QUANTIFYING AND MAPPING SOIL ORGANIC CARBON IN MALI, WEST AFRICA USING SPATIOTEMPORAL METHODS

A DISSERTATION SUBMITTED TO THE GRADUATE DIVISION OF THE UNIVERSITY OF HAWAI'I IN PARTIAL FULFILLMENT OF THE REQUIREMENTS OF THE DEGREE OF

DOCTOR OF PHILOSOPHY
IN
TROPICAL PLANT AND SOIL SCIENCES

MAY 2008

By
Antonio Luis Evora Ferreira Querido

Dissertation Committee:

Dr. Russell Yost, Chairperson
Dr. Goro Uehara
Dr. Everett Wingert
Dr. Jonathan Deenik
Dr. Tomoaki Miura

We certify that we have read this dissertation and that, in our opinion, it is satisfactory in scope and quality as a dissertation of the degree of Doctor of Philosophy in Tropical Plant and Soil Sciences.

DISSERTATION COMMITTEE

Chairperson

DEDICATION

I gratefully dedicate this dissertation to my loving wife Sandra Ribeiro and my precious daughters Nadira R. Querido and Malika R. Querido whom I truly hope to inspire.

This dissertation is also dedicated to my dear mother, Carolina (Bébé) Evora, for unconditional love and for instilling the importance of higher education.

To all my sisters and brothers, I also dedicate this dissertation as a token of my love and appreciation.

TABLE OF CONTENTS

ACKNOWLEDGMENTS

This dissertation is the culmination of years of intense work, with usual ups and downs, on a continuous quest to improve my scientific foundations. It is also a testimony of the many generous and inspiring contributions of so many people. Indeed, the successful completion of this work could not be possible without the personal and practical engagement of these individuals.

I am appreciative and thankful to my advisor, Dr. Russell Yost, for providing the opportunity for me to pursue my academic and scientific goals. I remember vividly the painstaking decision process to disengage from my institutional responsibilities at INIDA and fully accept Dr. Yost's offer to be part of his research group. I am fortunate to have the mentorship of a gracious and brilliant scholar whose encouraging words, thoughtful criticisms, time and tireless commitment have contributed immensely to my PhD program.

The members of my dissertation committee, Dr. Goro Uehara, Dr. Everett Wingert, Dr. Jonathan Deenik, and Dr. Tomoaki Miura have generously given their time and expertise to enhance the quality of my work. I thank them for their relevant contribution and guidance during the course work and throughout the dissertation program.

To my colleagues Aminata Diarra, Gaoussou Diarra, Hamidou Konare, Kyle Barber, Laura Delisle, Rosalin Pattnaik, Rowena Valencia-Gica, Hu Li (Tiger) for sharing the enthusiasm for soil science and for all the help and support through trying times. I would also like to express appreciation to Dr. Gordon Tsuji, Guy

Porter, Mieko MacLachlan, and Agnes Shimamura. Special thanks to Richard Kablan for his friendship, unconditional help, and support.

I'm in debt to IER-Mali and in particular to my colleagues Dr. M. Doumbia, P.C. Sibiry Traoré, Dr. K. Traoré, Dr. A. Berthe, A. Ballo, S. Traoré, S. Sissoko, M. Sissoko, O. Samaké, L. Dioni for all the help in data collection and fieldwork support. My gratitude is due to the local farmers that allowed us to study their fields and shared the experiences for a common cause.

The following institutional support was crucial for the success of my PhD program: SM-CRSP/ USAID for providing the financial support; the University of Hawai'i/Tropical Plant and Soil Sciences Department for institutional support; the Borlaug LEAP-fellowship for funding my field trip to Mali and important scientific equipment.

To INIDA/Cape Verde I would like to express my sincere gratitude for the institutional and financial support. Special thanks to all my colleagues at INIDA for their moral support and friendship.

To my wife Sandra Ribeiro and my two precious daughters (Nadira and Malika), to whom this dissertation is dedicated, I would like to express my heart-felt gratitude for their constant and unconditional love, support and source of strength all these years. Also, to my mother, my brothers and sisters and my extended family I would like to express my appreciation for their love and constant encouragement.

ABSTRACT

The Kyoto protocol recognized the importance of the terrestrial sink of carbon and proposed schemes that allow countries to treat sequestered carbon as a commodity that can be traded for global environmental benefit. Carbon sequestration can be a win-win scenario because it also introduces a set of new benefits into dryland farming communities particularly in Sub-Saharan Africa. The possibility, however, for agricultural producers to participate in the emerging market for tradable carbon-credits requires a reliable verification mechanism. Soil carbon inventories of many developing nations rely on a broad scale assessment. These approaches do not account for the spatial and temporal variability of soil carbon nor do they provide a measure of uncertainty associated with these assessments. This study proposed the use of Bayesian Maximum Entropy (BME) to quantify and map soil organic carbon at field scale in four agroecological zones of Mali, Sub-Saharan Africa. The prediction model comparisons using the mean error (ME) indicated that BME performed better than did the kriging methods (0.033, 0.41, respectively). BME prediction also provided a lower MSE representing a 25% reduction compared with Kriging, and 10% compared with cokriging. This study also demonstrated potential use of space – time covariances as tools to improve our understanding of spatial and temporal variability of soil organic carbon. Based on the temporal and spatial models maps were generated to predict mean trends. The estimation of tree biomass in

Sub-Saharan Africa is important for an accurate assessment of the potential of these systems to capture and store carbon. The results show that tree carbon represented as much as 34% of the amount of organic carbon stored in soil surface (0-20 cm). Data from 2000 to 2006 indicated a net increase of soil organic carbon, which varied between 2.6 to 13.9 Mg ha^{-1}. Despite the complexities that characterize the spatial and temporal distribution of most environmental processes, BME provides a framework to analyze both space and time components.

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ACN	Contour ridges technology(fr.) Aménagement en courbe des niveaux
ADOS	Permanent contour ridges
ADSC	Agricultural Diagnostic Service Center University of HAWAI'I
AVHRR	Advanced Very High Resolution Radiometer.
BME	Bayesian Maximum Entropy
ccdf	Conditional Cumulative Distribution Function
CCX	Chicago Climate Exchange
cdf	Cumulative Distribution Functions
CDM	Clean Development Mechanism
CFC's	Chlorofluorocarbons
CMDT	Compagnie Malienne pour le Développement des Textiles
COP	Conference of the Parties
ETS	Emission Trading Scheme
EU ETS	European Union Emission Trade System
FAO	United Nation Food and Agriculture Organization
GDP	Gross Domestic Product
GHGs	Greenhouse Gases
GIS	Geographical Information System
GPS	Global position system
GPS	Global Position System
IER	Institute d'Economie Rural

IPCC	Intergovernmental Panel on Climate Change
NSW	New South Wales
OHVN	Office de la Haut Vallée du Niger
PDF	Probability Density function
Ppmv	Part per million by volume
RTCM	Radio Technical Commission for Maritime Services
S/TRF	Spatiotemporal random field
SM-CRSP	Soil Management – Collaborative Research Support Program
UK ETS	United Kingdom Emission Trade System
UN	United Nations
UNFCCC	United Nations Framework Convention on Climate Change
UTM	Universal Transverse Mercator coordinates system
WGS-84	World Geodetic System, dating from 1984

CHAPTER 1. INTRODUCTION

A map of soil properties is an essential tool for natural resource and agricultural land management. Maps illustrating the spatial distribution of soil properties have long been generated using diverse approaches and methods (Mueller et al., 2000; Heuvelink and Webster, 2001). However, the accurate mapping of soil properties cannot be viewed as an easy task. The diversity and variability of soils have been mapped for centuries and a number of approaches have been introduced in soil science in an attempt to represent soil properties (Mueller et al., 2000; Heuvelink and Webster, 2001). The challenges associated with sampling in space and time to appropriately represent the diversity and variability of soils is substantial. The variability is such that a thorough sampling to effectively map soil properties of a particular area can be cost prohibitive and constitutes many practical challenges (McBratney et al., 2003). Hence relying on a finite number of sample points in space and time always provides an incomplete assessment and, by default, a less reliable map of the soil property.

In pedology, soils are classified in discrete classes based on taxonomic characterizations which rely on field observations and laboratory analysis (Brady and Weil, 2002; Heuvelink and Webster, 2001). The soil surveyor approach generates maps with sharp boundaries and, in most cases, providing a single or a range of values for a particular property (i.e. pH, texture) within a large spatial class (Mueller et al., 2000). Although this approach implements a criterion of consistency it does not aid in the prediction nor does it provide any measure of

uncertainty. The need to predict at unsampled locations created the necessity to develop models that could rely on the observations of a continuous variable (i.e. soil property) and predicted values with their associated uncertainties (Heuvelink and Webster, 2001). With the advent of geostatistics in 1960's (Webster and Oliver, 2004;Isaaks and Srivastava, 1989) and subsequent refinements of the geostatistics methods (Goovaerts, 1997, Christakos, 2000), soil properties (i.e. carbon content) can be predicted in space and time considering not only the measured data but also additional site specific information (Christakos et al., 2001).

Quantifying and mapping soil organic carbon is of paramount importance to soil fertility management considering the role of organic carbon in nutrient retention, release and supply. It's also an equally important management tool for monitoring and verification of carbon sequestration by projects aimed at reducing the net greenhouse gas (GHG) emissions. Soil carbon stocks can be increased by practices that increase residue inputs to soils and/or reduce soil carbon mineralization rates. In Sub-Saharan Africa, these practices should also be accompanied by soil and water conservation measures (i.e. Aménagement en courbes des niveaux (ACN)), which also increases the soil water availability (Kablan et al., 2008).

There is a global concern that the effect of global warming in various regions of the world can be even more disastrous than presently (IPCC, 2007; 2000 (b); 1997). The scientific consensus is that global warming is human

induced, caused by the net increase of greenhouse gases, such as carbon dioxide (CO_2), methane (CH_4), nitrous oxide (N_2O) chlorofluorocarbons (CFCs) and sulfate aerosols (IPCC, 1997; Brady and Well, 2001; Mooney et al., 2002; Robert, 2001). The IPCC 2007 special report stated that the "warming of the climate is unequivocal". The report indicated that the last decade (1995-2006) was the warmest in the recorded history of global surface temperature, since 1850. The IPCC special report on emission scenarios (2000) projected an increase of global GHG gases on the order of 25-90% (CO_2 equivalent) between 2000 and 2030. Many nations and corporations are actively seeking ways to reduce their net emissions of greenhouses (GHG). However, the inability of many countries to decrease their net emissions will continue to impede the efforts to meet the goals of the Kyoto protocol (Robert, 2001; IPCC 2007).

The Kyoto agreement was negotiated as an amendment to United Nations Framework Convention on Climate Change (UN-FCCC), where nations formally recognized climate as a shared resource whose stability was essential to the entire planet (United Nations, 1992). Article 3.1 called for a reduction in the overall carbon dioxide equivalent emission of the greenhouse gases "by at least 5 per cent below 1990 levels in the commitment period 2008-2012" (UN, 2001 (c)). Furthermore, Article 3.4 of the same agreement called for each party to provide a carbon stock assessment and established soil carbon levels of 1990 as the reference. Many countries submitted their carbon sequestration assessment, but faced a considerable challenge in doing so due to a lack of historical as well

as current spatially distributed data on soil organic carbon (SOC) (Steven et al., 2003).

The Kyoto agreement recognized the importance of soils in carbon sequestration and opened the possibility for worldwide projects to increase carbon sequestration (Rosenzweig et al., 2002; Mooney et al., 2002). Several scientists and policy makers questioned whether it would be feasible to include agricultural soil carbon (C) sequestration in either government programs or in a market for tradable C credits (Robert, 2001; Mooney et al., 2002). Environmental and agricultural goals are often seen as incompatible; however, this is certainly not the case in most subsistence agriculture systems in Sub-Saharan region. Common interests can be identified between farmers, striving to improve the fertility of their soil, and the increasing global concerns to mitigate the effects of climate change (Woomer et al., 2001; Steven et al., 2003).

The possibility for agricultural producers to participate in the emerging market for tradable carbon-credits also requires a reliable verification mechanism to ensure that the actions taken to increase carbon sequestration in the soil can last for the duration of a particular contract (Mooney et al., 2002). The goal of reducing atmospheric CO_2 concentrations by increasing carbon sequestration in African soils is a win-win scenario considering other regional challenges such as combating desertification, increased food security, and biodiversity preservation (Tiessen, 1998; Manlay 2002).

1.1. The Problem

The soil carbon inventories of many developing nations rely on a broad scale assessment (Tiessen et al., 2004). These approaches do not account for the spatial and temporal variability of soil carbon nor do they provide a measure of uncertainty associated with these assessments. The quantification accuracy of soil carbon at large scale is often reduced by: (i) the heterogeneity of soils; (ii) the climatic and biophysical factors causing variation in local biomass production; (iii) and land management practices.

Carbon trading could provide subsistence farming in Sub-Saharan Africa incentives for improving soil fertility (Jones et al., 2007). Consequently, a carbon accounting system is of paramount importance to allow soil carbon sequestration to be accepted as an effective mechanism for reducing atmospheric CO_2 (Jones et al., 2007; Antle and Uehara, 2002). The accounting system should be able to provide a spatial and temporal estimate of soil carbon over a large area relying on all available site specific data.

The direct methods for accounting the soil organic carbon are costly and time consuming, requiring a large number of samples (Woomer et al., 2001). Yost et al. (1993; 2002) demonstrated the application of geostatistical methods to aggregate soil carbon over large areas. The main challenges to an effective carbon accounting system reside in providing a quantitative estimate of the uncertainty associated with soil carbon measurement and ways to aggregate soil carbon estimates over large areas.

Traditional approaches (i.e. taxonomic classification, interpolations) of mapped soil properties often lack quality since they fail to capture the spatial complexity of a natural process (Goovaerts, 1997). Geostatistics provides a number of statistical procedures to account for the spatial variability and dependence of the variable being measured (Goovaerts, 1997; Webster and Oliver, 2004). However, kriging methods rely only on the experimental data and cannot account for additional information that can help to improve our understanding of the spatial distribution of a particular random variable (Christakos et al., 2002; Serre, 1999) Consequently, when considering the possibility of an entire field estimate of soil carbon for trading purpose, other geospatial methods should be tested to ensure the reliability of the soil carbon inventory.

1.2. Overall Goal

The overall goal of this study was to develop a reliable and scientifically sound approach to quantify the soil organic carbon in the semi-arid agro-ecological zone of Sub-Saharan Africa. This study aims to address the urgent need to provide a mechanism for the monitoring of the role of soil systems in the biophysical carbon cycle while providing an accurate assessment of the carbon sequestered and stored in the soil system.

1.3. Specific Objectives

- To apply Bayesian Maximum Entropy (BME) geostatistical method to quantify spatial and temporal estimates of soil carbon.
- To evaluate the potential of the BME approach to forecast the spatiotemporal changes of soil organic carbon in a particular field and to determine the uncertainty of the estimated changes.
- To map soil organic carbon stock in selected sites of the semi-arid Sub-Saharan Africa and provide a methodology for soil carbon monitoring and verification within the framework of a carbon emission market.
- To estimate the pool of carbon in trees and evaluate their contribution to carbon sequestration by semi-arid agroforestry systems.

CHAPTER 2. SOIL CARBON: A LITERATURE REVIEW

2.0. Soil Carbon Sequestration

The terrestrial ecosystem comprises a wide diversity of ecosystems important in the net removal of CO_2 from the atmosphere. Soils, in particular, play a major role in the global carbon budget. The amount of carbon found in the soil is larger than that in the atmosphere and biomass combined (Janzen, 2004; Brady and Weil, 2002; West, and Post, 2002). The importance of soils to humankind is, however, far more than its role in the global carbon cycle. Soils provide the essential conditions for food production and improve the environment by absorbing a wide range of pollutants resulting from human actions. However, the perception of soils in most societies around the world is dismissive of its importance (Hillel, 2004; Kimble et al., 2002). The generally negative views toward soil systems are detrimental to conservation efforts and tend to perpetuate practices and attitudes conducive to land and environmental degradation at a global scale. The definition of soil tends to vary according to its intended use; however, it can generally be described as a living membrane between bedrock and the atmosphere (Brady and Weil, 2002; Bohan et al., 2001). Soils are diverse ecosystems where a variety of vertebrates and invertebrate organisms live (Brady and Weil, 2002; Bohan et al., 2001).

Globally, it is estimated that currently 75% of carbon is stored in some form of organic carbon (Kimble et al., 2002; Brady and Weil, 2002). High levels

of soil carbon contribute to soil productivity and sustainability because it improves nutrient adsorption, water retention, and decreases the erodibility factor (Kimble et al., 2002; Bohan et al., 2001). Consequently, storing carbon in the belowground system "is the best long term option" to offset the CO_2 emissions (Kimble et al., 2002). Indeed, to offset the continuing carbon buildup in the atmosphere, the storage of organic carbon in soil is a feasible alternative (IPCC, 1997; 2000(c)). Carbon dioxide is removed from the atmosphere by plants through photosynthesis. In the presence of solar energy, plants use CO_2, water and nutrients to produce energy needed for biomass accumulation and seed production (Brady and Weil, 2002). Organic material is composed mainly of sugars, crude proteins, hemicellulose, cellulose, fats, and lignin (Brady and Weil, 2002). Plants are eventually consumed by animals or die and most of these organic materials return to soil and are broken down by soil organisms returning some of the stored carbon to the atmosphere. The mineralization of organic matter also releases essential nutrients for plant growth. Despite the dynamic processes associated with carbon cycle, it's in the soil that the carbon has the most residence time, hence making soil an ideal ecosystem where carbon sequestration may be effective.

Soil carbon is highly dynamic and according to Lal (2002), a reduction of 1 Pg of soil carbon presents an enrichment of 0.47 ppmv[1] of CO_2. Lal (2002) presented estimates from different authors (Jobbagy and Jackson, 2000; Batjes,

[1] Parts per million by volume, a unit used in climate change terminology to express concentrations of GHG.

1996; Eswaran et al., 1993) of the amount of organic carbon in the soil pool. The estimates in 1 m depth of soil varied from 1502-1642 Pg. The global estimates of soil carbon potential are approximate considering the difficulty associated with capturing the global variability and the particularity of diverse ecosystems.

The conversion of pristine land to agricultural production has been reported to sharply reduce the amount of carbon stored in soil (Isaac et al., 2005; Kieft et al., 1998). When soils are cultivated, the surface layers often lose 20-50% of their carbon (Schlesinger, 1995; 1999). The populations of microorganisms which participate in the mineralization of soil organic matter also increases dramatically in cultivated fields as compared with non-cultivated fields, which dramatically increases the oxidation rate of soil organic matter (Brady and Weil, 2002; Schlesinger, 2000).

Carbon sequestration in terrestrial ecosystems can be defined as the net removal of CO_2 from the atmosphere into long-lasting pools of carbon (IPCC, 2005). The pools can be living, aboveground biomass (e.g., trees), living biomass in soils (e.g., roots and microorganisms), or recalcitrant organic and inorganic carbon in soils and deeper subsurface environments (Brady and Weil, 2002; Eswaran et al., 1993). This carbon must be fixed into long-lived pools (Ingram and Fernandes, 2001). Otherwise, one may be simply altering the size of fluxes in the carbon cycle, not increasing carbon sequestration. Carbon sequestered is the difference between carbon gained through the photosynthetic

mechanism and the amount of carbon lost either by respiration or oxidation of organic matter. The overall gain or loss is referred to as the net ecosystem productivity (Montagnini and Nair, 2004).

Carbon sequestration can be a win-win scenario because it also introduces a set of new benefits into dryland farming communities particularly in Sub-Saharan Africa ((United Nations, 2001(b); Ringius, 2001). The benefits to local farmers involves the enhancement of soil quality, the conservation of the soil and water resources and the environment and, last but not least, the economic incentives that benefit the farmers and promote the best management practices while increasing the amount of carbon stored in the soil (Ringius, 2001).

2.2. Carbon Trading

Carbon trading is a concept that allows the trading of allowances to emit carbon dioxide and other greenhouse gases. Basically, two main types of carbon transactions are: i) allowance base and ii) project-based. In the allowance base transaction country A pays country B in exchange for a given quantity of GHG emissions "credits", that country A can use to meet its target under the climate mitigation (Capoor and Ambrosi, 2006). The Kyoto protocol recognized the importance of the terrestrial sink of carbon and proposed schemes that allow countries to treat carbon as a commodity that can be traded for global environmental benefit.

The carbon market structure stipulates that payments can be made in the following forms: cash, equity, debt, or in kind contributions associated with technology transfer to decrease GHG emissions (Capoor and Ambrosi, 2006). The project base transactions allow buyers to actually purchase emission credits from a project that can "credibly and verifiably demonstrate that it reduces GHG emissions" (Capoor and Ambrosi, 2006).

In Europe, the allowance transaction can take place over the counter by brokers, however it should be stated that trading platforms have already been established (Capoor and Ambrosi, 2006). According to the World Bank report (Capoor and Ambrosi, 2005) in 2004, the project-based transactions increased to 107 $MtCO_2e$ (metric ton of CO_2 equivalent), representing a 38% increase relative to the previous year. The European Directive 2003/87/2000 established the Emission Trading Scheme (ETS) that instructs and allocates to Member states an emission allowance. In project-based transactions, buyers can purchase emissions credit from a project that reduces GHG (Capoor and Ambrosi, 2006).

Developing nations could earn as much as $100 billion annually by 2050 from selling carbon credits, according to an analysis released by the World Bank at the United Nations conference on climate change in Nairobi, November 2006. The same report claimed that carbon credit trade has increased to about $5 billion over the past two years, a figure which could grow twenty-fold in the next 40 years as developing countries invest in renewable energy and sell their credits to developed countries (World Bank Report, 2006). The major buyers are the

European countries (56%) and Japan (38%), US, Canada, Australia and New Zealand all purchased 1% each and the rest (2%) from unspecified countries (Capoor and Ambrosi, 2006).

The primary sellers in the carbon market are in Asia (73%), with China and India as the principal leaders. Latin America countries account for 17% of all project based transactions that occurred in 2005. The joint implementation (JI) in economies in transition positioned last with only 3%. African countries continue to be largely marginalized in the carbon market, despite the huge potential of the continent in terms of carbon sequestration and historically lower emissions as compared with the other human-dominated regions of the globe.

The European Union Emission Trade System (EU ETS) is by far the largest carbon market in terms of value and volume. It's larger than New South Wales (NSW), and the Chicago Climate Exchange (CCX) markets combined (Capoor and Ambrosi, 2006). The World Bank report (2006) estimated that the EU ETS was worth US$8.2 billion in 2005 and traded US$6.6 billion just during the first three months of 2006 (Capoor and Ambrosi, 2006). This figure contrasts with US$57.2 million on NSW and US$2.8 for CCX. The EU ETS is playing a major role in helping the European Union achieve its obligation under the Kyoto agreement (Capoor and Ambrosi, 2006). According to the same report (Capoor and Ambrosi, 2006) the dynamics of the EU market contributed to raise the price expectation particularly for project-based markets. For instance in 2005, the report indicated that 374 million tCO_2e were transacted at an average price of

US$7.23 which totaled 2.7 billion USD. This price corresponded to more than a three-fold increase from the 2004 volume from project-based and over five times above the previous year's value (Capoor and Ambrosi, 2006).

2.3. Soil Carbon in Tropical Ecosystems

Tropical ecosystems store a large quantity of above and below ground carbon and play an important role in global carbon cycle. Soil organic carbon (SOC) in tropical systems is affected by processes involving environmental factors and human actions. Theng et al. (1989) concluded that the quality of tropical soil organic matter was similar to that found in temperate regions. They attribute the variation in quantity and composition to climatic and edaphic conditions (i.e. clay content, pH, soil mineralogy and moisture). However, studies on organic matter depletion and nutrient turnover of tropical soil have generally shown much greater rates of mineralization in the tropics than that observed in temperate zones (Coleman et al., 1989). Increasing temperature plays a major role in increasing the oxidation of carbon, which clearly increases the decomposition rate of the available plant material (Coleman et al., 1989).

In soil systems with low fertility, the mineralization of organic matter plays a critical role in the provision of nutrients (Piéri, 1989). Therefore, practices that destroy organic residues (i.e. slash and burn) have a negative impact on soil's nutrient availability, hindering the sustainability and productivity of agriculture systems (Manlay, 2000). In other words the fertility of tropical soils is strongly

linked to the rate of mineralization of organic carbon in the soil and residue management (Coleman et al., 1989; Piéri, 1989; Manlay, 2000; Lui et al., 2004). In 1991 Srivastava and Singh reported that some soils had lost 50% of their soil organic matter upon conversion from forest to an agricultural system. This result was corroborated by Tiessen et al. (1992) who also explained that a sudden decline in soil organic matter (SOM) was sufficient to cause drastic decreases in soil fertility, resulting from forest conversion to agriculture land in Africa and Latin America. Indeed complex interactions exist which go beyond the simple nutrient cycle and measurable nutrient pools. A report from Ghana focused on carbon and nitrogen dynamics in a 25 year chronosequence of cacao production, and concluded that within the first two years after land conversion, 16% of the original carbon had been lost mainly due to the increased decomposition rate (Isaac et al., 2005). Many studies have reported results that indicate a significant decrease of soil organic carbon following land conversion (Isaac et al., 2005).

The soil organic matter is, in most cases, the only nutrient input available to African farmers. Traditionally, these farmers have learned to value the role of litter and manure in crop production and in some cases deliberate efforts are made to replenish the soil of organic matter and to insure good production (Manlay, 2000). However, under subsistence farming the majority of harvested crop residue is not used to replenish the soil nutrient stocks. Manlay (2000) demonstrated that most plant biomass generated in mixed farming systems in

West African savannas has a multitude of uses, from a food source for livestock to domestic energy fuel. Consequently, most is not returned to the soil.

In the tropics and particularly in dry tropical climates, where the rainfall distribution varies both in space and time, the turnover of nutrients from soil organic matter is highly dependent on soil moisture(Coleman et al., 1989). The turnover rate of SOM depends not only on the amount of organic carbon present but it is also controlled by the biophysical conditions that characterize tropical areas. Bohan et al., (2001) reported that the average half life of fresh organic matter in the humid tropics is as little as 3-4 weeks as compared to 3-4 months in temperate regions. Therefore, to increase of soil organic matter content by simply adding organic material has proven to be a daunting task if not impossible. The more organic matter is added to a cultivated land, the more it is oxidized (Bohan et al., 2001). Bohan et al. (2001) described an experiment carried out in England since 1843 where 30 ton ha^{-1} of organic manure was added to cultivated soil annually but it failed to restore the soil organic matter to prior, uncultivated levels.

2.4. The Role of Organic Carbon in Soil and Water Conservation

Soil conservation measures are essential for the preservation of soils and often contribute to increased water availability to plants. In most African countries the idea of soil conservation is strongly associated with engineering approaches to control soil erosion and reduce overland run-off. A more comprehensive view of soil and water conservation that puts emphasis on soil productivity, sustainability and equitability is needed but is slowly being adopted by farmers. The shift from soil loss prevention (an engineering perspective only) has been combined with a number of agronomic, conventional and traditional practices aiming at improving soil properties and productivity of the biological systems. For example, the SM-CRSP has successfully tested the ACN technology and the preliminary results indicate a significant increase both in soil productivity and soil moisture (Gigou, 2006; Kablan et al., 2008). Other practices such as crop rotation, intercropping, and agroforestry have increased soil carbon and improved soil properties as well as the livelihood of the farmers (Bationo and Buerkert, 2001; Bationo and Kihara, 2007).

2.5. Soil Organic Carbon and Soil Properties

The organic carbon impact on soil properties is extensive. In terms of water retention, organic matter can hold up 20 times its weight in water (Bohan et al., 2001), which helps prevent drying and shrinking of soils and improves the moisture retention capability particularly in sandy soils. Another important

property of organic carbon in soil is that it helps maintain a uniform reaction in the soil, due to its capacity to buffer soil pH. Organic matter contributes to increase the exchange capacity of the soil from 20 to 70% (Boham et al., 2001). The mineralization of organic matter also yields several important elements (i.e. NH_4^+, NO_3^-, PO_4^{3-}, and SO_4^{2-}) to soil fertility. As indicated above, the return of crop residues is not a widespread practice in West Africa due to the following reasons: most crop residues are too often destroyed by termites, grazed by communal herds or simply burnt later in the season in preparation for the rain or as a source of energy for domestic use. Soil organic carbon is a critical factor in soil aggregate stability and its mineralization is particularly important to nutrient recycling in low input agricultural systems (Chivenge et al., 2006).

2.6. Agroforestry and Soil Carbon

Understanding the interactions of the woody species and other components of the ecosystem in the tropical and subtropical grass land and savanna is important when studying the potential of carbon sequestration of different ecosystems. Large areas of Africa, Australia and South America are comprised of savanna but it still remains the least studied terrestrial ecosystem (Huntley and Walker, 1982). According to the estimates of IPCC (2000), tropical savannas store one third as much carbon in vegetation as do tropical forests. Montagnini and Nair (2004) estimated that in small agroforestry systems of the tropics the potential for carbon sequestration ranges from 1.5 to 3.5 Mg C ha^{-1}.

The interaction of savanna trees and shrubs with understory grasses and soils is important for the sustainability of these systems (Woomer et al., 2004b). Some savanna trees (i.e. *Vitellaria paradoxa*) annually shed huge amounts of leaves contributing to increased soil litter and nutrient recycling (Traoré, 2003). Trees, thus are likely to contribute more to the soil organic carbon pool than most rainfed agricultural systems were most crops are annually harvested and a substantial portion of both grain and residue are transported out of the production system, resulting in minor and in some cases a negative carbon budget.

Within the context of the Kyoto protocol and climate change mitigation efforts, the IPCC (Intergovernmental Panel on Climate Change) recognizes agroforestry systems as having a high potential for carbon sequestration (Garrity, 2004; Flandez, 1998). Despite the positive outlook, an important question to be addressed is how small stakeholders can benefit from carbon sequestration projects.

2.7. Carbon Accounting

Soil carbon stock assessment often is based on assigning a carbon density to a class of soil, based on data collected from soil profiles for each soil group (Brown et al., 1989; Woomer et al., 2001; Mooney et al., 2002; Steven et al., 2003). The uncertainty associated with this type of carbon inventories results from the extrapolation of soil profile data (Garnett et al., 2001). Problems with

soil and vegetation classification, inaccurate area estimates and unrepresentative values of carbon content for soil types are listed among one of the major limitations of this approach (Steven et al., 2003; Garnett et al., 2001). Woomer et al. (2001) organized a workshop in Senegal West Africa on "landscape carbon sampling and biogeochemical modeling". The course provided guidelines for field and laboratory measurement carbon. Carbon measurements of woody biomass (200 m^2), understory biomass (1 m^2), surface litter and roots (0.25 m^2), soil bulk density and carbon were inventoried.

Vagen et al. (2005) provided a review of the potential for soil carbon sequestration in Sub-Saharan Africa. The study concluded that there was a high potential for soil organic carbon increase through the establishment of improved fallow systems with an estimated rate of carbon sequestration from 0.1 – 5.3 Mg C ha^{-1} yr^{-1}. A similar study by Ringius (2002) called for long-term (greater than 10 years) field experiments, demonstration and pilot projects for soil carbon sequestration in Africa. While Africa's contribution to global CO_2 emissions from fossil burning and other industrial activity is low (World Resources Institute, 1998) the entire Sub-Saharan Africa can contribute to offset CO_2 emissions on a global scale. Woomer et al. (2004) estimated Senegal's terrestrial carbon stocks in 1965, 1985, and 2000 using an inventory procedure involving satellite images revealing historical landuse change, and field measurements of standing carbon stocks occurring in soil and plants. The average of C stock was calculated for eleven ecological region and landuse combinations. The different ecological

regions and landuses were combined into 62 distinct areas and their respective carbon stock was computed with an adjustment for woody biomass removal. Based in this study (Woomer et al., 2004) carbon stocks in Senegal ranged from 9 t C ha^{-1} (degraded savannas) to 113 t C ha^{-1} (Senegal River Valley). The carbon estimates showed a declining trend from 55% in 1965 to 38% in 2000. The study concludes that 292 Mt of carbon were lost from Senegal between 1965 and 2000. The main cause of this decrease was attributed to a steady decline in woody biomass due to overexploitation and climate.

Feller (1993) concluded that in West Africa, as a whole, soil carbon contents are correlated with silt and clay content and landuse. Chivenge et al. (2006) reported on a study conducted in Zimbabwe and concluded that long term residue management was affected by conservation practices and differed with clay soil and sandy soils. The results suggested that tillage disturbance was the dominant factor in reducing C stabilization in a clayey soil. The report recommended different conservation practices for the two types of soil. In fine textured soils the focus should be on reducing soil organic carbon decomposition. On sandy soils, however, maintaining the carbon input could be more beneficial to long-term sustainability of the agroecosystem (Chivenge et al., 2006).

The total stock of organic carbon for Africa was predicted to be 90–96 Pg C (1Pg = 10^{15} g) for the upper 0.3 m of soil and 170–179 Pg C for the first 1m of soil (Batjes, 1996; 2003). About 12 per cent of global carbon is to be found in Africa (Batjes, 2001(b)). These estimates correspond to an area-weighted mean

content of about 3 kg C m^{-2} in the top 0.3 m, and 6 kg C m^{-2} in the 1m depth. Batjes (2003) estimates indicated an increase in soil carbon from about 3 kg C m^{-2} to 1m depth in the arid zone to about 9 kg C m^{-2} in the humid tropics.

Birch-Thomsen et al. (2007) used a geographical information system to map the temporal changes in soil organic carbon stock in the semi-arid region of Tanzania, East Africa. This study suggested that the original stock of native carbon has been reduced by more than 50% within the upper 0.25 m during the last 50 years of cultivation. Carbon estimates at a country or regional scale always pose a major scientific challenge both in terms of the methodological approach and the resources necessary for reliable estimates (Garnett et al., 2001). One important consideration of the scientific approach for estimating carbon stock at large scale should be the sampling density and the spatial representation of the samples. There will always be a tradeoff between sampling density and accuracy; however, one important aspect to be considered is the spatial autocorrelation of the soil carbon. Yost et al. (1993) pointed out that extrapolation on the basis of Soil Taxonomy and geostatistical extrapolation provided similar estimates of uncertainty in soil organic carbon. The study also indicated that soil C is likely to be correlated along large distances. Geostatistics techniques provide a framework for spatial and temporal analysis of soil carbon. It relies on a spatial model and provides an estimate of the error associated with prediction. Geostatistics may help meet some of the challenges of spatial and temporal estimation of soil carbon (Webster and Oliver, 2004).

CHAPTER 3. THE STUDY SITE

3.1. Site Characteristics

The republic of Mali is the largest West African country (1.24 million km^2) with a rich and illustrious history and culture. Located in the Northern West Africa, Mali is a land-locked country sharing its borders with several other nations (Figure 3.1). Mali is extremely flat, and the landscape is characterized by low plains broken occasionally by rocky hills. The country extends from the Sahara to the tropical humid zone and comprises three agroclimatic zones: (i) the Sahelian zone which includes a pastoral region and has a rainfall of less than 200 mm: (ii) a fragile agricultural region, with a rainfall of 400 mm and a dry savanna type of natural vegetation; and (iii) a zone that covers the Sudano-Sahelian, Sudanese and Sudano-Guinean region, with a shrub and tree savanna vegetation and a single rainy season which brings between 500 and 1200 mm of rainfall per year (Flandex, 1989).

The major rivers of Mali are the Niger and the Senegal (Fig. 3.1). The Niger is the principal West African river. It springs out the Fouta Djallon plains in Guinea and then it runs east into the Sahel, cutting through Bamako, towards Mopti and Tombouktou were it creates a large inland delta with several channels and lakes. Just before the ancient city of Gao the river turns south into Niger and Nigeria and finally into the Gulf of Guinea. Before reaching the sea, the Niger brings life to a stretch of 4,180 km in Guinea, Mali, Niger and Nigeria. The

Senegal River also originates in Fouta Djallon and runs in west Mali towards Kayes before entering Senegal. Despite these two rivers, water is still a scarce resource and only 1,380 km^2 of the land is irrigated. Out of the vast land surface that comprises Mali only 3.82% is considered arable land (CSLP, 2001).

Deforestation, soil erosion, desertification, inadequate supplies of potable water and poaching are identified as current environmental issues in Mali (CSLP, 2001). The accelerated desertification of Mali is due to on-going droughts, over-grazing, soil erosion, harsh desert winds, and the cutting of trees for firewood. Mali is party to the following international agreements on the environment: the Conventions on Biodiversity and Climate Change, the Kyoto Protocol, the Conventions on Desertification and on Endangered Species, Hazardous Wastes, the Law of the Sea, Ozone Layer Protection and Wetlands (CSLP, 2001). Recently, Mali developed its National Environmental Action Plan (NEAP) a Multi and inter-sectorial plan designed to address the main constraints associated with the natural environment. One key aspect of this plan is the decentralization and involvement of local communities in the design and implementation of the action plan (CSLP, 2001).

The natural hazards Mali faces are hot, dust-laden "harmattan" hazes common during dry seasons and especially during recurring droughts, and occasionally, flooding from the Niger River (CSLP, 2001). The extreme climate and random seasonal rainfall pattern makes productivity of rainfed agriculture highly unpredictable. The climate of the Western Africa sub-region varies greatly

from North to South and is mainly governed by the seasonal migration of the ITCZ (Intertropical Convergence Zone). Mali's climate is subtropical to arid characterized by a hot and dry period (February to June), rainy, humid, and mild season (June to November); cool and dry period (November to February). However, as expected, there is remarkable climate variability within Mali. The northern part is almost entirely arid desert or semi-desert. In the central region, known as the Sahel, most agriculture activities follow the Niger River's annual flood cycle, with high water between August and November. In the Southwestern area, rainfall is more abundant compared with the rest of the country.

The average temperatures range from 24° to 32°C (about 75° to 90°F) in the south. Temperatures are higher in the north. The hottest period of the year is registered before the rainy season of June to September. Annual rainfall declines from about 1,400 mm (about 55 in) in the south to some 1,120 mm (some 44 in) at Bamako and less than 127 mm (5 in) in the Sahara of the north (Traoré, 2003). Despite the drought and severe shortage of water that causes considerable hardship on the local population, Mali is predominantly an agricultural country. According to CSLP (Strategic Framework for Poverty Reduction) document (CSLP, 2001) 80% of the Mali active population participates in agriculture which contributes to 45% of the GDP. The country's most valuable resource is the Niger River, which abounds in fish and its waters are used for irrigation. Mali's mineral resources include gold, salt, phosphate

rock, iron ore, diamonds, and uranium. Gold is the most important mineral being mined (CSLP, 2001). The economy is predominantly agricultural, and crops depend almost entirely on irrigation or flooding from the Niger River, as 65 percent of the land is in desert or semi-desert region. Droughts in 1969-74 and 1981-83 devastated the crops and cattle herds of the northern region and set off waves of migration to urban areas, mainly to Bamako. Mali is heavily dependent on foreign aid and is vulnerable to fluctuations in world prices for cotton, its main export (CSLP, 2001).

The 1998 census, reported that Mali's population was 9.79 million people (Census, 1998). In 2000, the population was estimated at 10.6 million (CSLP, 2001). The greatest population growth is concentrated in the Southern provinces of Mali, particularly in the capital, Bamako, which is situated on the banks of Niger River. Given the overall country size, the population density is still considered very low (8.7 persons per square kilometer); however, there are wide disparities in regional concentration (CSLP, 2001).

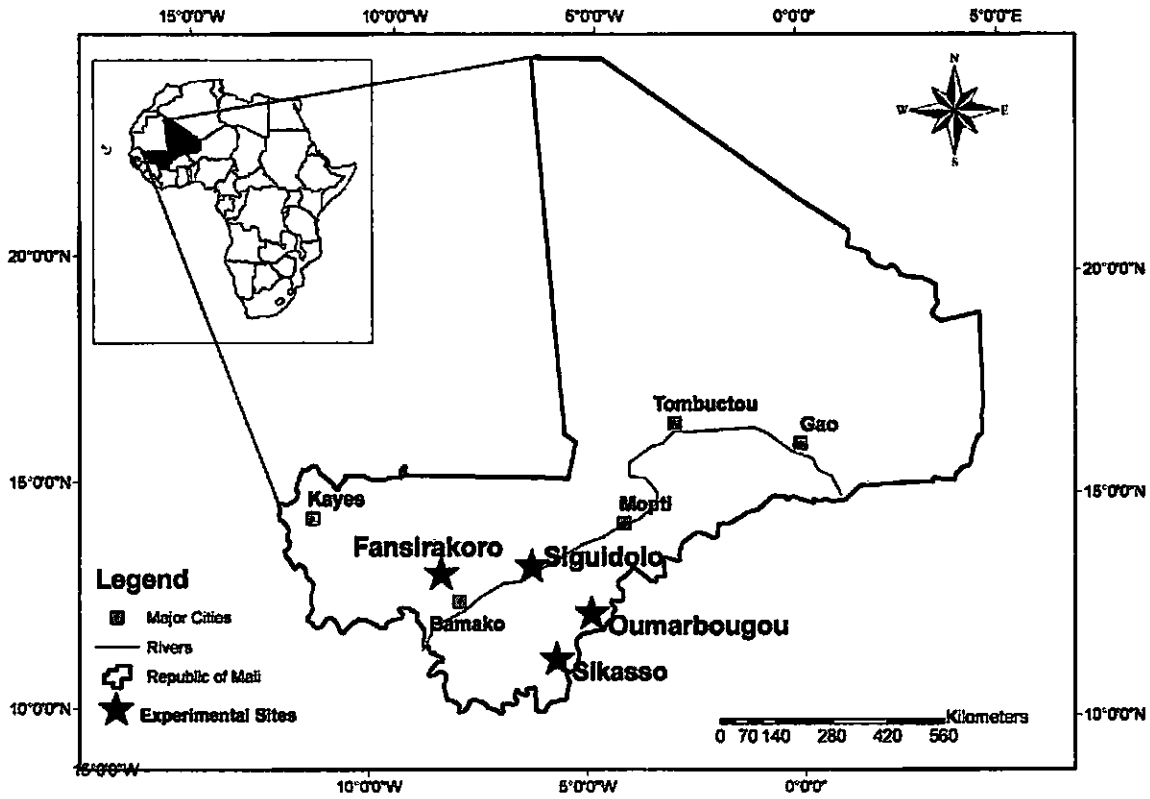


Figure 3.1. Location of the experimental sites in Mali. Four main experimental sites (Fansirakoro, Siguidolo, Oumarbougou and Sikasso) were selected by IER and SM-CRSP scientists and a total of 11 farm fields are monitored.

3.1.1. The Experimental Sites

The selected sites represent three distinct agro-ecological zones (see section 3.1). Table 3.1 below shows a brief summary of each experimental field in terms of rainfall, the cultivated crops and the farmer's field area. The Soil Management Collaborative Research Support Program (SM-CRSP) monitored over 11 farmer's fields and collected site specific data throughout the duration of the project (2000-2006). Each field is characterized by its own specific bio-physical characteristics and management practices. At each site a soil and water conservation practice known as "Aménagement en courbes des niveaux" (ACN) was implemented and several indicators monitored (i.e. soil carbon, soil moisture, and tree diversity). All sites were characterized as traditional agricultural production systems and in some cases minor inputs (fertilizer) were used. Most of the semi-arid zones of Mali are covered with a diversity of trees that play a critical role in the sustainability of these ecosystems.

It has been estimated (Traoré, 2003; Cissé, 1995) that almost all rainfed agriculture in Mali occurs in agroforestry systems, where the major trees such as *Vitellaria paradoxa* (Karité), *Faidherbia albida* (balanzan), and *Parkia biglobosa* (néré) are important to the biochemical cycle of the essential mineral nutrients. Through their extensive root system, these trees are able to remobilize a substantial portion of leached nutrients to the soil surface due to their seasonal leaf shedding. The agroforestry systems in Mali are representative of much of

the Sahel, and are very diverse. The complexity is enhanced by an intense human intervention in terms of tree selection and agricultural practices.

Table 3.1. Summary table of the sites considered in this study and the estimated area of each farmer field where carbon data were collected.

Sites	Farmer's Field	Rainfall Range (mm)	Main Crop	Field Area (ha)
Fansirakoro	Baracoro Ballo	~ 800-1000	Sorghum, Cotton Maize	4.77
Siguidolo	Zan Diarra	~ 600-800	Sorghum, Millet, Cotton	18.18
Oumarbougou	Mory Konate	~ 800-1000	Cotton Millet/Sorghum	20.15
Sikasso	Yaya Diassa	~ 900-1100	Cotton, Maize Millet/Sorghum	44.89

Fansirakoro

Fansirakoro is located in the Commune of Yélékebougou, 64 km from Bamako in the Upper Niger River watershed. This village belongs to the OHVN (Office de la Haut Vallée du Niger) "Cercle » of Kati (see Figure 3.2). At Fansirakoro two farmer's fields were monitored (Drissa Traoré and Baracoro Ballo) for soil carbon change and the impact of soil conservation on productivity and groundwater. The Fansirakoro experimental sites were subject to severe erosion and land degradation. At Baracoro Ballo and Drissa Traoré's fields small to medium size gullies were observed giving a clear indication of active erosion processes. These gullies tend to be continuous and part of a more complex

drainage network of the entire area of Fansirakoro. Most of agricultural practices involve labor intensive traditional cultivation with very minor mechanization. The main crops include millet, sorghum, and maize.

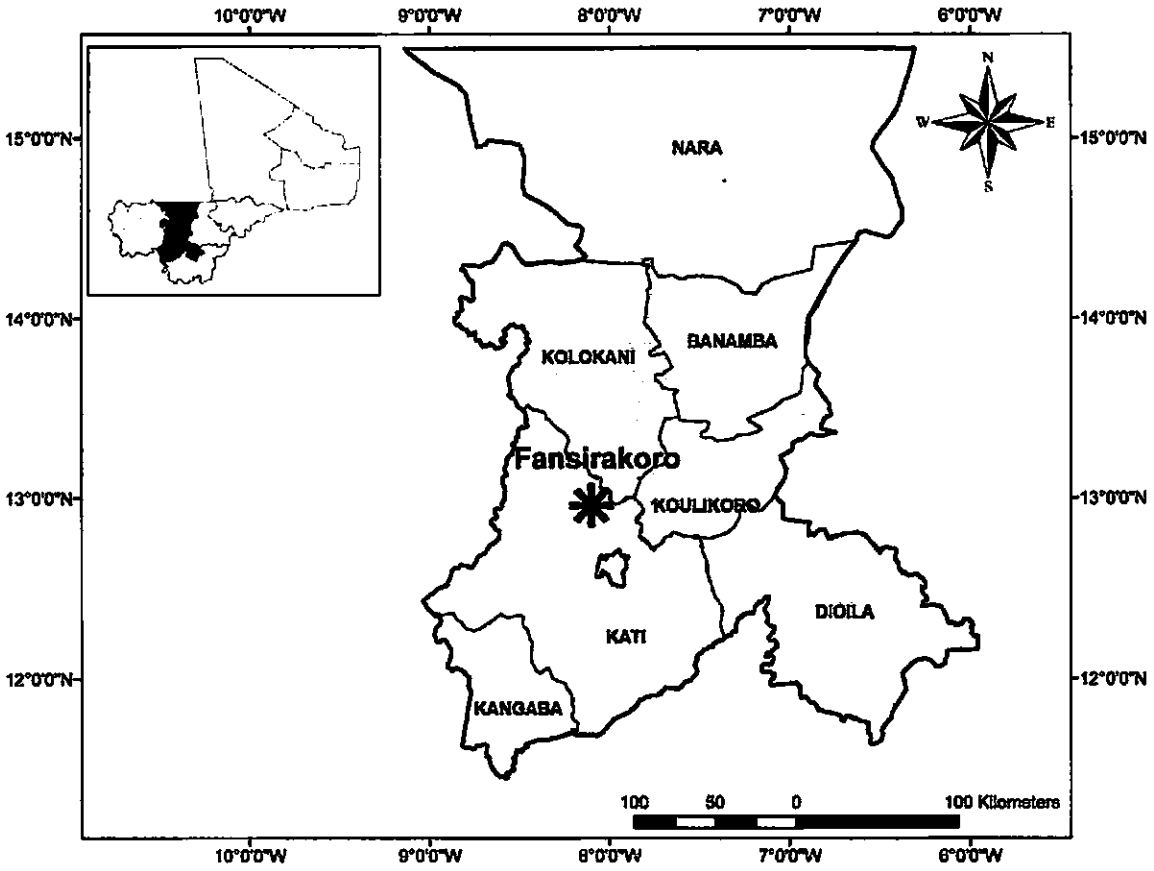


Figure 3.2. The Koulikoro Region is comprised of seven "Cercles". Fansirakoro is the selected site for the entire region.

