Institute of Plant Production and Agroecology in the Tropics and Subtropics
University of Hohenheim
Field: Global Food Security
Prof. Dr. Georg Cadisch (Supervisor)

# Prediction of soil properties for agricultural and environmental applications from infrared and X-ray soil spectral properties

Dissertation
Submitted in fulfillment of the requirements for the degree
"Doktor der Agrarwissenschaften"
(Dr.sc.agr. / Ph.D. in Agricultural Sciences)

to the
Faculty of Agricultural Sciences

presented by
**Towett, Erick Kibet**
Born on October 3rd 1981 in Kericho, Kenya

2013

This thesis was accepted as a doctoral dissertation in fulfillment of the requirements for the degree "Doktor der Agrarwissenschaften" by the Faculty of Agricultural Sciences at University of Hohenheim on 31/10/2013

Date of oral examination: 09/12/2013

**Examination Committee**

| | |
|---|---|
| Supervisor and Review: | Prof. Dr. Georg Cadisch |
| Co-Reviewer: | Prof. Dr. Torsten Müller |
| Additional examiners: | Prof. Dr. Karl Stahr |
| Head of the Committee: | Prof. Dr.-Ing. Stefan Böttinger. |

## Contents

## Abbreviations and acronyms

AAS – atomic absorption spectroscopy
AFSIS – Africa Soil Information Service
AGRA – Alliance for a Green Revolution in Africa
ANN – artificial neural networks
BGS – British Geological Survey
BMGF – Bill and Melinda Gates Foundation
BT – boosted trees
CEC – cation exchange capacity
CIAT – International Center for Tropical Agriculture
CV – coefficient of variability
DAAD – Deutscher Akademischer Austausch Dienst
DRIFTS – diffuse reflectance infrared Fourier transform spectroscopy
EQG – environmental quality guidelines
EthioSIS – Ethiopian Soil Information Service
FAO – Food and Agriculture Organization of the United Nations
FSC –Food Security Centre
GIS – geographic information systems
HF – Hydrofluoric acid
HWSD – Harmonized World Soil Database
ICP-AES – inductively coupled plasma – atomic emission spectroscopy
ICP-MS – inductively coupled plasma – mass spectroscopy
ICP-OES – inductively coupled plasma – optical emission spectroscopy
ICRAF – World Agroforestry Centre
IIASA – International Institute for Applied Systems Analysis
IR – infrared diffuse reflectance spectroscopy
ISO – International Organization for Standardization
ISRIC – World Soil Information (formerly International Soil Reference and Information Centre)
ISSCAS – Institute of Soil Science – Chinese Academy of Sciences
IUSS – International Union of Soil Sciences
JRC – Joint Research Centre of the European Commission
LDPSA – laser diffraction particle size anlyzer
LDSF – land degradation surveillance framework
LLD –lower limit of detection
MARS – multivariate adaptive regression splines
MARS – multivariate adaptive regression splines
MIR – mid infrared reflectance
MIRS – mid infrared reflectance spectroscopy
MLR – multiple linear regression
$Na_2O_2$ – sodium peroxide

NIR – near infrared reflectance
NIRS – near infrared reflectance spectroscopy
OOB – Out-of-bag validation
PC – principal component
PCA – principal component analysis
PLS – partial least squares
PLSR – partial least squares regression
PTF – pedotransfer functions
REML – restricted maximum likelihood
RF – random forests
RPD – root mean square error of prediction
SOC – soil organic carbon
SOTER – World Soils and Terrain Database
SQG – Soil quality guidelines
SSA– Sub-Saharan Africa
SSD – silicon drift detector
SVM – support vector machines
TXRF – total X-ray fluorescence spectroscopy
UN – United Nations
UNEP – United Nations Environment Program
UNESCO – United Nations Educational, Scientific and Cultural Organization
US.EPA – U.S. Environmental Protection Agency
VNIRS – visible near infrared reflectance spectroscopy
WD-XRF– wavelength dispersive X-ray fluorescence
WRB – World Reference Base for Soil Resources
XRD – X-ray diffraction spectroscopy
XRF – X-ray fluorescence spectroscopy
YES – Young Excellence Scholars

# CHAPTER 1

## Introduction to the thesis

## 1.0 Introduction to the thesis

### 1.1 Background and rationale

Many of today's most pressing problems facing developing countries, such as food security, climate change, and environmental protection, require large area data on soil functional capacity – the capacity of land to sustain delivery of essential ecosystem services, such as soil fertility and carbon sequestration. Livelihoods and economies in most developing countries depend critically on the ecosystem services that land provides, however, current information on land health and degradation is grossly inadequate (UNEP, 2012a). The lack of reliable data poses a fundamental bottleneck to the development of sound policies and for assessing progress towards goals throughout the developing world (UNEP, 2012a). Many Sub-Saharan Africa (SSA) landscapes are now characterized by a combination of poor soil health, poor crop health, poor water quality, and consequently contributing to poor human health and low levels of economic development (Shepherd and Walsh, 2007). African smallholder farmers are locked into poverty traps that are preventing urgently needed investments to maintain soil resources, and thus likely to result in further decline in agricultural productivity and provision of ecosystem services (AfSIS, 2012-2013; Nziguheba et al., 2010, Shepherd and Walsh, 2007).

In January 2005, the UN Millennium project released a plan on meeting the UN Millennium Development Goals by 2015 and one of the key recommendations was on soil nutrient replenishment (UN Millenium Project, 2005; Nziguheba et al., 2010; Sachs and McArthur, 2005). Another major component was the Hunger Task Force recommendation to focus on soil health as an essential part of the synergistic intervention to fight malnutrition (Sanchez and Swaminathan 2005) and to increase food production. The Alliance for a Green Revolution in Africa (AGRA) was launched in 2007, together with major programs in improved soil health with the overall vision to eliminate hunger and poverty in SSA (Sanchez et al., 2009a; Nziguheba et al., 2010). The first step towards this vision of AGRA is increased crop yields through rapid, sustainable agricultural growth based on smallholders, followed by a multisector approach that exploits the synergies among improved crop production, nutrition, health, and education (AGRA, 2013; Nziguheba et al., 2010). Achieving this major vision and other future plans will require reliable up-to-date information about soil health. However, existing gaps in knowledge about the condition and trend of SSA soils is highly fragmented hence the urgent need for accurate, up-to-date, geo-referenced soil information that will provide the basis for a sound decision-making in the implementation of soil management strategies for Africa and other core investments in infrastructure for development such as soil fertility.

## 1.2 Overview of soils in African context

The capacity of soils to deliver key ecosystem services - provisioning and regulating character - largely depends on the underlying soil properties which result from soil formation and management (which aims at changing soil properties for improving the soil's capacity to deliver services). Information on soil properties and how to manage them is of key importance for improving the soil's services delivery capacity and has been subject to large efforts of soil research and soil mapping (Leenars, 2013). In Sub-Saharan Africa, soil research started in the late 1880s with an initial focus on soil fertility for commodity crops for export and from the 1950's onwards, food crops received research attention (Leenars, 2013). Soil mapping started started in 1920's but very few countries were mapped prior to World War II, but since then soil survey organizations have carried out some detailed surveys and reconnaissance (Leenars, 2013). However, since the 1980's, after publication of the first soil map of the world (FAO-UNESCO, 1981) soil survey and mapping in Africa has diminished, and soil data collection carried out were sporadically in the context of soil fertility research (Leenars, 2013). The existing soil maps of the world (and in particular Africa), referred to as legacy soil data, are in large parts no longer reflecting the actual state of the soil resources.

In the context of an urgent need to combine existing regional and national updates of soil information worldwide and incorporate these with the information contained within the 1:5,000,000 scale Food and Agriculture Organization of the United Nations (FAO) – United Nations Educational, Scientific and Cultural Organization (UNESCO) Soil Map of the World (FAO, 1971-1981), the FAO, International Institute for Applied Systems Analysis (IIASA), ISRIC-World Soil Information, Institute of Soil Science – Chinese Academy of Sciences (ISSCAS) and Joint Research Centre of the European Commission (JRC) have recently completed a Harmonized World Soil Database (HWSD) based on existing global and national soil polygon maps. The HWSD contributes sound scientific knowledge for planning sustainable expansion of agricultural production to achieve food security and provides information for national and international policymakers in addressing emerging problems of land competition for food production, bio-energy demand and threats to biodiversity (FAO/IIASA/ISRIC/ISS-CAS/JRC, 2012). However reliability of the information contained in the database is variable: the parts of the database that still make use of the Soil Map of the World such as West Africa are considered less reliable, while most of the areas covered by SOTER databases are considered to have the highest reliability (Central and Southern Africa) (FAO/IIASA/ISRIC/ISSCAS/JRC, 2012).

The Map in Figure 1.1, adapted from the JRC Africa Soil atlas (Jones et al., 2013), shows the distribution of the dominant World Reference Base (WRB) Reference Soil Groups for Africa (IUSS Working Group WRB, 2006). The central, wetter part of the tropical and subtropical Africa is dominated by Ferralsols and they are associated with Acrisols but towards drier parts, Lixisols start to dominate (Jones et al., 2013). Large areas of Plinthosols occur in West Africa, while the desert regions in the north and the south are dominated by Calcisols, Leptosols, Regosols, Arenosols, and Gypsisols (Jones et al., 2013). Vertisols, Andosols and Nitosols are mostly associated with the African Rift Valley, with Vertisols mainly in Sudan and Ethiopia. Andosols are found along the Rift Valley in Eastern Africa, around Mount Cameroon, and in Madagascar (Jones et al., 2013). In the Mediterranean region, areas of Kastanozems and Phaeozems occur (Jones et al., 2013). Gleysols and Fluvisols are found throughout Africa, the latter associated with Africa’s flood plains, river fans, valleys, tidal marshes, deltas and mangroves while Solanchaks and Solanetz are mainly associated with coastal plains (Jones et al., 2013). Especially in southern Africa, Durisols occur locally. Alisols, Cambisols, Histosols, Luvisols, Planosols, Podzols and Umbrisols are reported to be scattered throughout the Africa map and to be locally important (Jones et al., 2013). Histosols are rather rare in Africa, occurring mostly in wetlands, isolated pockets in low-lying areas or depressions and in coastal regions where organic debris accumulates, and their distribution is limited by the rapid decomposition of organic material in tropical regions due to the permanently high temperatures (Jones et al., 2013). In urbanized areas and near large mines, Technosols may occur, however, most of these areas will be too small to be visible at the continental scale (Jones et al., 2013).

The continent of Africa contains all but one of the WRB Reference Soil Groups and illustrates a great soil diversity (Jones et al., 2013). Jones et al. (2013) reported that over 60% of the soil types represent hot, arid or immature soil assemblages: Arenosols (22%), Leptosols

(17%), Cambisols (11%), Calcisols (6%), Regosols (2%) and Solonchaks/ Solonetz (2%) and a further 20% or so are soils of a tropical or subtropical character: Ferralsols (10%), Plinthosols (5%), Lixisols (4%) and Nitisols (2%) while a considerable area (6%) is occupied by a further 16 reference groups that cover an area of less than 1% of the African land mass. This illustrates that a considerable number of soil types are associated with local soil-forming factors such as volcanic activity, accumulations of gypsum or silica, waterlogging, etc. (Jones et al., 2013).

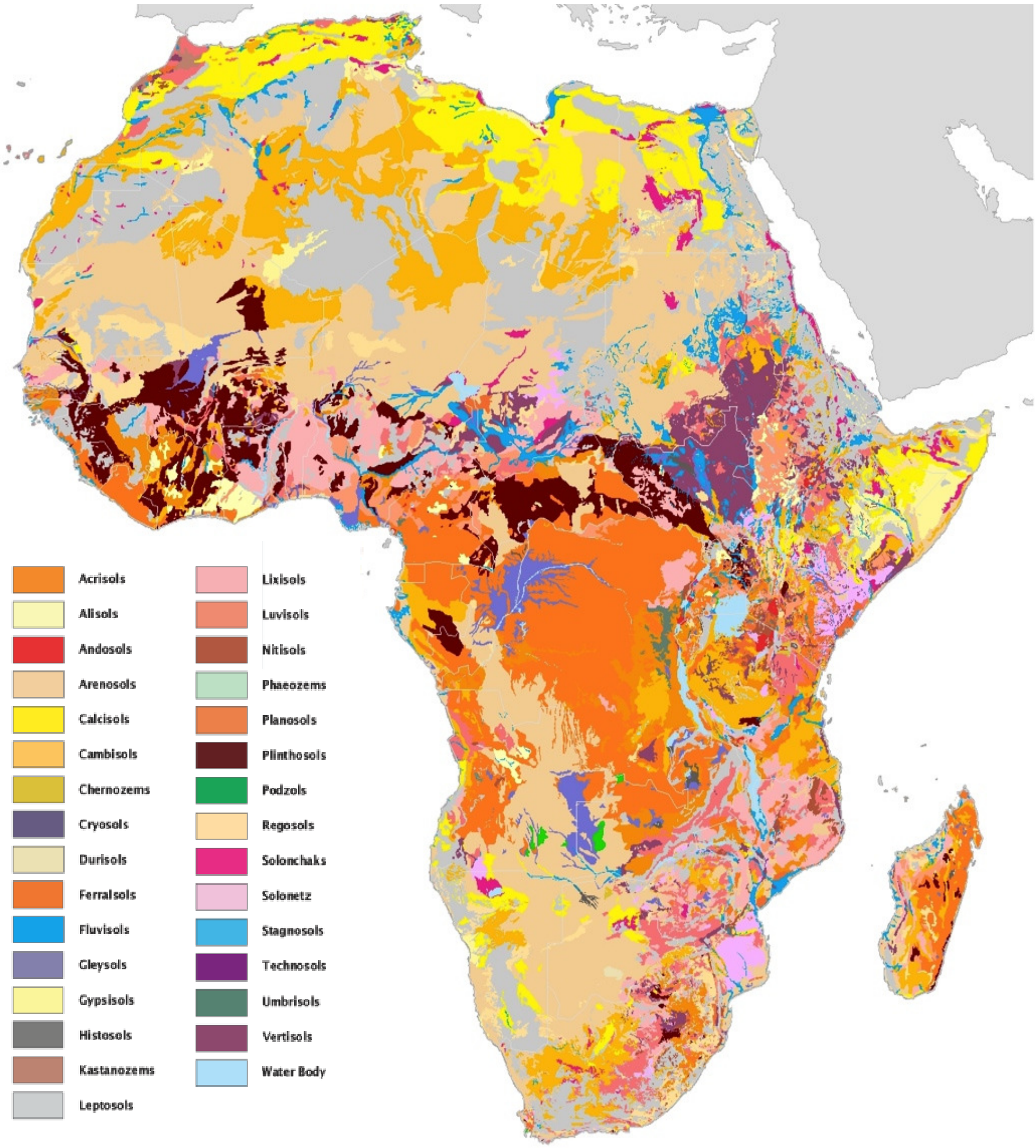


Figure 1.1: Map showing the distribution of the dominant World Reference Base (WRB) Reference Soil Groups for Africa (IUSS Working Group WRB, 2006) adapted from the JRC Africa Soil atlas (Jones et al., 2013).

### 1.3 Overview of the Africa Soil Information Service (AfSIS) project

Because knowledge about the condition and trend of African soils is highly fragmented and dated, there is an urgent need for accurate, up-to-date, and spatially referenced soil information to support agriculture in Africa. This coincides with developments in technologies that allow for accurate collection and prediction of soil properties. The Globally Integrated Africa Soil Information Service (AfSIS) is a large-scale, research-based project developing continent-wide digital soil maps for Sub-Saharan Africa using new types of soil analysis and statistical methods, the compilation and rescue of legacy soil profile data, new data collection and analysis, system development for large-scale soil mapping using remote sensing imagery and ground observations and conducting agronomic field trials in selected sentinel sites (AfSIS, 2012-2013). The AfSIS project is funded by Bill and Melinda Gates Foundation (BMGF) and Alliance for a Green Revolution in Africa (AGRA), and implemented by the International Centre for Tropical Agriculture (CIAT) in partnership with the World Agroforestry Centre (ICRAF), Earth Institute of Columbia University and ISRIC-World Soil Information Centre. The project area is proposed to include in the first Phase ~17.5 million km$^2$ of continental Sub-Saharan Africa (SSA), and 591,740 km$^2$ of Madagascar, giving a total area of ~18.1 M km$^2$, an area that encompasses more than 90% of Africa’s human population living in 42 countries (AfSIS, 2012-2013). This area excludes hot and cold desert regions based on the recently revised Köppen-Geiger climate classification (Kottek et al., 2006), as well as the non-desert areas of Northern Africa. AfSIS has just completed the process of surveying and sampling this area using a spatially stratified, random sampling approach consisting of 60, 100-km$^2$ sentinel landscapes, which are statistically representative of the variability in climate, topography and vegetation of the project area (AfSIS, 2012-2013). New data collection approaches used a hierarchical sampling approach that replicates soil and other biophysical measurements at different spatial scales, linking consistent, geo-referenced ground observations to laboratory measurements, agronomic field trials as well as remote sensing data (AfSIS, 2012-2013).

For over a decade ICRAF has been working on soil infrared spectral methods for rapid prediction of soil functional properties which are now being widely applied in land health surveillance schemes that employ a standardized protocol (the Land Degradation Surveillance Framework (LDSF)) for landscape level measurement and mapping of soil conditions. The framework is being applied throughout SSA under the AfSIS Project as well as in an increasing number of land management projects such as the Ethiopian Soil Information System (EthioSIS). The essence of the AfSIS project is that the collection and interpretation of data on soil health becomes an integral part of planning, monitoring, and impact assessment process at different scales and for different audiences in a globally consistent scientific framework (AfSIS, 2012-2013). The baseline for such a system is the LDSF which is used for soil health surveillance, defined as the ongoing, systematic collection, analysis and interpretation of data important for planning, implementation, and evaluation of soil management policy and practice, and is closely integrated with the timely dissemination and application of the data that is used in prevention and control of soil degradation (Vågen et al., 2013). The concepts used in the LDSF are similar to those used in the public health sector. The AfSIS project aims to develop a practical, timely, and cost-effective soil health surveillance service to map soil conditions, set a baseline for monitoring changes, develop global standards and methodologies, and provide options for improved soil and land management in SSA (AfSIS, 2012-2013).

A number of databases have been developed and extended to support the new data collection efforts, new techniques in laboratory analysis, and continent-wide digital soil mapping under the AfSIS project (AfSIS, 2012-2013). For example, the Africa Soil Profiles Database now contains over 12,000 geo-referenced legacy soil profile records for 37 countries (Leenars, 2013). The ICRAF-ISRIC visible near infrared (VNIR) spectral library of world soils is another valuable resource for research and applications for sensing soil quality both in the laboratory and from space and the Soil Survey (LDSF) Field Database which comprises of georefenced data and images collected using the LDFS will further ensure that sampling locations can be revisited at later points in time to quantify where specific changes have occurred (AfSIS, 2012-2013). In addition, the Soil Analytical Database provides reference laboratory measurements for one of 10 plots for all 16 clusters in each sentinel landscape, for a total of 32 reference samples (that is, top- and sub- soil from 16 plots) per sentinel landscape (AfSIS, 2012-2013). The samples from the remaining nine plots are analyzed using spectral diagnostics. The Soil Spectral Database contains both spectral measurements from the reference samples and the remaining nine samples from each cluster for all of the AfSIS sentinel landscapes (AfSIS, 2012-2013). The present study supported ICRAF's methods development for the AfSIS Project.

## 1.4 Soil health and degradation

Soil health has been defined as: “the capacity of soil to function as a living system, with ecosystem and land use boundaries, to sustain plant and animal productivity, maintain or enhance water and air quality, and promote plant and animal health. Healthy soils maintain a diverse community of soil organisms that help to control plant disease, insect and weed pests, form beneficial symbiotic associations with plant roots; recycle essential plant nutrients; improve soil structure with positive repercussions for soil water and nutrient holding capacity, and ultimately improve crop production" (FAO, 2008). Many of Africa's soils are derived from ancient granite rocks and thus they are inherently low in plant nutrients and compounding this natural deficit, nutrients leach and are taken away from the soil and fields with cultivation, with wind and water erosion, and with every harvest (AGRA, 2013).

Pressure on land resources and ecosystems has intensified greatly over the past several decades due to land use changes created by e.g. increasing population, economic development and global markets, exacerbated locally by land governance issues (UNEP, 2012b). Increased population pressure in Africa has resulted in surging demands for food and livestock feed are due to factors such as urbanization and changing diets that include more animal products (UNEP, 2012b) and consequently continuous cropping without soil conservation practices or fallow periods. This has, in turn, caused soil degradation and nutrient depletion across much of the continent. The Agricultural Production and Soil Nutrient Mining in Africa report highlights the continent's ‘soil health crisis’, revealing that three-quarters of Africa’s farmlands are severely degraded (Henao and Baanante, 2006). Restoration of soil fertility is necessary to increase crop yields and food production in order to combat the worsening food security situation in Africa, however, information about the extent and intensity of soil nutrient mining and a better understanding of its main causes are essential to the design and implementation of policy measures and investments to reverse the mining and subsequent decline in soil fertility (Henao and Baanante, 2006). In this context, rapid screening of soil properties using spectroscopic techniques should be viewed as key a contributor to the joint goals of increased agricultural production, food security, economic development, land conservation, and environmental protection.

### 1.4.1 Major nutrient constraints

Soil fertility decline is perceived to be widespread in the soils of the tropics, particularly in soils of SSA and most studies have used nutrient balances (in which fluxes and pools were estimated from published data, data derived from pedotransfer functions, or some other method) to assess the degree and extent of nutrient depletion. These approaches have created awareness but suffer methodological problems as several of the nutrient flows and stocks are not measured (Hartemink, 2006). Soil fertility decline includes nutrient depletion (larger removal than addition of nutrients), nutrient mining (large removal of nutrients and no inputs), acidification (decline in pH and/or an increase in exchangeable Al), the loss of organic matter, and an increase in toxic elements such as aluminum (Hartemink, 2006). Organic carbon is together with pH, reported to be the best simple indicator of the health status of a soil with moderate to high amounts of organic carbon associated with fertile soils with a good structure (FAO/IIASA/ISRIC/ISS-CAS/JRC, 2012).

Soil nutrient mining, the consumption of a key component of the soil's natural capital, is the result of overexploitation of agricultural land and this continued nutrient mining of soils would mean a future of even increased food insecurity and environmental damage (Henao and Baanante, 2006). Escalating rates of soil nutrient mining make nutrient losses highly variable in agricultural areas in the sub-humid and humid savannas of West and East Africa, and in the forest areas of Central Africa but in general depletion rates range from moderate, about 30 to 40 kg of nitrogen, phosphorus, and potassium (NPK)/ha yearly in the humid forests and wetlands of southern Central Africa and Sudan to more than 60 kg NPK/ha yearly in the sub-humid savannas of West Africa and the highlands and sub-humid areas of East Africa (Henao and Baanante, 2006). A review by Cobo et al. (2010) confirmed that soil nutrient mining results from 57 selected studies in Africa commonly showed most systems had negative N and K balances (i.e. 85 and 76% of studies showed negative means, respectively) while the trend for P was less severe (i.e. only 56% of studies presented means below zero). Due to low input use in Africa soil nutrient balances are often negative. The review results by Cobo et al. (2010) were genarally consistent with the claim of nutrient mining across the continent at least for N and K (e.g. Hartemink, 2006).

### 1.4.2 Total versus plant-available nutrients

A concept that has been emphasized is that soil is not a static system but the soil chemical elements are in a state of flux: there are additions, losses, crop removal, and internal cycling processes (Wong et al., 1991). These processes are often driven by temperature, rainfall, and additions of organic matter to the soil, any or all of which can be extreme in the tropical environments of SSA. The variations in soil nutrients are also derived from differences in the composition of the parent material and from fluxes of matter and energy into or from soils over geologic time or management (Helmke, 2000, Rawlins et al., 2012). Soil nutrient levels can either be described as total nutrients or plant-available nutrients and both forms give a much better indication of the nutrients that a particular soil type is likely to contribute to plants over the crop cycle.

Total nutrient content is not a satisfactory index for measuring nutrient availability due to the different and complex distribution patterns of the elements among various chemical species or phases in soils (Chen et al., 1996). In addition, the total levels of nutrients in the soil are of less interest from an agronomic viewpoint, as they are often poorly correlated with plant-availability. Another reason is also because not all of the total nutrients in the soil are immediately available for use by plants and microorganisms. It is thus desirable to determine bioavailable nutrients from the total content and this is achived by chemical extraction methods (single extraction or sequential extraction). There are reports of the use correlation coefficients between extractable nutrients and total contents as a criterion for bioavailability (Chen et al., 1996). The higher the correlation coefficient is, the more suitable the extraction method should be, however, different extractants differ in their reaction modes and there is a great variation of the amount of nutrients extracted, meaning that the various extractable fractions could differ largely in available nutrients (Chen et al., 1996). In addition, element bioavailability is regulated by many factors, such as soil chemical and physical properties and plant types (Chen et al., 1996). The efficiency of a production system depends on the importance of crop uptake versus the total supply of nutrients and high losses of nutrients limit the efficiency. Differences among soils are difficult to distinguish and thus the correlation method is not capable of investigating the specific nature of the bioavailabilty of elements in a particular soil (Chen et al., 1996). However, there is a relationship between total and available nutrients for some elements since plant-available element composition is related to what is totally available and to those of parent materials. Thus, there is a need to assess the concentration of total elemental composition in the soil since it influences the availability of a range of essential and potentially toxic elements, which consequently has implications for their uptake by crops (Rawlins et al., 2003).

## 1.5 Implications for food security

Monitoring the dynamics of the total and plant-available nutrients would promote their efficient use by crops and prolong the productive life of the soils (Wong et al., 1991). Low soil fertility is one of the major factors responsible for depressed yields on small-scale farms across Africa and for Africa's low agricultural productivity relative to other regions. Thus, African countries today face the challenges of increasing agricultural production with scarce overall resources and raising productivity in a way that conserves the natural resource base and prevents further degradation that has characterized Africa soils for generations (Henao and Baanante, 2006). In addition, information on the variation in soil chemistry, distinct from soil fertility, at different sites is desperately needed for, e.g. planning land use and management, and in spite of intensive land use, such information concerning element concentrations of Africa soils is still scarce. Thus, this study forms part of SSA-wide efforts through the AfSIS Project to provide accurate, up-to-date information about Africa soil condition to support policy and action on food security, production, regulation and supporting ecosystem services. The soil information will be essential to increase land productivity and food production, arrest hunger and ecosystem degradation, and to adapt to climate change in Africa. AfSIS will provide options for improved soil and land management through the dissemination of data on soil functional properties that will benefit farmer communities, public and private extension services, national agricultural research and soil survey organizations, the fertilizer sector, project and local planners, national and regional policymakers, and scientists (AfSIS, 2013).

## 1.6 Guidance approaches for agricultural and environmental management

In Sub Saharan Africa, data availability and quality are far from optimal and thus are important constraints on the potential to carry out agricultural and environmental health management. Presently, environmental quality guidelines (EQGs) are not available for tropical SSA soils. The available EQGs have been developed for soil element concentration values in various continents other than Africa in attempts to determine and predict concentrations above which effects occur and below which effects do not occur (Chapman et al., 2003), but these values vary by jurisdiction, land use and by proponent. A growing recognition of the need for reliable environmental and health data is emerging in many countries, while the development of remote sensing technologies is greatly increasing the potential for environmental survey and monitoring (Briggs, 2000). Because SSA is yet able to establish comprehensive systems of environmental health mapping, opportunity to develop prototype EQGs systems does exist in many areas (Briggs, 2000). In addition, because problems of inconsistencies and uncertainties in diagnosis could occur, considerable effort is needed in capturing suitably georeferenced element concentration data (Briggs, 2000). Thus, considerable scope does exist to obtain relevant data, at least in some parts of SSA, and the possibility of developing routine systems for data collection is improving thanks to new rapid methods for soil assessments reviewed below.

It should be noted that the role of EQGs in environmental quality assessments should be restricted to assisting in determining whether element concentrations pose relatively low or very high potential for significant toxicity to animal and plants (Chapman et al., 2003). Thus, the assessment of soil quality for naturally occurring elements in SSA must take into consideration regional variations in background concentrations, which strongly depend on geological and biological characteristics, as well as recent management in natural environment. This is because much of the variation in the concentration of major and trace elements in the soil is accounted for by the parent material from which the soil formed (Rawlins et al., 2012), or by non-anthropogenic sources, including weathering, volcanic, and hydrothermal activities (Chapman et al., 2003).

## 1.7 Promising methods for rapid soil assessment

Conventional assessments (methods and measurements) of soil capacity to perform specific agricultural and environmental functions are time consuming and resource intensive, limiting the use of large number of samples (Shepherd and Walsh, 2002). Dense sampling is often required to adequately characterize spatial variability in an area (Shepherd and Walsh, 2002) and, in addition, repeatability, reproducibility and accuracy of conventional soil analytical data are major challenges. New, rapid methods to quantify soil properties are needed to support national soil health surveillance systems especially in African developing countries where reliable data on soil properties is sparse, and to take advantage of new opportunities for digital soil mapping. Spectroscopic techniques, some of which are reviewed below, have shown promise as rapid highly reproducible methods of characterizing soils.

### 1.7.1 Infrared spectroscopy (IR)

Infrared diffuse reflectance spectroscopy (IR) has already shown promise as a rapid analytical tool with the use of visible (vis), near-infrared (NIR) and mid-infrared (MIR) diffuse reflectance Fourier transform spectroscopy (DRIFTS) in soil analyses having received much attention with an exponential increase in publications over the last 20 years (Guerrero et al., 2010). The increase in the potential for soil analysis is attributed to the large amount of information that the spectra hold, as well as recent advances in computation, instrument manufacturing, developments in multivariate statistics (chemometrics) and the great number of potential applications in soil science, including: soil colour, organic and mineral composition of soil and the amount of water present (hydration, hygroscopic, and free pore water), nutrient retention capacity, iron form and amount, carbonates, soluble salts, and aggregate and particle size distribution simultaneously (Guerrero et al., 2010; Shepherd, 2010).

Although soil scientists have investigated reflectance spectroscopy for several decades, the technology has not been widely taken up and routinely applied in soil studies in the African context. Thus, a roundtable of experts, which met in Nairobi in February 2006, proposed an approach that aims to provide reliable data on the condition of the soil resource base and degradation trends; application of the latest scientific and technological advances, including remote sensing and geographic information systems (GIS) as well as IR for rapid soil analysis (Swift and Shepherd, 2007). These techniques are now being applied to a new digital soil map of the world (Sanchez et al., 2009b). The ability to rapidly characterize large numbers of samples with IR opens up possibilities for soil evaluations that consider uncertainty in predictions and interpretations of soil properties. However, IR has some limitations in that it cannot predict extractable P and K well (Ludwig et al., 2002; Malley et al., 2009), which in addition to N are the main limiting nutrients in African soils. There are also uncertainties over how many samples are needed to provide robust global calibrations given the extreme variability in soil characteristics (Brown et al., 2006), especially in parent material. In addition, calibrations have to be adjusted for different soil types (Shepherd, 2010). Therefore, other new high-throughput spectral techniques using laser and X-ray technology could be valuable tools to supplement IR and stabilize calibrations across different soil types.

### 1.7.2 Total X-Ray Fluorescence Spectroscopy (TXRF)

One alternative spectral technique using X-ray technology that could be a valuable tool to supplement IR is total X-ray fluorescence spectroscopy (TXRF). The main principle of this technique is that atoms, when irradiated with X-rays, emit secondary X-rays – the fluorescence radiation. In TXRF, an X-ray beam is directed onto the sample at a very small angle, less than the critical angle of total reflection for X-rays thus causing a total reflection of the beam's photons after touching the sample prepared as thin film on a sample support. Since the wavelength and energy of the fluorescence radiation is specific for each element, TXRF analysis is possible because the concentration of each element can be calculated using the intensity of fluorescence radiation (Bruker, 2007; Towett et al., 2013). Advantages of the technique include minimal sample preparation, and low matrix interference, removing the need for external calibration (Stosnach, 2005). Standardization is internal and only requires addition of an element that is not present in the sample for quantification purposes. X-rays diffracted off soil samples are used to simultaneously quantify most of the elements present from sodium to uranium in the periodic table (Stosnach, 2005, Towett et al., 2013). TXRF can also be used as a versatile technique to investigate heavy metal pollution in soils (Stosnach, 2005) as well as trace elements in soil-water extracts (Shepherd, 2010). Lower limits of detection (LLD) are in the parts per million-concentration range for suspended soil and parts per billion levels in soil-water. There are possibilities to correlate extractable nutrient analysis with total element analysis and also to measure element concentrations in soil extracts. The total element concentration spectra can be used to capture key mineralogical differences in soils and as an input to pedotransfer functions. Thus TXRF could provide a powerful complement to IR, especially for predicting nutrient supply capacity, which is most important when considering soils as the substrates for plant growth and subsequent benefit to human health.

The total concentration of different elements in the soil has implications for human, plant and animal health, e.g., soil geochemistry influences the availability of a range of essential and potentially toxic elements which has implications for their uptake by grazing animals and crops (Rawlins et al., 2003). Therefore, understanding the nature of the key variables explaining diversity of total and water extractable concentrations of elements in the soil using TXRF and relating these to the mineralogy of parent rock can help to determine whether, and the extent to which, soil may have been contaminated by anthropogenic activities and thereby contribute to the protection of environmental health (Rawlins et al., 2003; Voortman, 2011). TXRF could also provide a particularly useful tool for prediction of soil properties in data sparse regions, especially in Africa where variations in soil mineralogy and nutrient balance critically determine vegetation composition and agricultural potential (Voortman, 2011). The presence of different vegetation types is reported to be a reliable indicator of differences in soil chemical properties and various properties of the exchange complex, micronutrient levels, and interactions among plant nutrients, significantly explain differences in vegetation and also the distribution of vegetation types (Voortman, 2011).

### 1.7.3 X-Ray Diffraction (XRD)

Soil mineralogy is a key determinant of many soil functions, for example nutrient quantities and intensities, pH and buffering, anion and cation exchange capacity, aggregate stability, soil carbon protection, dispersion, and resistance to erosion. These properties in turn determine soil agricultural and environmental qualities. However, soil mineralogy is not routinely used to predict soil functional properties because of the expense and nature of available instrumentation (e.g. large equipment with high power and cooler requirements). The recent availability of XRD technology and improved software for mineral identification and quantification could enable routine analysis of soil mineralogy as well as for prediction of soil properties (Shepherd, 2010). Results of X-ray diffraction analysis can be used to identify the main crystalline phases present in soil samples (Manhães et al., 2002). X-ray diffraction data could potentially be used to stabilize IR calibrations across soil types and as an input into pedotransfer functions. Because the environmental regulatory framework relies on total element concentration to delineate environmental contamination, XRD has the potential to identify in situ contaminant speciation, not achievable via other conventional chemical analyses, and thus could constitute an analytical step toward the reliable prediction of contaminant geochemistry (Dermatas et al., 2007).

## 1.8 The link between variability of soil properties and soil forming factors

Most of the soils in SSA are characterized by spatial diversity and local homogeneity (Voortman et al., 2003, Voortman, 2011), and, although in this study we anticipate having some knowledge of the average differences in soil nutrient concentration among sites, it is difficult to predict, accurately, the sites where nutrient deficiencies severely limiting crop growth is likely to develop. Thus, the quantification of the variability and patterns in soil total element concentrations at finer scales will require detailed analysis of site-specific variability of soils and of the soil chemical properties using rapid analysis techniques such as TXRF, XRD and MIR. However, in order to be able to make a sound interpretation of soil spectral results using TXRF, XRD and MIR and the variation and patterns among the soil properties, it is required to establish the linkages between variability of soil properties and climate, parent material, vegetation types, land use patterns, management and, other soil-forming factors.

In Africa a large fraction of cultivated soils developed from crystalline parent material under conditions of Precambrian Basement Complex and this has important implications for soil chemistry and the type, level, and spatial diversity of the nutrient deficiencies in these soils (Voortman, 2011). In particular, Voortman (2011) argued that under the conditions of Precambrian Basement Complex of Mozambique, the ecological diversity of land types might be more closely related to mineralogy of parent rock than is commonly acknowledged. However, the findings by the author from a Mozambique case study generate new questions regarding their general applicability in a wider context of the SSA in terms of key factors, or unifying principles for African land resource ecology (Voortman, 2011).

Parent materials, topography, organisms and climate are factors of soil formation (Jenny, 1941) that could be held responsible for the variability of total element concentration. The variation of total element concentration in soils of SSA could thus be related to some relatively easily measurable site characteristics, such as parent material type, slope, mean annual temperature, rainfall, etc. Therefore, regressing the site characteristics with patterns of variation in soil total element concentration will help examine the relationships between soils properties (e.g. total element composition), parent materials, climate, vegetation composition, mineralogy and many other site factors as well as soil-forming factors. For example, clay minerals are the most reactive inorganic components of soils and since they help to determine soil properties and largely govern their behaviors and functions, regressing the soil chemistry to mineralogy will help foster the interest of plant and food production (Viscarra Rossel, 2011). In addition, the type of clay mineral dominantly present in the soil often characterizes a specific set of pedogenetic factors in which the soil has developed (FAO/IIASA/ISRIC/ISS-CAS/JRC, 2012). Although SSA soils are generally of low fertility, the question whether this is causally related to high rainfall (the effect of leaching) was investigated by Voortman (2011) who provided evidence in his study of the Miombo soil ecosystem that the inverse relationship between high rainfall and soil fertility does not hold systematically.

Soil chemistry research through total element concentration and total nutrient fixing capacity of a soil could be studied apart from soil fertility and still have implications for soil fertility. The total nutrient fixing capacity of a soil is well expressed by its CEC and soils with low CEC have little resilience and cannot build up nutrient stores (FAO/IIASA/ISRIC/ISS-CAS/JRC, 2012). Currently, we do not know precisely whether the site variation in the total element concentrations of some nutrients, such as K, would affect crop or vegetation growth. According to Voortman (2011), the macronutrients, N, P and K play a modest role only in explaining land resource ecology on a Precambrian Basement Complex. Their results indicated that soil fertility and land resource ecology on soils derived from Precambrian Basement Complex rocks is a more complex issue than only N, P and K levels in the soil. Voortman (2011) suggested that research in both agricultural and natural systems that would provide fundamental insights, should, next to P and K (total and available) consider Ca, Mg, K, Na, S, and also Al and its saturation percentage if the pH drops below 5.5. The suggested essential micronutrients to be considered also include Cu, Fe, Mn, Zn, Cr, Ni, Si, and Se. Even though knowledge on actual soil conditions is very weak, micronutrient deficiencies are most likely to be widely prevalent on the basis of the parent rock in which soils have developed (Nubé and Voortman, 2006). The authors also noted that key ecological factors primarily correspond with properties of the cation exchange complex and with micronutrients and in particular the ratios of nutrients discriminate well between different ecological conditions. Thus, because of these relationships, it has been concluded that patterns of land use and vegetation can be used as a first indicator for the nutrient status of a soil and for a preliminary identification of potential plant nutrient deficiencies, so as to enable the structuring of well-targeted agronomic research (Voortman, 2011). Thus, with data on element concentration, land use planning and management and consequently linkages to the quality of human nutrition can be made (e.g. Nubé and Voortman, 2006; Voortman, 2011).

### 1.9 Justification and opportunities

Key research gaps exist on the optimal use and combination of the spectroscopic analytical techniques in Sub-Saharan Africa soils for predicting properties such as soil macro- and micro-nutrient content and soil nutrient retention capacity. In the current study, TXRF and XRD were tested in conjunction with IR to potentially provide powerful soil functional diagnostic capabilities for the direct prediction of key soil functional properties for agricultural and environmental applications. Optimal combinations of these spectral methods for use in pedotransfer functions for low cost, rapid prediction of properties of Africa soils also need to be investigated to establish the added value or redundancy when IR is complemented with TXRF, and XRD data (Shepherd, 2010). Prediction models for soil organic carbon, soil fertility properties (soil extractable nutrients, pH and exchangeable acidity) and soil physical properties (water holding capacity, soil stability and soil texture) need to be developed using IR. Despite the importance of soil mineralogy in determining soil functional properties, there have been few attempts to link the two quantitatively. High throughput XRD performed on neat soil samples provides new opportunities to use mineralogical information in pedotransfer functions that are expensive and time-consuming to measure, and can solve specific agricultural and environmental problems. TXRF, in addition to providing chemical fingerprinting that relates to mineralogy, can also directly determine total and extractable nutrients, thus complimenting IR. These techniques thus present new opportunities to revolutionize the way in which agronomy and soil science is done, greatly enhancing the potential for providing evidence-based decision support at multiple scales. There is now real possibility to harness such developments to enable science-based diagnostic surveillance approaches to agricultural and environmental management (Shepherd and Walsh, 2007).

## 1.10 Hypotheses

The overall goal of this current study was to help develop and test rapid, timely, cost-effective, and high throughput spectroscopic methods for soil health surveillance service to analyse soil properties, soil conditions, and set a baseline for monitoring changes. The hypotheses were that:

a) There are wide variations and coherent patterns in relationships among the total element concentrations measured using TXRF attributable to differences in parent materials between sites, and to local pedologic and hydrological factors within sites or due to differences in management in Sub-Saharan Africa soils.

b) Fingerprinting of soil element composition using TXRF may form a useful basis for classifying soils in a way that relates to soil-forming factors and inherent soil properties.

c) Element fingerprinting picks up differences in mineralogy that have diagnostic value and could account for differences in relationships between spectra (e.g. TXRF) and soil test values caused by mineralogical differences.

d) There are strong relationships between soil functional properties (carbon pools, extractable nutrients, exchangeable acidity, P sorption, water holding capacity, soil stability) and patterns in TXRF, XRD, and MIR data of Sub-Saharan Africa soils.

e) There is redundancy in information among soil TXRF, XRD, and MIR data, and only a subset of these techniques (e.g. MIR plus XRD or TXRF) is needed to provide robust pedotransfer functions for predicting soil functional properties.

## 1.11 Objectives

The overall objective of this research was to develop, test, and combine different infrared and X-ray spectral diagnostic techniques for the direct prediction of key soil properties for agricultural and environmental applications particularly for Sub-Saharan Africa. The specific objectives of this study were to:

i. develop and test an improved analytical method for the direct quantification of total element concentrations in soils using the S2 PICOFOX$^{\text{TM}}$ TXRF spectrometer,

ii. test the accuracy of the TXRF using a determined instrument sensitivity curve by comparing results with those measured using total acid dissolution ICP-MS analysis (international standard method) for a range of elements,

iii. quantify the variability in total element composition of soils from a diverse set of Africa soils ,

iv. explore the patterns in total element composition of soils analysed using a principal component analysis,

v. test the relationships between total element concentrations and nutrient supply capacity by relating Mehlich-3 soil tests (acid-extractable nutrients) to total element analysis patterns in soil,

vi. examine relationships between soil element fingerprints and site characteristics including mineralogy, climate, landform, vegetation type, plant material and management (cultivation),

vii. test whether TXRF can improve MIR predictions of soil test values, especially for those variables for which MIR tends to give poor predictions (e.g. extractable P and K; some micronutrients), and

viii. examine the extent to which the TXRF technique added value to the MIRS technique for improving global predictions of soil total and extractable nutrients.

## 1.12 Outline of the study

This thesis consists of five chapters. Following the current introductory chapter, chapter 2 provides an overview of the possibility to extend the use of TXRF for the direct quantification of total element concentrations through developing a method applicable for soil analysis. In the chapter, we develop and test an improved method for the direct quantification of total element concentrations in soils using S2 PICOFOX$^{TM}$ TXRF spectrometer, and determine the instrument sensitivity curve and compared the accuracy of the TXRF results with those measured total acid dissolution ICP-MS analysis (international standard method) for a range of elements. In chapter 3 we quantified the variability and explored patterns in total element composition using total X-ray fluorescence spectroscopy of soils from 34 randomly-located 100-km$^2$ sentinel sites distributed across Sub-Saharan Africa: Ghana (3 sites), Tanzania (8 sites), Ethiopia (4 sites), Mali (3 sites), Burkina Faso (1 site), Mozambique (4 sites), Nigeria (3 sites), Zambia (1 site), Kenya (3 sites), Guinea (2 sites), and Malawi (2 site). In the same chapter, we also explore the relationships between total element concentrations and nutrient supply capacity by relating Mehlich-3 soil tests (acid-extractable nutrients) to total element analysis patterns in soil and, in addition, examine relationships between element fingerprints and site characteristics including mineralogy, climate, landform, vegetation type, plant material and management (cultivation). The fourth chapter explores the possibility of combining mid infrared reflectance spectroscopy (MIRS) and total X-ray fluorescence spectroscopy (TXRF) for the prediction of soil total nutrient concentrations and soil texture. Finally, the fifth chapter reviews the answers provided to the research questions and considers the innovative aspects of the findings. The fifth chapter also discusses the methodological recommendations for future research in soil science, environmental and agricultural applications using spectroscopy, and present implications for food security in Sub-Saharan Africa.

### 1.13 References

1. AfSIS, 2012-2013. Africa Soil Information Service (AfSIS). Internet: cited 2013 June 17]. Available from:www.africasoils.net/
2. AGRA, 2013. Alliance for a Green Revolution in Africa (AGRA). Internet: Cited 2013 June 17]. Available from: http://www.agra.org/.
3. Awiti, A.O., Walsh, M.G., Shepherd, K.D., and Kinyamario, J., 2007. Soil condition classification using infrared spectroscopy: A proposition for assessment of soil condition along a tropical forest-cropland chronosequence. Geoderma 143 (1-2): 73-84. doi:10.1016/j.geoderma.2007.08.021.
4. Bilgili, A.V., van Es, H.M., Akbas, F., Durak, A., and Hively, W.D., 2010. Visible-near infrared reflectance spectroscopy for assessment of soil properties in a semi-arid area of Turkey. Journal of Arid Environments 74: 229–238. doi:10.1016/j.jaridenv.2009.08.011.
5. Briggs, D.J., 2000. Environmental Health Hazard Mapping for Africa. WHO-AFRO, Harare, Zimbabwe. [Internet]. [cited 2013 April 12]. [140 p.]. Available from: www.who.int/hac/techguidance/tools/healthhazard.pdf.
6. Brown, D.J., Shepherd, K.D., Walsh, M.G., Mays, D.M., and Reinsch, T.G., 2006. Global soil characterization with VNIR diffuse reflectance spectroscopy. Geoderma 132: 273–290.

7. Bruker, 2007. Lab report XRF 426: S2 PICOFOX total reflection X-ray fluorescence spectroscopy – working principles. Bruker AXS Microanalysis GmbH, Berlin Germany.
8. Chapman, P.M., Wang, F., Janssen, C.R., Goulet, R.R., Kamunde, C.N., 2003. Conducting Ecological Risk Assessments of Inorganic Metals and Metalloids: Current Status. Human and Ecological Risk Assessment 9 (4): 641-697.
9. Chen, B., Shan, X., and Qian, J., 1996. Bioavailability index for quantitative evaluation of plant availability of extractable soil trace elements. Plant and Soil 186:275-283.
10. Cheng, C., Laird, D.A., Mausbach, M.J., and Hurburgh, C.R., Jr., 2001. Near-infrared reflectance spectroscopy - principal components regression analyses of soil properties. Soil Science Society of American Journal 65 (2): 480-490.
11. Cobo, J.G., Dercon, G., and Cadisch, G., 2010. Nutrient balances in African land use systems across different spatial scales: A review of approaches, challenges and progress. Agriculture, Ecosystems and Environment 136: 1–15.
12. Dermatas, D., Chrysochoou, M., Pardali, S., and Grubb, D.G., 2007. Influence of X-ray diffraction sample preparation on quantitative mineralogy: implications for chromate waste treatment. Journal of Environmental Quality 36:487–497.
13. FAO, 2008. An international technical workshop. Investing in sustainable crop intensification. The case for improving soil health. Integrated Crop Management 6-2008. FAO, Rome: 22-24 July 2008.
14. FAO/IIASA/ISRIC/ISS-CAS/JRC, 2012. Harmonized World Soil Database (version 1.2). FAO, Rome, Italy and IIASA, Laxenburg, Austria. pp 50.
15. Guerrero, C., Rossel, R.A.V., and Mouazen, A.M., 2010. Special issue 'Diffuse reflectance spectroscopy in soil science and land resource assessment'. Geoderma 158:1–2.
16. Hartemink, A.E., 2006. Assessing soil fertility decline in the tropics using soil chemical data. Advances in Agronomy 89: 179–225.
17. Henao, J., and Baanante, C., 2006. Agricultural Production and Soil Nutrient Mining in Africa: Implications for Resource Conservation and Policy Development. IFDC – An International Center for Soil Fertility and Agricultural Development, Muscle Shoals, AL, U.S.A. pp 95.

18. Janik, L.J., Skjemstad, J.O., Shepherd, K.D., and Spouncer, L.R., 2007. The prediction of soil carbon fractions using mid-infrared-partial least square analysis. Australian Journal of Soil Research 45: 73–81.
19. Jenny, H., 1941. Factors of soil formation: a system of quantitative pedology. McGraw-Hill, New York.
20. Jones, A., Breuning-Madsen, H., Brossard, M., Dampha, A., Deckers, J., Dewitte, O., Gallali, T., Hallett, S., Jones, R., Kilasara, M., Le Roux, P., Micheli, E., Montanarella, L., Spaargaren, O., Thiombiano, L., Van Ranst, E., Yemefack, M., and Zougmoré, R. (eds.), 2013. Soil Atlas of Africa. European Commission, Publications Office of the European Union, Luxembourg. 176 pp.
21. Kottek, M., Grieser, J., Beck, C., Rudolf, B., Rubel, F., 2006. World Map of the Köppen-Geiger climate classification updated. Meteorologische Zeitschrift 15, 259-263. doi: 10.1127/0941-2948/2006/0130.
22. Leenars, J.G.B., 2013. Africa Soil Profile Database, Version 1.1. A compilation of georeferenced and standardized legacy soil profile data for Sub-Saharan Africa (with dataset). ISRIC Report 2013/03. Africa Soil Information Service (AfSIS) Project. ISRIC-World Soil Information, Wageningen, the Netherlands. 160 pp.
23. Ludwig, B., Khanna, P.K., Bauhus, J., and Hopmans, P., 2002. Near infrared spectroscopy of forest soils to determine chemical and biological properties related to soil sustainability. Forest Ecology and Management 71:121-132.
24. Malley, D., Akinremi, O.O., and Ige, D., 2009. Feasibility of predicting several forms of phosphorus and phosphorus retention capacity in Canadian Prairie soils by NIRS. Manitoba Soil Science Society Conference Feb.5-6, 2009. Winnipeg MB.
25. Manhães, R.S.T., Auler, L.T., Sthel, M., Alexandre, J., Massunaga, M.S.O., Carrió, J.G., dos Santos, D.R., da Silva, E.C., Garcia-Quiroz, A., and Vargas, H., 2002. Soil characterisation using X-ray diffraction, photoacoustic spectroscopy and electron paramagnetic resonance. Applied Clay Science 21:303–311.
26. Minasny, B., and McBratney, A.B., 2008. Regression rules as a tool for predicting soil properties from infrared reflectance spectroscopy. Chemometrics and Intelligent Laboratory Systems 94: 72–79.

27. Nubé, M., and Voortman, R.L., 2006. Simultaneously addressing micronutrient deficiencies in soils, crops, animal and human nutrition: opportunities for higher yields and better health. SOW-VU Working Paper (WP 06-02, 48 pp.). Centre for World Food Studies, VU University, Amsterdam.
28. Nziguheba, G., Palm, C.A., Berhe, T., Denning, G., Dicko, A., Diouf, O., Diru, W., Flor, R., Frimpong, F., Harawa, R., Kaya, B., Manumbu, E., McArthur, J., Mutuo, P., Ndiaye, M., Niang, A., Nkhoma, P., Nyadzi, G., Sachs, J., Sullivan, C., Teklu, G., Tobe, L., and Sanchez, P.A., 2010. The African Green Revolution: Results from the Millennium Villages Project. Advances in Agronomy 109:75-115.
29. Rawlins, B.G., Webster, R., and Lister, T.R., 2003. The influence of parent material on top soil geochemistry in eastern England. Earth Surface Processes and Landforms 28:1389–1409.
30. Rawlins, B.G., McGrath, S.P., Scheib, A.J., Breward, N., Cave, M., Lister, T.R., Ingham, M., Gowing, C., and Carter, S., 2012. The advanced soil geochemical atlas of England and Wales. British Geological Survey, Keyworth. www.bgs.ac.uk/gbase/advsoilatlasEW.html.
31. Sachs, J.D., and McArthur, J.W., 2005. The Millennium Project: a plan for meeting the Millennium Development Goals. Lancet 365: 347–53. Published online http://image.thelancet.com/extras/04art12121web.pdf.
32. Sanchez, P.A., and Swaminathan, M.S., 2005. Hunger in Africa: the link between unhealthy people and unhealthy soils. Lancet 365: 442–44.
33. Sanchez, P.A., Denning, G.L., and Nziguheba, G., 2009a. The African Green Revolution moves forward. Food Security 1:37-44. doi: 10.1007/s12571-009-0011-5.
34. Sanchez, P.A., Ahamed, S., Carré, F., Hartemink, A.E., Hempel, J., Huising, J., Lagacherie, P., McBratney, A.B., McKenzie, N.J., Mendonça-Santos, M.L., Minasny, B., Montanarella, L., Okoth, P., Palm, C.A., Sachs, J.D., Shepherd, K.D., Vågen, T., Vanlauwe, B., Walsh, M.G., Winowiecki, L.A., and Zhang, G., 2009b. Digital Soil Map of the World. Science 325 (5941): 680-681. doi:10.1126/science.1175084
35. Shepherd, K.D., and Walsh, M.G., 2002. Development of reflectance spectral libraries for characterization of soil properties. Soil Science Society of America Journal 66:988-998.

36. Shepherd, K.D., and Walsh, M.G., 2004. Diffuse reflectance spectroscopy for rapid soil analysis. In: R Lal (ed) Encyclopedia of Soil Science. Marcel Dekker Inc, New York. Online Published by Marcel Dekker: 04/26/2004.
37. Shepherd, K.D., and Walsh, M.G., 2007. Infrared spectroscopy – enabling an evidence-based diagnostic surveillance approach to agricultural and environmental management in developing countries. Journal of Near Infrared Spectroscopy 15:1–20.
38. Shepherd, K.D., 2010. Soil spectral diagnostics – infrared, X-ray and laser diffraction spectroscopy for rapid soil characterization in the Africa Soil Information Service. 19th World Congress of Soil Science, Soil Solutions for a Changing World 1–6 August 2010, Brisbane, Australia.
39. Stosnach, H., 2005. Environmental trace-element analysis using a benchtop total reflection X-ray fluorescence spectrometer. The Japan Society of Analytical Chemistry. Analytical Sciences 21:873-876.
40. Summers, D., Lewis, M., Ostendorf, B., Chittleborough, D., 2011. Visible near-infrared reflectance spectroscopy as a predictive indicator of soil properties. Ecological Indicators 11:123–131.
41. Swift, M.J., and Shepherd, K.D., (Eds), 2007. Saving Africa's Soils: Science and Technology for Improved Soil Management in Africa. Nairobi: World Agroforestry Centre. pp 28.
42. Terhoeven-Urselmans, T., Spaargaren, O., Vagen, T-G., and Shepherd, K.D., 2010. Prediction of Soil Fertility Properties from a Globally Distributed Soil Mid-Infrared Spectral Library. Soil Science Society of America Journal 74:1792–1799. doi:10.2136/sssaj2009.0218.
43. Towett, E.K., Shepherd, K.D., and Cadisch, G., 2013. Quantification of total element concentrations in soils using total X-ray fluorescence spectroscopy (TXRF). Science of the Total Environment 463–464: 374–388. http://dx.doi.org/10.1016/j.scitotenv.2013.05.068.
44. UN Millennium Project, 2005. Investing in Development: A Practical Plan to Achieve the Millennium Development Goals. New York. pp 356.
45. UNEP, 2012a. Land Health Surveillance: An Evidence-Based Approach to Land Ecosystem Management. Illustrated with a Case Study in the West Africa Sahel. United Nations Environment Programme, Nairobi. pp 211.

46. UNEP, 2012b. Environmental Accounting of National Economic Systems: An Analysis of West African Dryland Countries within a Global Context. United Nations Environment Programme, Nairobi. pp 139.
47. Vågen T-G., Winowiecki, L.A., Tondoh, J.E., Desta, L.T., 2013. Africa Soil Information Service (AfSIS) - Soil Health Mapping. http://hdl.handle.net/1902.1/19793 V2 [Version].
48. Viscarra Rossel, R.A., 2011. Fine resolution multiscale mapping of clay minerals in Australian soils measured with near infrared spectra. Journal of Geophysical Research 16: F04023. doi: 10.1029/2011JF001977.
49. Voortman, R.L., Sonneveld, B.G.J.S., and Keyzer, M.A., 2003. African land ecology: opportunities and constraints for agricultural development. AMBIO: A Journal of the Human Environment 32(5):367-373.
50. Voortman, R.L., 2011. Explorations into African land resource ecology on the chemistry between soils, plants and fertilizers. PhD Dissertation. Faculty of Economics and Business, Vrije Universiteit. pp. 264.
51. Wong, M.T.F., Wild, A., and Mokwunye, A.U., 1991. Overcoming soil nutrient constraints to crop production in West Africa: Importance of fertilizers and priorities in soil fertility research. Fertilizer Research 29: 45-54.

# Chapter 2

# Quantification of total element concentrations in soils using total X-ray fluorescence spectroscopy (TXRF)

# 2.0 Quantification of total element concentrations in soils using total X-ray fluorescence spectroscopy (TXRF)[1]

Erick K. Towett[1,2], Keith D. Shepherd[2], and Georg Cadisch[1]

[1]Institute for Plant Production and Agroecology in the Tropics and Subtropics, Universität Hohenheim, Garbenstr. 13, 70599 Stuttgart, Germany.
[2]World Agroforestry Centre (ICRAF), P.O. Box 30677, Nairobi, 00100, Kenya.
*Corresponding author at:* World Agroforestry Centre (ICRAF), P.O. Box 30677, Nairobi, 00100, Kenya.
Tel: +254207224163; Fax: +254207224000
*E-mail address: k.shepherd@cgiar.org*

## 2.1 Abstract

Total X-ray fluorescence spectroscopy (TXRF) determines concentrations of major and trace elements in multiple media. We developed and tested a method for the use of TXRF for direct quantification of total element concentrations in soils using a S2 PICOFOX$^{TM}$ spectrometer (Bruker AXS Microanalysis GmbH, Germany). We selected 15 contrasting soil samples from across Sub-Saharan Africa for element analysis to calibrate the instrument against concentrations determined using the inductively coupled plasma - mass spectroscopy (ICP-MS) standard method. A consistent underestimation of element concentrations using TXRF compared to ICP-MS reference analysis occurred, indicating spectrometer recalibration was required. Single-element recalibration improved the TXRF spectrometer's sensitivity curve. Subsequent analysis revealed that TXRF determined total element concentrations of Al, K, Ti, V, Cr, Mn, Fe, Ni, Cu, Zn, and Ga accurately (model efficacy/slope close to 1:1 line, and $R^2 > 0.80$) over a wide range of soil samples. Other elements that could be estimated with an acceptable precision ($R^2 > 0.60$) compared with ICP-MS although generally somewhat under- or overestimated were P, Ca, As, Rb, Sr, Y, Pr, Ta and Pb. Even after recalibration, compared to ICP-MS the TXRF spectrometer produced underestimations for elements Na, Mg, Ba, Ce, Hf, La, Nd, W and Sm and overestimations for elements Bi, Tl and Zr. We validated the degree of accuracy of the TXRF analytical method after recalibration using an independent set of 20 soil samples. We also tested the accuracy of the analysis using 2 multi-element standards as well as the method repeatability on replicate samples. The resulting total element concentration repeatability for all

---

[1] This chapter has been reprinted from:

Erick K. Towett, Keith D. Shepherd, Georg Cadisch. 2013. Quantification of total element concentrations in soils using total X-ray fluorescence spectroscopy (TXRF). *Science of the Total Environment*, 463–464: 374–388. http://dx.doi.org/10.1016/j.scitotenv.2013.05.068. Copyright © 2013, Elsevier.

elements analysed were within 10% coefficient of variability after the instrument recalibration except for Cd and Tl. Our findings demonstrate that TXRF could be used as a rapid screening tool for total element concentrations in soils assuming sufficient calibration measures are followed.

