Fertility Change with Soil Properties and Crop Performance In Beseku, Ethiopia. Land Degradation & Development, 24(3), 228-235.

Kettler, T.A., J.W. Doran, and T.L. Gilbert. (2001). Simplified method for soil particle-size determination to accompany soil-quality analyses. Soil Sci. Soc. Am. J. 65:849–852. doi:10.2136/sssaj2001.653849x

Knutson, C. L., Haigh, T., Hayes, M. J., Widhalm, M., Nothwehr, J., Kleinschmidt, M., & Graf, L. (2011). Farmer perceptions of sustainable agriculture practices and drought risk reduction in Nebraska, USA. Renewable Agriculture and Food Systems, 26(03), 255-266.

Kuit, M., Jansen, D. M., & Nguyen, V. T. (2004). Manual for Arabica cultivation. Tan Lam Agricultural Product Joint Stock Company, Quang Tri.

Läderach, P., Oberthür, T., Cook, S., Estrada Iza, M., Pohlan, J.A., Fisher, M., Rosales, Lechuga, R., 2011. Systematic agronomic farm management for improved coffee quality. Field Crops Research 120, 321-329

LHSP. Labeling Similar Groups. (n.d.). Retrieved March 20, 2016, from http://www.jerrydallal.com/LHSP/similar_prog.htm

Malavolta, E., & Netto, A. V. (1989). Nutrição mineral, calagem, gessagem e adubação dos citros. Associação Brasileira para Pesquisa da Potassa e do Fosfato. p198.

McDonald, J. H. (2009). Handbook of biological statistics (Vol. 2, pp. 173-181). Baltimore, MD: Sparky House Publishing.

Melke, A., & Ittana, F. (2015). Nutritional Requirement and Management of Arabica Coffee (Coffea arabica L.) in Ethiopia: National and Global Perspectives. American Journal of Experimental Agriculture, 5(5), 400.

Mitchell, A. E., Hong, Y. J., Koh, E., Barrett, D. M., Bryant, D. E., Denison, R. F., & Kaffka, S. (2007). Ten-year comparison of the influence of organic and conventional crop management practices on the content of flavonoids in tomatoes. Journal of Agricultural and Food Chemistry, 55(15), 6154-6159.

Moebius-Clune, B. N. (2010). Applications of integrative soil quality assessment in research, extension, and education (Doctoral dissertation, Cornell University).

Moebius-Clune, B.N., D.J. Moebius-Clune, B.K. Gugino, O.J. Idowu, R.R. Schindelbeck, A.J. Ristow, H.M. van Es, J.E. Thies, H. A. Shayler, M. B. McBride, D.W. Wolfe, and G.S. Abawi, (2016). Comprehensive Assessment of Soil Health – The Cornell Framework Manual, Edition 3.0, Cornell University, Geneva, NY.

Moebius-Clune, B. N., van Es, H. M., Idowu, O. J., Schindelbeck, R. R., Kimetu, J. M., Ngoze, S., Lehmann, J., and Kinyangi, J. M. (2010). Development and evaluation of scoring functions for integrative soil quality assessment and monitoring in western Kenya. In Applications of Integrative Soil Quality Assessment in Research, Extension, and Education. Ph.D. Dissertation, Cornell University, Ithaca NY.

Moebius, B.N., H.M. van Es, R.R. Schindelbeck, O.J. Idowu, J.E. Thies, and D.J. Clune. (2007). Evaluation of Laboratory-Measured Soil Properties as Indicators of Soil Physical Quality.Soil Science 172:895-912.

Montavon, P., Duruz, E., Rumo, G., & Pratz, G. (2003). Evolution of green coffee protein profiles with maturation and relationship to coffee cup quality. Journal of agricultural and food chemistry, 51(8), 2328-2334.

Munyuli, T. (2011). Farmers' perceptions of pollinators' importance in coffee production in Uganda. Agricultural Sciences, 2(03), 318.

Oberthür, T., Läderach, P., Posada, H., Fisher, M.J., Samper, L.F., Illera, J., Collet, L., Moreno, E., Alarcón, R., Villegas, A., Usma, H., Perez, C., Jarvis, A., 2011. Regional

Ogden, C. B., Van Es, H. M., & Schindelbeck, R. R. (1997). Miniature rain simulator for field measurement of soil infiltration. Soil Science Society of America Journal, 61(4), 1041-1043.

Prichard, T. (2004). Imposing water deficits to improve wine quality and reduce costs. Avaiable at (verified 2 Jan. 2009). Univ. of California–Davis, Davis, CA. Imposing water deficits to improve wine quality and reduce costs. Avaiable at http://ucce. ucdavis. edu/files/filelibrary/2019/1564. pdf (verified 2 Jan. 2009). Univ. of California–Davis, Davis, CA.

QGIS Development Team, <YEAR>. QGIS Geographic Information System. Open Source Geospatial Foundation Project. http://www.qgis.org/

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

Rahman, S. (2003). Environmental impacts of modern agricultural technology diffusion in Bangladesh: an analysis of farmers' perceptions and their determinants. Journal of environmental management, 68(2), 183-191.