Siguidolo

Siguidolo is located in the commune of Konobougou, "Region" of Segou (see Figure 3.3). In terms of agricultural extension, this area belongs to the CMDT (Compagnie Malienne pour le Developpement des Textiles) sector of Konobougou. In this particular region is characterized by extensive cultivation of cotton, with the CMDT subsidizing fertilizer loans that farmers have to repay with a portion of their cotton harvest. Despite the intense cultivation of a cash crop (cotton), all farmers still use rudimentary tools relying almost exclusively on animal traction and manpower.

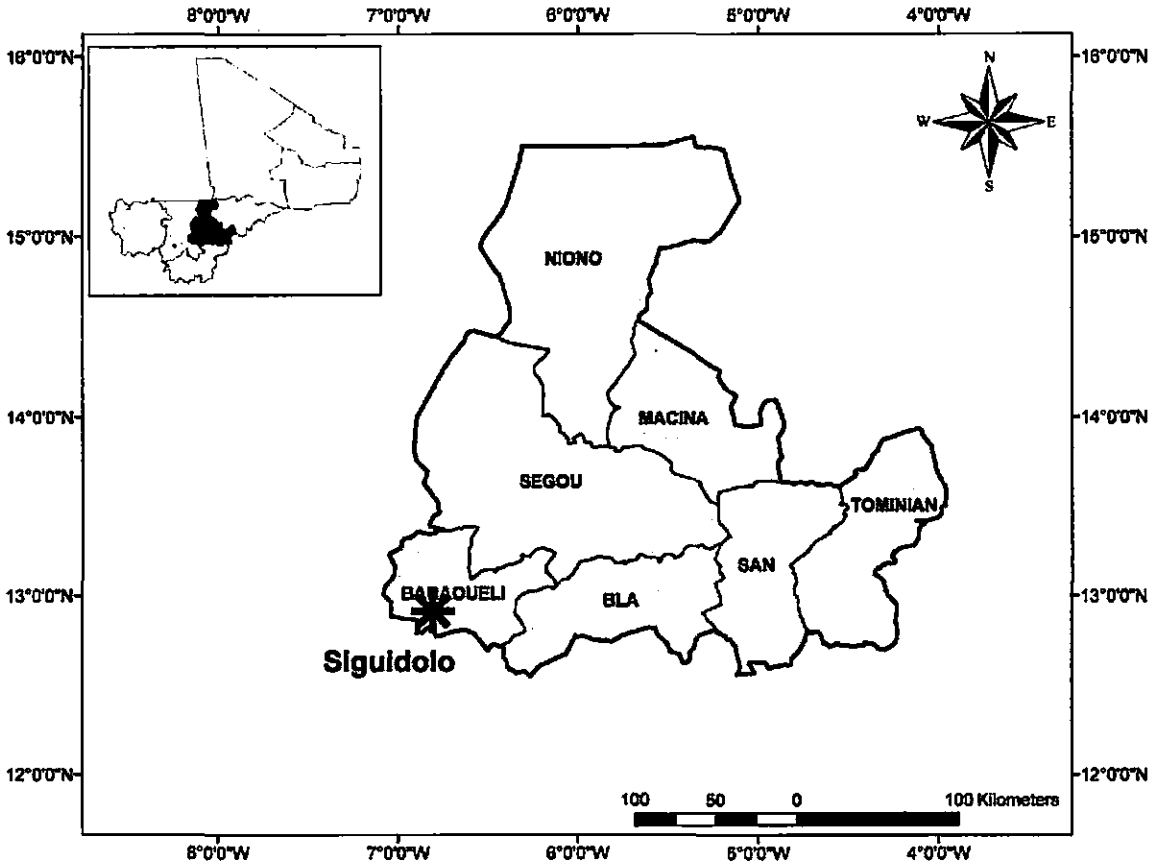


Figure 3.3. Segou Region is comprised of seven "Cercles". Siguidolo (Zan Diarra) is the representative site for the entire region.

Sikasso

Sikasso is the most southern "Region" of Mali, which is comprised of seven administrative units called "Cercles" and 148 communes (see Figure 3.4). Rainfall is relatively more abundant in Sikasso than the other sites (see Figure 3.5).

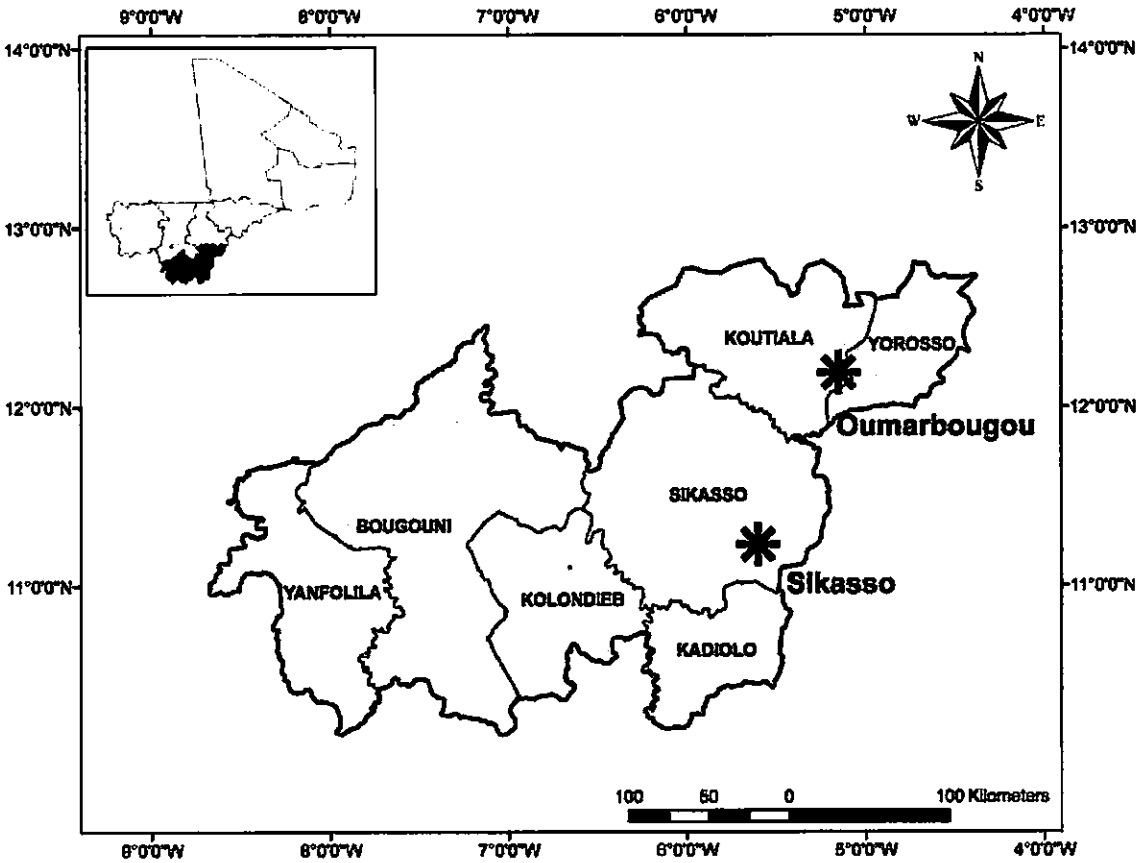


Figure 3.4. The Sikasso Region, with seven administrative units ("Cercles"). Oumarboughou and Sikasso sites are also indicated.

In this "Region" two major sites were selected: one in the "Cercle" of Koutiala (Oumarbougou) and the second site is located in Sikasso "Cercle", which has the same name as the region and the main city. All sites were associated with a local community where the land is normally managed by the family or the

village. The SM-CRSP and local institutions (e.g. IER) have introduced soil and water conservation measures to help increase food security through improving soil productivity and sustainability.

3.2. Climate

Although the rainfall amount varies from site to site (Table 3.1), the climate is characterized as Sudanian tropical dry (Manlay, 2000). The annual rainfall at these sites ranges between 600 mm to 1100 mm per year. Although, the beginning of the rainy season is highly unpredictable, most rainfall is expected between June and October (see Figure 3.5). Data from the Mali national service of meteorology indicated that there has been a major decline in rainfall and clear shift downwards of the rainfall isohyets since 1971 (Traoré et al., 2005). Nicholson (2001) presents a summary of rainfall patterns across Africa based on the records registered in 20th century. This study demonstrated that climate change was most significant in the semi-arid regions of West Africa. Nicholson (2001) studied the meteorological factors as well as desertification, albedo and dust and concluded that changes were more significant in Sahel between periods of 1931-1960 and 1968 – 1997.

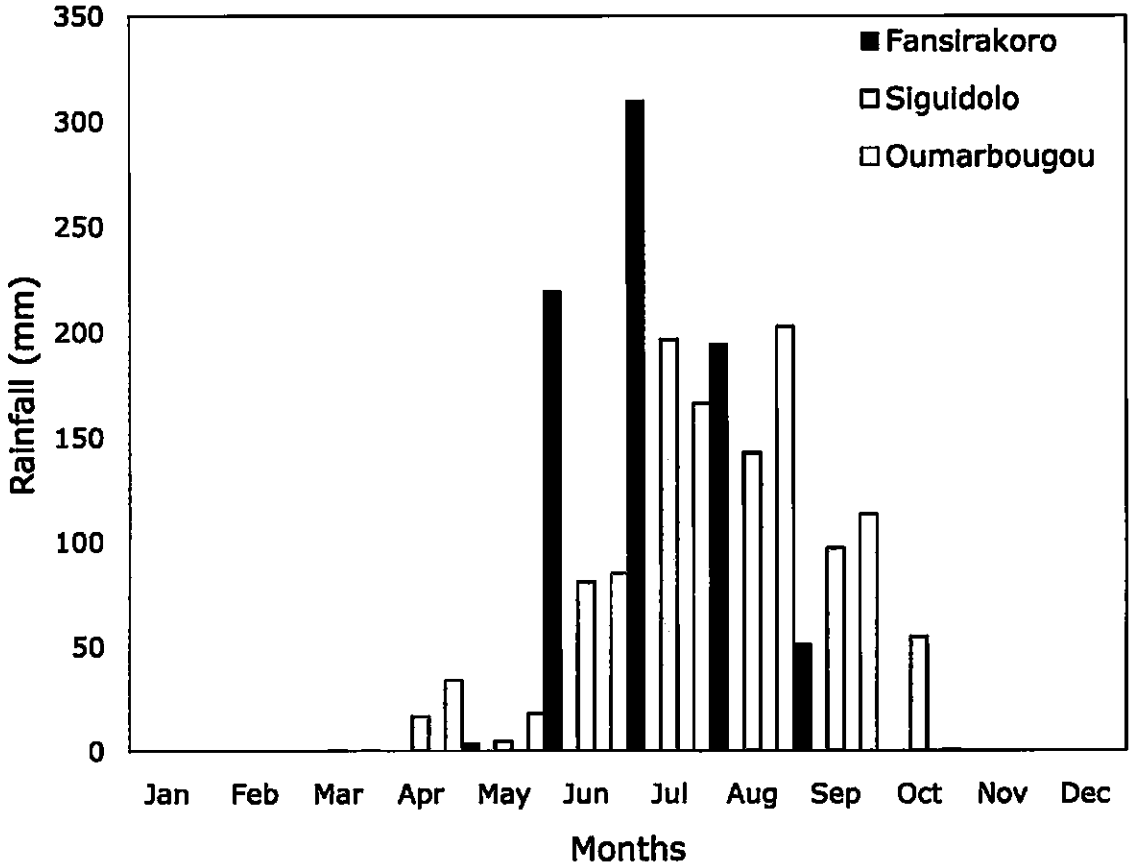


Figure 3.5. Monthly distribution of rainfall, registered in 2004 at Fansirakoro, Siguidolo and Oumarbougou. At Siguidolo and Oumarbougou the yearly total rainfall was 619 mm and 593 mm. The highest total rainfall amount was registered at Fansirakoro (778 mm).

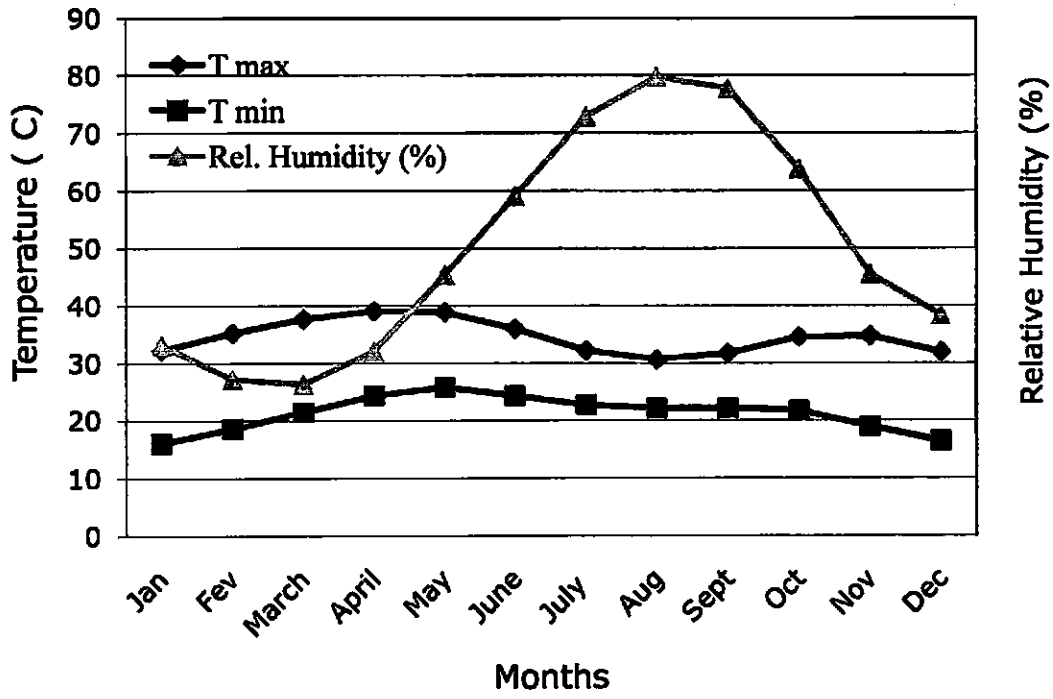


Figure 3.6. Climatic data from 1960-2002 at Konobougou, (adapted from Traoré, 2003). The relative humidity is low for most part of year and it peaks during the month of July to October. The beginning of the year is the hottest with temperatures reaching 40 °C in April and May.

The Figures 3.5 and 3.6 provide a brief glimpse of the climatic conditions at some of the sites. High relative humidity and lower temperature occur during the rainy season. The total monthly rainfall for 2004 at three sites did not exceed 400mm. A comparison of Figures 3.7 and 3.5 shows that only for a short period of the year does the rainfall exceed the evapotranspiration potential (ETP), indicating for most part of the year the water balance is negative.

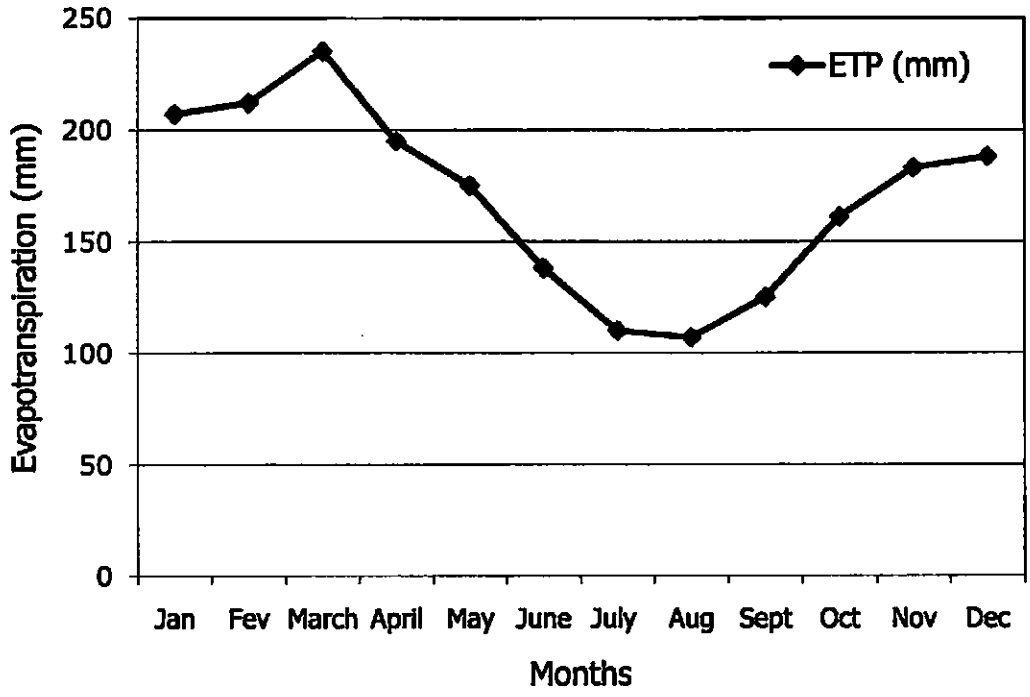


Figure 3.7. Evapotranspiration potential from 1960-2002 at Konobougou, (adapted from Traoré, 2003).

3.3. Soil

Most Sub-Saharan African soils have been subjected to extensive weathering processes that have leached most of their mineral nutrients (Piéri, 1989). These weathered soils are mainly Ultisols and Alfisols with low fertility and sandy surface texture (see Table 3.3 below). In Sub-Saharan Africa Typic Plinthustalfs and Plinthic Paleustalfs are the predominant types of soils. These are highly weathered soils with ustic moisture regime (some plant available moisture during the growing season followed by long periods of drought) contain a plinthite horizon (Brady and Weil, 2002). Plinthite (from Greek, plinthos, brick)

is an "iron-rich, humus poor mixture of clay with quartz and other minerals" (Soil Survey Staff, 2006). Generally the plinthic horizon when subjected to repeated wetting and drying may change irreversibly to an ironstone hardpan (Figure 3.8). The hardening of the plinthic layer reduces rainfall infiltration with negative consequences on erosion, runoff and crop yield.

These soils are usually low in organic carbon (less than 10 g kg^{-1}) and cation exchange capacity (CEC) is equally low due to low clay content and the predominance of low activity kaolinitic clays (Bationo et al., 2006; Piéri, 1989). The average soil organic carbon concentration for West Africa was estimated to be 4.2 – 4.5 Kg C m^{-2}, these values were lower than the average value for whole Africa (6.4 – 6.7 kg C m^{-2} (Batjes, 2001). The macronutrients such as nitrogen (N), potassium (K) and phosphorus (P) are extremely low (Traoré, 2003).

Table 3.2. Soil properties of the Siguidolo field site (Kablan et al., 2008).

Soil Classification: Plinthic Paleustalfs

Soil Depth	pH H_2O	Bulk Density (g cm^3)	Org. C (g kg^{-1})	Org. N (g kg^{-1})	Sand (%)	Silt (%)	Clay (%)
0-20	5.89	1.58	2.4	n.d.	79	16	4.9
20-40	5.36	1.52	2.2	n.d.	66	20	13.5

Table 3.3. Average and range in soil properties from the Oumarbougou study area (Kablan et al., 2008)

	Soil pH	Sand (%)	Silt (%)	Clay (%)	Soil C (%)
Mean*	5.3	72	22	6	0.23
Max	6.2	83	40	14	0.64
Min	4.5	55	10	3	0.12

*(n=17)

Generalized land degradation due to adverse agroecological conditions, erosion and runoff is a major problem affecting all experimental sites. Water erosion is a major problem contributing to severe soil degradation (Traoré, 2003; Gigou, 2007). Excessive overland flow at the onset of the cropping season can pose serious risks to plants since the increased runoff is associated with reduced soil water availability and crop production.

Casenave and Valentin (1990) showed that surface crusting plays a key role in reduced infiltration and the increased runoff and erosion in Sub-Saharan Africa. Nine forms of surface crusts and the factors responsible for their formation were described by Casenave and Valentin (1990). The soil at all sites was characterized by a generalized soil crusting and in most cases farmers could only work the land with the traditional implements after the first rainfall.

Figure 3.8 illustrates a toposequence and the relative intensity of soil erosion and runoff associated with geomorphological topology of the region. High tree densities are usually found near the village (Traoré, 2003; Manlay, 2000). Karité (*V. paradoxa*) and balanzan (*F. albida*) are among the most common tree species. The tree composition changes with the toposequence and farmers selectively allow trees species with high economic and spiritual value to

thrive (ex. Karité, néré (*P. biglobosa*), tamarind (*Tamarindus indica*), Baobab (*Adansonia digitata*), *Lannea acida, Lannea microcarpa, Sclerocarya birrea, Prosopis juliflora, Ficus* spp.) while undesirable ones are weeded out at the onset of every rainy season.

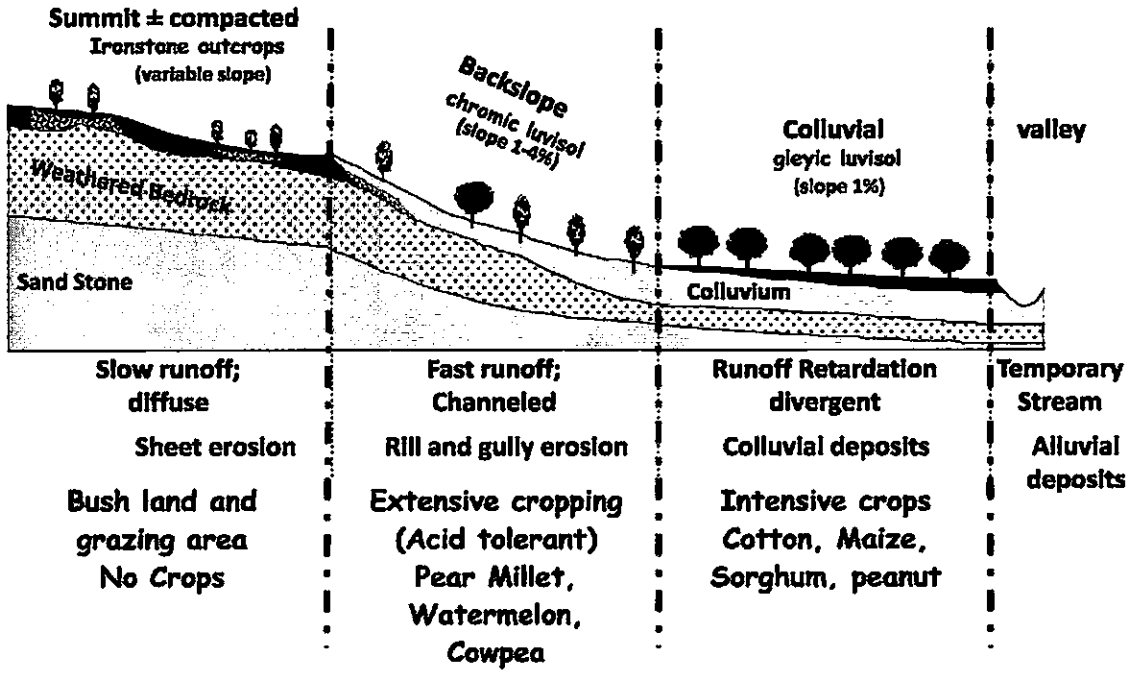


Figure 3.8. An illustration of the landscape form in Mali. The interaction between the topography, landuse, soil erosion and runoff. Note that composition of tree species also changes in the toposequence (Adapted from Gigou, 2007 and Maraux et al, 2007).

The summits are characterized by an indurated surface often resulting from erosion (exposed plinthite that has hardened to stone). Shrubs and grazing pastures are the main landcover found at the summits. These are essentially unsuitable areas for agricultural activity (Figure 3.8). The backslope (slope gradient between 1-4 %) is normally cultivated with acid tolerant crops such as pearl millet, watermelon and cowpea (Maraux et al., 2007). On the colluvial

deposits, where more hydromorphic soils are found, an intensive rainfed cultivation of cotton, maize, pearl millet, sorghum and peanut dominate the landscape (Gigou, 2007; Maraux et al., 2007).

Traoré (2003) reported that in soil derived from colluvial deposits the trees have an average C content of 22 Mg C ha^{-1}. According to the same author, isotopic carbon studies indicate that an increase in soil carbon was probably due to the tree influence on soil C, which was estimated to extend up to 2.5 m beyond the canopy. Traoré (2003) reported that with a density 24 tree ha^{-1}, 35 kg CaO ha^{-1}, 8 kg MgO ha^{-1}, 4.5 kg K_2O ha^{-1} and 9 kg N ha^{-1} was returned to the soil through the leaves in one year.

In the cotton area, farmers often receive fertilizer loans from the Compagnie Malienne pour le Developpement des Textiles (CMDT) to ensure cotton production. Through crop rotation, the local farmers have learned to fully utilize the expensive fertilizer that is subsidized for the cotton. The following crop rotation sequence: cotton (*Gossyplum hirsutum,* L.) – subsistence crop (maize (*Zea mays*), sorghum, millet) - cotton, is an example of one of the management practices used by the farmers to take advantage of the residual effect of fertilizers normally applied to cotton. Through the extension services of CMDT, farmers receive fertilizer credit that corresponds to 150 kg ha^{-1} of Nitrogen-Phosphorus-Potassium-Sulfur-Boron (NPKSB) fertilizer (applied at planting) and 50 kg ha^{-1} after the emergence. The NPKSB is a formulation rich in phosphorus (14-22-12, + 5.5 S + B) (Traoré, 2003) and its application has contributed to a

significant increase of cotton production and according to Traoré (2003) the application of inorganic fertilizer has become a common practice in cotton production.

Traditionally crop residues are used for feeding the livestock and also as an energy source for cooking. Most rural families in Sub-Saharan Africa still largely use wood as their main source of energy (Manlay, 2000; Traoré, 2003). Farmers often burn the stalks and twigs to prepare the land for tilling as soon as the rains come. There is a wide spread resistance to incorporating the remaining residue in the soil. However, farmers are becoming increasingly aware of the importance of soil organic matter and soil and water conservation practices. This is certainly the case in the SM-CRSP selected sites at Sikasso, Oumarbougou, and Fansirakoro.

According to FAO (1998), 10% of the world's population lives in Sub-Saharan Africa, despite the poor soil fertility, and generalized food insecurity. Africa as a whole makes very little use of inorganic fertilizers (less than 1% of the cultivated land receives fertilizer) (Bationo et al., 2006; Kwesiga et al., 1997). The agricultural systems selected by the SM-CRSP C project were mainly of subsistence with little or no use of inorganic fertilizer input. In the selected experimental sites the main agricultural inputs are from organic sources, occurring without a concerted effort to replenish soil fertility.

The soils of Sub-Saharan Africa are low in organic matter. Considering the fact that inorganic fertilizers are not widely used (except in cotton production

regions), the increase in soil carbon can improve the overall nutrient status of these impoverished soils while increasing soil water holding capacity and contributing to increased land productivity. Increased buildup of organic carbon in the soil has also a positive environmental impact, since it contributes to offset the CO_2 emissions.

CHAPTER 4. GEOSPATIAL ANALYSIS

4.1. Introduction

Almost all natural processes, soils in particular, develop over a large area or region and their spatial extent depends on climate, parent material, organisms, time, and topography. Data collected in one region show that neighboring samples tend to have similar values and the similarity decreases with increased distance between samples. This spatial correlation is used to investigate whether point values are randomly distributed or whether the values have a relation to each other with distance. The spatial autocorrelation structure can be modeled by the variogram (Bailey and Gatrell, 1995; Goovaerts, 1997; Isaaks et al., 1989).

Geospatial analysis has many applications in soil science for the prediction of soil properties using a small dataset (Anctil et al., 2002). Despite the fact that geostatistics is relatively new compared with other classical statistics several methods have been proposed, each of which designed to address a specific condition. This chapter discusses a few classical geostatistics methods and presents the theory of Bayesian Maximum entropy (BME). The accuracy and precision of BME and classical geostatistics methods are compared in chapter 5.

4.2. Classical Geostatistics Methods

Goovaerts (1997) defined geostatistics as a term applied to a group of statistical techniques used to describe the correlation in space of spatially distributed random variables and their prediction at unsampled locations. Geostatistics relies on spatial models (i.e. variograms) and the predictions are based on the weighted sums of the data, the spatial autocorrelation and the configuration of the data points (Webster and Oliver, 2004). Webster and Oliver (2004) traced the history of geostatistics and recognized the contribution of R. A. Fisher in 1919 that rightly recognized the implications of spatial variation in his experiment. By blocking his experiment (in 1919) R. A. Fisher developed the well known analysis of variance to estimate the effects of both short and long range variation (Webster and Oliver, 2004). Other scientists such as Youden and Mehlich (1937) were also credited for their relevant work in trying to reveal and estimate spatial variation. Kolmogorov, a Russian scientist, also developed a set of functions to represent the spatial correlation and provided ways to interpolate, according to Webster and Oliver (2004), similar to kriging. However, it was only in 1960 that a French mathematician (George Matheron) provided a complete solution for the theory of regionalized variables. Matheron later coined the term kriging after a South African engineer (D.G. Krige) who had observed that his estimates of ore grade improved when considering the grades of the neighbor samples. Like many other statistical techniques, geostatistics has evolved and a number of techniques have been proposed to improve the prediction and to

circumvent the limitations associated with ordinary kriging. In the sections below a few kriging methods and their relevance to this study are discussed.

4.2.1. Simple Kriging

Kriging is often referred to as the best linear unbiased estimator because it minimizes the error variance under constraints of unbiasedness and the estimates are weighted linear combinations of observed values (Isaaks and Srivastava, 1989; Goovaerts, 1997). At any given measured point, kriging estimates return the same value (exact interpolator) (Isaaks and Srivastava, 1989; Goovaerts, 1997). The weights are determined to minimize the error variance under the unbiasedness constraints (Goovaerts, 1997). Using the spatial variance structure given by the semivariogram model, simple kriging provides an estimate at an unmeasured location and the magnitude of the probable error of the estimated value (Nielsen and Wendroth, 2003).

$$\hat{Z}_{SK}(u) = \sum_{\alpha=1}^{N} \lambda_\alpha z(u_\alpha) + \{1 - \sum_{\alpha=1}^{N} \lambda_\alpha\}\mu \ , \qquad \text{(Eq. 4.1)}$$

Where Eq. 4.1 is the simple kriging estimator of the random variable Z_{sk}, at the prediction location u. The λ_α are the weights at location u computed such as to minimize the error variance. The μ is the known mean which assumes second order stationarity (i.e. the variance can be assumed constant across the region). At every unsampled location, the value is predicted using a weighted linear combination of the available data plus a weighted mean. The unbiasedness is guaranteed by the second term on the right-hand side of the equation (Eq. 4.1).

In simple kriging the sum of the weights are not constrained to sum to 1 (Webster and Oliver, 2004; Goovaerts, 1997).

The expected estimation error of a random variable is defined as:

$$E[\hat{Z}(u) - Z(u)] = 0 , \qquad \text{(Eq. 4.2)}$$

Where $\hat{Z}(u)$ is the predicted value at point (u) and $Z(u)$ the measured value at the same location. Since the estimator is unbiased, the expected error ($E[\hat{Z}(u) - Z(u)]$) is equal to zero. The stationarity of the means implies that the residual covariance function $C_R(h)$ is equal to the stationary covariance function $C(h)$ (Goovaerts 1997), thus simple kriging can be presented in terms of Z-covariance as:

$$\sum_{\beta=1}^{n(u)} \lambda_\beta^{SK}(u) C(u_\alpha - u_\beta) = C(u_\alpha - u) \qquad \alpha = 1, \dots n(u) , \qquad \text{(Eq. 4.3)}$$

Where $C(u_\alpha - u_\beta)$ is the covariance between data points u_α and u_β. The $C(u_\alpha - u)$ represents the covariance between locations (u_α) and prediction point(u). The ($\sum_{\beta=1}^{n(u)} \lambda_\beta^{SK}(u)$) is the sum of all weights. In other words, the sum of the weighted covariances between locations u_α and u_β is equal to the covariance between u_α and prediction location(u). The solution of simple kriging equation (Eq. 4.3) and its corresponding variance is positive if the covariance matrix is positive definite. This requires that no two data are colocated ($u_\alpha \neq u$), for a ≠ b and also that the covariance model *C(h)* must be permissible (examples of permissible models are spherical, exponential, Gaussian, and power models).

The minimum error variance also called the simple kriging variance, $\sigma^2_{SK}(u)$, is given by the following equation:

$$\sigma^2_{SK}(u) = C(0) - \sum_{\alpha=1}^{n(u)} \lambda_\alpha^{SK}(u) C(u_\alpha - u) \quad , \qquad \text{(Eq. 4.4)}$$

Where λ_α^{SK} represents the weights, C(0) is the covariance value for distance (h)=0. The other terms are as defined in Eq. 4.3.

Kriging yields not only a least squares estimate of the random variable of interest but at the same time it presents an error variance that, according to Goovaerts (1997), is dependent on the covariance model, which also implies that the estimation precision depends on the complexity of the spatial variability of the variable as modeled by the covariance (Goovaerts, 1997; Webster and Oliver, 2004). Simple kriging does not take into account any additional information that might exist in a particular area. The solutions presented by simple kriging rely exclusively on measured data. Because a simple kriging estimator always assumes a Gaussian distribution, the predicted value is the mean of the distribution.

4.2.2. Indicator Kriging

The indicator form of kriging begins with the coding of the information, based on a number of thresholds designed to create classes of approximately equal frequency. One advantage of indicator kriging is that it allows the use of both hard and soft data regardless their origin (Goovaerts, 1997). The main objective of indicator kriging is to evaluate, at any location, a set of cumulative

probability distributions (cdfs) or posterior probabilities of the random variable in order to estimate the threshold values Z_k. The function $F(u;Z_k|(n))$ is hereby defined by the following posterior probabilities (Goovaerts, 1997; Serre, 1999):

$$F(u;Z_k|(n) = \text{Prob}\{Z(u) \leq Z_k|(n)\}\ , \qquad \text{(Eq. 4.5)}$$

The $F(u;Z_k|(n)$ denotes the conditional or posterior probability that given the data Z(u) the true value is 1. In other words the indicator approach is based on the interpretation of the conditional probability (Eq. 4.5) as the conditional expectation of an indicator (Eq. 4.6) given the information (n).

The $i(u_\alpha;z_k)$ is the indicator value for the random variable Z(u) at a threshold value z_k as defined by the following:

$$i(u_\alpha;z_k) = \begin{cases} 1, & if\ z(u_\alpha) \leq Z_k \\ 0, & otherwise \end{cases}\ , \qquad \text{(Eq. 4.6)}$$

Values below or equal to the threshold receive the code 1 while any other value greater than the threshold will be coded as 0.

However, the conversion of a continuous variable into an indicator implies that a significant amount of information can be lost (Goovaerts, 1997, Webster and Oliver, 2004). Goovaerts (1997) exemplified the use of indicator kriging to determine the estimates of the probability of exceeding an established threshold. Like simple kriging, the indicator kriging estimator is also an exact estimator. That is, the posterior probability of exceeding a critical threshold is one at data locations where the values are smaller than the threshold and the probability is also equal to 0 at the location where the data values exceed the critical threshold (Goovaerts 1997; Webster and Oliver, 2004).

4.2.3. Ordinary Cokriging

Geostatistics provides a number of statistical procedures to account for the spatial variability and dependence of the variable being measured (Webster et al., 2002; Goovaerts, 1997; Isaaks et al., 1989). Unlike simple, ordinary and indicator kriging, cokriging allows the integration of a secondary variable that can be colocated or not-colocated. For instance, by allowing the incorporation of a secondary variable (typically the covariable is a low cost, more densely measured variable), through a procedure known as cokriging, the prediction of the primary variable (i.e. carbon) can sometimes be improved (Lark, 2002; Ishida and Ando, 1998). This procedure also ensures unbiasedness in making predictions at unsampled locations and provides an estimation of the error (Goovaerts, 1997; Webster et al., 2002). The estimated value at point x_0 can, with the integration of one or more co-variables, be defined as follows (Webster and Oliver, 2004):

$$\hat{Z}_u(u) = \sum_{l=1}^{V} \sum_{\alpha=1}^{nl} \lambda_{\alpha l} z_l(u_\alpha) , \qquad \text{(Eq. 4.7)}$$

Where, $\hat{Z}_u(u)$ is the predicted value at point u. The subscript l refers to variables, of which there are V, and the subscript l refers to the sites where variable l was measured (Webster and Oliver, 2004), and u denotes the variable we wish to predict. The $\lambda_{\alpha l}$ are weights that must satisfy the following conditions: The sum of the weights ($\sum_{\alpha=1}^{nl} \lambda_{ll}$) equals to 1 if l=u, or zero in case l ≠u. The $z_l(u)$ are measured data.

This procedure incorporates not only spatial correlation but also inter-variable correlation (Webster and Oliver, 2004). The cokriging equations are

normally obtained and expressed in terms of the variograms and cross-variograms, or, alternatively, the covariances and cross-covariances. Soil properties are spatially auto-correlated and show spatial dependence of their correlation with other variables. In cokriging the estimates of the temporal changes in the variable are coherent with the estimates of the variable on different dates (Goovaerts, 1997).

4.2.4. Variograms

The spatial correlation is usually represented by the variogram or covariance models (Goovaerts, 1997; Webster and Oliver, 2001; Lark, 2002b). Variograms and covariance functions are the fundamental tools for modeling dependent data observed over time, space, or space-time. (Isaaks and Srivastava, 1987; Goovaerts, 1997; De Cesare et al., 2001; Chunsheng, 2005). The variogram model provides an expected semivariance for a given distance. When the distance between sampled points is zero, the semi-variance is expected to be zero. In practice the semi-variance is not zero, but there exists a nugget effect, which is a measure of short scale variability and the experimental error. The experimental variogram is calculated as one-half of the observed average square difference in data values for every pair of data locations at a specified separation distance (see Eq.4.8).

$$\gamma(h) = \frac{1}{2N(h)} \sum_{\alpha=1}^{N(h)} \left[z(u_\alpha) - z((u_\alpha) + h) \right]^2, \qquad \text{(Eq. 4.8)}$$

Where, $[z(u_\alpha) - z(u_\alpha + h)]$ is an h-increment of attribute z, and N(h) represents the number of pairs at data locations, at h vector of distance (Goovaerts, 1997). The calculated semivariance values for all paired combinations, within the maximum distance allowed, are grouped into distance classes, also known as lag distances (see Figure 5.12). The number of pairs per bin is an indicator of the relative importance of a particular class on the model structure.

The variogram is a measure of spatial dependence or, in other words, how quickly the measurements of a particular phenomena change, on the average, with changes in distance (Isaaks and Srivastava, 1989; Goovaerts, 1997; Webster and Oliver, 2001; Cassiani and Christakos, 1998). This is based on the principle that in nature two observations (ex. soil carbon) sampled at a short range are more similar than those observations farther apart. The variogram is a two or three dimensional function with two independent variables (direction θ, the separation distance *h*) and one dependent variable, the semi-variance ($\gamma(\theta,h)$). The variogram modeling exercise produces a sill, range, nugget as well as the anisotropic information. Variogram models are, therefore, mathematical models that quantify the spatial variability of the data set.

The cross-semivariance, $\gamma(ij)$, is a measured of joint variability between two random variable and it can be defined as half the non-centered covariance between h-increments (Goovaerts, 1997):

$$\gamma(ij) = \frac{1}{2N(h)} \sum_{\alpha=1}^{N(h)} \left[(z_i(u_\alpha) - z_i(u_\alpha) + h)(z_j(u_k) - zj(u_\alpha) + h) \right], \qquad \text{(Eq. 4.9)}$$

where N refers to the number of data pairs that are separated by the same distance h. Unlike semivariance (Eq. 4.8), the cross-semivariance can be negative because the value of one property may be increasing while the other in the pair is decreasing. The cross-semivariance is used for modeling the experimental cross-variogram in cokriging operations.

Natural processes tend to have a preferred orientation, hence carbon data in particular field may change faster in one direction than others. Hence, the spatial distribution, trend, form and direction are examined, the anisotropy modeled in order to detected spatial variability with change in direction. The anisotropy can be corrected by transforming the original vector of coordinates in to a new vector, using trigonometric functions that are used in modeling the anisotropic semivariogram (Goovaerts, 1997).

4.3. Bayesian Statistics

Bayesian statistics relates the conditional and prior probabilities of two random events.

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{\sum_k^n P(A_k)P(B|A_k)} \qquad \text{(Eq. 4. 10)}$$

The Bayesian theorem combines prior probability $P(A_i)$ and likelihood information $P(B|A_i)$, in ways that the prior probability is updated and a posterior probability ($P(A_i|B)$, results (Sivia, 1996). The prior probability refers to the fact that no information from B is considered. ($P(A_i|B)$ is therefore known as conditional probability or posterior probability since the value of A depends upon B.

4.3.1. Bayesian inference in a Gaussian model

Ribeiro (1999) presented Bayesian inference in Gaussian model-based geostatistics. Basically this particular approach regards the variogram model parameters as random variables and the predictive distribution is obtained over the different parameters values with respect to their posterior distribution (Ribeiro, 1999). In this respect the particular approach differs from classical geostatistics techniques where the model parameters are considered as known. As referred to above, Bayes' theorem combines prior and likelihood information. The prior knowledge of the parameters is updated as follows:

$$pr(\vartheta|Y) \propto pr(\vartheta)\, pr(Y|\vartheta). \qquad \text{(Eq. 4.11)}$$

In the Bayesian approach both variable Y and model parameter ϑ are considered to be random quantities with a joint distribution pr $(y,\vartheta) = pr(y|\vartheta)\, pr(\vartheta)$ (Ribeiro, 1999). The information about the model parameters not related to the data is reflected in the prior distribution $pr(\vartheta)$. The distribution for ϑ conditional to variable Y is proportional to the product of the prior distribution of ϑ and the prior distribution of Y conditional to the set of parameters of the model.

The predictive distribution $(pr(y_o|y))$ takes into account the parameter uncertainty by averaging over the parameter space the conditional distribution $pr(y_o|y,\vartheta)$, the weights are given by the posterior distribution for the model parameters $pr(\vartheta|y)$ (Ribeiro, 1999):

$$pr(y_o|y) = \int pr(y_o, \vartheta|y)d\vartheta$$

$$= \int pr\,(y_o|y, \vartheta)pr(\vartheta|y)d\vartheta \qquad \text{(Eq. 4.12)}$$

In other words, the posterior distribution for the model ($pr(y_o|y)$ is an integration of the probability of the predicted (y_o) conditioned on data (y) and the model parameter (ϑ) combined with the prior probability ($pr(\vartheta|y)$). The Bayesian estimation is based on the predictive distribution while kriging methods normally estimate all model parameters to be inserted in the equation to predict at an unsampled location. The following distribution represents the kriging estimator:

$$(Y_o|Y) \sim pr(y_o|y, \hat{\vartheta}), \qquad \text{(Eq. 4.13)}$$

where pr($y_o|y$) is the density of predicted values given the data (y). The pr($y_o|y, \vartheta$) represents the probability density of y_0 given y (data values) and the $\hat{\vartheta}$ (predicted model parameters). When comparing the Equations 4.12 and 4.13 it's apparent that the Bayesian method (Eq. 4.11) provided a complete account of the likelihood surface and unlike kriging (Eq. 4.13) that accounts only for the maximum likelihood estimates of the covariance parameter.

Delisle (2006) applied the above described Bayesian approach to estimate soil carbon in selected fields of Mali. This study concluded that this particular Bayesian approach presented major improvement over the traditional geostatistics. Bayesian inference provided an elegant way to incorporate parameter uncertainty in the predictions.

4.4. Bayesian Maximum Entropy

Christakos (1990; 2000) argued that all kriging techniques presented a number of limitations. Classical geostatistics is mainly designed to fit into a purely inductive framework, where the experimental data are used to fit a mathematical model (variogram), which is then used to transform experimental data. The kriging weights depend only on the shape of the variogram, not on its global sill or any other factor multiplying the semivariogram or covariance model (Goovaerts, 1997). When a particular data point is screened by another its weight might become negative, yielding in some cases unacceptable results, such as values outside the data range or even negative (i.e. – 3 g kg^{-1} organic carbon). Even though the error variance of kriging methods relies on the covariance model and data configuration, studies have shown that kriging variance is independent of data values (Goovaerts, 1997).

Unlike modern spatiotemporal geostatistics, the kriging techniques do not have the capability to account for important knowledge such as expert opinion (Christakos et al., 2002). Christakos, (2000) concluded that kriging lacks "epistemic" content, and is concerned only with how to deal with data, rather than how to interpret and integrate the data to understand a process. Modern geostatistics, as defined by Christakos et al. (2002) appears as a normal construct and a matter of scientific progress of the techniques of spatial prediction.

4.3.2. Bayesian Maximum Entropy Theory

Bayesian Maximum Entropy (BME) is a novel method of geospatial analysis proposed by Christakos (1990; 2000). This method applies the concepts of Bayesian conditionalization and entropy (see Section 4.3), hence the acronym "BME" (Christakos, 1990). The entropy is a measure of the uncertainty associated with a random variable (Eq. 4.14). The amount of information about a random variable provided by the general knowledge ($\mathcal{G}$) and carried in the model and expressed by Eq 4.14 is called the entropy function. BME considers the physical knowledge and distinguishes between two main knowledge bases: (i) the general knowledge ($\mathcal{G}$) and (ii) the case specific knowledge (S). The general knowledge ($\mathcal{G}$) is considered as background information that characterizes a random variable (i.e. soil carbon). For instance, the statement about the mean and covariance of soil organic in a particular field constitute general knowledge.

The case specific knowledge in BME is defined in terms of hard data and soft data. In this study the hard data vector (x_{hard}) refers to a set of soil organic carbon measurements. The hard data are data collected during an experimental procedure of which there is a high degree of confidence on their accuracy. The soft data (x_{soft}) might include qualitative, categorical data, which imply a considerable amount of uncertainty. Depending on the overall objective of the study same data may be classified to a variety of categories (Christakos, 2000). Case specific knowledge may become available from a variety of sources, hence

encoding soft data is important to maximize the available knowledge. The type of hard and soft data depends on experimental and scientific considerations. By integrating a wide range of knowledge, BME can generate spatial-temporal maps that help our understanding of the underlying processes associated with the random variable of interest.

The BME annotation follows the following conventions, the Latin lower (x) denotes random variable and the Greek letters (χ) their realizations (data values). Uppercase Roman letters (X) denotes random fields.

The general principles of BME formalism can be summarized into four major stages: prior, metaprior, posterior and predictive.

4.3.2.1. The Prior Stage: General Knowledge

The general knowledge $\mathcal{G}$ may consist of statistical correlation functions (mean, covariances, variograms, multiple point moments, non-linear statistics) and scientific models (physical laws, scientific theories, logical principles) that can be collected on the site or elsewhere (Bogaert and D'Or, 2002b; Bogaert, 2002; Christakos, 2000; 2002). The main objective of this stage is to gather the prior probability function ($f_G(\chi_{map})$) given the general knowledge $\mathcal{G}$.

The prior distribution is computed by maximizing the entropy (in other words minimizing the amount of information added to the general knowledge) under the constraint that honors the prior characteristics of the general knowledge $\mathcal{G}$ (Bogaert and D'Or, 2002; Christakos, 2000; 2002). The entropy of a PDF is the

degree the uncertainty, consequently the narrower the PDF, the smaller the entropy (Christakos et al., 2002; Christakos, 2000).

