

SMALLHOLDER COFFEE TERROIRS OF RWANDA
LINKING LOCAL TO GLOBAL TRADE THROUGH
GEOGRAPHICAL CHARACTERIZATION

Adrie Mukashema

Graduation committee:

Chairman/secretary

Dean ITC

Supervisor

Prof.dr.ir. A. Veldkamp

University of Twente

Members

Prof.dr.ir. A. Stein

University of Twente

Prof.dr. V.G. Jetten

University of Twente

Prof.dr.ir. A. Verdoodt

Ghent University

Prof.dr. T. Dawson

King's College London, UK

Prof.dr.ir. A.K. Bregt

Wageningen University

ITC dissertation number 310

ITC, P.O. Box 217, 7500 AE Enschede, The Netherlands

ISBN 978-90-365-4409-2

DOI 10.3990/1.9789036544092

Cover designed by Benno Masselink

Printed by ITC Printing Department

Copyright © 2017 by Adrie Mukashema



UNIVERSITY OF TWENTE.

ITC

FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

SMALL HOLDER COFFEE TERROIRS OF RWANDA
LINKING LOCAL TO GLOBAL TRADE THROUGH
GEOGRAPHICAL CHARACTERIZATION

DISSERTATION

to obtain
the degree of doctor at the University of Twente,
on the authority of the rector magnificus,
prof.dr. T.T.M. Palstra,
on account of the decision of the graduation committee,
to be publicly defended
on Thursday 19 October 2017 at 14:45 hrs

by

Adrie Mukashema
born on 24 December 1972
in Nyaruguru, Rwanda

This thesis has been approved by
Prof.dr.ir. A. Veldkamp, supervisor

To my parents,
I wish you were around on this earth to witness about your big dream to your
daughter which now becomes true!

Acknowledgements

I would not have been able to start and complete this piece of work without support and commitment of different individuals and organizations that I would like to express my sincere thanks for their valuable contributions.

First and foremost, I would like to express my sincere gratitude to the Kingdom of the Netherlands through its Netherland fellowship Program (NFP) for higher education. Without the NFP financial support, my PhD study would not have been possible. Special thanks go Mrs. Hélène Bernot Ullerø, from NUFFIC who was the first to remind me that it was time to go for PhD study, and encouraged me to apply for NFP scholarship. Mrs. Hélène did not stop from application and award, but she kept on following my progress, and her encouragement especially during my qualifier period feed me with hope and determination. I am also thankful to the National University of Rwanda, my former employer to have strongly supported my application to NFP and granted me study leave and all associated benefits to be able to pursue my study without any problem. A note of thanks is addressed to Emmanuel Twarabamenye, professor at University of Rwanda for his constant encouragement throughout my PhD studies.

I would like to express my sincere and heartfelt gratitude to my promoter and supervisor Tom Veldkamp who has been a constant source of support and inspiration throughout my PhD journey. Tom took over from Eric and David Rossiter after the qualifier period. He has been a perfect mentor during my PhD journey, regardless his present heavy administration workload as Dean of ITC. I am also thankful to Professor Eric Smaling, my supervisor during the initial stage of my PhD study. I keep memories of my qualifier period that challenged me when I wanted to take-off this segment of my growth path. During this time, Eric played an amazing role, lifted me when I was almost giving up, but then he left me where I needed his hand the most. Surprisingly, before he left ITC for other duties, Eric made sure he leaves me in the right hands. Since from that time, my PhD journey became a worthy path to explore and complete.

This thesis would not have been possible without support of the National Agricultural Export development Board of Rwanda (NAEB). I would like to sincerely thank Alex Kanyankore, former Director General of NAEB, currently CEO of the Development Bank of Rwanda (BRD) and Celestin M. Gatarayiha for their commitment to provide all available data I required to write this thesis. Coffee yield and sensory quality data used in this study were provided by NAEB. Beside data, they also spent their time reviewing and validating the research outputs. With this support, there is no doubt that my research recommendations shall also be considered in future developments of coffee sector in Rwanda. I am indebted to NAEB technical staff Potien, Maurice,

Malliavin and Eric for organizing various data I acquired from their respective offices. High resolution map of coffee fields would not have been possible without availability of very high resolution orthophotos covering the entire Rwanda. I'm thankful to Government of Rwanda through Rwanda Natural Resource Authority (RNRA) to have availed this dataset for research and education purpose at the University of Rwanda from which I benefited.

Special thanks go to ITC staff: Sherif Amer, Marga Koelen, Frans Gollenbeek, Anton Vrieling, and Tom Loran, for their unvarying support to my PhD study. Sherif and Anton specifically contributed also to my scientific outputs. Their guidance and inputs tremendously improved the quality of my research articles. I cannot find words to express my gratitude to Sherif for being close throughout my study at ITC. He took care of any obstacle arising on my way until I complete this PhD. I am grateful for the great administration support I received from PhD management staff especially Paul Van Dijk, Loes Colenbrander and Marion Pierik, to my promoter's administration assistant Marja Verburg, and to the Natural resource department assistant Esther Hondebrink. Many thanks to ITC helpdesk and library unit for solving all IT related problems I faced along with my study at ITC. Thanks to ITC hotel staff especially Saskia and Patrick for their regular warm welcome many times I visited ITC.

Many thanks to Rwandan ITC students, PhD and Masters for the period 2010-2017; they made Enschede a second home, away from home. A special note of gratitude to Alphonse, Coco, Clarisse, Christine, Elias, Christine and Vestine; we shared good and challenging moment of our long PhD journey in the Netherlands. Office 4010 of ITC was a ladies' room full of life and inspiration. Special thanks to my officemates Claudia, Sisi, Tyas, Maitreyi, Anahita and Zhihui for the support and friendship. I will always remember good time we spent together, lunches, tea and coffee breaks we had together and experience we shared during our stay at ITC. Almino and Abel, we crossed in Enschede as they were almost at the end of their PhD journey, but the little time we spent together gave me courage that I will also make it.

I would like to extend my special gratitude to my family and friends in Rwanda and in Europe for always being on my side through various support including prayers. Many thanks go to my husband for his endless patience, and care. He took full control of our family, without whom, I would not have found strength to complete this PhD. Finally, dear my son Jackson and my daughter Liliane, thank you indeed for your patience, I know you have been very often deprived your mother's full care since 1998, when I decided to resume my very long study journey, you were little babies. 20 years are now gone but brighter years are ahead; you have also started your university journey; I wish you both the very best and success with your studies.

Table of Contents

Acknowledgements	i
Table of Contents.....	iii
List of figures	vi
List of tables.....	x
1 Introduction	1
1.1 Coffee provenance: “a novel business model” for coffee producing countries	2
1.2 Overview of the coffee sector in Rwanda	3
1.2.1 Geographic location and landscapes	3
1.2.2 Coffee sector governance	5
1.2.3 Current coffee production and quality trends	6
1.2.4 Link between coffee product and territory	7
1.2.4.1 Climate	7
1.2.4.2 Soils	8
1.2.4.3 Topography	9
1.2.4.4 Management of coffee farms	9
1.2.4.5 Coffee harvesting and post-harvesting processing chain	10
1.3 Challenges in zoning coffee terroirs	10
1.4 Research objectives.....	12
1.5 Thesis outline.....	13
2 Automated high resolution mapping of coffee in Rwanda using an expert Bayesian network.....	15
2.1 Introduction.....	16
2.2 Materials and methods.....	18
2.2.1 Data and variables	18
2.2.1.1 Remote sensing data.....	18
2.2.1.2 Evidence variables for coffee.....	18
2.2.1.3 Training and validation dataset.....	19
2.2.2 Preliminary spectral analysis	19
2.2.3 Coffee classification with a Bayesian network.....	20
2.2.3.1 Principles and definitions	20
2.2.3.2 Building an expert Bayesian network classifier.....	21
2.2.4 Country-wide implementation of the expert Bayesian network	26
2.3 Results.....	26
2.3.1 Bayesian network results for a test area	26
2.3.2 Accuracy of the Rwanda coffee map	29
2.4 Discussion	32
2.4.1 Effect of individual Bayesian steps on the scale and the quality of the final coffee map	32
2.4.2 Performance of the expert Bayesian network.....	33
2.5 Conclusions and outlook.....	34
3 Sixty percent of small coffee farms have suitable socio-economic and environmental locations in Rwanda	37
3.1 Introduction.....	38
3.1.1 Historical background of coffee distribution in Rwanda	39
3.1.2 Characteristics of coffee locations in Rwanda	40
3.2 Materials and methods.....	41

3.2.1	Biophysical and socio-economic data	41
3.2.2	Statistical analysis.....	42
3.2.2.1	Extracting biophysical and socio-economic data.....	42
3.2.2.2	Determining location drivers of coffee fields	42
3.2.2.3	Explaining location variability of coffee fields.....	43
3.2.2.4	Predicting suitable areas for growing coffee	43
3.2.3	Model validation.....	43
3.3	Results and Discussion	44
3.3.1	Location factors of coffee patterns.....	44
3.3.2	Explaining the current spatial distribution of coffee areas	46
3.3.2.1	Generalised national model.....	46
3.3.2.2	Regionalized agroecological zone (AEZ) specific models.....	46
3.3.3	Analysing potential locations for coffee expansion	50
3.3.4	Exploring spatial coffee distribution	51
3.3.4.1	Role of environmental and socio-economic factors in coffee field density	51
3.3.4.2	Modelling coffee distribution.....	53
3.3.4.3	Potential areas for future coffee expansion.....	53
3.4	Conclusion.....	54
4	Arabica coffee yields in Rwandan smallholder systems: future perspectives of climate change effect and improved soil organic carbon management	57
4.1	Introduction.....	58
4.2	Material and Methods.....	60
4.2.1	Coffee production and yield data.....	60
4.2.2	Biophysical properties of coffee farms.....	63
4.2.2.1	Climate	64
4.2.2.2	Topography	64
4.2.2.3	Soils	65
4.2.3	Data processing and statistical analysis	66
4.3	Results.....	68
4.3.1	Biophysical characteristics of coffee farms and yield levels	68
4.3.2	Factors determining of coffee yield trends in current farming systems	70
4.3.3	Modeling achievable coffee yields.....	74
4.4	Discussion	76
4.4.1	Coffee yield trends and biophysical relationships.....	76
4.4.2	Impact of climate change.....	77
4.4.3	Impact of improved land management.....	80
4.5	Conclusions	81
5	Understanding geographical origins of sensory quality for specialty coffees of Rwanda	83
5.1	Introduction.....	84
5.2	Materials and Methods	86
5.2.1	Coffee production and post harvesting processing chain.....	86
5.2.2	Coffee sample preparation and sensory test	87
5.2.3	Explaining the variation of coffee sensory quality	90
5.2.3.1	Biophysical data and coffee sensory data.....	90
5.2.3.2	Statistical analysis	91
5.3	Results.....	93

5.3.1	Spatial variability of Arabica coffee sensory in Rwanda	93
5.3.2	Biophysical factors of coffee sensory variation	95
5.3.3	Geographic areas for special coffee production in Rwanda	98
5.4	Discussion	99
5.4.1	Effect of soil fertility management on coffee sensory	99
5.4.2	Impact of climate change on coffee quality.....	100
5.4.3	Coffee quality and production trade-off	103
5.5	Conclusions	105
6	Synthesis: Integrating land suitability, productivity and sensory quality in 'Coffee terroirs' for sustainable coffee production systems in Rwanda....	107
6.1	Introduction.....	108
6.2	Identifying coffee terroir zones and geographical indicators of origin 111	
6.2.1	Conceptual model of coffee terroir.....	111
6.2.2	Spatial distribution of coffee fields, location and size of the farm 113	
6.2.3	Analysing coffee terroir subsets	115
6.2.4	Integrating land suitability, yield and quality subsets into coffee terroir sets.....	117
6.2.5	Weighting the land suitability, yield and quality factors in the terroir	118
6.2.6	Geographical indicators of origin	123
6.2.7	Coffee terroir dynamics in relation to climate change.....	124
6.2.8	Effect of scale on terroir mapping and uncertainty.....	126
6.3	Conclusions and outlook.....	126
7	Future directions	129
	Bibliography	131
	Summary	147
	Samenvatting.....	149
	Author's Biography.....	153
	ITC Dissertation List	155

List of figures

- Figure 1-1: Main topographic features of Rwanda with Agro-Ecological Zones: KS = Kivu lake Shore, IMP = Impala, IMB = Imbo, CND = Congo Nile watershed Divide, MPB = Mayaga plateau & Central Bugesera; CP = Central Plateau, ERP = Eastern Ridge and Plateau, EL = Eastern Lowland, BH = Buberuka Highlands, VHP = Virunga Summit and High plains. The black dashed outline delineates national parks i.e. Birunga (NW), Nyungwe (SW) and Akagera (E). The elevation map is derived from a 10-m resolution digital terrain model..... 5
- Figure 2-1: Transformed Divergence (TD) index values (-) show the spectral separability between coffee and other crops of Rwandan agro-ecosystems (mixed seasonal crops, banana and forest). The Agro-Ecological Zones are ordered in west–east direction. KS = Kivu lake Shore, IMP = Impala, IMB = Imbo, CND = Congo Nile watershed Divide, SP = South Plateau, MP = Mayaga plateau; CP = Central Plateau, CB = Central Bugesera, ERP = Eastern Ridge and Plateau, EL = Eastern Lowland. 20
- Figure 2-2: A Directed Acyclic Graph (DAG) used to train the expert Bayesian network model for coffee mapping. DAG 1 represents the naive Bayesian network classifier that distributes the probability of coffee based on the spectral properties of pixel candidates and the pattern of coffee fields used as training samples. DAG 2 and 3 represents the expert knowledge in the Bayesian network model. DAG 2 introduces tree and field rules whereas DAG 3 introduces flow accumulation and forest data to model the location of the coffee field in the toposequence and land use systems..... 22
- Figure 2-3: Intermediate and final coffee classification results after each step of the expert Bayesian network model. The location of the potential coffee ($P \geq 0.49$), is shown by the probability maps (left) and the number of pixels and their likelihood to coffee (C) or no-coffee (N) is shown by the boxplot (right). Top-down: a. Pixel probability result of spectral analysis; b. Object probability result of four-neighbor classification. c. Coffee objects are clustered by the Tree Crown Classifier. d. The Field size classifier aggregates the tree objects into coffee field objects and adjusts probability of coffee field. e. Final coffee probability map after the forest and flow accumulation data are added in the model to filter out coffee within the forest and in the channels. O* represent the number of pixels that are deemed to be outliers. 28
- Figure 2-4: The percentage of pixels that falls within each probability class for five classification steps of the expert Bayesian network model. These steps are (1) Pixel Probability Classification (PPC); (2) Object Probability classification (OPC), (3) Tree Crown Classification (TCC), (4) Field size classification (FSC), and (5) filtering by the forest and flow accumulation map that filter out coffee within the forest area and in the water channels. The test site is about 0.4 km² and located in Huye district..... 29
- Figure 2-5: Aggregated national coffee cover results following nation-wide application of the e-BN model to aerial orthophotos: (a) percentage of coffee cover within a 5 km × 5 km grid cell. (b) Coffee areas aggregated per district and overall accuracy for each district calculated based on 1362 GPS locations of coffee and no-coffee in the 30 districts of Rwanda. The graph (c) plots

mapped coffee (Cm) vs. to reported coffee area (Cc) by the national coffee census of 2009. The coffee map explains 92% of the spatial variation of the census data. 31

Figure 3-1: Coffee density (%) within 5x5 km grid. Ten main agroecological zones of Rwanda: (1) KS Kivu lake shore, (2) IMP Impala, (3) IMB Imbo, (4) CND Congo-Nile watershed divide, (5) MPB Mayaga plateau and Central Bugesera; (6) CP Central Plateau, (7) ERP Eastern Ridge and Plateau, (8) EL Eastern Lowland, (9) BH Buberuka Highlands, (10) VHP volcanoes and high plains. The country size is 26,338 km². 41

Figure 3-2: Explanatory power (β = standardised 'b' values) of biophysical and human variables loading in the five main components (factors) at country level and at agro-ecological zone level (factor label see legend table 3-2 and figure 3-2). National R² = 0.09, Agro-ecological based models: KI R² = 0.52; KS R² = 0.10; IMP R² = 0.17; IMB R² = 0.54; CND R²=0.25; CP R²=0.34; BH R²=0.28; MPB R² = 0.11; ERP R²=0.08; EL R²=0.19. 49

Figure 3-3: Actual coffee field density (a), predicted coffee density maps by the generalised (b) and regionalized (c) models compared with the actual density of coffee (a) mapped using high resolution imagery (Mukashema et al. 2014). The difference of actual and predicted are (d) explicitly visualises potential areas change. The scale from negative (-) to positive (+) change indicates where actual coffee area densities are larger (+) to lower (-) than expected by location factors. The coffee suitability (from S1 to N2) map generalised over AEZs (Verdoodt & Van Ranst, 2003a) is overlaid to calculated map. The country size is 26,338 km². For presentation, the resolution has been reduced to the grid cell size of 7 by 7 acres which correspond to the mean variance of coffee plots. 51

Figure 4-1: Map showing examples of coffee fields supplying coffee cherries to (selected) nearest coffee washing station (CWS). Different supply area touch each other following the density of coffee fields, the size and productivity of coffee farms as well as the processing capacity of coffee washing stations. By 2016, there were 245CWS. The diameter of the supply area ranges between 0.5km and 10km with an average diameter of 3km. Agroecological zones: KS Kivu lake shore, IMP Impala, IMB Imbo, CND Congo-Nile watershed divide, CP Central Plateau, MPB Mayaga plateau and Bugesera, ERP Eastern Ridge and Plateau. 61

Figure 4-2: Map showing the average total production supplied to 166 coffee washing stations (CWS) by individual farmers for the period from 2011 to 2014. Agroecological zones: KLS Kivu lake shore, IMP Impala, IMB Imbo, CND Congo-Nile watershed divide, MPB Mayaga plateau and Bugesera, CP Central Plateau, ERP Eastern Ridge and Plateau, EL Eastern Lowland, BH Buberuka Highlands, VHP Volcanoes and high plains. 62

Figure 4-3: a) Histogram showing the coffee yield gradient from Western mountains to Eastern lowlands. Yield data are presented at District level using the total production reported at the National coffee export board (NAEB) for the season 2015 divided by the number of hectares per district. b) The trends in annual total coffee production for Rwanda 2007-2015. 63

Figure 4-4: Achievable coffee yield map (a) calculated using the aggregated soil-based sub models. The predicted achievable yield is assessed against the actual yield for the season 2015 (b). The model performed relatively well (adjusted $R^2=0.6$). The overall mean difference between achievable yield and actual yield is estimated to 473 kg ha ⁻¹	75
Figure 4-5: Projected coffee yields by 2020 and 2050 using mean annual total rainfall and average temperature trends estimated by McSweeney and Semafara (2011) and Muhire and Ahmed (2015) using available climate data of Rwanda from 1931-2010.....	79
Figure 4-6: Projected coffee yields through increasing soil organic carbon by 0.6% from the minimum SOC of 1.2% required for coffee production.....	81
Figure 5-1: Map showing the provenance of coffee sample lots (used in the CoE) within the coffee washing station (CWS) supply areas. To avoid the loss in coffee quality, each coffee washing station is supplied by coffee farms located within 3 km radius on average. Where there are limited numbers of coffee washing stations, there are collection centres where farmers bring coffee cherries, and CWS arranges the transport. Agroecological zones: KS Kivu lake shore, IMP Impala, IMB Imbo, CND Congo-Nile watershed divide, MPB Mayaga plateau and Central Bugesera, CP Central Plateau, ERP Eastern Ridge and Plateau.	88
Figure 5-2: Map showing the overall sensory quality distribution and coffee washing stations (CWS) where coffee samples for sensory analysis are prepared. In total 503 coffee samples were brought by CWS to the national cupping laboratory for analysis of the sensory quality during the Cup of excellence (CoE) competitions 2008-2015. These samples represent coffee fields in the CWS supply area. Agroecological zones: KS Kivu lake shore, IMP Impala, IMB Imbo, CND Congo-Nile watershed divide, MPB Mayaga plateau and Central Bugesera, CP Central Plateau, ERP Eastern Ridge and Plateau, EL Eastern Lowland, BH Buberuka Highlands, VHP Volcanoes and high plain. Winning coffee washing stations and farmer's representatives are available online at https://www.allianceforcoffeexcellence.org/en/cup-of-excellence/country-programs/rwanda-program/	90
Figure 5-3: Spatial autocorrelation test and Moran's index showing the spatial dependence between coffee sensory qualities measured at 207 coffee washing stations using 503 coffee sample lots used in this study.	93
Figure 5-4: Coffee sensory quality potential (overall grade in percentage) as estimated by (a) the national model and (b) the agro-ecological based models. The scatter plot (C) presents the validation results of the national model (a), and the graph (d) presents the validation of the regional model (b).	99
Figure 5-5: Maps showing the coffee quality potential as result of projected increase in temperature by 0.35°C and rainfall by 40 mm in 2020 and by 1.4 oC and by 160 mm in 2050 (McSweeney & Semafara, 2011; Muhire & Ahmed, 2015).	103
Figure 5-6: The value (US\$/lb) of speciality coffees during the annual Cup of Excellence (CoE) auction: the value presented is the high bid price	

(1lb=0.454kgs). Ordinary coffees (coffee with score below 85%) are not sold at CoE auction. Of the total 245 coffee washing stations (CWS) in 2015, CWS winners were only 24 CWS in CoE 2015, 28 CWS in CoE 2014, 15 CWS in CoE 2013, and 26 CWS in CoE 2012. 104

Figure 5-7: The total production of speciality coffee (coffee with score above 85%) in relation to their quality as measured at the Cup of Excellence (CoE) 2012-2015. Of the total 245 CWS in 2015, CWS winners were only 24 CWS for CoE 2015, 28 CWS for CoE 2014, 15 CWS for CoE 2013, and 26 CWS for CoE 2012. 105

Figure 6-1: A methodological flowchart showing steps followed in developing Coffee Terroir Model 112

Figure 6-2 Coffee field distribution in agro-ecological zones of Rwanda; SW-NW: IMB Imbo, IMP Impala, KS Kivu lake shore, VHP Volcano high plain. Central S-N: CND Congo-Nile watershed divide, CP Central Plateau, BH Buberuka Highlands. SE-NE: MPB Mayaga plateau and Central Bugesera, ERP Eastern Ridge and Plateau, EL Eastern Lowlands. The black dots represent coffee washing stations. By 2016 coffee washing stations were operational across the country. For presentation purpose, we present for all agro-ecological zones an area of 3Km² only. The full map can only fit visually to A4 paper size at a scale of 1: 1,000,000..... 114

Figure 6-3 Coffee terroir based on suitability predicted by biophysical factors of coffee field location. The original data used in the analysis had a grid size of 10m, however for integration with other data and presentation purposes, the resolution has been reduced to a grid cell size of 1 x 1km..... 115

Figure 6-4 Coffee terroir areas predicted using relationship between actual yields and biophysical factors at coffee field location..... 116

Figure 6-5 Predicted terroir based on sensory quality data and biophysical conditions of coffee fields..... 117

Figure 6-6 Average value (US\$/lb) of speciality coffee lots sold at the annual Cup of Excellence (CoE) auction 2012-2015. The value presented is the average high bid price (1lb=0.454kgs). Ordinary coffees (coffee with score below 85%) are not sold at CoE auction. In this trend analysis, we considered 93 lots CoE winners of 421 total sample lots submitted in the 2012-2015 CoE competition 119

Figure 6-7 Coffee terroirs map produced by combining fuzzy field location, yield and quality subsets produced using field location suitability, yield levels and sensory quality models. 120

Figure 6-8 Land area in percentage of the total country land (i.e. 2.3 million hectares) for the range of fuzzy terroir subsets: Coffee field location suitability (a), Yield levels (b), Sensory quality (c), and overall terroir index (d). The last two classes are considered to contribute to terroir zones. 122

Figure 6-9 Coffee terroirs with only geographic areas that are expected to produce special coffee (Sensory Grade \geq 90%) and excellent coffees (Sensory Grade = 87% - 90%)..... 123

Figure 6-10 Projected coffee terroirs shifts in relation to climate change by 2050 as a result of increased average temperature by 1.4°C and total annual rainfall by 160mm. The shift in coffee terroirs is generally negative and East-West and South-North oriented causing a reduction of special coffee grade ($\geq 90\%$) to excellent coffees (87% - 90%), and excellent coffees to fine coffee (80-85%).125

List of tables

Table 2-1: Error matrix of observed coffee fields vs. classified coffee fields by the expert Bayesian network (e-BN) for the test site (0.4 km ²) in Huye, South plateau of Rwanda. The matrix values were calculated on the final coffee polygons. Overall accuracy of the test model were obtained by considering all pixels with $P > 0.49$ as coffee and $P < 0.49$ as no-coffee. Overall accuracy is 87.6% with Kappa value of 0.82.	29
Table 3-1: Summary of principal component analysis results for the 29 variables at coffee locations	45
Table 3-2: Coefficients (b) quantifying the contribution each variable to the current coffee areas (ha) distribution.....	48
Table 4-1: Percentage of coffee farms under biophysical factors and actual yields (t ha ⁻¹) categories. The variables used in this analysis have earlier revealed to be significant factors controlling coffee field location (Mukashema et al.2016).....	70
Table 4-2: Coffee yield models stratified based on agroecological zones with parameters: b-values of biophysical factors (EM) are presented with importance of each factor in the model (standardized coefficients Beta: β -values) in the brackets. All factors are significant at 0.1% level ($P \leq 0.001$). 73	
Table 4-3: Coffee yield models stratified based on lithology (soil types) with parameters: b-values of biophysical factors (EM) are presented with importance of each factor in the model (standardized coefficients Beta: β -values) in the brackets. All factors are significant at 0.1% level ($P \leq 0.001$). 74	
Table 4-4: Spatial autocorrelation results showing the spatial trend of coffee yields across the country: Positive value of Moran's index, higher than the expected index indicates that coffee yields are not randomly distributed, instead present a clustering trend.	76
Table 5-1: Summary statistics of environmental related parameters calculated on randomly selected 59, 520 coffee fields (i.e. 50% of the national coffee fields). Average values of each environmental variable are presented per agroecological zones. Standard deviations are in bracket (see chapter 4 for more detailed description).	91
Table 5-2: Separation of coffee beverage quality means between agroecological zones in homogeneous subsets at $\alpha = 0.05$ according to Tukey HSD test.	94
Table 5-3: The mean difference ($\alpha = 0.05$) in coffee sensory quality between agroecological zones of Rwanda.....	95

Table 5-4: Coffee sensory regression models: country model and regionalised models based on agro-ecological zones. Model coefficients (b-values) of biophysical factors and their respective importance (i.e. Beta: β -values) are presented between brackets. All factors are significant at 0.1% level ($P \leq 0.001$)..... 97

Table 6-1: Geographic indicators of origins and thresholds for 56 identified coffee terroirs.....124

1 Introduction

1.1 Coffee provenance: “a novel business model” for coffee producing countries

Coffee is an important crop that provides a livelihood to millions of people living in developing countries. It is the second most traded commodity in the world after petroleum (Tran, Lee, Furtado, Smyth, & Henry, 2016). Wine and coffee are two important and iconic examples of markets where origin plays a significant role. Both products are derived from perennial crops that engage producers in long-term relations with place and local resources that have a great impact on land use and landscape (Belletti, Marescotti, & Touzard, 2015).

Coffee value chains present some relevant features regarding the development of valorisation strategies based on geographical origin, which differentiate them from wine value chains. Compared to wine, coffee is an example of a buyer-driven (Ponte, 2002) global chain with globalization affecting every tie, while wine represents producer-driven value chains rooted in territories which produce a recognized and specifically labelled wine that can be marketed globally. In the wine value chain all the key activities are usually managed and executed in the territory of origin. Grapes are produced, harvested, processed and bottled in the same region. The coffee value chain is divided into two separate parts. The bulk of world coffee production is grown and harvested in the global south from where it is exported as green coffee to developed countries (mainly Europe and US). In the developed countries the most value adding steps are made such as blending, roasting, grinding and marketing. As a consequence, the reputation of coffee is mainly a competition between brands (processing and marketing part) and the origin of the coffee (often only labelled at country level) plays a secondary role.

With time, the coffee market has gradually shifted from a producer-driven to a more buyer-driven commodity supply chain, and today the coffee trade in global markets is concentrated in the hands of a few big multinational companies (Gereffi, Humphrey, & Sturgeon, 2005). Only recently, with the wave of differentiation strategies in agrifood consumption markets, the “turn to quality” is gradually taking place in the coffee industry (Belletti, Marescotti, & Touzard, 2015). The concept of differentiation by territorial origin is a recent concept in the coffee value chain and location of production is emerging as a new tool in the hands of producers. Geographical origin of coffee is of great interest for both consumers and producers since it provides decisive criteria of acceptability in terms of guaranteed quality (Gonzalez-Centeno et al., 2015). It also allows better checks for more honest trade supported by fair trade certification efforts.

In Rwanda Coffee Arabica (var. Bourbon) grown by smallholder farmers, produces different sensory qualities among the different regions where they

are produced. It is not easy to verify if this distinction is actually linked to their geographical origin or, conversely if it is caused by coffee roasting practices. According to Folmer (2014), high quality coffees are linked to the origin of beans and not to the post harvesting process and roasting. Hence, there is a need for methods that separate the effects of the geographical location from the effects of coffee processing (Gonzalez-Centeno et al., 2015). Such a method could potentially be guided by the concept of 'terroir'. Terroir is an area, usually rather small, whose soil and micro-climate impart distinctive qualities on food products (Gangjee, 2017). Terroirs are strongly related to quality since at purchasing points (cooperatives) higher quality is rewarded with higher prices (Quiñones-Ruiz et al., 2016) and higher quality has therefore direct links with the specific places where it is produced. Such a relationship between quality of a product and location makes a geographic area, a terroir (Foroni et al., 2017) a profitable concept. Such commodity localization via niche i.e. terroir and specialty agro-food markets holds particular promise for rural development (Barham, 2003; Hinrichs, 2003).

The aim of this study is to explore and develop different methodologies in order to derive the relevant coffee terroir characterization for Rwanda using existing data and information as much as possible. This could potentially provide the government of Rwanda with a scientifically based information system to earmark locations for small holder coffee production with high revenues. By deriving and mapping coffee terroirs for the whole country, multiple actors and coffee partners of Rwanda can become more directly engaged in more actively managing specific locations and/or regions for a more sustainable coffee market. This would certainly contribute to rural development for small holder coffee growers. In general, the development of a coffee terroir mapping methodology could advance other similar agricultural commodities with similar value chains such as tea and cacao.

1.2 Overview of the coffee sector in Rwanda

1.2.1 Geographic location and landscapes

Rwanda is located in the Central and Eastern Albertine Rift of Sub-Saharan Africa (1°–3°S, 29°–31°E) and covers 26,338 km². The country consists of mountains, rounded hills and valleys that have been formed by rifting and uplifting of Pre-Cambrian rocks followed by erosion since the Miocene and by volcanic activity that initiated during the Plio–Pleistocene (Huhndorf, Kerbis Peterhans, & Loew, 2007). The main topographic features (Figure 1-1) are a mountain chain from volcanoes in the north (VHP) to mountains in the south west, which makes up the Congo-Nile watershed divide (CND) at the altitude between 2000 and 4500 m a.s.l. Lake Kivu shore (KS), Impala (IMP) in the south west; central plateaus and ridges (CP, ERP) at altitudes of 1500–2000

m. Mayaga, Bugesera and Umutara lowlands (MBP, EL) in East at altitudes of 1000–1500 m, and Imbo plain (IMB) in the South West at altitude below 1000 m a.s.l. The abbreviations mentioned above refer to the ten agro-ecological zones of Rwanda and are also shown in (Figure 1-1).

The agriculture sector in Rwanda is dominated by small-scale farming systems consisting of mixed seasonal, annual and perennial crops. The total agricultural land (coffee included) is estimated at 55% of the country territory followed by forest ecosystems natural and planted forest currently close to 28%, urban areas (6%), wetlands (6%) and water bodies (5%).

The Rwanda coffee export industry plays a particularly important role in the country's development. For many years, coffee remains Rwanda's top export and main source of foreign exchange income (Boudreaux & Ahluwalia, 2009). It provides a livelihood for more than 355,000 Rwandan families (NAEB, 2016), over 2 million people, close to 25% of the total population (OCIR-CAFE, 2005), many of whom work in cooperatives and grow coffee on small plots (Boudreaux & Ahluwalia, 2009). A total of 35,891 hectares (NAEB, 2016) are located on the country's rural areas. *Coffea Arabica* is the dominant coffee species cultivated in Rwanda, with a share of more than 99% of the national coffee area. *Coffea Robusta* occupies less than 1% of the total production area and is only cultivated along Nyabarongo wetlands in Muhanga and Gakenke Districts. The most cultivated Arabica variety is Bourbon Mayaguez (BM 71, BM 139) and Jackson 2/1257. Coffee plantations in Rwanda are generally found in hilly terrain and largely consist of small monoculture fields. Very few plantations are shaded with selected leguminous trees, and sometimes in association with other perennial crops like banana. Coffee is grown on a range of different soil types, some more fertile than others. The age of coffee trees ranges from months to more than 30 years. Coffee field sizes depend on the available land of individual households. Loveridge, Nyarwaya, and Shingiro (2003) demonstrated that individual coffee fields in Rwanda are typically small consisting of about 155 to 200 coffee trees. The recent national coffee census (2016) also reported relatively small coffee field sizes consisting of 252 trees on average (NAEB, 2016). These individual fields are generally located on steep slopes, and near forests and bushes. Farm management practices (mulching, fertiliser application, pruning, etc.) also differ depending on socio-economic status of coffee farmers. The plant spacing varies from 1.5 to 2.5 m depending on agricultural zones.

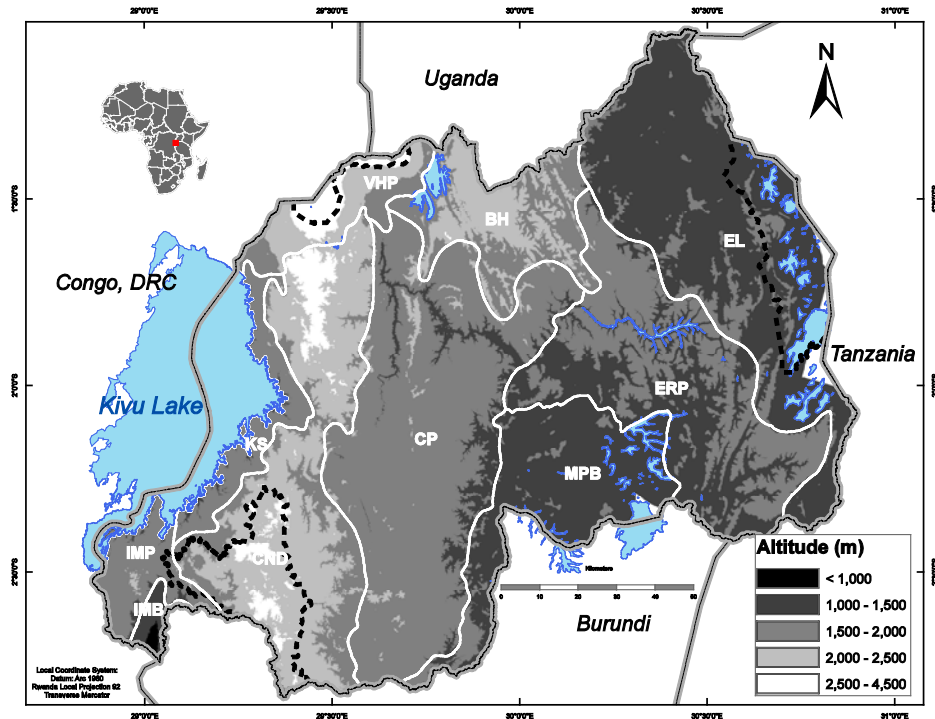


Figure 1-1: Main topographic features of Rwanda with Agro-Ecological Zones: KS = Kivu lake Shore, IMP = Impala, IMB = Imbo, CND = Congo Nile watershed Divide, MPB = Mayaga plateau & Central Bugesera; CP = Central Plateau, ERP = Eastern Ridge and Plateau, EL = Eastern Lowland, BH = Buberuka Highlands, VHP = Virunga Summit and High plains. The black dashed outline delineates national parks i.e. Birunga (NW), Nyungwe (SW) and Akagera (E). The elevation map is derived from a 10-m resolution digital terrain model.

1.2.2 Coffee sector governance

Coffee was introduced in Rwanda by missionaries in the early years of the twentieth century (Murekezi, 2009). Official government involvement began in the 1930s with the Belgium colonial government's coffee campaigns (Boudreaux & Ahluwalia, 2009; Murekezi, 2009). Under these policies, government authorities built nurseries and supplied seeds, but they also required Rwandan farmers to plant coffee trees.

After independence (1962), the government of Rwanda (GOR) continued the policy of requiring farmers to grow coffee. In 1964, a state-run agency, known as OCIR-café was created to be in charge of buying all coffee produced by farmers and exporting it. In the same year, RWANDEX, another company partially owned by the Government was created to be in charge of exporting coffee (Murekezi, 2009). In 2010, OCIR-Café was merged with OCIR-Thé (a parallel national authority in charge of tea production and

marketing), and formed the current National Agricultural Export Board (NAEB).

With the involvement of coffee cooperatives and private sector entrepreneurs, the Government of Rwanda has to date liberalized the coffee industry, along with other sectors of national economy (Boudreaux & Ahluwalia, 2009). The reforms in the sector included the privatisation of the State-owned coffee export company (RWANDEX) in 2006. Following its privatisation, a large number of new private domestic and international coffee exporters are active in Rwanda.

To date, NAEB continues providing valuable technical assistance to coffee growers and processors (MINAGRI, 2008). NAEB works closely with producer cooperatives, where the cooperatives are responsible for post-harvest activities up to green coffee beans. NAEB sets the farm gate prices, licenses producers, controls the quality, and regulates export and inputs along coffee production chain (seeds, fertilizer, pesticides, materials for harvests storages, and transport up to auctions). NAEB has the role of official “brander” of Rwandan coffee.

1.2.3 Current coffee production and quality trends

The average coffee productivity in Rwanda is about 0.7 ± 0.4 ton per hectare. This yield is low as compared the expected yield with regard to land capability analysis (Verdoodt & Van Ranst, 2003b). Generally, coffee yield of Rwanda is lower than the average (1.2 t ha^{-1}) obtained in small scale farming systems optimized for coffee growing (Wintgens, 2009). Compared with world-wide coffee growing countries, Rwandan coffee farmers obtain yield levels well below the levels (1.7 t ha^{-1}) obtained by small scale farmers in India (Jayakumar, Rajavel, & Surendran, 2016), and 1.8 t ha^{-1} in Brazil (de Muner, Masera, Fornazier, de Souza, & de Loreto, 2015). Very high yields (3 t ha^{-1}) are reported in Brazil on commercial Arabica coffee farms (Teixeira, Souza, Pereira, de Oliveira, & Rocha, 2015). In East Africa, Rwanda has similar yield ranges as observed in Uganda (0.7 to 0.9 t ha^{-1}) (N. Wang et al., 2015).

Concerning the quality, Rwanda is the first country, followed by Burundi in 2012, in Africa to adopt quality as the main strategy of coffee market through participation in World Cup of Excellence competitions since from 2008. As compared to other country in the world market, in 2015 CoE auction, the best coffee of Rwanda was sold at US\$ 30.60 per lb (1lb=0.454kgs) (the 2nd price after Costa Rica coffees sold at US\$42.10/lb, Burundi (US\$28.70 per lb), Guatemala (US\$25.80 per lb), El Salvador (US\$23.50 per lb), Honduras (US\$23.30 per lb), Nicaragua (\$21.60 per lb), Brazil (US\$18.90 per lb), and Colombia (North) coffees sold at US\$14.50 per lb (Refer to: <https://www.allianceforcoffeeexcellence.org>).

1.2.4 Link between coffee product and territory

As reported by Descroix and Soneck (2009), there are six basic environmental factors that need to be taken into account while selecting the best suitable places to plant coffee. i.e. temperature, water availability, sunshine, intensity, wind, type of soil and topography of the land (Wintgens, 2009). In case of Rwanda, optimal environmental conditions for coffee arabica have been registered on the shore of Kivu Lake (KS), in Impala (IMP) and in the central plateau (CP) agricultural regions. The Mayaga and peripheral Bugesera (MPB) are characterised by a very suitable soil conditions, however the climate especially low rainfall is a limiting factor to the growth of coffee in this region. A further decrease in rainfall towards the east, gives a further reduction in suitability, with the eastern plateau being moderately to marginally suitable, whereas the climate is marginal or too dry for coffee in the eastern lowland and Bugesera. Unsuitable climate conditions are recorded in the cool high altitude (Buberuka) of the North western part of Rwanda (Verdoodt & Van Ranst, 2003a).

1.2.4.1 Climate

Temperature values and their fluctuations have significant impact on the behaviour of the coffee tree. The optimum mean temperature for coffee arabica is given as 18°C during the night and 22°C during the day whereas optimal average annual temperatures are given between 22 and 28°C. Tolerated extremes extend to 15°C during the night and 25-30°C during the day. Temperatures higher than 25°C cause reduced photosynthesis and prolonged exposure to the temperatures above 30°C incur leaf chlorosis and generate "star flowers", or blossom wilting, as well as defective fruit set. High temperatures also favour the development of Coffee Leaf Rust (*Hemileia vastatrix*) and fruit blight (*Cercospora Cafeicola*) in coffee plants and accelerate fruit maturation whilst low temperatures lead to coffee berry diseases (*Colletotricum coffeanum*) (Wintgens, 2009)

The rainfall pattern also matters. It must include a few months with little or no rain because such a dry period is necessary to induce flowering. A total annual rainfall between 1400 and 2000 mm is favourable for arabica growing. Even well-distributed precipitation amounts below 800-1000 mm for Arabica can be hazardous to the productivity of a coffee plantation. This is particularly valid for Rwanda where no artificial irrigation is applied by coffee farmers. The average annual water requirement over 12 years for coffee was found to be 951 mm when the dry season is normal and the coffee is grown on soils with a high water retention capacity (Wintgens, 2009). Therefore, overall rainfall distribution throughout the year is a decisive factor for scheduling the cultivation practices and harvesting. For example, where the rainfall is bimodal, there will be two periods of blossoming and, consequently,

two harvest periods, known as the “early crop” and the “late crop”. Coffee from the “late crop” is of better quality.

Atmospheric humidity (or relative Humidity, RH) of the air also has a marked influence on the behaviour of the coffee plant. The best RH of Arabica is around 60%. A high level of atmospheric humidity will reduce the water loss by the plant, whereas a low level will increase evapotranspiration. In case of Arabica, persistent levels above 85% may affect the quality of the coffee.

Coffee flowering is significantly stimulated by direct sunlight and shading and by adequate fertilisation, in particular nitrogen, sunshine can boost productivity. For best results, coffee requires an average of 2200-2400 hours of sunlight per year (Wintgens, 2009). In Rwanda, most coffee plantations are established in direct sun. In Rubona, for example, the average annual exposure to sunlight between 7 a.m. and 5 p.m. is about 58% during the rainy season and 75% in dry season.

1.2.4.2 Soils

According to Wintgens (2009), coffee plants prosper on well drained alluvial and colluvial soils with a favourable texture as well as on deep soils developed in volcanic parent materials. Soil depth should be at least 2 m to allow tap root development to ensure the necessary water supply to the plant during the dry seasons of the year. However, soils where the water table is permanently high should never be used for coffee growing. The water table should be at least 1.5 m below the soil surface. Moreover, soils which are subject to occasional flooding should also be avoided as coffee tree roots are easily asphyxiated. Poorly drained soils or heavy clay soils are not recommended.

The radicular system of the coffee tree develops in 30 cm of the upper soil layer, which means that the physical properties of the topsoil are more important to the coffee tree than those of the deeper subsoil. Favourable soils for coffee production are considered to have a porosity of 50-60% and an organic matter content ranging from 2-5%. These relatively high percentages of organic matter are recommended as they make the soil more fertile, less prone to erosion, and offer an improved water and nutrient retention capacity (Cyamweshi et al., 2014). Generally, suitable coffee soils should not contain more than 20 – 30% of coarse sand and 70% of clay in the upper 30-50cm. In Rwanda, the best soil production environments have been recorded in Mayaga, and around Muhazi and Mugesera Lakes of Eastern plateau (I. Nzeyimana, Hartemink, & Geissen, 2014; Verdoodt & Van Ranst, 2003a).

1.2.4.3 Topography

Altitude is also an important factor because it is directly related to temperature. This is why the highest quality Arabica beans are produced at an altitude above 1000 m. Flat lands or slightly rolling hills are considered best suited for coffee production because in this landscape run-off is limited and erosion can be well controlled. This type of topography is generally associated with deep soils with a good water retention capacity.

1.2.4.4 Management of coffee farms

Coffee farming is composed of a set of management practices ranging from the nursery to the harvest via preparation of the parcel, fertilisation, weeding, pruning, as well as regular pest and diseases management. In Rwanda, the sowing in nursery takes place in October-November; one kg of seed produces about 2500 to 3000 seedlings. About 2500 seedlings per ha with 2m between lines and 1m between holes are planted. Organic fertilisation is also applied prior to the planting as well as at the beginning of each rainy season. About 5 kg of organic manure per coffee tree and 100 grams of NPK 22.6.12 (in Eastern ridges and lowland, and Impala region) or NPK 17.17.17 (in Kivu lake shore, Congo-Nile divide, and Central Plateau) is applied in two applications per year (October-November, and March-April). Managing N:K ratio in coffee soils is of significant importance to avoid the effect of excess potassium on the efficiency of nitrogen fertilizers. Nitrogen is essential for the vegetative growth of the trees. It boosts the development of branches and leaves while excess of potassium shortens leaf life and accelerates leaf fall. During fruit filling, coffee plants have higher nitrogen requirements. Potassium is of significant importance as well, particularly for the development and the maturation of the fruit. However, the excess of potassium inhibits positive effects of nitrogen (Wintgens, 2009).

Farm management practices (mulching, fertilizer application, pruning, etc.) differ depending on socio-economic status of coffee farmers. Coffee plantations in Rwanda are generally found in hilly terrain and largely consist of small monoculture systems sometimes in association with other perennial crops like banana due to land scarcity to enable farmers to achieve a better combination of food and cash crops (I. Nzeyimana, Hartemink, & de Graaff, 2013). Most of coffee trees are sunny-exposed fields; very few plantations are shaded with selected leguminous trees. The mulching material used by most coffee farmers is mostly dominated by *Eragrostis blepharoglossis* K. Schum, vegetation available on acid soils (Cyamweshi et al., 2014; Loveridge, Mpyisi, & Weber, 2002). In Rwanda, mulch is applied in coffee fields to control soil erosion (Innocent Nzeyimana et al., 2017).

