Developing increased EO capacity

for better agriculture and forestry management in Africa

LIVESTOCK SYSTEMS – TECHNICAL REPORT

Deliverable 32.1

Julien Minet[1], Abdoul Aziz Diouf[1,2], Issa Garba[1,3], Marie Lang[1,4,5], Bakary Djaby[1,3], Jason Sircely[4,6], Bernard Tychon[1], and Richard Conant[4,6]

[1]Université de Liège, Arlon Campus Environnement, Arlon, Belgium

[2]Centre de Suivi Ecologique, Dakar, Senegal

[3]Centre Régional AGRHYMET, Niamey, Niger

[4]International Livestock Research Institute, Nairobi, Kenya

[5]Vlaamse Instelling voor Technologisch Onderzoek, Mol, Belgium

[6]Natural Resource Ecology Laboratory, Colorado State University, Fort Collins, Colorado, USA

Date	2015-02-12	**Contract delivery date**	T0+36
Deliverable No	D32.1		
WP title	WP3.2 Livestock Systems		
WP leader	ILRI		
Status (F: Final, D: Draft, RD: Revised Draft)			F
Dissemination level (PU=public, RE=restricted, CO=confidential)			PU

CHANGE RECORD

Issue	Date	Page(s)/ Section(s)	Description of Change	Editor(s)/ Reviewer(s)
0.1	2014-07-11		First issue	Julien Minet
0.2	2014-07-14			Carolien Toté
0.3	2014-08-08		Adding several subsections	Julien Minet
0.4	2014-10-29		Addition global G-range use case, post-doc and related training efforts	Jason Sircely, Richard Conant
0.5	2015-01-27		Review and completing last missing subsections	Julien Minet
1.0	2015-02-12		Final Review	Qinghan Dong

TABLE OF CONTENTS

1 BACKGROUND ... 12
1.1 Scope and objectives ... 12
1.2 Related documents ... 12
1.3 About this report ... 13
2 INTRODUCTION ... 14
2.1 Basic concepts ... 14
2.2 Livestock in the use case areas ... 14
2.2.1 Senegal use case ... 14
2.2.2 Niger use case ... 14
2.2.3 Kenya use case ... 15
2.2.4 Global use case ... 15
2.3 Objectives ... 16
3 STARTING POINT ... 18
3.1 Senegal use case ... 18
3.2 Niger use case ... 18
3.3 Kenya use case ... 18
3.4 Global use case ... 18
4 CAPACITY BUILDING ACTIVITIES ... 20
4.1 Training workshop on forage biomass modelling, Belgium, Dec 2012 ... 20
4.2 Workshop on yield forecasting using remote sensing data, Niamey, Niger, Feb. 2013 ... 20
4.3 Workshop on fodder production forecasting using remote sensing data, Niamey, Niger, Feb. 2013 21
4.4 Training workshop on the G-Range ecosystem model, USA, Oct 2013 ... 21
4.5 Regional training on forecasting of forage resources in Ouagadougou, Burkina Faso, Jan. 2014 ... 22
4.6 Regional training on the G-range model in Nairobi, Kenya, Feb. 2014 ... 22
4.7 G-Range ecosystem model training in Khartoum, Sudan, Feb. 2014 ... 23
4.8 G-Range ecosystem model training in Arlon, Belgium, May 2014 ... 24
5 RESEARCH ACTIVITIES ... 25
5.1 Senegal use case ... 25
5.1.1 Research objectives and context ... 25
5.1.2 Materials and methods ... 25
5.1.3 Results and discussion ... 28
5.1.4 Conclusions ... 34
5.2 Niger use case ... 36
5.2.1 Research objectives and context ... 36
5.2.2 Materials and methods ... 36

5.2.3 Results and discussion ... 39
5.2.4 Conclusions and perspectives ... 47
5.3 Kenya use case ... 48
5.3.1 Research objectives and context ... 48
5.3.2 Study area and data set ... 49
5.3.3 Linear regression models using SPOT Vegetation indicators ... 52
5.3.4 Generalised linear models ... 62
5.3.5 Influence of location and season on drought related livestock mortality ... 67
5.3.6 General conclusions and perspectives ... 69
5.4 Global use case ... 69
5.4.1 G-Range ... 69
5.4.2 Uses and publications ... 74
6 IMPACT OF USE CASE DEVELOPMENT ... 75
6.1 Biomass estimation in Senegal and Niger ... 75
6.2 Index Based Livestock Insurance implementation ... 75
7 CONCLUSIONS ... 76
7.1 Conclusions on capacity building activities ... 76
7.2 Scientific conclusions ... 76

LIST OF FIGURES

Figure 1: Location of study sites in administrative departments of interest 27

Figure 2: NDVI profile of a pixel in the study area in 2011 (blue) and smoothed curve with the Savitzky-Golay method (brown) 28

Figure 3 : Herb biomass estimated (kg/ha) with NDVIi and NDVIpk metrics from 2006 to 2010......... 30

Figure 4 : Total biomass estimated (kg/ha) with NDVIi and NDVIpk metrics from 2006 to 2010 31

Figure 5 : Accuracy of model estimation for herb biomass and total biomass using NDVIi and NDVIpk between 2006 and 2010... 32

Figure 6: Diagram for the determination of variables from VAST 37

Figure 7: Observed values as a function of predicted values of biomass [kg DM/ha] for the global adjustment. 41

Figure 8: Residuals distribution, per lines (left) and histogram (right) 42

Figure 9: Observed values of biomass production as a function of predicted values, Azawak 43

Figure 10: Residuals analysis, Azawak. 44

Figure 11: Observed values of biomass production as a function of predicted values, Manga2 45

Figure 12: Residuals analysis, Manga2 45

Figure 13: The study area: District of Marsabit, Northern Kenya. 49

Figure 14 : Temporal characteristics of the study. 50

Figure 15: Constructed seasonal vegetation indicators. 52

Figure 16: Probability plot (up) and histogram of overall livestock mortality (down), normal distribution (red), exponential distribution (green). 54

Figure 17: Histogram of livestock mortality for bad (red) and good (green) vegetation conditions, normal distribution (black), exponential distribution (blue).. 55

Figure 18: Means comparisons for mortality classified according to the season (top) and the geographical cluster (bottom) (0: Laisamis cluster and 1: Chalbi cluster) 55

Figure 19: Mean comparison for livestock mortality classified according to the condition of vegetation (0: bad, 1: good. Red line: general mean, green line: sample mean). 56

Figure 20: Seasonal livestock mortality versus CZNDVIpos.. 56

Figure 21: Comparison of means for correlations classified by indicator, results of Student's T and Tukey-Kramer tests (green line = sample mean, blue lines = standard deviation lines and red line = general mean) 60

Figure 22: Plot for linear fit (left) and residuals in overall case (right).. 61

Figure 23: Distribution of residuals for the overall case and goodness of normal distribution fit. 62

Figure 24: Coefficients estimates and their associated Wald 95% confidence interval for model 1.... 66

Figure 25: Coefficients estimates and their associated Wald 95% confidence interval for model 2.... 66

Figure 26: G-range model evaluation results for field biomass harvest data. 71

Figure 27: G-Range model evaluation results for MODIS TNPP and Century soil carbon.................... 72

Figure 28: G-Range model evaluation results for MODIS VCF. 73

LIST OF TABLES

Table 1: Summary of the national workshop on forage biomass modelling in Arlon, Belgium, December 2012 20

Table 2: Summary of the AGRICAB workshop on yield estimation in Niamey, Niger, February 2013.. 20

Table 3: Summary of the AGRICAB workshop on fodder production in Niamey, Niger, February 2013 21

Table 4: Summary of the national workshop on the G-range ecosystem model, USA, October 2013. 21

Table 5: Summary of the AGRICAB regional training on forecasting of forage resources in Ouagadougou, Burkina Faso, January 2014 22

Table 6: Summary of the AGRICAB regional training on the G-range model in Nairobi, Kenya, February 2014 23

Table 7: Summary of the G-range ecosystem model training in Khartoum, Sudan, February 2014 23

Table 8: Summary of the G-range global rangelands model training in Arlon, Belgium, May 2014..... 24

Table 9: Basic equation of adjustment functions 26

Table 10 : Herb biomass estimation models with NDVIi and NDVIpk metrics for the years 2006 to 2010 29

Table 11: Total biomass estimation models with NDVIi and NDVIpk metrics for the years 2006 to 2010 30

Table 12 : Most important explanatory variables for herb biomass multilinear modelling in Senegal rangelands 32

Table 13 : Original set : selected models and performance criteria 33

Table 14 : Resampled set : selected models and performance criteria 34

Table 15 : Mean Absoulte Error (MAE, kg/ha), standard deviation (STDD, kg/ha) and Relative Root Mean Squared Error (RRMSE, %) of best models obtained by selection on bootstrap sample (n=number of observation) 34

Table 16: Exploratory and bootstrap analysis 40

Table 17: Four best global models 40

Table 18: Estimation of the coefficients of the global model with 4 variables 41

Table 19: Exploratory and bootstrap analysis of residuals 41

Table 20: Number of sampling sites by ecoregions 42

Table 21: Four best models - Azawak 42

Table 22: estimation of the model coefficients, Azawak 43

Table 23: Estimation of coefficients, residuals analysis, Azawak. 43

Table 24: Statistics of the regression – manga2 44

Table 25: Coefficients of the regression – manga2 44

Table 26: Number of observations sites per facies 46

Table 27: Models per facies 46

Table 28: Models by year 47

Table 29: Summary of data set........ 50

Table 30: Distribution analysis for seasonal livestock mortality. 53

Table 31: Coefficients of correlation between seasonal livestock mortality and NDVI-, DMP-, and fAPAR-derived variables. Highlighted cells represent the significant coefficients. 58

Table 32: Coefficients of correlation between NDVI-derived variables. 59

Table 33: Results of Student's T and Tukey-Kramer tests for mean comparison between NDVI-, DMP- and fAPAR-Mortality correlations 59

Table 34: Results of linear regression between livestock mortality and NDVI-derived variables. 60

Table 35: Coefficient for significant linear models. 61

Table 36: Description of the models and parameters. 63

Table 37: Results of Generalized linear models. 64

Table 38: Coefficients of the Generalized linear models. 65

Table 39: Results of one variable ANOVA (* mean squares = sum of squares/DF). 67

Table 40: Effect of the location and season on the global mortality mean. 68

ACRONYMS

AGRHYMET	Centre agro-hydrométrique
AMS	AgroMetShell
ASAL	Arid and Semi-Arid Lands
CSE	Centre de Suivi Ecologique
CILSS	Comité inter-États de lutte contre la sécheresse au Sahel
DMP	Dry Matter Productivity
DPSIR	Driving forces – Pressures – States – Impacts – Responses
DVI	Difference Vegetation Index
ECMWF	European Centre for Medium-Range Weather Forecasts
EVI	Enhanced Vegetation Index
fAPAR	fraction of Absorbed Photosynthetically Absorbed Radiation
GIS	Geographical Information System
GLM	Generalised linear models
HDR	Habilité à Diriger la Recherche
ILRI	International Livestock Research Institute
ILBI	Index Based Livestock Insurance
LAI	Leaf Area Index
LOOCV	Leave-One-Out Cross Validation
LULC	Land Use Land Cover
MAE	Mean Absolute Error
MEIA	Ministère de l'Elevage et des Industries Animales (Niger)
MODIS	Moderate-Resolution Imaging Spectroradiometer
MOSS	Microsoft Office SharePoint Server
NDVI	Normalized Difference Vegetation Index
REE	Relative Estimated Error
RFE	Rainfall Estimates
RMSE	Root Mean Square Error
RRMSE	Relative Root Mean Square Error
SAVI	Soil-adjusted Vegetation Index
SPOT	Satellite Pour l'Observation de la Terre
TAMSAT	Tropical Applications of Meteorology using SATellite data and ground-based observations
OSS	Observatoire du Sahara et du Sahel

UAM	University Abdou Moumouni
ULg	Université de Liège
VAST	Vegetation Analysis in Space and Time
VGT	VEGETATION sensor(s) on board SPOT4 (VGT1) and SPOT5 (VGT2)
VIP	Variable Importance in the Prediction
VITO	Vlaamse Instelling voor Technologisch Onderzoek
WP	Work Package

EXECUTIVE SUMMARY

This document reports on the main research and capacity building activities that were conducted in the AGRICAB WP32 "Livestock systems" since the beginning of the AGRICAB project. The Livestock system work package (WP) is led by ILRI with participations of ULg, CSE and OSS. The WP aims at estimating the fodder production in rangelands and monitoring livestock by using models fed by biophysical variables from remote sensing data. These models can then be used as basis for operational Early Warning System. Another research activity within the WP is the development of an Index Based Livestock Insurance (ILBI) based on remote sensing data in order to cover livestock losses due to fodder scarcity caused by dry meteorological conditions.

Capacity building activities (see section "4 Capacity building activities") implied eight training workshop in Belgium, Niger, Burkina Faso, Kenya, Sudan and USA, which were all dedicated to fodder production assessment methods in Sahel.

The research activities in this WP mainly rely on three PhD researchers and one post-doc in three African countries:

- in Senegal (Abdoul Aziz Diouf): Parametric and non-parametric forecasting of fodder production using remote sensing data
- in Niger (Issa Garba): Parametric and non-parametric forecasting of fodder production in Niger
- in Kenya (Marie Lang): Development of a pastoral insurance tool from Earth Observation data for arid and semi-arid lands
- in Kenya (Jason Sircely – Post-doc): Development of a global grazinglands model.

The section "5 Research activities" summarizes the research activities conducted in the context of these PhD projects and forms the main part of this report. The first two PhD projects are similar in terms of topic and methods but they are within two different countries. First results of Abdoul Aziz Diouf showed that new models linking the fodder production to biophysical variables from remote sensing data are worth being developed, (1) using multilinear regression instead of simple linear regression model and (2) by performing zoning of the model by phenological zones. Similarly, the work of Issa Garba focused on the development of new regression-based models for linking in situ measured fodder production to biophysical variables from remote sensing data. First results from Marie Lang showed that livestock mortality is poorly explained by vegetation indicators from remote sensing data using multiple linear regressions or generalized linear models. Characterization of the geographical areas and of seasons might help for the development of a livestock mortality model. The post-doctoral project has led to the development of a new grazinglands model that has been applied globally to forecast the impacts of climate change on grazingland productivity and regionally to assess the impacts of different management strategies. The three PhD research are now at their mid-term and most of the results and research outputs are to be expected in the next few years.

1 Background

1.1 Scope and objectives

The scope of this document is to report on the main research activities that were conducted in the AGRICAB WP32 "Livestock systems" that is led by ILRI and with participations of ULg, CSE and OSS. The objectives of the report are twofold: first, to report on capacity building activities and, second, to summarize the main scientific research activities that were conducted within this work package. The description of the research work is mainly present in the section "5 Research activities", is a summary of the methods and results gained by the researchers during the project and is the largest part of the report.

1.2 Related documents

Related documents that are publicly available are listed here below:

RD1 Fonctions d'ajustement pour l'estimation de la production fourragère herbacée des parcours naturels du Sénégal à partir du NDVI S10 de SPOT-VEGETATION http://www.agricab.info/Publications/Documents/Article1_Abdoul_Aziz_diouf.pdf

RD2 Estimation de la production fourragère au Sénégal à partir de l'indice de végétation par différence normalisée (NDVI) du satellite SPOT-Vegetation http://www.agricab.info/Publications/Documents/Poster_Abdoul_Aziz_Diouf_Vf.pdf

RD3 Protocole d'Analyse des données satellitaires de végétation et de pluviométrie pour le suivi de la campagne agro-pastorale au Sénégal http://www.agricab.info/Publications/Documents/Poster_Suiv_Camp_CSE_AGRICAB_V6.pdf

RD4 Analyse de la performance du modèle d'estimation de la biomasse du Ministère de l'élevage et des industries Animales (MEIA) du Niger http://www.agricab.info/Publications/Documents/modele_MEIA_NIGER_ARS_UAM_final.pdf

RD5 Suivi de la végétation par satellite : cas de l'utilisation des images ICN, VCI et SNDVI pour la prévision qualitative des productions végétales http://www.agricab.info/Publications/Documents/Suivi_culture_paturage_images_satellitaires_final.pdf

RD6 Influence of location, season and vegetation on drought related livestock mortality assessment from SPOT NDVI data in Marsabit District, Northern Kenya http://www.agricab.info/Publications/Documents/AGRICAB_WS_Paper_LivestockMortality_MLang_final.pdf

RD7 Prediction of Drought Related Livestock Mortality in Arid and Semi Arid Lands by remote sensing: Study of the impact of land cover classes http://www.agricab.info/Publications/Documents/AGRICAB_MLang_poster.pdf

RD8 D3.1 Use case and capacity building requirements [AGRICAB_D3.1.pdf] http://www.agricab.info/Achievements/Documents/AGRICAB_D3.1.pdf

1.3 About this report

This report describes the main activities that were conducted in the Work Package, mainly through the work done by the three PhD students and a post-doctoral researcher involved: Abdoul Aziz Diouf, Issa Garba, Marie Lang, and Jason Sircely. This report was coordinated by Julien Minet, ULg. It is mainly based on contributions of by the above researchers as well as their supervisors, in particular Djaby Bakary and Richard Conant. Abdoul Aziz Diouf specifically contributed on the Senegal use case (2.2.1 Senegal use case; 5.1 Senegal use case). Issa Garba specifically contributed on the Niger use case (2.2.2 Niger use case; 5.2 Niger use case). Marie Lang specifically contributed on the Kenya use case (2.2.3 Kenya use case; 5.3 Kenya use case). Jason Sircely's contributions are described in the Global use case (5.4 Global use case).

2 Introduction

2.1 Basic concepts

Livestock herding is a key agricultural activity in the Sahel. Largely relying on pastoralism in rangelands, livestock productivity is closely linked to fodder biomass production. In that respect, remote sensing data can help to evaluate and monitor the state of vegetation and to estimate biomass production in rangelands. Monitoring vegetation state through satellite imagery is often using vegetation indices such as the NDVI (Normalized Difference Vegetation Index), fAPAR (fraction of Absorbed Photosynthetically Absorbed Radiation) or LAI (Leaf Area Index). Satellite imagery must be available at a sufficient temporal resolution (~ few days) and at a reasonable spatial resolution (100 m – 1 km).

Several models of rangeland productivity have been developed based on biophysical variables coming from remote sensing data. These models can be based on an empirical relationship linking the state of vegetation (e.g., biomass measured in situ) to the biophysical variables that are remotely-sensed. Models can range from simple linear regression models, linking the biomass to a satellite index, to multiple regression models based on the combination of several biophysical variables from remote sensing, meteorological data and ecoregions characteristics. A seasonal characterization of the biomass – satellite indices models could also improve the predictive results. These models can be calibrated at the national or regional scales, or might be improved at the local scale taking into account the ecophysiological characteristics of smaller areas.

Calibrated and validated models are useful tools for assessing the fodder production in a given area thanks to satellite imagery. Operational tools can be derived to make biomass production forecasting and early warning systems.

2.2 Livestock in the use case areas

The Livestock System research is carried out on specific sites in three different focus countries, Senegal, Niger and Kenya, and complemented by a global-scale model development activity. The three PhD projects: Senegal (Abdoul Aziz Diouf), Niger (Issa Garba) and Kenya (Marie Lang) and a post-doc (Jason Sircely) lead these efforts. .

2.2.1 Senegal use case

Livestock is the primary renewable resource in the Sahel (Dicko et al, 2006), particularly in West Africa. In Senegal, livestock plays a very important role in socioeconomic development of societies. It provides about one third of the national agricultural affluence (CIRAD, 2010), particularly in rural areas, where livestock affects 30% of the population for which it provides food security, savings, labor-force and fertilizing of fields (ISRA, 2003). It is a pastoral breeding system with large herds of cattle and small ruminants (RPCA, 2010), for which more than 90% of the dry matter eaten by livestock comes from pasture (Carrière, 1996). Thus, natural rangelands are an almost indispensable component for the satisfaction of needs of livestock production. They also play important ecological roles, through soil conservation, carbon sequestration, biodiversity conservation and ecotourism development.

2.2.2 Niger use case

Livestock plays an economic role of a paramount importance in Niger, as it is the second country's export income, it counts 36 millions of animals and is practiced by 80% of the households (SRP, 2007). Livestock herding in Niger is traditionally managed with a weak level of investment. Livestock mainly relies on the natural fodder production with large intra-seasonal variations of productivity. In addition to its first economic role, livestock plays also a crucial role for food security of the population.

2.2.3 Kenya use case

The Kenya use case focuses on Northern Kenya that is characterized by a dry climate with bimodal rainfall pattern (one short dry and short rain season and one long dry and long rain) (Chantarat et al., 2013). As in the other arid and semi-arid lands (ASALs) of Africa, Northern Kenya is quite regularly facing severe droughts. For example, within the last 100 years, 28 major droughts where recorded in that area and 4 of them happened in the last 10 years (Adow, 2008).

In ASALs, poor households rely mostly on livestock as a source of livelihood (Mude et al., 2010). In Northern Kenya, livestock accounts for more than two thirds of the income for more than three million people (Chantarat et al., 2013; Mude et al., 2010). In these regions, drought is the predominant cause of livestock mortality. Livestock losses due to drought can be particularly high during drier seasons (severe weather shocks), such as during the extremely poor rainfall period in 2000 (Mude et al., 2010). In those regions, characterized by rain-fed agriculture and livestock based economy, people are highly vulnerable to extreme weather events (extreme temperature or excessive or insufficient rainfall) (UN Department of Economic and Social Affairs, 2007) and it is now known that, in the event of such a weather shock, households without risk transfer mechanisms are more likely to drive into permanent poverty (Barret & Mc Peak, 2006).

2.2.4 Global use case

Temperate and tropical rangelands occur throughout the world and are the most extensive land type on earth. These lands are biologically diverse, and support the livelihoods of millions of households. In most rangelands, there is not sufficient precipitation for agriculture, and so families use livestock to essentially turn sunlight into food. Historically, the inhabitants of these areas had to contend with droughts, fires, livestock raids, and other stressors and shocks. Great variability in precipitation and the climatic extremes that characterize semi-arid and arid areas led to coupled natural and human systems that were inherently flexible; the adaptive capacity of the system components – their ability to adapt to a changing environment – was high. Ungulates evolved physical and behavioral adaptations such as high heat tolerance and migration, and people adopted behaviors such as transhumant movement and developed cultural norms that minimized exposure to stresses.
Thus, several design objectives directed the further development and final parameterization of the G-Range ecosystem model (v1.1):

- A simulation tool for global rangelands that captures main primary production, and its dynamics;
- A tool of moderate complexity, one that could be useful to a new user in a week or less;
- A structure that includes simulating changes in all rangelands across the globe within asingle executed process;
- A monthly time-step in simulations;
- Representation of global vegetation at least at the scale of herbaceous, shrubs, and trees;
- The ability for the proportions of those different kinds of plant types to change over time;
- Simulations that may span from about 5 to 100 or more years;
- The ability to include natural or management modifications to rangelands, such as through fire or fertilization;

- Programming structures that will allow the software to be parallelized for use on multiprocessor clusters or networks, although making the software run parallel was not part of this effort;
- Greater concern for clarity in logic and ease of use than in conserving hard drive space or algorithmic elegance;
- Portability in the G-Range code, allowing simulations to be done on a variety of platforms (e.g., Windows, Linux cluster);
- A graphical user interface (GUI) that is weakly linked to G-Range. G-Range should be able to be run in batch model, without input from a user, and related to that;
- Parameter files used, rather than inputs from a GUI, allowing simulations to be made without using the interface;

Output should be straightforward spatial surfaces, without complex summary analyses.

2.3 Objectives

The main objective of this WP is to implement fodder biomass production models using remote sensing data in the Sahel, through case studies in three African countries: Senegal, Niger and Kenya and a global use case using the newly developed G-range model. The expected outputs are refined models of fodder productivity and tools for implementing early warning systems for the livestock sector. Specific objectives imply (1) the calibration of fodder production model using archive and newly collected field data and remote sensing data and (2) the improvement of models linking the fodder production and the biophysical variables derived from remote sensing data. At the same time, index-based livestock insurance (IBLI) schemes are being studied in the Kenya use case with the objective of insuring pastoralists during drought periods. Using the models developed in this work package together with the insurance products being developed will help farmers make better decisions on when to sell or buy animals in the following season.