The BME uses Shannon's[2] (Papoulis, 1991; Christakos, 2000) entropy (H) defined as follows:

$$H(\chi_{map}) = -\int \log f_G(\chi_{map}) f_G(\chi_{map})\, d\chi_{map}\,, \qquad \text{(Eq. 4.14)}$$

Where, $H(\chi_{map})$ is the expected information about the random variable, χ_{map} and $f_G(\chi_{map})$ is the probability density function (PDF) of the χ_{map}. The logarithm is used in order to provide the additivity characteristic for uncertainty (Schneider, 2000).

Incorporating the Lagrange multiplier μ_α into Eq. 4.13 results in the following equation:

$$\begin{aligned} L[f_G(\chi_{map})] = & -\int \log f_G(\chi_{map}) f_G(\chi_{map})\, d\chi_{map} \\ & -\textstyle\sum_\alpha \mu_\alpha \left[\int g_\alpha(\chi_{map}) f_G(\chi_{map})\, d\chi_{map}\right] \\ & -E[g_\alpha(\chi_{map})]\,, \end{aligned} \qquad \text{(Eq. 4.15)}$$

L is the loss function that denotes the minimum risk expected when maximizing the probability density function of the random variable χ_{map}. Where $g_\alpha(\chi_{map})$ is a set of functions (mean and covariance, higher-order moments) of χ_{map}. Eq. 4.14 basically allows the integration of the general knowledge $\mathcal{G}$. The maximum entropy solution can be obtained by setting the partial derivatives to zero and solving the system of equations with respect to μ_α (Christakos, 2000).

The prior probability density function:

[2] Claude Elwood Shannon, is considered as the "father of information theory"

$$f_G(\chi_{map}) = \frac{1}{Z}\exp\left(\sum_\alpha \mu_\alpha g_\alpha(\chi_{map})\right), \qquad \text{(Eq. 4.16)}$$

Z is the partition function.

$$Z = \int \exp\left\{\sum_\alpha \mu_\alpha g_\alpha(\chi_{map}\right\} d\chi_{map}, \qquad \text{(Eq. 4.17)}$$

The prior joint PDF $f_G(\chi_{map}) = \chi_K, f_G(\chi_{hard}, \chi_{soft})$ is general and ensures that all available knowledge prior to any measurement is properly taken into consideration (Bogaert and D'Or, 2002). In most cases the mean and covariance are used as the general knowledge G, representing the system of equations $g_\alpha(\chi_{map})$.

4.3.2.2. The Metaprior Stage: Site Specific Knowledge

The metaprior stage involves the collection and organization of specific knowledge in appropriate quantitative forms that can be incorporated into BME formulation in a posterior stage. The specific knowledge considers data collected from the site. The case specific knowledge (S) is comprised of all relevant data collected at the study site. Such data are further characterized into two groups (hard and soft data) and are expressed as follows:

$$S: \chi_{data} = (\chi_{hard}, \chi_{soft}). \qquad \text{(Eq. 4.18)}$$

The hard data at a particular space (x, y) and time (t_i) represent measurements considered accurate. The hard data set is represents as follows:

$$S: \chi_{hard} = (\chi_1, \ldots, \chi_{mh}), \qquad \text{(Eq. 4.19)}$$

The χ_1 includes single valued data with a space and time component (i.e. soil organic carbon collected at each farm). Examples of soft data (χ_{soft}=($\chi_{mh+1,\ldots}$, χ_m) at points p_i (i=mh+1) may include uncertain evidence, indirect measurements, fuzzy information, and intuitive information, about the field of study (Christakos, 2000; Christakos et al., 2002; Bogaert, 2003; Bogaert and Wibrin, 2004; Bogaert, 2004).

$$S: \chi_{soft} = (\chi_{mh+1}, \ldots, \chi_m), \qquad \text{(Eq. 4.20)}$$

The soft data (Eq. 4.21) can be expressed in terms of intervals of possible values x_i for i=mh+1,... m:

$$S: x_{soft} \in I, \; x_i \in I_i = [l_i, u_i], \qquad \text{(Eq. 4.21)}$$

Where I is the domain of x_{soft}, l_i, is the lower and u_i the upper limits of the intervals I_i. The section 5.2.4.2 discusses the encoding of clay and carbon (20-40 cm) content into intervals classes of soft data. Other forms of expressing soft data in terms of interval probabilities is discussed in Christakos et al. (2000).

$$S: P_s [x_{soft} \in I] \in [0, 1], \qquad \text{(Eq. 4.22)}$$

4.3.2.3. The Posterior Stage: Integration of Specific Knowledge

The third stage of BME consists of the integration of general knowledge $\mathcal{G}$ and site specific knowledge S to produce the posterior or the combined knowledge $\mathcal{K}$ ($\mathcal{K}= \mathcal{G} \cup S$). The covariance matrix associated with the vector of random variable x_{map} is written as:

$$C_{map} = (x_{map} - m_{map})(x_{map} - m_{map}) = \begin{bmatrix} C_{x(p1,p1)} & \cdots & C_{x(p1,pk)} \\ \vdots & \ddots & \vdots \\ C_{x(pk,p1)} & \cdots & C_{x(pk,pk)} \end{bmatrix} \quad \text{(Eq. 4.23)}$$

Where C_{map} is the covariance matrix of the x_{map}, and m_{map} is the mean. The prior PDF $f_G(\chi_{map})$ associated with the general knowledge is given by ϕ:

$$\phi\ (x, \bar{x}, c) = (2\pi)^{-n/2}|C|^{-1/2}\exp[-(x - \overline{x})^T C^{-1}(x - \overline{x})/2 \quad \text{(Eq. 4.24)}$$

The n-point Gaussian PDF (ϕ) of the random vector x has $\bar{x}$ mean and covariance matrix denoted as C. The posterior PDF is obtained by inserting the Gaussian PDF, $f_G(\chi_{\text{map}}) = \phi(\chi_{\text{map}}, 0, C_{\text{map}})$ into the Equation (4.24) assuming that $m_{\text{map}}= 0$.

The BME posterior integration (PDF) is given by:

$$f_k\left(\chi_k \middle| \chi_{hard}, f_s(\chi_{soft})\right) \quad \text{(Eq. 4.25)}$$

$$= (A^{-1}) f_S^{(k)}(\chi_k)\phi(\chi_{kh}, 0, C_{kh,kh}) \int d\chi_{soft}\, f_s(\chi_{soft})\phi(\chi_{soft}, B_{soft|kh}\chi_{kh}, C_{soft,kh})$$

Where, χ_k is the value at point k, the $\chi_{kh} = [\chi_k, \chi_{hard}]$ represent the hard data points, A= $\phi(\chi_{hard}, 0, C_{hard,hard})$ is a normalization constant. The terms $B_{\text{soft|kh}}$ and $C_{\text{soft|kh}}$ are defined as the conditional covariance vector and matrix, respectively. The equation above describes the PDF, $(f_k(\chi_k|\chi_{hard}, f_s(\chi_{soft}))$, at every estimation point

p_k (Christakos, 2000; Bogaert and D'Or 2002b; Christakos et al., 2002). The probability distribution (Eq. 4.25) of an unknown value, χ_k, at estimation location is conditional on all hard data values and all soft data probability distributions. Equation 4.25 provides a complete update of the hard and soft data and the general knowledge denoted by the Gaussian PDF (ϕ) with mean (m) and covariance (C) at each estimation point (χ_k).

4.3.2.4. The Predictive Stage:

The predictive stage provides the moments of the posterior PDF. The mean ($\overline{x_{k|k}}$), is the estimate of the PDF, the error variance ($\sigma^2_{k|K}$) provides a measure of the uncertainty associated with the estimated value. The third order moment of the posterior PDF can also be calculated providing the skewness of the random field at the estimation point (Serre, 1999).

4.3.2.4.1. BME Mode

The mode estimate $\hat{x}_k$ is the most probable value occurring within a spatial temporal random field at the estimation point p_k (Christakos, 2000, 2002; Serre, 2001). The BME mode is obtained by maximizing the posterior PDF (Eq. 4.25).

$$\hat{\chi}_{K,\mathrm{mode}} : \max_{\chi_k} f_K(\chi_k) \qquad \text{(Eq. 4.26)}$$

Where mode $f_K(\chi_k)$ is the posterior PDF, Equation 4.25. This function provides the most probable values of the random field at the estimation location.

Additional information can also be derived from the statistical moments of the BME posterior PDF, such as the mean, the variance and the skewness. In this way the mode of the predicted distribution at the estimation point is obtained and can be used to map the predicted values considering with hard, soft and general information.

4.3.2.4.2. Mean of BME Posterior PDF

The "BME mean" estimate is calculated based on the posterior PDF (Christakos, 2000, 2001, Serre, 1999) and can be written as follows:

$$\overline{x_{k|k}} = (A'^{-1}) \int d\chi_{ks}\, f_S^{(ks)}(\chi_{ks})\phi(\chi_{ks|h}, B_{ks|h}\chi_{hard}, C_{ks|hard}) \qquad \text{(Eq. 4.27)}$$

Where the subscripts *h*, *S*, and *K* denote hard data points, soft points and estimation points, respectively.

$A' = \int d\,\chi_{ks} f_S^{(ks)}(\chi_{ks})\phi(\chi_{ks}, B_{ks|hard}\chi_{hard}, C_{ks|hard})$ is the normalization parameter. The BME PDF (Eq. 4.25) is also used to obtain an assessment of the uncertainty associated with the estimates.

4.3.2.4.3. Variance of BME Posterior PDF

The BME associated error variance is computed as follows at every estimation grid point. Equation 4.28 shows the error variance considering hard and soft data.

$$\sigma_{k|K}^2 = A'^{-1} \int d\,\chi_{ks}(\chi_k - \bar{x}_{k|K})^2 f_S^{(ks)}(\chi_{ks})\phi(\chi_{ks}, B_{ks|hard}\chi_{hard}, C_{ks|hard}) \qquad \text{(Eq. 4.28)}$$

Where A′ presented above is the normalization parameter, $B_{ks|hard}$ and $C_{ks|hard}$) are the conditional covariance vector and matrix, respectively. χ_{ks} , represents the vector estimates at each location. BME $\sigma^2_{k|K}$ is hard data dependent but also it integrates the PDF of the soft data ($f_S^{(ks)}$) while considering the covariance information between pairs of points.

Traditionally, the geostatistical methods provide the variance of the estimation as a measure of uncertainty (Goovaerts, 1997, Webster and Oliver, 2004). The variance can be an adequate assessment of the estimation error, when the shape of posterior PDF is Gaussian (Serre, 1999). Besides the variance the BME also provides a more sophisticated and accurate assessment of the estimation error by means of the posterior PDF $f_K(\chi_k)$ derived on the basis of the total knowledge (Christakos et al., 2002; D'Or et al., 2001; Bogaert et al., 1999; Serre, 1999). The PDF allows the calculation of confidence intervals, which is a more appropriate description of the uncertainty associated with mapping of regionalized variable based on estimation points.

CHAPTER 5: QUANTIFYING AND MAPPING SOIL ORGANIC CARBON USING BAYESIAN MAXIMUM ENTROPY

5.1. Introduction

Soil organic carbon quantification is increasingly gaining relevance in the context of carbon sequestration as a mitigation effort to reduce CO_2 concentration in the atmosphere. In Sahel the overall potential of soil carbon sequestration still remains to be accurately estimated. According to Ringius (2002) the Sahel soil stores about 7.5 to 9.9 t C ha^{-1} (2.5 – 3.3 g C kg^{-1}) in the upper 20 cm layer. In non-degraded savanna, with higher plant productivity and organic matter, the soil carbon levels range from 7.5 – 18 t C ha^{-1} (Ringius, 2002). Soil organic carbon buildup is important in the semi-arid and sub-humid African region where land degradation and accelerated desertification is the main cause of food shortage and extreme poverty (Vagen et al., 2005). Despite the harsh climatic conditions that limit the productivity of these ecosystems, the Sahelian region still has the potential for carbon dioxide sequestration and storage (Woomer et al., 2004; Traoré, 2003; Manlay, 2000; Doumbia et al., in review).

Mapping soil properties has long been an important objective of soil scientists and environmentalists. It's one of the most effective ways of conveying information about the spatial representation of a particular natural phenomenon. However, quantifying and mapping soil carbon is not an easy task, in part due to

uncertainties and errors associated with the measurements and methodological procedures that often result in an incomplete or deficient accounting of soil organic carbon (Vagen et al., 2004; Houghton, 2003; Muller and Pierce, 2003; Chen et al., 2000).

Quantifying and mapping soil organic carbon is a useful tool aimed at providing an accurate accounting of CO_2 equivalent stored in the soil system. The IPCC (1997) methodological guidelines discussed the importance of soil carbon accounting. In this guideline total soil carbon is calculated as the product of soil carbon measurements and the soil weight per unit of area.

Most carbon inventory reports at the field and regional level have relied on average value of carbon measurements (Harvey, 2004; Heath et al., 2000). In general, the spatial variability is not considered and changes over time are invariably not taken into account. The spatial and temporal monitoring and verification of the amount of soil carbon sequestered provide an assessment of the permanence of the carbon storage and its potential impact on the global carbon cycle.

5.1.1. Prior regional assessment of soil carbon

Measuring soil organic carbon in terrestrial ecosystems is labor-intensive and a large scale assessment of carbon inventory is cost prohibitive when considering the traditional methods that involve analyzing soil sampling and measuring tree diameter (Yost et al., 2002).

The study of Ardö and Olsson (2003) presented a spatially explicit approach to quantify soil organic carbon in the semi-arid region of Sudan. The study integrated thematic layers extracted from a geographical information system (GIS) with CENTURY model to estimate historical and future soil organic carbon for an area of 262,000 km^2. The CENTURY is a "lumped parameter" ecosystem model initially developed to study soil organic matter dynamics in grassland soils (Parton et al., 1988), but has been modified to also accommodate agricultural and savanna systems (Ardö and Olsson, 2003). CENTURY model requires several parameters (i.e. climate, soil texture, litter N, lignin content, just to name a few), that are often difficult to obtain (Ardö and Olsson, 2003). Furthermore, this study relied on the FAO soil map, which does not account for the variability within the soil units. Hence, the uncertainties associated with parameterization of the model and the data are certain to occur but difficult to quantify. Despite the GIS and time component used to estimate soil organic carbon at each site, the Ardö and Olsson (2001) approach did not consider soil organic carbon as a regional variable nor did it fully account for the spatial and temporal distribution of soil carbon.

Yost et al. (1993) offered a procedure to upscale soil carbon estimates based on finite data set. The study relied on kriging techniques to interpolate and extrapolate soil carbon throughout the Pearl Harbor watershed in Hawaii. Based on data collected in Mali, Yost et al. (2002) also presented a geospspatial protocol for carbon accounting and mapping in selected sites in Mali. Considering

the systematic variability associated with soil units, this methodological approach also proposed the use of remotely sensed imagery for hierarchical stratification of sampling efforts and the up scaling issues as they impact the prediction error at a regional scale (Yost et al., 2002). Delisle (2006) expanded the geospatial approach of soil carbon quantification by working with Bayesian inference to better handle the variogram modeling and compared the spatial and non spatial methods(i.e. sample mean) of carbon accounting.

Woomer et al. (2004) estimated carbon stock on the basis of historical land use change within different ecological zones in Senegal. At each ecological region the above and below ground carbon stock was estimated for the years 1965, 1985 and 2000. The study by Woomer et al. (2004) attempted to address the spatial variability by clustering the samples of eight ecological zones and relying on 67 data sample locations. The spatial distribution of the sample locations are critical to evaluate whether a spatially explicit approach is able to capture large scale variability of soil carbon of an entire region. One limitation of the approach presented by Woomer et al. (2004) is that there is a lack of uncertainty assessment associated with spatial and temporal estimates and the use of remotely sensed imagery.

Several studies have demonstrated that the levels of soil carbon in West Africa soils are extremely low (Kablan et al., 2008; Traoré, 2003; Manlay, 2000). Although carbon accounting cannot be seen as totally new, most previous assessments relied on methods which largely neglect the spatial and temporal

variability of soil organic carbon across the landscape. In almost all reported cases the authors relied on finite set of samples and using non-spatial explicit approach. Consequently, the main objective of this chapter is to propose a practical method of carbon accounting using Bayesian Maximum Entropy (BME) (described in Chapter 4).

Given the fact that the available geostatistical techniques rely on a finite set of measured data and cannot integrate prior information, the reliability of the predictions of classical methods is exclusively dependent on the data set and its configuration (Goovaerts, 1997; Serre, 1999). In order for the classical geostatistical approach to accurately capture the spatial variability of a particular variable of interest it typically requires a huge, expensive sampling effort. The BME approach is based on the Bayesian axiom that the posterior probability can be improved if prior knowledge is taken into account (Serre, 1999).

Therefore, we hypothesize that by considering the measured data and the site specific soft data the accuracy and precision of soil organic carbon estimates improve significantly. Consequently, the geospatial prediction of soil organic carbon at the field and regional level can effectively be mapped using BME approach.

5.2. Materials and Methods

5.2.1. The Sampling Procedure

Soil samples were collected in 2000, 2002, 2004 and 2006 in most of the project sites as indicated in Table 5.1 (see also Figures 5.1, 6.1-6.3). All soil samples were georeferenced using a GPS (global position system) receiver with submeter accuracy. The GPS measurements were differentially corrected in real time using the Trimble RTCM_base station[3] mounted at IER-Sotuba (Bamako) and at Koutiala. Measurements without active base station support were post processed to ensure maximum accuracy and precision. Map projection and datum (UTM, Zone 29, WGS-84) were selected to facilitate field work data collection and data integration with locally available maps. The spacing between samples was considered in the field, aiming at understanding the process while capturing both long and short range spatial variance of soil carbon measurements. Cluster sampling at stratified locations (soil, landuse and relief form) was performed to estimate the short-range variation (see Figures 5.1).

[3] RTCM standards are used internationally for Differential Global Navigation Satellite Systems and Electronic Chart Systems. Trimble is a brand name of GPS receivers and base stations.
Disclaimer: The mention of a specific vendor or brand in no way implies endorsement or recommendation by the University of Hawai'i.

5.2.2. Laboratory procedure

Soil sampling and subsequent laboratory analysis is still by far the most reliable method of assessing the soil organic carbon. Soil carbon was analyzed at the University of Hawai'i Agricultural Diagnostic Service Center (UH-ADSC) using the total carbon dry combustion method developed by Nelson and Sommers (1982). In the dry combustion method, organic carbon and other carbonates minerals are oxidized in furnace in a stream of purified oxygen (O_2) (Nelson and Sommers, 1982). All organic carbon measurements were conducted using high-temperature resistance furnace (temperature at 1350 °C). With high temperature furnaces the organic and elemental carbon (C) is oxidized to CO_2 by O_2 without any special help of catalyst. The soil was ground and sieved to 100 mesh (0.149 mm opening / 0.0059 in). The soil carbon measurements were calculated in SI units, (grams of carbon per kilogram of soil (g/kg)).

5.2.3. Data set

In order to examine the potential of BME to quantify and map the spatial and temporal distribution of soil carbon, data sets of soil carbon were collected at the Fansirakoro (Baracoro Ballo), Siguidolo (Zan Diarra), Oumarbougou (Mory Konate) and Sikasso (Yaya Diassa) sites. Multiple datasets based on the sampling period were used in this study (See Table 5.1).

Table 5.1. Data sets from Mali used to quantify soil organic carbon.

Sites/Ecological Regions	Rainfall (mm)	Area(ha)	Sampling Period	N. of Samples*
Fansirakoro	800 - 1000	4.77	2002; 2004; 2006	166
Siguidolo*	600 - 800	18	2000; 2004; 2006	132
Oumarbougou*	800 - 1000	20	2000; 2002; 2004; 2006	286
Sikasso*	900 - 1100	44	2000; 2004; 2006	308

* includes samples from 0-20 cm and 20-40 cm depth.

Fansirakoro / Baracoro Ballo

The Baracoro Ballo field area was 4.77 ha, and was located in the Region of Koulikoro, north of the capital Bamako. The central coordinates of this field were 8° 5'16.64" W and 12° 57' 22.86"N (See Figure 5.1).

The soil samples that comprise the Fansirakoro/Baracoro Ballo data sets were collected in 2002, 2004 and 2006. A total of 166 soil samples collected at two depths (0-20 and 20-40 cm) were analyzed for carbon and clay content. In 2002 samples were collected mostly from the center of the field.

The 2004 and 2006 sampling covered the entire field with a clear objective to obtain well-spaced sampling. At each distinct soil unit, samples were collected at varying distances. Short distances sampling provide a better assessment of the error variance component. The spatial coverage of the field improved with the 2006 sampling (See Figure 5.1).

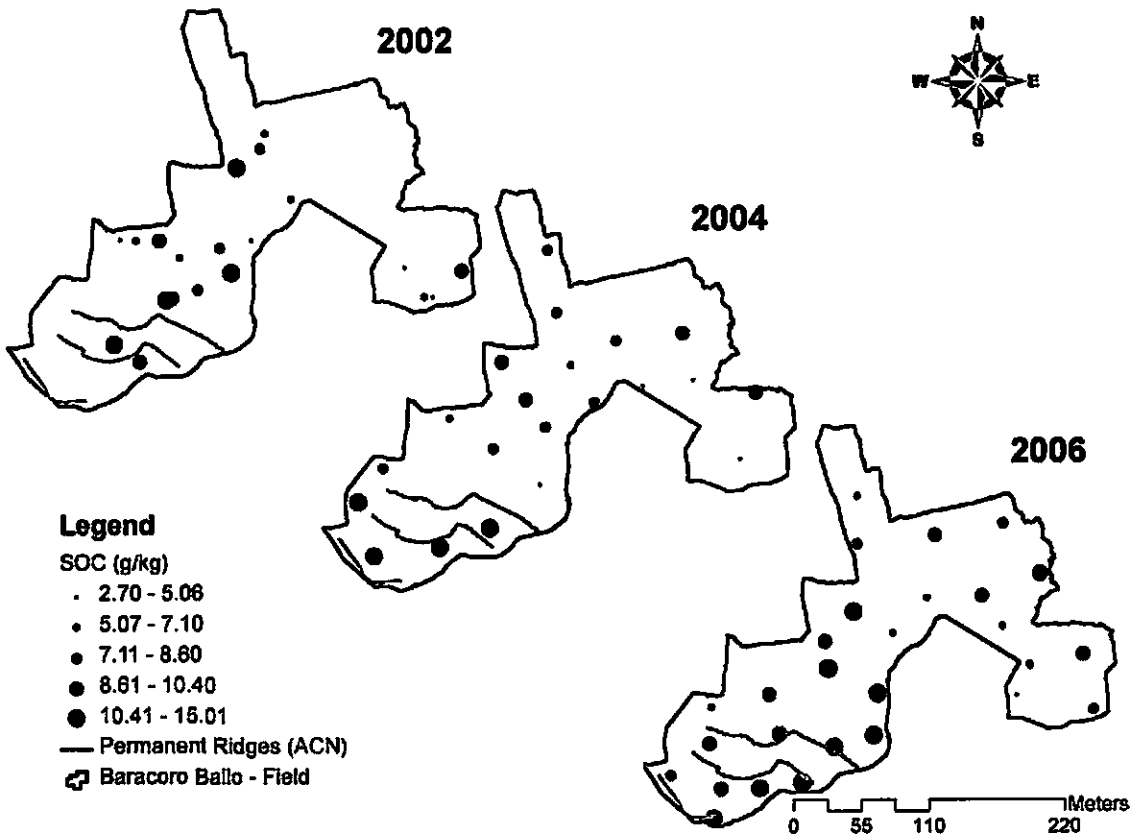


Figure 5.1. Soil carbon sampling in Baracoro Ballo's Field, 2002, 2004 and 2006. The spatial distribution and variation of soil organic carbon content (SOC) is illustrated. These results are based on data from 0 -20 cm depth. Note that the sampling layout varied for different time periods. Each sample point assumes a soil organic value, the spatial (x,y) and temporal (t) components.

5.2.4. Bayesian Maximum Entropy Approach

The mathematical formulation of BME approach was discussed in chapter 4. The numerical implementations of the BME method produce accurate estimates for space-time mapping and assessment of the uncertainties (Christakos and Serre, 2000). In BME the "total" physical knowledge ($\mathcal{K}$) may consist of two main knowledge bases: the general knowledge ($\mathcal{G}$) and the case specific or site knowledge ($\mathcal{S}$), where $\mathcal{K} = \mathcal{G} \cup \mathcal{S}$. Any information important to understanding the processes associated with the primary variable (i.e. hard data) that can be expressed in terms of intervals values (aggregated in discrete categories), probability statements, experts' opinions is referred to as soft data. The term soft data (as opposed to hard data) refers to the uncertainties associated with the use of these types of observations. The BME method maximizes the informative model f_k (Eq. 4.24) by integrating the prior/general knowledge (mean covariance), the hard data (measured) and the soft data (interval data, probabilistic) to generate a posterior probability distribution.

In BME the distribution of any random variable is represented in terms of spatio-temporal random field (S/TRF) $X_{(p)}$ (Christakos, 1992). Mathematically S/TRF is the collection of all possible realizations ($\chi_1 \ldots \chi_n$). In this study a realization is represented by a space/time sample of soil organic carbon (C). Each realization assumes a soil carbon value in space (s) and time (t) (Figure 5.1).

5.2.4.1. General Knowledge

The general knowledge $\mathcal{G}$ is the background knowledge that according to Christakos et al. (2002) and Serre (1999) might include physical laws of science and statistical moments. It is considered "general" because it is vague enough to characterize a large class of field data. The mean and covariance of the spatial temporal random field (S/TRF) are considered part of the general knowledge since they provide an indication of the statistical moments of the random field.

Mathematically, general knowledge ($\mathcal{G}$) can be expressed as follows (Serre, 1999; Christakos et al., 2002):

$$\mathcal{G} = \int d\chi_{map}\, g_\alpha(\chi_{map}) f_g(\chi_{map}) = \overline{g_\alpha}\,,\ a=1,\ldots,N_a \qquad \text{(Eq. 5.1)}$$

Where χ_{map} includes all the measured data, $f_g(\chi_{map})$ is the corresponding probability density function and g_α are the expectations of the fields involved. By convention $g_0 = 0$ and $\overline{g_1} = 1$. The knowledge of the statistical moments of order q of X(p) at location p_i can be calculated by letting $g_\alpha(\chi_i) = \chi_i^q$. The covariance and any other higher order statistical moments can be computed with the appropriate g_α functions (Serre, 1999).

5.2.4.2. Hard Data – Soil Organic Carbon

The soil organic carbon measured from top layer (0 – 20 cm) was processed as hard data. Data points were characterized by their measured value (i.e. carbon, g C kg^{-1} of soil), as well as the spatial (x, y) and temporal (t) components. Each data point is referred as a realization and a collection of realizations is the spatiotemporal random field (S/TRF). Figure 5.1 illustrates the locations of all hard data collected over a period of time (2000 – 2006) at Baracoro Ballo's field. The temporal lag for this dataset was 2 years. All data were compiled in GeoEAS format, which requires a title in the first line, the number of column in the second line followed by data lines (Serre, 1999; Christakos et al., 2002) (See Appendix I). This particular format is also used for BME read/write functions.

5.2.4.3. Soft data – Clay and Soil Organic Carbon

Soft data can be extracted from various sources (Christakos et al., 2002; Christakos and Serre, 2002). However, the selection of case specific soft data is important for the overall knowledge acquisition. Two important variables related to surface soil organic carbon buildup were modified to be used as soft data: (I) The clay content, measured as a weight fraction, and (ii) soil organic carbon content from samples collected at 20-40 cm depth. These variables were also used as covariables in cokriging technique.

The clay and carbon data were grouped into discrete classes, which permitted testing BME capability to integrate soft information. These procedures are illustrated in Table 5.2 below. All measured clay data in Baracoro Ballo's field were grouped in to 5 classes between 0 to 100% (Table 5.2).

Table 5.2. Transforming measured clay into soft data, based on 2006 dataset. Note that in BME interval data function only coordinates (x,y) and class boundaries are taken into account.

X (m)	Y (m)	SOC (g kg^{-1})	Clay (%)	Clay Classes	
				Low Boundary	Upper Boundary
598932.404	1432607.292	3.6	17.2	0	20
598932.225	1432569.337	5.9	23.28	20.01	40
598995.380	1432576.386	7.1	20.88	20.01	40
599050.201	1432586.005	6.6	19.44	0	20
599080.296	1432545.915	8.3	20.88	20.01	40

Table 5.2 shows only a small portion of the data set composed of 42 data points. Only 28 points were used as soft data (see Figure 5.12), given the fact most were cluster samples with similar amounts of clay. The same approach was used for carbon (0-40 cm), where carbon values (g kg^{-1}) were grouped into 6 classes (0-2; 2.01-4; 4.01-6; 6.01-8; 8.01-10; 10.01-20) of soil organic carbon. For each soft data variable a soft interval data file was created indicating the coordinates of the each soft data point and the upper and lower boundary of each class (see Appendix II).

Soil clay is often positively correlated with soil carbon, that is, soil with high clay tends to have high carbon (Bohan et al., 2001). Several authors have reported that clay particles effectively prevent or reduce the decomposition of soil carbon (Bohan et al., 2001; Brady and Weil, 2002). In sub-Saharan

agricultural production systems most cultural practices are conducted manually and without heavy mechanization, therefore, it's reasonable to assume that carbon content at greater depths is less susceptible to changes due to management practices. In other words, the amount of soil carbon at 20-40 cm is more influenced by other soil development processes than by anthropogenic actions. By grouping soil clay information into classes (interval soft data) we were able to test the capability of the BME approach with less accurate ranges of measurements.

Figure 5.2 illustrates the use of BME approach to estimate soil organic carbon in agricultural systems of Mali. The mathematical formulation that involves the integration general knowledge and site specific data to generate the posterior PDF is expressed by the Equation 4.24 (see Section 4.3.2.3). This approach allows the mapping of soil carbon based on measured data and additional site information that can help to improve the overall prediction.

The BME approach describes the soil carbon map by means of its posterior PDF (Eq. 4.24). The PDF's allow the extraction of relevant statistical moments such as, mode, mean, median and other portions of the pdf (see Section 4.3.2.4).

In addition to estimates of error variance normally associated with most geostatistical predictions, BME provides the estimation accuracy of the PDF, including the confidence intervals. BME variance of the posterior PDF (Eq. 4.27, section 4.3.2.4.3) provides the uncertainty assessment based on specific set of

values considered. In other words, both the prediction and error maps can be produced using the BME approach.

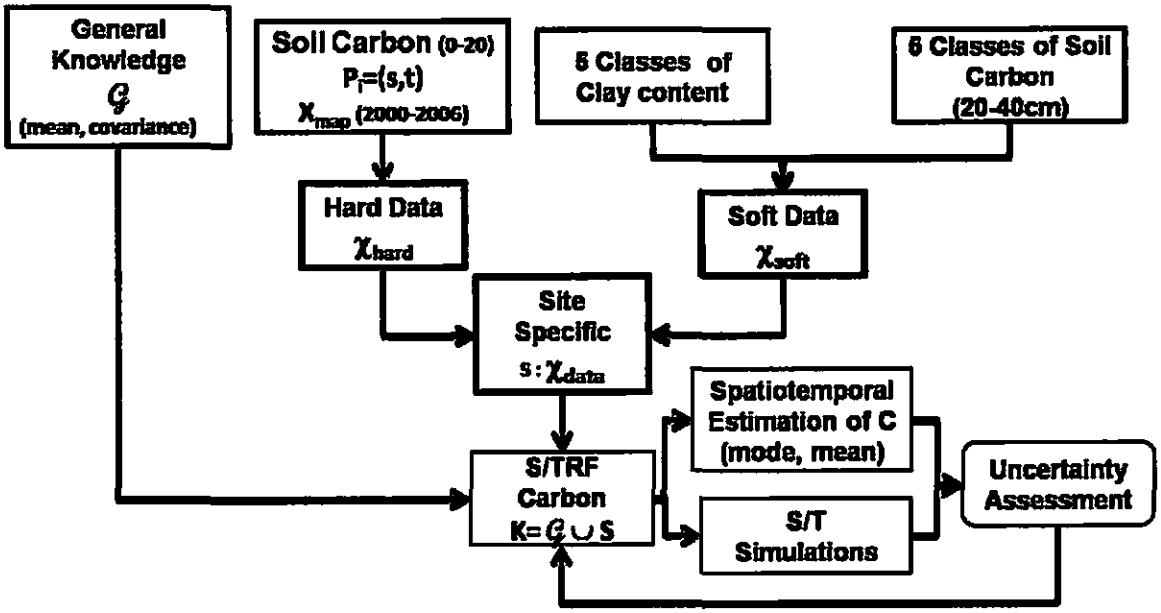


Figure 5.2. Proposed framework for estimating spatial and temporal distribution of soil organic carbon. BME considers all site specific information to predict the estimates of the variable of interest (ex. soil organic carbon).

5.2.5. Kriging and Cokriging Approach

Figure 5.3 demonstrates the workflow for mapping soil carbon using classical geostatistical tools. Both kriging and cokriging methods were discussed in Chapter 4. The primary variable was carbon, 0-20 cm, for kriging. Additional covariables (clay, carbon) were considered for cokriging (See Figure 5.3). In this study several cokriging combinations were tested to determine the value of additional variables in estimating soil organic carbon.

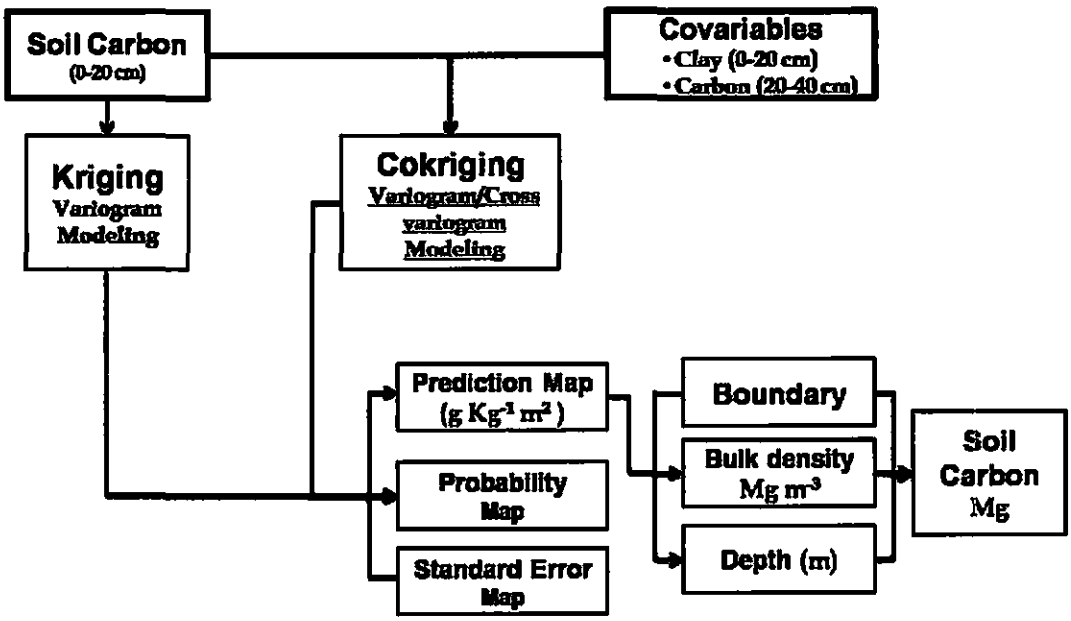


Figure 5.3. Proposed framework for quantifying and mapping soil organic carbon, using classical geostatiscal methods, kriging and cokriging.

Using a cross correlation between the primary and secondary variables, cokriging can increase the prediction accuracy provided that the relationship between the primary and secondary variable is strong. The prediction maps from BMELib, in Matlab, were exported to ArcGIS 9.2 (ESRI, 2006) for further calculation and mapping.

5.2.6. Analysis Summary

The following steps were carried out in this analysis:

1. Datasets of soil organic carbon for different sampling periods were characterized. Exploratory statistical analysis were performed and the normality assumption tested;
2. Hard Data – All measured soil organic carbon from surface layer (0-20 cm) were compiled;
3. Soft interval classes of clay content (0-20 cm) and carbon (20-40 cm) were constructed. The lower and upper boundary for each class were assigned to corresponding individual values of clay and carbon;
4. A space-time covariance model of soil carbon was constructed;
5. The BME posterior distribution (PDF) at each estimation grid intersection was calculated.
6. The BME interval mode function was used to compute the most likely values at a particular node, and the BME interval PDF was computed at a single estimation location provided the complete posterior PDF function using both hard and soft data;
7. The predictive power of BME, kriging and cokriging were compared in terms of bias and mean square error based;
8. Predict soil carbon change over time;
9. Maps were produced to display carbon tonnage in a field.

5.3. Results

5.3.1. Baracoro Ballo Case Study

5.3.1.1. Exploratory Statistical Analysis

Exploratory analysis of soil carbon data at each site was performed. This was a preliminary step to the spatial and temporal analysis of the data. Considering that statistical analysis is most efficient when variables are normally distributed, a normality test was also conducted and where appropriate, the data were transformed. Basic statistics, including histograms and their respective probability distribution functions (PDF) and cumulative distribution functions (cdf) were generated for each variable in the dataset.

Both short and long range spatial variability was calculated and modeled and the covariance structure of the data and the variograms for all sites were plotted. Short distance sampling helped capture the short range spatial variability of soil carbon. The spatial autocorrelation of soil organic carbon variability was analyzed.

Table 5.3 gives the summary statistics of carbon (g kg^{-1}) for Baracoro Ballo's farm. The sample mean between different sampling periods (2000-2006) did not vary significantly (Table 5.3). Even though the difference between surface (0-20 cm) and subsurface carbon (20-40 cm) was not significant, the latter was usually less.

Table 5.3. Summary statistics of measured soil organic carbon at Fansirakoro/Baracoro Ballo.

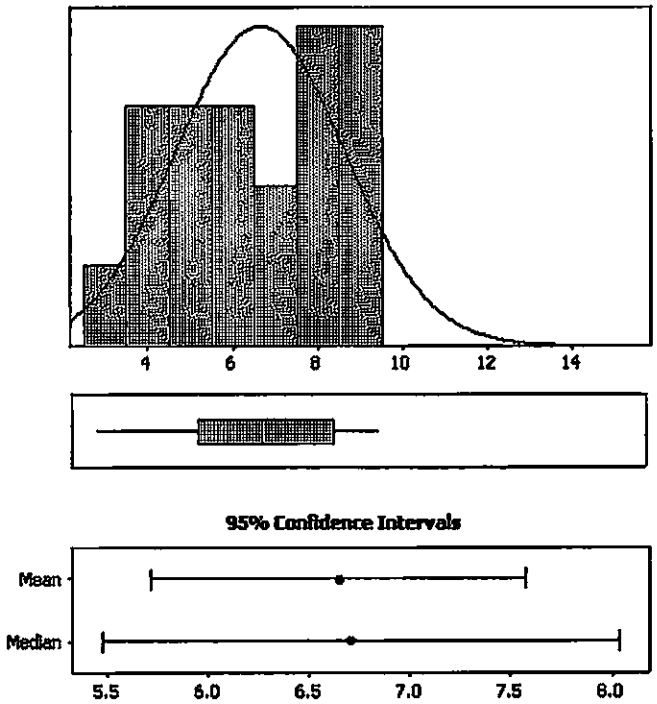
	N	Mean	Median	Std. Dev	σ^2	Skewness	Kurtosis
Fansirakoro/Baracoro Ballo Carbon (g kg^{-1}) [0-20cm]							
2002	20	6.65	6.71	1.99	3.90	-0.20	-0.90
2004	21	7.11	6.29	3.22	10.30	1.15	0.64
2006	42	7.56	8.00	2.16	4.68	-0.44	-0.60
Fansirakoro/Baracoro Ballo Carbon (g kg^{-1}) [20-40cm]							
2002	20	5.19	5.01	1.61	2.56	0.54	0.34
2004	21	6.33	5.88	2.53	6.37	0.65	-0.49
2006	42	5.19	4.70	1.75	3.06	1.08	0.88

N is the number of observations, Std. Dev is the standard deviation, and σ^2 is the variance. Skewness is a measure of symmetry while Kurtosis characterizes the distribution.

The distribution of surface (0-20 cm) soil carbon collected in 2002 was approximately normal. The initial statistical analysis did not show a pronounced skewness (Table 5.3, Figure 5.4). The kurtosis indicated that the variance was due to frequent and small deviations. A logarithm transformation did not yield significant improvement, therefore all analysis were conducted using non-transformed data.

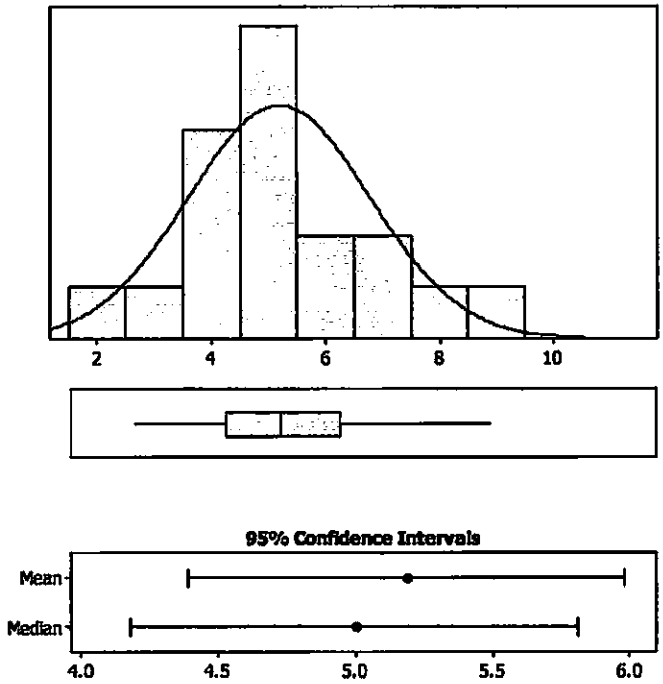
Overall soil organic carbon was low and the maximum carbon content did not exceed 10 g kg^{-1}. The interquartile range, shown below the graph in Figure 5.4 did not identify any potential outliers. Looking at the distribution, the box plot and the histogram of carbon in 2002, the observations seem to be evenly distributed around the mean.

Year = 2002 [0-20 cm]



Anderson-Darling Normality Test	
A-Squared	0.23
P-Value	0.784
Mean	6.6450
StDev	1.9937
Variance	3.9749
Skewness	-0.195154
Kurtosis	-0.901646
N	20
Minimum	2.7300
1st Quartile	5.1625
Median	6.7050
3rd Quartile	8.3800
Maximum	9.4800
95% Confidence Interval for Mean	
5.7119	7.5781
95% Confidence Interval for Median	
5.4747	8.0353
95% Confidence Interval for StDev	
1.5162	2.9120

Year = 2002 [20-40 cm]



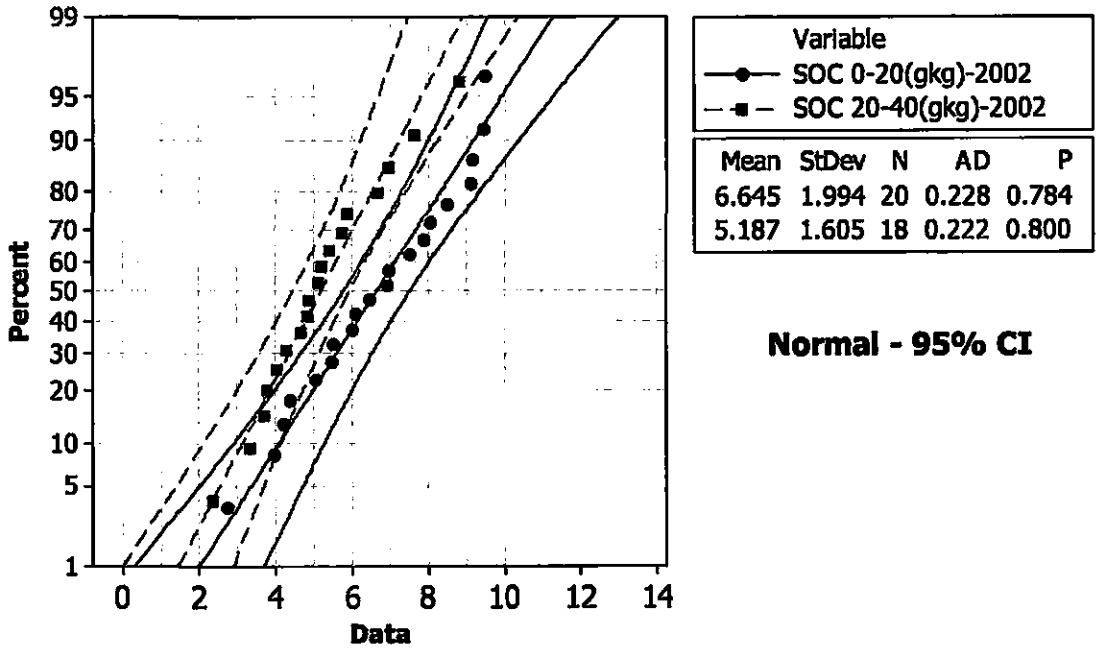
Anderson-Darling Normality Test	
A-Squared	0.22
P-Value	0.800
Mean	5.1872
StDev	1.6047
Variance	2.5751
Skewness	0.542107
Kurtosis	0.337949
N	18
Minimum	2.3500
1st Quartile	3.9875
Median	5.0050
3rd Quartile	6.0775
Maximum	8.8100
95% Confidence Interval for Mean	
4.3892	5.9852
95% Confidence Interval for Median	
4.1791	5.8130
95% Confidence Interval for StDev	
1.2042	2.4057

Figure 5.4. Statistical summary of soil organic carbon collected from Baracoro Ballo's field – 2002 (0-20 cm [top] and 20-40 cm [botton]).

The mean of carbon in 2002 (depth 0-20cm) was 6.645 (95% confidence intervals of 5.47 and 7.58). The standard deviation was 1.99 (95% confidence intervals of 1.52 and 2.91 g kg^{-1}). Using a significance level of 0.05, the Anderson-Darling Normality Test (A-Squared = 0.23, P-Value = 0.784) indicates that the carbon data follow a normal distribution (See Figure 5.4). Similar results were observed for the 20-40 cm data. The Anderson-Darling normality test (A-Squared=0.22 and P-value= 0.8) also confirmed that the skewness was not significant. The mean carbon for the lower depth was 5.19 g kg^{-1} (95% confidence intervals of 4.39 and 5.98 g kg^{-1}).

However, Figure 5.4 shows a slightly higher percentage of soil organic carbon was found in the top layer of the soil. These results are consistent with the well documented fact that carbon content tends to decrease with soil depth (Brady and Weil, 2002; Coleman et al., 1989).

The relationship between carbon in the top 20 cm and the subsurface (20-40 cm) was also examined to determine the possibility of using the subsurface as a collocated covariable for cokriging. The results presented in Figure 5.5 indicate that in the 2002 data set the correlation was positive and significant (r=0.573, P-Value=0.013). The plotted probability points follow the fitted lines closely, and the p-values for each Anderson-Darling (AD) test are greater than 0.10, suggesting that the data were normally distributed.