Ramírez, R., & Morales, D. (1989). Comparación de cuatro métodos de análisis del fósforo del suelo para estimar el requerimiento de P2O5 por el tomate (Lycopersicon esculentum). Agronomía Trop, 39(1-3), 79-94.

Reganold, J. P., Andrews, P. K., Reeve, J. R., Carpenter-Boggs, L., Schadt, C. W., Alldredge, J. R., ... & Zhou, J. (2010). Fruit and soil quality of organic and conventional strawberry agroecosystems. PLoS One, 5(9), e12346.

Riul, A., Soto, A. G., Mello, S. V., Bone, S., Taylor, D. M., & Mattoso, L. H. C. (2003). An electronic tongue using polypyrrole and polyaniline. Synthetic Metals, 132(2), 109-116.

Rodriguez-Suarez, J. A., Arias, M., Lopez, E., & Soto, B. (2007). Comparison of Multi-element to Single-Element Extractants for Macro-and Micronutrients in Acid Soils from Spain.Communications in soil science and plant analysis, 39(1-2), 231-240.

Romig, D. E., Garlynd, M. J., Harris, R. F., & McSweeney, K. (1995). How farmers assess soil health and quality. Journal of soil and water conservation, 50(3), 229-236.

Ryder, R. (2003). Local soil knowledge and site suitability evaluation in the Dominican Republic. Geoderma, 111(3), 289-305.

Schindelbeck, R. R., van Es, H. M., Abawi, G. S., Wolfe, D. W., Whitlow, T. L., Gugino, B. K., ... & Moebius-Clune, B. N. (2008). Comprehensive assessment of soil quality for landscape and urban management. Landscape and Urban Planning, 88(2), 73-80.

Silva, E. A. D., Mazzafera, P., Brunini, O., Sakai, E., Arruda, F. B., Mattoso, L. H. C., ... & Pires, R. C. M. (2005). The influence of water management and environmental conditions on the chemical composition and beverage quality of coffee beans. Brazilian Journal of Plant Physiology, 17(2), 229-238.

Topp, G.C., Y.T. Galganov, B.C. Ball, and M.R. Carter. (1993). Soil water desorption curves. In: M.R. Carter, editor, Soil sampling and methods of analysis. Can. Soc. Soil Sci., Lewis Publ., Boca Raton, FL. p. 569–579.

Vaast, P., Bertrand, B., Perriot, J.J., Guyot, B., Genard, M., 2006. Fruit thinning and shade improve bean characteristics and beverage quality of coffee (Coffea arabica L.) under optimal conditions. Journal of the Science of Food and Agriculture 86, 197-204

Vaast, P., Kanten, R.v., Siles, P., Dzib, B., Franck, N., Harmand, J., Genard, M., 2005. Shade: a key factor for coffee sustainability and quality, ASIC 2004. 20th Conference on Coffee Science, Bangalore, India, 11-15 October 2004. Association Scientifique Internationale du Café (ASIC), pp. 887-896.

Walker, J.M., (2002). The bicinchonic acid (BCA) assay for protein quantitation, in: Walker, J.M. (Ed.), The Protein Protocols Handbook. Humana Press, Totowa, NJ.

Walker, J. M. (2009). The bicinchoninic acid (BCA) assay for protein quantitation. The Protein Protocols Handbook, 11-15.

Weibel, F. P., Bickel, R., Leuthold, S., & Alföldi, T. (1998, August). Are organically grown apples tastier and healthier? A comparative field study using conventional and alternative methods to measure fruit quality. In XXV International Horticultural Congress, Part 7: Quality of Horticultural Products 517 (pp. 417-426).

Weil, R. R., Islam, K. R., Stine, M. A., Gruver, J. B., & Samson-Liebig, S. E. (2003). Estimating active carbon for soil quality assessment: A simplified method for laboratory and field use. American Journal of Alternative Agriculture, 18(01), 3-17.

Winston, E., Op de Laak, J., Marsh, T., Lempke, H., Chapman, K., Aung, O., & Nyunt, T. (2005). Arabica coffee manual for Myanmar. FAORAP.

Wright, S. F., & Upadhyaya, A. (1996). Extraction of an abundant and unusual protein from soil and comparison with hyphal protein of arbuscular mycorrhizal fungi. Soil Science, 161(9), 575-586.