In particular, this WP relies on three PhD students and one post-doc in the three study case countries and a global use case:

- Mr Abdoul Aziz Diouf, Senegal
- Mr Issa Garba, Niger
- Ms Marie Lang, Kenya
- Dr. Jason Sircely, global application

The first two PhD, Abdoul Aziz Diouf and Issa Garba, focus on the implementation of the fodder biomass production models based on remote sensing in Senegal and Niger, respectively, while Marie Lang works on the implementation of an index-based livestock insurance scheme in Kenya. Abdoul Aziz Diouf works partly at the Centre de Suivi Ecologique (CSE), Dakar, Senegal, where the biomass production in Senegal is monitored since 1987, and partly at the Université de Liège – Arlon (ULg). Issa Garba works at the Agrhymet Regional Centre, Niamey, Niger, which covers several countries of West Africa in agrometeorological studies and environmental monitoring. Marie Lang works in Belgium at Université de Liège – Arlon and at Vlaamse Instelling voor Technologisch Onderzoek (VITO) and closely collaborates with the International Livestock Research Institute (ILRI) in Nairobi, Kenya. The global use case is focused on developing a mechanistic grazingland model applicable globally.

A strong methodological objective of this WP is to share knowledge, tools and experiences between the three PhD students and the post-doc and their respective institutions, to compare methodologies and produce joint outputs such as joint publications.

3 Starting point

3.1 Senegal use case

To answer needs of information on resources management and food security with effects of drought and agricultural pressure on grazing land since the early 1970s (ISRA 2003), the national estimate of the fodder yield from remotely sensed data is carried out by the Centre de Suivi Ecologique (CSE) of Dakar since 1987. The method used was proposed by Tucker et al (1983). It is a technique based on empirical relationships between remotely sensed indices and biomass data collected on the ground sites each year. The biomass data are collected at the end of growing season (towards the end of October), through 36 sites distributed across the whole pastoral domain of Senegal. Then, biomass collected is used to calibrate the NDVI (Normalized Difference Vegetation Index) obtained actually from seasonal integration of SPOT-VEGETATION imageries provided by VITO. The result is a fodder production map with a 1 km * 1 km resolution on the entire country or Pastoral Units studied. This information on biomass production enables establishment of fodder balances and thereby, constitutes a guiding tool for annual strategies of pastoral resources backup.

However, this method has a number of constraints as especially the cost of ground observation devices (sampling across whole of Senegal), the high time between the collection and publication of results (shift with needs of information systems on food security) and imprecision which could be improved. This situation led to the formulation of the thesis research project of Abdoul Aziz Diouf ("Parametric and non-parametric forecasts of herbaceous fodder production with remote sensing data") on Senegal for development of models (reliable and responsive to the context of the West African Sahel) to estimate and forecast the fodder yields.

3.2 Niger use case

In Niger, the inventory of the field biomass time series has been done on the period 1999 to 2011 by the livestock department unit of the AGRHYMET Regional Centre (Niger). Biomass production prediction is made using an approach using the the integral of the NDVI curve which becomes an explanatory variable for biomass.

3.3 Kenya use case

Information about the index based livestock insurance (IBLI) program can be found in the AGRICAB deliverable D3.1 (see RD8). The IBLI program was launched in 2010 in a pilot program conducted by the International Livestock Research Institute (ILRI) in the Marsabit District, Northern Kenya (Mude et al., 2010). The contract was designed to manage important livestock losses caused by the kind of severe droughts that this area regularly faces. The index used in that contract is derived from the decadal Normalized Difference Vegetation Index (NDVI) obtained from the Advanced Very High Resolution Radiometer (AVHRR) sensor (Chantarat et al., 2009).

3.4 Global use case

G-Range was not intended to be programmed 'from scratch.' Completing such an effort would take more time than we had available to dedicate to the project, and would duplicate the efforts of many researchers. Numerous grassland and rangeland simulation models have been developed. We explored a variety of models, to different degrees (i.e., SimSAGS, MAPSS, IBIS, the Hurley Pasture

Model, GEM-1, Biome-BGC, GENDEC, Grazing Lands Application, GRAZPLAN, PHYGROW, and Pasture Quality Model, SAVANNA, and Century). Some models are quite complex, making global application unwieldy. Other models are too simple, not allowing for scenario analyses of the types we intend. Some simulation models are pointbased, inappropriate for a spatially explicit global simulation model. Some models use simple rules to infer changes, which can be unrealistic. For example, logic in a model may dictate that a drought of a given severity causes a given percentage of herbaceous plants to die. These models are inappropriate for our use. More practically, some were judged out-of-date based on their Web sites. Others appeared ‘closed source’ rather than ‘open source’ packages, such that the suitability of the software for our use would be difficult to judge without requests to the authors. Some were commercial products and were excluded from consideration.

4 Capacity building activities

Eight workshops were organized in different countries about the research activities that are conducted in the work package. These workshop were often tailored-made for the needs of the participants, especially the workshop with few participants. Four out of the eight workshops were about a hands-on training with the G-range model.

4.1 Training workshop on forage biomass modelling, Belgium, Dec 2012

Date	December 2012, 2 weeks
Location	Arlon, Belgium
Topic	Prévision des ressources fourragères et pastorales
Participants	2
Diffusion	All AGRICAB partners

Table 1: Summary of the national workshop on forage biomass modelling in Arlon, Belgium, December 2012

This small training workshop was tailored for two scientists from Senegal, working respectively in the Centre de Suivi Ecologique (CSE) and at the Direction de l'Analyse, de la Prévision et des Statistiques agricoles (DAPSA), in Senegal. The workshop was organized and given by Djaby Bakary in Arlon, Belgium.

4.2 Workshop on yield forecasting using remote sensing data, Niamey, Niger, Feb. 2013

Date	18-22 February 2013
Location	Niamey, Niger
Topic	Estimation des rendements à partir des images satellitaires
Participants	15
Participant PhD students	Issa Garba
Diffusion	All AGRICAB partners

Table 2: Summary of the AGRICAB workshop on yield estimation in Niamey, Niger, February 2013

This training workshop has gathered 15 participants from technical department of the Ministry of Agriculture of the CILSS countries and from the university Abdou Moumouni (UAM). It was held at the AGRHYMET regional centre in Niamey, Niger. Five sessions were given by four trainers from AGRHYMET and ULg (Djaby Bakary).

4.3 Workshop on fodder production forecasting using remote sensing data, Niamey, Niger, Feb. 2013

Date	25 February – 1 March 2013
Location	Niamey, Niger
Topic	Prévision des productions fourragères par utilisation des images satellitaires
Participants	~20
Participant PhD students	Issa Garba
Diffusion	All AGRICAB partners

Table 3: Summary of the AGRICAB workshop on fodder production in Niamey, Niger, February 2013

This training workshop, held at the AGRHYMET regional centre in Niamey, Niger, has gathered about 20 participants from technical department of the Ministry of Agriculture and from producers organisations of the CILSS countries and from the university Abdou Moumouni (UAM). Sessions were given by trainers from AGRHYMET, UAM and ULg (Djaby Bakary).

4.4 Training workshop on the G-Range ecosystem model, USA, Oct 2013

Date	October 2013, 1 month
Location	Fort Collins, Colorado, USA
Topic	Basic and advanced extended G-Range model training workshop
Participants	2
Diffusion	All AGRICAB partners

Table 4: Summary of the national workshop on the G-range ecosystem model, USA, October 2013

This workshop was designed to train 2 colleagues from Agricultural Research Corporation (Wad Madani, Sudan), one senior researcher and one master's student, in basic through advanced application, evaluation, and parameterization of the G-Range global rangeland ecosystem model: http://www.nrel.colostate.edu/projects/grange/index.php. Basic training was conducted on running the G-Range model at global and site scales, and processing of output. Advanced training was conducted on applying the model in two study areas selected by the participants, the Greater Horn of Africa and the Butana region of Sudan. Model application consisted of analyzing the model output for the study areas, evaluating model output using field biomass data and remote sensing layers, and conducting sensitivity analysis for several parameters of probable significance in the study areas. The researchers trained in this workshop currently seek to apply G-Range in forecasting forage production under climate change in the respective study areas, with the goal of assessing Sudanese government livestock feed supplementation programs.

4.5 Regional training on forecasting of forage resources in Ouagadougou, Burkina Faso, Jan. 2014

Date	13-25 January 2014
Location	Ouagadougou, Burkina Faso
Topic	Atelier AGRICAB: Prévision des ressources fourragères et pastorales
Participants	30
Participant PhD students	Abdoul Aziz Diouf
Diffusion	All AGRICAB partners

Table 5: Summary of the AGRICAB regional training on forecasting of forage resources in Ouagadougou, Burkina Faso, January 2014

As part of AGRICAB project activities, Abdoul Aziz Diouf participated at the regional training workshop held in Ouagadougou from 13 to 25 January 2014. The trip was funded by the University of Liège (Flight) and CSE (subsistence). This participation was part of the livestock systems work package (WP 3.2) coordinated by University of Liege and its partner in Senegal: the CSE. The purpose of the participation was: i) to make presentation on remote sensing based biomass estimation used in Senegal currently, but also the doctoral research of Abdoul Aziz Diouf which should lead to proposals for improvement of this method, especially for the estimation of herbaceous biomass, ii) to benefit from exchanges with pastoral experts invited in this session.

This workshop helped the participants to deepen their knowledge on the SPIRITS tool and the statistical modelling software CST. It was also an opportunity to discover method of estimating areas and agricultural statistics currently applied in Senegal, Kenya and Mozambique under the AGRICAB project. For the session on livestock systems, a presentation of Abdoul Aziz Diouf was given and all questions of participants were answered.

The workshop was globally a great contribution for researchers because they were able to meet and exchange especially with national experts in charge of pastoral issues in some countries of West Africa Sahel such as Burkina Faso (Mrs. Djara, Monitoring Pastoral Resources Agent), Ivory Coast (Mr. Bi Tré Tré, Agronomist-Animal Scientist), Niger (Boureima, Head of Pastoral Resources Management Division), Mali (Dr Konaté, Ing Agro-pastoraliste) and Chad (Dr Oueddo, Researcher- Agro-pastoraliste). Representatives of Action Against Hunger (ACF) Organisation from Dakar and Mali did also participate to the workshop.

4.6 Regional training on the G-range model in Nairobi, Kenya, Feb. 2014

Date	3-14 February 2014
Location	Nairobi, Kenya
Topic	AGRICAB Regional Thematic Workshop at RCMRD
Participants	~30
Participant PhD students	Marie Lang
Diffusion	All AGRICAB partners

Table 6: Summary of the AGRICAB regional training on the G-range model in Nairobi, Kenya, February 2014

In the 2.5 days allotted, participants were trained in basic application and evaluation of the G-Range global ecosystem model for rangelands: http://www.nrel.colostate.edu/projects/grange/index.php. Training was conducted on running the G-Range global rangeland ecosystem model at global and site scales, processing of output, and site-scale evaluation using field biomass data from Nairobi National Park over the first two days of the workshop. In the final half-day of the workshop, participants provided research questions and the group worked together to formulate hypotheses, determine the spatial and time scales involved, identify appropriate data sources for model evaluation, and identify the G-Range output variables that would be required for model evaluation and application. Proposed research questions of interest to the participants ranged widely, most focusing on land degradation and management effects on forage production among other ecosystem services, in the over 10 countries the workshop participants hailed from. Together with Jason Sircely, Marie Lang participated also as a trainer in this workshop.

4.7 G-Range ecosystem model training in Khartoum, Sudan, Feb. 2014

Date	4-7 February 2014
Location	Khartoum, Sudan
Topic	G-Range ecosystem model
Participants	9
Participant PhD students	N/A
Diffusion	All AGRICAB partners

Table 7: Summary of the G-range ecosystem model training in Khartoum, Sudan, February 2014

This workshop was in part a follow-up on and strengthen collaborations built during an earlier workshop in Fort Collins, CO, USA (October 2013), as well as to train additional researchers from Agricultural Research Corporation (Wad Madani, Sudan). Workshop participants included mostly master's and PhD students, as well as 2 senior researchers, who were trained in basic application and evaluation of the G-Range global ecosystem model for rangelands: http://www.nrel.colostate.edu/projects/grange/index.php. Training was conducted on running the G-Range global rangeland ecosystem model at global and site scales, processing of output, and site-scale evaluation using field biomass data from Nairobi National Park over the first two days of the workshop. In the final 1.5 days of the workshop, participants provided research questions and the group worked together to formulate hypotheses, determine the spatial and time scales involved, identify appropriate data sources for model evaluation, and identify the G-Range output variables that would be required for model evaluation and application, and participants with data from their study sites acted as the centers of group projects focused on model application and evaluation using their data.

4.8 G-Range ecosystem model training in Arlon, Belgium, May 2014

Date	19-23 May 2014
Location	Arlon, Belgium
Topic	G-Range ecosystem model
Participants	7
Participant PhD students	Issa Garba, Marie Lang [1]
Diffusion	All AGRICAB partners

Table 8: Summary of the G-range global rangelands model training in Arlon, Belgium, May 2014

The workshop aimed at providing a hands-on experience in G-Range, a global rangelands ecosystem model for simulating and forecasting production and carbon fluxes in rangelands: http://www.nrel.colostate.edu/projects/grange/index.php. It was intended for PhD students and early-stage researchers in ecological modeling, agriculture and remote-sensing. Initially, it was tailored for the three PhD students involved in AGRICAB at ULg: Abdoul Aziz Diouf, Issa Garba and Marie Lang. Other people interested in the theme of crop/ecological modeling subscribed to the meeting. Unfortunately, Abdoul Aziz Diouf could not participate because of the rejection of his visa. A webpage was made for information about the workshop: www.eed.ulg.ac.be/g-range-training/

Basic training was conducted on running the G-Range model at global and site scales, and processing of output, and site-scale evaluation using field biomass data from Nairobi National Park over the first two days of the workshop. The remaining 3 days of the workshop focused on applying the model in two study areas selected by the participants, at local scales in Morocco and at the national scale in Niger. Participants divided into two groups, with the Morocco group using remotely sensed dry matter productivity (DMP) data for evaluation, and the Niger group using field biomass harvest data for evaluation. Model application consisted of analyzing the model output for the study areas, and evaluating model output using the field biomass data and remote sensing layers.

[1] Aziz Diouf could not participate due to a visa issue.

5 Research activities

5.1 Senegal use case

5.1.1 Research objectives and context

The overall objective of the PhD research is to propose forecasting models of fodder and/or animal resources with the use of remote sensing data and ground biomass. The first inscription for the PhD was accepted for the academic year 2012-2013 with obligation to follow a doctoral training program in Science and Environmental Management at University of Liege. Research is carried out in a joint effort between Senegal (CSE, Dakar) and Belgium (ULg, Arlon) where Abdoul Aziz Diouf stays around 3 months per year.

The thesis committee is listed here below:

- Pr Bernard TYCHON (ULg, Supervisor),
- Dr Jacques André NDIONE, HDR (CSE, Co-supervisor)
- Dr Bakary DJABY, (ULg, Member of committee).

5.1.2 Materials and methods

Acquisition, preprocessing and data organization

All the information required, namely remote sensing and biomass data, were acquired. Remote sensing data are NDVI and DMP from SPOT VGT, MODIS09Q1 and MODIS09QA, rainfall of the ECMWF, and RFE of TAMSAT and FewsNet. NDVI and DMP pictures were downloaded on DevCoCast portal, MODIS image on the NASA-USGS LP DAAC and EFR Fews portal on the website of FewsNet. Biomass data were acquired from CSE database between 1999 and 2013 on 24 sites over the pastoral domain of Senegal. All data were pre-processed except MODIS and Tamsat images and implemented into access database. At this stage of our research, exploratory analysis of agro-meteorological data from ECMWF, particularly rainfall, has revealed a large number of outliers values in the study sites causing replacement of these data by RFE Fews in the next steps.

Statistical analysis and modelling of biomass

The work is mainly performed with SAS software. Most statistical analysis programs were coded (Simple regression, multiple, partial least squares method, Principal Component Analysis…).

Simple regression models

According to the research objectives, the estimation method of biomass based on simple linear regression was diagnosed through several adjustment variables based on NDVI and simple regression functions.

The issue here is that the relationship between aboveground biomass and NDVI is not always linear (Santin-Janin et al, 2009 and Bégué et al, 2011.), because of NDVI saturation when the vegetation becomes dense (e.g. Box et al., 1989, Xiaoping et al., 2011; Vescovo et al., 2012) as in the North-Soudanian zones of Senegal. The purpose is to find the most suitable function(s) for biomass estimation in Senegal rangelands.

- Study sites:

51 sites were used in this study: 36 for the model calibration and 15 for the validation.

- Remote sensing and biomass data

NDVIS10 satellite images from SPOT-VGT were used and derived variables were calculated. These are the integrated NDVI (NDVIi) and the maximum NDVI (NDVIpk) during the growing season, respectively used in Senegal and Niger. Biomass data concern herb and leaf of woody species collected from 2006 at 2010. This period was used especially for having maximum sites (and data) that can permit to test this empirical year-to-year approach (based only on annual data).

- Adjustment functions and model calibration

As presented in Table 1, six adjustment functions with the linear function commonly applied in Senegal were tested.

Model	Equation
Cubic	$Y = b0 + (b1 * X) + (b2 * X^2) + (b3 * X^3)$
Power	$Y = b0 * (X^{b1})$
Exponential	$Y = b0 * (e^{(b1 * X)})$
Linear	$Y = b0 + (b1 * X)$
Logarithmic	$Y = b0 + (b1 * \ln(X))$
Quadratic	$Y = b0 + (b1 * X) + (b2 * X^2)$
Y= biomass production (kg/ha) et X= NDVIi or NDVIpk	

Table 9: Basic equation of adjustment functions

- Consistency and accuracy of regression models

The typical NDVI values (from SPOT-VGT) range between 0.1 and 0.7 for vegetated areas (Jarlan et al., 2009). This implies that calibrated models were tested with values (by 0.15 steps) in this interval to depict their behavior for biomass estimation (consistent or no: estimated biomass must increase with NDVI values).

The model accuracies were calculated using data from 15 sites that represent about 30% of the total annual sample (from 51 sites) and were randomly selected with the "ALEA.ENTRE.BORNES" using MS EXCEL 2010. Precision calculation was done using the Relative Estimation Error (REE) (e.g. Jin et al, 2014). The following formulae were used:

$$REE = \sqrt{\frac{\Sigma[(O_i - P_i)/P_i]^2}{N}}$$

$$\text{Précision} = (1 - \text{REE}) * 100$$

where O_i represents the observed biomass production in the field, P_i, the biomass estimated by the model and N, the number of observations in the set.

Multiple linear regression models

Some tests were performed on multilinear forecasting models of herbaceous biomass from phenological metrics of NDVI at site scale.

The issue here is that the simple regression approach remains largely inaccurate (Diouf et al., 2014; Crépeau et al., 2003). Accuracy of Exponential and Power models is in average 50% for estimating herb biomass and 40% for the total biomass in the Senegal rangelands (five years period: 2006-2010). The purpose is to assess whether other approaches as the multilinear regression can they give models with best accuracy.

- Study sites:

Study was carried out with data from the 20 sites as represented in the Figure 1.

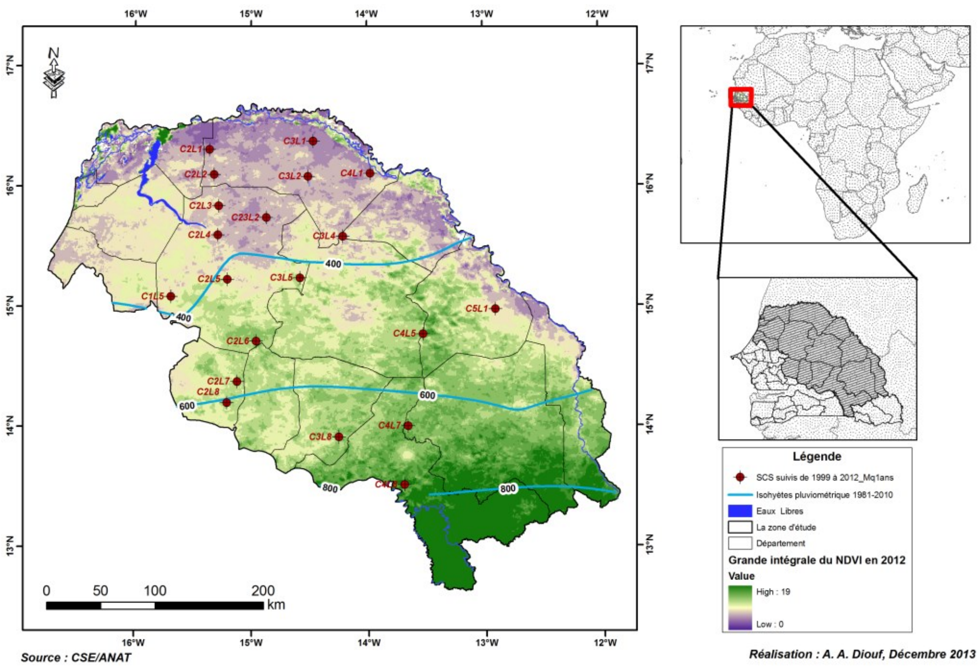


Figure 1: Location of study sites in administrative departments of interest

Remote sensing data correspond to NDVIS10 images from SPOT-VGT. It was used for calculation of phenological metrics using TIMESAT software. Some of the 11 metrics derived from NDVI profiles are presented in the Figure 2 (more information is available in Eklundh and Jönsson, 2011). Start and end of season are considered respectively at 20% and 50% after qualitative analysis of some pixels in the study area. Herb biomass was collected from 1999 to 2012 (except 2004 where data were not collected). Thus we have between 9 and 13 observations for calibrating site models since the monitoring was not done regularly for all sites.

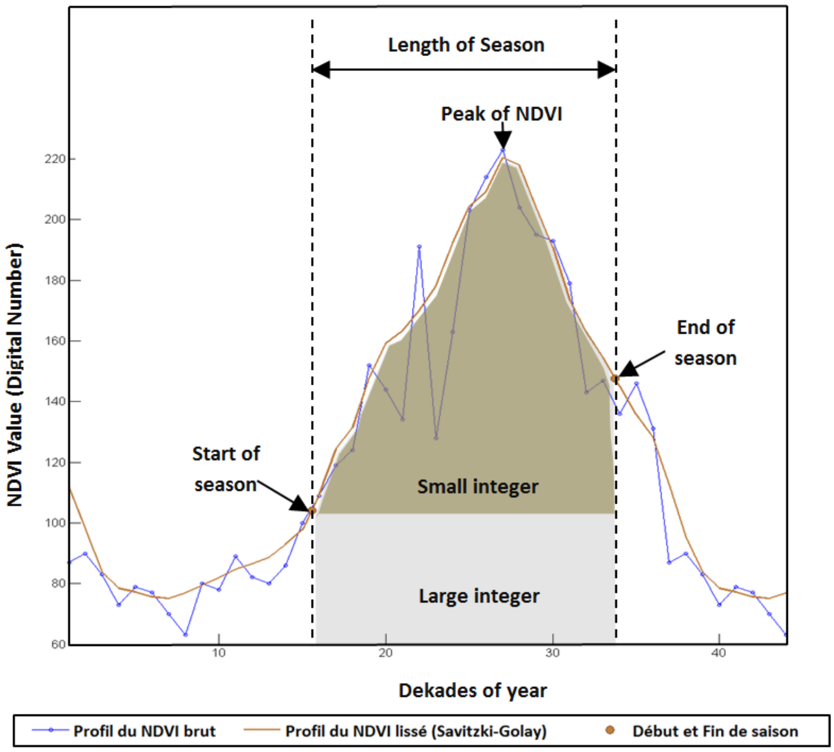


Figure 2: NDVI profile of a pixel in the study area in 2011 (blue) and smoothed curve with the Savitzky-Golay method (brown)

Determination of the explanatory variables was done using the Partial Least Squares (PLS) method that was developed in the 1960s by Herman Wold as an econometric technique (Tobias et al., 1995). This one is useful to reduce the number of variables when observations are not enough and thereby to avoid the over-fitting of model parameters.

- Criteria for selecting models

Criteria used are especially the adjusted Rsquare (R^2 aj), the mean square error (RMSE) and the relative mean square error (RRMSE). The R^2 aj is a corrected value of R ^{2}. It takes into account the influence of the number of predictive variables in a given model (Kouadio et al., 2012) and the size of sample.

- Resampling, calibration and validation of models

Due to the small number of observations per site for calibration of models, the data set was resampled for increase of observations number and thereby avoid the "over-fitting". Resampling was performed using the BOOTSTRAP method (see Saporta, 2006).

Validation was also performed on the resampled set with calculation of RMSE, RRMSE, MAE (Mean Absolute Error) for appreciating model performance. Specially, the RRMSE is useful since after Jamieson et al., (1991) the simulation is considered excellent with a normalized RMSE less than 10%, good if the normalized RMSE is greater than 10% and less than 20%, fair if normalized RMSE is greater than 20 and less than 30%, and poor if the normalized RMSE is greater than 30% .