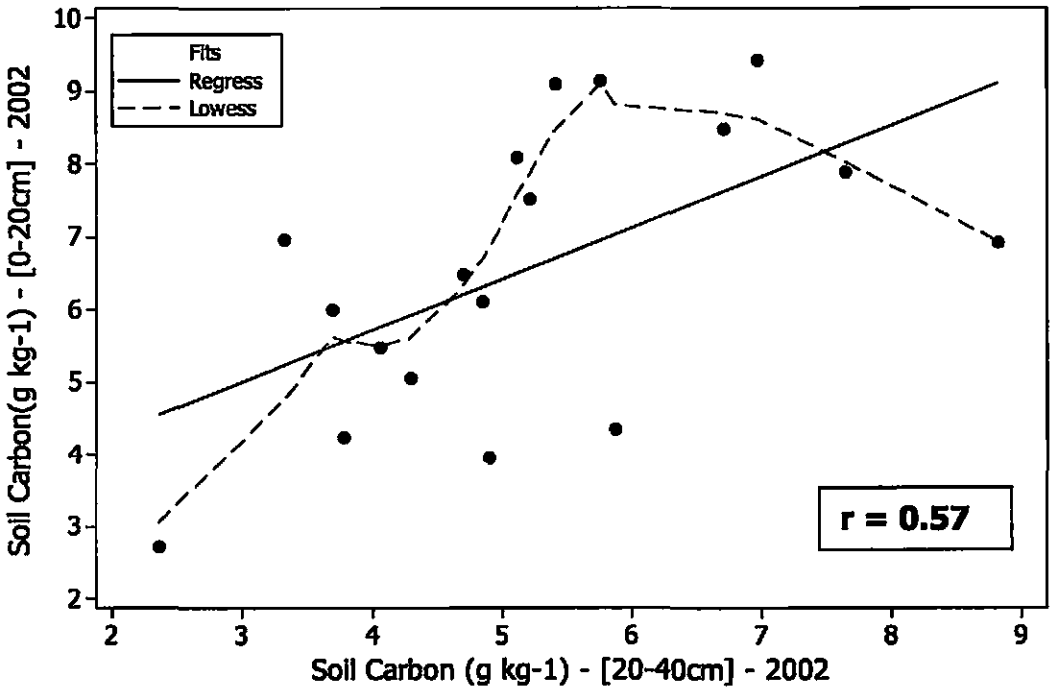


Figure 5.5. Probability plot of soil organic carbon (SOC) (top) and scatter plot (bottom) of soil organic carbon (g kg^{-1}) between surface and lower depths, 2002. The correlation between soil organic carbon content in these two layers is 0.573 and a P-Value = 0.013 show that it was significant.

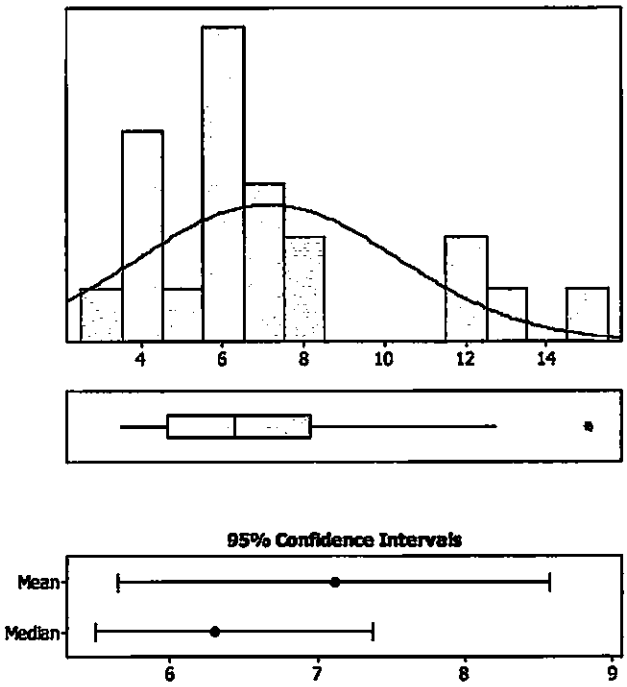
The analysis of the 2002 samples collected in the Baracoro field shows that the distribution is slightly skewed to left with a right tail (Figure 5.4). The mean is higher as compared with the previous year, and most of the observed values center between 5.64 and 8.57 g kg^{-1} carbon.

The 2004 summary statistics of soil organic carbon collected from Baracoro Ballo's farm is presented in Figure 5.6. This figure shows that the measure of central tendency of the distribution, median, differs significantly from the mean and the interquartile range indicating the possible existence of one outlier. After additional data inquiry it was concluded that this value was a true measurement and within the acceptable range, representing areas of the field where management practices had likely contributed to increased carbon.

The correlation of carbon values, at two depths, shows a significant correlation with an r= 0.847 and a P-value less than 0.001 (Figure 5.7). The lowess[4] smoother explores the relationship between two variables without fitting a specific model, such as regression line or theoretical distribution (Minitab 2004). The lowess routine fits a smoothed line to the data.

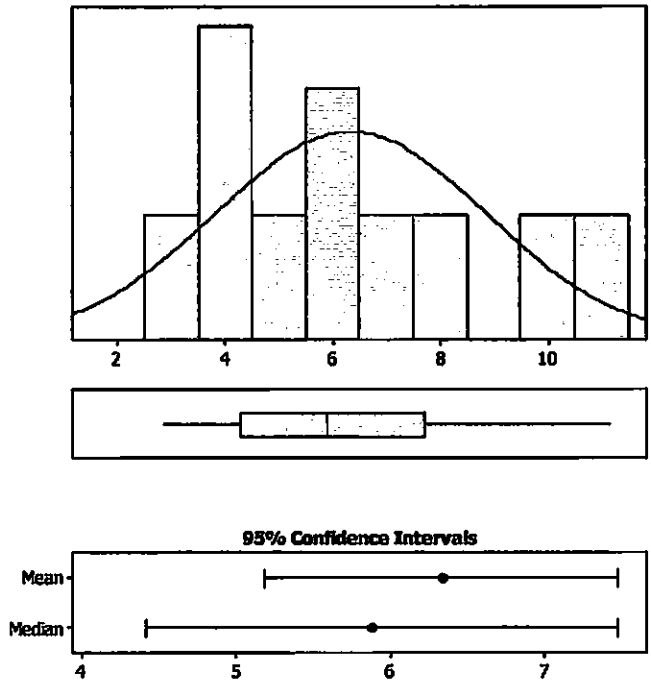
[4] Lowess stands for LOcally-WEighted Scatterplot Smoother(Minitab, 2004)

Year = 2004 [0-20 cm]



Anderson-Darling Normality Test	
A-Squared	1.07
P-Value	0.006
Mean	7.1052
StDev	3.2146
Variance	10.3335
Skewness	1.14705
Kurtosis	0.64001
N	21
Minimum	3.4400
1st Quartile	4.6250
Median	6.2900
3rd Quartile	8.1250
Maximum	15.0100
95% Confidence Interval for Mean	
5.6420	8.5685
95% Confidence Interval for Median	
5.4892	7.3695
95% Confidence Interval for StDev	
2.4593	4.6421

Year = 2004 [20-40 cm]



Anderson-Darling Normality Test	
A-Squared	0.50
P-Value	0.186
Mean	6.3338
StDev	2.5245
Variance	6.3732
Skewness	0.651501
Kurtosis	-0.494029
N	21
Minimum	2.8600
1st Quartile	4.2900
Median	5.8800
3rd Quartile	7.7100
Maximum	11.1400
95% Confidence Interval for Mean	
5.1847	7.4830
95% Confidence Interval for Median	
4.4200	7.4816
95% Confidence Interval for StDev	
1.9314	3.6456

Figure 5.6. Statistical summary of soil organic C collected from Baracoro Ballo's field – 2004 (0-20 cm [top] and 20-40 cm [bottom]).

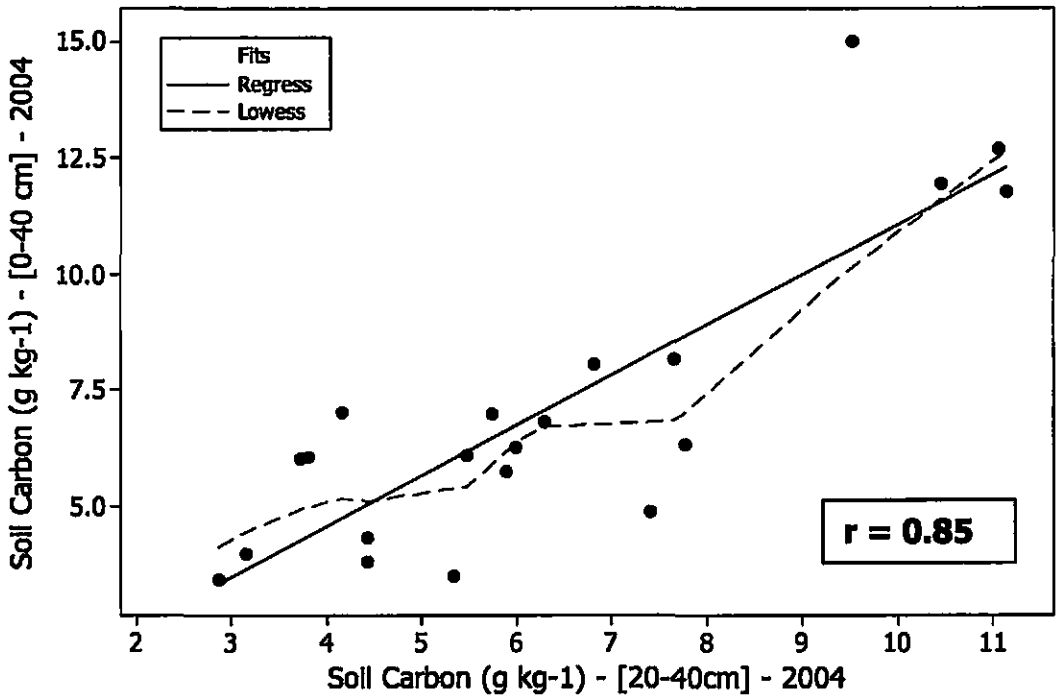
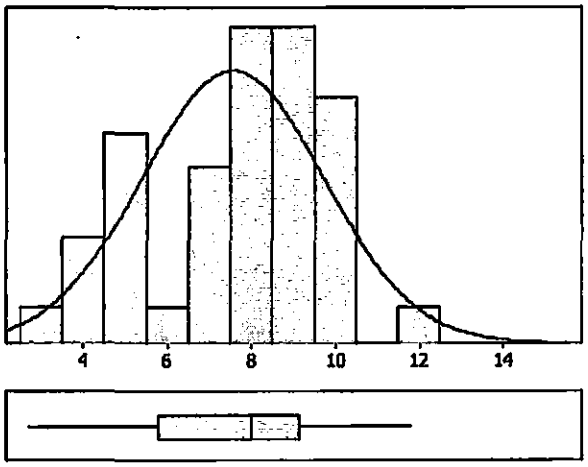
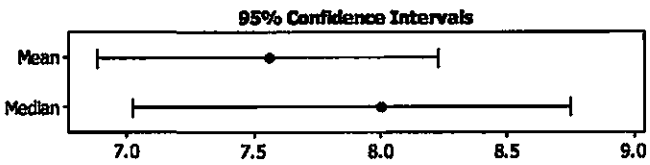


Figure 5.7. Scatter plot of soil organic carbon (g kg^{-1}) between surface and lower depth, 2004.

The normality test for the 2006 observations indicates that both data sets (Figure 5.8) were normally distributed. The interquartile range signaled one possible outlier. However, the maximum value measured in 2006 for upper level was 11.8 g kg^{-1} carbon. The top soil layer shows a higher overall mean. The median was 8 with 95% confidence interval ranging from 7 to 8.75 g kg^{-1} carbon (see Table in Figure 5.8). Similarly to other sites the data collected at Fansirakoro (Baracoro Ballo) showed a small increase in the overall carbon over years. The variance for 2006 was (4.68 and 3.06 for the 0-20 and 20-40 cm, respectively) which indicates the measure of dispersion, or the spread of the data, about the mean.

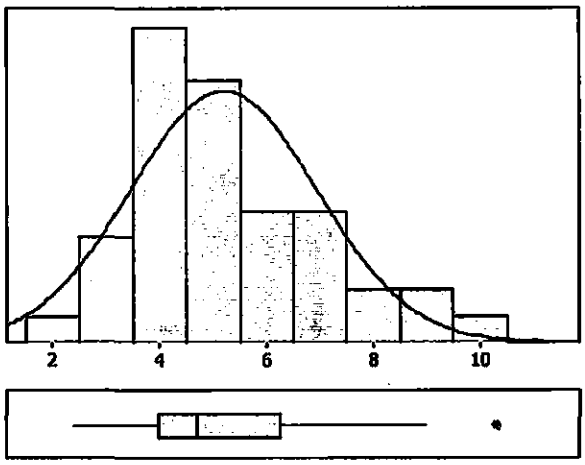
Year = 2006 [0-20 cm]

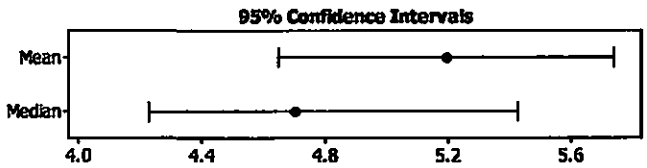




Anderson-Darling Normality Test	
A-Squared	0.73
P-Value	0.052
Mean	7.5571
StDev	2.1629
Variance	4.6781
Skewness	-0.436458
Kurtosis	-0.602782
N	42
Minimum	2.7000
1st Quartile	5.7750
Median	8.0000
3rd Quartile	9.1500
Maximum	11.8000
95% Confidence Interval for Mean	
6.8831	8.2311
95% Confidence Interval for Median	
7.0232	8.7536
95% Confidence Interval for StDev	
1.7796	2.7581

Year = 2006 [20-40 cm]





Anderson-Darling Normality Test	
A-Squared	1.21
P-Value <	0.005
Mean	5.1929
StDev	1.7495
Variance	3.0607
Skewness	1.07855
Kurtosis	0.88069
N	42
Minimum	2.4000
1st Quartile	3.9750
Median	4.7000
3rd Quartile	6.2750
Maximum	10.3000
95% Confidence Interval for Mean	
4.6477	5.7380
95% Confidence Interval for Median	
4.2232	5.4305
95% Confidence Interval for StDev	
1.4395	2.2309

Figure 5.8. Statistical summary of soil organic carbon collected at Baracoro Ballo – 2006 (0-20 cm and 20-40 cm).

The lower depth organic carbon content is characterized by a lower mean and median compared with the surface measurements. The Anderson-Darling test's p-value indicates that, at α levels greater than 0.005, there is evidence that the data do not follow a normal distribution (Figure 5.8).

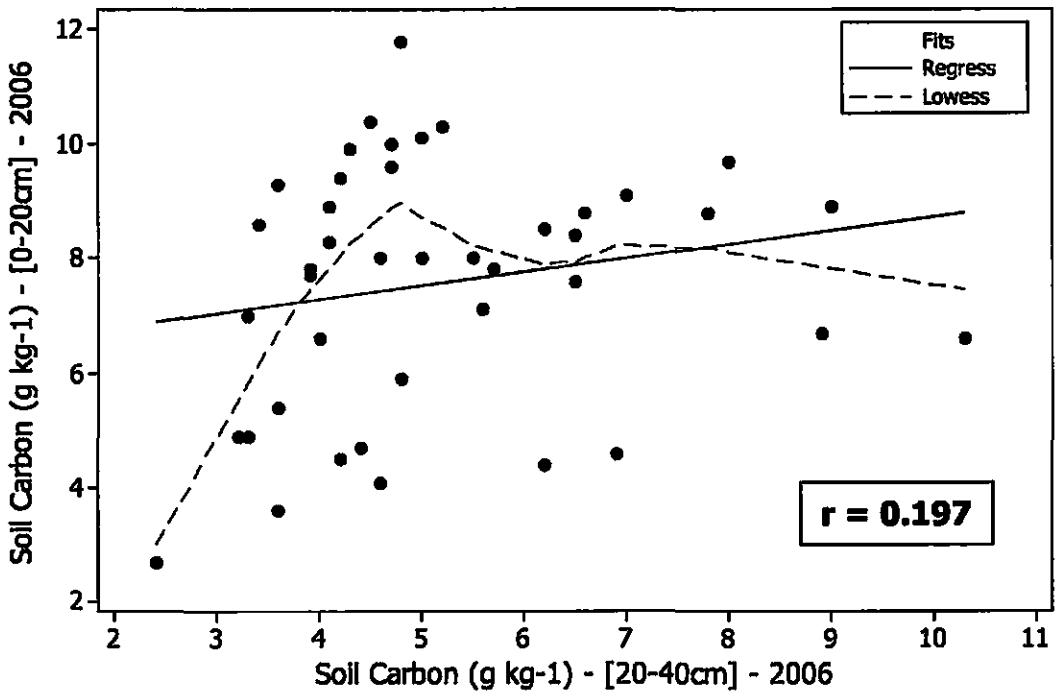


Figure 5.9. Scatter plot of soil organic (g kg^{-1}) between surface and lower depth, 2006. Pearson correlation r= 0.197 and the P-Value = 0.210.

The correlation coefficient between the top and lower layer soil carbon was very low (Figure 5.9). The lowess smoother, however, indicates that there was a sharp break in the relationship between the lower and top layer. This may be explained by the fact that in some locations of the field high values of soil organic carbon at 0-20 cm were matched with lower soil carbon at the 20-40 cm depth creating an initial strong positive correlation (Figure 5.9). However, as the soil organic carbon content, 20-40 cm, exceeded 5.5 g kg^{-1}, the

corresponding values in the surface layer (0-20 cm) were similar or lower affecting the overall correlation coefficient (Figure 5.9).

5.3.1.2. Modeling Spatial Variograms

The variogram model shapes and parameters for data collected in Fansirakoro, Baracoro Ballo's field, are summarized in Table 5.4. The 2002 data was fitted with a Gaussian model with a range of 91.30 m and sill of 5.18. The nugget variance (error variance + micro variance) represented a small fraction of the sill in 2002. Figure 5.10 and Table 5.4 present the spatial dependence information of surface soil carbon from the Baracoro Ballo field. Gaussian models provided the best fit to the semivariances for 2002 and 2004. A spherical model, however, best described the 2006 data set. The range, sill and nugget are presented in Table 5.4. Note that the 2002 model has an r^2= 0.938 and a range of 91 meters. The 2004 carbon data for the soil surface had a longer range compared with 2002 data and the r^2= 0.972. A spherical model with a range of 156.5 meters, a sill of 5.46 described the spatial dependence observed in the 2006 data. As the nugget increases the estimation variance also tends to increase due to the increase in the uncertainty of estimates (Webster and Oliver, 2004, Goovaerts, 1997).

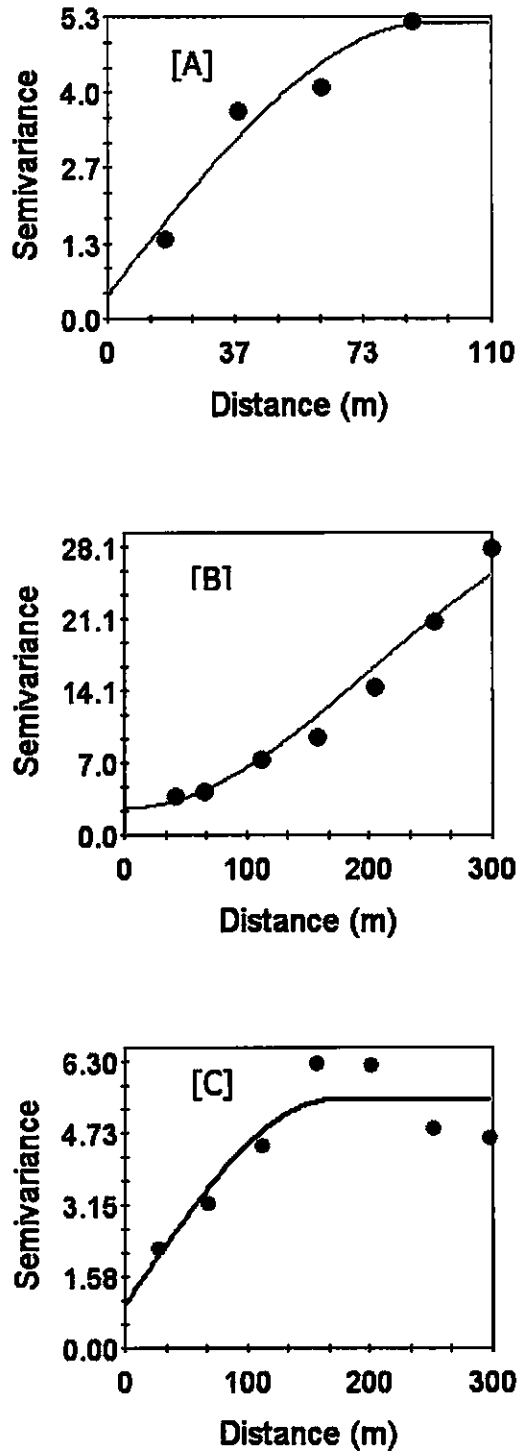


Figure 5.10. Experimental variograms, for surface (0-20 cm) soil carbon (g kg^{-1}) – 2002(A), 2004 (B) and 2006 (C)- Baracoro Ballo's field. The experimental semivariances of carbon are plotted as point symbols and the solid line represents the variogram model.

Table 5.4. Summary of the variogram model parameters for different sampling periods from Baracoro Ballo's field.

Year	Model	Nugget (c_0)	Sill (c)	Range (r) m	r^2
2002	Gaussian	0.42	5.18	91.30	0.938
2004	Gaussian	2.60	36.2	281.00	0.972
2006	Spherical	0.96	5.46	156.50	0.803

The variogram models (Figure 5.8 A & B) do not provide a clear sill and break on the range of the spatial dependence as can be seen in Figure 5.8 C. The sampling layout and dataset size did not adequately capture the spatial variability.

5.3.1.3. Modeling the Experimental Covariance

The covariance model is used for spatial and temporal analysis, since it provides important model parameters (range and sill) for estimation of the primary variable at unsampled locations in the field. Spatial dependence quantifies the property that as the distance between samples of soil carbon increases the correlation decreases.

The experimental covariance and model for subsets of data collected at Barracoro Ballo's field in 2002, 2004 and 2006 are presented in Figure 5.11. In 2002 a much shorter range of spatial dependence was observed compared to the 2004 and 2006 datasets (Figure 5.11). The 2006 experimental covariance was better fit by the covariance model and its covariance classes were comprised of a sizable number of pairs (>15).

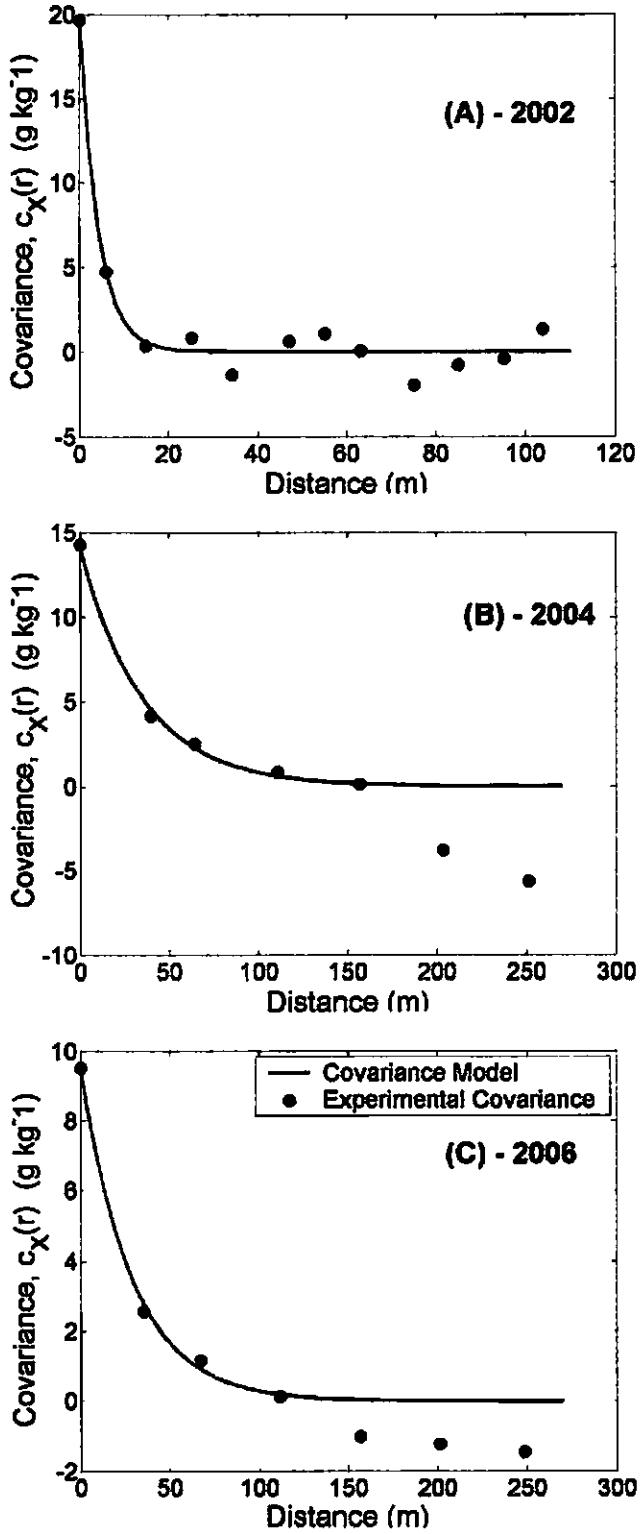


Figure 5.11. Spatial dependence of soil organic carbon as characterized by covariance models of three sampling periods from Baracoro Ballo's field (2002, 2004 and 2006).

5.3.1.4. Comparison between BME, cokriging and kriging

The locations of hard and soft data in Baracoro Ballo's field are indicated below (Figure 5.12). In this case study two additional types of information about Baracoro Ballo's field were considered: (i) soil carbon from a surface (0-20 cm) and deep sample (20-40 cm); and (ii) clay. The site specific soft data used were clay and soil carbon from deeper sample (20-40 cm). Through a discretization technique these data were assigned into classes and used as soft data.

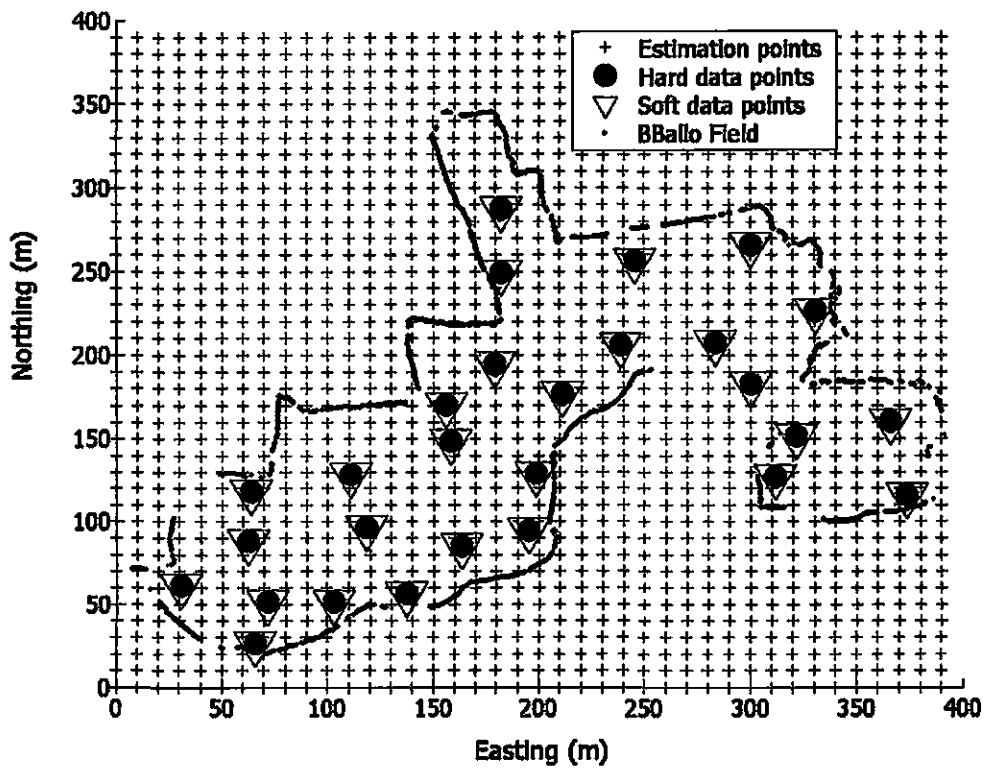


Figure 5.12. The location of the 2006 hard data (carbon 0-20 cm), Soft data (clay, carbon 20-40 cm) and the 2 dimensional grid defined as estimation points. The estimation points total 1600, based on a 10x10 meter grid. Twenty eight hard data and 28 soft data locations were used to implement the BME approach. Note that the soft data points are collocated with the hard data in this case.

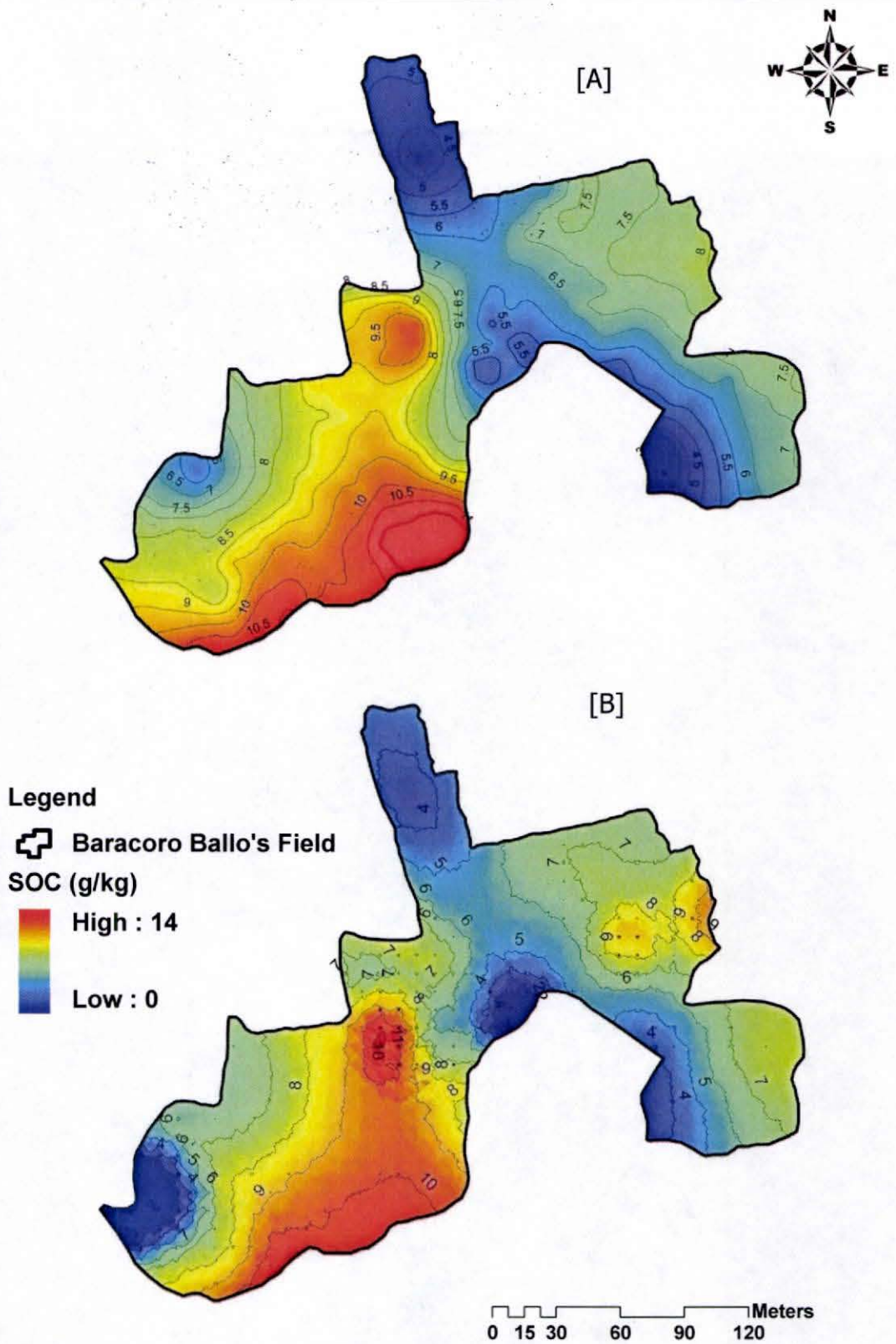


Figure 5.13. Prediction maps for Baracoro Ballo's Field, dataset of 2006, with BME - soft data (clay, carbon 20-40cm) and hard data (carbon 0-20 cm) [A]. BME – soft data (clay) and hard data (carbon 0-20 cm) [B].

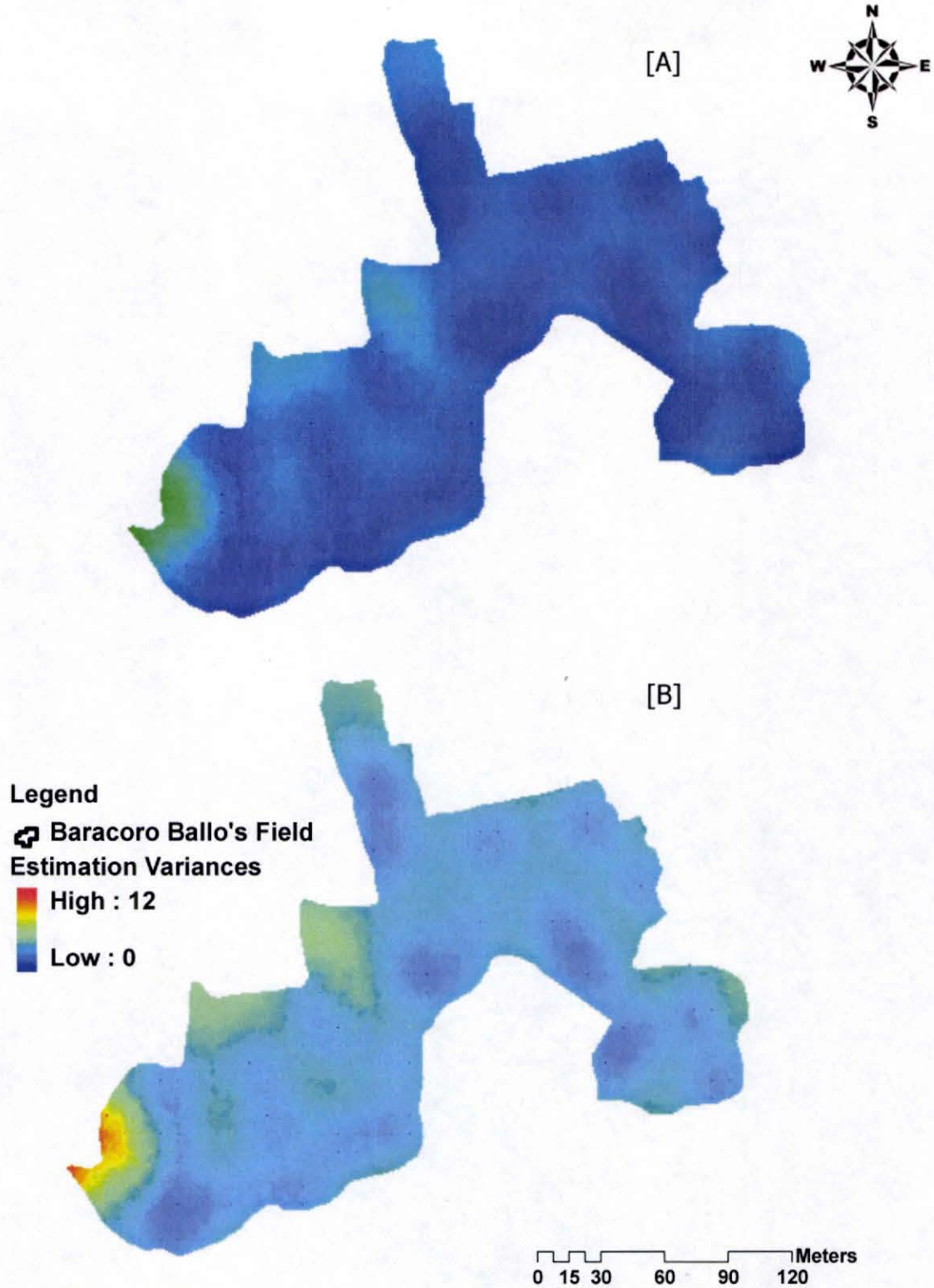


Figure 5.14. BME estimation variances of the prediction maps shown in Figure 5.16. Top map shows the estimation error of BME (w/ clay and carbon (20-40 cm as soft data) – [A]. The estimation error of BME (w/clay as soft data) is presented in Figure 5.16 – [B].

Figure 5.13 demonstrates a multivariate BME, which integrates classes of soil carbon (lower depth) and clay (0-20cm) with surface carbon (0-20 cm) to generate a prediction map for the entire grid surface. Figure 5.13-A illustrates an estimation map based on two additional variables with soft data (clay and soil carbon (20-40 cm). Figure 5.16-B is a BME prediction map with soil carbon (20-40 cm) as the only soft data.

Figure 5.13 provides information about the spatial distribution of soil carbon in this particular field. This figure gives an indication of the range of soil organic carbon values as well as its spatial distribution. Soil organic carbon levels over 8 g kg^{-1} were registered mainly in the southwestern maizeer of the Baracoro Ballo's field (Figure 5.13). For most Sub-Saharan African soil carbon levels seldom exceed 10 g kg^{-1} (Piéri, 1989; Bationo, et al., 2001). A soil organic carbon prediction surface can be an important management tool since it provides the spatial representation of areas of a region or field below certain levels (Figure 5.13).

Figure 5.14 shows the error map for the predictions presented in Figure 5.16. Unlike the classical geostatistics methods (i.e. kriging, cokriging) the BME variance computation includes the contribution of interval soft data in the error variance. Where only hard data is used, the variance of BME posterior PDF is the same as simple kriging error variance.

Cokriging

Cokriging takes into account covariance and cross covariance between the estimation point and the data points of a primary variable and the covariable(s). Figure 5.15 shows an example of the cross-covariance that includes the primary variable soil carbon (0-20cm) and two covariables (or secondary variables), clay, and soil organic carbon 20-40cm. Soil organic carbon 20-40 cm data did not show a clear experimental structure of the variance as did the primary variable. The variogram of carbon (20-40 cm) is almost pure nugget effect. The cross-variogram of carbon (0-20) vs carbon (20-40) indicates that correlation between the primary variable and the secondary variable was weak. For clay the cross-variograms show a better fit suggesting that using clay as a covariable could improve the prediction of carbon (0-20 cm) (Figure 5.15 and Figure 5.17).

The cokriging analysis (Eq. 4.6) compared carbon at the 20-40 cm depth and clay content (0-20 cm) as covariables to predict carbon at soil surface. The surface prediction was restricted to an area 400 x 400m with 10 x10m prediction grid. Figure 5.16 shows a comparison of three carbon prediction surfaces at Baracoro Ballo using cokriging techniques. Soil carbon (20-40cm) and surface clay (0-20 cm) served as covariables. The use of covariable(s) is expected to improve the spatial prediction of the primary variable. When clay alone was used the range of soil organic carbon in the field increased. In all scenarios the high and low areas of predicted soil organic carbon remained the same.

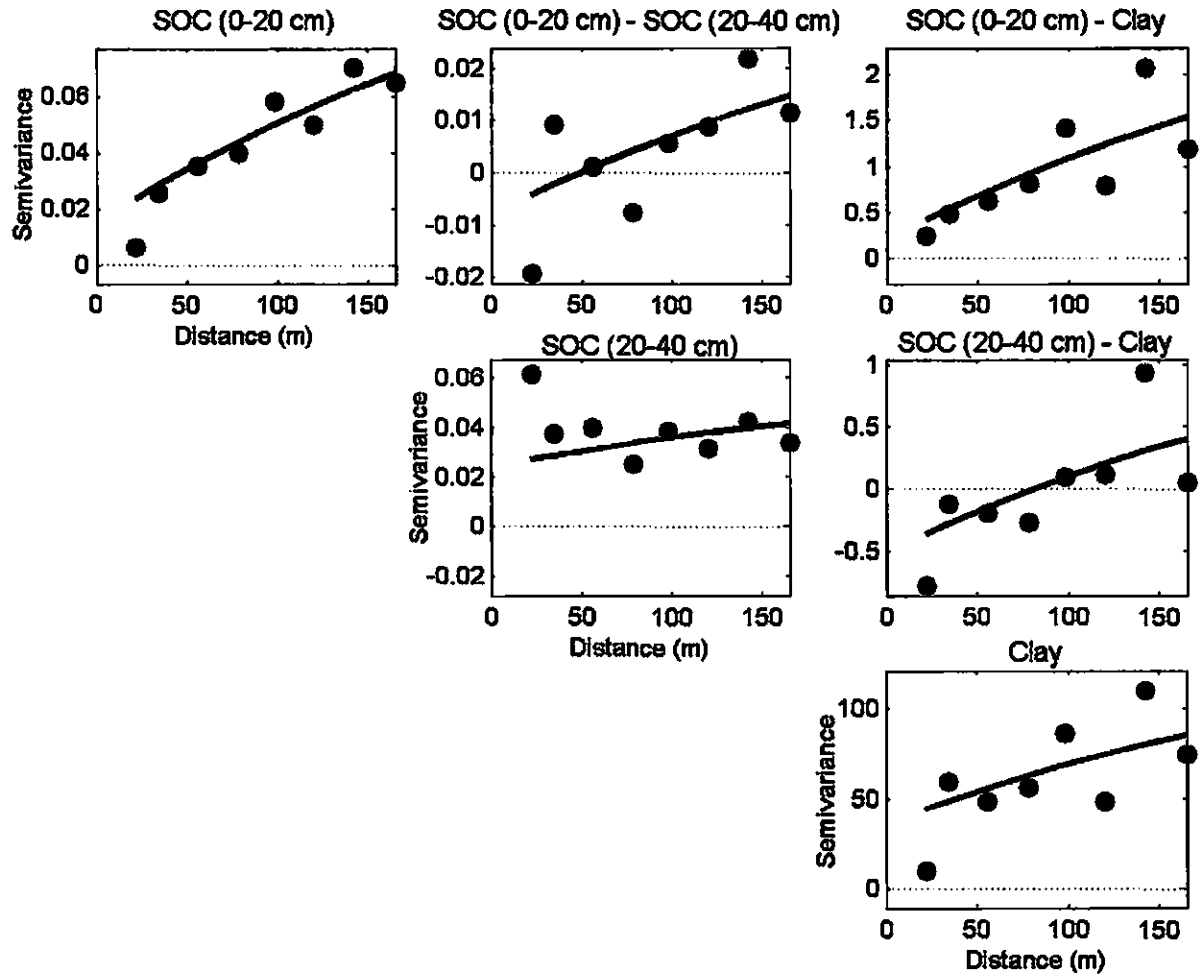


Figure 5.15. Matrix of variograms diagonal and cross-semivarigrams of soil organic carbon (SOC 0-20cm, SOC 20-40 cm) and clay (0-20 cm), for the 2006 dataset from Baracoro Ballo's Field. Soil organic carbon (SOC 0-20 cm) is the primary variable, the covariables are SOC (20-40 cm) and clay (0 -20 cm).

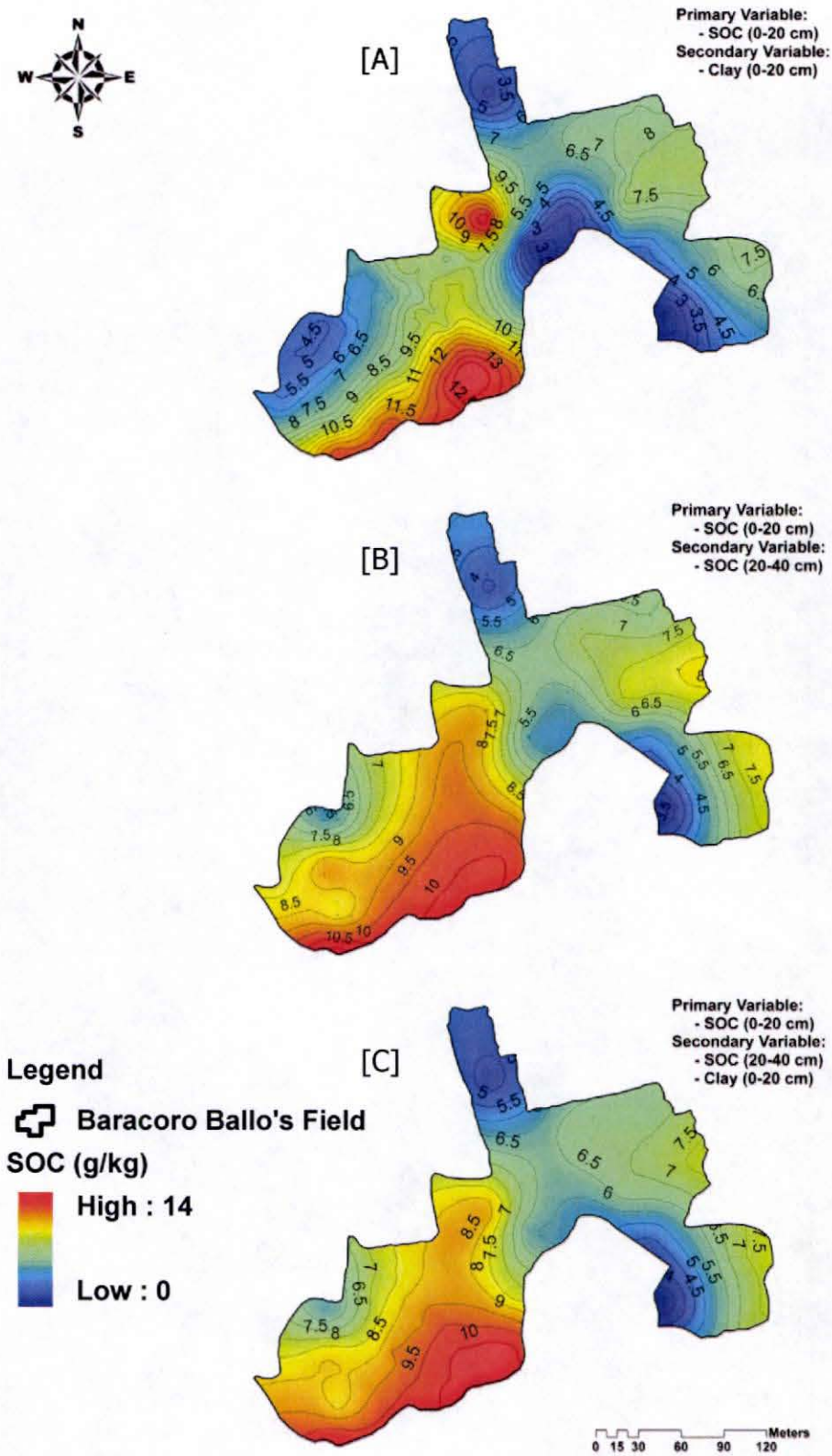


Figure 5.16. Comparing three cokriging predictions for surface soil carbon. [A] cokriging with clay (0 – 20 cm); [B] Cokriging with SOC (20-40 cm); [C] Cokriging with both clay (0 – 20 cm) and SOC (20-40 cm).