APPENDIX

CHAPTER 1. UNDERSTANDING SOIL HEALTH AND ASSOCIATED FARMERS' PERCEPTIONS IN COLOMBIAN COFFEE SYSTEMS

To compute the averages and the standard deviations of indicator values and scores by municipality

```{r}
CompleteDataSet <- read.csv("C:/Users/Fatma Rekik/Desktop/My Research/CompleteDataSet.csv")
```

```{r}
subsetCauca<-CompleteDataSet[1:223,]
```

```{r}
#For Cauca (General)
summary(subsetCauca)
names(subsetCauca)
newdataCau<-subsetCauca[,c(14,16,19,22,25,28,31,34,37,40,42,44,46,49,50,51)]

names(newdataCau)
sapply(newdataCau, sd)

#For Morales
subsetMorales<-subsetCauca[subsetCauca$municipality=="Morales",]
summary(subsetMorales)
newdataMor<-subsetMorales[,c(14,16,19,22,25,28,31,34,37,40,42,44,46,49,50,51)]
sapply(newdataMor, sd)

#For Cajibio
subsetCajibio<-subsetCauca[subsetCauca$municipality=="Cajibio",]
newdataCaj<-subsetCajibio[,c(14,16,19,22,25,28,31,34,37,40,42,44,46,49,50,51)]
sapply(newdataCaj, sd)

#For Popayan
subsetPopayan<-subsetCauca[subsetCauca$municipality=="Popayán",]
summary(subsetPopayan)
newdataPop<-subsetPopayan[,c(14,16,19,22,25,28,31,34,37,40,42,44,46,49,50,51)]
sapply(newdataPop, sd)

#For Piendamo
subsetPiendamo<-subsetCauca[subsetCauca$municipality=="Piendamó",]
summary(subsetPiendamo)
```

```
newdataPien<-subsetPiendamo[,c(14,16,19,22,25,28,31,34,37,40,42,44,46,49,50,51)]
sapply(newdataPien, sd)

#For Rosas
subsetRosas<-subsetCauca[subsetCauca$municipality=="Rosas",]
summary(subsetRosas)
newdataRos<-subsetRosas[,c(14,16,19,22,25,28,31,34,37,40,42,44,46,49,50,51)]
sapply(newdataRos, sd)

#For Timbio
subsetTimbio<-subsetCauca[subsetCauca$municipality=="Timbio",]
summary(subsetTimbio)
newdataTimb<-subsetTimbio[,c(14,16,19,22,25,28,31,34,37,40,42,44,46,49,50,51)]
sapply(newdataTimb, sd)

For ANOVA
a1<-aov(OverQ~GROUP*GENDER.OF.PRODUCER*Municipality, data=newdata3)
summary(a1)
lsmeans(a1,pairwise~GENDER.OF.PRODUCER*Municipality)
lsmeans(a1,pairwise~GROUP*GENDER.OF.PRODUCER)
lsmeans(a1,pairwise~GROUP*Municipality)

For Boxplots
newdata3$Municipality<-droplevels(newdata3$Municipality)
boxplot(OverQ~Municipality,data=newdata3, main="Soil Health in different Municipalities")
means<-tapply(newdata3$OverQ, newdata3$Municipality, mean)
means
points(means, col="red")

summary(newdata3$GENDER.OF.PRODUCER)
newdata3$GENDER.OF.PRODUCER<-droplevels(newdata3$GENDER.OF.PRODUCER)
boxplot(OverQ~GENDER.OF.PRODUCER,data=newdata3, main="Soil Health by Gender of
Producer", na.omit(newdata3))
means2<-tapply(newdata3$OverQ, newdata3$GENDER.OF.PRODUCER, mean)
means2
points(means2,col="red")

newdata3$GROUP<-droplevels(newdata3$GROUP)
boxplot(OverQ~GROUP,data=newdata3, main="Soil Health of Perceived Fertile and Non
Fertile Soils")
means3<-tapply(newdata3$OverQ, newdata3$GROUP, mean)
means3
points(means3,col="red")


newdata3$Status<-droplevels(newdata3$Status)
```

```
boxplot(OverQ~Status,data=PerceptionData, main="Soil Health of Farms Associated and not Associated with Local Cooperative")
means4<-tapply(PerceptionData$OverQ, PerceptionData$Status, mean)
means4
points(means4,col="red")

For Table1.2
subsetCauca<-CompleteDataSet[CompleteDataSet$department=="Cauca",]
install.packages("lsmeans")
library(lsmeans)
cld(lsmeans(lm(WAS~municipality,data = subsetCauca), pairwise~municipality))
cld(lsmeans(lm(AWC~municipality,data = subsetCauca), pairwise~municipality))
cld(lsmeans(lm(OM~municipality,data = subsetCauca), pairwise~municipality))
cld(lsmeans(lm(AC~municipality,data = subsetCauca), pairwise~municipality))
cld(lsmeans(lm(protein~municipality,data = subsetCauca), pairwise~municipality))
cld(lsmeans(lm(respiration~municipality,data = subsetCauca), pairwise~municipality))
cld(lsmeans(lm(pH~municipality,data = subsetCauca), pairwise~municipality))
cld(lsmeans(lm(P~municipality,data = subsetCauca), pairwise~municipality))
cld(lsmeans(lm(K~municipality,data = subsetCauca), pairwise~municipality))
cld(lsmeans(lm(Mg~municipality,data = subsetCauca), pairwise~municipality))
cld(lsmeans(lm(Fe~municipality,data = subsetCauca), pairwise~municipality))
cld(lsmeans(lm(Mn~municipality,data = subsetCauca), pairwise~municipality))
cld(lsmeans(lm(Zn~municipality,data = subsetCauca), pairwise~municipality))
cld(lsmeans(lm(OverQ~municipality,data = subsetCauca), pairwise~municipality))

For Table 1.4
coopmembers<-subsetCauca[subsetCauca$bdb=="",]
cld(lsmeans(lm(OverQ~municipality,data = coopmembers), pairwise~municipality))

The PCA to narrow down the indicators to those with high variance and affect farmers ranking the most
```{r}
library(FactoMineR)