5.1.3 Results and discussion

a. Simple regression models

Biomass estimation models

The coefficients of determination (R^2) and Fisher F-Value vary greatly from one year to another. This interannual variation in the relationship between NDVI and plant biomass is specific to this study area. As already notified by Prince (1991), this relationship was more variable in the Ferlo compared to other West African rangeland. Indeed, this variation is the rule rather than the exception for this area and could be attributed to natural variations in primary production or the sampling method used to estimate biomass (Diouf and Lambin, 2001). However, Exponential and Power models have the highest R^2 and F-Value ($p<0.05$) with values generally above 0.80 for the second model (Table 4 and 5). These two models correspond to the most plausible among the six studied either for herb biomass or total biomass estimation and any variables.

		2006		2007		2008		2009		2010	
Metric	**Model**	**R²**	**F-Value**	**R²**	**F-Value**	**R²**	**F-Value**	**R²**	**F-Value**	**R²**	**F-Value**
NDVIi	**Cubic**	0.66	20.43	0.55	13.19	0.52	11.41	0.46	9.01	0.26	3.82
	Exponential	**0.81**	**73.68**	**0.88**	**126.21**	**0.87**	**109.40**	**0.82**	**75.69**	**0.85**	**94.17**
	Linear	0.46	29.54	0.52	36.78	0.50	33.41	0.21	9.25	0.09	3.20
	Logarithmic	0.40	23.04	0.48	31.12	0.46	29.28	0.18	7.49	0.07	2.64
	Power	**0.80**	**66.09**	**0.88**	**125.16**	**0.86**	**107.95**	**0.81**	**70.62**	**0.84**	**91.99**
	Quadratic	0.59	23.92	0.55	20.12	0.51	17.49	0.35	8.76	0.17	3.31
NDVIpk	**Cubic**	0.72	27.92	0.70	25.05	0.58	15.01	0.44	8.47	0.19	2.55
	Exponential	**0.82**	**78.06**	**0.91**	**182.72**	**0.88**	**128.69**	**0.82**	**77.47**	**0.85**	**94.25**
	Linear	0.46	28.62	0.59	48.76	0.53	38.17	0.23	9.98	0.09	3.27
	Logarithmic	0.39	21.85	0.52	36.23	0.49	32.03	0.19	8.05	0.07	2.71
	Power	**0.80**	**67.95**	**0.91**	**166.60**	**0.88**	**123.55**	**0.81**	**72.41**	**0.84**	**92.42**
	Quadratic	0.60	24.95	0.67	33.89	0.58	22.40	0.38	9.92	0.17	3.40

Table 10 : Herb biomass estimation models with NDVIi and NDVIpk metrics for the years 2006 to 2010

		2006		2007		2008		2009		2010	
Metric	**Model**	**R²**	**F-Value**	**R²**	**F-Value**	**R²**	**F-Value**	**R²**	**F-Value**	**R²**	**F-Value**
NDVIi	**Cubic**	0.58	14.47	0.55	12.97	0.53	11.85	0.49	10.25	0.26	3.83
	Exponential	**0.84**	**86.28**	**0.85**	**98.30**	**0.87**	**116.62**	**0.89**	**134.09**	**0.86**	**107.69**
	Linear	0.28	13.48	0.44	26.48	0.43	25.14	0.29	13.63	0.13	5.25
	Logarithmic	0.22	9.43	0.38	20.68	0.39	21.51	0.24	10.73	0.12	4.61
	Power	**0.82**	**76.44**	**0.84**	**90.89**	**0.87**	**109.01**	**0.88**	**123.38**	**0.86**	**105.22**
	Quadratic	0.54	19.55	0.54	19.59	0.47	14.82	0.45	13.72	0.18	3.75
NDVIpk	**Cubic**	0.46	9.03	0.56	13.71	0.50	10.63	0.44	8.32	0.16	2.06
	Exponential	**0.80**	**69.15**	**0.84**	**88.10**	**0.87**	**115.20**	**0.88**	**122.06**	**0.86**	**105.75**
	Linear	0.19	7.83	0.38	21.03	0.42	25.08	0.24	10.72	0.13	4.93
	Logarithmic	0.13	5.25	0.32	16.07	0.39	21.79	0.19	8.20	0.12	4.42
	Power	**0.79**	**63.21**	**0.83**	**80.42**	**0.87**	**109.87**	**0.87**	**113.77**	**0.86**	**104.34**

Quadratic	0.43	12.57	0.49	15.99	0.47	14.35	0.44	12.79	0.16	3.15

Table 11: Total biomass estimation models with NDVIi and NDVIpk metrics for the years 2006 to 2010

Simulations profiles and model consistency

Exponential and Power models are also the most consistent compared, firstly, to Quadratic and Cubic models whose simulations does not always in the same direction as NDVI values in the range [0.1 to 0.7], and secondly, to the logarithmic and linear models which give negative values of biomass when the NDVI value is less than 0.3 (Figure 3 and 4). This means that linear model should be used with great care as it is not uncommon to observe these values (<0.3) in some parts of the Senegal rangeland, especially in the North, where vegetation cover can be relatively low. The negative values of biomass in the resulting raster did not match well with the calculation of the fodder balance.

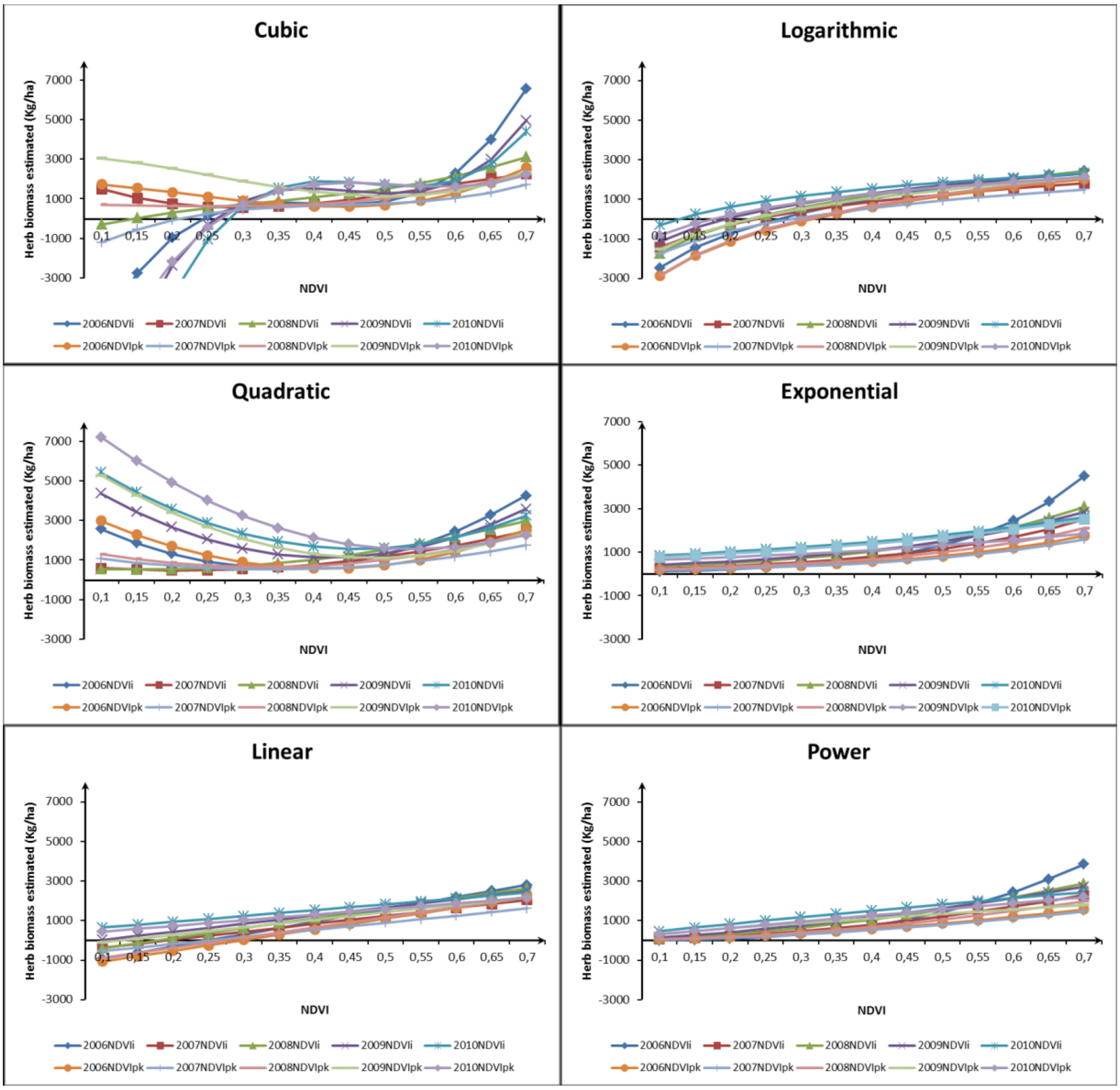


Figure 3 : Herb biomass estimated (kg/ha) with NDVIi and NDVIpk metrics from 2006 to 2010

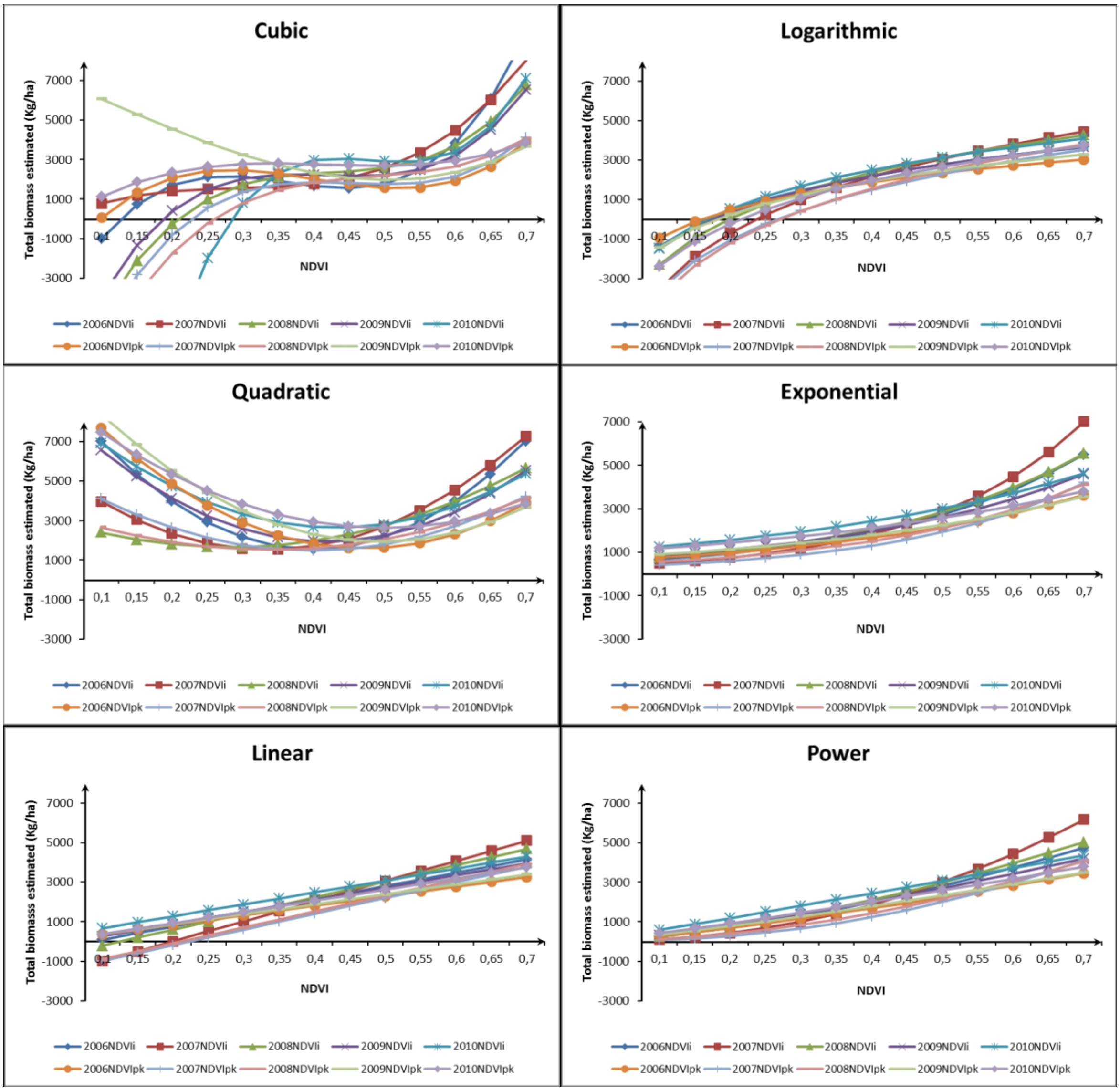


Figure 4 : Total biomass estimated (kg/ha) with NDVIi and NDVIpk metrics from 2006 to 2010

Models accuracy

The accuracy of models varies from year-to-year but generally can be related to the production type considered. Accuracy is in average higher for herb biomass estimation with values that can exceed 50%, whereas it is around 40% when estimating the total biomass. This situation could be explained by impact of foliage on the values of the index that is slightly lower than the herb layer for equivalent cover or biomass (Hiernaux and Justice, 1986). So because of difference in the relationship between NDVI and the leaf and herb biomass, it might be better to consider these two parameters separately (Diallo et al., 1991) during the development of simple regression models in this approach with data from the same year.

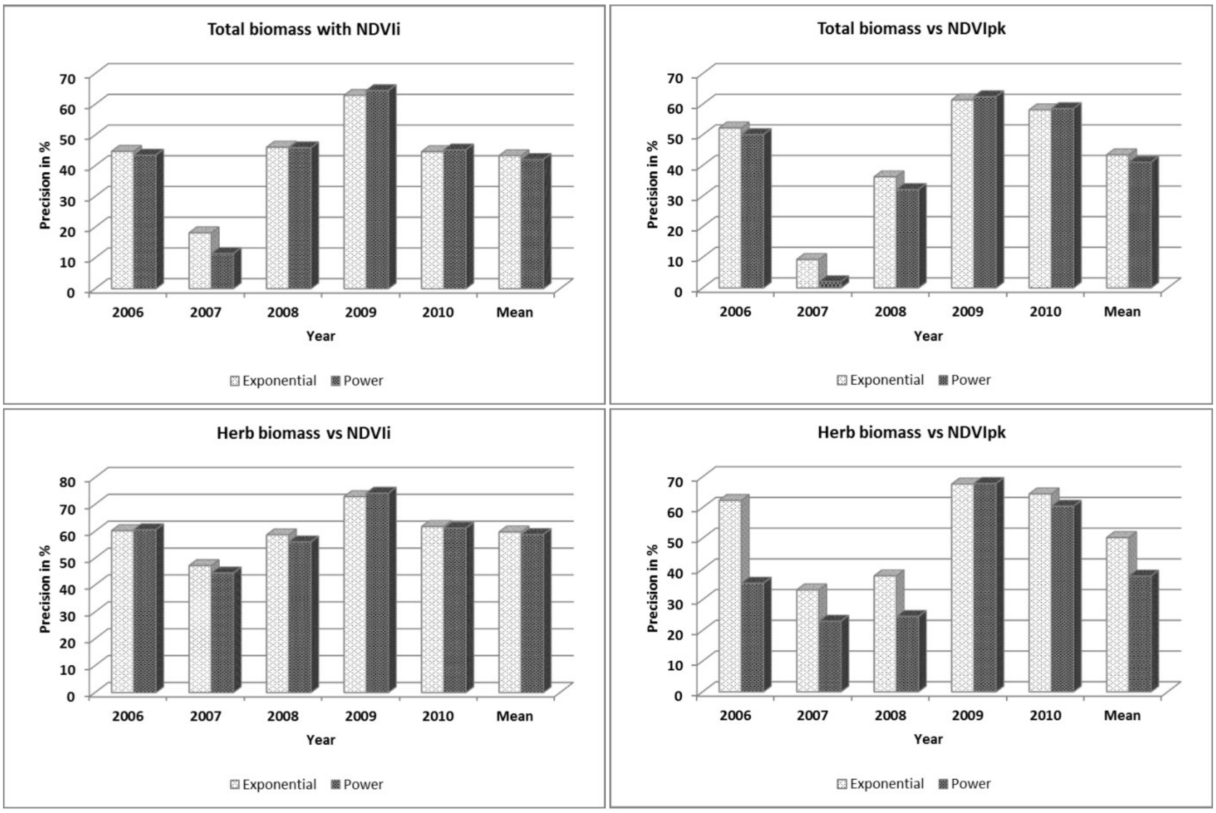


Figure 5 : Accuracy of model estimation for herb biomass and total biomass using NDVIi and NDVIpk between 2006 and 2010.

b. Multiple regression models at site scale

Models were calibrated in this study only for herb biomass estimation.

Determination of explanatory variables

After PLS regression, the most important variables in the model using the VIP (Variable Importance in the Prediction) are presented in the Table 6.

Variables	VIP
Sintg (Small Integral)	1.40
Peak (Maximum value of NDVI)	1.38
Lintg (Large Integral)	1.34
Ampl (Range value of NDVI)	1.31
Rderiv (Right Derivative)	1.01
Los (Length of the season)	0.87

Table 12 : Most important explanatory variables for herb biomass multilinear modelling in Senegal rangelands

Correlation analysis reveals that Sintg, Lintg, Peak and Ampl were greatly correlated between them. To avoid models saturation, these 4 variables were associated at Rderiv and Los separately to finally give eight sets of explanatory variables used for the calibration. Each set thus contained two variables.

Estimation models

After regression only the best model between the eight is maintained for each site (table 7 and 8). Criteria used are adjusted R^2 and RMSE. R^2 aj is generally less than 0.50 for all of models selected on original and resampled sets. But the resample with Bootstrap method allows improving a bit R^2 aj and RMSE for models at a given site.

Site	Variables	Adjusted R^2	R^2	RMSE	N
c3l8	Peak. Los	-0.05	0.16	615.31	11
c4l8	Peak. Los	-0.00057	0.20	490.60	11
c2l8	Lintg. Rderiv	0.01	0.17	1172.52	13
c2l6	Sintg. Rderiv	0.09	0.26	271.35	12
c2l5	Sintg. Rderiv	0.10	0.28	301.05	11
c2l3	Sintg. Rderiv	0.18	0.34	195.72	11
c2l4	ampl. Los	0.19	0.34	177.97	12
c3l5	ampl. Los	0.22	0.35	204.69	13
c2l7	Sintg. Rderiv	0.27	0.39	944.88	13
c5l1	Sintg. Rderiv	0.30	0.41	430.57	13
c2l2	Peak. Los	0.32	0.44	252.79	12
c4l7	Lintg. Rderiv	0.33	0.45	753.04	12
c4l5	Lintg. Rderiv	0.45	0.57	218.73	10
c3l1	Sintg. Rderiv	0.46	0.56	304.27	12
c2l1	Peak. Los	0.46	0.56	241.23	12
c4l1	ampl. Los	0.51	0.59	452.25	13
c3l2	ampl. Los	0.53	0.62	322.01	11
c23l2	Peak. Los	0.55	0.64	256.37	11
c1l5	Lintg. Rderiv	0.67	0.73	209.30	12
c3l4	Peak. Los	0.82	0.86	117.78	9

Table 13 : Original set : selected models and performance criteria

Site	Variables	Adjusted R^2	R^2	RMSE	N
c3l8	Peak. Los	0.17	0.17	525.52	4500
c2l8	Lintg. Rderiv	0.17	0.17	1019.30	5000
c4l8	Peak. Los	0.21	0.21	417.23	4500
c2l6	Sintg. Rderiv	0.25	0.25	234.69	4500
c2l5	Sintg. Rderiv	0.27	0.27	255.48	4500
c2l3	Sintg. Rderiv	0.34	0.34	167.60	4500
c2l4	ampl. Los	0.34	0.34	155.34	4500
c2l7	Sintg. Rderiv	0.38	0.38	834.41	5000
c3l5	ampl. Los	0.38	0.38	174.59	5000
c5l1	Sintg. Rderiv	0.41	0.41	376.52	5000
c2l2	Peak. Los	0.45	0.45	220.18	4500
c4l7	Lintg. Rderiv	0.45	0.45	659.13	4500
c2l1	Peak. Los	0.56	0.56	208.91	4500
c3l1	Sintg. Rderiv	0.56	0.56	263.90	4500
c4l5	Lintg. Rderiv	0.57	0.57	183.11	4000
c4l1	ampl. Los	0.59	0.59	388.40	5000

c3l2	ampl. Los	0.62	0.62	274.69	4500
c23l2	Peak. Los	0.64	0.64	218.76	4500
c1l5	Lintg. Rderiv	0.72	0.72	183.36	4500
c3l4	Peak. Los	0.86	0.86	96.68	3500

Table 14 : Resampled set : selected models and performance criteria

As mentioned in table 9, results are quite interesting, since only three sites (C3L5, C3L1 and C4L1; among twenty) show RRMSE values greater than 30% that has been considered as the maximum acceptable level. These results give hope for future of this research which will consist to develop models for eco-geographical and/or climatic zones.

Sites	MAE	STDD	RMSE	RRMSE	n
C4L5	110.85	84.67	139.48	7.95	4000
C3L4	57.77	49.50	76.07	13.20	3500
C1L5	117.56	88.23	146.97	13.32	4500
C2L4	100.51	77.04	126.63	14.02	4500
C23L2	138.48	100.03	170.82	15.28	4500
C2L6	152.80	120.99	194.89	15.88	4500
C2L3	103.72	86.12	134.81	16.88	4500
C4L8	247.47	250.63	352.20	19.06	4500
C3L8	323.49	287.10	432.49	21.11	4500
C2L2	132.49	110.04	172.22	22.03	4500
C2L5	168.10	127.33	210.87	22.49	4500
C4L7	424.13	361.71	557.40	23.13	4500
C2L1	130.00	109.56	170.00	23.77	4500
C3L2	155.50	153.43	218.44	23.95	4500
C5L1	252.35	203.40	324.10	25.02	5000
C2L8	661.51	538.54	852.97	28.42	5000
C2L7	532.74	459.32	703.38	29.60	5000
C3L5	114.08	102.58	153.42	30.06	5000
C3L1	174.32	129.86	217.37	32.63	4500
C4L1	256.06	218.27	336.46	32.72	5000

Table 15 : Mean Absoulte Error (MAE, kg/ha), standard deviation (STDD, kg/ha) and Relative Root Mean Squared Error (RRMSE, %) of best models obtained by selection on bootstrap sample (n=number of observation)

5.1.4 Conclusions

a) Simple regression models

1. Exponential and Power models correspond statistically to the most plausible models among the six studied models either for herb biomass or total biomass estimation and any variables.

2. Exponential and Power models are also the most consistent compared especially to the Linear model which can gives negative values of biomass when NDVI values are less than 0.3.

3. Leaf and herb biomass, must be considered separately during the development of simple regression models in this approach with data from the same year.

4. Accuracy is globally low linked to the approach and may be it will improved with multitemporal and/or multilinear models using time-series of remote sensing and ground biomass data.

Part of these results were presented in the 27th conference of the International Association of climatology in Dijon (France) between 02 and 05 July 2014, in oral format with an extended summary published in seminar acts (see RD1) and in a poster (see RD2).

b) Multiple regression models at site scale

1. The more important variables (phenological metrics) to predict herb biomass with multilinear regression models are Sintg (Small Integer), Peak (Maximum value of NDVI), Lintg (Large Integer), Ampl (Range value of NDVI), Rderiv (Right Derivative) and Los (Length of season). But these could not be used in the same set of selection because of partial correlation between the first four variables.

2. The Bootstrap method can be used to avoid the “over-fitting” and improve the models accuracy at site scale.

3. This multilinear regression approach could be used to develop more accurate models at the zonal scale (eco-geographical and/or climatic).Moreover, for these needs, zonal classification of the study area was done from the phenological metrics time-series. Also characterization of the sites by statistical analysis based on some variables of primary production and species composition was conducted as part of a Master dissertation study. This study has allowed firstly, to participate in the supervision of a Master student (that will be recorded as part of the doctoral training of Abdoul Aziz Diouf) helped by Dr Jacques André Ndione and secondly, to identify groups of sites and driving parameters as rainfall, soil type and topography. This work was successfully presented on Thursday, March 13, 2014 at the University Gaston Berger of Saint-Louis (North Senegal) by the student Adiouma Fall, after 6 months of internship in CSE (to September 1, 2013 at February 29, 2014).