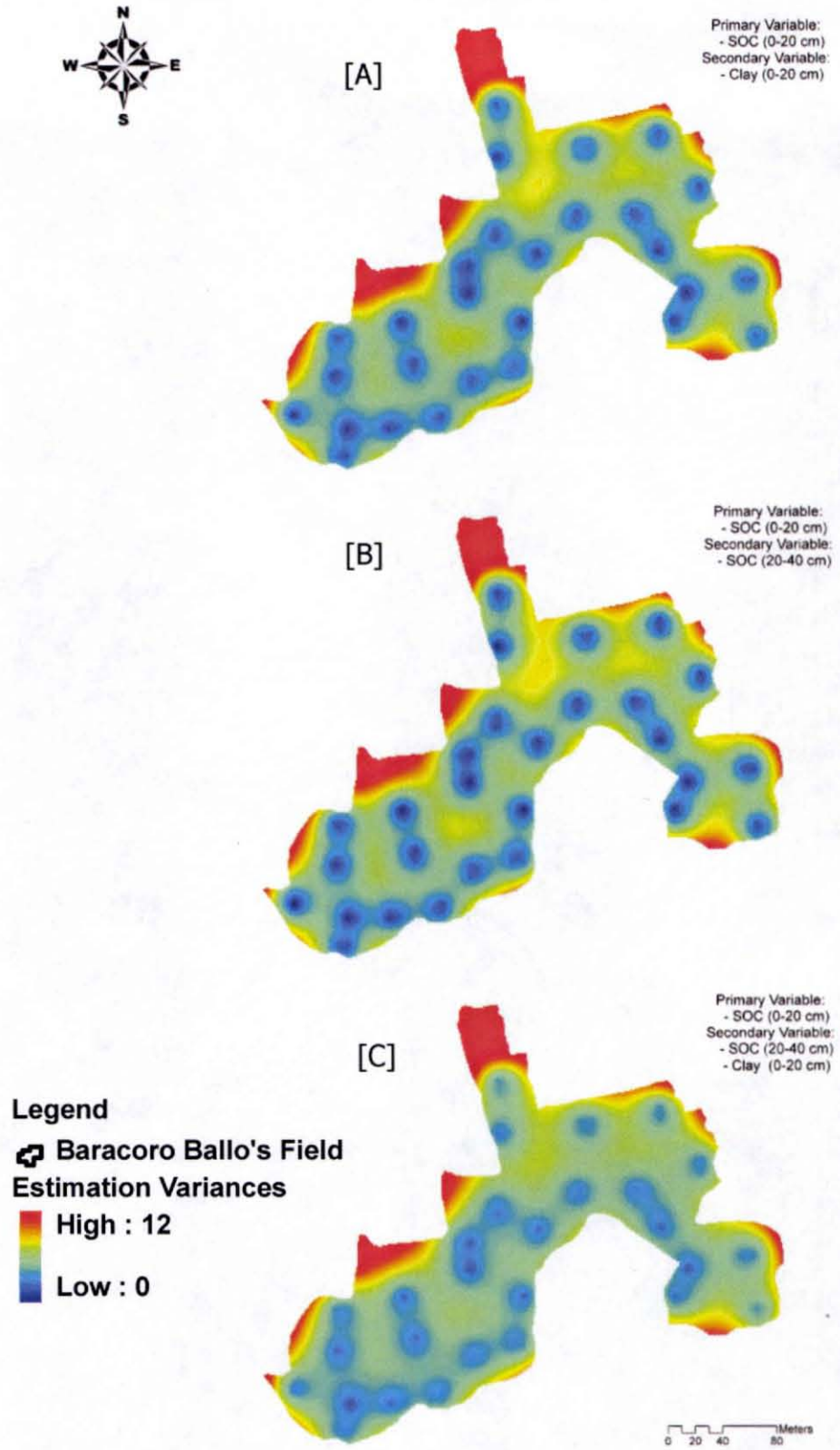


Figure 5.17. Three cokriging estimation variances of the prediction maps for Baracoro Ballo's field. [A] variance error - cokriging w/ clay, [B] Variance error - cokriging w/ SOC(20-40 cm), [C] Variance error - cokriging w/ clay and SOC(20-40 cm).

Figure 5.16 shows the cokriging estimates, the prediction ranges are similar to kriging estimates presented above (Figure 5.18).

The estimation variance of the cokriging maps display a similar pattern to kriging, that is, near the measured data the variance is zero or near zero and tends to increase away from the measured data (Figure 5.17 and Figure 5.19 - B).

The results show that adding covariables (clay and SOC (20-40 cm) reduced the prediction error. Furthermore, it appears that clay as a secondary variable reduced prediction error more than SOC (20-40 cm). Figure 5.17 – A shows the efficiency of the clay covariable in decreasing the estimation error beyond the range of the data points. The cokriging prediction map (w/ clay) has a lower error variance compared with the cokriging (w/ carbon (20-40 cm). The figures illustrate that when the covariance between primary and secondary variable is strong the prediction error decreases (Figure 5.7).

Kriging

The kriging estimates were based on hard data (soil carbon 0-20cm) only. As described in section 4.2.1 kriging is a linear estimator that relies exclusively on the measured data. The spatial neighborhood consists of defining the maximum number of hard data to be considered in the estimation, and the maximum distance allowed between estimation points. At each estimation location, there was a maximum of 10 soil carbon points used for the estimation and these values were located not further than 300 meters from the estimation location.

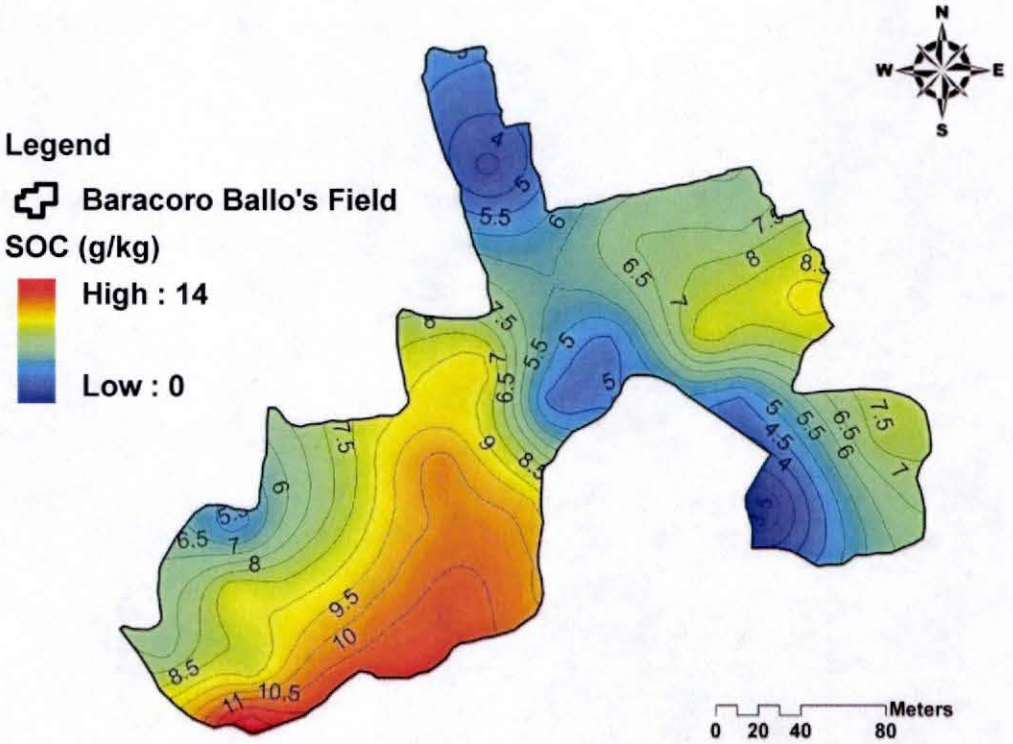


Figure 5.18. Kriging estimate of soil carbon for Baracoro Ballo's Field, based on hard data only, 2006.

Figure 5.18 illustrates a prediction map of soil carbon in Baracoro Ballo's field. The prediction was based on 28 data samples (Figure 5.12). Increased soil organic carbon seemed associated with subareas where important soil and water conservation measures (contour ridges, ACN) were put in place by the SM-CRSP project with the help of local farmers. The carbon content ranged from 3 to 12 g kg^{-1} and the mean soil carbon was 7.75 g kg^{-1}. Baracoro Ballo's field was heterogeneous (Figures 5.13, 5.18). Kriging minimizes the error variance at the sample location. Kriging variances were plotted, indicating the prediction error at each estimation location (Figure 5.19-B). The error variance near the sample location was lower than 0.05. As the prediction distance increases away from the hard data the variance increased up to 1.2 g kg^{-1}. Inside Baracoro Ballo's field the estimation variance was low. For model comparison the analyses were restricted to estimation points inside the field's perimeter.

The error variance is an important indicator of the reliability of the soil organic carbon prediction. For soil carbon quantification and mapping purposes it is important that the prediction model yields the lowest error variance. Figure 5.19 compares the error variance generated by BME (Clay, SOC (20-40 cm – interval soft data) with kriging.

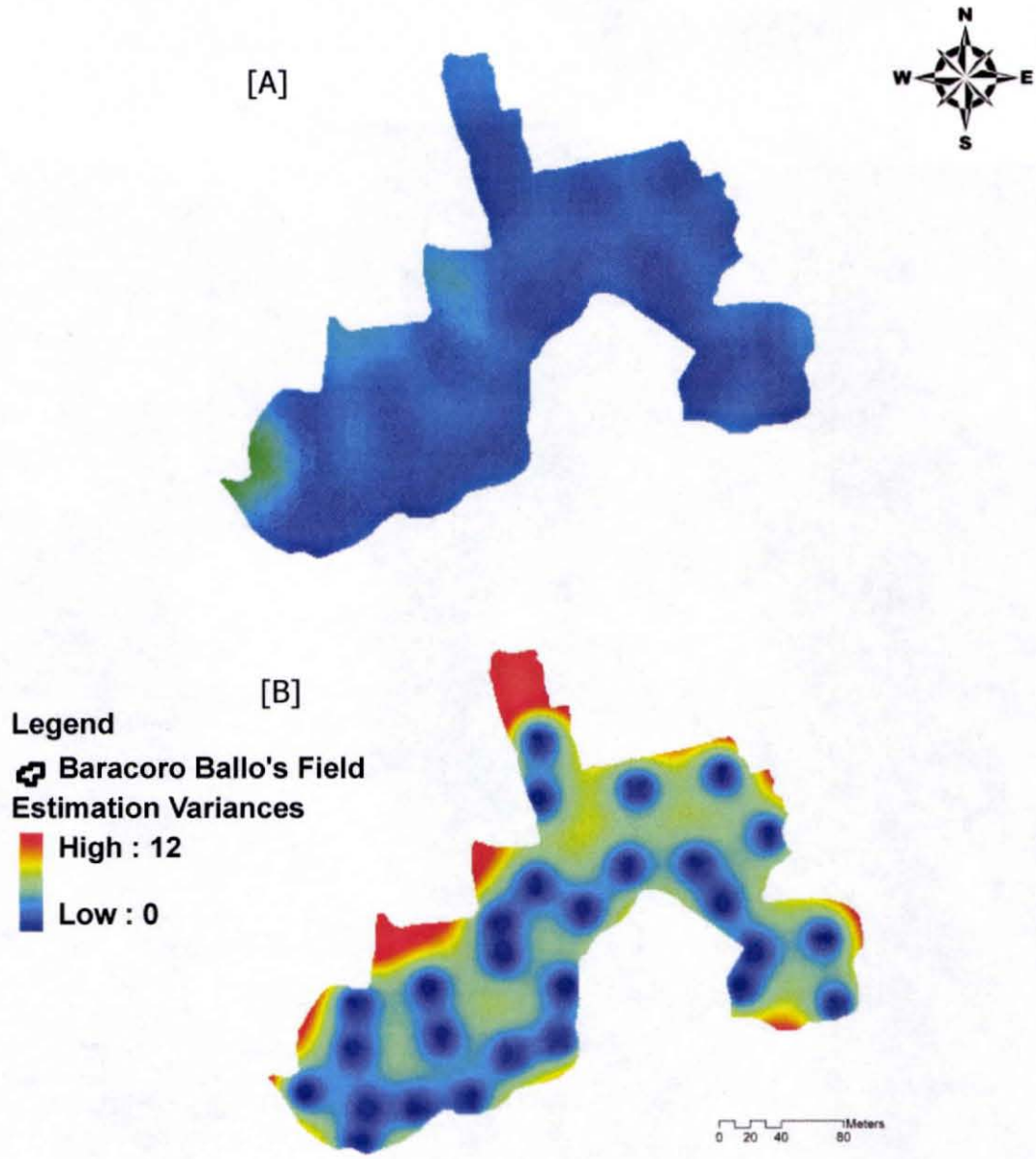


Figure 5.19. The estimation error of BME (carbon, clay) prediction (A) is presented and it indicates an overall decrease in error variance. Kriging variance (B) based on hard data (carbon, 0-20 cm). The error variance is lower near the data points.

These estimation error maps could be contrasted with the cokriging estimation error illustrated in Figure 5.17-C. Kriging and cokriging produced similar outcomes, due to the fact that both interpolators are linear and unbiased. Consequently, near the data points the error variances were lower, close to zero.

As described in Eq. 4.22, when using the soft data the error variance is based on the estimation error for the entire PDF, which accounts not only for the hard data but also for the probability distribution of the soft data.

Table 5.5. A comparison of total soil organic carbon (Mg) predictions using different methods. The variance for each prediction maps is also presented.

Quantification Method	Total Soil Organic Carbon (Mg)	Variance
Kriging	118.68	0.20
Cokriging (clay)	117.73	3.93
Cokriging (SOC)	119.15	3.41
Cokriging (clay, SOC)	119.49	0.29
BME (clay,SOC)	122.22	0.03

The results presented in Table 5.5 show the total soil organic carbon computed for the entire field. The difference between kriging calculations and BME is only 3.5 Mg of carbon; however, the variance for the surface prediction was considerably lower when the BME method (with clay, SOC) was applied to the data. The use of two covariables in cokriging did not reduce the estimation variance. Furthermore, the use of either clay or SOC as a covariable increased the estimation error variance considerably when compared with cokriging with both clay and soil organic carbon (20-40 cm).

5.3.1.5. Simulation of S/TRF and Method Comparison

The stochastic simulation is the process of generating alternative realizations of the space/time random field (Serre, 1999). Each realization represents a stochastic image, which reflects the probability distribution function and the spatial dependence that have been imposed on the S/TRF. Whereas cross-validation provides a robust method to test kriging and cokriging it is not suitable for comparing stochastic simulation models. One method of comparing stochastic simulation is to compare the resulting probability density functions of alternative simulation methods with a reference probability density function (PDF). This approach was used in this study to compare BME with other methods.

Similarly, in order to compare the mapping accuracy between the estimation methods, a stochastic Gaussian sequential conditional simulation of the S/TRF was generated with the hard and soft data of interval type.

The maps and probability distribution functions (PDF) generated using standard spatial stochastic simulation were used as reference values and paired with prediction maps (kriging, cokriging and BME) for accuracy assessment.

The bias of each method is reported in Table 5.6. The bias of an estimator is the difference between the expected value of one method and the true or reference value of the parameter being estimated. The bias or the mean error (ME) parameter was computed based on the mean of predicted by each method minus the mean of simulated results.

Table 5.6. Mean estimation error (ME) for simple Kriging, cokriging and BME methods The bias of each method as compared with a Gauassian conditional simulated prediction.

Method	Bias (ME)
Kriging	0.41
Cokriging (clay, SOC)	0.60
BME(SOC)	0.29
BME (clay, SOC)	0.033

An unbiased estimator should have an ME equal to zero. In other words, the predicted should equal the reference value. Consequently, the lower the mean error (ME) the better is the performance of the estimator. Table 5.6 shows that BME (clay, SOC) and BME (SOC) have the lowest ME as compared to kriging and cokriging.

The mean squared error (MSE) can be used to compare the resulting PDFs as is defined as follow:

$$MSE = \frac{1}{m_{Xh}} \sum_{i=1}^{m_{Xh}} (X_i^* - X_i)^2 \qquad \text{(Eq. 5. 2)}$$

Where the X_i^* is the reference simulated surface data, X_i are the predicted maps (kriging, cokriging and BME). In this study m_{Xh} corresponds to the 1600 prediction locations over the entire grid/field (Figure 5.12). The MSEs and percent change for each method are presented in Table 5.7 below. The hypothesis was that BME may provide more accurate estimates than other geostatistics methods (Christakos et al., 2000; Christakos et al., 2002). In this case, more accurate indicates a closer approximation to the reference spatial simulation as quantified by Eq. 5.2.

Table 5.7. Comparisons of BME prediction accuracy with that of kriging and cokriging with one or two variables.

Methods	MSE	% Change
Kriging	0.338	
Cokriging (SOC)	0.463	36.98
Cokriging (Clay,SOC)	0.304	- 10.06
BME (Soft :Clay, SOC)	0.225	- 24.56

The MSE should be small and if the model is accurately predicting, the MSE should be equal to kriging variance and hence the relative mean square error (RMSE) is equal to 1. Table 5.7 shows that the BME prediction provided a lower MSE representing a 25% reduction compared with Kriging, and 10% compared with cokriging. In this particular analysis cokriging with only one covariable (SOC) fails to improve the accuracy of prediction and the comparison shows that the cokriging MSE was, in fact, 37% greater than that of kriging.

5.4. Discussion of Spatial/Model Comparison

As shown above with BME, the estimates of soil organic carbon were more accurate than estimates generated using kriging methods (Table 5.6). The prediction model comparisons using the mean error (ME) indicated that BME performed better than did the kriging methods (Table 5.7).

The contribution of the soft data to prediction accuracy and precision can be large (see Tables 5.6 and 5.7). The major difference between BME and classical geostatistics methods is that the latter lacks the capability to incorporate soft data with other site specific information. BME on the other hand, facilitates the integration of both hard and soft data (ex. texture, landuse, slope classes) and the prediction is obtained based on the entire posterior probability distribution.

The integration of hard and soft data improved the accuracy and precision of soil organic carbon estimation at the field level. BME provides a framework to maximize use of the auxiliary site specific data. Site specific soft data that relates to the main variable of interest (hard data) is often abundant or easy to acquire. Because mapping soil properties often requires intensive soil sampling, the use of either interval classes or probabilistic soft data to improve prediction of the primary variable of interests is of paramount importance for carbon accounting at both field and regional scales.

CHAPTER 6: SPATIO-TEMPORAL ESTIMATES OF SOIL AND TREE CARBON

6.1. Introduction

Time is an important factor in soil development and in actual fact all soil properties vary in time, and soil organic carbon is no exception. However, some changes in soil properties occur very slowly relative to the human time scale. For this reason there is a tendency to consider spatial variability assuming no change in time (Heuvelink and Webster, 2001). More recently, several studies applying geostatistical tools have attempted to extend modeling of spatial phenomena to the time domain (Serre, 1999; Christakos et al., 2002). Jost et al. (2005) used a spatio-temporal kriging for mapping soil water storage in a forest ecosystem. Comegna and Vitale (1993) also modeled the soil water moisture in space and time using an autoregressive model. Walter et al. (2003) estimated the evolution of organic carbon over the landscape using a spatio-temporal simulation method. Since changes occurring in soil are dynamic, the challenges to soil scientists are to model them simultaneously in space and time (Goovaerts and Chiang, 1993; Bogaert et al., 1999; Christakos et al., 2002). "Temporal prevalence" is the successive existence of a particular spatial pattern which necessarily implies that the spatial variation is smaller than the temporal, (Goovaerts and Chiang, 1993).

Short term forecasting of air pollution and other atmospheric parameters have been reported using the spatiotemporal models (i.e. Bayesian Kriged

Kalman model) (Sahu, 2005). Several references (Cressie, 1994; Wikle et al., 1998; West and Harrison, 1997; Jost et al., 2005) demonstrated models for spatiotemporal data. A recent approach of Serre (1999) and Christakos et al. (2002) showed that the Bayesian maximum entropy approach facilitates the modeling of spatial and temporal point measurements.

The space and time estimates of soil organic carbon provide an understanding of the dynamics of soil carbon. This approach gains particular relevance as the Kyoto protocol Clean Development Mechanism (CDM) allows developing nations to capitalize on their sequestration effort with concrete economic benefits that might derive from carbon trading. When considering carbon trading, time becomes a critical parameter both in terms of establishing a reference baseline and analysis of temporal trends of a particular spatial pattern.

Temporal trend mapping can also give an indication of the proper time between sampling to avoid redundancy and to improve sampling efficiency. For carbon compliance monitoring, a temporal trend analysis might indicate the departure from the initial mean value over a specified period of time.

Most traditional soil carbon inventories map the spatial variation but ignore the temporal component. Combining space-time to map soil organic carbon is a more robust approach than purely spatial mapping. The challenge involving space time estimates of soil organic carbon resides in the selection of the appropriate covariance model. The theoretical and computational advantages of a composite space/time mapping were presented by Christakos et

al. (2002). The work of Christakos (2000, 2002), Cressie and Huang (1999), Jones and Zhang (1997) discussed a series of new classes of stationary space-time covariance functions. Given the difficulty to simultaneously think of the spatial and temporal variations, it is often easier to assess how the covariance of a particular place changes in time and how the covariance at a particular time varies across space (Stein, 2005). Separable models, which are a product of permissible spatial and temporal models, allow the two covariance models to be analyzed independently. These classes of covariance models are convenient but do not model the space-time interaction (Cressie and Huang, 1999). The nonseparable spatiotemporal covariances are obtained with parabolic-type partial differential equations (Christakos, 2000; Stein 2005). The spatio-temporal approach in BMElib provides a flexible method of defining space-time covariance models. The BME covariance equation can include separable and nonseparable functions associated to homogeneous/stationary or nonhomogeneous/ nonstationary random fields (Christakos, 2000).

In the context of mapping soil carbon for trading and compliance, the assumption of separability of the covariance is desirable since it facilitates the study of spatial and temporal components separately. The main criticism of this approach is that it separates dependence of space and time which might be unrealistic for most soil proprieties (Jost et al., 2005; Walter et al., 2003; Christakos et al., 2002).

This study demonstrates potential use of space – time covariances as tools to improve our understanding of spatial and temporal variability of soil organic carbon. Based on the temporal and spatial models structures maps can be generated to predict mean trends into the future.

6.2. Materials and Methods

6.2.1. The study area

6.2.1.1. Oumarbougou - Mory Konate

The Mory Konate farm is located in the Eastern most portion of the Sikasso "Region" (Figures 3.1 and 6.1). The central coordinate location for this field was 5° 8′ 54.375″ W and 12°11′ 8.939″ N. The farmer area was approximately 20 hectares subdivided into five fields with different crop rotations and management practices. As implemented there were no with and without ACN fields. Hence direct estimates of the ACN effect were not possible. Comparisons of ACN effect on soil C and are presented in Doumbia et al. (2008, in review). The data set from this field consisted of 286 soil samples collected from the top soil (0-20 cm) and subsoil (20-40 cm).

The temporal component of this data covers the following sampling periods: 2000, 2002, 2004 and 2006. The 2000 data were collected in November (just after the cropping season) while the all subsequent sampling was conducted off season during March-April. The sampling strategy was similar to that described for Fansirakoro.

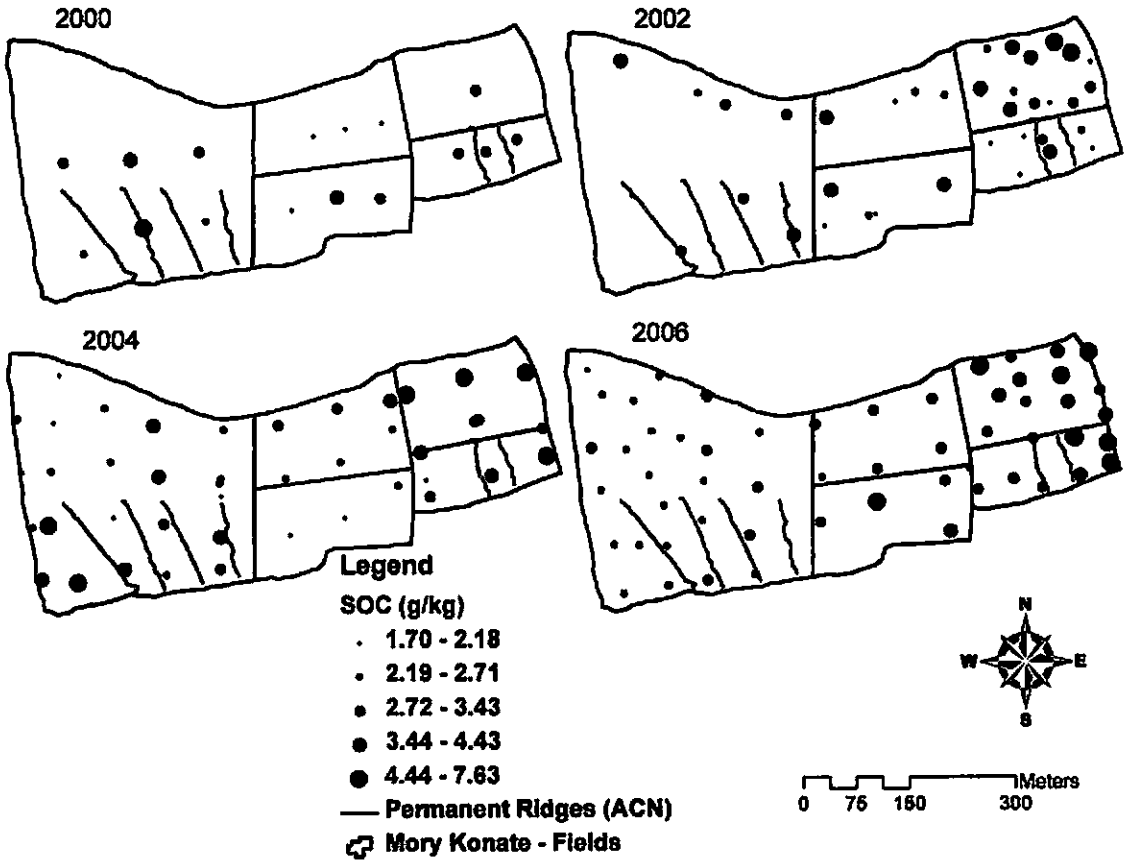


Figure 6.1. Dataset from Mory Konate's field, 2000, 2002, 2004 and 2006. Sample collection scheme and soil organic carbon (SOC), 0-20 cm depth. Mory Konate subdivides his field into five main subplots for management as a coping strategy against drought and potential crop failure. Each sample point has a separate space and time location.

In this region cotton is a major cash crop. Cotton production is intensive and additional inputs (i.e. fertilizers) are required to guarantee a successful crop. Traditionally farmers rotate either sorghum or millet after cotton. Since cotton is the only crop that receives fertilizer, the farmers aim to capitalize on the residual effects of added nutrients.

6.2.1.2. Siguidolo - Zan Diarra

The data set from Zan Diarra's field is comprised of 132 samples collected in 2000, 2004 and 2006. The area of this field is 18 ha and the sampling strategy was similar to that applied in other sites. The central coordinates of this field are 6° 47′ 38.468″ W and 12° 54′ 54.342″ N. Zan Diarra's field is located in the Region of Segou (Figure 3.1). The first sampling was essentially exploratory with a few samples taken at selected locations in the field. This was the first field with the ACN technology, which was implemented in 1994. In 2004 and 2006 the number of samples collected was increased to provide information for the entire field. In 2006 the soil sampling went beyond the farmer limits to include an adjacent non-ACN field (see Figure 6.2). This was an attempt to test the effect of the ACN technology and C sequestration.

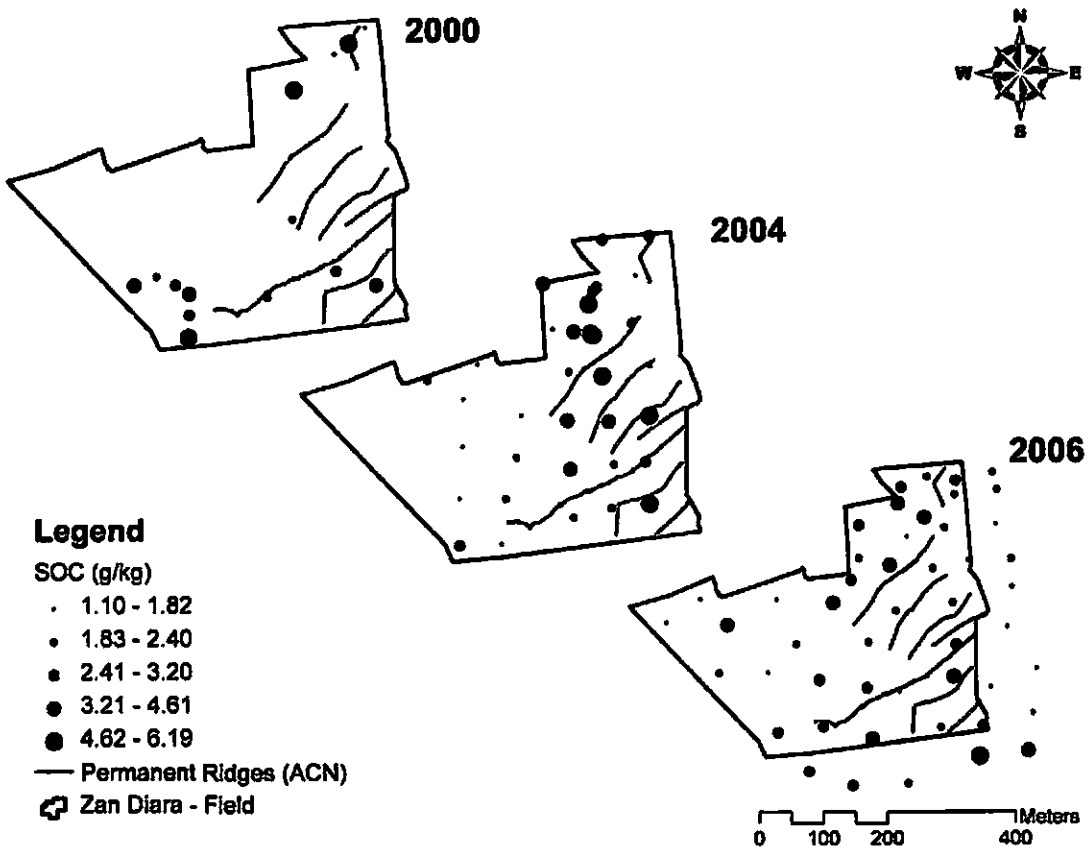


Figure 6.2. Dataset from Zan Diarra's field, 2000, 2004 and 2006. Sample collection scheme and magnitude of soil organic carbon (SOC) of top layer (0–20 cm) is here represented. The 2000 dataset has less spatial coverage as compared to 2004 and 2006. Each sample point has a separate space and time location.

6.2.1.3. Sikasso - Yaya Diassa

Yaya Diassa's field is located in the southern Region of Sikasso (Figure 3.1 and 6.3). The central coordinates of this field are 5° 35′ 10.916″ W and 11° 14′ 5.185″ N. The soil carbon in this field was measured in 2000, 2004 and 2006. The sampling location follows the same approach already applied in other sites. In this field the ACN technology was installed prior to sampling. New and improved ridges as well as renewed interest in soil water conservation have contributed to a continuous expansion of the farm's perimeter and hence the sampling area. The total data set for Yaya Diassa field was 308 soil samples, representing two subsets, one taken at the 0-20 cm and one at the 20-40 cm depth (Figure 6.3).

As indicated for the sites of Mory Konate and Zan Diarra, no direct control of no ACN was present in this field. Thus direct comparisons of ACN versus no ACN on C sequestration were not possible. Such comparisons are presented in Doumbia et al., 2008, in review.

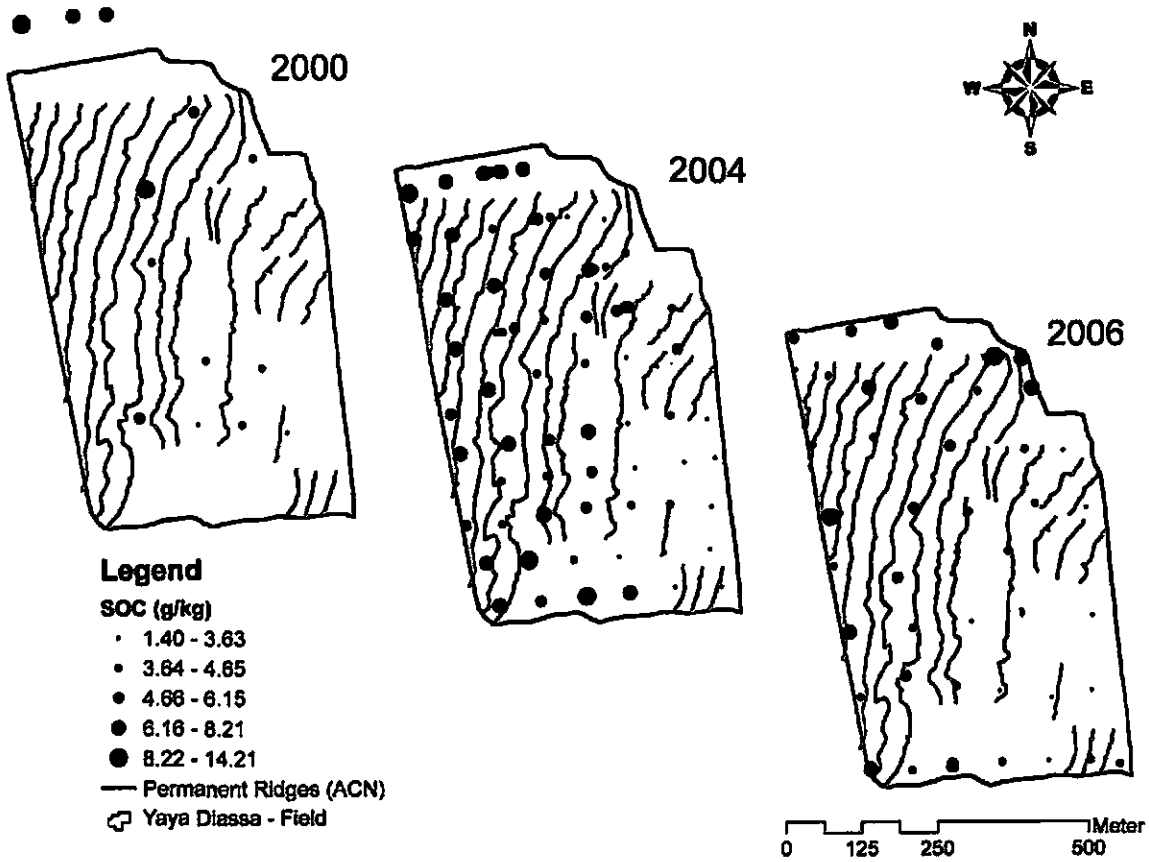


Figure 6.3. Dataset from Yaya Diassa's field, 2000, 2004 and 2006. Sample collection scheme and magnitude of soil organic carbon (SOC) is here represented. Similar to previous datasets the 2004 and 2006 have better spatial coverage of the entire Yaya Diassa field. Each sample point has a separate space and time location.

6.2.2. Conditional Simulation

Sequential conditional simulation is widely used and accepted as a method of creating stochastic models (Goovaerts, 1997; Goovaerts, 2001). This type of simulation includes the usual random generation of a realization from an underlying statistical model, but in this case the underlying model is one that describes the spatial relations as defined by covariance or variogram models (Goovaerts, 1997; Deutsch and Journel, 1998). Unlike the traditional geostatistics tools (i.e. kriging), sequential conditional simulation does not aim at minimizing a local error variance. In addition to honoring the data value, the predictions generated reproduce the sample or experimental statistics such as the histogram and the semivariogram. The output of the simulation is a visual and quantitative measure of the spatial uncertainty (Goovaerts, 1997; 2001; Deutsch and Journel, 1998; Sanchez et al., 2004). Consequently, spatial stochastic simulation is of particular interest where the spatial variation of the measured field must be preserved (Srivastava, 1996). Conditional simulation is based on Bayes's axiom (Goovaerts, 1997) whereby any two points ccdf (conditional cumulative distribution function) can be represented as a product of a one-point ccdf:

$$F(u_1', u_2'; z_1, z_2|(n)) = \text{Prob}\ \{Z(u_1') \leq z_1;\ Z(u_2';) < z_2|(n)\}\ , \qquad \text{(Eq. 6.1)}$$

Where | (n) represents conditioned to n data set $z(u_\alpha)$ and to the realization $Z(u_1')$. In other words the ccdf at the initial location u_1' is conditional to n original data z_1, z_2. The derived ccdf also becomes conditioned to all previous simulated values (Goovaerts, 1997). In other words, in sequential simulation, a ccdf is

modeled at each location and sampled along the random sequence. In order to ensure reproduction of the z-covariance model, at each point a ccdf is created that is conditional not only to the hard data but also all values simulated at locations previously visited (Goovaerts, 1997). At the end of this process, the simulated normal score (raw value minus sample mean divided by standard deviation of the sample) values are back transformed to the original variable. When multiple realizations are needed, the above algorithm is repeated, starting at random initial grid locations and revisiting grid nodes randomly each time. In other words, simulation is a local estimator, which obeys a global spatial correlation function while it attempts to reproduce the global pattern of spatial continuity and global statistics (histogram, covariance) (Goovaerts, 1997).

A sequential simulation of organic carbon was carried out based on data collected at different sites (Figures 6.1, 6.2 and 6.3). The simulated carbon values were constrained to a range of 1 to 15 g kg^{-1}, which were considered physically possible. For each simulated soil carbon surface (Figure 6.7) 276 points were randomly extracted to construct a dataset with carbon values repeated in time. In other words the approach generated a new dataset composed of fixed locations and varying soil organic carbon prediction for selected years (2000 – 2006). Each estimation point represents an average value calculated from a distribution of 1000 realizations. The data set structure is further explained in the next section.

6.2.3. Spatiotemporal Model

Most spatiotemporal approaches have the following general model (Eq. 6.1) (Tilmann et al., 2002).

$$Z(s,t) = m_Z(s,t) + X(s,t), \quad \forall(x,t) \in D \times T, \qquad \text{(Eq. 6.2)}$$

Where Z(s,t) is the variable estimate over space (s) and time (t), $m_Z(s,t)$ the global mean trend and $X(s,t)$ is a zero-mean spatiotemporal random field, which represents the space-time fluctuations around $m_Z(s,t)$. The space and time domain is presented by D and T, respectively. This approach also computes the residuals as defined below:

$$X(s,t) = Z(s,t) - m_Z(s,t), \qquad \text{(Eq. 6.3)}$$

BME functions provide the following additive model for space/time mean trend:

$$m_Z(s,t) = m_s(s) + m_t(s,t), \qquad \text{(Eq. 6.4)}$$

The $m_s(s)$ is the spatial component and $m_t(t)$ is the temporal component. Equation 6.4 models the mean trend $m_Z(s,t)$ of $Z(s,t)$. Equation 6.4 computes the spatial ($m_s(s)$) and temporal mean ($m_t(t)$) components of space/time random field $Z(s,t)$, using measurements at fixed sampling points and time.

The spatial mean component ($m_s(s)$) is obtained by averaging the measurements at each sampling location. Then a smoothed spatial mean component is obtained by applying an exponential spatial filter to ($m_s(s)$). In the same way ($m_t(t)$) is calculated by averaging the measurement for each

measuring event, and a smoothed temporal mean component is obtained by applying an exponential temporal filter to ($m_t(t)$ (Serre, 1999).

6.2.3.1. Space/time covariance models

The covariance function describes important characteristics of a S/TRF, such as its variability, correlation in space and time and its continuity (Journel, 1978; Serre, 1999). For instance a S/TRF is continuous in the mean square sense at the p_o, if and only if, its covariance $c_x(p,p')$ is continuous at the $p=p'=p_o$.

All covariance functions must satisfy the following properties:

$c_x(p,p') = c_x(p',p)$, (Eq. 6.5)

Another property of the covariance function, also known as Schwartz inequality, provides an upper bound for the covariance between any two points in space and time in terms of the local variances:

$c_x(p,p') \leq \sigma_x(p)\sigma_x(p')$ (Eq. 6.6)

A continuous covariance function must be a non-negative definite function that also satisfies the properties defined by Eq. 6.5, and 6.6. A function that verifies the non-negative definite properties is said to be permissible and can be used to model a random field (Serre, 1999; Christakos, 2000).

Depending on the understanding of a particular process the spatial temporal random field (S/TRF) models can be further classified as ordinary and generalized (Christakos et al., 2002; Christakos, 1992). An ordinary S/TRF can be spatially homogeneous and temporally stationary or

nonhomogenous/nonstationary (Christakos et al., 2002; Serre, 1999; Kolovos et al., 2004). By stationarity is meant that the distribution of the random process has certain attributes, usually either the mean or variance, are the same everywhere in the field.

A homogenous/stationary S/TRF has the following properties: (i) mean trend is constant, $m_X(p) = m_X$; (ii) covariance functions are based only on the temporal and spatial lag (r, t, respectively). The nonhomogenous/non-stationary S/TRF model is more realistic for mapping purpose; however, the choice of heterogeneity conditions should be based on the understanding of the processes associated with variable of interest.

Spatial homogeneity (stationarity) is required for making inferences from a model that characterizes the process of the spatial structure of data at unsampled locations (Christakos et al., 2002; Serre, 1999).

The space/time covariance of a homogenous/stationary S/TRF is represented as follows:

$$c_x(s,t\,;s',t') = E[\,(X(s,t) - m_x(s,t))\;\;(X(s',t') - m_x(s',t'))\,], \qquad \text{(Eq. 6.7)}$$

Where, $X(s,t)X(s',t')$ are the expected values associated with the random variable X(s,t) at location s - s' and t - t' and m_x is the mean trend constant.

Several covariance models can be used to model the distribution of a random variable. The first set of covariance functions implemented in this analysis was associated with vector format data. This covariance function was an exponential space-time model with space/time separable shape.

$$c_x(r,t) = c_1 \exp(-3r/a_{r1} - 3t/a_{t1}), \qquad \text{(Eq. 6.8)}$$

Equation 6.8 is defined by the sill (c_1), which represents the positive variance contribution, the spatial range in meters (a_{r1}) and the temporal range in years (a_{t1}). Other separable and non-separable space/time covariance models are also part of the BMElib functions (Christakos et al., 2002).

The experimental space/time cross-covariance for a given lag distance r and temporal lag τ can be computed with the following equation:

$$\hat{c}_x(r,\tau) = \frac{1}{N\,(r,\tau)} \sum_{i=1}^{N(r,\tau)} (X_{-(r,\tau),i}\, X_{+(r,\tau),i}) - m_{-(r,\tau)}\, m_{+(r,\tau)}\,, \qquad \text{(Eq. 6.9)}$$

Where $N(r,\tau)$ is the number of pairs of points with values $(X_{-(r,\tau)}\, X_{+(r,\tau)})$ separated by a distance of r and a time of τ and $m_{-(r,\tau)}$ is the mean of the $X_{-(r,\tau),i}$ values and $m_{+(r,\tau)}$ is the mean of the $X_{+(r,\tau),i}$. A tolerance associated with spatial (dr) and temporal ($d\tau$) lag was established as follows:

$$r - dr \leq ||s_{-(r,t)} - s_{+(r,t)}|| \leq r + dr \text{ and } \tau - d\tau \leq ||t_{(r,t)} - s_{+(r,t)}|| \leq \tau + d\tau, \qquad \text{(Eq. 6.10)}$$

Where, $||.||$ denotes the Euclidean distance or the ordinary absolute distance between two spatial or temporal classes of experimental covariances. The tolerances dr and $d\tau$ are based on absolute differences between spatial ($s_{-(r,t)}$-$s_{+(r,t)}$) and ($t_{(r,t)}$-$s_{+(r,t)}$) classes of experimental covariances. In other words, for each classes (i.e. $s_{-(r,t)}$-$s_{+(r,t)}$) a +/- range is established based on the classes distance r and a time of τ. Eq. 6.9 describes the cross-covariance estimates corresponding to different spatial and temporal lags.

6.2.3.2. Space/time data formats in BMElib

There are two basic approaches of dealing with spatiotemporal data in BMElib depending on the spatial component of the sampling structure:

(i) Space/time vector format: Data collected irregularly in space and time is organized in BMElib as space/time vector format. That is, each data point is characterized by a specific spatial and temporal component. In space time vector format the coordinates (x,y) and time are stored in one vector and the measured values (z) are associated with locations and time. The data set can be represented in BME with two matrices (i.e. ch[x,y,t] and z[]) as shown in the following example (Figure 6.4).

ch[x,y,t]			z[soc]
x	y	t	SOC
739270.8	1428593	1	3.8
739186.5	1428672	1	3.2
739220.8	1428686	1	2
739250.6	1428673	1	2.8
739533.6	1429083	3	2.29
739510.2	1429022	3	1.22
739507.3	1428950	3	1.87
739535.6	1428879	3	1.39
739683.2	1428585	4	3.01
739599	1428576	4	6.19
739474	1428529	4	1.96
739379.1	1428525	4	2.18

Figure 6.4. Example of matrices of space-time vector format dataset in BMElib.

(ii) Space/time grid format: The grid format has fixed spatial locations and a specific time for every sampling period. In grid format the coordinates of the sampling points are stored by a matrix (x, y), the time is stored in a vector (1 x number of sampling period, and the space/time data is stored in a Z matrix (# of sampling points by # sampling period) as shown bellow (Figure 6.5):

ch[]	
x	y
771.705	459.7756
717.705	449.7756
781.705	449.7756
671.705	443.7756

z[]			
SOC			
1.4	1.2	1.5	1.8
1.5	2.5	1.5	2
1.4	1.1	1.5	2
1.6	2.3	1.8	2.5

t[]	T	1	2	3	4

Figure 6.5. Example of matrices of space-time grid format dataset in BMElib.

6.2.3.3. Soil Organic Carbon Accounting

The calculation of the total organic carbon in the field was done as follows:

[Total organic carbon] = [BME prediction map]x[soil bulk density]x[soil depth]

The amount of soil organic carbon was predicted for the entire field based on soil samples collected between 0 – 20 cm. A bulk density of 1.6 g cm^{-3} was assumed for all locations based on estimates of Kablan et al. (2008). This calculation procedure was produced using the "map calculator" algebraic operations in ArcGIS. The same procedure can be carried out in Matlab.

6.3. Results

Table 6.1 summarizes the soil organic carbon measurements made at Siguidolo, Oumarbougou and Sikasso between 2000 and 2006. The median soil carbon for the upper and lower layer (0-20 and 20-40 cm, respectively) decline with time at Siguidolo as compared with the other sites (Oumarbougou and Sikasso).

Table 6.1. Summary statistics for measured soil organic carbon, 0-20 and 20-40 cm.

	N	Mean	Median	Std. Dev	σ^2	Skewness	Kurtosis
Siguidolo/Zan Diarra Carbon (g kg^{-1}) [0-20cm]							
2000	14	2.69	2.80	0.91	0.84	-0.33	-1.10
2004	35	2.34	1.98	1.05	1.10	1.90	4.07
2006	67	2.16	2.01	0.71	0.50	3.23	15.81
Oumarbougou/Mory Konate Carbon (g kg^{-1}) [0-20cm]							
2000	16	2.85	2.80	0.68	0.46	2.19	6.85
2002	38	3.15	3.00	0.74	0.54	0.63	0.15
2004	49	3.24	2.85	1.24	1.53	1.61	2.77
2006	42	7.56	8.00	2.16	4.68	-0.44	-0.60
Sikasso/Yaya Diassa Carbon (g kg^{-1}) [0-20cm]							
2000	17	4.26	4.20	1.85	3.43	0.90	0.77
2004	75	5.06	4.88	1.50	2.22	0.72	0.39
2006	62	5.45	4.82	2.23	4.98	1.23	2.45

	N	Mean	Median	StDev	σ^2	Skewness	Kurtosis
Siguidolo/Zan Diarra Carbon (g kg^{-1}) [20-40cm]							
2000	14	2.69	2.70	0.71	0.50	-0.30	1.22
2004	35	2.50	2.31	0.74	0.53	1.24	1.19
2006	67	2.26	2.05	0.76	0.57	4.84	30.07
Oumarbougou/Mory Konate Carbon (g kg^{-1}) [20-40cm]							
2000	16	2.53	2.25	0.77	0.59	1.70	2.31
2002	38	2.79	2.47	0.81	0.65	1.00	0.31
2004	49	3.06	2.73	1.03	1.06	1.04	0.06
2006	42	5.13	4.70	1.75	3.06	1.08	0.88
Sikasso/Yaya Diassa Carbon (g kg^{-1}) [20-40cm]							
2000	17	3.43	3.00	1.37	1.86	0.86	- 0.25
2004	75	4.30	4.13	1.03	1.06	0.61	0.59
2006	62	5.12	4.23	6.83	46.66	7.63	59.38

N is the number of observations, Std. Dev is the standard deviation, and σ^2 is the variance. Skewness is a measure of symmetry while Kurtosis characterizes the distribution.

A sharp increase in soil organic carbon is reported at Mory Konate's field from an average of 3 g kg^{-1} in 2004 to 8 g kg^{-1} in 2006. A similar trend was also measured in Yaya Diassa's farm field; however, the changes were not statistically significant. As expected carbon in the lower depth (20-40 cm) was lower corroborating the results found in the literature (Brady and Weil, 2002).

6.3.1. Space-time modeling

Figure 6.1 displays the soil sampling locations in Mory Konate's field and the respective magnitude of soil carbon at those locations over four different time periods. Soil carbon is modeled to estimate the spatial/temporal random field. The estimated values of $c_x(r, t)$ (Eq. 6.9) are shown in Figure 6.6 as a function of the spatial lag (r) and temporal lag (t).

The range of spatial dependence from Mory Konate's data set was relatively smaller (80 m) which denotes that for sample points collected at distance greater than the specified range are independent (Figure 6.6). However, the temporal covariance of this dataset extends to 6 years, a period longer than other sites.

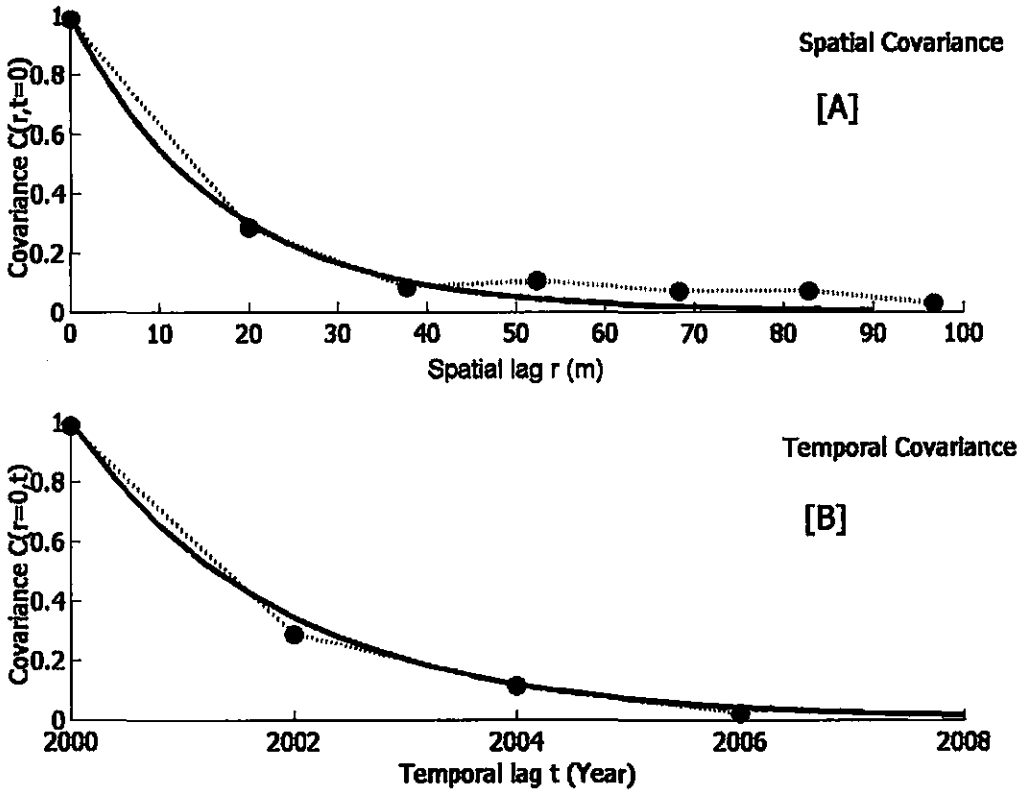


Figure 6.6. Space/Time covariance of soil carbon as a function of spatial lag r – [A] and time lag t (2000, 2002, 2004, 2006) - [B] for Mory Konate's field. The solid lines are the fitted covariance models.The calculated space/time covariance values from actual measurements are plotted with markers.

Yaya Diassa's field results suggest that the spatial covariance remained the same for a period of four years with a spatial dependence range of 300 meters (Figure 6.7). Yaya Diassa's field is located in Southern Mali and this data set covered the period of 2000, 2004 and 2006. The first sampling period is characterized by a small data set (Figure 6.3). The last two data sets (2004 and 2006) have a higher intensity sampling, hence a better representation of the spatial distribution of the soil carbon (Figure 6.3).

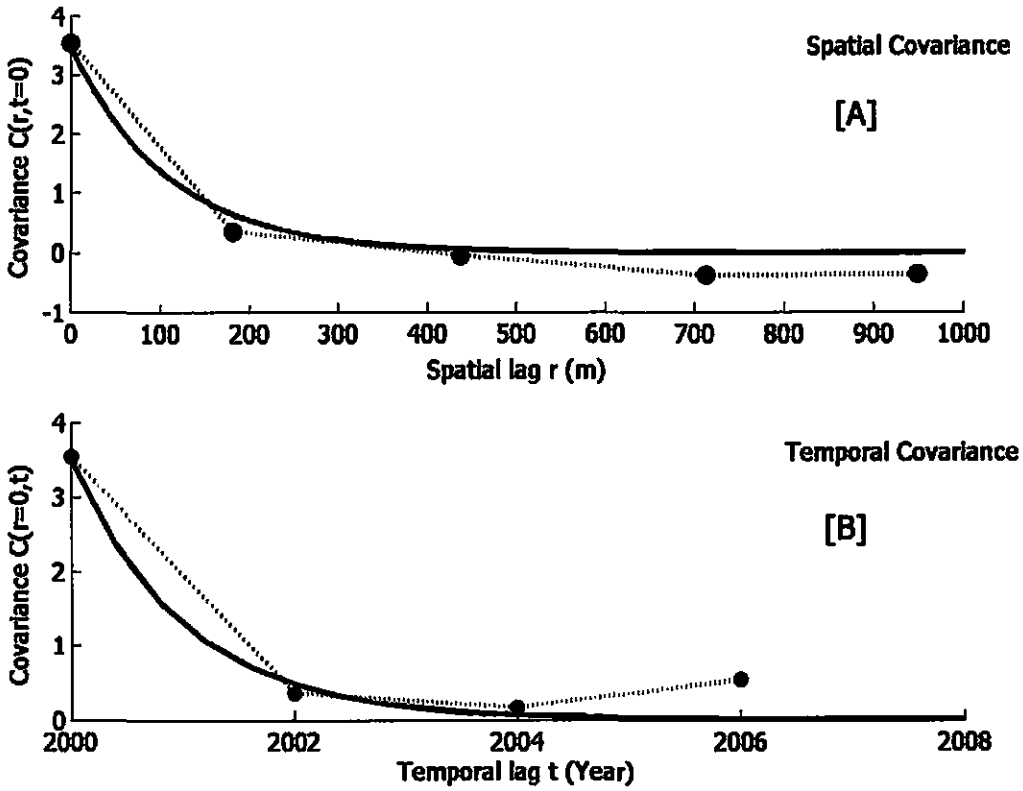


Figure 6.7. Space/Time covariance of soil carbon from Yaya Diassa's Field as a function of spatial lag r – [A] and time lag t (2000, 2002, 2004, 2006) – [B]. The solid lines are the fitted covariance models. The calculated space/time covariance values from actual measurements are plotted with markers.

The spatial dependence at Zan Diarra's field has a range of 100 meters. The temporal covariance indicates that the temporal correlation continues for four years (Figure 6.8). Once the temporal covariance reach zero the spatial covariance can no longer be regarded as a reliable model to describe soil carbon in this field. Based on these results the sampling frequency for this site should not be greater than 4 years (Figure 6.8).

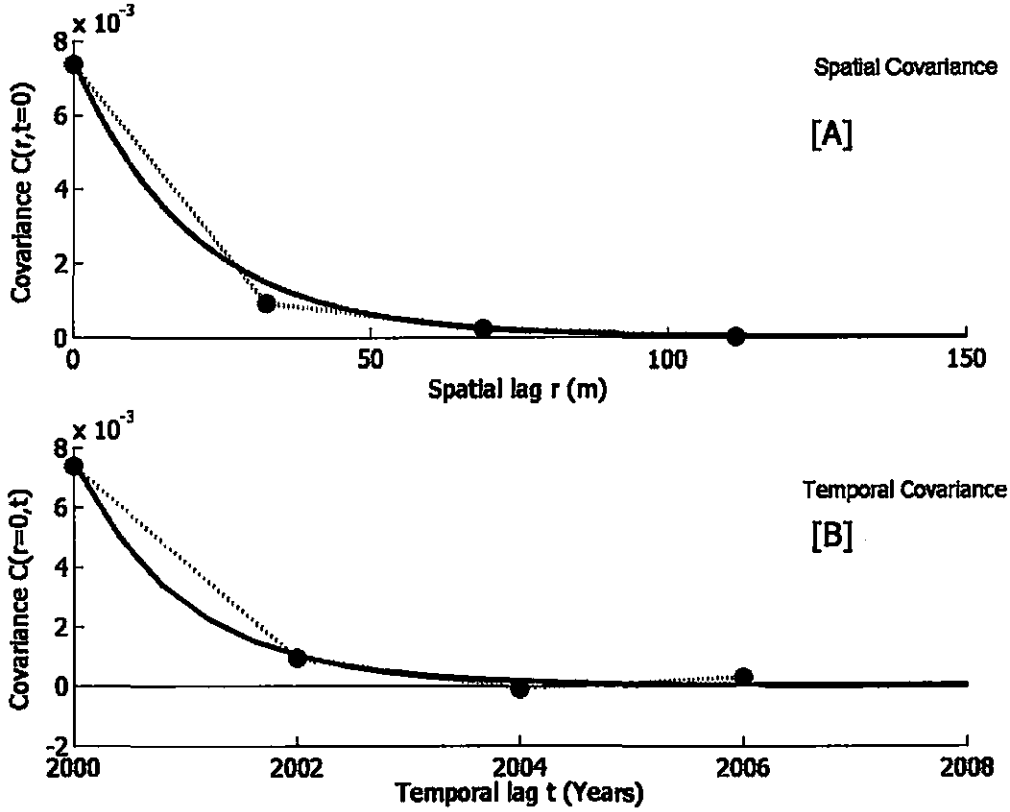


Figure 6.8. Space/Time covariance of soil carbon in Zan Diarra's field as a function of spatial lag r – [A] and time lag t – [B]. The solid lines are the fitted covariance models. The calculated covariance values from actual measurements are plotted with markers.

The simultaneous modeling of both spatial and temporal covariance of soil organic carbon in three distinct zones of Mali, West Africa was reported above (Figures 6.6, 6.7, 6.8). The temporal trends indicate that in all studied sites there were dynamic spatial changes of soil carbon and that the spatial pattern revealed during the first sampling period was mostly likely to change after four years. This trend analysis is critical for management and for verification and monitoring of soil carbon for carbon trading. The temporal covariances provided a relevant insight on the persistence of the spatial model.

6.3.2. **Space–time prediction**

This section investigates the temporal changes of soil carbon in Mory Konate's field. The boxplot (Figure 6.9) shows the basic statistics for each of the simulated data sets. A few observations fell beyond the interquartile range and hence are marked as potential outliers.

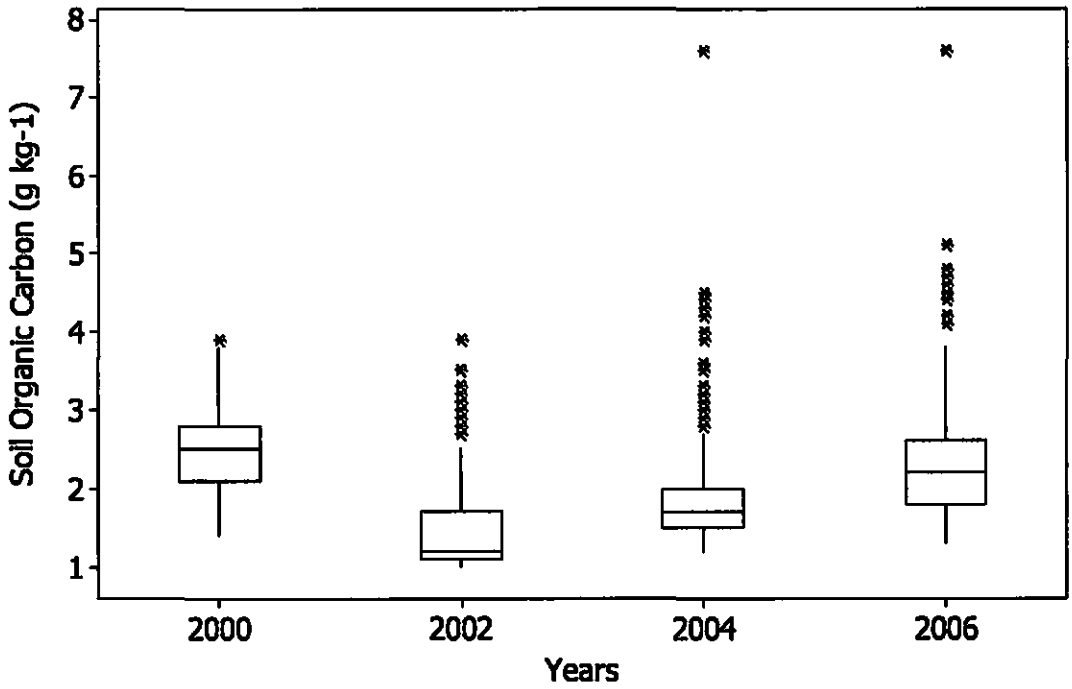


Figure 6.9. Boxplot of the simulated values of soil carbon in Mory Konate's field.

The conditional simulation generates a map or realizations of z-values which tend to reproduce a more realistic representation of a regionalized variable than kriging estimates. Note that the simulation yielded many values beyond the interquartile range and considered as outliers (Figure 6.9). Despite the positive skewness of the data set our decision was to work with raw data instead of carrying the analysis with the log transformation data. Data transformation

reduces the influence of the extreme values and may unduly reduce the influence of measured data (Goovaerts, 1997).

The simultaneous modeling of space and time considers the spatial lag variation but also takes into account the temporal covariance associated with the soil carbon in this particular field. The space/time covariance provides an indication of the spatial and temporal correlation of soil organic carbon (Figure 6.6). The covariance model is of interest for soil sampling purposes, both in terms of sampling density and sample configuration. The temporal covariance, on the other hand, provides the information regarding the duration of a particular spatial pattern. The maximum temporal lag of 6 years as suggested by Figure 6.6-[B] indicates that beyond this time frame one might need to collect new samples in order to get a reasonable assessment of soil organic carbon.

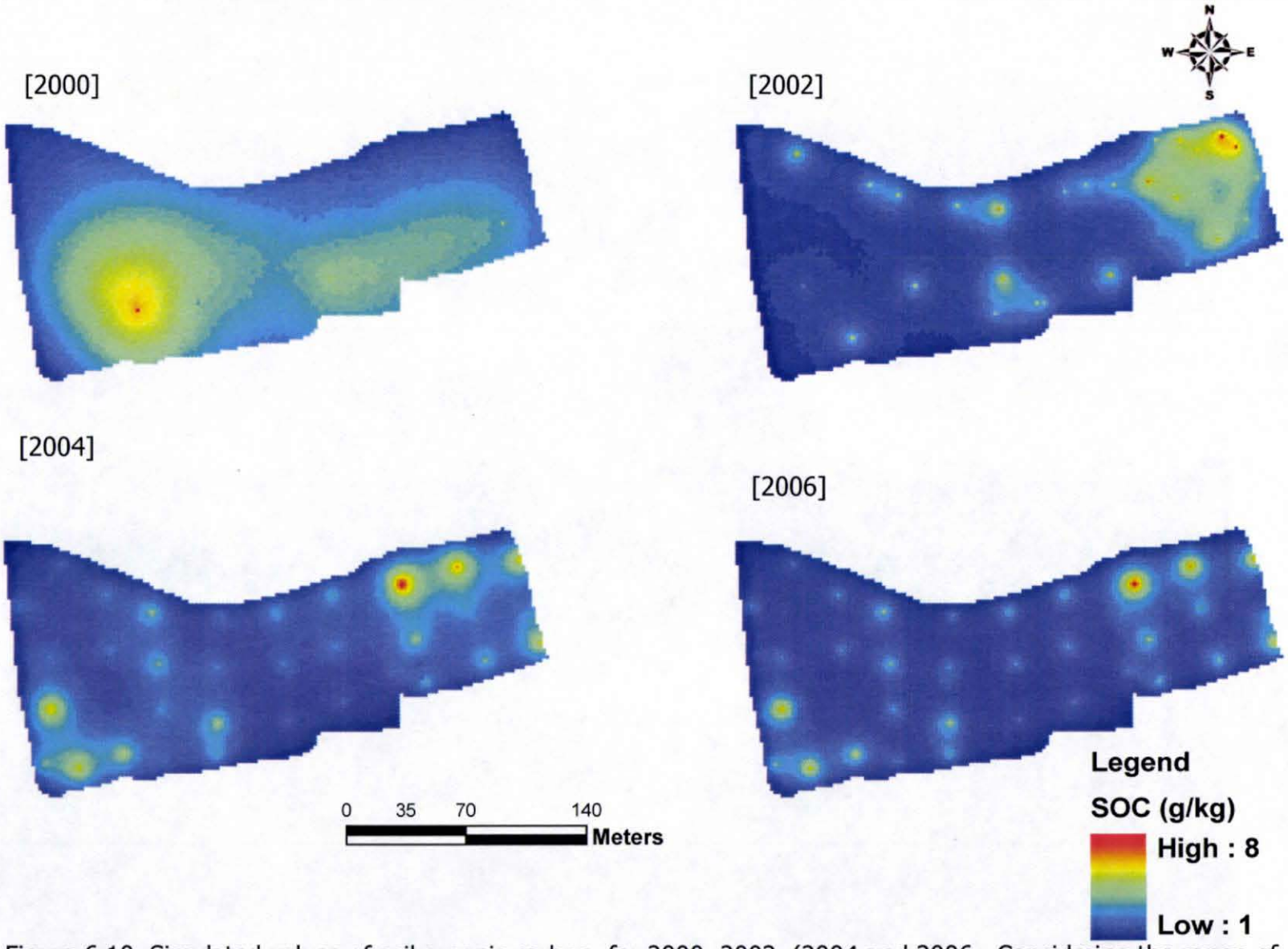


Figure 6.10. Simulated values of soil organic carbon, for 2000, 2002, (2004 and 2006. Considering the range of carbon normally found in the region the simulations were constrained to range between 1 to 15 g kg^{-1}. The markers show the location were samples were collected.

The raw spatial mean is obtained by averaging carbon values for each year. The mean trend (Eq. 6.4) is calculated and removed from the data. The smooth temporal mean component is generated with the help of an exponential temporal filter. Although the overall trend (Fig. 6.11) indicates a sharp decrease of soil carbon after two years (2000 - 2002), the spatial analysis of mean trend at each location indicates that the decrease in soil carbon is particularly accentuated in certain areas of the field (Fig. 6.12). The area with soil organic carbon exceeding 2.2 g kg^{-1} is closer to the homesteads, however the overall increase is attributed to ACN practices implemented to help improve soil water availability and land productivity.

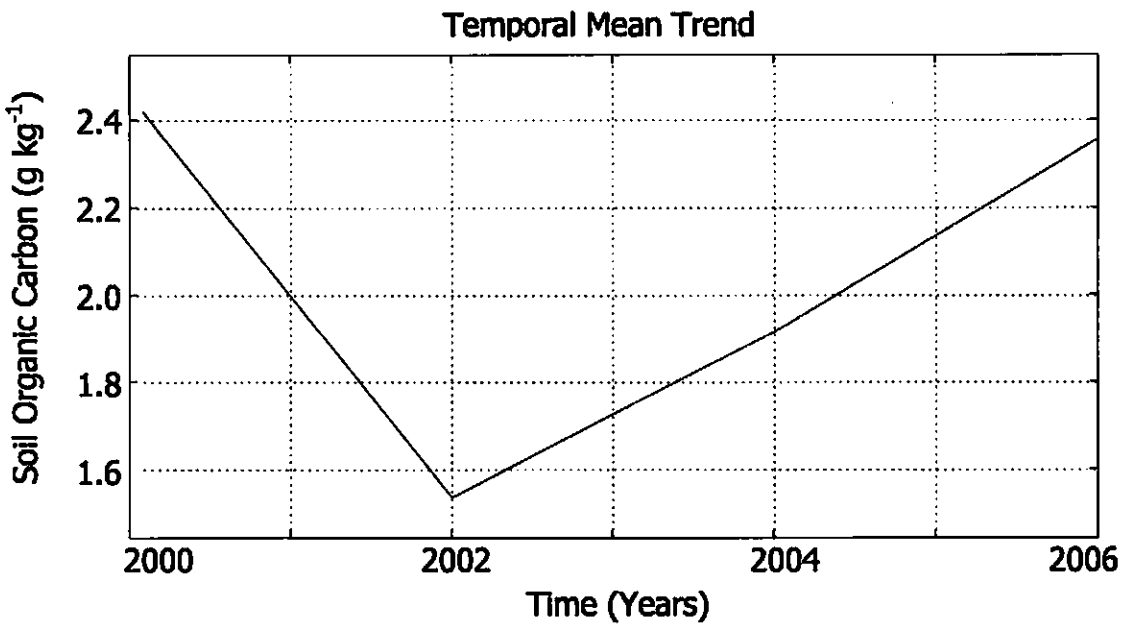


Figure 6.11. Temporal mean trend on soil carbon in Mory Konate field. Note that from 2000 to 2006 C is increasing, while the spatial mean trend suggests that the soil carbon build up is higher in certain areas.

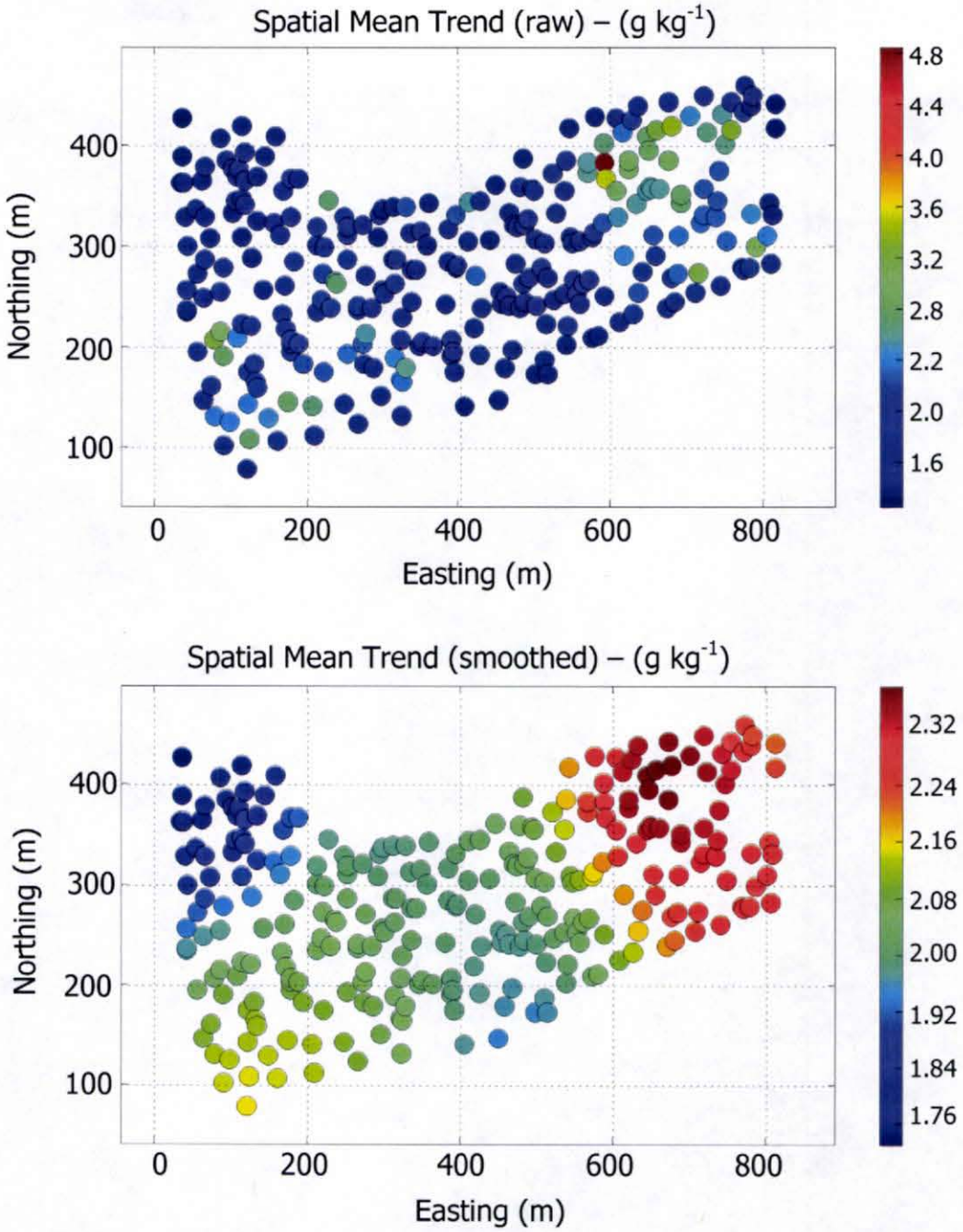
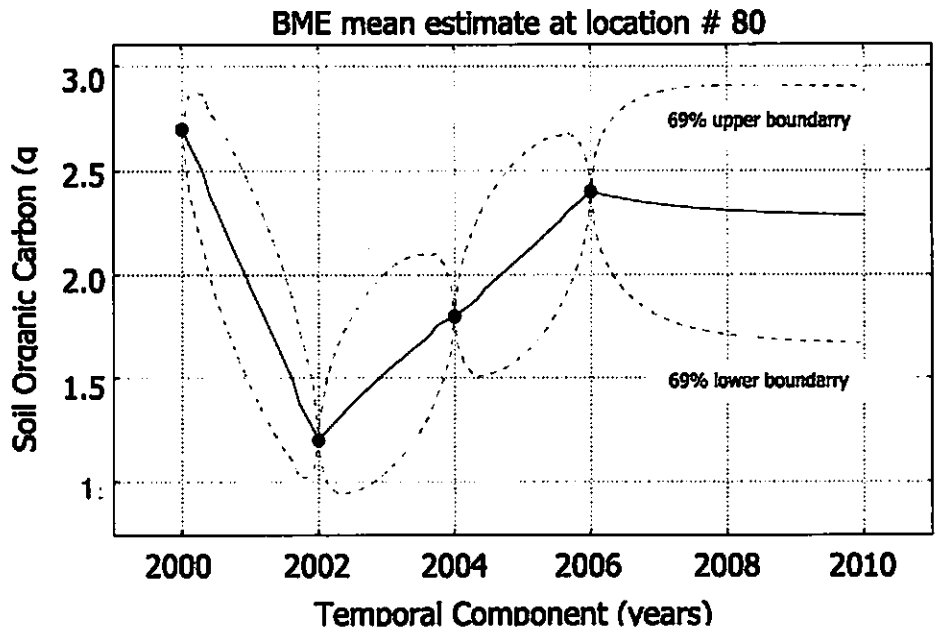


Figure 6.12. Modeling space time mean trends, in Mory Konate (2000-2006). Spatial mean trend (raw) is presented on top, followed by smoothed spatial mean trend.



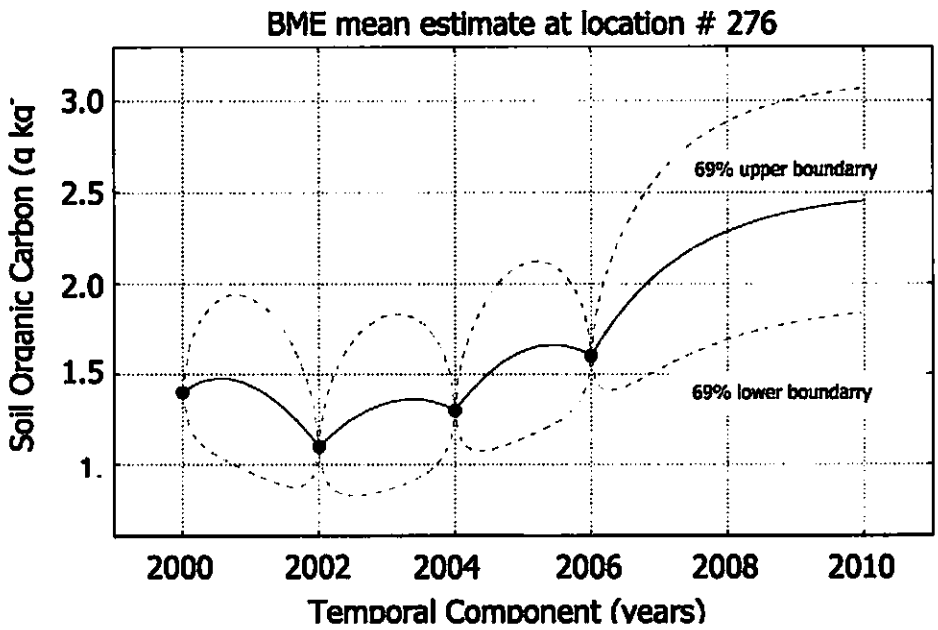


Figure 6.13. Spatial and temporal trend of soil carbon in Mory Konate's field. Temporal profiles at two locations. The solid line represents the BME estimate and dotted lines are the 69% lower and upper bound of the BME mean estimate. The marks (circle) indicates measured data.

The temporal analysis shows the changes occurring at two locations between 2000 and 2012 (Figure 6.13). Note that soil carbon was monitored with sampling interval of two years from 2000 to 2006 (Fig. 6.13). The trend suggests a soil carbon build up over time. At each location a confidence interval was also plotted indicating the +/_ one standard deviation around the BME mean estimate over time. Figure 6.13 indicates that for predicted values beyond sampling period the variance increases as does the confidence interval. This is due to the usually high uncertainty associated with forecasting future events.

The capability to extrapolate and produce estimates of soil carbon at an unsampled time in the future demonstrate the relevance and applicability of this approach in monitoring long-term projects of carbon sequestration (Figures 6.14 and 6.15). The approach also provides the estimates of the uncertainty associated with forecasting results based on past measurement trends.

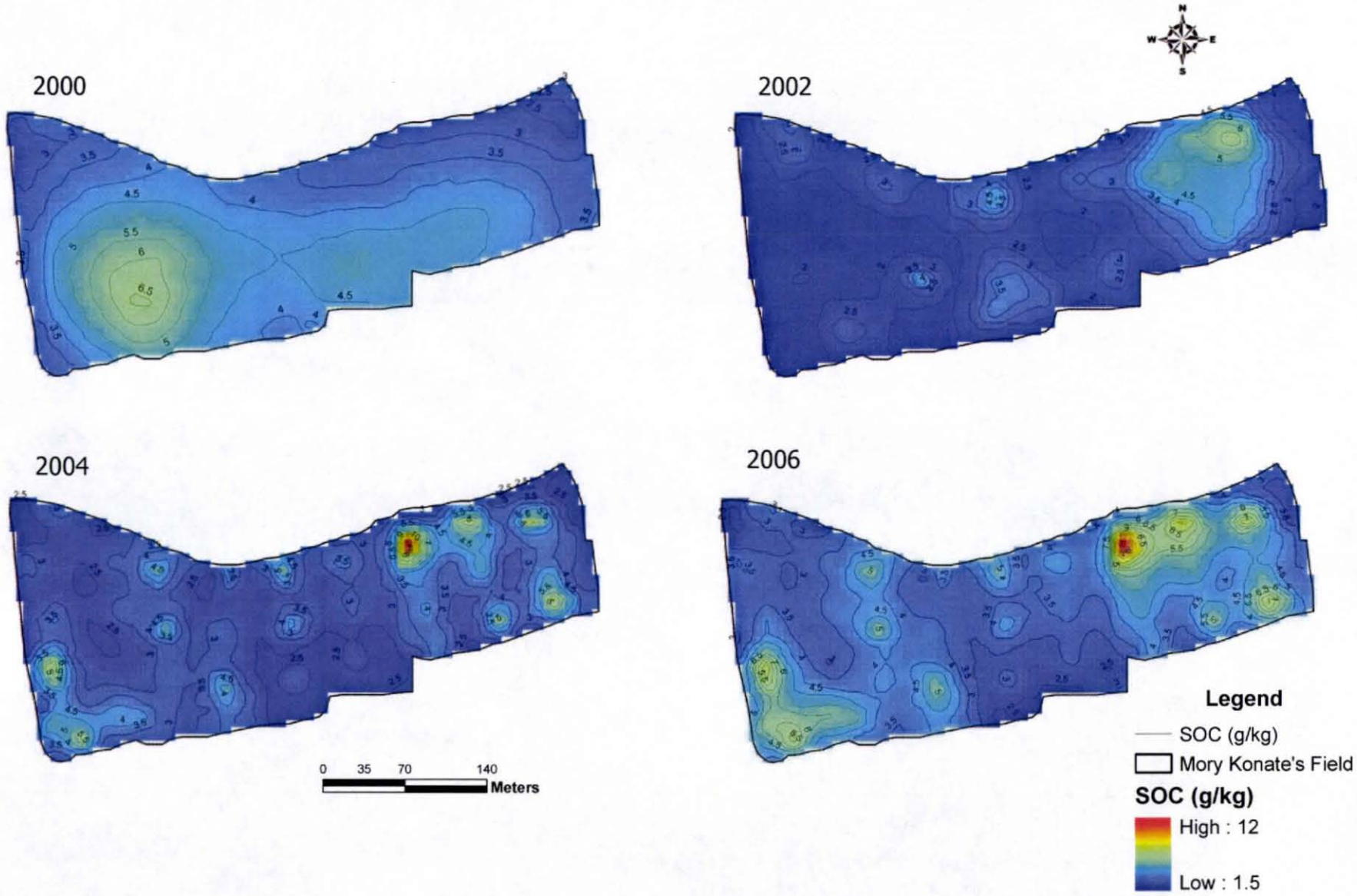


Figure 6.14. BME estimates of soil organic carbon at Mory Konate's Field in 2000, 2002, 2004 and 2006 ranging from 1.5 – 12 g kg^{-1}.

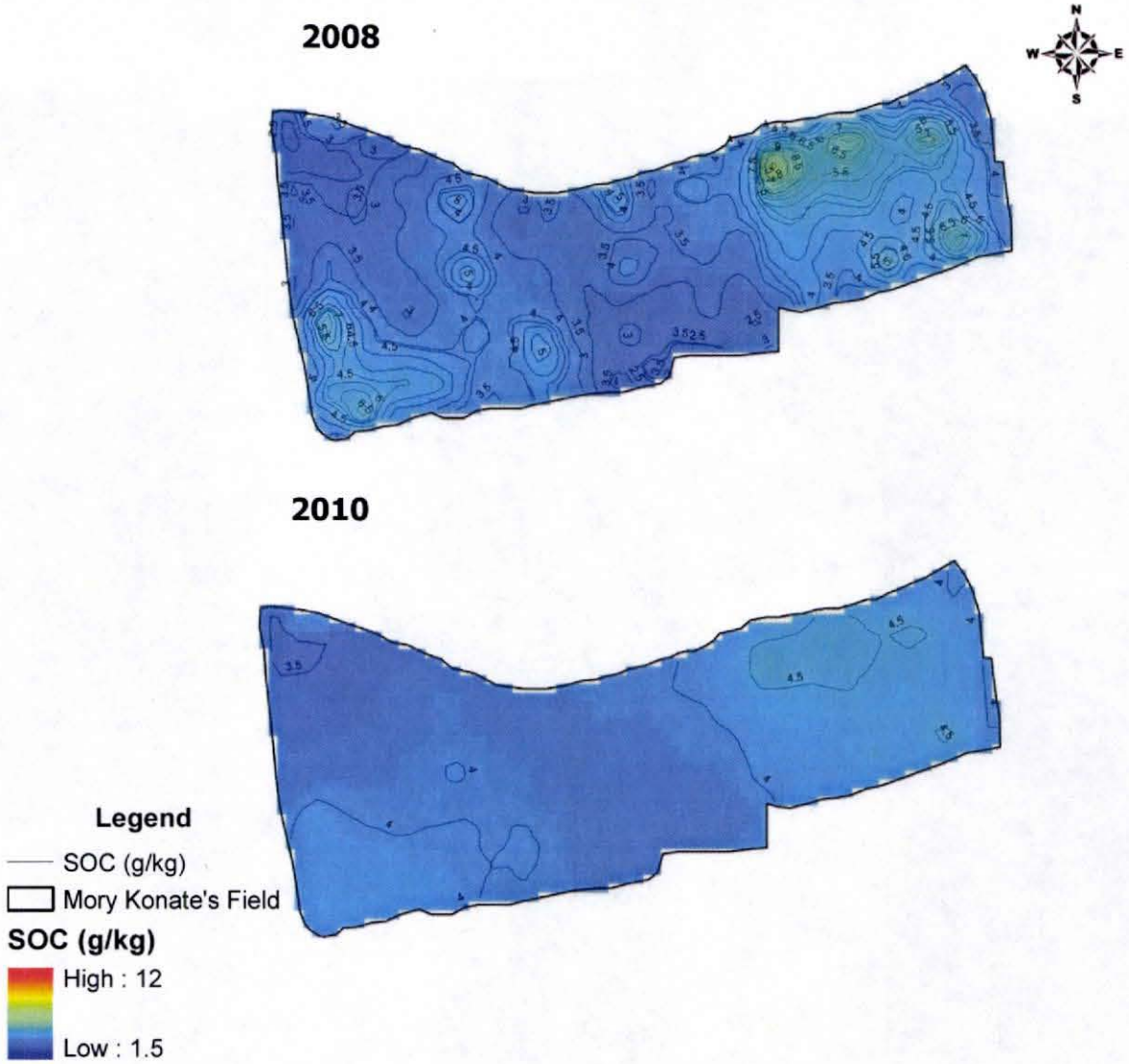


Figure 6.15. BME spatiotemporal forecasting of soil organic carbon, 2008 and 2010. The forecasted maps provided a global assessement based on the spatiotemporal mean trend. There is an overall increase with a substantial smoothing of predicted values towards the local mean.

The BME prediction of spatial temporal soil carbon in Mory Konate's fields is presented below. The spatial pattern of high soil carbon maybe be due to agricultural management practices (i.e. crop rotation, residue management and soil conservation practice) (Figures 6.14 and 6.15). The estimates of overall carbon means and ranges increase during the study period (Figure 6.13).

As expected, the carbon build up as shown by the spatial distribution is not uniform. Certain regions of the field appeared to accumulate more carbon than others (Figure 6.14 and Figure 6.15). For management purposes,the mapping of soil carbon using spatial distributed approach provides relevant information for soil fertility management (Fig. 6.14 and 6.15).

The forecasted soil carbon for 2008 and 2010 (Fig. 6.15) is based on space-time covariance model of soil carbon (Figure 6.6). The mean and confidence intervals of the average soil organic carbon estimates for the field suggest an initial decrease and then an increase and stabilization (Fig.6.15)

The extrapolation beyond the time range shows no change in overall carbon, even though the maps indicate spatial changes (Fig. 6.16). This might represent a "steady state" soil carbon status as a result of management practices over time (i.e. ACN).

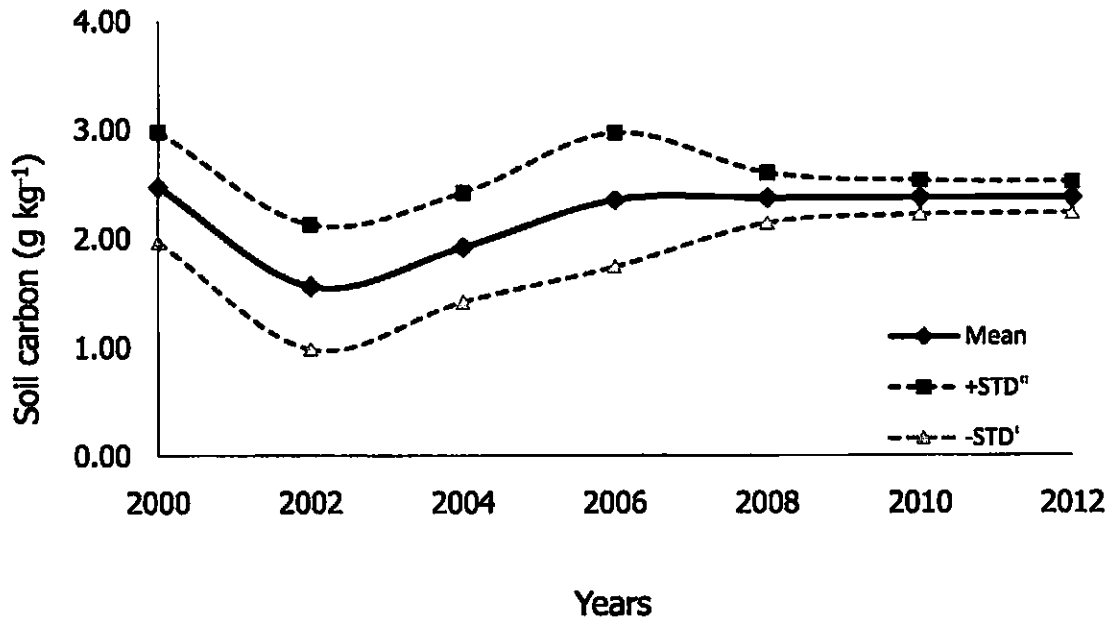


Figure 6.16. BME mean estimate for the entire Mory Konate field. This is the global mean for prediction maps from year 1 (2000) to year 7 (2012).

The BME method also provides an assessment of the error variance of the predicted value (Table 6.2). The average of the variances for the forecasted years are the same and the range between the maximum and minimum values as well as the standard deviation indicate less variability in the forecasted mean estimate for the entire field.

Table 6.2. Summary Statistics of Variance estimates maps (g kg^{-1}) shown in Figures 6.12 & 6.13.

	BME Variance Estimates / Year						
	2000	2002	2004	2006	2008	2010	2012
Min	0.14	0.17	0.21	0.23	0.32	0.34	0.34
Max	0.39	0.64	1.12	1.16	0.61	0.48	0.46
Range	0.25	0.48	0.91	0.93	0.29	0.14	0.11
Mean	0.25	0.26	0.32	0.40	0.40	0.40	0.40
St. Dev.	0.05	0.01	0.09	0.10	0.04	0.03	0.02

6.3.3. Quantifying soil organic carbon

The 2008 and 2010 estimates are forecasted values based on the spatial-temporal trends (Figures 6.16 and 6.17). The results for Mory Konate indicate a net gain of soil organic carbon from 2002 to 2006 (Figure 6.17). The 2000 prediction appears to be an overestimate. This overestimate may be due to sampling configuration and the measured soil organic carbon. The increase in soil organic carbon from 2002 to 2004 was estimated at 126 Mg, representing a 23% increase. From 2004 to 2006 the gain was evaluated as 153 Mg (23% increase).

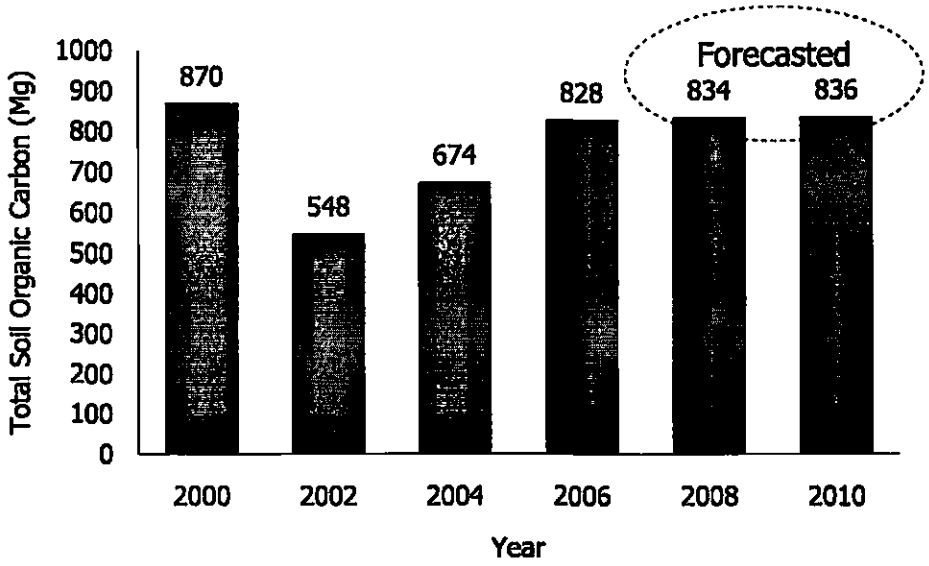


Figure 6.17. The total organic carbon calculated for Mory Konate's field. It should be pointed out that the 2000 dataset yielded far greater total carbon the subsequent years. This could be an artifact caused by limited number of samples and a different sample configuration.

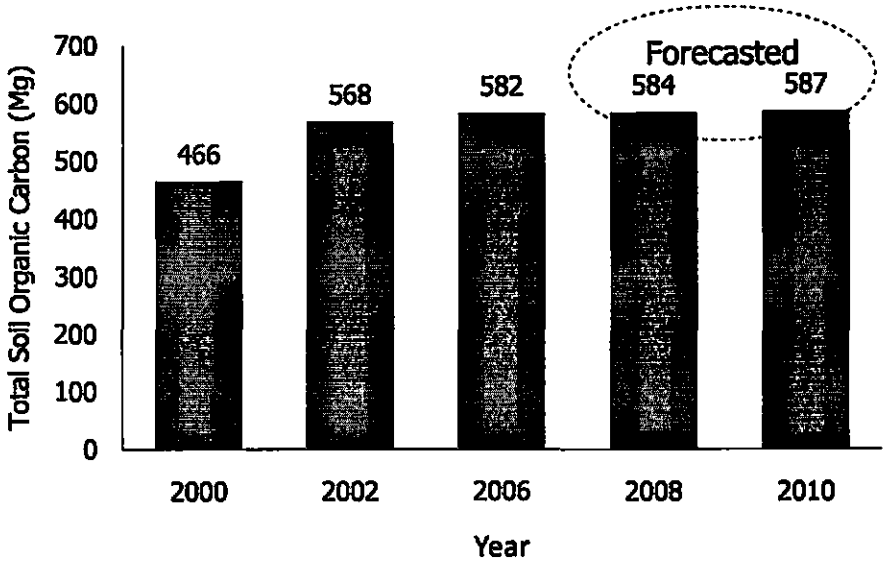


Figure 6.18. The total organic carbon calculated for Yaya Diassa's field.

The total amounts of soil carbon computed in 2000, 2002 and 2006 were 466, 568, 582 Mg, respectively (Figure 6.18). At Yaya Diassa's field an increase of 103 Mg of soil organic carbon was measured for the period of 2000 to 2002. The net increase from 2002 to 2004 was 14 Mg.

6.3.4. Accounting for Tree Biomass in Agroforestry Systems

The estimation of tree biomass in Sub-Saharan Africa is important for a complete assessment of the potential of these systems to capture and store carbon. In the Sahel trees play a key role in nutrient recycling. Several authors have reported the high importance of agroforestry systems in the Sub-Sahara region (Picard et al., 2006; Traoré, 2003; Boffa et al., 1998; Berthe, 1997). Certain trees are selectively allowed to grow, due to their economic, social and even spiritual importance while other trees may be spared for shading or

removed and used as fuel. The spatial heterogeneity of different farm fields can have a significant impact on carbon stocks and productivity of the land. Woomer et al. (2001) proposed laboratory and field guidelines to measure soil and aboveground biomass in West Africa.

One particular characteristic of the Sudano-Sahelian region is the dispersed spatial occurrence of trees and a diversity of shrubs (Traoré, 2003; Picard et al., 2006). The low input production systems that include mainly millet, sorghum, maize and cotton coupled with livestock grazing places huge pressure on crop residue. According to Tiessen et al. (1998) the average value of soil carbon in tropical savanna system is around 25 Mg C ha^{-1}. In general these fields are characterized by low carbon, and their potential to act as a major sink of soil carbon is always conditioned by the management practices and the climatic conditions. Not surprisingly, the lack of water tends to limit land productivity and tends to reduce soil carbon buildup.