newdata <- subsetCauca[c(14,16,19,22,25,28,31,34,37,40,42,44,46)]
names(newdata)
perception<-newdata[68:223,]

pca1 = PCA(perception,graph=TRUE)
summary(pca1)
pca1$var$coord

#Testing prcomp
pr<-prcomp(perception, center = TRUE, scale. = TRUE)
print(pr)
plot(pr, type = "l")
```
```

```
plot(PC1%*%PC2)
summary(pr)
```
The Logistical Regression Function for determining which indicators influence farmers' ranking of A (fertile) Vs. B (non-fertile)
```{r}
PercAndFertt <- read.csv("D:/Research/PercAndFertt.csv")

logit<-glm(FERTILITY.RANKING~OM+AC+AWC+RESPIRATION+P+WAS+PROTEIN, data = PercAndFertt, family = "binomial")

summary(logit)
```

The Perception accuracy study based on farmers' gender and their municipality
```{r}
Perception_Gender=matrix(c(37,13,19,6),nrow=2, ncol=2, byrow=TRUE)
fisher.test(Perception_Gender)
chisq.test(Perception_Gender)

Perception_Municipality=matrix(c(11,2,10,3,11,2,9,4,8,5,9,4),nrow=6, ncol=2, byrow=TRUE)
fisher.test(Perception_Municipality)
chisq.test(Perception_Municipality)
```

CHAPTER 2. CHARACTERIZATION OF SOIL HEALTH STATUS ON COLOMBIAN COFFEE FARMS

#For Doing the Shapiro-Wilk Test
```{r}
attach(subsetCauca)
shapiro.test(WAS)
shapiro.test(AWC)
shapiro.test(OM)
shapiro.test(AC)
shapiro.test(protein)
shapiro.test(respiration)
shapiro.test(pH)
shapiro.test(P)
shapiro.test(K)
shapiro.test(Mg)
shapiro.test(Fe)
shapiro.test(Mn)
shapiro.test(Zn)
```

```{r}
# Composite graph: 15 figures arranged in 3 rows and 5 columns
attach(subsetCauca)
layout(matrix(c(1,2,3,4,5,6,7,8,9,10,11,11,12,13,13),3,5, byrow = TRUE))
#WAS
hist(WAS, breaks=30, freq = FALSE, main = "WAS")
lines(density(WAS), col="blue", lwd=1)
legend("topleft",bty = "n", c("m= 94.3", "s= 8.3", "p <0.01"))
#AWC
hist(AWC, breaks=20, freq = FALSE, main = "AWC")
lines(density(AWC), col="blue", lwd=1)
legend("topright",bty = "n", c("m= -1.4", "s= 0.4", "p< 0.01"))
#OM
hist(OM,breaks=30, freq = FALSE, main = "OM")
lines(density(OM), col="blue", lwd=1)
legend("topright",bty = "n", cex = 0.8, c("m= 17.4", "s= 6.0", "p <0.01"))
#AC
hist(AC, breaks=30, freq = FALSE, main = "AC")
lines(density(AC), col="blue", lwd=1)
legend("topleft",bty = "n",cex = 0.85, c("m= 818.3", "s= 237.7", "p <0.01"))
#Protein
hist(protein, breaks=30, freq = FALSE, main = "Protein")
lines(density(protein), col="blue", lwd=1)
legend("topright",bty = "n", c("m= 9.2", "s= 2.5", "p= 0.82"))
#respiration
hist(respiration, breaks=30, freq = FALSE, main = "Respiration")
lines(density(respiration), col="blue", lwd=1)
legend("topright",bty = "n",cex = 0.87, c("m= 1", "s= 0.2", "p= 0.15"))
#pH
hist(pH, breaks=30, freq = FALSE, main = "pH")
lines(density(pH), col="blue", lwd=1)
legend("topright",bty = "n", c("m= 4.8", "s= 0.3", "p= 0.09"))
#P
hist(P, breaks=30, freq = FALSE, main = "P")
lines(density(P), col="blue", lwd=1)
legend("topright", bty = "n", c("m= 9.6", "s= 4.3", "p <0.01"))
#K
hist(K, breaks=20, freq = FALSE, main = "K")
lines(density(K), col="blue", lwd=1)
legend("topright", bty = "n", c("m= 4.6", "s= 0.6", "p <0.01"))
#Mg
hist(Mg, breaks=40, freq = FALSE, main = "Mg")
lines(density(Mg), col="blue", lwd=1)
legend("topright", bty = "n", c("m= 4.2", "s= 1.3", "p <0.01"))
#Fe
```

```
hist(Fe, breaks=30, freq = FALSE, main = "Fe")
lines(density(Fe), col="blue", lwd=1)
legend("topright", bty = "n", c("m= 2.9", "s= 0.4", "p <0.01"))
#Mn
hist(Mn, breaks=20, freq = FALSE, main = "Mn")
lines(density(Mn), col="blue", lwd=1)
legend("topright", bty = "n", c("m= 1.6", "s= 0.7", "p<0.01"))
#Zn
hist(Zn, breaks=30, freq = FALSE, main = "Zn")
lines(density(Zn), col="blue", lwd=1)
legend("topright",bty = "n", c("m= -0.4", "s= 0.8", "p<0.01"))
```