5.2 Niger use case

5.2.1 Research objectives and context

The overall objective of the PhD research is to improve the fodder forecasting methods using remote sensing data by using two methods, parametric and non-parametric. The approaches developed in Senegal (Abdoul Aziz Diouf) and Niger (Issa Garba) are similar. Specific objectives within this research are:

- To improve the currently-used approaches of fodder production estimation.
- To test the efficiency of other types of vegetation indices.
- To test the metrics derived from time series of remote sensing images of vegetation indices as input parameter of the model, in order to determine the most efficient.
- To test the multiple linear regression approach by using metrics derived from the vegetation indices, weather parameters and in situ measured data.
- To test the similarity approach by using both vegetation and agrometeorological indices.

The thesis committee is listed here below:

- Pr Bernard TYCHON (ULg, Supervisor),
- Dr Bakary DJABY (ULg, Co-supervisor),
- Dr Ibra TOURE (Member of the committee).

A part of these results was also presented at a scientific and technical research seminar on livestock and food security at the university Abdou Moumouni in Niamey, Niger, in September 2009 (see RD4 and RD5).

5.2.2 Materials and methods

a. Data

Various data sources were used in this work

- Biophysical data, i.e., time-series of SPOT VEGETATION images, that can provide explanatory variables
- Meteorological and agriculture data: precipitation, evapotranspiration and natural vegetation phenology. There were used as input variables in the software AMS (see hereafter)
- Annual sum of RFE-FEWSNET
- Sum of the NDVI as computed by the MEIA
- In-situ measured biomass that is the dependent variables.

Tools

In this work, the following softwares were used:

1. For remote sensing data treatment and analysis (VGTExtract, WINDISP, VAST)
2. For computing agrometeorological variables

3. For statistical analysis (SAS-JMP, SPSS)

These softwares are explained in the following:

VGTExtract is a free software than can be used in batch mode. It was developed by VITO for the uncompression and the extraction according to a spatial frame of SPOT images and their conversion into an appropriate format (ILWIS, ENVI, RST, GeoTiff, RAW and WINDISP). The software and its documentation can be downloaded without identification on http://www.vgt4africa.org/

VAST (Vegetation Analysis in Space and Time) was used in our work for extracting biophysical parameters derived form time series of NDVI images (from 1998 to 2012). VAST was developed in the 1990's by M Felix Lee that was a technical assistant of FewsNet in Chad, for the systematic analysis of series of NDVI images. The software analyzes the annual time series of NDVI images in order to derive the following phonological parameters (Figure 6):

- PEAK : the decade at which the NDVI is maximal ;
- SDAT : the decade of the start of the agricultural season ;
- HORZ = PEAK – SDAT ;
- SVAL : the value of NDVI at SDAT ;
- PVAL : the value of NDVI et PEAK ;
- VERT = PVAL – SVAL ;
- EVAL : the value of NDVI at the time PEAK + 4 (namely, approximately the end of the season) ;
- DROP = PVAL – EVAL ;
- SLOP : The slope of the line going from (SDAT, SVAL) to (PEAK, PVAL) ;
- CUMM : The sum of the values of NDVI from SDAT to PEAK ;
- SKEW : the ratio between the three values of NDVI following PEAK (from PEAK + 1 to PEAK + 3) and the sum of the 7 values of PEAK from PEAK — 3 to PEAK + 3.

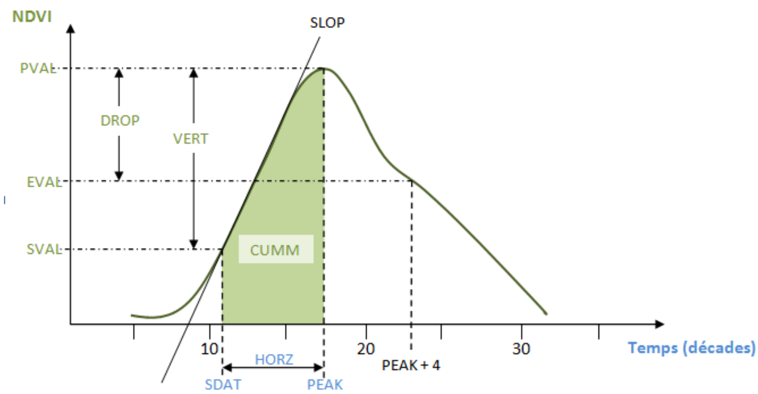


Figure 6: Diagram for the determination of variables from VAST

WINDISP is a free software developed by the FAO for the Global Information and Early Warning System. It has some modules that allow to display and analyses images, shapefiles and databases; to make graphs of the evolution of time series of images (NDVI, precipitation); to superimpose images and administrative maps in order to extract statistics; and to compute statistics for each pixel of a time series of images. It is most renowned for its great visualization capacity and NDVI images processing, precipitations estimated by satellite and outputs from the VAST software. The software and its documentation are available and downloadable at http://www.fao.org/giews/english/windisp

AGROMETSHELL 1.57 (AMS) is a tool developed by the FAO (2007) for the crop monitoring and yield forecasting. It can simulate the water balance and the risk of production shortage. In this work, AMS was used to compute the agrometeorological parameters that will be the input variables for the forecasting model of fodder production. AMS hold a database that can be regularly updated for ensuring an up-to-date production of variables. Modules of the software, which are based on the computation of the water balance, can be used to analyze the impact of climatic factor on the different crops. AMS is a model based on the assumption that yields can be explained by agrometeorological conditions. AMS is an important early warning tool for food security, as it allows to evaluate the effect of climatic conditions on crops and to forecast the agricultural yields through statistical modeling. The model and its documentation can be downloaded at http://www.hoefsloot.com/agrometshell.htm.

b. Methodology

Minimal vegetation detection

To determine the threshold of minimal vegetation, we relied on bibliographic studies and on the NDVI profiles of each site. According to H. Pierre, 1985 ; AGRHYMET, 2000, the real value of NDVI giving the minimal threshold for vegetation detection is 0.1. The formula for the NDVI is:

$$NDVI_{SPOT\ VGT} = (DN*0.004) – 0.1$$

where DN is the digital number. Series of images of NDVI from SPOT VEGETATION from the years 2000 to 2012 were used to extract phenological parameters of the vegetation at the sites where biomass was measured in situ. June was chosen as the minimal date for the beginning of the vegetation season, while October was chosen for the maximal date for the end of the vegetation season, the value 110 as the minimal value of the presence of vegetation and 5 as the minimal variation between two decades.

The same processing was made for the images eMODIS. The equation for retrieving NDVI is

$$NDVI_{eMODIS} = (DN*0.01) – 1$$

If the real value of NDVI representing the minimal threshold for vegetation detection is 0.1, then the value of the DN is equal to 1.1/0.01 = 50. The value 50 as the minimal value of the presence of vegetation and 5 as the minimal variation between two decades

Planting dekades

There are 2 options available in the AMS software in order to determine the planting dekades: the first is based on a threshold of efficient rainfall (to be fixed) followed by more rainfall in the next two dekades, the second is based on a fraction of the total water need of the vegetation. In this work, since we deal with natural vegetation with several species having different germination rates, we chose the second option, by fixing a fraction of 10% of the total water need.

Vegetation length

There is little literature about the phenology of natural fodder species in Sahel. Studies conducted by B. Traoré (1978) on the phonological cycle of some graminaes and legumes showed that the vegetation length vary according to species and water conditions of the spot. We took an average value of 7 dekades knowing that some species can complete their phonological cycle in only 5 dekades while others can need more than 8 dekades (F.W.T. Penning de Vries et M.A. Djitèye, 1999).

Statistical analysis

In this work, we made a regression between the in situ measured fodder yields (dependent variable) and biophysical variables (independent variables) derived from NDVI images and from the AMS model outputs. The specific objectives of this work are to (1) explain the fodder production as a function of the independent variables and (2) to predict the values of fodder production from new values of the independent variables. Data quality was checked before the statistical analysis. Among other tasks, the variances of the independent variables were computed in order to check if these variables cover a sufficient range of values.

Variables selection

To select the independent variables, we used the descendant step-by-step method, by progressively eliminating variables. 20 variables were initially available. The method was used with the software SAS.

Model selection

To select the best models, we used the R^2 and RMSE criteria, and a small number of variables (4). Since the number of variables was small, all possibilities of model combinations were tested.

Cross validation

The models were validated using the Leave-One-Out Cross Validation (LOOCV) method. In this study, the LOOCV was applied to determine the model with the best RMSE and a number of variables below or equal to 4. This validation was made for the global dataset but also by facies, zone, LULC and years.

Residuals analysis

The model hypothesis, i.e., normality, homoscedasticity and the absence of cross-linearities were checked, using the tools available in the software SAS.

Forecasting

The best model that is chosen at the end of the process can then be used for predicting new response values (i.e., biomass fodder production) from measured independent variables.

5.2.3 Results and discussion

a. Exploratory analysis

From the exploratory analysis, the samples of the measured biomass data has an average of about 700 kg DM/ha with a standard deviation of 531 kg DM/ha. Results from a bootstrap analysis on the year 2000 show that the average evolves from 642 to 762 kg DM/ha in the 95% confidence interval, with bias on the average and the standard deviation being, respectively, 0.96 and -0.85 (Table 16). There are 26 independent variables from the outputs of AMS and VAST, the seasonal sum of precipitations and the two variables used by MEIA (INT and MAX).

Terms			Statistic	Standard error	Bootstrap			
					Bias	Standard error	95 % confidence interval	
							Lower	Upper
Breel (Real biomass [kg DM/ha])	Average		699.11	30.12	0.96	30.20	642.94	762.26
	95 % confidence interval for the average	Lower bound	639.85					
		Upper bound	758.37					
	Variance		282097		-165.40	28872	228150	343006
	Standard deviation		531.13		-0.85	27.18	477.65	585.67

Table 16: Exploratory and bootstrap analysis

Global adjustment for fodder biomass production

The variables chosen from the variables selection step are the following: MAX, DRO, EVA, HOR, PEA, PVA, and SLO. All possible models using these 8 variables were tested, namely, 2^8 (256) models, which were automatically classified by decreasing RMSE and according to the number of variables. Results (Table 17) show the four best models among the four first clusters of independent variables according to the RMSE criterion.

Models	N°	Cal. R^2	Cal. adjusted R^2	Val. R^2	RMSE cal. Kg DM/ha	RMSE val. Kg DM/ha	Dif RMSE
Y= -603.1347+4590.8119 MAX	1	0.57	0.57	0.57	327.47	328.42	1.39
Y= -1193.012 + 2822.3077 MAX + 15.510667 DRO	2	0.62	0.62	0.61	308.44	310.65	2.21
Y=-388.0145+3133.0981MAX - 15.41595 DRO+17.625787VER	3	0.66	0.66	0.65	294.16	297.26	3.10
Y= -2190.82+3344.13 MAX -20.46 DRO + 74.06PEA +20.78 VER	4	0.69	0.68	0.67	285.22	288.94	3.72

Table 17: Four best global models

The best model according to the RMSE is the N°4, with a relative RMSE of 40%. The four variables that were selected for that model are MAX, DRO, PEA, and VER. The significance of the coefficients of the model gave highly significant results (Table 18). Figure 7 show the observed values as a function of predicted values for the global adjustment.

Variables	Estimation	Error standard	t-ratio	Prob. > |t|

Constant	-2190.82	407.51	-5.38	<. 0001
max	3344.13	513.06	6.52	<. 0001
DRO	-20.46	2.17	-9.44	<. 0001
PEA	74.06	16.51	4.49	<. 0001
VER	20.78	2.80	7.44	<. 0001

Table 18: Estimation of the coefficients of the global model with 4 variables

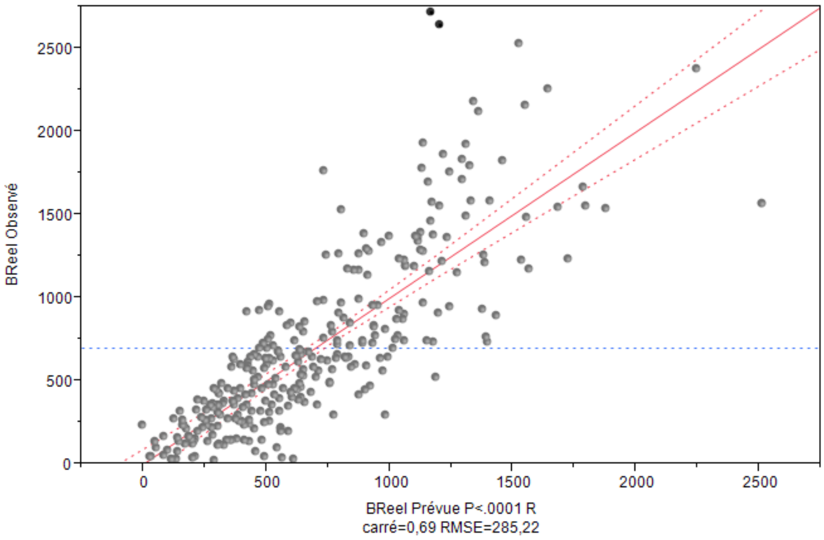


Figure 7: Observed values as a function of predicted values of biomass [kg DM/ha] for the global adjustment.

The residual analysis showed that the average of the residuals is 0 kg DM/ha with a standard deviation of 288,33 kg DM / ha. Bootstrap analysis based on 2000 showed that this average varies from 642 to 762 kg DM / ha within the 95% confidence interval. Bias on the average and the standard deviation being, respectively, 0.89 and -0.35 (Table 19)

Parameters		Statistics	Bootstrap			
			Bias	Standard error	95 % confidence interval	
					Lower	Upper
N	Valid	304	0	0	304	304
	Missing	0	0	0	0	0
Average		0.00	0.89	16.02	-32.49	33.16
Standard deviation		283.33	-0.35	14.86	255.30	313.02
Variance		80273.71	24.69	8440.26	65176.17	97979.15

Table 19: Exploratory and bootstrap analysis of residuals

The analysis of the residuals shows a regular distribution of the residuals (Figure 8). The distribution is Gaussian, with a Durbib-Watson index of 1.83 and a percentage of autocorrelation of 8%.

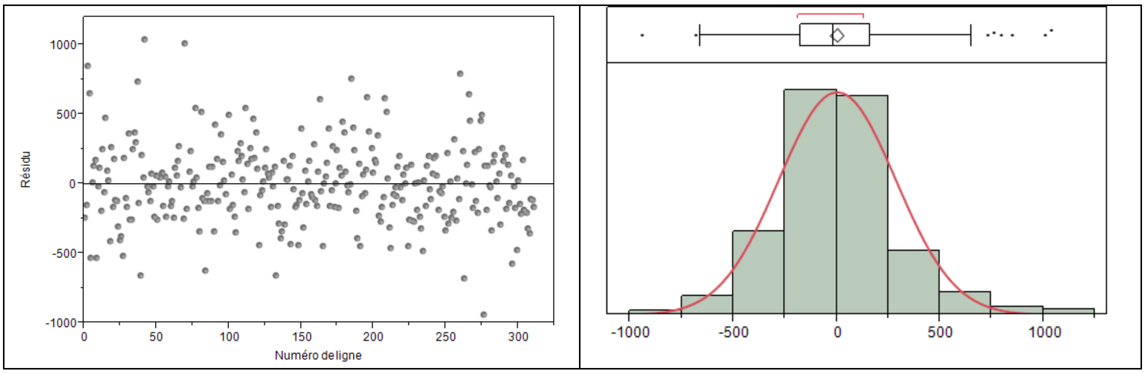


Figure 8: Residuals distribution, per lines (left) and histogram (right)

Analysis across the ecoregions

According to the number of sampling sites and their spreading across the ecoregions, it appears that only the ecoregions “azawak”, “maga1” and “maga2” have enough observations to realize adjustments of the fodder production as a function of the independents variables (Table 20).

ecoregions	ADM1	Air	AZ	BD	GR	LAC	LIP	MA1	MA2	TEN	VD
number	7	10	152	2	13	1	10	58	48	2	5

Table 20: Number of sampling sites by ecoregions

Azawak

Table 21 show that Azawak count 152 sites, with an average production of 684,69 kg DM / ha. The variable selection resulted in the following: VER, EVA, DRO, MAX, INDXNORMAL (Figure 9). The best model for the Azawak ecoregion according to the RMSE has 4 variables, that are MAX, DRO, PEA and VER. Parameter estimation show highly significant probabilities (Table 22). The scatterplot of the observed values as a function of the predicted values show a R^2 of 0.76 and a RMSE of 250 kg DM /ha (Figure 9). It has a relative RMSE of 36%.

Models	Cal. R^2	Cal. adjusted R^2	Val. R^2	RMSE cal KG MS/ha	RMSE val KG MS/ha	Dif RMSE
Y= -242.92 +22.84 VER	0.65	0.65	0.64	300.94	304.03	3.17
Y= -1122.41 -18.11 EVA +13.80 VER	0.72	0.72	0.71	270.71	274.96	3.91
Y= -997.40 -6.05 INDXNORMAL + 17.88 EVA +16.01 VER	0.74	0.73	0.72	262.23	267.76	4.6
Y = -2800.78 + 2979.32 max -23.84 DRO + 98.49 PEA +27.14 VER	0.77	0.76	0.75	250.50	255.25	5.21

Table 21: Four best models - Azawak

Term	Estimation	Error standard	t-ratio	Prob. > |t|
Constant	-2800.78	555.67	-5.04	<. 0001*
max	2979.32	714.62	4.17	<. 0001*
DRO	-23.84	2.9078	-8.20	<. 0001*
PEA	98.49	22.56	4.36	<. 0001*
VER	27.14	3.59	7.56	<. 0001*

Table 22: estimation of the model coefficients, Azawak

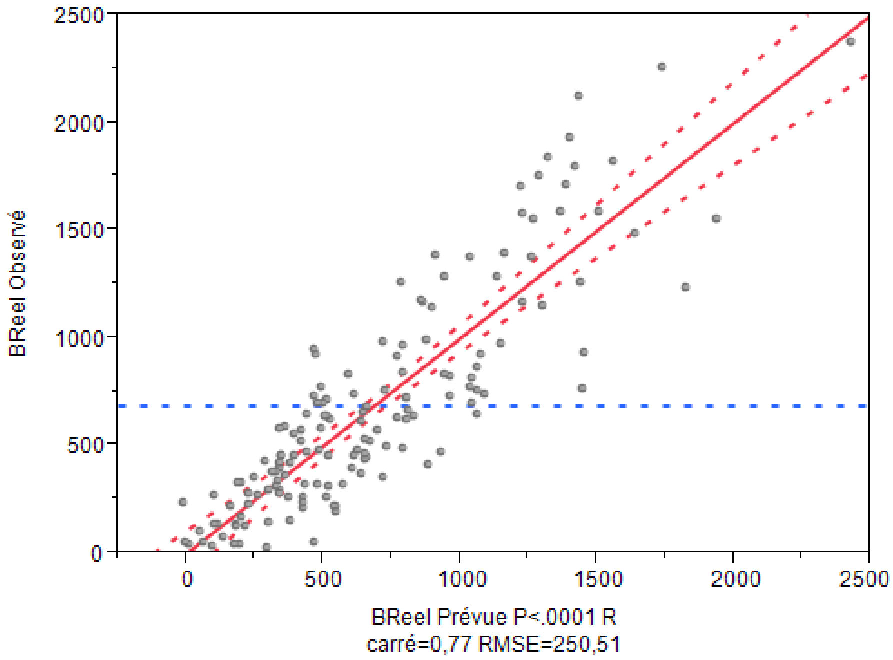


Figure 9: Observed values of biomass production as a function of predicted values, Azawak

Residual analysis in the Azawak ecoregion show an average of 0 kg DM / ha and a standard deviation of 247 kg DM / ha. There is a uniform repartition of residuals per line (Figure 10 (a)) and a Gaussian distribution of the residuals (Figure 10 (b)), with an index DW of 1.59.

Type	Coefficient	Estimation	< 95 %	> 95 %
Position	µ	0	-39.87	39.87
Dispersion	σ	247	221.96	278.76

Table 23: Estimation of coefficients, residuals analysis, Azawak.

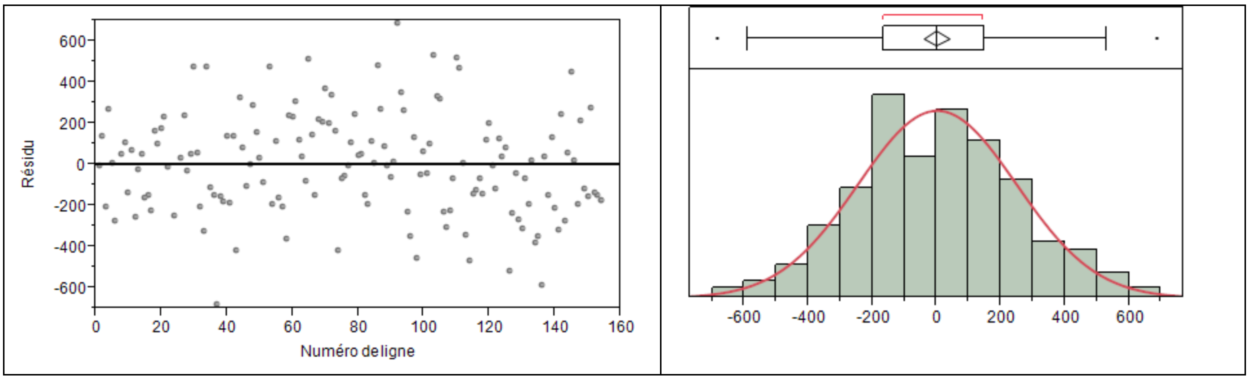


Figure 10: Residuals analysis, Azawak.

Manga2

Table 20 shows that manga2 count 48 sites. The average of the observations in biomass is 827.56 kg DM / ha. The variable selection on this ecoregion leads to the following equation, depending on only two variables that are INT and CUM:

Y= -1300,79 + 485,89 INT + 21,36 CUM

Representation of the observed variables as a function of predicted values showed a R^2 of 0.65 and a RMSE of 288.39 kg DM / ha (Figure 11: Observed values of biomass production as a function of predicted values, Manga2Figure 11). The model is characterized by a relative RMSE of 33% (Table 24). The parameter estimation of the model show highly significant probabilities (Table 25).

Parameters	values
R^2	0.65
Adjusted R^2	0.64
R^2 in validation	0.62
RMSE	33 %
RMSE validation	288.39
Average of the response	827.56
Observations (or weighted sums)	48

Table 24: Statistics of the regression – manga2

Term	Estimation	Error standard	t-ratio	Prob. > |t|
Constant	-1300.793	237.4195	-5.48	<. 0001*
Int	485.89253	96.22733	5.05	<. 0001*
CUM	21.360238	7.700754	2.77	0.0080*

Table 25: Coefficients of the regression – manga2

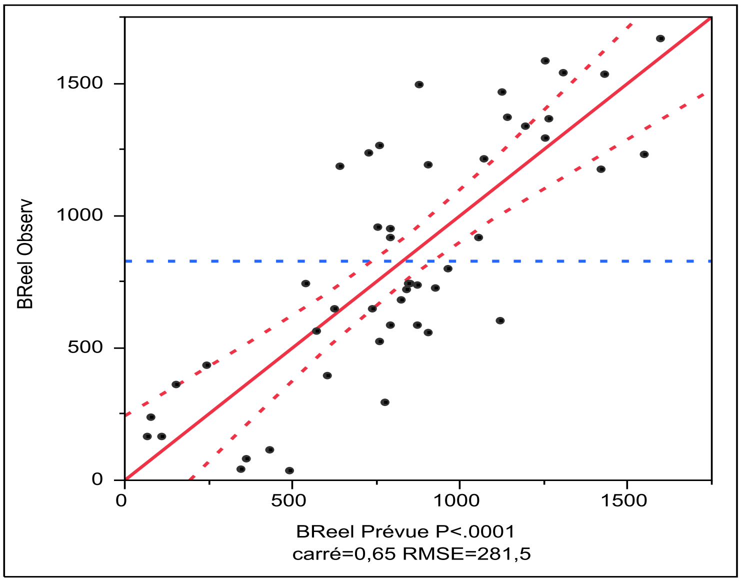


Figure 11: Observed values of biomass production as a function of predicted values, Manga2

Residual analysis in the Manga2 ecoregion showed an average of 0 kg DM / ha and a standard deviation of 275 kg DM / ha. There was a uniform repartition of residuals per line (Figure 12 (a)) and a Gaussian distribution of the residuals (Figure 12 (b)), with an index DW of 1.87 and a autocorrelation rate of 0.09.