6.3.4.1. Biodiversity and Evenness

Several indices can be used to characterize the biodiversity of the agroforestry systems of the farms in this carbon survey. Biodiversity indices are used depending on the objectives and circumstances of the area being studied. The Shannon index was computed as follows:

$$H' = -\sum_{i=1}^{s}(p_i \ \ln(p_i)) \ , \qquad \text{(Eq. 6.11)}$$

Shannon's index accounts for both abundance (S, number of species) and "evenness" of the species found at each site. Where "evenness" is a measure of how equal the populations are numerically. For example an evenness of 1 indicates that different species are in equal number. The proportion of species *i* relative to the total number of species (p_i) is calculated, and then multiplied by the natural logarithm of this proportion ($\ln p_i$). The resulting product is summed across species, and multiplied by -1 (Krebs, 1989) (Eq. 6.11).

Yaya Diassa's farm is considered the most humid of all sites with 22 different species and a Shannon index of 0.69 (Table 6.3); however the evenness is low (0.11) indicating that only a few species are the most dominant. Indeed at this site *V. paradoxa* and *P. biglobosa* makeup 93.3% of the total trees in the field. The combined inventory of all sites indicates that *Vitellaria paradoxa* (Karité) represents about 78.5% of the total trees surveyed.

Table 6.3. Tree diversity index and evenness in different fields

Fields	Shannon - Wienner Index (H'_{si})	Evenness (E_{si})	N^0. Trees	N^0. Different Species
Drissa Traoré	0.68	0.24	17	3
Baracoro Ballo	0.75	0.15	149	14
Andre Dembele	0.70	0.18	47	3
Yaya Diassa	0.69	0.11	465	22
Mory Konate	0.57	0.10	276	9
Zan Diarra	1.81	0.40	97	19

Zan Diarra's field has the highest diversity index (1.81), considering the total number of trees and species. The evenness of the species distribution is

still low but higher than at any other site. At this site the dominant species in decreasing order were *V. paradoxa* (karité), *Adansonia digitata* (Baobab), and *Faidherbia albida* (*Acacia albida*) (Table 6.4). Unlike the Sikasso site (Yaya Diassa), Zan Diarra's field is located in the Segou region (Sahel) were the rainfall is historically lower than at other site (600-800 mm). Nevertheless, 17 different species were recorded in this field (Table 6.1). Farmer's awareness in preserving the spontaneous growth of certain species may be the contributing factor for the increased biodiversity index in Zan Diarra field's.

At Mory Konate's the diversity index and evenness was the lowest (0.57 and 0.1, respectively). Out of 276 trees only nine different species were found at this site. Karité (*V. paradoxa*) was the most dominant species (240). Karité (vernacular name given in Mali), also known as the shea butter tree, has a significant impact on the economy of individual family. The shea butter nut extract is widely used in cosmetics and also as cooking oil in West Africa. In addition, the processing of shea butter oil is a village women's industry.

Table 6.4. Tree species list of individual farms.

Scientific Names	Zan Diarra	Mory Konate	Yaya Diassa	Andre Dembele	Baracoro Ballo	Drissa Traoré
Adansonia digitata	22	1	5	-	2	-
Pericopsis elata	-	-	1	-	-	-
Afzelia africana	2	-	-	-	-	-
Anogeissus leiocarpus	1	-	-	-	-	-
Annona senegalensis		-	1	-	-	-
Azadirachta indica	2	-	1	-	-	-
Aegle marmelos		-	-	-	2	-
Bombax constantum	1	-	-	-	-	-
Cordyla pinnata	1	-	-	-	-	-
Daniellia oliveri		-	1	-	-	-
Eucalyptus sinensis		-	6	-	-	-
Faidherbia albida	12	-	-	-	-	-
Ficus spp	1	-	2	-	1	-
Isobernia doka	-	-	1	-	-	-
Kaya senegalensis	-	1	1	-	-	-
Lannea acida	2	2	-	-	2	1
Lannea microcarpa	4			-	4	
Gmelina arborea	-		1	-	-	-
Mangifera indica	-		1	-	-	3
Mitragina specosia	-			-	3	-
Parkia biglobosa	-	21	49	14		-
Prosopis africana	1		-	-	-	-
Sclerocarya birrea	6	2	-	-	1	-
Sterculia setigera	3	4	-	-	-	-
Sterospermum kunthianum	7		-	-	-	-
Tamarindus indicus	1	4	-	-	-	-
Terminalia avicenodes	1	1	1	-	-	-
Terminalia erinaceus	1	-	1	-	-	-
Terminalia velutina	1	-	-	-	-	-
Vitellaria paradoxa	28	240	385	32	127	13
Ziziphus mauritiana	-	-	2	-	-	-
Species Number	19	9	16(22)*	2(3)*	8(14)*	3
Total N. of Trees	97	276	459(465)*	46(47)*	142(149)*	17

* Species with only known vernacular names were removed from this list. The numbers in parenthesis represent the total number of species and trees.

6.3.4.2. Allometric Models

Tree biomass measurement is normally carried out either by destructive methods, where trees are cut and weighed or by nondestructive techniques in which an allometric parameter related to weight (such as DBH, crown area) has been calibrated using a subsampling of destructively measured parameters and then related to biomass by an allometric equation. In either case, some destructive biomass samples are inevitable.

Allometric equations for tree biomass quantification of certain tropical forest species can be found in the literature. Brown (1997) proposed biomass regression equations for specific climatic range (see Eq. 6.13, 6.14). Louppe (1995) developed several equations for the main wood species trees in northern Cote d'Ivoire derived from the linear relationship between total volume and circumference (C) at breast height (Eq. 6.9). For example, for a common species found throughout the West Africa, *V. paradoxa* Gaertn. f. (Karité) the proposed equation was:

$$\text{Total Volume } (m^3) = -0.0735 + 0.7499\, C^2 , \qquad \text{(Eq. 6.12)}$$

Louppe's (1995) allometric equation provides an estimate of the wood volume (Eq. 6.12). Using a species specific wood density for each species the biomass was estimated.

Brown (1997) provided regression equations to estimate the biomass (kg) using only the DBH. The Brown equation was valid for a limited range of DBH

and climatic zones based on the amount of rainfall. The following equations were proposed:

$$Y = \exp\{-1.996 + 2.32\ln(D)\} \quad \text{(Dry, rainfall <900 mm yr}^{-1}\text{)}, \qquad \text{(Eq. 6.13)}$$

$$Y = 4.69 - 12.8008D + 1.241D^2 \quad \text{(Moist, rainfall >900mmyr}^{-1}\text{)}, \qquad \text{(Eq. 6.14)}$$

where Y is biomass per tree in kg, D is the diameter at breast height. The polynomial functions, however are not applicable to all range of diameter (4-40 cm) measurements (Eq. 6.13, 6.14).

Brown's (1997) equations account for critical factors that affect tree biomass yield (i.e. climate) and use the tree diameter as an input parameter.

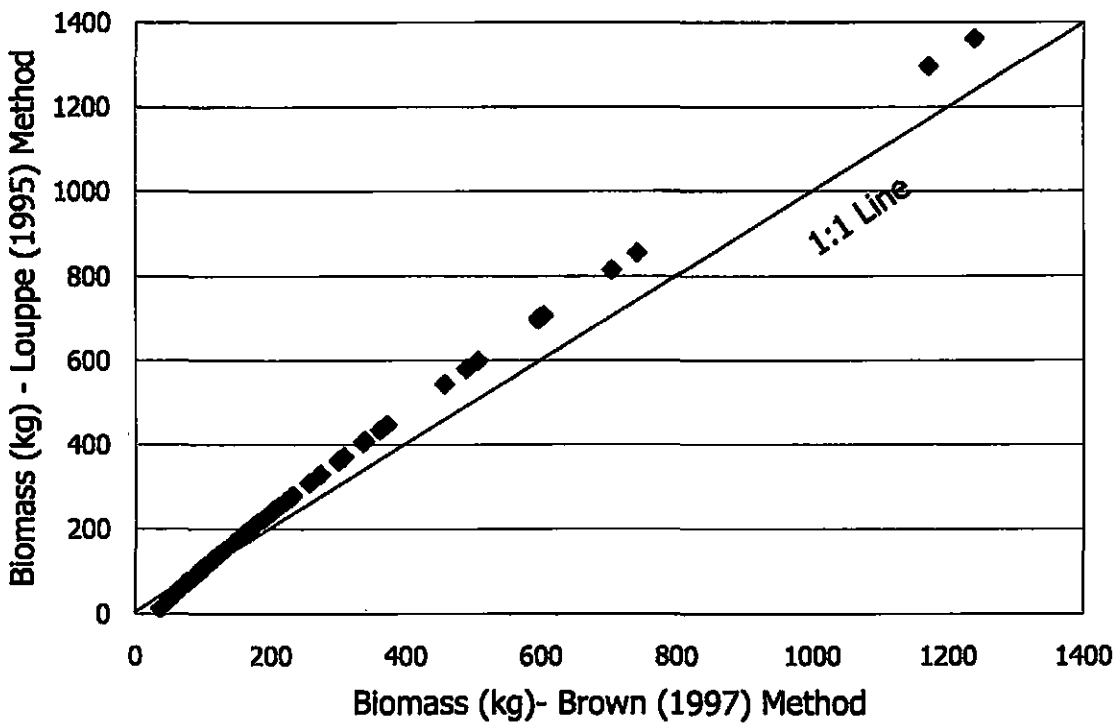


Figure 6.19. The relationship between diameter and tree biomass calculated using both the Brown, 1997 and the Louppe, 1995 allometric equations.

The 1:1 graphical comparison shows that Brown's equation provides more conservative estimates of the *V. paradoxa* biomass (Figure 6.19). Louppe's equation was calibrated for total volume of wood, and the conversion to biomass (weight) using a specific wood density (i.e. 748 kg m^{-3} for V. paradoxa) could contribute to the overestimation (Figure 6.19).

The crown area was calculated based the radius of the canopy measured in the field. The fit of the model for biomass shows a low r^2 indicating a weak predicting power based on crown area (Figure 6.20).

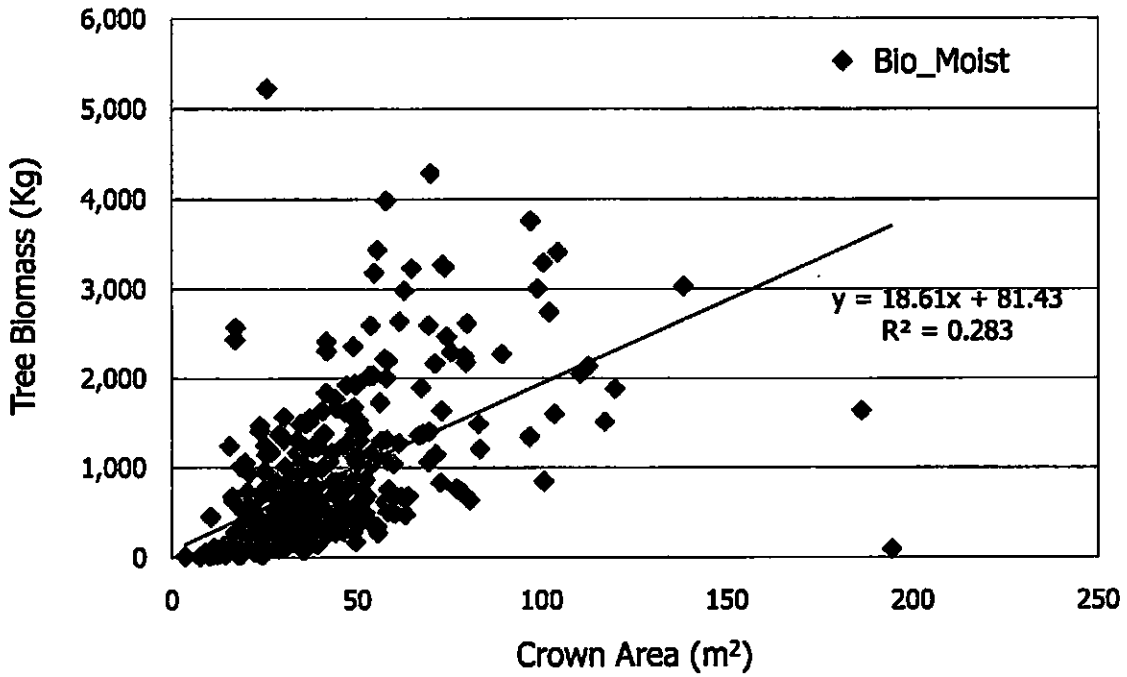


Figure 6.20. Crown Area versus biomass, Yaya Diassa field's. Considering the two equations for Sikasso.

Crown area estimates based on the radii measurements can contribute to the lack of good fit as observed in Figure 6.20. The field measurement of the crown edges is certainly prone to errors. Furthermore, the estimation always assumes the crown area to be either elliptical or circular. For most, Sahelian species the crown architecture is highly irregular making this allometric parameter a weak predictor of the tree biomass as a function of crown area (Figure 6.20).

6.3.4.3. Tree Inventory and Biomass Estimation

African savanna agroforestry systems have a significant impact on sequestration/soil organic carbon stock (Picard et al., 2006). Deciduous and non- deciduous trees have a wide range of spatial influence in the landscape that probably affects the carbon content in the soil. The analysis and results discussed in this section were to evaluate and quantify the organic carbon stored in trees at each experimental site.

The data suggest that the trees were, on the average, large (Fig. 6.21). The error bars provide an indication of the huge variability found in the data set. Likely much of the variability is due to varying developmental stages of the trees at each site.

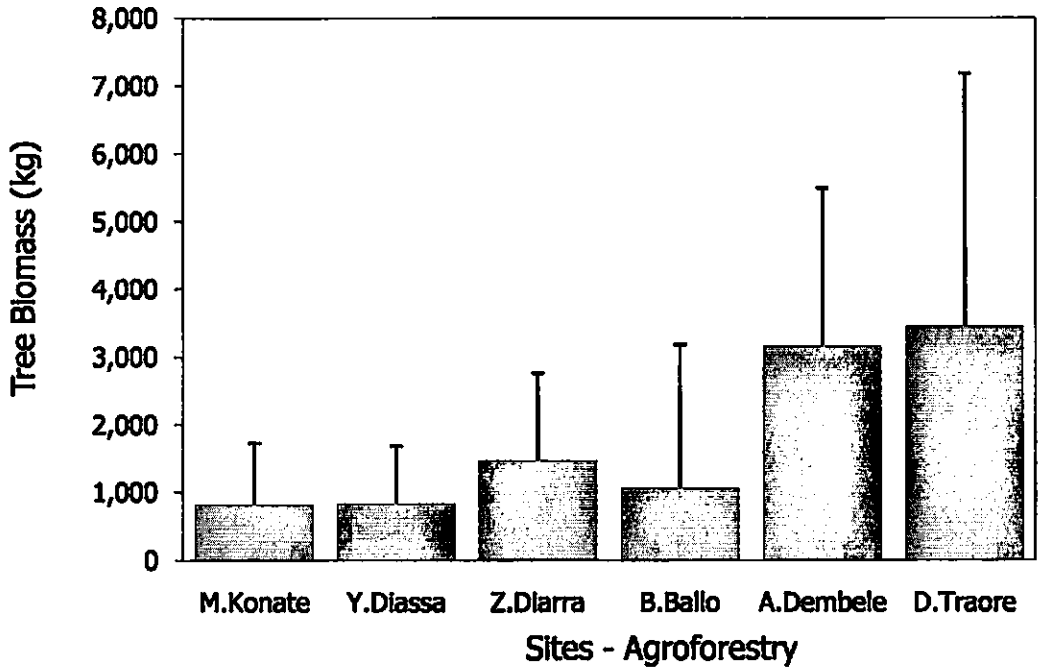


Figure 6.21. The mean and standard deviation of biomass found at each location. Tree biomass was estimated using Brown (1997) equations for moist and dry region.

Because the rainfall at these sites included both wet and dry, Brown's two allometric equations were used to compute the individual tree biomass. The highest biomass was reported at Fansirakoro, followed by Sikasso, Oumarbougou and Siguidolo sites (Figure 6.21). Sikasso receives the highest annual rainfall of the three (900-1100 mm); however, several other factors appear to determine the tree abundance and growth rates (i.e. farmers' preference and management practices, soil properties and depth, tree position in the toposequence). Furthermore, the total biomass computed based on DBH measurement depends on individual tree phenological stage. For example in Drissa Traoré's fields, 17

individual trees contributed a total of 13 Mg ha^{-1} of tree biomass. This particular site was characterized by larger and older trees.

The Baracoro Ballo field results indicated that 91 out of 149 trees weighted equal or less than 250kg (DBH << 24 cm). These results suggested two significant trends: (i) the soil and water conservation measures contributed to new and spontaneous regeneration of local species and (ii) the importance of farmer's awareness in preserving and protecting biodiversity.

During field preparation before planting, farmers clear the field to prevent the growth of undesirable species of trees. However, according to Traoré (2003), in recent past, species like *V. paradoxa* (karité), *Parkia biglobosa* (neré), and *Faidherbia albida* (balanzan) were protected by the village elders. This fact can help to explain the clear dominance of karité, neré and balanzan in almost all sites.

The statistical results for tree inventory and biomass indicated that at all sites the data distribution were extremely shewed due to the existence of a few big trees (Figures 6.22 – 6.24). The standard deviation and the 95% confidence limit show the spread of the tree biomass around the mean.

Based on the total biomass per tree, a conversion factor of 0.55 was used to estimate the amount of carbon equivalent stored in trees ($Carbon_{(in\ tree)}$= $0.55Biomass_{(total)}$). The coefficient of biomass conversion (0.55) was proposed by WIIAD (1997). In practice the biomass coefficient varies according to plant species. However, considering the practical limitations involved in separating

different biomass components and accounting for variation in carbon content, the proposed conversion factor is an average value commonly used in the literature (Hernandez, 2004).

The perceived gradient in biomass/tree from Sikasso (South) to Fansirakoro might be due to changes in soil clay content. Baracoro Ballo and Andre Dembele sites despite lower rainfall than Sikasso (Yaya Diassa) and Oumarbougou (Mory Konate) have the highest biomass per tree (Table 6.5).

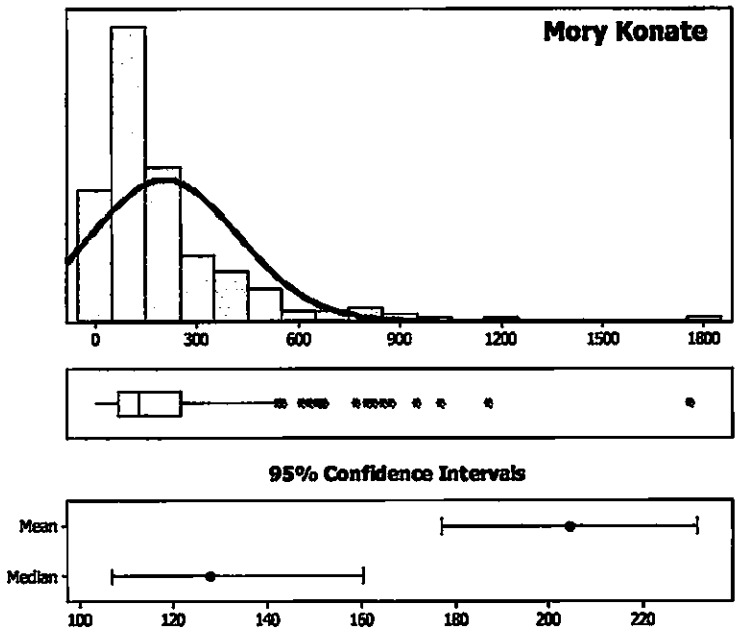
Table 6.5. Tree inventory and biomass based on Brown (1995) equations for moist and dry forest regions.

"Region"/Site	Field size	No. Trees	Total Tree Biomass	Mean	Std. Deviation	Median	95% Confidence Interval for Mean		Biomass	Carbon
Sikasso	ha		Kg/farm	Kg /tree					Mg ha^{-1}	$Mg.ha^{-1}$
YDiassa**	44.89	360	343,444	954	2,552	503	689.5	1218.5	7.65	4.21
Segou										
ZDiarra*	18.18	148	85,451	577	635	368	474.15	680.59	4.70	2.59
Oumarbougou										
MKonate*	20.15	247	100,934	409	436	256	354.01	463.27	5.01	2.76
ADembele*	18.76	47	72,325	1,539	1,244	1261	1173.7	1904	3.86	2.12
Fansirakoro										
DTraoré*	2.65	17	34,876	2,052	2,232	810	904	3199	13.16	7.24
BBallo*	4.77	149	75,430	506	1,101	188	327.99	648.49	15.81	8.70

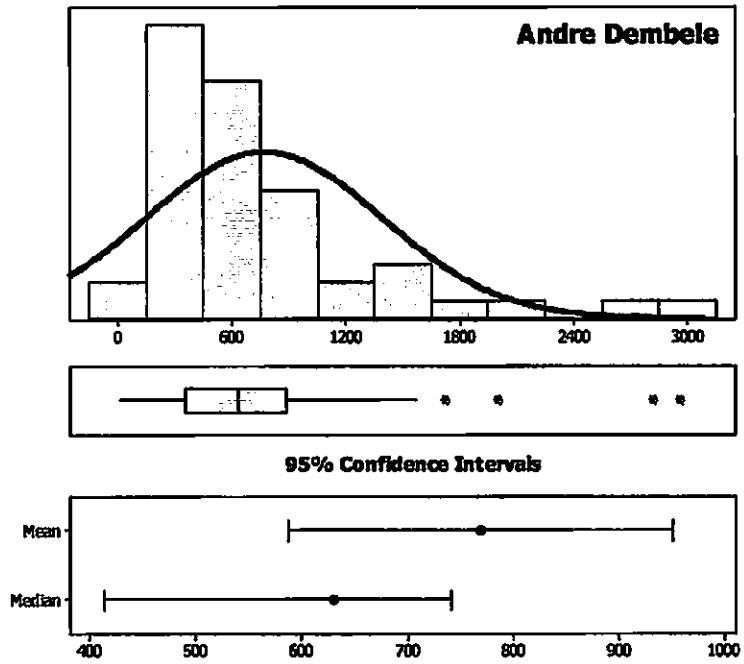
* Brown(1997) – allometric equation for dry regions (Eq. 6.13);
**- Brown(1997) allometric equation for moist regions (Eq. 6.14)

Biomass conversion:

$$\text{Carbon}_{(tree)} = 0.55\ \text{Biomass}_{(total)}\ \text{(WIIAD, 1997)}$$

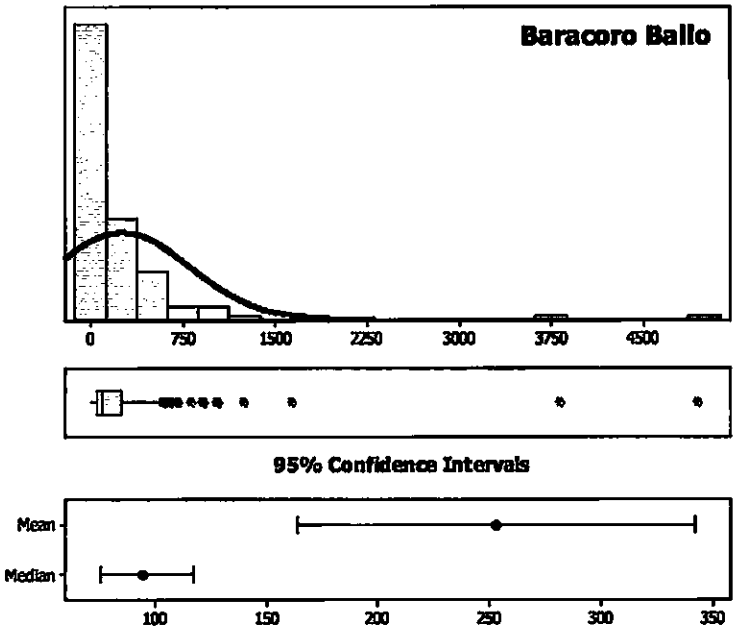


Anderson-Darling Normality Test	
A-Squared	17.43
P-Value <	0.005
Mean	204.32
StDev	217.94
Variance	47496.20
Skewness	2.8786
Kurtosis	12.6489
N	247
Minimum	0.34
1st Quartile	66.04
Median	127.96
3rd Quartile	254.50
Maximum	1756.97
95% Confidence Interval for Mean	
177.01	231.63
95% Confidence Interval for Median	
107.15	160.53
95% Confidence Interval for StDev	
200.26	239.06

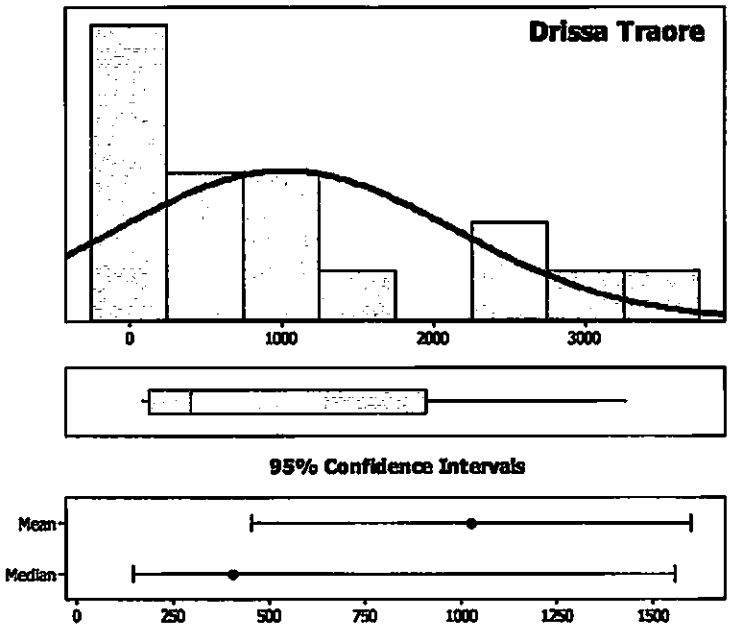


Anderson-Darling Normality Test	
A-Squared	2.92
P-Value <	0.005
Mean	769.42
StDev	621.88
Variance	386739.88
Skewness	1.96181
Kurtosis	4.29263
N	47
Minimum	11.40
1st Quartile	351.86
Median	630.36
3rd Quartile	889.20
Maximum	2951.90
95% Confidence Interval for Mean	
586.83	952.01
95% Confidence Interval for Median	
414.06	741.25
95% Confidence Interval for StDev	
516.77	781.08

Figure 6.22. Tree Carbon at Oumarbougou (Mory Konate and Andre Dembele).

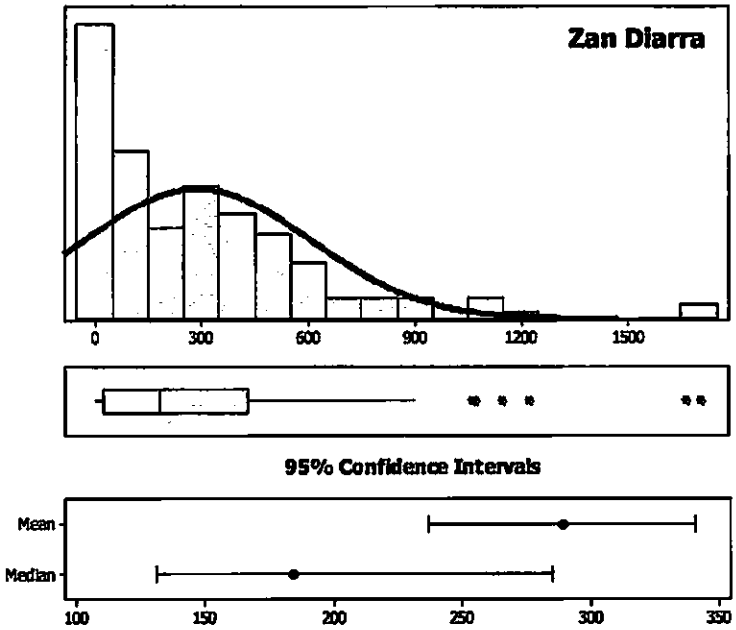


Anderson-Darling Normality Test	
A-Squared	25.41
P-Value <	0.005
Mean	253.12
StDev	550.52
Variance	303067.90
Skewness	6.2512
Kurtosis	46.3301
N	149
Minimum	0.52
1st Quartile	48.28
Median	94.11
3rd Quartile	251.14
Maximum	4930.73
95% Confidence Interval for Mean	
164.00	342.24
95% Confidence Interval for Median	
75.95	117.65
95% Confidence Interval for StDev	
494.31	621.26

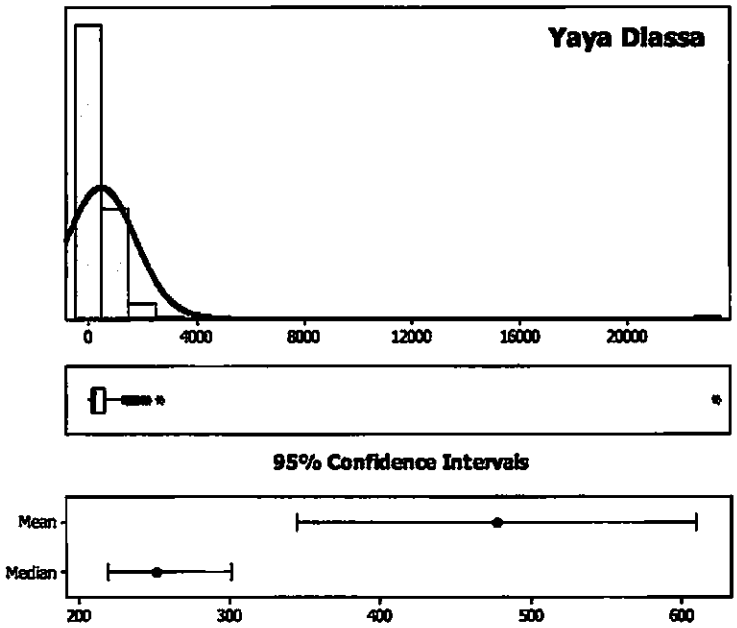


Anderson-Darling Normality Test	
A-Squared	1.32
P-Value <	0.005
Mean	1025.8
StDev	1115.9
Variance	1245200.2
Skewness	1.08718
Kurtosis	-0.13751
N	17
Minimum	87.5
1st Quartile	131.3
Median	405.1
3rd Quartile	1945.7
Maximum	3262.7
95% Confidence Interval for Mean	
452.0	1599.5
95% Confidence Interval for Median	
145.9	1557.3
95% Confidence Interval for StDev	
831.1	1698.3

Figure 6.23. Tree Carbon at Fansirakoro (Baracoro Ballo and Drissa Traoré)



Anderson-Darling Normality Test	
A-Squared	6.21
P-Value <	0.005
Mean	288.68
StDev	317.70
Variance	100935.79
Skewness	1.77282
Kurtosis	4.29195
N	148
Minimum	0.07
1st Quartile	24.18
Median	184.04
3rd Quartile	427.66
Maximum	1701.30
95% Confidence Interval for Mean	
237.07	340.29
95% Confidence Interval for Median	
130.99	285.06
95% Confidence Interval for StDev	
285.17	358.69



Anderson-Darling Normality Test	
A-Squared	68.16
P-Value <	0.005
Mean	477.0
StDev	1275.8
Variance	1627713.7
Skewness	16.086
Kurtosis	287.165
N	360
Minimum	2.3
1st Quartile	143.5
Median	251.5
3rd Quartile	582.9
Maximum	23301.1
95% Confidence Interval for Mean	
344.8	609.2
95% Confidence Interval for Median	
220.2	301.4
95% Confidence Interval for StDev	
1188.9	1376.5

Figure 6.24. Tree Carbon at Segou (Zan Diarra) and Sikasso (Yaya Diassa).

The tree inventory and biomass presented in Table 6.5 highlights the importance of trees in CO_2 removal from the atmosphere and its subsequent storage. It is well known fact that the savanna woodlands are intensively exploited as a source of energy for both rural and urban communities (Traoré, 2003; Nicolas et al., 2006). Fuelwood plays a major role in the economy of the household; therefore, farmers already recognize the value and importance of adequate management of their trees. Moreover, the preservation of trees is a best management practice that would ensure carbon sequestration and increase the potential impact of carbon trading in developing countries.

6.3.4.4. Combining Soil and Tree Carbon

The contribution of semi-arid agroforestry systems to carbon sequestration and the carbon market has largely been ignored. Although several approaches have been proposed to quantify the total biomass stored in a typical tropical forest, there is not yet a standard procedure to measure at a large scale the contribution of semi-arid agroforestry systems to carbon sequestration and carbon trading markets. This study provides an assessment of the potential contribution of semi-arid agroforestry on carbon stock.

The results indicated that the contribution of tree biomass to total organic carbon is substantial varying from 7 to 34% of the soil organic carbon measured at 0-20 cm (Table 6.6).

Table 6.6. The relative contribution of tree biomass to soil plus tree biomass carbon at the field level. The estimate of soil carbon was based on 2006 data.

Site	Total carbon (Mg)	Soil organic Carbon (Mg)	Tree carbon (Mg)	Tree + Soil carbon ($Mg.ha^{-1}$)	Tree carbon contribution (%)
Baracoro Ballo	163.5	122.0	41.5	34	34
Mory Konate	883.4	827.9	55.5	44	6.7
Yaya Diassa	773.1	584.2	188.9	17	32

The estimated soil organic carbon (0-20 cm) at Baracoro Ballo's field was 122 Mg (26 Mg ha^{-1}) and the aboveground tree carbon 41.5 Mg (Table 6.6). The results show that tree carbon represented 34% of the amount of organic carbon stored in the soil (Table 6.6). Mory Konate's field had the highest amount of (soil + tree) organic carbon (44 Mg ha^{-1}). The percentage of tree carbon to soil organic carbon is estimated to be 6.7% at Mory Konate's field.

At Yaya Diassa the estimated tree carbon represented 32 % of the predicted soil organic carbon in 2006. The also estimated soil and tree organic carbon was calculated to be 17 Mg ha^{-1}. Yaya Diassa's field, on the other hand, has the highest number of trees and amount of total tree biomass (Table 6.5).

6.4. Feasibility of Sequestering Soil Carbon

The establishment of a carbon trading system that promotes soil carbon sequestration has the potential to enhance farmer's incentive to store the organic residue in the soil and to preserve trees (Perez et al., 2007). This opportunity can bring substantial impact to Sub-Saharan Africa dry lands (Tschakert and Tappan, 2004; Stavins, 1999). This action can increase soil fertility and provide economic incentives that can help to perpetuate best conservation practices of organic matter, soil and water.

Antle (2000) indicated that the economic feasibility of carbon sequestration was conditioned by the profitability of the systems and the transaction costs associated with the contracts. Based on estimates of soil carbon sequestration rates, the study suggested that the market value of carbon could well exceed $10 per tonne (Antle, 2000). More recent studies of Capoor and Ambrosi (2006) reported an average price of carbon at $7.23 in 2005. According to Antle (2000) farmers in the developing world could have an annual income of $300 managing a 3 hectare land. This estimate was based on sequestration rates of 0.1 to 0.3 t ha^{-1} yr^{-1}.

Given the uncertainty associated with the market price of carbon and the challenges in monitoring the net gain of soil carbon in a large region, the prediction at a field scale provides an overall indication of the amount of land in a given region needed to attain the minimum tradable unit of carbon. The net gain of soil organic carbon and potential carbon at four locations in Mali

illustrates the variability and the need to assess each field and farm separately (Table 6.7). The net increase was calculated relative to the year 2000 for each site. The addition of soil carbon varied from site to site. Mory Konate's field showed a net increase of 13.9 Mg ha^{-1} , which was greater than that at any other experimental site (Figure 6.7).

Table 6.7. Net increase of soil organic carbon from 2000 to 2006.

Sites	Area (ha)	Net Increase of SOC (Mg/farm) 2000-2006	Net Increase (Mg ha^{-1})	Potential Revenue (USD)
BBallo	4.77	30.53	6.4	$305.28
Mkonate	20.15	279.86	13.9	$2,798.60
ZDiarra	18.18	103.25	5.7	$1,032.47
YDiassa	44.89	116.50	2.6	$1,165.00

Based on our estimates at $10 a tonne (Mg) the equivalent revenue would have been very attractive for most poor farmers (Figure 6.6). The net increase of soil organic carbon at the experimental sites was due to ACN technology, which aims to increase rainfall water harvesting and soil water availability (Kablan et al., 2008). Assuming that farmers will continue improve the land management practices and seek to maximize their return, the prospect of carbon trading can be an effective tool for poverty reduction and global environmental impact.

6.5. Conclusion

The spatiotemporal prediction of soil organic carbon in agroforesty systems of Sub-Saharan Africa was achieved with the Bayesian Maximum Entropy approach. Despite the complexities which characterize the spatial and temporal distribution of most environmental processes, BME provides a framework to analyze both space and time components. Carbon trading is increasingly becoming the market's response to slow the global warming. Countries with high CO_2 emissions can become carbon neutral by purchasing carbon credits from their counterparts with low emissions or by funding joint implementation projects aiming at reducing CO_2 from the atmosphere. As the carbon offsetting mechanism is established and monetary transactions begin to take place regularly, countries and carbon brokers will have to rely on a set of tools to estimate carbon amount to be traded. In this regard, the capability to monitor the impact of joint implementation actions in space and time will be of paramount importance.

Bayesian Maximum Entropy provided a solid framework that can be used for mapping spatial and temporal variability of soil carbon at field scale. One clear advantage of BME is based on the ability to map soil properties taking into account both soft and hard data. The integration of these two data types is seamlessly done using BMElib[5] in Matlab[6] (Christakos et al., 2002; The Mathworks, 2007). The "range" of the variogram over time in a particular region

[5] BMElib, a collection of numerical routines essential for computational BME analysis (Christakos et. a. 2002).
[6] This citation should not be perceived as an endorsement of this software package

is useful information for carbon sequestration projects. Nations or even corporations, interested in funding joint implementation programs to enhance CO_2 sequestration can promptly evaluate the impact of the project by examining both spatial and temporal covariance of soil carbon in a particular field. Similarly to other kriging methods the BME spatiotemporal covariance model suggests the necessary sampling frequency and density needed to achieve a specified precision. BME temporal means can be important in carbon verification and accreditation. The departure of mean carbon in a particular field from the initial or critical threshold over a period is important for compliance monitoring purposes.

The method of estimating tree biomass requires a costly survey of all trees present in the field. Hence, this approach may be the ideal for assessing tree biomass at small scales (i.e. field) but for regional scale quantification of tree biomass the method may be too costly in time and money. As a promising alternative, remote sensing imagery can provide high spatial resolution digital images that can facilitate the large scale inventory of savanna woodlands. This high spatial resolution image (60-70 cm/pixel) has the potential not only to delineate the crown area but to enable species recognition (Niemann et al., 1998).

The environmental conditions of Sub-Saharan Africa are harsh and characterized by low rainfall and annual vegetation. In these ecosystems the yield of organic material is very low. This study shows that agroforestry systems

play an important role in carbon sequestration. Despite the harsh climatic conditions the biodiversity and density of trees were considerable in all study sites. Consequently, carbon accounting in savanna ecosystems of Sub-Saharan Africa should not neglect the amount of carbon stored as tree biomass.

CHAPTER 7: CONCLUSIONS

7.1. Conclusions

The developing countries now, more than ever, are faced with an increasing dilemma to provide a livelihood for their populations while at the same time caring for and protecting the environment. Environmental protection is by no means a new concept; history teaches that ancient civilizations in Mesopotamia, threatened with environmental challenges, took action to combat the adverse effect. Today our generation is forced to mitigate the adverse impact of global climate change due to the steady increase in greenhouse gases caused by human actions. The scientific evidence of global warming is mounting, turning the most skeptical arguments mute and demanding immediate action to ensure that this destructive trend can be reversed or at least halted.

One of the global responses to increased greenhouse gases emissions came in form of an international treaty (commonly known as the Kyoto Protocol) that aimed at placing an emission cap to all nations. The treaty created a mechanism that would allow countries to also trade carbon. In order to trade carbon an inventory that takes into account the soil, a major carbon pool, is vital and necessary.

This study provides a practical approach that incorporates a promising new geostatistics methodology (Christakos, 2000; 2002) to estimate the amount of soil organic carbon. The Bayesian Maximum Entropy (BME) approach provides a general framework for spatial-temporal mapping of soil, relying not only on

measured hard data but also additional site specific information called "soft data". This study demonstrates that soil organic carbon estimates can be improved with use of soft data.

Spatiotemporal mapping also demonstrates BME's capability to provide valuable information about the spatial and temporal variability of soil organic carbon. This information is presented in terms of space/time soil carbon mapping and with probability distributions important for an informed decision regarding soil management as well as for carbon accounting and compliance verification.

By integrating interval soft data of other soil properties, the estimates of soil organic carbon were more accurate than the traditional kriging methods. BME prediction lowered the MSE compared with kriging and cokriging (24% and 10% reduction, respectively).

Furthermore, the potential to use soft information (i.e. soil texture, landuse, taxonomic range of soil properties) is a clear advantage for BME spatial-temporal estimation of soil organic carbon in West Africa, where precise hard data about a particular site is often sparse and costly. BME also provided a strong and relatively flexible set of functions in Matlab that allow the user to generate important output maps and information for uncertainty analysis.

This study also evaluated the spatiotemporal trend of soil organic carbon in different agro-ecological sites of Mali. Although the current data were collected over a short period of time for space/time analysis, preliminary

information indicates a positive trend in soil organic carbon. The increase of soil organic carbon at the project sites was directly associated with ACN technology of soil and water conservation measures and increased farmer's awareness of the impact of organic matter on soil productivity.

The temporal covariance models provided a quantification of the temporal component associated with the spatial variability and the autocorrelation between successive years. Our study showed that the temporal autocorrelation at all sites persisted for four years. Beside modeling the spatial and temporal dependence, covariance models also can be used to assess the adequacy of sampling density (space) and frequency (time). This information is important for compliance and verification of carbon sequestration projects. The temporal covariance quantifies the time lag of a particular measurement or practice and provides an indication of how long a particular spatial trend is expected to last. Consequently, this tool can aid the compliance and verification process over the duration of carbon sequestration contract.

Based on BME output maps of soil carbon, this study provided a simple framework to compute the total amount of carbon. This approach integrates the output of BMElib with ArcGIS and uses map calculation procedures to generate the carbon tonnage.

Finally, the study quantified the contribution of agroforestry systems in the total carbon estimates using allometric equations. The results confirm previous work carried out in Mali by Traoré (2003) that also describes the

importance of agroforestry systems in nutrient recycling and soil fertility. Our results indicated that the amount of carbon stored in tree biomass is substantial and should not be neglected in carbon inventories and tonnage in Sub-Saharan Africa. Furthermore, a preliminary calculation indicated attractive scenarios for poor farmers in this region if the concept and benefits of carbon trading can directly reach each farmer.

7.2. Recommendations and Suggestions for Carbon Stock Assessment

This work demonstrated the applicability of BME in predicting the amount of soil carbon in West Africa. The estimation of soil carbon is important considering the increasing potential for global carbon transactions. Our approach looked at soil organic carbon and the contribution of trees to total carbon estimates. However a more holistic approach that takes into account the soil, agriculture production systems as well as the amount of biomass stored in the woodland of the savanna region is deemed important.

BME integrates both general and site specific knowledge including expert opinions. Therefore, indigenous knowledge of the local farmer's should be considered as potential prior information specific for every site. Our analysis considered the mean and covariance of the soil carbon measured at each location as the general knowledge, future work should also consider the integration of other type of general information (i.e. other statistical moments). In order to account for the uncertainty associated with variogram modeling, the

Gaussian simulation method could be used as part of BME to help estimate the model parameters.

The Kyoto protocol allows countries to work for carbon neutrality in project-based transactions, in which buyers purchase emission credits from a project that sequesters carbon from the atmosphere. Some project-based transactions are conducted to meet voluntary targets, but most are ultimately intended for compliance with the Kyoto Protocol or other regulatory regimes. Consequently, a soil carbon estimates are relevant to all nations, in particular to the potential buyers and sellers of carbon.

For most West African nations carbon sequestration and trading is a win-win scenario when considering its impact on soil quality and productivity, direct benefit to improve the welfare of poor farmers, and to mitigate the global effects of increased CO_2 in the atmosphere. The Kyoto protocol binds the member countries to create conditions to reduce greenhouse gas emissions while providing an innovative approach that allows carbon to be traded as a commodity (UN, 2001). However, in order for most developing nations to adequately quantify the amount of carbon dioxide equivalent stored in their soil, a practical procedure to predict the amount of soil carbon in a given field, country or even region is of paramount importance.

7.2.1. Soil carbon assessment – BME approach

This section provides the sequence of tasks and procedures for estimating soil organic carbon using the Bayesian Maximum Entropy approach. Figure 7.1 illustrates the procedures and the relationships between key modules of this approach and should be taken as guide. There are a number of small, yet important, activities that are not fully explained. For instance, all data collection requires the use of GPS and it's equally essential that all satellite images are properly georeferenced and basic image processing has been performed.

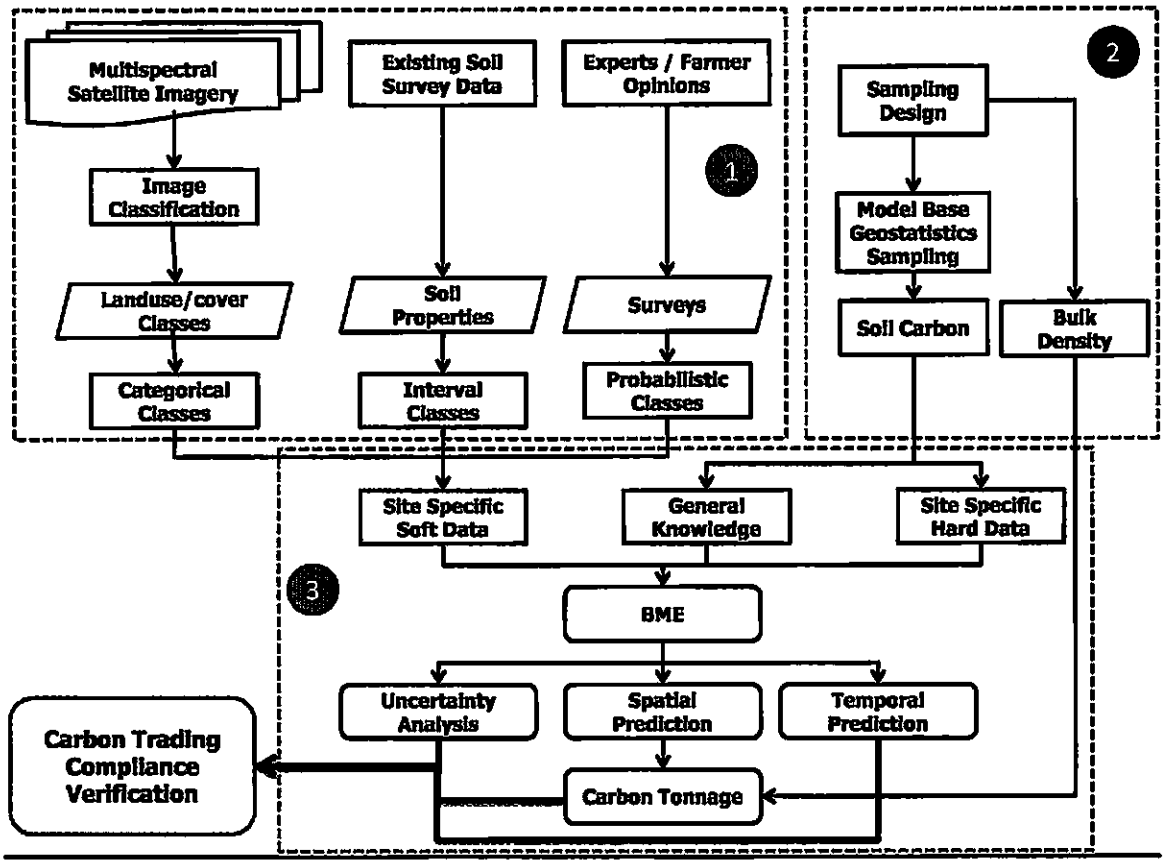


Figure 7.1. Procedure for quantifying soil organic carbon using BME.