#PCA for the New (Weighted) Soil HEalth Scoring Framework for Colombian Sampes

```{r}
Weight<- subsetCauca[c(14,16,19,22,25,28,31,34,37,40,42,44,46,50,51)]
names(Weight)
library(FactoMineR)
pca.weighted = PCA(Weight,graph=TRUE)
summary(pca.weighted)
pca.weighted$var$coord
```

#Optimal Range Piecewise functions

```{r}
#pH
x<-c(2,4.5,4.6,4.7,4.8,4.9,5,5.1,5.2,5.3,5.4,5.5,6.5,6.6,6.7,6.8,6.9,7,7.1,7.2,7.3,7.4,7.5,7.6,7.7,10)
y<-c(0,0,17,33,47,59,70,79,87,93,97,100,100,98,96,92,86,80,72,63,53,42,29,15,0,0)
plot(x,y, xlab="pH", ylab = "Score")
lines(x,y)

```

Composite graph for A-CSHSF scoring curves: 15 figures arranged in 3 rows and 5 columns

```{r}
attach(subsetCauca)
layout(matrix(c(1,2,3,4,5,6,7,8,9,10,11,11,12,13,13),3,5, byrow = TRUE))
#WAS
WAS<-c(0:100)
score1<-c(0:100)
plot(WAS, score1, type="l", xlab="WAS", ylab= "score")
#AWC
AWC<-c(0,0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8)
score2f<-c(0.62,10.56,50,89.44,99.38,99.99,100,100,100)
score2m<-c(0.13,2.28,15.87,50,84.13,97.72,99.87,99.99,100)
plot(AWC, score2f, type="l", xlab="AWC", ylab= "score", col="red")
```

```
lines(AWC,score2m,col="blue")
legend("bottomright",c("fine","medium"),lty=c(1,1),lwd=c(1,1),col = c("red","blue"),cex = 1,bty
= "n")

#OM
OM<-c(5,10,15,20,25,30,35,40,45,40)
score3f<-c(1.93,10.84,34.4,66.7,89.7,98.2,99.8,99.991,99.999,100)
score3f<-c(4.68,20.60,51.43,81.38,95.98,99.54,99.97,99.99,100,100)
score3m<-c(0.64,5.53,24.16,57.66,86.16,97.63,99.80,99.99,100,100)
plot(OM, score3f, type="l", xlab="OM", ylab= "score", col="red")
lines(OM,score3m, col="blue")
legend("bottomright",c("fine","medium"),lty=c(1,1),lwd=c(1,1),col = c("red","blue"),cex = 1,
bty = "n")

#AC
AC<-c(300,500,700,900,1100,1300,1500,1900,2100)
score4f<-c(5.12,21.40,51.90,81.28,95.81,99.49,99.97,99.99,100)
score4m<-c(0.18,2.67,17.13,51.32,84.50,97.71,99.86,99.99,100)
plot(AC, score4f, type="l", xlab="AC", ylab= "score", col="red")
lines(AC,score4m, col="blue")
legend("bottomright",c("fine","medium"),lty=c(1,1),lwd=c(1,1),col = c("red","blue"),cex = 1,
bty = "n")

#Protein
prot<-c(2,4,6,8,10,12,14,16,22,24)
score5<-c(0.19,1.8,9.75,30.99,61.94,86.52,97.15,99.66,99.99,100)
plot(prot, score5, type="l", xlab="protein", ylab= "score")

#respiration
resp<-c(0.2,0.4,0.6,0.8,1,1.2,1.4,1.6,2.2,2.4)
score6f<-c(0.13,1.39,8.08,27.43,57.93,84.13,96.41,99.53,99.99,100)
score6m<-c(0.03,0.43,3.66,16.89,45.03,76.06,93.84,99.12,99.99,100)
plot(resp, score6f, type="l", xlab="respiration", ylab= "score", col="red")
lines(resp,score6m, col="blue")
legend("bottomright",c("fine","medium"),lty=c(1,1),lwd=c(1,1),col = c("red","blue"),cex = 1,
bty = "n")

#pH
pH<-c(4.5,4.7,4.9,5.2,5.5,6.5,6.7,6.9,7.1,7.3,7.5,7.7)
score7<-c(0,45.22,48.4,53.19,100,100,54.49,52.25,50,47.75,45.51,0)
plot(pH, score7, type="l", xlab="pH", ylab= "score")
#P
P<-c(2,4,6,8,10,14,18,22,26,34)
score8<-c(3.68,9.35,10.77,35.19,53.55,84.81,97.54,99.82,99.99,100)
plot(P, score8, type="l", xlab="P", ylab= "score")
#K
```

```
K<-c(20,40,60,80,100,120,140,160,180,300)
score9<-c(6.68,15.87,30.85,50,69.15,84.13,93.32,97.72,99.38,100)
plot(K, score9, type="l", xlab="K", ylab= "score")
#Mg
Mg<-c(10,20,40,60,80,100,120,140,200,220)
score10<-c(5.09,10.16,29.27,57.21,81.83,94.91,99.10,99.90,99.99,100)
plot(Mg, score10, type="l", xlab="Mg", ylab= "score")
#Fe
Fe<-c(2,4,6,10,13,33,50,70,90,150,200,300)
score11<-c(8.31,22.09,43.89,85.92,100,100,90.61,85.18,77.97,48.09,23.26,1.81)
plot(Fe, score11, type="l", xlab="Fe", ylab= "score")
#Mn
Mn<-c(1,2,3,4,5,6,7,8)
score12<-c(93.32,84.13,69.15,50,30.85,15.87,6.68,0)
plot(Mn, score12, type="l", xlab="Mn", ylab= "score")
#Zn
Zn<-c(0.01,0.03,0.05,0.07,0.09,0.1,0.6,2,4,6,8)
score13<-c(5.48,21.19,50,78.81,94.52,100,100,85.76,55.55,21.46,4.26)
plot(Zn, score13, type="l", xlab="Zn", ylab= "score")
```