## 2.2 Introduction

Soil is a critical natural resource that plays a key role in determining human well-being, providing key ecosystem services, supporting food production, and the natural recycling of C and essential nutrients in the environment (Smith et al., 2009). Although soils are recognized to be critically important, our knowledge of the concentration of naturally-occurring elements in soils is limited (Smith et al., 2009). Many Sub-Saharan Africa landscapes are now characterized by a combination of poor soil and crop health, low water quality, and consequently contributing to poor human nutrition and low levels of economic development (Swift and Shepherd, 2007). Shepherd and Walsh (2007) noted the urgent need to develop rapid screening tools to characterize soils in support of agricultural development in developing countries, which are robust enough to be used under basic conditions and to be applied at scale.

Spectroscopic techniques have shown promise as rapid and highly reproducible methods of characterizing soil properties. Near- and mid infrared diffuse reflectance techniques have recently been successfully applied to characterize chemical properties of a range of soils (e.g. Janik and Skjemstad, 1995; Madari et al., 2006; Minasny and McBratney, 2008; Minasny et al., 2009; Kamau-Rewe et al., 2011; Shepherd and Walsh, 2002; 2004; and Viscarra Rossel et al., 2006), and have been employed at larger scale in combination with enhanced geostatistics (Cobo et al., 2010). However, their applicability is limited to a small range of elements and matrix interferences often require extensive calibration for different environments. X-ray fluorescence (XRF) spectrometry, a non-consumptive technique for multi-element determination, has been used to assess total elemental concentrations in soil and is applicable to a wide range of matrix types. Apart from demanding sample preparation requirements (e.g. pressed pellets), a major disadvantage of conventional XRF has been the poor elemental sensitivity. Compromised element sensitivity is, in large part, a consequence of high background noise levels, resulting from instrumental geometries and sample matrix effects (Mukhtar and Haswell, 1991). Total X-ray fluorescence spectroscopy (TXRF) was designed to reduce matrix effects by changing geometries. Sample preparation requirements are also reduced with TXRF compared with XRF.

In TXRF, a monochromatic X-ray beam is directed onto the sample at a very small angle (<0.1º), less than the critical angle of external total reflection for X-rays. This causes a total reflection of the beam's photons after touching the sample and thus the beam has a minimum interaction with the sample support over which a few microlitres/micrograms of the sample are deposited (Dhara and Misra, 2011; Stosnach, 2007). This unique geometry makes TXRF an advanced variant of conventional XRF and leads to excitation of the sample by incoming as well as totally reflected beams (Dhara and Misra, 2011). By illuminating the sample with a totally reflected beam, the absorption as well as the scattering of the beam in the sample matrix is reduced (Stosnach, 2007). The resulting benefits are a greatly reduced background noise, and consequently much higher sensitivities and a significant reduction of matrix effects eliminating the need for external calibration (Bruker, 2007a; Stosnach, 2005, 2007). In addition, the preparation of samples as a thin layer largely precludes matrix effects such as absorption or secondary excitation (Marguí, et al., 2010; Stosnach, 2005) and consequently it has been proposed that a single element can be used as internal standard for all elements (Dhara and Misra, 2011). These features, i.e., sample excitation by incoming and totally reflected beams and lower background noise, also result in comparatively better detection limits in TXRF than in XRF by several orders of magnitude (Dhara and Misra, 2011). Moreover, since the sample is deposited on a limited area of the TXRF sample support, sample contamination can be controlled (Dhara and Misra, 2011). The TXRF technique has been applied for elemental analysis for both the quantitative and qualitative determination of elements with $Z > 11$ ($Z$ = atomic number). Since the wavelength and energy of the fluorescence radiation is specific for each element, TXRF analysis is possible because the concentration of each element can be calculated using the intensity of fluorescence radiation (Bruker, 2007a). The radiation is labeled K-, L- or M-radiation depending on the ionized shell of an atom, and an additional indication e.g. $K_{\alpha 1}$, $K_{\beta 1}$, $L_{\alpha 1}$ etc. is made depending on the refilling shell (Bruker, 2007b). The fluorescence radiation emitted by the sample is then determined by an energy-dispersive detector and the intensity is measured by means of an amplifier coupled to a multichannel analyzer (Bruker, 2007a; Stosnach, 2007).

TXRF is efficient and fast and requires only minimum specimen quantities (Bennun and Sanhueza, 2009; Klockenkämper, 1997). It has been employed in a variety of disciplines, including biology, physics, biomedicine, chemistry, archaeology, medicine, and geology (Bennun and Sanhueza, 2009; West et al., 2011). TXRF analysis of soils using the S2 PICOFOX$^{TM}$ TXRF spectrometer (Bruker AXS Microanalysis GmbH, Germany) has been proposed to allow rapid and simultaneous determination of the concentrations of many elements from Na to U in the periodic table (Shepherd, 2010; Stosnach, 2005). However, relatively few papers exist on direct quantification of total element concentrations in soil samples using the TXRF method. Marguí et al. (2010) and Rousseau (2001) studied several rapid and simple analytical approaches that reduced matrix effects, improved detection limits and explored new sample preparation methods for total selenium determination in soils using TXRF. Although the S2 PICOFOX TXRF instrument comes factory calibrated, it is necessary to do quality control to ensure a high quality of the analytical results, particularly for soils with their varying matrix effects.

When testing the accuracy of the quantification, it is insufficient to assess an individual instrument parameter, but rather the performance of the complete analytical procedure, from sample preparation to analytical result. The accuracy is tested by means of multi-element standards with known concentrations: upon spectra evaluation, the nominal values are compared with the measured values, and in the case of a deviation of more than 10% an error analysis must be performed (Bruker, 2007b). One of the error sources that can be a reason for the deviation between nominal and measured values is error in the TXRF calibration. The remedy for calibration error is the comparative measurement with other reference standards and recalibration if necessary. The quality of the results obtained using the TXRF technique is greatly impacted by the accuracy of the instrument sensitivity curve as well as the results of the standards used to calibrate the instrument (Bennun and Sanhueza, 2010). The accuracy and detection limit of the technique are also affected by instrumental, sample and data processing errors (Mori and Uemura, 1999). Instrumental error factors originate from source X-ray stability, accuracy of glancing angle, position accuracy of sample stage and spurious peaks interferences, while the sample error factors arise from lateral and depth distribution of the analyte, surface roughness and diffraction of primary X-rays (Mori and Uemura, 1999).

The objectives of this study were to (i) develop and test an improved analytical method for the direct quantification of total element concentrations in soils using the S2 PICOFOX$^{TM}$ TXRF spectrometer, and (ii) using a determined instrument sensitivity curve, test the ability to calibrate/validate against total acid dissolution ICP-MS analysis (international standard method) for a range of elements. Soil samples from a wide range of soils in the Sub-Saharan Africa tropics and textures as well as elemental ranges were used to develop a generalizable approach.

## 2.3 Materials and Methods

The S2 PICOFOX spectrometer used in this study is a portable benchtop TXRF instrument, featuring an air-cooled low power X-ray metal-ceramic tube with a Molybdenum target, working at 50 W of max power, and a liquid nitrogen-free Silicon Drift Detector (SSD) (Bruker, 2007b). A summary of the technical specifications of the instrument is given in Table 2.1. The compact S2 PICOFOX TXRF system design allows even mobile use for measurements on-site or in the field because it is independent of any cooling media (Bruker, 2007b). Further benefits of the TXRF system are the absence of a vacuum chamber, the absence of matrix or memory effects and the ability for multi-element analysis (Bennun and Sanhueza, 2010). We purposely selected 15 representative soil samples with a wide range of properties and elemental concentrations to calibrate the S2 PICOFOX for total element analysis in soils and an additional 20 soil samples as an independent validation set. These samples were analysed by both the TXRF technique and the conventional standard method of total acid dissolution inductively coupled plasma – mass spectroscopy (ICP-MS) sodium peroxide ($Na_2O_2$) based fusion.

Table 2.1: Technical specifications of the S2 PICOFOX TXRF spectrometer.

| | |
|---|---|
| X-ray tube | 50 kV, 1 mA, Mo target |
| Element range | Na to U |
| Optics | Multilayer monochromator, 17.5 keV |
| Detector, area, resolution | Silicon drift, 10 mm$^2$, <160 eV |
| Carrier | Quartz or other material, 30 mm diameter |
| Sample station | Cassette for 25 discs |
| Control | PC, data transfer via serial interface |
| Size, weight | 590 x 450 x 300 mm, 37 kg |
| Power consumption | Max 150 W |
| Voltage, frequency | 100-230V ± 10%; 50-60 Hz |
| Manufacturer | Bruker AXS Microanalysis GmbH |

### 2.3.1 Sample selection and preparation

The 15 soil samples, for use in the testing of the applicability, accuracy and recalibration of the TXRF method, were selected from a set of 100 reference soils collected from various locations across Africa. They were obtained from the World Agroforestry Centre (ICRAF) archive Africa soil library. Selection of the 15 samples was based on principal components analysis (PCA) of the following variables from reference analytical data from both ICRAF and Crop Nutrition Services Laboratories in Nairobi, Kenya: Sand, Silt, Clay, Fe, Ca, pH and C. After corresponding transformation to achieve close to statistical normality Euclidian distances were calculated based on the number of principal components recommended by the PCA and thereafter percentiles were calculated and ranked in an ascending order. The 15 samples were randomly selected from each of the quartile ranges, with 4 samples from each quartile range except the first quartile range where 3 samples were selected, using all the variables to ensure that we had a wide variation in possible matrix influences from within the set of 100 samples. An independent validation set of 20 soil samples was selected based on the Kennard-Stone sample selection algorithm (Kennard and Stone, 1969) using a PCA of TXRF total element concentration data from a set of 603 samples associated with the Africa Soil information Service (AfSIS) project (www.africasoils.net), taken from nineteen 100-km$^2$ stratified random sites across Africa: Ghana (3 sites), Tanzania (8), Congo (2), Mali (2), Burkina Faso (1), Malawi (1) and Mozambique (2) (AfSIS, 2011). The soil samples were air-dried and passed through a 2-mm sieve before being oven-dried at 40 ºC overnight. A sub-sample 5 g of each soil sample was ground to a fine powder <200 µm using a Retsch RM 200 mill (Retsch GmbH, Haan, Germany) and thereafter, 3 g of each soil sub-sample was further milled to <50 µm using a micronising mill (McCrone, Westmont, U.S.A.).

### 2.3.2 Cleaning and preparation of TXRF sample carriers

The following tasks connected to cleaning of TXRF quartz glass sample carriers were done under a laboratory fume hood to avoid contamination: (i) mechanically wiping each sample carrier with a fluff-free tissue paper soaked in acetone (Panreac Quimica, Spain), (ii) mounting the sample carriers onto a washing cassette then transferring the cassette into a 1000 ml glass beaker containing hot alkaline cleaning solution (RBS 50 (Chemical Products S.A., Belgium) diluted to ratio 5:50, RBS:double distilled water), (iii) rinsing the sample carriers with deionized double distilled water, then transferring the washing cassettes into 1000 ml glass beaker filled with 10% nitric acid (Sigma-Aldrich, Germany) before heating for 2 hours on a hot plate previously set to 80 °C in a chemical fume chamber, (iv) thoroughly rinsing the sample carriers with deionized double distilled water before transferring into a 1000 ml glass beaker containing hot (80 °C) distilled water then rinsing the sample carriers with ultra-pure water prepared using a Centra R60, Purelab Flex 2 (Elga, UK), (v) drying the sample carriers in an oven set at 80 °C for 30 minutes before carefully wiping the sample carriers with tissue paper soaked in acetone, and finally (vi) adding 10 µl silicon solution (Serva Electrophoresis, Germany) at the center of the sample carriers and drying the sample carriers in an oven at 80 °C for another 30 minutes to leave a surface residue of silicon that helped homogenise the soil suspension for TXRF analysis.

### 2.3.3 TXRF measurements

We tested a method for the direct quantification of total element concentration in soils using TXRF and for this an amount of approximately 50 mg of each finely ground soil sub-sample (3 g) was mixed with 2.5 ml of Triton X-100 (Fischer Scientific, UK) solution (0.1 vol. %) to form a soil suspension and then spiked with 40 µl of 1000 mg $l^{-1}$ Selenium (Fluka Analytical, Germany) as the internal standard. Selenium, an ubiquitous natural constituent in soil, originating from parent materials in the earth's crust is widely variable but on a worldwide basis, Kabata-Pendias and Mukherjee (2007) estimated an average Se content of 0.3 mg $kg^{-1}$ in surface soils with levels typically ranging from 0.05 to 3.5 mg $kg^{-1}$. Even though there are enriched Se concentrations (up to above 100 mg $kg^{-1}$) in soils derived from volcanic dust (Kabata-Pendias and Mukherjee, 2007) data specific to Se concentration in Africa soils were determined for the set of 100 reference soils collected from various locations across Africa by TXRF at the ICRAF Spectral Diagnostics Laboratory as ranging from 0.01 to 0.3 mg $kg^{-1}$. We also determined concentrations of different elements in the set of 100 reference soils, but found virtually all elements in some soils and compared with elements Sc (2.5 - 5.8 mg $kg^{-1}$) and Y (5.7 to 79 mg $kg^{-1}$) recommended initially by the instrument manufacturers as internal standards, Se was thus appropriate for use as internal standard in TXRF analysis of soil due to the low natural background concentration. However, for future studies, if the soil/sediment samples have extraordinarily high Se concentrations the application of Se is not ideal, and we recommend that a quick qualitative scan be applied and if high Se-signals are detected, a different internal standard (e.g. Bismuth) can be applied.

Triton X-100, an organic compound, applied for TXRF sample preparation, enhances the homogeneity of samples (Stosnach, 2005). The solution was placed into an ultrasonic water bath at room temperature and sonicated in continuous mode for 15 minutes. The solution was then thoroughly homogenised using a digital shaker and 10 µl of the turbid soil solution immediately dispensed on to a clean siliconised quartz glass sample carrier, ensuring a homogenous sample. Each soil sample, prepared as a thin film on the quartz disk, was then dried for 10-15 minutes on a hot plate (Staurt® SD300) set at 52 ºC and placed in a clean laminar flow hood. The final samples for TXRF analysis contained thin, circular soil films on the quartz glass sample carriers. Samples were then measured with the spectrometer in triplicate, and the relative abundance of each line was determined, as depicted in Figure 2.1. The data acquisition time was 1000 s per sample, thus requiring about 7 hours for analysis of each batch of 25 samples. Each batch of samples included 1 gain correction mono-element standard sample, which, in our case was Gallium, used for spectra correction of certain detector artefacts such as background and escape peaks. A drift in the spectroscopic amplification was compensated or reset by means of a gain correction, a process in which a correction value is transferred to the spectroscopic amplifier after a duplicate measurement of a known fluorescence peak (Bruker, 2007b). The relative abundance of intensities of the different elements, as depicted in Figure 2.1, are processed by referring to one in particular (the internal standard) (Bennun and Sanhueza, 2010). The interpretation of the TXRF spectra and data evaluation was performed using the software program SPECTRA 6.3 (Bruker AXS Microanalysis GmbH).

### 2.3.4 Pile-up peak correction

Pile-up peaks (also called sum-peaks) occur in samples with high concentrations when two X-ray photons hit the TXRF detector at the same time causing the detection of a peak at double energy of the photon (Bruker, Personal communications). Three aqueous multi-element standards were used for pile-up peak corrections for some elements e.g. Fe, Ca, and K, including the following: (i) Multi-element solution “general”: 1 mg $l^{-1}$ each of Ca, Ti, V, Cr, Mn, Fe, Co, Ni, Cu, Zn, As, Se and Sr; (ii) Multi-element solution “light elements”: 100 mg $l^{-1}$ Mg and 10 mg $l^{-1}$ each of Al, P, S, K and Ca; as well as (iii) Multi-element solution “heavy elements”: 0.1 mg $l^{-1}$ each of Pb, and Bi and 10 mg $l^{-1}$ each of Pd, Cd and Sb. Pile-up peak corrections were done manually using the SPECTRA software utilizing the data from multi-element standards for the elements e.g. Fe, Ca and K. The X-ray fluorescence lines of the individual elements were stored in the software in the form of an atomic library and later identification of the elements was done by an interactive comparison of the observed spectra lines and measured spectrum (Figure 2.1). In order to allocate the correct intensities to the individual elements in the spectrum, a deconvolution routine (SuperBayes) of the SPECTRA software was applied which used measured mono-element profiles for the evaluation of peak intensities (Bruker, 2007b).

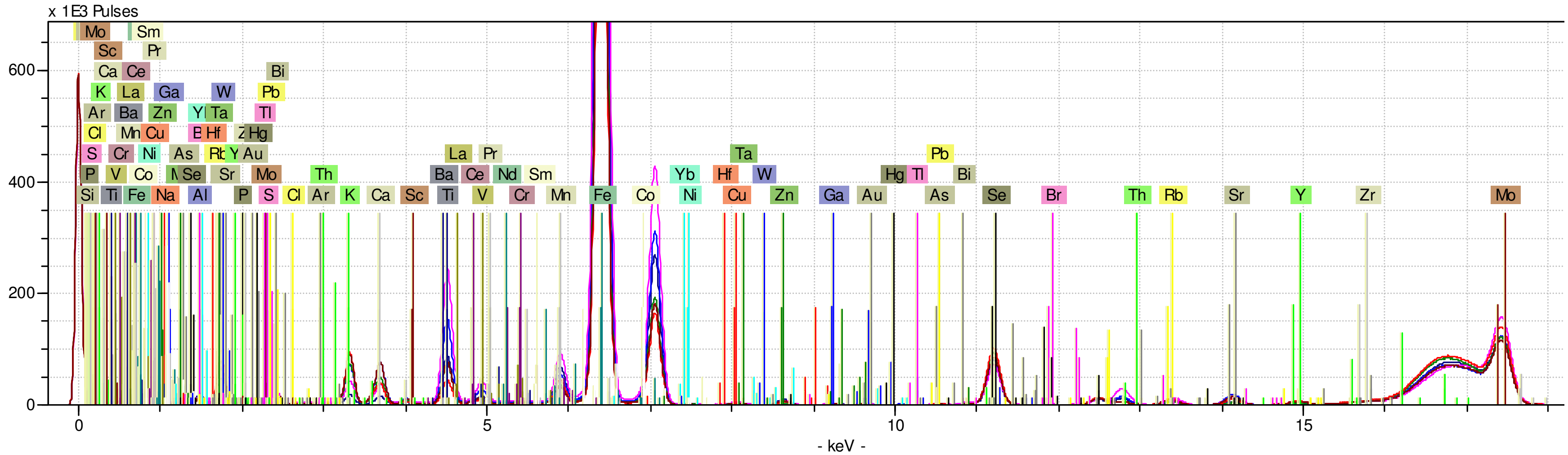


Figure 2.1: TXRF spectra of six of the calibration soil samples measured with the S2 PICOFOX spectrometer showing peaks and the abundance of each line determined for all elements.

### 2.3.5 ICP-MS reference measurements

The selected soil samples were sent to the Analytical Geochemistry Laboratory, British Geological Survey (BGS), UK, for conventional total acid dissolution ICP-MS analysis. Two further sample preparation and analysis routines were used by the BGS laboratory prior to analysis of the set of 15 soil samples (BGS, Documented In-House Methods, Personal communications):

(i) Fusion of 0.25 g of sample with 1 g of sodium peroxide ($Na_2O_2$) flux (used as the standard method in this study); and

(ii) Digestion of 0.25 g of sample with a mixture of 2.5 ml HF, 1 ml $HClO_4$ and 2.5 ml $HNO_3$ concentrated acids.

The independent validation set of 20 soil samples was prepared and analysed using only the first routine, i.e. fusion of 0.25 g of sample with 1 g of sodium peroxide ($Na_2O_2$) flux.

### 2.3.6 Evaluation of TXRF method precision

One of the TXRF method's most important source of errors that must be minimized is variation due to sample preparation (Rousseau, 2001), thus all samples and reference materials were prepared and presented to the instrument in an identical and reproducible manner. In slurry sampling, only a small amount of sample can be assayed, so samples must be very homogenous to ensure representative results (Sánchez-Moreno et al., 2010). In our method, homogenous soil sample suspensions were obtained by fine grinding of the sample and shaking the mixture with an ultrasonic bath for about 30 minutes. Materials with low density are prone to float or agglomerate while other materials present a hydrophobic behavior, so a dispersing agent, Triton X-100 was used in the slurry media to get a homogenous suspension. In addition, since the internal standard was added to the solution in a well-defined concentration, the uncertainty of the variations of the different depositions of the sample on the sample holder were avoided. In order to provide a quantitative estimate of the analytical precision for the TXRF analysis for each element, the coefficient of variability (CV) based on triplicate analysis of each sample was calculated.

### 2.3.7 TXRF method recalibration and determination of accuracy

A consistent systematic underestimation of all element concentrations using actual (non-adjusted) TXRF analysis compared to reference analysis methods (ICP-MS) was observed with the exception of Na, Co, Zr, La, Sm, and Bi which were overestimated, indicating spectrometer recalibration was required (Annex Table A2.1). The single-element recalibration method was used, involving the calculation of the calibration value for each element utilizing TXRF count statistics and concentration data of the 15 soil samples. For this, the Excel spreadsheet array function "linest", which is a linear least squares curve fitting routine that fitted a line in the form y=mx+b with a non-zero intercept and produced uncertainty estimates for the fit values, was utilized. The recalibration of the TXRF spectrometer then followed a sequence that determined the relative intensity of different elements in the measured spectrum, giving rise to a sensitivity curve as depicted in Fig. 2.2. The adjusted sensitivity values varied systematically with atomic number as would be expected. For the K-lines the element range goes from Na to Zr and for L-lines from Pd to U in the periodic table (Figure 2.2). The relative element sensitivities are always just valid for a particular system configuration (Bruker 2007a), and thus we suspect that changes of the excitation version or detector as was the case in our TXRF system led to new element sensitivities that were significantly different from the original ones (Figure 2.2). Recalibration was also necessary primarily because the prepared suspension samples did not fulfill the “no matrix approach” as ideal TXRF samples with a thickness of < 100 µm and thus, we have adsorption effects in the TXRF system which decrease with decreasing fluorescence energy (Stosnach, 2013, personal communications). Another minor reason was (especially for Fe) that the concentration range in the soil samples was by orders of magnitude higher than the original calibration range (Stosnach, 2013, personal communications).

The independent evaluation of the accuracy of the quantification results after recalibration was tested using the set of 20 independent soil samples. In addition, further evaluation of accuracy was done using Merck (Merck KGaA, Darmstadt, Germany) and Bernd Kraft (Bernd Kraft GmbH, Duisburg, Germany) multi-element standards, with known concentrations of 100 mg $l^{-1}$ and 10 mg $l^{-1}$ simultaneously, by comparing the nominal values with the measured values. All standard reference samples for investigations were prepared and presented to the spectrometer using the same experimental conditions as the soil samples investigated for this study.

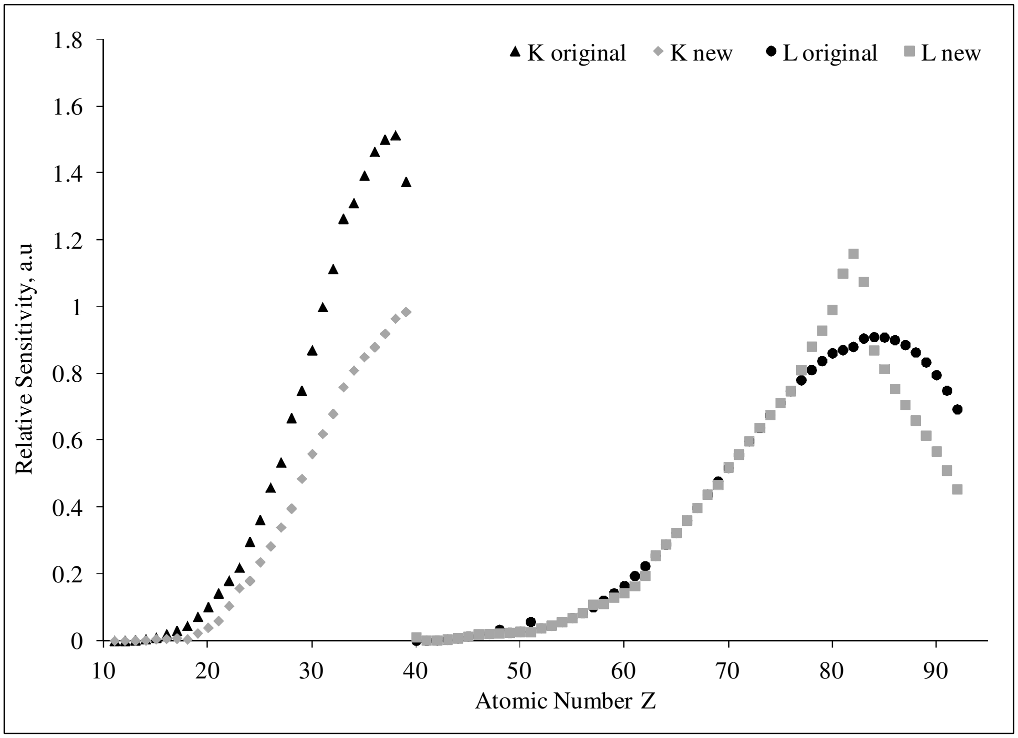


Figure 2.2: Set of relative element sensitivity curves of the TXRF spectrometer for original (▲) and new (♦) K-lines values as well as original (●) and new (■) L-lines values. One possible explanation why the primary calibration has changed significantly is due to a change in the X-ray detector of the S2 PICOFOX instrument.

## 2.4 Results and Discussion

### 2.4.1 Comparisons of the analytical results after recalibration

Recalibration was successful except for a few elements (Figure 2.3; Annex 2.1). Elements accurately determined by TXRF (model efficacy or slope close to 1:1 line, and $R^2 > 0.80$) compared to ICP-MS were Al, K, Ti, V, Cr, Mn, Fe, Ni, Cu, Zn, and Ga (Figure 2.3a). Most elements with high TXRF accuracy i.e. Al, K, V, Cr, Mn, and Ga also had a good analytical precision with CV's ranging from 6-10% (Table 2.2). Other elements estimated with an acceptable accuracy ($R^2 > 0.60$ to 0.80) although generally somewhat under- or overestimated were P, Ca, As, Rb, Sr, Y, Pr, Ta and Pb (Figure 2.3b). Compared to ICP-MS the TXRF spectrometer produced non-systematic underestimations for elements such as Na, Mg, Ba, Ce, Hf, La, Nd, W and Sm, and overestimations for elements Bi, Tl and Zr, even after recalibration (Figure 2.3c; Annex 2.1). However, the ranges in bias and differences between ICP-MS and TXRF results were reduced by the recalibration procedure used (Annex 2.1). Because of the significant improvements found, we suspect that using a larger set of calibration samples would further improve the accuracy of the instrument. The over- or underestimation for some elements that had TXRF values higher or lower than those of the ICP-MS analyses could be attributed to a number of factors during sample preparation and TXRF data processing such as sample matrix, acid dissolution methodology, and the methods reproducibility of the analyses (Lin, 2009). We speculate that TXRF data processing in a complex sample matrix such as soil was a likely cause, especially issues to do with line interference for some elements as well as concentrations close to the detection limit. We also hypothesize there were minimal problems of contamination due to working under a clean laminar flow hood chamber but there is need to confirm this in future studies e.g. contaminations from solution used during sample preparation or silicon on the sample holder.

Despite the fact that most elements could be quantified in our soil samples using TXRF, the technique appeared to be inappropriate for some elements due to other limitations. We suspect that line interferences for Na, La, Ce, Ba, Nd, Hf, Ti, W and Bi, and low concentrations close to the detection limits for Sm, made these elements difficult to analyse (Figure 2.3c, Table 2.2). We hypothesize that the conspicuously high TXRF estimates of Zr (Figure 2.3c) result from the extremely low energy of their fluorescence lines and in addition Zr has an interference with Mo making it difficult to accurately quantify this element in a spectrometer system that uses Mo X-ray tube for excitation (Figure 2.1). In further studies there is need to confirm sample (slurry) homogeneity before/during pipetting and in addition, confirm whether there are elements that are highly concentrated in minor mineral phases (e.g. Zr in Zirconia). Lanthanum and Ce were detected with difficulty in soil samples because of the line interferences with Ba, Sm and especially Ti (Figure 2.1) (Stosnach, 2010, personal communications). In the soil samples, the element Ag could not be detected using a molybdenum tube except by means of its low intensity L-line, but this line was strongly overlapped by the dominating K-lines of the matrix elements K and Ca, which prevented a quantitative determination using TRXF (Stosnach, 2005). In the particular case of Hg quantification, the procedure used (sample evaporation by drying) was unsuitable due to the high vapor pressure and low boiling point of this element that produced evaporation and loss of Hg from the surface of the reflector during the drying process. Additionally, Si was not detected due to the use of quartz glass sample carriers that give blank results for this element.

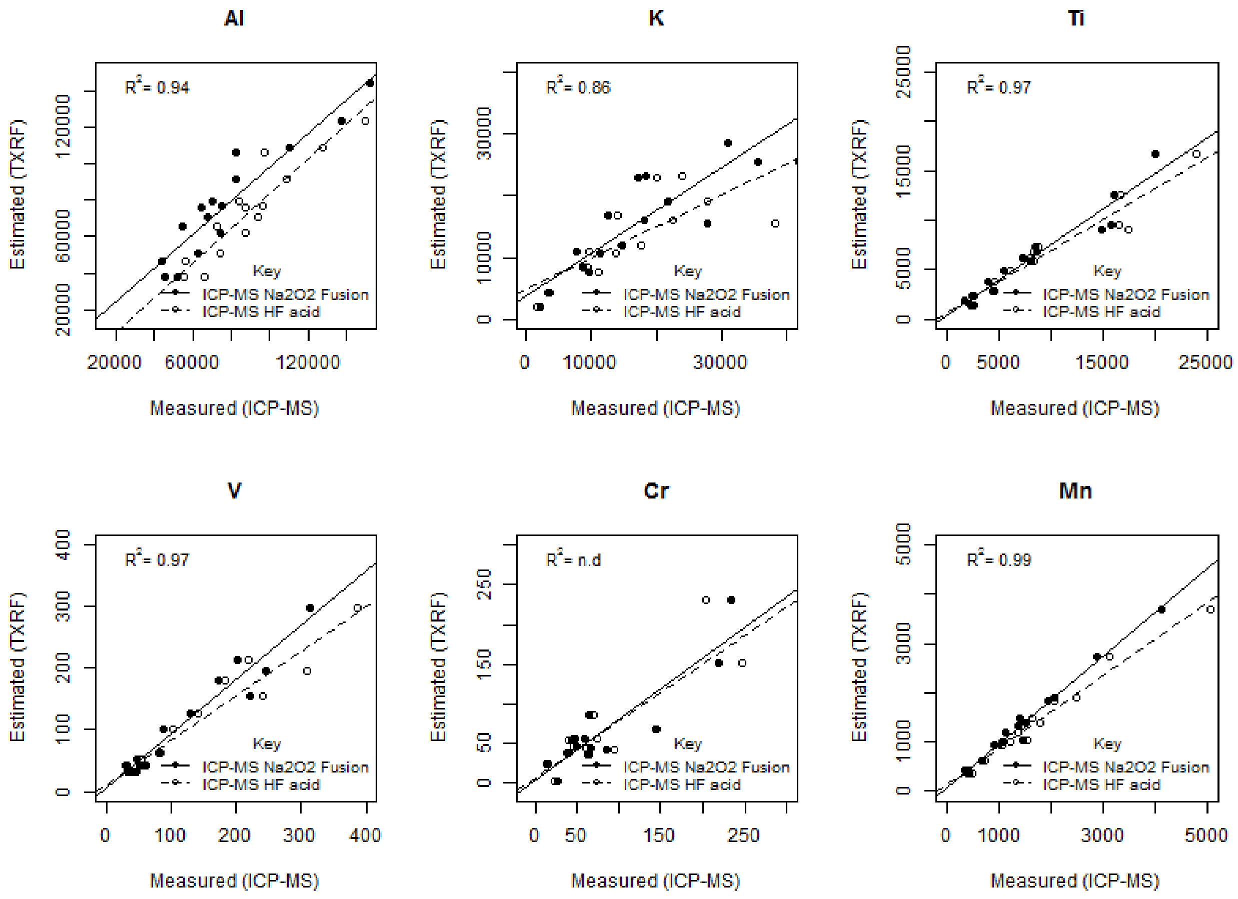


Figure 2.3a: Actual concentrations (in mg kg$^{-1}$) measured using conventional total acid dissolution ICP-MS versus concentrations estimated using TXRF for 15 soil samples after recalibration for the elements determined with high accuracy. Missing R$^2$ values were due to missing data either reference or estimated results.

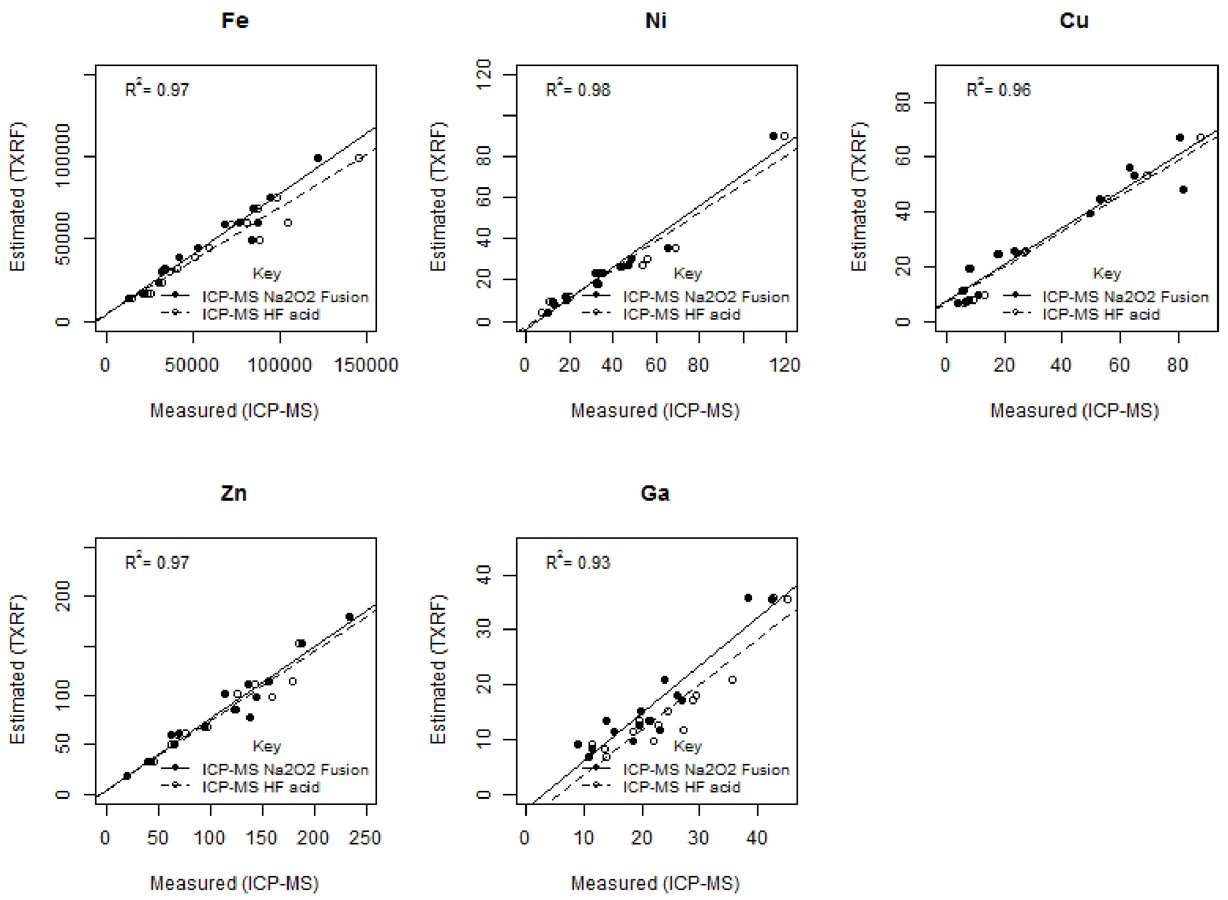


Figure 2.3a cont.: Actual concentrations (in mg kg$^{-1}$) measured using conventional total acid dissolution ICP-MS versus concentrations estimated using TXRF for 15 soil samples after recalibration for the elements determined with high accuracy. Missing R$^2$ values were due to missing data either reference or estimated results.

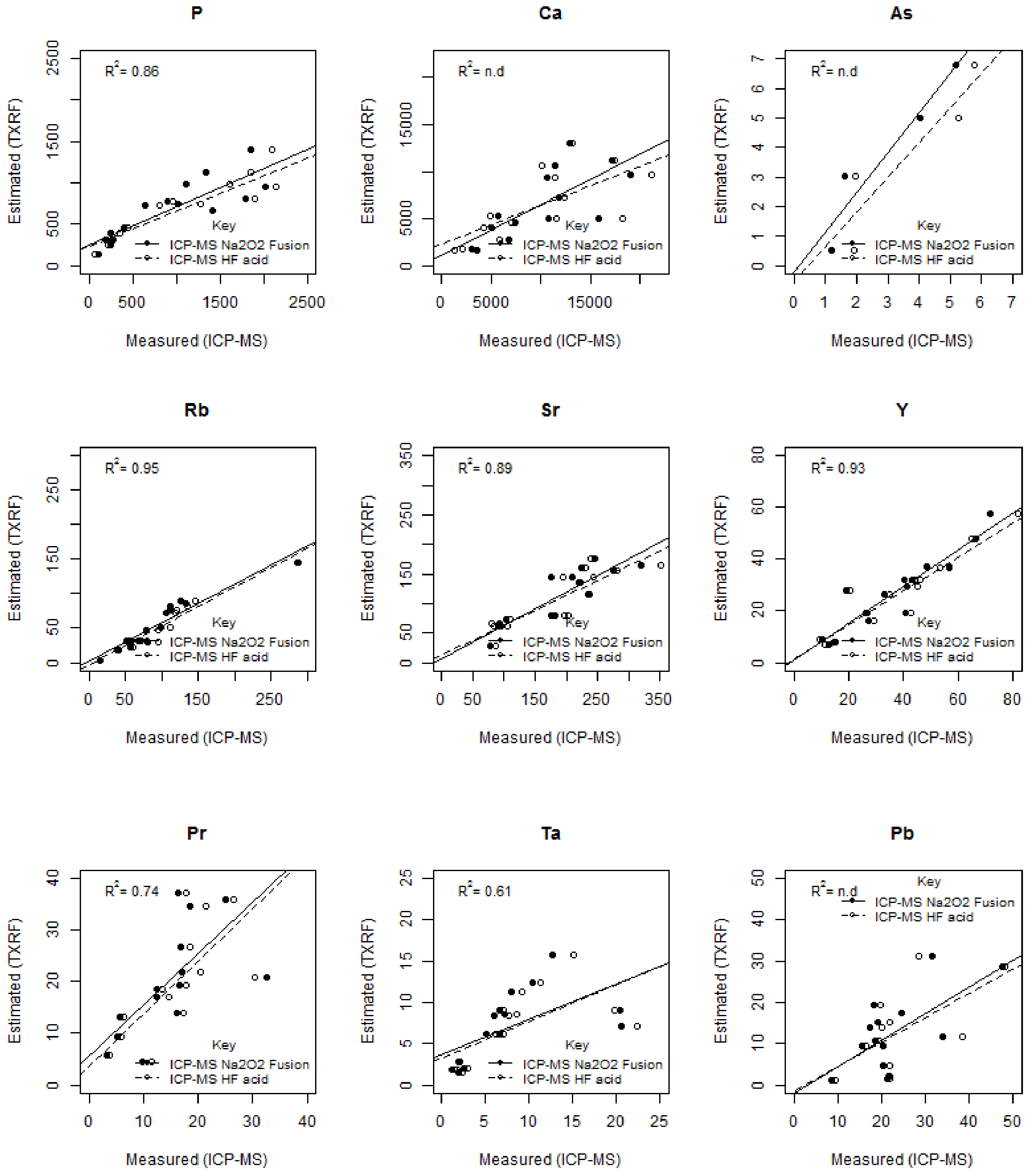


Figure 2.3b: Actual concentrations (in mg kg$^{-1}$) measured using conventional total acid dissolution ICP-MS versus concentrations estimated using TXRF for 15 soil samples after recalibration for the elements determined with acceptable accuracy. Missing R$^2$ values were due to missing data either reference or estimated results.

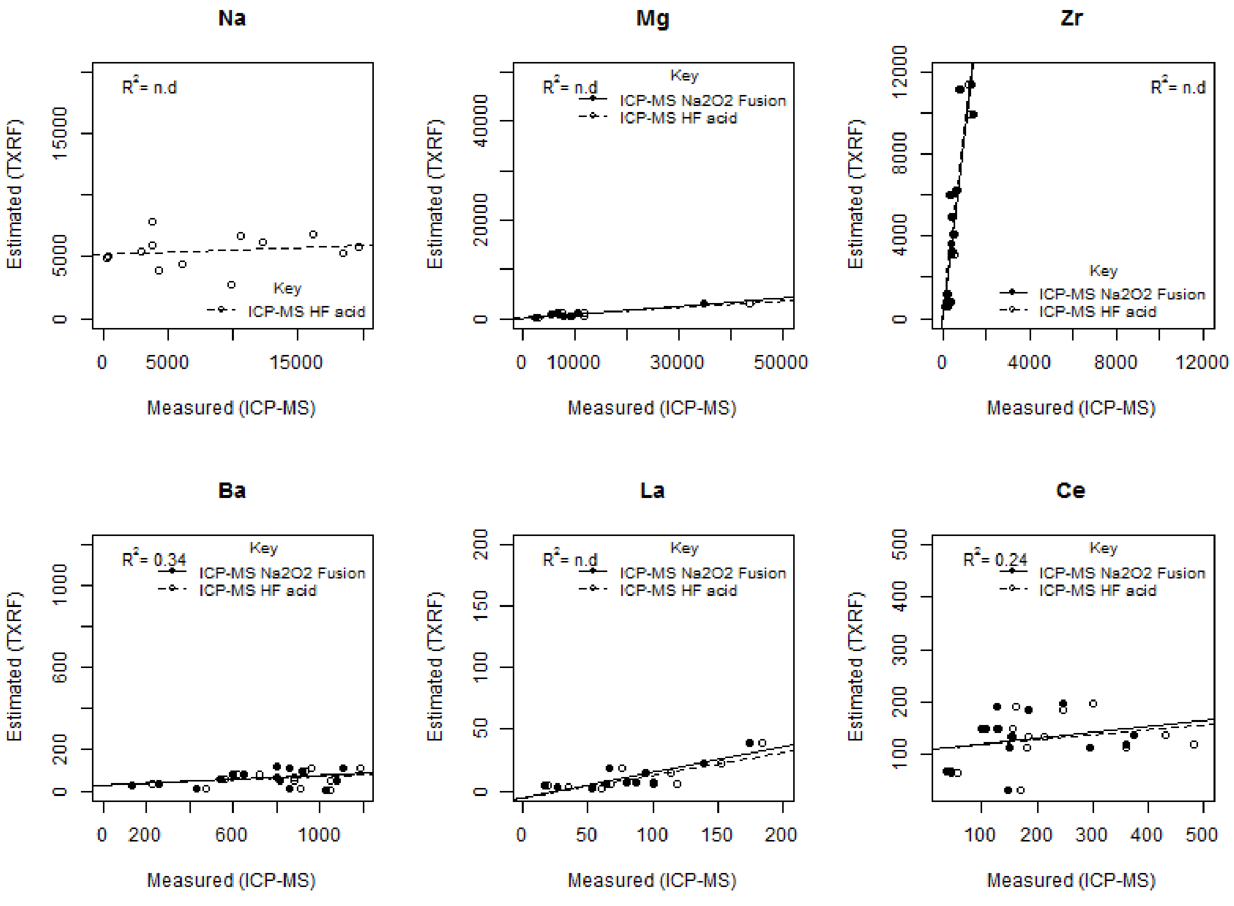


Figure 2.3c: Actual concentrations (in mg kg$^{-1}$) measured using conventional total acid dissolution ICP-MS versus concentrations estimated using TXRF for 15 soil samples after recalibration for the elements determined with poor accuracy. Missing R$^2$ values were due to missing data either reference or estimated results.

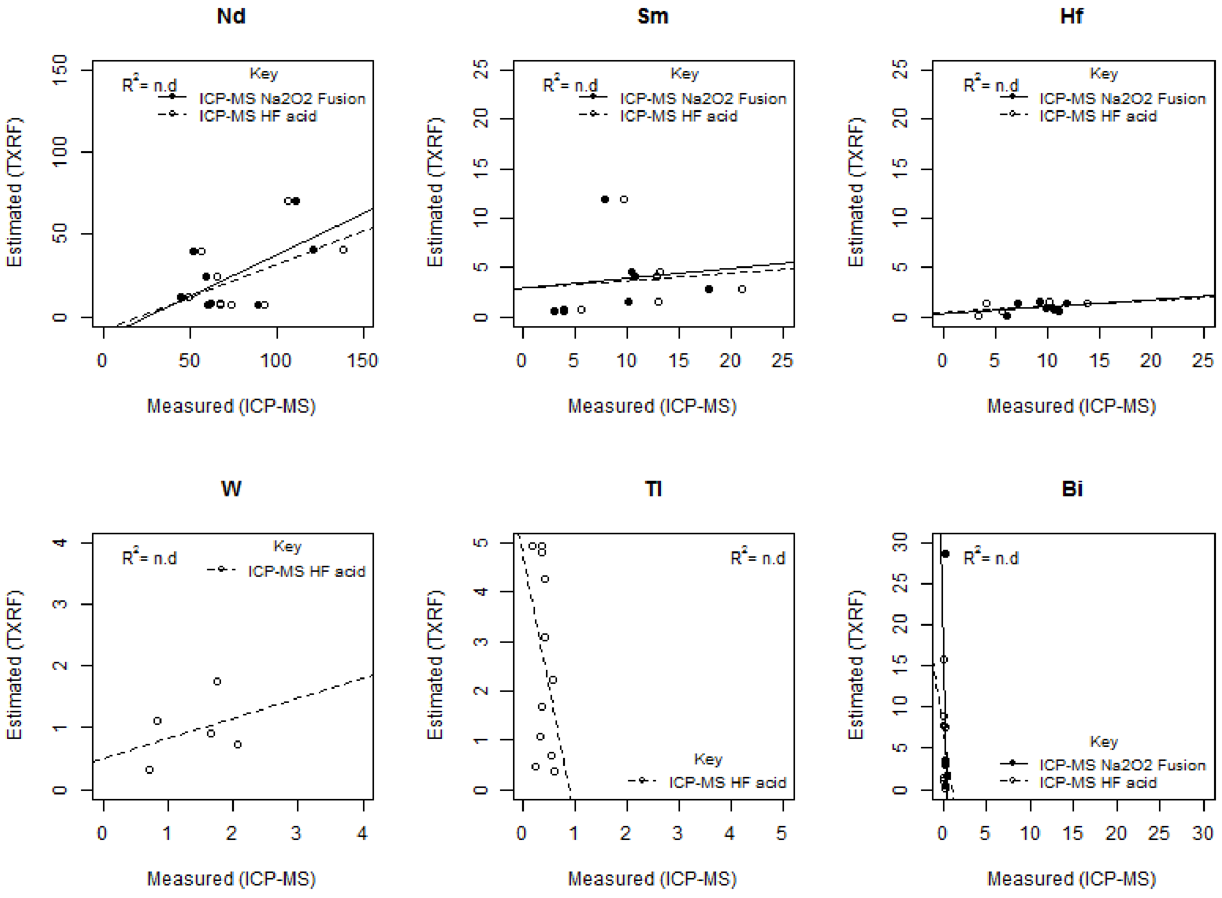


Figure 2.3c cont.: Actual concentrations (in mg kg$^{-1}$) measured using conventional total acid dissolution ICP-MS versus concentrations estimated using TXRF for 15 soil samples after recalibration for the elements determined with poor accuracy. Missing R$^2$ values were due to missing data either reference or estimated results.

Table 2.2: TXRF mean elemental concentrations (mg kg$^{-1}$), lower detection limits (LLD) (mg kg$^{-1}$) and coefficient of variability (CV %) for the calibration and validation sets. Elements are ordered according to increasing atomic number.

| | Calibration set (n=15) | | | Validation set (n=20) | | |
|---|---|---|---|---|---|---|
| Element | Mean Conc | Mean LLD | Mean CV | Mean Conc | Mean LLD | Mean CV |
| Na | 5500 | 2400 | 21 | 16000 | 4200 | 18 |
| Mg | 1100 | 450 | 9 | 17700 | 5000 | 18 |
| Al | 78000 | 300 | 12 | 40200 | 300 | 10 |
| Si | n.d. | n.d. | | n.d. | n.d. | |
| P | 700 | 35 | 23 | 700 | 50 | 19 |
| S | 200 | 30 | 11 | 1600 | 45 | 37 |
| Cl | 70 | 20 | 57 | 400 | 35 | 23 |
| K | 15000 | 10 | 6 | 16300 | 15 | 10 |
| Ca | 6000 | 4 | 5 | 69600 | 7 | 10 |
| Sc | 30 | 3 | 6 | 20 | 4 | 48 |
| Ti | 6000 | 2 | 5 | 4900 | 2 | 10 |
| V | 100 | 2 | 13 | 30 | 2 | 34 |
| Cr | 60 | 2 | 10 | 100 | 2 | 17 |
| Mn | 1400 | 1 | 6 | 1000 | 1 | 12 |
| Fe | 45900 | 1 | 6 | 37200 | 2 | 10 |
| Co | n.d. | n.d. | | 75 | 1 | 24 |
| Ni | 20 | 0 | 6 | 50 | 1 | 12 |
| Cu | 30 | 0 | 6 | 20 | 0 | 11 |
| Zn | 80 | 0 | 6 | 50 | 0 | 10 |
| Ga | 20 | 0 | 7 | 12 | 0 | 11 |
| As | 5 | 0 | | 2 | 0 | 34 |
| Se | 883 | 0 | 0 | 883 | 0 | 0 |
| Br | 5 | 0 | 6 | 20 | 0 | 11 |
| Rb | 50 | 0 | 8 | 70 | 0 | 11 |
| Sr | 110 | 0 | 7 | 200 | 0 | 10 |
| Y | 30 | 0 | 10 | 30 | 0 | 15 |
| Zr | 4800 | 317 | 14 | 170 | 25 | 19 |
| Mo | n.d. | n.d. | | 900 | 160 | 51 |
| Ba | 60 | 2 | 31 | 4700 | 22 | 27 |
| La | 110 | 2 | 7 | 1000 | 12 | 18 |
| Ce | 130 | 2 | 14 | 40 | 2 | 44 |
| Pr | 20 | 2 | 16 | 1 | 1 | 28 |
| Nd | 20 | 2 | 42 | 15 | 2 | 49 |
| Sm | 5 | 1 | 8 | 30 | 1 | 17 |
| Yb | n.d. | n.d. | | n.d. | n.d. | |
| Hf | 1 | 0 | 26 | 2 | 0 | 23 |
| Ta | 7 | 0 | | 3 | 0 | 37 |
| W | 1 | 0 | | 1 | 0 | 31 |
| Hg | 0 | 0 | 33 | 15 | 0 | 17 |
| Tl | 3 | 0 | | n.d. | n.d. | |
| Pb | 15 | 0 | 9 | 40 | 0 | 29 |
| Bi | 5 | 0 | 27 | 5 | 0 | 38 |
| Th | n.d. | n.d. | | 40 | 0 | 17 |

Key: Conc – concentration, n.d – not detected.

Since the intercept for some elements were not at zero, and based on the good linear relationship found in the regression lines of the measured (ICP-MS) vs estimated (TXRF) plots of the different elements (Figure 2.3a-c), we tested the possibility for alternative adjustments of the TXRF results using the regression equation of the ICP-MS $Na_2O_2$ fusion-based results. Here, the linear regression based equation of the ICP-MS $Na_2O_2$ fusion results (y=A+Bx), where A is the intercept and B is the slope, was thus used to perform a linear fit on our TXRF data. Recalculated TXRF results clearly showed that such recalculation has a large upward bias (overestimations) for all the elements with the exception of V, Ni and Rb (Table 2.3). It is also striking that this approach did not work for Bi (the recalculated Bi concentrations gave only negative values compared to the true measured (ICP-MS) value after recalibration) (Table 2.3). Thus this approach, has limitations, for instance the models may be over-adjusting the residuals seen in some of the individual element plots. Therefore, such correction will necessitate further investigation.

Table 2.3: Recalculated TXRF results (mg kg$^{-1}$) corrected to ICP-MS $Na_2O_2$ values using the regression equation for 15 soil samples (calibration set), and their comparison to mean nominal and estimated element concentrations. Elements are ordered according to increasing atomic number.

| | Recalculated TXRF results (corrected to ICP-MS $Na_2O_2$) | Estimated (TXRF) after recalibration procedure | Measured ICP-MS $Na_2O_2$ fusion value |
|---|---|---|---|
| Element | Mean concentration | Mean concentration | Mean concentration |
| Al | 91500 | 78400 | 78400 |
| P | 2000 | 700 | 900 |
| K | 26600 | 14900 | 16000 |
| Ca | 13300 | 6100 | 9600 |
| Ti | 9000 | 6200 | 8000 |
| V | 100 | 100 | 100 |
| Mn | 1600 | 1400 | 1400 |
| Fe | 67300 | 46000 | 57000 |
| Ni | 30 | 20 | 40 |
| Cu | 60 | 30 | 30 |
| Zn | 120 | 100 | 100 |
| Ga | 15 | | |
| Rb | 100 | 55 | 100 |
| Sr | 210 | 110 | 180 |
| Y | 40 | 30 | 30 |
| Ba | 2000 | 60 | 700 |
| Ce | 2200 | 120 | 150 |
| Pr | 30 | 20 | 20 |
| Ta | 25 | 10 | 10 |
| Bi | -1 | 5 | 0 |

### 2.4.2 Evaluation of TXRF method precision

The mean CV values for the three replications of all the soil samples, in general, indicated that the analytical precision was good (Table 2.2). An influence of ± 10% is normally acceptable but the value of the CV (in %) depends on the element (especially for the light main elements of interest (Na, P, S, and Al), there was a systematic underestimation due to adsorption effects, as samples prepared out of suspensions have no ideal thickness). The CV was <10% for many of the elements detected, with the exception of Na, Al, P, S, Cl, V, Zr, Ba, Ce, Pr, Nd, Hf, Au, Hg, and Bi in the calibration set, and Na, Mg, P, S, Cl, Sc, V, Cr, Mn, Co, Ni, As, Y, Zr, Mo, Ba, LA, Ce, Pr, Nd, Sm, Hf, Ta, W, Au, Pb and Bi in the validation set (Table 2.2). The low reproducibility values for some elements were connected to the low counting statistics caused by the low sensitivity of the TXRF instrument for light elements (Na, Al, and P), disturbances by overlaps between elements with high and comparatively low concentrations (V), elements present in concentrations close to their LLD, and also variations due to sample drying on the sample holder. Light elements are reported to be analysed inaccurately due to the absorption of their X-ray fluorescence signals in the air atmosphere of the instrument (Stosnach, 2005). The assumption that the absorption in air is more important compared to the absorption within that sample is true for conventional XRF systems with high distances between tube, sample and detector, however, in TXRF the distance between sample surface and detector is very short (1.5 – 2 mm) and the sample fluorescence is detected with high efficiency leading to a high absolute efficiency of the method (Bruker, 2007a). Thus, the adsorption in the air is negligible. The introduction of contaminants e.g. Na, P, Ca, Mg, Al, in the form of dust, during sample processing for TRXF analysis may not have had a serious consequence in our elemental analysis results due to working under a laminar flow hood, but their possible impact cannot be ruled out. But then any contaminations would have showed up in the standards. Our results showed good results with the slurry method used (Table 2.2), however, we hypothesize that the representativity of small amounts of inhomogenous sample slurry might have also had an effect. In addition, for quantitative TXRF analysis of particles, a size of <20 µm is recommended for best accuracy (Bruker, 2009), but will increase preparation costs and time (García-Heras et al., 1997), while in our case particles were below 50 µm.

### 2.4.3 Evaluation of TXRF method accuracy

TXRF recalibration resulted in accurate results for most elements of relevance to crop-livestock-human nutrition and they were estimated well in the independent validation set with the exception of a few (Table 2.2). Similar to what we found after recalibration of the instrument, the TXRF spectrometer still produced non-systematic underestimations for elements such as Al, S, V, Zr, Ce, Pr, Nd and Hf and overestimations for elements Mg, P, Co, Ni, Rb, Mo, Ba, La, Sm, W, Pb, Bi and Th compared to ICP-MS for the set of 20 independent soil samples (data not shown). Accuracy of the data was also determined by analysis of two multi-element standards, Merck and Bernd Kraft using the TXRF method. The Merck and Bernd Kraft multi-element standards had certified concentrations of 100 mg $l^{-1}$ and 10 mg $l^{-1}$ in nitric acid 0.1 mol $l^{-1}$ respectively. The nitric acid matrix is preferred because the elements are stable/soluble and also because of its oxidizing ability and the freedom from chemical and spectral interferences. Upon spectral data evaluation for the two multi-element standards, there were deviations of less than 10% for all elements when the nominal values were compared with the measured values, except for higher deviation in the case of Sr (-18%) and Tl (23%) (Table 2.4). However, these are not soil standards and hence their value is limited. One of the limitations with the TXRF technique is the determination of L-line elements due to line interferences with the strong K-lines. Therefore, Tl could be determined by their low intensity L-lines only, yet these lines were overlapped by the strong K-lines of the matrix elements K and Ca, which prevented accurate quantitative determination of the element using TRXF (Stosnach, 2005), hence the high % deviation of the measured value from the nominal value. Similarly, Sr was determined by its K-line but there could be other line overlaps that influence the accuracy of the peak deconvolution.