1. Procedure for generating site specific soft data:

Landuse/landcover, soil properties and experts opinions are three important groups of soft data that may improve the prediction of soil organic carbon. These are expected to be existing data specific to the site of interest. Although BME can integrate multivariate soft data it's not essential to use all available soft data. Preliminary trials should indicate which soft data maximizes the information and improves the posterior distribution of the random variable of interest.

Remote sensing – high resolution remotely sensed images can be used to generate landcover classes. The landuse / landcover maps are generated using supervised classification followed by a groundtruth check. The different classes can be coded in meaningful group and used as categorical soft data.

Existing soil survey/Data – information about soil properties and taxonomic units are generally available. Existing soil data collected from prior studies may be available. Because these data were collected by prior studies and often lack a detail description of the sampling procedures, and experimental errors, such data may be useful as interval soft data. For instance, a particular soil map unit maybe characterized by a range of soil pH, or a soil property can be discretized into interval classes (i.e clay content).

Expert's opinions – This group of soft data refers to survey data that exist or could be obtained out to provide relevant information about a process under

consideration. Expert or farmer's qualitative opinions can be transformed into probabilistic and interval soft data for use by the BME system.

The main goal of this step is to identify and properly code the soft data. All soft must be placed in a particular digital file format as required by BMElib software (Christakos et al., 2002).

2. Soil Sampling

Sampling design – the sampling design should address key questions concerning the scale (field, region, country), the temporal component and cost. It's important that the soil sampling be representative and randomly collected to adequately capture the spatial and temporal variability ensure accurate inference. The sample size is often determined by available funds; hence, it is essential to for the sampling scheme be well-designed to ensure proper spatial coverage. Another important consideration is the depth of soil column for which one wishes to predict the soil carbon (i.e. 0.2 m, 0.4 m, 1m?). For carbon tonnage estimates the bulk density must also be measured. A single bulk density value is often used for smaller fields. For large scale assessment representative bulk density measurements are needed. In this procedure the measured data is referred to as hard data. Both hard and soft data comprise the site specific data.

Geostatistical sampling – the aim is to optimize the sample size and sampling events using a model-based approach (Gruijter et al., 2006). As mentioned

before in many occasions the sample size is determined by available funds. Consequently, the use of variograms can help optimize sampling pattern.

3. Bayesian Maximum Entropy

The BME theoretical approach is presented in Chapter 4. For additional readings, mathematical proof and BMElib software manipulation the reader is referred to Christakos (2000) and Christakos et al. (2002). The general principles of BME formalism can be summarized into four major stages also known as prior, metaprior, posterior and predictive.

1. The prior stage, involves the gathering of general knowledge ($\mathcal{K}$) (mean, covariance).
2. Meta prior stage is the collection and organization of site specific data (S) (hard and soft) in appropriate quantitative forms that can be incorporated into the BME software.
3. The posterior stage is the integration of general knowledge $\mathcal{G}$ and site specific knowledge S to produce the posterior or the combined knowledge $\mathcal{K}$ ($\mathcal{K}= \mathcal{G} \cup S$).
4. The predictive stage provides the moments of the posterior PDF. In this stage the mean, the mode and the uncertainty associated with the predicted value (error variance) is computed.

Carbon Tonnage – carbon tonnage is the product of soil volume (m^{-3}), bulk density ($Mg\ m^{-3}$) and carbon content ($g\ kg^{-1}$). The spatial prediction of soil carbon is an important tool for carbon trading.

7.2.2. Procedures for tree biomass estimation

The estimation of tree biomass can be accomplished by allometric equations, which relates tree measured parameters to weight. This approach requires the following data: tree height (H); diameter at breast height (DBH), diameter of canopy into perpendicular directions (Length (L), Width (W)); height of the base of the crown (Hc).

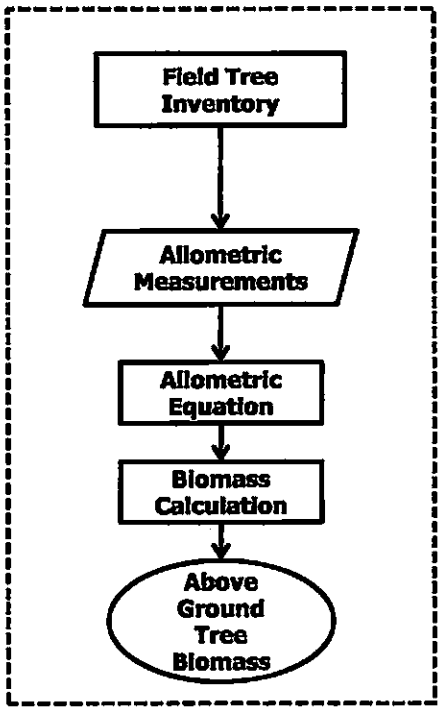


Figure 7. 2. Procedures for tree biomass estimation.

The allometric equations as a function of DBH (Brown, 1997) can be used to estimate the above ground tree biomass and carbon content. This procedure is time consuming requring a huge effort in data collection and might not be economically feasible for a large scale sampling.

APPENDIX I:

The GeoEAS file format – for Hard data

The GeoEAS[7] file format is a text file format used in BMElib to store data. This format has a title in the first line, the number of columns in the second line followed by rows with the name of each column.

EXAMPLE :

	A	B	C	D	E	F	G	H
1	balloCdata							
2	6							
3	x							
4	y							
5	OC20							
6	OC40							
7	year							
8	slopeClass							
9	598941	1432617	0.629	0.599	2004	6.497845		
10	598948.5	1432566	0.577	0.588	2004	7.271944		
11	598959.8	1432523	0.491	0.74	2004	23.19819		
12	598996.5	1432543	0.611	0.547	2004	20.90354		
13	599050.7	1432549	0.703	0.416	2004	11.29907		
14	599110	1432501	0.701	0.573	2004	9.979772		
15	599097.6	1432447	0.353	0.534	2004	13.35599		
16	599059.1	1432511	0.344	0.286	2004	11.00426		
17	599018	1432506	0.383	0.442	2004	14.61926		
18	598978.4	1432493	0.602	0.372	2004	15.06763		
19	598938.9	1432473	0.606	0.38	2004	9.175463		
20	598935.2	1432426	0.401	0.315	2004	10.37368		
21	598894.3	1432391	1.27	1.105	2004	8.088264		
22	598853.2	1432375	1.179	1.114	2004	11.00543		
23	598799.9	1432368	1.195	1.045	2004	5.864947		
24	598788.2	1432412	1.501	0.951	2004	6.095888		
25	598807.3	1432439	0.682	0.629	2004	8.03979		
26	598860.6	1432480	0.434	0.442	2004	1.133233		
27	598896.9	1432455	0.634	0.777	2004	9.074914		
28	598923	1432495	0.807	0.681	2004	10.54562		
29	598903.4	1432528	0.818	0.765	2004	6.375362		
30	598849	1432478	0.456	0.587	2002	2.071636		
31	598861.9	1432477	0.61	0.484	2002	1.106094		
32	598880.6	1432477	0.752	0.521	2002	8.231021		
33	598897	1432464	0.6	0.369	2002	8.542185		
34	598911.8	1432438	0.692	0.881	2002	14.14794		
35	598939.1	1432452	0.943	0.698	2002	1.304367		
36	598929.7	1432472	0.697	0.331	2002	19.19601		
37	598955.4	1432478	0.278	0.235	2002	1.52994		
38	598987.6	1432512	0.549		2002	18.95153		
39	599126.2	1432439	0.789	0.764	2002			
40	598962.4	1432552	0.649	0.469	2002	12.61745		
41	598943.6	1432537	0.909	0.541	2002	21.24551		
42	598864.9	1432379	0.808	0.512	2002	7.687588		
43	598844.3	1432398	0.915	0.576	2002	9.035888		

- The GeoEAS file format is very simple. It is an ASCI text file with header lines followed by the data as a space delimited ASCI file.
- The first header line is the title. We usually use the title to show the location of the MS Access file where the data originated.
- The second line is the number of columns of data. Here there are 4 columns of data.
- The next four lines are the variable names for the data in each column. You are allowed up to ten characters for the variable name.
- The data follow with a space between each field.

[7] GeoEAS file format was developed by Enlgund (1991) – US EPA.

APPENDIX II:

File format – for interval soft data

b06bbinterdata.txt

	A	B	C	D	E	F
1	Baracoro Ballo Interval Soft data - Carbon (20-40) and Clay					
2	x	y	a_clay	b_clay	a_c40	b_c40
3	598942.9276	1432563.281	40	60	4	6
4	599009.0092	1432575.553	60	80	4	6
5	599055.581	1432576.497	60	80	2	4
6	599067.224	1432546.603	80	100	4	6
7	599025.687	1432536.219	80	100	2	4
8	598981.9472	1432535.275	40	60	6	8
9	598938.8368	1432526.778	80	100	6	8
10	598894.1531	1432509.471	80	100	6	8
11	598969.9896	1432489.332	40	60	6	8
12	599060.3011	1432509.786	40	60	2	4
13	599079.8109	1432466.046	40	60	4	6
14	599103.7262	1432491.535	60	80	4	6
15	599107.5023	1432436.782	60	80	8	10
16	599066.2799	1432432.691	20	40	2	4
17	598967.7869	1432447.795	80	100	4	6
18	598888.1743	1432471.081	80	100	4	6
19	598857.0215	1432439.614	60	80	6	8
20	598821.1486	1432453.145	40	60	3	4
21	598796.9187	1432412.237	80	100	8	10
22	598873.0699	1432408.146	80	100	6	8
23	598903.908	1432426.712	80	100	4	6
24	598963.0668	1432416.013	80	100	6	8
25	598876.5313	1432386.748	80	100	4	6
26	598855.1335	1432355.91	80	100	4	6
27	598816.4285	1432387.378	60	80	2	4
28	598779.9263	1432360.63	60	80	8	10
29	598810.135	1432333.568	80	100	4	6

b06bbinterdata

- X, Y – UTM coordinates of each point
- a_clay – lower boundary of each class of softdata for clay
- b_clay – upper boundary of each class of softdata for clay
- a_c40 – lower boundary of each class of softdata for carbon (20-40 cm)
- b_c40 – upper boundary of each class of softdata for for carbon (20-40 cm)

Note: BMElib "load function" requires that the columns with label should be removed.

APPENDIX III:

Covariance estimation using data in S/T vector format

```
%%% Read the hard data from file
[val,valname,filetitle]=readGeoEAS('mkSTC_y.dat');
% [val,valname,filetitle]=readGeoEAS('mkSTC_y_rdvec.dat');
ch=val(:,1:3);
zh=val(:,4);
th=ch(:,3);

%%% Open a new graphic window and display the hard data that were measured
between 2000  and 2002, and plot the hard data proportional to the magnitude of
carbon data.

figure;
markerplot(ch(1<=th&th<2,1:2),zh(1<=th&th<2),[4 30]);
xlabel('Easting (m)');ylabel('Northing (m)');
title('Markerplot of data for Year 2000 ');

%%% Open a new graphic window and display the hard data that were measured
between 2002  and 2004, and plot the hard data proportional to the magnitude of
carbon data.

figure;
markerplot(ch(2<=th&th<3,1:2),zh(2<=th&th<3),[4 30]);
xlabel('x (m)');ylabel('Northing (m)');
title('Markerplot of data for Year 2002');

%%% Open a new graphic window and display the hard data that were measured
between 2004  2006, and plot the hard data proportional to the magnitude of carbon
data.

figure;
markerplot(ch(3<=th&th<4,1:2),zh(3<=th&th<4),[4 30]);
axis([0 800 0 500]);
xlabel('Easting (m)');ylabel('Northing (m)');
title('Markerplot of data for Year 2004');
```

```
figure;
subplot(2,2,1),
   markerplot(ch(1<=th&th<2,1:2),zh(1<=th&th<2),[4 30]);
   axis([0 800 0 500]);
   xlabel('Easting (m)');ylabel('Northing (m)');
   title('Soil Carbon for Year 2000 ');
  subplot(2,2,2),
   markerplot(ch(2<=th&th<3,1:2),zh(2<=th&th<3),[4 30]);
   xlabel('Easting (m)');ylabel('Northing (m)');
   axis([0 800 0 500]);
   title('Soil Carbon for Year 2002');
subplot(2,2,3),
   markerplot(ch(3<=th&th<4,1:2),zh(3<=th&th<4),[4 30]);
   axis([0 800 0 500]);
   xlabel('Easting (m)');ylabel('Northing (m)');
   title('Soil Carbon for Year 2004');
subplot(2,2,4),
   markerplot(ch(4<=th&th<5,1:2),zh(4<=th&th<5),[4 30]);
   axis([0 800 0 500]);
   xlabel('Easting (m)');ylabel('Northing (m)');
   title('Soil Carbon for Year 2006');

%%% Estimate the covariance by using pairs in all directions.

cls=[0 30 45 60 75 90 105]; clt=[0 1 2 3];
 [ds,dt,c,o]=crosscovarioST(ch,ch,zh,zh,cls,clt); %

variance=var(zh);

%%% Plot the estimated covariance on a new graphic window

figure;
subplot(2,1,1);hold on;
        plot([0 ds(:,1)'],[variance c(:,1)'],'.r:');
        plot([0 100],[0 0],'k-');
        ylabel('Covariance C(r,t=0)');
        xlabel('Spatial lag r (m)');
        itle('Spatial Covariance - MoryKonate [2000-2006]');
        legend([cm hCe],'Covariance Model','Experimental Covariance');
subplot(2,1,2);hold on;
        plot([0 dt(1,:)],[variance c(1,:)],'.r:');
        plot([0 4],[0 0],'k-');
        set(gca,'XTick',(0:1:4));
        ylabel('Covariance C(r=0,t)');
        xlabel('Temporal lag t (Year)');
        title('Temporal Covariance - MoryKonate [2000-2006]');
```

%%% The model is an separable space/time covariance with a sill (variance) cc=0.01, an exponential spatial component with a range aa=50 Km, an exponential temporal component with a range at=2.8 Year.

```
r=0:0.2:90;
covmodel='exponentialC';
covparam=[0.01 50];
subplot(2,1,1);
modelplot(r,covmodel,covparam);
ylabel('Covariance C(r,t=0)');
xlabel('spatial lag r (m)');

t=0:.2:4;
covmodel='exponentialC';
covparam=[0.01 2.8];
subplot(2,1,2);
modelplot(t,covmodel,covparam);
ylabel('Covariance C(r=0,t)');
xlabel('temporal lag t (Year)');
```

APPENDIX IV:

Estimation of covariance using space/time grid data

```
clear;
%close all;
clc;
echo on;

%%% Read the hard data from file

[val,valname,filetitle]=readGeoEAS('mkSTC_y.dat');
cMS=val(:,1:2);
Zh=val(:,4:end);
nME=size(Zh,2);
tME=1:nME;

%%% Estimate the covariance by using pairs in all directions.
rLag= [0:30:200]';
rLagTol=[25:25:180]';

rLag=[0.0,0.5,0.1,0.1525,0.2,0.254166666667,0.305,0.355833333333,0.406666666667,
0.4575];

rLagTol=[00.0,0.0254166666667,0.0254166666667,0.0254166666667,0.025416666666
7,0.0254166666667,0.02541666 ]';

[Cr npr]=stcov(Zh,cMS,tME,Zh,cMS,tME,rLag,rLagTol,0,0);

tLag=  [0:4];
% tLagTol=[0 1*ones(1,length(tLag)-1)];
tLagTol=[0:4];
[Ct npt]=stcov(Zh,cMS,tME,Zh,cMS,tME,0,0,tLag,tLagTol);

%%% Plot the estimated covariance values on a new graphic window
figure;
subplot(2,1,1);hold on;
   plot(rLag,Cr,'.r:');
   plot([0 10],[0 0],'k-');
   ylabel('Covariance C(r,t=0)');
   xlabel('spatial lag r (m)');
   title('Covariance');
   title('Spatial Covariance - MoryKonate [2000-2006]');
subplot(2,1,2);hold on;
   plot(tLag,Ct,'.r:');
   plot([0 4],[0 0],'k-');
   set(gca,'XTick',(0:1:4));
```

```
   ylabel('Covariance C(r=0,t)');
   xlabel('temporal lag t (Year)');
   title('Temporal Covariance - MoryKonate [2000-2006]');
%% A reasonable model is an separable space/time covariance with a sill (variance)
cc=1.3,
an exponential spatial component with a range aa=13 m an exponential temporal
component with a range at=11 Year.
```

```
figure;
subplot(2,2,1),
        markerplot(ch(1<=th&th<2,1:2),zh(1<=th&th<2),[4 30]);
        axis([0 800 0 500]);
        xlabel('Easting (m)');ylabel('Northing (m)');
        title('Simulated Soil Carbon for Year 2000 ');
subplot(2,2,2),
        markerplot(ch(2<=th&th<3,1:2),zh(2<=th&th<3),[4 30]);
        xlabel('x (m)');ylabel('Northing (m)');
        axis([0 800 0 500]);
        title('Simulated Soil Carbon for Year 2002');
subplot(2,2,3),
        markerplot(ch(3<=th&th<4,1:2),zh(3<=th&th<4),[4 30]);
        axis([0 800 0 500]);
        xlabel('Easting (m)');ylabel('Northing (m)');
        title('Simulated Soil Carbon for Year 2004');
subplot(2,2,4),
        markerplot(ch(4<=th&th<5,1:2),zh(4<=th&th<5),[4 30]);
        axis([0 800 0 500]);
        xlabel('Easting (m)');ylabel('Northing (m)');
        title('Simulated Soil Carbon for Year 2006');
```

```
r=0:.2:180;
covmodel='exponentialC';
covparam=[0.0027 290];
subplot(2,1,1);
   modelplot(r,covmodel,covparam);
   ylabel('Covariance C(r,t=0)');
   xlabel('spatial lag r (m)');
```

```
t=0:.2:4;
covmodel='exponentialC';
covparam=[0.003 2.5];
subplot(2,1,2);
   modelplot(t,covmodel,covparam);
   ylabel('Covariance C(r=0,t)');
   xlabel('temporal lag t (Year)');
```

REFERENCES

Anctil, F., R. Mathieu, L. E. Parent, A.A. Viau, M. Sbih, and M. Hessami. 2002. Geostatistics of near-surface moisture in bare cultivated organic soils. Journal of Hydrology (Amsterdam) 260:30-37.

Antle, J., 2000. Economic feasibility of using soil carbon sequestration policies and markets to alleviate poverty and enhance sustainability of the world's poorest farmers. Paper presented at the expert workshop on carbon sequestration, sustainable agriculture and poverty alleviation held at the World Meteorological Organization (WMO). (http://www.climate.montana.edu/pdf/geneva.pdf . retrieved on 10/12/07).

Antle, J.M., G. Uehara. 2002. Creating incentives for sustainable agriculture: defining, estimating potential and verifying compliance with carbon contracts for soil carbon projects in developing countries. In: A Soil Carbon Accounting System for Emissions Trading – Special Publication, Soil Management Collaborative Research Support Program. University of Hawaii, Honolulu, HI, pp. 1–12.

Ardo J., L. Olsson. 2003. Assessment of soil organic carbon in semi-arid Sudan using GIS and the CENTURY model."Journal of Arid Environments Vol 54:633-651.

Bailey, T.C., A.C. Gatrell. 1995. Interactive spatial data analysis. Addison Wesley Longman.

Bationo, A., A. Buerkert. 2001. Soil organic carbon management for sustainable land use in Sudano-Sahelian West African, Nutrient Cycling in Agroecosystems. 61:131–142.

Bationo, A., A. Hartemink, O. Lungu, M. Naimi, P. Okoth, E. Smaling, L. Thiombiano. 2006. African soils: Their productivity and profitability of fertilizer use. background paper prepared for the african fertilizer summit. Abuja Nigeira.

Bationo, A., J. Kihara. 2007. "Soil organic carbon dynamics, functions and management in West African agro-ecosystems." Agricultural Systems 94(1):13-25.

Batjes, N. H., 2003. Estimation of soil carbon gains upon improved management within croplands and grasslands of Africa. International Soil Reference and Information Centre (ISRIC)/World Data Centre for Soils

Batjes, N.H. 2001. Management options for reducing CO2 concentrations in the atmosphere by increasing carbon sequestration in the soil." ISRIC Technical Paper 30/ NRP Report No 410 200 031.

Batjes, N.H. 2001b. Options for increasing carbon sequestration in West African soils: an exploratory study with special focus on Senegal." Land Degradation and Development. 12:131-142.

Batjes, N.H., 1996. Total carbon and nitrogen in the soils of the world. Eur. J . Soil Sci. 47, 151 – 163.

Berthe, Y. 1997. The role of forestry in combating desertification. Direction Nationale des Ressources Forestières, Fauniques et Halieutiques, Bamako, Mali. XI World Forestry Congress. Antalya, Turkey.

Birch-Thomsen, T., B. Elberling, B. Fog, J. Magid. 2007. Temporal and spatial trends in soil organic carbon stocks following maize cultivation in semi-arid Tanzania, East Africa. Nutrient Cycling in Agroecosystems. 79(3): 1385-1314.

Boffa, J-M., G. Yaméogo, P. Nikiéma, D.M. Knudson. 1998. Shea nut (Vitellaria paradoxa) production and collection in agroforestry parklands of Burkina Faso. FAO - Non-Wood Forest Products 9, Rome. http://www.fao.org/docrep/w3735e/w3735e17.htm, retrieved on 09/26/07.

Bogaert, P. 2002. Spatial prediction of categorical variables: the Bayesian Maximum Entropy approach. Stoch. Env. Res. Risk A. 16:425-448.

Bogaert, P. 2004. Predicting and simulating categorical random fields : the BME approach. Proceedings of the 1st International Conference for Advances in Mineral Resources Management and Environmental Geotechnology (AMIREG 2004), 119-126, June 7-9, Chania, Crete.

Bogaert, P. and D.D'Or. 2002. Estimating soil properties from thematic soil maps: the Bayesian Maximum Entropy approach. Soil Sci. Soc. Am. J. 66:1492-1500.

Bogaert, P. 2003 (in press). Spatial prediction of categorical variables: the BME approach. In: Proceedings of the Fourth Conference on Geostatistics for Environmental Applications - geoEnV IV. Sanchez-Vila X., J. Carrera and J. Gomez-Hernandez Eds. Dordrecht, The Netherlands, Kluwer Academic Publishers. Quantitative geology and geostatistics.

Bogaert, P., M. L. Serre, G. Christakos. 1999. Bayesian Maximum Entropy methods using transformations. Lippard S.J., Naess, A. and Sinding-Larsen, R. Proceedings of the 5th Annual Conference of the International Association for Mathematical Geology Trondheim, Norway., pp. 57-62.

Bogaert, P., M.A. Wibrin. 2004. Combining categorical and continuous information within the Bayesian Maximum Entropy paradigm. Presented at GeoEnV V - Geostatistics for Environmental Applications, October 13-15, Neuchatel, Switzerland.

Bohan, H. L., McNeal, G. A. O'Connor. 2001. Soil Chemistry, 3rd. John Wiley & Sons. New York.

Brady, C. N., R.R. Weil. 2002. The Nature and properties of soils. 12ed. Prentice-Hall. New York.

Brown, S. 1997. Estimating Biomass and Biomass Change of Tropical Forests: a Primer. Forestry Paper 134, FAO, Rome. (http://www.fao.org/docrep/W4095E/w4095e00.htm. retrieved on 09/26/07).

Brown, S., A.J.R. Gillespie, A.E. Lugo. 1989. Biomass estimation methods for tropical forests with applications to forest inventory data. Forest Science 35:881-902.

Capoor, K., P. Ambrosi. 2006. State and trends of the carbon market 2006. World Bank.

Casenave, A., C. Valentin. 1990. Les etats de surface de la zone sahel. Influence sur l'infiltration. Institute Français de Recherche Scientifique pour le Développement. 212p.

Cassiani, G., G. Christakos. 1998. Analysis and estimation of natural processes with non-homogeneous spatial variation using secondary information. Math. Geol. 30:57-76.

Chen, F., D. E. Kissel, T. Larry, T. West, W. Adkins. 2000. Field-scale mapping of surface soil organic carbon using remotely sensed imagery. Soil Science Soc. Am. J. vol. 64:746-753.

Chivenge, P.P., H.K. Murwira, K.E.Giller, P. Mapfumo, J. Six. 2006. Long-term impact of reduced tillage and residue management on soil carbon stabilization: Implications for conservation agriculture and contrasting soils. Soil & Tillage Research. 94:328-237.

Christakos G., M.L. Serre. 2002. Soft data space/time mapping of coarse particulate matter annual arithmetic average over the US. GeoEnv2002, 4th European Confer. on Geostatistics for Envir. Appl's, Barcelona, Spain, Nov. 27-29.

Christakos, G. 1990. A Bayesian/maximum-entropy view to the spatial estimation problem. Math. Geol. 22:763-777.

Christakos, G. 2000. Modern spatiotemporal geostatistics. Oxford University Press, New York.

Christakos, G., 1992. Random field models in earth sciences: Academic Press, San Diego, California.

Christakos, G., P. Bogaert, M.L. Serre. 2002. Temporal GIS. Advanced functions for field-based applications. Springer-Verlag, New York NY.

Christakos, G., Serre M.L. 2000. BME analysis of spatiotemporal particulate matter distributions in North Carolina Atmospheric Environment. 34(20):3393-3406.

Chunsheng, M. 2005. Spatio-temporal variograms and covariance models. Advances in Applied Probability. 37 (3):706-725.

Cissé, M.I. 1995. Les parcs agroforestiers du Mali: état des connaissances et perspectives pour leur amélioration. Consultant's Report n° 93. Semi-arid Lowlands of West Africa (SALWA). DRFH/IER. 53 pp.

Coleman, D.C., J.M. Oades, G. Uehara, 1989. Dynamics of soil organic matter in tropical ecosystems, NifTAL Project, Honolulu.

Comegna V., C. Vitale. 1993. Space–time analysis of water status in a volcanic Vesuvian soil. Geoderma 60 (1993) (1-4), pp. 135–158.

Cressie, N. 1994. Comment on "An approach to statistical spatial-temporal modeling of meteorological fields" by M. S. Handcock and J. R. Wallis. J. Am. Statist. Ass., 89, 379–382.

Cressie, N., H. Huang. 1999. Classes of nonseparable, spatiotemporal stationary covariance functions. Journal of American Statistical Association. 94:1330-1340.

CSE (Le Centre du Suivi Ecologique). 2000. Sequestration du carbone dans les sols carbon sequestration in soils): Atelier International, Dakar. Ministere De L'Environnment.

De Cesare, L., D. Myers, D. Posa. 2001. Estimating and modeling space–time correlation structures. Statistics and Probability Letters 51 (1):9–14.

Delisle, L. 2006. A spatial approach to estimating soil carbon Stocks at the field level. MSc Thesis. 152 p. University of Hawaii.

Deutsch, C.V., A.G. Journel. 1998. GSLIB: Geostatistical software library and user's guide. 2nd ed. Oxford Univ. Press, New York.

D'Or, D., P. Bogaert, G. Christakos. 2001. Application of the BME approach to soil texture mapping. Stochastic Environmental Research and Risk Assessment (SERRA). 15(1):87-100.

Doumbia, M., A. Jarju, M. Sené, K. Traoré, K. Brannan, R. Yost, R. Kablan, C. Yamoah, P.C. S. Traoré, A. Ballo. (in review) Soil organic carbon increases associated with ridge-tilllage. Agronomy for sustainable development.

Economic Commission for Africa Poverty Reduction Strategy Paper (PRSP), Learning Group. 2001. Strategic Framework for Poverty Reduction (CSLP). The Experience of Mali.

Englund, E. 1991. GEO-EAS 1.2.1- Geostatistical environmental assessment software - user's guide. U.S. Environmental Protection Agency - Environmental Monitoring Systems Laboratory.

ESRI, 2006. ArcGIS 9.2. Redlands California.

Eswaran, H., Berg, E.V.D., Reich, P. 1993. Organic carbon in soil of the world. Soil Sci. Soc. Am. J. 57, 192–194.

FAO. 1998. Current world fertilizer situation and look. 1996/997–2002/2003. Food and Agriculture Organization, Rome.

Feller, C., 1993. Organic inputs, soil organic matter and functional soil organic compartments in low-activity clay soils in tropical zones. In: Mulongoy, K., Merckx, R. (Eds.), Proceedings of an International Symposium on Soil Organic Matter Dynamics and Sustainability of Tropical Agriculture. 4–6 November 1991, Leuven, Belgium. Wiley, New York, pp. 77–88.

Fiandez, M. S. 1998. Dry forest silviculture in the Sudano-Sahelian region: Burkina Faso's experience. FAO - Burkina Faso (FAO/BKF/89/0 11)

Garnett, M.H., Ineson, P. , Stevenson, A.C., 2000. Effects of burning and grazing on carbon sequestration in a Pennine blanket bog, UK. Holocene 10, 729 – 736.

Garrity, D. P. 2004. Agroforestry and the achievement of the Millennium Development Goals. Agroforestry Systems. 61: 5-17.

Gigou, J. 2007. Les stratégies de gestion conservatoire des eaux et des sols en zone Soudanienne et Sahélienne. SM-CRSP workshop, Mali.

Goovaerts, P. 1997. Geostatistics for Natural Resources Evaluation. Oxford University Press. NY.

Goovaerts, P. 2001. Geostatistical modeling of uncertainty in soil science. Geoderma. 103:3-26.

Goovaerts, P., C. N. Chiang. 1993. Temporal persistence of spatial patterns for mineralizable nitrogen and selected soil properties. Soil Sci. Soc. Am. J. 57:372–381.

Gruijter, J. D. Brus, M. Bierkens, M. Knotters. 2006. Sampling for natural resource monitoring. Springer-Verlag. The Netherlands.

Harvey, L. D. D., 2004. Declining temporal effectiveness of carbon sequestration: Implications for compliance with the United National Framework Convention on Climate Change. Climatic Change. 63:259-290.

Heath, L.S. J.E. Smith. 2000. Soil carbon accounting and assumptions for forestry and forest-related land use change. P. 89-1001. In Joyce, L.A. and R. Birdsey, eds. The Impact of Climate Change on America's Forest. USDAQ Forest Service, General Technical Report RMRS-GTR-59. 134 p.

Hernandez, P. R. 2004. Assessing carbon stocks and modelling win-win scenarios of carbon sequestration through land-use changes, FAO Rome.

Heuvelink, G.B.M., and R. Webster. 2001. Modeling soil variation: Past, present, and future. Geoderma 100:269-301.

Houghton, R.A. 2003. Why are estimates of the terrestrial carbon balance so different? Global Change Biology 9:500–509.

Huntley B.J., Walker, B.H. 1982. Ecology of Tropical Savanas. Springer-Verlag. New York.

Ingram, J.S.I., E.C.M. Fernandes. 2001. Managing carbon sequestration in soils: concepts and terminology. Agriculture Ecosystems and Environment. 87:111-117.

IPCC (Intergovernmental Panel on Climate Change). 1997. An introduction to simple climate models used in the ipcc second assessment report. IPCC Technical Paper II (eds. J. T. Houghton, L. G. M. Filho, D. J. Griggs and K. Masket). IPCC, Geneva, Switzerland. (www.ipcc.ch/ipccreports/special-reports.htm retrieved on 07/ 2007).

IPCC (Intergovernmental Panel on Climate Change), 2000 (b). Nebojsa Nakicenovic and Rob Swart (Eds.) Emissions Scenarios. Cambridge University Press, UK. pp 570 (Retrieved from http://www.cambridge.org/ipcc. retrieved on 07/2007)

IPCC (Intergovernmental Panel on Climate Change). 2000 (c). Robert T. Watson, Ian R. Noble, Bert Bolin, N. H. Ravindranath, David J. Verardo and David J. Dokken (Eds.) Land Use, Land-Use Change, and Forestry Cambridge University Press, UK. pp 375.

IPCC (Intergovernmental Panel on Climate Change). 2001. Climate Change 2001. The scientific basis. Cambridge University Press, New York.

IPCC (Intergovernmental Panel on Climate Change). 2005. IPCC Special Report on Carbon Dioxide Capture and Storage. Prepared by Working Group III of the Intergovernmental Panel on Climate Change [Metz, B.,O. Davidson, H. C. de Coninck, M. Loos, and L. A. Meyer (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 442 pp.

IPCC (Intergovernmental Panel on Climate Change). R.T.Watson, M.C.Zinyowera, R.H.Moss (Eds). 1997. The regional impacts of climate change: An assessment of vulnerability. Cambridge University Press, UK. pp 517. (www.ipcc.ch/ipccreports/special-reports.htm retrieved on 07/ 2007).

IPCC(Intergovernmental Panel on Climate Change). 2000 (a). Special report on land use change and forestry. Summary for policy makers. Geneva, Switzerland. 20pp.

IPCC(Intergovernmental Panel on Climate Change). 2007. IPCC fourth assessement report: Climate Change 2007: Synthesis Report

Isaac, M. E., A. M. Gordon. 2005. Temporal changes in soil carbon and nitrogen in west African multistrata agroforestry systems: a chronosequence of pools and fluxes. Springer NY. 65:23-31.

Isaaks, E., R. Srivastava. 1989. An introduction to applied geostatistics. Oxford University Press, New York.

Ishida, T. Ando, H. 1998. Use of disjunctive cokriging to estimate soil organic matter from Landsat thematic Mapper image. Int. J. Remote. 20:1549-1565.

Janzen, H. 2004. Carbon cycling in earth systems- a soil science perspective. Agriculture, Ecosystems and Environment. 104:399-417.

Jobbagy, E.G, R.B. Jackson. 2000. The vertical distribution of soil organic carbon and its relation to climate and vegetation. Ecological Applications. 10 (2):423-436.

Jones, J.W., J. Koo, J.B. Naab, W.M. Bostick, S. Traoré, W.D. Graham. 2007. Integrating stochastic models and in situ sampling for monitoring soil carbon sequestration Agricultural Systems. 94(1):52-62.

Jones, R.H., Y. Zhang. 1997. Models for continuous stationary space-time processes. In Modeling Longitudinal and spatially correlated data, eds. T.G. Gregoire, D.R. Brillinger, P.J. Diggle, E. Russek-Cohen, W.G. Warren, and R.D. Wolfinger, New York: Springer-Verlag. Pp. 289-298.

Jost, G., G.B.M. Heuvelink, A. Papritz. 2005. Analysing the space-time distribution of soil water storage of a forest ecosystem using spatio-temporal kriging. Geoderma. 128:258-273.

Journel, A., C. J. Huijbregts. 1978. Mining geostatistics, Academic Press, London, U.K.

Kablan, R., K. Brannan, R.S. Yost, K. Brannan, M.D. Doumbia, K. Traoré, A. Yoroté, Y. Toloba, S. Sissoko, O. Samaké M. Vaksman, L. Dioni, M. Sissoko. 2008. Amenagement en courbes de niveau", increasing rainfall capture, storage, and drainage in soils of Mali. Arid Lands Research and Management. 22:62-80.

Kieft, T. L., C. S. White. 1998. Temporal dynamics in soil carbon and nitrogen resources at a grassland-shrubland ecotone. JSTOR. 79:671-683.

Kimble et al., 2002 J.M. Kimble, R. Lal and R.F. Follett, Agricultural Practices and Policies for Carbon Sequestration in Soil, Lewis Pub., Boca Raton, FL, USA (2002) 512 p.

Kolovos, A., G. Christakos, D.T. Hristopulos, M.L. Serre. 2004. Methods for generating non-separable spatiotemporal covariance models with potential environmental applications. Advances in Water Resources. 27:815-830.

Krebs, C., 1989. Ecological methodology. HarperCollins, New York.

Kwesiga, F.R., Izac, A-M. N., Ndiritu, C.G. and Woomer, P.L. 1997. Soil fertility replenishment in Africa: An investment in natural resource capital. In: R.J. Bursch, P.A. Sanchez and F. Calhoon (eds.) Replenishing Soil Fertility in Africa. Soil Science Society of America Special Publication. 51:1-46. Madison, Wisconsin.

Lal, R. 2000. World cropland soils as a source or sink for atmospheric carbon, Advances in Agronomy 71:145–191

Lark R.M., 2002b. Optimized spatial sampling of soil for estimation of the variogram by maximum likelihood. Geoderma. 105 (1-2):49-80.

Lark, R, M. 2002 Robust estimation of the pseudo cross-variogram for Cokriging soil properties. European Journal of Soil Science. 53:253-270.

Liu, S. M. Kairé, E. Wood, O. Diallo, L.L. Tieszen. 2004. Impacts of land use and climate change on carbon dynamics in south-central Senegal. Journal of Arid Environments. 59(3):583-604.

Louppe, D. 1995. Tarifs de cubage pour les espèces de bois d'œuvre de la foret Badenou. IDEFOR_DFO, CIRAD Foret. Cote d'Ivoire.

Manlay, C.R.J, D. Masse, J.L. Chotte. 2002. Carbon, nitrogen, and phosphorus allocation in agro ecosystems of a West African savanna Agriculture, Ecosystems, and Environment. 88:233-248.

Manlay, R. 2000. Organic matter dynamics in mixed-farming systems of the West African Savanna. A village case study from South Senegal. IRD - ENGREF. Montpellier. PhD Dissertation. 190pp.

Maraux, F., P. Dugué, F. Ganry. 2007. Improving soil fertility and rehabilitating degraded soils: technical and social dynamics in Western and Central African drylands. In : CIRAD, FARA, CTA. AIDA Conference, Agricultural Innovation in Dryland Africa, Accra, Ghana. Personal communication. (http://publications.cirad.fr/auteur.php?mat=989, retrieved on July, 2007)

McBratney, A.B., M.L. M. Santos, B. Minasny. 2003. On digital soil mapping. Geoderma. 117:3-52.

Mooney, H., J. Canadell, F. S. Chapin, J. Ehleringer, Ch. Korner, R. McMurtrie, W. Parton, L. Pitelka, and D.-E. Schulze. 1999. Ecosystem physiology responses to global change. In B. H. Walker, W. L. Steffen, J. Canadell, and J. S. I. Ingram, editors. The terrestrial biosphere and global change. Implications for natural and managed ecosystems. Cambridge University Press, London, UK.

Montagnini, F., P. K. R. Nair. 2004. Carbon sequestration: An underexploited environmental benefit of agroforestry systems. Springer. 61:281-295.

Muller, T. G., F.J. Pierce. 2003. Soil Carbon maps: enhancing spatial estimates with simple terrain attributes and multiple scales. Soil Sci.Soc. Am. Journal. 67:258-267.

Nelson, D.W., L.E. Sommers. 1982. Total Carbon and organic matter, in methods of soil analysis, Part 2. Chemical and Microbiological Properties – Agronomy Monograph no. 9 (2nd Edition). Madison, USA.

Nicholson, S.E. 2001. Climatic and environmental change in Africa during the last two centuries. In: Desanker P (ed) Africa and global climate change. CR SPECIAL 8. Clim Res. 17:123–144.

Nielsen, D.R., O. Wendroth. 2003. Spatial and temporal statistics. GeoEcology textbook, Catena, Reiskirchen.

Niemann, K.O., S. Adams, and G. J. Hay. 1998. Automated tree crown identification using digital orthophoto mosaics. Proceedings on Automated Interpretation of High Spatial Resolution Imagery for Forestry. Victoria, B.C. 105-114.

Papoulis, A. 1991. Probability, random variables, and stochastic processes. McGraw-Hill International Editions, Singapore.

Parton, W.J., J.W.B. Stewart, C.V. Cole. 1988. Dynamics of C, N, P and S in grassland soils: a model, Biogeochemistry. 5:109–13.

Perez, C., C. Roncoli, C. Neely, J. L. Steiner. 2007. Can carbon sequestration markets benefit low-income producers in semi-arid Africa? Potentials and challenges. Agricultural Systems. 94(1):2-12.

Piéri, C. 1989. Fertilité des terres de savanes – Bilan de trente ans de recherche et de développement agricoles au sud du Sahara, Paris, Ministère de la Coopération et du Développement and CIRAD-IRAT, 444.

Ribeiro, Jr., P.J., 1999. Bayesian inference in Gaussian model-based geostatistics: Technical Report ST-99-08. Lancaster University/UFPR. UK.

Ringius, L. 2001. What prospects for soil carbon sequestration in the cdm. Cop-6 and beyond Energy & Environment. 12(4):275-285.

Ringius, L. 2002. Soil carbon sequestration and CDM: Opportunities and challenges for Africa. Climate Change. 54:471-495.

Robert, M. 2001. Soil carbon sequestration for improved land management. FAO. Rome.

Rosenzweig, R., M. Varilek, B. Feldman, R. Kuppalli, and J. Janssen. 2002. The Emerging International Greenhouse Gas Market. Pew Center for Global Climate Change.

Sahu, S. 2005. A Bayesian kriged Kalman model for short-term forecasting of air pollution levels. Applied Statistics. 54(1):223-244.

Sanchez, P.A., K.D., Shepherd, M.J., Soule, F.M., Place, A.U., Mukwunye, R.J.,Bursch, Sarac, C.; I. H. Demirel, O. I. Sen, N. Varol. 2004. Geostatistical simulation of the total organic carbon values: An example from petroleum source rocks on the coastal area of Western Taurus region, Turkey. Petroleum Science and Technology. 22(3-4): 367-379.

Schelsinger, W. H. 1995. An overview of the carbon cycle , in Lal, R., Kimble, J., Levine, E., and Stwart, B.A. (eds), Soil and Global Change, CRC Lewis Publishers, Boca Raton. 9-25.

Schlesinger, W. 1999. Carbon sequestration in soils. Science. 284:5423.

Schlesinger, W. 2000. Carbon sequestration in soils: some cautions amidst optimism. Agricultural Ecosystems and Environment. 82:121-127.

Schneider, T.D, 2000. Information theory primer with an appendix on logarithms, National Cancer Institute.

Serre, M.L. 1999. Environmental spatiotemporal mapping and ground water flow modeling using the BME and ST methods. Ph. D. diss. University of North Carolina - Chapel Hill.

Sivia, D. S. 1996. Data Analysis: A Bayesian Tutorial. Clarendon Press. Oxford. New York.

Soil Survey Staff. 2006. Keys to Soil Taxonomy, 10th ed. USDA-Natural Resources Conservation Service, Washington, DC.

Srivastava, M.R., 1996. An overview of stochastic simulation. In: Mowrer, H.T., Czaplewski, R.L., Hamre, R.H. (Eds.), Spatial accuracy assessment in natural resources and environmental sciences: Second Int. Symposium. U.S. Dept. of Agriculture, Forest Service, Fort Collins, General Technical Report RM-GTR. 277:13–22.

Stavins, R. N. 1999. The costs of carbon sequestration: A Revealed-Preference Approach. The American Economic Review. 89(4):994-1009.

Stein, M. L. 2005. Space-time covariance functions. Journal of the American Statistical Association. 100:310-321.

Steven, S. L., S. de Neve, G. Hofman P. Boeck, D. Beheydt, O. van Cleemput, I. Mestdagh, P. Lootens, L. Carlier, N. Van Camp, H. Verbeeck, I. Vande Walle, R. Samson, N.¨ Lust and R. Lemeur. 2003. Carbon stock changes and carbon sequestration potential of Flemish cropland soils. Global Change Biology. 9:1193–1203.

The Mathworks. 2007. Getting started with MATLAB. Natick, MA (www.mathworks.com. retrieved on 09/26/07).

Theng, K.G., K.R. Tate, P. Sollins. 1989. Constituents of organic matter in temperate and tropical soils p. 5-32. In Coleman, D.C., J.M. Oades, G. Uehara, 1989. Dynamics of soil organic matter in tropical ecosystems, NifTAL Project, Honolulu.

Tiessen, H, C. Feller, E.V.S.B. Sampaio, P.Garin. 1998. Carbon sequestration and turnover in semiarid savannas and dry forest. Climatic Change. 40: 105-117.

Tiessen, H., Salcedo I.H., Sampaio E.V.S.B. 1992. Nutrient and soil organic matter dynamics under shifting cultivation in semi-arid northeastern Brazil. Agric. Ecosyst. Environ. 38:139-151.

Tieszen, L.L.,G.G. Tappan, A. Toure. 2004. Sequestration of carbon in soil organic matter in Senegal: An Overview. Journal of Arid Environments. 59(3):409-425.

Tilmann G., M.G. Genton, P. Guttorp. 2002. Geostatistical space-time models, stationarity, separability and full symmetry. Technical Report no. 475. Department of Statistics University of Washington.

Traoré, K. 2003. Le parc à karité : sa contribution à la durabilité de l'agrosysteème. Cas d'une toposéquece à Konobougou (Mali-Sud).

PhD. Ecole Nationale Superieure Agronomique de Montepellier. Montepellier. France.

Traoré, S.B., F. Reyniers, M. Vaksman, B. Koné, A. Sidibé, A. Yoroté, K. Yattara, M. Kouressy. 2000. Adaptation à la sécheresse des ecotypes locaux de sorghos du Mali. Sécheresse. 11:227-237.

Tschakert, P., G. Tappan. 2004. The social context of carbon sequestration consideration from a multi-scale environmental history of the Old Peanut Basin of Senegal. Journal of Arid Environments. 59(3):511-533.

United Nations. 2001 (a). Report on the Conference of the Parties on the second part of its sixth section held at Bonn from 16-27 July. FCCC/CP/2001/5/Add.1. 57 pp.

United Nations. 2001(b). Report on the Conference of the Parties on the second part of its sixth section held at Bonn from 16-27 July. FCCC/CP/2001/5/Add.2. 78 pp.

United Nations. 2001(c). Kyoto Protocol to the United Nations Framework Convention on Climate Change. (www.un.org, retrieved on 01/2006).

United Nations. 1992. United Nations Framework Convention on Climate Change. New York.

Vagen, T. G., R. Lal, B. R. Singh. 2005. Soil carbon sequestration in Sub-Saharan Africa: A review. Land Degradation & Development. 16:53–71.

Walter, C. R. A. V. Rossel, A.B. McBratney. 2003. Spatio-temporal simulation of the field-scale evolution of organic carbon over the landscape. Soil Sci. Soc. Am. J. 67:1477-1486.

Webster, R., M. A. Oliver. 2001. Geostatistics for environmental scientists. John Willey and Sons, Ltd. England.

West, T.O., W.M. Post. 2002. Soil organic carbon sequestration rates by tillage and crop rotation: a global data analysis, Soil Science Society of America Journal, 66:1930–1946.

WIIAD (Winrock International Institute for Agriculture Development). 1997. A guide to monitoring carbon storage in forestry and agroforestry projects. Winrock, USA, Forest Carbon Monitoring and Verification Services.

Wikle, C. K., L. M. Berliner, N. Cressie. 1998. Hierarchical Bayesian space-time models. Environ. Ecol.Statist. 5:17–154.

Woomer, P.L. A. Touré, M. Sall. 2004. Carbon stocks in Senegal's Sahel transition zone. Journal of Arid Environments. 59(3):499-510.

Woomer, P.L., L. L. Tieszen, G. Tappan, A. Toure, M. Sall. 2004b. Land use change and terrestrial carbon stock in Senegal. Journal of Arid Environments. 59(3):625-642.

Woomer, P.L., L. L. Tieszen, W. J. Parton, A. Toure. 2001. Landscape Carbon Sampling and Biogeochemical Modeling. J. Arid Environments 59:499-510.

Yost, R.S., K. Loague, R. Green. 1993, Reducing variance in soil organic carbon estimates: soil classification and geostatistical approaches. Geoderma. 57:247–262.

Yost, R.S., P. Doraiswamy, Doumbia M. 2002. Defining the contract area: using spatial variation in land, cropping systems and soil organic carbon. In: A Soil Carbon Accounting System for Emissions Trading – Special Publication, Soil Management Collaborative Research Support Program (Ed.). University of Hawaii, Honolulu, HI, pp. 13–40.