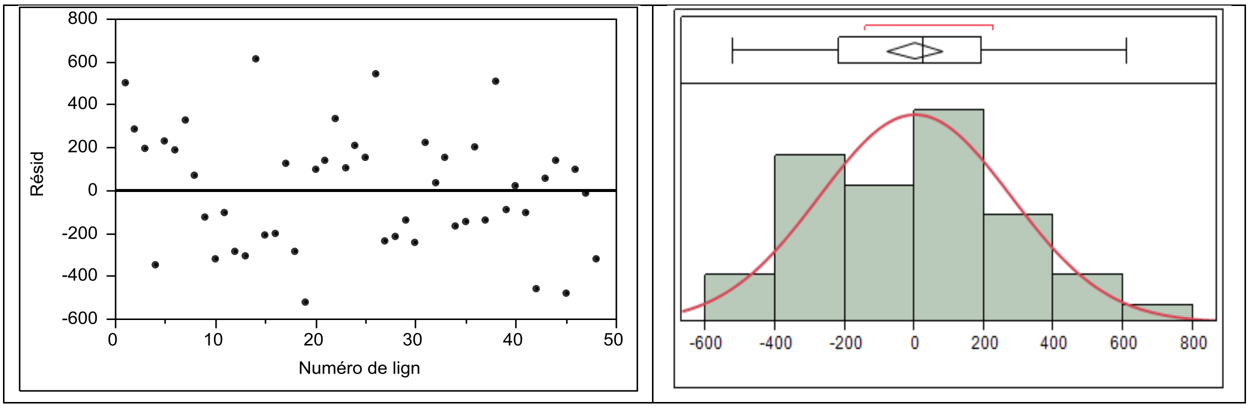


Figure 12: Residuals analysis, Manga2

Analysis of the biomass production by facies

The counting of the number of sites per facies shows that only 8 facies have a number of observations larger than 15, which was considered as the minimal number of observations to conduct the analysis of the biomass according to independent variables (Table 26).

Facies	Number	Facies	Number

	of sites		of sites
ADM1 Ql10-1a Nord sahélienne	7	LIP Ql6 sahélienne	9
Air Y4-1a saharienne	10	MA1 Qc1 Nord sahélienne	6
AZ Ge5-1a Nord sahélienne	22	MA1 Qc1-1a Nord sahélienne	9
AZ Ql10-1a Nord sahélienne	3	MA1 Qc7-1a Nord sahélienne	16
AZ Ql1-1a Nord sahélienne	82	MA1 Qc7-1a saharienne	10
AZ Ql1-1a saharienne	5	MA1 Ql1-1a Nord sahélienne	17
AZ Re35-a Nord sahélienne	21	MA2 Je33-1/3a Nord sahélienne	2
AZ Re35-a saharienne	20	MA2 Qc1 Nord sahélienne	21
BD Ql11-1a Sud sahélienne	2	MA2 Qc7-1a Nord sahélienne	17
GR Be28-1a Nord sahélienne	2	MA2 Qc7-1a sahélienne	4
GR Ql10 Nord sahélienne	4	MA2 Vc15 Nord sahélienne	4
GR Ql6-1a Nord sahélienne	7	TEN Re35-a saharienne	2
LAC Water Nord sahélienne	1	VD Bv7-a sahélienne	5
LIP Be28-1a sahélienne	1		

Table 26: Number of observations sites per facies

The variable selection per facies leads to the following equations (Table 27) with 3 or 4 variables, which are different according to the facies. The R^2 of the models vary from 0.72 to 0.93.

Facies	Models	Cal. R^2	Cal. adjusted R^2	Val. R^2l	RMSE cal. kg DM/ha	RMSE val. kg DM/ha
MA2 Qc7-1a Nord sahélienne	2913+1.02 INDXLATEST +11.26WDEFF - 22.25 ETAF -185.03 HOR	0.93	0.92	0.86	99.99	122.58
MA2 Qc1 Nord sahélienne	-2970.78-25.66 INDXHARVES +43.93 INDEXNORMAL +37.48 ETAI +36.45 EVA	0.82	0.77	0.65	231.45	282.00
(MA1 Ql1-1a Nord sahélienne	-4564.35+3580.39 MAX -10.06 DRO + 165.52 PEA + 2.12 RAI	0.86	0.81	0.74	125.90	141.37
MA1 Qc7-1a Nord sahélienne	5260.48 - 62.87 DRO +25.37 PVA +49.87 SLO + 176.36 SDA	0.78	0.70	0.42	168.31	224.76
AZ Re35-a saharienne	-3433.08 3154.10 MAX +6.15 TWR - 25.90 DRO + 26 .73 VER	0.86	0.82	0.74	121.01	139.60
AZ Re35-a Nord sahélienne	-2304.57 -28.79 DRO +65.18VER 82.91SDA	0.80	0.77	0.72	233.14	253.15
AZ Ql1-1a Nord sahélienne	-3218.28 +31.72 YEAR -22 .13 DRO +119.34 PEA +39.06 VER	0.78	0.77	0.75	269.17	278.66
AZ Ge5-1a Nord sahélienne	-269.31 + 3005.07 MAX – 13 .45 INDXNORMAL +2.39WDEFR +15.83 EVA	0.77	0.72	0.64	279.12	311.51

Table 27: Models per facies

Analysis of the biomass production by years

The regression was tested for each year of the observations, leading to year-specific models. Detailed analysis of these regressions are not presented here but summarized in Table 28.

Year	Models	Cal. R^2	Cal. adjusted R^2	Val. R^2	RMSE cal.	RMSE val.
2001	72.43 -20.58 INDXNORMAL+ 6.40 ETAT +32.11 EVA - 30.99 SVA	0.83	0.79	0.67	25 %	179
2002	-201.30 + 4006.70 MAX - 6.41 DRO -0.55 RAI	0.90	0.88	0.85	21 %	95
2003	-13.62 + 9475.86 MAX - 50.53 WDEFI - 342.48 HOR	0.96	0.94	0.89	0.11%	123
2004	-224.94+ 2357.8581 MAX	0.60	0.58	0.50	44 %	130
2005	-2639.25+ 698.67 INT + 1.91 WDEFI + 106.11 SDA	0.82	0.80	0.77	27 %	205
2006	-858.74 -686.75 INT + 76.18 CUM + 66.51 EVA - 44.45 SVA	0.87	0.85	0.80	29 %	244
2007	-1495.72 + 505.03 INT +2.16TWR	0.61	0.57	0.53	20 %	226
2008	-1125.59 + 14.50 PVA	0.59	0.57	0.53	58 %	292
2009	-1551.76+ 3.61 ETAF +31.13 EVA	0.82	0.80	0.75	22 %	120
2010	-4931.74+ 1057.45 INT +151.27 SDA	0.68	0.66	0.62	34 %	361
2011	1850.33+ 2329.23 MAX -27.77 WDEFI +6.63 ETAF - 89.84 PEA	0.76	0.73	0.65	29 %	155
2012	720.92 -16.19 INDXLATEST 6.95 WDEFR 19.71 PVA	0.78	0.74	0.65	24 %	348

Table 28: Models by year

5.2.4 Conclusions and perspectives

This work focused on the development of new regression-based models for linking in situ measured fodder production to biophysical variables from remote sensing data in Niger. A unique database of biomass field data was collected. Global multilinear model adjustments were tested using different biophysical variables from remote sensing data and outputs from agrometeorological models. A similar approach than the global adjustment was made by ecoregions, which resulted in a slight decrease in the relative RMSE from 40% (global) to 36% (Azawak ecoregion) and 33% (Manga2 ecoregion). Better results were also obtained when decomposing the model per year and per facies.

5.3 Kenya use case

5.3.1 Research objectives and context

a. Objectives

This thesis project focuses on developing a predictive drought related livestock mortality index from Earth Observation data in arid and semi-arid lands. Such an index could be used in various contexts such as famine early warnings, food security programs. The research is based on ILRI's Index Based Livestock Insurance program (IBLI), which was design to protect Kenyan pastoralists' assets against drought related livestock mortality.

The thesis committee is listed here below:

- Pr Bernard TYCHON (ULg, Supervisor),
- Dr Bakary DJABY, (CSE, Co-supervisor),
- Dr Andrew MUDE (ILRI, Member of the thesis committee)
- Dr Isabelle PICARD (VITO, Member of the thesis committee)

The Index-based Livestock insurance project (IBLI)

The Index-Based Livestock Insurance (IBLI) is a livestock insurance program designed by the International Livestock Research Institute (IRLI) in order to cover livestock (cows, camels, sheep and goats) losses due to forage scarcity caused by dry meteorological conditions. The IBLI contract was first sold in Marsabit district, Northern Kenya, in January and February 2010 and the first compensations were paid in October 2011, as the result of the dramatic drought that was experienced in the whole Horn of Africa region in the year 2011.

In the case of index insurance, such as the IBLI project, payments are made when an objectively measured index reaches a pre-defined strike level within a pre-defined area and spatiotemporal coverage (Skees, 2007; Chantarat et al., 2013). The most important advantage of this kind of insurance program is the savings in time and money since no individual verification has to be done on the field as it is the case in traditional insurance programs.

In the IBLI contract, the index that triggers payment is based on freely available NDVI data. In the current contract, livestock mortality is calculated at the end of each coverage season (end of September for the long rain-long dry and end of February for the short rain-short dry seasons respectively) for each of the Marsabit District divisions using NASA's MODerate resolution Imaging Spectrometer (MODIS) NDVI data (Chantarat et al., 2013). After collection, the 16 day composite NDVI is first standardized as regard to the long term average and standard deviation for the given period (z-score). In order to take into account not only the impact of the current but also the past vegetation conditions on livestock mortality, the ZNDVI (standardized NDVI) is cumulated over the previous season (pre-conditions) and over the current year (previous season plus current season). The IBLI mortality is then expressed in the form of a linear regression involving the different cumulative ZNDVI variables (Chantarat et al., 2013).

The research in this PhD project was presented at several conferences: at the African Association of Remote Sensing of the environement (AARSE) conference in El Jadida, Morocco, October 2012 and in the AfricaGIS conference in Addis-Abeba, Ethiopia, November 2013.

5.3.2 Study area and data set

a. Study area

The study area (Figure 13) consists in Marsabit District, located in Northern Kenya. The seven study sites are centred on seven villages and were constructed to represent the area where herds graze (Chantarat et al., 2013). In this IBLI project, mortality linear regressions are computed by geographical cluster (see Chantarat et al., 2013). These clusters are the Chalbi cluster, or Upper Marsabit cluster, which comprises the North Horr and Maikona divisions and the Laisamis cluster, or Lower Marsabit cluster, including the Laisamis, Loiyangalani and Central-Gadamoji divisions.

The District of Marsabit experiences bi-modal rainfall pattern, presented in Figure 14. It is characterized by a long rain and long dry season (LRLD), from March to September and a short rains and short dry season (SRSD), from October to February. In this report, we use the following season naming convention: the name of the season (LRLD or SRSD) will be followed by the year to which the end of the season belongs. For example, as it is represented in Figure 14, the SRSD2001 season starts in October 2000 and ends in February 2001 while the LRLD2001 season starts in March 2001 and ends in September 2001.

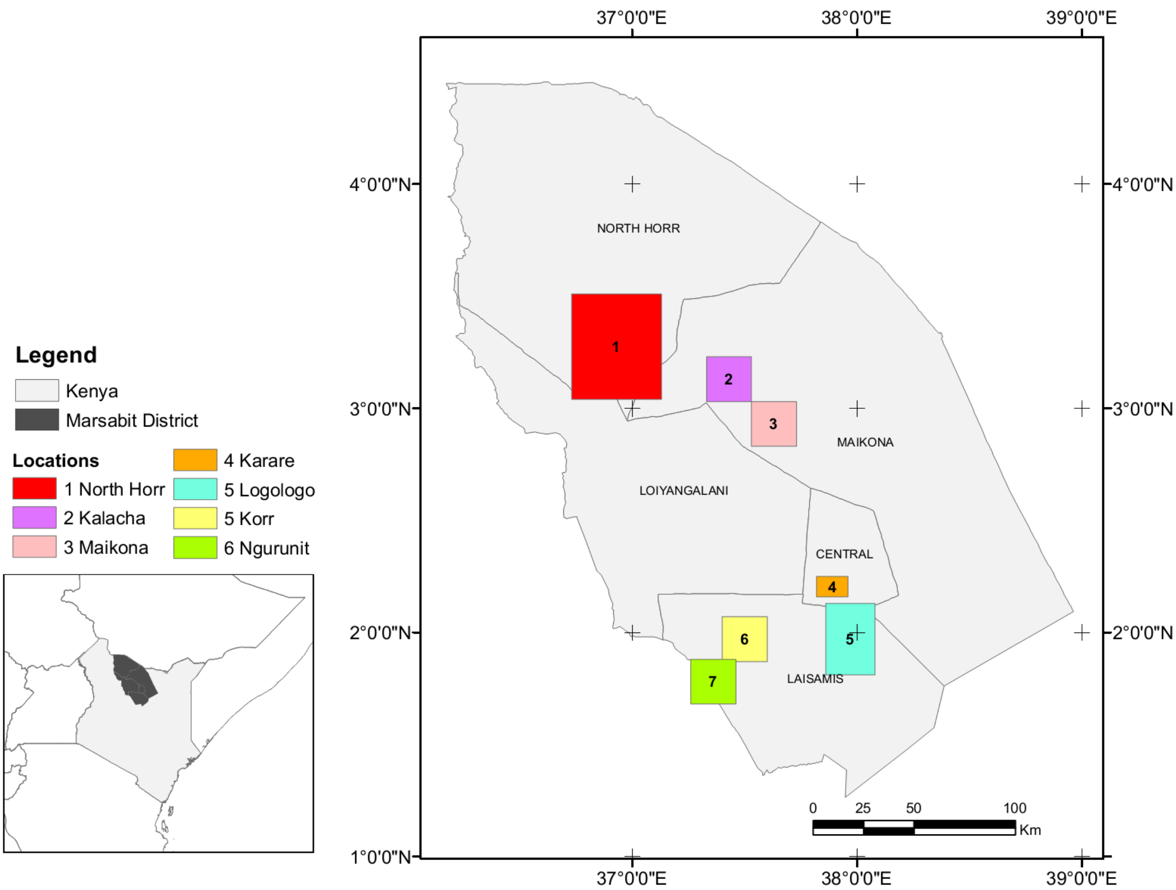


Figure 13: The study area: District of Marsabit, Northern Kenya.

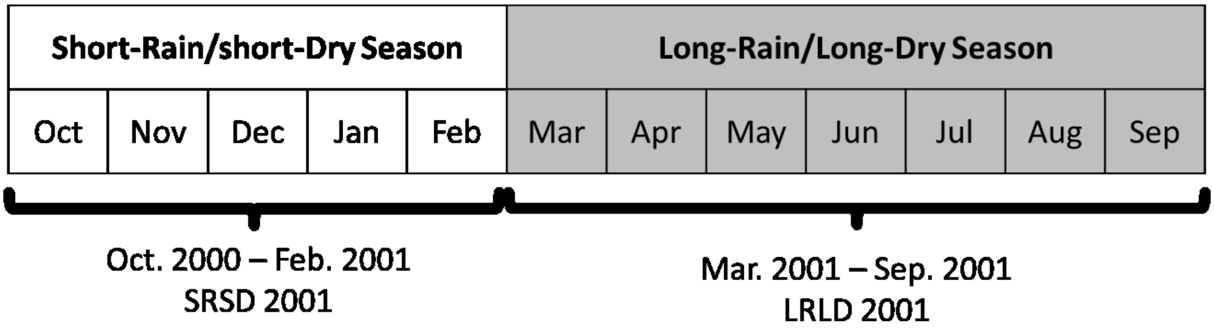


Figure 14 : Temporal characteristics of the study.

Description of the data sets

A summary of the data set is presented in Table 29.

Mortality	Seasonal constructed mortality	• 1999 Long-rains/Long-dry season to 2012 Short-rains/Short-dry season • Per location • Aggregated for the 4 animal species • In % of the herd size
Remote sensing	Raw data	• SPOT Vegetation NDVI, DMP and FAPAR, • dekadal (10-day maximum value composite), • from 1-10 April 1998 to 1-10 May 2012, • 1 km * 1 km spatial resolution
	Constructed vegetation indicators	• CZNDVIpos, CZNDVIpre, CZNDVIp and CZNDVIn • CZDMPpos, CZDMPpre, CZDMPp and CZDMPn • CZFAPARpos, CZFAPARpre, CZFAPARp and CZFAPARn
Other data	Location_ID	Name or ID number of the location (see Figure 13)
	Cluster	Name of cluster (Upper or Lower)
	Year	Year at the end of the season (see Figure 14)
	Season	Season: SRSD(Oct – Feb) LRLD (Mar – Sep) (see Figure 14)
	Veget_condition	Vegetation condition: Good (CZNDVIpos ≥ 2.5) BAD (CZNDVIpos < 2.5)

Table 29: Summary of data set.

- Mortality data

The livestock mortality dataset is constructed from ALMRP (Government of Kenya's Arid Lands Resource Management Program) monthly statistics (number of animals at the beginning of month, mortality, birth, sales and slaughter …) at the location level by animal (sheep, goat, cattle and camel). Seasonal livestock mortality is defined by Chantarat et al., 2013, to represent average livestock mortality by location, in percentage of the total herd size and for the four animals combined using the Tropical Livestock Unit. In this study, we use constructed seasonal livestock mortality for the seven locations, from the 1999 Long-rains/Long-dry season to the 2012 Short-rains/Short-dry season.

- Remote sensing data

In this study, we use SPOT Vegetation Normalized Difference Vegetation Index (NDVI), Dry Matter Productivity (DMP) and Fraction of Absorbed Photosynthetically Active Radiation (FAPAR). The dekadal indicators (10-day maximum value composite) cover the period from 1 to 10 April 1998 to 1 to 10 May 2012 at a 1km x 1km spatial resolution.

The spatialized forage availability is introduced in the model under the form of four variables constructed by accumulation over different periods of time of the standardized NDVI (z-score). The z-scored NDVI (ZNDVI) is given by equation (1) and was introduced to standardize NDVI values over space.

$$(1) \quad ZNDVI_{pdt} = \frac{NDVI_{pdt} - E_{pd}(NDVI_{pdt})}{S_{pd}(NDVI_{pdt})}$$

Where $ZNDV_{Ipdt}$ is the standardized NDVI for pixel p, dekade d and year t; $NDVI_{pdt}$ is the NDVI value for pixel p, period d and year t; $E_{pd}(NDVI_{pdt})$ is the long term mean of NDVI for pixel p and period d and Spd(NDVIpdt) is the long term standard deviation of NDVI for pixel p and period d.

The cumulated variables over the dekades (CZNDVI), presented in , were introduced to take into account the current vegetation condition (CZNDVIpos) but also the effect of the past periods (CZNDVIpre) as well as the current positive (CZNDVIp) and adverse (CZNDVIn) effects of the state of vegetation on livestock mortality (Chantarat et al., 2013). The seasonal constructed vegetation indicators are finally spatially averaged within the limits of the locations defined in Figure 13.

The same variables were calculated for the other two indicators. In order to simplify the writing, when a case applies for the three indicators, we will write the cumulated variables as following: CZ*VAR*pos, CZ*VAR*pre, CZ*VAR*p and CZ*VAR*n.

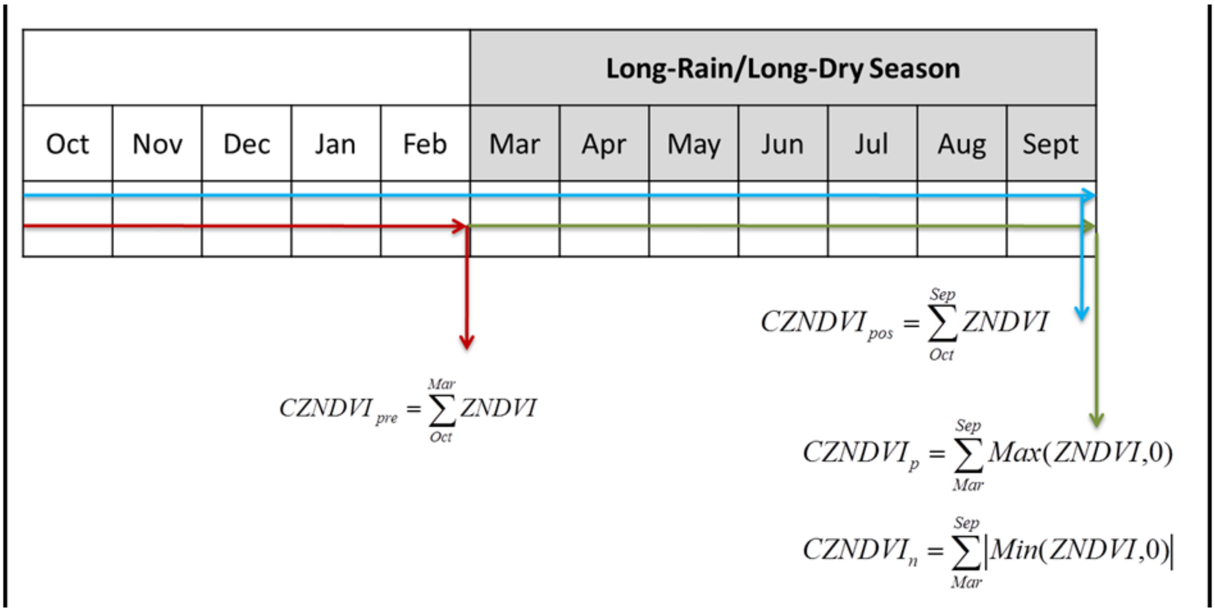


Figure 15: Constructed seasonal vegetation indicators.

5.3.3 Linear regression models using SPOT Vegetation indicators

The first set of results we will present is the linear modelling of livestock mortality from SPOT Vegetation indicators (NDVI, DMP and FAPAR). We tried to reproduce the method described in Chantarat et al., 2013, to design a livestock mortality index from these indicators.

a. Materials and methods

In this study, we used the data set described in section 5.3.2.

The tests that will be described hereafter were performed in the following cases:

- Overall case,
- by season,
- by geographical cluster and
- by vegetation condition.

In this study, we started by analysing the mortality and vegetation indicators distributions in different cases. We tested the normal distribution with associated density function (2) using Shapiro-Wilk test on one hand and Kolmogorov's test for exponential distribution of density function (3) on the other hand.

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

$$(3) \qquad f(x,\lambda) = \begin{cases} \lambda\, e^{-\lambda x} & if\ x \geq 0 \\ 0 & if\ x < 0 \end{cases}$$

Since the IBLI insurance contract differentiates between good and bad vegetation condition (based on the NDVI), we also studied the mortality repartition between the two cases by comparison of means and visually with box plots. During the comparison of means process, we first tested the equality of variances and performed Student's T test if the variance equality test was positive and Welch test if it was not.

We computed the correlations between the mortality and the vegetation constructed variables as well as between the different indicators. The later was studied in order to determine the behaviour of the indicators relative to each other and if one could serve as a back-up in case the principal one fails.

Finally, we computed linear regressions between mortality and the vegetation indicators in different cases (overall case, by vegetation condition, by season, by cluster). We first used mixed stepwise method with P-value thresholds of 0.25 for entering the model and 0.1 to leave for variable selections and computed linear models with the selected ones.

Results

Primary data analysis

Table 30 presents the results of the distribution analysis for seasonal livestock mortality.

Normal distribution

		μ estimate	σ estimate	W	p
Overall mortality		0.1025	0.1411	0.7255	< .0001
Mortality by season	LRLD	0.0945	0.1225	0.7612	< .0001
	SRSD	0.1112	0.1594	0.6974	< .0001
Mortality by cluster	Upper	0.1061	0.1505	0.6960	< .0001
	Lower	0.0998	0.1344	0.7409	< .0001
Mortality by Vegetation Condition	Bad	0.1385	0.1619	0.8114	< .0001
	Good	0.0351	0.0354	0.8515	< .0001
Exponential distribution					
		λ estimate	**D**	**p**	
Overall mortality		0.1025	0.2321	< 0.01	
Mortality by season	LRLD	0.0945	0.2310	< 0.01	
	SRSD	0.1112	0.2610	< 0.01	
Mortality by cluster	Upper	0.1061	0.2500	< 0.01	
	Lower	0.0998	0.2570	< 0.01	
Mortality by Vegetation Condition	Bad	0.1385	0.2150	< 0.01	
	Good	0.0351	0.0816	> 0.15	

Table 30: Distribution analysis for seasonal livestock mortality.

The histogram for overall mortality is displayed in Figure 16.