Table 2.4: Evaluation of accuracy: summary of comparison between TXRF measured and reference values (mg $l^{-1}$) for the Merck and Bernd Kraft multi-element standards after TXRF instrument recalibration.

| Merck data (n=48) | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Element | Ca | Ti | V | Cr | Mn | Fe | Co | Ni | Cu | Zn | As | Se | Sr | Sb | Tl | Pb |
| Measured | 96 | 101 | 100 | 98 | 95 | 105 | 97 | 97 | 100 | 97 | 100 | 100 | 95 | 102 | 77 | 91 |
| Reference | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| CV % | 4 | -1 | 0 | 2 | 5 | -5 | 3 | 3 | 0 | 3 | 0 | 1 | 5 | -2 | 23 | 9 |
| Benrd Kraft data (n=10) | | | | | | | | | | | | | | | | |
| Element | Ca | Ti | V | Cr | Mn | Fe | Co | Ni | Cu | Zn | As | Se | Sr | | | |
| Measured | 11 | 10 | 11 | 11 | 11 | 11 | 11 | 10 | 10 | 11 | 10 | 10 | 12 | | | |
| Reference | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | | | |
| CV % | -10 | -3 | -5 | -5 | -7 | -9 | -8 | -5 | 0 | -5 | 4 | -3 | -18 | | | |

Based on the performance of TXRF with the independent validation set (Table 2.2), we tested the combination of the calibration (n = 15) and independent validation (n=20) sets to obtained a wider range of soils and hence more reliable calibrations (Figure 2.4). Because of the slight improvements found in the model efficacy of these new calibrations, the results confirm our suspicion that using a larger set of calibration samples would further improve the instrument accuracy. However, the range of concentrations did not increase strongly using the added validation set suggesting (i) a good pre-selection of the original calibration set, and (ii) the need to widen the range with additional purposely selected samples. The current TXRF calibration and results will be considered for use in screening for, or selecting, any elements we bring forward during future direct quantification of total element concentrations in soils using TXRF.

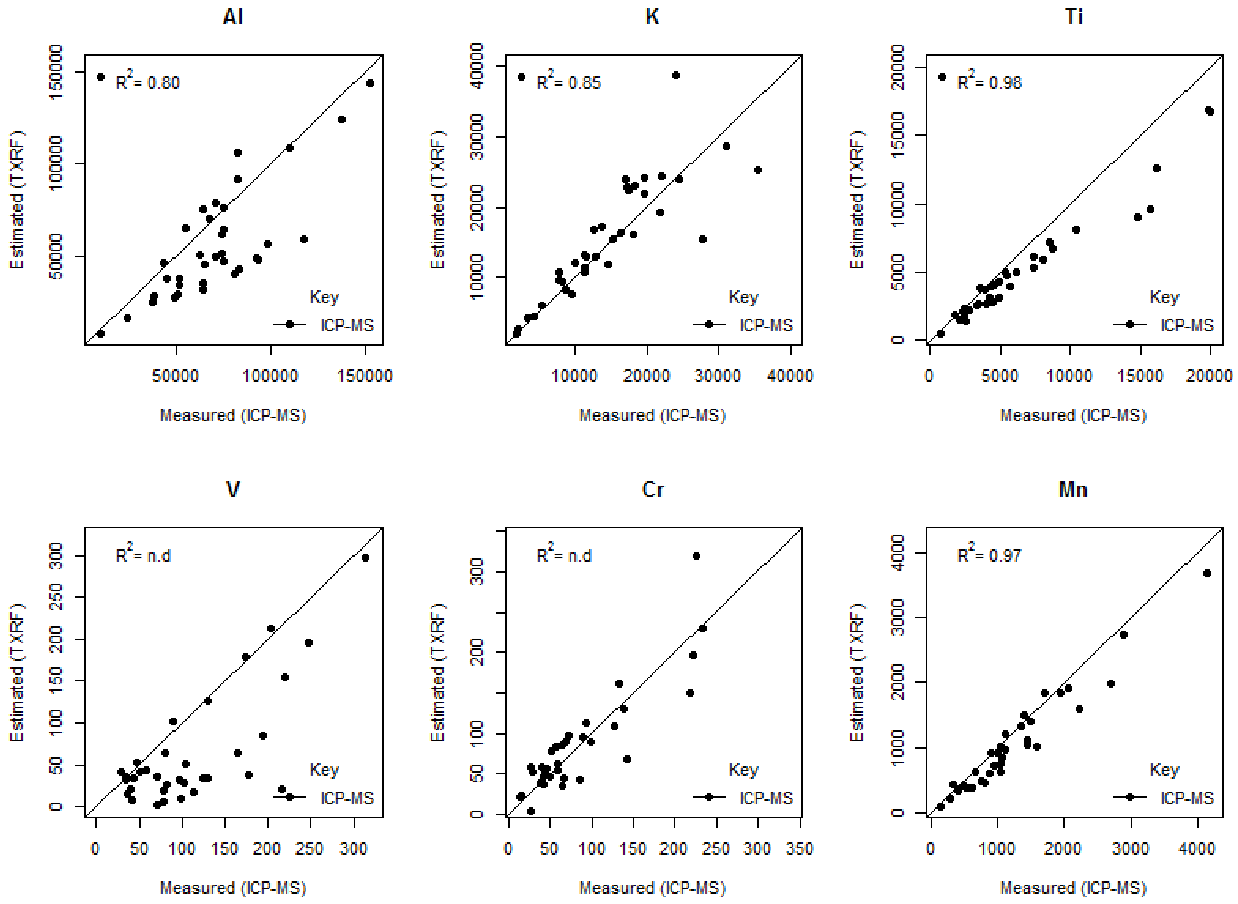


Figure 2.4: Actual concentrations (in mg kg$^{-1}$) measured using conventional total acid dissolution ICP-MS versus concentrations estimated using TXRF for 35 soil samples after recalibration for analyses performed on combined calibration (n = 15) and independent validation (n = 20) soil sample sets. Missing R$^2$ values were due to missing data either reference or estimated results.

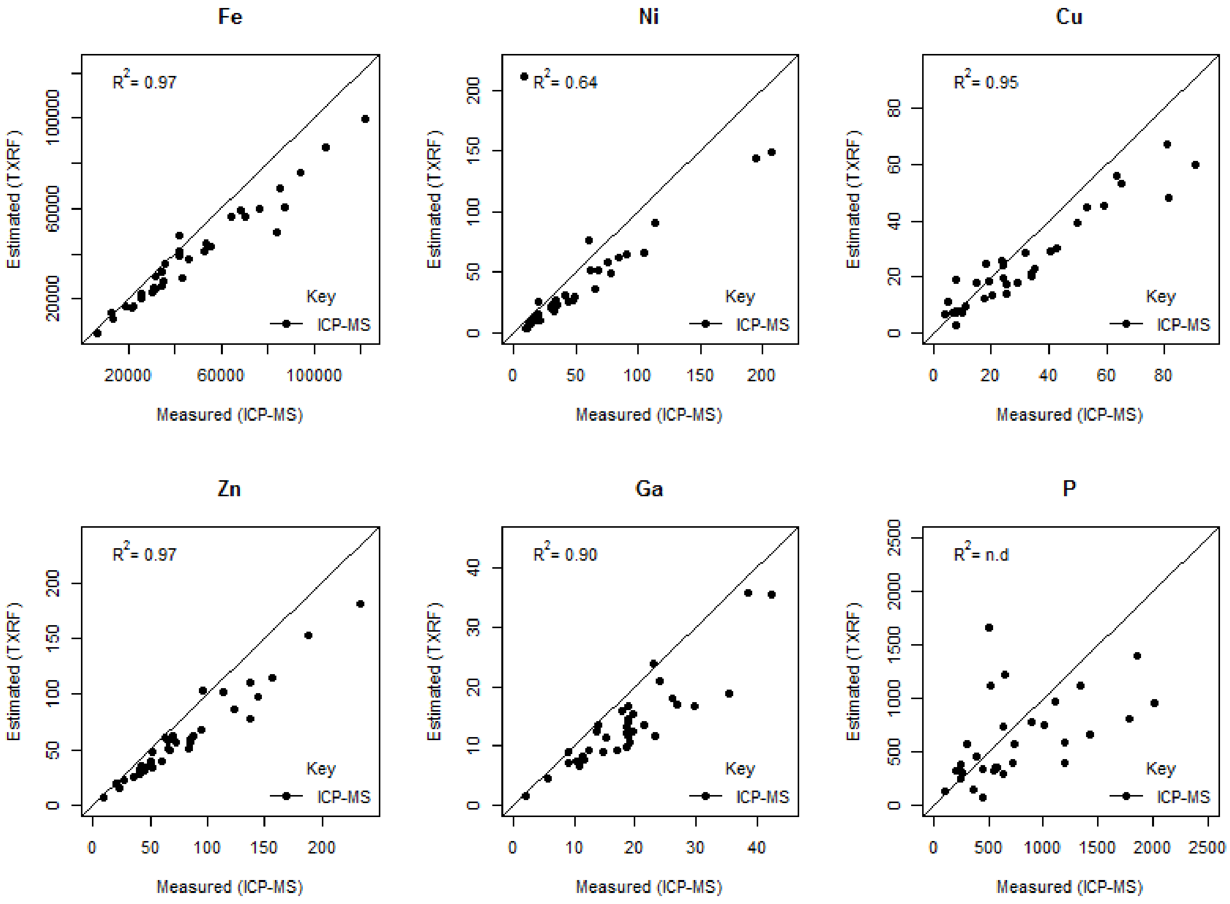


Figure 2.4 cont.: Actual concentrations (in mg kg$^{-1}$) measured using conventional total acid dissolution ICP-MS versus concentrations estimated using TXRF for 35 soil samples after recalibration for analyses performed on combined calibration (n = 15) and independent validation (n = 20) soil sample sets. Missing R$^2$ values were due to missing data either reference or estimated results.

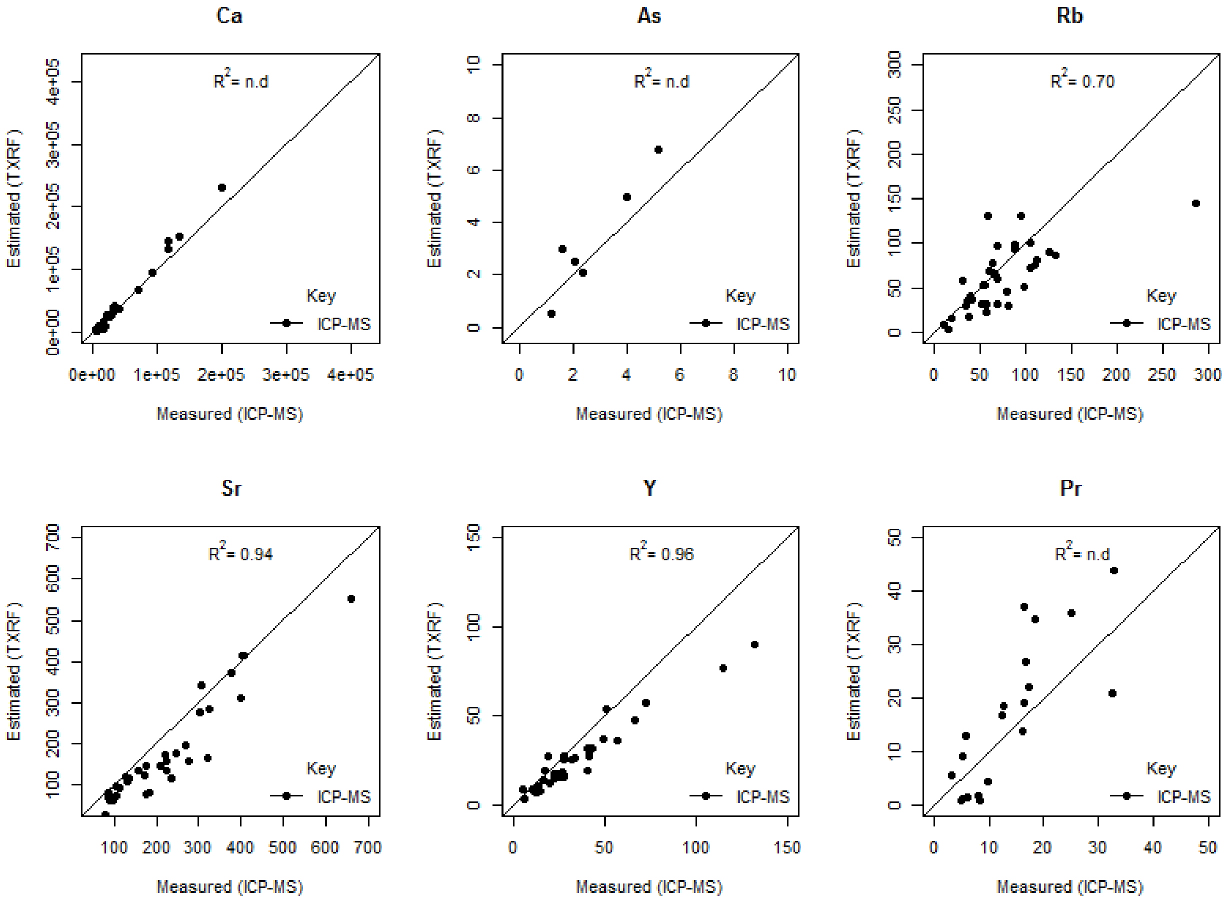


Figure 2.4 cont.: Actual concentrations (in mg kg$^{-1}$) measured using conventional total acid dissolution ICP-MS versus concentrations estimated using TXRF for 35 soil samples after recalibration for analyses performed on combined calibration (n = 15) and independent validation (n = 20) soil sample sets. Missing $R^2$ values were due to missing data either reference or estimated results.

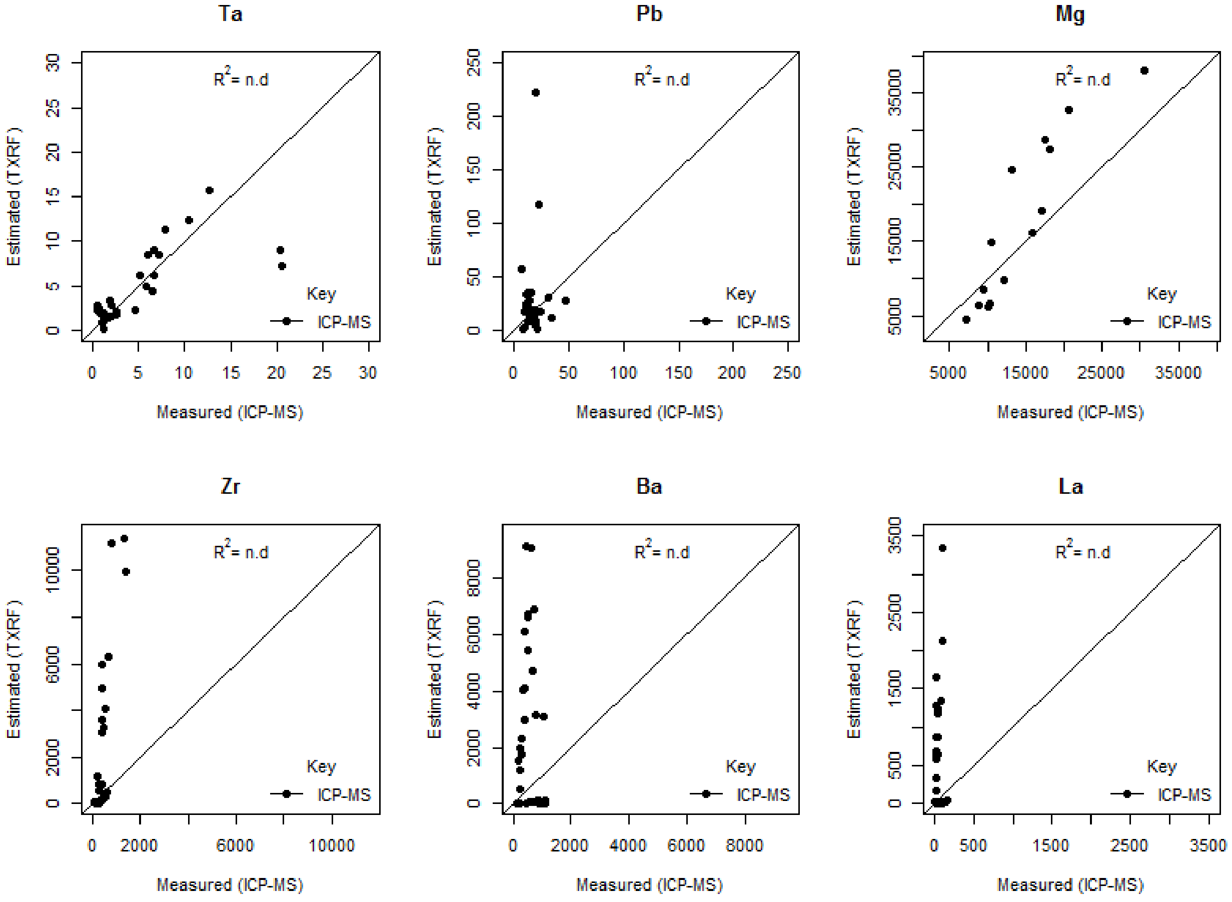


Figure 2.4 cont.: Actual concentrations (in mg kg$^{-1}$) measured using conventional total acid dissolution ICP-MS versus concentrations estimated using TXRF for 35 soil samples after recalibration for analyses performed on combined calibration (n = 15) and independent validation (n = 20) soil sample sets. Missing $R^2$ values were due to missing data either reference or estimated results.

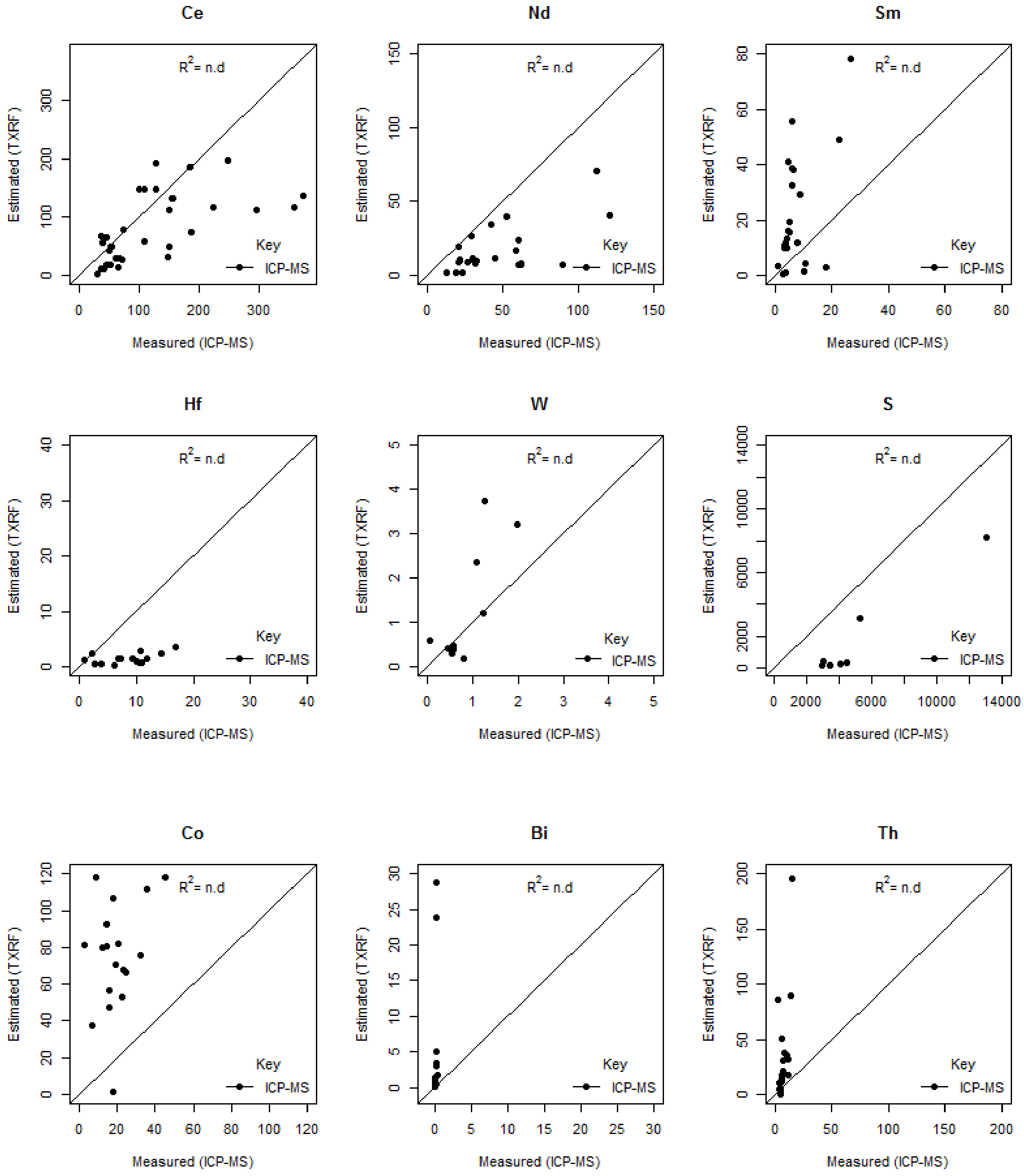


Figure 2.4 cont.: Actual concentrations (in mg kg$^{-1}$) measured using conventional total acid dissolution ICP-MS versus concentrations estimated using TXRF for 35 soil samples after recalibration for analyses performed on combined calibration (n = 15) and independent validation (n = 20) soil sample sets. Missing $R^2$ values were due to missing data either reference or estimated results.

## 2.5 Conclusions

The S2 PICOFOX$^{TM}$ spectrometer (Bruker AXS Microanalysis GmbH, Germany) has been reported as the world's first and only mobile bench-top system for a fast quantitative and semi-quantitative multi-element microanalysis of various types of samples applying the principles of TXRF (Stosnach, 2005). With detection limits in the ppb to ppm range, the S2 PICOFOX$^{TM}$ is also reported to be ideally applicable for trace element analysis and because of the instrument's total independence from cooling media, it is not just applicable in the laboratory but mobile on-site analysis (Bruker, 2007b). The suitability of this system for the mobile on-site analysis of heavy contaminated soils and sediments has been reported (Stosnach, 2005). This study has reported the possibilities and restrictions of TXRF for soil analysis application with regard to some elements (light and L-line elements). Our results show that TXRF is a promising technique for rapid and simultaneous determination of the concentrations of many elements in soil samples including low atomic number elements such as Al, P, K and Ca, if properly recalibrated. We developed and tested a procedure (single-element recalibration) to improve the sensitivity curve that provided accurate results <10% for most of the nominal or reference values for a set (n=20) of validation samples as well as two multi-element standards. Further, the TXRF method compared well with total acid dissolution ICP-MS analysis for a range of elements.

Errors introduced by line interferences depend on the concentration of the interfering element and, thus, are not systematic and cannot be corrected by the setup calibration. In general, the only factors that can be corrected by adjusting the calibration curve are systematic errors introduced by sample preparation. The most common factor with the correction for systematic errors is fluorescence radiation adsorption for the light elements (Na – Ca), which are usually not analysed accurately due to the absorption of their X-rays fluorescence signals in the air atmosphere of the instrument (Stosnach, 2005), however, because in TXRF the distance between sample surface and detector is very short (1.5 – 2 mm), the sample fluorescence is detected with high efficiency leading to a high absolute efficiency of the method (Bruker, 2007a) and thus, the adsorption in the air is negligible. We also tried the Bruker automatic pile-up correction procedure using the SPECTRA software, however, there were only very minor improvements in the results for some elements. However, we have realized that the pile up peaks are important error sources and the current corrections, both manual and automatic, are ineffective for some of the elements, requiring further research. Our results of analysis of an independent validation set as well as multi-element standards confirmed the extent of the improvement in the TXRF instrument sensitivity curve after recalibration. Despite some non-systematic underestimation for some elements even after recalibration, the TXRF method would be useful for estimation or measurements of total element concentrations of Al, P, Ca, K, Zn, etc. to reduce the current data uncertainty in Africa soils as confirmed by our analytical results.

The total concentration of different elements in the soil – its geochemistry – has implications for both human and animal health, for example, it influences the availability of a range of essential and potentially toxic elements which has implications for their uptake by both animals and crops (Rawlins et al., 2012). Understanding the natural concentrations of elements in the soil can help to determine whether, and the extent to which, soil may have been contaminated by anthropogenic activities (Rawlins et al., 2012). Environmental quality guidelines (EQGs) have been developed for soil element concentration values in attempts to determine and predict concentrations above which effects occur and below which effects do not occur (Chapman et al., 2003), but these values vary by jurisdiction, land use and by proponent as outlined for some elements in Table 2.5. However, presently EQGs are not available for tropical Africa soils. Certainly in Africa, data availability and quality are far from optimal and thus are important constraints on the potential to carry out environmental health mapping. Thus the role of EQGs in environmental quality assessments should, at present, be restricted to assisting in determining whether element concentrations pose relatively low or very high potential for significant toxicity to resident organisms (Chapman et al., 2003). Much of the variation in the concentration of major and trace elements in the soil is accounted for by the parent material from which the soil formed (Rawlins et al., 2012). Elements, other than those metals after uranium on the periodic table, are naturally occurring and can result from non-anthropogenic sources, including weathering, volcanic, and hydrothermal activities (Chapman et al., 2003). Thus, the assessment of soil quality for naturally occurring elements in Africa must take into consideration regional variations in background concentrations which strongly depend on geological and biological characteristics as well as recent management in natural environment.

As noted by Chapman et al. (2003) natural background concentrations of element in sediments or soils can vary greatly between sites and areas and as such determinations must be made in a geological context, accommodating both mineralogy and provenance (origin of soil parent materials) (Chapman et al., 2003). It should be noted that relatively high concentrations of elements can occur naturally in Africa soils thus making the distinction between pollution versus naturally occurring geological assessment difficult. This is because soils do reflect the natural composition of parent material and thus in some mining areas, naturally occurring elements in soils or rocks could be characterized as "contaminated". Thus no single guideline on concentrations as outlined in Table 2.5 can adequately represent the variance in the background concentrations across Africa. Whilst there may be areas that are yet able to establish comprehensive systems of environmental health mapping, opportunity to develop at least prototype systems does exist in many areas (Briggs, 2000). In addition, a growing recognition of the need for reliable environmental and health data is emerging in many countries, while the development of remote sensing technologies is greatly increasing the potential for environmental survey and monitoring (Briggs, 2000). Because problems of inconsistencies and uncertainties in diagnosis could occur, considerable effort may be needed in capturing suitably geo-referenced element concentration data (Briggs, 2000). Considerable scope does exist to obtain relevant data, at least in some parts of Africa, and the possibility of developing routine systems for data collection is undoubtedly improving. We thus propose that the TXRF technique be used as a rapid screening tool for assessment of total element concentrations and chemical fingerprints in soils. Element fingerprints using TXRF could be further tested for inferring soil chemical and physical functional properties which is of interest in the Africa soils context.

Table 2.5: Soil quality guidelines (SQG) values (mg $kg^{-1}$) for the protection of environmental and human health: summary of the compiled environmental guidelines and standards for Canada, Australia, Dutch and UK.

| | Canadian Land use[1] | | | | Australian[1] | | Dutch[1] | United Kingdom[1] | Australian (air-dry cultural soils)[2] | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Industrial | Commercial | Residential | Agricultural | SQG | SQG | | | | | |
| Element Name | Level | Level | Level | Level | Level | Level | Level | Level | Range | Maximum | Tolerable |
| Arsenic, As | 12 | 12 | 12 | 12 | 20 | 70 | 76 | 32 | 0 - 20 | < 8000 | 20 |
| Cadmium, Cd | 22 | 22 | 10 | 1 | 2 | 10 | 13 | 10 | 0 - 1 | < 200 | 3 |
| Chromium, Cr | 87 | 87 | 64 | 64 | 80 | 370 | 78 | | 2 - 50 | < 20000 | 100 |
| Copper, Cu | 91 | 91 | 63 | 63 | 65 | 270 | 190 | | 1 - 20 | < 22000 | 100 |
| Lead, Pb | 600 | 260 | 140 | 70 | 50 | 220 | 530 | | 0 - 20 | < 10000 | 100 |
| Mercury, Hg | 50 | 24 | 7 | 7 | 0 | 1 | 36 | | 0 - 1 | | 2 |
| Nickel, Ni | 50 | 50 | 50 | 50 | 21 | 52 | 100 | 130 | 2 - 50 | < 200 | 50 |
| Selenium, Se | 3 | 3 | 1 | 1 | | | | | 0 - 5 | | 10 |
| Thallium, Tl | 1 | 1 | 1 | 1 | | | | | | | |
| Vanadium, V | 130 | 130 | 130 | 130 | | | | | | | |
| Zinc, Zn | 360 | 360 | 200 | 200 | 200 | 410 | 720 | | 3 - 50 | < 1000 | 300 |
| Antimony, Sb | 40 | 40 | 20 | 20 | 2 | 25 | 22 | | | | |
| Cobalt, Co | 300 | 300 | 50 | 40 | | | 190 | | 1 - 10 | < 800 | 50 |
| Molybdenum, | 40 | 40 | 10 | 5 | | | 190 | | 0 - 5 | < 500 | 5 |
| Silver, Ag | 40 | 40 | 20 | 20 | 1 | 4 | | | | | |
| Tin, Sn | 300 | 300 | 50 | 5 | | | | | | | |
| Barium, Ba | 2000 | 2000 | 500 | 750 | | | | | | | |
| Sulphur, S | | | | 500 | | | | | | | |

[1] SQG values calculated from ESDAT Environmental Data Management Software, available online under http://www.esdat.com.au/Environmental_Standards.aspx.

[2] Adapted from Hock and Elstner (1995).

Annex 2.1: Actual mean concentration values (mg kg$^{-1}$) and the quantitative concentration ranges for 15 soils samples used in this study. Elements are ordered according to increasing atomic number.

| | Mean Concentration | | | | Quantitative concentration ranges | | | |
|---|---|---|---|---|---|---|---|---|
| Element | ICP-MS $Na_2O_2$ Fusion | ICP-MS HF acid digest | TXRF after instrument recalibration | TXRF before recalibration | ICP-MS $Na_2O_2$ Fusion | ICP-MS HF acid | TXRF after instrument recalibration | TXRF before instrument recalibration |
| Na | | 7960 | 5474 | 11562 | | 296 - 19688 | 2714 - 7835 | 8853 - 15093 |
| Mg | 6741 | 8059 | 1078 | 3532 | 950 - 34689 | 998 - 43636 | 306 - 3090 | 3532 – 3532 |
| Al | 78375 | 94596 | 78438 | 23469 | 43781 -152303 | 55710 - 161728 | 37662 - 143855 | 13222 – 39076 |
| P | 903 | 1042 | 669 | 407 | 109 - 2014 | 65 - 2149 | 141 - 1394 | 67 – 944 |
| K | 16033 | 19852 | 14899 | 6102 | 2074 - 35457 | 1740 - 42594 | 2024 - 28676 | 765 – 13831 |
| Ca | 10106 | 10073 | 6118 | 4443 | 3081 - 19065 | 1326 - 21135 | 549 - 12978 | 359 – 9720 |
| Ti | 8138 | 8811 | 6175 | 4317 | 1762 - 20016 | 1764 - 23894 | 1404 - 16784 | 1262 – 10309 |
| V | 117 | 133 | 108 | 30 | 29 - 313 | 33 - 387 | 31 - 298 | 2 – 94 |
| Cr | 79 | 80 | 67 | 40 | 15 - 234 | 15 - 247 | 4 - 230 | 2– 119 |
| Mn | 1442 | 1653 | 1356 | 835 | 337 - 4132 | 400 - 5069 | 348 - 3682 | 216 – 1865 |
| Fe | 57770 | 64293 | 45935 | 33730 | 12911 - 122158 | 14909 - 145849 | 14161 - 99224 | 10474 – 61893 |
| Co | 21 | 22 | | 40 | 4- 65 | 4 - 72 | | 40 - 40 |
| Ni | 36 | 37 | 23 | 15 | 10 - 114 | 7 - 119 | 4- 90 | 2– 67 |
| Cu | 33 | 35 | 30 | 21 | 4 - 82 | 6 - 88 | 7 - 67 | 55 – 53 |
| Zn | 109 | 114 | 83 | 65 | 20 - 234 | 21 - 233 | 19 - 180 | 16 – 107 |
| Ga | 21 | 25 | 16 | 12 | 9 - 42 | 11 - 45 | 7 - 36 | 7– 22 |
| As | 2 | 3 | 4 | | 0- 6 | 0 - 7 | 1 - 7 | |
| Br | | | 4 | 4 | | | 0 - 18 | 0 – 17 |
| Rb | 95 | 104 | 55 | 67 | 16 - 287 | 16 - 287 | 4 - 145 | 22 – 212 |
| Sr | 181 | 191 | 110 | 114 | 78 - 319 | 81 - 352 | 29 - 176 | 38 – 187 |
| Y | 37 | 39 | 27 | 26 | 10 - 72 | 9- 82 | 7 - 57 | 6 – 44 |
| Zr | 528 | 458 | 4803 | 5618 | 214 - 1357 | 97 - 1258 | 591 -11378 | 976 – 13296 |
| Mo | 2 | 2 | | | 0 - 7 | 1 - 8 | | |
| Ba | 727 | 760 | 62 | 518 | 137 - 1105 | 137 - 1187 | 6 - 123 | 7– 1902 |
| La | 85 | 97 | 12 | 283 | 16 - 174 | 19 - 206 | 2 - 39 | 42 – 634 |
| Ce | 174 | 216 | 128 | 161 | 36 - 374 | 44 - 483 | 31 - 197 | 9 – 308 |
| Pr | 16 | 18 | 21 | 9 | 3 - 33 | 4 - 38 | 4- 44 | 2 - 15 |
| Nd | 58 | 64 | 24 | 35 | 12 - 121 | 14 - 138 | 7 - 71 | 4 - 59 |
| Sm | 10 | 12 | 4 | 33 | 2 - 20 | 3 - 26 | 1 -12 | 10 - 74 |
| Yb | 4 | 4 | | | 1 - 8 | 1 - 8 | | |
| Hf | 13 | 13 | 1 | 4 | 6 - 31 | 3- 34 | 0 - 2 | 1 - 13 |
| Ta | 8 | 8 | 7 | | 1- 21 | 2 - 22 | 2 - 16 | |
| W | | 2 | 1 | | | 1 - 4 | 0 - 2 | |
| Tl | 1 | 1 | 3 | | 0 - 1 | 0 - 2 | 0 - 5 | |
| Pb | 22 | 23 | 13 | 15 | 9 - 48 | 9- 48 | 1- 31 | 6 - 40 |
| Bi | 0 | 0 | 6 | 2 | 0- 1 | 0 - 1 | 0 - 29 | 1 - 9 |

## 2.6 Acknowledgements

The research work undertaken for this study was financially supported by the World Agroforestry Centre (ICRAF) through the Africa Soil Information Service (AfSIS) project funded by the Bill & Melinda Gates Foundation, and by the Dr. Hermann Eiselen Ph.D. Grant from the Foundation fiat panis. Towett E., acknowledges the Young Excellence Scholars (YES) PhD Scholarship from the Food Security Centre at the University of Hohenheim, Germany, a travel grant as well as stipend within the framework of the “Global Food Security” funded by the Deutscher Akademischer Austausch Dienst (DAAD). The authors acknowledge Andrew Sila (ICRAF, Nairobi, Kenya) for assistance with data analysis, Robin Chacha (ICRAF, Nairobi, Kenya) for assistance with TXRF analysis of the soil samples and other technical staff of the ICRAF Soil-Plant Spectral Diagnostics Laboratory for assistance with sample preparation. The authors are also grateful to Hagen Stosnach (Bruker AXS GmbH, Berlin, Germany), for his valuable help and advice during the performance of the recalibration analytical work. Richard Coe (ICRAF-ILRI), Ermias Eynekulu (ICRAF), Gudrun Keding (Bioversity International, Nairobi), and Todd Rosenstock (ICRAF) provided very helpful advice and reviews that greatly improved the manuscript.

## 2.7 References

1. Africasoils.net [Internet]. Africa Soil information Service (AfSIS); c2012–13 [cited 2011 Oct 20]. Available from: www.africasoils.net/.
2. Bennun L, Sanhueza V. A procedure for the improvement in the determination of a TXRF spectrometer sensitivity curve. Anal Sci 2010; 26:331–5.
3. Briggs DJ. Environmental Health Hazard Mapping for Africa. WHO-AFRO, Harare, Zimbabwe. [Internet]. 2000 [cited 2013 April 12]. [140 p.]. Available from: www.who.int/hac/techguidance/tools/healthhazard.pdf.
4. Bruker. S2 PICOFOX Total Reflection X-ray Fluorescence Spectroscopy - Working Principles. Lab Report. Bruker AXS Microanalysis GmbH, Berlin, Germany; 2007a. Report No.: XRF 426.
5. Bruker. S2 Picofox user manual. Bruker AXS Microanalysis GmbH, Berlin, Germany; 2007b. Report No.: Doc-M81-EXX001-V1.

6. Bruker. S2 PICOFOX - Why Wait for Sample Prep! Shortest Time-to-Result through TXRF. Lab Report. Bruker AXS Microanalysis GmbH, Berlin, Germany; 2009. Report No.: XRF 435.
7. Chapman PM, Wang F, Janssen CR, Goulet RR, Kamunde CN. Conducting Ecological Risk Assessments of Inorganic Metals and Metalloids: Current Status. Human and Ecological Risk Assessment 2003; 9 (4): 641-697.
8. Cobo JG, Dercon G, Yekeye T, Chapungu L, Kaszere C, Murwira A, et al. Integration of mid-infrared spectroscopy and geostatistics in the assessment of soil spatial variability at landscape level. Geoderma 2010; 158:398–411.
9. Dhara S, Misra NL. Application of total reflection X-ray fluorescence spectrometry for trace elemental analysis of rainwater. Pramana-J Phys 2011; 76 (2):361–6.
10. Esdat.com.au [Internet]. Environmental Data Management Software; c2013 [cited 2013 April 12]. Available from: http://www.esdat.com.au/Environmental_Standards.aspx.
11. Garcia-Heras M, Fernández-Ruiz R, Tornero JD. Analysis of archaeological ceramics by TXRF and contrasted with NAA. J Archaeol Sci 1997; 24:1003–1014.
12. Hock B, Elstner EF. Schadwirkung auf Pflanzen. Spektrum Akademischer Verlag, Berlin, Oxford; 1995.
13. Janik LJ, Skjemstad JO. Characterization and analysis of soils using mid-infrared partial least squares. II. Correlations with some laboratory data. Aust J Soil Res, 1995; 33:637–650.
14. Kabata-Pendias A, Mukherjee AB. Trace elements from soil to human. Springer-verlag Berlin Heidelberg; 2007.
15. Kamau-Rewe M, Rasche F, Cobo J, Dercon G, Shepherd KD, Cadisch G. Generic prediction of soil organic carbon in Alfisols using diffuse reflectance Fourier-Transform Mid-Infrared spectroscopy. Soil Sci Soc Am J 2011; 75:2358–2360.
16. Kennard R, Stone L. Computer aided design of experiments. Technometrics 1969; 11:137–148.
17. Klockenkämper R. Total reflection X-ray fluorescence analysis. In: Winefordner, JD, editor. Chemical Analysis: A Series of Monographs on Analytical Chemistry and Its Applications. Vol. 140. John Wiley & Sons: New York; 1997.

18. Lin J. Performance of the Thermo Scientific Niton XRF Analyzer: The Effects of Particle Size, Length of Analysis, Water, Organic Matter, and Soil Chemistry. [Internet]. 2009 May [cited 2011 Nov 4]. [36 p.]. Available from: http://nature.berkeley.edu/classes/es196/projects/2009final/LinJ_2009.pdf.
19. Madari BE, Reeves JB, Machado PLOA, Guimarães CM, Torres E, McCarty GW. Mid- and near-infrared spectroscopic assessment of soil compositional parameters and structural indices in two Ferralsols. Geoderma 2006; 136:245–259.
20. Marguí E, Floor GH, Hidalgo M, Kregsamer P, Román-Ross G, Streli C, et al. Analytical Possibilities of Total Reflection X-ray Spectrometry (TXRF) for Trace Selenium Determination in Soils. Anal Chem 2010; 82:7744–7751.
21. Minasny B, McBratney AB. Regression rules as a tool for predicting soil properties from infrared reflectance spectroscopy. Chemometr Intell Lab 2008; 94: 72–79.
22. Minasny B, Tranter G, McBratney AB, Brough DM, Murphy BW. Regional transferability of mid-infrared diffuse reflectance spectroscopic prediction for soil chemical properties. Geoderma 2009; 153: 155–162.
23. Mori Y, Uemura K. Error factors in quantitative total reflection X-ray fluorescence analysis. X-ray Spectrom 1999; 28:421–6.
24. Mukhtar S, Haswell SJ. Application of total-reflection X-ray fluorescence spectrometry to elemental determinations in water, soil and sewage sludge samples. Analyst 1991; 116:333–8.
25. Rousseau RM. Detection limit and estimate of uncertainty of analytical XRF results. The Rigaku J 2001; 18 (2):33–47.
26. Rawlins BG, McGrath SP, Scheib AJ, Breward N, Cave M, Lister TR, Ingham M, Gowing C, Carter S. The advanced soil geochemical atlas of England and Wales. British Geological Survey, Keyworth. [Internet]. 2012 [cited 2013 April 12]. [228 p.]. Available from: www.bgs.ac.uk/gbase/advsoilatlasEW.html.
27. Sánchez-Moreno RA, Gismera MJ, Sevilla MT, Procopio JR. Direct and rapid determination of ultratrace heavy metals in solid plant materials by ET-AAS ultrasonic-assisted slurry sampling. Phytochem Anal 2010; 21:340–347.
28. Shepherd KD. Soil spectral diagnostics – infrared, X-ray and laser diffraction spectroscopy for rapid soil characterization in the Africa Soil Information Service. Proceedings of the 19th

World Congress of Soil Science, Soil Solutions for a Changing World; 2010 August 1–6; Brisbane, Australia. [Internet]. c2010 [cited 2011 October 21]. Available from: http://www.iuss.org/19th%20WCSS/Symposium/pdf/0738.pdf.

29. Shepherd KD, Walsh MG. Development of reflectance spectral libraries for characterization of soil properties. Soil Sci Soc Am J 2002; 66:988–998.
30. Shepherd KD, Walsh MG. Diffuse reflectance spectroscopy for rapid soil analysis. In: R Lal (ed) Encyclopedia of Soil Science 2004. Marcel Dekker Inc, New York. Online Published by Marcel Dekker: 04/26/2004.
31. Shepherd KD, Walsh MG. Infrared spectroscopy – enabling an evidence-based diagnostic surveillance approach to agricultural and environmental management in developing countries. J Near Infrared Spec 2007; 15:1–20.
32. Smith DB, Woodruff LG, O'Leary RM, Cannon WF, Garrett RG, JaKilburn JE, et al. Pilot studies for the North American soil geochemical landscapes project – site selection, sampling protocols, analytical methods, and quality control protocols. Appl Geochem 2009; 24:1357–1368.
33. Stosnach H. Environmental trace-element analysis using a benchtop total reflection X-ray fluorescence spectrometer. The Japan Society of Analytical Chemistry. Anal Sci 2005; 21:873–876.
34. Stosnach H. S2 PICOFOX Total Reflection X-ray Fluorescence Spectroscopy - Working Principles. Lab Report. Bruker AXS Microanalysis GmbH; 2007. Report No.: XRF 426.
35. Swift MJ, Shepherd KD. (Eds). Saving Africa's Soils: Science and Technology for Improved Soil Management in Africa. Nairobi: World Agroforestry Centre; 2007.
36. Viscarra Rossel RA, Walvoort DJJ, McBratney AB, Janik LJ, Skjemstad JO. Visible, near-infrared, mid-infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. Geoderma 2006; 131:59–75.
37. West M, Ellis AT, Potts PJ, Streli C, Vanhoof C, Wegrzynek D, et al. Atomic spectrometry update-X-ray fluorescence spectrometry. J Anal At Spectrom 2011; 26:1919–1963.

# Chapter 3

# Variability and patterns in total element composition of Sub-Saharan Africa soils using total X-ray fluorescence spectroscopy

## 3.0 Variability and patterns in total element composition of Sub-Saharan Africa soils using total X-ray fluorescence spectroscopy[2]

Erick K. Towett[a,b], Keith D. Shepherd[b,*], Jerome E. Tondoh[c], Leigh A.Winowieki[d], Desta Lulseged[e],
Mercy Nyambura[b], Andrew Sila[b,] and Georg Cadisch[a]

[a]Institute for Plant Production and Agroecology in the Tropics and Subtropics, Universität Hohenheim, Garbenstr. 13, Stuttgart, 70599, Germany.
[b]World Agroforestry Centre (ICRAF), P.O. Box 30677, Nairobi, 00100, Kenya.
[c]World Agroforestry Centre (ICRAF), WCA Sahel Node BPE 5118, Bamako, Mali.
[d]International Centre for Tropical Agriculture (CIAT), c/o ICIPE Duduville Campus, P.O. Box 823, Nairobi, 00621, Kenya.
[e]International Center for Tropical Agriculture (CIAT), c/o Chitedze Agricultural Research Station, P.O. Box 158 Lilongwe, Malawi.
*Corresponding author at:* World Agroforestry Centre (ICRAF), P.O. Box 30677, Nairobi, 00100, Kenya.
*Tel: +254207224163; Fax: +254207224000*
*Email:* k.shepherd@cgiar.org

### 3.1 Abstract

Conventional measurement methods for determination of total element concentrations in soils are time consuming and resource intensive and consequently are not widely used in routine soil analysis. In contrast, total X-ray fluorescence spectroscopy (TXRF) allows for rapid and simultaneous determination of the concentrations of many elements in multiple sample matrices and hence allows characterization of large areas. We examined the variability and patterns in total element composition by TXRF of soils from 34 randomly-located 100-km$^2$ sentinel sites distributed across Africa: Ghana (3 sites), Tanzania (8), Ethiopia (4), Mali (3), Burkina Faso (1), Mozambique (4), Nigeria (3), Zambia (1), Kenya (3), Guinea (2), and Malawi (2). The land degradation surveillance framework (LDSF), which is a hierarchical spatially stratified random sampling scheme with ten 100 m$^2$ plots nested within sixteen 1 km$^2$ clusters, nested within 100 km$^2$ sites was employed. Paired topsoil (0-20 cm) and subsoil (20-50 cm) samples were randomly selected from one plot from each of 16 clusters for each site and analysed for total element concentrations for 17 elements; Al, P, K, Ca, Ti, V, Cr, Mn, Fe, Ni, Cu, Zn, Ga, Sr, Y, Ta, and Pb. The within and between site patterns of variation in total element composition as well as reproducibility and accuracy of TXRF relative to conventional methods (inductively coupled plasma - mass spectroscopy (ICP-MS)) of total element analysis were explored. Total element concentration values were within the range reported globally for soil Cr, Mn, Zn, Ni, V, Sr, and Y and in the high range for Al, Cu, Ta, Pb, and Ga. There were significant variations ($P < 0.05$) in total element composition within and between the sites for all the 17 elements analysed.

---

[2] A version of this chapter has been submitted to the Geoderma Journal in June 2013.

Three elements had a large error of variance (Al, Ti and Ga). The greatest proportion of total variance and number of significant variance components occurred at the site (55-88%) followed by the cluster nested within site levels (10-40%). Our results also indicated that the strong observed within site as well as between site variations in many elements can serve to diagnose soil fertility potential. Our explorations of the relationships between element composition data and other site factors using "randomForest" statictics demonstrated that all soil soil-forming factors (e.g., parent material, climate, topography, management (landuse)) have important influence on total elemental concentrations in the soil. The most important variables explaining the main patterns of variation in total element concentrations were cluster, topography, landuse, precipitation and temperature, however the importance of cluster can be explained by spatial correlation at distances of <1 km. TXRF can provide efficient chemical fingerprinting which could be further tested for inferring soil chemical and physical functional properties which is of interest in the African soil contexts for agricultural and environmental management at large scale.

## 3.2 Introduction

Fingerprinting of soil elemental composition may form a useful basis for classifying soils in a way that relates to soil-forming factors and inherent soil functional properties (Kabata-Pendias and Mukherjee, 2007; Rawlins et al., 2012). Variation in the concentration of soil chemical elements is derived from differences in the composition of the parent material and from fluxes of matter and energy into or from soils over geologic time (Helmke, 2000; Rawlins et al., 2012) or recent management. In addition, soil-forming factors (e.g. parent material, climate, vegetation) are important drivers of total elemental concentrations in the soil. The nature of the key variables explaining ecological diversity of soils can be related to the mineralogy of parent rock and although these relationships have been inferred, mineralogy of parent rock is a principal factor determining spatial patterns of land resources (Voortman, 2011). For example, Rawlins et al. (2009) demonstrated use of element composition for the prediction of particle size distribution and their errors. Primary associations of elements reflect mineralogy and some elements can thus be used as proxies for mineralogy including heavy minerals, carbonates and clays (Grunsky et al., 2009). While total elemental concentrations in soil reflect parent material concentrations are also impacted by weathering rates and over the longer term by land use and climate. Soil mineralogy is a key determinant of many soil functions, for example nutrient quantities and intensities, pH and buffering, anion and cation exchange capacity, aggregate stability, soil carbon protection, dispersion, and resistance to erosion. The presence of different vegetation types is a reliable indicator of differences in soil chemical properties and various properties of the exchange complex and micronutrient levels; and together with interactions among plant nutrients, these significantly explain differences in vegetation and also the distribution of vegetation types (Voortman, 2011). Understanding the natural or geogenic concentrations of elements in the soil can help to determine whether, and the extent to which, soil may have been contaminated by anthropogenic activities (Rawlins et al., 2003) and thereby contribute to the protection of environmental health. Thus, the total concentration of different elements in the soil – its geochemistry – has implications for human, plant and animal health.

Recent advances in analytical chemistry and increasing environmental applications have added substantially to our knowledge of the biogeochemistry of trace elements (Kabata-Pendias and Mukherjee, 2007). Different chemical preparations of samples (e.g. total digestion) as well as different instrumental methods used for the determination of elements have an influence on final results but the use of reference certified materials have decreased uncertainties of analytical data (Kabata-Pendias and Mukherjee, 2007). Commonly used instrumental methods for soil analyses including inductively coupled plasma - mass spectroscopy (ICP-MS), inductively coupled plasma - optical emission spectrometry (ICP-OES), inductively coupled plasma - atomic emission spectrometry (ICP-AES), atomic absorption spectrometry (AAS), and electroanalytical techniques require total digestion and are often expensive and time-consuming and hence have been rarely applied in routine soil analysis.

XRF spectroscopy has been used over the last three decades as a non-consumptive technique for multi-element determinations, applicable to a wide range of matrix types, however, the major disadvantage of conventional XRF has been the poor elemental sensitivity, which is mainly a consequence of high background noise levels, resulting from instrumental geometries and sample matrix effects (Mukhtar and Haswell, 1991). In addition, with XRF spectrometry there is need to make pressed pellets, hence sample preparation is time-consuming. Total X-ray fluorescence spectroscopy (TXRF) is an upcoming but established commercially available technology which can provide rapid and simultaneous determination of the concentrations of many elements from Na to U in the periodic table with minimal sample preparation time (Stosnach, 2005; Towett et al., 2013).

The main principle of the TXRF technique is that atoms, when irradiated with X-rays, emit secondary X-rays – the fluorescence radiation. This basis facilitates TXRF analysis because the wavelength and energy of the fluorescence radiation is specific for each element and the concentration of each element can be calculated using the intensity of fluorescence radiation (Bruker, 2007a). A monochromatic X-ray beam is directed onto the sample at a very small angle (<0.1°), less than the critical angle of external total reflection for X-rays, causing total reflection of the beam photons after touching the sample, and consequently much higher sensitivities and a significant reduction of matrix effects (Bruker, 2007a; Stosnach, 2005; Stosnach, 2007). TXRF analysis only requires addition of an element that is not present in the sample as the internal standard for quantification purposes. Time from sample preparation to analysis for TXRF analysis is shorter than with conventional XRF using fused beads or pressed pellets or using acid digestion and ICP-MS or ICP-OES (Bruker, 2009), and hence TXRF provides rapid determination of the concentrations as opposed to the conventional methods.

To date, TXRF has not been widely used as an analytical technique for the total element composition of soils. However, a number of studies have used conventional XRF. For example, Marques et al. (2004), in an inventory of the trace element geochemistry of the Brazilian Cerrado soils, quantified the total elemental contents by wavelength-dispersive X-ray fluorescence spectroscopy (WD-XRF) and identified elements that accumulated versus those that were depleted over long periods of pedological weathering. The total element concentration fingerprints of TXRF, or spectra, may also be used to capture key mineralogical differences in soils and as an input to pedotransfer functions (Shepherd, 2010).

TXRF has also been tested as a versatile technique to investigate heavy metal pollution in soils as well as for basic research, particularly for locating trace elements in soils (Schulthess, 2011; Stosnach, 2005). Towett et al. (2013) developed and tested a method for the use of TXRF for direct quantification of total element concentrations in soils and demonstrated that TXRF could be used as a rapid screening tool for total element concentrations in soils assuming sufficient calibration measures are followed. Successful use of the TXRF technique would open up the possibilities for using total element composition to improve global predictions of soil properties, such as cation exchange capacity and extractable nutrients. Despite the importance of soil mineralogy in determining soil properties (Jenny, 1941), there have been few attempts to quantitatively link functional capacity to mineralogy or total element composition. Hence, we will evaluate how far TXRF total element derived patterns in soils relate to ‘available’ element results from Mehlich-3 soil tests (acid-extractable nutrients) and hence to soil nutrient supply capacity. Especially in Africa where variations in soil mineralogy and nutrient balance critically determine vegetation composition and agricultural potential (Voortman, 2011) TXRF could provide a particularly useful tool for prediction of soil properties in data sparse regions.

This work thus set out to test TXRF as an analytical technique for the total element composition of soils in the context of the Africa Soil Information Service (AfSIS) project, which was established to address the need for accurate up-to-date and spatially referenced soil information to support agriculture in Africa (AfSIS, 2013). The objectives were to; (i) quantify the variability in total element composition of soils from a diverse set of soils across Sub-Saharan Africa, (ii) explore the patterns in total element composition of soils analysed, (iii) assess the relationships between total element concentrations and nutrient supply capacity by relating Mehlich-3 soil tests (acid-extractable nutrients) to total element analysis patterns in soil, and (iv) examine relationships between element fingerprints and site characteristics including mineralogy, climate, landform, vegetation type, plant material and management (cultivation).

## 3.3 Materials and methods

### 3.3.1 Study area and sampling

Georeferenced samples associated with the Africa Soil Information Service (AfSIS) Project (www.africasoils.net) were taken from a set of sentinel sites already sampled during Phase 1 of the project in 2011-2012, which were randomized over Sub-Saharan Africa within the major Köppen-Geiger climatic zones (Kottek et al., 2006) (Figure 3.1). Field sampling was made based on the Land Degradation Surveillance Framework (LDSF) protocol (Vågen et al., 2013). The protocol is built around the use of “Sentinel Sites” or “Blocks” of 10 x 10 km in size. Hence, the current study was based on samples from thirty four 100-km$^2$ random sites across Sub-Saharan Africa-: Ghana (3 sites), Tanzania (8 sites), Ethiopia (4 sites), Mali (3 sites), Burkina

Faso (1 site), Mozambique (4 sites), Nigeria (3 sites), Zambia (1 site), Kenya (3 sites), Guinea (2 sites), and Malawi (2 site) (Figure 3.1). The sampling scheme was nested: within sites, sixteen 1 $km^2$ clusters were located and within each cluster ten 1,000 $m^2$ "Plots" (about 35 m diameter) were sampled. The plots within clusters, and clusters within sentinel sites, were randomly placed so that unbiased estimates of problem prevalence were obtained. Soil samples were collected from four 100 $m^2$ "Sub-plots" located within the ten 1000 $m^2$ plots at two depths, 0-20, and 20-50 cm and only one composite of the four subplots was taken for each soil depth making a total of 360 soil samples. For this study, we selected 10% of the 360 soil samples of each sentinel site for reference and spectral analyses and thus we only chose Plot 1 of the randomized plots within each cluster on condition that both top- and subsoil were available and if either depth was missing, Plot 2 was taken for analysis. Thus, a total of 1074 samples (16 samples per cluster x 2 soil depths x 34 sentinel sites) were used for exploring spectral (TXRF) patterns, after excluding 14 samples that we either unavailable (11) or considered as spectral outliers (3). A summarized description of the 34 sentinel sites from which the soil samples used for this study were collected, including the location, average elevation, annual total precipitation, temperature ranges, major soil classes (based on the IUSS Working Group WRB (2007) and the FAO/EC/ISRIC (2003) world soil resources map (Scale 1:30 000 000 approx)), major landforms, topography, percentage of cultivated area and vegetation structure is given in Table 3.1.

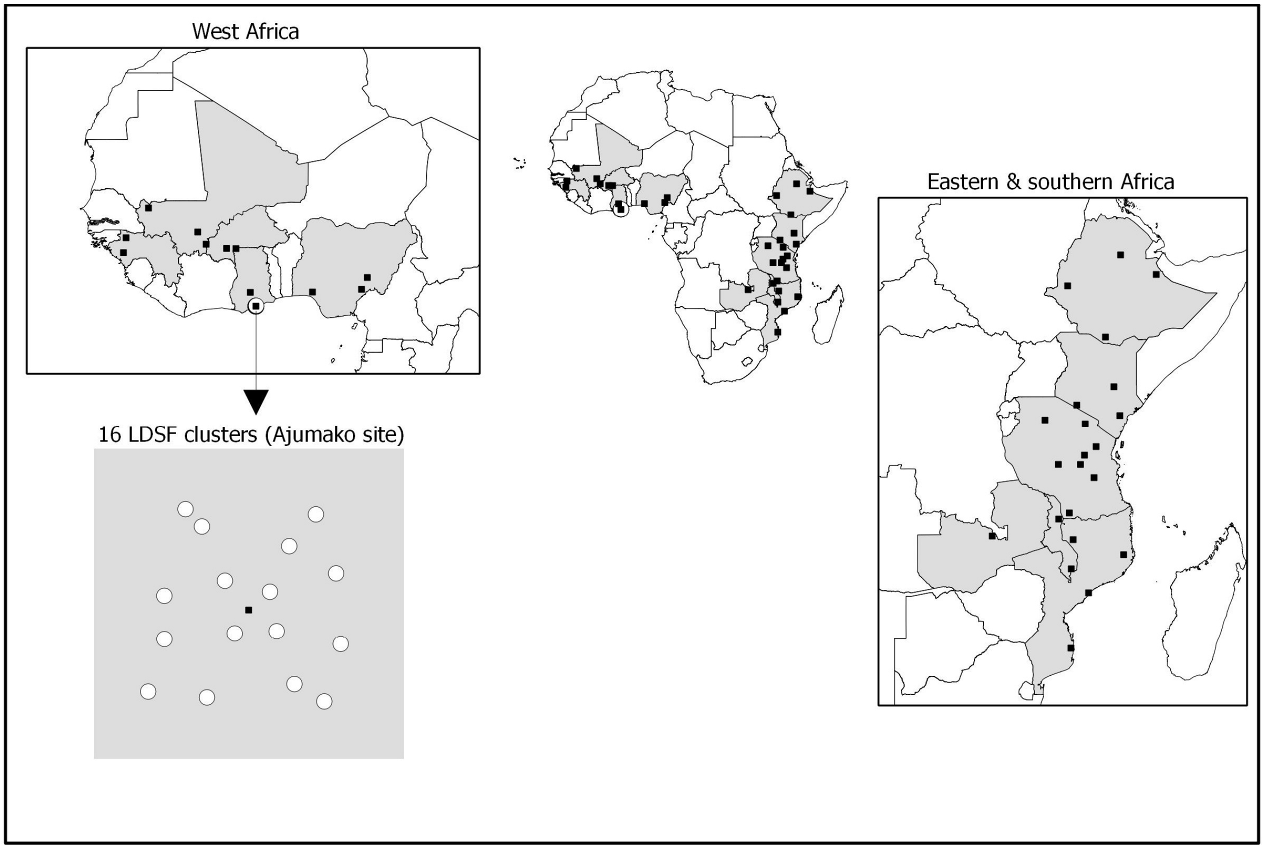


Figure 3.1: Graphics illustrate the centroid locations of 34 AfSIS sentinel sites sampled and available for this study. The basic sampling unit (sentinel site) is a 10 x 10 km block of land (black squares), within which there is a spatial hierarchy of 16 randomized LDSF clusters down to 1000 $m^2$ plots. The layout of each sentinel site sampled for this study consisted of 16 clusters (white circles).