Another management measure to maintain coffee production on the longer term is pruning. Coffee tree pruning is done three times along with the tree

growth. The first pruning which consists of bowing the coffee-tree at 10 months is required in order to get shoots leading to 3 stems. The second named production pruning aims at removing primary lower branches, branches which are dried out and secondary branches. The third named regeneration pruning takes place every 5 or 7 years of production. Diseases and pests control are part of the daily management of the coffee plantations.

1.2.4.5 Coffee harvesting and post-harvesting processing chain

Coffee cherries, once well ripe, are harvested usually manually by farmers. They transport the cherries by themselves to the nearest coffee washing station for washing, drying and hulling to produce green coffee beans. These stations are owned either by farmer's cooperatives or by local investors. The cherries are first pre-sorted, by floating them in water to separate the heavier cherries from the lighter ones. Heavier cherries are run through pulping machines that remove much of the fruit from the bean. Beans are fermented in holding tanks, as fermentation makes it easier to remove the remaining mucilage. After the fermentation process is complete, the beans are washed four to five times. The beans are sorted for the second time and left to float for 24 hours. After this 24-hour of floating, the beans are laid out on sorting tables, and workers pick out damaged or irregular beans. The good beans are dried in the sun until they reach a desired moisture level (about 12%), which may take between ten days to two weeks, depending upon the weather. Parchment coffee is then packed in burlap bags and transported to hulling centres. Hulling parchment coffee to green coffee beans is done by the local millers. No parchment coffee is allowed for export according to Rwanda regulations, only green or roasted coffees are exported.

1.3 Challenges in zoning coffee terroirs

Reliable spatial information on coffee production areas (coverage) is needed for terroir zoning. Though, coffee census inventories are regularly carried out (2009, 2015) specific information on the geographical location of coffee fields is lacking. The census only publishes the number of trees reported by farmers and the area (hectares) is estimated assuming an average line spacing (2m x 2m). To date, serious efforts are made by the Government of Rwanda to avail very high resolution aerial orthophotos (of 25cm pixel size) for land use planning and management. Hence with suitable tools and new methodologies the high resolution orthophotos could be used to produce a coffee base map. Ideally such a map could be regularly updated along with the planting of new coffee and/or removal of coffee for other land uses.

There are very few studies of predicting spatial coffee distribution (e.g. Cordero-Sancho & Sader, 2007; Gomez et al., 2010; Moreira, Adam, & Rudorff, 2004; Tardin, Deassuncao, & Soares, 1992). All of them dealt either

with large plots of coffee, or a small study area, or used very high resolution remote sensing data (such as Quick bird or Ikonos imagery). However, coffee plantations, in Rwanda, are spread over the countries' landscape and mainly cultivated in smallholder farming systems. An additional challenge is the varying age of the coffee plantations: some are more than 30 years old, while others are younger than 3 years. Some are very small plots of about 200 trees and very sparse (2mx2m) or close by 1.5mx1.5m, with or without mulch. Farm management practices are also not the same throughout the country, depending on socio-economic welfare of farm owners, expressed in health differences of coffee trees. All these lead to a complex spectral signature for coffee field canopy cover on multispectral imagery, which we generally use to detect crop cover with Remote Sensing techniques. Two recent attempts for directly mapping coffee systems therefore relied on visual image interpretation (Trabaquini, Miglioranza, de Franca, & Neto, 2011) or classification improved by visual interpretation (Moreira, Rudorff, Barros, De Faria, & Adami, 2010). However, visual image interpretation requires manpower and is time demanding, especially when dealing with small-scale coffee farming systems in mountainous areas. Hence an automated method is needed that is based on aerial orthophotos to produce a coffee base map of Rwanda. Once such novel method succeeds, it could be extended to the Eastern and central Africa region and elsewhere.

It is not only the lack of information on coffee field distribution that hampers coffee sector plans, but also the lack of information on the distribution of suitable environments for growing coffee in the region. We do have very generalised information on physical and natural environmental factors requirements for coffee production; i.e. the altitude, slope aspect, soil structure, nutrient level needed for high production and high quality coffee (Wintgens, 2009). However, these factors are site-specific and dynamic. Only topographic factors remain stable over several decades. In fact, the climate of Rwanda is expected to change towards a warmer and wetter climate by 2050 (McSweeney & Semafara, 2011). Future temperature projections for Rwanda were analysed by McSweeney and Semafara (2011) using daily temperature data between 1970 and 2010. Their analysis indicates that the mean temperatures have significantly increased in Rwanda over the past 40 years (an average of 0.35°C per decade, $P < 0.001$) taking average annual temperature towards 22°C in 2010. A similar increase was observed in minimum and maximum temperatures. Rainfall records, by comparison, showed no significant trend between 1931 and 1990 (there was not sufficient data to assess the most recent past). Increases in mean temperature are projected under all models and all emissions scenarios, while the majority of models also indicate increases in annual rainfall totals, though a few show small reductions. This has been confirmed for Rwanda by Muhire and Ahmed (2015) who used the same rainfall data for a period from 1935 to 1992 and indicated a significant incline over the northern region, and a decreased of

number of rainy days over the eastern and central plateau. The mean rainfall was found to increase sharply in most parts of the country. A significant increase of rainfall by 1.5 and 6.8 mm per year (4 mm per year on average) from October to December corresponding to short rains were observed. Rainfall gradually decreases in January and February in many parts of the country, then increases sharply from March to May which corresponds to the long rains in 90% of the country territory. Climate and soil are both factors that affect coffee yield and quality, but we do not know exactly how.

We also bear in mind that coffee productivity depends not only on good environment but, more importantly, on the management of the coffee farm. In a liberalised coffee industry, the land use and farming strategy is determined by decisions of the farmer household. Farmers choose what crops to plant, what investments to make, what share each crop should be given in the household farm and how to manage plots for the household livelihood and profitability. This tends to change the potential of a variety such as arabica coffee to yield as optimal possible as it could.

Once the geographic location where specific coffee is grown is known, it is possible to relate coffee characteristics to local production conditions. With detailed location information we can potentially identify where the best locations are found for high yields and high quality coffee. Such locations produce high value coffee sometimes referred to as “appellation” (Barham, 2003). ‘Appellation’ needs also precise knowledge on the entire production process of that product, particularly its geographical and environmental conditions. The geographic locations are sometimes also known for their unique pedo-climate conditions also referred to as “terroir” analogous to wine regions (Rotaru, Filipov, Mustea, & Stoleru, 2010). The product-based farming system should also be known in the defined geographical terroir (Berard & Marchenay, 2006). Such holistic conceptualisation of terroir has never been attempted so far in the coffee-based farming system and therefore we need to develop it. An additional challenge for Rwandese coffee is that it is grown by many small holder farmers.

1.4 Research objectives

The overall objective of research is to identify and delineate possible coffee ‘terroirs’ in Rwanda in a reproducible way. The main underlying research hypothesis is that there are existing process based empirical relationships between coffee production characteristics and location factors. This implies that when we know the different bio-physical and human location drivers we can derive/predict specific coffee system properties.

To derive quantitative coffee system information we need a spatially explicit analysis of current coffee farming systems in Rwanda. Because in Rwanda all coffee is produced in small fields by smallholder farmers a very detailed

inventory of current smallholder coffee fields is required. Step one of this research is the development of a methodology to allow a systematic automated detailed inventory of coffee fields in Rwanda.

Once we know at which locations coffee is grown, we can explore to what extent coffee production is related to location factors. There are three possible ways to make spatially explicit analyses of coffee system characteristics: 1) location suitability for coffee based on current coffee field distribution; 2) location factors that are associated with observed yield variability; 3) location factors that are related to observed coffee sensory quality variability.

A remaining challenge is that the three identified coffee system properties (location suitability, coffee yield and coffee quality) are not directly measured at field level. Coffee yield and quality are measured at different management units, coffee washing stations and lots based on a combination of different coffee fields respectively.

When we know the three empirical relationships of Rwandese coffee system properties and their environment we might combine them in an overall spatially explicit terroir characterization. Such a terroir map could have significant value for identifying, monitoring and developing smallholder coffee production in Rwanda. Once we have an empirical description of the system we can also explore which management measures might improve current coffee system performance or where suitable locations are found for new coffee farms. It will also allow evaluating foreseen changes in climate for example.

1.5 Thesis outline

This thesis consists of six main chapters, each addressing one of the challenges presented in section 1.3.

Chapter 1: In this chapter, we introduce the need for developing a coffee terroir model that can bring together the coffee producer and consumer for sharing economic benefits as well as for identifying and monitoring the proper locations of distinctive coffee of Rwanda.

Chapter 2: In this chapter, we will develop an automated methodology to extract small-scale coffee fields, including at least 200 coffee trees, for the whole of Rwanda using aerial orthophotos covering more than 99% of Rwanda and on one QuickBird image for the remaining part.

Chapter 3: In this chapter, we will investigate which of all available potential coffee related factors (bio-physical and socio-economic) are associated with observed coffee field distribution variability in Rwanda. This analysis could lead to location factor related coffee field suitability.

Chapter 4: In this chapter, coffee yield variability in relation with known location factors will be analysed. Coffee yield data is derived at coffee washing stations. Apart from a spatially explicit analysis we will also explore effects of known and expected climate changes.

Chapter 5: In this chapter, coffee sensory quality is also spatially explicitly investigated for the whole country based on spatially explicitly sampled coffee lots. Also, future trends in coffee quality are projected based on climate change scenarios for Rwanda.

Chapter 6: In this synthesis chapter, we integrate coffee location suitability (Chap.3), coffee yield level variability (Chap.4) and coffee sensory quality (Chap.5) into a coffee terroir model for the whole of Rwanda.

2 Automated high resolution mapping of coffee in Rwanda using an expert Bayesian network

This chapter is published as: Mukashema, A., Veldkamp, A., & Vrieling, A. (2014). Automated high resolution mapping of coffee in Rwanda using an expert Bayesian network *International Journal of Applied Earth Observation and Geoinformation*, 33, 331-340.

2.1 Introduction

Introduction Coffee farming delivers strong economic, social and ecological benefits to many tropical countries. Despite the relevance of coffee cultivation, reliable spatially explicit coffee system inventories are lacking. Such data are needed to gain better understanding of the spatial differences in coffee quality and quantity and to allow further monitoring of the long-term coffee sector performance (Loveridge et al., 2003). In most countries, the only spatial information available is collected during national agricultural censuses, and is reported by administrative unit. Systematic attempts to map coffee from remote sensing data are very limited and to our knowledge non-existing for the African highlands. Existing mapping studies concentrate on large-scale coffee plantations in Latin America (e.g. Cordero-Sancho & Sader, 2007; Gomez et al., 2010; Moreira et al., 2004; Moreira et al., 2010; Ortega-Huerta, Komar, Price, & Ventura, 2012; Trabaquini et al., 2011). Despite their importance for rural livelihoods as well as for the national economy of African highland countries, maps that depict the spatial distribution of coffee are lacking. This is because only high resolution images (0.25 m) can reveal enough detail for identification of individual mature coffee trees. Another complication is that hedgerow and scattered trees are common features in the coffee farming systems of the East African highlands, often impeding direct observation of coffee trees by direct coverage or through the shadows that the larger trees cast on the coffee.

Remotely sensed high-resolution imagery including high quality orthorectified aerial photos (orthophotos) have increasingly become available for Africa. However, automated extraction of small-scale coffee fields from such data is a challenge. In Costa Rica shade-grown coffee was mapped using supervised classification on a single Landsat Enhanced Thematic Mapper (ETM+) image with red, near-infrared, and mid-infrared bands (Cordero-Sancho & Sader, 2007). This study yielded a lower accuracy (56%) as compared to other studies that included more spectral bands and ancillary data (Langford & Bell, 1997). Better classification results (83%) were obtained for a case study in New Caledonia when high spatial resolution satellite sensors (QuickBird) were used and coffee tree crown sizes were considered in the classification (Gomez et al., 2010). (Moreira et al. (2004)) demonstrated that coffee fields have a high spectral variability due to differences in age, plant spacing and cultivars. In addition, topographical effects and spectral confusion between coffee and other tree crops may lead to poor classification accuracies (Cordero-Sancho & Sader, 2007; Gomez et al., 2010; Ortega-Huerta et al., 2012).

Therefore, two recent attempts for mapping coffee systems rely on visual image interpretation (Trabaquini et al., 2011) or classification improved by visual interpretation (Moreira et al., 2010). However, visual image interpretation requires manpower and is time demanding, especially when

dealing with small-scale coffee farming systems in mountainous areas. Hence an automated method is needed.

Object-based classification provides an alternative approach to per-pixel image analysis. It allows developing site-specific models that cater the main (local) features of coffee cultivation, such as field size (Chubey, Franklin, & Wulder, 2006; Kasper Johansen, Coops, Gergel, & Stange, 2007). Most object-based classification methods use segmentation to generate meaningful image objects (Brandtberg, 2007; Doxani, Karantzalos, & Strati, 2012; Gougeon & Leckie, 2006). Several algorithms, membership functions and thresholds are used for image segmentation (e.g. Brandtberg & Walter, 1998; Q. Chen, Baldocchi, Gong, & Kelly, 2006; Gong, Sheng, & Biging, 2002; K. Johansen, Phinn, Witte, Philip, & Newton, 2009; Kim et al., 2010; Pouliot, King, Bell, & Pitt, 2002; L. Wang, Gong, & Biging, 2004; Wulder, Niemann, & Goodenough, 2000). However, all of them reported the problem of overestimation or underestimation of surface areas due to the gap area in between trees. This has an effect on the delineation of field boundaries. A potential solution was suggested by Forster, Buehler, and Kellenberger (2009) by the use of "growth factors" such as number of trees per hectare, or distance between trees in order to accurately delineate tree crown areas. Recently, Bayesian networks have been proposed as an alternative to image segmentation. Instead of segmentation, Bayesian network techniques use probability functions to discretize image pixels into image objects (Mello et al., 2010).

This paper presents and implements an expert Bayesian network as a novel object-based classification technique to extract coffee fields from very high resolution (VHR) imagery. The method is developed and tested for ten agricultural zones of Rwanda using aerial orthophotos. By subsequently applying the automated method on 198 orthophotos and one QuickBird image, we produced a high resolution coffee map of Rwanda. We first assessed the spectral separability between coffee and other major land cover classes. Then, the expert Bayesian network model is developed for one site and tested in other nine sites. Finally, the model parameters are implemented to all aerial orthophotos at national level and assessed using field and census validation data sets.

2.2 Materials and methods

2.2.1 Data and variables

2.2.1.1 Remote sensing data

We used very high spatial resolution imagery covering the entire extent of Rwanda. Aerial orthophotos covered 99% of the country and a QuickBird image of June 2004 covered the remaining part in the south-west of Rwanda. Orthophotos were obtained from an aerial photography mission over Rwanda during the summertime (July–August) in 2008 and 2009. The images were acquired using a Vexcel Ultra Cam-X aerial digital photography camera at 3000 m altitude above ground level, with a mean ground resolution of 0.22 m (Swedesurvey, 2010). The orthophotos were processed by the image provider to a nominal 0.25 m pixel size and contain three visible bands (blue, green, and red). The QuickBird image has four multispectral bands recording radiation in the visible and near infrared domain that have a 2.44 m spatial resolution and a panchromatic band with a spatial resolution of 0.61 m at nadir. Using the panchromatic band, pan-sharpening was performed on the multispectral QuickBird bands. A total of 198 orthophotos and one QuickBird image were processed to map coffee at the national level.

Two additional spatial data sources were used in our study as an input to the classification. The first is a high-resolution digital terrain model (DTM) that was produced from the aerial photographs using stereoscopy (Swedesurvey, 2010). The second is an existing national forest cover map of 2007 (of over 90% overall accuracy), which was made based on image classification using ASTER, SPOT, and Landsat TM images (Schilling et al., 2007).

2.2.1.2 Evidence variables for coffee

We characterized the spatial pattern of coffee across the country based on spectral and spatial variables derived from the orthophotos and QuickBird image? Spectral variables considered in this study are homogeneity and dissimilarity within and between coffee fields. As these are relative spectral variables, further radiometric calibration was not justified as it would have minor effects on these variables. Therefore we derived these variables directly from the digital number (DN) values provided. Homogeneity is measured by the minimum and maximum DN within a window around a pixel, and dissimilarity is assessed from the standard deviation (DN) within that window. Spatial variables used for mapping coffee fields are the average diameter of the tree crown and the minimum size of the field. The diameter of the coffee tree crown is set to 2 m based on our knowledge of the coffee trees and their spacing across Rwanda whereas the size of the field is set to a

minimum of 200 coffee trees following Nsengiyumva (2009) giving it the same minimum criterion as used in the national coffee census. Land use and topography variables were also considered to optimize the location of coffee within the landscape. The existing national forest map and DTM-derived flow accumulation data were added in the process to exclude forest and channel/gully locations, therefore reducing possible errors related to location, the shape and the size of coffee fields.

2.2.1.3 Training and validation dataset

Two datasets were used for model development at small test site, and evaluation of the model performance in the rest of the country as follow: (1) 80 GPS sample points were randomly collected in coffee fields to train the model in the 0.4 km² test site in the South plateau (Huye district). (2) 137 GPS points were collected in the test site based on random sampling of which 79 were taken in coffee fields and 58 in no-coffee fields. These points were for validation of the model at the test site. (3) 469 GPS coffee locations were collected following random sampling during our field work to validate the model transferability in the ten agro-ecological zones that exist in Rwanda. (4) 640 GPS locations of coffee plantations where soil samples were taken by the National coffee authority (OCIR-CAFÉ) for lime and fertiliser recommendation (Cordingley, 2009). (5) 253 GPS locations of soil profiles for which only two profiles were located in coffee plantations and 251 in no-coffee fields (Verdoodt & Van Ranst, 2006). A total of 1362 sample locations were used for validation. (6) Reported coffee areas by the national coffee census of 2009. The census was conducted in all 30 districts of Rwanda by OCIR-CAFÉ together with Statistics Officers at the district and sector level (Nsengiyumva, 2009).

2.2.2 Preliminary spectral analysis

Before setting up our classification model, we first examined to what extent coffee can spectrally be distinguished from other crops on high resolution images. Pixel samples of coffee and no-coffee fields were randomly selected through visual interpretation of the aerial photographs from ten training subsets representing different agro-ecological zones of Rwanda. Transformed Divergence (TD) values were calculated to evaluate the separability between coffee, forest, banana, and mixed seasonal crops. TD values range between 0 and 2000, with higher values (≥ 1900) indicating good separability of class pairs. TD values below 1700 imply that classes are poorly separable, and therefore spectral classification techniques are not expected to produce accurate classification results (Kumar & Silva, 1974). This initial analysis (Figure 2-1) showed that coffee can hardly be separated from banana, mixed crops and forest based on their spectral signatures only. Particularly low TD values were obtained for the plateau area where coffee fields are small and

surrounded by mixed crops, banana and forest. This indicates that we cannot obtain a good accuracy if we try to map these coffee fields using merely spectral classification techniques on orthophotos.

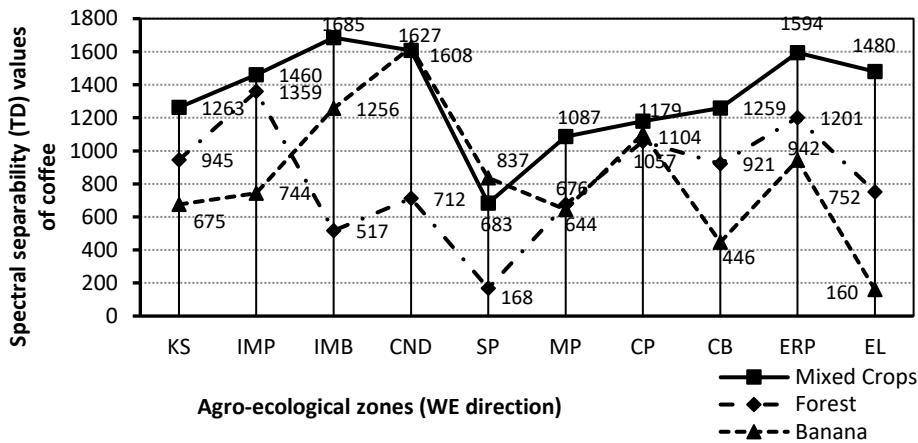


Figure 2-1: Transformed Divergence (TD) index values (-) show the spectral separability between coffee and other crops of Rwandan agro-ecosystems (mixed seasonal crops, banana and forest). The Agro-Ecological Zones are ordered in west–east direction. KS = Kivu lake Shore, IMP = Impala, IMB = Imbo, CND = Congo Nile watershed Divide, SP = South Plateau, MP = Mayaga plateau; CP = Central Plateau, CB = Central Bugesera, ERP = Eastern Ridge and Plateau, EL = Eastern Lowland.

2.2.3 Coffee classification with a Bayesian network

2.2.3.1 Principles and definitions

A Bayesian network can be considered as a graphical model of interactions among a set of data variables (Gambelli & Bruschi, 2010). It is used for knowledge representation and reasoning about a data domain (Bressan, Oliveira, Hruschka Jr, & Nicoletti, 2009; Cheng & Wang, 2010). Bayesian networks are based on the Bayes theorem on conditional probability between two events A and B. The probability of A given that B occurs $p(A/B)$ is given by:

$$p(A/B) = \frac{p(B/A) \times p(A)}{p(B)} \tag{2.1}$$

A Bayesian network structures a set of data (usually a finite set of random samples containing various variables) and analyses their relationships (Aguilera, Fernández, Fernández, Rumí, & Salmerón, 2011). Building a Bayesian network model requires two important steps. The first step is the analysis of causal relations between different variables for random samples (Gambelli & Bruschi, 2010) in the form of a network structure known as

“Directed Acyclic Graph (DAG)”. In a DAG, each variable is represented by a node N , and connected to other nodes by an arc A (Park & Stenstrom, 2006). The arcs represent the relationship between nodes. Once the DAG is formed, the second step consists of quantifying the relations between the connected nodes using conditional probability (Gambelli & Bruschi, 2010). A Bayesian network is consequently a function of N , A , and θ , where the components N and $\langle N, A \rangle$ denotes the DAG graph with children nodes $\{X_i, i = 1 \dots n\} \in N$. N consists of n variables referred as nodes in the DAG, and $a \in N$ representing probabilistic dependency between n nodes. The component $\langle \theta \rangle$ represents the conditional probability distribution of the sample value x_i of homogenous samples X_i in the attribute table. The conditional probability θ_{X_i} / Pa_{X_i} at each pixel is given by $p(X_i / Pa_{X_i})$ following the equation (2.2).

$$p(X_1, X_2, \dots, X_n) = \prod_{i=1}^n p(X_i / Pa_{X_i}) = \prod_{i=1}^n (\theta_{X_i} / Pa_{X_i}) \quad (2.2)$$

Where $n = |N|$ are parents; thus a Bayesian network can infer the spatial distribution of the parent Pa given X representative and homogenous samples. A Bayesian network can also be used as a classifier because it produces the posterior probability distribution of a class node, given their corresponding values (Aguilera, Fernández, Reche, & Rumí, 2010; Bressan et al., 2009; Cheng & Wang, 2010; Qu, Wang, Wan, Li, & Zhou, 2008).

2.2.3.2 Building an expert Bayesian network classifier

When building an expert Bayesian network classifier, a number of information about a target features and their corresponding values are needed to estimate probabilities of membership. Such information is structured in the DAG and learning algorithms are used calculate the conditional probability distribution from data (Bressan et al., 2009) according to predefined rules and thresholds by the subject specialist. Figure 2-2 shows the DAG structure that is developed to automate the delineation of coffee fields in orthophotos by the Bayesian network classifier.

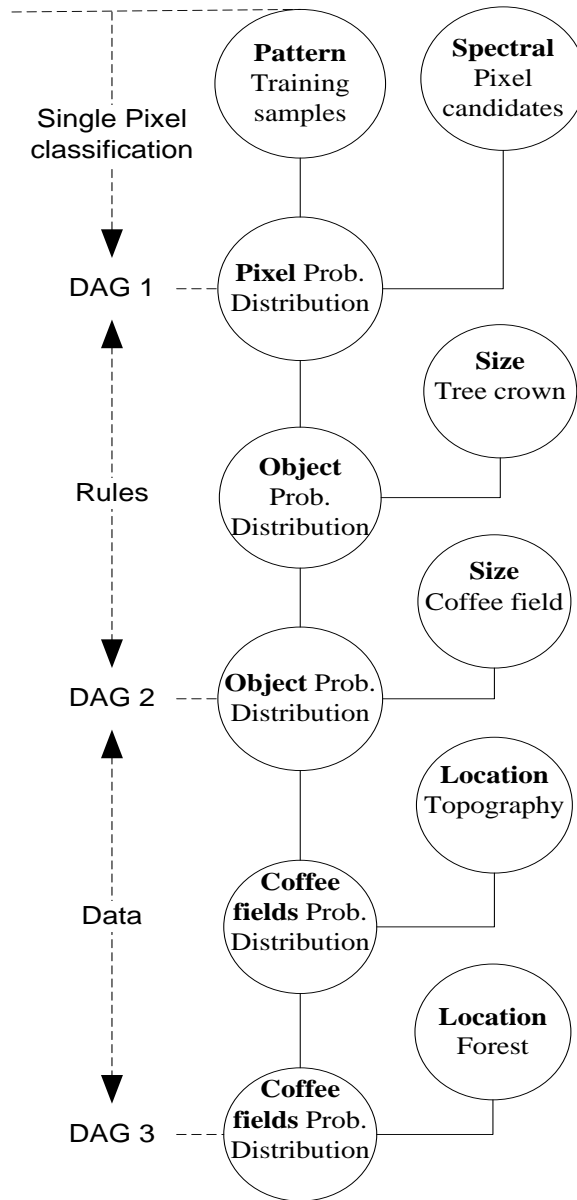


Figure 2-2: A Directed Acyclic Graph (DAG) used to train the expert Bayesian network model for coffee mapping. DAG 1 represents the naive Bayesian network classifier that distributes the probability of coffee based on the spectral properties of pixel candidates and the pattern of coffee fields used as training samples. DAG 2 and 3 represents the expert knowledge in the Bayesian network model. DAG 2 introduces tree and field rules whereas DAG 3 introduces flow accumulation and forest data to model the location of the coffee field in the toposequence and land use systems.

Step 1. Pixel-based classification: a naive Bayesian classifier

Let C_{x_i} be a set of coffee field and N_{x_j} a set of no-coffee field parents i.e. fields in which coffee and no-coffee samples are randomly selected. Pixel candidates for coffee and no-coffee samples are given by X_i and X_j . x_i, x_j are the DN values of the pixel candidates of the three bands for coffee and no-coffee pixels. The distribution of coffee field pixels is determined by the distribution of coffee pixels candidates $\{\theta_{x_i} / C_{x_i} = p(X_i / C_{x_i})\}$, given the distribution of no-coffee pixels candidates $\{\theta_{x_j} / N_{x_j} = p(X_j / N_{x_j})\}$ used to train the Bayesian model. N_{x_j} is also called the background class. This no-coffee class may contain banana, forest, mixed crops, bare soils, or buildings. The dependence relation between pixel samples values (x_i and x_j) allows the calculation of conditional probability $\langle \theta \rangle$ distribution. This conditional probability is deduced from the spectral characteristics (i.e. DN variability in the three bands) learned from coffee pixel samples. In this first level classification, a classifier assigns the probability of coffee pixel candidates $X_i = (X_1, X_2, \dots, X_n)$ to belong to coffee class C as shown in Equation (2.2). Higher probability values are assigned to pixels spectrally similar to coffee pixel samples. Lower probability values are assigned to no-coffee pixels or otherwise significantly different from coffee. The joint probability distribution is a product of the probability distribution of DN values ($0 \leq x_i \leq 255$) of the coffee pixel candidates given the probability distribution of DN values of the no-coffee ($c = 0 - 255$), as given by Equation (2.3):

$$p(X_i / C) = p(x_1, x_2, \dots, x_n) / c = \prod_i^n p(x_i / c) \quad (2.3)$$

Where x_i denotes DN values of the pixel candidates X_i , c representing DN values of coffee field samples C_{x_i} . The resulting Bayesian classifier is then:

$$f_{nb}(C) = \frac{p(c_i = C)}{p(c_i = N)} = \prod_{i=1}^n \frac{p(x_i / c_i = C)}{p(x_i / c_i = N)} \geq 1, x_i = C, \text{ otherwise, } N \quad (2.4)$$

Where $f_{nb}(C_i)$ is the naive Bayesian network classifier (n-BN). This classifier is referred throughout in this study as a single Pixel Probability Classifier (PPC). The output of $f_{nb}(C_i)$ is a pixel probability with values $\{x_i / c_i = 0 - 1\}$. The value of $f_{nb} = 1$ corresponds to the probability threshold above which the spectral dependence supports the classification of coffee. The second stage of

the naive Bayesian classifier consists of Object Probability Classification (OPC). In this stage, we used the probability threshold to convert the probability map to a binary coffee/no-coffee map. In order to classify a pixel as coffee or no-coffee, we assume that its direct neighborhood provides significant additional information. Following Bressan et al. (2009) we applied the four neighbor rule to eliminate single pixels of no-coffee within a cluster of coffee pixels. Single coffee pixels of coffee within a cluster of no-coffee are also eliminated accordingly. The four neighbor operation only retains class if its four neighbors have that class. The resulting image objects remain with pixels with only probability greater than the probability threshold ($P > 0.49$).

Step 2. Object-based classification: an expert Bayesian classifier

The second step introduces the spatial characteristics of coffee tree crown and the coffee field in the Bayesian network model as spatial evidence. Knowledge about the Rwandan coffee systems is used to build rules into the network structure. These rules are a set of conditions defined by the subject specialist in the DAG to optimize the conditional likelihood for coffee fields given surrounding no-coffee fields. The local dependence i.e. "direct neighbor" probability is derived from the conditional probability pixel outputs $\{x_i / c_i\}$ of step 1. Equation (2.3) is consequently extended to Equation (2.5).

$$p(a_i / X_i) = p(a_1, a_2, \dots, a_n) / p(x_i / c_i) = \prod_{i=1}^n p(a_i / p(x_i / c_i)) \quad (2.5)$$

Where a_i denotes the derivative probability value from the output image pixels probability of the DN value c_i , and a_i represent the spatial nodes predefined by the subject specialist. The joint probability distribution $\langle \theta_{a_i} \rangle$ is determined by the derivative probability $dp(x_i / c_i)$ translating the local dependence (i.e. local structure) of coffee C_{X_i} given the structure of no-coffee N_{X_i} in the Rwandan agro-ecosystems.

In the Bayesian network structure, a_1 introduces the Tree Crown rule-set Classifier (TCC). The coffee average crown diameter is used to set a threshold in the Bayesian network model in order to calculate the joint probability distribution θ_{a_1} as derivative of probability of OPC output. The average coffee tree crown size used in this study is thus 16 pixels corresponding to the average coffee tree crown diameter (2 m in our case) or in other words, the average of coffee plant spacing (2 m \times 2 m). The mean derivative probability for each cluster is calculated using 4 \times 4 kernel

functions. This procedure will discard young coffee trees and pruned coffee trees with smaller tree crowns from the analysis.

After the tree crown classifier, a node a_2 representing the Field Size rule-set Classifier (FSC) is introduced in the Bayesian network structure. Coffee tree probability output is aggregated into a coffee field probability. We set the threshold to 3200 pixels corresponding to the average minimum field size of 200 trees as reported by (Nsengiyumva, 2009). The field size rule therefore filters out coffee fields that are below the threshold and recomputed the new probability of coffee field object in the conditional probability attribute table. In some cases, pruned coffee trees may be recuperated by the field size classification as long as these pruned trees are located near the coffee trees with larger crowns. The boundary of the coffee field is delineated at this stage. Only if edge-rows contain pruned trees these will remain excluded. Unlike the rule-set classifiers, the location values (a_3, a_4) are added in the Bayesian network as data-set classifier. In this way the DAG2 is extended to DAG 3. We used flow accumulation data from the digital terrain model to remove the classified coffee fields in gullies and flood plains. The national forest map of 2007 (Westinga, Mukashema, & van Gils, 2013) was finally used to remove classified coffee fields in known large forest areas to reduce errors due to the poor separability between coffee and forest. These data do not only filter but also the final probability of the coffee polygon is accordingly adjusted to the total number of coffee pixels. The expert-based Bayesian network classifier for coffee becomes, after a series of a_i nodes and on basis of coffee and no-coffee pixel candidates, as follows:

$$f_{eb}(C) = f_{nb}(C) \prod_{i=1}^n \frac{dp(a_i / x_i = C)}{dp(a_i / x_i = N)} \geq 1; DF = \prod_{i=1}^n dp(a_i / x_i) \quad (2.6)$$

Where x_i represents spectral dependence variables, a_i represents the spatial dependence variables used to build the Bayesian model. The dependence factor (DF) is determined by the product of the derivative probability of object nodes a_i given the spectral nodes (x_i) used to build the model. The DF also reflects how each of the local dependences (local structure) distributes pixels to coffee class C and how all of them worked together to produce a better classification result for coffee. The dependence factor can be used as a measure of the model sensitivity since it quantifies percentage of the pixels changing probability/uncertainty at each stage of the Bayesian network classifier.

2.2.4 Country-wide implementation of the expert Bayesian network

The Bayesian network was applied to 198 orthophotos covering almost 99% of the entire country (26,338 km²) and one QuickBird image covering the remaining part in Impala region. For the QuickBird, the same steps were followed, but tree crown and coffee field threshold values were different due to the different pixel size. The tree crown threshold was set to 9 pixels corresponding to a circle with a diameter of 2m; the mean probability of the tree object was calculated using 3 × 3 kernel function. For the field classification a field size threshold of 1300 pixels was used, which corresponds to 200 trees. Unlike on orthophotos, we also used the near infrared (NIR) band for classifying the QuickBird image. The accuracy of our resulting national coffee map was assessed against the 1362 coffee and no-coffee locations using Kappa statistics, Root Mean Square error (RMSE) and the coefficient of determination (R²). Moreover, the mapped coffee area was aggregated to district (C_m) and compared to the district coffee areas reported by the detailed national coffee census of 2009 (C_c) as shown by Equation (2.7).

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^n (C_{m(i)} - C_{c(i)})^2} \quad (2.7)$$

Besides a large R² and a low RMSE, the slope of the regression should ideally be close to 1, and the offset close to 0.

2.3 Results

2.3.1 Bayesian network results for a test area

Bayesian network results for a test area Figure 2-3 presents the coffee probability distribution results in the test areas (0.4 km² in Huye district, Southern province). In these test areas, coffee is relatively more prevalent as compared to the country average. This is why random sampling of the test areas resulted in a higher proportion of coffee samples as compared to the proportion that would be obtained when randomly sampling for the whole country. The location of the potential coffee (P ≥ 0.49) is shown by the probability maps (left) and the number of pixels and their likelihood to coffee (C) or no-coffee (N) is shown by the box plot (right). With the pixel probability classifier (PPC) (Figure 2-3a), the mean probability threshold is first determined (PT = 0.49). The probability threshold constitutes a cut-off below which the pixel probability classifier considers pixels as no-coffee pixels. The results showed that only 25% of the total image pixels were likely to be

coffee ($P \geq 0.49$) with only 0.5% of the pixels having probabilities between 0.49 and 0.75, and 24.6% of the pixels over 0.75.

In Figure 2-3b, the object probability is assigned the mean probability to the coffee object produced after the four-neighbor analysis. This operation slightly increased the percentage coffee pixels from 25% to 27% of the total image pixels.

Figure 2-3c shows that the likelihood for coffee decreased when tree crown size rule was applied. About 27% of total image pixels previously classified as coffee by the object probability classifier decreased to 22% due to the tree crown size criterion. The probability for coffee also decreased from 0.75 to 0.50 showing that the use of tree crown size as an explanatory variable was an important step to distinguish coffee from other tree crops.

A clear distinction between coffee and no coffee was observed when a minimum field size rule was applied to the tree crown probability output. The resulting coffee probability map demonstrated an increase of coffee fields from 22% to 28% of the total image pixels with only 8% highly confirmed as coffee fields ($P \geq 0.75$). The level of detail however reduced and outliers were removed (Figure 2-3d). Forest and flow accumulation data are added in the final step to filter out coffee within known forest areas and in channels and flood plains. Coffee probability is also adjusted to remained coffee pixels. Only 15.5% of total image pixels remained as coffee fields as shown in Figure 2-3e. The size of the test sites used to build the model is about 0.4 km² located in Huye.

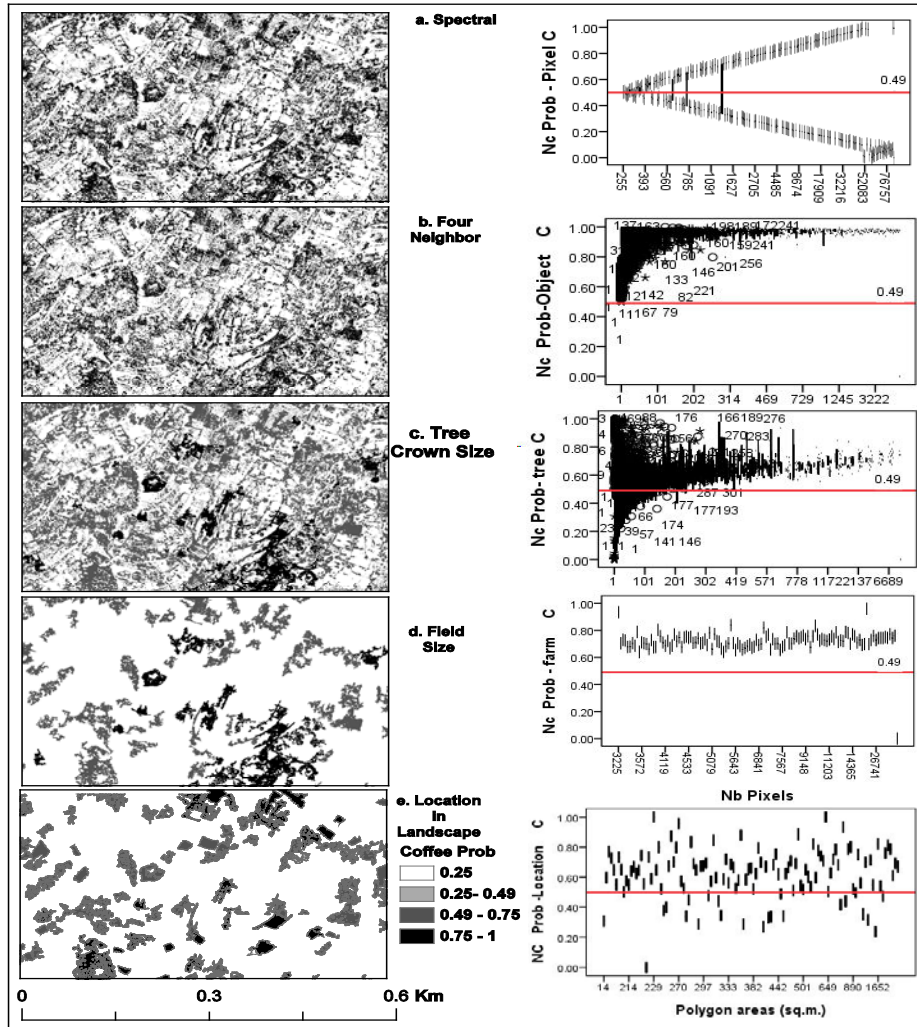


Figure 2-3: Intermediate and final coffee classification results after each step of the expert Bayesian network model. The location of the potential coffee ($P \geq 0.49$), is shown by the probability maps (left) and the number of pixels and their likelihood to coffee (C) or no-coffee (N) is shown by the boxplot (right). Top-down: a. Pixel probability result of spectral analysis; b. Object probability result of four-neighbor classification. c. Coffee objects are clustered by the Tree Crown Classifier. d. The Field size classifier aggregates the tree objects into coffee field objects and adjusts probability of coffee field. e. Final coffee probability map after the forest and flow accumulation data are added in the model to filter out coffee within the forest and in the channels. O* represent the number of pixels that are deemed to be outliers.

Step-wise addition of factors to the Bayesian network resulted in a gradual decrease in the percentage of image pixels identified as coffee for the test area (Figure 2-4). With the pixel-based classifier, an overall accuracy of only 48% was obtained. After application of the subsequent steps of the expert Bayesian network model, the overall

accuracy was 87.6% with a Kappa value of 0.82 for the test area (Table 2-1). This shows that for the test site the Bayesian network could accurately extract coffee fields from the ortho-images.

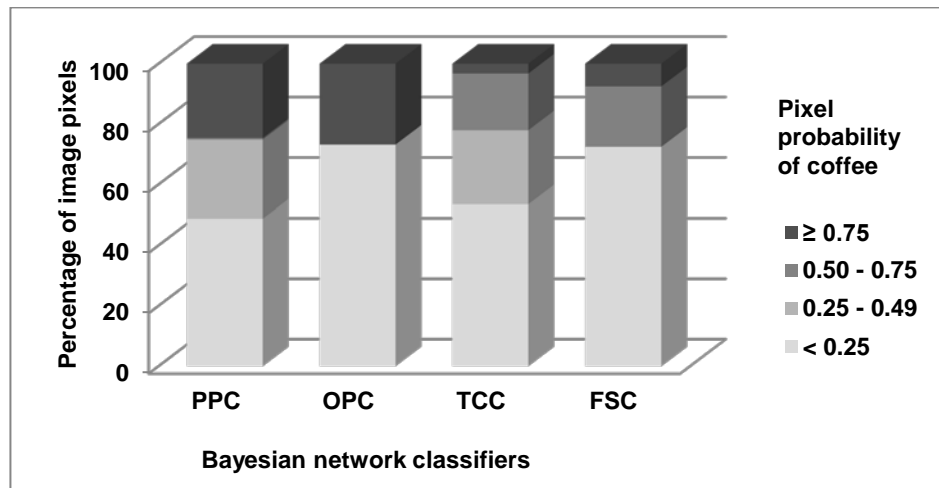


Figure 2-4: The percentage of pixels that falls within each probability class for five classification steps of the expert Bayesian network model. These steps are (1) Pixel Probability Classification (PPC); (2) Object Probability classification (OPC), (3) Tree Crown Classification (TCC), (4) Field size classification (FSC), and (5) filtering by the forest and flow accumulation map that filter out coffee within the forest area and in the water channels. The test site is about 0.4 km² and located in Huye district.

Table 2-1: Error matrix of observed coffee fields vs. classified coffee fields by the expert Bayesian network (e-BN) for the test site (0.4 km²) in Huye, South plateau of Rwanda. The matrix values were calculated on the final coffee polygons. Overall accuracy of the test model were obtained by considering all pixels with $P > 0.49$ as coffee and $P < 0.49$ as no-coffee. Overall accuracy is 87.6% with Kappa value of 0.82.

	Reference fields			Total
	Prob.	Coffee	No-coffee	
	≥ 0.75	24	0	24
	0.50 - 0.75	38	0	38
	0.25 - 0.49	14	0	14
	< 0.25	0	58	58
	0	3	0	3
Total		79	58	137

2.3.2 Accuracy of the Rwanda coffee map

Application of the trained expert Bayesian network to all images resulted in a total coffee area for the whole country of 25,148 ha, corresponding to 1.4% of the total arable land. The maximum coffee area (2979 ha) was found in Nyamasheke district, and the minimum area (5 ha) in Burera district. Figure 2-5 shows the coffee coverage map aggregated by 5 km × 5 km grid (Figure 2-5a) and by district boundary (Figure 2-5b). The maximum coffee coverage

within a single 5 km × 5 km pixel is about 13.5% and was found in Rusizi and Nyamasheke (west), Nyaruguru (south), Ruhango (central) and Gatsibo (east).

The district accuracy of the produced coffee map ranges from 56% to 100%; with an average accuracy of 83.4% for the whole country (Figure 2-5b). We observed agro-ecological dependent differences in the accuracy results which may be related to the level of separability between coffee and other land use classes. The accuracy was higher in Congo-Nile water-shed (97.1% in Rutsiro), Mayaga Plateau (94% in Ruhango) and in the Eastern Ridges and Plateaus (over 97% in Gatsibo and Gicumbi). This is likely because in these areas mostly older trees are present with a better developed tree crown. In fact, the Rusizi, Nyamasheke, Karongi and Rutsiro districts in west Rwanda are the first established (in the 1930s) coffee areas and have high production levels (Nsengiyumva, 2009). However, for Rusizi the model performed worse than for other western districts. In Rusizi QuickBird covered more than 50% of the district area. The time of the QuickBird acquisition (April 2004, older than orthophotos) and the lower resolution of QuickBird may explain the lower accuracy for Rusizi. The Bayesian model yielded an accuracy of 76.9% on QuickBird. Other low accuracies (68% and 70%) were observed in newly established (in 2000s) coffee fields of districts of the Eastern province (Nyagatare) and of Kigali (Nyarugenge). According to the national coffee census (Nsengiyumva, 2009), the highest rate of young coffee trees is observed in the Eastern Province (36%), followed by Kigali City (35.5%). Young coffee trees do not satisfy the tree crown rule in our expert Bayesian network. Moreover, in young coffee plantations line spacing i.e. large gaps between trees is common. When these gaps, classified as no-coffee by the tree crown classifier, are relatively small, they are included in coffee field by the field classifier, but for large gaps this is not the case.

The environment in which coffee fields are located also matters. For example, in high mountainous areas, accuracies were higher as compared to lowland areas. Except for coffee fields established in Mayaga, Bugesera and Imbo lowlands due to the coffee intensification policy during the period from 1960s to 1980s (Verwimp, 2003), more coffee fields are located at high altitudes (above 1500 m a.s.l.). These coffee areas experience relatively low temperatures (15–20°C), high rainfall (>1000 mm per year) during ¾ of the year and consequently develop very well their branches and leaves (i.e. a good crown development).

Another factor that may explain differences in accuracy is the farming system in place. Larger coffee fields are found in the western mountains and eastern lowlands as compared to the central plateaus and ridges. Very highly fragmented and over-exploited lands of the central plateaus inhibited classification of very small coffee fields as these were ruled out by the field size criterion. The health of coffee trees in such fields is also poor leading to

reduced accuracies in the plateau districts as opposed to other agro-ecological zones.