To test whether medium versus fine is significant for each indicator

```{r}
#First, we extract fine and medium rows to become subsets
fine<-subset(subsetCauca,TextureGroup=="fine")
medium<-subset(subsetCauca, TextureGroup=="medium")

#Now we do a t.test for each indicator
t.test(fine$WAS,medium$WAS)#p=0.8
t.test(fine$AWC,medium$AWC)#p=1.2e-14
t.test(fine$OM,medium$OM)#p=7.95e-7
t.test(fine$AC,medium$AC)#p=1.4e-09
t.test(fine$protein,medium$protein)#p=0.09
t.test(fine$respiration,medium$respiration)#p=0.03
t.test(fine$pH,medium$pH)#p=1.12e-5
t.test(fine$P,medium$P)#p=0.07
t.test(fine$K,medium$K)#p=0.13
t.test(fine$Mg,medium$Mg)#p=0.4
t.test(fine$Fe,medium$Fe)#p=0.03
t.test(fine$Mn,medium$Mn)#p=6.0e-5
t.test(fine$Zn,medium$Zn)#p=0.4

lm1<-lm(WAS~TextureGroup, data=subset(subsetCauca, TextureGroup!="coarse"))
summary(lm1)

#To give letter difference annotations
```

```
library(multcompView)
library(lsmeans)
lsmeans(lm1, pairwise~TextureGroup)
cld(lsmeans(lm1, ~TextureGroup))


lm2<-lm(AWC~TextureGroup, data=subset(subsetCauca, TextureGroup!="coarse"))
summary(lm2)
lsmeans(lm2, pairwise~TextureGroup)
cld(lsmeans(lm2, ~TextureGroup))
```

#Best Subsets Regression
```{r}
Colombian.Soil.health.scores.all.4.types <- read.csv("C:/Users/Fatma Rekik/Dropbox/My Research/2nd Manuscript/Colombian Soil health scores all 4 types.csv")

library(leaps)
#Using CASH
data=subsetCauca
BSR<-
regsubsets(OverQ.Cor~Sand+Silt+Clay+WAS+AWC+AC+OM+protein+respiration+pH+P+K+Mg+Fe+Mn+Zn, nbest = 2, method = "exhaustive",data=subsetCauca)
summary(BSR)
plot(BSR, scale="r2")

#Using A-CSHSF
data=Colombian.Soil.health.scores.all.4.types
subsets<-
regsubsets(OverQ.Col~WAS+AWC+AC+OM+protein+respiration+pH+P+K+Mg+Fe+Mn+Zn, nbest = 2, method = "exhaustive",data=Colombian.Soil.health.scores.all.4.types)
summary(subsets)
plot(subsets, scale="r2")

summary(lm(OverQ~AC, data=subsetCauca))
```

#checking if scoring functions yield different results

#Interclass Correlation
```{r}
ICC <- read.csv("C:/Users/Fatma Rekik/Dropbox/My Research/2nd Manuscript/ICC.csv")
library(irr)

ICCWAS<-ICC[,c(1,2,3)]
icc(ICCWAS,model = "oneway", type = "agreement")
```

```
ICCAWC<-ICC[,c(4,5,6)]
icc(ICCAWC,model = "oneway", type = "agreement")

ICCOM<-ICC[,c(7,8,9)]
icc(ICCOM,model = "oneway", type = "agreement")

ICCAC<-ICC[,c(10,11,12)]
icc(ICCAC,model = "oneway", type = "agreement")

ICCProt<-ICC[,c(13,14)]
icc(ICCProt,model = "oneway", type = "agreement")

ICCResp<-ICC[,c(15,16)]
icc(ICCResp,model = "oneway", type = "agreement")

ICCpH<-ICC[,c(17,18)]
icc(ICCpH,model = "oneway", type = "agreement")

ICCP<-ICC[,c(19,20)]
icc(ICCP,model = "oneway", type = "agreement")

ICCK<-ICC[,c(21,22)]
icc(ICCK,model = "oneway", type = "agreement")

ICCMinor<-ICC[,c(23,24)]
icc(ICCMinor,model = "oneway", type = "agreement")

ICCOverQ<-ICC[,c(25,26,27)]
icc(ICCOverQ,model = "oneway", type = "agreement")
```

CHAPTER 3. SOIL HEALTH; A KEY DETERMINANT OF COFFEE CUP QUALITY

```
#For Description Statistics table
summary(Coffee.Quality.to.SH)
sapply(Coffee.Quality.to.SH, sd)


#For Morales
summary(Coffee.Quality.to.SH[Coffee.Quality.to.SH$City=="Morales",])
MoralesSubset<-Coffee.Quality.to.SH[Coffee.Quality.to.SH$City=="Morales",]
sapply(MoralesSubset, sd)

#For Cajibio
subsetCajibio<-subsetCauca[subsetCauca$municipality=="Cajibio",]
```

```
sapply(subsetCajibio, sd)

#For Popayan
subsetPopayan<-subsetCauca[subsetCauca$municipality=="Popayán",]
summary(subsetPopayan)
newdataPop<-subsetPopayan[,c(14,16,19,22,25,28,31,34,37,40,42,44,46,49,50,51)]
sapply(newdataPop, sd)

#For Piendamo
subsetPiendamo<-subsetCauca[subsetCauca$municipality=="Piendamó",]
summary(subsetPiendamo)
newdataPien<-subsetPiendamo[,c(14,16,19,22,25,28,31,34,37,40,42,44,46,49,50,51)]
sapply(newdataPien, sd)

#For Rosas
subsetRosas<-subsetCauca[subsetCauca$municipality=="Rosas",]
summary(subsetRosas)
newdataRos<-subsetRosas[,c(14,16,19,22,25,28,31,34,37,40,42,44,46,49,50,51)]
sapply(newdataRos, sd)

#For Timbio
subsetTimbio<-subsetCauca[subsetCauca$municipality=="Timbio",]
summary(subsetTimbio)
newdataTimb<-subsetTimbio[,c(14,16,19,22,25,28,31,34,37,40,42,44,46,49,50,51)]
sapply(newdataTimb, sd)

#Let's see if cup quality is different by municipality and coop association
```{r}
#1. Municipality
boxplot(CoffeQuality2014~City, main="Municipality Cup Quality Differences",ylab="Cup
Quality",data = Coffee.Quality.to.SH)

means<-tapply(Coffee.Quality.to.SH$CoffeQuality2014, Coffee.Quality.to.SH$City, mean)
means
points(means, col="red")

library(lsmeans)

#Difference in Cup Quality
cld(lsmeans(lm(CoffeQuality2014~City,data = Coffee.Quality.to.SH), pairwise~City))#not
significantly different

#Difference in Overall SH Score
cld(lsmeans(lm(OverQ.Col~City,data = Coffee.Quality.to.SH), pairwise~City))

#Difference in WAS
```
```

```
cld(lsmeans(lm(WAS~City,data = Coffee.Quality.to.SH), pairwise~City))

#Difference in AWC
cld(lsmeans(lm(AWC~City,data = Coffee.Quality.to.SH), pairwise~City))

#Difference in OM
cld(lsmeans(lm(OM~City,data = Coffee.Quality.to.SH), pairwise~City))

#Difference in AC
cld(lsmeans(lm(AC~City,data = Coffee.Quality.to.SH), pairwise~City))

#Difference in Prot
cld(lsmeans(lm(Prot~City,data = Coffee.Quality.to.SH), pairwise~City))

#Difference in Resp
cld(lsmeans(lm(Resp~City,data = Coffee.Quality.to.SH), pairwise~City))

#Difference in pH
cld(lsmeans(lm(pH~City,data = Coffee.Quality.to.SH), pairwise~City))

#Difference in P
cld(lsmeans(lm(P~City,data = Coffee.Quality.to.SH), pairwise~City))

#Difference in K
cld(lsmeans(lm(K~City,data = Coffee.Quality.to.SH), pairwise~City))

#Difference in Mg
cld(lsmeans(lm(Mg~City,data = Coffee.Quality.to.SH), pairwise~City))

#Difference in Fe
cld(lsmeans(lm(Fe~City,data = Coffee.Quality.to.SH), pairwise~City))

#Difference in Mn
cld(lsmeans(lm(Mn~City,data = Coffee.Quality.to.SH), pairwise~City))

#Difference in Zn
cld(lsmeans(lm(Zn~City,data = Coffee.Quality.to.SH), pairwise~City))

#Difference in Clay
cld(lsmeans(lm(Clay~City,data = Coffee.Quality.to.SH), pairwise~City))

#Difference in Elevation
cld(lsmeans(lm(Elevation~City,data = Coffee.Quality.to.SH), pairwise~City))

#Difference in NumTrees
cld(lsmeans(lm(NumTrees~City,data = Coffee.Quality.to.SH), pairwise~City))
```