Table 3.1: Site description: location, average elevation, annual total precipitation, monthly temperatures ranges, major soil classes, topography (toposequence), percentage of cultivated area and landuse of the 34 sentinel sites from which the soil samples used for this study were collected.

| Site | Country | Lat | Long | Elev (m) | Precip_ (mm) | Temp range (°C) | Soil class* | PosToposeq | Cultivated area (%) | Landuse |
|---|---|---|---|---|---|---|---|---|---|---|
| Ajumako | Ghana | 5.408 | -0.745 | 63.1 | 1217.8 | 23.6-31.7 | acrisols, alisols, plinthisols | midslope | 75 | - |
| Bondigui | Burkina Faso | 10.913 | -3.546 | 329.3 | 1002.3 | 24.4-37.4 | lixisols | mid-footslope | 44 | wooded grassland |
| Boumeoul | Guinea | 11.935 | -13.144 | 100.5 | 1692.9 | 24.1-38.7 | plinthosols | upland footslope | 31 | woodland |
| Bukwaya | Tanzania | -3.024 | 33.050 | 1197.2 | 876.4 | 18.6-30.8 | vertisols | upland-bottomland | 59 | cropland/shrubland |
| Chica_b | Mozambique | -14.714 | 39.873 | 275.6 | 1025 | 20.2-32.2 | arenosols | midslope | 27 | thicket/bushland |
| Chiculecule | Mozambique | -22.836 | 35.302 | 70.3 | 911.7 | 20.5-31.8 | arenosols | upland | 52 | cropland/woodland |
| Chinyanghuku | Tanzania | -6.895 | 36.129 | 902.8 | 600.3 | 18.6-31.2 | lixisols | midslope-bottomland | 20 | cropland/bushland |
| Dambidolo | Ethiopia | 8.613 | 35.015 | 1348.8 | 1349.3 | 16.1-31.6 | nitisols, andosols | upland-bottomland | 66 | cropland/woodland |
| Finnkolo | Mali | 11.313 | -5.502 | 409.1 | 1083 | 23.8-37 | lixisols | footslope | 63 | - |
| Fisenge | Zambia | -13.094 | 28.477 | 1249.6 | 1239.3 | 16.7-31.6 | ferralsols, acrisols, nitisols | upland-midslope | 63 | cropland/wooded grassland |
| Fria | Guinea | 10.507 | -13.393 | 193.2 | 2619.2 | 23.5-37.8 | leptosols | midslope-footslope | 63 | cropland/wooded shrubland |
| Ibi | Nigeria | 8.138 | 9.894 | 133.5 | 1113.2 | 24.1-37.1 | fluvisols, gleysols, cambisols | midslope-bottomland | 47 | shrubland/cropland |
| Ihassunge | Mozambique | -18.013 | 36.832 | 7.9 | 1316.3 | 23.3-32.7 | calcisol, cambisol, luvisol | bottomland | 81 | cropland |
| Imorun | Nigeria | 6.753 | 4.658 | 113.4 | 1723.3 | 23.6-33.1 | acrisols, alisols, plinthosols | upland - midslope | 63 | forest |
| Itende | Tanzania | -6.891 | 34.208 | 1216.5 | 630.6 | 17.6-30.3 | acrisols, alisols, plinthosols | upland - midslope | 0 | woodland |
| Katsina Ala | Nigeria | 7.014 | 9.342 | 175.5 | 1643.1 | 23.3-35 | lixisols | upland-bottomland | 87 | - |
| Kiberashi | Tanzania | -5.346 | 37.482 | 1109.9 | 758.1 | 16-28.8 | lixisols | midslope | 25 | forest/woodland |
| Kidatu | Tanzania | -8.036 | 37.303 | 319.4 | 1368.1 | 23.4-30.5 | lixisols | upland | 0 | woodland |
| Kisongo | Tanzania | -3.355 | 36.541 | 1328.8 | 875.5 | 16.1-28.4 | nitosols, andosols | midslope-footslope | 44 | - |
| Koloko | Mali | 12.483 | -6.296 | 292.1 | 899.3 | 25.3-38.3 | lixisols | upland -bottomland | 81 | cropland |
| Kontela | Mali | 14.809 | -10.999 | 67.1 | 603.7 | 27-41 | lixisols | footslope/bottomland | 25 | shrubland/grassland |
| Kubeasi | Ghana | 6.723 | -1.277 | 229.3 | 1473.6 | 22.2-33.3 | acrisols, alisols, plinthosols | upland -footslope | 56 | - |
| Kutaber | Ethiopia | 11.296 | 39.607 | 2515.7 | 1091.4 | 10.3-25.5 | nitosols, andosols | midslope-footslope | 63 | cropland/shrubland |
| Lambussie | Ghana | 10.896 | -2.649 | 308.1 | 961.5 | 24.5-37.3 | lixisols | midslope | 87 | cropland/shrubland |
| Macassangila | Mozambique | -13.407 | 35.495 | 1109.4 | 1228.1 | 17.8-29.3 | ferralsols | midslope | 45 | woodland/grassland/cropland |
| Marafa | Kenya | -2.634 | 39.546 | 165.5 | 687.2 | 22.6-32.9 | calcisols, cambisols, luvisols | bottomland | 0 | shrubland |
| Mbalambala | Kenya | -0.126 | 39.027 | 217.4 | 321.1 | 22.1-36.3 | calcisols, cambisols, luvisols | upland | 0 | shrubland |
| Mbinga | Tanzania | -11.091 | 35.161 | 983.4 | 1328.1 | 19.1-29.5 | lixisols | midslope | 31 | woodland |
| Mega | Ethiopia | 4.182 | 38.296 | 1540.1 | 592.1 | 12.6-27.8 | vertisols | bottomland | 0 | bushland/shrubland/grassland |
| Merar | Ethiopia | 9.602 | 42.701 | 1897.6 | 789.1 | 14.2-26.3 | calcisols, cambisols, luvisols | upland | 58 | cropland/shrubland/grassland |
| Morijo | Kenya | -1.719 | 35.811 | 2213.2 | 869.5 | 11.1-23.9 | calcisols, cambisols, luvisols | midslope-footslope | 6 | shrubland/grassland |
| Nkhata Bay | Malawi | -11.626 | 34.239 | 551.1 | 1672.8 | 21-30.4 | ferralsols, acrisols, nitosols | midslope-bottomland | 69 | cropland/wooland/grassland |
| Pandambili | Tanzania | -6.084 | 36.474 | 1086.8 | 727.3 | 16.7-31.4 | lixisols | bottomland | 6 | grassland |
| Thuchila | Malawi | -15.940 | 35.329 | 708.9 | 1517 | 19.3-30.7 | lixisols | upland | 94 | cropland |

Lat – latitude; Long – longitude; °C – degrees centigrade; Elev-m – Elevation in meters; Precip-mm – Average total precipitation in mm; Temp range – monthly temperature ranges; PosTopoSeq – Position in the topography. *Soil classification according to IUSS Working Group WRB. 2007. Elevation, precipitation and temperature data were calculated using the WorldClim sets of climate layers (climate grids) for Africa with a spatial resolution of about 1 km$^2$ (Hijmans et al., 2005). The soil classifications are based on the IUSS Working Group WRB (2007) and the FAO/EC/ISRIC (2003) world soil resources map (Scale 1:30 000 000 approx), while the other site description data were extracted from the Land Degradation Surveillance Framework (LDSF) database used in AfSIS.

### 3.3.2 Sample preparation and analyses

Soil samples were air-dried and passed through a 2-mm sieve before sub sampling to 10 g by coning and quartering. The 10 g subsamples were oven-dried at 40 ºC and then 5 g were ground to a fine powder (<200 µm) using a Retsch RM 200 mill (Retsch, Düsseldorf, Germany). Approximately 3 g of each soil sample was then further milled to <50 µm using a McCrone micronising mill (McCrone, Westmont, U.S.A). Spectral TXRF analyses were done at the World Agroforestry Center (ICRAF) Soil-Plant Spectral Diagnostics Laboratory in Nairobi, Kenya, using a Bruker S2 PICOFOX TXRF instrument (Bruker AXS GmbH, Germany). The instrument has an air-cooled low power (50 kV, 1 mA, Mo target) X-ray metal-ceramic tube with a Peltier cooled high resolution XFlash® Silicon Drift Detector (SDD), with a 10 $mm^2$ drifts chamber, and an energy resolution <160 eV at 100 kcps (Mn Kα line, 5.9 keV) (Bruker, 2007b). The TXRF method for analysis of soils was previously optimized and recalibrated against a set of international reference soil standards, and 15 purposely selected representative soil samples with a wide range of properties and elemental concentrations collected from various locations across Africa (Towett et al., 2013).

Sample preparation was according to the optimized procedure described by Towett et al. (2013). Briefly, an amount of ~ 50 mg of the finely ground soil sample was mixed with 2.5 ml of Triton X100 (Fischer Scientific, UK) solution (0.1 vol. %) to form a soil suspension and then spiked with 40 µl of 1000 mg $l^{-1}$ Selenium (Fluka Analytical, Germany) as the internal standard. The solution was sonicated for 15 minutes and then mixed well using a digital shaker and 10 µl of the turbid soil solution was immediately dispensed on to a clean siliconized quartz glass sample carrier (30 mm diameter) containing a drop of silicon solution that helped homogenise the soil suspension for TXRF analysis. Each soil sample prepared on the quartz disk was then dried for 10-15 minutes at 52 ºC on a hot plate (Staurt® SD300) placed in a clean laminar flow hood. Samples were analysed in triplicate. A mono-element standard Gallium (Ga) sample recommended by the instrument manufacturer was used for background and escape peak correction (Bruker, 2009). In addition, a drift in the spectroscopic amplification was compensated or reset by means of a gain correction (Bruker, 2007b). The data acquisition time was 1000 s per sample. The evaluation of spectra and element quantification were performed using the software program SPECTRA 6.3 (Bruker AXS GmbH, Germany).

Conventional pH measurements (utilizing a soil: water ratio of 1:2 weight to volume basis) as well as Mehlich-3 soil tests (acid-extractable nutrients) for the elements Al, Mg, P, K, Ca, Mn, Fe, Cu, and Zn using ICP mass spectroscopy were done in an external certified (ISO 17025 Accredited) laboratory (Crop Nutrition Laboratory Services Ltd, Nairobi, Kenya). Mineral profiling was done using X-ray diffraction spectroscopy (XRD) where finely ground (<50 µm) sampls were loaded into steel sample holders and analysed using a Bruker D2 Phaser XRD spectrometer instrument that was equipped with a LYNXEYE compound silicon strip, 1-dimensional detector with Theta/Theta geometry. The instrument was intergrated with the DIFFRACplus TOPAS graphics based, non-linear least squares profile analysis programe.

### 3.3.3 Detection limits of the elements

An important statistical consideration in TXRF analysis is the capability of the instrument to detect whether an element is present or not in a specimen and to be able to show with some defined statistical certainty that a given element is present if its concentration is greater than a certain limit (Rousseau, 2001). The lower limit of detection (LLD), which in this study is assumed to be the concentration equivalent to three standard counting errors of a set of measurements of the background intensity (Bruker, 2007b), was calculated using the following formula (Klockenkämper, 1997):

$$LLD_i = \frac{3.C_i.\sqrt{N_{BG}}}{N_i}, \qquad (1)$$

where $LLD_i$ = LLD of the element i; $C_i$ = concentration of the element i; $N_i$ = area of fluorescence under peak in counts; and $N_{BG}$ = background area subjacent the fluorescence peak.

In TXRF analysis, the concentration of an element is calculated from the net peak intensity which is equal to the difference between the measured peak intensity and the background intensity and the decision "detected" or "not detected" is thus established by comparison to a limit of detection, which is the combination of the mean background intensity and the background noise (Rousseau, 2001). Since knowing that an element concentration is below a given minimum detectable value is useful, we developed a procedure for assigning values to the LLD. In order to develop the analysis, we first of all made scatter plots of the total element concentration values against the LLD values calculated according to equation (1). We took the 25th percentile of the LLD values as the LLD limit (Annex Table A3.1). Taking the median value would have overestimated the LLD in many cases, and taking the 2.5th percentile would be sensitive to outliers and to changes as new samples were added. We developed a script in R statistical software (R Development Core Team 2013) that substituted the missing values with the 25th percentiles of average LLD values of the respective elements. Our approach was similar to that used by Reeves and Smith (2009) in their examination of mid- and near-infrared diffuse reflectance spectroscopy as possible tools for the determination of major and trace elements in soils where they replaced any value below the method detection limit with one-half the detection limit prior to statistical calculations.

### 3.3.4. Data analysis

We excluded from our dataset elements that were not accurately estimated using the TXRF technique as outlined by Towett et al. (2013), leaving us with the following 17 elements (Al, P, K, Ca, Ti, V, Cr, Mn, Fe, Ni, Cu, Zn, Ga, Sr, Y, Ta and Pb) out of 38 elements analysed by the technique. Concentrations of these elements detected in the soils were in a broad range and the data showed positive skewness in frequency distributions for most of the elements. Therefore, the appropriate transformation to achieve close to statistical normality of each variable was done before proceeding with analysis. We then established total concentrations of elements for soils occurring within particular sites sampled in our study, and documented systematic variation in their concentration and explored the possibility of finger-printing complete element profiles.

We used a principal component analysis (PCA) based on a correlation matrix using the R package “FactoMineR” (Lê et al., 2008) to explore the variation and view the interrelationships between different element concentrations and to detect sample patterns and variable relationships within and between the sites. The PCA transformed the measured TXRF element concentration variables into new, uncorrelated principal components (PCs) with each principal component being a linear combination of all the original measurement variables. To explore the proportion of variance for each element between sites, between clusters within site, depth, site and depth interaction and residual variance, we performed a mixed model analysis of variance using the following PROC MIXED statement:

*model log(element)= /ddfm=kr residual outp=resid; random site; random depth; random site*depth; random site*cluster.* (2)

where log = natural log; ddfm = specifies the method for computing denominator degrees of freedom; kr = Kenward-Roger; outp = output; resid = residuals.

The correlations among the five covariance parameters (sites, between clusters within site, depth, site and depth interaction and residual variance) were modeled using compound symmetry covariance matrix. The model was fitted by the Restricted Maximum Likelihood (REML) method in the SAS MIXED procedure (SAS Institute, 2011). A summary of some of the important options invoked in the SAS PROC MIXED statement by function and the class level information for three classes (site, cluster and depth), are summarized in Annex Table A3.2. We also explored the relationships between total element concentrations and nutrient supply capacity by correlating Mehlich-3 soil tests (acid-extractable nutrients) to total element analysis patterns in soil using a Spearman's rank correlation analysis in SAS (SAS Institute, 2011). In order to examine relationships between element fingerprints (principal components (PCs)) and site characteristics including elevation, climate, major landform, landform designation, vegetation type, management (cultivation) and mineralogy, we used the R package "randomForest" (based on Breiman, 2001) with an 'out-of-bag' (OOB) test sample being held out and used to estimate model error and for the calculation of variable importance. Thus we accessed whether any of the site and soil-forming factors could explain the variation in the element composition, while clarifying the questions:

(i) are element concentrations' main patterns of variation related to the site factors and if so, which are dominant?

(ii) do element concentration patterns relate to mineralogy?

For the first question, we performed a Random Forests regression (R-version 2.15.3; R Development Core Team 2013) of the first 5 PCs of the TXRF element concentration data against site factors while for the second question, we explored the Random Forests regression of elements PCs against mineral composition. We included and excluded different variables to test their effects on the overall model and variable importance.

## 3.4 Results and discussion

### 3.4.1 TXRF Method repeatability

The coefficient of variability (CoV in %) values for sample replicates calculated for those samples that had no missing values (before substituting the missing values with the 25th percentile value of the LLD) in the data after log transformation were within 10% for most of the elements except for P (20%), V (21%), Cr (11%), Y (15%), Ta (18%) and Pb (15%) (Figure 3.2). The mean CoV values calculated by site are given in Annex Table A3.3. We expected that the normalization by log transformation before calculation of the CoVs would be a better measure of the variation of the replicates than the actual values since most elements had left-skewed distributions, but this was not the case for some elements e.g. V, Y and Ta. Some of the elements being close to LLD caused high CoVs e.g. Ta (Tables A3.1, A3.3 and A3.4). One source of error with TXRF that must be minimized is variation due to sample preparation (Rousseau, 2001) and thus we attempted to minimize the variations caused by sampling the slurry (Sánchez-Moreno et al. 2010), ensuring that all samples and reference materials were prepared and presented to the instrument in an identical and reproducible manner (Towett et al., 2013). In addition, due to the assumptions that materials with low density were prone to float or agglomerate while other materials present a hydrophobic behavior, a dispersing agent, Triton X-100 was used in the slurry media and the shaking of the mixture in an ultrasonic bath for 15 minutes to obtain a homogenous suspension. Since an internal standard was added to the solution, any effects due to variation in deposition patterns of the sample on the sample holder would be expected to be small.

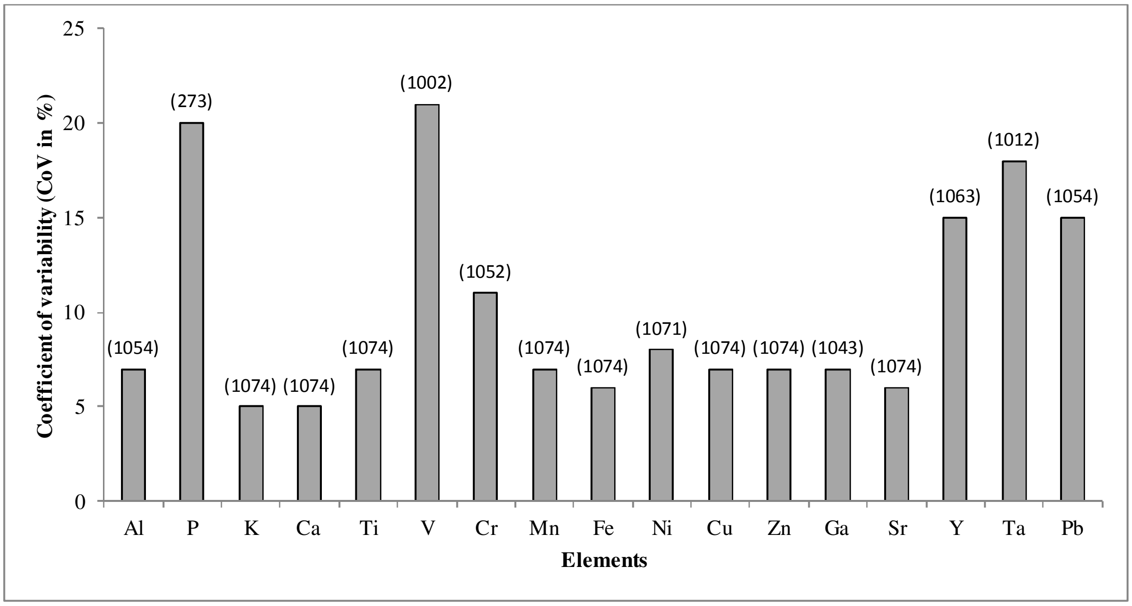


Figure 3.2: Bar charts of coefficient of average variation (CoV in %) of total element concentration calculated from the 3 laboratory replicates and the number of samples used to calculate the CoV values (in brackets) based on the samples that had no missing values (before substituting the missing values with the 25th percentile value of the LLD) in the data after log transformation.

### 3.4.2 Total element concentration in soil samples

As expected, there was wide variation within and between total element concentrations (Table 3.2). For example, the total concentration of P varied from 25-2360 (mean 140) mg kg$^{-1}$ across the 34 sentinel sites, while K and Ca values ranged between 290-77900 and 80-426460 mg kg$^{-1}$, respectively. This high variation was presumably attributable to differences in parent materials between sites and to local pedologic and hydrological factors within sites or due to differences in management (Table 3.1). A comparison between the results of analysis of soils in the different sentinel sites selected across SSA showed that they were within the reported ranges of worldwide soils (Table 3.2). However, our mean values obtained for soils of SSA exceeded the mean worldwide concentrations by 52 (12%), 15 (60%), and 18 (42%) mg kg$^{-1}$ for Mn, Pb, and Cr respectively, while the mean values for V, Zn and Sr were below the mean worldwide contents (Kabata-Pendias and Murkhejee; 2007) by 20 (33%), 32 (51%) and 27 (18%) mg kg$^{-1}$ respectively. Some of the elements of interest with regards to soil fertility are discussed below in order of decreasing abundance.

The concentration of Fe observed in SSA soils analysed in this study varied within the worldwide reported range (Table 3.2). In contrast, the total Al contents (90-89070 mg kg$^{-1}$) ranged wider than the worldwide range (10000 - 40000 mg kg$^{-1}$) reported by Kabata-Pendias and Murkhejee (2007). Aluminium is not completely extracted from the soil matrix during aqua regia digestion (Rawlins et al., 2012), hence the values reported here (using the TXRF technique) are higher than those reported for the worldwide soils in reported in Kabata-Pendias and Murkhejee (2007). However, Al may be abundant in soils, but most of this is not biologically available being present in the matrix of clays and other silicate minerals (Rawlins et al., 2012) and thus total Al has no direct agronomic implications. There are no reported compiled ranges for total Ca contents worldwide except for low contents of 200 mg kg$^{-1}$ reported for Brazilian Cerrado soils (Marques et al., 2004). However, the overall Ca content of soils analysed in this study had a much wider range and lower values than these reported for Brazilian Cerrado soils.

The concentration of Cr in surface soils is reported to vary widely but the amounts of Cr in the soil studied were within the reported range with a median value of 45 mg kg$^{-1}$ close to the worldwide averages of 54 mg kg$^{-1}$ (Haluschak et al., 1998) and 42 mg kg$^{-1}$ (Kabata-Pendias and Mekherjee, 2007) (Table 3.2). The median Pb content of soils in our study was similar to those reported for Africa complied from Ghana and South Africa soils (Herselman, 2007; Kabata-Pendias and Mekherjee, 2007). Nickel concentrations in the soils sampled for this study ranged within the worldwide range for soils reported (Haluschak et al., 1998; Hooda, 2010; Kabata-Pendias and Mekherjee, 2007; and US.EPA, 2006). The highest concentration of Cu measured in our sample set was 110 mg kg$^{-1}$ (Table 3.2) and this was within the reported range from 1 to 250 mg kg$^{-1}$ in surface soils. Although spectral interferences between peaks for Cu and Zn may affect the detection limits and accuracy of the TXRF analysis (US.EPA, 2006; Hooda, 2010), the detection limits for these elements obtained using the TXRF instrument in this study were 0.22 and 0.20 mg kg$^{-1}$ respectively (Annex Table A3.1) and these were lower than the typical detection limits reported for field-portable XRF instruments ranging from 10 to 50 mg kg$^{-1}$ (US.EPA, 2006; Hooda, 2010).

Table 3.2: Summary statistics of the median, mean, minimum, percentiles and maximum values of the total element concentrations (mg kg$^{-1}$) of 1074 soil samples collected from 34 sentinel sites and the reported concentration mean ranges of background contents of elements in the world soils.

| | Values compiled from this study (mg kg$^{-1}$) | | | | | | | | Reported mean and ranges of background contents of elements in crust and worldwide soils (mg kg$^{-1}$) | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Element | Median | Mean | Min | 2.5th %ntile | 25th %ntile | 75th %ntile | 97.5th %ntile | Max | Crustal Average[1] | A | B | C | D | E | F | Reported worldwide ranges[2] |
| Al | 34190 | 33930 | 90 | 250 | 21370 | 45720 | 68240 | 89070 | - | - | - | - | - | - | - | 10000-40000 |
| P | 40 | 140 | 25 | 40 | 40 | 45 | 1170 | 2360 | - | - | - | - | - | - | - | - |
| K | 7900 | 10890 | 290 | 650 | 2750 | 16800 | 34370 | 77900 | - | - | - | - | - | - | - | - |
| Ca | 1910 | 9780 | 80 | 170 | 680 | 8580 | 68640 | 426430 | - | - | - | - | - | - | - | - |
| Ti | 3190 | 4260 | 3 | 30 | 1920 | 5530 | 13070 | 25610 | 4400 | - | 3700 | - | 15480 | 2900 | - | 200-24000 |
| V | 20 | 40 | 0.7 | 1.1 | 7 | 40 | 160 | 390 | 135 | 60 | 69 | 180 | 320 | 80 | - | 5.0-500 |
| Cr | 45 | 60 | 0.7 | 1.4 | 20 | 80 | 250 | 600 | 100 | 42 | 22 | 58 | 86 | 54 | 72 | 1-1500 |
| Mn | 280 | 470 | 2 | 13 | 140 | 640 | 1600 | 6575 | 900 | 418 | 411 | - | 535 | 550 | - | <7->9000 |
| Fe | 20790 | 27950 | 20 | 230 | 9570 | 39070 | 91010 | 181690 | - | - | - | - | - | - | - | 1000-550000 |
| Ni | 10 | 20 | 0.3 | 0.9 | 5 | 25 | 70 | 360 | 20 | 18 | 13 | 26 | 25 | 19 | 39 | 0.2-500 |
| Cu | 10 | 20 | 0.3 | 1 | 6 | 20 | 60 | 110 | 55 | 14 | 17 | 48 | 109 | 25 | 17-29 | 1.0-250 |
| Zn | 20 | 30 | 0.3 | 1 | 10 | 40 | 90 | 140 | 70 | 62 | 65 | 89 | 73 | 60 | 45-47 | 10-602 |
| Ga | 8 | 8 | 0.2 | 0.2 | 4 | 10 | 20 | 30 | 15 | 1.2 | 8.9 | 20 | 31 | 17 | - | 0.4-70 |
| Sr | 50 | 120 | 1 | 3 | 20 | 130 | 810 | 1985 | 375 | 147 | 163 | 190 | - | 240 | - | 32->1000 |
| Y | 9 | 10 | 0.2 | 0.5 | 5 | 18 | 40 | 110 | 33 | 12 | 27 | 89 | 73 | 60 | - | 16-33 |
| Ta | 3 | 3 | 0.1 | 0.1 | 1.1 | 5 | 9 | 15 | 2.0 | 1.1 | 1.1 | 1.7 | 2.3 | - | - | 0.8-5.3 |
| Pb | 20 | 40 | 0.3 | 0.4 | 8 | 40 | 190 | 640 | 14 | 25 | 18 | 24 | 22 | 19 | 18-22 | 2.0-16338 |

[1] Values compiled by Kabata-Pendias and Murkhejee (2007); [2] Values compiled from Kabata-Pendias and Murkhejee (2007); Hooda, 2002; U.S. EPA Verification report, 2006; and Haluschak et al., (1998). Given are mean values for various soils: A - worldwide mean contents after Kabata-Pendias and Murkhejee (2007); B - Sweden soils after Eriksson (2001) (cited by Kabata-Pendias and Murkhejee (2007)); C - Japan agricultural soils, after Takeda (2004) cited by Kabata-Pendias and Murkhejee (2007); D -Median for soils in Brazil, after Licht (2005) (cited by Kabata-Pendias and Murkhejee (2007)); E - data for U.S. and Canada soils, after Burt et al., 2003 (cited by Kabata-Pendias and Murkhejee (2007)) and Haluschak et al., (1998); F - Reported median values for Africa compiled from Ghana soils, after Antwi-Agyei et al., (2009) and South African soils, after Herselman – 2007).

### 3.4.3 Principal component analysis of total element concentrations

The PCA of element concentrations revealed that patterns in total element concentrations between sites appeared to relate to differences in mineralogical ‘functional groups’ (Figures 3.3a-b). The pattern of clustering of the individual minerals and sorting of heavy minerals (V, Pb, Ni, Cr, Cu Ti, and Fe) along the positive Dim1 axis is apparent (Figure 3.3a). The first two PCA axes together accounted for 46.6% (Dim1 - 32.5%, and Dim2 - 14.1%) of the total variance within the data set (Figure 3.3a). The elements K, Ca, Ti, Fe and Sr were predominantly associated with Dim 1 while Dim 2 was dominated by elements Al, Ni and Ga. Elements Cr, Mn and Zn were predominant in the Dim 3 (Figure 3.3b, Annex Figure A3.1), while only V was predominant in Dim 4 and the elements P, Cu, Y, Ta and Pb were predominant in Dim 5 (Annex Figure A3.1).

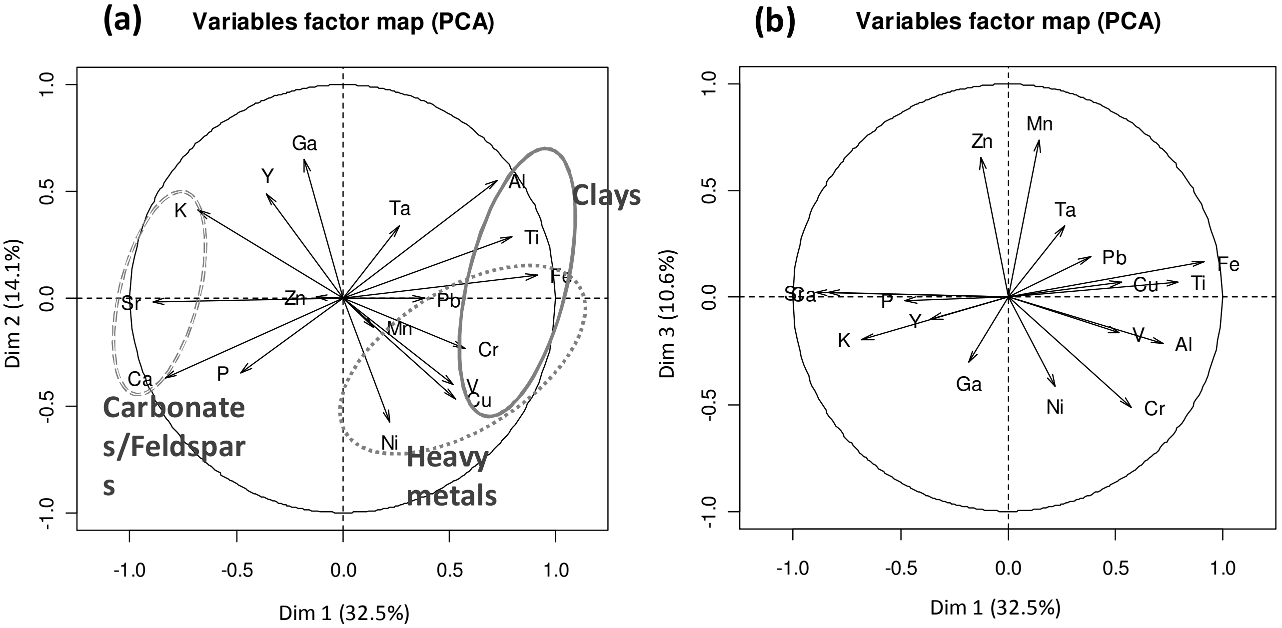


Figure 3.3: (a-b) Biplots (arrow sizes are proportional to the “initial” variability in the elements present) based on the principal component Dim 1 vs Dim 2 and Dim 1 and Dim 3, on the log transformed data of the soil total element concentration from all sites analysed.

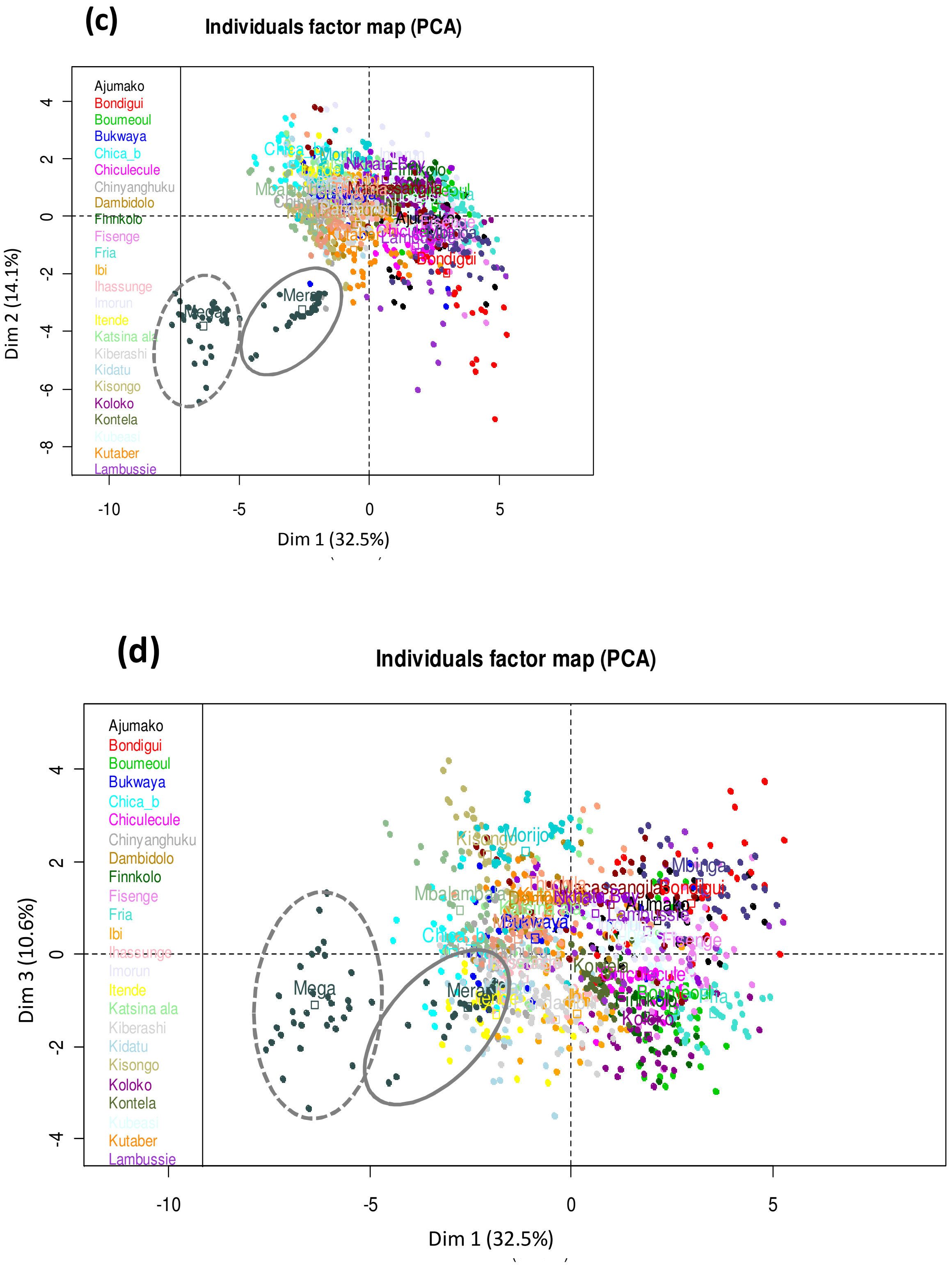


Figure 3.3 cont: (c-d) Scatter plots, based on the principal component Dim 1 vs Dim 2 and Dim 1 and Dim 3, on the log transformed data of the soil total element concentration from all sites analysed.

Clay minerals in the PCA plot are represented by some elements, e.g. Al, Cr, V, Cu, Fe, and Ti, most of which fell along the positive Dim1 axis (Figure 3.3a), similar to results reported by Cannon and Horton (2009) and Grunsky et al. (2009), i.e. that the clay factor commonly includes, although not exclusively, the elements Al, Cr, Cu, Ni, V, and Fe, all elements that are common in trace amounts in clay minerals or adsorbed to them. Qian et al. (1996) reported that although the heavy metals Ni, Cu and Pb were found to be enriched in the clay fraction, the extractable behaviour of Pb differed from that of Ni and Cu by being more susceptible to extraction. In our results, Ni and Cu had similar co-occurrence near the elements typical of clays and this was in agreement with reports by Cannon and Horton (2009) and Grunsky et al. (2009). Elements typical of feldspars and carbonates minerals (Ca, K and Sr) (Figure 3.3a) indicate the strong association of Ca and Sr and these are reported to be highly correlated ($r = 0.98$, $p = 0.0001$) (Marzecová et al., 2011). Having similar chemical properties, Sr readily substitutes for Ca in the structure of carbonate minerals. In our study, the sites with a strong Ca-Sr relationship included Ajumako, Dambidolo, Finnkolo, Fisenge, Ihasunge, Imorun, Itende, Kiberashi, Kisongo, Koloko, Kutaber, Lambussie, Macassangila, Mbalambala, Mega, Merar, Morijo, Nkhata Bay and Pandambili (Figure 3.4). The soils in these sites are classified as: Calcisols, Cambisols, Luvisols (Mbalambala, Merar and Morijo); Lixisols (Finnkolo, Kiberashi, Koloko, Lambussie, and Pandambili); Acrisols, Alisols (Ajumako, Ihansunge); Nitisols, Andosols (Dambidolo, Kisongo, and Kutaber); Ferralsols, Acrisols, Nitisols (Fisenge, and Nkhata Bay) and Vertisols (Mega) (Table 3.1). Hence, the distribution of Sr in minerals is largely controlled by Ca and is also strongly controlled by parent rocks and climate but is reported to likely concentrate in mafic igneous rocks and in carbonate sediments (Kabata-Pendias and Mekherjee, 2007; Marzecová et al., 2011). In addition, clay minerals have a large capacity to adsorb Sr, and most argillaceous sediments are enriched in this element (Kabata-Pendias and Mekherjee, 2007). Calcium, Sr and K co-occurred together and these could be associated with weathering of silicate bearing minerals (Figure 3.3a).

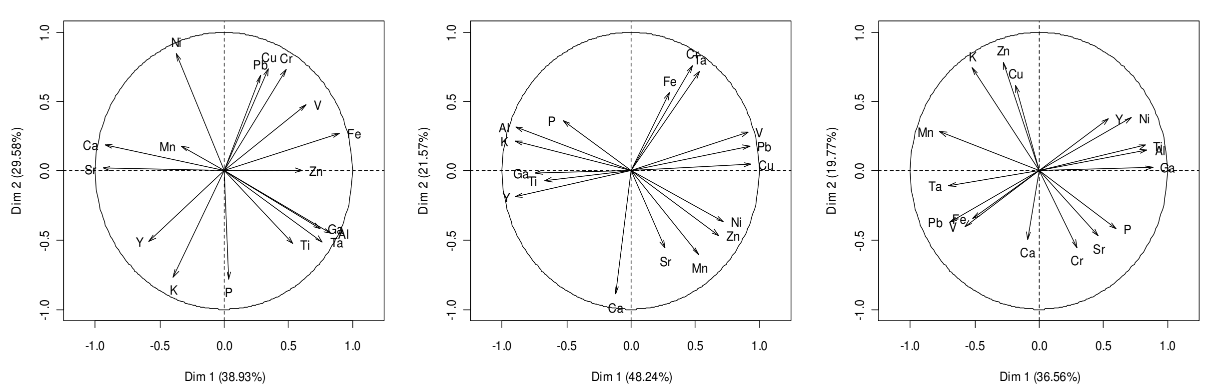
Similar to what was reported by Acosta et al. (2011), since K occurred in opposite direction to elements like Cu and Fe in our PCA results, an increase in the concentration of K is often followed by a decrease in the contents of Fe and Cu in soils. At the opposite extreme, K is also readily leached in highly weathered soil for example those with high Al and Fe concentrations. The positive relationship between Pb and Zn in our study can also be explained because Pb was found to coexist with Zn in the internal growth of a crystal lattice (Acosta et al., 2011). The cluster co-association of occurrence between Cu, Zn and Pb may also be the result of weathering of sulphide minerals similar to the results reported by Adebiyi et al. (2005). Lead and Zn are reported to be associated with Al- and/or Mn-bearing minerals such as albite, anorthite and biotite while MnO minerals are the main adsorbent for Pb in carbonate rocks under weathering (Acosta et al., 2011).

The TXRF technique has the potential to detect heavy metal pollution in soils, but our data did not contain levels considered as polluted. The cluster co-occurrence of heavy minerals (Ni, Cu, Zn, V, Pb, Cr, Fe, Ti, and Fe) along the positive Dim1 axis is apparent (Figure 3.3a). Similar to what is observed in the heavy metal clustering in this study, the accumulation of heavy metals with the clay fraction could be attributed to the high surface area and the presence of clay minerals, organic matter, and Fe-Mn oxides (Qian et al., 1996). Rodríguez et al. (2008) made an association of heavy metals with the factors in a PCA to indicate the hypothetical sources of these elements (lithogenic, anthropogenic, or mixed). We thus also infer that the concentrations of Cu, Zn, Cr, Pb, and Ni in our study (Figures 3.3a-b) could serve as proxies relating to differences arising from lithogenic or anthropogenic origin (Table 3.1). According to Hooda (2010), a high level of Cu is typical in soils of various parent materials for example loams developed on basalt rock (Cambisols) or some tropical soils (e.g. Ferralsols), however in our study the highest level of Cu was found in a soil classified as a Lixisol from Bondigui site (Table 3.1, A3.7). Further evidence for this inference was tested using the Random Forests algorithm where we explored the relationship between the total element concentrations with mineralogy and other site characteristics, discussed below.

### 3.4.4 Elemental variation between and within site

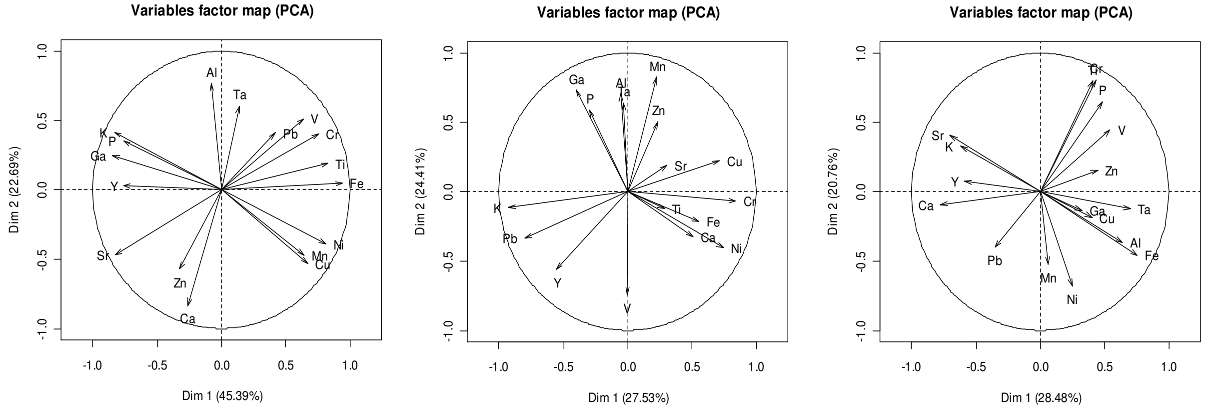
There were substantial variations in total element composition both within and between sites (Figures 3.3c-d and Figure 3.4). Elements clustered out differently in the different sample sets, indicating a wide variation in associations, but some elements are poorly represented (short arrows in the PCA) (Figure 3.4). The first two principal components explained between 49-84%

of the total variability for the different sentinel sites (Figure 3.4). As an example of the site differences, Figures 3.3b, 3.4ac and 3.4ad present the scatter and biplots of the Mega and Merar sentinel sites, both in Ethiopia, showing different elements clustering out separately indicating their different functional properties. The two sites also have similar precipitation amounts and temperature ranges but different parent materials (Table 3.1). The PCA of total element concentration confirms the fact that there are known specific differences e.g. between the Mega site, which is a gently undulating bottomland lying in a level plain with a mix of bushland, shrubland and grassland used for extensive grazing located on flat surfaces in central mountainous range around 1540 meters above sea level. On the other hand, the Merar site has a major landform that is also designated as level plain but located in an upland topography lying at an average altitude of 1898 meters above sea level used for a mix of cropland, shrubland and grassland (Table 3.1). The results provide further evidence to the apparent relation in the PCA to differences in mineralogical “functional groups” and to the fact that both soils have different parent materials (i.e. Vertisols and Calcisols/Cambisols for Mega and Merar respectively), however, the clearly distinct patterns and differences could also be explained by management and vegetation at the different sites (Table 3.1). For example, the Mega site which lies on an extensive rangeland area representing major management constraints brought about by the physical properties and the soil moisture regime of Vertisols and since they are used only for extensive grazing, they are consequently poor in nutrients because of no external inputs. Thus, the total element concentration analysis of the soils from the Mega site revealed very low concentrations of all elements with the exception of Ca (Tables A3.4-A3.7). Figures 3.3c-d acknowledges the fact that soil-forming factors are important drivers of element concentrations in soils since it shows sentinel sites with similar characteristics clustering together. The PCA patterns in total element concentrations within and between the sites, (Figures 3.3-3.4), could thus be used as proxies for natural nutrient status and the processes that might control the variation and differences in mineralogical “functional groupings”, and increase the wealth of knowledge of concentrations and spatial distribution of naturally occurring minerals in the soils of SSA.



(a) Ajumako (b) Bondigui (c) Boumeoul

(d) Bukwaya (e) Chica_b (f) Chiculecule



(g) Chinyanghuku (h) Dambidolo (i) Finnkolo

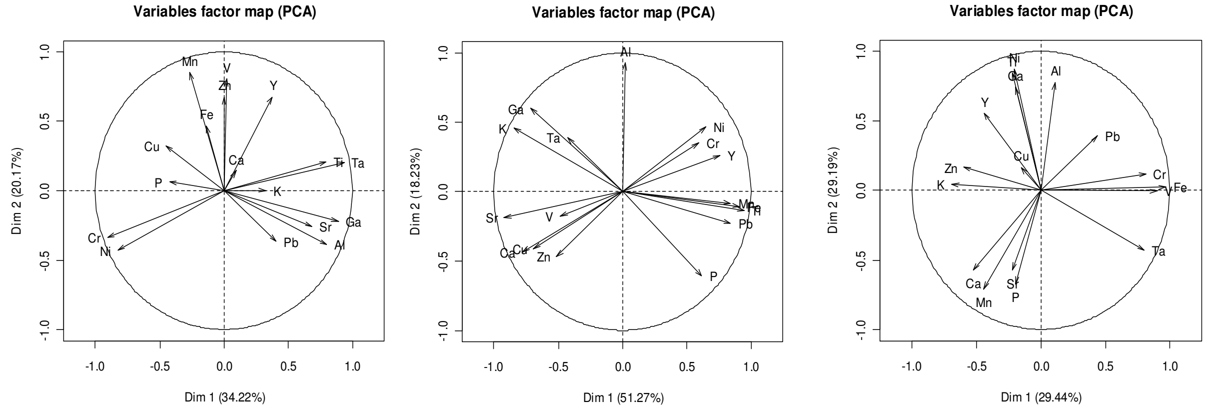
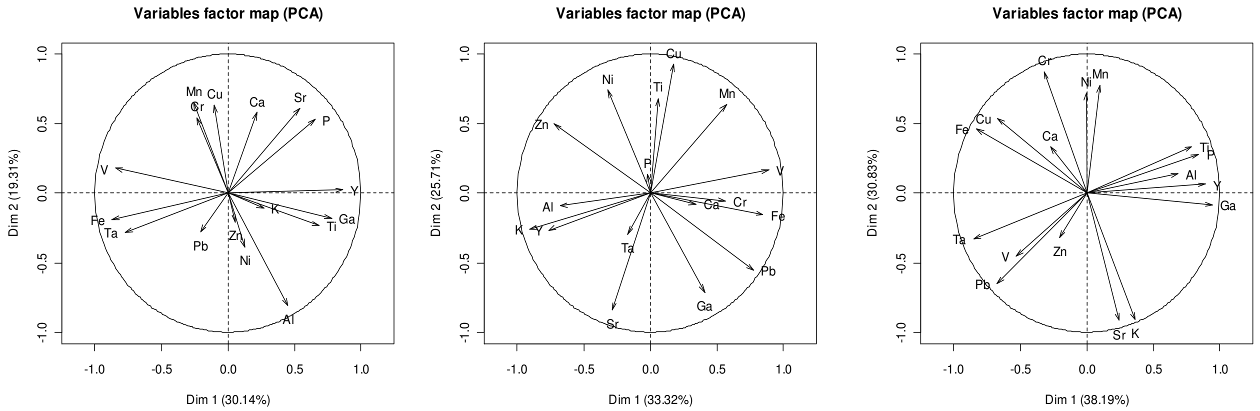
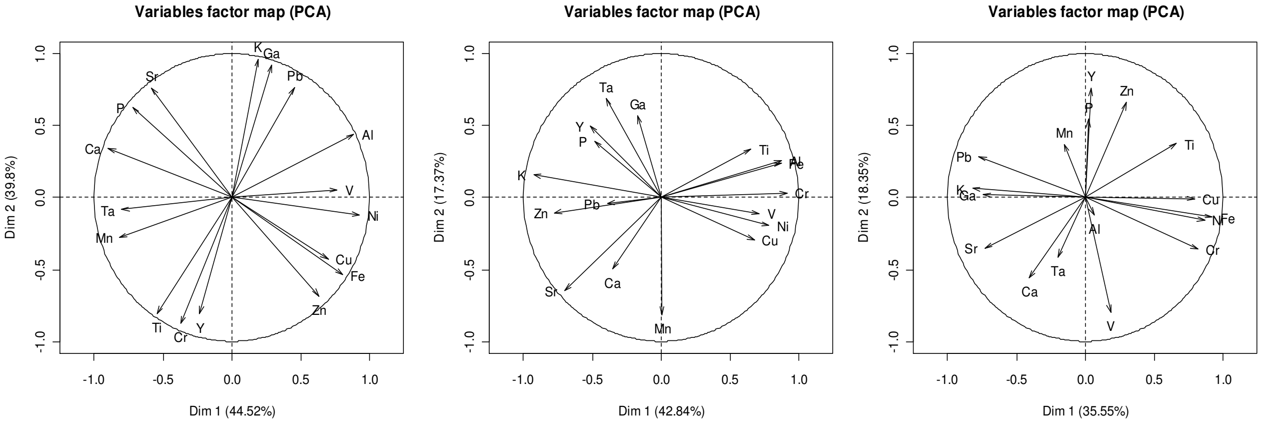


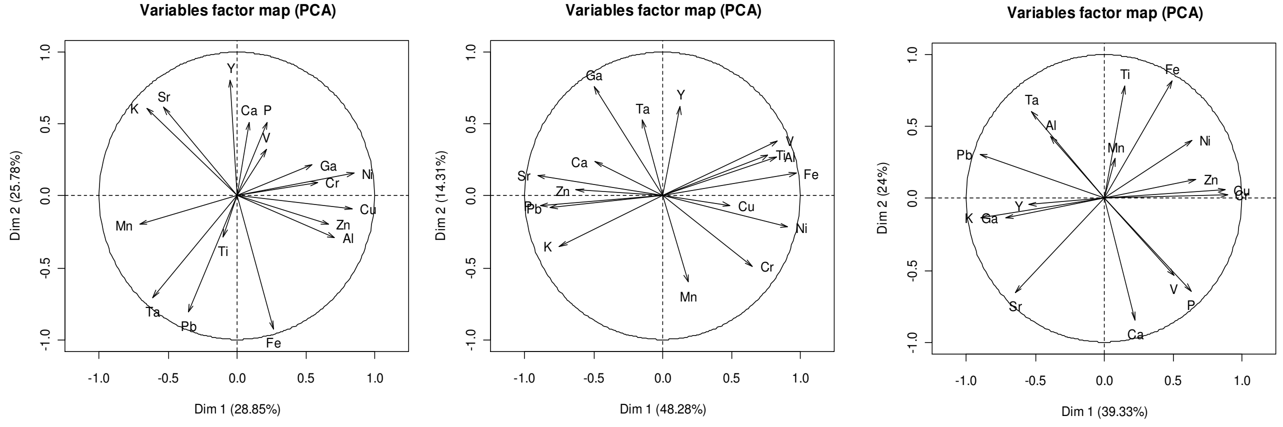
Figure 3.4: Biplots based on PCA of element concentration for each of the 34 sentinel site sample sets. Sentinel sites arranged according to alphabetical order.



**(j) Fisenge** **(k) Fria** **(l) Ibi**

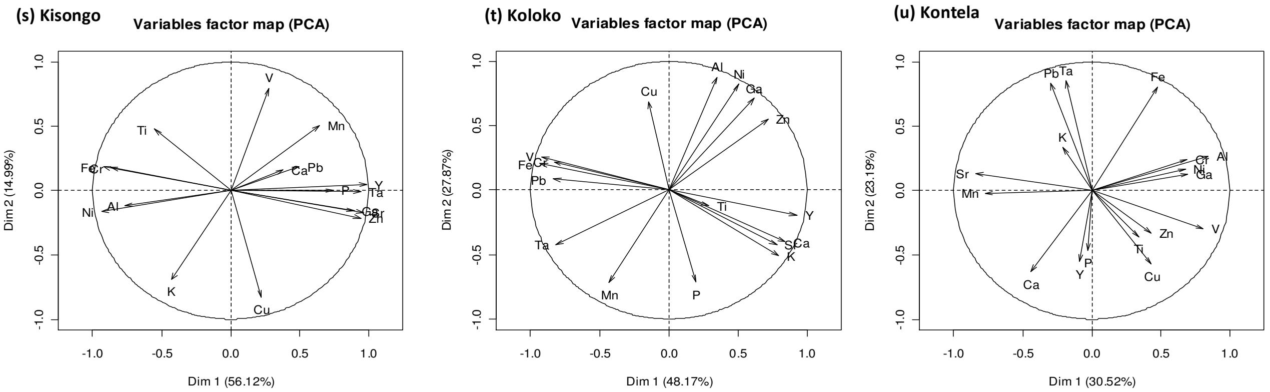


**(m) Ihassunge** **(n) Imorun** **(o) Itende**



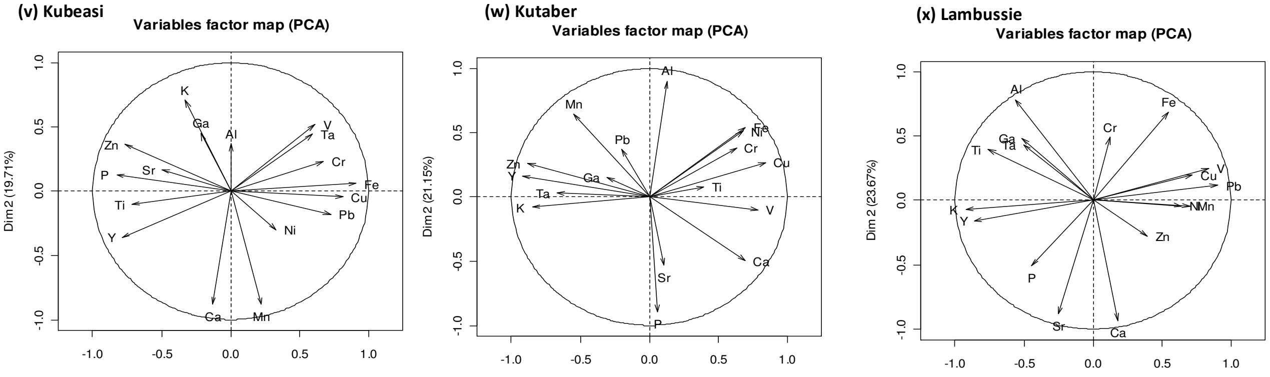
**(p) Katsina Ala** **(q) Kiberashi** **(r) Kidatu**

Figure 3.4 cont: Biplots based on PCA of element concentration for each of the 34 sentinel site sample sets. Sentinel sites arranged according to alphabetical order.