We compared the mapped coffee fields from the orthophotos of 2008/2009 with the national coffee census of 2009 (Figure 2-5c). The 2004 QuickBird based inventory was omitted in this comparison because of the large time difference between image acquisition and census. Our spatially aggregated coffee map shows a strong agreement with the census data ($R^2 = 0.92$). This shows that both methods agree in most districts despite the underestimated coffee areas of Rusizi due to the earlier described problems of QuickBird data. The root mean square error (RMSE) revealed a total of 256 ha reported by the census that could not be explained by the mapping results. This difference may be attributed partly by the fact that pruned trees and young trees were not accounted for with our mapping approach to loss of coffee tree by the tree crown classifier due to the pruning of mature trees and stage of development for the younger trees. It can also be a result of average plant spacing used to estimate the area in the census. In fact coffee spacing varies from 1.5 m \times 1.5 m to 2.5 m \times 2.5 m with an average of 2 m \times 2 m but for the estimation of the area, OCIR-CAFE used the average (2 m \times 2 m) for the whole country.

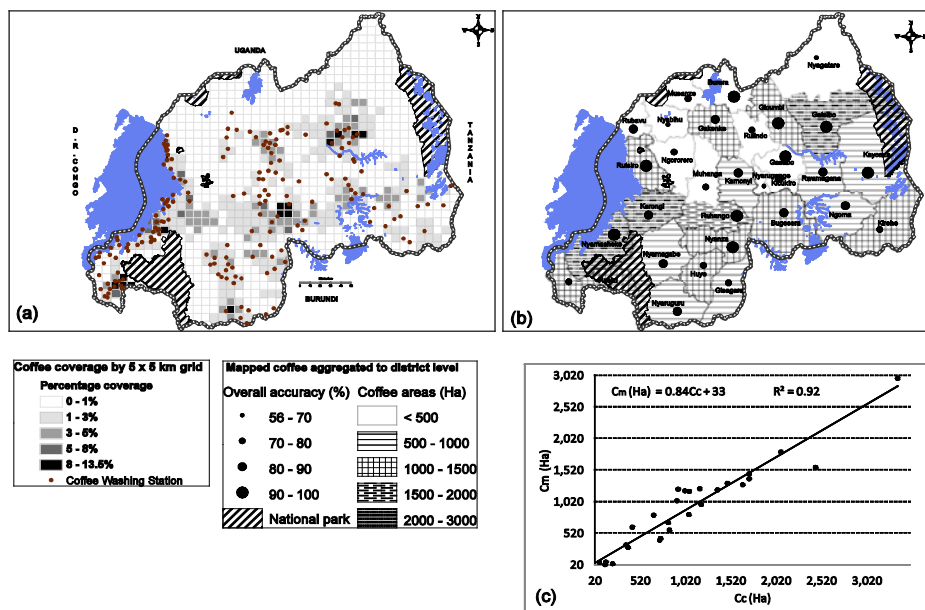


Figure 2-5: Aggregated national coffee cover results following nation-wide application of the e-BN model to aerial orthophotos: (a) percentage of coffee cover within a 5 km \times 5 km grid cell. (b) Coffee areas aggregated per district and overall accuracy for each district calculated based on 1362 GPS locations of coffee and no-coffee in the 30 districts of Rwanda. The graph (c) plots mapped coffee (Cm) vs. to reported coffee area (Cc) by the national coffee census of 2009. The coffee map explains 92% of the spatial variation of the census data.

2.4 Discussion

Discussion Coffee trees and fields show a large spectral variability on orthophotos and QuickBird imagery. The Bayesian network structure proposed in this paper provided a possibility to handle such complexity through a stepwise object-based classification by modeling the probability of coffee trees and fields based on their surroundings. The conversion of expert knowledge on key spatial characteristics of coffee cultivation in Rwanda into evidence variables is the key to the model. In our case, spectral and spatial evidence of coffee occurrence were combined to gain a better representation of coffee farming systems design in Rwanda.

2.4.1 Effect of individual Bayesian steps on the scale and the quality of the final coffee map

The sequence of classifiers as a result of the number of evidences for coffee occurrence is an important process in the modeling of coffee distribution. From a single unit (pixel) to an object unit (field) through a subunit (tree), a number of aggregation steps are involved. The hierarchy of the classifiers is carefully defined to reach a desired accuracy. Dropping a step or reversing the order of classifiers will affect the accuracy. For example, when the four-neighbor coffee pixels is ignored, the tree crown classifier will include about 34% of the total image pixels of no-coffee in coffee pixels.

Aggregation is an important component in building the e-BN structure. During the subsequent steps of a Directed Acyclic Graph, the unit size of the resulting probability map increased. For example, scattered coffee pixels in the pixel probability classification output when the minimum Coffee Mapping Unit (CMU) is set to 1 pixel tend to decrease their abundance in the OPC output (i.e. CMU = 4 pixels). When the tree crown rule is set to the average tree crown size (i.e. CMU = 16 pixels), the abundance of coffee is reduced. Yet, when the field size rule is included to allow only the coffee field with 200 coffee trees and above (i.e. CMU \geq 3200 pixels) the number of coffee pixels decreased and the uncertainty also reduced. The CMU is therefore aggregated to a coarser scale (1:3200). The level of detail in field size output is significantly reduced but the outlier pixels are also reduced accordingly. The effect of aggregations depends on the original resolution of the image used and the steps as predefined by the expert in the BN model. It is known that aggregation of spatial data usually obscures internal variability and uncertainty (Kok & Veldkamp, 2001; Veldkamp, Kok, et al., 2001).

In this work we did not consider the possibility of polyculture systems in model training. Coffee grown in polyculture fields (if any) may have been mapped with low probability (i.e. high uncertainty) because of generalization. To better account for differences in field types and field sizes within the country, thresholds could be adapted across the country. Due to the choice of

the minimum field size, very small plots of less than 8 acres were lost. Field observation indicated that such small plots are rare and associated with abandoned fields. Depending on the purpose and scale, for highly fragmented zones, one could consider changing thresholds for adapting the method for local context. However, for nation-wide coffee mapping, we considered the average thresholds sufficient to capture more than 90% of the coffee plantations, and this has been proven by a strong agreement (92%) between our mapping and the National census results of the same year. In fact the census counted the number of trees regardless the size of the field (Nsengiyumva, 2009).

2.4.2 Performance of the expert Bayesian network

Firstly, our results confirmed that spectral variables alone are not sufficient to accurately identify coffee fields from high resolution images. With pixel probability classification, each coffee pixel is considered to be independent. However, this assumption is unrealistic given that we deal with coffee farms that are spread over a large area (more than 17,000 km² of arable land) across different environments. Information on the key characteristics of farming systems in place is an important step to the design an automated approach to classification of coffee. For a small area of 1200 km², Gomez et al. (2010) obtained a good accuracy when predictive variables were used together with object oriented image analysis using Artificial Neural Network (ANN). In general, the accuracy of the results of Bayesian network models (and of other object-oriented models) likely increases when more relevant variables can be provided. Another observation in our study is that the spatial resolution (i.e. pixel size) seems more important than the spectral resolution (i.e. the number of bands) in separating coffee from other land covers. Although QuickBird contained the additional NIR band, we obtained a relatively low accuracy (76%). An additional cause of this might however be that the QuickBird imagery are acquired five years before the 2009 census. Given the poor spectral separability between low-density woodlands and coffee fields (Cordero-Sancho & Sader, 2007), we improved coffee field identification in mountain forest areas (in west) and in the plateaus and ridges of the east (in Gatsibo) by the use of an existing forest map. The use topographic data in the Bayesian network model was also relevant to exclude coffee from clear flow accumulation areas (streams and gullies).

Although we obtained overall high classification accuracy with our approach, several factors may cause sub-optimal performance. (1) Topographical effects may have contributed to the spectral confusion between east-facing and west-facing coffee trees that are differently illuminated due the sun's position during the time of image acquisition. This effect also caused a reduction of accuracy in other studies; (2) The Bayesian network model

classification was not corrected for shadow of coffee trees; (3) Differences in coffee age and health also account for a loss in accuracy.

In the last decade remotely sensed data is often combined or compared with census data to study environmental dynamics. The main challenge of this integration is the different temporal and spatial scales involved. This is often solved by aggregating the two data sources to the coarsest resolution of the two sources. This has as main disadvantage that some of the detail and variability is lost but it allows to convert typical land cover information into more in land use characterization. So the integration and/or combination of census and remotely sensed data always come with a quality trade-off. The strength of rigorously collected census data is that they are based on on-site inventories and farmer interviews. Nonetheless, spatial detail and verification of all farmer-reported coffee trees is lacking in the census outcomes. Remote sensing, i.e. our high-resolution national coffee map, can therefore offer an important complement to the census data. However, as discussed above, it has its own limitations in terms of accuracy. Despite existing limitations for both sources, in our view they offer important complementary information.

The contribution of our method to existing object oriented classification techniques is the stepwise analysis that allows the creation of intermediate classification outputs. This means the flexibility of the method in shortening or extending rules and data towards improving quality of the map. Potentially more use could be made of the probability value that was derived for each coffee field, for example as an indicator of the health and productivity of the coffee field.

2.5 Conclusions and outlook

In this study we developed a stepwise probabilistic approach for coffee systems classification using an expert Bayesian network. Using the case of Rwandan coffee, the methodology proved to be effective in extracting small-scale coffee fields from very high resolution images. We are convinced that the presented method can be extended (1) to other perennial small scale cropping systems (e.g. banana, cassava systems) that are difficult to map with spectral-based pixel classification techniques. As such the method has great potential to be applied for coffee and other perennial crops (like banana, cacao) across the East African highlands, but also for mapping of perennial crops in small-scale farming systems elsewhere for which the detailed spatial distribution is currently largely unknown.

The method provided an accurate representation of the spatial distribution of coffee cultivation areas of Rwanda despite the spectral diversity of coffee trees and the field sizes. We produced the first nation-wide coffee map for Rwanda. This automated high resolution coffee map contains the producing coffee fields at an overall accuracy of 87%. However, some districts have a

lower accuracy, which could potentially be improved with local adaptation of decision rules in the Bayesian network and interactive phase especially for the young or regenerated plantations after cuttings. The produced high spatial resolution coffee map is an important input for further environmental and socio-economic studies regarding the productivity and quality of Rwandan coffee.

3 Sixty percent of small coffee farms have suitable socio-economic and environmental locations in Rwanda

This chapter is published as: Mukashema, A., Veldkamp, T., & Amer, S. (2016). Sixty percent of small coffee farms have suitable socio-economic and environmental locations in Rwanda. *Agronomy for Sustainable Development*, 36(2).

3.1 Introduction

Sustainable production of perennial crops such as coffee, tea and cacao in small holder systems, is often a function of human decisions based on location factors, local habits and traditions. These decisions are also influenced by economic demographic and infrastructural factors that can be very dynamic in space and time (Lambin et al., 2001). Quite often perennial crop distribution has a historical component, reflecting past conditions including competition with other crops and land use functions (Veldkamp, Verburg, et al., 2001). Therefore, simplistic relationships explaining spatial crop distribution are rarely found. Interestingly, land evaluation approaches using only agro-ecological zones and soil information are still often used to identify locations for crop expansion (Verdoodt & Van Ranst, 2006). In order to allow more adaptive land use planning it is essential to understand all underlying location factors that explain perennial crop distribution patterns, including the socio-economic factors (Verburg & Veldkamp, 2001). Perennial cash crops in particular are important because they are often the least dynamic cropping system for small holders, generating revenues for farmers and countries.

One of the primary cash crops in Rwanda is coffee, but the question which factors govern its precise geographic pattern remains unanswered. There are of course general relationships with climate, soil and local management (I. Nzeyimana et al., 2013). But these simple relationships do not explain the current coffee yields nor field density distribution across the country. There are two important coffee area inventories for Rwanda, one is the national census (Nsengiyumva, 2009), and the other is an image-based inventory (Mukashema, Veldkamp, & Vrieling, 2014) for all fields with more than 200 coffee trees. At the district level, these inventories have a satisfactory 92% fit. Rwanda is the most densely populated country in Africa with 415 pers./km² and a total population of 10.5 million in 2012 (NISR, 2012). Coffee competes with food crops for the limited land available. Agricultural land (including coffee) take up 54.2% of the total country territory, the remaining 45.8% comprise of forest (28.8%), wetlands and water bodies (11 %) and urban areas (6%). As with other cropping systems, the spatial distribution of coffee is a function of environmental drivers (soil, climate and topography) and human drivers (demography, infrastructure, technology and institutions)(Hernández-Martínez, Manson, & Hernández, 2009; Veldkamp & Fresco, 1997; Verbist, Dinata Putra, & Budidarsono, 2005). Up to now, suitability for coffee cultivation in Rwanda was only assessed at the level of agro-ecological zone which is a highly aggregate spatial scale (Verdoodt & Van Ranst, 2006). Recently, a more detailed soil-based approach using kriging and multi-criteria analysis was developed (I. Nzeyimana et al., 2014). Both approaches did include neither socio-economic factors nor the potential dynamic climate factors making both approaches not sensitive for the rapid

changing Rwandan socio-economic conditions and the expected climate change. The later can be significant in the nearby future as illustrated by Nyandwi, Veldkamp, and Amer (2016) who demonstrated that an increase of 1% in annual temperature can cause a net wetland area decline by 12 %, making these areas suitable for other land use. It is our objective to analyse the spatial coffee field distribution using the new high resolution coffee map (Mukashema et al., 2014) and using available biophysical and socioeconomic factors that co-determine coffee cultivation locations in Rwanda. When these relationships are known, we will explore how sustainable the current coffee location distribution is and identify potential new suitable locations for coffee production.

3.1.1 Historical background of coffee distribution in Rwanda

According to local reports, the first coffee fields were established by missionaries in Nyundo (northwest) in 1903 and in Mibilizi (southwest) in 1904. Official government involvement began in the 1930s with the Belgian colonial government's coffee campaigns (Boudreaux & Ahluwalia, 2009; Murekezi, 2009). Under these policies, government authorities built nurseries and supplied seeds, but they also forced farmers to plant coffee trees. Coffee cultivation slowly became accepted by the population and gradually spread over the country. Farmers established small coffee fields often surrounded by larger banana fields. By 1930, the expansion slowed down due to famine. After World War II, the coffee area expanded again due to the higher coffee price on the global market. In 1946, an estimated total of eight million coffee trees (about 3200 ha) was reached. After independence (1962), the government of Rwanda continued the policy of requiring farmers to grow coffee and with the implementation of a 10-year coffee plan in the period 1950–1960, together with a resettlement plan, coffee areas expanded from 12,710 ha in 1964 to 23,930 ha by 1970. Farmers were also forbidden to interplant other crops with coffee (Boudreaux & Ahluwalia, 2009). In 1974, the World Bank supported the expansion of coffee growing alongside Lake Kivu, causing an increase of coffee acreage increased to 31,900 ha by 1981. The Government built coffee nurseries and supplied seeds all over the country to facilitate farmers' access to the selected variety of Bourbon coffee Arabica (Verwimp, 2003). During the worldwide coffee crisis of the late 1980s, however, prices tumbled and caused a slowdown of coffee expansion in Rwanda. With their drop in income, farmers were tempted to shift towards production of other cash crops such as bananas, but this was forbidden by Rwandan law, and with an extensive local control system in place, only a few thousand coffee trees were uprooted. The expansion of coffee cultivation areas resumed from 2000 onwards and continues to date mainly at the expense of forest land (UNEP, 2011).

Based on recent coffee mapping (Mukashema et al., 2014) and the agricultural census (Figure 3-1), we know that the western region is the main coffee-growing zone covering 4.4% of the arable land at the Kivu lake shore and Imbo zones followed by Impala zone (2.9%), then the central and eastern plateaus (1.5%) and Mayaga and Bugesera in southeast (1.2%). The Buberuka highlands and high plains of the volcano zones are currently not used for growing coffee. In fact, there are very few coffee fields located at the edge of Buberuka highlands near the north-eastern plateau and ridges.

3.1.2 Characteristics of coffee locations in Rwanda

Coffee in Rwanda is currently grown under temperature conditions ranging from 14.2 to 23.8°C in areas with annual rainfall varying between 700 mm in the East and 2120 mm in the Congo-Nile watershed divide and the Buberuka highlands. Coffee areas are characterised by a relative humidity varying between 67 and 82 % at altitudes varying between 970 and 2575 m. Ideally, coffee trees have to be planted at altitudes below 2000 m (Wintgens, 2009). Coffee fields are typically located in areas with slopes varying between 8 and 89%. With regard to lithology, 63% of the coffee fields are established on soils developed on shale, and mica schist, 23.5% on granite and 9.7% on basic rocks. Very few fields (1.4%) are established on volcanic soils with high base saturation and high pH. Such soils provide considerable challenges for coffee production. The relationships between coffee growing and soil properties have been described by I. Nzeyimana et al. (2014). There are two main problems with soil fertility, (1) aluminium toxicity which is very often associated with low pH and low organic carbon content, (2) low levels of available phosphorous.

Coffee farmers require access to public and/or private transport facilities for the transport of both inputs (fertilisers, pesticides, mulch etc.) as well as outputs (coffee berries). About 94% of coffee fields are located within 1 km from the primary and secondary roads. A different situation, however, unfolds if farmers have to travel to bring their coffee cherries to the nearest coffee washing station for processing. Only 9.3% of the coffee fields are located within 1 km distance from a coffee washing station. Even within the same region, cultural and social differences, and family size have been identified as important factors which influence farmers to grow coffee (Wintgens, 2009). In Rwanda, due to land scarcity, the food security of a household generally takes priority over cash cropping.

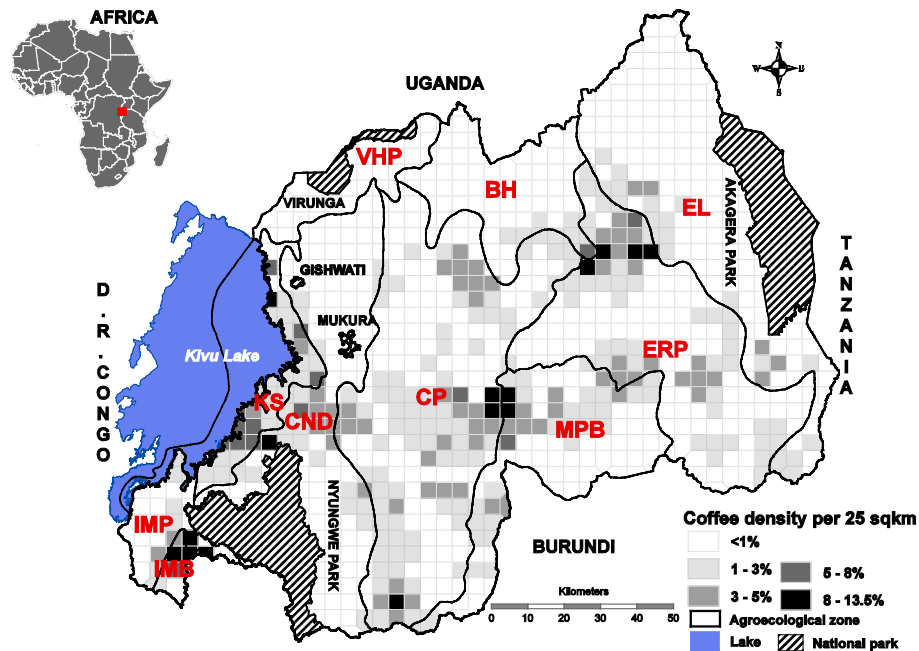


Figure 3-1: Coffee density (%) within 5×5 km grid. Ten main agroecological zones of Rwanda: (1) KS Kivu lake shore, (2) IMP Impala, (3) IMB Imbo, (4) CND Congo-Nile watershed divide, (5) MPB Mayaga plateau and Central Bugesera; (6) CP Central Plateau, (7) ERP Eastern Ridge and Plateau, (8) EL Eastern Lowland, (9) BH Buberuka Highlands, (10) VHP volcanoes and high plains. The country size is 26,338 km².

3.2 Materials and methods

3.2.1 Biophysical and socio-economic data

The biophysical and socio-economic data sources used in this study are briefly described:

1. Coffee fields' data (i.e. location and size of fields with more than 200 coffee trees) was produced by (Mukashema et al., 2014). The total of 119,513 coffee fields observed were randomly split into a training set (59,757 coffee plots) and a validation set (59,756 coffee plots). The randomising was done to avoid spatial autocorrelation and duplications of variables as advised by Overmars, de Koning, and Veldkamp (2003).
2. Soil data were acquired from the soil geo-database of Rwanda located at the Ministry of Agriculture and Animal Resources. This soil database is a result of a semi-detailed soil survey of Rwanda which started in 1981 and was finalised in 1994. A physical and chemical soil analysis was conducted for 1833 soil profiles spread over the country. Cation exchange capacity (CEC) was measured in the B horizon or at 0.50 m depth, the sum of the basic cations Ca²⁺, Mg²⁺ and K⁺, pH in water and organic

carbon and nitrogen content of the upper 0.25 m of mineral soil. This soil survey resulted in the elaboration of 43 soil maps at the scale of 1:50,000 covering the Rwanda territory. More details on the process of producing soil maps and database can be found in Verdoort and Van Ranst (2006).

3. Climate data i.e. temperature, rainfall, relative humidity and potential evapotranspiration were obtained from the National Meteorological Service at the Ministry of Infrastructure. Average annual data were calculated for 183 weather stations using, for most stations, 60 years records from 1950 to 2010. We interpolated these data using the thin-plate smoothing spline algorithm as proposed by Hijmans, Cameron, Parra, Jones, and Jarvis (2005).
4. Topographic data i.e. altitude, hill shade, slope and aspect were generated by processing 10 m-resolution digital terrain data produced using stereoscopy (See Mukashema et al., 2014).
5. Distances from each coffee field to infrastructural facilities were obtained by processing the data of coffee fields, GPS locations of coffee washing stations and the road network. The road network was earlier digitised using the topographic map of 1988 at a scale of 1:50,000 and updated by CGIS.
6. Population data of 2012 i.e. the total, density and age group originally recorded at the level of a "cell" which is the smallest administrative unit in Rwanda (the size of a cell varies considerably (mean 11 ± 18 km²)).

3.2.2 Statistical analysis

Having information on the distribution of coffee plots (Mukashema et al., 2014), the identification of the environmental and human factors at coffee-growing locations becomes possible.

3.2.2.1 Extracting biophysical and socio-economic data at coffee locations

The biophysical data and demographical data layers, available at national scale, were rasterised into 10 m cells to allow comparison with the high resolution coffee field data. In total, 29 independent variables of which 22 are biophysical and nine socio-economic variables were extracted at each coffee field location and compiled in a GIS database for further spatial analysis.

3.2.2.2 Determining location drivers of coffee fields

We first describe the 29 variables available at coffee field level. In order to reduce the number of variables, we applied principal component analysis (PCA) using the orthogonal Varimax rotation. Only variables with a PCA score of 0.5 and above are considered for further steps. Kaiser's criterion and communality measures are used to retain only factors with eigenvalues

greater than 1 and variable communalities greater than 0.5. Variables scoring higher than 0.7 of the total correlation are considered to indicate a fairly good consistency with the factor into which they are loaded. Only 17 variables remained in five independent factors for further analysis.

3.2.2.3 Explaining location variability of coffee fields

We then evaluated with multiple linear regression as described by Austin et al. (1998) the extent to which the 17 remaining variables that explain the variability of current location (x, y) of area (ha) occupied by coffee. This procedure was done for the country as a whole and separately for each agro-ecological zone. The area (C ha) used for coffee cultivation is therefore determined by ($V_{i(x,y)}$) a number of biophysical and socioeconomic variables and their respective contribution (b_i) as follows:

$$C(ha) = b_0 + \sum_{i=1}^n b_i V_{i(x,y)} + \varepsilon \quad (3.1)$$

Where, $b_i, i=0\dots n$ are coefficients obtained by stepwise (forward) regression model and ε is the error associated with the model. Variables V_i with b values which do not differ significantly from zero were dropped. Although the 'best' model is usually selected on the basis of goodness-of-fit (R^2), we limited the maximum number of independent variables to ten in order to minimise noise in the models thus facilitating interpretation. Although the b values and their significance are important estimates, standardised b values (β , beta) are more explicit because they allow inter-comparison of the of factors within the regression model. Before the analysis, all used variables were first checked for normal distribution and during the multiple regressions modelling, models were checked for collinearity. By using randomly selected subsets, the effect of spatial autocorrelation was addressed (Overmars et al., 2003).

3.2.2.4 Predicting suitable areas for growing coffee

We used the weighted coefficient variables from the specific regression models for the individual agro-ecological zones (AEZ) to determine potential areas for expanding or intensification of coffee cultivation. The model can also be used to identify areas where coffee is currently grown but which may not very suitable for the cultivation of coffee. Actual and predicted areas were standardised (in %) using the maximum predicted area estimated by the AEZ specific models. The actual and predicted areas were compared using the validation dataset (i.e. 50% of the total coffee fields earlier separated from the training set).

3.2.3 Model validation

The performance of the models (national and regional) was assessed by comparing the predicted suitable coffee locations with actual coffee-growing

locations. First, spatial coffee data were randomly split into a training set (59,757 coffee fields) and a validation set (59,756 coffee fields). Second, the predicted outcomes were compared with the observed input and the difference map was produced, with mean error and R^2 calculated. The difference map shows where and in how far the model over- or underestimates suitable coffee-growing locations.

3.3 Results and Discussion

3.3.1 Location factors of coffee patterns

Principal component analysis (PCA) revealed that 17 variables out of 29 load into five factors which explain 86% of the total variance in the total dataset. The factor loadings after rotation are shown in Table 3-1 including the not included variables. Population distribution and age constitute the first factor explaining 22.1% of the total variance in the dataset. The second factor which relates to climate and topography explains 20.6%. The third factor represents the topsoil characteristics and fertility explains 17.8% of total variance. The fourth factor stands for the subsoil characteristics (i.e. soil lithology) and accounts for 14.8% of the total variance. The fifth factor is a mixture of physical soil properties and infrastructure and accounts for 10.7% of the total variance in the coffee dataset.

The reliability test revealed that 14 of the above 17 contributing variables are highly reliable ($r > 0.7$) in their respective factor loadings. Temperature and relative humidity are fairly reliable in the climate factor ($r_{t,RH} < 0.2$), while the distance from the field to the nearest coffee washing station is only marginally reliable in the factor combining soil texture and infrastructure ($r_{CWS} = 0.04$) indicating that soil texture and coffee washing station facilities should be treated separately in the analysis.

Table 3-1: Summary of principal component analysis results for the 29 variables at coffee locations

Variable items*	Rotated factor loadings				
	1	2	3	4	5
1. Total population (PT)	0.97	0.18	0.05	0.04	0.07
2. Population under 16 years old (P16)	0.94	0.20	0.06	0.07	0.04
3. Population between 16 and 18 years (P16-18)	0.93	0.15	0.02	0.05	-0.03
4. Population with more than 18 years old (P18)	0.95	0.15	0.03	0.02	0.10
5. Potential evapotranspiration (PET mm.yr ⁻¹)	0.18	0.92	0.04	0.10	-0.02
6. Annual average temperature (T°C)	0.17	0.91	0.09	0.14	-0.10
7. Relative humidity (%)	-0.17	-0.88	-0.15	-0.08	-0.02
8. Altitude (Alt m)	-0.17	-0.83	-0.14	-0.23	-0.02
9. Soil total nitrogen (N %)	0.08	0.14	0.94	0.00	0.18
10. Nitrogen-potassium ratio (N/K)	0.03	0.03	0.90	-0.33	0.12
11. Soil organic Carbon (SOC %)	-0.03	-0.17	-0.86	-0.31	-0.08
12. Soil pH (water)	0.06	0.27	0.61	0.71	-0.07
13. Potassium content (K)	0.09	0.19	-0.12	0.86	-0.04
14. Calcium content (Ca) (cmol+.kg ⁻¹)	0.04	0.23	0.08	0.84	-0.26
15. Clay content (Cy. %)	-0.04	0.04	-0.23	0.26	-0.84
16. Sand content (Sd. %)	0.07	0.08	0.27	-0.25	0.81
17. Distance farm-coffee washing station (CWS)	0.07	-0.13	-0.15	0.38	0.56
Eigen values	3.8	3.5	3.0	2.5	1.8
% of Variance explained	22.1	20.6	17.8	14.8	10.7
Cumulative percentages	22.1	42.6	60.5	75.3	86.0

*Significant variables are rotated by Varimax method. Five factors and 17 variables (1) population, (2) climate and topography, (3) topsoil soil fertility, (4) subsoil, (5) soil physical property and infrastructure are loaded with eigenvalues greater than 1 and variable communality greater than 0.7 except for the distance of farm to coffee washing station (CWS) variable which has a communality of 0.5. The percentage of variance explained by each factor is also presented. Valid N list wise = 53,255 coffee field Not included variables: population density people/km², rainfall (mm.year⁻¹), Hill shade (-), slope aspect (degree), slope (%), pH (KCl), C/N ratio, CEC clay (cmol+.kg⁻¹), Al (cmol+.kg⁻¹), Mg (cmol+.kg⁻¹), P (ppm) road distance to coffee field (m)

3.3.2 Explaining the current spatial distribution of coffee areas

3.3.2.1 Generalised national model

The generalised linear model generated from the composite national coffee dataset explained only 9% of the total coffee area variability. This clearly indicates that a single model cannot properly explain the spatial distribution of coffee fields for Rwanda as a whole. Table 3-2 shows in detail the relationship between the current coffee areas and the biophysical and socio-economic factors. The coefficients (b values) of the generalized linear (GLM) model, standardised (β values, Figure 3-2) variables and coefficients of determination (R^2) are presented in Table 2. As shown by the graphs of the standardised β values (see Figure 3-2), the national model is a composite of all five factors. The standardised β values indicate the relative amount of change in coffee area size compared to an independent variable change. The graphs can be viewed as AEZ factor fingerprints of coffee-growing zones. They indicate which factors are important in which region. The generalised model has highest contributions from population youth ($\beta_{P16} = -0.4$), and soil acidity ($\beta_{pH} = 0.3$). Lesser contributions are made by atmospheric humidity and evapotranspiration ($\beta_{RH,PET} = -0.2$), soil organic carbon ($\beta_{SOC} = 0.2$), and soil texture ($\beta_{Cy,Sd} = -0.2$). For the national model, the figures indicate that all the factors are of equal importance. This confirms that the coffee field pattern is a function of complex interacting biophysical and socio-economic factors.

3.3.2.2 Regionalized agroecological zone (AEZ) specific models

The country was subsequently stratified into ten agroecological zones for which separate multiple linear regression models were made (see Table 3-2 and Figure 3-2). Better model fits were achieved. For the Imbo (IMB) and Kivu islands (KI) the derived model explains up to 54 % of the total coffee area variability. The model of the central plateau explained 34%. In other zones, 28% of coffee areas found in the Buberuka Highlands (BH), 25% in the Congo-Nile watershed divide and 17 % in Impala (IMP) were explained. Coffee cultivation along the lake Kivu shore (KLS), and the Eastern plateau and lowlands are poorly explained by biophysical and population factors ($R^2 < 10\%$).

Figure 3-2 summarises the standardised beta (β) values of the significantly contributing variables at country level as well as at AEZ level. From west to east, the total population and the household size and age are an important factor in Kivu lake Shore (KS), and Impala (IMP) ($\beta = 0.4-0.9$). For the Kivu Islands (KI) with their highly fertile soils, the N/K ratio is the only location factor explaining the pattern of coffee distribution. Similarly, in Imbo (IMB)

and the Congo-Nile watershed Divide (CND), soil fertility is highly important ($\beta=0.4-1.3$). In the Central Plateau (CP), topsoil fertility and subsoil properties are equally important. In CP the N/K ratio has a higher degree of importance than other soil parameters ($\beta= -0.8$). Climate has more influence on coffee location in the Buberuka Highlands (BH). In BH, the negative effect of the atmospheric humidity ($\beta_{RH}= -0.7$) slows down the expansion of coffee growing. In the Mayaga, Bugesera (MPB), in the southeast, and the Eastern Lowlands (EL), the population factor is more important than climate and soil factors ($\beta_P=0.4-0.8$).

Legend and explanations for the Table 3-2 and Figure 3-2:

Each agro-ecological zone specific linear models for each (AEZ) has all b values significant at $P < 0.001$. The adjusted goodness of fits (Adj. R^2) is also presented. The standardised betas are listed in Figure 3-2. The beta standardised (β) allow inter-comparison of relative contribution of various factors such climate, soil, fertility, demographic and infrastructure development in respective agricultural zones. It gives factor 'fingerprint' of the location factors. KS Kivu lake shore, IMP Impala, IMB Imbo, CND Congo-Nile watershed divide, MPB Mayaga plateau and Central Bugesera, CP Central Plateau, ERP Eastern Ridge and Plateau, EL Eastern Lowland, BH Buberuka Highlands, VHP volcanoes and high plains, PT total population, P16 population under 16 years old, P16-18 population between 16 and 18 years old, P18 population with more than 18 years old, PET potential evapotranspiration, T annual average temperature, RH annual relative humidity, Alt altitude, N soil total nitrogen, N/K nitrogen-potassium ratio, SOC soil organic carbon, pHw soil pH water, K potassium content, Ca calcium content, Cy clay content, Sd sand content, CWS distance farm-coffee washing station

Table 3-2: Coefficients (b) quantifying the contribution each variable to the current coffee areas (ha) distribution

Factors & Variables	Agro-ecological zones										
	Nation	KI	KS	IMP	IMB	CND	CP	BH	MPB	ERP	EL
Constant	32,6	-0,7	-8,4	21,4	10,9	32,9	18,3	101,6	60,5	-2,9	-39,1
P16	$4 \cdot 10^{-3}$				$-5 \cdot 10^{-3}$		0,01	-13	$4 \cdot 10^{-4}$		
P16-18	-0,02		-0,01		0,02	0,02	-0,04			-0,02	-0,01
P18			3,10	-5,1		-5,1				$1 \cdot 10^{-3}$	$2 \cdot 10^{-3}$
PT				10^{-3}						-10^{-3}	-10^{-3}
Alt				$-5 \cdot 10^{-3}$	0,01					$3 \cdot 10^{-3}$	0,01
T			0,43	-0,63	3,57	-1,01	1,14				0,42
RH	-0,52						-1,3	-1,01			0,16
PET	-0,01				-0,04	-0,01	-0,01	-0,01	-0,01	-0,01	-0,01
OC	0,58			-0,4	-4,47			-1,26			
N						-1,11	1,07		-0,24		
N/K		0,13			-2,58	0,25	-0,55	-0,15		0,05	
pH	4,78						21,06		-1,77		
Ca	-0,44		-0,07	-0,06	-0,37		-5,55	-1,15			-0,08
K					-28,43		22,8				
Sd	-0,13		0,03			-0,14	-0,29			-0,04	
Cy	-0,19				0,45		-0,31		-0,21		
Infrastructure	CWS		$-2 \cdot 10^{-4}$	$2 \cdot 10^{-4}$	10^{-3}	-10^{-4}			$4 \cdot 10^{-4}$		
	Adj.R²	0,09	0,52	0,1	0,17	0,54	0,25	0,34	0,28	0,11	0,08
										0,19	

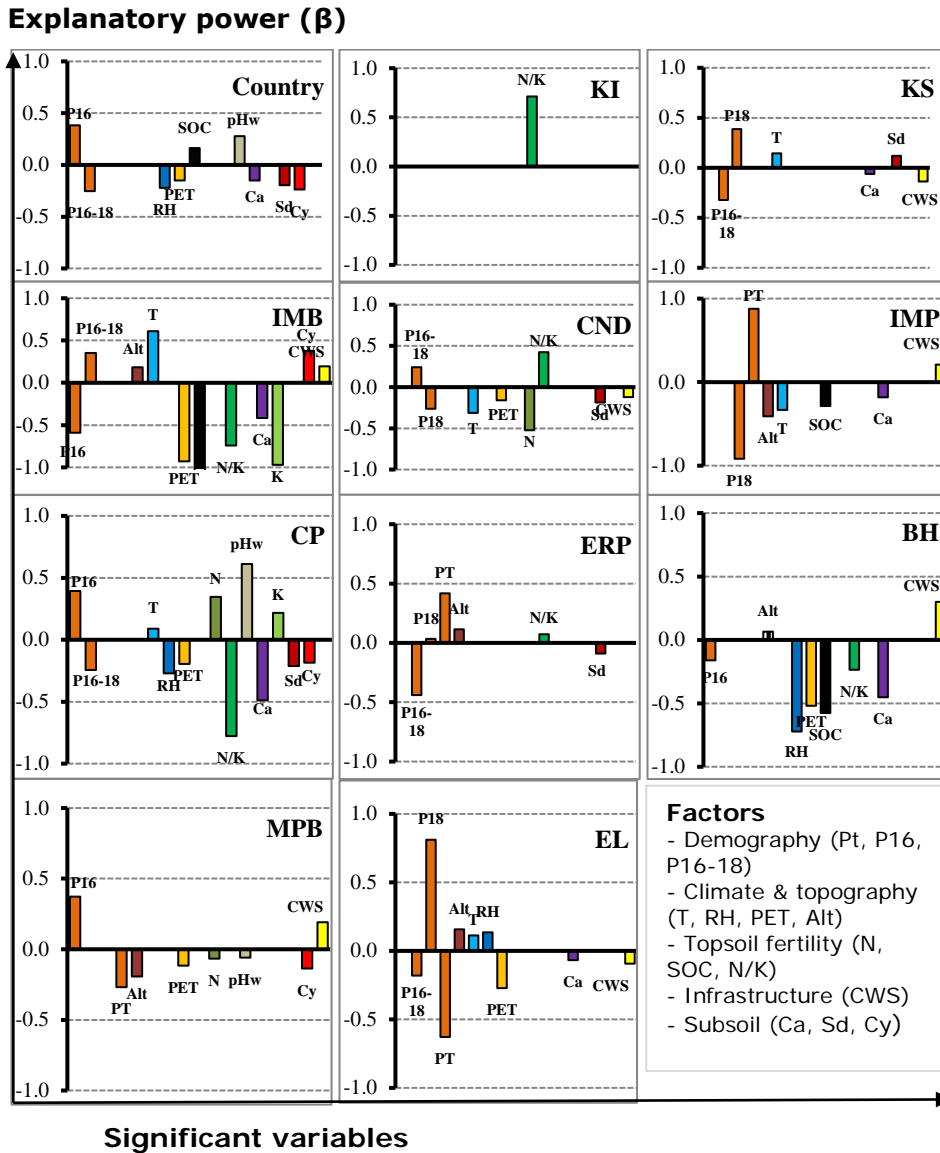


Figure 3-2: Explanatory power (β = standardised 'b' values) of biophysical and human variables loading in the five main components (factors) at country level and at agro-ecological zone level (factor label see legend table 3-2 and figure 3-2). National $R^2 = 0.09$, Agro-ecological based models: KI $R^2 = 0.52$; KS $R^2 = 0.10$; IMP $R^2 = 0.17$; IMB $R^2 = 0.54$; CND $R^2=0.25$; CP $R^2=0.34$; BH $R^2=0.28$; MPB $R^2 = 0.11$; ERP $R^2=0.08$; EL $R^2=0.19$.

3.3.3 Analysing potential locations for coffee expansion

Potential areas for coffee cultivation for the whole country were identified using the regional multiple regression models. This was done under the assumption that when the model based on local biophysical and socio-economic conditions, indicates higher coffee field densities than currently observed, there is potential for more coffee fields within the current land use systems. In Figure 3-3, the calculated potential coffee location areas are presented as percentage of the total arable land. There are five classes of occurrence: 0% no coffee, 0–25%, 25–50%, 50–75% and 75–100% coffee occurrence. Assuming that a lower calculated coffee area percentage indicates a lower suitability with respect to the analysed location factors we can also indicate where coffee is grown under suboptimal conditions. Figure 3-3 presents the actual coffee fields (a), potential coffee-growing areas according to the national model (b) and according to the regional models (c) and the areas where according to the regional models, coffee area increase is possible (d). These changes include expanding (+) or reducing (–) of coffee field density as a consequence of good or poor environment and socio-economic conditions.

For comparison purposes and to link to the still commonly used land evaluation classes, we propose to label the following: 0 % calculated coffee predicted area as unsuitable and not remediable (N2), 0–25% as unsuitable but remediable (N1), 25–50% as marginally suitable (S3), 50–75% as moderately suitable (S2) and 75–100% as suitable (S1). Strictly speaking this is not correct but it allows some grounds for comparison. The map (Figure 3-3) shows few areas where coffee suitability appears to be very high. Those areas are Kivu lake Shore, central and south Plateau and Mayaga. Smaller suitable areas are visible in the eastern ridge and plateau as well as the Buberuka highlands (in Burera-Ruhondo region). The least suitable areas are the Congo-Nile watershed divide. The actual coffee field occurrence (Figure 3-3a) is compared with the predicted coffee area suitability (Figure 3-3c). Of the actual coffee fields 6.6% is located in 'unsuitable' areas (N2 and N1), 33.4% in marginally suitable (S3) areas, 57.3% in moderately suitable (S2) areas and 2.7% in suitable areas (S1). So altogether, 60% of the existing coffee fields are located in moderately suitable to suitable areas.

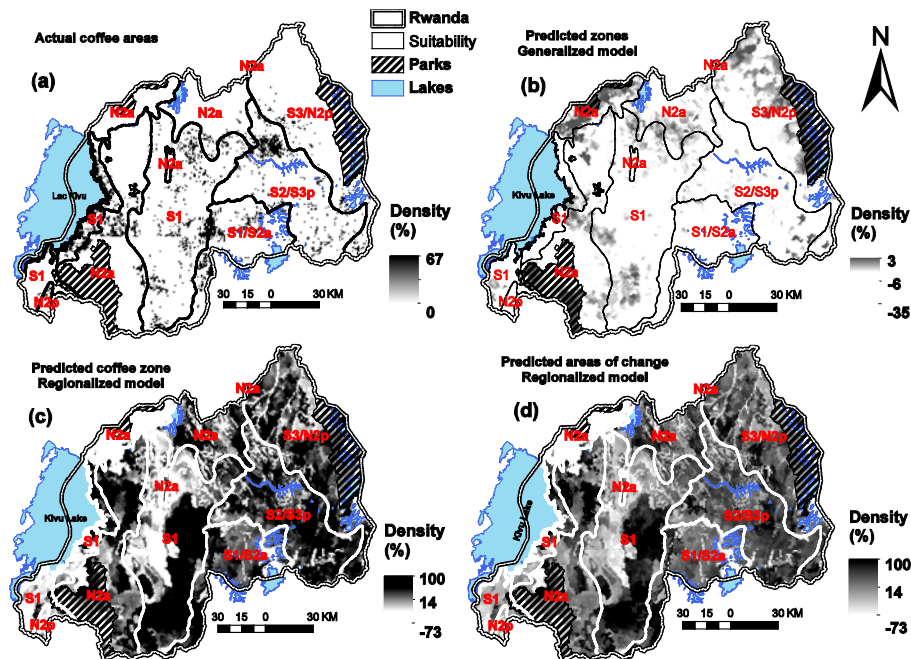


Figure 3-3: Actual coffee field density (a), predicted coffee density maps by the generalised (b) and regionalized (c) models compared with the actual density of coffee (a) mapped using high resolution imagery (Mukashema et al. 2014). The difference of actual and predicted are (d) explicitly visualises potential areas change. The scale from negative (-) to positive (+) change indicates where actual coffee area densities are larger (+) to lower (-) than expected by location factors. The coffee suitability (from S1 to N2) map generalised over AEZs (Verdoodt & Van Ranst, 2003a) is overlaid to calculated map. The country size is 26,338 km². For presentation, the resolution has been reduced to the grid cell size of 7 by 7 acres which correspond to the mean variance of coffee plots.

3.3.4 Exploring spatial coffee distribution

3.3.4.1 Role of environmental and socio-economic factors in coffee field density

Rwandan coffee systems are clearly driven by both regional socio-economic and environmental factors. These factors play out differently in the ten agro-ecological zones. Climatic (temperature) and topographic (altitude) factors play a prominent role in Imbo and Kivu lake shore in the western part of Rwanda. In this area, an increase of 3 ha of coffee area is observed as the elevation decreases by 200 m and the temperature increases by 1.2°C (towards the border of lake Kivu). Potential evapotranspiration plays a modest role in the Congo-Nile divide, the central plateau and Buberuka highland, the Mayaga plateau and eastern lowlands. It indicates that in these

regions the coffee system appears to be more climate sensitive than in other regions.

Soils play also a prominent role in determining coffee location within the whole country. Although we expected aluminium toxicity and phosphorous levels to play an important role, it was not included in any of the regressions, but the related pH was. The most important soil factor is the nitrogen-potassium ratio of the soil. This ratio reflects the interface between nutrient availability in the topsoil and the inherent chemical properties of the subsoil. The more significant contribution is found in central western part of the country. In many regions (Kivu lake shore, Impala, Imbo, central plateau, Buberuka highlands and eastern low lands), the soil acidity pH and calcium depletion of the subsoil has a negative effect on the occurrence of coffee. A combination of lime, agroforestry, organic and mineral fertiliser could improve soil fertility for such highly depleted acidic soils (Bucagu, Vanlauwe, & Giller, 2013).

Soil texture (sand and clay content) plays only a minor role in the spatial distribution of coffee. Except in the Kivu lake shore area where sand and clay content has a positive effect on coffee occurrence; a negative effect is observed in the eastern Nile divide zones.

The importance of soil derived properties gives some merit to the soil-based approach used by I. Nzeyimana et al. (2014). Their methodology, however, clearly leaves out relevant socioeconomic coffee factors that are most dynamic or prone to change.

The coffee production infrastructure plays a modest contribution in most regions with the exception of the eastern ridge and plateaus and Buberuka highlands. In the area where there are more coffee washing stations (see Figure 3-2), a positive contribution to coffee area can be observed while in the more remote areas with less coffee washing stations, a negative contribution is observed. This suggests that coffee washing stations can stimulate coffee production in their neighbourhood. This is related to a governmental project for initiating coffee washing stations that started in 2002 with pilots with cooperatives in South province (example of Maraba coffee). Government of Rwanda provided cooperatives with advisory work on the development of business plans, strengthened organisation and building linkages with coffee buyers and financial institutions. By 2010 this had become a programme with 183 washing stations. Population density and associated labour availability have a strong association with the occurrence of coffee in Kivu lake shore area, Impala, Imbo, Mayaga, eastern ridge and plateau. The other central regions have more modest contribution of demography. When in the future other types of socio-economic data, such as transport distances/costs for fertilisers and coffee prizes become available, we expect an improved system understanding.