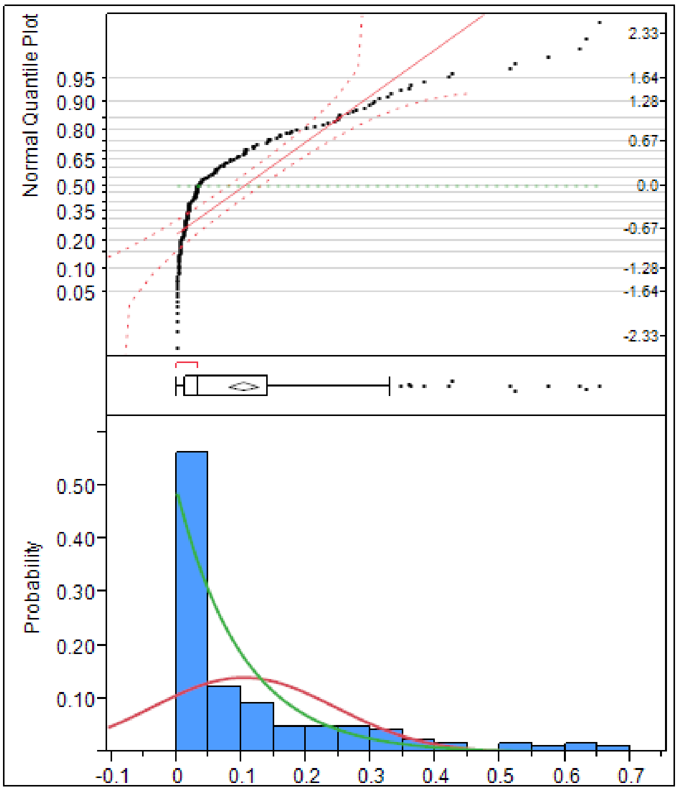


Figure 16: Probability plot (up) and histogram of overall livestock mortality (down), normal distribution (red), exponential distribution (green).

The result of Shapiro-Wilk test for normality confirms that the distribution is not normal ($w = 0.73$ and $p < 0.0001$). According to the result of the Kolmogorov's test, the exponential distribution has to be rejected as well, although the value of p is slightly higher ($D = 0.23$, $p < 0.01$). We then classified the mortality values by season and cluster but we did not observe any improvement in the quality of the fits, as shown by the results in Table 30.

Finally, we considered mortality by vegetation condition. The histogram for bad and good vegetation condition is presented in Figure 17. In bad vegetation conditions, the result of normal and exponential is still negative ($p < 0.0001$ and $p < 0.01$ respectively). However, in the case of good vegetation conditions, we can see that Kolmogorov’s test is significant. With a coefficient D equal to 0.08 and a probability $p > 0.15$, livestock mortality in good vegetation conditions follows an exponential distribution of scale parameter equal to 0.0351.

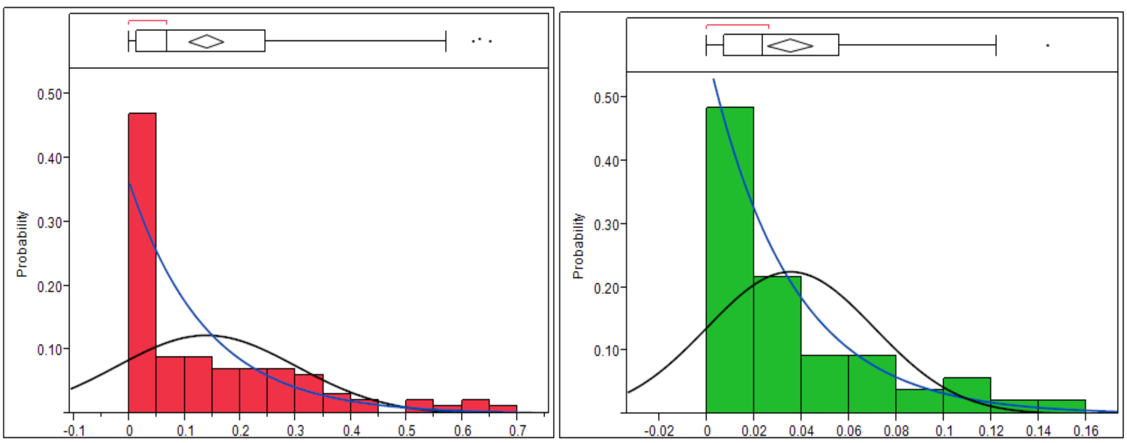


Figure 17: Histogram of livestock mortality for bad (red) and good (green) vegetation conditions, normal distribution (black), exponential distribution (blue)

From the results of the analysis of distributions of livestock mortality, we can already deduce that the use of a multiple linear regression model might not be the best option for modelling livestock mortality from vegetation indicators. According to the future results of linear modelling, we will consider or not alternative methods, such as Generalised Linear Models (GLM).

In this primary analysis of data, we compared mortality means between different categories, namely cluster, geographical location, season and vegetation condition.

The analysis of means shows that there is no statistically significant difference between mean mortalities in any of the first three cases. When classifying constructed mortality according to various parameters, we could observe that there is no significant difference between the samples, except for the vegetation condition. Figure 18 shows, for example, the comparison between mortality samples classified according to season and geographical cluster.

However, regarding clustering according to the condition of the vegetation, we can observe that, in some cases, bad vegetation conditions are associated with low mortality values (see Figure 19). This is also visible on Figure 20, where a regime, characterized by increasing livestock mortality when the vegetation state is improving (increasing CZNDVIpos), can be observed for very low CZNDVIpos values (red circle). We can already assume that this feature will lead to some difficulties when modelling livestock mortality from remote sensing vegetation indicators.

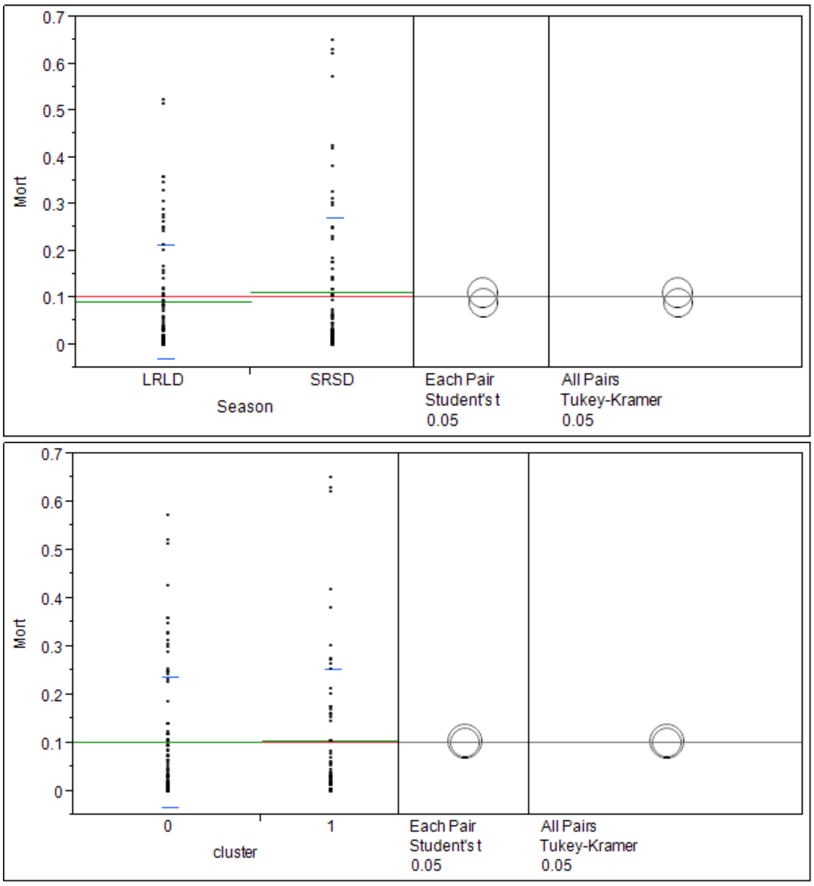


Figure 18: Means comparisons for mortality classified according to the season (top) and the geographical cluster (bottom) (0: Laisamis cluster and 1: Chalbi cluster)

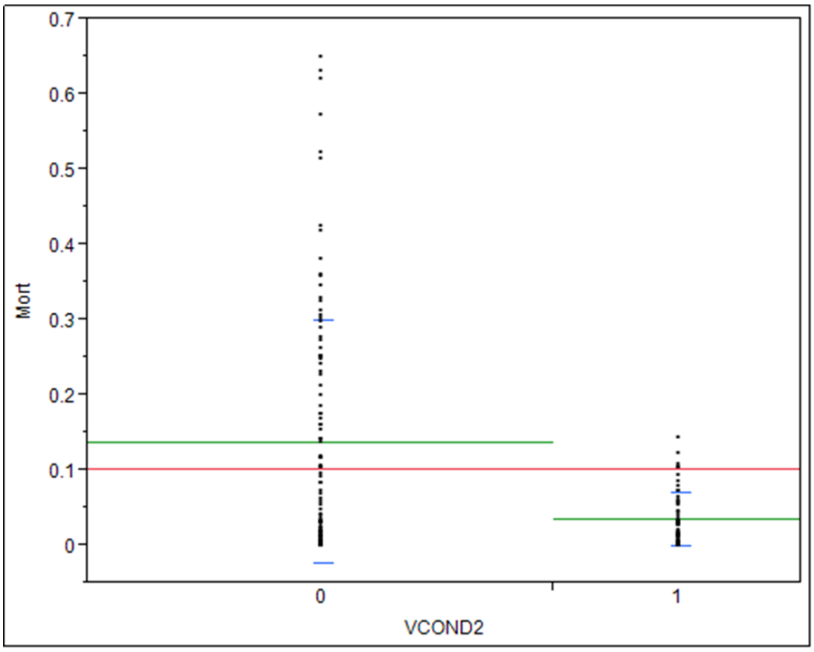


Figure 19: Mean comparison for livestock mortality classified according to the condition of vegetation (0: bad, 1: good. Red line: general mean, green line: sample mean).

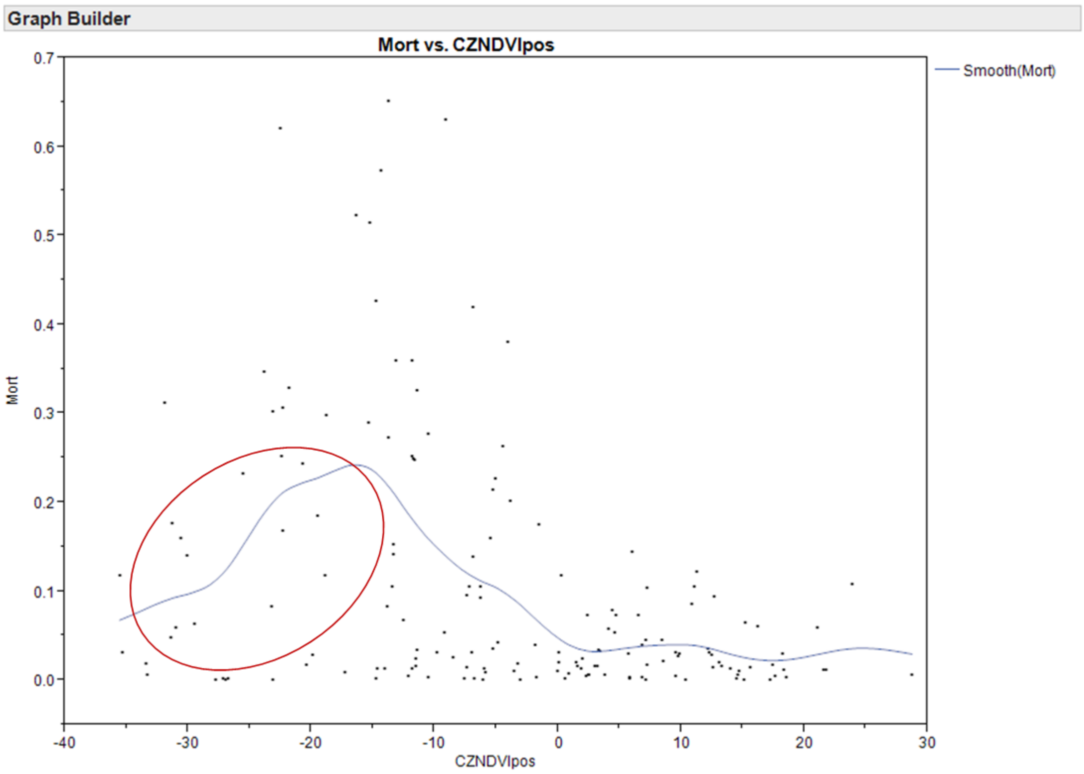


Figure 20: Seasonal livestock mortality versus CZNDVIpos.

Finally, we computed correlations between mortality and the four cumulated variables for the three different indicators in the usual cases.

The coefficients of correlation between mortality and the NDVI-, DMP-, and fAPAR-derived variables are presented in Table 31. Highlighted cells represent the significant coefficients.

In every significant case, and for every variable, we can observe that the mortality is negatively correlated to CZVARpos, CZVARpre and CZVARp and positively correlated with CZVARn, which makes sense given the definition of the different cumulative variables. It is, however, not always the case when the coefficients are not significant.

For the 3 variables, the highest correlations can be observed during the short season in the lower cluster for CZVARpos and CZVARpre (high coefficients can also be found for these variables in the lower cluster with no distinction between the seasons and for the short season or with no distinction between the clusters). For CZVARn, the highest correlations are found in the lower cluster, during the long season and in the lower cluster, when seasons are pooled for NDVI, DMP and fAPAR.

In the case of CZVARp, correlations present different behaviour according to the vegetation indicator. If we consider NDVI, the highest correlations with mortality occur in the upper cluster. For DMP, correlation between livestock mortality and CZDMPp is higher in the lower cluster, during the long rain-long dry season. Finally, concerning fAPAR, we can observe that the correlations between CZfAPARp and mortality are more important in the upper cluster, with no distinction between the seasons.

			CZNDVIpos	CZNDVIpre	CZNDVIp	CZNDVIn	CZDMPpos	CZDMPpre	CZDMPp	CZDMPn	CZfAPARpos	CZfAPARpre	CZfAPARp	CZfAPARn
Overall			-0.37	-0.25	-0.27	0.33	-0.40	-0.28	-0.25	0.32	-0.39	-0.30	-0.23	0.31
Cluster	**Lower**		-0.42	-0.31	-0.22	0.39	-0.46	-0.33	-0.24	0.38	-0.44	-0.34	-0.21	0.37
	Upper		-0.35	-0.19	-0.40	0.29	-0.37	-0.18	-0.33	0.32	-0.40	-0.23	-0.35	0.29
Season	**LRLD**		-0.32	-0.18	-0.30	0.34	-0.37	-0.23	-0.32	0.37	-0.38	-0.25	-0.29	0.39
	SRSD		-0.41	-0.29	-0.23	0.36	-0.43	-0.30	-0.19	0.33	-0.43	-0.32	-0.20	0.30
Vcond	**Bad**		-0.17	-0.08	-0.10	0.16	-0.26	-0.16	-0.12	0.18	-0.25	-0.17	-0.08	0.15
	Good		-0.17	0.05	-0.17	0.23	-0.10	0.08	-0.16	0.22	-0.12	0.05	-0.14	0.28
Cluster + Season	**Lower**	**LRLD**	-0.35	-0.21	-0.29	0.38	-0.39	-0.24	-0.35	0.38	-0.39	-0.27	-0.29	0.39
		SRSD	-0.49	-0.40	-0.13	0.36	-0.54	-0.43	-0.10	0.38	-0.53	-0.43	-0.12	0.37
	Upper	**LRLD**	-0.26	-0.10	-0.36	0.23	-0.31	-0.18	-0.32	0.29	-0.36	-0.21	-0.34	0.32
		SRSD	-0.23	-0.22	-0.04	0.30	-0.38	-0.20	-0.23	0.44	-0.46	-0.23	-0.39	0.46

Table 31: Coefficients of correlation between seasonal livestock mortality and NDVI-, DMP-, and fAPAR-derived variables. Highlighted cells represent the significant coefficients.

In the overall case, we can see that all the coefficients are significant and that the correlations are comprised, in absolute value, between 0.25 and 0.37 for NDVI, 0.24 and 0.40 for DMP and 0.22 and 0.39 for fAPAR. When the mortality data is classified by geographical cluster or season, almost every correlation is significant.

However, when data is classified according to the condition of vegetation, we found that most of the correlations are not significant. This is not a surprise given the patterns observed in Figure 17 where, even if the means are significantly different, there is a superimposition of the histograms (i.e. we can find low values of mortality in the two categories). Since we find similar values of mortality in both classes, it is impossible to determine to which of them belongs a certain value and therefore, to find significant correlations between vegetation indicators and mortality according to the vegetation condition.

We also computed correlations between the cumulated variables of a same indicator. In most cases, we found highly correlated variables (for example, see Table 32 for NDVI). This suggests that adding more than one variable to the model should not improve it. However, since every variable has a different signification, they still contain valuable information and could be exploited in another way.

	CZNDVIpos	CZNDVIpre	CZNDVIp	CZNDVIn
CZNDVIpos	1			
CZNDVIpre	0.7861	1		
CZNDVIp	0.6283	0.0902	1	
CZNDVIn	-0.8209	-0.393	-0.6162	1

Table 32: Coefficients of correlation between NDVI-derived variables.

Finally, in order to determine if one indicator could perform better than the other at modelling livestock mortality, we compared the correlations between mortality and the cumulated variables for the three indicators. Comparison of means show that the samples' means are not significantly different and are extremely close to each other (see Table 33 and Figure 21). This suggests that none of the three indicators is better correlated to mortality than the others and that therefore; all the indicators should the same ability to predict livestock mortality due to drought.

Student's T test			
Indic 1	**Indic 2**	**Difference**	**p-Value**
NDVI	fAPAR	0.015	0.693
NDVI	DMP	0.017	0.652
DMP	fAPAR	0.002	0.955
Tukey-Kramer test			
Indic 1	**Indic 2**	**Difference**	**p-Value**
NDVI	fAPAR	0.015	0.918
NDVI	DMP	0.017	0.894
DMP	fAPAR	0.002	0.998

Table 33: Results of Student's T and Tukey-Kramer tests for mean comparison between NDVI-, DMP- and fAPAR-Mortality correlations

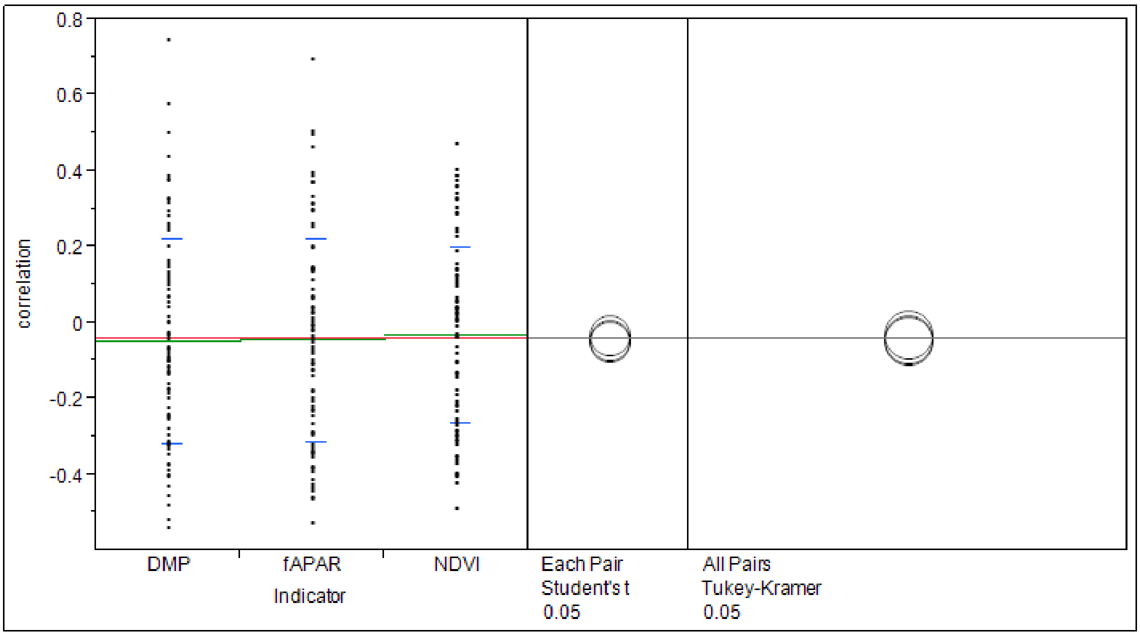


Figure 21: Comparison of means for correlations classified by indicator, results of Student's T and Tukey-Kramer tests (green line = sample mean, blue lines = standard deviation lines and red line = general mean)

Multiple linear regressions

We performed linear regression using the least squares method on NDVI-, DMP- and FAPAR-derived variables and livestock mortality in the different cases that we previously analysed. The results for NDVI are summarized in Table 34.

Case	N	Variable	R^2	Adjusted R^2	RMSE	Sig Prob	Normality of residuals
Overall	161	CZNDVIpos	0.14	0.13	0.13	0	No
Lower	93	CZNDVIpos	0.19	0.18	0.12	0	No
Upper	68	CZNDVIp	0.10	0.08	0.14	0.0094	No
LRLD	83	CZNDVIn	0.13	0.12	0.12	0.0009	No
SRSD	78	CZNDVIpos	0.16	0.15	0.15	0.0003	No
Bad	105	CZNDVIpos	0.03	0.02	0.16	0.088	No
Good	56	-					-
Lower_LRLD	48	CZNDVIn	0.16	0.14	0.13	0.0049	No
Lower_SRSD	45	CZNDVIpos	0.26	0.24	0.11	0.0004	No
Upper_LRLD	35	CZNDVIp	0.14	0.11	0.09	0.0245	No
Upper_SRSD	33	CZNDVIn	0.14	0.11	0.02	0.0311	No

Table 34: Results of linear regression between livestock mortality and NDVI-derived variables.

We first notice that, when data are clustered according to the state of vegetation, CZNDVIpos is a significant parameter for the model in bad vegetation conditions while no parameter is kept in good vegetation conditions.

In all the other cases, the models are significant (Sig Prob is close to 0) but, in every case, we observe very low R^2. As we predicted from the results of the correlation analysis, only one variable is included in the model for each case. Finally, we never observe normally distributed residuals. The coefficients of the significant models are presented in Table 35.

Case	Variable	Intercept	Coefficient
Overall	CZNDVIpos	0.086659	-0.00356
Lower	CZNDVIpos	0.083322	-0.00387
Upper	CZNDVIp	0.144851	-0.01049
LRLD	CZNDVIn	0.040871	0.007227
SRSD	CZNDVIpos	0.093013	-0.00445
Bad2	CZNDVIpos	0.104314	-0.00265

Table 35: Coefficient for significant linear models.

Detailed results for the overall cases are displayed hereafter (Figure 22 and Figure 23).

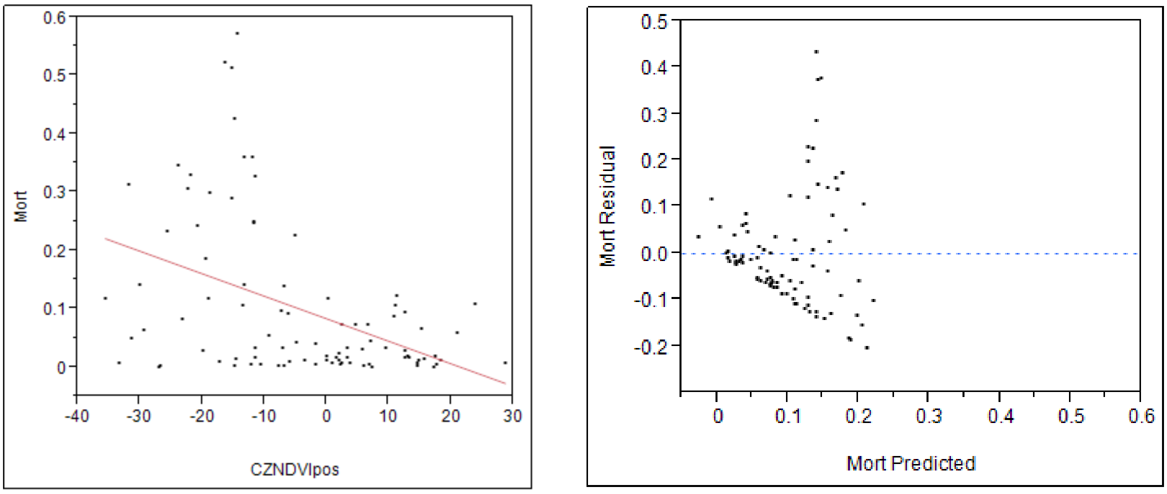
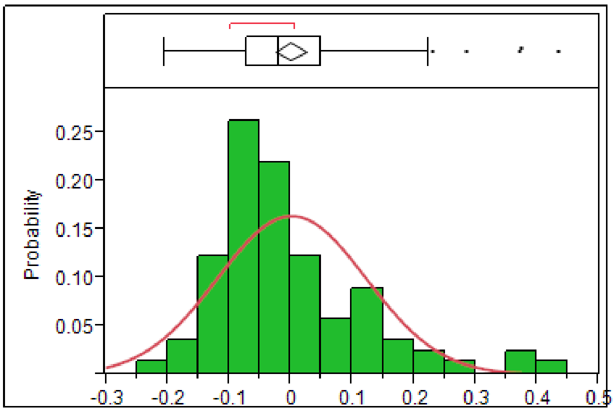


Figure 22: Plot for linear fit (left) and residuals in overall case (right).