(s) Kisongo

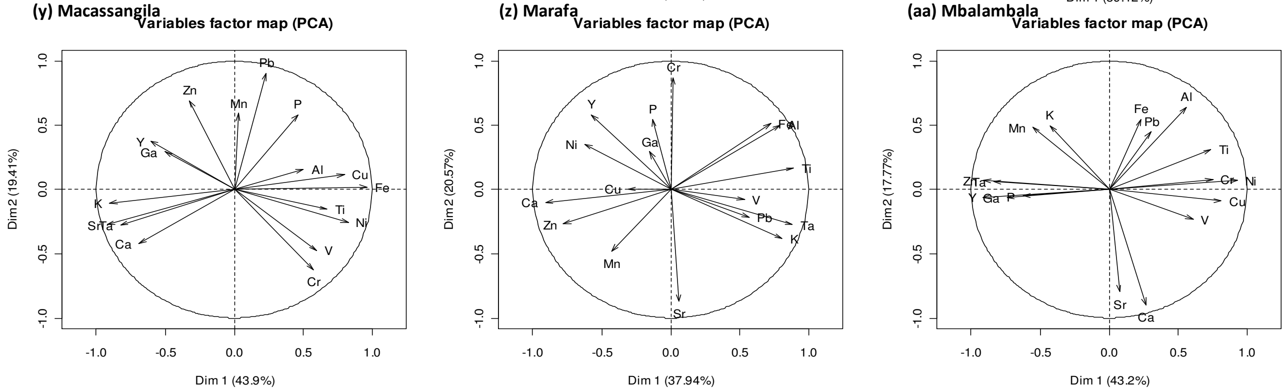
(t) Koloko

(u) Kontela


(v) Kubeasi

(w) Kutaber

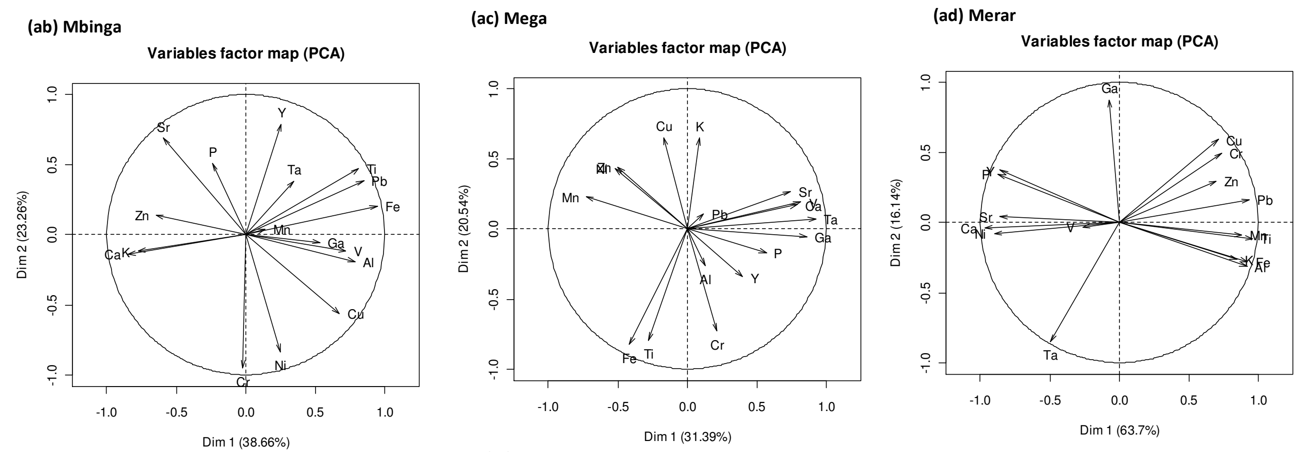
(x) Lambussie


(y) Macassangila

(z) Marafa

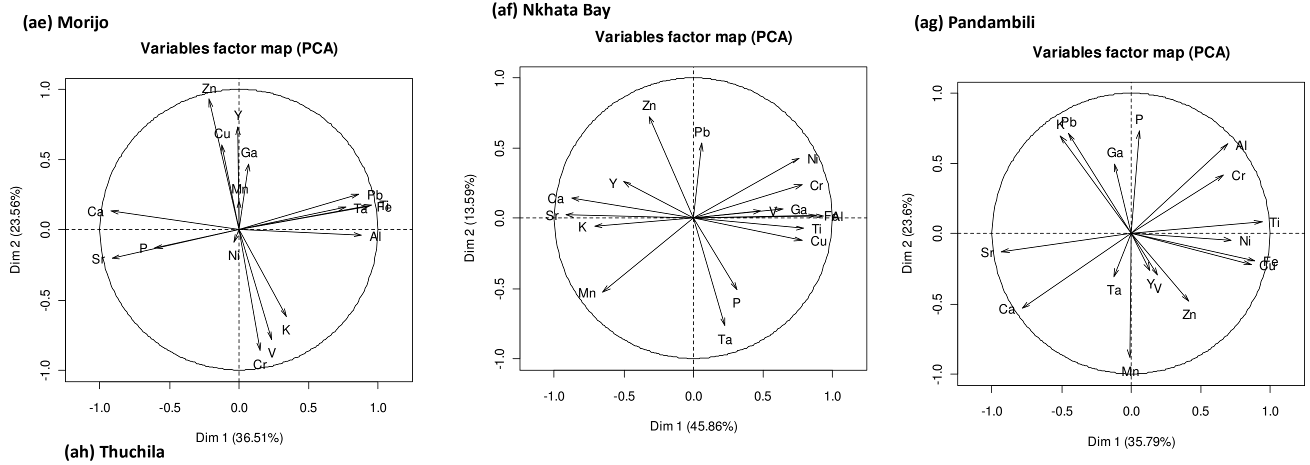
(aa) Mbalambala

Figure 3.4 cont: Biplots based on PCA of element concentration for each of the 34 sentinel site sample sets. Sentinel sites arranged according to alphabetical order.


(ab) Mbinga

(ac) Mega

(ad) Merar


(ae) Morijo

(af) Nkhata Bay

(ag) Pandambili

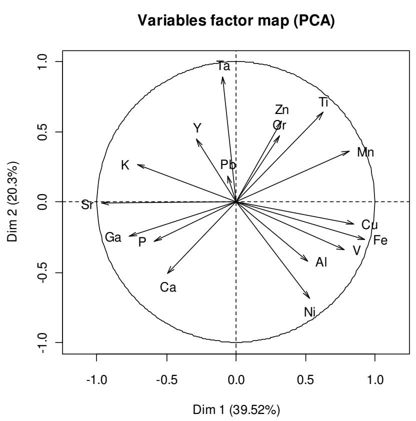

(ah) Thuchila

Figure 3.4 cont: Biplots based on PCA of element concentration for each of the 34 sentinel site sample sets. Sentinel sites arranged according to alphabetical order.

### 3.4.5 Restricted maximum likelihood analysis of the proportion of variance.

There were significant ($P < 0.05$) variance proportions for site, cluster nested within site, as well as depth nested within cluster and site and the interactions between site and depth (Table 3.3). The greatest proportion of total variance and number of significant variance components occurred at the site (55-88%) followed by the cluster nested within site levels (10-40%). This was not surprising given that we would expect lower variation within depths at a given sampling location. However, for the element P, a large proportion of the samples (75%) were at or below the 25th percentile LLD concentration of the analytical measurements. Consequently, our results concerning variation of P between and within sentinel sites should be interpreted with caution. Nevertheless, this may have implications for natural production systems and agriculture, as it may indicate also low potential P supply from minerals because TXRF is capable of detecting minerals rich in P. There were also 3 elements with a large error of variance (Al, Ti and Ga). Overall site contributed to more than 55% of the total variance of all the 17 elements analysed (Table 3.3). Our results differed with those reported by Laiho et al. (2004) who examined the proportion of variance of the elements K, Zn and Mn in soils from 11 peatland forest sites in Central Finland and reported strong variation among the sites, but even for these elements, the within-site variation contributed the biggest proportion of the total variance for the depth 0–30 cm. However, the cluster nested within site and site interaction contributed to at least 10% of the total variance of all the 17 elements analysed (Table 3.3). The observed strong within site as well as between site variations in the elements in our study were expected to be highly diagnostic of soil fertility potential because site and the interaction of cluster nested within site contributed more than half of the total variance for all elements (Table 3.3).

Table 3.3: Covariance parameter estimates and their percentage contribution to the total variance for site, cluster nested within site and depth nested within cluster and site and their interactions.

| | | | Site | | Site*Cluster | | Site*Depth | | Depth | | Residual | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Element | %LLD# | n | Estimate | %Tot var | Estimate | %Tot var | Estimate | %Tot var | Estimate | %Tot var | Estimate | %Tot var |
| Al | 2 | 1068 | 0.966 | 88 | 0.112 | 10 | 0.004 | 0.4 | 0.005 | 0.45 | 0.016 | 1.4 |
| P | 75 | 1059 | 0.718 | 76 | 0.198 | 21 | 0.002 | 0.2 | $1.4*10^{-21}$ | <0.01 | 0.025 | 2.6 |
| K | 0 | 1065 | 0.913 | 71 | 0.354 | 28 | 0.003 | 0.2 | $6.8*10^{-21}$ | <0.01 | 0.010 | 0.8 |
| Ca | 0 | 1068 | 2.186 | 79 | 0.480 | 17 | 0.034 | 1.2 | 0.017 | 0.60 | 0.051 | 1.8 |
| Ti | 0 | 1067 | 1.398 | 87 | 0.199 | 12 | 0.001 | 0.1 | 0.001 | 0.04 | 0.014 | 0.9 |
| V | 7 | 1067 | 1.463 | 77 | 0.379 | 20 | 0.009 | 0.5 | 0.008 | 0.39 | 0.053 | 2.8 |
| Cr | 2 | 1068 | 0.808 | 65 | 0.384 | 31 | 0.005 | 0.4 | 0.006 | 0.46 | 0.039 | 3.2 |
| Mn | 0 | 1067 | 1.007 | 68 | 0.393 | 27 | 0.023 | 1.6 | 0.008 | 0.51 | 0.040 | 2.7 |
| Fe | 0 | 1066 | 1.459 | 80 | 0.335 | 18 | 0.005 | 0.3 | 0.009 | 0.47 | 0.026 | 1.4 |
| Cu | 0 | 1066 | 0.705 | 69 | 0.285 | 28 | 0.002 | 0.2 | 0.003 | 0.27 | 0.027 | 2.6 |
| Ni | 0 | 1067 | 0.745 | 55 | 0.540 | 40 | 0.014 | 1.1 | 0.008 | 0.56 | 0.049 | 3.6 |
| Zn | 0 | 1067 | 0.924 | 79 | 0.217 | 19 | 0.001 | 0.1 | 0.001 | 0.12 | 0.026 | 2.2 |
| Ga | 3 | 1067 | 0.677 | 77 | 0.164 | 19 | 0.006 | 0.7 | 0.009 | 1.01 | 0.024 | 2.7 |
| Sr | 0 | 1069 | 1.554 | 83 | 0.299 | 16 | 0.004 | 0.2 | 0.000 | 0.01 | 0.017 | 0.9 |
| Y | 1 | 1066 | 0.721 | 66 | 0.309 | 28 | 0.004 | 0.3 | 0.000 | 0.00 | 0.061 | 5.6 |
| Ta | 6 | 1069 | 1.015 | 73 | 0.313 | 23 | 0.005 | 0.4 | 0.004 | 0.28 | 0.050 | 3.6 |
| Pb | 2 | 1069 | 1.103 | 57 | 0.689 | 36 | 0.011 | 0.6 | 0.009 | 0.45 | 0.117 | 6.0 |

%LLD# = percentage of the total number of samples (n=1074) with values at or below the LLD concentration; n= number of observations used after removing extreme studentized residuals; Estimate = log value estimated by the model; % Totvar = percentage contribution of the parameter estimate to the total variance. Model fitted with the SAS Sytem Mixed Procedure with the following parameters: covariance structure = variance components; estimation method = REML; residual variance method= profile; fixed effects SE method = Kenward-Roger.

### 3.4.6 Relationship with Mehlich-3 soil tests

The results of the exploratory analysis using PCA of the relationships between Mehlich-3 soil tests (acid-extractable nutrients) and total element analysis patterns in soil (Figure 3.5) were expected to relate to nutrient supply capacity. For example, the “inverse” relationship between pH (along the positive Dim 2) and elements total and extractable Ca and extractable Mg (co-occuring along the negative Dim 2) shown in Figure 3.5, could relate to nutrient supply capacity and could also be explained because calcitic lime (containing Ca) and dolomitic lime (containing Ca and Mg) can be used to raise the pH of very acid soils (Voortman, 2011). Extractable Cu, Al, and Ca tended to show a similar trend in the cluster co-occurrence to total Cu, Al, and Ca (Figure 3.5). In addition, these trends were confirmed by a Spearman’s rank correlation analysis results which showed that the extractable elements Mn, Cu, Ca, and Al were highly correlated to their total element concentration counterparts (Table 3.4) with more than 69% of the variation accounted for. This result was also in agreement with a study by Franklin et al. (2007) in which the correlations of Mehlich and total concentrations were found to be significant for Cu and Mn with more than 50 % of the variation accounted for. In our study, total Mn was closely associated with Fe implying that higher Mn levels are often found in soils rich in Fe. Mehlich-3 Mn also revealed a cluster co-occurrence of association with total Ca and Sr as well as Mehlich-3 Mg, Ca, and K (Figure 3.5). This was in agreement with the reported binding capability of Mn to bivalent cations such as Ca, Mg and Sr by Kabata-Pendias and Murkejee (2007). Kabata-Pendias and Murkejee (2007) also reported that the geochemistry of Ga is similar to that of Al and that Ga distribution in soils reflects the positive correlation with clay fraction and Fe hydroxides. In this study, there was a cluster co-occurrence of association between Ga, total Al and Mehlich-3 Al (Figure 3.5), and the clay factor commonly includes the total Al (Figure 3.3) as reported by Cannon and Horton (2009) and Grunsky et al. (2009). The results of spearman’s rank correlation outlined in Table 3.4 showed that there were strong correlations between Ga and total Al and Mehlich-3 extractable Al ($r = 0.88$ and 0.77 respectively). Our exploratory analysis using PCA presented here with a view to look for associations within multiple variables and to narrow down the predictor variables was only used in a quantitative way to show how far TXRF total element derived patterns in soils relate to ‘available’ element results from Mehlich-3 soil tests (acid-extractable nutrients) and hence to soil nutrient supply capacity. The associations presented in Figure 3.5 and Table 3.4 look promising and future studies should go ahead to develop predictive models as to what extent and for which elements total element concentrations relate to nutrient supply/availability.

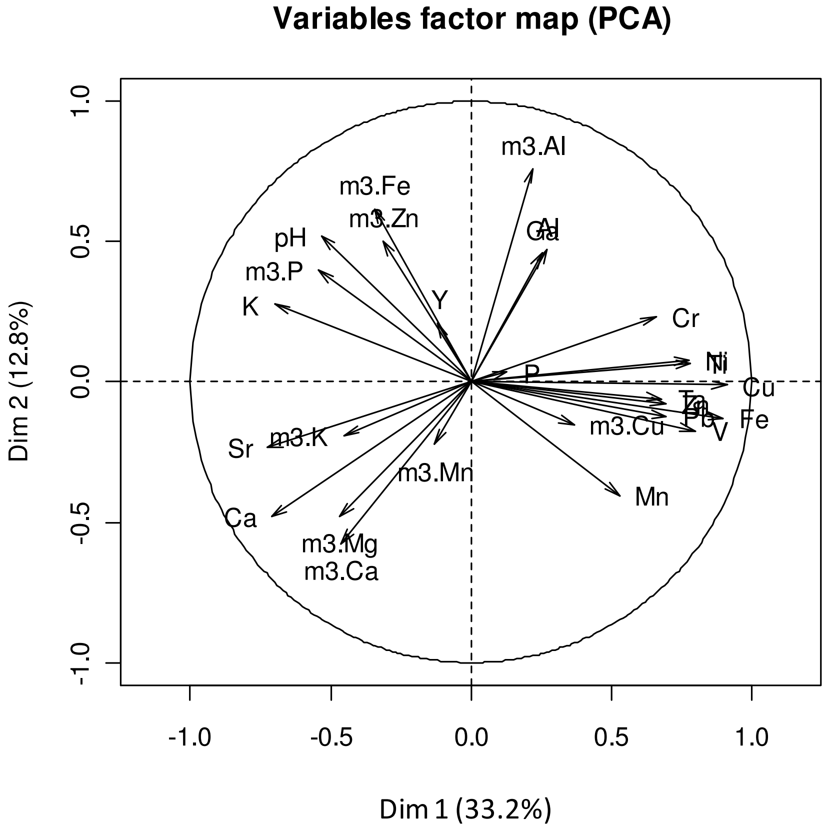


Figure 3.5: Biplot based on the principal component Dim 1 vs Dim 2, on the log transformed data of the soil total element concentration and Mehlich-3 extractable micronutrients data from all sites analysed.

Table 3.4: Correlation matrix of pH, TXRF water-extractable element concentrations and Mehlich 3 acid-extractable (coupled with ICP-MS) element concentrations.

| | Al | P | K | Ca | Ti | V | Cr | Mn | Fe | Ni | Cu | Zn | Ga | Sr | Y | Ta | Pb | pH | m3.Ca | m3.K | m3.Mg | m3.P | m3.Al | m3.Cu | m3.Fe | m3.Mn | m3.Zn |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Al | 1 | 0.48 | -0.15 | 0.11 | 0.74 | 0.71 | 0.50 | 0.61 | 0.78 | 0.70 | 0.72 | 0.81 | 0.88 | 0.03 | 0.30 | 0.69 | 0.72 | -0.25 | 0.21 | 0.30 | 0.31 | -0.32 | 0.78 | 0.53 | 0.01 | 0.12 | 0.18 |
| P | | 1 | -0.22 | 0.06 | 0.60 | 0.51 | 0.18 | 0.58 | 0.55 | 0.37 | 0.49 | 0.62 | 0.40 | 0.06 | 0.26 | 0.41 | 0.49 | -0.03 | 0.17 | 0.32 | 0.14 | 0.09 | 0.45 | 0.47 | 0.11 | 0.27 | 0.31 |
| K | | | 1 | 0.51 | -0.49 | -0.44 | -0.51 | -0.25 | -0.46 | -0.42 | -0.53 | -0.25 | -0.03 | 0.70 | 0.14 | -0.25 | -0.32 | 0.31 | 0.13 | 0.30 | 0.19 | 0.27 | -0.29 | -0.26 | -0.01 | -0.08 | -0.03 |
| Ca | | | | 1 | -0.07 | 0.03 | -0.14 | 0.25 | 0.02 | -0.03 | -0.04 | 0.21 | 0.16 | 0.82 | 0.18 | 0.04 | -0.04 | 0.55 | 0.73 | 0.51 | 0.71 | 0.19 | -0.14 | 0.21 | 0.11 | 0.21 | 0.21 |
| Ti | | | | | 1 | 0.84 | 0.63 | 0.73 | 0.89 | 0.71 | 0.87 | 0.80 | 0.57 | -0.18 | 0.35 | 0.81 | 0.78 | -0.14 | 0.26 | 0.20 | 0.28 | -0.29 | 0.68 | 0.68 | 0.14 | 0.27 | 0.24 |
| V | | | | | | 1 | 0.67 | 0.70 | 0.92 | 0.73 | 0.88 | 0.76 | 0.53 | -0.13 | 0.24 | 0.81 | 0.78 | -0.13 | 0.26 | 0.15 | 0.29 | -0.39 | 0.66 | 0.65 | 0.11 | 0.20 | 0.20 |
| Cr | | | | | | | 1 | 0.41 | 0.70 | 0.80 | 0.77 | 0.51 | 0.36 | -0.26 | -0.05 | 0.55 | 0.53 | -0.26 | 0.12 | 0.05 | 0.19 | -0.42 | 0.50 | 0.55 | 0.07 | 0.11 | 0.20 |
| Mn | | | | | | | | 1 | 0.82 | 0.55 | 0.77 | 0.80 | 0.45 | 0.06 | 0.31 | 0.64 | 0.69 | 0.12 | 0.39 | 0.30 | 0.36 | -0.14 | 0.49 | 0.69 | 0.04 | 0.69 | 0.28 |
| Fe | | | | | | | | | 1 | 0.74 | 0.94 | 0.85 | 0.56 | -0.14 | 0.21 | 0.85 | 0.85 | -0.15 | 0.24 | 0.20 | 0.29 | -0.36 | 0.69 | 0.70 | 0.09 | 0.33 | 0.26 |
| Ni | | | | | | | | | | 1 | 0.83 | 0.68 | 0.59 | -0.14 | 0.20 | 0.54 | 0.61 | -0.27 | 0.19 | 0.25 | 0.32 | -0.35 | 0.75 | 0.63 | 0.03 | 0.21 | 0.21 |
| Cu | | | | | | | | | | | 1 | 0.80 | 0.51 | -0.22 | 0.22 | 0.74 | 0.79 | -0.16 | 0.26 | 0.17 | 0.29 | -0.36 | 0.67 | 0.77 | 0.09 | 0.35 | 0.24 |
| Zn | | | | | | | | | | | | 1 | 0.73 | 0.05 | 0.28 | 0.73 | 0.69 | -0.11 | 0.33 | 0.38 | 0.40 | -0.18 | 0.69 | 0.71 | 0.16 | 0.33 | 0.33 |
| Ga | | | | | | | | | | | | | 1 | 0.11 | 0.33 | 0.49 | 0.47 | -0.30 | 0.13 | 0.35 | 0.26 | -0.24 | 0.77 | 0.42 | -0.03 | 0.02 | 0.14 |
| Sr | | | | | | | | | | | | | | 1 | 0.12 | -0.06 | -0.15 | 0.44 | 0.46 | 0.46 | 0.51 | 0.28 | -0.19 | 0.03 | 0.09 | 0.08 | 0.14 |
| Y | | | | | | | | | | | | | | | 1 | 0.20 | 0.24 | 0.19 | 0.27 | 0.17 | 0.28 | -0.06 | 0.30 | 0.23 | -0.03 | 0.16 | -0.05 |
| Ta | | | | | | | | | | | | | | | | 1 | 0.72 | -0.05 | 0.28 | 0.23 | 0.35 | -0.36 | 0.57 | 0.59 | 0.15 | 0.20 | 0.19 |
| Pb | | | | | | | | | | | | | | | | | 1 | -0.17 | 0.18 | 0.12 | 0.20 | -0.33 | 0.60 | 0.59 | 0.06 | 0.29 | 0.24 |
| pH | | | | | | | | | | | | | | | | | | 1 | 0.65 | 0.27 | 0.49 | 0.13 | -0.48 | 0.08 | -0.15 | 0.38 | 0.03 |
| m3.Ca | | | | | | | | | | | | | | | | | | | 1 | 0.49 | 0.85 | 0.03 | -0.02 | 0.51 | 0.14 | 0.41 | 0.23 |
| m3.K | | | | | | | | | | | | | | | | | | | | 1 | 0.61 | 0.25 | 0.26 | 0.37 | 0.01 | 0.28 | 0.38 |
| m3.Mg | | | | | | | | | | | | | | | | | | | | | 1 | -0.06 | 0.11 | 0.48 | 0.16 | 0.31 | 0.21 |
| m3.P | | | | | | | | | | | | | | | | | | | | | | 1 | -0.22 | -0.18 | 0.23 | 0.04 | 0.32 |
| m3.Al | | | | | | | | | | | | | | | | | | | | | | | 1 | 0.47 | 0.05 | 0.05 | 0.17 |
| m3.Cu | | | | | | | | | | | | | | | | | | | | | | | | 1 | 0.25 | 0.50 | 0.36 |
| m3.Fe | | | | | | | | | | | | | | | | | | | | | | | | | 1 | -0.07 | 0.28 |
| m3.Mn | | | | | | | | | | | | | | | | | | | | | | | | | | 1 | 0.24 |
| m3.Zn | | | | | | | | | | | | | | | | | | | | | | | | | | | 1 |

### 3.4.7 Relationship with mineralogy and other site characteristics

In order to confirm whether site or soil-forming factors (e.g. mineralogy, climate, topography and vegetation) are important drivers of total elemental concentrations in the soil, we performed a Random Forests (RF) regression of the factors against the first 5 principal components of the TXRF element concentration. The RF multiple regression forests allowed us to view the importance of the predictor variables and below we discuss only the top four variables in decreasing order of importance for the first 3 principal components or dimensions which explained as much as 57% of the variability of the total element concentrations.

#### 3.4.7.1 Element concentration fingerprints *versus* site factors

First, a regression of 6 predictor variables namely soil class, climate zones, cluster, plot flooding, depth and position in the topography showed that the OOB-models had an overall accuracy ($R^2$) of 0.59, 0.44 and 0.43 for dimensions 1, 2 and 3 respectively (Figure 3.6a). The most important predictor variables explaining the variation observed in the first principal component or dimension 1 of the total element concentration data, were soilclass (based on the IUSS Working Group WRB (2007) and the FAO/EC/ISRIC (2003) world soil resources map (Scale 1:30 000 000 approx)), climate zone (Köppen-Geiger climatic zones), position in the topography (e.g. bottomland, footslope, midslope, upland) followed by whether the plot is flooded or not (Figure 3.6a). For the dimension 2, the most important variables were soilclass, climate zone, position in the toposequence, and plot flooding, while for dimension 3 the four most important variables were similar to those in the second dimension (Figure 3.6a). When we excluded the climate zone predictor variable in the regression model and in place of this included total annual precipitation and average montly temperature, the regression results showed an overall improved fit ($R^2$) of the models' overall accuracy of 0.87, 0.74 and 0.70 for dimensions 1, 2 and 3 c (Figure 3.6b). In addition, the order of the most important predictor variables explaining the variation observed were also changed when climate zone was removed from the model such that for dimension 1 they were soil class, precipitation, average temperature, followed by position in the topography (Figure 3.6b). Topography is the most commonly considered soil-forming factor and expresses the variation of total element concentration pattern due to modification of the water relationships in soils and it influences soil erosion to a considerable extent (Jenny, 1941).

Since the soil class and climate zone ranked the top most important variables explaining the varaiation of the total element concentrations' main pattern of variation in Dim 1, these were both excluded in the RF model and the predictor variables precipitation, average temperature retained and consequently the model accuracies were slightly decreased to a $R^2$ of 0.79, 0.69, and 0.64 for dimension 1, 2 and 3 respectively (Figure 3.6c). For the dimensions 1, 2 and 3, the two most important variables were then precipitation and temperature followed by position in the topography and cluster (Figure 3.6c). When the new variable 'landuse' was introduced in addition to the soil classification according to the FAO/EC/ISRIC (2003) and IUSS Working Group WRB (2007), the overall model accuracies were further marginally improved to $R^2$ values of 0.87, 0.77 and 0.70 for dimensions 1, 2 and 3 respectively (Figure 3.6d). Excluding the soil class from the model used in Figure 6d the fit of the models in dimensions 1-3 were 0.81, 0.70 and 0.64 $R^2$ respectively (Figure 3.6e). The top most important variable for all the dimensions was temperature followed by precipitation, topography and landuse in dimensions 1 and 2 while in dimensions 3, after temperature, the most important variables were precipitation, landuse followed by topography (Figure 3.6e). Our results that landuse was one of the important variables, e.g. ranking fourth, fourth, and third in dimensions 1, 2 and 3 respectively (Figure 3.6e), supported results by Voortman (2011) that the presence of different vegetation types is a reliable and precise indicator of differences in soil chemical properties and that interactions among nutrients significantly explain differences in vegetation and also the distribution of vegetation types. These results are interesting in that they imply that >70% of variation in soil elemental composition patterns can be predicted using information in existing databases or readily observable features. The predictor 'soil depth' was of least importance for total element concentration pattern of variation prediction, suggesting that neither top- nor subsoils influence much the variation of element concentration probably due to expected correlations of element concentrations between the two depths.

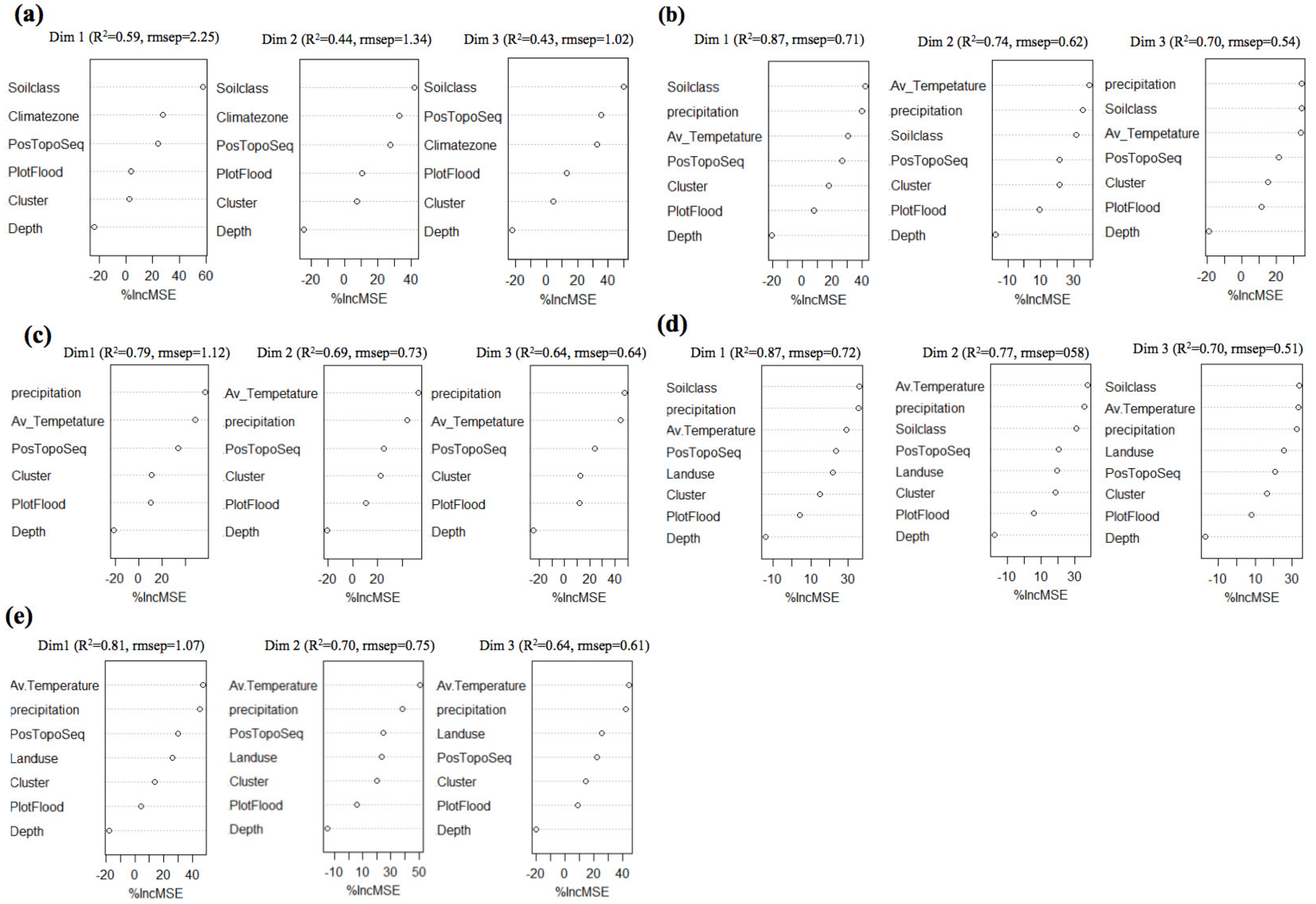


Figure 3.6: Variable importance plots for the first 3 dimensions of the Random Forests regression of TXRF element concentrations against mineralogy and site/soil-forming factors showing the model accuracies and mean decrease in accuracy (%IncMSE) for site or soil-forming predictor variables (a) excluding precipitation and temperature (b) excluding climate zone but including precipitation and temperature (c) excluding climate zones and soil class but including precipitation and temperature (d) including soil classification and landuse (e) excluding climate zones and soil class but including precipitation, temperature and landuse (see Table 3.1 for predictor parameter description).

### 3.4.7.2 Element concentration *versus* mineralogy composition

Mineralogy data from XRD analysis were extracted for each site and for our analysis we used both raw semi-quantitative mineralogy data and the dominant mineralogy grouping. In general RF model performances were acceptable with prediction accuracies for dimensions 1-3 ranging from 0.71-0.46 and 0.81-0.56 for raw mineralogy data and dominant mineralogy groupings respectively (Figure 3.7a-b). These results suggest that patterns of variation in total element concentrations are largely due to high variability in mineralogy. The most important predictor variables explaining the variation in the dominant mineralogy grouping for dimension 1 were K-Feldspars followed by kaolinite/1:1 clays, then quartz and plagioclase (Figure 3.7a). Thus, since K, Ca, Ti, Fe and Sr were predominantly associated with Dim 1, the contents of these elements are presumed to be originating primarily from parent material as reported by Kabata-Pendias and Murkhejee (2007) and US.EPA (2006). Lower concentrations of certain elements, e.g. Ca and K in some soils in our study, could also be explained by a reported breakdown of primary minerals, particularly K-feldspars and plagioclase (Acosta et al., 2011) or depletion in well-drained soils over long periods of pedogenic weathering since they are either divalent or monovalent (Marques et al., 2004). For dimension 2 the most important variables were quartz, kaolinite/1:1 clays, K-feldspars, and oxides (Figure 3.7a). Because Al, Ni and Ga were dominant in Dim 2 (Figure 3a, Annex Figure A3.1), our results are in agreement with those reported by Acosta et al. (2011) that Al in soils has been attributed to the formation of clay minerals. Total Al in soil reflects the type of soil and the underlying geology being present in the matrix of clays and other silicate minerals and highly weathered soils are often high in Al concentrations (Rawlins et al., 2012). However, at the opposite extreme, low Al is a marker of organic-rich soils, which contain a smaller proportion of aluminosilicate minerals (Rawlins et al., 2012). For dimension 3 the most important variables were oxides followed by K-feldspars, plagioclase and kaolinite/1:1 clays (Figure 3.7a). While our results showed that Cr, Mn and Zn were predominant in the Dim 3 (Figure 3.3b, Annex Figure A3.1), Zn is reported to be generally associated with Al- and Fe-containing minerals such as feldspars, micas, pyroxenes and amphiboles (Acosta et al., 2011).

The most important predictor variables explaining the variation in the raw mineralogy for dimension 1 were microcline, followed by hematite, kaolinite and albite; for dimension 2 were microcline, calcite, hematite and hornblende; while for dimension 3 were hematite, tridymite, albite and nacrite (Figure 3.7b). Our findings were also in agreement with a report by Voortman (2011) that the nature of the key variables explaining the variability of soils are related to the mineralogy of parent rock and thus we also infer that the mineralogy of parent rock is a principal factor determining spatial patterns of soil total element concentrations.

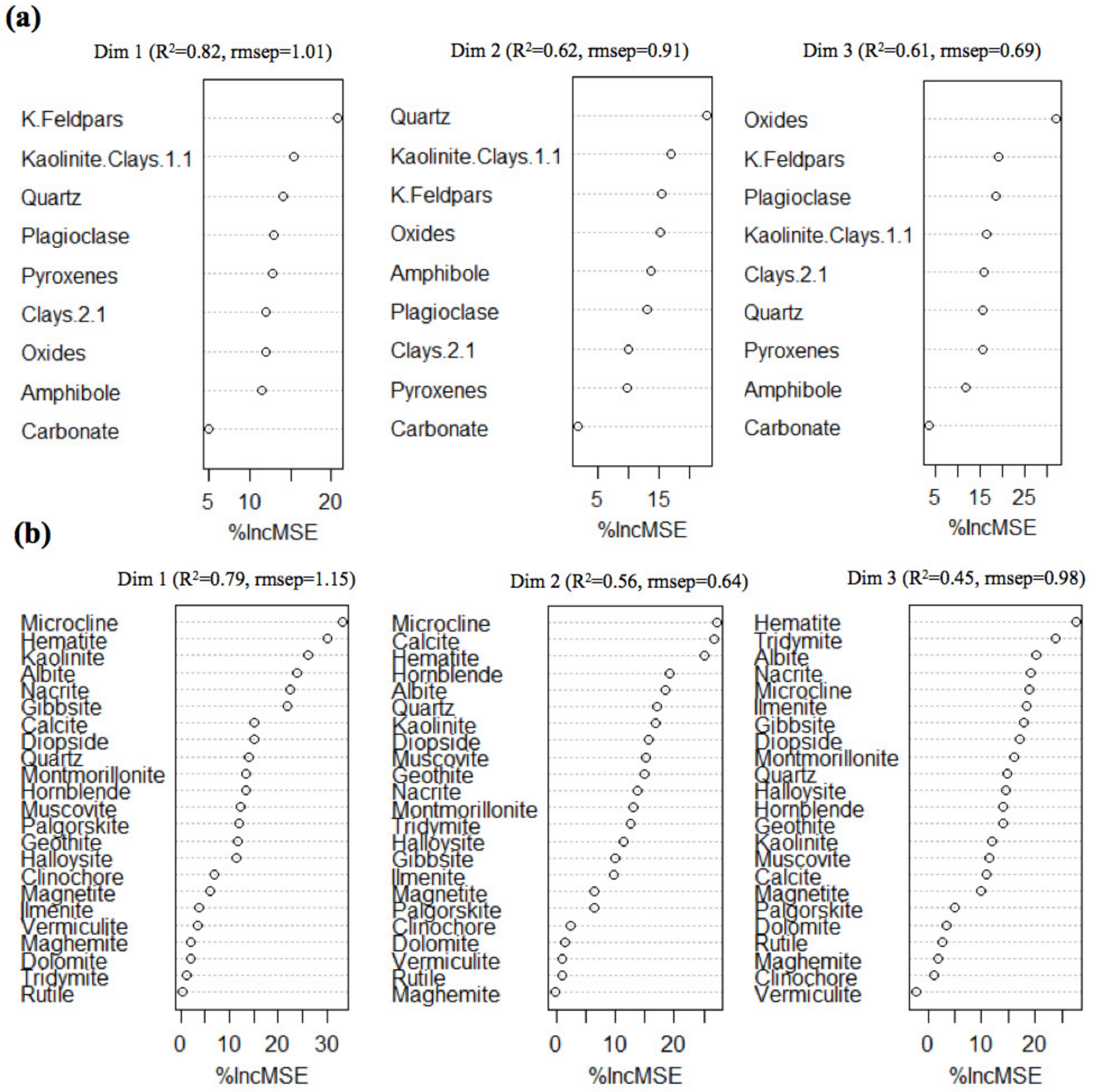


Figure 3.7: Variable importance plots for the first 3 dimensions of the Random Forests regression of TXRF element concentrations against mineralogy and site/soil-forming factors showing the model accuracies and mean decrease in accuracy (%IncMSE) for (a) dominant mineralogy groupings, (b) raw semi-quant mineralogy data.

### 3.4.8 Element concentration *versus* mineralogy composition plus site and soil-forming factors

Since some site factors and mineralogy investigated in this study separately had a high explanatory power of the patterns of variations of the total elementsl concentrations we tested how much variation could be explained when they were combined. In general RF model performances were improved compared to separate models with prediction accuracies of 0.92, 0.84, and 0.79 for dimensions 1, 2 and 3 (Figure 3.8a). The most important variables explaining the main patterns of variation in total element concentrations were cluster, landuse, topography, temperature and precipitation (Figure 3.8a). Thus all soil-forming factors (e.g., parent material, climate, topography, management (landuse)) were demonstrated to have an important influence on total elemental concentrations in the soil. The importance of cluster can be explained by spatial correlation at distances of less than 1 km. Since cluster was the only non-readily observable variable, we tested how predictive performance was reduced or not by leaving out cluster. In general, the model performances were slightly reduced with prediction accuracies of 0.90, 0.80 and 0.75 for dimensions 1, 2 and 3 respectively (Figure 3.8b). Thus, our exploratory analyses using RF regression of the factors against the first 5 principal components of the TXRF element concentration confirmed that site or soil-forming factors (e.g., mineralogy, climate, topography, vegetation and landuse) are important drivers of total elemental concentrations in the soil.

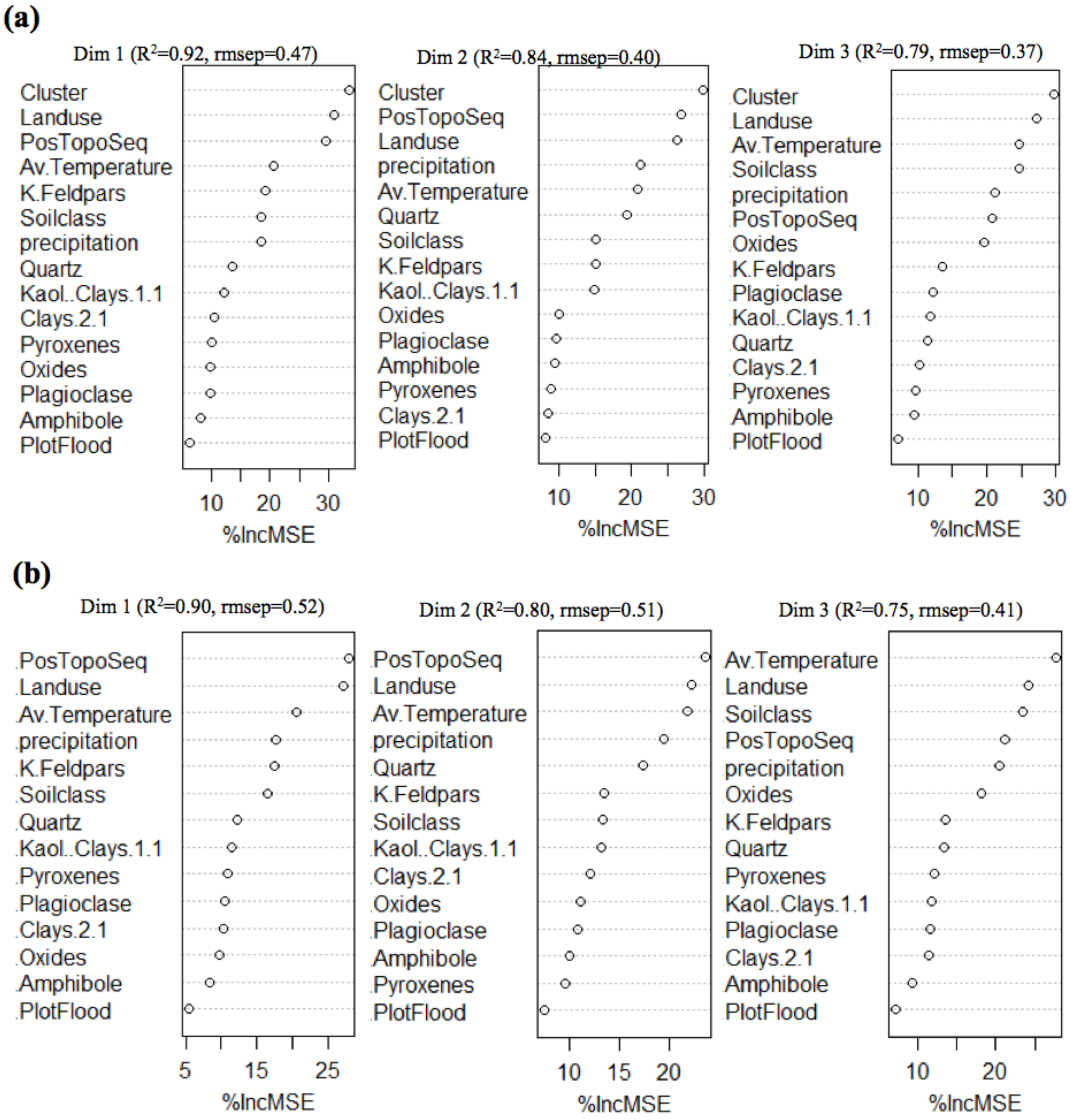


Figure 3.8: Variable importance plots for the first 3 dimensions of the Random Forests regression of TXRF element concentrations against mineralogy and site/soil-forming factors showing the model accuracies and mean decrease in accuracy (%IncMSE) for combination of mineralogy and site factors (a) including cluster and (b) excluding cluster.

### 3.5 Conclusions

Knowledge of the concentration and spatial distribution of naturally occurring elements in the soils of Sub-Saharan Africa (SSA) is limited and there is a need for a better understanding of the factors that might regulate this variation. This study helped to establish the concentrations of 17 elements for soils occurring within 34 sentinel sites across SSA and document patterns in total element concentrations within and between sites, which appeared to relate to differences in mineralogical 'functional groups'. We observed strong within site and between site variations in many elements and which were expected to be highly diagnostic of soil fertility potential. TXRF can provide chemical fingerprinting which could be further tested for inferring soil chemical and physical functional properties, which is of interest in the African soil contexts for agricultural and environmental management. Our exploratory analysis using PCA with a view to test for associations within multiple variables and to narrow down the predictor variables was only used in a quantitative way to show how far TXRF total element derived patterns in soils relate to 'available' element results from Mehlich-3 soil tests (acid-extractable nutrients) and hence to soil nutrient supply capacity. The associations presented in the present study look promising and future studies should go ahead to develop predictive models as to what extent and for which elements total element concentrations relate to nutrient supply/availability.

Our exploratory analyses of the relationships between element composition data and other site factors using Random Forests have demonstrated that all site or soil-forming factors (e.g., mineralogy, climate, topography, vegetation and landuse) have an important influence on total elemental concentrations in the soil. The fact that the soil-forming factors are related to the concentration of the naturally coccuring elements in the soil gives rise to the notion that they might be predicted from the soils' element composition. Our results also implied that >70% of variation in soil element composition patterns can be predicted using information in existing databases or readily observable features. Thus, future studies should investigate the feasibility of quantitatively predicting soil functional properties from concentrations of elements and soil-forming factors e.g. for digital soil mapping.

This work shows the utility of TXRF to provide improved capabilities for analytical services on soil total element concentrations, due to the high-throughput rates of analyses with regards to costs, time and the number of samples that can be analyzed per day. Successful use of the TXRF technique would open up the possibilities for using total element composition to improve global predictions of soil properties, such as cation exchange capacity and extractable nutrients as demonstrated here with some acceptable relations with Mehlich-3. Therefore, it is worth exploring the possibility of using the TXRF method to measure element concentrations in other media such as soil water extracts and testing whether extractable data could be used as a complementary input to pedotransfer functions (PTFs) for low cost, rapid prediction of soil fertility properties. In particular, after careful testing TXRF could be a powerful complement to other rapid analytical techniques such as infrared spectroscopy for science-based diagnostic surveillance in agricultural and environmental management.

### 3.6 Acknowledgement

The research work was supported by the World Agroforestry Centre (ICRAF) through the Africa Soil Information Service (AfSIS) project funded by the Bill &Melinda Gates Foundation, and by the Dr. Hermann Eiselen Ph.D. Grant from the Foundation Fiat Panis. E. Towett acknowledges the Young Excellence Scholars (YES) PhD Scholarship from the Food Security Centre at the University of Hohenheim, Germany, a travel grant as well as stipend within the framework of the “Global Food Security” funded by the Deutscher Akademischer Austausch Dienst (DAAD). The authors acknowledge the AfSIS field crew who provided the soil samples, Robin Chacha and Hezekiah Nyandika for assistance with TXRF analyzing the soil samples, Beatrice Oware and Bella Kauma for assistance with XRD data semi-quantification and analysis, as well as other technical staff of the ICRAF Soil-Plant Spectral Diagnostics Laboratory for assistance with sample preparation. The authors are also grateful to Ric Coe, Tor-Gunnar Vågen, Anja Gassner, Hans-Peter Piepho, Joseph Ogutu, Hagen Stosnach, Michael Gatari, Ermias Aynekulu, Juan Laso, Kefyalew Sahle and Todd Rosenstock for their valuable advice and help with data analyses that helped improve this manuscript.

### 3.7 References

1. Acosta, J.A., Martínez-Martínez, S., Faz, A., Arocena, J., 2011. Accumulations of major and trace elements in particle size fractions of soils on eight different parent materials. Geoderma 161, 30-42.
2. Adebiyi, F.M., Olise, F.S., Owoade, O.K., Obiah, I.B., Ajayi, T.R., Olaniyi, H.B., Asubiojo, I.O., 2005. TXRF characterization of EDTA extractable metals from soils of bituminous sands occurrence area, Agbabu, South Western Nigeria. Journal of Applied Sciences 5 (7), 1151-1157.
3. AfSIS, 2013. Africasoils.net [Internet]. Africa Soils Information Service (AfSIS). Cited 2013 Jun 20. Available from: www.africasoils.net/.
4. Antwi-Agyei, P., Hogarh, J.N., Foli, G., 2009. Trace elements contamination of soils around gold mine tailings dams at Obuasi, Ghana. African Journal of Environmental Science and Technology 3 (11), 353-359.
5. Breiman, L., 2001. Random Forests, Machine Learning 45(1), 5-32.
6. Bruker, 2007a. Lab Report XRF 426: S2 PICOFOX Total Reflection X-ray Fluorescence Spectroscopy - Working Principles. Bruker AXS Microanalysis GmbH, Berlin, Germany.
7. Bruker, 2007b. S2 Picofox user manual. Doc-M81-EXX001-V1. Bruker AXS Microanalysis GmbH, Berlin, Germany.
8. Bruker, 2009. Lab Report XRF 435: S2 PICOFOX - Why Wait for Sample Prep! Shortest Time-to-Result through TXRF. Bruker AXS Microanalysis GmbH, Berlin, Germany.
9. Cannon, W.F., Horton, J.D., 2009. Soil geochemical signature of urbanization and industrialization – Chicago, Illinois, USA. Applied Geochemistry 24, 1590-1601.
10. FAO/EC/ISRIC, 2003.World Soil Resources Map.
11. Franklin, R.E., Duis, L., Smith, B.R., 2006. Mehlich extractable and total element concentrations in South Carolina Soils. Communications in Soil Science and Plant Analysis 37 (5-6), 679-691.
12. Grunsky, E.C., Drew, L.J., Sutphin, D.M., 2009. Process recognition in multi-element soil and stream-sediment geochemical data. Applied Geochemistry 24, 1602-1616.
13. Helmke, P.A., 2000. The chemical composition of soils, in: Sumner, M.E., (ed.), Handbook of Soil Science. CRC Press, Boca Raton, FL, pp. B3-B24.
14. Hooda, P.S., 2010. Trace elements in soils. John Wiley and Sons, UK.

15. Haluschak, P., Eilers, R.G., Mills, G.F., Grift, S., 1998. Status of Selected Trace Elements in Agricultural Soils of Southern Manitoba. Technical Report 1998-6E Land Resource Unit, Brandon Research Centre, Research Branch, Agriculture and Agri-Food Canada.
16. Herselman, J.E., 2007. The concentration of selected trace metals in South African soils. PhD Thesis, University of Stellenbosch, South Africa.
17. 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, 1965-1978.
18. IUSS Working Group WRB, 2007. World reference base for soil resources 2006, first update 2007. World Soil Resources reports No. 103. FAO, Rome.
19. Jenny, H., 1941. Factors of soil formation: a system of quantitative pedology. McGraw-Hill, New York.
20. Kabata-Pendias, A., Mukherjee, A.B., 2007. Trace elements from soil to human. Springer-Verlag Berlin Heidelberg.
21. Klockenkämper, R., 1997. Total reflection X-ray analysis. Wiley, New York.
22. Kottek, M., Grieser, J., Beck, C., Rudolf, B., Rubel, F., 2006: World Map of the Köppen-Geiger climate classification updated. Meteorologische Zeitschrift 15, 259-263. DOI: 10.1127/0941-2948/2006/0130.
23. Laiho, R., Penttilä, T., Laine, J., 2004. Variation in soil nutrient concentrations and bulk density within peatland forest sites. Silva Fennica 38 (1), 29-41.
24. Lê, S., Josse, J., Husson, F., 2008. FactoMineR: An R Package for Multivariate Analysis. Journal of Statistical Software 25 (1), 1-18.
25. Marques, J.J., Schulze, D.G., Curia, N., Mertzman, S.A., 2004. Trace element geochemistry in Brazilian Cerrado soils. Geoderma 121, 31-43.
26. Marzecová, A., Mikomägia, A., Koff, T., Martmac, T., 2011. Sedimentary geochemical response to human impact on Lake Nõmmejärv, Estonia. Estonian Journal of Ecology 60 (1), 54-69. doi: 10.3176/eco.2011.1.05.
27. Misra, N.L., Dhara, S., Ovari, M., Zaray, G., Aggarwal, S.K., Varga, I., 2010. Determination of low atomic number elements at trace levels in uranium matrix using vacuum chamber total reflection X-ray fluorescence. Spectrochimica Acta Part B 65 (6), 457-460.

28. Mukhtar, S., Haswell, S.J., 1991. Application of total-reflection X-ray fluorescence spectrometry to elemental determinations in water, soil and sewage sludge samples. Analyst 116, 333-338.
29. Polgári, Z., Szoboszlai, N., Óvári, M., Záray, G., 2011. Possibilities and limitations of the total reflection X-ray fluorescence spectrometry for the determination of low Z elements in biological samples. Microchemical Journal 99, 339-343.
30. Qian, J., Shan, X., Wang, Z., Tu, Q., 1996. Distribution and plant availability of heavy metals in different particle-size fractions of soil. Science of the Total Environment 187, 131-141.
31. R Development Core Team, 2013. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. Available from: http://www.R-project.org.
32. Rawlins, B.G., Webster, R., Lister, T.R., 2003. The influence of parent material on top soil geochemistry in eastern England. Earth Surface Processes and Landforms 28, 1389–1409.
33. Rawlins, B.G., Webster, R., Tye, A.M., Lawley, R., O'hara, S.L., 2009. Estimating particle-size fractions of soil dominated by silicate minerals from geochemistry. European Journal of Soil Science 60, 116-126.
34. Rawlins, B.G., McGrath, S.P., Scheib, A.J., Breward, N., Cave, M., Lister, T.R., Ingham, M., Gowing, C., Carter, S., 2012. The advanced soil geochemical atlas of England and Wales. British Geological Survey, Keyworth. www.bgs.ac.uk/gbase/advsoilatlasEW.html.
35. Reeves III, J.B., Smith, B.D., 2009. The potential of mid- and near-infrared diffuse reflectance spectroscopy for determining major- and trace-element concentrations in soils from a geochemical survey of North America. Applied Geochemistry 24 (8), 1472-1481.
36. Rodríguez, J.A., Nanos, N., Grau, J.M., Gil, L., López-Arias, M., 2008. Multiscale analysis of heavy metal contents in Spanish agricultural topsoils. Chemosphere 70, 1085-1096.
37. Rousseau, R.M., 2001. Detection limit and estimate of uncertainty of analytical XRF results. The Rigaku Journal 18 (2), 33-47.
38. SAS Institute, 2011. SAS 9.3 Software. SAS Institute Inc., USA.
39. Sánchez-Moreno, R.A., Gismera, M.J., Sevilla, M.T., Procopio, J.R., 2010. Direct and rapid determination of ultratrace heavy metals in solid plant materials by ET-AAS ultrasonic-assisted slurry sampling. Phytochemical Analysis 21, 340-347.

40. Schulthess, C.P., 2011. Historical perspective on the tools that helped shape soil chemistry. Soil Science Society of America Journal 75, 1-28. doi:10.2136/sssaj2011.0142
41. Shepherd, K.D., 2010. Soil spectral diagnostics – infrared, X-ray and laser diffraction spectroscopy for rapid soil characterization in the Africa Soil Information Service. 19th World Congress of Soil Science, Soil Solutions for a Changing World 1-6 August 2010, Brisbane, Australia.
42. Stosnach, H., 2005. Environmental trace-element analysis using a benchtop total reflection X-ray fluorescence spectrometer. The Japan Society of Analytical Chemistry. Analytical Sciences 21, 873-876.
43. Stosnach, H., 2007. Lab Report XRF 426. S2 PICOFOX Total Reflection X-ray Fluorescence Spectroscopy - Working Principles. Bruker AXS Microanalysis GmbH.
44. Tarsoly, G., Ovari, M., Zaray, G., 2010. Determination of fluorine by total reflection X-ray fluorescence spectrometry. Spectrochimica Acta Part B 65 (4), 287-290.
45. Towett, E.K., Shepherd, K.D., Cadisch, G., 2013. Quantification of total element concentrations in soils using total X-ray fluorescence spectroscopy (TXRF). Science of the Total Environment 463–464: 374–388. http://dx.doi.org/10.1016/j.scitotenv.2013.05.068.
46. US.EPA, 2006. Innovative technology verification report. XRF Technologies for measuring trace elements in soil and sediment, Rontec PicoTAX XRF analyzer. United States Environmental Protection Agency, Office of Research and Development, Washington DC, EPA/540/R-06/005.
47. Vågen, Tor-G., Winowiecki, L.A., Tondoh, J.E., Desta, L.T., 2013-01-01. Africa Soil Information Service (AfSIS) - Soil Health Mapping. http://hdl.handle.net/1902.1/19793 V2 [Version].
48. von Bohlen, A., 2009. Total reflection X-ray fluorescence and grazing incidence X-ray spectrometry - tools for micro- and surface analysis. A review. Spectrochimica Acta Part B 64, 821-832.
49. Voortman, R.L., 2011. Explorations into African land resource ecology on the chemistry between soils, plants and fertilizers. PhD Dissertation. Faculty of Economics and Business, Vrije Universiteit, pp. 264.

## 3.8 Annex

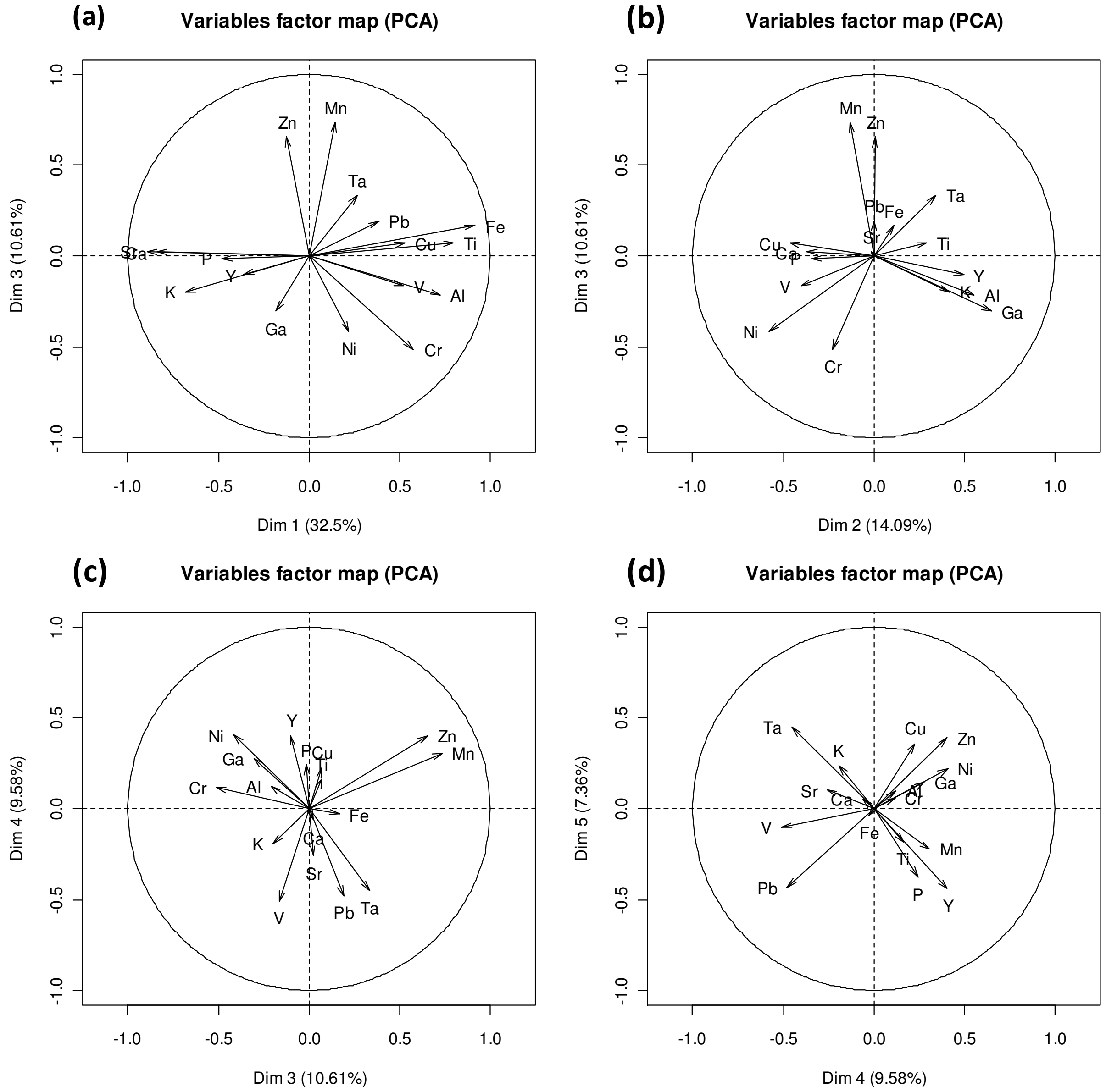


Figure A3.1: Biplots (arrow sizes are proportional to the “initial” variability in the elements present) based on the principal component for (a) Dim 1 vs Dim 3 (b) Dim 2 and Dim 3 (c) Dim 3 and Dim 4 (d) Dim 4 and Dim 5, on the log transformed data of the soil total element concentration from all sites analysed.