#Is coffee quality different between men and women?
cld(lsmeans(lm(CoffeQuality2014~Gender,data = Coffee.Quality.to.SH), pairwise~Gender))

Full Linear Model
```{r}
summary(lm(CoffeQuality2014~WAS+ AWC+ AC+ OM+ Prot+
Resp+pH+P+K+Mg+Mn+Fe+Zn+NumTrees+Elevation, data = Coffee.Quality.to.SH ))
```

Best Subsets Regression For Dimension Reduction
```{r}
library(leaps)

summary(regsubsets(CoffeQuality2014~WAS+ AWC+ AC+ OM+ Prot+
Resp+pH+P+K+Mg+Mn+Fe+Zn+NumTrees+Elevation, nbest = 1, method =
"exhaustive",nvmax=5,data=Coffee.Quality.to.SH))

plot(regsubsets(CoffeQuality2014~WAS+ AWC+ AC+ OM+ Prot+
Resp+pH+P+K+Mg+Mn+Fe+Zn+NumTrees+Elevation, nbest = 1, method =
"exhaustive",nvmax=5,data=Coffee.Quality.to.SH),scale="r2")
```

Reduced Model
```{r}
summary(lm(CoffeQuality2014~OM+Prot+ Resp+pH+Fe, data = Coffee.Quality.to.SH ))
```

Correlation Matrix

```{r}
library(Hmisc)
rcorr(as.matrix(Coffee.Quality.to.SH[c(29:44,47,49)]))
```

```{r}
library(leaps)

#Aroma
plot(regsubsets(Aroma~WAS+AWC+OM+AC+Prot+pH+P+K+Mg+Fe+Mn+Zn+Sand+Elevatio
n+NumTrees, nbest = 1, method =
"exhaustive",nvmax=3,data=Coffee.Quality.to.SH),scale="r2")

#Flavor
```

```
plot(regsubsets(Flavor~WAS+AWC+OM+AC+Prot+pH+P+K+Mg+Fe+Mn+Zn+Sand+Elevatio
n+NumTrees, nbest = 1, method =
"exhaustive",nvmax=3,data=Coffee.Quality.to.SH),scale="r2")

#Aftertaste
plot(regsubsets(Aftertaste~WAS+AWC+OM+AC+Prot+pH+P+K+Mg+Fe+Mn+Zn+Sand+Elev
ation+NumTrees, nbest = 1, method =
"exhaustive",nvmax=3,data=Coffee.Quality.to.SH),scale="r2")

#Acidity
plot(regsubsets(Acidity~WAS+AWC+OM+AC+Prot+pH+P+K+Mg+Fe+Mn+Zn+Sand+Elevati
on+NumTrees, nbest = 1, method =
"exhaustive",nvmax=3,data=Coffee.Quality.to.SH),scale="r2")

#Body
plot(regsubsets(Body~WAS+AWC+OM+AC+Prot+pH+P+K+Mg+Fe+Mn+Zn+Sand+Elevation
+NumTrees, nbest = 1, method = "exhaustive",nvmax=3,data=Coffee.Quality.to.SH),scale="r2")

#Uniformity
plot(regsubsets(Uniformity~WAS+AWC+OM+AC+Prot+pH+P+K+Mg+Fe+Mn+Zn+Sand+Ele
vation+NumTrees, nbest = 1, method =
"exhaustive",nvmax=3,data=Coffee.Quality.to.SH),scale="r2")

#Sweetness
plot(regsubsets(Sweetness~WAS+AWC+OM+AC+Prot+pH+P+K+Mg+Fe+Mn+Zn+Sand+Elev
ation+NumTrees, nbest = 1, method =
"exhaustive",nvmax=3,data=Coffee.Quality.to.SH),scale="r2")

#Bowl.cleaning
plot(regsubsets(Bowl.clenaing~WAS+AWC+OM+AC+Prot+pH+P+K+Mg+Fe+Mn+Zn+Sand+
Elevation+NumTrees, nbest = 1, method =
"exhaustive",nvmax=3,data=Coffee.Quality.to.SH),scale="r2")

#Balance
plot(regsubsets(Balance~WAS+AWC+OM+AC+Prot+pH+P+K+Mg+Fe+Mn+Zn+Clay+Elevati
on+NumTrees, nbest = 1, method =
"exhaustive",nvmax=3,data=Coffee.Quality.to.SH),scale="r2")
```