Our systematic analysis clearly demonstrates that coffee field distribution is part of a complex land use system where not only many different factors play a significant role, but where locally different trade-off decisions are made. From an example of the effect of demography on the spatial distribution of coffee fields, we demonstrated that coffee in Rwanda is often not planted on the biophysical optimal location due to the equal importance of socio-economic factors and the competition with other land use functions. This difference explains why current coffee density (Figure 3-1) and the modelled coffee density maps (Figure 3-3) do not resemble the predicted potential coffee map (Fig.3, Nzeyimana et al. 2014).

3.3.4.2 Modelling coffee distribution

The regional models all have different variable contributions from the environmental and human-related factors. This reflects the regional diversity as discussed earlier. The coefficients of determination (R^2) range between 10 and 54%. The fit of the regression models improves when using a more detailed spatial stratification of the country. This confirms that coffee system analysis is a very scale and context sensitive exercise (Veldkamp, Verburg, et al., 2001).

Regional models were used to calculate coffee occurrence for the whole country. This resulted in a probability map displaying areas ranging from 0 to 100% coffee occurrence (Figure 3-3). The existing coffee suitability map (Verdoodt & Van Ranst, 2003a) is a highly simplified agro-ecological zones of Rwanda, therefore cannot be used to conclude on the details produced by our probability map. It is also important to realise that the probability map is based upon factors which in reality can change quickly over time. Dynamic processes such as population mobility (within 5–10 years), soil management (5–10 years), climate change (decadal), infrastructure and coffee washing station construction (5–10 years) can be easily accommodated by the methodology presented here by recalculating a new probability map using updated factor data. It is also important to mention that it is possible to include additional factors in the analysis if such data becomes available.

3.3.4.3 Potential areas for future coffee expansion

Potential areas for coffee expansion according to the regression models are shown in (Figure 3-3d). South of the Congo-Nile watershed divide, east of Nyungwe forest, a potentially suitable area is visible. However, this area is used for tea production since many decades. This represents a good example of how the land use history and land use causes suboptimal utilisation of land from one commodity perspective. An additional suitable coffee area is observed in the northern part of the Congo-Nile watershed divide. In this area, protected forest is the predominant land use (Westinga et al., 2013). Another suitable area for coffee growing is located in the east and north east.

This area used to be within the Akagera National park. After 1995, this area was rezoned to allow human resettlement, pasture and large-scale maize cultivation. In the western part of Buberuka highlands, another suitable area is identified. However, this is a traditional wheat growing area. Also in some central parts of the country, highly suitable lands are visible. In these areas, the absence of coffee is explained by the vicinity of the rapidly growing city of Kigali and its satellite cities. In a similar fashion highly suitable locations are found near to urbanising areas such as Rwamagana and Bugesera (located in the southeast) and Muhanga and Huye (located in the central and southern parts of the country). This matches the global trend that cities eat up very suitable agricultural land (Haller, 2014).

In summary, all potentially suitable coffee areas outside current coffee regions have other rational reasons than poor suitability to grow coffee. Causes are either related to competition with other land uses or a specific land use history. Despite of the above, the regionalized model fit (Figure 3-3) is satisfactory (60%). The model outputs also demonstrate that a significant proportion of coffee cultivation (40 %) is actually grown in less suitable locations (40 %). Translocating this coffee to more suitable locations might increase coffee production. Given the previously discussed limitations on coffee expansion to new areas, it might be a more realistic option to intensify coffee production. We propose to reanalyse the coffee system on a regular basis to allow more sustainable recommendations about planned changes in coffee growing areas.

3.4 Conclusion

Biophysical and socio-economic factors together determine the spatial pattern of coffee growing in Rwanda. The main potential explanatory factors of coffee can be statistically grouped into five main components: (1) socio-economic demographic factors, (2) local climate and topography, (3) topsoil fertility, (4) subsoil properties and (5) physical soil properties and coffee production infrastructure. Multiple regression analysis using these variables for the whole country explains only 9% of the total variability of coffee field occurrence. Spatially disaggregate models using different regional dependent combinations of explanatory variables resulted in a much improved explanatory power of spatial coffee field variability (adj. R^2 arising to 54%). In general, the smaller the region the better the model fit. This suggests a high diversity in local factors determining coffee field distribution.

The regionalized models were subsequently used to model coffee occurrence for the whole country. This demonstrated that 60% of current coffee fields are located at suitable locations. The remaining fields are situated in less suitable locations. The modelling exercise suggests, given all model limitations, that there are still potentially suitable coffee areas available in

the country. But many highly suitable areas are already occupied by other competing uses. Given the very limited land resources in Rwanda, it appears to be more realistic to increase future coffee production not so much by expansion of the coffee growing area as a whole but rather via a process of substitution whereby coffee cultivation is intensified at suitable locations. In less suitable locations, coffee growing could be reduced or converted into more suitable land uses.

4 Arabica coffee yields in Rwandan smallholder systems: future perspectives of climate change effect and improved soil organic carbon management

This chapter is resubmitted after revision as: Mukashema, A., Veldkamp, A., and Gatarayihira C.M. Arabica coffee yields in Rwandan smallholder systems: future perspectives of climate change effect and improved soil organic carbon management. In *Agriculture, Ecosystems and Environment*

4.1 Introduction

Coffee (*Coffea Arabica*, bourbon variety) farming in Rwanda is mainly a small holder activity. Coffee farms have an average field size of 0.24 hectare (Mukashema et al., 2014) and are part of diverse Rwandan small-scale cropping systems which are on average 0.59 hectare in size. Arabica coffee (Bourbon variety) was introduced in Rwanda at the beginning of the twentieth century. Most Rwandan smallholder farmers grow bourbon Mayaguez and Jackson varieties which seem to be well adapted to the conditions in Rwanda (Wintgens, 2009). The Bourbon varieties are planted at medium density (2,000-2,500 plants per hectare) compared to other varieties such as Caturra planted at 5,000-10,000 plants per hectare in Colombia and Costa Rica.

Coffee yield levels in Rwanda (average 0.7 ± 0.4 t ha⁻¹) are generally lower than averages (1.2 t ha⁻¹) obtained in small scale farming systems optimized for coffee growing (Wintgens, 2009). Compared with other coffee growing countries, Rwandan coffee farmers obtain yield levels well below the levels (1.7 t ha⁻¹) obtained by small scale farmers in India (Jayakumar et al., 2016), and 1.8 t ha⁻¹ in Brazil (de Muner et al., 2015). Very high yields (3t ha⁻¹) are reported in Brazil on commercial Arabica coffee farms (Teixeira et al., 2015). In East Africa, Rwanda has similar yield ranges as observed in Uganda (0.7 to 0.9 t ha⁻¹) (N. Wang et al., 2015).

Coffee yield is context and location dependent and usually controlled by the local interaction of environmental and socio-economic factors (Castro-Tanzi, Dietsch, Urena, Vindas, & Chandler, 2012; Veldkamp, Kok, et al., 2001). Location factors which determine suitable coffee areas in Rwanda have already been analysed in detail by Mukashema, Veldkamp, and Amer (2016) and found that 60% of coffee farms are located in suitable environmental and socio-economic locations. However, this analysis did not include the yield levels obtained. Nzeyimana et al., (2014) used a 2.3% sub-sample field measured yield approach for 2005 to assess potential productivity zones and yields for the ten agroecological zones using ordinary kriging. A strong limitation of their multi-criteria evaluation study was the use of highly aggregated spatial units of ten AEZ only. This makes it difficult to make concrete recommendations for small holder systems.

The population in Rwanda is expected to grow from 11 million in 2012 to 15 million in 2022 (NISR, 2012) and arable land is getting increasingly fragmented and is less available. In this context, productivity improvement is the obvious way for coffee farmers to increase their income from current coffee areas. Reallocating coffee to more suitable areas while releasing less suitable areas to other land uses as identified by (Mukashema et al., 2016) might also increase net yield but this will take many years to become

effective and is a very sensitive process. Therefore, the by far simplest solution is to increase coffee yields of current coffee fields.

Not all coffee areas have a similar potential to support higher yields in a sustainable manner, due to local climatic, topographic, soil quality constraints, and in some cases due to limited access to coffee infrastructures and input markets. An often used approach to understand yield increase potential is yield gap analysis using crop growth models (van Ittersum et al., 2013). Biophysically constrained maximum yields (usually water limited yield potentials) are simulated assuming optimal fertilizing and crop management with no pests and diseases (Tittonell & Giller, 2013). Such a yield gap analysis cannot be done for coffee because there are no reliable crop growth models available for this perennial crop, because existing models are mainly developed for annual crops (e.g. Affholder, Poeydebat, Corbeels, Scopel, & Tittonell, 2013; Tittonell & Giller, 2013; Xiao et al., 2016). The only comparable crop for which a type of crop growth model exists is Cacao (Diaz-Jose, Diaz-Jose, Mora-Flores, Rendon-Medel, & Tellez-Delgado, 2014; Ofori-Boateng & Insah, 2014). However, these studies focused on identifying regional yield differences only. For coffee, limited attempts have been undertaken to predict yields, usually only by focussing on climate parameters (e.g. Craparo, Van Asten, Laderach, Jassogne, & Grab, 2015; Jayakumar et al., 2016; Ovalle-Rivera, Laderach, Bunn, Obersteiner, & Schroth, 2015; Picini, Camargo, Ortolani, Fazuoli, & Gallo, 1999; Santos & Camargo, 2006). In Uganda, N. Wang et al. (2015) monitored coffee yields and production factors for both Arabica and Robusta coffee and found that lower yield levels were associated with unfavourable climate and poor soil nutrient status. These constraints varied strongly and were location specific, calling for site-specific management approaches. In Rwanda, Bucagu et al. (2013) investigated the potential of agroforestry by *Tephrosia vogelii* intercropped in coffee plantations and Pinard, Boffa, and Rwakagara (2014) investigated the role of shade trees on coffee productivity. Both studies registered location specific increase of yield levels but they included only effects of factors that can be directly influenced by farm management. There are no studies found for Africa that combine climate, topography and soil properties to determine the environmental and management factors controlling coffee yields. This is probably because the spatial data needed to perform such analysis are difficult to collect in Africa. Due to relative scarcity of field data, many studies recommend simulation models in estimating potential yield and related gap with actual yields (e.g. Affholder et al., 2013) but such results would have mainly theoretical value in coffee farming because day to day crop management in small holder farms is known to be sub-optimal due to the competition of different crops for the limited available farm resources such as inputs and labour (Murekezi, Jin, & Loveridge, 2014). Thus probably more than anywhere else, actual yields under African smallholder conditions are the result of a tight interaction between the crop genotype, the local

environment and existing farming management practices (G×E×M) model (Tiftonnell & Giller, 2013).

An alternative approach to understand what is affecting current coffee yield levels, is by analysing how location characteristics relate to actual yield (Y_{act}) variability within the current system using a holistic approach with known potential yield constraints (Veldkamp et al., 2001). Using a multiple annual yield average will reduce the temporal variability (Kok and Veldkamp, 2011) resulting in more general (multi-annual) valid relationships. Regression analysis can then be used to understand which location factors are associated with higher and lower yield levels. The developed empirical multiple regression yield model can calculate an 'achievable' yield level for all locations given the uncertainties and limitations involved. This allows the identification of locations where current yields are below the 'achievable' yield. Because the significant physical constraints are already taken into the developed yield model it can be assumed that this lower yield level is mainly related to sub-optimal local land management within the current system due to time and means competition with other land use practices within the small holder farm (Verburg et al., 2000).

The advantage of this approach is that it can indicate locations with potential yield gains without the need for major system restructuring or reforms. This is for the short term a much more feasible approach than calculating theoretical yield gaps which require, in order to be bridged, irrigation, optimal fertilizer applications and pest control measures which in most cases requires major investments that can be only afforded by large scale farmers (e.g. Liu, Li, Zhang, & Yang, 2016; Teixeira et al., 2015).

For policy making it is highly relevant to know the current yield controlling factors instead of the theoretical ones, to improve the current sub-optimal agronomic practices to increase yield towards the levels defined as the 2020 targets for Rwanda.

The objective of this paper is hence to explore the spatially explicit coffee yield variability and its association with local topographic, climatic and soil conditions within current smallholder farming systems in Rwanda. Identified yield constraints can then be used to locate areas where a higher yield can be achieved and to recommend future development options to improve coffee yields.

4.2 *Material and Methods*

4.2.1 Coffee production and yield data

National production data from the National Agriculture export board of Rwanda (NAEB) was used in the analysis. Coffee yield data was collected at

166 coffee washing stations (CWS) for the period of 2011-2014 while the production of 2015 was collected at 245 CWS (see Figure 4-2 for their locations). These CWS are mainly owned by coffee farmers organized in cooperatives. Only a few CWS are owned by private investors. Ripe coffee cherries are manually harvested, screened and transported to coffee washing stations for washing and drying. Coffee growers bring their cherries to the nearest CWS. Each coffee washing station has a well-defined supply area (Figure 4-1). The diameter of the supply area ranges between 0.5km and 10km with an average diameter of 3km.

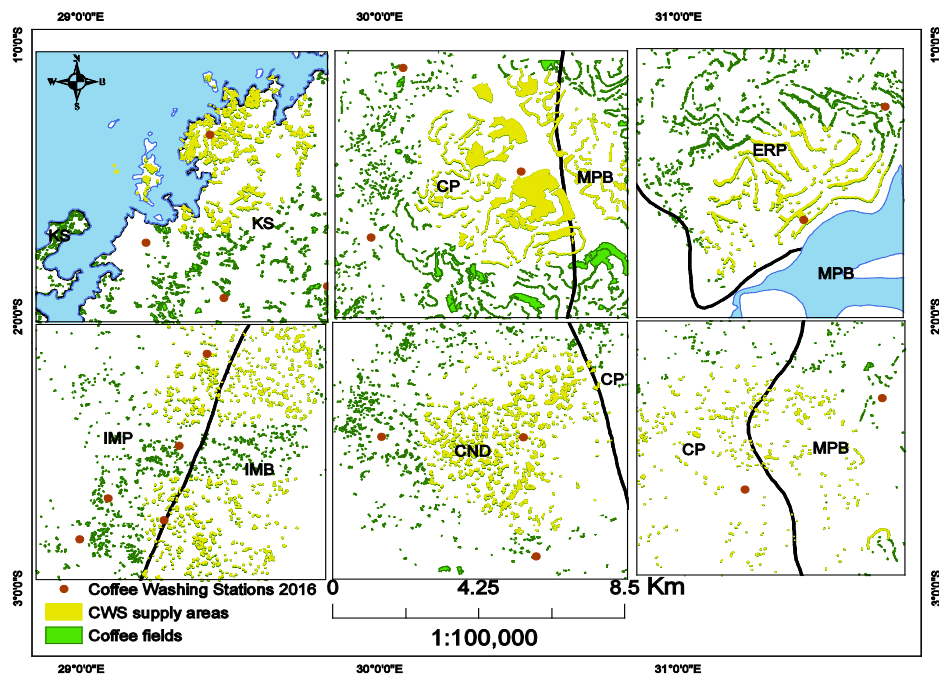


Figure 4-1: Map showing examples of coffee fields supplying coffee cherries to (selected) nearest coffee washing station (CWS). Different supply area touch each other following the density of coffee fields, the size and productivity of coffee farms as well as the processing capacity of coffee washing stations. By 2016, there were 245CWS. The diameter of the supply area ranges between 0.5km and 10km with an average diameter of 3km. Agroecological zones: KS Kivu lake shore, IMP Impala, IMB Imbo, CND Congo-Nile watershed divide, CP Central Plateau, MPB Mayaga plateau and Bugesera, ERP Eastern Ridge and Plateau.

At the end of the season (March-July), the CWS owners report the total production to the National coffee board (NAEB). This procedure implies that the production data report coffee cherries quantities that passed farm gate screening and taken to a coffee washing station (Figure 4-2) which is estimated to 95% of total cherry production on the coffee trees. The 5% are process by farmers at home and sold to ordinary market (NAEB unpublished technical reports 2011-2015).

The total national production varied from 12, 928 tons in 2011 to 17,829 tons in 2015 of green coffee beans. A coffee field map containing 119,513 individual coffee fields with a total area of 25,148 ha, (Mukashema et al. (2014), was used to convert CWS yields into coffee yield per hectare. Coffee yield varies highly across the country, 62.8% of coffee farms have low yields (0.5t ha^{-1}), 14.3% have yields between $0.5 - 1\text{ t ha}^{-1}$, 6.3% have yields ranging from 1t ha^{-1} to 1.5 t ha^{-1} , only 16,6% reach yields above 1.5 t ha^{-1} .

Figure 4-3a shows that the coffee yield (t ha^{-1}) decreases as we move from western highlands to Eastern lowlands. This already indicates that there is a relationship between agro-ecological conditions and coffee yield. Figure 4-3b demonstrates the national annual production from 2007 to 2015. From 2007 up to 2012 total production levels remained relatively constant. There has been a stepped increase since 2013 due to increased NPK fertiliser applications. Nationally, NPK fertilizer use in coffee increased from 1000t in 2011 to 5000t in 2015. About 5 kg of organic manure per coffee tree and 100 grams of NPK 22.6.12 (in Eastern ridges and lowland, and Impala region) or NPK 17.17.17 (in Kivu lake shore, Congo-Nile divide, and Central Plateau) is applied in two applications per year (October/November, and March/April). Another overall production gain factor is related to the production of newly developed coffee areas that were planted from 2010 onwards. In the analysis we used the 2011-2014 average a combination of two low and two high production years. The production level of 2015 was used as validation year.

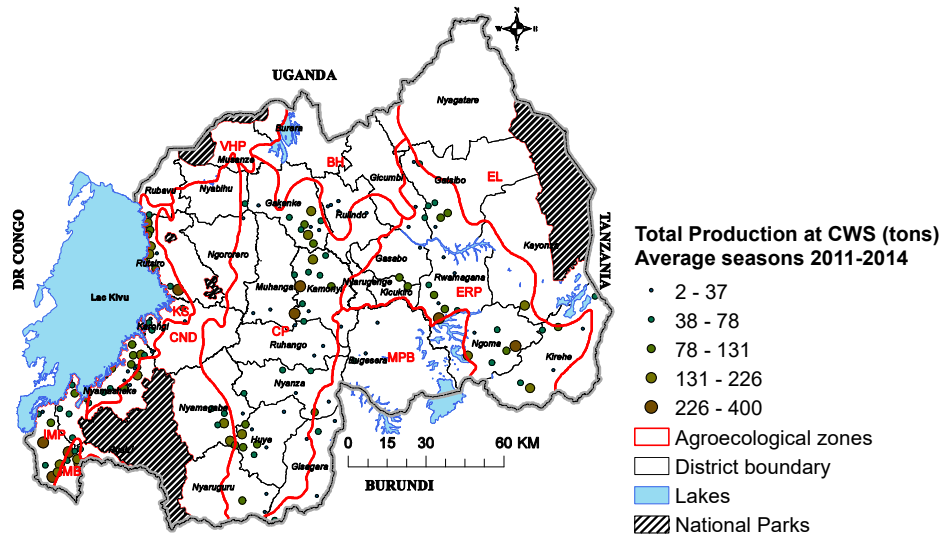
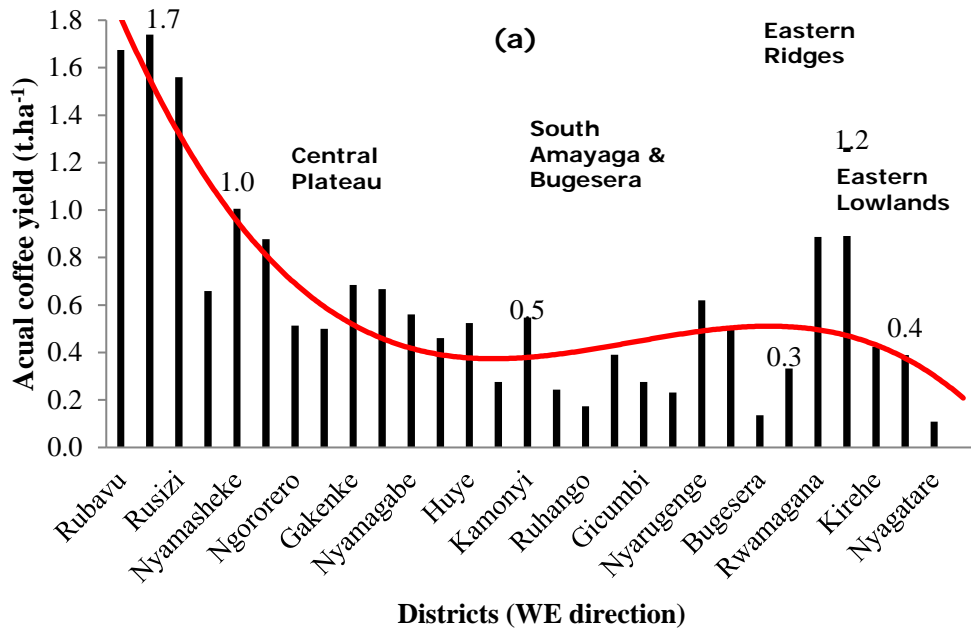


Figure 4-2: Map showing the average total production supplied to 166 coffee washing stations (CWS) by individual farmers for the period from 2011 to 2014. Agroecological zones: KLS Kivu lake shore, IMP Impala, IMB Imbo, CND Congo-Nile watershed divide, MPB Mayaga plateau and Bugesera, CP Central Plateau, ERP Eastern Ridge and Plateau, EL Eastern Lowland, BH Buberuka Highlands, VHP Volcanoes and high plains.

North & Western Highlands



Production (tons)

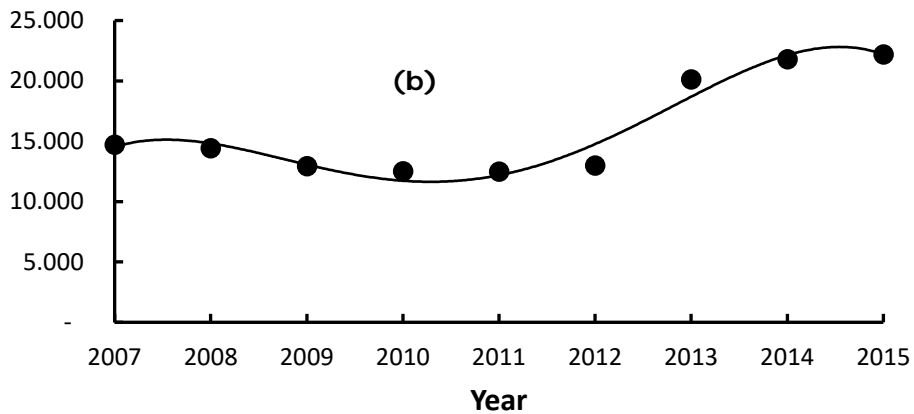


Figure 4-3: a) Histogram showing the coffee yield gradient from Western mountains to Eastern lowlands. Yield data are presented at District level using the total production reported at the National coffee export board (NAEB) for the season 2015 divided by the number of hectares per district. b) The trends in annual total coffee production for Rwanda 2007-2015.

4.2.2 Biophysical properties of coffee farms

The source of biophysical data used in this study are described in Mukashema et al. (2016).

4.2.2.1 Climate

In Rwanda, coffee is currently grown under temperature conditions ranging from 14.2 to 23.8°C. The named areas and regions are indicated in the map of Fig. 1 by their abbreviations. The vast majority of coffee fields (97.9%) is situated in the central (CP) and eastern plateau (ERP) areas where optimal temperature (16-22°C) are found for growing coffee according to Wintgens (2009). A minority of 1.7% of the coffee areas is found in Congo-Nile water divide (CND) and Buberuka highlands (BH) with relatively low temperatures (<16°C). Only 0.4% of all coffee fields, mainly located in the eastern lowlands (EL), have a higher than optimal temperature (>22°C).

Coffee fields are mainly located in areas with annual rainfall varying between 700 mm in the east and 2,120 mm in west and north. The eastern lowlands (EL) and central Bugesera (MPB) are significantly drier, receiving rainfall less than 900 mm annually which is slightly lower than the minimum of 951 mm required for coffee growing (Wintgens, 2009). Coffee fields of the Impala zone (IMP) and south-west Congo-Nile water divide (CND) receive rainfall exceeding 1,600mm annually. The potential evapotranspiration in coffee fields vary greatly from 440 and 1,550 mm following the NW-SE gradient as the temperature increases. Coffee is grown in areas with a relative humidity varying between 67 and 82% with an average of 73±3% which is quite suitable compared to the optimum required ranging from 60 to 85% (Wintgens, 2009).

In the south and west Rwanda, coffee is exposed to 1,480 hours of sunlight during the rainy season and around 1,900 hours during the dry season which between 58% and 75% of the required sunlight exposure. The optimum amount of sunlight exposure for coffee cultivation is considered to range between 2,200 to 2,400 hours of sunlight per year (Wintgens, 2009).

4.2.2.2 Topography

Arabica coffee (*coffea arabica*) is generally adapted to the tropical highlands, however, considerable topographic diversity can be observed. Flat land or gently rolling hills are generally considered best suited for coffee cultivation (Wintgens, 2009). In Rwanda, coffee is cultivated at altitudes varying between 970 and 2,575m. About 28% of the coffee cultivation area is located at altitudes below 1,500m, 67% at altitudes between 1,500 and 2,000 m, and only 4% is at altitudes above 2,000m. Areas below 1,750 m found in the Lake Kivu shore (KS) area, are more favourable and the growth season is shorter than in the areas with high altitude in the upper catchment of Kivu Lake where flowering and fruits ripening require more time.

Concerning field slope, Rwanda coffee fields are located in areas with slopes varying between 8% and 89%. In the western steep areas of the Congo Nile water divide (CND) as much as 23.2% of the coffee fields are found on slopes

steeper than 40%. About 24.2% of all coffee fields are located on slopes between 25 and 40% and those are mostly situated in the southern central plateau (CP). In the Mayaga and eastern ridges (MBP) area about 28.3% of all fields is found on slopes between 12 and 25%. Only 24.4% of the coffee fields are found on the gentle slopes (<12%) in the Bugesera (MPB) and Eastern lowlands (EL).

4.2.2.3 Soils

With regard to lithologies and related soils, 63% of the coffee fields are established on soils developed on shale, and mica schist, 23.5% on granitic soils and 9.7% on soils developed on basic rocks such as basalt and amphibolites. Very few fields (1.4%) are established on volcanic soils which have a high base saturation and a high pH. More than 75% of coffee fields are found on loamy sand soils which are depleted of essential plant nutrients. Total soil nitrogen, which mainly depends on field management, greatly varies from very low (0.07%) to very high (0.83%). Such soils provide considerable challenges for coffee production in Rwanda (Rutunga, Janssen, Mantel, & Janssens, 2007).

The top soil pH in coffee fields varies from 4.2 to 6.8. Only 8% of coffee fields are established on soils, which have what is considered an optimum coffee soil pH (between 5.8 and 6.8) along Kivu lake shore (KS) and Impala (IMP). About 59% of the total coffee fields are established on strongly weathered acidic soils (pH 5.0-5.8) of the central plateau; and 35% of the total coffee fields are on very acid soils (pH<5.0) mainly located in upper Congo-Nile water divide (CND) which suffer from Aluminium (Al) toxicity and highly deficiencies in Phosphorus (P) (I. Nzeyimana et al., 2013). In Rwanda Aluminium toxicity is very often associated with low soil pH and low organic carbon content (<1.5%).

Soil organic carbon (SOC) content in the upper layer (0-30cm) of the coffee fields is not considered a limiting factor because most coffee fields are mulched. Only 0.1% of coffee fields have a low SOC content (<0.8%), 9.2% are classified moderately humic (0.8-1.2%), 49.4% are humic (1.2-2.8%) and 41.3% very humic soils (SOC> 2.4%). The cation exchange capacity (CEC) is generally low ($14.3 \pm 6.8 \text{ cmol+Kg}^{-1}$) in coffee fields due to highly depleted 1:1 clay soil type (kaolinite), characteristic of soils developed on granites. About 73.5% of coffee fields have a very low CEC (<16 cmol+Kg^{-1}), 19.3% moderate CEC (16-24 cmol+Kg^{-1}) and only 7.2% with very high CEC (>24 cmol+Kg^{-1}). Coffee fields with poor CEC conditions are mainly found in granitic soils of the central plateau.

The available phosphorous for coffee plants is very low (less than 3ppm) in 49%, and low (from 3 to 20 ppm) in 39.9% of total coffee fields. Only 9.3% of coffee field has a satisfactory level of available phosphorous content

between 20 and 50 ppm. A maximum of 257ppm of available phosphorous are observed in few coffee fields which is still very low compared to the recommended optimum of 440ppm (Wintgens, 2009). Potassium (K) and calcium (Ca) are generally deficient in coffee fields with 92.4% of fields deficient in potassium (K^+ less than $0.2 \text{ cmol}^+\text{Kg}^{-1}$) and 93.6% of coffee fields deficient in calcium (with Ca^{2+} less than $4 \text{ cmol}^+\text{Kg}^{-1}$). Similarly, Magnesium (Mg^{2+}) content although highly variable across coffee cultivated soils ranges from 0.4 to $10.7 \text{ cmol}^+\text{Kg}^{-1}$ which are still below the recommended norm for coffee fields.

4.2.3 Data processing and statistical analysis

Spatial data of environmental variables related to climate, soil and topography were processed at 10m resolution. Coffee data (fields and areas, geographic locations of coffee washing stations (CWS) and total coffee production supplied to CWS for the period from 2011 to 2015) were divided in a 2011-2014 average yield value and 2015 was separated to be used as validation set. For each CWS the supplying area was delineated using the near-distance analysis tool of ArcGIS 10.4 version. According to the National Agricultural Export development board (NAEB) of Rwanda, each coffee washing station (CWS) is given a supply area (Fig. 1) corresponding to its processing capacity depending on the productivity of nearest coffee farms and the reasonable travel distance by coffee growers to the nearest CWS. The average travel distance to the nearest coffee washing stations by the farmer is about 3 km. With the current number of CWS (2015-2016), the maximum distance between the individual farms and the nearest CWS is 10 km, the nearest being at 0.5 on average. A cooperative may possess more than one CWS depending on the terrain and availability of transport infrastructure.

In order to identify biophysical factors underlying the variability of coffee yields, the total production for each CWS was redistributed to the coffee fields in its supply area (Fig.1), proportionally to the supplying field size in order to estimate the yields for every coffee field. The national dataset of coffee fields ($N=119,513$) was then randomly split into two equal subsets: one for production of coffee yield model ($n=59,520$) and model validation set. Biophysical data values were then extracted at coffee field location.

We first explored the difference in yield levels at country level in relation to average biophysical conditions in coffee fields by categorizing yield levels and biophysical requirement as recommended by Wintgens (2009) and by Verdoodt and Van Ranst (2006) for Rwanda. To determine if the difference in proportion of coffee farms allocated to biophysical categories are significantly different between pair category of yields we examined the results of column proportion tests (Z tests) using Bonferron method. This test assigns a subscript letter to the categories of column variable. If a pair of values is

significantly different, the values have different subscript letters assigned to them.

We also analysed the spatial variation coffee yield across the country agro-ecological zones using spatial autocorrelation and global Moran's index for the average coffee yield measured from 2011 to 2014. Moran's Index indicates the behaviour of autocorrelation, in which smaller value than the expected index demonstrates a trend for dispersion, whereas positive value indicates a clustering trend for contiguous areas (Silva, de Queiroz, Pinto, & Santos, 2014). The z-score and p-value are measures of statistical significance which tell you whether or not to reject the null hypothesis. In this study, the null hypothesis is that coffee yields are randomly distributed among coffee washing stations and between coffee field supply areas.

In order to identify biophysical factors with significant effect on coffee yields, all available topographic, climate and soil variables were analysed using stepwise multiple regression method (equation 1). All statistical analyses were performed using the software IBM SPSS statistics (22nd edition).

$$Y_{ach} = b_0 + \sum_{i=1}^n b_i (EM)_{i(x,y)} + \varepsilon_i \quad (4.1)$$

Where, Y_{ach} (kg ha⁻¹) is the achievable yield under site-specific environmental and management conditions (EM), b_i , $i=0...n$ are coefficients obtained by forward stepwise regression (criterion is $p < 0.05$), and ε is the error associated with the model. EM represents environmental variables related to climate, soil and topography selected for their direct influence on the coffee growing. The 'best' yield model is selected on the basis of goodness-of-fit (adjusted R^2). For each EM variable, a coefficients b-value, and its associated standardized beta β -value is calculated to allow a comparison among EM factors. The model was first developed for the whole country. In a sub sequential step it was explored if stratification by agroecological zones and lithological based soil units improved observed model explanatory.

The parameters (b-values) of the selected yield model and parameters were later used to calculate the 'achievable' yield map of Rwanda using the non-used data subset. The goodness of fit between achievable coffee yield map and actual farm yield was confirmed by adjusted coefficient of determination (R^2). Standardized beta β -values were also used to compare model parameters in order to identify yield constraints. Before the analysis all used variables were first checked for normal distribution, no transformation was applied to original data, and during the multiple-regression modelling models were checked for collinearity. By using randomly selected subsets the effect of spatial autocorrelations was addressed (Overmars et al., 2003).

The best achievable yield models and related predicted achievable yield map were validated using random coffee fields from the validation set, randomly selected within 229 coffee washing station service areas. Using the validation set, the mean yield difference $\overline{Y_{diff}}$ was also calculated by comparing predicted achievable yield values at field location $Z^{\wedge}(Y_{ach})$ with actual coffee yield as reported by the coffee washing stations $Z^*(Y_{act})$. Similarly, the change in coffee yields as a result of change in one of the co-interacting factors for a given period of time is calculated as follow:

$$\overline{Y_{diff}} = \frac{1}{n} \sum_{j=1}^1 [Z^{\wedge}(Y_{ach})_j - Z^*(Y_{act})_j] \quad (4.2)$$

Beside adjusted R^2 , we also assessed the spatial autocorrelation (Global Moran's I) on the model residuals (Yact-Yach) to see if the under/over-predictions are clustered or not. When a model fairly predicts the dependant variable (achievable coffee in this case), the residuals are spatially random. That means large residuals are intermixed with small residuals; large residuals do not cluster together spatially.

The climate of Rwanda is expected to change towards a warmer and wetter climate by 2050. An increase of temperature by 0.35°C per decade and increase of rainfall by 4mm/yr (i.e. 40mm per decade) were projected (McSweeney & Semafara, 2011; Muhire & Ahmed, 2015). Using this projected decadal increase of temperature and rainfall, we recalculated the achievable coffee yield map for future conditions by 2020 and 2050 using the coffee yield models with increased temperature and rainfall, holding constant other significant biophysical parameters of the model. Under this climate change scenario prediction, we assume that coffee will continue to be grown in smallholder systems in Rwanda without adaptation measures.

4.3 Results

4.3.1 Biophysical characteristics of coffee farms and yield levels

Table 4-1 indicates the number of coffee farms under biophysical factors and actual mean yield. The highest mean yield (2.1-2.3 t ha⁻¹) was obtained at elevation between 1,000 and 1,400m and with annual rainfall above 1,200 mm. Other optima (1.7-1.8 t ha⁻¹) were found at slopes between 16 and 30% at temperatures ranging between 16 and 24°C, on soils with pH above 5.3, and with soil organic carbon above 2.4%.

Table 4-1 also shows the proportion (in percentage) of coffee farms in each yield category for each category of biophysical factors. At altitudes between 1,000 and 1,400m, the proportion of coffee farms with higher yields increases significantly. However, above 1,400m, the proportion of coffee farms with higher yield decreases. The slope categories demonstrate a similar behaviour. On slopes below 16%, the proportion of coffee farms with higher yield decreases, while above 16%, the proportion of coffee farms increase as the yield increase from 1t ha⁻¹. The upward trend of yields with slopes above 16% is significant up to 30% slope gradient, and then declines significantly as the slopes steepen up above 30%.

For areas with an annual rainfall below 1,000 mm, there is not a consistent trend of precipitation with proportion of coffee farms. Instead, areas with annual rainfall ranging between 1,000mm and 1,200mm the proportion of coffee farms is significantly reduced with higher yields while a positive trend is observed for areas with annual rainfall above 2,000mm. This demonstrates that a positive relationship between rainfall and yield exists although this factor interacts with the topographic factors that tend to reduce the rainfall impact on yield at slopes above 30%, probably due to run-off effects. For temperature, there is a clear increase in proportion of coffee farms in all temperature categories with higher yield levels, but this relationship is also not entirely consistent indicating that other factors may have interacted. Concerning coffee field soil quality, a significant variability in proportion of coffee farms is observed with higher yields associated with higher pH and soil organic carbon (SOC), but less consistent for the cations exchange capacity (CEC) of the soils. Under the category of pH below 4.5, the proportion of coffee farms significantly increases with increasing yields; while in the category of soil pH above 5.3 the proportion of coffee farms reduces with increase of yields. This area is used for staple crops. For the other soil pH categories an upward trend in proportion of coffee farms are observed with higher yields. This suggests that in current Rwandese coffee farming systems, coffee has relatively good yields even at lower pH range. The cation exchange capacity (CEC) values below 16 cmol⁺kg⁻¹ shows a downward trend in proportion of coffee farms with increased yields, while other categories (16 to 24 cmol⁺kg⁻¹ category and above) did not show a consistent trend with increasing coffee yield. Soils with CEC values in the low and high yield categories have the same subscript (a), which means they cannot be distinguished from each other. For soil organic carbon, with the exception of the first category (<0.8%) which did not significantly differ with the yield increase, a strong and positive relationship between soil organic carbon and coffee yields can be observed.

Table 4-1: Percentage of coffee farms under biophysical factors and actual yields (t ha⁻¹) categories. The variables used in this analysis have earlier revealed to be significant factors controlling coffee field location (Mukashema et al.2016).

Biophysical factors unsuitable (N) to highly suitable (S1)	Mean yield tha ⁻¹	Coffee yield (t ha ⁻¹) [low →high]				n= 59,520
		< 0.5	0.5-1	1-1.5	>1.5	
<1000(S3)	3.0	0.0 _a		0.1 _{a, b}	0.1 _b	11
1000–1400(S2)	2.3	5.5 _a	1.6 _b	18.9 _c	13.4 _d	4,216
1400–2700(S1)	1.7	94.5 _a	98.4 _b	81.0 _c	86.5 _d	55,273
<8(N)	1.7	33.0 _a	26.2 _b	24.9 _b	31.1 _c	18,580
8–16(S3)	1.7	31.8 _a	34.0 _b	28.1 _c	28.4 _c	18,623
16–30(S2)	1.8	30.6 _a	34.5 _b	40.1 _c	35.9 _b	19,400
>30(S1)	1.8	4.6 _a	5.3 _a	6.9 _b	4.7 _a	2,896
<800(N)	1.3	0.0 _{a, b}	0.1 _b			23
800–1000(S3)	1.2	11.8 _a	2.9 _b	0.2 _c	3.1 _b	4,974
1000–1200(S2)	1.5	50.9 _a	32.5 _b	10.2 _c	29.0 _d	25,044
>2000(S1)	2.1	37.2 _a	64.6 _b	89.6 _c	68.0 _d	29,459
14–15(S3)	1	0.1 _a				41
15–16(S2)	1.2	0.5 _a	0.2 _{b, c}	0.3 _{a, c}	0.0 _b	198
16–24(S1)	1.7	99.4 _a	99.8 _{b, c}	99.7 _{a, c}	100 _b	59,261
<4.5(N)	2.2	1.8 _a	3.7 _b	7.3 _c	4.5 _d	1,707
4.5-5.0(S3)	1.8	35.6 _a	28.7 _b	36.1 _{a, c}	37.5 _c	20,794
5.0-5.3(S2)	1.6	23.0 _a	20.0 _b	11.7 _c	16.9 _d	12,373
5.3-6.8(S1)	1.8	39.6 _a	47.7 _b	45.0 _c	41.1 _d	24,586
<16(S3)	1.6	79.4 _a	62.2 _b	44.3 _c	59.8 _d	42,533
16–24(S2)	1.8	17.7 _a	35.4 _b	16.8 _a	18.6 _a	12,089
24–75(S1)	2.9	2.9 _a	2.5 _a	38.8 _b	21.7 _c	4,878
<0.8(N)	2.3	0.5 _a	0.8 _b	0.3 _a	1.5 _c	408
0.8-1.2(S3)	1.4	9.0 _a	13.3 _b	3.2 _c	1.8 _d	4,807
1.2-2.4(S2)	1.8	22.3 _a	28.1 _b	15.7 _c	30.6 _d	14,338
>2.4(S1)	1.8	68.2 _a	57.7 _b	80.8 _c	66.1 _d	39,847

4.3.2 Factors determining of coffee yield trends in current farming systems

Table 4-2, and 3 show the significant (with criterion <0.05) multiple regression models using biophysical variables to explain coffee yield variability for the whole country. The stratified regression models developed

for the nine major agro-ecological zones of Rwanda are given in (Table 4-2) and stratification for the seven main lithological soil units are displayed in (Table 4-3).

The country scale multiple regression coffee yield model explains 21% of the total variability in coffee yields across the country. The independent explanatory location factors were annual rainfall ($\beta_{\text{rain}}=0.31$), average temperature ($\beta_{\text{T}}=0.18$), slope gradient ($\beta_{\text{slope}}=-0.18$), and soil cation exchange capacity CEC ($\beta_{\text{CEC}}=0.152$). Altitude, Soil pH and organic carbon did not have an independent significant contribution in the national model.

When the national data was stratified by agroecological zones, the model performance only improved in Impala (IMP) zone ($R^2=0.55$) in south west of Rwanda (See Figure 4-1 for coffee fields location). An alternative stratification by lithology related soil types resulted in sub-models explaining up to 93% of the yield variability (Table 4-3). More in-depth analysis therefore focussed on the by soil type stratified models only. The effect of size of specific soil units was noted. In general the smaller more homogenous units have multiple regression models with more predictive power than the larger more complex soil units. For example, for volcanic soils (P) unit the coffee yield model has the highest prediction ($R^2= 0.93$) but in this unit only 1,537 hectares of coffee (about 6% of the country coffee areas) is grown. The same applies for the yield models obtained for coffee planted on alluvium and down slopes ($R^2 = 0.79$) and valley colluviums ($R^2 = 0.73$), containing 3,689 hectares of coffee (14.6% of the country coffee areas). Coefficients closer to the national model (17-25%) were found for larger more heterogeneous soil units also containing more coffee area. Table 4-3 summarizes the obtained significant yield models with biophysical parameters and model coefficients (b) values and associated standardized coefficient beta β -values (in bracket).

Taking into account soil types (Table 4-3) and agro-ecological zones (Table 4-2), some factors tend to be more important in one region and less in other regions. The locations of the discussed areas are indicated on Figures 4-1&2. On basaltic soils mainly found in low elevation of Impala (IMP) zone in south west of Rwanda, Climate and soil factors are equally important in determining high yields. In these zones, gentle slopes ($\beta_{\text{slope}}=-0.31$), high rainfall ($\beta_{\text{rain}}=0.21$), low soil acidity ($\beta_{\text{pH}}=-0.46$), high CEC ($\beta_{\text{CEC}}=0.24$) and high soil organic carbon ($\beta_{\text{SOC}}=0.27$) were factors of high achievable yields obtained in these agricultural regions.

On Kivu shoreline (KS) in west climate factors appear more important than soil factors in explaining coffee yield variability. In this zone, high altitude ($\beta_{\text{Alt}}=0.40$), together with high rainfall ($\beta_{\text{rain}}=0.28$) and low temperature ($\beta_{\text{T}}=1.4$) produce high yields ($R^2=0.93$).

Coffee planted on more acid soils (low pH) developed on granite, shales, and amphibolites are located in Congo-Nile watershed divide (CND) in west, in South, central plateau (CP) and in eastern ridges (ERP). These zones contain more than 50% of the country's coffee areas and have a broad range of different soil types from low (Inceptisols/Cambisols) to high weathered acidic soils (Oxisols/Ferralsols). In these zones, soil factors are the most important determinant of yield outputs. Soil acidity ($\beta_{\text{pH}}=-0.25$), low capacity of cation exchange ($\beta_{\text{CEC}}=0.19$), together with low soil organic carbon content ($\beta_{\text{SOC}}=-0.16$) at high slopes ($\beta_{\text{slope}}=-0.22$) are the major co-interacting factors determining the current yield variability. Temperature ($\beta_{\text{T}}=0.29$) also has a somewhat stronger effect on yield levels than altitude ($\beta_{\text{Alt}}=0.06$) in the eastern lowlands.

Table 4-2: Coffee yield models stratified based on agroecological zones with parameters: b-values of biophysical factors (EM) are presented with importance of each factor in the model (standardized coefficients Beta: β -values) in the brackets. All factors are significant at 0.1% level ($P \leq 0.001$)

Biophysical factors	Country	CND	MPB	BH	CP	EL	ERP	IMB	IMP	KLS	
Constant		-6.69	-8.265	0.986	-0.992	7.806	-1.499	-6.873	3.471	27.403	-16.489
Altitude (m)		0.001 (0.12)	0.002 (0.16)	0.000 (0.07)		0.001 (0.14)	-0.001 (-0.13)	0.001 (0.43)		0.002 (0.24)	
Slope (%)		-0.024 (-0.18)									
Annual Rainfall (mm)		0.002 (0.31)	-0.005 (-0.26)	0.001 (0.23)	0.001 (0.09)		0.002 (0.09)		0.007 (0.47)		
Temperature (oC)		0.246 (0.18)	0.429 (0.30)		-0.302 (-0.18)	0.065 (0.39)	0.381 (0.27)			0.656 (0.33)	
pH _{H2O}					-0.438 (-0.14)			-1.031 (-0.42)	-5.210 (-0.96)	-0.091 (-0.08)	
CEC (cmol ⁺ /kg)		0.029 (0.15)	-0.032 (-0.16)	0.164 (0.39)	0.010 (0.13)		0.007 (0.18)			0.077 (0.24)	
Soil Organic Carbon (%)					-0.138 (-0.23)			-0.723 (-0.64)	-2.152 (-0.73)		
Adjusted R ²	0.21	0.11	0.31	0.10	0.07	0.18	0.18	0.20	0.55	0.14	
Std. Error	1.18	1.31	0.99	0.26	1.13	0.12	0.93	1.5	1.5	1.07	
Nb. coffee fields (n)	59,520	4,598	4,398	1,319	21,259	2,922	14,633	1,492	2,633	6,156	

Agroecological zones: KLS Kivu lake shore, IMP Impala, IMB Imbo, CND Congo-Nile watershed divide, MPB Mayaga plateau and Central Bugesera, CP Central Plateau, ERP Eastern Ridge and Plateau, EL Eastern Lowland, BH Buberuka Highlands.

Legend for Table 4-3: Lithology: *A:* Alluvium; *B:* Basic rocks: gabbro, basalt, dolerite, and amphibolites; *Bv:* Basaltic rocks; *G:* Acids rocks: granite, tonalite and gneiss; *I:* shales, slates, phyllites, sandstone schist, phyllite; *O:* Down slopes and Valley colluviums; *P:* Trachytes: volcanic ash deposit granite

Table 4-3: Coffee yield models stratified based on lithology (soil types) with parameters: b-values of biophysical factors (EM) are presented with importance of each factor in the model (standardized coefficients Beta: β -values) in the brackets. All factors are significant at 0.1% level ($P \leq 0.001$)

Biophysical factors	Lithology disaggregated models of coffee yields						
	(A)	(B)	(Bv)	(G)	(I)	(O)	(P)
Constant	0.16	0.757	0.635	6.248	-9.555	7.397	-38.295
Altitude (m)	-0.001 (-0.269)		0.001 (0.117)	0.001 (0.061)		-0.007 (-0.389)	0.002 (0.402)
Slope (%)	-0.020 (-0.440)	-0.027 (-0.309)		-.021 (-0.155)	-0.025 (-0.200)	-0.037 (-0.249)	
Annual Rainfall (mm)	0.001 (0.268)		0.003 (0.214)		0.003 (0.405)	0.003 (0.184)	0.004 (0.278)
Temperature (oC)		0.084 (0.085)			0.355 (0.291)		1.711 (1.407)
pH _{H2O}		-0.264 (-0.456)		-1.126 (-0.251)			
CEC (cmol ⁺ kg ⁻¹)	0.021 (0.171)	-0.024 (-0.126)	0.043 (0.244)	0.016 (0.048)	0.040 (0.188)		
Soil Organic Carbon (%)		0.421 (0.274)	-1.952 (-0.224)	-0.021 (-0.162)		0.442 (0.662)	
Adjusted R2	0.79	0.25	0.72	0.17	0.20	0.73	0.93
Std. Error	0.25	0.76	0.71	1.2	1.16	0.83	0.37
Nb. coffee fields (n)	609	3,333	2,836	13,876	38,299	454	113

4.3.3 Modeling achievable coffee yields

Figure 4-4a shows 'achievable' yields as estimated using soil stratified multiple regression coffee yields models. These yield levels are valid for local biophysical conditions and the current average small holder management. Of course one would expect higher yields in commercial-scale fields, where it is feasible to ensure adequate farm management by controlling erosion, weeds and pests. Paired observations between 'achievable' yield (Yach) and 'actual' yields 2015 (Yact) (Figure 4-3b) gives with systematic yield difference (Yach - Yact) of 473 kg ha⁻¹, which represent 40.2% of the actual national average yield (703 kg ha⁻¹). This suggests that the coffee yield could be higher on average. The yield model map (Figure 4-4a) also depicts where lower yields occur as result of local climate and soil management constrains (Yach < 0.5 t ha⁻¹). The yield model performed relatively well in predicting achievable yield (R²=0.6 or 0.5 assuming that the intercept is zero). Furthermore, the yield model demonstrates that not all regions have equal potential for producing high yields within the current smallholder coffee system. The map (Figure 4-4a) highlights areas of high coffee yield across the country. In West, Impala (IMP) and Kivu Lake shoreline (KS) remain the areas with highest achievable yields (2 t ha⁻¹). The Congo-Nile watershed divide (CND) presents very few local areas with high achievable yields. In central plateau (CP) area the high achievable yields are on the lower edges.

In the eastern ridges and plateau (ERP) and lowland areas (EL), high yields are only envisaged in two hotspots, one in the upper Muvumba catchment in north east Rwanda and in central eastern ridges and plateau area.

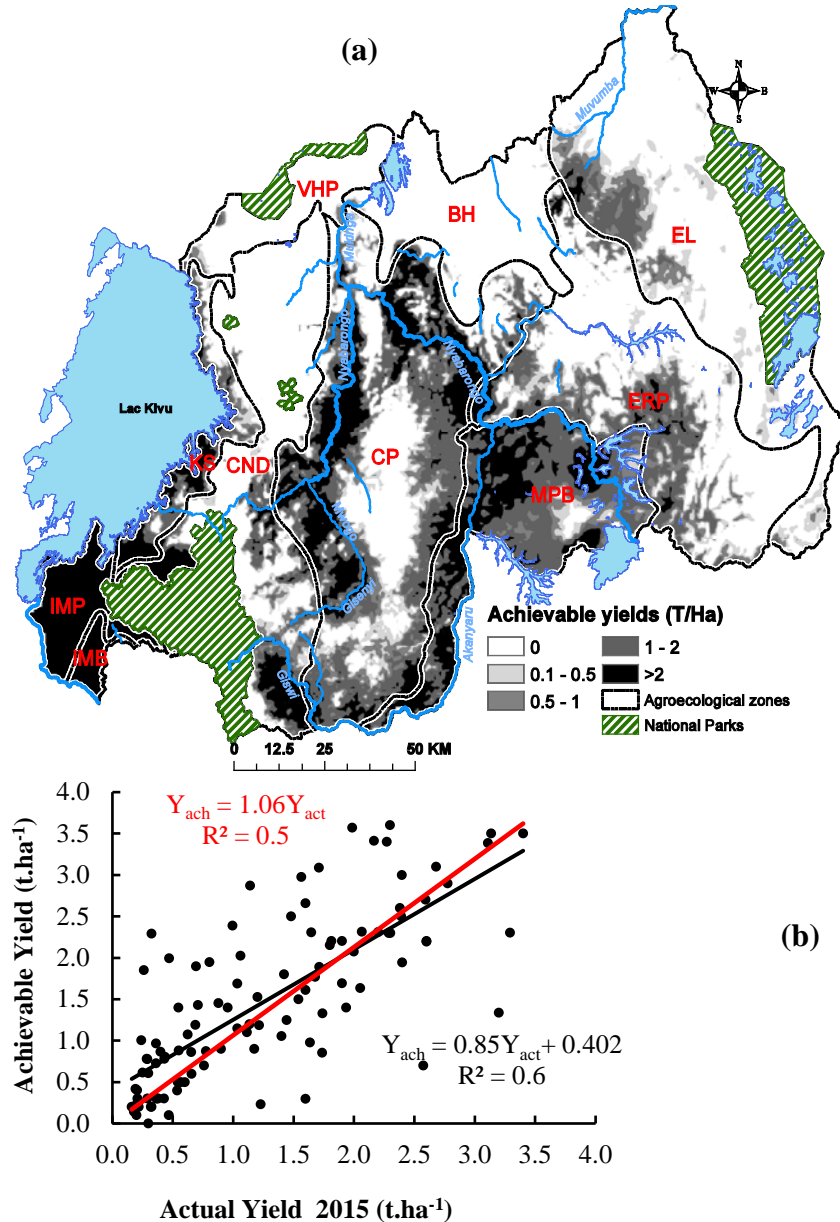


Figure 4-4: Achievable coffee yield map (a) calculated using the aggregated soil-based sub models. The predicted achievable yield is assessed against the actual yield for the season 2015 (b). The model performed relatively well (adjusted $R^2=0.6$). The overall mean difference between achievable yield and actual yield is estimated to 473 kg ha⁻¹.

The results of spatial autocorrelation confirmed the trend in both actual and achievable yields (Table 4-4). The distribution of coffee yields follows the environmental factors described in section 4.3.2 with specified weights. The yield models performed relatively well in predicting achievable yield (Adjusted $R^2=0.6$). The results of spatial autocorrelation of the model residuals (table 4) also show that the residuals are spatially random. Moran's Index was 0.126 while Expected Index is -0.006 (Z-score=1.272, P-value=0.203) meaning residuals are not significant, and that larger values of residuals are intermixed with small values confirming the fairly prediction of achievable yields.

Table 4-4: Spatial autocorrelation results showing the spatial trend of coffee yields across the country: Positive value of Moran's index, higher than the expected index indicates that coffee yields are not randomly distributed, instead present a clustering trend.

	Moran'S Index	Expected Index	Variance	Z-score	P-value
Actual yield	0.815	-0.000277	0.000948	26.50	0.000
Yield trend	0.234	-0.000017	0.000016	58.91	0.000
Residuals	0.126	-0.006	0.011	1.272	0.203

4.4 Discussion

4.4.1 Coffee yield trends and biophysical relationships

We have observed that coffee yields in Rwanda vary widely across the country's agro-ecosystems and landscapes. We used the average yield of Coffee Arabica for 4 years from 2011 to 2014 and environmental related variables at coffee field locations to determine which biophysical factors are associated with coffee yield variability in Rwanda. Multiple linear regressions models for coffee yield and local biophysical conditions of coffee fields clearly indicate that no single variable explains the yield variability across the country. Rather, soils, climate and topography interact in regionalized models to explain up to 70% of the observed coffee yield variability.

These models indicate that topography and specifically altitude has a positively relationship with coffee yield in most areas with the exception of the areas with soils developed in basic and shale or slate rocks in the central plateau. On the other hand, slope demonstrates a negative relation with coffee yield in most regions. Slope reduces achievable yields but as it is an inherent property of the undulating highlands of Rwanda it cannot be easily modified except by terracing existing coffee farms, which will take them out of production for many years.

The relationship between coffee yield and the climatic factors annual rainfall and temperature is more subtle as already demonstrated by the categorical comparisons made in Table 1. Low temperatures make the Congo-Nile

watershed divide (CND) and Buberuka highlands (BH) less suitable for coffee growing (Mukashema et al., 2016; I. Nzeyimana et al., 2014). This relationship reinforces the notion that increasing temperature can have a significant influence (Jayakumar et al., 2016) on coffee production (Craparo et al., 2015). Current average coffee yields in the south and central plateau (CP) are mainly limited by suboptimal soil conditions (pH, CEC and SOC). Soils developed in schistose and granitic parent materials on moderately sloping hillsides (I. Nzeyimana et al., 2014) are characterized by high acidity which seems an important limitation for better performance of individual coffee plants (Castro-Tanzi et al., 2012). This case is in south and central plateau with soil pH between 5.0-5.8. Although our findings (table 4-2, 4-3) seem to contradict the general principle, this is not the case because the pH of central plateau is within the tolerable limits (pH between 4.5 and 5.5). And yield in this agricultural region (figure 4-3a) is not high (0.5-1 t ha⁻¹). In western highlands (Impala (IMP), Imbo (IMB) and Kivu lake shore (KL)), high achievable yields are obtained in upper Kivu lake catchment where soil pH is optimal to coffee growing. This region has high favourable soils with clayey soils developed on basalt, together with abundant rainfall brought by wet winds from the South Atlantic that pass over Congo basin and Lake Kivu (Muhire & Ahmed, 2015) and moderate temperatures, produces high coffee yields (I. Nzeyimana et al., 2014; Verdoodt & Van Ranst, 2003a). In eastern ridges (ERP) and lowlands (EL), coffee is grown under relatively low annual precipitation (<900mm). More importantly, is however the rainfall distribution in this region and specifically the length of the dry season of four to five consecutive dry months (May-September) caused by dry anticyclones of saint Helena and Azores (Muhire & Ahmed, 2015). As a result of overall increasing temperatures, evapotranspiration potential (ETP) during the dry season also increases and this counteracts the benefit of possible shorter dry seasons that could be induced by the positive impact of Indian Ocean winds on water availability in eastern savannah.

4.4.2 Impact of climate change

Figure 4-5a,b shows that future increase in temperature by 0.35°C per decade will potentially cause an increase in coffee yield by 559 kg ha⁻¹ (i.e. 44.3% of the current average yield), while an expected increase of 1.4°C by 2050 would cause an increase of coffee yield by 818 kg ha⁻¹ (about 53.8% of the current yield) . These increases are all relative to the benchmark of the annual average temperature of 22°C in 2010. Similar trends are related to projected increase of rainfall by 4mm/yr (i.e. 40mm per decade). Rainfall increase is predicted to increase yield by 553 kg ha⁻¹ (44%) in 2020 and 793 kg ha⁻¹(53% of the current average yield) by 2050 (Figure 4-5 c,d). Both rainfall and temperature increase on average about 639 kg ha⁻¹ (47.6%) by 2020 and 1.1 t ha⁻¹ by 2050, about 61.8% increase from the current average

yield (Figure 4-5e,f). These increases seem to be confirmed by van Soesbergen et al. (2016) who envisage on average a yield increase of approximately 100% by 2050 due to projected climate change in Albertine rift region including Rwanda. It seems that this is in line with results of Ovalle-Rivera et al. (2015) who found that by 2050 Arabica coffee regions in Ethiopia and Kenya are projected to become more suitable with higher expected yields. These climate scenarios assume no further land conversion for coffee and no shift of coffee cultivation is envisaged; instead highlight the necessity to intensify and adapt the current farming systems to climate change while avoiding further land degradation (Pretty, Toulmin, & Williams, 2011). However, It should be noted that with temperature increase and changing precipitation patterns, pest and disease pressure will increase potentially causing yield level reduction (Ovalle-Rivera et al., 2015) a factor not directly accounted for in our projection.

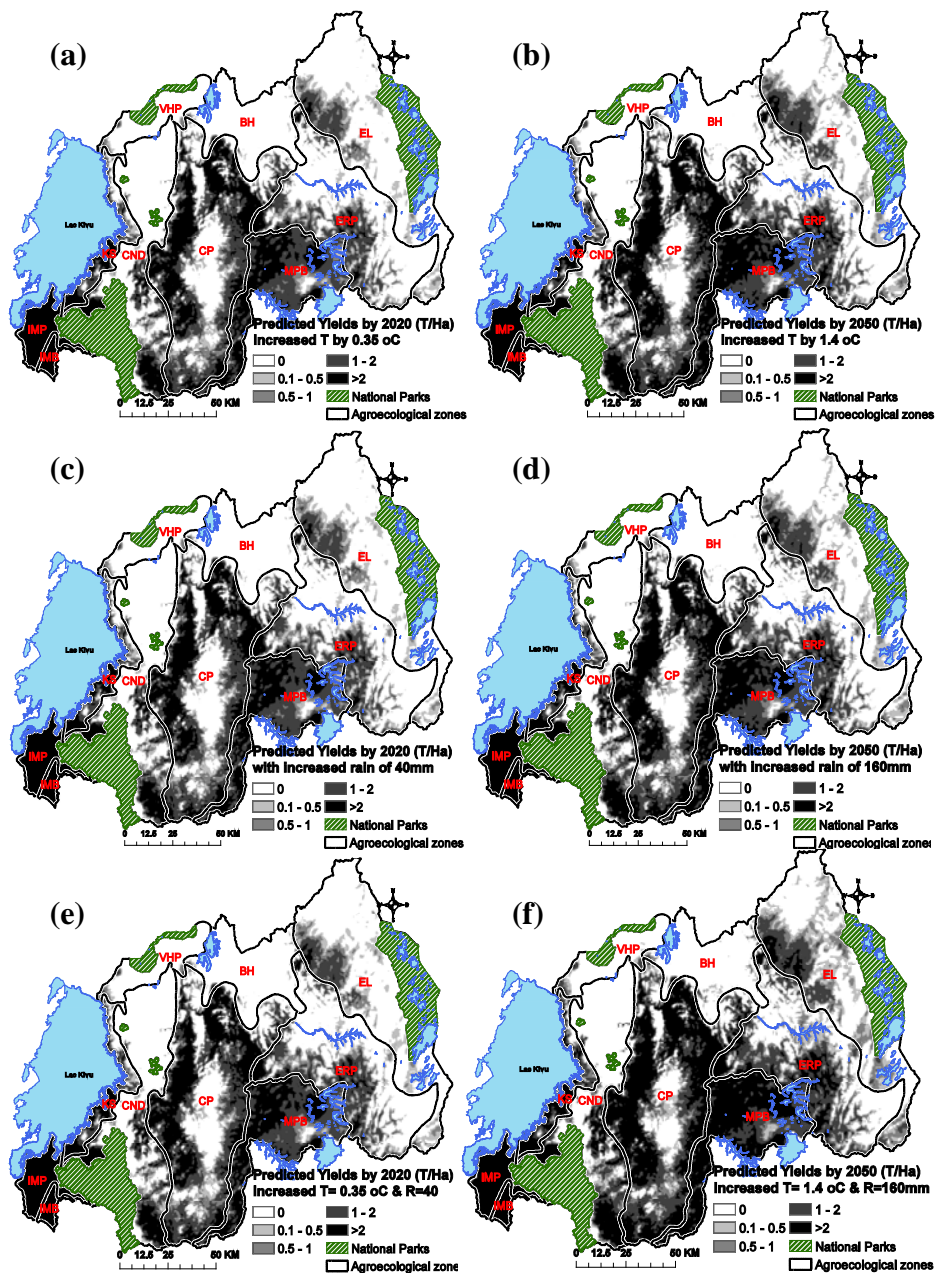


Figure 4-5: Projected coffee yields by 2020 and 2050 using mean annual total rainfall and average temperature trends estimated by McSweeney and Semafara (2011) and Muhire and Ahmed (2015) using available climate data of Rwanda from 1931-2010.

4.4.3 Impact of improved land management

Identification of key factors that determine coffee yield variability could support the government of Rwanda to provide applicable site-specific adaptation strategies to boost yields to calculated 'achievable' levels within in current smallholder agricultural systems. Both topographic factors (altitude and slope) of a coffee field cannot be easily changed. Lack of precipitation can potentially be compensated by irrigation, while temperature can be influenced by using shade trees. The use of agroforestry systems may be an economically feasible way to protect coffee plants from extremes microclimate conditions (Lin, 2007) and sustain coffee production (Souza et al., 2012) of Eastern province of Rwanda.

Increasing soil pH or CEC is also difficult to achieve because it would require lots of lime which has proven to be cost inefficient (Cyamweshi et al., 2014). These measures are either too expensive or long term and therefore are not considered to be a feasible option for the current small-scale coffee farming systems the coming decades.

The only factor affecting yield levels that can be feasibly managed at decadal time scales by smallholder farmers is soil organic matter content. Agroforestry could be a solution to produce enough biomass to mulch the coffee field (Roose & Ndayizigiye, 1997) and increase soil organic matter. Such mulch would not only control soil erosion but also prevent fertility depletion (Cyamweshi et al., 2014) by adding nutrients in soil and consequently improving coffee yields (Bucagu et al., 2013). A drawback of agroforestry is that it takes many years for shade trees to grow and the resulting competition between coffee and shade trees for light, water and nutrients. More standard organic agriculture methods could also achieve higher soil organic matter levels (Rutunga et al., 2007).

We have therefore explored potential gains in achievable coffee yields by increasing soil organic matter content levels in coffee field soils (Figure 4-6) without identifying which management strategy is required to achieve this increase. The projected coffee yield increase by increasing soil organic carbon by 0.6% by using improved land management could potentially boost yield levels by approximately 50%.

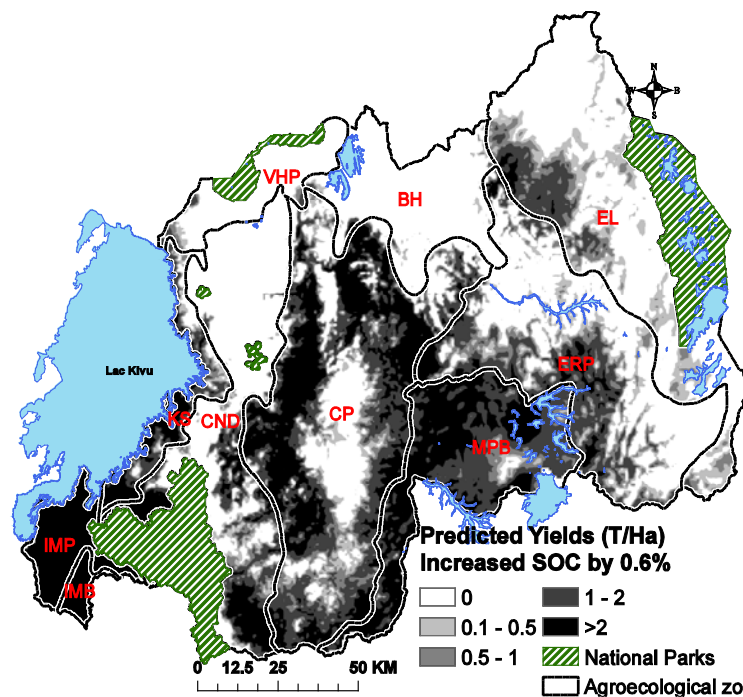


Figure 4-6: Projected coffee yields through increasing soil organic carbon by 0.6% from the minimum SOC of 1.2% required for coffee production

4.5 Conclusions

Arabica coffee yield of smallholder farms in Rwanda is currently suboptimal compared what can be achieved under improved management conditions given current biophysical constraints. The mean national difference between achievable and actual yield is estimated to 473 kg ha^{-1} which represents a gap of 40.2% nationally. This study aims to better understand which location factors are associated with current yield level variability. Topography, climate and soil data at coffee field location were analysed with average coffee yields for the period from 2011 to 2014 to examine the relationship between environmental factors and coffee yield in Rwanda. These factors were used to determine potential areas of high and low yields. The results show that in the western regions altitude, temperature and rainfall, and high nutrient retention capacity are the major determinants of high achievable coffee yield, while in eastern lowlands coffee yield is constrained by shortage of rain or water availability in depleted soils. Extreme low temperatures make the Buberuka highlands marginal for coffee production. In South and central part of Rwanda, soil acidity, nutrient and organic carbon depletion are the most limiting factor of higher coffee yields. It clearly emerges that coffee yield is not only based on soil fertility factor but on a combination of topographic, climate and soil factors which are all relevant to policy making and

development planning of the coffee sector. Based on our results, we recommend next to the ongoing distributing of mineral fertiliser, to aim for stimulating improved water and organic matter management in coffee farms.

The study also indicated that expected climate change for Rwanda will have an overall positive impact on coffee yields. Projected combined increase of temperature and rainfall would potentially increase yield levels by almost 50% in 2020 and by approximately 60% in 2050. An increase of average soil organic carbon content by 0.6% by using improved land management could potentially boost yield levels by another 50%. Implementation of more erosion control measures, agro-forestry and integrated small-scale coffee and livestock farming has the capacity to boost the productivity of coffee by elevating soil organic carbon contents. Smallholder coffee farmers can implement these proposed actions themselves to ensure higher yields and net income (assuming a stable coffee price) in the long term, making them less dependent of external inputs.

5 Understanding geographical origins of sensory quality for specialty coffees of Rwanda

This chapter is submitted as: Mukashema, A., Veldkamp, A., and Gatarayiha C.M. Understanding geographical origins of sensory quality for specialty coffees in Rwanda. In *PLOS One*

5.1 Introduction

Coffee is the most important traded agricultural commodity providing a living to more than 125 million people globally (Tran et al., 2016) and to over 355 thousand Rwandan families in particular. In Rwanda Bourbon arabica coffee produced in highlands can obtain premium prices on auction markets as high as US\$67.5 per kilogram in the 2015 Cup of Excellence competition. The most recent data from the National Agricultural Export development Board (NAEB) indicate that over 22,185 tons of green coffee beans have been produced of which 18,750 tons were exported in 2015. Arabica coffee has the potential to produce coffees that result in a high-quality beverage (Ramos, Ribeiro, Cirillo, & Borem, 2016; Wintgens, 2009) and in Rwanda, has proofed to produce flavour and quality that is widely sought by the consumer market.

However, coffee business development is lagging behind other industries such as for example wine and oil in terms of product differentiation (Antoine, Hoo Fung, & Grant, 2016; Belletti, Marescotti, Sanz-Cañada, & Vakoufaris, 2015). Despite its potential for export growth in Rwanda, the sector faces a challenge of serving the specialty coffee market because traders demand a consistent high quality. Coffee in Rwanda is grown exclusively by smallholder farmers. Such small-scale farming system is a challenge for specialty coffee market because it involves collecting and mixing coffee from many different producers to meet the quantity demand without compromising the quality. This problem is common for other African countries growing specialty coffee in smallholder farming systems. Mixing a poor quality coffee lot into a better lot tends to reduce the quality of be better lot much more than it improves the quality of the poorer lot.

In Rwanda, once coffee cherries are well ripe, they are harvested by farm owners. Subsequently they are transported by them to the nearest coffee washing station, typically owned either by farmer cooperatives or by local investors, for fully washing, drying and processing into green coffee beans. With such a standardized post-harvesting processing chain, the influence of off-farm factors on quality is assumed to be controlled and comparable. This post-harvest processing technique was introduced in 2002 when Rwanda began to put into action a strategy for enhancing the quality of coffee production so that it could be sold at the speciality coffee market. It has progressively replaced a traditional processing that used to be done at household level which was negatively affecting the quality of the processed coffee. The traditionally produced coffee was typically sold at the 'ordinary' coffee market. Through public-private partnerships the government of Rwanda supported coffee farmers to establish coffee washing stations which amounted up to 245 in 2016. This strategy tremendously improved the overall quality, enabling Rwanda to access specialty coffee markets since 2008. Since then smallholder coffee farmers are connected with the

worldwide market for specialty coffee through annual Cup of Excellence (CoE) competitions and other promotional programs of Rwanda coffee.

However, much of the effort that Rwanda invests in assisting farmers to produce high quality coffee, sensory analysis through CoE has revealed that arabica coffees in Rwanda produces different tastes, and although some have never been able to win the CoE competitions, they were as well of good quality and taste. Very little is known about the source of these quality differences, until recently where there has been increased interest in associating the flavour and quality of coffee product with its geographic origins (Antoine et al., 2016). Only a few studies are known on the relationships between different coffee genotypes their production environments and sensory quality (Ramos et al., 2016). For African coffee growing countries, these studies are completely lacking due to scarce spatial data on coffee farms and their sensory quality. This understanding is important for sustainable production of specialty coffees and further protection of special land uses (Alves, Vieira, Volpato, Lacerda, & Borem, 2016) thus securing economic returns to coffee growers (Bosselmann et al., 2009). Yet, in order to satisfy the speciality coffee market, farmers and processors need to know the geographical area, production factors and product quality characteristics (Quiñones-Ruiz et al., 2016). According to Rolle et al. (2012) adding information about the geographical origin of an agricultural product can also facilitate its acceptance in the global market.

For specialty crops, Zou et al. (2012) state that the location where crops are grown generally co-determines their final quality. Coffee is a specialty crop whose diversity in production systems allows products with defined attributes for different markets (Silva, de Queiroz, Ferreira, Correa, & Rufinod, 2016). Production of high quality Arabica coffee is considered to depend on three main factors: the genetic resource, environmental conditions, and management (Bosselmann et al., 2009). It has been noted before that coffee quality highly depends on geographical factors (de Toledo et al., 2017) and that specific coffee sensory attributes are influenced by location-specific environmental factors (Vasconcelos, Franca, Gloria, & Mendonca, 2007). The association of arabica coffee genotypes with specific environmental conditions is known to produce outstanding coffees. But the same genotypes grown under different environmental conditions can also produce poor sensory quality coffee (Wintgens, 2009). In Rwanda only bourbon arabica coffee is grown in smallholder systems and they produce highly variable sensory qualities despite the uniform genotype and post-harvest processing. We therefore hypothesize that the high variability in coffee sensory quality is related to the local growing conditions on the small holder coffee fields where the coffee beans are produced.

To produce a specialty coffee, it is important to take into consideration location attributes such as topography, soil type, and climate such as rainfall

distribution and seasonal temperature variability (Chemura, Kutywayo, Chidoko, & Mahoya, 2016; Joët et al., 2010; Ovalle-Rivera et al., 2015). Altitude determines microclimate in which coffee is grown and for the southern hemisphere (where Rwanda is located), fields with a face slope oriented to the south receive a smaller amount of solar radiation than do those oriented to the north (Bertrand et al., 2012). This lower amount of energy received affects the process of fruit ripening and thus the product quality (Bertrand et al., 2012; Silva et al., 2016).

Topographical factors such as altitude and slope exposure (Avelino et al., 2005; Bosselmann et al., 2009; Decazy et al., 2003), climate, specifically both precipitation (da Silva et al., 2005; Decazy et al., 2003) and temperature (Bertrand et al., 2012), and soil (Marquetti et al., 2016) are known to be able to determine the characteristics of coffee flavor profiles. Interaction of these factors gives the coffee a unique identity (i.e. profile) that defines the final quality as they jointly determine the chemical composition of the coffee beans (Marquetti et al., 2016). Coffee sensory quality has been evaluated by criteria such as flavor, taste, bean size, colour, shape, roast potential, processing method, and crop year, among others (Choi, Choi, Park, Lim, & Kwon, 2010). However, among these parameters, flavor and taste are the most important criteria (Vasconcelos et al., 2007) and they are potentially influenced by environmental factors such as soil and climate, as well as genetic factors such as species (Marquetti et al., 2016).

In order to test our hypothesis, that coffee field characteristics are a main determinant of coffee quality we will analyse to what extent a set of topographic, climatic and soil factors can statistically explain the observed variability in coffee sensory results of the cup of excellence competitions from 2008 to 2015. Once we know which factors are relevant we can gain some insight in the local constraints of specialty coffee production in Rwanda. With this knowledge we can potentially identify new areas that could be considered for future speciality coffee production in Rwanda. Given known future scenarios we can also explore how sensitive the speciality coffee production in Rwanda is for a plausible projected climate change scenario.

5.2 *Materials and Methods*

5.2.1 Coffee production and post harvesting processing chain

Rwanda traditionally grows *coffea arabica*, bourbon variety in small holder farming systems farms. Coffee is cultivated on over 36,000 hectares (NAEB, 2016) i.e. over 1% of the total country land, which produces on about 22,185 tons of green coffee per year on average. Coffee in Rwanda is grown in sunny monoculture. The plantations are located on steep slopes and at various

positions relative to the sun. Plantations are established at different altitudes, climate regimes and on different soil types. The plantations are usually made in contour lines, with line spacing of 2m × 2m and are conducted for many years (over 50 years undergoing several pruning for regeneration). Detailed information on the distribution of coffee fields in Rwanda can be found in chapter 2 (Mukashema et al., 2014) and environmental and socio-economic conditions in which coffee is grown is described in chapter 3 (Mukashema et al., 2016).

As soon as the coffee 'cherries' are ripe, farmers usually harvest them by hand, which they do every few days since they only pick the red (ripe) fruit. To extract the bean, the fruit hull has to be removed in a rigidly-defined multi-step process. Firstly, growers bring their cherries to the nearest coffee washing station (CWS) to get them weighed and de-pulped. Farmer is paid directly in cash for his unprocessed cherry. The cherries are then pre-sorted by floating them on water, which separates the heavier cherries from the lighter ones (which is at the same time a classification of quality). Next, the cherries are run through pulping machines that remove much of the fruit from the bean. The beans are then fermented in holding tanks, as fermentation makes it easier to remove the remaining mucilage. After the fermentation process is accomplished, the beans are washed up to four or five times. The beans are sorted again and left to float for 24 hours. After this period, the beans are laid out on sorting tables, and workers pick out damaged or irregularly-shaped beans. After that, the beans are dried in the sun until they reach the desired moisture level of about 12 per cent, which requires between ten days to two weeks depending on the weather and sunshine intensity. This 'parchment' coffee is then packed in burlap bags and stored for shipment by an exporter. Hulling coffee to green coffee beans is done by the local millers. No parchment coffee is allowed in export according to Rwanda export regulations. Only green or roasted coffee is exported.

5.2.2 Coffee sample preparation and sensory test

The coffee washing stations (CWS) play the primary role in ensuring that the small amounts of different farm produces are properly combined during the above described post-harvest processing chain so that uniform lots can be provided at the cupping stage. Cupping lots varied increasingly from 2008, and have reached 207 coffee lots in 2015 collected from 245 CWS. Not all CWS participate in the cupping exercise (Figure 5-1). There are CWS without coffee lots, as there are other CWS presenting more than one sample lot. Lot samples are prepared by coffee washing stations for the evaluation of beverage quality, knowing the geographic location (farm owner, latitude, longitude) where lots are harvested.

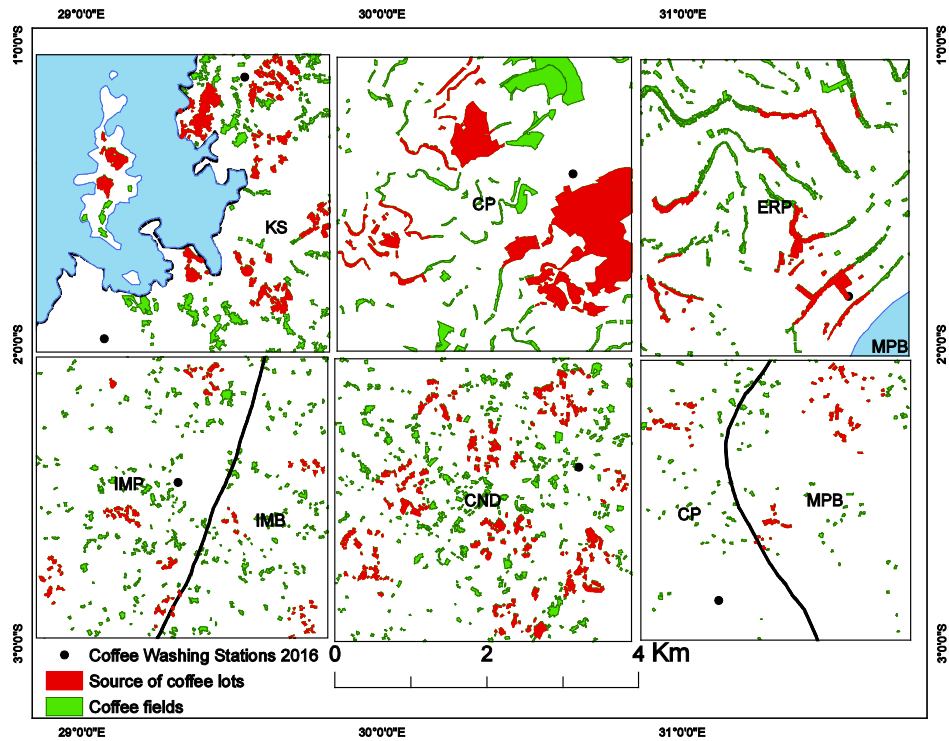


Figure 5-1: Map showing the provenance of coffee sample lots (used in the CoE) within the coffee washing station (CWS) supply areas. To avoid the loss in coffee quality, each coffee washing station is supplied by coffee farms located within 3 km radius on average. Where there are limited numbers of coffee washing stations, there are collection centres where farmers bring coffee cherries, and CWS arranges the transport. Agroecological zones: KS Kivu lake shore, IMP Impala, IMB Imbo, CND Congo-Nile watershed divide, MPB Mayaga plateau and Central Bugesera, CP Central Plateau, ERP Eastern Ridge and Plateau.

The samples brought to the national laboratory are analysed for their quality using a sensory quality test process popularly known as the cup test. Samples are roasted separately. According to Cup of Excellence (CoE) protocol, for the composition to be cupped and evaluated for effusion, a ratio of 12 g ground coffee per 180 ml hot water is used. Special care is given to the water used for tasting to ensure that it is clean and odourless and that it has a suitable pH. Water is boiled and subsequently added to the ground coffee. This water is added directly to the ground coffee in a circular motion, filling the cup until it is completely full. After this hydration process, the samples remain undisturbed for 3–5min prior to proceeding to the review.

There are three step of evaluation: the pre-selection (round I), the cupping by national jury (round II), and the cupping by the international jury (round III) giving the final grade which determine the final quality and the price at auction. Members of pre-selection and national jury are well trained local

staff in coffee cupping organized by the National Agricultural Export Board (NAEB). International jury is recruited through the Alliance for Coffee Excellence (ACE). Some local experts are also selected from the large group of National cupping experts from different parts of the country, to work with the international cupping expert. National cuppers to be part of the international jury are rated on their score correlation, range of scoring and confidence for the samples cupped during rounds I and II. The Cup of Excellence protocol is used: 90–100 points are special coffees (and receive the presidential award); 85–89.99 are excellent coffees; 80–84.99 are fine coffees; and below 80 are coffee of less than premium quality, which is not classified (ordinary coffee). According to this classification, a specialty coffee cannot be defective; it must also have at least one well-defined attribute, such as aroma, flavor, body or acidity. Coffees that have grades below 80 are not classified. This designation means that a coffee that does not possess features that reflect a unique and particular flavor nor has a quality that is defective will be removed from classification.

From each coffee sample, five cups are prepared and analyzed according to the sensory characteristics described above. The following sensorial attributes are evaluated: drink, sweetness, acidity, body, flavor, aftertaste, and balance with grades ranging from 0 to 8. By combining these scores, the overall grade for each sample is determined. According to the Cup of Excellence protocol, each sample begins with a pre-set score of 36 points, to which will be added the scores of each attribute. Scoring the coffee lots is on a scale from 0 to 100 points. Those samples that presented scores over 80 are classified as specialty coffee.

Figure 5-2 shows the geographic location of coffee washing stations where coffee beans from surrounding coffee fields are processed. Coffee washing stations prepare and supply coffee samples to the national cupping laboratory, for analysis of the sensory quality. In total 503 sample lots were analysed in annual Cup of Excellence competitions 2008-2015.

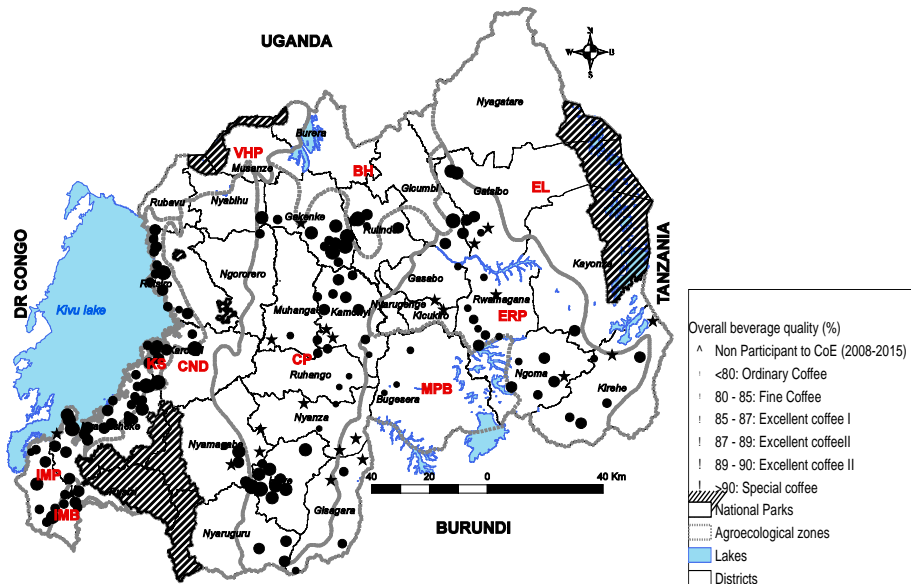


Figure 5-2: Map showing the overall sensory quality distribution and coffee washing stations (CWS) where coffee samples for sensory analysis are prepared. In total 503 coffee samples were brought by CWS to the national cupping laboratory for analysis of the sensory quality during the Cup of excellence (CoE) competitions 2008-2015. These samples represent coffee fields in the CWS supply area. Agroecological zones: KS Kivu lake shore, IMP Impala, IMB Imbo, CND Congo-Nile watershed divide, MPB Mayaga plateau and Central Bugesera, CP Central Plateau, ERP Eastern Ridge and Plateau, EL Eastern Lowland, BH Buberuka Highlands, VHP Volcanoes and high plain. Winning coffee washing stations and farmer’s representatives are available online at <https://www.allianceforcoffeexcellence.org/en/cup-of-excellence/country-programs/rwanda-program/>

5.2.3 Explaining the variation of coffee sensory quality

5.2.3.1 Biophysical data and coffee sensory data

To investigate the associated factors with the observed Arabica coffee beverage quality, we used the Cup of Excellence (CoE) competition results from 2008 to 2015 and environmental characteristics at coffee field locations where the tasted lot was harvested. The available CoE results were in total 503 samples of which 207 samples are from CoE 2015, 173 samples from CoE 2014, 15 from CoE 2013, 26 from CoE 2012, 36 from CoE 2011, 22 from CoE 2010 and 24 samples from CoE 2008. We had to aggregate several years in order to increase the sample size. These samples represent about 14,286 coffee fields (see Figure 5-1) used in sampled coffee lots of 207 CWS participated in the CoE. All 503 samples were considered in the study to ensure the country coverage. Using spatial join tool of ArcGIS10.3 software, the coffee map of Rwanda (Mukashema et al., 2014) was used to join the sensory quality to the coffee fields where it was produced. Data on

topography, Climate and Soil at coffee field location were obtained from various source described by Mukashema et al. (2016). Table 5-1 shows the situation in biophysical properties of coffee fields.

Table 5-1: Summary statistics of environmental related parameters calculated on randomly selected 59, 520 coffee fields (i.e. 50% of the national coffee fields). Average values of each environmental variable are presented per agroecological zones. Standard deviations are in bracket (see chapter 4 for more detailed description).

Environ. variables	IMB	IMP	KLS	CND	CP	MPB	ERP	EL	BH
Altitude (Alt m)	1,269 (125.4)	1,608 (126.1)	1,650 (114.3)	1,955 (144.7)	1,702 (127.6)	1,461 (86.0)	1,604 (132.2)	1,437 (76.5)	1,925 (124.8)
Rainfall (R mm)	1,581 (179.5)	1,601 (160.6)	1,269 (72.3)	1,439 (121.1)	1,256 (109.6)	1,133 (69.1)	1,050 (64.7)	988 (83.2)	1,247 (100.4)
Temperature (ToC)	20.8 (0.9)	19.8 (0.6)	19.6 (0.6)	17.9 (1.0)	18.9 (0.7)	20.5 (0.7)	19.7 (0.7)	19.9 (0.8)	18.2 (0.9)
Soil Sand (Sd %)	25.0 (7.1)	28.1 (10.9)	45.9 (7.2)	42.8 (5.7)	52.3 (8.0)	51.7 (5.0)	37.7 (7.6)	53.7 (8.0)	37.2 (7.1)
Soil Clay (Cy %)	47.1 (4.3)	52.9 (12.4)	34.7 (6.2)	38.5 (5.2)	30.9 (6.8)	36.0 (5.8)	43.7 (6.6)	30.2 (5.5)	33.9 (7.4)
Soil pH (pH-KCl)	3.9 (0.7)	4.6 (0.4)	4.2 (0.3)	3.5 (0.8)	4.2 (0.3)	4.6 (0.3)	4.6 (0.4)	4.7 (0.5)	4.1 (0.3)
Soil CEC (cmol*kg ⁻¹)	34.4 (7.1)	32.4 (10.7)	16.3 (3.6)	18.2 (7.1)	11.2 (3.6)	9.5 (2.9)	15.2 (3.5)	12.3 (3.7)	13.4 (3.5)
Soil org. Carbon (%)	4.0 (1.6)	3.3 (0.8)	3.6 (2.6)	5.0 (2.6)	3.2 (2.0)	2.5 (0.9)	3.3 (1.4)	2.2 (1.2)	4.3 (1.6)
C/N ratio (C/N)	12.7 (1.2)	12.4 (0.7)	12.7 (2.9)	13.8 (1.2)	13.5 (0.9)	13.2 (1.1)	13.8 (0.8)	13.4 (1.7)	13.7 (1.0)
Phosphorous (P ppm)	22.4 (13.7)	15.8 (15.2)	10.7 (9.4)	7.6 (8.6)	8.3 (12.5)	2.8 (4.6)	4.0 (11.3)	11.4 (11.0)	2.4 (3.7)
Potassium (cmol*kg ⁻¹)	0.5 (0.2)	0.5 (0.2)	0.3 (0.1)	0.3 (0.1)	0.3 (0.1)	0.4 (0.1)	0.4 (0.1)	0.4 (0.1)	0.3 (0.1)
Calcium (cmol* kg ⁻¹)	8.4 (7.3)	8.1 (3.6)	2.9 (1.7)	2.6 (1.0)	2.9 (0.9)	4.2 (1.3)	4.3 (2.2)	5.1 (3.2)	3.0 (1.2)
Aluminium (cmol*kg ⁻¹)	3.0 (2.7)	2.0 (1.9)	1.8 (0.9)	2.4 (0.8)	1.5 (0.8)	1.1 (0.9)	1.0 (0.7)	0.3 (0.3)	2.1 (0.9)
N=59,520	1,492	2,580	6,103	3,752	20,339	4,398	14,633	2,922	1,318

Agroecological zones: KLS Kivu lake shore, IMP Impala, IMB Imbo, CND Congo-Nile watershed divide, MPB Mayaga plateau and Central Bugesera, CP Central Plateau, ERP Eastern Ridge and Plateau, EL Eastern Lowland, BH Buberuka Highlands

5.2.3.2 Statistical analysis

We first analysed the spatial variation of coffee sensory quality using spatial autocorrelation and global Moran's Index applied to average coffee beverage quality measured from 2008 to 2015 samples submitted to the national cupping laboratory for CoE competition. Except defective samples which were automatically disqualified in the analysis, other samples available were used in this study to ensure the national coverage (see Figure 5-2). Moran's Index

indicates the behaviour of spatial autocorrelation, in which smaller value than the expected index demonstrates a trend for dispersion, whereas positive value indicates a clustering trend for contiguous areas (Silva et al., 2014). Second, we analysed the variance in the coffee sensory quality between agro-ecological zones of Rwanda by comparison of means and separation by Tukey tests ($P \leq 0.05$). To understand the local field source of variation of coffee quality, we used multiple (stepwise) regression analysis using a randomly selected subset of 80%. For each variable used the distribution variability was checked for a normal distribution. We identified relevant environmental properties (topography, climate and soil related) at coffee field location, to explain the coffee sensory quality variability across the country. Significant biophysical (topography, climate and soil) model parameters, model coefficients (b-values) explaining the variance in sensory quality and the level of importance (beta β -values) were estimated by the equation 1:

$$Q(\%) = b_0 + \sum_{i=1}^n b_i (EM)_{i(x,y)} + \varepsilon_i \quad (5.1)$$

Where, Q is sensory quality model and ε is the error associated with the model; b_i , $i=1 \dots n$ are model coefficients and EM are environmental (topographic, climate and soil) factors. Statistical analysis was performed in SPSSv.22 software. The quality model is selected on the basis of goodness-of-fit (adjusted R^2). Spatial autocorrelation of the residuals was checked with Moran's Index.

We then used raster data of significant biophysical parameters ($P \leq 0.001$) of the significant sensory quality model(s) to calculate a potential coffee sensory quality map. This map estimates the geographic origins of specialty coffees and predicts potential new areas where speciality coffees are most likely to be produced.

The validation of the predicted quality map was done by calculating the model fit between measured quality and modelled sensory quality expressed by the adjusted coefficient of determination (adj. R^2) and the minimized root mean square prediction error (RMSPE) per equation 2. For the model calculation 70% (CWS $n=144$) of the of the total CWS sample lots was used and the remaining (CWS $n=63$) unused 30% was used for quasi-validation purpose. This validation was done for both the national model and the regional model.

$$RMSPE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Q_m - Q_p)^2} \quad (5.2)$$

Where Q_m is the measured sensory quality at the cup of excellence competition, Q_p is the quality predicted by the sensory quality model. Beside

adjusted R^2 and RMSPE, we also assessed the spatial autocorrelation (Global Moran's I) on the model residuals (Q_m-Q_p) to see if the under/over-predictions are clustered or not. Once the residuals are spatially random, it means that a model fairly predicted the dependant variable in our case the coffee quality.

5.3 Results

5.3.1 Spatial variability of Arabica coffee sensory in Rwanda

Figure 5-3 presents the calculated Moran's Index value and the spatial autocorrelation significance tests for the overall quality scores at the 207 CWS participating in Cup of Excellence competitions 2008-2015. A Moran's Index of 0.23 demonstrates that the coffee sensory quality in Rwanda has a slightly clustered pattern. Given the z-score of 2.008, there is less than 5% likelihood that this clustered pattern could be the result of random chance ($P < 0.05$). This means that the spatial variation of sensory quality observed in 503 sample lots collected in all agro-ecological zones of Rwanda is slightly clustered. Hence there is a need to understand the driving location factors of this clustering. We therefore continued with a randomly selected 70% sub sample.

The quality scale was divided according to the Cup of Excellence (CoE) protocol for sensory classification of specialty coffees. As shown earlier in Figure 5-2 the plateaus and mountains regions produced coffee of good to excellent quality that are classified as 'special' coffees. The lowlands and Bugesera region do not produce premium quality coffee and therefore are classified as ordinary coffees.

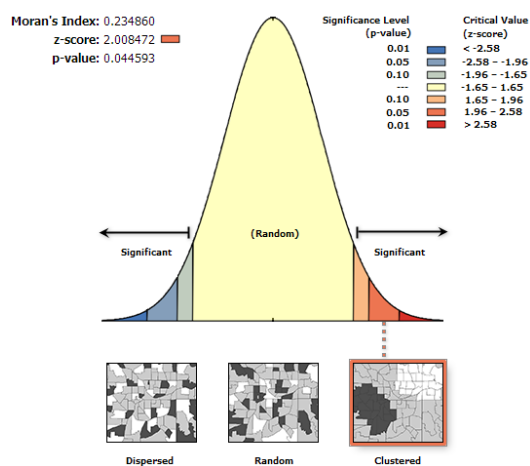


Figure 5-3: Spatial autocorrelation test and Moran's index showing the spatial dependence between coffee sensory qualities measured at 207 coffee washing stations using 503 coffee sample lots used in this study.

When the coffee sensory quality variability between agro-ecological zones (Table 5-2) is analysed and compared the pattern becomes clearer. We observed that the overall quality values were significantly different ($P < 0.05$) in comparison between central plateau and other AEZs. The overall quality scores ranged from 75.2 to 92.7% between agro-ecological zones, while their variability within AEZ regions ranges from 0.48 to 1.95 scores. Amayaga and Bugesera (MPB) regions have the lowest quality (75.2%) while the central plateau (CP) had the highest sensory quality (92.07%). Higher grades ($>90\%$) were mainly found in South and the northern plateau (CP). Although the analysis of coffee quality variance revealed a significant difference in sensory quality between agro-ecological zones ($F=375$, $P < 0.001$), Tukey post hoc test clustered them into two homogeneous subsets ($\alpha=0.05$): Impala, Kivu lake shore, Congo-Nile divide and Buberuka highlands in one set, central plateau and eastern ridges and lowlands in other set with the central plateau scoring high as compare to other agro-ecological zones. The mean difference between the central plateau and other agro-ecological zones ranged between 0.58 and 1.89% (see Table 5-3).

Table 5-2: Separation of coffee beverage quality means between agroecological zones in homogeneous subsets at $\alpha= 0.05$ according to Tukey HSD test.

AEZ	IMB	IMP	KLS	CND	CP	MPB	ERP	EL	BH
Min	86.2	86	83.2	82.2	78.2	75.2	80.1	83.9	85.4
Max	87.9	88.2	89.1	89.1	92.7	87.9	88.5	88.5	89.5
Mean	87.1 ^{ab}	86.9 ^a	86.4 ^a	86.8 ^a	88.3 ^b	86.8 ^a	87.5 ^{ab}	87.7 ^{ab}	86.5 ^a
SD	0.48	0.60	1.22	1.95	1.33	1.86	1.39	1.66	1.18
N	598	611	3,222	1,982	3,974	856	2,454	331	256

Agroecological zones (AEZ): IMB Imbo, IMP Impala, KLS Kivu lake shore, CND Congo-Nile watershed divide, CP Central Plateau, MPB Mayaga plateau and Central Bugesera, ERP Eastern Ridge and Plateau, EL Eastern Lowland, BH Buberuka Highlands. SD: Standard deviation, N is the number of coffee fields where coffee samples were collected.

Table 5-3: The mean difference ($\alpha=0.05$) in coffee sensory quality between agroecological zones of Rwanda

(I)	Agro-ecological zones	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval		
					Lower Bound	Upper Bound	n=14,286
	KLS	1.89 [*]	0.03	0.000	1.78	2.00	3,222
	IMP	1.37 [*]	0.06	0.000	1.17	1.56	611
	IMB	1.22 [*]	0.06	0.000	1.02	1.42	598
	CND	1.43 [*]	0.04	0.000	1.30	1.55	1,982
	MPB	1.45 [*]	0.05	0.000	1.27	1.62	856
	ERP	0.79 [*]	0.04	0.000	0.67	0.91	2,454
	EL	0.58 [*]	0.08	0.000	0.33	0.85	331
	BH	1.80 [*]	0.09	0.000	1.51	2.09	256

5.3.2 Biophysical factors of coffee sensory variation

Table 5-4 shows which analysed biophysical factors significantly contributed to the explanation of the coffee sensory quality variability for the whole country and specific agro-ecological zones using multiple linear regression analysis (stepwise). All studied variables were normally distributed, so no transformations were necessary. Spatial autocorrelation of the residuals was checked with Moran's Index. None of the developed regression model residuals display significant spatial autocorrelation according to (Global Moran's test). This demonstrates that the residuals are spatially random, suggesting that we were able to capture the underlying relationships that contributed to the spatial autocorrelation in the coffee quality samples.

For each model all significant biophysical factors contributing to the explaining the coffee sensory quality variability is listed (Table 5-4) with their coefficients (b-values), standardized beta's (β -values), coefficient of determination (R^2) and minimized error of the model (ϵ). β -value varies between 0 and 1, and provides a scale of the relative importance among biophysical factors combined within each developed sensory quality regression model. Overall the results of the multivariate analysis indicate that sensory quality is generally explained by the location topography (altitude and slope) where coffee beans are produced, local climate (average temperature and total annual rainfall), and soil properties (structure, acidity and nutrients). The study revealed that the soil nutrients (nitrogen, phosphorous and potassium) are important factors in explaining the difference in sensory quality of coffee. To a lesser extent are the contributions of soil structure (soil organic carbon and sand content of the topsoil) and the soil acidity.

The national model explained 21% of the overall sensory quality variability. The stratified the country dataset (503 samples) based on agroecological zones, produced regionalised models with improved explanatory power with R^2 ranging from 21% to 47-98%. The regionalised models show that the sensory quality of coffee in Impala (IMP), and Kivu lake shore (KLS) is positively influenced by a favourable climate ($\beta_{rain, t}=0.6$) and a reduced soil acidity ($\beta_{C/N, pH, Ca}=0.3-0.4$), and regionally affected by increased sand content and high potassium content ($\beta_{Sd, K}=-0.3-0.4$) at a lower elevation of Imbo zone (IMB). In the Congo Nile divide (CND) watershed, very high altitudes ($1,955m \pm 145m$) and soil quality are negatively associated with coffee sensory quality. Unfortunately, these two parameters are not easy to change, making a big part of Congo-Nile Divide zone less suitable for producing high coffee quality coffee. In the central and south plateau and eastern ridges, soil properties have a stronger positive contribution in determining quality of coffee than the contributions of climate and the topography. In the South and Central plateau (CP), soil structure (clay soils) are positively associated coffee beverage quality ($\beta_{Cy}=0.4$), however soil potassium has a negative contribution ($\beta_K=-0.2$). In Mayaga and Bugesera region, high temperature and carbon-nitrogen ratio ($\beta_{t,C/N}=0.1-0.2$) seem to have a slightly positive contribution to the quality of coffee produced, however other parameters were included in the model with equal and weaker relationships with coffee sensory quality.

In the eastern lowlands (EL), soil properties have a positive contribution to sensory quality together with topography than climate. Altitude ($\beta_{Alt}=0.24$) has a positive association with coffee quality variability produced in eastern lowlands. An increase of calcium ($\beta_{Ca}=0.5$) and clay content ($\beta_{Cy}=0.4$) have a positive effect on the coffee sensory quality produced in East. However high temperature ($\beta_t=-0.13$) and high soil cation exchange capacity (CEC) ($\beta_{CEC}=-0.7$) in this region have a negative association with coffee quality. In Buberuka highlands (BH) in the north of Rwanda, are almost not used for coffee growing. The 256 farms producing in Buberuka highlands indicated that the quality is affected by extreme climate (low temperature: $\beta_t=-0.5$), and high precipitation: $\beta_t=-0.4$ which is associated with an increased soil acidity and aluminium toxicity ($\beta_{pH, Al}=-0.3$) due to the process of nutrient leaching.

Table 5-4: Coffee sensory regression models: country model and regionalised models based on agro-ecological zones. Model coefficients (b-values) of biophysical factors and their respective importance (i.e. Beta: β -values) are presented between brackets. All factors are significant at 0.1% level ($P \leq 0.001$).

Biophysical factors	Country model	Agroecological zone based models								
		IMB	IMP	KLS	CND	CP	MPB	BH	ERP	EL
Constant	105.18	93.45	67.26	77.78	93.35	100.6	81.44	102.8	96.58	83.35
Altitude (Alt m)	-0.001 (-0.16)	-0.001 (-0.12)		0.002 (0.19)	-0.001 (-0.12)		0.001 (0.04)	0.001 (0.125)	-0.002 (-0.12)	0.004 (0.24)
Slope (S %)	-0.018 (-0.12)	-0.430 (-0.17)				-0.008 (-0.06)				
Total rainfall (rain mm)	-0.001 (-0.11)	0.003 (0.17)	0.002 (0.63)	0.005 (0.24)	0.003 (0.17)		0.001 (0.04)	-0.009 (-0.39)		
Avg Temperature (t °C)	-0.692 (-0.49)		0.412 (0.62)			-0.513 (-0.21)	0.277 (0.14)	-0.500 (-0.46)	-0.657 (-0.22)	-0.270 (-0.13)
Soil sand content (Sd %)	-0.035 (-0.20)	-0.139 (-0.34)		-0.080 (-0.48)	-0.139 (-0.34)	-0.022 (-0.10)	-0.012 (-0.03)		-0.031 (-0.21)	0.053 (0.28)
Soil clay content (Cy %)		-0.053 (-0.17)	0.081 (0.38)	-0.051 (-0.24)	-0.050 (-0.16)	0.079 (0.44)		0.026 (0.16)	0.038 (0.22)	0.180 (0.34)
pH KCl (pH)	-0.644 (-0.30)		0.245 (0.18)	0.942 (0.30)		-0.904 (-0.17)		-1.652 (-0.30)	1.587 (0.25)	
Soil CEC (cmol*kg ⁻¹)		0.072 (0.22)			0.072 (0.22)	-0.091 (-0.18)	-0.018 (-0.02)		-0.088 (-0.21)	-0.350 (-0.72)
Soil organic carbon (%)	-0.261 (-0.33)		-0.029 (-0.06)		-0.015 (-0.02)	-0.107 (-0.15)			0.133 (0.08)	
C/N ratio	0.293 (0.36)		0.304 (0.42)			0.112 (0.09)	0.938 (0.19)			
Soil available P (P ppm)	0.018 (0.15)	0.031 (0.14)		0.009 (0.07)	0.031 (0.14)		-0.016 (-0.02)		0.011 (0.14)	0.018 (0.13)
Soil K (cmol*kg ⁻¹)	1.811 (0.14)	-5.210 (-0.18)		-3.765 (-0.37)	-4.941 (-0.17)	-2.890 (-0.21)			3.081 (0.24)	
Soil Al (cmol*kg ⁻¹)		0.340 (0.17)	-0.054 (-0.28)	0.434 (0.34)	-0.437 (-0.17)	0.306 (0.15)	-0.056 (-0.03)	-0.351 (-0.29)		-2.389 (-0.25)
Soil Ca (cmol*kg ⁻¹)				0.253 (0.34)	0.303 (0.16)				-0.270 (-0.19)	0.729 (0.46)
Adj. R ²	0.21	0.56	0.51	0.47	0.60	0.52	0.98	0.81	0.27	0.62
Std. Error (ε)	1.4	0.32	0.42	0.88	1.2	0.91	0.12	0.51	1.1	1.0
N farms	14,286	598	611	3,199	1,982	3,974	856	256	2,454	331

Agroecological zones: KLS Kivu lake shore, IMP Impala, IMB Imbo, CND Congo-Nile watershed divide, MPB Mayaga plateau and Central Bugesera, CP Central Plateau, ERP Eastern Ridge and Plateau, EL Eastern Lowland, BH Buberuka Highlands.

The effect of the CoE sample size and the spatial distribution of coffee fields on the model performance were observed. For example, in Mayaga and Bugesera (MPB) region, limited coffee farms (n=856 coffee fields located in consolidated land) participated in the CoE competition (6 CWS participating in CoE). The MPB model looks very accurate ($R^2=0.98$) due to homogeneous samples unlike in Eastern ridge and plateau where coffee samples are from scattered farms (n=2,454 coffee farms) with 25 CWS scattered across ERP

zone. The quality model of Eastern ridge and plateau (ERP) explained up to 27% of the variation in coffee sensory quality (closer to the national model). ERP zone is highly heterogeneous in term of relief (altitude: 1,604 ±132 m) and associated soil and climate parameters making the sensory quality more heterogeneous. For other agro-ecological zones, resulting significant models explained coffee sensory variation ($R^2=0.47-0.81$) with medium to high explanatory power by significant sensory factors and display a low standard error of the model ($\epsilon=0.12-1.2\%$).

5.3.3 Geographic areas for special coffee production in Rwanda

Figure 5-4 shows potential areas of coffee sensory quality production as estimated by both national (a) and regionalised (b) sensory models. The national model (Figure 5-4a) suggests that the coffee sensory quality in Rwanda could be higher on average. It generally overestimated the potential quality offered by current natural environment especially in the east and the north is probably due to biased distribution of available CoE samples. Instead, the regional Agro-ecological models highlight a number of environmental constraints that negatively contribute to the sensory quality of coffee. Some constraints are manageable, others are not feasible or their management is not cost efficient. Figure 5-4b highlights geographic areas where specialty coffee (i.e. $Q_p > 85\%$) may occur according to the regional regression models based on local climate and soil properties. The regional models performed relatively well in predicting speciality coffee potentials up to 64% (or 55% assuming that the intercept is zero). We also calculated Moran's I for the model residuals to check for spatial autocorrelation. Moran's Index was 0.003 while Expected Index is -0.007 (Z-score=0.629, P-value=0.529) demonstrating that the residuals are spatially randomly distributed. The model thus satisfactorily predicts the coffee quality with good precision (RMSPE=0.98%) and without any remaining autocorrelation.

Furthermore, these models demonstrate that not all regions have equal potential for producing specialty coffee within the current smallholder farming system. According to these models, the south and central plateau (CP) remain areas of special coffees (>90%). However, in the Congo-Nile watershed divide (CND) some local hotspots with special coffee are highlighted. More hotspots of excellent coffees (87-90%) are also found in the north and south eastern ridges and plateau (ERP) while in lowland areas, very few hotspots produce fine coffees which are located in upper Muvumba catchment area in the Northern East of Rwanda. Although Buberuka highlands (BH) is not generally used for coffee, our analysis suggests that in this area could potentially produce more coffee of excellent quality.

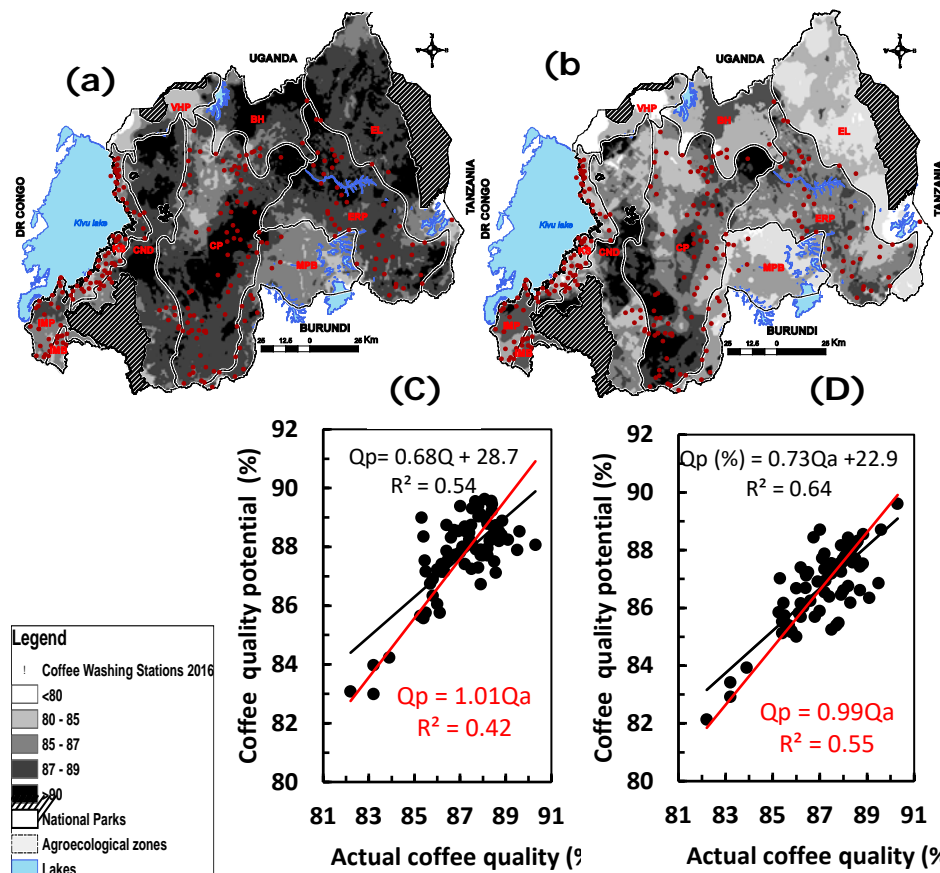


Figure 5-4: Coffee sensory quality potential (overall grade in percentage) as estimated by (a) the national model and (b) the agro-ecological based models. The scatter plot (c) presents the validation results of the national model (a), and the graph (d) presents the validation of the regional model (b).

5.4 Discussion

5.4.1 Effect of soil fertility management on coffee sensory

Bourbon coffee arabica, grown in the various agro-ecological zones of Rwanda, produces coffee with different tastes and flavours. Valued good quality coffees are found in Kivu lake shore (Nyamasheke, Rutsiro and Rubavu districts), and in Congo-Nile watershed divide (Karongi district), in south and central plateau (southern Province districts). Although, little is known about the direct influence of soil on the sensory quality, existing studies indicated that Arabica coffee planted in different soils produces coffee with different bean size and cup quality (Sanchez, Marques, Siqueira, Camargo, & Pereira, 2013). The types and texture of soils influence the root systems and soil-water holding due to high bulk density (Innocent Nzeyimana

et al., 2017). In Rwanda, soils with sand content less than 46.2%, soil acidity ($\text{pH} < 5.8$) and organic carbon content greater than 1.2% produce high sensory quality of coffee. This is in line with A. Youkhana and Idol (2009) who found that soils with less coarse and with clay content above 30% are good for improving coffee cherries quality.

Fertile soils produce larger beans often resulting in a higher final quality. Experiments have shown that a high levels of Nitrogen fertilizer applications increases total production but also reduces beans density (Wintgens, 2009), hence reduced overall coffee quality. However, our study shows that an increase of Carbon-Nitrogen ratio could improve the quality of coffee; adding nitrogen fertilizer without increasing soil organic carbon content would have a negative effect on sensory quality. In general can an excess of nitrogen increases the caffeine content, resulting to a more bitter taste of the brew which in not desired for Arabica coffee, lowering is sensory quality (Wintgens, 2009). These examples indicate that under certain conditions there is a trade-off between increase of yield and the sensory quality. We advocate to improve the soil organic carbon content and quality as demonstrated by Bucagu et al. (2013) and A. H. Youkhana and Idol (2016) as a sustainable way to enhance coffee quality. Except in the eastern ridges, more potassium on the exchange complex is negatively associated with the sensory quality of coffee. The coffees which won the CoE competitions came from the shores of Kivu Lake, Congo-Nile divide and central plateau where potassium content in soils is at minimum level ($\text{K} \leq 0.3 \text{ cmol}^+\text{kg}^{-1}$). There is moderate correlation and positive contribution of phosphorous content on the quality of coffee, similar with other studies (Wintgens, 2009). Cup quality responded to soil cation exchangeable calcium (Ca) which is inline of what has been found in Costa Rica by Castro-Tanzi et al. (2012). Calcium is the abundant element identifier of coffee beans and depth-profile testing has shown that Ca is homogenous inside the beans (Anggraeni, Nasution, & Suyanto, 2016). Therefore, improving calcium content in soils of Kivu lake shore, Congo-Nile divide, and central plateau, to around $4 \text{ cmol}^+\text{kg}^{-1}$ of soil could potentially improve the quality of coffee in these regions. However, it is important to note that doubling the concentration of calcium in soils can result in a bitter taste in the cup. This is the case of moderate quality of coffee produced by Impala and Imbo zones where soils are rich in calcium ($\text{Ca} > 8 \text{ cmol}^+\text{kg}^{-1}$). Improving calcium will not only improve the quality, but will potentially boost the yield through lowering of acidity and aluminium toxicity which has been a major constraint to productivity of acid soils of Rwanda (Cyamweshi et al., 2014).

5.4.2 Impact of climate change on coffee quality

Climatic variables are the most factors controlling the yield and quality of coffee (Aparecido, Rolim, Lamparelli, de Souza, & dos Santos, 2017). The

appropriateness of climate for the cultivation of coffee depends on the latitude and height above sea level (Wintgens, 2009). Among the climatic factors, the mean air temperature during seed development greatly influenced the sensory profile (Bertrand et al., 2012). In Rwanda, the sensory attributes such as fine acidity and flavour quality are found in coffee produced in the cool climate of the south plateau area (Songer, 2008). This is in line with what found by Bertrand et al. (2012) that sensory attributes such as acidity, fruity character and flavour quality of arabica coffee produced in Reunion Island were correlated with typical of coffees produced under cool climates. Rainfall is also correlated with bitterness while high evapotranspiration is reported to induce 'green' taste (Bertrand et al., 2012). However, extreme high rainfall influences the growth of the coffee trees and the development of pests. Frost at high altitudes can damage the plants and cherries, and affect the quality of coffee beans (Murugan, Shetty, Ravi, Anandhi, & Rajkumar, 2012). This is the case of Northern Buberuka highlands where the sensory quality model shows that an excessive increase of rainfall ($1,247 \pm 100\text{m}$) coupled with an excessive low temperature ($<18 \pm 0.9^\circ\text{C}$) cause a decrease of coffee quality. Amayaga and Bugesera region suffer from a prolonged period drought which causes diseases and insects induced damage to coffee cherries, resulting in lower quality beans. Besides defects, high temperature and solar radiations of Mayaga and Bugesera region could reduce aromatic quality of green coffee which would be detected as off-flavours even after roasting (Bertrand et al., 2012). In fact coffee washing stations of Bugesera and many of Mayaga zone (Ruhango and Nyanza districts) have not been able to participate in CoE and those who participated were not able to win ($Q < 85\%$).

An accelerated maturation in hot and humid environment result to a strong intensity, roasted flavour and bitter (Steen, Waehrens, Petersen, Munchow, & Bredie, 2017). This seems to be the case of Impala and Imbo in western Rwanda. On the other hand, has it been reported that beans produced at too high altitude (low temperature) have a silvery, greenish skin and low acidity (Bertrand et al., 2012) making Congo-Nile watershed divide and highlands definitely less suitable for producing high quality coffee. Bourbon Arabica coffee produced at altitude higher than 1400m of central plateau and eastern ridges display more acidity, aroma and flavours. This is the case of Muyongwe coffee that has been performing remarkably well and has won several prizes including Cup of Excellence awards at several occasions. Besides winning the CoE competitions, Muyongwe coffee won an appreciation at East Africa Fine Coffee Association "EAFCA" in 2009. The price winning tasting described aromatics that included nuts, spices, fruits, and flowers. It also had several descriptors applied to the mouth feel and the cleanest, sweetest, and longest finishes of the set of samples.

Coffee requires sufficient and well-distributed rainfall (Wintgens, 2009). In Rwanda, the dry months of June-August correspond to the summer. The flowering starts with the first rain of September and October. When this period is prolonged to September-October, the flowering is delayed; the rain season is shortened, thus causing losses of quality because of the development of empty beans due to lack of sufficient water. Furthermore the remaining beans are smaller due to physiological deficiency known as “die-back”(Wintgens, 2009). This phenomenon occurs in Amayaga and Bugesera region as well as in Eastern lowlands. In this case it is necessary to improve conditions by mulching coffee plants, adapting nutrients inputs, as well as applying suitable drainage or appropriate shade.

Rwandan climate is expected to change towards a slightly warmer and wetter climate (McSweeney & Semafara, 2011; Muhire & Ahmed, 2015). The mean temperature is projected to increase by 0.35°C per decade (McSweeney & Semafara, 2011), whereas the annual rainfall is expected to increase by 4 mm per year (Muhire & Ahmed, 2015). Using this projected increase in temperature and rainfall, we recalculated the map of sensory quality for the expected climate conditions in 2020 (Figure 5-5a) and 2050 (Figure 5-5b) using quality models earlier determined with model parameters and their coefficients (b-values) presented in Table 5-4. The underlying assumption is that empirical relationships of the multiple regressions are still valid from 2020 to 2015. Fig 4 shows that projected future increase in temperature and rainfall will cause a decrease in coffee quality up to 4.8% in some parts of the country. Most affected coffees areas are Huye Mountain coffees in south plateau, Gakenke: Muhondo and Muyongwe coffees in north plateau, Nyamasheke and Rutsiro coffees along Kivu lake shore and Rwimbogo coffees in Impala, in south west.

An increase in coffee quality up to 10.4% is however expected in other regions which are presently marginal in terms of coffee sensory quality. For example, Buberuka highlands will become more suitable (Burera district) as result of increase in temperature by 1.4°C by 2050. In eastern lowlands (Gatsibo, Kirehe and Ngoma districts), coffee quality potential will improve by 10.4% by 2050 due to expected increase of rainfall by 160 mm in 2050. The net overall balance in quality change is limited to about 0.3% in 2020 and 1.1% in 2050.

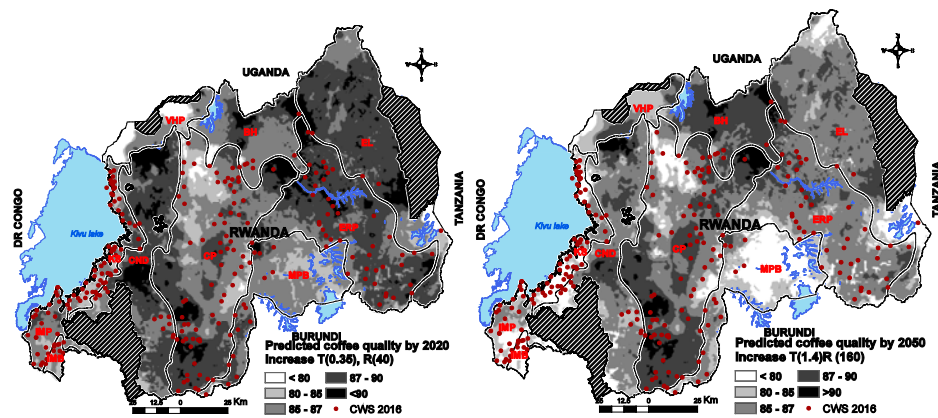


Figure 5-5: Maps showing the coffee quality potential as result of projected increase in temperature by 0.35°C and rainfall by 40 mm in 2020 and by 1.4 oC and by 160 mm in 2050 (McSweeney & Semafara, 2011; Muhire & Ahmed, 2015).

5.4.3 Coffee quality and production trade-off

The international market values speciality coffees with much higher prices (Alves et al., 2016). As illustrated in the Figure 5-6, using four consecutive annual Cup of Excellence (CoE) Competitions, it is clear that as soon as Rwanda started to participate in the CoE, the price of coffee increased. Consumers show an increasing interest in local and quality coffees than ordinary coffees at world coffee market. Since 2013, special coffees is commanding as higher as US\$20 per kilogram (minimum domestic price) at CoE auction in Rwanda while ordinary the commercial grade coffees are hardly sold at US\$5per kilogram of green coffee. This clearly demonstrates to coffee growers that though the production quantity is important, a higher quality can easier generate a much higher income.

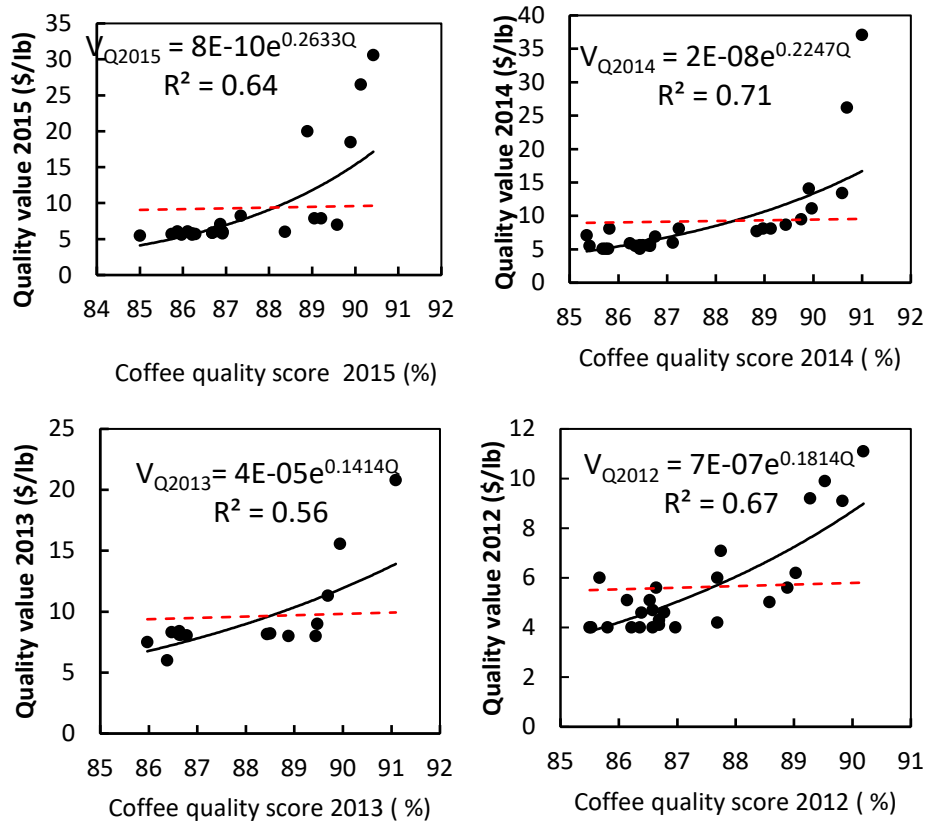


Figure 5-6: The value (US\$/lb) of specialty coffees during the annual Cup of Excellence (CoE) auction: the value presented is the high bid price (1lb=0.454kgs). Ordinary coffees (coffee with score below 85%) are not sold at CoE auction. Of the total 245 coffee washing stations (CWS) in 2015, CWS winners were only 24 CWS in CoE 2015, 28 CWS in CoE 2014, 15 CWS in CoE 2013, and 26 CWS in CoE 2012.

Usually, good practices to increase the productivity also result in producing good quality cherries. But we also saw earlier with N fertilizer and soil Ca properties a trade-off between quantity and quality management. Figure 5-7 shows that overall there is no meaningful correlation (from 2012-2015) between the quality of the coffee and the production levels. Farmers could gain a lot if they understand their local coffee environments and if they should invest in either higher coffee yields or in higher coffee quality. Ideally, if possible they could opt for both high yield and high quality to have maximum returns. Otherwise they have to make a smart trade-off between yield quantities versus coffee quality.

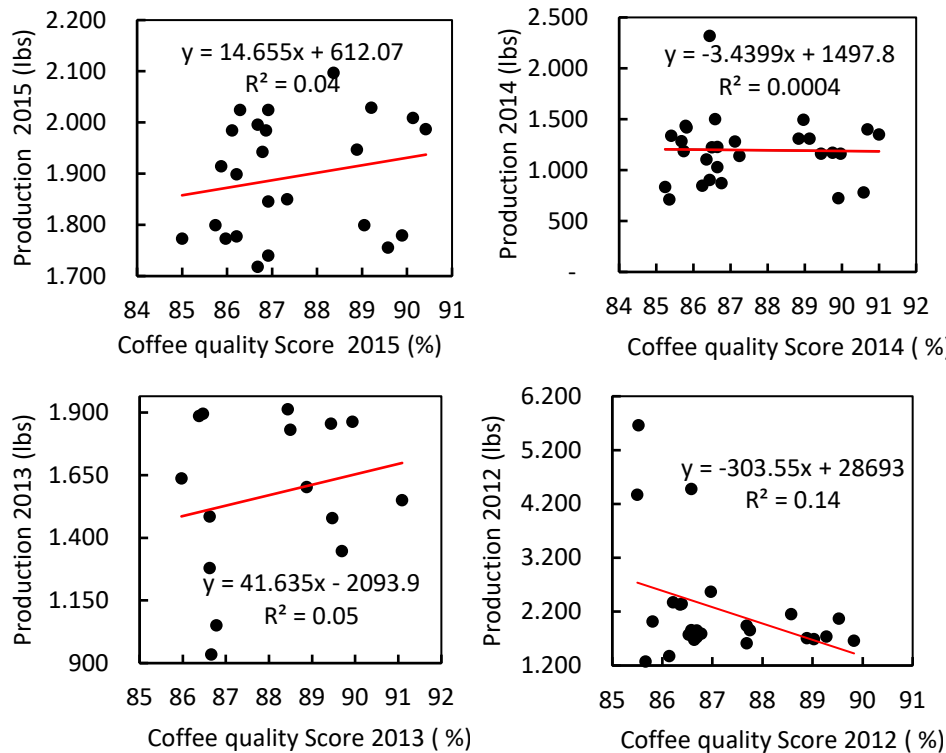


Figure 5-7: The total production of speciality coffee (coffee with score above 85%) in relation to their quality as measured at the Cup of Excellence (CoE) 2012-2015. Of the total 245 CWS in 2015, CWS winners were only 24 CWS for CoE 2015, 28 CWS for CoE 2014, 15 CWS for CoE 2013, and 26 CWS for CoE 2012.

5.5 Conclusions

This study demonstrated that the sensory quality of Arabica coffee is partially explained by the geographic origins where coffee was produced. Overall coffee quality grade is statistically correlated with the biophysical conditions of coffee origins using multiple regression analysis. Significant geographical factors and their co-interactions were used to predict geographic locations of specialty coffees.

The study showed that although the topography and climate of Rwanda play an important role in the spatial distribution of coffee quality in Rwanda, local soil conditions are often a predominant factor.

The study demonstrated also that overall coffee quality is not very sensitive to projected climate change (nationally only net 1% increase in 2050). There will be regional shifts of the potential specialty coffee locations.

There are multiple soil properties contributing to high sensory quality. Unfortunately, they are in general not easy to improve.

A smart soil organic matter management (both in terms of quality and C/N) could help to improve Rwandan coffee quality. Alternatively, one could explore new potential high quality producing coffee locations which will change coming decades.

6 Synthesis: Integrating land suitability, productivity and sensory quality in 'Coffee terroirs' for sustainable coffee production systems in Rwanda

This chapter is submitted as: Mukashema, A., and Veldkamp, A. Integrating land suitability, productivity and sensory quality in Coffee terroirs for sustainable coffee production systems in Rwanda. In *Land Use Policy*.

6.1 Introduction

Terroir is a French concept which finds its origin in viticulture. A terroir refers to a specific area where the interactions between local environment and viticulture practices leads to wines with distinctive characteristics (Costantini, Lorenzetti, & Malorgio, 2016). Gangjee (2017) defines “terroir” as a limited area or terrain whose specific soil and micro-climate impart distinctive qualities to food products. Most Geographical Indicators (GI) of terroirs which are linked to a specific origin and protection of a product are found in Europe and USA (e.g. S. Bowen, 2010; S. Bowen & Zapata, 2009; Vaudour, 2002), for mainly wine (e.g. Berard & Marchenay, 2006; Foroni et al., 2017; Rotaru et al., 2010) and olive oil (Belletti, Marescotti, Sanz-Cañada, et al., 2015). Some terroirs are found outside Europe, such as the example of Taquila cheese in Mexico (S. Bowen, 2010; S. Bowen & Zapata, 2009), and wine of specific British Columbia valleys (P. A. Bowen et al., 2005). A few terroir like examples are found in the third world (e.g. Schlesier et al., 2009; Smeyers-Verbeke et al., 2009) and all are for wine.

However, for coffee, no specific terroirs exist yet, although the origin and provenance of coffee beans becomes more important in the global coffee trade. Nowadays the origin is often linked to the country of origin but coffee buyers tend to focus on those areas that produce the highest quality coffee. These highest quality coffees are also predominantly linked to the origin of beans and to a lesser extent to the post harvesting process or roasting. When a green coffee bean is roasted, thousands of compounds are formed via chemical reactions. These molecules originate from precursors present in the green coffee bean, which are transformed into a large variety of volatile and non-volatile compounds during roasting (Folmer, 2014). This growing awareness that the location of coffee production matters makes coffee “terroirs” a logical next step.

The concept of “terroir” was first used in coffee studies to determine the effect of climate on beverage quality of coffee beans in large scale production of Brazil (da Silva et al., 2005), and Costa Rica (Avelino et al., 2005; Decazy et al., 2003; Vaast, Bertrand, Perriot, Guyot, & Genard, 2006). Both studies found that high altitude and annual rainfall under 1500 mm were favourable factors of coffee sensory quality (Decazy et al., 2003), east-facing slopes gave coffee beverages generally superior attributes (Avelino et al., 2005), and shade also positively affected coffee beans size and coffee beverage quality (Vaast et al., 2006). This was confirmed by Silva et al. (2016) who recently found that the position and altitude of the fields are the main variables that influence coffee quality. Sanchez et al. (2013) assessed the effect of soil properties on yield and quality of coffee and found that coffee grown on soils with high clay and available iron contents, and low pH and base saturation, provided the highest yield and best sensory quality. They

further concluded that soil–relief relationships contained in a digital elevation model (DEM) can be effectively used for the hybrid mapping of areas of varying suitability for coffee production and quality. However, none of these studies considered analysing a “terroir” in a holistic manner in order to discern collective environmental characteristics and co-interactions for distinctive terroirs as defined by Costantini et al. (2016) and Gangjee (2017).

Although many scholars explain the terroir concept (e.g. Barham, 2003; Belletti, Marescotti, Sanz-Cañada, et al., 2015; S. Bowen, 2010), very limited work has been done on systematic mapping of terroirs (e.g. Marciniak, Reynolds, & Brown, 2009; Reynolds, Senchuk, van der Reest, & de Savigny, 2007). Most of the terroir studies have found effective sensory-differentiation of wines from experimental plots of commercial wineries, but they often lack a clear spatial delineation (Vaudour, 2002). A recent attempt to map terroirs aimed to reconstruct the past biological condition of vines and their mortality (Vaudour, Leclercq, Gilliot, & Chaignon, 2017). The study found strong relationships between vine mortality, soil type, past land use, and long term soil management, particularly in terms of soil organic carbon. Reynolds et al. (2007) used GPS and GIS for terroir mapping and concluded that sandy soils reduce wine acidity. Geospatial technologies were also used to understand the relationship between wine terroir and local conditions in the vineyards (Marciniak, Reynolds, & Brown, 2013; Marciniak, Reynolds, Brown, Jollineau, & Kotsaki, 2017). They observed for their case study that soil-water content was correlated with vine water status which allowed them to establish spatial relationships between soil, yield and wine flavours. However, in these studies, topo-climatic spatial variability was not considered. One reason is that the relationship between land characteristics and product quality has been mostly studied at the detailed or semi-detailed scales like in vineyards (Vaudour, 2002; Vaudour et al., 2017). In such small scale case studies the topographic and climatic variability will be very limited.

This approach could potentially be applied to large scale commercial coffee farms but not on small-scale farming systems where coffee is mostly grown mixed with many other crops by smallholder farmers. Another challenge is that relating coffee yield and sensory quality to environmental conditions requires a large amount of coffee field scale, spatially explicit, data about location suitability, yield and sensory taste variables for the whole country. For Rwanda, such detailed data on coffee field locations became only available during this research after high resolution mapping (Chapter 2) (Mukashema et al., 2014), yield (Chapter 3) and sensory quality (Chapter 4) through annual cup of Excellence competitions between farmers, as well as national biophysical and socio-economic data at coffee location level (Chapter 3) (Mukashema et al., 2016) potentially allowing the delineation of coffee terroirs for the whole country territory.

Mapping possible coffee terroirs requires relating coffee yield and sensory data to its local production environment. Many spatial multi-criteria fuzzy-based land suitability models have been developed, mainly in the area of suitability analysis of land use (e.g. dos Santos et al., 2017; Taleai, Sharifi, Sliuzas, & Mesgari, 2007), infrastructure development and planning (e.g. Makropoulos & Butler, 2006), and human settlement (Chow & Sadler, 2010). Recently, in agriculture, fuzzy based spatial models have been used in irrigated agriculture planning (Y. Chen, Khan, & Paydar, 2010), crop suitability (e.g. Perveen, Nagasawa, Ahmed, Uddin, & Kimura, 2008), forestry (e.g. Bustillos-Herrera, Valdez-Lazalde, Aldrete, & Gonzalez-Guillen, 2007; dos Santos et al., 2017), natural resource management (Sporcic, Landekic, Lovric, Bogdan, & Segotic, 2010) and biodiversity conservation (e.g. Breceda, Arriaga, Bojorquez, & Rodriguez, 2005).

All of these multi-criteria methods use analytic hierarchy process (AHP) combined with linguistic ordered weighted average (OWA) approach and are based on three-steps algorithms (Makropoulos & Butler, 2006) described by Despic and Simonovic (2000). Step 1 consists of decomposing a complex fuzzy set (A) into less complex subsets (Bi). Step 2, the fuzzy sets Bi is constructed, representing the set $A = f(B_1, B_2, \dots, B_n)$; and Step 3, consists of aggregation of fuzzy subsets Bi into a fuzzy set A.

We propose to use a similar logical framework to generate coffee terroirs. The primary issue is how to combine geographical indicators (e.g. natural environment, social, and know-how factors etc. into a composite fuzzy set representing coffee terroir set.

It is the objective of this synthesis chapter to develop a methodology to delineate Rwandese coffee terroirs using a fuzzy logic based spatial multi-criteria evaluation of coffee production of smallholder farming systems of Rwanda. These terroirs should identify where high quality specialty coffee can be produced in relatively large quantities. For our purpose, we define a coffee terroir as a three-dimensional set of highly suitable land, producing high yields and high sensory quality. All terroir components are analyzed independently and used to reconstruct and map coffee terroirs. The generated coffee terroir zones could be used to identify, protect and further develop the specialty coffees geographical origins in Rwanda.

6.2 Identifying coffee terroir zones and geographical indicators of origin

6.2.1 Conceptual model of coffee terroir

Establishing relationships between smallholder coffee production and its local production conditions was already done in Chapter 2 to 5. All relevant data is already described in those chapters and will not be repeated here.

For consistency purposes, we have rasterized all data to 3 * 3 km grids. The Coffee Washing station supply units are the spatial units for which the yield data is collected. The average travel distance to the nearest coffee washing stations by the farmer is about 3 km. The lots used for the sensory coffee determination are usually also the aggregate of fields within a radius of 3 km. The accuracy of the biophysical data and coffee field locations is much more accurate but we use the coarser spatial resolution as support of the terroir identification and mapping.

Three main subsequent steps were used to build the coffee terroir model:

1. Coffee distribution (Chapter 2) and environmental data preparation and processing,
2. Identification of underlying factors of coffee location suitability (Chapter 3), yield level (Chapter 4) and sensory quality variation (Chapter 5),
3. Predicting suitability to coffee growing, achievable yields and potential quality (Chapter 3,4, and 5),
4. Zoning coffee terroirs through integration of coffee location suitability, achievable yield and quality potentials.

Figure 6-1 summarizes the hierarchical steps used to develop the coffee terroir model that allowed prediction of coffee terroirs and identification of geographical indicators of origin in Rwanda.