[redacted]

Figure 23: Distribution of residuals for the overall case and goodness of normal distribution fit.

In addition, we tested different variables transformations (log, ln, sqrt) but observed no improvement in the linear relations and still a non-normality of residuals.

Conclusions

The study of the distribution of mortality data showed that it does not follow a normal distribution. In some cases (good vegetation conditions), it however shows an exponential distribution. We therefore conclude that the use of simple linear regression models might not give satisfying results. This was confirmed when we tried to implement linear regression models. Indeed we did not observe normally distributed residues. The conclusion is that, in this case, linear regressions are not best suited to model livestock mortality from vegetation indicators.

During the analysis of correlations, we observed that:

- Adding a variable should not improve the model.
- Data clustering does not give better results.
- None of the indicators is better than the others at modelling livestock mortality.

These results were confirmed when we realized linear regression models study, where, although most of the models were significant, none of the combinations gave satisfying R^2.

Finally, since none of the vegetation indicators outperform the others in predicting livestock mortality, one indicator could be used to model livestock mortality while the others could be considered as backups in case the primary indicator would not be available or its quality would become poorer.

Given the non-normality of the residuals, our next step is to compute Generalised linear models (see section 5.3.4).

Regarding the poor R^2 values, this might be linked to the fact that we sometimes find low mortality values associated with very bad conditions. Our hypothesis is that, in very bad vegetation conditions, the grazing area might be extended and that the polygons used to delineate grazing areas do not represent the actual grazing area anymore.

5.3.4 Generalised linear models

The choice of using Generalised linear models (GLM) was made because the dependent variable did not fulfil all the requirements for linear regression models. Indeed, our livestock mortality sample does not follow a normal distribution; neither does it satisfy the homoscedasticity condition. The Generalised linear models were introduced as an extension of the linear models and can be applied to a set of independent observations yi which distribution belongs to the exponential family (including normal, Poisson & Binomial distributions) (Jiang, 2008). In this case, GLM were also used to confirm the results of the ANOVA analysis of mortality.

a. Materials and methods

Generalised linear modelling was performed using seasonal constructed NDVI-derived variables and livestock mortality. The data set characteristics specified in section 5.3.2 apply here.

In Generalised linear modelling, the relation between the expected value of y_i ($\mu_i = E(y_i)$) and the linear predictor can be expressed as follows (Equation 4) (Lynch, 2007):

$$(4) \qquad \mathbf{F}(\boldsymbol{\mu}_\mathbf{i}) = \boldsymbol{\beta}\, \mathbf{X}_\mathbf{i}^\mathbf{T}$$

Where F is called the “link function” and depends on the distribution of the variable y_i.

The models that were designed are presented in more details in Table 36.

Element of the model	Value
Link function (F)	• Identity: F(μ)=μ (SAS Institute Inc., 2008) • Log: F(μ)=log(μ) (SAS Institute Inc., 2008)
Mortality derived variables (μ)	• Mortality (in percent of the total herd size) • Log(Mortality) • Mortality*100
Models	$F(\mu) = \beta_0 + \sum_{i=1}^{n} \beta_i X_i$ with X_i=$CZNDVI_{pos}$, $CZNDVI_{pre}$, $CZNDVI_p$, $CZNDVI_n$, Season, Location, VCOND

Table 36: Description of the models and parameters.

For this part of the study the GENMOD procedure of the SAS software (version 9.1.3) was used. We started by testing the goodness of fit of the different variable distributions with their associated link functions combined to three variable transformations. The parameters taken into account for assessing goodness of fit were the scaled deviance D and Pearson’s Chi-square χ2. Once the best combination was identified, we proceeded with the modelling phase. The significance of the different variables was assessed using the Chi-square test and its associated probability $p(\chi^2 \leq \chi_c^2)$, with low values of probabilities implying significant chi-square).

Concerning mortality derived variables, the *log(mortality)* function was introduced because the GENMOD procedure did not include the exponential distribution. We therefore transformed the variable and considered a normal distribution. On the other hand, the Mortality*100 variable was introduced to get rid of the small numbers that appears when the mortality, expressed in %, is low.

Results

Analysis of the goodness of fit, using the maximum likelihood method, showed that all the combinations of distributions and link functions fit the mortality distribution quite well. With values of the scaled deviance varying between 1.088 and 1.092, we could not determine the best case and computed Generalised linear models for each of them.

The best results regarding the Generalised linear models for the continuous vegetation derived variables where obtained when we used a logarithm link function:

- Model 1: variable = mortality, link function = Log
- Model 2: variable = mortality*100, link function = Log

The chi-square χ_c^2 and associated probability $p(\chi^2 \leq \chi_c^2)$ for those models are presented in Table 37.

Parameter	Value	$\chi_c 2$		p	
		Mod1	Mod2	Mod1	Mod2
Intercept		24.78	5.78	***<.0001***	0.016
CZNDVIpos		4.75	3.46	***0.029***	***0.063***
CZNDVIpre		4.61	3.34	***0.032***	***0.068***
CZNDVIp		3.95	2.64	***0.047***	0.104
CZNDVIn		4.47	3.22	***0.035***	***0.073***
Season	LRLD	3.07	2.04	***0.080***	0.154
LOC	1	1.17	0.9	0.654	0.342
LOC	2	2.33	2.75	0.359	0.097
LOC	3	0.26	0.2	0.855	0.658
LOC	4	4.01	3.58	0.140	0.058
LOC	5	0.22	0.09	0.517	0.764
LOC	6	0.96	0.85	0.280	0.356
VCOND	Good	4.31	3.68	***0.038***	***0.055***

Table 37: Results of Generalized linear models.

The chi-squares and probability values show that, after the intercept term, the NDVI derived variables and the vegetation condition representing the state of vegetation, are the most significant parameters. Models 1 and 2 are the only ones where the 10% significance level is achieved. Within those models, that threshold is reached for the variables describing forage availability but not the geographic and seasonal variables, which suggest that the state of vegetation is the factor that impact livestock mortality the most.

The β_i coefficients and their associated standard error for the two models are presented in Table 38.

		Model 1		Model 2	
Parameter		**Estimate**	**Standard Error**	**Estimate**	**Standard Error**
Intercept		-3.22	0.65	1.51	0.63
CZNDVIpos		-0.50	0.23	-0.42	0.23
CZNDVIpre		0.46	0.22	0.39	0.21
CZNDVIp		0.45	0.23	0.37	0.22
CZNDVIn		-0.49	0.23	-0.42	0.23
Season	LRLD	-0.41	0.23	-0.32	0.22
Season	SRSD	0.00	0.00	0.00	0.00
LOC	1	0.42	0.39	0.37	0.39
LOC	2	0.56	0.37	0.58	0.35
LOC	3	0.20	0.40	0.17	0.39
LOC	4	0.71	0.36	0.68	0.36
LOC	5	-0.22	0.47	-0.14	0.45
LOC	6	0.37	0.37	0.34	0.37
LOC	7	0.00	0.00	0.00	0.00
VCOND	0	1.07	0.52	0.97	0.50
VCOND	1	0.00	0.00	0.00	0.00
Scale	1	0.01	0.11	0.72	11.28

Table 38: Coefficients of the Generalized linear models.

For the variables describing the state of vegetation, we can observe quite high standard error, varying from 45 to 50% for model 1 and 52 to 62% for model 2. This is also reflected in Figure 24 and Figure 25, representing the Wald 95% confidence interval, where large confidence intervals are observed with respect to the estimates' values.

Conclusions

The results of our Generalised linear models showed that, even if we obtained significant models, the uncertainty on the parameters is still quite important since we obtained large confidence intervals;

As for the classical linear models, our conclusion is that the situation where low mortality is associated to very bad vegetation conditions prevents us to compute predictive models for assessing livestock mortality.

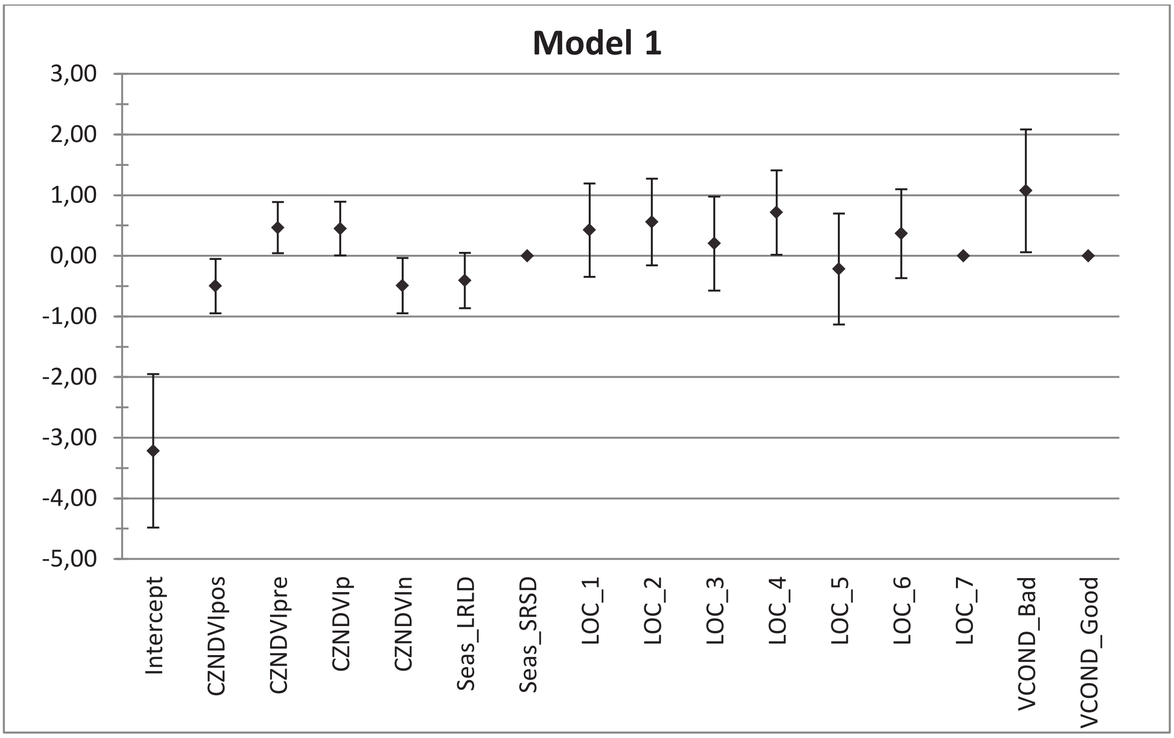


Figure 24: Coefficients estimates and their associated Wald 95% confidence interval for model 1.

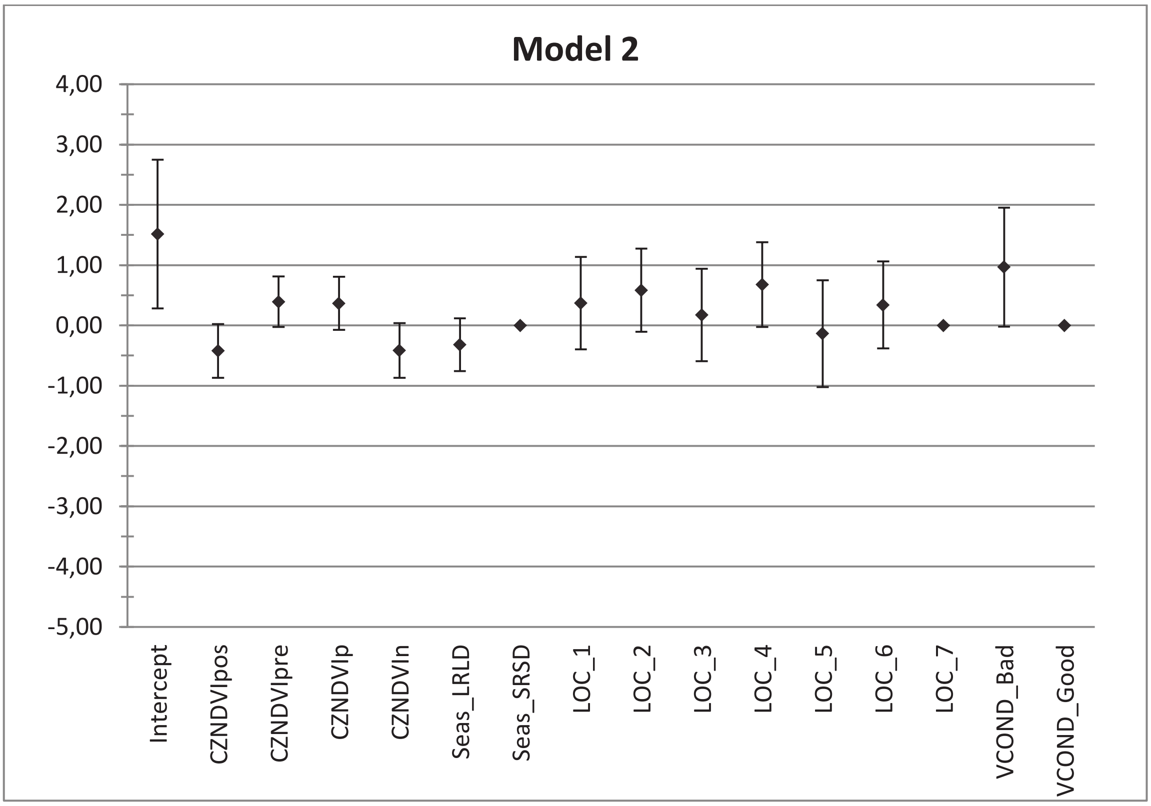


Figure 25: Coefficients estimates and their associated Wald 95% confidence interval for model 2.

5.3.5 Influence of location and season on drought related livestock mortality

The work that will be presented in this section focusses in particular on the study of the influence of geographical location and season on seasonal livestock mortality. The question we tried to answer here is: could there be a geographical and seasonal influence on livestock mortality (directly or through NDVI)? In other words, do location and season explain significantly the trend in mortality?

As it was previously mentioned, this study was conducted because of some particularities observed during the attempt of linear modelling, i.e. low mortality associated with very poor vegetation conditions.

This research was partly presented in the AfricaGIS conference in Addis Abeba, Ethiopia, November 2013 (see RD6) and in the African Association of Remote Sensing of the Environmentconference in El Jadida, Morocco, November 2012 (see RD7).

a. Materials and methods

In this study, we use seasonal constructed NDVI-derived variables and livestock mortality. The data set characteristics specified in section 5.3.2 apply here.

In preparation of the livestock mortality modelling, we tried to determine if the season and geographical parameters have an influence on livestock mortality. We therefore performed an analysis of variance (ANOVA) using one of the two categorical variables. We performed an F test to obtain an indication of the contribution of the parameters to the total variance of the livestock mortality model through the F ratio, representing the ratio between the variance explained by the variable in our case (or model in general) and the model unexplained variance. We then combined the two factors to study the influence of each pair i on overall mean livestock mortality ($\overline{M}$) through the coefficients of the equation (5):

$$(5) \qquad \overline{M} = \sum_{i=1}^{n} \alpha_i \overline{M_i}$$

Where $\overline{M}$ is the general mortality mean, $\overline{M_i}$ is the mean of mortality due to parameter i (e.g. mean of mortality in location 2, during the LRLD season) and α_i gives the contribution of this parameter to the general mean.

Results

The results of the analysis of variance (ANOVA) of livestock mortality according to the locations and seasons are presented in Table 39.

Param.	Degrees of Freedom (DF)		Mean squares *		F	Prob > F
	Variable	Error	Variable	Error		
Location	6	154	0.018	0.020	0.889	0.505
Season	1	159	0.011	0.020	0.564	0.454

Table 39: Results of one variable ANOVA (* mean squares = sum of squares/DF).

The analysis of variance (ANOVA) of livestock mortality according to the location, the F test showed a ratio of explained to unexplained variance of 0.889, with an associated probability of 0.505. In the case of the seasonal parameter, the values are lower, with an F ratio h of 0.564, and an associated p value of 0.454. We can conclude from this that the geographical parameter, and to a lesser extent the seasonal parameter, explain partially the variance and trend observed in livestock mortality but that their influence is not significant.

We then studied the effect of the combinations of location and season on livestock mortality, compared to a reference location (7) and season (SRSD). As we can see in Table 40, the amplitude of the coefficients can vary by a factor 10 between the locations. We conclude that some combinations of location and season have a stronger impact on mean livestock mortality.

Parameter value (i)	Coeff. value (α_i)
Intercept	0.103
Season[LRLD]	-0.008
Location[1]	-0.007
Location[2]	0.027
Location[3]	-0.007
Location[4]	0.040
Location[5]	-0.036
Location[6]	0.013
Season[LRLD]*Location[1]	-0.008
Season[LRLD]*Location[2]	-0.043
Season[LRLD]*Location[3]	-0.005
Season[LRLD]*Location[4]	0.004
Season[LRLD]*Location[5]	0.016
Season[LRLD]*Location[6]	0.015

Table 40: Effect of the location and season on the global mortality mean.

This table also reveals that locations can be grouped according to their behaviour with respect to mortality as following: 1-3-7, 6-2 and 4-5.

From the results of the analysis of variance, we can conclude that livestock mortality does not vary significantly between the observation sites and seasons, which would suggest that a global model should be appropriate. However, if we analyse the data in more detail, by studying the impact of the combined factors, we can observe that there could be some differences at the individual level that were previously hidden.

Conclusions

From the results of the analysis of variance, we can conclude there is no significant global effect of location and season on the mortality trend, i.e. livestock mortality does not vary significantly between the observation sites and seasons. This would suggest that a global model should be appropriate. However, a more detailed analysis of the data, in the form of a study of the impact of the combined factors (ANOVA with two variables), revealed that all the different combinations of geographical and seasonal parameters do not contribute equally to global livestock mortality. Finally, analysis of variance also showed that we could group locations that exhibited a similar behaviour regarding mortality (1-3-7, 6-2 and 4-5).

This study showed that, in addition to forage availability, other effects; that are reflected in the geographical parameter; should be considered in drought related livestock mortality modelling.

Various factors may have an influence on mortality; some were already highlighted in Chantarat et al., 2013. The new elements that will be examined include:

- variations in local meteorological and vegetation conditions,
- differences in herding techniques and systems of production: size of the herd, species of animals kept, mobility of the herders …
- adaptation of herders to drought …

5.3.6 General conclusions and perspectives

Multiple linear regressions computed between seasonal livestock mortality and vegetation indicators exhibited very low R^2 and non-normal residuals. Although we could fix this second problem by considering Generalized linear models, our models still show very poor predictive results (large errors and confidence interval on our model parameters).

The primary reason might be the presence of cases where low mortality is observed, despite severe conditions on the field (reflected in the vegetation indicators). Our hypothesis is that, in severe drought cases, pastoralists move further away from their location and that, in this case, the boundaries used to delineate the grazing areas do not represent the actual grazing areas.

An adaptation of the method regarding this aspect is therefore needed.

We also investigated the spatial and temporal aspects of our system and found that different combinations of geographical and seasonal parameters do not contribute equally to global livestock mortality. We conclude that, in addition to forage availability, other effects; that are reflected in the geographical parameter; should have an impact on drought related livestock mortality modelling. Comparison of the characteristics of the locations and seasons that show similar and opposite behavior should be performed in order to identify these effects and later integrate them to the modelling process.

5.4 Global use case

5.4.1 G-Range

G-Range is a global model that simulates generalized changes in rangelands through time, created with support from the International Livestock Research Institute. Spatial data and a set of parameters that control plant growth and other ecological attributes in landscape units combine with computer code to represent ecological process such as soil nutrient and water dynamics, vegetation growth, fire, and wild and domestic animal offtake. The model is spatial, with areas of the world divided into square cells. Those cells that are rangelands have ecosystem dynamics simulated. A graphical user interface allows users to explore model output.

The G-Range application captures main primary production and its dynamics. It is of moderate complexity, and of a nature that a user may learn its use in a week or less. A monthly time step is used to simulate herbaceous plants, shrubs, and trees, and those plant types can change in their covers within each landscape cell through simulated time. The model represents all rangelands within a single computer process, which simplifies the logistics involved in analyses. Simulations may span a few to thousands of years. Detailed information about G-Range and the reason for its creation are described in Boone et al. (2011).

Extensive sensitivity analyses were then conducted to establish the highest degree of agreement with simulated output from the Century model (Conant et al., *in prep.*), which has been vetted many times, as well as several additional biophysical surfaces. For sensitivity analyses we began with a list of the 100+ responses from G-Range, and then selected from those the types of responses that would be most helpful in comparisons. We then compiled from existing Century simulations or materials from web sites spatial data at 0.5 degree x 0.5 degree resolution, and reported the results in a report posted online (Boone et al. 2014; http://warnercnr.colostate.edu/~rboone/pubs/GRange_II_Sensitivity_Report_Rotated.pdf). Following the sensitivity analyses, we made several significant changes to the structure of the model, improving its internal consistency and debugging the code of the program.

Next, we evaluated G-Range using data at global and site scales, from a variety of sources. At global scales, total net primary productivity (TNPP) was evaluated against TNPP from MODIS satellite imagery (Zhao et al. 2006), and cover of herbs, shrubs, and trees was evaluated against vegetation continuous fields (VCF), also from MODIS (Hansen et al. 2006).

At site scales, field biomass harvest data from 39 independent studies from the literature and online databases was compiled, yielding 317 site-years of aboveground net primary production (ANPP), and 95 site-years of belowground net primary production (BNPP). ANPP was quantified as peak standing crop (maximum live+dead biomass), and BNPP was from a variety of methods, primarily climate-adjusted means of peak standing crop (maximum live+dead root biomass) and maximum-minimum root biomass, a global regression equation (Gill and Jackson 2000), as well as root ingrowth, ^{14}C radioactive isotope labeling, and minirhizotron methods where these superior methods were available. These results are presented in Figure 1. To achieve consistency across spatial scales and methodologies, the 0.5° grid cells containing sites with field biomass harvest data were also used to evaluate model agreement with global surfaces at site scales, specifically MODIS TNPP, MODIS VCF, and Century total soil carbon (Figures 2-3). Performance benchmarks were established as absolute difference less than 100 g m^{-2} for ANPP, 200 g m^{-2} for BNPP, 300 g m^{-2} for TNPP, 10% herb/shrub and tree cover, and 200 g m^{-2} for total soil carbon, or less than 100% difference (to accommodate sites with large values). Once all data confrontations at all scales exhibited performance within the established benchmarks for a majority of biomes, the parameterization was deemed final for G-Range v.1.1, enabling re-release of the model for application purposes. In addition to improving the consistency of the model across scales, the new parameterization provided substantial gains in model fit for several output variables, including soil carbon and TNPP, at global scales.