Table A3.1: Lower limit of detection (LLD) values (mg kg$^{-1}$) calculated for each element detected using the TXRF technique.

| Element | Al | P | K | Ca | Ti | V | Cr | Mn | Fe | Ni | Cu | Zn | Ga | Sr | Y | Ta | Pb |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| LLD | 208 | 42 | 8. | 4.14 | 1.62 | 1.13 | 1.10 | 0.88 | 0.87 | 0.31 | 0.22 | 0.20 | 0.22 | 0.18 | 0.18 | 0.12 | 0.35 |

Table A3.2. Summary of some of the important options invoked in the SAS PROC MIXED statement by function and the class level information for the three classes (site, cluster and depth) in the model of the data used in this study.

| Model Information | |
|---|---|
| Data Set | TXRF |
| Dependent Variable | log (element) |
| Covariance Structure | Variance Components |
| Estimation Method | REML |
| Residual Variance Method | Profile |
| Fixed Effects SE Method | Kenward-Roger |
| Degrees of Freedom Method | Kenward-Roger |

Class Level Information

| Class | Levels | Values |
|---|---|---|
| Site | 34 | Chinyanghuku, Dambidolo, Finnkolo, Fisenge, Fria, Ibi, Ihassunge, Imorun, Itende, Katsina Ala, Kiberashi, Kidatu, Kisongo, Koloko, Kontela, Kubeasi, Kutaber, Lambussie, Macassangila, Marafa, Mbalambala, Mbinga, Mega, Merar, Morijo, Nkhata Bay, Pandambili, Thuchila. |
| Cluster | 16 | 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16. |
| Depth | 2 | Subsoil, Topsoil. |

Table A3.3: The coefficient of variability (CoV in %) values for sample replicates calculated for each site using samples that had no missing values (before substituting the missing values with the 25th percentile value of the LLD) in the data after log transformation.

| Site | Country | Al | P | K | Ca | Ti | V | Cr | Mn | Fe | Ni | Cu | Zn | Ga | Sr | Y | Ta | Pb |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Ajumako | Ghana | 5 |  | 4 | 4 | 5 | 27 | 9 | 5 | 4 | 6 | 7 | 6 | 7 | 4 |  |  |  |
| Bondigui | Guinea | 8 |  | 7 | 6 | 11 | 12 | 11 | 9 | 10 | 10 | 11 | 10 |  | 11 |  | 16 | 18 |
| Boumeoul | Tanzania | 6 |  | 4 | 5 | 5 | 19 | 5 | 6 | 4 | 8 | 4 | 6 | 4 | 4 | 12 | 24 | 10 |
| Chica - b | Mozambique | 6 |  | 6 | 6 | 6 |  | 19 | 6 | 7 |  | 8 | 10 | 8 | 8 | 17 |  |  |
| Chiculecule | Mozambique | 7 |  | 5 | 6 | 5 | 33 | 11 | 5 | 5 | 6 | 7 | 7 | 5 | 10 | 34 |  |  |
| Chinyanghuku | Tanzania | 7 |  | 6 | 7 | 6 | 11 | 8 | 7 | 5 | 6 | 6 | 6 | 6 | 7 | 11 | 11 | 23 |
| Dambidolo | Ethiopia | 7 |  | 6 | 7 | 7 | 14 | 16 | 8 | 8 | 7 | 8 | 8 | 7 | 9 | 10 | 16 | 22 |
| Finkolo | Mali | 5 |  | 3 | 4 | 5 | 9 | 4 | 4 | 3 | 5 | 7 | 5 | 4 | 3 | 16 | 11 | 11 |
| Fisenge | Zambia |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Fria | Guinea | 8 |  | 4 | 7 | 8 | 7 | 8 | 10 | 6 | 6 | 6 | 5 | 5 | 5 |  | 15 | 12 |
| Ibi | Nigeria | 7 |  | 4 | 4 | 5 | 9 | 7 | 4 | 4 | 8 | 5 | 8 | 5 | 5 | 17 | 13 | 8 |
| Ihassunge | Mozambique | 8 |  | 5 | 7 | 7 | 36 | 16 | 8 | 5 | 9 | 11 | 7 | 6 | 7 | 16 | 17 | 11 |
| Imorum | Nigeria | 12 |  | 5 | 4 | 5 | 24 | 6 | 5 | 4 | 6 | 7 | 5 | 4 | 4 | 21 | 13 | 7 |
| Itende | Tanzania | 7 |  | 6 | 6 | 6 | 30 | 8 | 5 | 5 | 8 | 8 | 8 | 6 | 6 | 13 |  |  |
| Katsina ala | Nigeria | 7 |  | 4 | 5 | 5 |  | 13 | 5 | 4 | 8 | 5 | 5 | 7 | 6 | 16 |  |  |
| Kiberashi | Tanzania | 9 |  | 6 | 5 | 4 | 22 | 8 | 5 | 4 | 5 | 6 | 4 | 5 | 5 | 15 | 16 | 18 |
| Kidatu | Tanzania | 9 |  | 6 | 5 | 6 | 29 | 12 | 6 | 5 | 14 | 8 | 10 | 6 | 6 | 16 |  |  |
| Kisongo | Tanzania | 8 | 25 | 5 | 5 | 6 | 28 |  | 7 | 5 | 10 | 6 | 5 | 7 | 5 | 9 | 16 | 13 |
| Koloko | Mali | 9 |  | 7 | 5 | 12 | 16 | 9 | 11 | 12 | 11 | 11 | 12 | 10 | 13 | 19 | 23 | 22 |
| Kontela | Mali | 9 |  | 5 | 6 | 7 | 9 | 9 | 8 | 8 | 8 | 7 | 8 | 9 | 7 | 14 | 10 | 13 |
| Kubeasi | Ghana | 7 |  | 4 | 4 | 5 | 17 | 8 | 4 | 4 | 5 | 5 | 5 | 5 | 4 | 19 | 15 |  |
| Kutaber | Ethiopia | 5 |  | 3 | 3 | 4 | 12 | 9 | 5 | 4 | 4 | 4 | 4 | 5 | 5 | 12 | 17 | 20 |
| Lambussie | Ghana | 8 |  | 6 | 5 | 8 | 12 | 8 | 6 | 6 | 8 | 6 | 8 | 10 | 7 | 23 |  | 15 |
| Macassangila | Mozambique | 5 |  | 4 | 5 | 7 | 37 |  | 5 | 5 |  | 5 | 7 | 5 | 6 | 10 | 15 | 16 |
| Marafa | Kenya | 8 |  | 7 | 6 | 9 | 40 | 14 | 8 | 8 | 7 | 8 | 8 | 10 | 10 | 12 | 28 | 21 |
| Mbalambala | Kenya | 13 |  | 12 | 11 | 12 |  |  | 10 | 12 | 13 | 12 | 13 | 11 | 11 | 13 | 32 | 33 |
| Mbinga | Tanzania | 6 |  | 5 | 5 | 12 | 36 | 12 | 9 | 6 | 8 | 5 | 8 | 5 | 6 | 8 |  | 16 |
| Mega | Ethiopia |  |  | 11 | 5 | 93 |  |  | 43 | 75 | 50 | 52 | 58 |  | 6 |  |  |  |
| Merar | Ethiopia | 7 | 17 | 5 | 5 | 7 | 15 | 10 | 7 | 5 | 6 | 6 | 6 | 6 | 6 | 13 |  |  |
| Morijo | Kenya | 5 |  | 4 | 3 | 5 | 66 | 20 | 4 | 4 | 6 | 6 | 4 | 7 | 3 | 6 | 20 | 8 |
| Nkhata bay | Malawi | 6 |  | 5 | 7 | 6 |  |  | 7 | 6 | 8 | 7 | 7 | 7 | 7 | 13 | 15 |  |
| Pandambili | Tanzania | 6 |  | 6 | 5 | 6 | 19 | 6 | 5 | 6 | 7 | 6 | 5 | 6 | 7 | 11 | 17 | 19 |
| Thuchila | Malawi | 8 |  | 5 | 5 | 8 |  | 12 | 7 | 7 |  | 6 | 9 | 8 | 6 | 12 | 12 | 15 |

Table A3.4: 2.5th percentiles of the total element concentration values (mg $kg^{-1}$) of 1074 soil samples across 34 sentinel sites sampled for this study.

| Site | Country | Al | P | K | Ca | Ti | V | Cr | Mn | Fe | Ni | Cu | Zn | Ga | Sr | Y | Ta | Pb |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Ajumako | Ghana | 12756 | 42 | 1197 | 405 | 1985 | 8.1 | 22 | 97 | 8077 | 2.5 | 4.6 | 7.9 | 2.1 | 13 | 0.9 | 0.8 | 7.1 |
| Bondigui | Burkina Faso | 9872 | 42 | 1597 | 377 | 1786 | 13 | 23 | 97 | 6461 | 2.9 | 4.5 | 5.7 | 0.2 | 10 | 0.2 | 1.2 | 7.9 |
| Boumeoul | Guinea | 14576 | 42 | 1056 | 144 | 1900 | 10 | 34 | 37 | 6654 | 6.5 | 6.3 | 4.3 | 3.1 | 5.9 | 3.9 | 0.9 | 4.9 |
| Bukwaya | Tanzania | 9793 | 42 | 5634 | 317 | 468 | 5.2 | 3.6 | 39 | 2104 | 0.9 | 1.7 | 3.7 | 1.4 | 17 | 1.3 | 0.3 | 4.0 |
| Chica - b | Mozambique | 11229 | 42 | 4682 | 922 | 269 | 1.0 | 1.4 | 56 | 874 | 0.3 | 1.1 | 3.6 | 2.0 | 24 | 2.4 | 0.1 | 0.3 |
| Chiculecule | Mozambique | 4002 | 42 | 311 | 101 | 1062 | 2.7 | 42 | 43 | 1355 | 2.6 | 1.8 | 2.9 | 0.7 | 1.2 | 0.8 | 0.1 | 1.5 |
| Chinyanghuku | Tanzania | 32851 | 42 | 11158 | 11875 | 1448 | 20 | 21 | 196 | 13844 | 4.8 | 5.4 | 17 | 11 | 246 | 3.5 | 0.9 | 3.1 |
| Dambidolo | Ethiopia | 36695 | 42 | 4467 | 4016 | 1359 | 8.4 | 13 | 295 | 13887 | 4.7 | 6.5 | 16 | 9.1 | 51 | 7.4 | 1.7 | 8.3 |
| Finnkolo | Mali | 9404 | 42 | 625 | 102 | 799 | 5.9 | 13 | 34 | 5879 | 2.8 | 3.2 | 3.3 | 1.8 | 7.1 | 2.8 | 0.8 | 0.7 |
| Fisenge | Zambia | 17156 | 42 | 1622 | 178 | 2590 | 4.2 | 28 | 89 | 5074 | 7.4 | 19 | 10 | 0.7 | 2.7 | 1.2 | 0.4 | 0.7 |
| Fria | Guinea | 47489 | 42 | 1364 | 375 | 5976 | 60 | 116 | 155 | 34722 | 11 | 14 | 15 | 9.5 | 13 | 2.6 | 3.6 | 30 |
| Ibi | Nigeria | 7602 | 42 | 1766 | 273 | 1606 | 6.4 | 12 | 26 | 3576 | 1.8 | 3.2 | 2.9 | 1.7 | 10 | 7.9 | 0.5 | 3.6 |
| Ihassunge | Mozambique | 16149 | 42 | 11287 | 3014 | 1543 | 8.8 | 22 | 93 | 7092 | 3.8 | 2.8 | 6.7 | 2.4 | 90 | 4.3 | 1.3 | 9.4 |
| Imorun | Nigeria | 19641 | 42 | 868 | 295 | 1573 | 3.1 | 4.6 | 98 | 8558 | 2.5 | 3.1 | 12 | 4.9 | 9.2 | 3.2 | 1.7 | 8.8 |
| Itende | Tanzania | 9702 | 42 | 3671 | 335 | 409 | 1.2 | 9.2 | 32 | 2459 | 1.6 | 1.9 | 4.2 | 2.1 | 12 | 2.1 | 0.1 | 3.9 |
| Katsina ala | Nigeria | 7098 | 36 | 1904 | 332 | 1107 | 1.1 | 7.2 | 70 | 3456 | 0.8 | 2.1 | 4.2 | 1.6 | 10 | 3.8 | 0.2 | 0.3 |
| Kiberashi | Tanzania | 18738 | 42 | 6830 | 1983 | 843 | 4.2 | 2.8 | 34 | 3550 | 1.1 | 2.0 | 10 | 4.0 | 76 | 1.1 | 0.7 | 9.8 |
| Kidatu | Tanzania | 4503 | 41 | 1510 | 234 | 286 | 1.9 | 5.9 | 12 | 1329 | 0.5 | 1.7 | 1.3 | 0.9 | 10 | 1.0 | 0.1 | 0.3 |
| Kisongo | Tanzania | 48373 | 634 | 12990 | 17055 | 8553 | 41 | 1.1 | 956 | 42128 | 1.8 | 14 | 71 | 11 | 674 | 17 | 6.3 | 37 |
| Koloko | Mali | 21841 | 42 | 2153 | 302 | 4409 | 26 | 44 | 84 | 9307 | 12 | 9.3 | 12 | 5.0 | 17 | 10 | 1.7 | 11 |
| Kontela | Mali | 26010 | 42 | 7479 | 1020 | 3305 | 25 | 44 | 130 | 16086 | 12 | 11 | 12 | 4.5 | 37 | 12 | 3.2 | 14 |
| Kubeasi | Ghana | 7588 | 42 | 1039 | 211 | 1268 | 2.7 | 11 | 21 | 1654 | 2.3 | 2.1 | 9.3 | 1.9 | 5.7 | 0.7 | 0.2 | 1.4 |
| Kutaber | Ethiopia | 36364 | 42 | 1989 | 9671 | 5316 | 31 | 30 | 604 | 31946 | 15 | 15 | 57 | 11 | 132 | 10 | 1.7 | 8.2 |
| Lambussie | Ghana | 10541 | 42 | 1366 | 349 | 1624 | 3.7 | 23 | 86 | 5317 | 2.0 | 5.3 | 5.7 | 1.6 | 8.6 | 1.7 | 0.5 | 0.8 |
| Macassangila | Mozambique | 12987 | 41 | 1508 | 331 | 1579 | 1.6 | 3.6 | 152 | 7183 | 0.5 | 2.6 | 11 | 2.6 | 15 | 3.5 | 1.5 | 2.4 |
| Marafa | Kenya | 19079 | 42 | 6270 | 2657 | 1686 | 5.3 | 20 | 239 | 9958 | 8.4 | 5.9 | 12 | 3.6 | 64 | 3.5 | 0.6 | 3.7 |
| Mbalambala | Kenya | 32027 | 42 | 14509 | 10858 | 2317 | 1.1 | 18 | 626 | 10691 | 2.7 | 6.9 | 17 | 6.3 | 185 | 10 | 0.5 | 2.7 |
| Mbinga | Tanzania | 38802 | 80 | 441 | 123 | 3225 | 45 | 18 | 793 | 36799 | 23 | 27 | 36 | 8.8 | 7.8 | 12 | 1.7 | 33 |
| Mega | Ethiopia | 102 | 27 | 787 | 6460 | 4 | 0.8 | 0.7 | 1.8 | 25 | 0.6 | 0.3 | 0.3 | 0.2 | 32 | 0.2 | 0.1 | 0.3 |
| Merar | Ethiopia | 8677 | 305 | 2884 | 64727 | 670 | 6.8 | 28 | 124 | 6475 | 37 | 3.2 | 8.2 | 2.1 | 80 | 9.3 | 0.1 | 0.5 |
| Morijo | Kenya | 17008 | 42 | 5314 | 1743 | 1418 | 2.2 | 10 | 66 | 12355 | 4.4 | 4.6 | 22 | 4.4 | 33 | 13 | 2.7 | 7.6 |
| Nkhata Bay | Malawi | 31888 | 42 | 2270 | 188 | 2096 | 14 | 11 | 147 | 15740 | 6.8 | 7.7 | 21 | 7.8 | 4.0 | 6.0 | 2.7 | 18 |
| Pandambili | Tanzania | 23221 | 42 | 7412 | 748 | 908 | 10 | 42 | 53 | 6920 | 10 | 4.3 | 11 | 5.1 | 40 | 2.1 | 1.1 | 3.4 |
| Thuchila | Malawi | 18037 | 42 | 6821 | 1060 | 5459 | 9.1 | 17 | 211 | 12370 | 1.0 | 4.4 | 17 | 2.2 | 65 | 5.6 | 3.1 | 12 |

Table A3.5: 25th percentiles of the total element concentration values (mg kg$^{-1}$) of 1074 soil samples across 34 sentinel sites sampled for this study.

| Site | Country | Al | P | K | Ca | Ti | V | Cr | Mn | Fe | Ni | Cu | Zn | Ga | Sr | Y | Ta | Pb |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Ajumako | Ghana | 32251 | 42 | 2146 | 754 | 3253 | 22 | 70 | 354 | 27554 | 12 | 17 | 25 | 7.7 | 15 | 3.6 | 2.5 | 23 |
| Bondigui | Burkina Faso | 29671 | 42 | 2186 | 694 | 2976 | 56 | 56 | 192 | 31388 | 7.8 | 18 | 13 | 2.4 | 15 | 1.0 | 3.5 | 32 |
| Boumeoul | Guinea | 20444 | 42 | 1633 | 256 | 2940 | 19 | 51 | 55 | 14376 | 8.3 | 10 | 7.2 | 4.4 | 9.0 | 8.3 | 1.9 | 12 |
| Bukwaya | Tanzania | 16361 | 42 | 6934 | 856 | 622 | 8.5 | 7.6 | 88 | 4295 | 2.3 | 3.4 | 7.8 | 2.9 | 31 | 3.8 | 1.0 | 8.6 |
| Chica - b | Mozambique | 22293 | 42 | 11989 | 2598 | 613 | 1.4 | 5.3 | 146 | 3338 | 1.4 | 2.1 | 6.7 | 5.1 | 74 | 6.0 | 0.3 | 4.5 |
| Chiculecule | Mozambique | 7083 | 42 | 419 | 169 | 1323 | 3.6 | 52 | 113 | 2860 | 4.2 | 2.6 | 3.8 | 1.4 | 1.5 | 1.4 | 0.3 | 3.6 |
| Chinyanghuku | Tanzania | 37621 | 42 | 15767 | 15415 | 1895 | 29 | 30 | 304 | 19548 | 6.7 | 9.0 | 24 | 13 | 289 | 5.4 | 2.4 | 7.5 |
| Dambidolo | Ethiopia | 45458 | 141 | 5865 | 5319 | 3553 | 18 | 45 | 753 | 37160 | 10 | 20 | 33 | 11 | 80 | 15 | 2.6 | 30 |
| Finnkolo | Mali | 14818 | 42 | 899 | 187 | 1378 | 10 | 20 | 59 | 8300 | 5.2 | 4.6 | 5.3 | 3.4 | 9.0 | 6.3 | 1.3 | 5.9 |
| Fisenge | Zambia | 26026 | 42 | 2155 | 208 | 3201 | 8.9 | 39 | 114 | 11082 | 18 | 22 | 13 | 3.0 | 3.8 | 3.1 | 2.3 | 2.5 |
| Fria | Guinea | 60130 | 42 | 1961 | 675 | 7594 | 93 | 182 | 229 | 55794 | 18 | 19 | 19 | 17 | 22 | 7.3 | 4.5 | 78 |
| Ibi | Nigeria | 9009 | 42 | 2572 | 435 | 1861 | 11 | 21 | 70 | 4696 | 2.8 | 3.7 | 3.7 | 2.2 | 14 | 9.1 | 0.7 | 5.5 |
| Ihassunge | Mozambique | 18132 | 42 | 17058 | 6812 | 2500 | 12 | 34 | 127 | 8800 | 4.8 | 4.2 | 11 | 4.3 | 102 | 6.8 | 1.7 | 14 |
| Imorun | Nigeria | 26700 | 42 | 1099 | 528 | 2427 | 4.5 | 23 | 126 | 14942 | 4.8 | 7.1 | 15 | 7.7 | 12 | 5.2 | 2.8 | 13 |
| Itende | Tanzania | 17874 | 42 | 16982 | 889 | 735 | 2.4 | 14 | 88 | 2881 | 2.6 | 2.6 | 5.4 | 3.7 | 35 | 4.3 | 0.2 | 8.4 |
| Katsina ala | Nigeria | 12690 | 42 | 5159 | 617 | 2197 | 1.1 | 19 | 145 | 4600 | 2.0 | 4.6 | 6.6 | 2.6 | 17 | 5.7 | 0.4 | 1.1 |
| Kiberashi | Tanzania | 30739 | 42 | 11117 | 3659 | 1320 | 11 | 14 | 138 | 7025 | 6.6 | 5.1 | 14 | 7.2 | 102 | 3.7 | 1.4 | 17 |
| Kidatu | Tanzania | 7626 | 42 | 5377 | 566 | 484 | 2.3 | 8.5 | 25 | 2080 | 1.0 | 2.3 | 2.3 | 1.2 | 38 | 2.2 | 0.2 | 0.7 |
| Kisongo | Tanzania | 52441 | 825 | 14881 | 25980 | 9542 | 73 | 15 | 1167 | 48325 | 2.9 | 17 | 77 | 12 | 789 | 19 | 7.8 | 54 |
| Koloko | Mali | 35968 | 42 | 3479 | 791 | 5356 | 42 | 66 | 142 | 14857 | 20 | 14 | 16 | 8.2 | 24 | 19 | 2.2 | 18 |
| Kontela | Mali | 30674 | 42 | 9915 | 1481 | 3632 | 30 | 47 | 220 | 19577 | 16 | 13 | 15 | 5.7 | 46 | 14 | 4.1 | 19 |
| Kubeasi | Ghana | 27059 | 42 | 2241 | 698 | 2543 | 12 | 42 | 119 | 12924 | 6.8 | 9.3 | 17 | 5.0 | 14 | 2.5 | 2.2 | 11 |
| Kutaber | Ethiopia | 40879 | 137 | 5094 | 16355 | 8618 | 71 | 44 | 750 | 53821 | 32 | 24 | 66 | 13 | 174 | 16 | 3.7 | 21 |
| Lambussie | Ghana | 16781 | 42 | 2599 | 607 | 2406 | 14 | 54 | 194 | 14191 | 5.6 | 9.2 | 9.3 | 2.5 | 16 | 6.2 | 2.0 | 5.3 |
| Macassangila | Mozambique | 31874 | 42 | 2713 | 614 | 3651 | 14 | 30 | 293 | 16980 | 5.3 | 8.9 | 23 | 7.7 | 29 | 5.0 | 2.4 | 8.5 |
| Marafa | Kenya | 23007 | 42 | 8973 | 4804 | 1999 | 10 | 22 | 368 | 12155 | 11 | 8.7 | 15 | 4.5 | 116 | 5.0 | 1.2 | 5.9 |
| Mbalambala | Kenya | 40541 | 42 | 21780 | 24005 | 3451 | 1.1 | 35 | 770 | 26785 | 10 | 12 | 36 | 12 | 308 | 19 | 1.0 | 10 |
| Mbinga | Tanzania | 43437 | 247 | 560 | 220 | 7646 | 75 | 35 | 1141 | 50499 | 31 | 31 | 42 | 11 | 12 | 22 | 2.9 | 69 |
| Mega | Ethiopia | 146 | 39 | 1027 | 7627 | 6.5 | 0.9 | 0.9 | 3.7 | 47 | 1.0 | 0.5 | 0.6 | 0.2 | 41 | 0.3 | 0.1 | 0.4 |
| Merar | Ethiopia | 27798 | 458 | 13012 | 83054 | 2728 | 28 | 92 | 439 | 21438 | 60 | 18 | 40 | 7.3 | 105 | 16 | 0.1 | 18 |
| Morijo | Kenya | 26562 | 42 | 8750 | 2559 | 2132 | 5.7 | 16 | 682 | 21080 | 11 | 7.5 | 54 | 8.5 | 47 | 29 | 3.9 | 11 |
| Nkhata Bay | Malawi | 46226 | 42 | 10076 | 633 | 2754 | 26 | 24 | 301 | 23092 | 11 | 13 | 38 | 11 | 13 | 10 | 3.7 | 30 |
| Pandambili | Tanzania | 28249 | 42 | 9924 | 1157 | 1490 | 13 | 56 | 107 | 11579 | 18 | 11 | 17 | 7.7 | 54 | 3.4 | 1.5 | 5.7 |
| Thuchila | Malawi | 25563 | 42 | 13466 | 2249 | 6741 | 16 | 41 | 402 | 17175 | 4.9 | 11 | 24 | 4.9 | 140 | 10 | 4.3 | 19 |

Table A3.6: 75th percentile of the total element concentration values (mg $kg^{-1}$) of 1074 soil samples across 34 sentinel sites sampled for this study.

| Site | Country | Al | P | K | Ca | Ti | V | Cr | Mn | Fe | Ni | Cu | Zn | Ga | Sr | Y | Ta | Pb |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Ajumako | Ghana | 46637 | 42 | 4185 | 2344 | 5377 | 54 | 215 | 790 | 50362 | 24 | 33 | 58 | 12 | 44 | 9.4 | 4.9 | 70 |
| Bondigui | Burkina Faso | 39864 | 42 | 3699 | 4267 | 7720 | 228 | 162 | 1216 | 113587 | 32 | 57 | 38 | 6.2 | 38 | 12 | 8.1 | 291 |
| Boumeoul | Guinea | 41828 | 42 | 4799 | 783 | 5942 | 63 | 90 | 244 | 45439 | 18 | 22 | 21 | 10 | 17 | 16 | 4.6 | 52 |
| Bukwaya | Tanzania | 32885 | 42 | 17600 | 4503 | 2101 | 21 | 30 | 338 | 21629 | 16 | 22 | 26 | 5.8 | 84 | 7.9 | 2.6 | 20 |
| Chica - b | Mozambique | 40132 | 42 | 33425 | 9585 | 1671 | 8.1 | 13 | 315 | 9891 | 4.0 | 5.6 | 15 | 12 | 182 | 29 | 1.1 | 26 |
| Chiculecule | Mozambique | 9658 | 42 | 830 | 391 | 1727 | 5.1 | 69 | 189 | 4068 | 6.7 | 3.3 | 4.7 | 1.9 | 3.1 | 2.6 | 0.4 | 5.5 |
| Chinyanghuku | Tanzania | 41332 | 55 | 22444 | 22904 | 2241 | 37 | 53 | 416 | 26154 | 15 | 14 | 34 | 14 | 412 | 8.5 | 3.3 | 12 |
| Dambidolo | Ethiopia | 57284 | 559 | 21708 | 18969 | 11803 | 54 | 110 | 1614 | 71084 | 47 | 31 | 62 | 15 | 268 | 49 | 4.2 | 144 |
| Finnkolo | Mali | 35834 | 42 | 2256 | 442 | 3302 | 24 | 57 | 126 | 21047 | 16 | 10 | 11 | 8.9 | 18 | 16 | 2.9 | 17 |
| Fisenge | Zambia | 35489 | 42 | 3296 | 408 | 4283 | 15 | 62 | 178 | 15810 | 37 | 35 | 26 | 5.4 | 5.0 | 6.0 | 10 | 6.9 |
| Fria | Guinea | 74560 | 42 | 4168 | 1354 | 9193 | 129 | 302 | 385 | 83044 | 29 | 30 | 28 | 21 | 27 | 12 | 5.5 | 139 |
| Ibi | Nigeria | 14947 | 42 | 8231 | 758 | 2629 | 16 | 58 | 152 | 8189 | 4.2 | 5.2 | 5.7 | 3.7 | 29 | 13 | 1.3 | 11 |
| Ihassunge | Mozambique | 33606 | 42 | 23314 | 9729 | 4115 | 24 | 84 | 231 | 27928 | 13 | 15 | 39 | 8.3 | 144 | 23 | 3.7 | 32 |
| Imorun | Nigeria | 39211 | 42 | 5216 | 1173 | 3246 | 16 | 59 | 249 | 28476 | 10 | 12 | 21 | 11 | 22 | 12 | 4.4 | 22 |
| Itende | Tanzania | 27405 | 42 | 26974 | 1707 | 1080 | 5.7 | 32 | 143 | 5137 | 7.8 | 4.3 | 7.5 | 5.8 | 79 | 7.2 | 0.8 | 19 |
| Katsina ala | Nigeria | 22930 | 46 | 18652 | 1058 | 5230 | 1.1 | 39 | 541 | 18179 | 3.5 | 9.0 | 14 | 5.6 | 55 | 11 | 2.4 | 18 |
| Kiberashi | Tanzania | 55042 | 42 | 19364 | 7341 | 2965 | 30 | 49 | 263 | 26749 | 25 | 12 | 27 | 13 | 163 | 8.1 | 3.5 | 24 |
| Kidatu | Tanzania | 24351 | 42 | 23253 | 1053 | 1536 | 7.5 | 21 | 201 | 6344 | 3.0 | 5.4 | 5.3 | 4.5 | 98 | 5.7 | 1.1 | 11 |
| Kisongo | Tanzania | 58856 | 1384 | 18568 | 38360 | 12033 | 124 | 58 | 1398 | 56393 | 33 | 26 | 88 | 15 | 881 | 22 | 8.9 | 73 |
| Koloko | Mali | 57526 | 42 | 5303 | 1756 | 6380 | 71 | 123 | 243 | 28924 | 46 | 25 | 28 | 18 | 37 | 31 | 4.0 | 35 |
| Kontela | Mali | 43279 | 42 | 12482 | 2712 | 4174 | 42 | 70 | 305 | 27039 | 23 | 19 | 20 | 8.3 | 59 | 17 | 5.5 | 28 |
| Kubeasi | Ghana | 39958 | 42 | 4967 | 1449 | 3808 | 22 | 97 | 282 | 27795 | 13 | 17 | 25 | 9.5 | 23 | 4.9 | 3.7 | 33 |
| Kutaber | Ethiopia | 47599 | 500 | 14715 | 26169 | 10841 | 104 | 101 | 1078 | 62212 | 47 | 47 | 95 | 16 | 379 | 27 | 6.6 | 38 |
| Lambussie | Ghana | 37801 | 42 | 4272 | 4441 | 6047 | 162 | 136 | 810 | 86583 | 22 | 45 | 38 | 6.7 | 61 | 9.0 | 5.3 | 174 |
| Macassangila | Mozambique | 54714 | 247 | 8221 | 1375 | 11320 | 38 | 76 | 791 | 47528 | 34 | 30 | 51 | 13 | 62 | 14 | 4.3 | 47 |
| Marafa | Kenya | 29820 | 42 | 18245 | 15437 | 2628 | 14 | 30 | 530 | 15337 | 16 | 12 | 20 | 6.0 | 172 | 8.4 | 2.5 | 10 |
| Mbalambala | Kenya | 50165 | 359 | 26918 | 39407 | 4339 | 3.8 | 65 | 1000 | 34668 | 23 | 19 | 68 | 16 | 435 | 30 | 3.2 | 20 |
| Mbinga | Tanzania | 48979 | 404 | 2754 | 736 | 11732 | 119 | 68 | 1529 | 62161 | 48 | 48 | 57 | 12 | 19 | 32 | 3.7 | 104 |
| Mega | Ethiopia | 227 | 60 | 1956 | 9609 | 25 | 1.1 | 1.3 | 8.7 | 173 | 1.7 | 0.8 | 1.1 | 0.2 | 55 | 0.7 | 0.1 | 0.7 |
| Merar | Ethiopia | 37002 | 1076 | 16731 | 153354 | 3453 | 35 | 126 | 656 | 31975 | 71 | 24 | 54 | 10 | 126 | 19 | 0.4 | 35 |
| Morijo | Kenya | 42018 | 42 | 12442 | 10852 | 4059 | 11 | 30 | 1119 | 35300 | 14 | 14 | 83 | 13 | 92 | 44 | 7.6 | 38 |
| Nkhata Bay | Malawi | 54961 | 42 | 21305 | 5631 | 3917 | 41 | 45 | 555 | 39209 | 19 | 18 | 65 | 16 | 44 | 24 | 4.7 | 52 |
| Pandambili | Tanzania | 36031 | 42 | 15362 | 6735 | 2069 | 20 | 88 | 250 | 17171 | 29 | 15 | 23 | 10 | 161 | 5.1 | 2.3 | 10 |
| Thuchila | Malawi | 42100 | 57 | 30420 | 5740 | 10634 | 46 | 57 | 653 | 38435 | 16 | 28 | 43 | 9.5 | 426 | 16 | 6.3 | 44 |

Table A3.7: 97.5th percentiles of the total element concentration values (mg kg$^{-1}$) of 1074 soil samples across 34 sentinel sites sampled for this study.

| Site | Country | Al | P | K | Ca | Ti | V | Cr | Mn | Fe | Ni | Cu | Zn | Ga | Sr | Y | Ta | Pb |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Ajumako | Ghana | 62822 | 43 | 12510 | 16782 | 6298 | 93 | 311 | 1155 | 86693 | 78 | 67 | 111 | 15 | 156 | 18 | 6.5 | 210 |
| Bondigui | Burkina Faso | 48586 | 57 | 6514 | 28659 | 10479 | 384 | 363 | 3351 | 180612 | 50 | 91 | 68 | 13 | 85 | 25 | 11 | 617 |
| Boumeoul | Guinea | 62062 | 91 | 12158 | 5152 | 6426 | 111 | 197 | 1738 | 71641 | 70 | 33 | 47 | 16 | 37 | 23 | 6.9 | 154 |
| Bukwaya | Tanzania | 49111 | 42 | 36328 | 15290 | 5898 | 111 | 104 | 1934 | 79553 | 58 | 68 | 40 | 11 | 201 | 14 | 6.2 | 180 |
| Chica -b | Mozambique | 46971 | 49 | 49068 | 23892 | 2229 | 25 | 106 | 465 | 23061 | 37 | 18 | 40 | 15 | 410 | 107 | 2.0 | 38 |
| Chiculecule | Mozambique | 12095 | 42 | 1507 | 1600 | 1994 | 6.2 | 84 | 317 | 5482 | 10 | 4.4 | 6.0 | 2.3 | 4.8 | 4.5 | 0.7 | 9.3 |
| Chinyanghuku | Tanzania | 46258 | 226 | 24767 | 33348 | 2744 | 58 | 535 | 733 | 37362 | 99 | 24 | 49 | 15 | 512 | 13 | 3.8 | 16 |
| Dambidolo | Ethiopia | 65575 | 1363 | 29443 | 30539 | 19139 | 101 | 157 | 2105 | 96911 | 72 | 57 | 86 | 19 | 362 | 94 | 6.7 | 233 |
| Finnkolo | Mali | 57298 | 42 | 4209 | 741 | 7537 | 41 | 86 | 221 | 37088 | 36 | 17 | 18 | 20 | 30 | 30 | 4.6 | 40 |
| Fisenge | Zambia | 42599 | 42 | 5256 | 652 | 6582 | 66 | 117 | 552 | 41598 | 51 | 54 | 63 | 8.8 | 7.1 | 10 | 15 | 14 |
| Fria | Guinea | 85209 | 64 | 11104 | 2297 | 11186 | 153 | 466 | 527 | 95746 | 44 | 44 | 32 | 25 | 42 | 18 | 6.9 | 182 |
| Ibi | Nigeria | 21735 | 58 | 23890 | 2112 | 3431 | 74 | 142 | 197 | 73299 | 10 | 21 | 16 | 5.5 | 68 | 18 | 9.7 | 119 |
| Ihassunge | Mozambique | 52094 | 42 | 25897 | 15298 | 17286 | 47 | 206 | 612 | 46513 | 41 | 29 | 60 | 13 | 166 | 44 | 5.3 | 64 |
| Imorun | Nigeria | 46900 | 42 | 27410 | 2324 | 3941 | 31 | 100 | 763 | 39978 | 17 | 21 | 34 | 14 | 52 | 21 | 5.0 | 36 |
| Itende | Tanzania | 41536 | 42 | 51989 | 4509 | 1644 | 15 | 81 | 223 | 11010 | 16 | 9.2 | 11 | 7.9 | 148 | 37 | 1.3 | 34 |
| Katsina ala | Nigeria | 74252 | 170 | 33326 | 2794 | 14413 | 9.0 | 53 | 1013 | 46007 | 10 | 19 | 28 | 20 | 247 | 15 | 5.1 | 59 |
| Kiberashi | Tanzania | 65933 | 68 | 35346 | 10180 | 4071 | 42 | 63 | 428 | 32566 | 39 | 17 | 35 | 18 | 244 | 10 | 4.4 | 37 |
| Kidatu | Tanzania | 31494 | 45 | 54018 | 1829 | 2452 | 14 | 30 | 268 | 13342 | 10 | 8.1 | 9.0 | 6.5 | 390 | 12 | 2.0 | 18 |
| Kisongo | Tanzania | 66185 | 1861 | 20678 | 47122 | 13812 | 144 | 271 | 1542 | 65357 | 98 | 45 | 94 | 16 | 1056 | 24 | 10 | 98 |
| Koloko | Mali | 75125 | 43 | 6210 | 2238 | 7742 | 167 | 267 | 321 | 74820 | 73 | 30 | 39 | 25 | 44 | 41 | 7.5 | 152 |
| Kontela | Mali | 51744 | 42 | 17968 | 5318 | 4741 | 60 | 98 | 811 | 47394 | 32 | 26 | 40 | 11 | 97 | 22 | 8.7 | 85 |
| Kubeasi | Ghana | 47870 | 42 | 6966 | 2932 | 4853 | 43 | 120 | 435 | 39385 | 19 | 21 | 64 | 11 | 34 | 7.1 | 4.4 | 56 |
| Kutaber | Ethiopia | 53393 | 769 | 26800 | 30240 | 14421 | 123 | 177 | 1477 | 66358 | 75 | 62 | 120 | 23 | 1134 | 41 | 7.4 | 55 |
| Lambussie | Ghana | 63485 | 1049 | 5582 | 22200 | 7522 | 250 | 331 | 6018 | 127683 | 224 | 74 | 63 | 12 | 214 | 13 | 8.0 | 364 |
| Macassangila | Mozambique | 65043 | 1261 | 34886 | 5498 | 25182 | 107 | 122 | 1443 | 93618 | 51 | 68 | 123 | 18 | 178 | 31 | 7.6 | 166 |
| Marafa | Kenya | 34163 | 42 | 22208 | 26559 | 3122 | 21 | 41 | 1026 | 20373 | 19 | 17 | 26 | 7.9 | 217 | 12 | 3.2 | 12 |
| Mbalambala | Kenya | 62660 | 641 | 33942 | 55628 | 5854 | 31 | 97 | 1539 | 48560 | 30 | 35 | 102 | 20 | 574 | 49 | 4.4 | 40 |
| Mbinga | Tanzania | 61972 | 597 | 6096 | 2664 | 13377 | 157 | 122 | 1797 | 76194 | 114 | 62 | 70 | 13 | 29 | 47 | 5.1 | 132 |
| Mega | Ethiopia | 403 | 107 | 3356 | 10894 | 39 | 1.7 | 4.3 | 72 | 313 | 31 | 2.7 | 2.8 | 0.3 | 62 | 1.1 | 0.1 | 1.0 |
| Merar | Ethiopia | 39948 | 2271 | 18227 | 396478 | 3733 | 53 | 137 | 747 | 34317 | 198 | 26 | 62 | 10 | 149 | 21 | 0.7 | 42 |
| Morijo | Kenya | 52733 | 476 | 19631 | 23109 | 5241 | 19 | 143 | 1913 | 49509 | 23 | 21 | 95 | 17 | 124 | 52 | 9.6 | 58 |
| Nkhata Bay | Malawi | 71264 | 155 | 26651 | 11926 | 5661 | 53 | 64 | 1374 | 45951 | 27 | 22 | 87 | 21 | 86 | 29 | 6.0 | 75 |
| Pandambili | Tanzania | 43015 | 46 | 26953 | 9080 | 2752 | 32 | 116 | 435 | 25350 | 47 | 22 | 40 | 13 | 246 | 8.2 | 3.0 | 20 |
| Thuchila | Malawi | 50777 | 470 | 73356 | 18356 | 17524 | 78 | 80 | 945 | 57330 | 21 | 39 | 61 | 13 | 1290 | 22 | 8.9 | 73 |

# Chapter 4

# The potential of combining mid infrared (MIR) and total X-ray fluorescence (TXRF) spectroscopy for the prediction of soil properties

# 4.0 Combined mid infrared and total X-ray fluorescence spectroscopy for predicting soil properties[3]

Erick. K. Towett[1,2], Keith D. Shepherd[1], Andrew Sila[1], Ermias Aynekulu[1], and Georg Cadisch[2,*]

[1]World Agroforestry Centre (ICRAF), P.O. Box 30677-00100 GPO, Nairobi, Kenya.
[2]Institute of Plant Production and Agroecology in the Tropics and Subtropics, Universität Hohenheim, Garbenstr. 13, 70599 Stuttgart, Germany.
**Corresponding author at*: Institute of Plant Production and Agroecology in the Tropics and Subtropics, Universität Hohenheim, Garbenstr. 13, 70599 Stuttgart, Germany
Tel: +49 711459 23538; Fax: +49 711 459 22304
Email: georg.cadisch@uni-hohenheim.de

## 4.1 Abstract

Mid-infrared diffuse reflectance spectroscopy (MIR) is increasingly being used to predict a number of soil properties but some properties, notably extractable nutrients, are often predicted poorly. This paper evaluated the potential of total X-ray fluorescence spectroscopy (TXRF) to predict MIR calibration residuals. A total of 700 soil samples from 44 stratified randomly-located 100-$km^2$ sentinel sites distributed across sub-Saharan Africa (SSA), were analysed for physico-chemical composition using conventional reference methods as well as by both MIR and TXRF. The reference methods were calibrated to MIR and TXRF data using Random Forests (RF) regression models. MIR spectra resulted in good prediction models using RF out-of-bag validation ($R^2 > 0.80$) for organic and total C and N, Mehlich-3 Ca and Al, and pH. Also predicted well ($R^2 > 0.60$) were Ca/Mg ratio, exchangeable Mg, phosphorus sorption index, and particle size distribution. Calibration models were less satisfactory ($R^2 < 0.60$) for Mehlich-3 extractable K, Mn, Fe, Cu, B, Zn, P, S, and Na, exchangeable acidity, electrical conductivity (Ecd), exchangeable sodium percentage, and air-dispersed particle sizes. Including total element concentration data from TXRF analysis in the RF models significantly reduced root mean square error of prediction by 63% for Ecd, 54% for Mehlich-3 S, and 53% for Mehlich-3 Na. The prediction improvement from including TXRF was due to detection of a few outliers that did not appear as MIR spectral outliers. MIR showed remarkable ability to capture total elemental composition effects on physico-chemical soil properties but TXRF may have potential for outlier detection in large studies.

---

[3] A version of this chapter has been submitted to the Soil Science Society of America Journal in Dec 2013.

## 4.2 Introduction

The need for new, rapid methods to quantify soil chemical properties is particularly acute in developing countries of Sub-Saharan Africa (SSA) where reliable data on soil properties is sparse and dated. Given the multi-component nature of soil systems, the variability of properties, as well as their spatial distribution provides a challenge for the choice of appropriate methods to characterize soil properties (Kibblewhite et al., 2008). The Africa Soil Information Service (AfSIS) project (www.africasoil.net) to which this present study contributes, employs an approach known as land health surveillance, defined as the ongoing, systematic collection, analysis, and interpretation of data essential to the planning, implementation, and evaluation of land management policy and practice, and application of these data to promote, protect, and restore land and ecosystem health (UNEP, 2012). Innovations in land health surveillance include the use of infrared spectroscopy for low cost, rapid soil characterization, which has enabled large area soil assessment (Shepherd and Walsh, 2007), and linkage of spectrally derived soil fertility indicators to conventional soil assessment (chemical) data through statistical models.

Spectroscopic techniques have shown promise as rapid and accurate methods for characterizing soil properties. Infrared Spectroscopy (IR) in the visible to near-infrared (0.35 to 2.5 μm) and mid-infrared (2.5 to 25 μm) wavelength regions have been investigated for non-consumptive analysis of soils and simultaneous estimations of several soil physical, chemical and biological properties such as total C, total N, cation exchangeable change capacity (CEC), exchangeable Ca, exchangeable Mg, clay content, sand content, K, P, S and soil pH (e.g. Awiti et al., 2007; Brown et al., 2006; Du and Zhou, 2009; Minasny and McBratney, 2008; Shepherd and Walsh, 2002; 2007; and Terhoeven-Urselmans et al., 2010).

IR has limitations in predicting weakly extractable nutrients (Janik et al., 1998) such as extractable P and K, which are main limiting nutrients in soils of SSA (Ludwig et al., 2002). However, results by Kleinebecker et al. (2013) demonstrated that near infrared reflectance spectroscopy (NIRS) does have the potential to reliably measure total concentrations, NaCl- and oxalate-extractable element fractions of Al, Ca, Fe, K, Mg, N, Na, P, S, Si and Zn across a wide range of aquatic sediments. Similarly, Soriano-Disla et al. (2013) obtained good results for Ca, Mg, Al, Fe, Ga, Si and Na in their development of partial least squares (PLS) regression models using diffuse reflectance Fourier transform mid-infrared (MIR) spectroscopy for the prediction of the concentration of elements in soil determined by X-ray fluorescence (XRF). Another challenge with IR is in its sensitivity to interferences, requiring extensive calibration for different environments. Prediction of particle size distribution with IR is variable and in addition, calibrations have to be adjusted for different soil types (Kamau-Rewe et al., 2011; Shepherd, 2010). A current challenge is to build calibration libraries that are large enough to sufficiently capture the large variation in global soils (Brown et al., 2006).

The concept of using light to analyse soils can be extended to the X-ray region using new instrumentation developments that may complement IR in prediction of some soil properties. Total X-ray fluorescence spectroscopy (TXRF) is a relatively new commercially available technique that can provide for rapid and simultaneous determination of the total concentrations of many elements in the periodic table of elements. Advantages of the technique compared to conventional methods include minimal sample preparation and low matrix interference (Stosnach, 2005). The total element concentration spectra can be used to capture key mineralogical differences in soils (Acosta et al., 2011; Towett et al., Submitted) and there are possibilities to correlate extractable nutrient analysis with total element analysis. This is due to the fact that all soils contain some of all the naturally occurring chemical elements and the variations in the concentration of elements is derived from differences in the composition of the parent material and from fluxes of mater and energy into or from soil over geologic time (Helmke, 2000; Kabata-Pendias and Mukherjee, 2007) or recent management. Thus TXRF provides chemical fingerprinting and “functional mineral groupings” that could relate to potential nutrient supply capacity (Towett et al., Submitted). The objectives of this study were thus to evaluate whether TXRF can complement MIR for predicting soil test values, especially for tests that are poorly predicted by MIR (e.g. extractable P and K; and some micronutrients).

## 4.3 Material and Methods

### 4.3.1 Study area, soil sampling and processing

Georeferenced soil samples associated with the Africa Soil Information Service (AfSIS) (www.africasoils.net) were taken from a set of sentinel sites randomized over Sub-Saharan Africa (SSA) (AfSIS, 2013). A total of 700 soil samples from 44 random 100-km$^2$ sentinel sites, stratified according to Köppen-Geiger climatic zones (Kottek et al., 2006), distributed across SSA: Burkina Faso (1 site, 16 samples), Cameroon (1 site, 16 samples), Ethiopia (4 sites, 64 samples), Ghana (3 sites, 48 samples), Guinea (2 sites, 32 samples), Kenya (3 sites, 48 samples), Malawi (2 sites, 32 samples), Mali (3 sites, 48 samples), Mozambique (6 sites, 96 samples), Nigeria (3 sites, 47 samples), South Africa (3 sites, 47 samples), Tanzania (8 sites, 127 samples), Uganda (2 sites, 32 samples) and Zambia (3 sites, 47 samples) were used in exploring spectral patterns in this study. A summary of the average co-ordinates of the sites from which the samples came from is given in Annex Table A1. Field sampling was made based on the Land Degradation Surveillance Framework (LDSF) protocol (Vågen et al., 2013). The protocol is built around the use of "Sentinel sites" or "Blocks" of 10 x 10 km in size. The basic sampling unit used in the LDSF is called a Cluster. Each Sentinel site was stratified into sixteen 1-km$^2$ clusters containing 10 randomized plots of 1000 m$^2$ each. Soil samples were collected from four 100 m$^2$ "sub-plots" located within the ten 1000 m$^2$ plots at 0-20 cm depth and combined into one composite sample per plot, giving a total of 160 soil samples collected per "Sentinel Site". We selected 10% of soil samples from each sentinel site for reference and spectral analyses based on one plot taken at random from each cluster, giving 16 samples per sentinel site, and a total of 700 samples. Soil samples were initially air-dried and passed through a 2-mm sieve before sub sampling to 10 g by coning and quartering. The 10 g subsamples were oven-dried at 40 ºC and then ground to pass a 75 µm sieve using a Retsch RM 200 mill (Retsch, Duesseldorf, Germany) for MIR analysis and further ground using a McCrone micronising mill (McCrone, Westmont, U.S.A) to <50 µm for TXRF analysis.

### 4.3.2 Spectral analyses method

Samples were analysed at the World Agroforestry Centre (ICRAF)'s Soil-Plant Spectral Diagnostics Laboratory in Nairobi, Kenya, using a high-throughput Bruker Tensor 27 Fourier-Transform MIR spectrometer (Bruker Optik GmbH, Germany) attached to a High-Throughput Screening (HTS-XT) accessory. Fine ground soil samples (approx. 20 to 30 mg) were loaded onto micro-titre plates in four replicates (Shepherd and Walsh, 2007). Infrared absorbance spectra were recorded after co-adding 32 scans at 4 cm$^{-1}$ intervals in the range of 400 to 4000 cm$^{-1}$. The average of the spectra for the four replicates was taken.

The TXRF methodology was used to analyse total elemental concentrations in each soil sample using a S2 PICOFOX$^{TM}$ TXRF spectrometer (Bruker AXS Microanalysis GmbH, Germany). Based on a recently developed and tested method (Towett et al., 2013), 50 mg of the finely ground (20-50 µm) sample was mixed with 2.5 ml of Triton X100 (Fischer Scientific, UK) solution (0.1 vol.%) to form a soil suspension and spiked with 40 µl of 1000 mg l$^{-1}$ Selenium

(Fluka Analytical, Germany) as the internal standard. The suspension was placed into an ultrasonic water bath at room temperature and sonicated in a continuous mode for 15 min; mixed well using a digital shaker and 10 µl of the turbid soil solution immediately dispensed on to a clean siliconized quartz glass sample carrier and dried for 10-15 min at 52 ºC on a hot plate (Staurt® SD300) in a clean laminar flow hood.

### 4.3.3 Reference soil analysis

Total and organic C and N analyses were determined based on the flash dynamic combustion method (Skjemstad and Baldock, 2008) using the Flash EA 1112 Elemental Analyzer (Thermo Scientific, Milan, Italy) at the ICRAF Soil-Plant Spectral Diagnostic Laboratory. Additionally, conventional Mehlich-3 soil tests (Mehlich, 1984) were analyzed for extractable P, K, Ca, Mg, Na, Mn, Fe, Cu, Zn, B, Mo, S, and Al using inductively-coupled plasma spectroscopy (ICP). Conventional pH measurements were done using a standard method utilizing a soil: water ratio of 1:2 weight to volume basis. Exchangeable acidity was determined on 1N KCl extracts of soils using a 1:10 soil to extractant volume ratio, and titration with a standardized base (sodium hydroxide; Anderson and Ingram, 1993). Phosphorus sorption index (PSI) was determined, based closely on the method by Bache and Williams (1971), by weighing 1.5 g soil into a 50 mL extraction bottle, adding 30 mL of diluted phosphate solution (75 mg P $L^{-1}$ in 0.03M KCl), then two drops of toluene was added to prevent microbial growth. The solution was shaken briefly with a vortex mixer to disperse the soil then the bottles were put on a shaker in a room maintained at 25° C for 20 hours (overnight) and later centrifuged for 10 minutes at rcf=1500g to make a clear supernatant. The supernatant was filtered though medium-fast paper (Whatman No.1) and if the filtered supernatant was cloudy then it was re-filtered with a slow paper (e.g. Whatman No.2) and P was measured in the filtered supernatant. All chemical analyses were done by Crop Nutrition Laboratory Services in Nairobi (ISO 17025 accredited).

Samples were analysed for their particle sizes using a detectable size range of 0.01-3000 µm utilizing a Horiba (Model: LA-950V2) (Horiba Ltd., Kyoto, Japan) Laser Diffraction Particle Size Analyser (LDPSA) at the ICRAF Laboratory. The LDPSA instrument allowed continuous flow of a soil sample suspended in (i) a dry air stream or (ii) a water stream, to which different sonification cycles were applied using an in-built ultrasonic probe. The protocol began with measurement of particle size distribution of dry soil suspended in the air stream to provide a measure of micro-aggregation without wetting and particle size distribution was then measured in water, followed by a second reading one minute later, and finally after full dispersion using Calgon and sonification (130 W, 20 kHz).

### 4.3.4 Chemometric analyses

Quantitative analyses based on MIR spectra required the development of calibrations that related the first derivative of the spectral information to the reference analysis data using the entire spectra. TXRF element concentration data was included as predictor along with the first derivative of MIR spectral data using the Random Forests (Breiman, 2001) algorithm. The RF multivariate algorithm is reported to be resistant to over-fitting and usually performs well in problems with a low sample-to-features ratio, such as spectrometric data (Wei et al., 2012; Ghasemi and Tavakoli, 2013). RF also handles complex data types well and obviates the need for transformation of predictors to approximate normal distributions, which were advantages for this study. The 'randomForest' library in R (Breiman, 2001) was used.

Prior to calibration, it was necessary to perform spectra pre-treatment to eliminate the physical effect of light scattering, which can be due to particles of different sizes and shapes (Minasny and McBratney, 2008). Thus, a first-order derivative of the spectral range 601.7-4001.6 $cm^{-1}$ was calculated using the Savitsky-Golay algorithm using the soil.spec package in R (R-version 2.15.3; R Development Core Team 2013) available under the Packages link at www.cran.r-project.org/ (Accessed 13 September 2013). The Savitsky-Golay smoothing degree depended on two parameters: the frame size and the order of the polynomial used for smoothing; the frame size of the Savitzky Golay filter was set to 21 data points and the polynomial order was set to 3. RF calibration models with an out-of-bag validation (RF-OOB) were then developed using the entire data set of 700 samples. Additionally, we compared the performance of the RF algorithm against the partial least squares (PLS) as a standard algorithm on the same data set but using a hold-out validation on the simultaneous determination of soil properties using MIRS. Selection of calibration (70%) and validation (30%) samples for the PLS was done following a procedure adapted from Kennard and Stone (1969) applied to the score values of the first eight principal components of the first derivative MIR spectra. For our analysis, comparing the independent hold-out validation and the "out-of-bag" validation and using each validation method to estimate model errors (root mean square errors) for all variables, the RF out-performed the PLS algorithm on simultaneous prediction of soil properties on the same sample set. Thus, RF models were also built to predict the reference properties from the TXRF total element composition using the raw total element concentration data as ‘spectra’. Finally, we used the raw TXRF spectra in conjunction with first derivative MIR spectra to predict the reference soil properties. We also tested an approach whereby we used the TXRF data to test whether it can improve MIR predictions of soil test values, especially for those variables for which MIRs tends to give poor predictions (e.g. extractable P, K, and some micronutrients). Thus, we used RF to calibrate the residuals of the predictions from the MIR spectral data for individual reference properties to the raw TXRF total element data, as mixing different data types in the predictor variables might affect the variable importance weights in the fitted models.

## 4.4 Results and discussion

### 4.4.1 Statistical description of data

A wide variation was found in individual soil physico-chemical properties in the soil sample sets using conventional chemistry methods and particle size analysis using laser diffraction (Table 4.1). Soil pH in $H_2O$ ranged from 4 to 10, with a mean pH 6. Exchangeable acidity (unbuffered KCl extraction) values in soils varied from 0.01 to 4 cmolc kg$^{-1}$. Similarly, exchangeable bases varied considerably from 1 to 180 cmolc kg$^{-1}$. Soils had total organic C and total N contents ranging from 0.1 to 10% and from 0.003 to 1%, respectively. Average clay contents ranged from 0% (air-dispersed particles measured using laser diffraction) to 100% (calgon dispersed particles after 4 min of ultrasonification), average sand contents ranged from 0.1% (calgon dispersed particles after 4 min of ultrasonification) to 100% (air-dispersed particles) while the average silt fraction varied between 0% (air-dispersed particles) and 60% (water dispersed particles after 4 min of ultrasonification).

There was a wide variation in MIR spectral data used in this study, ranging from 0.5 – 2.5 in maximum absorbance in each of the four spectral ranges (1) fingerprint (for example, O–Si– O stretching and bending), (2) double-bond (for example, C=O, C=C, C=N), (3) triple bond (for example, C≡C, C≡N) and (4) X–H stretching (for example, O–H stretching) (Figure 4.1).

There was also a wide variation in the total element concentrations measured using TXRF in the samples used in this study (Table 4.2). For example, the total concentration of Na varied from 7,260-43,820 (mean 20,080) mg kg$^{-1}$ across the sentinel sites sampled, while total P, K, Ca, Mg and Al values ranged between 25-2,240, 355-77,900, 150-386,490, 945-49,350 and

130-77,130 mg kg$^{-1}$ respectively. This high variation was attributable to differences in parent materials between sites and to local pedologic and hydrological factors within sites or recent management as discussed in Towett et al. (Submitted).

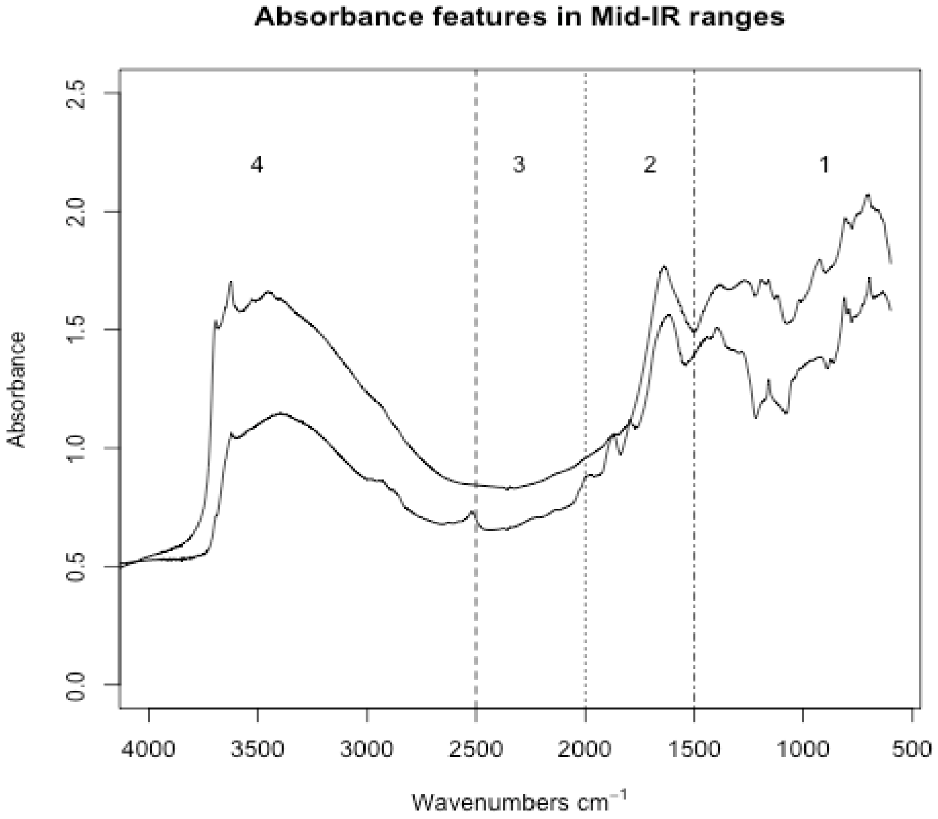


Figure 4.1: Diffuse reflectance spectra of two Africa soils sampled for this study showing absorbance features in the mid-IR wavelength ranges. Mid-IR spectroscopy regions: (1) fingerprint (for example, O–Si– O stretching and bending), (2) double-bond (for example, C=O, C=C, C=N), (3) triple bond (for example, C≡C, C≡N) and (4) X–H stretching (for example, O–H stretching).