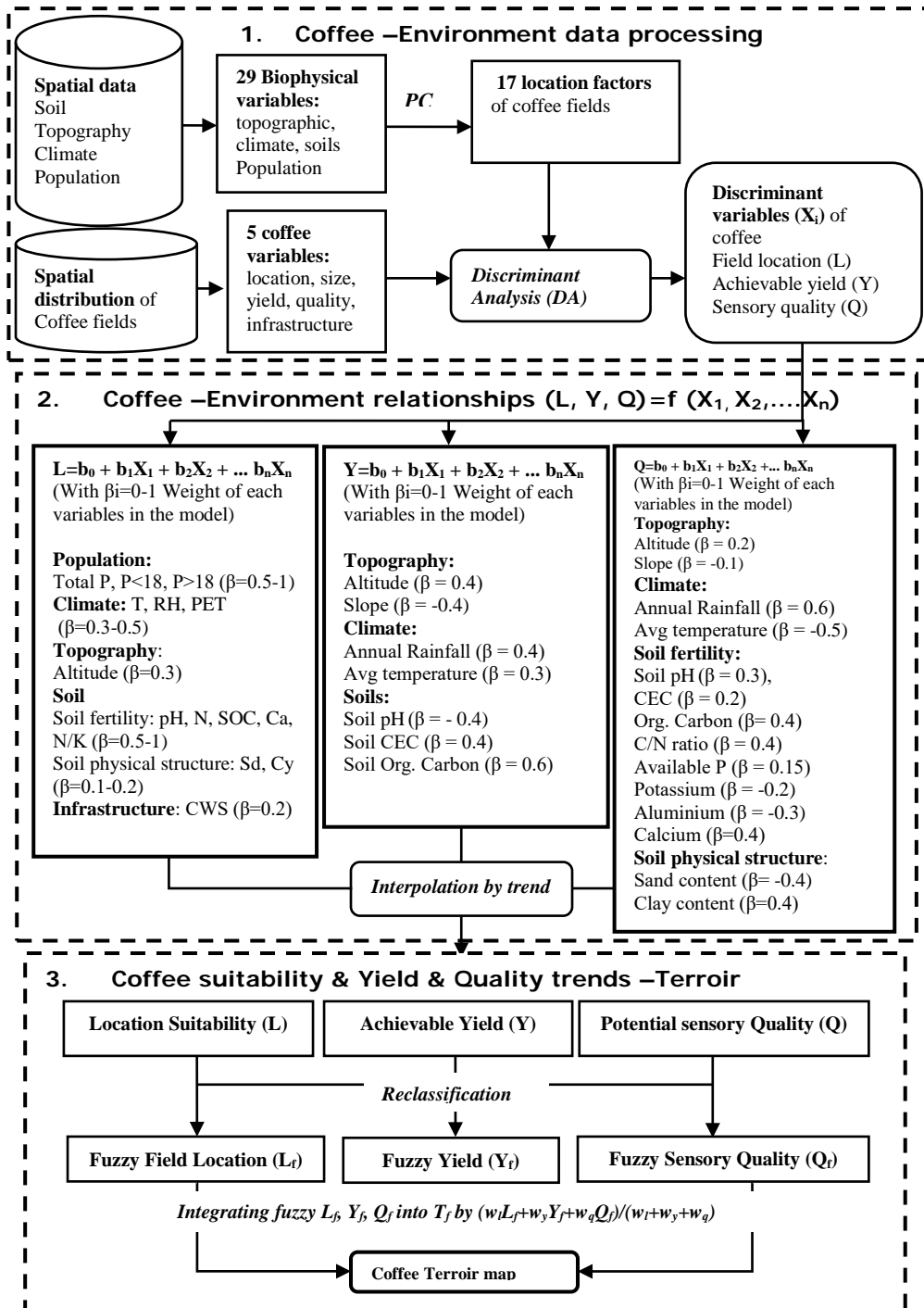


Figure 6-1: A methodological flowchart showing steps followed in developing Coffee Terroir Model

6.2.2 Spatial distribution of coffee fields, location and size of the farm

Relating coffee to its geographical origin (terroir) requires a reliable description of the origin as a physical environment (Vaudour, 2002). Hence mapping coffee fields was a primary step in the terroir model development. The coffee field map (Chapter 2) provided geographic coordinates (x,y) of coffee production locations in mixed smallholder farming systems, and allowed collecting the data of environmental factors at coffee field location for further development of the terroir model and terroir components. Figure 6-2 illustrates the different sizes and distributions of the coffee fields within the different agro-ecological zones for 3 km². It clearly makes the point that there is still a lot of local variability in the 3*3 km grid cells used to derive the terroirs.

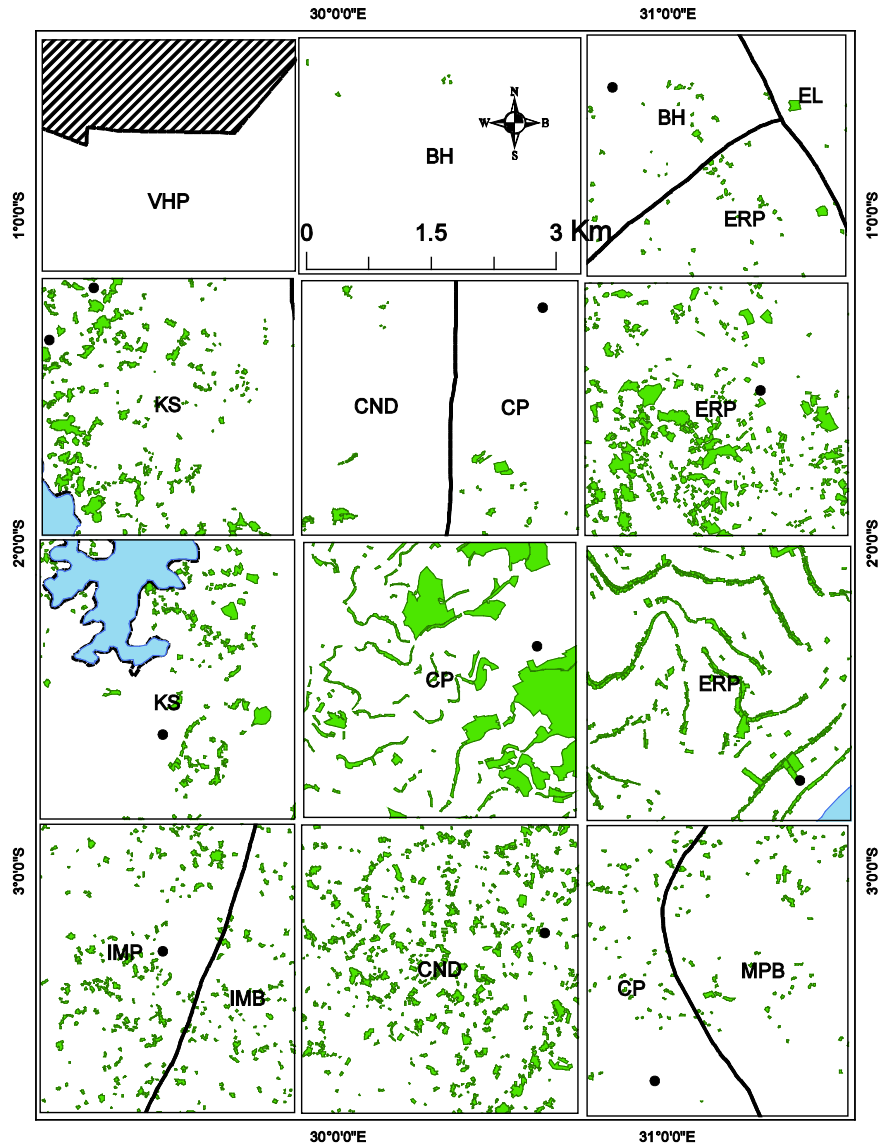


Figure 6-2 Coffee field distribution in agro-ecological zones of Rwanda; SW-NW: IMB Imbo, IMP Impala, KS Kivu lake shore, VHP Volcano high plain. Central S-N: CND Congo-Nile watershed divide, CP Central Plateau, BH Buberuka Highlands. SE-NE: MPB Mayaga plateau and Central Bugesera, ERP Eastern Ridge and Plateau, EL Eastern Lowlands. The black dots represent coffee washing stations. By 2016 coffee washing stations were operational across the country. For presentation purpose, we present for all agro-ecological zones an area of 3Km² only. The full map can only fit visually to A4 paper size at a scale of 1:1,000,000.

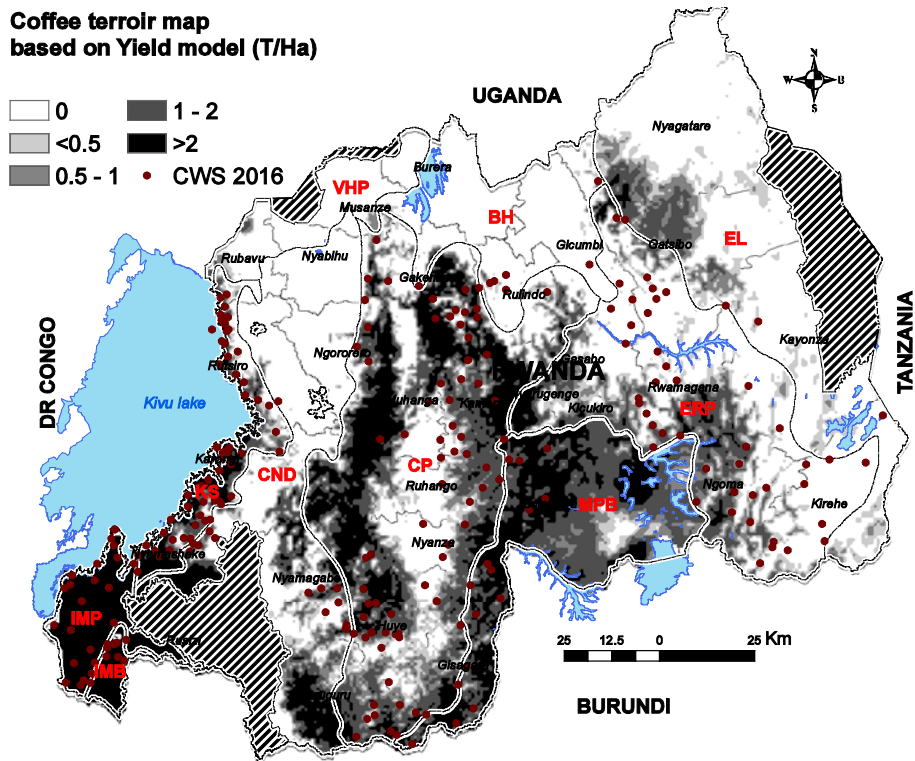


Figure 6-4 Coffee terroir areas predicted using relationship between actual yields and biophysical factors at coffee field location.

Figure 6-5 indicates the geographic origins of high quality coffees produced and new potential areas where there is likelihood that specialty coffees would be produced if any expansion takes place in that zone. Chapter 5 describes how the subset was derived.

through above linear relationships. One question for example is “does the coffee field C belongs to terroir T?” and the answer can be for example: yes if C plantation is located on soil with a pH greater than 4.5 or less than 6.5. Obviously, soil pH is not a fuzzy variable. What is fuzzy is the “suitability for location (L_{x,y}) where coffee is grown to belong to coffee terroir” from a pH point of view. Simple fuzzy linguistic is the location L_{x,y} to be suitable for terroir membership IF pH is between 4.5 and 6.5. This location L_{x,y} should produce yield (Y_{x,y}) greater than 1.5 tons per hectare and overall sensory grade (Q_{x,y}) above 87%. The main assumption here is that either terroir set (T_{x,y}) or coffee field location suitability (L), yield (Y) and quality (Q) subsets are fuzzy sets or are reclassified into new fuzzy sets and vary from one location to another. This means fuzzy members are utilized to transform the variables in which the criteria are measured into a common [1-5] scale following fuzzy membership as proposed by Makropoulos and Butler (2006). The terroir is indicated by a 5 when the actual value of L,Y,Q subsets represent the best area for producing special coffees and a 1 for the worst areas when the actual value of the subsets L,Y,Q produce ordinary coffee or are not productive at all.

The spatial reclassification function of ArcGIS 10.4 software by Jenks optimization method was applied to the matrix L,Y,Q fuzzy layers for determining the terroir areas, to represent excellent (5), good (4), regular (3), bad (2), and inappropriate (1) terroir. The Jenks optimization method, also called the Jenks natural breaks classification method, is a data clustering method designed to determine the best arrangement of values into different classes. This is done by seeking to minimize each class's average deviation from the class mean, while maximizing each class's deviation from the means of the other groups (dos Santos et al., 2017).

6.2.5 Weighting the land suitability, yield and quality factors in the terroir

The individual terroir subsets link individual environmental criteria to a final terroir index model through a terroir weighted additive function

$T = w_L L + w_Y Y + w_Q Q$, giving the quality four times more importance than other subsets $w_L < w_Y < w_Q$.

The importance of the quality subset in the terroir definition is proven by the continuous increase of coffee value (\$/lb) as a result of increased quality from ordinary coffee to specialty coffee production as shown by Figure 6-6. It shows that when the farmer aims to produce ordinary coffee with quality below 87%, a 2% grade increase from 85% corresponds to an increase of 2.5 US\$/lb (1:1 V/Q ratio). Any effort of the farmer to improve quality one unit grade above 87%, will lead to a quadruple gain in value to 4 US\$/lb (4:1V/Q

ratio). Therefore, we considered this general relationship between the value and quality of the produce in determining the quality and yield (Q/Y) ratio for terroir zoning both depending on land quality (L) as basis for terroir modelling. By using the global trade value the developed terroir is dependent on the global coffee market.

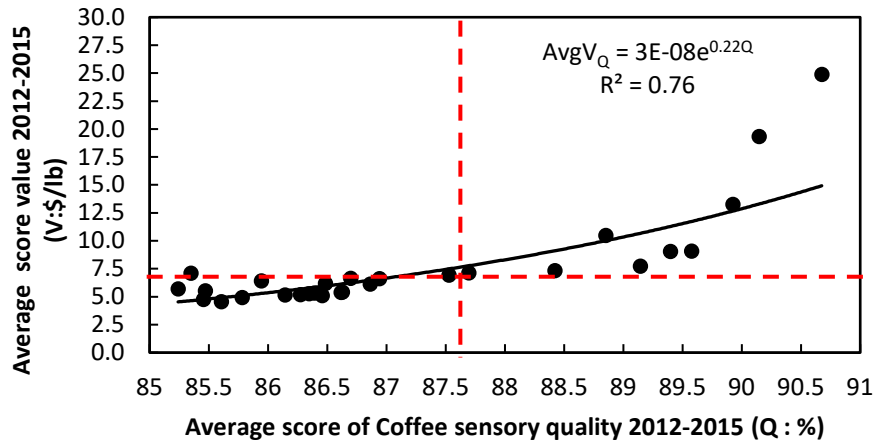


Figure 6-6 Average value (US\$/lb) of speciality coffee lots sold at the annual Cup of Excellence (CoE) auction 2012-2015. The value presented is the average high bid price (1lb=0.454kgs). Ordinary coffees (coffee with score below 85%) are not sold at CoE auction. In this trend analysis, we considered 93 lots CoE winners of 421 total sample lots submitted in the 2012-2015 CoE competition

In the modelling of best coffee terroirs, L,YQ matrix fuzzy subsets are combined by GIS-based overlay analysis to calculate the fuzzy terroir index shown in Figure 6-7.

high quality and high yields. Based on this analysis, therefore, possible land reallocation can be suggested to contribute to better land use planning.

Figure 6-8a shows that land suitability satisfies the terroir criteria at 26% of which 19% fall in the range between 50-75% and 7% in range of 75-100% suitable. About 34% of total country land (i.e. 831,087 hectares) satisfies the terroir condition from the yield point of view (Figure 6-8b) of which 16% would have yield ranges of 1.5 - 2 t ha⁻¹ and 19% with yields greater than 2t ha⁻¹. Figure 6-8c shows that the terroir set from the sensory quality criterion is limited (24%), but due to its high weight (4-times) in the terroir model, the final terroir feature is more strongly related to quality than to yield and location suitability (see Figure 6-5 & 7). About 20% of the country land (i.e. 432,648 ha) are classified as terroir to produce excellent coffee with grade between 87 and 90%, only 4% land is classified as a terroir of special coffee with grade above 90%. The overall terroir (Figure 6-8d) output shows that only 9% of the country land has the potential to consistently produce very high yield, and special coffee, while 18% of the country land can be used as terroir at the next level. Figure 6-9 shows that terroir hotspots with special coffee occur for the majority of pixels in The Central Plateau, followed by Impala, Amayaga and Buberuka Highlands zones. At the edge of Congo-Nile agro-ecological zone and eastern ridges and plateau also shows areas with excellent coffees. These two zones can be considered for the process of appellation and protection of geographic indications of origin. However, the model finds very few areas of high quality coffee along the Kivu lake shoreline and Eastern Lowlands. It is important to note that this analysis did not exclude land already occupied by other land uses. To date about 50% (i.e. 13,131 hectares) of the country coffee areas belongs to the prospected terroirs, and 56 of 66 terroirs units already have coffee. Only 10 new terroir units are projected by the terroir model as best candidate areas for the future coffee expansion.

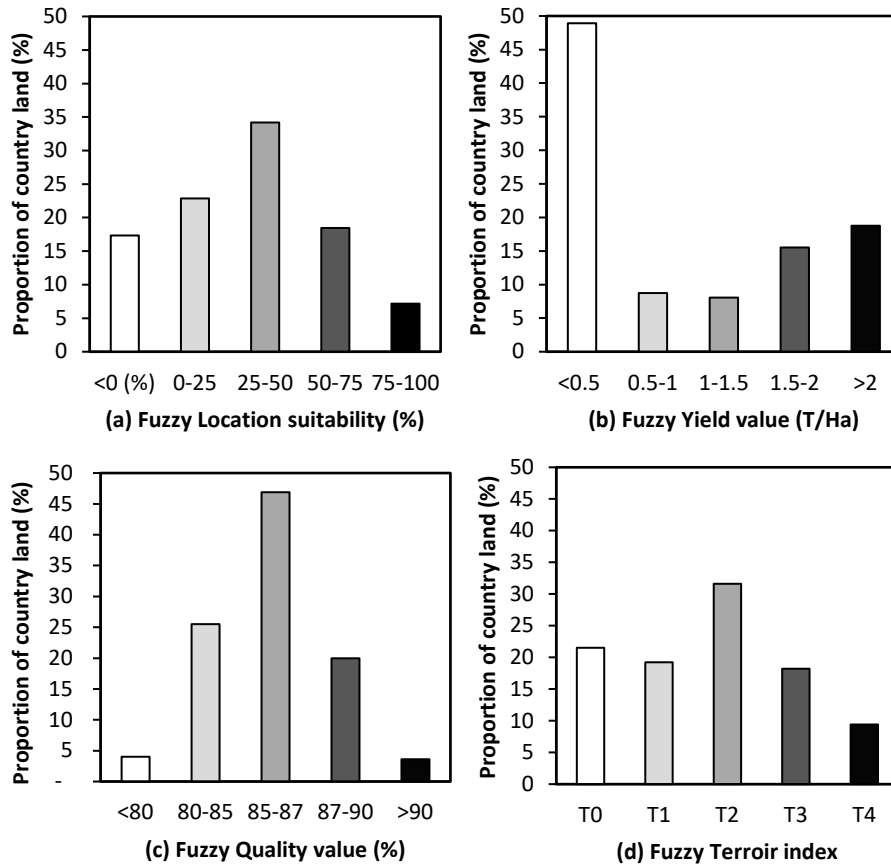


Figure 6-8 Land area in percentage of the total country land (i.e. 2.3 million hectares) for the range of fuzzy terroir subsets: Coffee field location suitability (a), Yield levels (b), Sensory quality (c), and overall terroir index (d). The last two classes are considered to contribute to terroir zones.

6.2.6 Geographical indicators of origin

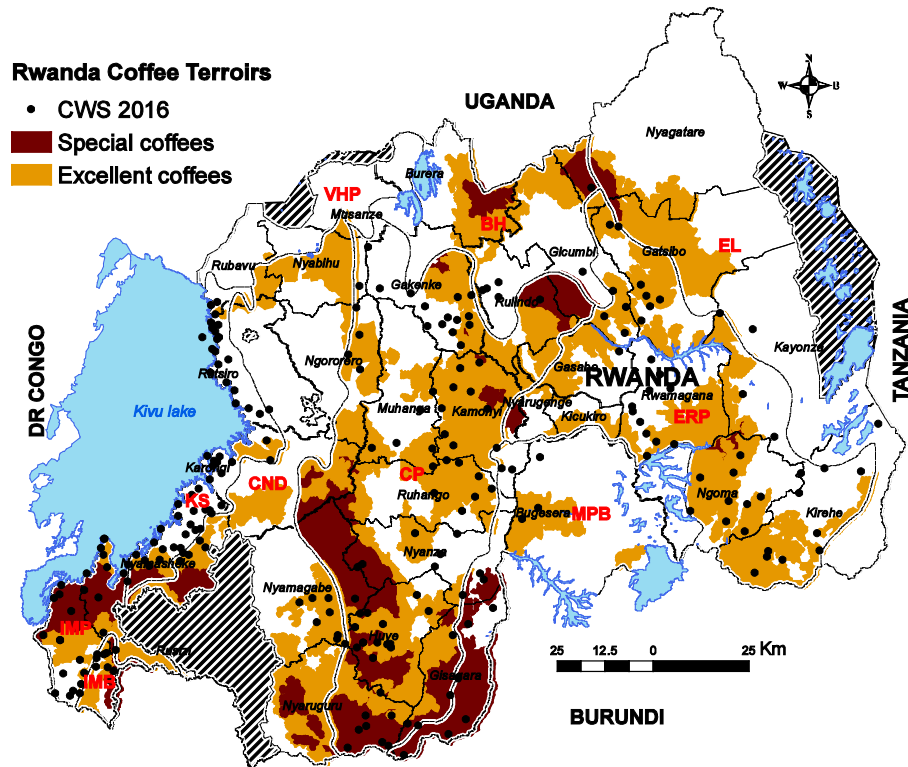


Figure 6-9 Coffee terroirs with only geographic areas that are expected to produce special coffee (Sensory Grade \geq 90%) and excellent coffees (Sensory Grade = 87% - 90%).

Table 6-1: Geographic indicators of origins and thresholds for 56 identified coffee terroirs

Biophysical Factors	Biophysical indicators	Min	Max	Mean	Std. Dev.
Climate	Altitude (Alt m)	1,400	2,149	1,676	181
	Slope (degrees)	6	24	14	4
	Aspect (degrees)	90	226	161	27
	Annual rainfall (mm)	926	1,975	1,276	203
	Avg temperature (oC)	18	21	19	2
Soils	pH _{KCl}	3.0	5.3	4.3	0.5
	PH _{H2O}	3.7	5.8	5.0	0.3
	Cation exchange capacity (CEC cmol ⁺ kg ⁻¹)	5.1	36.5	14.7	5.5
	Avail. Phosphorous (P ppm)	0.1	29.1	6.4	6.8
	Sand content (S _d %)	19.7	58.4	43.0	9.2
	Clay content (C _y %)	25.6	64.8	37.8	7.7
	C/N ration	10.6	14.8	13.5	0.8
	Organic carbon (SOC %)	1.2	8.9	3.7	1.6
	Calcium (Ca cmol ⁺ kg ⁻¹)	1.8	10.3	3.5	1.4
	Magnesium (cmol ⁺ kg ⁻¹)	0.7	6.5	1.5	0.8
	Potassium (K cmol ⁺ kg ⁻¹)	0.2	0.6	0.3	0.1
Aluminium (Al cmol ⁺ kg ⁻¹)	0.3	4.6	1.7	1.0	

6.2.7 Coffee terroir dynamics in relation to climate change

The Rwandan climate is projected to change towards a warmer and wetter climate (McSweeney & Semafara, 2011). An increase of 0.35°C per decade has been registered over the past 40 years (McSweeney & Semafara, 2011). An increase of rainfall between 1.5 and 6.8 mm per year (4 mm yr⁻¹ in average) was also noted for the period from 1931-2010 (Muhire & Ahmed, 2015). Using these projections, we recalculated coffee terroirs under a future temperature and rainfall change scenario (holding constant other matrix variables of terroir model). Figure 6-10 presents the current coffee terroir layer (2015) and the future terroir in 2050. Superimposed, the two layers show that extensive areas of current coffee terroirs in Bugesera, and eastern ridges and plateau in Ngoma, and Rwamagana districts will become ordinary to marginal due to projected climate change (high change or disappearance). The greatest victim seems to be Bugesera District, for which the terroir model predicts that it will become climatically unsuitable for both coffee yield and quality production by 2050. Even along the Kivu lake shoreline (KLS) where predictions show that the total area will not change significantly, there will be shifts in the terroir belt towards upper Kivu Lake Catchment and the Congo-Nile watershed (CND). Buberuka highlands (BH) will experience a possible terroir gain due to an increase in temperature by 2050. The changes

in climatic variables will affect coffee terroirs which means that it will become more difficult and expensive to achieve high quality and high yields (Chemura et al., 2016) in coffee terroirs predicted to become marginal, while it will become easier for those future terroirs that will become climatically favoured. It is important that farmers in current coffee terroirs that will turn marginal revise their strategies in the face of climate change. This will include adopting new coffee varieties and agro-economic models that are suited to climate change (Chemura et al., 2016) with greater tolerance of high temperatures and altered pest and disease pressures (Schroth et al., 2009). The Government of Rwanda should consider renewing the national policy that promotes weather-based insurance (Akter, Krupnik, Rossi, & Khanam, 2016; Rao, Gopinath, Prasad, Prasannakumar, & Singh, 2016). This does not apply to the Rwandan coffee sector alone as other studies also pointed that coffee areas will reduce or will undergo transformative adaptation due to climate change (Chemura et al., 2016; Davis, Gole, Baena, & Moat, 2012; Gay, Estrada, Conde, Eakin, & Villers, 2006; Jayathilaka, Soni, Perret, Jayasuriya, & Salokhe, 2012; Laderach et al., 2017).

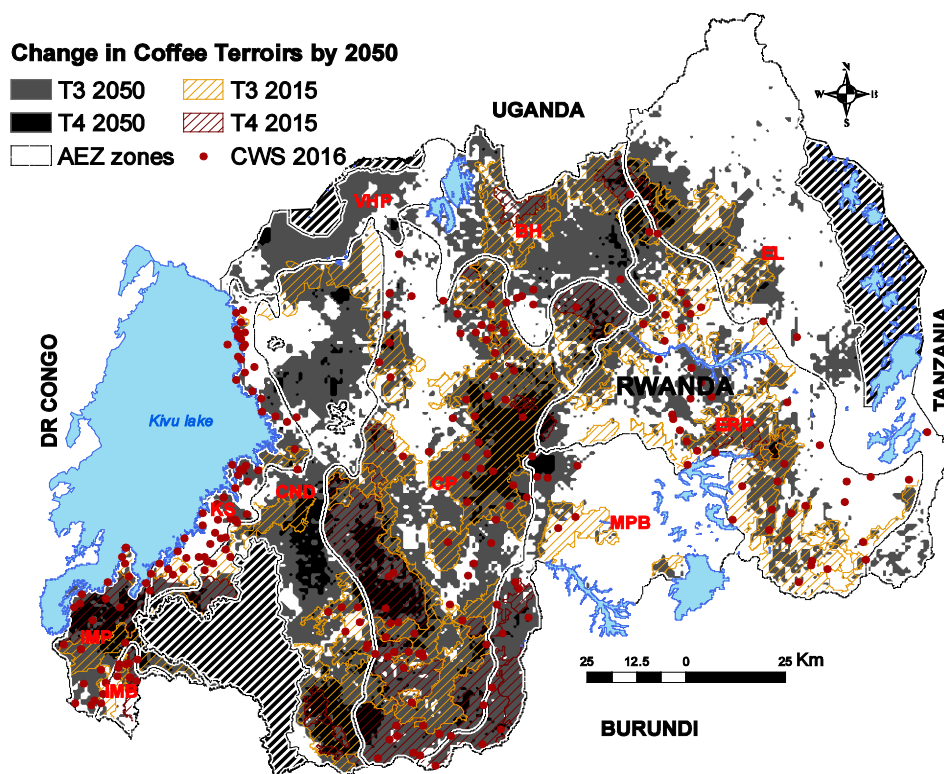


Figure 6-10 Projected coffee terroirs shifts in relation to climate change by 2050 as a result of increased average temperature by 1.4°C and total annual rainfall by 160mm. The shift in coffee terroirs is generally negative and East-West and South-North oriented causing a reduction of special coffee grade ($\geq 90\%$) to excellent coffees (87% - 90%), and excellent coffees to fine coffee (80-85%).

6.2.8 Effect of scale on terroir mapping and uncertainty

While spatially explicit quantification of coffee yields and quality allows the identification of areas of coffee terroirs for specialty coffee production, the accuracy and usefulness of the terroir information varies considerably depending on the spatial scale and methods used. Terroir analysis required to draw together data from numerous sources. The manipulation of data from the multiple disciplines, scales and locations raises the question of the choice of the appropriate spatial level of analysis, or 'scale'. Such a question is particularly relevant for terroir (Vaudour, 2002). As previously explained we used the spatial resolution of the coarsest data source (3 km * 3 km) to identify terroirs. This is automatically the smallest terroir possible in the used methodology. We also illustrated in Figure 6-2 that this resolution is too coarse to capture the coffee field diversity within Rwanda.

When we consider the uncertainties of the calculated terroirs we have to take into account that the source factors have all been quantified at a similar national extent with varying spatial resolution and for the terroir mapping all aggregated to 3*3km². The underlying multiple regression models have also been (quasi) validated yielding model fits of 60% for location suitability, 60% for coffee yield and 64% for coffee quality. All three were independently derived and demonstrate that 60% accuracy at 3 km * 3 km grids seems realistic given the data sources used. This still leave a large margin of uncertainty but the reconstructed relationships point us towards the locations where Rwandese coffee terroirs are most likely to be found. Since terroirs as such cannot be measured in the field we do not propose a sampling scheme to come to independent confirmation of our results. We therefore suggest that the produced terroirs maps, especially the extrapolated maps for 2050, should be used with caution and treated as locations of higher probability of high yields and sensory quality and its possible directions of change in the future.

6.3 Conclusions and outlook

While determining land suitability for speciality coffee production in Rwanda, we applied the concept "terroir" usually known in viticulture to represent an eco-geo-pedological unit that can be distinguished as a special or excellent coffee-producing entity.

The most commonly used criteria are climatic, geological, topographical and pedological land properties. Yet the appropriate resolution for the spatial modelling of terroir is not clearly established so far in previous studies, moreover a conceptual methodology for terroir analysis and mapping is not straight forward. Thresholds and rules are involved and the methods should be crop specific. The present study demonstrated a stepwise approach to determine potential coffee terroirs in small-scale farming systems. The

modelling approach proposed in this study is flexible and can incorporate additional variables whenever needed and dynamic enough to accommodate the ever-changing environment. While it is too early to conclude about the reality of the delineated terroirs, we suggest there is a need for more independent confirmations such as, for example, biochemical fingerprinting of coffee produced in these specific terroirs. A simple first confirmation of the derived terroirs is the observation that all permanent winners of Cup of Excellence competitions (2008-2015) are located within the delineated terroirs for special coffees. But we should not forget that the current underlying models only have an accuracy of 60%. We consider the proposed methodology to be usable in other small-scale coffee producing country, or adaptable to other high earning agricultural commodity crops in Rwanda and elsewhere. Specifically, cacao seems to be a very similar crop as coffee where large gains can be made by applying the terroir concept. The coffee terroirs could be directly applied in a similar fashion in Ethiopia and East Africa.

7 Future directions

In this study, we developed a stepwise probabilistic approach for coffee mapping using an expert Bayesian network model applied to national orthophotos dataset. The method provided an accurate representation of the spatial distribution of coffee plantations of Rwanda at an overall accuracy of 87% despite the spectral diversity of coffee trees and the field sizes. We demonstrated that spectral-based image classification is not sufficient to accurately map small-scale agricultural fields, but when combined with object-based classification yields higher accuracy. We anticipated that the method has potential for mapping other perennial small scale cropping systems in the East African Highlands and elsewhere. This line of research consists of a probabilistic approach to mapping mixed and small-scales land use systems. Such information is a prerequisite for new land use policies and management in countries where land is becoming a scarcer resource every year. Such data are also useful for land suitability assessment for agricultural value chain improvement.

Evaluating crop suitability is usually based on traditional land approaches in many countries, using only agroecological zoning and soil data. However, land use is also influenced by socio-economic and other biophysical factors. This study demonstrated how coffee location suitability is influenced by socio-economic factors as well as biophysical factors, and determined the strength of each factor and relationships between factors in locating some coffee in sub-optimal growing areas. This study clearly demonstrated that coffee distribution is part of a complex land use system where not only many different factors play a significant role, but where locally different trade-off decisions are made. We demonstrated that coffee cropping systems are not only environmentally driven as often assumed. Sometimes human factors overtake biophysical land suitability factors, which implies that land evaluation should consider human decisions as an equally strong factor as physical and bio-climatic factors in shaping present and future land use systems and related productivity. Future work could explore linkages between natural and human systems in increasing resilience to coffee stressors such as increased risk of extreme weather events, and decreasing coffee suitability at lower altitudes. The study could be scaled up to a regional level to explore the possibility of shifting coffee cultivation following the change in environmental niches. Such a study could use Agent Based Modelling to simulate human decision making at different levels. Farmers make decisions based on local conditions, while the National coffee board makes policy decisions based on national plans and ambitions. It would be relevant to understand how these different levels interact and lead to new patterns of coffee production. This is specifically important given the imminent climate changes.

At a more detailed level would it be interesting to investigate how coffee competes with other crops within a smallholder environment for farmer's attention and investments. A household has only a limited amount of capital and labour available that needs to be divided among all farm activities. It is very well possible that in specific farm/household configurations coffee management is far from optimal due to competing claims on available farmers' resources. Such a research could combine farmer interviews with Agent Based Modelling to better understand decision making and emerging patterns at farm level.

Studying crop suitability requires further analysis of local yields. Most unsuitable/marginal classified land produces low yields while high suitable lands are confirmed by production of high yields. However, this study revealed that 60% of current coffee fields are located in appropriate socio-economic and environmental locations. We also found that coffee yield depends on a range of environmental conditions (topography, climate and soils) in which coffee is grown and this is not correlated to the site suitability. More research is required to better understand how site suitability and coffee yield and coffee quality relate. We saw already in Chapter 5 that for some management options there is a trade-off between high yields and high quality. Only in some cases are both are possible. We identified such win-win locations and delineated them as terroirs for sustainable coffee land use.

Finally, a more quantified study of all uncertainties involved could be executed. A systematic Monte Carlo sensitivity analysis using the known data variability could give us more realistic estimates of the uncertainties involving the coffee terroir map. Some effects of uncertainty propagation might also play a role in our methodology of terroir delineation. A good example on how this could be done is well explained in Nol, Heuvelink, Veldkamp, de Vries, and Kros (2010). A proper execution of such an uncertainty propagation analysis goes beyond the scope of this thesis.

Bibliography

- Affholder, F., Poeydebat, C., Corbeels, M., Scopel, E., & Tiftonell, P. (2013). The yield gap of major food crops in family agriculture in the tropics: Assessment and analysis through field surveys and modelling. *Field Crops Research*, *143*, 106-118. doi: 10.1016/j.fcr.2012.10.021
- Aguilera, P. A., Fernández, A., Fernández, R., Rumí, R., & Salmerón, A. (2011). Bayesian networks in environmental modelling. *Environmental Modelling & Software*, *26*(12), 1376-1388.
- Aguilera, P. A., Fernández, A., Reche, F., & Rumí, R. (2010). Hybrid Bayesian network classifiers: Application to species distribution models. *Environmental Modelling & Software*, *25*(12), 1630-1639.
- Akter, S., Krupnik, T. J., Rossi, F., & Khanam, F. (2016). The influence of gender and product design on farmers' preferences for weather-indexed crop insurance. *Global Environmental Change-Human and Policy Dimensions*, *38*, 217-229. doi: 10.1016/j.gloenvcha.2016.03.010
- Alves, H. M. R., Vieira, T. G. C., Volpato, M. M. L., Lacerda, M. P. C., & Borem, F. M. (2016). Geotechnologies for the characterization of specialty coffee environments of Mantiqueira de Minas in Brazil. In L. Halounova, V. Safar, P. L. N. Raju, L. Planka, V. Zdimal, T. S. Kumar, F. S. Faruque, Y. Kerr, S. M. Ramasamy, J. Comiso, Y. A. Hussin, P. S. Thenkabail, S. Lavender, A. Skidmore, P. Yue, P. Patias, O. Altan & Q. Weng (Eds.), *Xxiii Isprs Congress, Commission VIII* (Vol. 41, pp. 797-799).
- Anggraeni, K., Nasution, A., & Suyanto, H. (2016). Recognition of Spectral Identifier from Green Coffee Beans of Arabica and Robusta Varieties Using Laser-Induced Breakdown Spectroscopy. In A. M. Hatta & A. Nasution (Eds.), *Second International Seminar on Photonics, Optics, and Its Applications* (Vol. 10150).
- Antoine, J. M. R., Hoo Fung, L. A., & Grant, C. N. (2016). Geographic determination of the growing origins of Jamaican and international coffee using instrumental neutron activation analysis and other methods. *Journal of Radioanalytical and Nuclear Chemistry*, *309*(2), 525-534. doi: 10.1007/s10967-015-4666-4
- Aparecido, L. E. D., Rolim, G. D., Lamparelli, R. A. C., de Souza, P. S., & dos Santos, E. R. (2017). Agrometeorological Models for Forecasting Coffee Yield. *Agronomy Journal*, *109*(1), 249-258. doi: 10.2134/agronj2016.03.0166
- Austin, E. J., Willock, J., Deary, I. J., Gibson, G. J., Dent, J. B., Edwards-Jones, G., . . . Sutherland, A. (1998). Empirical models of farmer behaviour using psychological, social and economic variables. Part I: linear modelling. *Agricultural Systems*, *58*(2), 203-224.

- Avelino, J., Barboza, B., Araya, J. C., Fonseca, C., Davrieux, F., Guyot, B., & Cilas, C. (2005). Effects of slope exposure, altitude and yield on coffee quality in two altitude "terroirs" of Costa Rica, Orosi and Santa María de Dota. *Journal of the Science of Food and Agriculture*, 85(11), 1869-1876.
- Barham, E. (2003). Translating terroir: the global challenge of French AOC labeling. *Journal of Rural Studies*, 19(1), 127-138.
- Belletti, G., Marescotti, A., Sanz-Cañada, J., & Vakoufaris, H. (2015). Linking protection of geographical indications to the environment: Evidence from the European Union olive-oil sector. *Land Use Policy*, 48, 94-106. doi: <http://doi.org/10.1016/j.landusepol.2015.05.003>
- Belletti, G., Marescotti, A., & Touzard, J.-M. (2015). Geographical Indications, Public Goods, and Sustainable Development: The Roles of Actors' Strategies and Public Policies. *World Development*. doi: 10.1016/j.worlddev.2015.05.004
- Berard, L., & Marchenay, P. (2006). Local products and geographical indications: taking account of local knowledge and biodiversity. *International Social Science Journal*, 58(187), 109-+.
- Bertrand, B., Boulanger, R., Dussert, S., Ribeyre, F., Berthiot, L., Descroix, F., & Joet, T. (2012). Climatic factors directly impact the volatile organic compound fingerprint in green Arabica coffee bean as well as coffee beverage quality. *Food Chemistry*, 135(4), 2575-2583. doi: 10.1016/j.foodchem.2012.06.060
- Bosselmann, A. S., Dons, K., Oberthur, T., Olsen, C. S., Raebild, A., & Usma, H. (2009). The influence of shade trees on coffee quality in small holder coffee agroforestry systems in Southern Colombia. *Agriculture Ecosystems & Environment*, 129(1-3), 253-260. doi: 10.1016/j.agee.2008.09.004
- Boudreaux, K., & Ahluwalia, P. (2009). Cautiously optimistic: economic liberalization and reconciliation. *MACRO_BOUDREAUX*, 54.
- Bowen, P. A., Bogdanoff, C. P., Estergaard, B. F., Marsh, S. G., Usher, K. B., Snaith, C. A. S., & Frank, G. (2005). Geology and wine 10: Use of geographic information system technology to assess viticulture performance in the Okanagan and Similkameen valleys, British Columbia. *Geoscience Canada*, 32(4), 161-176.
- Bowen, S. (2010). Embedding Local Places in Global Spaces: Geographical Indications as a Territorial Development Strategy. *Rural Sociology*, 75(2), 209-243. doi: 10.1111/j.1549-0831.2009.00007.x
- Bowen, S., & Zapata, A. V. (2009). Geographical indications, terroir, and socioeconomic and ecological sustainability: The case of tequila. *Journal of Rural Studies*, 25(1), 108-119. doi: 10.1016/j.jrurstud.2008.07.003
- Brandtberg, T. (2007). Classifying individual tree species under leaf-off and leaf-on conditions using airborne lidar. *Isprs Journal of*

- Photogrammetry and Remote Sensing*, 61(5), 325-340. doi: 10.1016/j.isprsjprs.2006.10.006
- Brandtberg, T., & Walter, F. (1998). Automated delineation of individual tree crowns in high spatial resolution aerial images by multiple-scale analysis. *Machine Vision and Applications*, 11(2), 64-73. doi: 10.1007/s001380050091
- Breceda, A., Arriaga, L., Bojorquez, L., & Rodriguez, M. (2005). Defining critical areas for conservation and restoration in a Mexican biosphere reserve: A case study. *Natural Areas Journal*, 25(2), 123-129.
- Bressan, G. M., Oliveira, V. A., Hruschka Jr, E. R., & Nicoletti, M. C. (2009). Using Bayesian networks with rule extraction to infer the risk of weed infestation in a corn-crop. *Engineering Applications of Artificial Intelligence*, 22(4-5), 579-592.
- Bucagu, C., Vanlauwe, B., & Giller, K. E. (2013). Managing Tephrosia mulch and fertilizer to enhance coffee productivity on smallholder farms in the Eastern African Highlands. *European Journal of Agronomy*, 48, 19-29. doi: 10.1016/j.eja.2013.02.005
- Bustillos-Herrera, J. A., Valdez-Lazalde, J. R., Aldrete, A., & Gonzalez-Guillen, M. D. (2007). Land suitability for eucalyptus (*Eucalyptus grandis* Hill ex Maiden) plantations: Definition through the analytic hierarchy process and gis. *Agrociencia*, 41(7), 787-796.
- 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. doi: <http://dx.doi.org/10.1016/j.agee.2012.04.013>
- Chemura, A., Kutwayo, D., Chidoko, P., & Mahoya, C. (2016). Bioclimatic modelling of current and projected climatic suitability of coffee (*Coffea arabica*) production in Zimbabwe. *Regional Environmental Change*, 16(2), 473-485. doi: 10.1007/s10113-015-0762-9
- Chen, Q., Baldocchi, D., Gong, P., & Kelly, M. (2006). Isolating individual trees in a savanna woodland using small footprint lidar data. *Photogrammetric Engineering and Remote Sensing*, 72(8), 923-932.
- Chen, Y., Khan, S., & Paydar, Z. (2010). To retire or expand? A fuzzy GIS-based spatial multi-criteria evaluation framework for irrigated agriculture. *Irrigation and Drainage*, 59(2), 174-188. doi: 10.1002/ird.470
- Cheng, H., & Wang, R. (2010). Semantic modeling of natural scenes based on contextual Bayesian networks. *Pattern Recognition*, 43(12), 4042-4054.
- Choi, M. Y., Choi, W., Park, J. H., Lim, J., & Kwon, S. W. (2010). Determination of coffee origins by integrated metabolomic approach of combining multiple analytical data. *Food Chemistry*, 121(4), 1260-1268. doi: 10.1016/j.foodchem.2010.01.035

- Chow, T. E., & Sadler, R. (2010). The consensus of local stakeholders and outside experts in suitability modeling for future camp development. *Landscape and Urban Planning, 94*(1), 9-19. doi: 10.1016/j.landurbplan.2009.07.013
- Chubey, M. S., Franklin, S. E., & Wulder, M. A. (2006). Object-based analysis of Ikonos-2 imagery for extraction of forest inventory parameters. *Photogrammetric Engineering and Remote Sensing, 72*(4), 383-394.
- Cordero-Sancho, S., & Sader, S. A. (2007). Spectral analysis and classification accuracy of coffee crops using Landsat and a topographic-environmental model. *International Journal of Remote Sensing, 28*(7-8), 1577-1593. doi: 10.1080/01431160600887680
- Cordingley, J. (2009). Soil Fertility Survey of Rwanda for Developing Lime and Fertilizer Recommendations for coffee (pp. 62). Kigali: OCIR Cafe.
- Costantini, E. A. C., Lorenzetti, R., & Malorgio, G. (2016). A multivariate approach for the study of environmental drivers of wine economic structure. *Land Use Policy, 57*, 53-63. doi: 10.1016/j.landusepol.2016.05.015
- Craparo, A. C. W., Van Asten, P. J. A., Laderach, P., Jassogne, L. T. P., & Grab, S. W. (2015). Coffea arabica yields decline in Tanzania due to climate change: Global implications. *Agricultural and Forest Meteorology, 207*, 1-10. doi: 10.1016/j.agrformet.2015.03.005
- Cyamweshi, R., Nabahungu, N., Mukashema, A., Ruganzu, V., Gatarayiha, M., Nduwumuremyi, A., & Mbonigaba, J. (2014). Enhancing nutrient availability and coffee yield on acid soils of the central plateau of Southern Rwanda. *Glob J Agric Res, 2*(2), 44-554.
- da Silva, E., 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*, 229-238.
- Davis, A. P., Gole, T. W., Baena, S., & Moat, J. (2012). The Impact of Climate Change on Indigenous Arabica Coffee (*Coffea arabica*): Predicting Future Trends and Identifying Priorities. *Plos One, 7*(11). doi: 10.1371/journal.pone.0047981
- de Muner, L. H., Maser, O., Fornazier, M. J., de Souza, C. V., & de Loreto, M. D. S. (2015). Energetic sustainability of three Arabica coffee growing systems used by family farming units in Espírito Santo state. *Engenharia Agricola, 35*(3), 397-405. doi: 10.1590/1809-4430-Eng.Agric.v35n3p397-405/2015
- de Toledo, P. R. A. B., de Melo, M. M. R., Pezza, H. R., Toci, A. T., Pezza, L., & Silva, C. M. (2017). Discriminant analysis for unveiling the origin of roasted coffee samples: A tool for quality control of coffee related

- products. *Food Control*, 73, Part B, 164-174. doi: <http://doi.org/10.1016/j.foodcont.2016.08.001>
- 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. doi: 10.1111/j.1365-2621.2003.tb05772.x
- Despic, O., & Simonovic, S. P. (2000). Aggregation operators for soft decision making in water resources. *Fuzzy Sets and Systems*, 115(1), 11-33.
- Diaz-Jose, J., Diaz-Jose, O., Mora-Flores, S., Rendon-Medel, R., & Tellez-Delgado, R. (2014). Cacao in Mexico: Restrictive factors and productivity levels. *Chilean Journal of Agricultural Research*, 74(4), 397-403. doi: 10.4067/s0718-58392014000400004
- dos Santos, A. R., Paterlini, E. M., Fiedler, N. C., Ribeiro, C. A. A. S., Lorenzon, A. S., Domingues, G. F., . . . Iwakiri, S. (2017). Fuzzy logic applied to prospecting for areas for installation of wood panel industries. *Journal of Environmental Management*, 193, 345-359. doi: <https://doi.org/10.1016/j.jenvman.2017.02.049>
- Doxani, G., Karantzalos, K., & Strati, M. T. (2012). Monitoring urban changes based on scale-space filtering and object-oriented classification. *International Journal of Applied Earth Observation and Geoinformation*, 15(0), 38-48.
- Folmer, B. (2014). How can science help to create new value in coffee? *Food Research International*, 63, 477-482. doi: 10.1016/j.foodres.2014.03.020
- Froni, F., Vignando, M., Aiello, M., Parma, V., Paoletti, M. G., Squartini, A., & Rumiati, R. I. (2017). The smell of terroir! Olfactory discrimination between wines of different grape variety and different terroir. *Food Quality and Preference*, 58, 18-23. doi: <https://doi.org/10.1016/j.foodqual.2016.12.012>
- Forster, D., Buehler, Y., & Kellenberger, T. W. (2009). Mapping urban and peri-urban agriculture using high spatial resolution satellite data. *Journal of Applied Remote Sensing*, 3. doi: 033523
10.1117/1.3122364
- Gambelli, D., & Bruschi, V. (2010). A Bayesian network to predict the probability of organic farms' exit from the sector: A case study from Marche, Italy. *Computers and Electronics in Agriculture*, 71(1), 22-31.
- Gangjee, D. S. (2017). GIs Beyond Wine: Time to Rethink the Link? *IIC - International Review of Intellectual Property and Competition Law*, 48(2), 129-133. doi: 10.1007/s40319-017-0556-3
- Gay, C., Estrada, F., Conde, C., Eakin, H., & Villers, L. (2006). Potential impacts of climate change on agriculture: A case of study of coffee production in Veracruz, Mexico. *Climatic Change*, 79(3-4), 259-288. doi: 10.1007/s10584-006-9066-x

- Gereffi, G., Humphrey, J., & Sturgeon, T. (2005). The governance of global value chains. *Review of International Political Economy*, 12(1), 78-104. doi: 10.1080/09692290500049805
- Gomez, C., Mangeas, M., Petit, M., Corbane, C., Hamon, P., Hamon, S., . . . Despinoy, M. (2010). Use of high-resolution satellite imagery in an integrated model to predict the distribution of shade coffee tree hybrid zones. *Remote Sensing of Environment*, 114(11), 2731-2744.
- Gong, P., Sheng, Y., & Biging, G. S. (2002). 3D model-based tree measurement from high-resolution aerial imagery. *Photogrammetric Engineering and Remote Sensing*, 68(11), 1203-1212.
- Gonzalez-Centeno, M. R., Adrover-Obrador, S., Simal, S., Frau, M. A., Femenia, A., & Rossello, C. (2015). A multivariate methodology to distinguish among wine Appellations of Origin. *Agronomy for Sustainable Development*, 35(1), 295-304. doi: 10.1007/s13593-014-0242-5
- Gougeon, F. A., & Leckie, D. G. (2006). The individual tree crown approach applied to Ikonos images of a coniferous plantation area. *Photogrammetric Engineering and Remote Sensing*, 72(11), 1287-1297.
- Haller, A. (2014). The "sowing of concrete": Peri-urban smallholder perceptions of rural-urban land change in the Central Peruvian Andes. *Land Use Policy*, 38, 239-247. doi: 10.1016/j.landusepol.2013.11.010
- Hernández-Martínez, G., Manson, R. H., & Hernández, A. C. (2009). Quantitative classification of coffee agroecosystems spanning a range of production intensities in central Veracruz, Mexico. *Agriculture, Ecosystems & Environment*, 134(1-2), 89-98. doi: <http://dx.doi.org/10.1016/j.agee.2009.05.020>
- Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25(15), 1965-1978. doi: 10.1002/joc.1276
- Hinrichs, C. C. (2003). The practice and politics of food system localization. *Journal of Rural Studies*, 19(1), 33-45.
- Huhndorf, M. H., Kerbis Peterhans, J. C., & Loew, S. S. (2007). Comparative phylogeography of three endemic rodents from the Albertine Rift, east central Africa. *Molecular Ecology*, 16(3), 663-674. doi: 10.1111/j.1365-294X.2007.03153.x
- Jayakumar, M., Rajavel, M., & Surendran, U. (2016). Climate-based statistical regression models for crop yield forecasting of coffee in humid tropical Kerala, India. *International Journal of Biometeorology*, 1-10. doi: 10.1007/s00484-016-1181-4
- Jayathilaka, P. M. S., Soni, P., Perret, S. R., Jayasuriya, H. P. W., & Salokhe, V. M. (2012). Spatial assessment of climate change effects on crop

- suitability for major plantation crops in Sri Lanka. *Regional Environmental Change*, 12(1), 55-68. doi: 10.1007/s10113-011-0235-8
- Joët, T., Laffargue, A., Descroix, F., Doubeau, S., Bertrand, B., Kochko, A. d., & Dussert, S. (2010). Influence of environmental factors, wet processing and their interactions on the biochemical composition of green Arabica coffee beans. *Food Chemistry*, 118(3), 693-701. doi: <http://dx.doi.org/10.1016/j.foodchem.2009.05.048>
- Johansen, K., Coops, N. C., Gergel, S. E., & Stange, Y. (2007). Application of high spatial resolution satellite imagery for riparian and forest ecosystem classification. *Remote Sensing of Environment*, 110(1), 29-44.
- Johansen, K., Phinn, S., Witte, C., Philip, S., & Newton, L. (2009). Mapping Banana Plantations from Object-oriented Classification of SPOT-5 Imagery. *Photogrammetric Engineering and Remote Sensing*, 75(9), 1069-1081.
- Kim, S.-R., Kwak, D.-A., Lee, W.-K., Son, Y., Bae, S.-W., Kim, C., & Yoo, S. (2010). Estimation of carbon storage based on individual tree detection in *Pinus densiflora* stands using a fusion of aerial photography and LiDAR data. *SCIENCE CHINA Life Sciences*, 53(7), 885-897. doi: 10.1007/s11427-010-4017-1
- Kok, K., & Veldkamp, A. (2001). Evaluating impact of spatial scales on land use pattern analysis in Central America. *Agriculture Ecosystems & Environment*, 85(1-3), 205-221. doi: 10.1016/S0167-8809(01)00185-2
- Kumar, R., & Silva, L. (1974). Statistical separability of spectral classes of blighted corn. *Remote Sensing of Environment*, 3(2), 109-115.
- Laderach, P., Ramirez-Villegas, J., Navarro-Racines, C., Zelaya, C., Martinez-Valle, A., & Jarvis, A. (2017). Climate change adaptation of coffee production in space and time. *Climatic Change*, 141(1), 47-62. doi: 10.1007/s10584-016-1788-9
- Lambin, E. F., Turner, B. L., Geist, H. J., Agbola, S. B., Angelsen, A., Bruce, J. W., . . . Xu, J. (2001). The causes of land-use and land-cover change: moving beyond the myths. *Global Environmental Change*, 11(4), 261-269. doi: [http://dx.doi.org/10.1016/S0959-3780\(01\)00007-3](http://dx.doi.org/10.1016/S0959-3780(01)00007-3)
- Langford, M., & Bell, W. (1997). Land cover mapping in a tropical hillsides environment: A case study in the Cauca region of Colombia. *International Journal of Remote Sensing*, 18(6), 1289-1306. doi: 10.1080/014311697218421
- Lin, B. B. (2007). Agroforestry management as an adaptive strategy against potential microclimate extremes in coffee agriculture. *Agricultural and Forest Meteorology*, 144(1-2), 85-94. doi: 10.1016/j.agrformet.2006.12.009

- Liu, X., Li, F., Zhang, Y., & Yang, Q. (2016). Effects of deficit irrigation on yield and nutritional quality of Arabica coffee (*Coffea arabica*) under different N rates in dry and hot region of southwest China. *Agricultural Water Management*, 172, 1-8. doi: <http://dx.doi.org/10.1016/j.agwat.2016.04.007>
- Loveridge, S., Mpyisi, E., & Weber, M. T. (2002). Farm-Level Perspectives in Rwanda's Coffee Supply Chain Coordination Challenge (R. F. S. R. Project/MINAGRI-RWANDA, Trans.) *Agricultural Policy Synthesis* (Vol. 2E): Michigan State University
- Loveridge, S., Nyarwaya, J. B., & Shingiro, E. (2003). *Decaffeinated? Situation, Trends, and Prospects for Coffee Production in Rwanda*. (Number 5E). Kigali, Rwanda.
- Makropoulos, C. K., & Butler, D. (2006). Spatial ordered weighted averaging: incorporating spatially variable attitude towards risk in spatial multi-criteria decision-making. *Environmental Modelling & Software*, 21(1), 69-84. doi: 10.1016/j.envsoft.2004.10.010
- Marciniak, M., Reynolds, A. G., & Brown, R. (2009). Use of GIS, GPS, and Remote Sensing to Understand the Terroir of the Niagara Peninsula. *American Journal of Enology and Viticulture*, 60(4), 555A-555A.
- Marciniak, M., Reynolds, A. G., & Brown, R. (2013). Influence of water status on sensory profiles of Ontario Riesling wines. *Food Research International*, 54(1), 881-891. doi: 10.1016/j.foodres.2013.08.030
- Marciniak, M., Reynolds, A. G., Brown, R., Jollineau, M., & Kotsaki, E. (2017). Applications of Geospatial Technologies to Understand Terroir Effects in an Ontario Riesling Vineyard. *American Journal of Enology and Viticulture*, 68(2), 169-187. doi: 10.5344/ajev.2016.16083
- Marquetti, I., Link, J. V., Lemes, A. L. G., Scholz, M. B. d. S., Valderrama, P., & Bona, E. (2016). Partial least square with discriminant analysis and near infrared spectroscopy for evaluation of geographic and genotypic origin of arabica coffee. *Computers and Electronics in Agriculture*, 121, 313-319. doi: <http://dx.doi.org/10.1016/j.compag.2015.12.018>
- McSweeney, R., & Semafara, J. N. (2011). Green Growth and Climate Resilience (National Strategy for Climate Change and Low Carbon Development). In M. Cole (Ed.), *Rwanda's Climate: Observations and Projections* (pp. 44). Rwanda's Climate: Observations and Projections: Smith School of Enterprise and the Environment, University of Oxford.
- Mello, M. P., Rudorff, B. F. T., Adami, M., Rizzi, R., Aguiar, D. A., Gusso, A., . . . Ieee. (2010). *A simplified Bayesian Network to map soybean plantations*. New York: Ieee.
- MINAGRI. (2008). Strategic Plan for the Transformation of Agriculture in Rwanda-Phase II (PSTA II) (M. a. E. Planning, Trans.) (pp. 126). Kigali: Ministry of Agriculture and Animal Resources.

- Moreira, M. A., Adam, M., & Rudorff, B. F. T. (2004). Spectral and temporal behavior analysis of coffee crop in Landsat images. *Pesquisa Agropecuaria Brasileira*, 39(3), 223-231.
- Moreira, M. A., Rudorff, B. F. T., Barros, M. A., De Faria, V. G. C., & Adami, M. (2010). Geotechnologies to map coffee fields in the states of Minas Gerais and Sao Paulo. *Engenharia Agricola*, 30(6), 1123-1135.
- Muhire, I., & Ahmed, F. (2015). Spatio-temporal trend analysis of precipitation data over Rwanda. *South African Geographical Journal*, 97(1), 50-68. doi: 10.1080/03736245.2014.924869
- Mukashema, A., Veldkamp, A., & Vrieling, A. (2014). Automated high resolution mapping of coffee in Rwanda using an expert Bayesian network. *International Journal of Applied Earth Observation and Geoinformation*, 33, 331-340. doi: 10.1016/j.jag.2014.05.005
- Mukashema, A., Veldkamp, T., & Amer, S. (2016). Sixty percent of small coffee farms have suitable socio-economic and environmental locations in Rwanda. *Agronomy for Sustainable Development*, 36(2). doi: 10.1007/s13593-016-0363-0
- Murekezi, A. (2009). *Essays on the Effects of Coffee Market Reforms, Supply Chains, and Income Improvement in Rwanda*. (PhD dissertation), Michigan State University. Retrieved from http://www.aec.msu.edu/theses/fulltext/murekezi_phd.pdf
- Murekezi, A., Jin, S. Q., & Loveridge, S. (2014). Have coffee producers benefited from the new domestic cherry market? Evidence using panel data from Rwanda. *Agricultural Economics*, 45(4), 489-500. doi: 10.1111/agec.12101
- Murugan, M., Shetty, P. K., Ravi, R., Anandhi, A., & Rajkumar, A. J. (2012). Climate change and crop yields in the Indian Cardamom Hills, 1978-2007 CE. *Climatic Change*, 110(3-4), 737-753. doi: 10.1007/s10584-011-0115-8
- NAEB. (2016). National Coffee Census 2015 (C. Division, Trans.) (pp. 33). 2016: Rwanda National Agricultural Export Board.
- NISR. (2012). Rwanda Fourth Population and Housing Census (NISR, Trans.) (Vol. Thematic Report: Characteristics of households and housing, pp. 256). Kigali: National Institute of Statistics of Rwanda, Ministry of Finance and Economic Planning (MINECOFIN).
- Nol, L., Heuvelink, G. B. M., Veldkamp, A., de Vries, W., & Kros, J. (2010). Uncertainty propagation analysis of an N2O emission model at the plot and landscape scale. *Geoderma*, 159(1-2), 9-23. doi: 10.1016/j.geoderma.2010.06.009
- Nsengiyumva, F. (2009). *Rwanda National Coffee Census*. Kigali, Rwanda.
- Nyandwi, E., Veldkamp, T., & Amer, S. (2016). Regional climate sensitivity of wetland environments in Rwanda: the need for a location-specific approach. *Regional Environmental Change*, 16(6), 1635-1647. doi: 10.1007/s10113-015-0905-z

- Nzeyimana, I., Hartemink, A. E., & de Graaff, J. (2013). Coffee farming and soil management in Rwanda. *Outlook on Agriculture*, 42(1), 47-52. doi: 10.5367/oa.2013.0118
- Nzeyimana, I., Hartemink, A. E., & Geissen, V. (2014). GIS-Based Multi-Criteria Analysis for Arabica Coffee Expansion in Rwanda. *Plos One*, 9(10). doi: 10.1371/journal.pone.0107449
- Nzeyimana, I., Hartemink, A. E., Ritsema, C., Stroosnijder, L., Lwanga, E. H., & Geissen, V. (2017). Mulching as a strategy to improve soil properties and reduce soil erodibility in coffee farming systems of Rwanda. *Catena*, 149, Part 1, 43-51. doi: <http://doi.org/10.1016/j.catena.2016.08.034>
- OCIR-CAFE. (2005). New Action Plan 2006-2008 for the development of Rwabda Coffee Sector (pp. 23). Kigali: OCIR-CAFE.
- Ofori-Boateng, K., & Insah, B. (2014). The impact of climate change on cocoa production in West Africa. *International Journal of Climate Change Strategies and Management*, 6(3), 296-314. doi: 10.1108/ijccsm-01-2013-0007
- Ortega-Huerta, M. A., Komar, O., Price, K. P., & Ventura, H. J. (2012). Mapping coffee plantations with Landsat imagery: an example from El Salvador. *International Journal of Remote Sensing*, 33(1), 220-242. doi: 10.1080/01431161.2011.591442
- Ovalle-Rivera, O., Laderach, P., Bunn, C., Obersteiner, M., & Schroth, G. (2015). Projected Shifts in Coffea arabica Suitability among Major Global Producing Regions Due to Climate Change. *Plos One*, 10(4). doi: 10.1371/journal.pone.0124155
- Overmars, K. P., de Koning, G. H. J., & Veldkamp, A. (2003). Spatial autocorrelation in multi-scale land use models. *Ecological Modelling*, 164(2-3), 257-270. doi: 10.1016/s0304-3800(03)00070-x
- Park, M.-H., & Stenstrom, M. K. (2006). Using satellite imagery for stormwater pollution management with Bayesian networks. *Water Research*, 40(18), 3429-3438.
- Perveen, M. F., Nagasawa, R., Ahmed, A. O. C., Uddin, M. I., & Kimura, R. (2008). Integrating biophysical and socio-economic data using GIS for land evaluation of wheat cultivation: A case study in north-west Bangladesh. *Journal of Food Agriculture & Environment*, 6(2), 432-437.
- Picini, A. G., Camargo, M. B. P. D., Ortolani, A. A., Fazuoli, L. C., & Gallo, P. B. (1999). Test and analysis of agrometeorological models for predicting coffee yield. *Bragantia*, 58, 157-170.
- Pinard, F., Boffa, J. M., & Rwakagara, E. (2014). Scattered shade trees improve low-input smallholder Arabica coffee productivity in the Northern Lake Kivu region of Rwanda. *Agroforestry Systems*, 88(4), 707-718. doi: 10.1007/s10457-014-9712-7

- Ponte, S. (2002). The 'latte revolution'? Regulation, markets and consumption in the global coffee chain. *World Development*, 30(7), 1099-1122. doi: 10.1016/s0305-750x(02)00032-3
- Pouliot, D. A., King, D. J., Bell, F. W., & Pitt, D. G. (2002). Automated tree crown detection and delineation in high-resolution digital camera imagery of coniferous forest regeneration. *Remote Sensing of Environment*, 82(2-3), 322-334.
- Pretty, J., Toulmin, C., & Williams, S. (2011). Sustainable intensification in African agriculture. *International Journal of Agricultural Sustainability*, 9(1), 5-24. doi: 10.3763/ijas.2010.0583
- Qu, Y., Wang, J., Wan, H., Li, X., & Zhou, G. (2008). A Bayesian network algorithm for retrieving the characterization of land surface vegetation. *Remote Sensing of Environment*, 112(3), 613-622.
- Quiñones-Ruiz, X. F., Penker, M., Belletti, G., Marescotti, A., Scaramuzzi, S., Barzini, E., . . . Samper-Gartner, L. F. (2016). Insights into the black box of collective efforts for the registration of Geographical Indications. *Land Use Policy*, 57, 103-116. doi: <http://doi.org/10.1016/j.landusepol.2016.05.021>
- Ramos, M. F., Ribeiro, D. E., Cirillo, M. A., & Borem, F. M. (2016). Discrimination of the sensory quality of the *Coffea arabica* L. (cv. Yellow Bourbon) produced in different altitudes using decision trees obtained by the CHAID method. *Journal of the Science of Food and Agriculture*, 96(10), 3543-3551. doi: 10.1002/jsfa.7539
- Rao, C. S., Gopinath, K. A., Prasad, J., Prasannakumar, & Singh, A. K. (2016). Climate Resilient Villages for Sustainable Food Security in Tropical India: Concept, Process, Technologies, Institutions, and Impacts. In D. L. Sparks (Ed.), *Advances in Agronomy, Vol 140* (Vol. 140, pp. 101-214).
- Reynolds, A. G., Senchuk, I. V., van der Reest, C., & de Savigny, C. (2007). Use of GPS and GIS for elucidation of the basis for terroir: Spatial variation in an Ontario Riesling vineyard. *American Journal of Enology and Viticulture*, 58(2), 145-162.
- Rolle, L., Siret, R., Segade, S. R., Maury, C., Gerbi, V., & Jourjon, F. (2012). Instrumental Texture Analysis Parameters as Markers of Table-Grape and Winegrape Quality: A Review. *American Journal of Enology and Viticulture*, 63(1), 11-28. doi: 10.5344/ajev.2011.11059
- Roose, E., & Ndayizigiye, F. (1997). Agroforestry, water and soil fertility management to fight erosion in tropical mountains of Rwanda. *Soil Technology*, 11(1), 109-119. doi: [http://dx.doi.org/10.1016/S0933-3630\(96\)00119-5](http://dx.doi.org/10.1016/S0933-3630(96)00119-5)
- Rotaru, L., Filipov, F., Mustea, M., & Stoleru, V. (2010). Influence of some "Terroir Viticole" Factors on Quantity and Quality of Grapes. *Notulae Botanicae Horti Agrobotanici Cluj-Napoca*, 38(1), 176-181.

- Rutunga, V., Janssen, B. H., Mantel, S., & Janssens, M. (2007). Soil use and management strategy for raising food and cash output in Rwanda. *Journal of Food Agriculture & Environment*, 5(3-4), 434-441.
- Sanchez, M. G. B., Marques, J., Siqueira, D. S., Camargo, L. A., & Pereira, G. T. (2013). Delineation of specific management areas for coffee cultivation based on the soil-relief relationship and numerical classification. *Precision Agriculture*, 14(2), 201-214. doi: 10.1007/s11119-012-9288-z
- Santos, M. A. d., & Camargo, M. B. P. d. (2006). Calibration of an agrometeorological model for predicting coffee (*Coffea arabica* L.) productivity in Sao Paulo State, Brazil. *Bragantia*, 65, 173-183.
- Schilling, M., Mukashema, A., Westinga, E., Lasry, F., Mlotha, J., & Iraguha, G. (2007). *Forest Mapping of Rwanda*. Synthesis. Consultancy, MINIRENA-RWANDA. Centre for Geographic Information Systems and Remote Sensing (CGIS), National University of Rwanda (NUR). Huye.
- Schlesier, K., Fahl-Hassek, C., Forina, M., Cotea, V., Kocsi, E., Schoula, R., . . . Wittkowski, R. (2009). Characterisation and determination of the geographical origin of wines. Part I: overview. *European Food Research and Technology*, 230(1), 1-13. doi: 10.1007/s00217-009-1140-y
- Schroth, G., Laderach, P., Dempewolf, J., Philpott, S., Haggard, J., Eakin, H., . . . Ramirez-Villegas, J. (2009). Towards a climate change adaptation strategy for coffee communities and ecosystems in the Sierra Madre de Chiapas, Mexico. *Mitigation and Adaptation Strategies for Global Change*, 14(7), 605-625.
- Silva, S. D., de Queiroz, D. M., Ferreira, W. P. M., Correa, P. C., & Rufinod, J. L. D. (2016). Mapping the potential beverage quality of coffee produced in the Zona da Mata, Minas Gerais, Brazil. *Journal of the Science of Food and Agriculture*, 96(9), 3098-3108. doi: 10.1002/jsfa.7485
- Silva, S. D., de Queiroz, D. M., Pinto, F. D. D., & Santos, N. T. (2014). Characterization and delimitation of the terroir coffee in plantations in the municipal district of Araponga, Minas Gerais, Brazil. *Revista Ciencia Agronomica*, 45(1), 18-26.
- Smeyers-Verbeke, J., Jager, H., Lanteri, S., Brereton, P., Jamin, E., Fahl-Hassek, C., . . . Romisch, U. (2009). Characterization and determination of the geographical origin of wines. Part II: descriptive and inductive univariate statistics. *European Food Research and Technology*, 230(1), 15-29. doi: 10.1007/s00217-009-1142-9
- Songer, P. (2008). Development of Appellation Designations based upon Flavor Profiles for Rwanda Regional Coffees (pp. 83): SONGER & ASSOCIATE, INC.

- Souza, H. N. d., de Goede, R. G. M., Brussaard, L., Cardoso, I. M., Duarte, E. M. G., Fernandes, R. B. A., . . . Pulleman, M. M. (2012). Protective shade, tree diversity and soil properties in coffee agroforestry systems in the Atlantic Rainforest biome. *Agriculture, Ecosystems & Environment*, 146(1), 179-196. doi: <http://dx.doi.org/10.1016/j.agee.2011.11.007>
- Sporcic, M., Landekic, M., Lovric, M., Bogdan, S., & Segotic, K. (2010). Multiple criteria decision making in forestry - methods and experiences. *Sumarski List*, 134(5-6), 275-286.
- Steen, I., Waehrens, S. S., Petersen, M. A., Munchow, M., & Bredie, W. L. P. (2017). Influence of serving temperature on flavour perception and release of Bourbon Caturra coffee. *Food Chemistry*, 219, 61-68. doi: 10.1016/j.foodchem.2016.09.113
- Swedesurvey. (2010). *Report for Production of ortho photo in Rwanda*. Rwanda National Land Use and Development Master Plan. Technical report. Rwanda National Land Centre, Ministry of Natural Resources. Kigali, Rwanda.
- Taleai, M., Sharifi, A., Sliuzas, R., & Mesgari, M. (2007). Evaluating the compatibility of multi-functional and intensive urban land uses. *International Journal of Applied Earth Observation and Geoinformation*, 9(4), 375-391. doi: 10.1016/j.jag.2006.12.002
- Tardin, A. T., Deassuncao, G. V., & Soares, J. V. (1992). Preliminary analysis of TM data for coffee, citrus and sugarcane discrimination in furnas region, mg, Brazil. *Pesquisa Agropecuaria Brasileira*, 27(9), 1355-1361.
- Teixeira, A. L., Souza, F. D., Pereira, A. A., de Oliveira, A. C. B., & Rocha, R. B. (2015). Selection of arabica coffee progenies tolerant to heat stress. *Ciencia Rural*, 45(7), 1228-1234. doi: 10.1590/0103-8478cr20130317
- Tittonell, P., & Giller, K. E. (2013). When yield gaps are poverty traps: The paradigm of ecological intensification in African smallholder agriculture. *Field Crops Research*, 143(0), 76-90. doi: <http://dx.doi.org/10.1016/j.fcr.2012.10.007>
- Trabaquini, K., Miglioranza, E., de Franca, V., & Neto, O. C. P. (2011). Characterization of coffee crops using techniques of geoprocessing and remote sensing in the city of Umuarama - PR. *Ciencia E Agrotecnologia*, 35(1), 35-44. doi: 10.1590/s1413-70542011000100004
- Tran, H. T. M., Lee, L. S., Furtado, A., Smyth, H., & Henry, R. J. (2016). Advances in genomics for the improvement of quality in coffee. *Journal of the Science of Food and Agriculture*, 96(10), 3300-3312. doi: 10.1002/jsfa.7692
- UNEP. (2011). *Rwanda: From Post-Conflict to Environmentally Sustainable Development*. Nairobi: United Nations Environment Programme.

- 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(2), 197-204. doi: 10.1002/jsfa.2338
- van Ittersum, M. K., Cassman, K. G., Grassini, P., Wolf, J., Tittonell, P., & Hochman, Z. (2013). Yield gap analysis with local to global relevance-A review. *Field Crops Research*, 143, 4-17. doi: 10.1016/j.fcr.2012.09.009
- van Soesbergen, A., Arnell, A. P., Sassen, M., Stuch, B., Schaldach, R., Göpel, J., . . . Palazzo, A. (2016). Exploring future agricultural development and biodiversity in Uganda, Rwanda and Burundi: a spatially explicit scenario-based assessment. *Regional Environmental Change*, 1-12. doi: 10.1007/s10113-016-0983-6
- Vasconcelos, A. L. S., Franca, A. S., Gloria, M. B. A., & Mendonca, J. C. F. (2007). A comparative study of chemical attributes and levels of amines in defective green and roasted coffee beans. *Food Chemistry*, 101(1), 26-32. doi: 10.1016/j.foodchem.2005.12.049
- Vaudour, E. (2002). The Quality of Grapes and Wine in Relation to Geography: Notions of Terroir at Various Scales. *Journal of Wine Research*, 13(2), 117-141. doi: 10.1080/0957126022000017981
- Vaudour, E., Leclercq, L., Gilliot, J. M., & Chaignon, B. (2017). Retrospective 70 y-spatial analysis of repeated vine mortality patterns using ancient aerial time series, Pléiades images and multi-source spatial and field data. *International Journal of Applied Earth Observation and Geoinformation*, 58, 234-248. doi: <https://doi.org/10.1016/j.jag.2017.02.015>
- Veldkamp, A., & Fresco, L. O. (1997). Reconstructing land use drivers and their spatial scale dependence for Costa Rica (1973 and 1984). *Agricultural Systems*, 55(1), 19-43. doi: 10.1016/s0308-521x(95)00080-o
- Veldkamp, A., Kok, K., De Koning, G. H. J., Schoorl, J. M., Sonneveld, M. P. W., & Verburg, P. H. (2001). Multi-scale system approaches in agronomic research at the landscape level. *Soil & Tillage Research*, 58(3-4), 129-140. doi: 10.1016/s0167-1987(00)00163-x
- Veldkamp, A., Verburg, P. H., Kok, K., de Koning, G. H. J., Priess, J., & Bergsma, A. R. (2001). The need for scale sensitive approaches in spatially explicit land use change modeling. *Environmental Modeling & Assessment*, 6(2), 111-121. doi: 10.1023/a:1011572301150
- Verbist, B., Dinata Putra, A. E., & Budidarsono, S. (2005). Factors driving land use change: Effects on watershed functions in a coffee agroforestry system in Lampung, Sumatra. *Agricultural Systems*, 85(3), 254-270.

- Verburg, P. H., & Veldkamp, A. (2001). The role of spatially explicit models in land-use change research: a case study for cropping patterns in China. *Agriculture Ecosystems & Environment*, 85(1-3), 177-190. doi: 10.1016/S0167-8809(01)00184-0
- Verdoodt, A., & Van Ranst, E. (2003a). *Land Evaluation for agricultural production in the tropics: A large-scale Land Suitability classification in Rwanda*. Gent University, Belgium.
- Verdoodt, A., & Van Ranst, E. (2003b). *Land Evaluation for agricultural production in the tropics: A two-level Crop Growth Model for annual crops*. Gent University, Belgium.
- Verdoodt, A., & Van Ranst, E. (2006). Environmental assessment tools for multi-scale land resources information systems - A case study of Rwanda. *Agriculture Ecosystems & Environment*, 114(2-4), 170-184. doi: 10.1016/j.agee.2005.10.006
- Verwimp, P. (2003). The political economy of coffee, dictatorship, and genocide. *European Journal of Political Economy*, 19(2), 161-181.
- Wang, L., Gong, P., & Biging, G. S. (2004). Individual tree-crown delineation and treetop detection high-spatial-resolution aerial imagery. *Photogrammetric Engineering and Remote Sensing*, 70(3), 351-357.
- Wang, N., Jassogne, L., van Asten, P. J. A., Mukasa, D., Wanyama, I., Kagezi, G., & Giller, K. E. (2015). Evaluating coffee yield gaps and important biotic, abiotic, and management factors limiting coffee production in Uganda. *European Journal of Agronomy*, 63, 1-11. doi: 10.1016/j.eja.2014.11.003
- Westinga, E., Mukashema, A., & van Gils, H. (2013). A comparison of fine resolution census and image-based national forest inventories: a case study of Rwanda. *Forestry*, 86(4), 453-461. doi: 10.1093/forestry/cpt016
- Wintgens, J. N. (2009). *Coffee: Growing, Processing, Sustainable Production. A guidebook for Growers, Processors, Traders, and Researchers* (Second, updated ed.). Weinheim: WILEY-VCH VerlagGmbH&Co.KGaA.
- Wulder, M., Niemann, K. O., & Goodenough, D. G. (2000). Local Maximum Filtering for the Extraction of Tree Locations and Basal Area from High Spatial Resolution Imagery. *Remote Sensing of Environment*, 73(1), 103-114.
- Xiao, D. P., Shen, Y. J., Zhang, H., Moiwo, J. P., Qi, Y. Q., Wang, R. D., . . . Shen, H. T. (2016). Comparison of winter wheat yield sensitivity to climate variables under irrigated and rain-fed conditions. *Frontiers of Earth Science*, 10(3), 444-454. doi: 10.1007/s11707-015-0534-3
- Youkhana, A., & Idol, T. (2009). Tree pruning mulch increases soil C and N in a shaded coffee agroecosystem in Hawaii. *Soil Biology and Biochemistry*, 41(12), 2527-2534.

Bibliography

- Youkhana, A. H., & Idol, T. W. (2016). Leucaena-KX2 mulch additions increase growth, yield and soil C and N in a managed full-sun coffee system in Hawaii. *Agroforestry Systems*, *90*(2), 325-337. doi: 10.1007/s10457-015-9857-z
- Zou, J. F., Peng, Z. X., Du, H. J., Duan, C. Q., Reeves, M. J., & Pan, Q. H. (2012). Elemental Patterns of Wines, Grapes, and Vineyard Soils from Chinese Wine-Producing Regions and Their Association. *American Journal of Enology and Viticulture*, *63*(2), 232-240. doi: 10.5344/ajev.2012.11087

Summary

Coffee is an important crop that provides a livelihood to millions of people living in developing countries and to over 355 thousand Rwandan families in particular. It is the second most traded commodity in the world after petroleum. However, coffee business development is lagging behind other industries such as for example wine in terms of product differentiation. No specific terroirs exist yet, although the origin and provenance of coffee beans becomes more important in the global coffee trade. For Rwanda, this constitutes a challenge of serving the specialty coffee market because traders demand a consistent high quality. Arabica coffee is grown exclusively by smallholder farmers. Such small-scale farming system is a challenge for specialty coffee market because it involves collecting and mixing coffee from many different producers to meet the quantity demand without compromising the quality. Hence the aim of this study was to identify specialty coffee terroirs of Rwanda, characterize them so that they will further serve as tool for policy-making towards protection of geographical origins of specialty coffee in Rwanda.

Mapping coffee terroirs requires relating coffee yield and sensory quality to its local production environment. The first activity was to be able to avail coffee map which will allow us to link coffee with its geographical production. We first developed an expert Bayesian network model to extract the small-scale coffee fields from aerial orthophotos covering more than 99% of Rwanda and on one QuickBird image for the remaining part. The method consists of a stepwise adjustment of pixel probabilities, which incorporates expert knowledge on size of coffee trees and fields, and on their location. The combination of spectral and digital elevation and forest data allowed mapping of coffee fields with an overall accuracy of 87%.

We secondly linked coffee fields with biophysical and socio-economic location factors to identify significant factors underlying coffee field location. We identified 29 potential factors, including demography, and environmental factors such as climate, soil and topography. These factors were reduced to 17 variables explaining 86 % of the total dataset variability, by factor analysis. Using multiple regression analysis, we produced agro-ecological based models used to predict suitable location for growing coffee in Rwanda. These regional models show that 60 % of the actual coffee farms are located in suitable locations.

Current coffee yields in Rwanda are about $0.7 \pm 0.4 \text{ t ha}^{-1}$ which is generally below the average yield (1.2 t ha^{-1}) obtained in small scale farming systems worldwide. An empirical analysis of coffee yields was made to determine which local environmental factors are associated with observed coffee yield levels. We third analysed the relationship between current yield level and the production environment. The developed yield models were used to estimate

the achievable coffee yields within the current smallholder farming systems. The resulting yield map shows that within Rwandan smallholder coffee systems, yield levels could be up to 40% higher if land management improves given local biophysical conditions. We further analysed the potential effect of projected near future climate change. Combined projected increase of temperature and rainfall revealed that yield levels would potentially increase by almost 60% in 2050.

Fourth, based on the results of annual Cup of Excellence competitions (2008-2015), we determined the geographical origins of specialty coffees. Multiple regression analysis was applied to identify significant biophysical location factors determining coffee sensory quality in Rwanda. The developed regression models were used to identify the geographic origins, their extent and their level of sensitivity to climate change. The study showed that the inherent coffee sensory quality in Rwanda is always determined by specific topographic and soil properties. Direct improvement of these soil factors is often not possible or cost ineffective. Furthermore, measures such as N-fertilizer applications, intended to increase yield levels lead to a reduced coffee sensory quality. Climate, specifically extreme temperatures, also reduces coffee quality regionally. A plausible climate change scenario for 2050 applied to the developed coffee quality regressions models demonstrated that coffee quality will probably have both regional increases as well as decreases. The net national change is a slight increase of only 1%, suggesting a limited overall climate sensitivity of Rwandan coffee quality for projected climate change.

By integrating Land suitability, achievable yields, and sensory quality potentials, in a fuzzy logic based spatial multi-criteria model, we delineated speciality coffee terroirs of Rwanda and their geographical indicators. We demonstrated also that coffee terroirs of Rwanda are climate sensitive. It is important that farmers in current coffee terroirs that will turn marginal revise their strategies in the face of climate change. This will include adopting new coffee varieties and agro-economic models that are suited to climate change with greater tolerance of high temperatures and altered pest and disease pressures. The government of Rwanda should consider renewing the national the policy accordingly that promote climate change in coffee sector developments and applied research.

Samenvatting

Koffie is een belangrijk gewas dat voor miljoenen mensen in ontwikkelingslanden een belangrijke bron van inkomsten is. In Rwanda voorziet koffie productie 355 duizend Rwandese families in hun levensonderhoud. Het is op olie na de meest verhandelde grondstof in de wereld. De ontwikkeling van de koffie-industrie blijft echter achter bij andere industrieën, zoals bijvoorbeeld wijn in termen van productdifferentiatie. Er bestaan nog geen specifieke 'terroirs', herkomst gebieden, al hoewel de oorsprong en herkomst van koffiebonen steeds belangrijker wordt in de wereldwijde koffiehandel. Voor Rwanda vormt dit een extra uitdaging om de koffiemarkt te bedienen, omdat handelaren een consistente hoge kwaliteit eisen. In Rwanda wordt Arabica koffie uitsluitend geteeld door kleine boeren. In dergelijke kleinschalig landbouwsystemen is extra moeilijk om consequent genoeg hoge kwaliteit koffie te produceren. Vaak wordt koffie van verschillende herkomst gemengd om aan de vraag te voldoen terwijl dit kan leiden tot een afname in de geleverde kwaliteit.

Het doel van deze studie is daarom om de locaties van hoogwaardige koffie, zogenaamde 'terroirs' in Rwanda te identificeren, ze te karakteriseren en te inventariseren, zodat ze als hulpmiddel kunnen dienen voor het verder ontwikkelen van de koffie productie in Rwanda.

Het karteren van koffie terroirs vereist een relatie tussen locatie geschiktheid, koffieopbrengst de sensorische kwaliteit en de lokale productiefactoren. Om deze relaties te kunnen onderzoeken is er een koffiekaart nodig om koffie karakteristieken te kunnen koppelen aan de locatie karakteristieken. We hebben daarom een nieuw expert Bayesiaans netwerkmodel ontwikkeld om de kleinschalige koffievelden te kunnen karteren op basis van bestaande orthofoto's welke meer dan 99% van Rwanda beslaan en een QuickBird-afbeelding voor het overige deel van het land. De combinatie van spectrale en digitale expert regels en waarneembare veldkarakteristieken maakte het mogelijk om koffievelden te inventariseren met een totale nauwkeurigheid van naar schatting 87%.

Vervolgens zijn alle koffievelden gekoppeld aan hun biofysische en sociaaleconomische locatie karakteristieken. Er zijn 29 mogelijke factoren geïdentificeerd welke de geschiktheid van de locatie voor koffie productie kan bepalen. Hieronder zijn demografische en omgevingsfactoren zoals klimaat, bodem en topografie. Deze factoren werden gereduceerd tot 17 variabelen die 86% van de totale dataset variabiliteit verklaarden, door middel van factoranalyse. Met behulp van meerdere regressieanalyses, zijn significante agro-ecologische modellen gemaakt welke gebruikt zijn om de geschikte locaties voor koffie in Rwanda te voorspellen. Deze regionale

geschiktheidsmodellen tonen aan dat 60% van de huidige koffievelden zich op geschikte locaties bevinden.

De gemeten koffieopbrengsten in Rwanda liggen ongeveer $0,7 \pm 0,4$ t ha⁻¹. Dit ligt onder het mondiale gemiddelde (1,2 t ha⁻¹) wat in kleinschalige landbouwsystemen wordt verkregen. Een empirische analyse van koffieopbrengsten is gemaakt om te bepalen welke lokale omgevingsfactoren geassocieerd zijn met waargenomen opbrengstniveaus van koffie. We hebben de relatie tussen het huidige opbrengstniveau en de productieomgeving geanalyseerd. De ontwikkelde opbrengstmodellen werden gebruikt om de haalbare koffieopbrengsten te beoordelen binnen de huidige kleinschalige landbouwsystemen. De resulterende opbrengstkaart laat zien dat binnen Rwandese kleinschalige koffiesystemen de opbrengstniveaus tot 40% hoger kunnen zijn indien plaatselijke biofysische omstandigheden verbeterd worden door de boeren. Het potentiële effect van geprojecteerde nabije toekomstige klimaatverandering zijn ook bekeken. Gecombineerde geprojecteerde toename van de temperatuur en de regenval bleek volgens de gemaakte opbrengstmodellen te leiden tot opbrengstniveaus in 2050 welke zo'n 60% hoger liggen.

Op basis van de resultaten van de jaarlijkse Cup of Excellence competities in Rwanda (2008-2015), hebben we de geografische herkomst van hoogwaardige koffie bonen bepaald. De kwaliteitsbeoordelingen van de Cup of Excellence competitie is hierbij leidend. Multiple regressieanalyses werden wederom toegepast om significante biofysische locatie factoren te bepalen die sensorische koffie kwaliteit in Rwanda bepalen. De ontwikkelde regressiemodellen werden gebruikt om de geografische oorsprong, hun omvang en hun gevoeligheid voor klimaatverandering te identificeren. Uit de studie bleek dat de inherente koffie sensorische kwaliteit in Rwanda altijd bepaald wordt door specifieke combinatie van topografische en bodemeigenschappen. Directe verbetering van deze bodemfactoren is vaak niet mogelijk of niet lonend. Bovendien kunnen maatregelen zoals hogere N-kunstmest giften, welke bedoeld zijn om de opbrengst te verhogen, juist leiden tot een verminderde sensorische kwaliteit van de geproduceerde koffie. Klimaat, met name extreme temperaturen, kunnen ook de regionale koffiekwiteit reduceren. Een aannemelijk klimaatverandering scenario voor 2050 toegepast op de ontwikkelde modellen van koffiekwiteit laat zien dat de koffiekwiteit waarschijnlijk zowel regionale toenames als afnames zal kennen. De netto nationale verandering laat een lichte stijging zien van slechts 1%, wat in tegenstelling tot de opbrengsten een beperkte algemene klimaatgevoeligheid voor Rwandese koffiekwiteit suggereert.

Door het integreren van koffie locatie geschiktheid, koffie opbrengsten en de sensorische koffie kwaliteit in een fuzzy logic gebaseerde multi-criteria

model, hebben we geprobeerd de koffie 'terroirs' van Rwanda af te bakenen. Deze exercitie laat zien waar de Rwandese koffie 'terroirs' zijn en hoe klimaatgevoelig hun locaties zijn. Het is daarom belangrijk dat de boeren in de huidige koffie 'terroirs' hun management strategieën aanpassen aan de komende verwachte klimaatverandering. Dit kan door het introduceren van nieuwe koffie variëteiten met een grotere tolerantie van hogere temperaturen en neerslag regimes. Ook wordt aangeraden dat de Rwandese overheid zijn nationale het beleid moet vernieuwen waarbij de informatie van dit onderzoek wordt meegewogen om de koffie productie meer toekomstbestendig te maken.

Author's Biography

Adrie Mukashema was born in Nyaruguru District, South province of Rwanda, on 24th December 1972. She obtained her Bachelor of Science in Agricultural Engineering in 2003 at the National University of Rwanda (NUR). She was recruited as assistant researcher at PEARL project, a NUR outreach and development program in coffee sector. In 2005, she was recruited as an Assistant Lecturer at NUR and was immediately sent to the Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, the Netherlands, to pursue a Master of Science in Natural Resource Management with specialisation in Soil information for Land Management. After completing this degree in March 2007, she returned to teaching at NUR and coordinating NUFFIC/NPT/RWA/071 project, a Dutch funded capacity building program at the NUR Centre for GIS and Remote Sensing (CGIS-NUR). In 2010, She was awarded a fellowship from NUFFIC to pursue her PhD studies at the University of Twente. In February, 2013, Adrie was appointed by the Cabinet of Rwanda as Deputy Director General of Rwanda Natural Resources Authority (RNRA), in charge of Department of Forestry and Nature conservation. In March 2016, she returned to University of Twente, to complete her PhD studies. Her research career focuses on land and coffee sector. She currently has the following scientific outputs:



ISI Journal Articles

Mukashema, A. Veldkamp, A., Gatarayiha M. C., Integrating land suitability, productivity and sensory quality in Coffee terroirs for sustainable coffee production systems in Rwanda In: *Land Use Policy* (in prep).

Mukashema, A. Veldkamp, A., Gatarayiha M. C., Understanding geographical origins of specialty coffees in Rwanda In: *PLOS One* (in review).

Mukashema, A. Veldkamp, A., Gatarayiha C., Arabica coffee yields in Rwandan smallholder systems: future perspectives of climate change effect and improved soil organic carbon management In: *Agriculture, Ecosystems and Environment* (in review).

Mukashema, A., Veldkamp, A., & Vrieling, A. (2014). Automated high resolution mapping of coffee in Rwanda using an expert Bayesian network. *International Journal of Applied Earth Observation and Geoinformation*, 33, 331-340. doi: 10.1016/j.jag.2014.05.005

Mukashema, A., Veldkamp, T., & Amer, S. (2016). Sixty percent of small coffee farms have suitable socio-economic and environmental locations in Rwanda. *Agronomy for Sustainable Development*, 36(2). doi: 10.1007/s13593-016-0363-0

Westinga, E., **Mukashema, A.**, & van Gils, H. (2013). A comparison of fine resolution census and image-based national forest inventories: a case study of Rwanda. *Forestry*, 86(4), 453-461. doi: 10.1093/forestry/cpt016

Other Refereed Journal Articles

Cyamweshi, R., Nabahungu, N., **Mukashema, A.**, Ruganzu, V., Gatarayiha, M., Nduwumuremyi, A., & Mbonigaba, J. (2014). Enhancing nutrient availability and coffee yield on acid soils of the central plateau of Southern Rwanda. *Glob J Agric Res*, 2(2), 44-554

Twarabamenye, E., **Mukashema, A.** (2012). Long-run Trend of Cities' Informal Housing: a Solution for the Majority of Kigali Urban Dwellers and a Challenge to Urban Development in Rwanda in *Journal of Environmental Management and Safety* Vol. 3, No. 2, 35 – 56

Twarabamenye, E., **Mukashema, A.** (2010). The future of City Planning in Rwanda: Contribution of the National University of Rwanda. *Journal of Building and Land Development* Vol.17 (2), 2010/19

ITC Dissertation List

https://www.itc.nl/Pub/research_programme/Research-review-and-output/PhD-Graduates