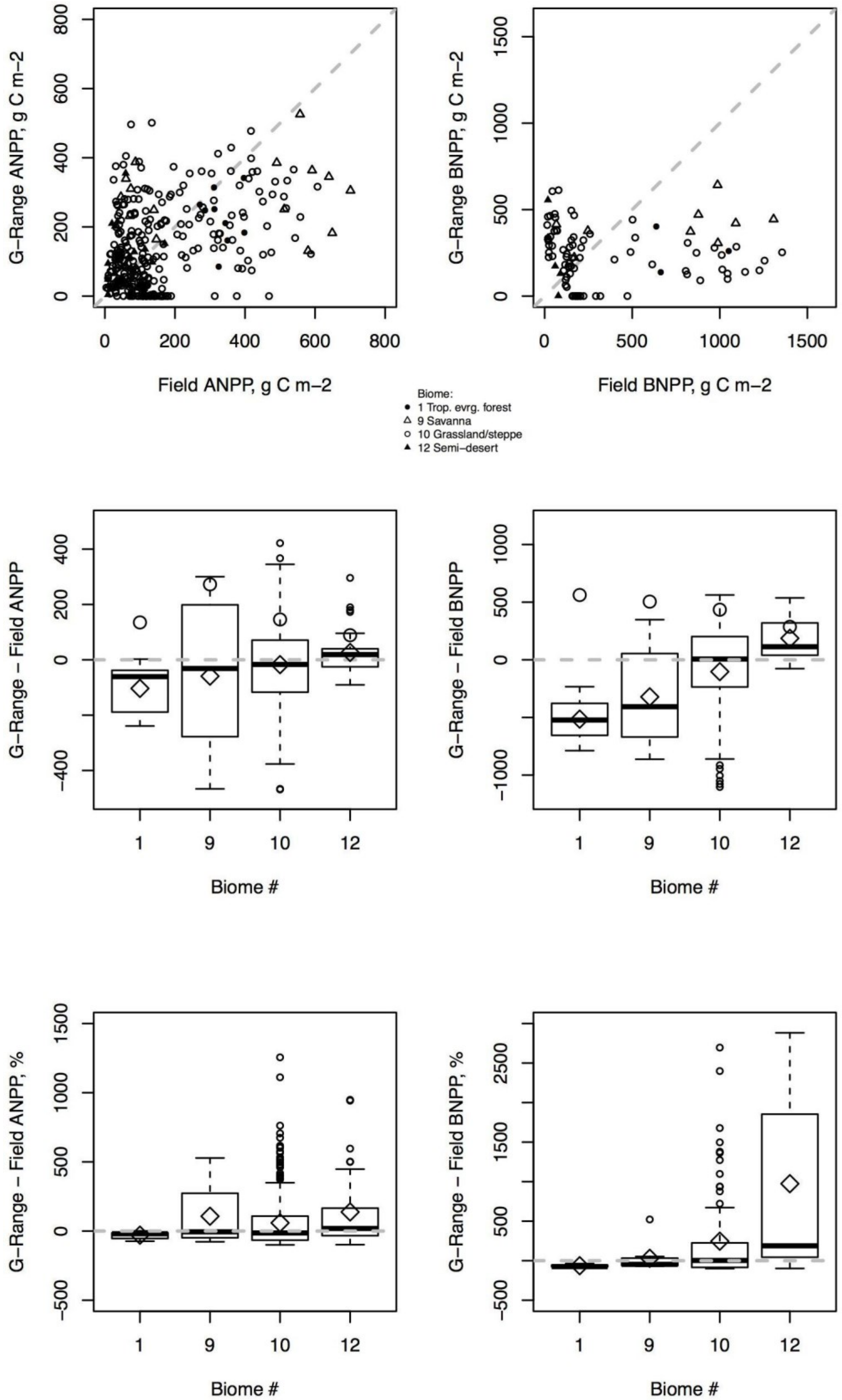


Figure 26: G-range model evaluation results for field biomass harvest data.

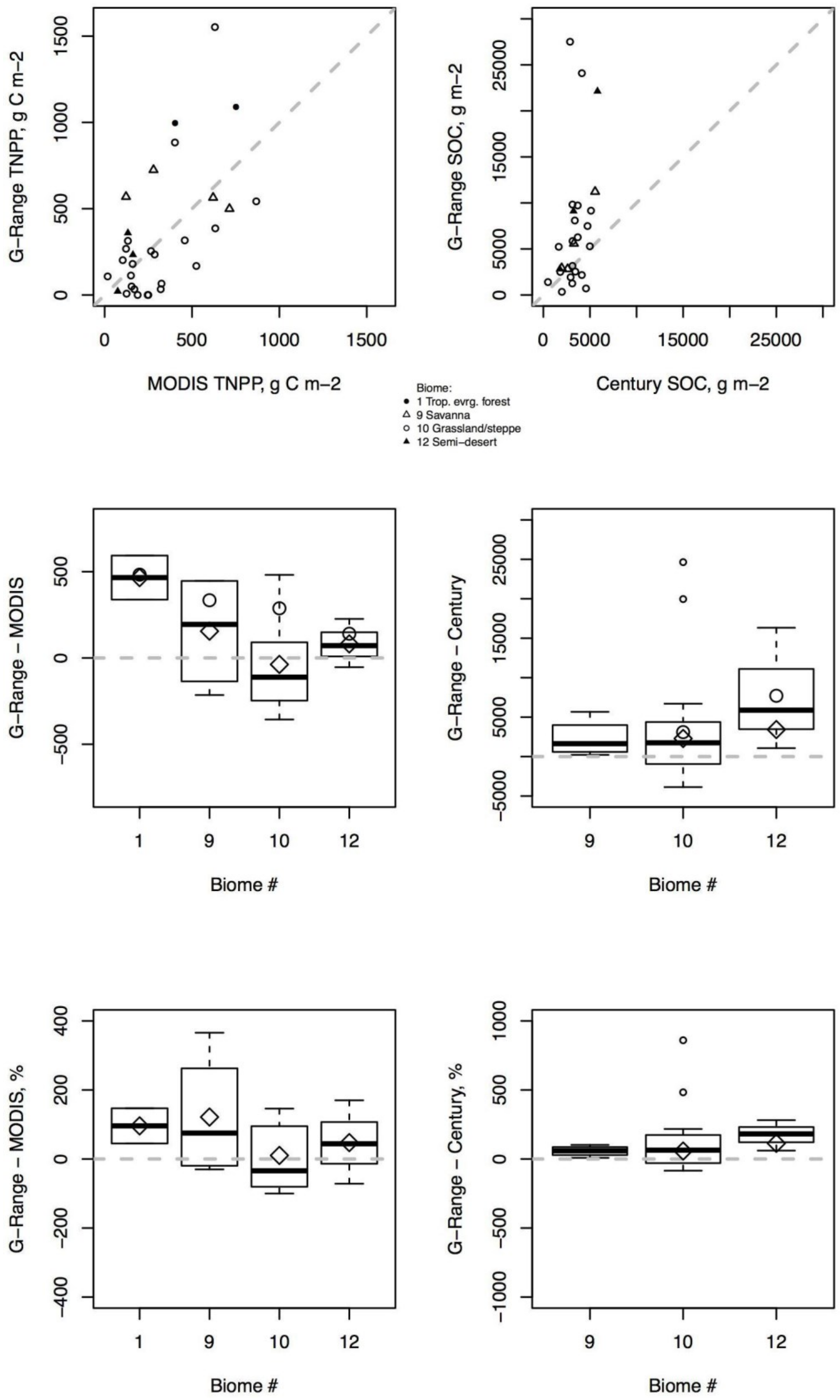


Figure 27: G-Range model evaluation results for MODIS TNPP and Century soil carbon.

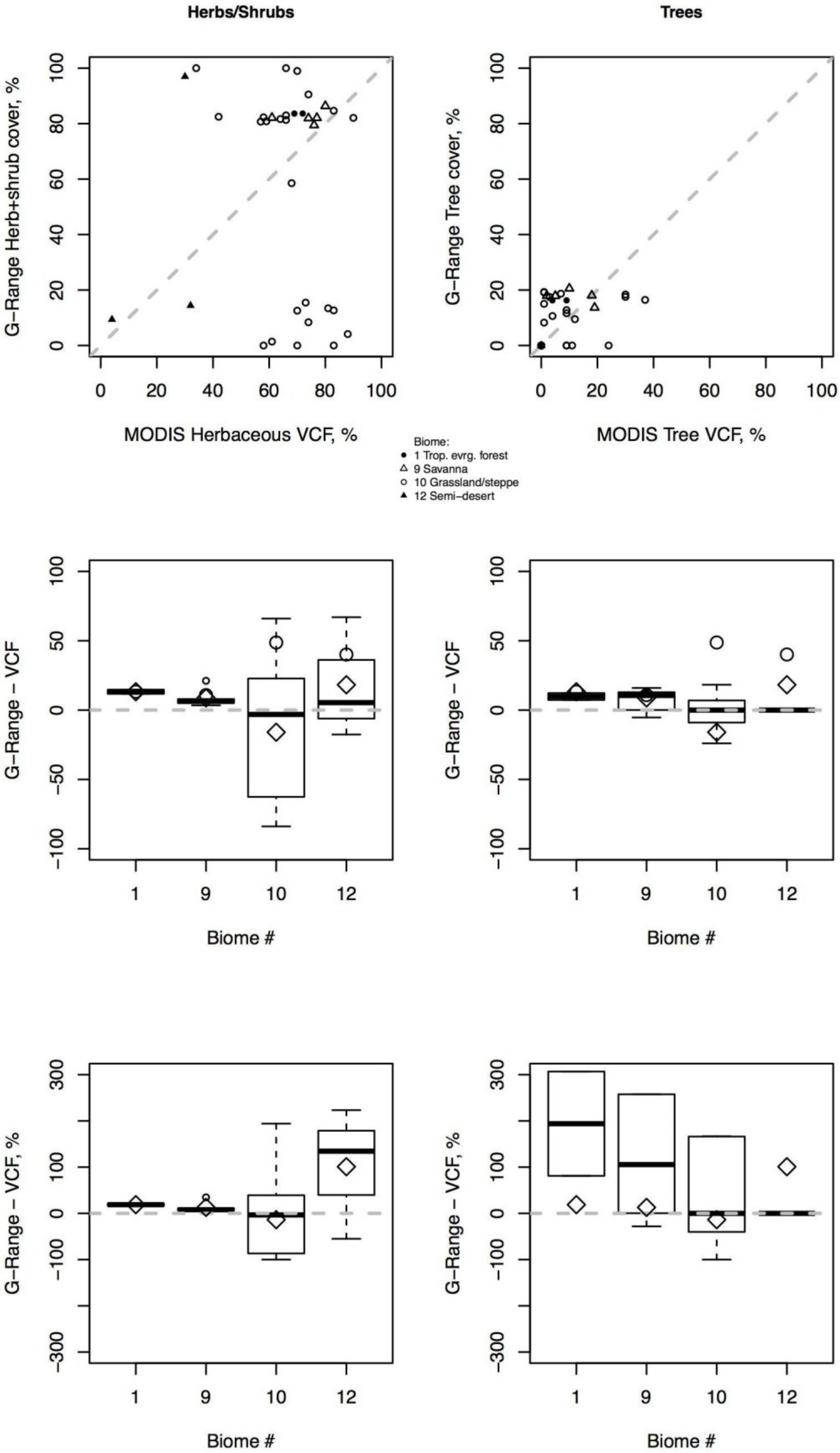


Figure 28: G-Range model evaluation results for MODIS VCF.

5.4.2 Uses and publications

Multiple applications and projects using the G-Range model are either completed or in-progress. Now that the model is ready for application, we are beginning several projects. First, the compilation of the biomass dataset led to a review of biomass harvest methods (Sircely, *in review*). A formal model description is being prepared presently, and will be published with the field observation and global dataset evaluation presented here. We are conducting forecasting of rangeland vegetation and ecosystem services under climate change through the year 2070, submission of which is also imminent.

6 Impact of use case development

6.1 Biomass estimation in Senegal and Niger

It is foreseen that the activities in this WP will allow to implement operational forage biomass models based on remote sensing data in West Africa, i.e., in Senegal and Niger. The two countries Senegal and Niger have operational network for data collection on forage resources for several years. The models can be designed to be the starting point of early warning systems for livestock vulnerability. The technical implementation of this activity would be led by the livestock department unit of the AGRHYMET Regional Centre (Niger), which covers several countries of West Africa, and the Senegal Centre de Suivi Ecologique (CSE, Senegal).

The refined models of biomass prediction that are currently designed through the works of Abdoul Aziz Diouf in Senegal and Issa Garba in Niger will permit to improve the biomass prediction using remote-sensing derived indices, such as NDVI and fAPAR. It has been already shown by Abdoul Aziz Diouf that, in Senegal, simple linear regression models are outperformed by power and exponential models, that also lead to more consistency, i.e., negative value of biomass prediction are avoided. Multilinear regression approach may also improve the prediction performance with using additional information such as phenological parameters from remote sensing. The multilinear regression approach could be also used to develop more accurate models at the zonal scale (eco-geographical and/or climatic). In Niger, the analysis of the biomass production by ecophysiological facies showed also the advantage of making the approach at the zonal scale.

6.2 Index Based Livestock Insurance implementation

Activities of this WP related to the index based livestock insurance (IBLI) are implemented in Northern Kenya. The IBLI was initially designed by ILRI and was already tested with the first IBLI contract that was sold in Marsabit district, Northern Kenya, in January and February 2010. This first implementation of the IBLI used freely available NDVI data (MODIS NDVI from USGS).

Multiple linear regressions and generalized linear models were both tested in order to improve the prediction of livestock mortality with the remotely-sensed biomass indices. Not single values of NDVI were used but standardized values of NDVI over space (ZNDVI) as well as cumulative values of the latter index. As similarly for the biomass prediction in West Africa, explanation of the mortality by ecophysicological zoning was attempted, as well as taking into account the seasonality of the mortality. The IBLI tool that has to be developed must account for these geographical and seasonal variations and should be general enough to be adapted to local conditions.

7 Conclusions

7.1 Conclusions on capacity building activities

Eight capacity building workshops were organized in the framework of WP32 from December 2012 to May 2014: two in Belgium, two in Niger, and in Burkina Faso, Kenya, Sudan, and USA. The activities were related to models of fodder biomass production in African rangelands, with an emphasis on the use of remote sensing data. Capacity building activities that were made within the WP were particularly designed for the needs of the users and showed a close collaboration with the trainers. In particular, the three PhD students and the post-doc involved in the WP closely collaborate in these workshops and provided their own dataset and/or results for feeding the workshops. For instance, Abdoul Aziz Diouf assisted the trainer in the explanations to the users, in the Ouagadougou workshop, and Issa Garba provide field data to the other participants in the Arlon workshop.

7.2 Scientific conclusions

The scientific developments within the WP32 were made by the work of the three PhD students and post-doc involved in the project, Abdoul Aziz Diouf, Issa Garba, Marie Lang, and Jason Sircely. The first two PhD researches are about the development of fodder biomass models using remote sensing and field data in Senegal (Abdoul Aziz Diouf) and Niger (Issa Garba), in the view of implementing early warning systems for the livestock sector. The last PhD research (Marie Lang) is about the implementation of an index-based livestock insurance (IBLI) scheme in Kenya with the objective of insuring pastoralists during drought periods. The post-doctoral researcher has led efforts to validate the G-Range model at the site level.

Results of Abdoul Aziz Diouf in the Senegal case study were based on a huge database of fodder biomass field data with historic (since 1999) and newly collected data. Simple regression models between NDVI and fodder biomass were tested using the database. Then, multilinear regression models using different phenological metrics of NDVI were developed and led to relative RMSE below 30%. The perspectives are to test alternatives sources of data and vegetation indices, as well as to develop the model by ecoregions (phenoclimatic zoning)

Similarly, the work of Issa Garba focused on the development of new regression-based models for linking in situ measured fodder production to biophysical variables from remote sensing data, but in Niger. A unique database of biomass field data was collected. Global multilinear model adjustments were tested using different biophysical variables from remote sensing data and outputs from agrometeorological models. A similar approach than the global adjustment was made by ecoregions, which resulted in a slight decrease in the relative RMSE from 40% (global) to 36% (Azawak ecoregion) and 33% (Manga2 ecoregion). Better results were also obtained when decomposing the model per year and per facies.

The work of Marie Lang focused on making a model of livestock mortality forecasting using biophysical variables from remote sensing data, in order to implement an index-based livestock insurance. First results showed that multiple linear regressions computed between seasonal livestock mortality and vegetation indicators exhibited low R^2 and non-normal residuals. The non-normal residuals problem was resolved by using generalized linear models, but still, the predictability of the livestock mortality by the model was poor. The main reason might be an incorrect delineation of the grazing areas in the data and that pastoralist do migrate in case of severe conditions, leading to low mortality values while severe conditions occur. As similarly to some results in Senegal and Niger, taking into account some geographical patterns (i.e., zoning by ecoregions) might help to improve the predictive power of the models.

The work of Jason Sircely has engaged with partners at Agricultural Research Corporation (Wad Madani, Sudan), who received G-Range training under AGRICAB, and who seek to conduct forecasting of forage production under climate change, especially to inform livestock supplementation policies of the Sudanese government. With colleagues at ILRI, the model was applied to quantify forage variability in the Greater Horn of Africa, to inform drought preparation and response policies of the IGAD regional government, and will soon begin working with Kenyan government researchers to provide forage projections useful in drought preparation and response, several of whom also received G-Range training via AGRICAB.

REFERENCES

Adow, M. (2008). Pastoralists in Kenya, Forced Migration Review, 31, 34

Barret, B.C. & McPeak, J.C (2006). Poverty Traps and Safety Nets, In Silber J. (Ed), Poverty, Inequality and Development – Economic Studies in Inequality, Social Exclusion and Well-Being (pp. 321-332). US.: Springer.

Bégué, A., Vintrou, E., Ruelland, D., Claden, M., Dessay N. (2011). "Can a 25-year trend in Soudano-Sahelian vegetation dynamics be interpreted in terms of land use change? A remote sensing approach." Global Environmental Change 21: 413–420.

Boone, R.B., Conant, R.T., Hilinski, T.E. (2011). G-Range: Development and use of a beta global rangeland model. Final report to the International Livestock Research Institute, Nairobi, Kenya.

Boone, R., Conant, R., Sircely, J. (2013). Adjustment and Sensitivity Analyses of a Beta Global Rangeland Model. CGIAR Research Program on Climate Change, Agriculture and Food Security (CCAFS), Montpellier, France.

Box, E., Holben, B., Kalb, V. (1989). Accuracy of the AVHRR vegetation index as a predictor of biomass, primary productivity and net CO2flux. Vegetatio 80:71–89.

Carrière, M. (1996) Impact des systems d'élevage pastoraux sur l'environnement en Afrique et en Asie tropicale et sub-tropicale aride et sub-aride. CIRAD, Montpellier.

Cesaro J.D., Magrin G., Ninot O. (2010) Atlas de l'élevage au Sénégal, CIRAD, Paris.

Chantarat, S., Mude, A.G., Barret, C.B. & Carter, M.R. (2013). Designing Index-Based Livestock Insurance for Managing Asset Risk in Northern Kenya, The Journal of Risk and Insurance, 80(1), 205-237.

Crépeau C, Bennouna T, Bicheron P, et al., (2003). Suivi de la végétation en zone sèche par des produits d'information dérivés de l'imagerie spatiale. Sécheresse en ligne 2003 ; 1E (1).

Diallo, O., Diouf, A., Hanan, N. P., Ndiaye, A., and Prévost, Y., (1991). "AVHRR monitoring of savanna primary production in Senegal, West Africa: 1987-1988." International Journal of Remote Sensing 12(6): 1259-1279.

Dicko, M. S., Djitèye, M. A., Sangaré, M. (2006) Les systèmes de production animale au Sahel. Science et changements planétaires / Sécheresse, 17(1), 83-97.

Diouf, A. A., Djaby, B., Diop, M. B., Wele, A., Ndione, J.-A., Tychon, B. (2014). Fonctions d'ajustement pour l'estimation de la production fourragère herbacée des parcours naturels du Sénégal à partir du NDVI s10 de SPOT-vegetation. XXVIIème Colloque de l'Association Internationale de Climatologie, 2-5 juillet 2014 –Dijon (France).

Diouf, A., Lambin, E.F., (2001). Monitoring land-cover changes in semi-arid regions: remote sensing and field observations in the Ferlo, Senegal. Journal of Arid Environments 48, 129–148.

Eklundh, L., and Jönsson P., (2011). Timesat 3.1, Manuel software. 79 pages.ISRA, (2003). Rapport national sur l'état des ressources zoo-génétiques au Sénégal. 54pages.

Gill, R.A., Jackson, R.B. (2000). Global patterns of root turnover for terrestrial ecosystems. New Phytologist 147: 13-31.

Hansen, M., DeFries, R.S., Townshend, J.R.G., Carroll, M., Dimiceli, C., and Sohlberg, R.A. (2006). Vegetation Continuous Fields MOD44B, Collection 4. University of Maryland, College Park, Maryland, USA.

Hiernaux, P. H. Y. et Justice, C. O. (1986): Suivi du développement végétal au cours de l'été 1984 dans le Sahel Malien, International Journal of Remote Sensing, 7:11, 1515-1531.

Jamieson, P.D., Porter, J.R., Wilson, D.R., (1991). A test of computer simulation model ARC-WHEAT1 on wheat crops grown in New Zealand. Field Crops Res. 27, 337–350.

Jarlan, J., Mangiarotti, S., Mougin, E., Mazzega, P., Hiernaux, P.H.Y., & Le Dantec, V. (2008). Assimilation of SPOT/VEGETATION NDVI data into a sahelian vegetation dynamics model. REMOTE SENSING OF ENVIRONMENT, 112, 1381–1394.

Jiang, J. (2008). Linear and Generalised Linear Mixed Models and Their Applications, New York: Springer.

Jin, Y.X.; Yang, X.C.; Qiu, J.J.; Li, J.Y.; Gao, T.; Wu, Q.; Zhao, F.; Ma, H.L.; Yu, H.D.; Xu, B. (2014). Remote sensing-based biomass estimation and its spatio-temporal variations in temperate grassland, Northern China. Remote Sens. 6, 1496–1513.

Kouadio, L., Djaby, B., Duveiller, G., El Jarroudi, M. and Tychon, B. (2012). Cinétique de décroissance de la surface verte et estimation du rendement du blé d'hiver. Biotechnol. Agron. Soc. Environ. 16(2), 179-191.

ISRA (2003). Rapport national sur l'Etat des ressources zoogénétiques au Sénégal, Dakar, Sénégal.

Lynch, S.M. (2007). Introduction to Applied Bayesian Statistics and Estimation for Social Scientists, New York: Springer.

Mude, A.G., Barrett, C.B., Carter, M.R., Chantarat, S., Ikegami, M. & McPeak J.G (2010). Project Summary: Index Based Livestock Insurance for Northern Kenya's Arid and Semi-Arid Lands: The Marsabit Pilot. ILRI Project Reports, Papers and Documents, Retrieved August 27, 2013, from http://mahider.ilri.org/handle/10568/494

Prince, S. D. (1991). Satellite remote sensing of primary production: comparison of results for Sahelian grasslands 1981–1988. International Journal of Remote Sensing, 12, 1301–1311.

RCPA (2010), L'élevage au Sahel et en Afrique de l'Ouest, Proceedings of the 26th annual meeting of the « Réseau de Prévention des Crises Alimentaires » (RPCA), Accra, Ghana.

Santin-Janin, H., Garel, M., Chapuis, J.-L., Pontier, D., (2009). "Assessing the performance of NDVI as a proxy for plant biomass using non-linear models: a case study on the Kerguelen archipelago." Polar Biology 32: 861-871.

Saporta, G. (2006). Probabilités, analyse des données et statistique. Deuxième édition révisée et augmentée. Éditions Technip, Paris, 622 pages ; ISBN 2-7108-0814-5.

SAS Institute Inc. (2008). SAS/STAT 9.2 User's Guide, Cary, NC: SAS Institute Inc.

Sircely, J. In review. Accuracy and reliability of biomass harvest methods for net primary production in grasslands. In review, Ecosystems.

Skees, J.R. (2008). Challenges for use of index-based weather insurance in lower income countries. Agricultural Finance Review, 68(1), pp.197 – 217.

Tobias, R. (1995), "An Introduction to Partial Least Squares Regression," in Proceedings of the Twentieth Annual SAS Users Group International Conference, Cary, NC: SAS Institute Inc., 1250–1257.

Tucker, C.J., Vanpraet, C., Boerwinkel, E., Gaston, A. (1983). Satellite remote sensing of total dry matter production in the Senegalese Sahel, Remote Sensing of Environment, Volume 13, Issue 6, Pages 461-474.

United Nations Department of Economic and Social Affairs, Division for Sustainable Development (2007). Developing Index-Based Insurance for Agriculture in Developing Countries, Sustainable Development Innovation Briefs, 2, 1-7.

Vescovo, L., Wohlfahrt, G., Balzarolo, M., Pilloni, S., Sottocornola, M., Rodeghiero, M. & Gianelle, D. (2012). New spectral vegetation indices based on the near-infrared shoulder wavelengths for remote detection of grassland phytomass, International Journal of Remote Sensing, 33:7, 2178-2195.

Xiaoping, W., Ni, G.; Kai, Z.; Jing, W. (2011). Hyperspectral Remote Sensing Estimation Models of Aboveground Biomass in Gannan Rangelands. Procedia Environnemental Sciences Vol. 10 (2011). p. 697-702.

Zhao, M., Running, S. W. and Nemani, R. R. (2006). Sensitivity of Moderate Resolution Imaging Spectroradiometer (MODIS) terrestrial primary production to the accuracy of meteorological reanalyses. Journal of Geophysical Research 111, G01002.