Table 4.1: Statistical description of the reference soil data analysed using conventional laboratory (reference) analyses.

| Property[†], units | n | Mean | Std Dev | Min | Max | 2.5th per | 25th per | 50th per | 75th per | 97.5th per |
|---|---|---|---|---|---|---|---|---|---|---|
| pH, units | 700 | 6 | 1 | 4 | 10 | 5 | 6 | 6 | 7 | 9 |
| ExAc, cmolc $kg^{-1}$ | 700 | 0.4 | 1 | 0.01 | 4 | 0.0 | 0.1 | 0.2 | 1 | 2 |
| ECd, dS/m | 700 | 0.1 | 0.3 | 0.01 | 5 | 0.0 | 0.0 | 0.1 | 0.1 | 1 |
| ExCa, cmolc $kg^{-1}$ | 700 | 10 | 18 | 3 | 170 | 1 | 2 | 5 | 10 | 70 |
| ExK, cmolc $kg^{-1}$ | 700 | 1 | 1 | 0.01 | 10 | 0.1 | 0.2 | 0.3 | 1 | 3 |
| ExMg, cmolc $kg^{-1}$ | 700 | 3 | 4 | 0.14 | 40 | 0.3 | 1 | 2 | 4 | 12 |
| ExNa, cmolc $kg^{-1}$ | 700 | 1 | 7 | 0.00 | 140 | 0.0 | 0.1 | 0.1 | 0.2 | 3 |
| ExBas, cmolc $kg^{-1}$ | 700 | 15 | 20 | 1 | 180 | 1 | 4 | 7 | 17 | 80 |
| CaMg, ratio | 700 | 4 | 6 | 0.2 | 120 | 1 | 2 | 3 | 4 | 14 |
| ESP, % equivalent | 700 | 4 | 6 | 0.02 | 80 | 0.3 | 1 | 2 | 4 | 15 |
| ESR, cmolc $kg^{-1}$ | 700 | 0.1 | 0.2 | 0.001 | 4 | 0.0 | 0.0 | 0.0 | 0.0 | 0.2 |
| m3.P, mg $kg^{-1}$ | 700 | 20 | 30 | 0.3 | 340 | 1 | 5 | 8 | 20 | 120 |
| m3.Al, mg $kg^{-1}$ | 700 | 820 | 460 | 6 | 2630 | 170 | 480 | 740 | 1050 | 1940 |
| m3.B, mg $kg^{-1}$ | 700 | 1 | 1 | 0.001 | 10 | 0.0 | 0.1 | 0.2 | 1 | 3 |
| m3.Cu, mg $kg^{-1}$ | 700 | 2 | 2 | 0.001 | 20 | 0.0 | 1 | 1 | 3 | 7 |
| m3.Fe, mg $kg^{-1}$ | 700 | 130 | 90 | 8 | 780 | 30 | 70 | 101 | 150 | 400 |
| m3.Mn, mg $kg^{-1}$ | 700 | 120 | 110 | 1 | 670 | 5 | 42 | 100 | 175 | 380 |
| m3.Zn, mg $kg^{-1}$ | 700 | 2 | 2 | 0.001 | 40 | 0.3 | 1 | 1 | 2 | 6 |
| m3.S, mg $kg^{-1}$ | 699 | 25 | 180 | 2 | 3940 | 3 | 6 | 10 | 15 | 50 |
| PSI, units | 684 | 70 | 80 | -30 | 450 | -6 | 20 | 50 | 100 | 330 |
| psa_asand, % | 699 | 90 | 10 | 0 | 100 | 50 | 80 | 90 | 90 | 100 |
| psa_asilt, % | 699 | 8 | 7 | 0 | 40 | 1 | 4 | 7 | 10 | 30 |
| psa_aclay, % | 699 | 5 | 7 | 0 | 50 | 0.2 | 1 | 3 | 7 | 25 |
| psa_c4clay, % | 699 | 40 | 20 | 1 | 100 | 4 | 20 | 38 | 58 | 90 |
| psa_c4silt, % | 699 | 20 | 10 | 1 | 60 | 3 | 10 | 18 | 30 | 40 |
| psa_c4sand, % | 699 | 40 | 25 | 0.1 | 100 | 4 | 20 | 40 | 60 | 90 |
| psa_w4clay, % | 699 | 40 | 20 | 2 | 95 | 6 | 20 | 40 | 50 | 80 |
| psa_w4silt, % | 699 | 20 | 10 | 1 | 60 | 3 | 15 | 20 | 30 | 45 |
| psa_w4sand, % | 699 | 40 | 20 | 2 | 100 | 4 | 20 | 35 | 60 | 90 |
| Total N, % | 651 | 0.1 | 0.1 | 0.01 | 1 | 0.0 | 0.0 | 0.1 | 0.1 | 0.4 |
| Total C, % | 651 | 2 | 2 | 0.1 | 10 | 0.3 | 1 | 1 | 2 | 6 |
| Acidified N, % | 651 | 0.1 | 0.1 | 0.003 | 1 | 0.0 | 0.0 | 0.1 | 0.1 | 0.4 |
| Acidified C, % | 651 | 2 | 1 | 0.1 | 10 | 0.3 | 1 | 1 | 2 | 6 |

[†] pH = Soil reaction (1:2 volume water extract); ExAc = Exchangeable acidity (unbuffered KCl extraction); Ecd = Electrical conductivity (1:2 volume water extract); ExCa = Exchangeable Ca; ExK = Exchangeable K; ExMg = Exchangeable Mg; ExNa = Exchangeable Na; ExBas =Exchangeable bases (sum of Mehlich exch Ca, Mg, K, Na); CaMg = Ca/Mg ratio; ESP = Exchangeable sodium percentage (100.ExNa/(ExCa+ExMg+ExNa); ESR = Exchangeable sodium ratio; m3.P = Mehlich 3 extractable P; m3.Al = Mehlich 3 extractable Al; m3.B = Mehlich 3 extractable B; m3.cu = Mehlich 3 extractable Cu; m3.Fe = Mehlich 3 extractable Fe; m3.Mn = Mehlich 3 extractable Mn; m3.Zn = Mehlich 3 extractable Zn; m3.S = Mehlich 3 extractable S; PSI = Phosphorus sorption index; psa_asand = sand content (>50 microns) for air-dispersed particles by laser diffraction; psa_asilt = silt content (8 - 50 microns) for air-dispersed particles by laser diffraction; psa_aclay = clay content (< 8 microns) for air-dispersed particles by laser diffraction; psa_c4clay = clay content (<8 microns) for calgon dispersed particles after 4 min of ultrasonification; psa_c4silt = silt content (8 - 50 microns) for calgon dispersed particles after 4 min of ultrasonification; psa_c4sand = sand content (> 50 microns) for calgon dispersed particles after 4 min of ultrasonication; psa_w4clay = clay content (<8 microns) for water dispersed particles after 4 min of ultrasonification; psa_w4silt = silt content (8 - 50 microns) for water dispersed particles after 4 min of ultrasonification; psa_w4sand = sand content (> 50 microns) for water dispersed particles after 4 min of ultrasonication; n = number of samples; Std Dev = standard deviation; Min = Minimum; max = maximum; per = percentile.

Table 4.2: Statistical description of the total element concentration measured using TXRF. The number of samples that had complete data was 555.

| Property | Mean | Std Dev | Min | Max | 2.5th per | 25th per | 50th per | 75th per | 97.5th per |
|---|---|---|---|---|---|---|---|---|---|
| Na | 20080 | 4550 | 7260 | 43820 | 7850 | 17610 | 20800 | 23370 | 26780 |
| Mg | 5400 | 4050 | 945 | 49350 | 1240 | 4265 | 4265 | 4265 | 17830 |
| Al | 31960 | 16100 | 130 | 77130 | 150 | 20104 | 31045 | 43000 | 64900 |
| P | 140 | 280 | 25 | 2240 | 30 | 42 | 42 | 42 | 1169 |
| S | 75 | 440 | 15 | 8460 | 15 | 40 | 40 | 40 | 139 |
| Cl | 380 | 3650 | 20 | 74380 | 20 | 74 | 111 | 170 | 400 |
| K | 10920 | 9780 | 355 | 77900 | 390 | 2930 | 8440 | 16730 | 33480 |
| Ca | 9780 | 25120 | 150 | 386490 | 160 | 830 | 2240 | 8980 | 68610 |
| Sc | 7 | 15 | 2 | 170 | 3 | 3.9 | 3.9 | 4 | 35 |
| Ti | 4230 | 3645 | 3 | 25610 | 5 | 1903 | 3186 | 5380 | 13070 |
| V | 40 | 45 | 1 | 358 | 1 | 7.2 | 18 | 40 | 150 |
| Cr | 60 | 60 | 1 | 530 | 1 | 23 | 45 | 80 | 240 |
| Mn | 470 | 500 | 4 | 5800 | 4 | 150 | 298 | 630 | 1570 |
| Fe | 26420 | 24245 | 20 | 181690 | 40 | 8760 | 19880 | 35450 | 87270 |
| Co | 6 | 20 | 0.4 | 110 | 1 | 1 | 1 | 1 | 70 |
| Ni | 20 | 20 | 0.3 | 190 | 0.3 | 5 | 10 | 20 | 70 |
| Cu | 15 | 10 | 0.4 | 90 | 1 | 6 | 10 | 20 | 50 |
| Zn | 30 | 25 | 0.4 | 140 | 1 | 11 | 20 | 40 | 90 |
| Ga | 8 | 5 | 0.2 | 30 | 0.2 | 4 | 7 | 10 | 20 |
| As | 0.3 | 1 | 0.1 | 10 | 0.2 | 0.2 | 0.2 | 0.2 | 1 |
| Br | 4 | 10 | 0.2 | 250 | 0.2 | 0.2 | 1 | 3 | 30 |
| Rb | 50 | 35 | 1 | 260 | 1 | 20 | 40 | 70 | 135 |
| Sr | 115 | 180 | 1 | 1380 | 1 | 20 | 50 | 120 | 810 |
| Y | 10 | 10 | 0.2 | 105 | 0.2 | 5 | 9 | 20 | 40 |
| Zr | 80 | 120 | 18 | 950 | 18 | 20 | 20 | 80 | 390 |
| Ba | 2175 | 3050 | 3 | 48160 | 5 | 270 | 1470 | 3015 | 8365 |
| La | 1030 | 940 | 8 | 4840 | 8 | 330 | 690 | 1470 | 3440 |
| Ce | 60 | 60 | 1 | 380 | 1 | 15 | 40 | 80 | 250 |
| Pr | 1 | 2 | 0.4 | 30 | 0.4 | 1 | 1 | 0.9 | 7 |
| Nd | 10 | 10 | 1 | 50 | 1 | 2 | 6 | 10 | 40 |
| Sm | 10 | 12 | 1 | 80 | 1 | 2 | 7 | 15 | 40 |
| Hf | 4 | 5 | 0.2 | 40 | 0.2 | 1 | 2 | 5 | 18 |
| Ta | 3 | 2 | 0.1 | 15 | 0.1 | 1 | 3 | 4 | 9 |
| W | 0.4 | 1 | 0.1 | 6 | 0.2 | 0.2 | 0.2 | 0.2 | 2 |
| Pb | 33 | 55 | 0.3 | 640 | 0.3 | 7 | 20 | 35 | 190 |
| Bi | 2 | 4 | 0.1 | 30 | 0.1 | 0.1 | 0.1 | 1 | 11 |
| Th | 35 | 64 | 0.2 | 745 | 0.3 | 3 | 10 | 30 | 210 |

### 4.4.2 Prediction of soil properties by MIRS and TXRF

The details of predictive results of the RF models of the different spectra and combinations are shown in Table 4.3. For example, MIR spectra on their own resulted in good prediction models using RF out-of-bag validation ($R^2 > 0.80$) for organic and total C and N, Mehlich-3 Ca and Al, and pH. Also predicted well ($R^2 >= 0.60$) were Ca/Mg ratio, exchangeable bases, exchangeable Mg, phosphorus sorption index (PSI), and some particle size distribution variables (Table 4.3). Calibration models were less satisfactory ($R^2<0.60$) for Mehlich-3 extractable K, Mn, Fe, Cu, B, Zn, P, S, and Na, exchangeable acidity, electrical conductivity (Ecd), exchangeable sodium percentage (ESP), exchangeable sodium ratio (ESR), and air-dispersed particle sizes (Table 4.3). Our poor results for MIR prediction of extractable P ($R^2 = 0.10$) and K ($R^2 = 0.51$) were in agreement with the findings of other researchers e.g. Janik et al. (1998), Shepherd and Walsh (2002), Ludwig et al. (2002). The ability to predict levels of extractable cations is reported to vary with the extraction method (Cheng et al., 2001), and because soil supply of nutrients to plants depends on many interrelated soil factors, further work should investigate whether plant response to N, P, and K can be better predicted from soil IR than from soil extractions (Shepherd and Walsh, 2002). Also air-dispersed particle sizes were not satisfactory ($R^2 < 0.60$) predicted by MIR alone (Table 4.3). The latter results were in agreement with reports that prediction of particle size distribution with IR is variable due to different soil types (Kamau-Rewe et al., 2011 and Shepherd, 2010).

The TXRF raw data matrix was also able to partly predict the wet chemistry reference data, although the predictions were not better than those obtained using MIR first derivative spectra with the exception of exchangeable Mg and Na, Mehlich-3 Mn, B, Cu, Zn and S, ESP, Ecd and ESR (Table 4.3). As expected TXRF was poorer ($R^2 < 0.72$) in predicting organic C, total C, and organic N compared to MIR ($R^2 < 0.85$) as these elements are not directly determined with TXRF, however the variance explained is still quite high and may be attributable to TXRF signatures relating to soil fertility potential and mineralogy-related carbon protection mechanisms.

Table 4.3: Results of Random Forests model accuracy obtained for reference analysis data regression to first derivative MIR spectra, TXRF raw spectra on its own, raw TXRF spectra combined with first derivative MIR spectra. Results are for out-of-bag validation.

| Property[†] | n | MIR 1st derivative | | n | TXRF raw on its own | | n | MIR 1st derivative + TXRF raw | |
|---|---|---|---|---|---|---|---|---|---|
| | | $R^2$ | RMSE | | $R^2$ | RMSE | | $R^2$ | RMSE |
| Organic carbon | 700 | 0.90 | 0.2 | 555 | 0.70 | 0.6 | 555 | 0.92 | 0.2 |
| Total carbon | 651 | 0.90 | 0.2 | 555 | 0.72 | 0.6 | 555 | 0.93 | 0.2 |
| M3.Al | 700 | 0.86 | 29701 | 555 | 0.81 | 42689 | 555 | 0.86 | 31023 |
| Total nitrogen | 651 | 0.86 | 0.001 | 555 | 0.63 | 0.003 | 555 | 0.86 | 0.001 |
| Organic nitrogen | 651 | 0.85 | 0.001 | 555 | 0.65 | 0.003 | 555 | 0.86 | 0.001 |
| ExCa | 700 | 0.84 | 52 | 555 | 0.79 | 73 | 555 | 0.86 | 48 |
| pH | 700 | 0.82 | 0.2 | 555 | 0.72 | 0.3 | 555 | 0.81 | 0.2 |
| ExBas | 700 | 0.79 | 104 | 555 | 0.44 | 111 | 555 | 0.81 | 105 |
| PSI | 684 | 0.77 | 1465 | 555 | 0.73 | 1959 | 555 | 0.76 | 1731 |
| psa.c4sand | 699 | 0.75 | 168 | 555 | 0.71 | 201 | 555 | 0.73 | 186 |
| psa.c4clay | 699 | 0.74 | 142 | 555 | 0.68 | 180 | 555 | 0.70 | 166 |
| psa.w4sand | 699 | 0.74 | 150 | 555 | 0.71 | 177 | 555 | 0.74 | 162 |
| ExMg | 700 | 0.73 | 3.4 | 555 | 0.77 | 2.6 | 555 | 0.76 | 2.8 |
| psa.w4clay | 699 | 0.73 | 109 | 555 | 0.64 | 155 | 555 | 0.72 | 121 |
| CaMg | 700 | 0.60 | 15 | 555 | 0.41 | 22 | 555 | 0.48 | 19 |
| psa.w4silt | 699 | 0.60 | 46 | 555 | 0.43 | 64 | 555 | 0.56 | 49 |
| psa.c4silt | 699 | 0.59 | 46 | 555 | 0.43 | 62 | 555 | 0.55 | 49 |
| ExAc | 700 | 0.58 | 0.1 | 555 | 0.44 | 0.1 | 555 | 0.56 | 0.1 |
| M3.Mn | 700 | 0.56 | 4887 | 555 | 0.60 | 3838 | 555 | 0.63 | 3617 |
| ExK | 700 | 0.51 | 0.4 | 555 | 0.51 | 0.5 | 555 | 0.48 | 0.5 |
| M3.B | 700 | 0.51 | 0.4 | 555 | 0.66 | 0.3 | 555 | 0.69 | 0.3 |
| psa.aclay | 699 | 0.51 | 20 | 555 | 0.41 | 16 | 555 | 0.39 | 16 |
| psa.asilt | 699 | 0.49 | 23 | 555 | 0.45 | 25 | 555 | 0.51 | 22 |
| M3.Cu | 700 | 0.43 | 2.7 | 555 | 0.61 | 1.9 | 555 | 0.53 | 2.3 |
| M3.Fe | 700 | 0.41 | 5211 | 555 | 0.30 | 5423 | 555 | 0.40 | 4664 |
| psa.asand | 699 | 0.33 | 517 | 555 | 0.32 | 605 | 555 | 0.36 | 568 |
| ESP | 700 | 0.20 | 30 | 555 | 0.72 | 11 | 555 | 0.61 | 15 |
| ECd | 700 | 0.18 | 0.1 | 555 | 0.72 | 0.03 | 555 | 0.68 | 0.03 |
| ESR | 700 | 0.17 | 0.04 | 555 | 0.76 | 0.01 | 555 | 0.63 | 0.02 |
| M3.P | 700 | 0.10 | 966 | 555 | 0.11 | 1050 | 555 | 0.02 | 1155 |
| ExNa | 700 | 0.07 | 45 | 555 | 0.68 | 19 | 555 | 0.65 | 21 |
| M3.Zn | 700 | 0.04 | 4.5 | 555 | 0.13 | 4.5 | 555 | 0.10 | 4.2 |
| M3.S | 699 | 0.13 | 38251 | 555 | 0.61 | 16691 | 555 | 0.58 | 17767 |

[†] pH = Soil reaction (1:2 volume water extract); ExAc = Exchangeable acidity (unbuffered KCl extraction); ECd = Electrical conductivity (1:2 volume water extract); ExCa = Exchangeable Ca; ExK = Exchangeable K; ExMg = Exchangeable Mg; ExNa = Exchangeable Na; ExBas =Exchangeable bases (sum of Mehlich exch Ca, Mg, K, Na); CaMg = Ca/Mg ratio; ESP = Exchangeable sodium percentage (100.ExNa/(ExCa+ExMg+ExNa); ESR = Exchangeable sodium ratio; m3.P = Mehlich 3 extractable P; m3.Al = Mehlich 3 extractable Al; m3.B = Mehlich 3 extractable B; m3.cu = Mehlich 3 extractable Cu; m3.Fe = Mehlich 3 extractable Fe; m3.Mn = Mehlich 3 extractable Mn; m3.Zn = Mehlich 3 extractable Zn; m3.S = Mehlich 3 extractable S; PSI = Phosphorus sorption index; psa_asand = sand content (>50 microns) for air-dispersed particles by laser diffraction; psa_asilt = silt content (8 - 50 microns) for air-dispersed particles by laser diffraction; psa_aclay = clay content (< 8 microns) for air-dispersed particles by laser diffraction; psa_c4clay = clay content (<8 microns) for calgon dispersed particles after 4 min of ultrasonification; psa_c4silt = silt content (8 - 50 microns) for calgon dispersed particles after 4 min of ultrasonification; psa_c4sand = sand content (> 50 microns) for calgon dispersed particles after 4 min of ultrasonication; psa_w4clay = clay content (<8 microns) for water dispersed particles after 4 min of ultrasonification; psa_w4silt = silt content (8 - 50 microns) for water dispersed particles after 4 min of ultrasonification; psa_w4sand = sand content (> 50 microns) for water dispersed particles after 4 min of ultrasonication; n = number of samples; RMSE = root mean square error.

### 4.4.3 Prediction of soil properties combining MIRS and TXRF

As hypothesised, including total element concentration data from TXRF analysis in the RF models significantly reduced the root mean square errors of prediction by 63% for Ecd, 54% for Mehlich-3 S, 53% for exchangeable Na, 50% for ESP, 50% for ESR, 29% for total C, 28% for Mehlich-3 B, 26% for Mehlich-3 Mn, 17% for exchangeable Mg, 15% for Mehlich-3 Cu, 11% for Mehlich-3 Fe, 10% for organic C, 6% for Mehlich-3 Zn, and 4% for air-dispersed particle sizes for silt content (8-50 microns) by laser diffraction (Table 4.3). However, the prediction improvement from including TXRF was due to detection of a few outlier samples that did not appear as MIR spectral outliers (Figure 4.2). In addition, there was good ability of TXRF to predict the residuals in e.g. ExNa and the derived ESP test and when 3 outliers were omitted the advantage of adding TRXF data was lost. The outliers had exceptionally high total Cl concentrations (>21,500 mg $kg^{-1}$) and were all from the same sentinel site, Ihassunge in Mozambique. It is likely that these were sodic or saline sites with pH values ranging from 3.6-5.3 (mean 4.6) and electrical conductivity (ECd) values in the range of 3.1-4.8dS/m (mean 3.9 dS/m) for these 3 samples. However, MIR models for most soil fertility properties with the exception of ESP, ESR, exchangeable bases, exchangeable acidity, exchangeable Mg and Na, as well as Mehlich-3 B and S, did not improve when these 3 outliers were also omitted from the MIR spectral data indicating that these were not MIR spectral outliers. Thus, we could hypothesise that TXRF may have potential for outlier detection in large data sets or even potential use as a site stratification tool.

Our hypothesis that supplementing MIR data with TXRF fingerprints would improve overall prediction of reference data did not hold. The lack of relation of extractable P and K to either IR spectra or total element patterns leads us to question the efficacy of the reference tests themselves. We propose future studies should examine direct relationships between crop response in pot and field studies and MIR or TXRF patterns, as especially in African soils that have not received large fertilizer additions we would expect crop response to P and K fertilizers to relate to mineralogy and organic characteristics detected by these methods. IR detects factors that control P availability (mineralogy) and P sorption (carbonates, Fe and Al oxides). In addition as P occurs as inorganic phosphate composed of a $PO_4^{-3}$ anion and a metallic cation and because of the charges on these ions, the dipole moments of these bonds are large and consequently the MIR bands are characteristically strong (Smith, 1998). The important MIR vibrations of the phosphate group are phosphorus-oxygen stretching and bending vibrations, with the stretch being very intense and broad and appearing between 1100 and 1000 $cm^{-1}$ but unfortunately, sulfates and silicates also have strong broad bands in this region (Smith, 1998). Thus, one must make use of secondary bands, such as bending vibrations to be able to distinguish these molecules, such as the important secondary band for phosphates which is a bending vibration found on the edge of the MIR spectral range between 600 and 500 $cm^{-1}$ (Smith, 1998). A somewhat similar approach has been suggested by Cobo et al. (2010) showing that MIRS data could be directly integrated, after principal component analyses, in geostatistic assessments without the necessity of extensive calibration/validation steps.

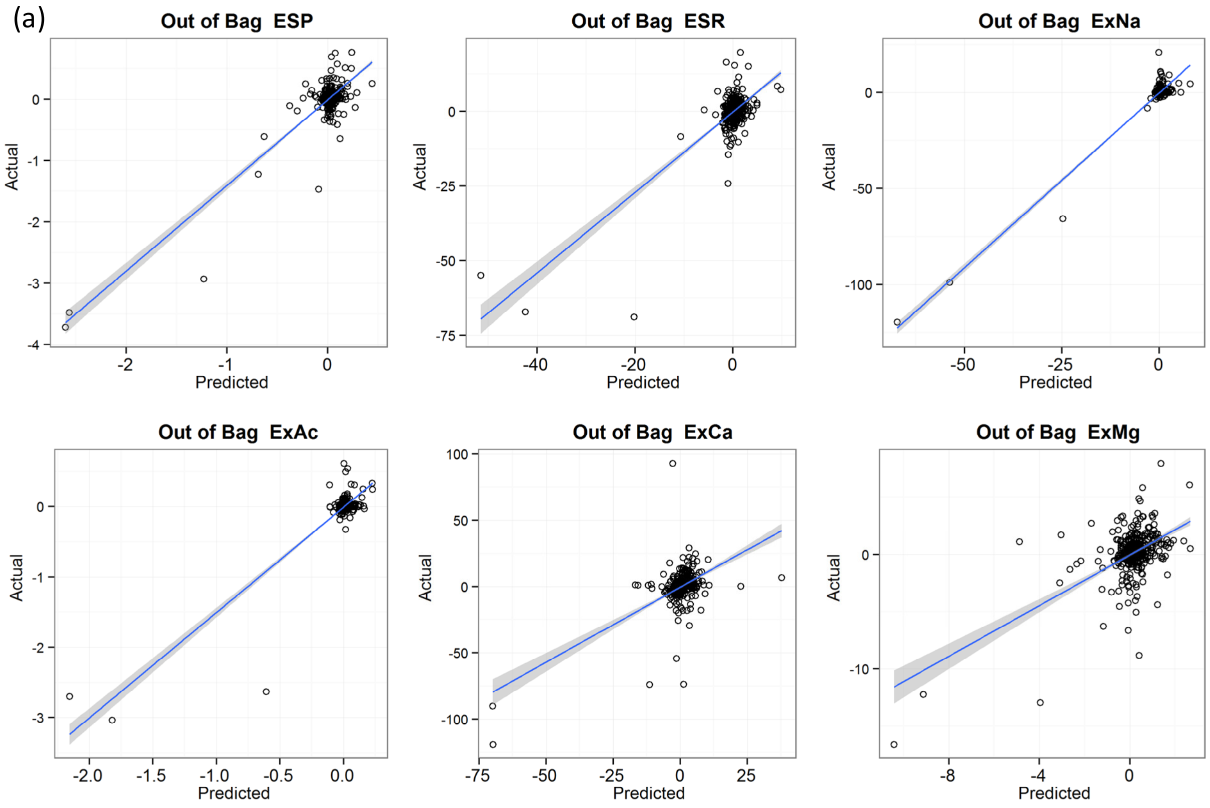


Figure 4.2a: Scatterplots of prediction models developed from MIR residuals combined with TXRF spectra for some soil properties that TXRF was able to pick out differences in samples during the calibration of residuals of prediction of MIR first derivative reflectance spectra to TXRF spectra using Random Forests with an out-of-bag validation.

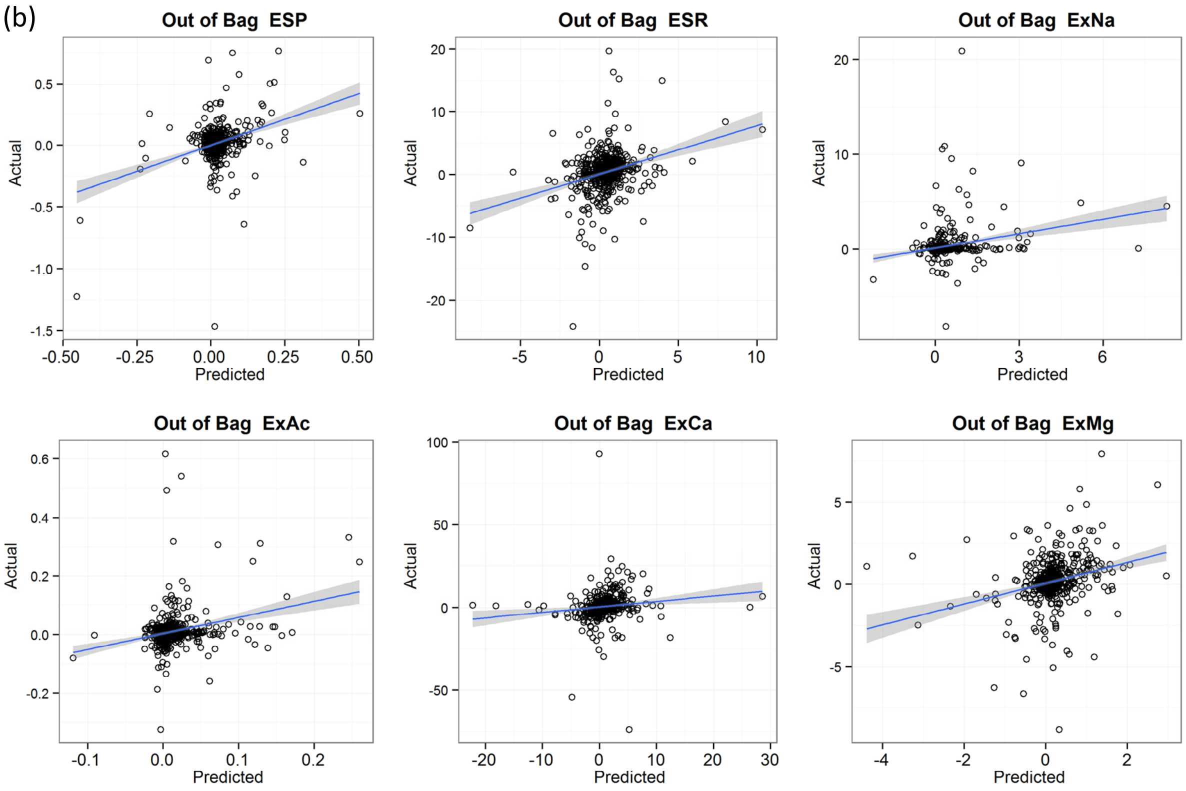


Figure 4.2b: Scatterplots of prediction models developed from MIR residuals combined with TXRF spectra for some soil properties that TXRF was able to pick out differences in samples during the calibration of residuals of prediction of MIR first derivative reflectance spectra to TXRF spectra using Random Forests with an out-of-bag validation for the entire sample set but excluding three outlier samples.

### 4.4.4 Comparison of results of RF with those of PLS on same data set

Comparing the results of the RF algorithm with an out-of-bag validation (RF-OOB) with results of PLS on the same data set but using a hold-out validation suggested that RF out-performed the PLS regression for the simultaneous determination of soil fertility properties (Table 4.4). These results were in agreement with the report by Ghasemi and Tavakoli (2013) that RF models have a better performance than PLS regression models. RF regression results were also computationally much faster and did not rely on data transformations or any assumptions about the data distributions compared to PLS. In addition, RF was relatively robust to outliers and noise in a data set as well as having a more interpretable algorithm than PLS and is hence an alternative approach in multivariate calibration. Since $R^2$ is reported to be sensitive to asymmetric distribution of the data and density of the spread of points in the plots of predicted versus experimental data, Ghesemi and Tavakoli (2013) recommended the root-mean-square errors (RMSE) as a better measure for comparison of the quality of the models. In addition, $R^2$ is not robust to outliers and thus the RMSE values were used in the current study for comparison of the models. The $R^2$ results of a 70% holdout validation were similar to those of the RF-OOB validation, however the RMSE for the same soil fertility properties predicted using MIRS spectra were lower in RF-OOB that their PLS holdout counterparts indicating that RF procedure was not overfitting (Table 4.4). This result was in agreement with the report of Wei et al. (2012) that the RF multivariate algorithm does not over-fit and usually performs well in situations with a low sample-to-features ratio, such as spectrometric data. However, Ghesemi and Tavakoli (2013) suggested that in general, the selection of a suitable method for multivariate calibration should consider Wolpert's No Free Lunch theorem that says that there is no single best algorithm that works for all problems. Thus, we take the output of RF analysis as the smarter algorithm because it has been proven to be helpful in leading to a deeper understanding of the spectrometric data problems (such as low sample-to-features ratio) (Wei et al., 2012). We recommend to use the PLS method for small data stes, and RF for large data sets because we have found these to be optimal on average over many data sets. The emphasis must always be on what approach provides the best predictive ability on independent samples taken from a target area one is developing calibrations for.

Table 4.4: Model summary results comparing calibration of soil properties to MIRS first derivative reflectance spectra using Random Forests with an out-of-bag (RF OOB) validation for the entire sample set, and PLS 70% holdout calibration and prediction using 30% independent training set.

| Property[†] | n | RF OOB models using all data | | n | PLS 70% Holdout models | | n | PLS 30% testing set predictions | |
|---|---|---|---|---|---|---|---|---|---|
| | | $R^2$ | RMSE | | $R^2$ | RMSE | | $R^2$ | RMSE |
| Acidified carbon | 651 | 0.90 | 0.20 | 448 | 0.87 | 0.26 | 203 | 0.81 | 0.6 |
| Total carbon | 651 | 0.90 | 0.22 | 448 | 0.87 | 0.27 | 203 | 0.81 | 0.7 |
| M3.Al | 700 | 0.86 | 29701 | 490 | 0.83 | 30838 | 210 | 0.86 | 201 |
| Total nitrogen | 651 | 0.86 | 0.001 | 448 | 0.86 | 0.001 | 203 | 0.81 | 0.04 |
| Acidified nitrogen | 651 | 0.85 | 0.001 | 448 | 0.86 | 0.001 | 203 | 0.81 | 0.04 |
| ExCa | 700 | 0.84 | 52 | 490 | 0.76 | 51 | 210 | 0.83 | 10 |
| pH | 700 | 0.82 | 0.19 | 490 | 0.80 | 0.20 | 210 | 0.79 | 0.5 |
| ExBas | 700 | 0.79 | 104 | 490 | 0.73 | 85 | 210 | 0.74 | 15 |
| PSI | 684 | 0.77 | 1465 | 483 | 0.71 | 1775 | 201 | 0.71 | |
| Psa.c4sand | 699 | 0.75 | 168 | 489 | 0.75 | 166 | 210 | 0.67 | 47 |
| Psa.c4clay | 699 | 0.74 | 142 | 489 | 0.69 | 161 | 210 | 0.69 | 15 |
| Psa.w4sand | 699 | 0.74 | 150 | 489 | 0.73 | 152 | 210 | 0.72 | 13 |
| ExMg | 700 | 0.73 | 3.4 | 490 | 0.60 | 3.0 | 210 | 0.50 | 13 |
| Psa.w4clay | 699 | 0.73 | 109 | 489 | 0.69 | 119 | 210 | 0.67 | 3. 6 |
| CaMg | 700 | 0.60 | 15 | 490 | 0.41 | 22 | 210 | 0.68 | 12 |
| Psa.w4silt | 699 | 0.60 | 46 | 489 | 0.56 | 48 | 210 | 0.51 | 3.6 |
| Psa.c4silt | 699 | 0.59 | 46 | 489 | 0.52 | 50 | 210 | 0.57 | 7.8 |
| ExAc | 700 | 0.58 | 0.10 | 490 | 0.55 | 0.11 | 210 | 0.53 | 7.4 |
| M3.Mn | 700 | 0.56 | 4887 | 490 | 0.54 | 5504 | 210 | 0.47 | 0.4 |
| ExK | 700 | 0.51 | 0.4 | 490 | 0.67 | 0.2 | 210 | 0.30 | 70.6 |
| M3.B | 700 | 0.51 | 0.4 | 490 | 0.48 | 0.39 | 210 | 0.25 | 1.0 |
| Psa.aclay | 699 | 0.51 | 20 | 489 | 0.54 | 20 | 210 | 0.27 | 0. 8 |
| Psa.asilt | 699 | 0.49 | 23 | 489 | 0.43 | 23 | 210 | 0.26 | 5.2 |
| M3.Cu | 700 | 0.43 | 2.7 | 490 | 0.47 | 2.05 | 210 | 0.48 | 6.3 |
| M3.Fe | 700 | 0.41 | 5211 | 490 | 0.39 | 6064 | 210 | 0.24 | 1.9 |
| Psa.asand | 699 | 0.33 | 517 | 489 | 0.36 | 542 | 210 | 0.23 | 70 |
| ESP | 700 | 0.20 | 30 | 490 | 0.16 | 28 | 210 | 0.03 | 22 |
| ECd | 700 | 0.18 | 0.08 | 490 | -0.02 | 0.07 | 210 | 0.18 | 6.9 |
| ESR | 700 | 0.17 | 0.04 | 490 | -0.04 | 0.04 | 210 | 0.09 | 0.4 |
| M3.P | 700 | 0.10 | 966 | 490 | 0.11 | 832 | 210 | 0.08 | 0.3 |
| ExNa | 700 | 0.07 | 45 | 490 | -0.12 | 33 | 210 | 0.13 | 36 |
| M3.Zn | 700 | 0.04 | 4.48 | 490 | 0.17 | 4.42 | 210 | 0.02 | 8.9 |
| M3.S | 699 | -0.13 | 38251 | 489 | -0.08 | 40217 | 210 | 0.10 | 1.8 |

[†] pH = Soil reaction (1:2 volume water extract); ExAc = Exchangeable acidity (unbuffered KCl extraction); Ecd = Electrical conductivity (1:2 volume water extract); ExCa = Exchangeable Ca; ExK = Exchangeable K; ExMg = Exchangeable Mg; ExNa = Exchangeable Na; ExBas =Exchangeable bases (sum of Mehlich exch Ca, Mg, K, Na); CaMg = Ca/Mg ratio; ESP = Exchangeable sodium percentage (100.ExNa/(ExCa+ExMg+ExNa); ESR = Exchangeable- sodium ratio; m3.P = Mehlich 3 extractable P; m3.Al = Mehlich 3 extractable Al; m3.B = Mehlich 3 extractable B; m3.cu = Mehlich 3 extractable Cu; m3.Fe = Mehlich 3 extractable Fe; m3.Mn = Mehlich 3 extractable Mn; m3.Zn = Mehlich 3 extractable Zn; m3.S = Mehlich 3 extractable S; PSI = Phosphorus sorption index; psa_asand = sand content (>50 microns) for air-dispersed particles by laser diffraction; psa_asilt = silt content (8 - 50 microns) for air-dispersed particles by laser diffraction; psa_aclay = clay content (< 8 microns) for air-dispersed particles by laser diffraction; psa_c4clay = clay content (<8 microns) for calgon dispersed particles after 4 min of ultrasonification; psa_c4silt = silt content (8 - 50 microns) for calgon dispersed particles after 4 min of ultrasonification; psa_c4sand = sand content (> 50 microns) for calgon dispersed particles after 4 min of ultrasonication; psa_w4clay = clay content (<8 microns) for water dispersed particles after 4 min of ultrasonification; psa_w4silt = silt content (8 - 50 microns) for water dispersed particles after 4 min of ultrasonification; psa_w4sand = sand content (> 50 microns) for water dispersed particles after 4 min of ultrasonication; n = number of samples; RMSE = root mean square error.

## 4.5 Conclusion

This paper validated that low cost high-throughput IR and TXRF spectroscopy can predict a number of soil properties in diverse African soils and could be used both as a front line screening technique for sample selection for more expensive tests and for development of calibrations for direct prediction of soil properties. Globally applicable calibrations to predict standard soil properties based on infrared spectra may increase the use of this low cost technique especially in developing countries of Sub-Saharan Africa (SSA) where reliable data on soil chemical and physical information needed to give advice on land management is sparse and dated. RF was a computationally fast and robust to analyse complex spectra of multi-component analytical methods (TXRF and MIR). MIR showed remarkable ability to capture total element composition effects on physic-chemical soil properties. However, TXRF data used as a predictor did not add value to MIR beyond identifying outlying samples, but these outliers did not appear as MIR spectral outliers, hence, TXRF may have potential for outlier detection and as a site stratification tool. TXRF may be a useful tool for simultaneous determination of elemental composition and a predictor of conventionally measured soil properties to a medium level of accuracy. This may be an advantage in environmental monitoring where absolute element concentration levels are of interest as well as their relations to other soil properties. The lack of any relation of extractable P and K with IR and TXRF data in contrast with their strong relationships with soil mineral and organic composition calls into question the value of the weak extraction soil tests for advising on soil fertility management. We thus recommend further studies to relate crop response to fertility management to IR and TXRF signatures as they provide and integrated measure of soil functional properties.

## 4.6 Acknowledgement

The research work undertaken for this study was financially supported by the World Agroforestry Centre (ICRAF) through the AFSIS project and the Dr. Hermann Eiselen Ph.D. Grant from the Foundation fiat panis. E.K. Towett acknowledges the Young Excellence Scholars (YES) PhD Scholarship from the Food Security Centre at the University of Hohenheim, Germany and a travel grant as well as stipend within the framework of the “Global Food Security” funded by the Deutscher Akademischer Austausch Dienst (DAAD). The authors are grateful to Robin Chacha, Hezekiah Nyandika, Elvis Weullow, Dickens Ateku and Jane Ndirangu for their assistance with TXRF and MIR analyses and thank all the technical staff at the ICRAF Soil Plant Spectral Diagnostics Laboratory in Nairobi for their assistance with sample processing for spectroscopic analyses.

## 4.7 References

1. Acosta, J.A., Martínez-Martínez, S., Faz, A., Arocena J., 2011. Accumulations of major and trace elements in particle size fractions of soils on eight different parent materials. Geoderma 161, 30-42.
2. AfSIS, 2013. Africa Soil Information Service. www.africasoils.net. Assessed on 17.09.2013.
3. Anderson, J.M., Ingram, J.S.I., (eds), 1993. Tropical Soil Biology and Fertility. A Handbook of Methods. 2nd Ed. CABI, Wallingford, UK, 221 pp.
4. Awiti, A.O., Walsh, M.G., Shepherd, K.D., Kinyamario, J., 2007. Soil condition classification using infrared spectroscopy: A proposition for assessment of soil condition along a tropical forest-cropland chronosequence. Geoderma 143 (1–2), 73–84. doi:10.1016/j.geoderma.2007.08.021.
5. Bache, B.W., and Williams, E.G., 1971. A phosphate sorption index for soils. Journal of Soil Science 22, 289-301.
6. Brown, D.J., Shepherd, K.D., Walsh, M.G., Mays, D.M., Reinsch, T.G., 2006. Global soil characterization with VNIR diffuse reflectance spectroscopy. Geoderma 132, 273–290.
7. Breiman, L., 2001. Random Forests, Machine Learning 45(1), 5-32.
8. Cheng, C., Laird, D.A., Mausbach, M.J., Hurburgh, C.R., Jr., 2001. Near-infrared reflectance spectroscopy - principal components regression analyses of soil properties. Soil Science Society of American Journal 65 (2), 480.
9. Cobo, J.G., G. Dercon, T. Yekeye, L. Chapungu, C. Kadzere, A. Murwira, R. Delve, and G. Cadisch., 2010. Integration of mid-infrared spectroscopy and geostatistics in the assessment of soil spatial variability at landscape level. Geoderma 158, 398-411.
10. Du, C., Zhou, J., 2009. Evaluation of soil fertility using infrared spectroscopy: a review. Environmental Chemistry Letters 7 (2), 97-113.
11. Ghasemi, J. B., Tavakoli, H., 2013. Application of random forest regression to spectral multivariate calibration. Analytical Methods 5, 1863-1871.
12. Helmke, P.A., 2000. The chemical composition of soils, in: Sumner, M.E., (ed.), Handbook of Soil Science. CRC Press, Boca Raton, FL, pp. B3-B24.

13. Janik, L.J., Merry, R.H., Skjemstad, J.O., 1998. Can mid infrared diffuse reflectance analysis replace soil extractions? Australian Journal of Experimental Agriculture 38, 681-696.
14. Kabata-Pendias, A., Mukherjee, A.B., 2007. Trace elements from soil to human. Springer-verlag Berlin Heidelberg. pp 550.
15. Kamau-Rewe, M., Rasche, F., Cobo, J., Dercon, G., Shepherd, K.D., Cadisch, G., 2011. Generic prediction of soil organic carbon in Alfisols using diffuse reflectance Fourier-Transform Mid-Infrared spectroscopy. Soil Science Society of America Journal 75, 2358–2360.
16. Kennard, R., Stone, L., 1969. Computer aided design of experiments. Technometrics 11, 137-148.
17. Kibblewhite, M.G., Ritz, K., and Swift, M.J., 2008. Soil health in agricultural systems. Philosophical Transactions of the Royal Society B 363:685–701. doi:10.1098/rstb.2007.2178.
18. Kleinebecker, T., Poelen, M.D.M., Smolders, A.J.P., Lamers, L.P.M., and Hölze, N., 2013. Fast and Inexpensive Detection of Total and Extractable Element Concentrations in Aquatic Sediments Using Near-Infrared Reflectance Spectroscopy (NIRS). PLoS ONE, 8 (7): e70517. Doi:10.1371/journal.pone.0070517.
19. Kottek, M., Grieser, J., Beck, C., Rudolf, B, Rubel, F., 2006. World Map of the Köppen-Geiger climate classification updated. Meteorologische Zeitschrift 15, 259-263. doi: 10.1127/0941-2948/2006/0130.
20. Ludwig, B., Khanna, P.K., Bauhus, J., Hopmans, P., 2002. Near infrared spectroscopy of forest soils to determine chemical and biological properties related to soil sustainability. Forest Ecology and Management 71:121-132.
21. Mehlich, A., 1984. Mehlich 3 soil test extractant: A modification of the Mehlich 2 extractant. Commun. Soil Science Plant Analysis 15(12), 1409-1416.
22. Minasny, B., McBratney, A.B., 2008. Regression rules as a tool for predicting soil properties from infrared reflectance spectroscopy. Chemometrics and Intelligent Laboratory Systems 94, 72–79.
23. R Development Core Team, 2013. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. Available from: http://www.R-project.org. Accessed 23 Mar 2013.

24. Skjemstad, J., Baldock, J.A., 2008. Total and organic carbon. In 'Soil Sampling and Methods of Analysis. 2nd Edition' (MR Carter and EG Gregorich, eds.), pp. 225-238. Soil Science Society of Canada.
25. Shepherd, K.D., Walsh, M.G., 2002. Development of reflectance spectral libraries for characterization of soil properties. Soil Science Society of America Journal 66, 988-998.
26. Shepherd, K.D., Walsh, M.G., 2007. Infrared spectroscopy – enabling an evidence-based diagnostic surveillance approach to agricultural and environmental management in developing countries. Journal of Near Infrared Spectroscopy 15, 1–20.
27. Shepherd, K.D., 2010. Soil spectral diagnostics – infrared, x-ray and laser diffraction spectroscopy for rapid soil characterization in the Africa Soil Information Service. 19th World Congress of Soil Science, Soil Solutions for a Changing World 1–6 August 2010, Brisbane, Australia.
28. Smith, B.C., 1998. Infrared spectral interpretation: a systematic approach. CRS press LCC, Florida. pp 265.
29. Soriano-Disla, J.M., Janik, L., McLaughlin, M.J., Forrester, S., Kirby, J., Reimann, C., and the EuroGeoSurveys GEMAS Project Team, 2013. The use of diffuse reflectance mid-infrared spectroscopy for the prediction of the concentration of chemical elements estimated by X-ray fluorescence in agricultural and grazing European soils. Applied Goechemistry 29, 135-143.
30. Stosnach, H., 2005. Environmental trace-element analysis using a benchtop total reflection X-ray fluorescence spectrometer. The Japan Society of Analytical Chemistry. Analytical Sciences 21, 873-876.
31. Terhoeven-Urselmans, T., Spaargaren, O., Vågen, T-G., Shepherd, K.D., 2010. Prediction of Soil Fertility Properties from a Globally Distributed Soil Mid-Infrared Spectral Library. Soil Science Society of America Journal 74, 1792–1799.
32. Towett, E.K., Shepherd, K.D., Cadisch, G., 2013. Quantification of total element concentrations in soils using total X-ray fluorescence spectroscopy (TXRF). Science of the Total Environment 463–464: 374–388. http://dx.doi.org/10.1016/j.scitotenv.2013.05.068.
33. Towett, E.K., Shepherd, K.D., Tondoh, J.E., Winowieki, L.A., Lulseged, D., Sila, A., Cadisch, G., Submitted. Quantification of the variability and pattern in total element composition of sub-Saharan Africa soils using total X-ray fluorescence spectroscopy. Geoderma Journal.

34. UNEP, 2012. Land Health Surveillance: An Evidence-Based Approach to Land Ecosystem Management. Illustrated with a Case Study in the West Africa Sahel. United Nations Environment Programme, Nairobi. pp. 211.
35. Vågen, Tor-G., Winowiecki, L.A., Tondoh, J.E., Desta, L.T., 2013. Africa Soil Information Service (AfSIS) - Soil Health Mapping. http://hdl.handle.net/1902.1/19793 V2 [Version].
36. Wei, Y., Li, B., Xi, L., Yao, X., Wu, X., 2012. Discrimination of Radix Angelicae Sinensis according to harvest time and geographical origin with mid-infrared diffuse reflectance spectroscopy and chemometric methods. Spectroscopy Letters 45, 430–437.

Annex Table A4.1. Location of AfSIS sites and their average coordinates from which soil samples (top soil, 0-20 cm) were collected from and the number of samples from each site collected (n = 700) for this study.

| Country | Sentinel Site | Latitude | Longitude | # of samples |
|---|---|---:|---:|---:|
| Burkina Faso | Bondigui | 10.91340417 | -3.546170833 | 16 |
| Cameroon | Bana | 5.268653853 | 10.25876531 | 16 |
| Ethiopia | Dambidolo | 8.613306761 | 35.01514411 | 16 |
| | Kutaber | 11.29683918 | 39.60666704 | 16 |
| | Mega | 4.18128863 | 38.29669666 | 16 |
| | Merar | 9.602970243 | 42.7006824 | 16 |
| Ghana | Ajumako | 5.407900313 | -0.745416042 | 16 |
| | Kubeasi | 6.722705209 | -1.276720625 | 16 |
| | Lambussie | 10.89595744 | -2.649256444 | 16 |
| Guinea | Boumeoul | 11.93431708 | -13.14245333 | 16 |
| | Fria | 10.50737167 | -13.39262708 | 16 |
| Kenya | Marafa | -2.634250729 | 39.54624796 | 16 |
| | Mbalambala | -0.125638376 | 39.02689099 | 16 |
| | Morijo | -1.719174498 | 35.81122851 | 16 |
| Malawi | Nkhata Bay | -11.62575948 | 34.23945417 | 16 |
| | Thuchila | -15.94035406 | 35.32925635 | 16 |
| Mali | Finnkolo | 11.3128025 | -5.501523229 | 16 |
| | Koloko | 12.48264011 | -6.295524688 | 16 |
| | Kontela | 14.8088874 | -10.99925917 | 16 |
| Mozambique | Chica_b | -14.71417489 | 39.87284344 | 16 |
| | Chiculecule | -22.83744229 | 35.30298042 | 16 |
| | Ihassunge | -18.01346133 | 36.83151658 | 16 |
| | Macassangila | -13.40821094 | 35.49424813 | 16 |
| | Martinho | -16.14472333 | 38.97313823 | 16 |
| | Massuque | -24.20596906 | 34.36484344 | 16 |
| Nigeria | Ibi | 8.138450555 | 9.894032444 | 15 |
| | Imorun | 6.753241041 | 4.658440209 | 16 |
| | Katsina ala | 7.01415698 | 9.34168625 | 16 |
| South Africa | Hopetown | -29.6167 | 24.0833 | 16 |
| | Madadeni | -27.57691875 | 30.0537875 | 16 |
| | Prieska | -29.94247333 | 22.74732667 | 15 |
| Tanzania | Bukwaya | -3.023720875 | 33.05021692 | 16 |
| | Chinyanghuku | -6.895035052 | 36.12872167 | 15 |
| | Itende | -6.891425669 | 34.20797873 | 16 |
| | Kiberashi | -5.346139641 | 37.48160624 | 16 |
| | Kidatu | -8.035865843 | 37.30332637 | 16 |
| | Kisongo | -3.355084375 | 36.54144375 | 16 |
| | Mbinga | -11.09117067 | 35.16088915 | 16 |
| | Pandambili | -6.084341229 | 36.47446958 | 16 |
| Uganda | Hoima | 1.4167 | 31.0833 | 16 |
| | Namasuba | 0.524841943 | 32.29991484 | 16 |
| Zambia | Chilende | -9.172582021 | 29.66256313 | 16 |
| | Fisenge | -13.09518131 | 28.47630013 | 15 |
| | Monga | -16.07885156 | 28.29504042 | 16 |

# Chapter 5

# Prediction of soil functional properties from infrared and X-ray soil spectral properties: synthesis and outlook

# 5.0 Prediction of soil properties from infrared and X-ray soil spectral properties: synthesis and outlook

## 5.1 Review of the answers provided to the research's primary objectives

The current research study has analyzed soil samples gathered at the various randomized sentinel sites and proven that MIR DRIFT spectroscopy, XRD, TXRF technology methods are useful for rapid, non-destructive characterization of the composition of materials based on the interaction of electromagnetic energy with matter. The analytical methods (MIR, XRD and TXRF) have results that partially overlap and confirm each other and hence give a promising result to our examination and thus, all the techniques in combination can achieve reliable, definite, and accurate results, and provide additional information about the mineralogical, chemical, and physical properties of soils. The findings, therefore, present opportunity to improve soil assessments using the high-throughput spectral methods using information revealed by the various spectral methods. For instance, XRD information on soil mineralogy can be combined with information from infrared and total X-ray fluorescence spectroscopy, which characterize soil physico-chemical properties, to provide powerful diagnostic capabilities, and be used as complementary inputs to pedo-transfer functions for low-cost and rapid prediction of soil functional properties for agricultural and environmental applications, particularly in the case of SSA where reliable data is inadequate.

All the soil samples used in the current study were analysed for total element concentrations using TXRF. The study also compared the results of analysis with literature values and the results showed that total element concentration values were within the range reported globally for soil Cr, Mn, Zn, Ni, V, Sr, and Y and in the high range for Al, Cu, Ta, Pb, and Ga. We also established total concentrations of elements for soils occurring within particular sites sampled in our study, and documented systematic variation in their concentration and explored the possibility of finger-printing complete element profiles. This study has demonstrated that there are significant variations ($P < 0.05$) in total element composition within and between the sites for all the elements analysed, with the greatest proportion of total variance and number of significant variance components occurring at the site (55-88%) followed by the cluster nested within site levels (10-40%). We have concluded that the substantive variations between and within the sites analysed can inform on the soil fertility potential. The soils exhibited a wide range of relationships between the physico-chemical and mineralogical properties, in a principal component analysis done using ‘FactoMineR’ package in R and in addition, the cluster co-occurrence of individual element concentrations and the soil fertility variables identified across the different sentinel sites is expected to relate to differences in mineralogical and site or soil-forming factors. This was confirmed by our explorations of the relationships between element profiles and mineral data as well as other site and soil-forming factors which emphasized the control exerted by weathering of parent materials and all soil forming factors on the variation of the total element concentrations. In addition, this thesis has explored the relationships between the total element concentrations in the soils analysed and other site and soil-forming factors using ‘randomForest’ package in R and revealed that all soil soil-forming factors (e.g., parent material, climate, topography, management (landuse)) have an important influence on total elemental concentrations in the soil.

## 5.2 Present implications of the outcomes of this study for food security in Sub-Saharan Africa

Successful use of the spectroscopic techniques tested in this study such as the TXRF technique would open up the possibilities for using total element composition to improve global MIRS predictions of soil properties, such as cation exchange capacity and extractable nutrients. Especially in Africa where variations in soil mineralogy and nutrient balance critically determine vegetation composition and agricultural potential (Voortman, 2011) TXRF could provide a particularly useful tool for prediction of soil properties in data sparse regions. Despite the importance of soil mineralogy in determining soil properties (Jenny, 1941), there have been few attempts to quantitatively link functional capacity to mineralogy or total element composition. Hence, this study has evaluated how far TXRF total element derived patterns in soils relate to ‘available’ element results from Mehlich-3 soil tests (acid-extractable nutrients) and hence to soil nutrient supply capacity.

In addition, this study has shown that an important soil capital, in particular phosphorous, is inherently low in all soils analysed and this basic problem needs to be addressed in Africa. Going forward, the trend in African soils fast becoming nutrient deficient and thus spelling a bleak future for many due to low crop yields also needs to be considered in further studies, as this was beyond the scope of the current study. However, not all is lost because the outcomes of this study will contribute to ICRAF’s projects aimed at replenishing the diminishing African soil capital for better yields and improved livelihoods, especially through the AfSIS Project, thus enabling stakeholders to get better information on problems and opportunities relating to soil management in Africa. The AfSIS project is also conducting crop testing trials to see how the soils respond to fertilizers, and the outcomes of the tests together with those from this study will help to give advice to different stakeholders on appropriate soil management interventions for different types of soils and locations. Use of the state-of-the-art spectral diagnostic methods described here for large-area soil health measurement and monitoring will accelerate economic development in developing SSA countries with regards to climate change, increasing water scarcity and impacts on local and global food security as well as sustainable agricultural production and ecosystem resilience in the tropics.

## 5.3 Implication of spectral approaches for soil diagnosis

Because soil is heterogeneous and highly variable due to the complex interaction of soil-forming factors such as parent material, climate organisms, topography and human impact that are interacting through space and time, the characterization of soil properties' variability normally represent a serious challenge when an accurate data is needed (Knadel et al., 2013). Spectroscopic approaches for soil diagnosis e.g. MIRS are proven technology for rapid, non-destructive characterization of the composition of materials based on the interaction of electromagnetic energy with matter can be applied to predict a number of important soil properties as outlined in the fourth chapter of this thesis. Infrared spectroscopy has been a key technology in enabling the development of soil health surveillance systems by providing a rapid and reliable tool for soil health screening (Shepherd & Walsh, 2007). Nevertheless, because soil spectra such as MIRS spectra can contain information on the fundamental composition of soils, they can be used alone to describe the soil type and how it varies across SSA landscapes. By testing other new spectroscopic techniques, this study has demonstrated an increased speed with which accurate and detailed soil information can be made accessible, especially by the AfSIS Project. AfSIS is building spectral diagnostic libraries of all reference soil samples taken from the sentinel sites as well as from the soil management (diagnostic trial) experiments (AfSIS, 2013). In addition to the central ICRAF facility located in Nairobi, other new regional laboratories e.g. in Mali at the Institut d'Economie Rural (IER) in Bamako, in Tanzania at the Agricultural Research Institute (ARI) in Mlingano, near Arusha, and in Malawi at the

Department of Agricultural Research and Services in Chitedze, near Lilongwe, and many others have now been equipped to carry out spectral analyses. Capacity building support to the national research institutions in Africa and projects e.g. the EthioSIS Project, in addition to advisory services, such as on multivariate calibration model development and instrumentation troubleshooting, is provided on a continuous basis by ICRAF. In addition, the AfSIS field offices are also being equipped with portable field spectrometers to ensure that soils can be rapidly analyzed in situations where exporting physical samples to a centralized laboratory would be difficult, to select optimal subsets of samples for shipping for more expensive analyses, to develop field-based, spectrometric diagnostic tests, and to radiometrically calibrate remote sensing images (AfSIS, 2013). Spectral and reference analyses are being done at ICRAF's Spectral Diagnostics Laboratory to ensure consistency in methods, and quality control is conducted to ensure spectral quality across the spectral diagnostic laboratory network (AfSIS, 2013). The spectral diagnostics instruments will also provide the capacity to analyze a wide range of agricultural inputs (organic resources) and can be an important additional contribution to validation of integrated soil fertility management practices (e.g. monitoring improvements in manure/compost nutrient status).

In chapter 3, this thesis has also attempted to provide the missing information on the total elemental compositions of African soils and to link element concentration fingerprints of soil mineralogy to soil function with the recognition of key soil forming factors. In order to summarize the content in the spectra and to help with their interpretation, we performed a PCA e.g. of the TXRF data and then by implementing the PCA, relationships among the samples and the individual spectra were explained, giving a general overview of the main variability and pattern in total element concentrations of SSA soils using the TXRF technique. In addition, the resulting principal component scores were then combined with MIRS spectral data for the prediction of different soil properties. In chapter 3 it has been shown that the X-ray diffraction instrument, with capability for high throughput, allows quantitative soil mineralogy profiling and coupled with TXRF data as described in the present study could provide the missing link for evaluation of soil fertility and functional capacity and the mineral and element profile complements with infrared spectral analysis as a diagnostic screening tool. In Chapter 4 of this thesis the AfSIS Project's 32 soil samples per sentinel site for a sample set of 44 sentinel sites sites randomized over Sub-Saharan Africa have been analyzed using standard laboratory methods through the ICRAF Soil-Plant Spectral Diagnostics Laboratory for e.g. pH, electrical conductivity, exchangeable acidity, Mehlich extractable elements (Al, P, K, Ca, Mg, Na, S, Fe, Mn, Cu, B, Zn), total C and N, organic C, and P sorption index and soil texture properties (analysed using laser diffraction particle size anlyzer (LDPSA)) and prediction models developed using both MIRS and TXRF spectral techniques.

In chapter 4, this thesis has described how low cost high-throughput spectroscopy methods can be used both as a front line screening technique for development of calibration models and for the direct development of indicators of soil properties. Globally applicable calibrations to predict standard soil properties based on infrared spectra may increase the use of this low cost technique especially in developing countries of Sub-Saharan Africa (SSA) where reliable data on soil chemical and physical information needed to give advice on land management is sparse and dated. In the current study, a major step forward in the quantitative analysis based on mid-IR spectra was the use of RF statistics requiring the development of calibrations that related the first derivative of the spectral information to the reference data using the entire spectra as opposed to only a few wavelengths. Results also indicated that besides its better accuracy, RF was computationally fast and robust to analyse complex spectra of multi-component analytical methods (TXRF and MIRS). In addition, our results demonstrated the applicability of TXRF as a useful supplement to improve prediction of properties that were not well predicted by MIR, such as extractable nutrients Mehlich-3 S and exchangeable Na, exchangeable acidity (unbuffered KCl extraction), exchangeable sodium percentage, and exchangeable sodium ratio. We have proven that TXRF data can predict reference data directly and provide more predictive power when combined with the first derivative MIRS spectra. These techniques (TXRF and MIR) thus open up possibilities for using element profiling to improve global predictions of soil properties. Hence, the results of this study showed that spectral diagnostic approaches can be a valuable tool for qualitative and quantitative assessment of soils and a more reliable alternative to the conventional soil characterization methods. Spectroscopic techniques have a high potential and have shown promise in the present study as rapid and accurate methods of characterizing soil properties.

### 5.4 Implications for soil mapping in Africa and other parts of the world

In chapter 3, it has been concluded that topsoil chemical properties have the same diagnostic value as the subsoil but in combination with subsoil values they may provide extremely valuable insight in soil functional properties for both agricultural and environmental applications especially in the SSA context where there is deficiency in data. Therefore, the collection of subsoil samples, and the inclusion of their chemical composition in the analysis, is essential for in-depth soil chemical properties research, but care need to be taken to avoid spatial auto-correlation. Digital soil mapping is the creation of spatial soil information systems using field and laboratory methods coupled with spatial and non-spatial soil inference systems and it uses statistical models to predict soil functional properties and degradation prevalence at unobserved locations in the landscape (AfSIS, 2013). AfSIS is currently producing digital soil maps using legacy data e.g. from the existing databases, new legacy data collection as well as the first round of these maps, using the Africa Soil Profiles Database, now available (Leenaars, 2013). Thus production of new digital soil maps using the sentinel site soil data analysed in the current study and characterization of additional sentinel sites would further reduce the statistical uncertainties in the spatial models that will be developed under digital soil mapping activity for the AfSIS project. The new data available from the current study would help the digital soil map provide for example (i) information on a soil's properties; (ii) a geographical representation of soil constraints (such as element toxicity, carbon deficits) with known confidence, (iii) spatial targeting of management recommendations, and (iv) a baseline for change detection and impact assessment as outlined by AfSIS (AfSIS, 2013). Results from the current study could be mapped to visualize different soil characteristics derived from the spectral data and to generate digital soil maps that clearly reflect the general patterns of variability in SSA, including making extrapolations some areas that that are yet to be sampled.

## 5.5 Innovative aspects of the findings and recommendations for future research in soil science, environmental and agricultural applications using spectroscopy.

This section relates the innovative research approach of this thesis to general methodological considerations for soil analyses and it formulates recommendations for data to be considered and methods to be used, and provides an outlook on the future research needs that follow from it. This thesis has shown the added value of using MIR, XRD and TXRF and recent advances in multivariate computational statistics, and proved that these techniques have a great potential for improving the way in which soils are evaluated, while significantly reducing the costs to do so. These techniques have also shown promising results that present new opportunities to revolutionize the way in which agronomy and soil science is done, and thus there is now real possibility to further harness such methods to enable science-based prediction of soil properties for agricultural and environmental management from soil spectral data especially for SSA.

This thesis emphasizes and has built on the premise that science-based approaches to agricultural and environmental management such as the use of a land degradation diagnostic surveillance framework could accelerate economic development in developing countries (Shepherd and Walsh, 2007). The current study has supported ICRAF's methods development for the Africa Soils Information Service (AfSIS) Project (www.africasoils.net), which will over the next years develop a practical, timely, cost-effective, soil health surveillance service to map soil conditions, set a baseline for monitoring changes and provide options for improved soil management in SSA. Results for soil analysis using the all techniques in this study could form a consistent baseline for soil analysis in AfSIS project against which changes in the soil element composition can be monitored and evaluated over time as opposed to previously existing data on soils in the form of maps and soil profile data. In addition, there is every reason to believe that the current fundamental information on the element composition of SSA soils will immediately stimulate practical phases of soil management. In this study sentinel sites are 10 x 10 km blocks, within sampling strata were used for characterization of the soils, and were designed to provide accurate baseline data and monitoring of land health and factors affecting it. The future of Africa Soil Information Service is that, even though the sentinel site data reported here, which consists of a set of 100-km$^2$ samples of land within which soil were measured using a spatially stratified randomized sampling scheme, can be extended or extrapolated to provide information on soil constrains at continental, national or local scales in SSA.

### 5.6. References

1. Africasoils.net [Internet]. Africa Soil information Service (AfSIS); c2012–13 [cited 2013 Oct 25]. Available from: www.africasoils.net/.

2. Jenny H. 1941. Factors of soil formation: a system of quantitative pedology. McGraw-Hill, New York.

3. Knadel M, Viscarra Rossel RA, Deng F, Thomsen A, and Greve MH. 2013. Visible-near infrared spectra as a proxy for top-soil texture and glacial boundaries. Soil Science of America Journal 77: 568-579.
4. Leenaars JGB. 2013. Africa Soil Profiles Database, Version 1.1. A compilation of geo-referenced and standardized legacy soil profile data for Sub Saharan Africa (with dataset). ISRIC report 2013/03. Africa Soil Information Service (AfSIS) project. ISRIC – World Soil Information, Wageningen, the Netherlands.
5. Shepherd KD, Walsh MG. 2007. Infrared spectroscopy – enabling an evidence-based diagnostic surveillance approach to agricultural and environmental management in developing countries. Journal of Near Infrared Spectroscopy 15:1–20.
6. Voortman RL. 2011. Explorations into African land resource ecology on the chemistry between soils, plants and fertilizers. PhD Dissertation. Faculty of Economics and Business, Vrije Universiteit, pp. 264.

## 6.0 Summary

## Prediction of soil properties for agricultural and environmental applications from infrared and X-ray soil spectral properties

Many of today's most pressing problems facing developing countries, such as food security, climate change, and environmental protection, require large area data on soil functional capacity. Conventional assessments (methods and measurements) of soil capacity to perform specific agricultural and environmental functions are time consuming and expensive. In addition, repeatability, reproducibility and accuracy of conventional soil analytical data are major challenges. New, rapid methods to quantify soil properties are needed, especially in developing countries where reliable data on soil properties is sparse, and to take advantage of new opportunities for digital soil mapping. Mid infrared diffuse reflectance spectroscopy (MIR) has already shown promise as a rapid analytical tool and there are new opportunities to include other high-throughput techniques, such as total X-ray fluorescence (TXRF), and X-ray diffraction (XRD) spectroscopy. In this study TXRF and XRD were tested in conjunction with IR to provide powerful diagnostic capabilities for the direct prediction of key soil properties for agricultural and environmental applications especially for Sub-Saharan Africa (SSA) soils. Optimal combinations of spectral methods for use in pedotransfer functions for low cost, rapid prediction of chemical and physical properties of African soils as well as prediction models for soil organic carbon and soil fertility properties (soil extractable nutrients, pH and exchangeable acidity) were tested in this study. These state-of-the-art methods for large-area soil health measurement and monitoring will aid in accelerating economic development in developing sub-Saharan Africa countries with regards to climate change, increasing water scarcity and impacts on local and global food security as well as sustainable agricultural production and ecosystem resilience in the tropics.

This study has developed and tested a method for the use of TXRF for direct quantification of total element concentrations in soils using a TXRF (S2 PICOFOX$^{TM}$) spectrometer and demonstrated that TXRF could be used as a rapid screening tool for total element concentrations in soils assuming sufficient calibration measures are followed. The results of the current study have shown that TXRF can provide efficient chemical fingerprinting which could be further tested for inferring soil chemical and physical functional properties which is of interest in the African soil context for agricultural and environmental management at large scale. Further, this thesis has helped to improve understanding of the variation and patterns of element concentration data for 1034 soil samples from 34 stratified randomly-located 100-km$^2$ "sentinel" sites across SSA and explored the link between variability of soil properties and climate, parent material, vegetation types and land use patterns with the help of Random Forests statistics. Our results of total element concentration were within the range reported globally for soil Cr, Mn, Zn, Ni, V, Sr, and Y and in the high range for Al, Cu, Ta, Pb, and Ga. There were significant variations ($P < 0.05$) in total element composition within and between the sites for all the elements analysed. In addition, the greatest proportion of total variance and number of significant variance components occurred at the site (55-88%) followed by the cluster nested within site levels (10-40%). Our results also indicated that the strong observed within site as well as between site variations in many elements can serve to diagnose their soil fertility potential. Explorations of the relationships between element composition data and other site factors using "randomForest' statistics have demonstrated that all site and soil-forming factors have important influence on total elemental concentrations in the soil with the most important variables explaining the main patterns of variation in total element concentrations being cluster, topography, landuse, precipitation and temperature. However, the importance of cluster can be explained by spatial correlation at distances of <1 km.

This study has also analysed the potential of combining analyses undertaken using MIR spectroscopy and TXRF on 700 soil samples from 44 "sentinel" sites distributed across SSA. MIR prediction models for soil organic carbon, and other soil fertility properties (such as soil extractable nutrients, pH, exchangeable acidity and soil texture) were developed using Random Forests (RF) regression and the current study has added total element concentration data to the residuals of the MIRS predictions to test how they can improve the MIR prediction accuracies. The RF approach out-perfomed the conventional partial least squares regression (PLSR) on simultaneous determination of soil properties; and in addition, RF results were also easily interpretable, computationally much faster and did not rely on data transformations or any other assumptions about data distributions compared to PLSR. With respect to the potential of combining TXRF and MIR spectra, including total element concentration data from TXRF analysis in the RF models significantly reduced root mean square error of prediction by 63% for Ecd, 54% for Mehlich-3 S, and 53% for Mehlich-3 Na. Thus, TXRF spectra were a useful supplement to improve prediction of soil properties not well predicted by MIRS. The prediction improvement from including TXRF was due to detection of a few outliers that did not appear as

MIR spectral outliers. MIR showed remarkable ability to capture total elemental composition effects on physico-chemical soil properties but TXRF may have potential for outlier detection in large studies. This study has also helped to develop high-throughput spectral analytical methods and provided recommendations on optimal spectral analytical methods for the Globally Integrated Africa Soil Information Service (AfSIS) Project. Successfully developed methods in this study will become part of the standard AfSIS procedures.

## 7.0 Zusammenfassung

## Vorhersage physikalischer und chemischer Bodeneigenschaften für landwirtschaftliche und umwelttechnische Anwendungen mittels Infrarot und Röntgen-Spektral Methoden

Viele der heutigen dringendsten Problemfelder der Entwicklungsländer wie Gewährleistung der Ernährungssicherheit, Anpassung an Klimawandel und verbesserter Umweltschutz erfordern umfangreiche, flächendeckende Daten über die funktionelle Kapazität von Böden. Herkömmliche Verfahren (Methoden und Messungen) zur Bestimmung von spezifischen landwirtschaftlichen und ökologischen Bodenfunktionen sind zeitaufwendig und teuer. Neben den Kosten sind die Wiederholbarkeit, Reproduzierbarkeit und Genauigkeit von herkömmlichen analytischen Methoden große Herausforderungen. Neue, schnelle Methoden zur Quantifizierung von Bodeneigenschaften sind notwendig, vor allem in Entwicklungsländern, wo zuverlässige Daten über Bodenqualität schwer zu beschaffen sind, und um die Vorteile der neuen Möglichkeiten einer digitalen Bodenkartierung auszunutzen. Infrarot-Spektroskopie mit diffuser Reflexion (IR) hat bereits gute Ergebnisse als ein schnelles Analyse-Instrument gezeigt und es gibt neue Möglichkeiten, um andere Hochdurchsatz-Techniken wie die Total-Röntgenfluoreszenz (TXRF) und Röntgenbeugungs-Spektroskopie (XRD) einzusetzen. In dieser Studie wurden TXRF und XRD in Verbindung mit IR getestet, um leistungsstarke Diagnosefunktionen für die direkte Vorhersage der wichtigsten funktionellen Eigenschaften von Böden für Landwirtschaft und Umwelt-Anwendungen besonders für die Böden Afrikas südlich der Sahara zur Verfügung zu stellen. In dieser Studie wurden optimale Kombinationen von spektralen Methoden getestet, die für den Einsatz in Pedotransferfunktionen mit niedrigen Kosten, einer schnellen Vorhersage der chemischen und physikalen Eigenschaften der afrikanischen Böden, sowie in Prognosemodellen für organischen Kohlenstoff im Boden und die Bestimmung von Bodenfruchtbarkeitsparametern (extrahierbare Nährstoffe, pH-Wert und austauschbare Säuren) geeignet sind. Diese aktuellen Methoden zur großflächigen Messung und Überwachung der Bodengesundheit können dazu beitragen, die wirtschaftliche Entwicklung in den Ländern Afrikas südlich der Sahara positiv zu fördern, besonders in Bezug auf den Klimawandel, die lokale und globale Ernährungssicherheit sowie die Nachhaltigkeit der landwirtschaftlichen Produktion und der Stabilität der Ökosysteme.

In diese Studie wurde zunächst ein Verfahren zur Verwendung von TXRF zur direkten Quantifizierung der gesamten Elementkonzentration in 15 Bodenproben unter Verwendung eines

TXRF (S2 PICOFOX$^{TM}$) Spektrometers entwickelt und mit 20 weiteren Bodenproben getestet. Die Ergebnisse zeigten, dass bei ausreichender Kalibrierung TXRF als ein schnelles Screening-Werkzeug für die meisten Elemente verwendet werden kann. Die Ergebnisse der aktuellen Studie haben ausserdem gezeigt, dass TXRF effiziente chemische Fingerabdrücke liefern kann, die zum Ableiten von chemischen und physikalischen Bodeneigenschaften dienen können.

Diese Arbeit hat weiter dazu beigetragen, den Zusammenhang zwischen Variabilität der Bodeneigenschaften und Klima, Bodenausgangsmaterial, Vegetationstypen und Landnutzung mit Hilfe von TXRF, XRD und IR-spektralen Methoden zu verstehen. Dafür wurden 1034 Bodenproben analysiert, die im Rahmen des ‘Africa Soil Information Service‘ (AfSIS) Projektes von 34 randomisiert ausgewählten stratifizierten Standorten von jeweils 100 km$^2$ in zahlreichen Länders Afrikas südlich der Sahara entnommen wurden. Die Ergebnisse der Gesamt-Elementkonzentrationen dieser Bodenproben lagen im Bereich der dokumentierten Konzentrationen für die Elemente Cr, Mn, Zn, Ni, V, Sr und Y, lagen aber höher als gewöhnlich für die Elemente Al, Cu, Ta, Pb, and Ga. Signifikante Unterschiede ($P < 0{,}05$) der Gesamt-Elementkonzentrationen wurden sowohl innerhalb als auch zwischen den beprobten 34 Standorten gefunden. Die Variabilität war deutlich grösser zwischen den 34 Standorten (55-88 % Varianz) als innerhalb der Standorte (10-40 % Varianz). Mit Hilfe von ‘Random Forests‘-Regressionen konnte gezeigt werden, dass die Gesamt- Elementkonzentrationen der untersuchten Bodenproben von umweltbezogenen Standortvariablen wie Topographie und Landnutzungstyp als auch Klimafaktoren wie Temperatur und Niederschlag wesentlich beeinflusst werden.

In einem weiteren Schritt wurde die Aussagekraft einer Kombination von MIR und TXRF-Methoden und der ‘Random Forests‘-Regression anhand von 700 Bodenproben von 44 Standorten in Afrika südlich der Sahara getestet. Dazu wurden zunächst MIR-Vorhersagemodelle für organischen Bodenkohlenstoff und andere Bodenfruchtbarkeitsparameter (extrahierbare Nährstoffe, pH-Wert und austauschbare Säuren) mit Hilfe von ‘Random Forests‘ (RF)-Regressionen entwickelt. Durch Einbringen der Gesamtelement-Daten zu den Residuen der IR-Vorhersagen konnten die MIR-Regressionsmodelle signifikant verbessert werden. Im Vergleich zu der gewöhnlich benutzten ‘partial least square‘-Regression (PLSR) zeigte die entwickelte RF-Regression deutlich bessere Ergebnisse, war schneller anzuwenden und einfacher zu interpretieren und war nicht auf zeitaufwändige und fehleranfällige Datentransformationen wie die PLSR angewiesen. Durch die Kombination von TXRF- und MIR-Spektren konnte ausserdem die Vorhersage-Genauigkeit der Bodenparameter deutlich verbessert werden, z.B. für Ecd um 63%, Mehlich-3 S um 54%, Mehlich-3 Na um 53% verglichen zur alleinigen Nutzung der MIRS-Spektren.

Zusammenfassend hat die vorliegende Studie dazu beigetragen, neue spektrale Bodenanalysemethoden mit hohem Durchsatz zu entwickeln und Empfehlungen für die optimierte Anwendung dieser Methoden zu erarbeiten, die bereits erfolgreich von dem oben erwähnten AfSIS-Projekt übernommen und in die Standard-AfSIS Verfahren integriert worden sind.

# Acknowledgements

Now thanks be to God, Who in Christ always leads me in triumph (2 Corinthians 2:14), and I am assured and know that [God being a partner in my labor] all things work together and are [fitting into a plan] for good to and for those who love God and are called according to [His] design and purpose (Romans 8:28). I am particularly grateful to the Lord for all He helped me achieve during the current study. But then looking at the next phase of research, I am even more thrilled and thankful to God for the glory, victories and successes He's already planned and positioned for me throughout my career and future. For His divine power has bestowed upon us all things that [are requisite and suited] to life and godliness, through the [full, personal] knowledge of Him who chose and called me by and to His own glory and excellence (2 Peter 1:3). I thus look forward with faith, courage and excitement to a glorious and fruitful further studies, because of making books (publications) there is no end (Ecclesiastes 12:12).

During the course of the research work for this PhD study I received notable support by the World Agroforestry Centre (ICRAF) through the Africa Soil Information Service (AfSIS) project funded by the Bill &Melinda Gates Foundation, and by the Dr. Hermann Eiselen Ph.D. Grant from the Foundation Fiat Panis. I would also like to acknowledge the Young Excellence Scholars (YES) PhD Scholarship from the Food Security Centre at the University of Hohenheim, Germany, travel grants as well as a stipend within the framework of the "Global Food Security" funded by the Deutscher Akademischer Austausch Dienst (DAAD) under the EXCEED program.

I enormously learnt and flourished at all stages of my research under the guidance of two renowned scientists as my supervisors; Prof. Dr. Georg Cadisch and Dr. Keith D. Shepherd. It is impossible to acknowledge adequately and specifically all the assistance, supervision and encouragement rendered by them. I tender to both of you a sincere 'Thank You' for improving the manuscripts and thesis logically and technically.

It is a pleasure to acknowledge the cooperation and contribution of the personnel of the ICRAF Soil Spectral Diagnostics Laboratory, who assisted in the sample preparation work and furnished spectral data analysis and interpretation from the different spectrometers. I would like to thank Mr. Andrew Sila for his help with the statistical analysis of complex spectral data. I wish to acknowledge the contribution of Mr. Robin Chacha and Mr. Hezekiah Nyandika especially their assistance with TXRF analyzing the soil samples, and Mrs. Mercy Nyambura, Mrs. Beatrice Oware and Miss. Bella Kauma for assistance with XRD data analysis. All the

ICRAF Soil Spectral Diagnostics Laboratory technical staff are acknowledged for their support with sample processing for spectral analysis. Mr. Samuel Gaturu and Mrs. Linda Kidweye of the ICRAF Soil Spectral Diagnostics Laboratory are also acknowledged for their administrative assistance.

I gratefully appreciate the AfSIS field crew, especially to Dr. Jerome Tondoh, Dr. Leigh Winowieki, and Dr. Desta Lulsege, who provided the soil samples, and other persons involved in the project for providing the data used in this study in particular Dr. Tor-Gunnar Vågen. I am also very gratefully to Dr. Ric Coe, Dr. Anja Gassner, Prof. Dr. Hans-Peter Piepho, Dr. Joseph Ogutu, Dr. Hagen Stosnach, Dr. Michael Gatari, Dr. Ermias Aynekulu, Dr. Gudrun Keding, Mr. Juan Laso, Mr. Kefyalew Sahle, Dr. Katja Kehlenbeck and Dr. Todd Rosenstock for their valuable advice and help with data analysis, interpretations, and reviewing the manuscripts.

The administrative assistance of the ICRAF Training Unit staff, as well as at Uni-Hohenheim Institute 380a in Stuttgart Germany is highly appreciated. Special thanks to Mrs. Gabriele Kircher for her support in a number of administrative issues related to the university and arranging office and accommodation in Stuttgart. I would also like to acknowledge the administrative support from Dr. Brigitte Kranz, Mrs. Hélène Stauss, and all the current and former members of staff of FSC for administrative support with the university, visa application as well as for facilitating a conformable stay in Stuttgart, Germany and travel for workshops in Germany and field research in Kenya. Many thanks go to Dr. Detlef Virchow for his continued support as the Executive Manager of FSC, as well as our continued formal and informal learning opportunities he organized.

I would also like to thank all my office mates and colleagues both at ICRAF and Institute 380a at the University of Hohenheim who have given much time to long and profitable discussions and interactions.

Finally, I wish to extend my profound appreciation and deepest gratitude to my beloved wife Mrs. Agnes Towett, my sons Ian Kiprop and Karsten Kipkemoi and my entire family whose endless love, encouragement and tremendous sacrifice have inspired and encouraged me to accomplish this PhD study.

## Appendix: Additional publications

(related to this thesis and published during the same time frame of doctoral studies)

### 1. Extended Synopsis submitted and accepted for oral presentation at the FAO/IAEA International Symposium on Managing Soils for Food Security and Climate Change Adaptation and Mitigation, to be held in Vienna, Austria from 23 to 27 July 2012.

The number assigned to the paper is IAEA-CN-191-108

# The Potential of Combining Infrared and Total X-Ray Fluorescence Spectroscopy for the Prediction of Soil Functional Properties

E.K. Towett[a], K.D. Shepherd[b], G. Cadisch[a], E. B. Eynekulu[b], A. Sila[b]

[a] Institute of Plant Production and Agroecology in the Tropics and Subtropics, University of Hohenheim, Garbenstr. 13, 70599 Stuttgart, Germany.

[b] World Agroforestry Centre (ICRAF), P.O. Box 30677-00100 GPO, Nairobi, Kenya.

E-mail address of main author: etowett@gmail.com or e.towett@uni-hohenheim.de

## Synopsis

Conventional soil assessments are expensive and dense sampling is often required to adequately characterize spatial variability in an area [1]. New, rapid methods to quantify soil properties are needed to support soil health surveillance systems. Infrared Spectroscopy (IR) has shown promise as a simple and non-destructive analytical method that is now routinely used to predict several soil properties simultaneously, including soil organic carbon, nutrient retention capacity, and water holding capacity [2, 3]. The ability to rapidly characterize large numbers of samples with IR opens up new possibilities for soil evaluations that consider uncertainty in predictions and interpretations of soil properties. However, IR has some limitations in that it cannot predict extractable P and K well, which in addition to N are often the main limiting nutrients in African soils. Prediction of particle size distribution with IR is variable and in addition, calibrations have to be adjusted for different soil types. A new spectral technique using Total X-ray Fluorescence (TXRF) could be a valuable tool to supplement IR and stabilize calibrations. TXRF provides for rapid and simultaneous determination of the concentrations of most elements from sodium to Uranium with minimal sample preparation time [4]. Advantages of the technique compared to conventional methods include minimal sample preparation, and low matrix interference, removing the need for external calibration. The total element

concentration spectra can be used to capture key mineralogical differences in soils and there are possibilities to correlate extractable nutrient analysis with total element analysis and also to measure element concentrations in soil extracts. Thus TXRF could provide a powerful complement to IR, especially for predicting nutrient supply capacity, which is most important when considering soils as the substrates for plant growth. The technique opens up possibilities for using total element profiling to improve global predictions of soil functional properties, such as soil organic carbon, cation exchange capacity, extractable nutrients, P sorption, water holding capacity, and soil stability.

Optimal combinations of IR and TXRF techniques for prediction of soil functional properties have yet to be evaluated. The objectives of this study were to quantify the variability and patterns in soil mineralogy, total element composition with TXRF and mid infrared (MIR) spectra of soils from a random sample set of Sub-Saharan Africa soils. Georeferenced samples associated with the Africa Soil Information Service (AfSIS) (www.africasoils.net), taken from a set of sentinel sites randomized over Sub-Saharan Africa were used for characterization. A total of 348 soil samples from eleven 100-km$^2$ sentinel sites across Sub-Saharan Africa: Tanzania (6 sites), Congo (2 sites), Mali (2 sites), Burkina Faso (1 site) were used in exploring spectral patterns. Paired topsoil and subsoil samples taken from 32 randomized sample points at each site were analysed. Soils were air-dried and passed through a 2-mm sieve for IR analysis, but were further ball-milled to less than 75 µm for TXRF and MIR analyses. MIR (2.5-25 µm) wavelength regions respond to a number of important soil properties including mineral composition, iron oxides, water (hydration, hygroscopic, free), carbonates, soluble salts, and particle size distribution [3]. Fine ground samples were analysed with MIR (Bruker, Tensor 27 MIR spectrometer) using a robotic high-throughput system employing micro-titre plate. Spectral TXRF analyses were done using a Bruker S2 PICOFOX TXRF instrument (Bruker AXS, Germany). Standardisation was internal and only required addition of an element that was not present in the sample (Se) for quantification purposes. Samples were suspended in detergent (Triton X-100), spiked with a known quantity of Se as an internal standard, pipetted onto carriers, and dried. We explored the within and between site patterns of variation in total element composition using scatter plots and principal component analysis in R statistical software. Quantitative analysis based on mid-IR spectra required the development of calibrations that related the spectral information to total element concentration using the entire spectra as opposed to only a few wavelengths, utilizing partial least squares (PLS) and principal component analysis (PCA). This was a precursor to exploring relationships with directly measured soil properties.

MIR spectral data related well to those of TXRF, and most of the elements detected using TXRF could be predicted well from MIR using principal component scores extracted from the TXRF data. Total elements Na, Al, Cr, S, Sc, and Zn gave good calibration models ($R^2 > 0.5$, RMSE < 0.6) in less than 8 principal components while total elements Mg, P, K, Ca, Ti, V, Mn, Fe, Ni and Ga had relatively good calibration models ($R^2 > 0.5$, RMSE < 0.5) but with more than eight principal components whereas Cl and Co gave poor calibration models (Fig. 1). PCA of TXRF data (Fig. 2) revealed that patterns in total element concentrations between sites appeared to relate to differences in mineralogical ‘functional groups’. The pattern of clustering of the individual minerals and sorting of heavy minerals along the positive Dim1 axis was apparent (Fig. 2). Al, K, Ga, Li, Sc and V have been reported to have especially strong correlations in soils caused by their mutual occurrence in clay minerals [5]. The clay factor commonly includes the elements As, Cu, Ni and Cr, all elements that are common in trace amounts in clay minerals or adsorbed to them [5]. Here, the clay minerals in the scatter plot are represented by Al and Cr and they lied along the positive Dim1 axis (Fig. 2). However, As, Cu and Fe were not included in the clay mineral cluster. Elements typical of feldspars and carbonates minerals (Ca, K, Sr and Ba) lied in the negative Dim1 and negative Dim2 axes. TXRF thus provides chemical fingerprinting and ‘functional’ mineral groupings that relate to potential soil nutrient supply capacity and other properties. Patterns in total element concentrations within and between sites appeared to relate to differences in mineralogical ‘functional groups’. It is thus worth testing whether TXRF in conjunction with MIR can be used as a complementary input to pedotransfer functions for low cost, rapid prediction of soil functional properties. The next step of this research will test combined use of MIR and TXRF for prediction of soil functional properties, and whether MIR spectra can predict clustering of samples in relation to total element profiles.

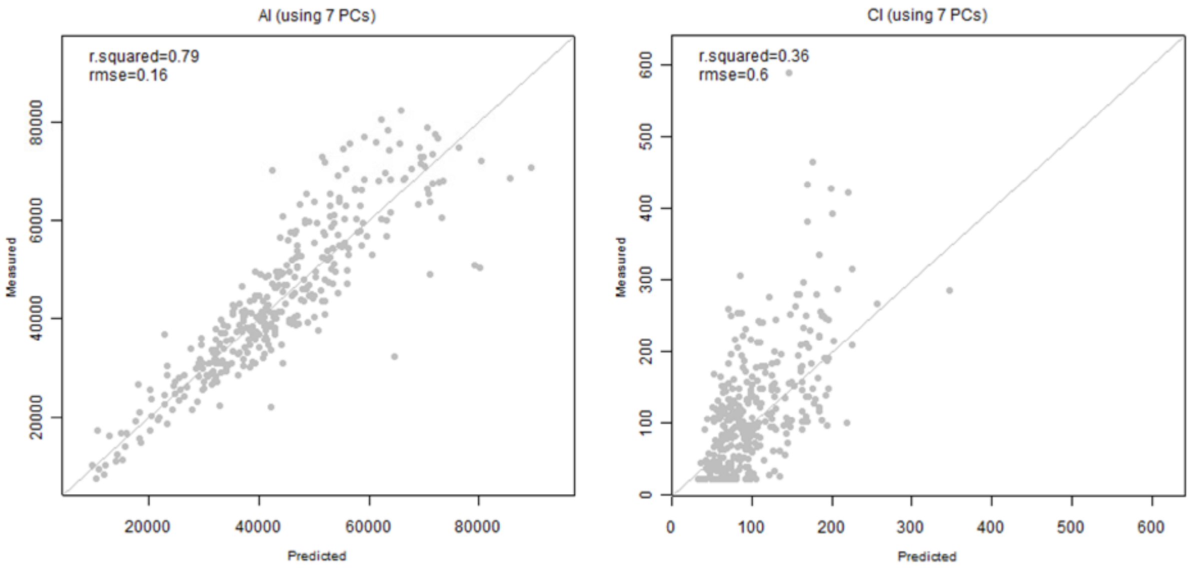


*Fig. 1. Scatter plots of the PLS regression models developed for elements Al and Cl using MIR (predicted and TXRF (measured) spectral data.*

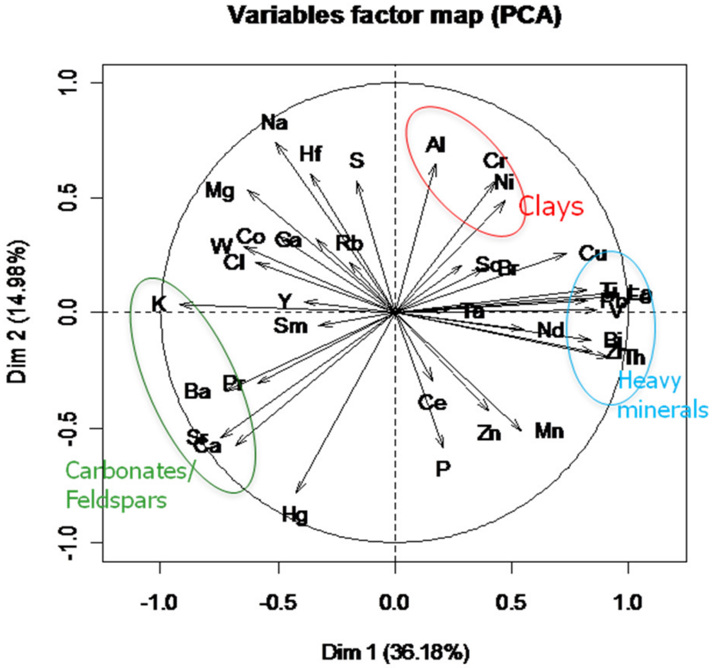


*Fig. 2. Biplot of principal components 1 and 2 for the soil total element concentration values with complete data matrix after computation of the 25th percentile of lower limit of detection (LLD) values for each element to replace the "non-detectable" values in the data set.*

## References

[1] Shepherd KD and Walsh MG. 2002. Development of reflectance spectral libraries for characterization of soil properties. Soil Science Society of America Journal 66:988-998.

[2] Brown DJ, Shepherd KD, Walsh MG, Mays DM, and Reinsch TG. 2006. Global soil characterization with VNIR diffuse reflectance spectroscopy. Geoderma 132: 273–290.

[3] Shepherd KD and Walsh MG. 2007. Infrared spectroscopy – enabling an evidence-based diagnostic surveillance approach to agricultural and environmental management in developing countries. J. Near Infrared Spectroscopy 15:1–20.

[4] Stosnach H. 2005. Environmental trace-element analysis using a benchtop total reflection X-ray fluorescence spectrometer. The Japan Society of Analytical Chemistry. Analytical Sciences 21:873-876.

[5] Cannon W.F. , & Horton J.D. 2009. Soil geochemical signature of urbanization and industrialization – Chicago, Illinois, USA. Applied Geochemistry 24 :1590–1601.

## 2. Abstract submitted and accepted for a poster presentation at the Tropentag, October 5-7, 2011, Bonn; "Development on the margin".

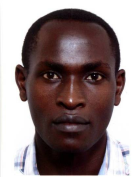
# Quantification of the Variability and Pattern in Total Element Composition of Sub-Saharan Africa Soils

Erick Kibet Towett[1], Keith Shepherd[2], Georg Cadisch[3]
[1]University of Hohenheim, Dept. of Plant Production and Agroecology in the Tropics and Subtropics, Germany
[2]World Agroforestry Centre (ICRAF), Research, Kenya
[3]University of Hohenheim, Dept. of Plant Production and Agroecology in the Tropics and Subtropics, Germany

## Abstract

Measuring total element concentration of soils using conventional methods is time-consuming. Total X-ray Fluorescence Spectroscopy (TXRF) provides for rapid and simultaneous determination of the concentrations of most elements from sodium to Uranium with minimal sample preparation time. The technique opens up the possibilities for using total element profiling to improve global predictions of soil functional properties, such as soil organic carbon, cation exchange capacity, extractable nutrients, P sorption, water holding capacity, and soil stability. In this paper we present our investigations of the quantification of the variability and patterns in total element composition of soils from eight 100-km$^2$ sites across Sub-Saharan Africa: Tanzania (3 sites), Congo (2 sites), Mali (2 sites), Burkina Faso (1 site). Paired topsoil and subsoil samples taken from 32 randomised sample points at each site were analysed. We explored the within and between site patterns of variation in total element composition and their relationships with directly measured soil functional properties and TXRF soil spectral properties using scatter plots, principal component analysis, and classification and regression trees in R statistical software. The results indicate that TXRF provides chemical fingerprinting that relates to potential soil nutrient supply capacity. There were also relative variations in total element composition within and between the sites analysed. Thus TXRF can be used as a complementary input to pedotransfer functions for low cost, rapid prediction of soil functional properties. TXRF could also provide improved capabilities, for improving advisory services on soil constraints to plant growth with subsequent benefit to food security and human health.

# Curriculum vitae

| | |
|---|---|
| **Curriculum Vitae** |  |
| **Personal information** | |
| Surname(s) / First name(s) | **Towett Erick Kibet** |
| Current Address | World Agroforestry Centre, United Nations Avenue, Gigiri, 30677-00100, Nairobi, Kenya |
| Telephone(s) | + 254 20 7224412 (office) Mobile: +254 715034333 |
| Alternative Address | University of Hohenheim, Institute 380a, Garbenstr. 13, 70599 Stuttgart, Germany |
| Telephone(s) | + 49 71145923473 (office) Mobile: +49 17670362897 |
| E-mail | etowett@gmail.com (private), kibettowett@yahoo.com (private), e.towett@uni-hohenheim.de (professional), e.towett@cgiar.org (professional) |
| Nationality | Kenyan |
| Date of birth | 03.10.1981 |
| Gender | Male |
| Marital Status | Married |
| **Desired employment / Occupational field** | **Soil-Plant Spectral Diagnostics Scientist** |
| **Education and training** | |
| Dates | From October 2010 – December 2013 |
| Title of qualification awarded | PhD (Dr.Sc.agr.) –date of oral defence 9th December 2013. |
| Name and type of organisation providing education and training | University of Hohenheim, 70599, Stuttgart, Germany |
| Principal subjects/occupational skills covered | I have also covered the following individual subjects: Methods of scientific working; Tropical soils and land evaluation; Interdisciplinary Aspects of Food Security; Crop production affecting hydrological cycle as well as a series of seminars organized by the Food Security Centre (FSC) – University of Hohenheim some jointly with Center for Development Research (ZEF) - University of Bonn. |
| Research title | Prediction of soil properties for agricultural and environmental applications from infrared and x-ray soil spectral properties. |
| Objectives of research | The overall objective of the research is to test different infrared and x-ray spectral diagnostic techniques for the direct prediction of key soil properties for agricultural and environmental applications. |
| Specific objectives of research | The specific objectives of the study are to:<br>(i) develop and test an improved analytical method for the direct quantification of total element concentrations in soils using the S2 PICOFOX$^{TM}$ TXRF spectrometer,<br>(ii) test the ability of the total X-ray fluorescence spectroscopy (TXRF) to calibrate/validate against total acid dissolution ICP-MS analysis (international standard method) for a range of elements using a determined instrument sensitivity curve,<br>(iii) test the ability of the total X-ray fluorescence spectroscopy (TXRF) to calibrate/validate against total acid dissolution ICP-MS analysis (international standard method) for a range of elements using a determined instrument sensitivity curve,<br>(iv) quantify the variability in total element composition of soils from a diverse set of sub-Saharan Africa Soils, |

| | |
|---|---|
| Specific objectives of research contd. | (v) explore the patterns in total element composition of soils analysed and the relationships between total element concentrations and nutrient supply capacity by relating Mehlich-3 soil tests (acid-extractable nutrients) to total element analysis patterns in soil,<br>(vi) examine relationships between element fingerprints and site characteristics including mineralogy, climate, landform, vegetation type, plant material and management (cultivation),<br>(vii) test whether TXRF can improve on mid infrared spectroscopy (MIR) predictions of soil test values, especially for those variables for which MIR tends to give poor predictions (e.g. extractable P and K; some micronutrients),<br>(viii) test whether TXRF analysis of water extractable elements could replace acid extraction coupled with the ICP-MS analysis, and<br>(ix) examine the extent to which the TXRF technique added value to the MIRS technique for improving global predictions of soil total and extractable nutrients. |
| Analytical and spectral methods used | For most of the statistical analyses of reference (chemical) and spectral data, extensive use of R statistical package has been employed, utilizing R packages such as “soil.spec”, “ggplot2”, “gtools”, “latticist”, “Kernsmooth”, “FactoMineR”, “Princomp”, “som”, “Hmisc”, “fields”, “rpanel”, “reshape”, “rattle”, ‘DiagnosisMed”, “gregmisc”, “randomForests” (RF), and “lme4” (REML). I have also used KNIME (Konstanz Data Miner), SAS, SPECTRA and OPUS software for analysis of the spectral data, and GRASS and QGIS for analysis of GIS data. |
| Research outputs | I have one paper published on: “Quantification of total element concentrations in soils using total X-ray fluorescence spectroscopy (TXRF)” (Science of the Total Environment Journal).<br>Another paper is under review on: “Variability and pattern in total element composition of sub-Saharan Africa soils using total X-ray fluorescence spectroscopy” (Submitted to the Geoderma Journal, June 2013).<br>One manuscript is also submitted which analyse: “Combining mid infrared and total X-ray fluorescence spectroscopy for predictinf soil properties” (submitted to the Soil Science Society of America Journal, Dec 2013). |
| Level in national or international classification | Grade 1.0 (Very good) |
| Dates | From October 2006 – to October 2008 |
| Title of qualification awarded | Master of Science (M.Sc.) |
| Name and type of organisation providing education and training | Georg-August University of Göttingen, 37075, Göttingen, Germany |
| Principal subjects/occupational skills covered | Study Program was “Tropical and International Agriculture” and the examination was in the subject area “Tropical Agriculture”.<br>I have also covered the following individual subjects: Crop production systems in the Tropics; Tropical animal husbandry systems; Socioeconomics of rural development; Remote sensing image processing with open source software; Propagation techniques and ecophysiology in the tropics; Agrobiodiversity and plant genetic resources in the tropics; pests and diseases of tropical crops; Quality and processing of tropical plant products; Monitoring and evaluation of rural development policies and projects; Quantitative research methods in rural development economics as well as several colloquia on research.<br>My thesis was titled “Optimizing the use of Near Infrared Reflectance Spectroscopy (NIRS) to predict nutritional quality in cowpea (*Vigna unguiculata*) leaves for human consumption”.<br>I have one paper published on “Applicability of near infrared reflectance spectroscopy (NIRS) for determination of crude protein content in cowpea (*Vigna unguiculata*) leaves” with the Food Science & Nutrition journal in 2013. |
| Level in national or international classification | Grade 1.4 (Very good). |
| Dates | From October 2000 – to July 2004 |
| Title of qualification awarded | Bachelor of Science (B.Sc.) |
| Principal subjects/occupational skills covered | Chemistry and zoology |
| Name and type of organisation providing education and training | University of Nairobi, P.O. Box 30197-00100, Nairobi Kenya |

| | |
|---|---|
| Level in national or international classification | Second class honours, upper division (equivalent to grade 2.0) |
| **Work experience** | |
| Dates | July 2013 – To date |
| Occupation or position held | Research Analyst |
| Main activities and responsibilities | 1. Assist in preparing a manuscript on spectral band interpretations towards calibration free soil spectral indicators using Random Forests.<br>a. Literature research on spectral band assignments<br>b. Statistical calibration of soil properties to mid-infrared spectra using Random Forests in R<br>c. Preparation of graphics using ggplot2 in R and tables.<br>d. Contribution to writing of methods section and interpretation of results.<br>2. Develop soil calibrations for a Bruker Handheld XRF spectrometer.<br>3. Assist with preparation and submission of published data to Dataverse repositories.<br>4. Contribute to preparation of training materials and teaching at ICRAF's Soil Spectroscopy Training Course for Scientists and Technicians.<br>5. Assist with analysis and interpretation of AfSIS soil reference and spectral data.<br>6. Assist with other tasks related to furthering the work of ICRAF's Soil-Plant Spectral Diagnostics Laboratory as requested. |
| Name and address of employer | World Agroforestry Centre (ICRAF), UN Avenue, Gigiri, P. O. Box 30677 – 00100, Nairobi, Kenya. |
| Type of business or sector | Research |
| Dates | Jan 2009 – May 2009 |
| Occupation or position held | Intern and Research Fellow |
| Main activities and responsibilities | - Spectroscopic analysis of soil and plant materials<br>- Developing near infrared reflectance spectroscopy (NIRS) calibrations for analysis of organic resource quality. |
| Name and address of employer | World Agroforestry Centre (ICRAF), UN Avenue, Gigiri, P. O. Box 30677 – 00100, Nairobi, Kenya. |
| Type of business or sector | Research |
| Dates | Sep 2005 – Sep 2006 |
| Occupation or position held | Consultant |
| Main activities and responsibilities | - Liaising with various Bioversity consultants and partners in the 'African Leafy Vegetable' (ALV) and 'Dietary Diversity' projects to fill in gaps, ensure correct data entry to computer database, and quality control of output. Major documents produced: 1. ALV recipe book for Sub-Saharan Africa (SSA); 2. Leafy vegetables picture book for Tanzania; 3. Foods of the Nairobi People; 4. Recipes common for Kenyan Foods.<br>- Organizing, editing and gathering additional data and information from partners for bringing several books, booklets and calendars on vegetables to publishable standards together with IPGRI scientists.<br>- Organizing the final workshop of the 'Dietary Diversity' Project scheduled for end of October, 2006. |
| Name and address of employer | Bioversity International [formerly, the International Plant Genetic Resources Institute, sub-Saharan Africa Office (IPGRI-SSA)], c/o World Agroforestry Centre (ICRAF), UN Avenue, Gigiri, P. O. Box 30677 – 00100, Nairobi, Kenya. |

| | |
|---|---|
| Dates | Jun 2005 – Aug 2005 |
| Occupation or position held | Volunteer |
| Main activities and responsibilities | - Gather, document and analyse information on indigenous knowledge on plant uses, production and consumption patterns;<br>- Scientific input into and management of the web-based database SEPASAL (Survey of Economic Plants of Arid and Semi-arid lands); and<br>- Writing public awareness material, e.g. posters, booklets, leaflets. |
| Name and address of employer | National Museums of Kenya – Kenya Resource Centre for Indigenous Knowledge (KENRIK), P.O. Box 40658-00100, Nairobi, Kenya. |
| Dates | Jul 2003 - Jul 2004 |
| Occupation or position held | Research Assistant |
| Main activities and responsibilities | - Steam distillation of essential oil from medicinal plant material.<br>- Application of Gas Chromatography (GC) and Mass Spectroscopy (GC-MS), thin layer chromatography (TLC) and high performance liquid chromatography (HPLC) to crude plant extracts and volatiles/pheromones from insects.<br>- Report writing on anti-malarial effect of *Turraea nilotica* root-bark extracts.<br>- General laboratory management. |
| Name and address of employer | International Centre for Insect Physiology and Ecology (ICIPE), P.O. Box 30772-00506, Nairobi, Kenya. |
| Dates | Jul 2002 – Oct 2002 |
| Occupation or position held | Research Assistant |
| Main activities and responsibilities | - Comprehensive surveys of sites of cultural and biological importance in Central and North Rift regions of Kenya and traditional farming systems in West Pokot district of Northern Kenya;<br>- Compiling reports and writing public awareness material, e.g. posters, booklets, leaflets. |
| Name and address of employer | National Museums of Kenya – Kenya Resource Centre for Indigenous Knowledge (KENRIK), P.O. Box 40658-00100, Nairobi, Kenya. |
| Type of business or sector | Research |

**Personal skills and competences**

Language(s)

| Self-assessment | **Understanding** | | | | **Speaking** | | | | **Writing** | |
|---|---|---|---|---|---|---|---|---|---|---|
| *European level (*)* | Listening | | Reading | | Spoken interaction | | Spoken production | | | |
| **English** | C2 | Proficient user | C2 | Proficient user | C2 | Proficient user | C2 | Proficient user | C2 | Proficient user |
| **Swahili** | C2 | Proficient user | C2 | Proficient user | C2 | Proficient user | C2 | Proficient user | C2 | Proficient user |
| **German** | B1 | Independent user | B2 | Independent user | B2 | Independent user | B2 | Independent user | B2 | Independent user |

| | |
|---|---|
| Social skills and competences | o I able and willing to work well in an interdisciplinary, multicultural environment with individuals from different scientific, societal and cultural backgrounds and with organizations from the public, private and civil society sector, both national and international.<br>o I have good leadership, management and communication skills in multicultural settings.<br>o Currently, I am providing teaching assistance to students on lab equipment, chemometrics, spectral databases and statistical analysis of spectral data, including giving presentations at training courses.<br>o During my M.Sc. studies in Goettingen, Germany, I was a member of the City's Athletics club LGGoettingen and currently during my PhD studies in Stuttgart, I am a semi-professional runner for a local running team. |
| Organisational skills and competences | o Gave a presentation of Africa Soil Information Service (AfSIS) Soil Spectral Diagnostics and Characterization during an International Training Program on the Use of Decision Support Tools for Fertilizer Recommendations and Climate Variability Assessment held at the International Center for soil Fertility and Agricultural Development (IFDC), Nairobi, Kenya March 8-19, 2010.<br>o During the 2nd World Congress of Agroforestry (23-28 Aug 2009, Nairobi, Kenya), I assisted in a study-tour of the ICRAF Soil-Plant Spectral Diagnostics Laboratory as a Congress side event.<br>o During my Internship at ICRAF, gave 2 seminars presentations on my previous research work on the use of NIRS for quality analysis (one at Bioversity International's sub-Sahara Office in Nairobi and another at the University of Nairobi, Dept. of Nuclear Science and Technology).<br>o During my M.Sc. I organised a seminar series on my research work "Optimising the use of Near Infrared Reflectance Spectroscopy (NIR) to Predict Nutritional Quality of Cowpea Leaves"; three were at the University of Goettingen, Germany, and one at the World Vegetable Centre-Regional Centre for Africa in Arusha, Tanzania. |
| Technical skills and competences | o I have good working knowledge, advanced skills, hands-on experience and training in proximal (spectral) soil sensing methods such as Near and Middle Infrared Reflectance Spectroscopy as well as modern soil analytical techniques including x-ray and laser diffraction technologies for soil and plant analysis at ICRAF.<br>o I have PhD-level experience in soil-plant spectral methods and a sound knowledge of tropical soil science and land management.<br>o I also have some strong statistical analytical skills in multivariate analysis of soil/plant data as well as database management.<br>o I have assisted ICRAF's Soil-Plant spectral Diagnostics Lab with development of TXRF soil analytical method, review of spectral data; performed comparisons of TXRF and reference data as well as correction of the ICRAF TXRF results for the "1st round robin for TXRF in environmental & biological analysis" (organized in the frame of VAMAS TWA2) which were submitted for the report "Interlaboratory comparison of total-reflection X-ray fluorescence (TXRF) spectroscopy for environmental analysis", Brescia (Italy), 27th March 2012.<br>o Contributed to reviewing of 2 submitted scientific manuscripts to Geoderma and Applied Geochemistry journals.<br>o Participated in preparation of funding proposals jointly with ICRAF SD4, the University of Nairobi and the University of Manitoba which was submitted to the Canadian research fund, IDRC, to support the spectral diagnostics research.<br>o Attended a ZEF-FSC PhD Workshop "Working Within Political Contexts: Strategies and methods for implementation-oriented research" held at: Hostel am Flußbad Der Cöpenicker e. V. Gartenstraße 46-48,12557, Berlin,Germany, from 12-17 March 2013.<br>o Attended a Near Infrared Reflectance Spectroscopy training course "Getting the best out of light" held during the International Conference on NIRS Onsite at the Federal Agricultural Research Centre (FAL), Braunschweig, Germany, 12-13 June 2007. |
| Computer skills and competences | Competent with most Microsoft Office programmes, Internet and email, Photoshop, OpenOffice suite, and some experience in statistical packages (R, KNIME, SPECTRA, SAS, SPSS, Genstat, WinISI, OPUS, Unscrambler, QGIS, GRASS). |
| Artistic skills and competences | During my work at ICRAF, I developed 7 posters highlighting new soil-plant spectral diagnostic techniques and their application namely Near Infrared Reflectance Spectroscopy (NIR), Mid Infrared Reflectance Spectroscopy (MIR), Visible Near Infrared Reflectance Spectroscopy (VisNIR), Bench-top X-ray Diffraction Spectroscopy, and Laser Diffraction Particle Size Analysis |

| | |
|---|---|
| Other skills and competences | Ability to complete tasks in a timely manner and to a high standard. |
| Driving licence | Kenyan driving licence - category B, C and E vehicles |
| Extracurricular Interests and Activities: | I am a semi-professional middle & long distance runner and love to travel - interest to learn about other cultures and countries. |

## PUBLICATIONS

***Journal Articles***

1. Towett EK, Shepherd KD and Cadisch G. Quantification of total element concentrations in soils using total X-ray fluorescence spectroscopy (TXRF). Science of the Total Environment 463–464 (2013) 374–388.
2. Towett EK, Merle A, Shepherd KD, Polreich S, Eynekulu E, and Maass BL. Applicability of near infrared reflectance spectroscopy (NIRS) for determination of crude protein content in cowpea (Vigna unguiculata) leaves. Food Science & Nutrition 2013; 1(1): 45–53. DOI: 10.1002/fsn3.7
3. Towett EK, Shepherd KD, Tondoh JE, Winowieki LA, Lulseged D, Nyambura M, Sila A and Cadisch G. Variability and pattern in total element composition of Africa soils using total X-ray fluorescence spectroscopy (Submitted to the Geoderma Journal, 30 June 2013).
4. Towett EK, Shepherd KD, Cadisch G, Eynekulu E, and Sila A. The potential of combining mid infrared (MIR) and total X-ray fluorescence (TXRF) spectroscopy for the prediction of soil properties. (Submitted to Soil Science Society of America Journal, Dec 2013).

***Books and theses***

1. Alex Oduor, Kipruto Cherogony, Jane Mutune, Maimbo Malesu, Anne Karuma, Maurice Cherogony, Douglas Nyolei, Jacqueline Kandagor & Duncan Onyango with Chris Ajele, Wilfred Wafula, Samuel Njalale, Alphus Lusweti, Peter Maina, Erick Towett, Arjen de Vries,David Tanui, Francis Mwangi, William Kimosop, Robert Kibunja and Saina Kiptoo. 2012. Food Security Framework for Turkana County. Diocese of Lodwar, P.O. Box 101-30500, Lodwar, Kenya. 217 pp.
2. Patrick Maundu, Yasuyuki Morimoto, Erick Towett, Julia A. Ombonya and Elizabeth Obel-Lawson (Eds). Mboga za Watu wa Pwani. Kilifi Utamaduni Conservation Group. Bioversity International 2011. ISBN: 978-92-9043-881-6.
3. Towett, E.K. 2008. Optimizing the use of Near Infrared Reflectance Spectroscopy (NIRS) to predict nutritional quality in cowpea (*Vigna unguiculata*) leaves for human consumption. MSc Thesis at Georg-August-University Göttingen, Germany. 109 pp. PDF-file - 4.5 MB – Available online: https://www.uni-goettingen.de/de/document/download/810d13d380e2d08b6709a151493fed23.pdf/Towett_2008_MScThesis_01.pdf
4. Johns, T., Maundu, P., Lawson, E., Kimiywe, J., and Towett E. 2006. Cookbook of African leafy vegetables. Bioversity International, Maccarese, IT. PD-file - 1.18 MB - Available online: http://idl-bnc.idrc.ca/dspace/handle/123456789/42275.

***Technical Reports and Miscellaneous Publications***

1. Near infrared reflectance spectroscopy (NIRS) calibrations for analysis of organic resource quality. Report submitted to the World Agroforestry Centre (ICRAF). April, 2009. PDF file-Available online: http://africasoils.net/sites/default/files/icrafOrgResCal.pdf
2. Isolation, characterization and larvicidal activity of *Turraea nilotica* against *Anopheles gambiae* larvae. Report submitted to the International Centre of Insect Physiology and Ecology (ICIPE). 2004.
3. Report on the status and use of sites of geographical, cultural and biological importance in the Central and North Rift regions of Kenya. Submitted to the National Museums of Kenya. 2002.
4. Traditional farming and livestock production systems in West Pokot, Kenya and their role in sustainable development (available as 40 pp and 109 pp full document by Imbumi Maryam). Submitted to the National Museums of Kenya. 2002.
5. Industrial attachment report submitted to the Tea Research Foundation of Kenya (TRFK). 2001.

***Conference posters and presentations***

1. Towett, E.K., Alex, M., Paul, C., Polreich, S. and Maass, B.L. 2008. Optimizing the use of near infrared reflectance spectroscopy (NIRS) to predict nutritional quality in cowpea (*Vigna unguiculata*) leaves for human consumption. Presented at Tropentag 2008, 7-9 Oct. 2008, University of Hohenheim, Stuttgart, Germany. Book of Abstracts p. 170. Available online: http://www.tropentag.de/2008/abstracts/links/Towett_deXjHwFz.php
2. Quantification of the variability and patterns in total element composition of sub-Saharan Africa soils. Erick Kibet Towett*, Keith D. Shepherd, and Georg Cadisch. 2011. Tropentag 2011 Book of Abstracts: Development on the Margin. Cuvillier Verlag, Goettingen, 2011. Available online: http://www.tropentag.de/2011/abstracts/links/Towett_YYA3uQry.php
3. The Potential of Combining Infrared and Total X-Ray Fluorescence Spectroscopy for the Prediction of Soil Functional Properties. E.K. Towett*, K.D. Shepherd, G. Cadisch, E. Aynekulu, and A. Sila. An extended synopsis submitted for the International Conference on Managing Soils for Food Security and Climate Change Adaptation and Mitigation, held on 23-26 July 2012, in Vienna, Austria.
4. Potential of MIR, TXRF and XRD as complementary techniques for assessment of soil properties. E.K. Towett*, A. Sila, M. Nyambura, E. Aynekulu, G. Cadisch, and K.D. Shepherd. Presented at the FAO-JRC International Workshop: “Soil spectroscopy: the present and future of soil monitoring”, held on 4-6 Dec 2013, at FAO HQ, Rome, Italy.

## Statutory declaration

I herewith declare that I have composed and completed this thesis independently without having used any other sources or resources than stated therein.

The author also acknowledges that the copyright of published manuscripts contained within this thesis resides with the copyright holder(s) of these works.

Place, Date: Stuttgart, 12/12/2013 Signature: